Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls – University

Heidelberg

Presented by

Alexander Stratford Cook, MSci

Born in: Hemel Hempstead, United Kingdom

Oral examination: 4th December 2024

Towards photoacoustic neuroimaging in mice: a chemigenetic approach for acoustogenic calcium sensors

Referees:
Dr. Richard Wombacher
Dr. Robert Prevedel

Inaugural - Dissertation
zur
Erlangung der Doktorwürde
der Gesamtfakultät
für Mathematik, Ingenieur- und Naturwissenschaften
der
Ruprecht – Karls – Universität
Heidelberg

vorgelegt von

Alexander Stratford Cook, MSci

aus: Hemel Hempstead, Vereinigtes Königreich

Tag der mündlichen Prüfung: 4. Dezember 2024

Photoakustische Neuroimaging bei Mäusen: ein chemigenetischer Ansatz für akuptogene Kalziumsensoren

Gutachter:
Dr. Richard Wombacher
Dr. Robert Prevedel

Acknowledgements

Whilst this page is not a scientific part of the thesis, I think that it is the most important of all the 265 pages because without all these people, none of the work in this project would have been possible. So I will try to put my typically British understated demeanour aside for a minute, and in a few short words, which can never do full justice, thank everyone for all their support.

First, I must express my gratitude to my defence committee members, Dr. Richard Wombacher, Dr. Robert Prevedel, Dr. Georgia Rapti and Prof. Dr. Alexis Maizel, for your time and efforts evaluating my TAC reports, presentations and finally this thesis. I hope it is at least an interesting read to see how far we have come with this project and what the future potentially holds.

Thank you, Claire, for giving me the opportunity to start this project in your lab, and for all your guidance and support on this journey. Thank you for your scientific enthusiasm and rigour, always helping to find a solution to any problems along the way, without a doubt you have helped enormously in my development as a scientist over the past four years.

Now to the rest of the lab, such a great team. Thank you, Bego, for always looking after us like we are your extended family, providing medicines, celebrating our birthdays! Thank you for showing me the ways of the "duck" and being so patient with us meddling chemists in your pond, there's no way the lab could function without you (and we wouldn't have spotted all the Doppelgängers!). Thank you, Franzi, for your infectious energy, your excitement for seemingly random things (at least in my eyes!) from Gluggle Jugs to Okonomiyaki! Thank you being the primary source of my knowledge for all things Heidelberg, especially in those early days, and of course the delicious Plätzchen every Weihnachten. Thank you, Ana Clara, unfortunately you joined in those last months when my energy and motivation levels were dipping, but your friendly, positive, good-natured personality has really helped to keep me going, and I know that the lab is in good hands and you won't find "a better place to play" (to steal some Oasis lyrics)! Oh don't worry Magda, I haven't forgotten you, how could I with all that TikTok knowledge you imparted and a movie list which I don't think I will ever finish! Thank you for your bubbly enthusiasm, those spontaneous chats by our fume hoods, and teaching me to always smile no matter the situation. Thank you also to the interns who have stayed with us for a few months, Morgane, Anna and Iulia, for adding your own character to the group.

The second excellent team I have been lucky enough to be a part of is the photoacoustic team. Thank you, Robert, for kickstarting and leading this project and your calm, measured approach for resolving the different challenges we have faced. Thank you, Juan and Gretel, for making time for our experiments and the invaluable discussions, helping me to gain an understanding of neurobiology. I gain quite some comfort that the mice were treated well in your careful hands and you always ensured they were only used when completely necessary. My biggest thanks must be saved for my accomplice and friend, Nikita. Thank you for your unwavering patience at firing lasers at tubes and the random dye solutions I delivered, and your creative suggestions for trouble-shooting. Thank you for your courage and resilience, especially in the darkest hours when the Fabry-Pérot cracked the night before the crucial brain imaging experiment. Thank you for all our chats about bikes and adventures, showing me that we can still explore the world during our PhDs, and giving me a taste of the Berlin lifestyle even though we're 600 km away.

I would also like to thank everyone else at EMBL and beyond who contributed to this project: the generous Lavis, Schreiter and Johnsson groups for sharing dyes, plasmids and cell lines; Dr. Dieter Schollmeyer for the X-ray crystallography; Dr. Lucia Alvarez (Hennig group) for the stopped-flow tutorial; Federico Marotta (Bork group) for setting up Colabfold on the cluster and patiently explaining how to use it to a computational novice like myself; Dr. Judith Weber (Spark Berlin) for valuable photoacoustic probe discussions at the beginning of the project; the EMBL communications team for the nice article on our project (Ivy Kupec, Isabel Romero Calvo); Murat Kucukdisli and Florian Braun for sharing group meetings, chemicals, and saving us when our instruments were broken (Chemical Synthesis Core Facility); and all the other EMBL Core Facilities for their services and support, Protein Expression and Purification (especially Karine Lapouge), Metabolomics (Bernhard Drotleff), Advanced Light Microscopy Facility (Beate Neumann and Manuel Gunkel), Mechanical Workshop, Scientific Instruments Maintenance, and the Gene Editing & Virus Facility in Rome.

I think I am very fortunate to have a wonderful support network and group of friends who I can always rely on and who have supported me throughout this PhD from near and far (yes, I know the UK is only a 90-minute flight away but it does feel a long way at times!), providing consolation when necessary, and so much enjoyment and unforgettable memories. Unfortunately, I can't name all of you here as I think the binding limit for the thesis might be exceeded, but I would like to thank specifically a few people, and I can't wait to celebrate with you all after the defence or in the near future.

Thank you, Manu, whilst the olive oil with basil was the anointed "game changer" in the kitchen in our first few recipes, you were the real game changer to my life in Heidelberg in those first Covid-stricken months. Thank you for those wild parties, expanding my culinary repertoire (although I'm not sure Cécile will ever forgive you for all those Tuesday nights when I returned stinking of onion), our adventures from Italy to India, and I'm certain our non-stop banter will continue as you cycle between HBS and HDW. Thank you, Kavan, for your enforced coffee breaks, always being an open ear to ease our collective frustrations and keeping me on my toes with your teasing jokes. Thank you also for your adventurous spirit, despite all the hurdles of this PhD, keeping us alive in the Himalayas and man-eating tiger territory, and I'm sure your determination will carry you far in your next steps. Thank you to everyone at TSG Heidelberg 78 for providing the perfect escape, tearing around the track and roads together, and ensuring I made the most of the beautiful Heidelberg trails in my time here. Thank you to all the cricket guys at TSG Rohrbach: to be honest I did not expect to play at all when moving to Germany, but I couldn't have asked for a more friendly, welcoming team and will miss those sunny summer Sundays chasing after a leather ball with you. Thank you, Nick, for visiting me so often, I loved those weekends exploring Germany with you, and your positive, dreamy, outlook on life is always a breath of fresh air.

A big thank you to all of my family back home for your unconditional love and support. It's so great to have such a warm, welcoming environment to return to every time I come back for holidays, always wonderful to see you, and catch up with relaxed, lovely conversations. In particular, I would like to thank Grandad in particular for his curiosity and tenacity trying to understand my work from the unit seminars and papers! Now it's time for an even bigger thank you to my parents. Thank you, Mum and Dad, for instilling in me a love for exploration and adventure, and for your strength in encouraging your only child to move abroad to pursue his career. Thank you for all the time invested in my education, not just academically, but as a person and I for sure wouldn't be standing here today at the end of this PhD if it was not for all the lessons you have taught me. Thank you for all of your endless love and affection, always welcoming me back home with open arms, or embarking together on Tracey Tours to tick off all the castles in Germany!

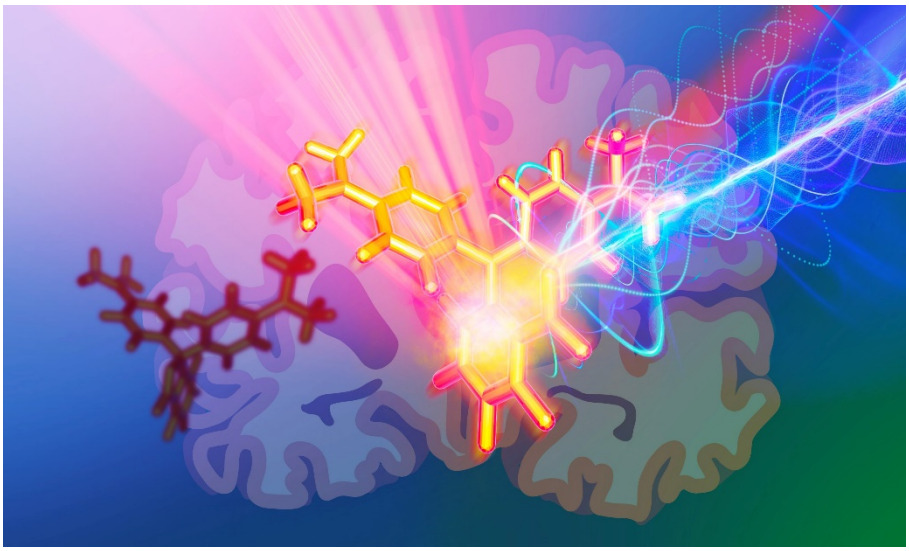
Enfin, mais certainement pas le moindre, je voudrais remercier Cécile d'être à mes cotes, de m'avoir soutenue dans tous les hauts et les bas. Merci pour ton amour inconditionnel et débordant, qui pouvait essuyer immédiatement mes frustrations après une longue journée de travail, et merci d'avoir célèbre mes succès comme si c'étaient les tiens. Merci de m'avoir inspiré, avec ta persévérance, ton courage, en restant fidèle à tes valeurs admirables et en ne reculent jamais malgré toute l'opposition que tu as rencontrée: te rendre fière de moi a été une grande motivation. Merci d'avoir vu la beauté dans toutes les petites choses de la vie avec ton sourire radieux et ton rire chaleureux, les idées folles que tu m'as convaincu d'essayer, ton désir insatiable d'expérimenter et d'essayer quelque chose de nouveau, pour toujours être prêt pour l'aventure et me suivre aveuglément parce que tu as confiance en moi. J'ai hâte de célébrer ta défense de doctorat, lorsque on sera libre d'explorer le monde et commencer notre prochain chapitre ensemble.

Abstract

Photoacoustic imaging is an attractive imaging modality for *in vivo* applications such as neuroimaging, able to image large fields of view, deep in tissue, at high resolution. However, robust molecular reporters are required for this modality, to label specific targets of interest and visualise dynamic biochemical events. Various genetically-encoded or synthetic reporters have been implemented, but suffer from poor optical properties or lack of labelling specificity. Here, we report the development of novel "chemigenetic" contrast agents and calcium sensors for photoacoustic imaging, based on synthetic dyes and the self-labelling HaloTag protein.

I designed and synthesised new photoacoustic reporters based on quenched near-infrared rhodamine-like derivatives, tailored for use with HaloTag to exploit the open-closed equilibrium of the lactone ring. These were characterised *in vitro* for high photoacoustic signal and turn-on upon binding to HaloTag. The most promising ligands were used for the design of calcium indicators, with the HaloTag-based HaloCaMP protein, which show an absorption-modulated change in calcium-dependent photoacoustic signal. Cell permeability of these ligands was confirmed by efficient labelling of live cells cytosolically expressing HaloTag. Our resulting "acoustogenic" calcium sensors show large turn-ons in response to calcium, up to 8-fold, high photoacoustic signal and photostability, outperforming existing far-red sensors for this modality such as NIR-GECO1 in tissue-mimicking phantoms. *Ex vivo* brain slices with neuronal expression of HaloTag, labelled with our ligands, could be consistently visualised via photoacoustic tomography. *In vivo* labelling experiments in mice, however, showed mixed results with the occasional specific labelling of HaloTag-expressing neurons suggesting that the bioavailability of our ligands needs to be improved.

In summary, we developed a novel approach for the design of photoacoustic reporters, with the first chemigenetic probes for this modality. Our acoustogenic dyes and first-generation photoacoustic calcium sensors, show superior *in vitro* performance to existing sensors, and we can label mice brain tissues to produce a localised, strong PA signal. To reach our long-term goal of whole-brain neuroimaging in mice, we have started protein engineering to improve the dynamic range, calcium binding affinity, and used established methods to assess dye bioavailability to optimise this key parameter for future *in vivo* applications. In general, with this hybrid design, we hope to stimulate photoacoustic probe and sensor development which can enable this powerful modality to uncover its true potential in the field of neuroimaging and beyond.



Artistic illustration of our photoacoustic probes for neuroimaging.

Zusammenfassung

Die photoakustische Bildgebung ist eine attraktive Bildgebungsmodalität für *In-vivo*-Anwendungen, wie z.B. das Neuroimaging, da sie große Bereiche tief im Gewebe mit hoher Auflösung abbilden kann. Um spezifische Ziele von Interesse zu markieren und dynamische biochemische Ereignisse sichtbar zu machen, werden robuste molekulare Reporter benötigt. Verschiedene genetisch kodierte oder synthetische Reporter wurden bereits entwickelt und getestet, zeigen jedoch oft schlechte optische Eigenschaften oder mangelnde Spezifität bei der Markierung.

In dieser Thesis berichte ich über die Entwicklung neuartiger „chemigenetischer" Kontrastmittel und Kalziumsensoren für die photoakustische Bildgebung, die auf synthetischen Farbstoffen und dem selbstmarkierenden HaloTag-Protein basieren. Ich habe neuartige photoakustische Reporter entwickelt und synthetisiert, die auf gequenchten rhodaminähnlichen Derivaten im nahen Infrarotbereich basieren. Diese Reporter wurden speziell für die Interaktion mit HaloTag optimiert, um die dynamischen Eigenschaften des offen-geschlossenen Gleichgewichts des Lactonrings gezielt auszunutzen. Die Liganden wurden *in vitro* hinsichtlich ihrer hohen photoakustischen Signale und ihrer Aktivierung bei der Bindung an das HaloTag-Protein charakterisiert. Die vielversprechendsten dieser Liganden wurden zur Entwicklung von Kalziumindikatoren eingesetzt, indem sie mit dem HaloTag-basierten HaloCaMP-Protein kombiniert wurden. In diesem Zusammenhang wurde untersucht, ob sie eine absorptionsmodulierte Änderung des kalziumabhängigen photoakustischen Signals aufweisen. Zusätzlich konnte die Zellpermeabilität dieser Liganden durch die effiziente Markierung lebender Zellen, die HaloTag zytosolisch exprimieren, nachgewiesen werden. Die von mir entwickelten „akustogenen" Kalziumsensoren zeigen eine bis zu 8-fache Erhöhung des photoakustischen Signals in Reaktion auf Kalzium und zeichnen sich durch hohe Photostabilität aus. Sie übertreffen bestehende Rotlichtsensoren für diese Modalität, wie beispielsweise NIR-GECO1, in gewebeähnlichen Phantomen. *Ex-vivo*-Gehirnproben mit neuronaler Expression von HaloTag, die mit meinen Liganden markiert wurden, konnten konsistent mittels photoakustischer Tomographie visualisiert werden. *In-vivo*-Markierungsexperimente in Mäusen ergaben jedoch unzuverlässige Ergebnisse, da nur gelegentlich eine spezifische Markierung von HaloTag-exprimierenden Neuronen beobachtet wurde. Dies deutet darauf hin, dass die Bioverfügbarkeit meiner Liganden verbessert werden muss.

Zusammenfassend präsentiere ich einen neuartigen Ansatz zur Entwicklung von photoakustischen Reportern, der die erste Generation chemigenetischen Sonden für diese Modalität darstellt. Meine akustogenen Farbstoffe und die ersten Generationen von photoakustischen Kalziumsensoren zeigen in vitro eine überlegene Leistung im Vergleich zu bestehenden Sensoren. Zudem sind sie in der Lage, Mäusegehirngewebe erfolgreich zu markieren, um ein lokalisiertes und starkes photoakustisches Signal zu erzeugen. Um mein langfristiges Ziel der Ganzhirn-Neurobildgebung bei Mäusen zu erreichen, habe ich mit dem Protein-Engineering begonnen, um sowohl den dynamischen Bereich als auch die Kalziumbindungsaffinität zu verbessern. Darüber hinaus habe ich Methoden zur Bewertung der Bioverfügbarkeit des Farbstoffs entwickelt, um diesen entscheidenden Parameter für zukünftige *In-vivo*-Anwendungen zu optimieren. Insgesamt strebe ich an, mit diesem Hybriddesign die Entwicklung photoakustischer Sonden und Sensoren voranzutreiben, sodass das volle Potenzial dieser leistungsstarken Modalität im Bereich des Neuroimaging und darüber hinaus ausgeschöpft werden kann.

List of abbreviations and acronyms

2P: two-photon microscopy

AAV: adeno-associated virus

ACSF: artificial cerebrospinal fluid

ADPA: anthracene-9,10-dipropionic acid

AM: acetoxymethyl

AR-PAM: acoustic resolution photoacoustic microscopy

BAPTA: 1,2-bis(2-aminophenoxy)ethane-*N,N,N′,N′*-tetraacetic acid

BBB: blood-brain barrier

BODIPY: boron dipyrromethene

BphP: bacteriophytochromes

BRET: bioluminescence resonance energy transfer

BV: biliverdin IXα

CaM: calmodulin

CHAPS: 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate

CKK: calmodulin-dependent kinase kinase

cp: circularly permutated

CT: X-ray computed tomography

DAPKP: death-associated protein kinase 1

DIEA: *N,N*-diisopropylethylamine

DMF: *N,N*-dimethylformamide

DMSO: dimethyl sulfoxide

DNA: deoxyribonucleic acid

DREADD: designer receptors exclusively activated by designer drugs

EGFP: enhanced green fluorescent protein

EGTA: ethylene glycol-bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid

EMA: European Medicines Agency

ENOSP: endothelial nitric oxide synthase

eq: equivalents

ER: endoplasmic reticulum

FAP: fluorogen activating protein

FDA: Food and Drug Administration

fMRI: functional magnetic resonance imaging

FMT: fluorescence molecular tomography

FOV: field of view

FRET: Förster/ fluorescence resonance energy transfer

fUS: functional ultrasound imaging

GECI: genetically-encoded calcium indicator

GFP: green fluorescent protein

GPCR: G-protein-coupled receptors

HEPES: 2-(4-(2-hydroxyethyl)-1-piperazinyl)-ethanesulfonic acid

hNQO1: human NAD(P)H: quinone oxidoreductase isozyme 1

HT7: HaloTag7 protein

HT9: HaloTag9 protein

HTL: HaloTag ligand

IC: internal conversion

ICG: indocyanine green

ICV: intracerebroventricular

iRFP: near-infrared fluorescent protein

ISC: intersystem crossing

IV: intravenous

JF: Janelia Fluor dye

LB: lysogeny broth

LED: light-emitting diode

M13: skeletal muscle myosin light chain kinase peptide

MAC: masked acyl cyanide

MaP: Max Planck dye

MIP: maximum intensity projection

MLCK: myosin light chain kinase

MOBHA: 2-(2′-morpholino-2′-oxoethoxy)-*N,N*-bis(hydroxycarbonylmethyl)aniline

MOM: methoxymethyl

MOPS: 3-(*N*-morpholino)propanesulfonic acid

NIR: near-infrared (650 - 900 nm)

NIR-II: second near-infrared window (1000 - 1700 nm)

OD: optical density

OR-PAM: optical resolution photoacoustic microscopy

PA: photoacoustic

PAI: photoacoustic imaging

PAM: photoacoustic microscopy

PAMPA: parallel artificial membrane permeability assay

PAT: photoacoustic tomography

PBS: phosphate-buffered saline

PEG: polyethylene glycol

PeT: photoinduced electron transfer

PET: positron emission tomography

PFA: paraformaldehyde

Pfr: far-red absorbing state of phytochromes

Pnr: near-infrared absorbing state of phytochromes

POI: protein of interest

Pr: red absorbing state of phytochromes

PTZ: pentylenetetrazole

RBN: number of rotatable bonds

RESOLFT: reversible saturable optical fluorescence transition

ROI: region of interest

S_0, S_1, S_2: ground electronic state, first singlet excited state, second singlet excited state

SDS: sodium dodecyl sulphate

SLP: self-labelling protein

S_NAr: nucleophilic aromatic substitution

SNR: signal-to-noise ratio

SPECT: single-photon emission computed tomography

T_1: triplet excited state

TBS: *tert*-butyldimethylsilyl

TFA: trifluoroacetic acid

TFE: 2,2,2-trifluoroethanol

THP: tetrahydrofuran

TICT: twisted intramolecular charge transfer

TMR: tetramethylrhodamine

TPSA: topological polar surface area

Tris: tris(hydroxymethyl)aminomethane

U2OS: human osteosarcoma-derived cell line

UST: ultrasound transducer

List of notations for photophysical properties

D_{50}: dielectric constant value which results in half of the absorbance amplitude for a dioxane: H_2O titration

K_d: dissociation constant (nM, µM, mM)

$K_{L\text{-}Z}$: open-closed equilibrium constant

k_{obs}: observed rate constant (s^{-1})

k_{off}: dissociation rate constant (s^{-1})

k_{on}: association rate constant ($M^{-1}\cdot s^{-1}$)

$\Delta A/A_0$: absorbance turn-on

$\Delta F/F_0$: fluorescence turn-on

$\Delta PA/PA_0$: photoacoustic signal turn-on

ε: extinction coefficient ($M^{-1}\cdot cm^{-1}$)

ε_{sat}: extinction coefficient for the calcium-saturated state ($M^{-1}\cdot cm^{-1}$)

λ_{em}: emission wavelength (nm)

λ_{ex}: excitation wavelength (nm)

λ_{max}: maximum absorption wavelength (nm)

λ_{PA}: photoacoustic excitation wavelength (nm)

Φ: fluorescence quantum yield

Φ_{apo}: fluorescence quantum yield for the calcium-free state

Φ_{sat}: fluorescence quantum yield for the calcium-saturated state

Table of Contents

Chapter I: Photoacoustic imaging – seeing into the depths of the brain

I.1 Neuroimaging

One of the major objectives of neuroscience is to determine how the underlying neural activity is converted into the multifaceted functions of the brain including behaviour, emotion, cognition, perception and memory. Crucial to this understanding are brain structure-function relationships, although the complex interconnectivity of this organ dictates that specific brain regions often have several diverse functions (Figure 1).[1] The olfactory bulb sits at the front of the brain, and is the initial processing centre for olfactory information from sensory neurons.[2] This region neighbours the largest part of the brain, the cerebral cortex, which is involved in many cognitive functions including memory, language and movement. In the middle of the brain lies the thalamus which integrates signals from sensory stimuli and coordinates motor responses with the cortex, whilst at the bottom, the hypothalamus is the major homeostatic regulator via the neuroendocrine system.[3] These two structures also contribute to the limbic system for processing emotions, primarily in the amygdala, and additionally the hippocampus is involved, which plays an important role in long-term memory.[4] At the back of the brain, the cerebellum is principally responsible for movement and motor coordination.[5] Below this the brain stem, containing the pons and medulla, joins to the nerves in the spinal cord, and controls essential functions such as autonomous breathing and cardiac rhythm.[6] The function of each brain region is determined by their connectivity patterns which are dependent on the activity of constituent neural circuits or ensembles.[7] These are, in turn, organised groups of neurons responsible for processing particular types of information and initiating the appropriate response. Therefore, improving our understanding of brain function is a technological challenge across scales, which requires novel correlative imaging modalities, and in parallel, the development of molecular tools to enable visualisation of the neural activity.

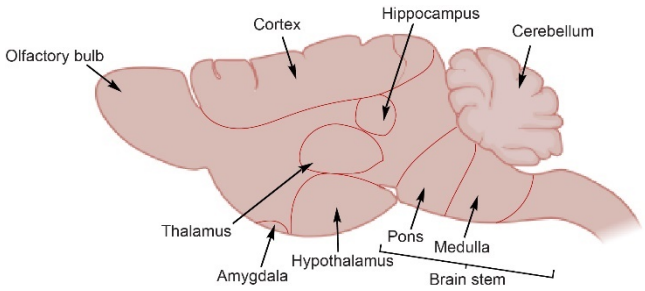


Figure 1: Simplified diagram of the mouse brain with important brain regions highlighted. Created with Biorender.

Information encoded by the brain is transmitted in many forms between neurons leading to diverse readouts of neural activity, which can be monitored using functional reporters and fluorescence imaging.[8] The rapid, transient changes in membrane potential responsible for the interneural electronic transmission can be directly reported with membrane-localised voltage indicators (Figure 2). This electrical activity induces spikes in neuronal cytosolic calcium ion concentrations of 10−100-fold via Ca^{2+} influx through voltage-gated calcium channels and ionotropic glutamate receptors, for example.[9, 10] Calcium homeostasis is quickly restored by efflux pumps like the plasma membrane calcium ATPase and the sodium-calcium exchanger. Many fluorescent calcium indicators have therefore been developed because of the importance of this "second messenger" in signalling, the discovery of calcium-binding moieties for sensor design,[11-13] and simultaneous technological advancements in fluorescence microscopy. Another chemical messenger, glutamate, is one of several neurotransmitters essential for neuronal signalling and can be sensed directly, usually with genetically-encoded indicators possessing specific neurotransmitter binding sites.[14] In addition, on a slower timescale, downstream changes in gene expression within specific neurons in brain regions associated with learning and memory can also be followed.[15] Moreover, during elevated neuronal activity, astrocytes mediate increases in cerebral blood flow, known as neurovascular coupling, via relaxation of arteriole smooth muscle and vasodilation of local blood vessels, in order to match the metabolic demand.[16] These haemodynamic changes can be monitored across the brain by macroscale imaging modalities such as functional magnetic resonance imaging (fMRI).[17]

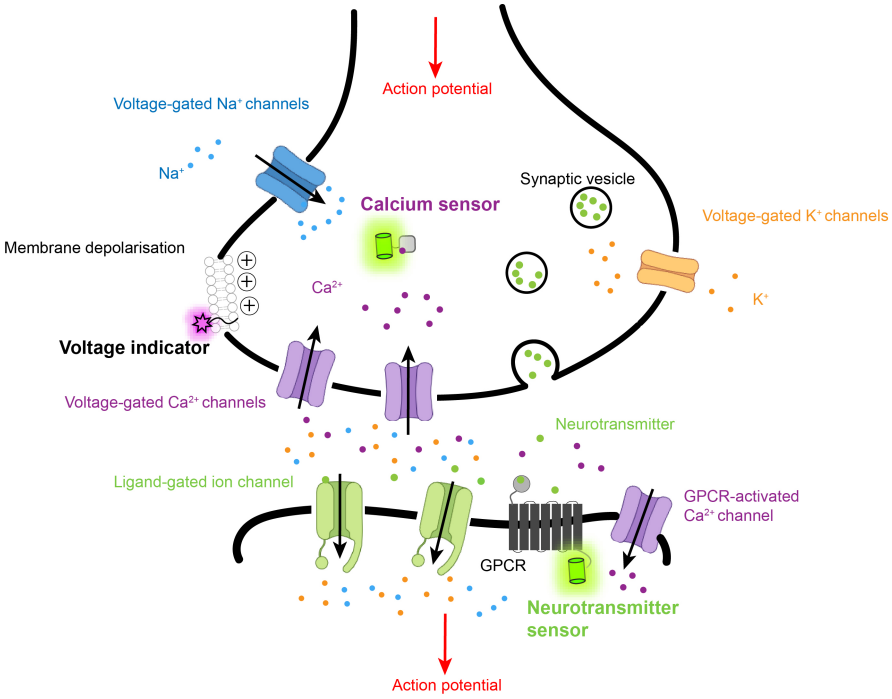


Figure 2: Simplified diagram of the molecular mechanism of action potential transmission between neurons at a synapse and the chemical changes which can be monitored using sensors. An activated neuron undergoes membrane depolarisation, largely via Na^+ influx through fast-opening voltage-gated Na^+ channels, followed by Ca^{2+} entry through voltage-gated Ca^{2+} channels. Exocytosis of excitatory neurotransmitters such as glutamate, stimulates ligand-gated ion channels either directly or via G-protein-coupled receptors (GPCR) activation, facilitating propagation of the action potential. Repolarisation to restore the resting state of a neuron is principally achieved via K^+ efflux.[18] Membrane-localised voltage sensors, cytosolic calcium sensors and neurotransmitter sensors are the three major fluorescent reporters of neuronal activity. Adapted from ref.[19]

Currently, the combination of multiphoton microscopic techniques and fluorescent calcium indicators is very powerful for functional neuroimaging at the cellular or subcellular level, whilst the whole-brain imaging capability of fMRI and positron emission tomography (PET) makes them the modality of choice for larger animals or human studies.[17, 20] However, despite the wide range of imaging tools, there is still a need for a modality which can bridge this gap between the microscale and the macroscale, correlating the neural circuit activity to entire brain regions, by melding high spatiotemporal resolution with imaging depth. Photoacoustic imaging (PAI) is an emerging, hybrid biological imaging technique which can fill this void and make a significant contribution to the neuroimaging field.

I.2 Photoacoustic Imaging

I.2.1 General principle and fundamentals of photoacoustic imaging

The principle behind PAI, the photoacoustic (PA) effect, has been known for over a century[21] but it has taken relatively recent technological advancements[22] to lead to an imaging modality applicable for biological research and suitable for clinical use. In PAI, a sample is irradiated by short laser pulses (Figure 3a), resulting in the excitation of chromophores. These excited absorbers predominantly undergo non-radiative relaxation (Figure 3b), in contrast to the competing fluorescence process, which leads to a local temperature rise and thermoelastic expansion. This energy is dissipated by the propagation of acoustic waves, which are detected by a single or multiple transducer(s) to reconstruct an image. A wide range of both endogenous and exogenous reporters can provide contrast for this modality (see section I.3).[23] The hybrid nature and the intrinsic versatility of PAI has

made it a valuable addition to the existing imaging toolbox for a diverse range of molecular imaging, preclinical and clinical applications.[24]

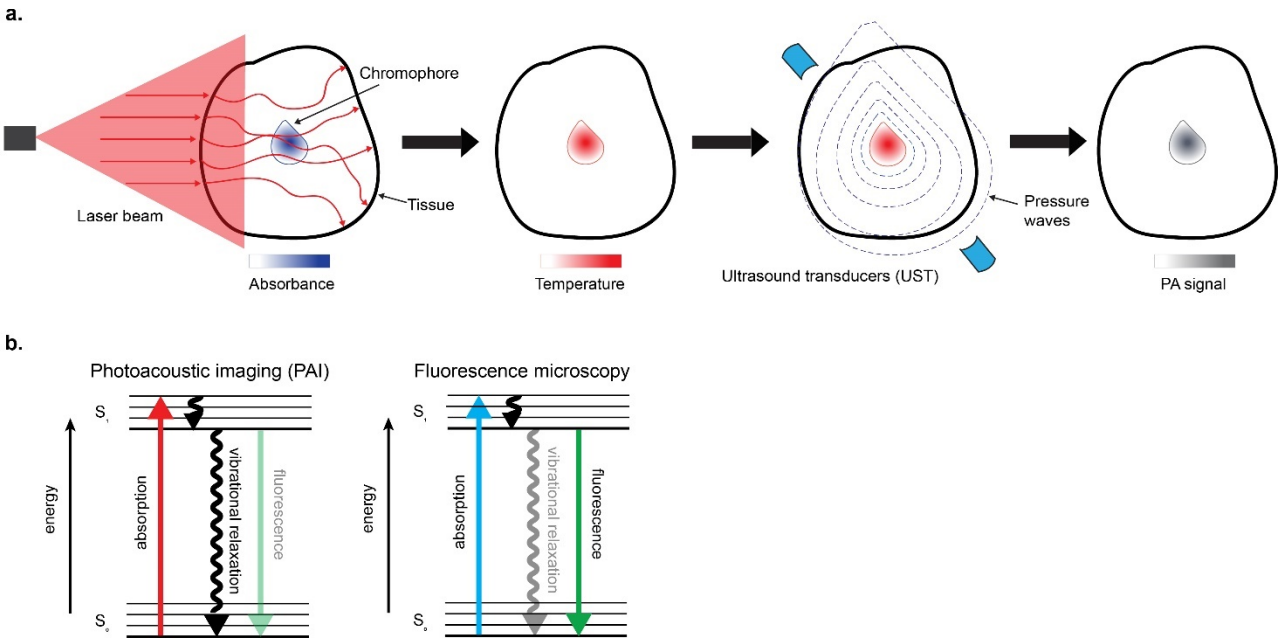


Figure 3: The fundamental principle and features of photoacoustic imaging (PAI). (**a**) Schematic of the origin of the photoacoustic effect; adapted from ref.[25] (**b**) Simplified Jablonski diagrams comparing PAI and fluorescence imaging.

I.2.2 Photoacoustic imaging techniques

There are different configurations of PAI which make this technique highly scalable, by tuning imaging speed, penetration depth and spatial resolution (Figure 4). Point-by-point scanning of a focused excitation laser beam, results in optical-resolution photoacoustic microscopy (OR-PAM), which is able to achieve single capillary level resolution through intact skulls,[26] and even image single red blood cells with resolutions below 1 µm.[27] However, this technique is also diffraction-limited with depth restricted to ~1 mm. To image deeper brain tissue up to 5 mm in depth, acoustic resolution photoacoustic microscopy (AR-PAM) can be used with less focused light and raster scanning of focused ultrasound detectors across the focal volume (20-50 µm resolution).[28, 29] For centimetre penetration depths and imaging whole brains or small animals, photoacoustic tomography (PAT) combines diffused optical excitation with tomographic scanning of single or multiple transducers (>50-100 µm resolution). Signal is acquired simultaneously from several voxels, which requires image reconstruction methods.[30] Volumetric PAT is also possible with a fixed spherical array of detectors enabling whole volume reconstruction independent of raster scanning.[31] In addition to its multidimensionality, PAI can provide high multiplexing capability, using multi-wavelength illumination and spectral unmixing through the wide range of endogenous and exogenous contrast agents (see section I.3), another advantage over non-optical techniques.[32]

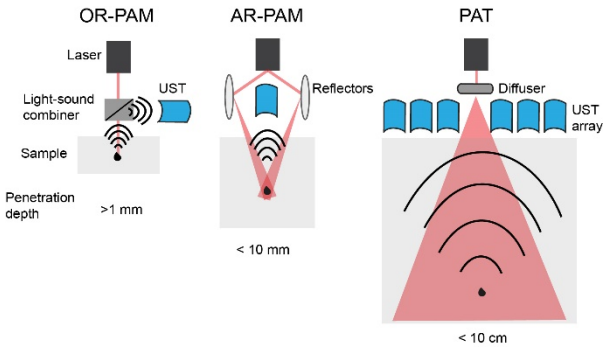


Figure 4: Configurations of PAI. Schematic of the major configurations of PAI and their penetration depths: optical-resolution photoacoustic microscopy (OR-PAM), acoustic-resolution photoacoustic microscopy (AR-PAM) and photoacoustic tomography (PAT). Adapted from ref.[33]

I.2.3 Comparison with other imaging modalities

PAI combines the high contrast of optical imaging with the high spatial resolution and penetration depth of ultrasound, thanks to 1000-fold weaker scattering in tissue than light,[34] and thus represents an ideal mesoscopic bridging modality, offering a compromise between these key parameters, which cannot be matched by other modalities (Figure 5).[22] Purely optical methods, for example confocal and multiphoton microscopy, enable monitoring of neural activity with high spatiotemporal resolution in individual neurons and their subcellular compartments (Figure 5a), facilitated by a plethora of functional, fluorescent sensors, particularly genetically-encoded calcium indicators (GECIs). However, the resolution rapidly deteriorates beyond 1 mm depth due to light scattering, so fluorescence-based modalities are limited to imaging superficial regions such as the cortex and hippocampus.[35] Moreover, for mammalian models, a cranial window, skull thinning or optical clearing techniques are required and the small accessible field-of-view (FOV) makes it impossible to visualise the full connectivity of neural circuits in one image (Figure 5b). Whilst fluorescence molecular tomography (FMT) using near-infrared (NIR) fluorophores can image larger volumes, the challenge of light diffusion cannot be avoided.[36] As a result, non-optical macroscale modalities such as fMRI, PET, single-photon emission computed tomography (SPECT) and X-ray computed tomography (CT) are optimal for imaging large volumes such as whole brains (or entire model organisms) because of their deep tissue imaging capability, large FOV and fast volumetric imaging rates (Figure 5a, b). However, these features come with a severe trade-off in both spatial and temporal resolution, and use ionising radiation or strong magnetic fields. Functional ultrasound (fUS) imaging has developed rapidly in recent years bridging this spatiotemporal resolution gap in the mesoscale. However, the modality is still largely restricted to imaging vasculature, because blood flow is the sole contrast mechanism, which can be enhanced by the use of microbubbles, and the development of functional ultrasound reporters is just in its infancy.[37, 38]

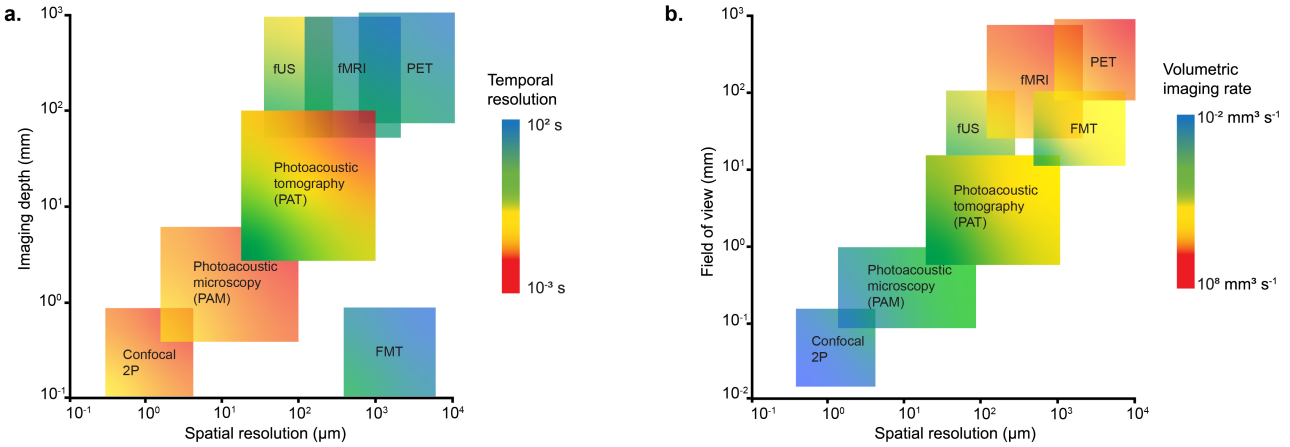


Figure 5: Comparison of PAI with other imaging modalities. (**a**) Summary diagram of the imaging depth, spatial and temporal resolution of the major biological imaging modalities: confocal and two-photon (2P) microscopy, PAM, PAT, functional ultrasound (fUS), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), fluorescence molecular tomography (FMT). (**b**) Summary diagram of the field of view, spatial resolution and volumetric imaging rate of the same modalities from (**a**). Adapted from ref.[22, 39, 40]

PAI has already demonstrated its potential to image across the rodent brain, detecting dynamic neural activity from haemodynamic and exogenous contrast sources, primarily calcium indicators, with higher spatial and temporal resolution than competing modalities including fMRI, fUS and PET. In the following section, I will discuss the existing contrast agents for PA and their application to neuroimaging, including their achievements and limitations which provide the motivation behind this project.

I.3 Contrast agents for photoacoustic imaging

I.3.1 Endogenous contrast agents

Given the origin of the photoacoustic effect, any molecule with a high absorption coefficient can provide a source of PA contrast. This is ideal for label-free PAI, enabling visualisation of structure and/or function without delivery of exogenous agents. The strong PA contrast of haemoglobin and its high concentration in the blood, enables imaging of vascular morphology including microvasculature (Figure 6a).[27, 41] *Ex vivo* imaging of perfused brains allows imaging of other chromophores. Major brain structures can be visualised via the cytochrome-lipid distribution[42] and at a more local level dopaminergic neurons, rich in neuromelanin, and myelinated bundles are particularly visible.[43] Histological imaging of brain slices with PAM has also achieved a spatial resolution down to 250 nm.[44]

Aside from anatomical and structural information, haemoglobin is the only intrinsic chromophore which can indirectly report on neural activity via neurovascular coupling. The unique oxygen-dependent spectral signatures of haemoglobin, arising from coordination of Fe^{2+}/Fe^{3+} to the protoporphyrin IX chromophore of the haem cofactor (Figure 6a, b), and the multispectral functionality of PA enables quantitative imaging of haemodynamics in small animals' brains *in vivo*. Key parameters including oxygenated, deoxygenated and total haemoglobin concentrations, cerebral blood flow, and hence oxygen metabolism can be determined.[45] For example, using OR-PAM after scalp removal, Yao *et al*. investigated cortical haemodynamic responses in mice, at capillary-level resolution, to electrical stimulation of the hindlimbs.[41] PAT has facilitated imaging of whole mouse and rat brains through intact skulls to elucidate functional connectivity via haemodynamic changes in different brain regions in response to physical or chemical stimulation.[46-48] Recently, simultaneous recording of haemodynamic and calcium responses, via PAT and the fluorescence of GCaMP6f (see section I.3.2.3) respectively, to electrical paw stimulation investigated the correlation between neural activity and different haemodynamic parameters.[49] There was a significant trend between the calcium dynamics and rises in

oxygenated and total haemoglobin levels, and oxygen saturation, with a corresponding decrease in deoxyhaemoglobin. However, haemodynamic changes are very slow compared to the original neural activity and so exogenous contrast agents are required to report on more closely coupled dynamics such as calcium fluxes. The advantage of endogenous contrast now becomes a hindrance for imaging, as the PA responses from sensors must be detectable over the background haemoglobin signals, imposing stringent requirements for performant PA contrast agents.

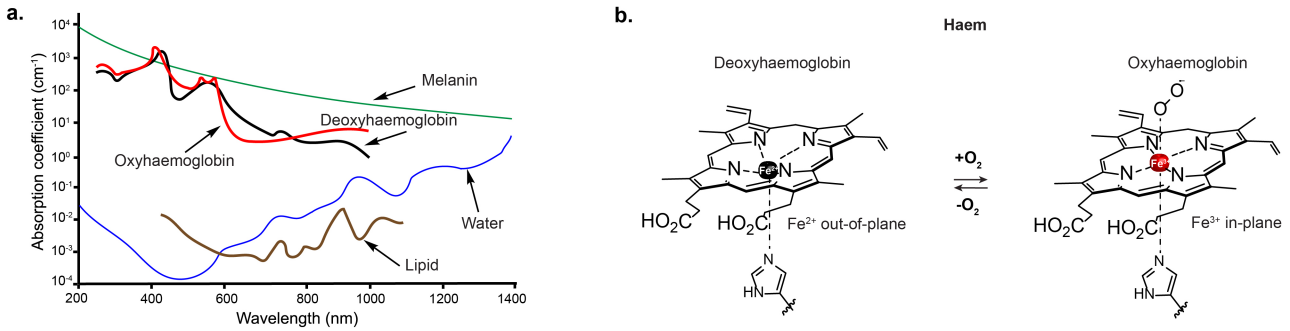


Figure 6: Endogenous photoacoustic contrast. (**a**) Absorption spectra of the predominant sources of endogenous contrast in the mammalian brain; adapted from ref.[30] (**b**) Structure of the haem cofactor responsible for the oxygen-dependent absorption properties of haemoglobin; oxygen-binding pulls the $Fe^{2+/3+}$ ion into the plane of the porphyrin ring resulting in a change in structure and electronic configuration.

I.3.2 Exogenous contrast agents

I.3.2.1 Key properties of a PA contrast agent

There are a number of essential properties for extrinsic probes to be suitable for photoacoustic neuroimaging.[23] First, these reporters must absorb strongly in the far-red/near-infrared (NIR) region of the spectrum to minimise background absorption from haemoglobin in particular (Figure 6a). Second, they need to possess a low quantum yield to maximise photoacoustic generation efficiency (Figure 3b), and show high photostability to maintain PA signal during the imaging window. Furthermore, the chromophores must be compatible with cellular labelling strategies, targetable and be amenable to the design of biosensors, which can report on changes in concentration of cations such as calcium, or neurotransmitters, physiologically relevant to neuronal activity. Finally, there are biochemical considerations which must be met for *in vivo* applications, including cell permeability, non-toxicity and pharmacokinetics. Here, I will introduce the existing PA probes and sensors developed for neuroimaging, grouped into the two major classes of synthetic and genetically-encoded reporters.

I.3.2.2 Small-molecule dyes

Structurally diverse synthetic dyes have been exploited as contrast agents for PAI due to their high NIR extinction coefficients, photostability and ease of functionalisation.[50] I will focus on three major families: cyanines, BODIPYs and xanthenes as PA labels because they have been used for sensor development and neuroimaging, but many others such as squaraines and porphyrins have also been shown to provide PA contrast.

Cyanines

Cyanines have been the most extensively studied dye family for PAI, in particular the FDA-approved indocyanine green (**ICG**; Figure 7a), because of their strong absorption at NIR wavelengths. This began with the development of a theranostic agent for early-stage cancer with **ICG** encapsulated in silicate-based nanoparticles and targeted via human epidermal growth factor receptor 2 antibodies conjugated on the surface.[51]

Although mostly applied to cancer biology rather than neuroimaging, **ICG** enabled the first photoacoustic angiography of the rat cortex *in vivo*.[52] Incorporation of a hydrophilic polyethylene glycol (PEG) group into the structure was necessary for systemic delivery and improving bioavailability. Aside from low water solubility and poor pharmacokinetics, the other major limitations are the photostability of these scaffolds,[53] and the lack of labelling specificity, usually relying on the enhanced permeability and retention effect of tumour cells, or active targeting to receptors.[54, 55] Nevertheless, there have been recent attempts to make activatable cyanine probes. Fluorescent hemicyanine dyes can be repurposed for PAI through replacement of the endocyclic oxygen atom with sulphur and *ortho*-chlorination to maintain the deprotonated form at physiological pH.[56, 57] This provided a desired bathochromic shift into the NIR, greater absorption and reduced quantum yield enabling generation of novel ratiometric PA probes by functionalisation of the phenol. These included a H_2O_2-sensitive probe (Figure 7b), which was injected retro-orbitally into an Alzheimer's disease mouse model, able to cross the blood-brain barrier (BBB) and resulted in elevated PA signal compared to wild-type mice, due to the oxidative stress associated with such neurological disorders.[58]

For measuring neuronal activity, two voltage-sensitive cyanine dyes have been used in PAI. In polarised cells these positively charged dyes form nonfluorescent aggregates, which separate into fluorescent monomers upon cell depolarisation, decreasing the PA signal.[59] The NIR-absorbing **PAVSD800-2** (Figure 7c) shows up to 24% changes in PA intensity in lipid vesicles. With a very similar structure, **IR780** perchlorate showed a response at submillimetre spatial resolution in the motor cortex of living rats upon chemically stimulated neural seizure, which was validated by electrical recordings.[60] The only other published example of a PA voltage indicator uses dipicrylamine, which has a strongly blue-shifted absorption compared to cyanines so required imaging through a cranial window at λ_{PA} = 500 nm.[61] The sensitivity of these first probes is very low, which has also been a challenge for fluorescent voltage sensors, especially with the limited achievable sensor concentration in a cell membrane. Moreover, this sensing mechanism of dye distribution occurs in the order of seconds, so is even slower than calcium dynamics.

Photoacoustic sensing of calcium ions has been led by cyanines, and the approach follows the engineering principles of fluorescent sensors. These were pioneered by Roger Tsien, by attaching a selective calcium-chelating moiety, usually 1,2-bis(2-aminophenoxy)ethane-*N,N,N′,N′*-tetraacetic acid (BAPTA), to a chromophore, with suppression of photoinduced electron transfer (PeT) transfer upon calcium binding changing the absorption and/or fluorescence spectral features.[13] The first synthetic PA calcium indicator, "**L**" (Figure 7d), uses a calcium-chelating group attached to the heptamethine cyanine **IR780**, resulting in a strongly absorbing NIR probe (ε = 194000 $M^{-1}\cdot cm^{-1}$ at λ_{max} = 765 nm).[62] Calcium binding results in a 2-fold decrease in absorbance and fluorescence, which translates to a similar decrease in PA signal, visualised in agar phantoms. To resolve the cell impermeability of this sensor, the same group synthesised a similar probe, **CaSPA_550-AM** (Figure 7e), which consists of a hemicyanine chromophore coupled to a BAPTA group with the carboxylates protected as acetoxymethyl (AM) esters.[63] On cell entry, endogenous intracellular esterases remove the protecting group, leaving the active probe. **CaSPA_550** did not show improved sensitivity, also showing a 2-fold reduction in PA signal upon calcium binding, and is blue-shifted (λ_{max} = 550 nm), but labels mammalian cells at low micromolar concentrations. Intracellular calcium dynamics could be recorded in heart organoids, showing up to 10% signal changes from spontaneous contractions, which correlated with the responses from the fluorescent calcium indicator **Fluo-4-AM**. Injection into larval zebrafish brains facilitated *in vivo* detection of calcium fluctuations, where PA signal decreased upon addition of neurostimulants. Nevertheless, these synthetic dyes are still far from optimal as their acoustogenic response is relatively small and negative-going (i.e. lower signal in the calcium-bound state), their large molecular weight makes *in vivo* delivery difficult, and their lack of labelling specificity prevents targeted neuroimaging applications.

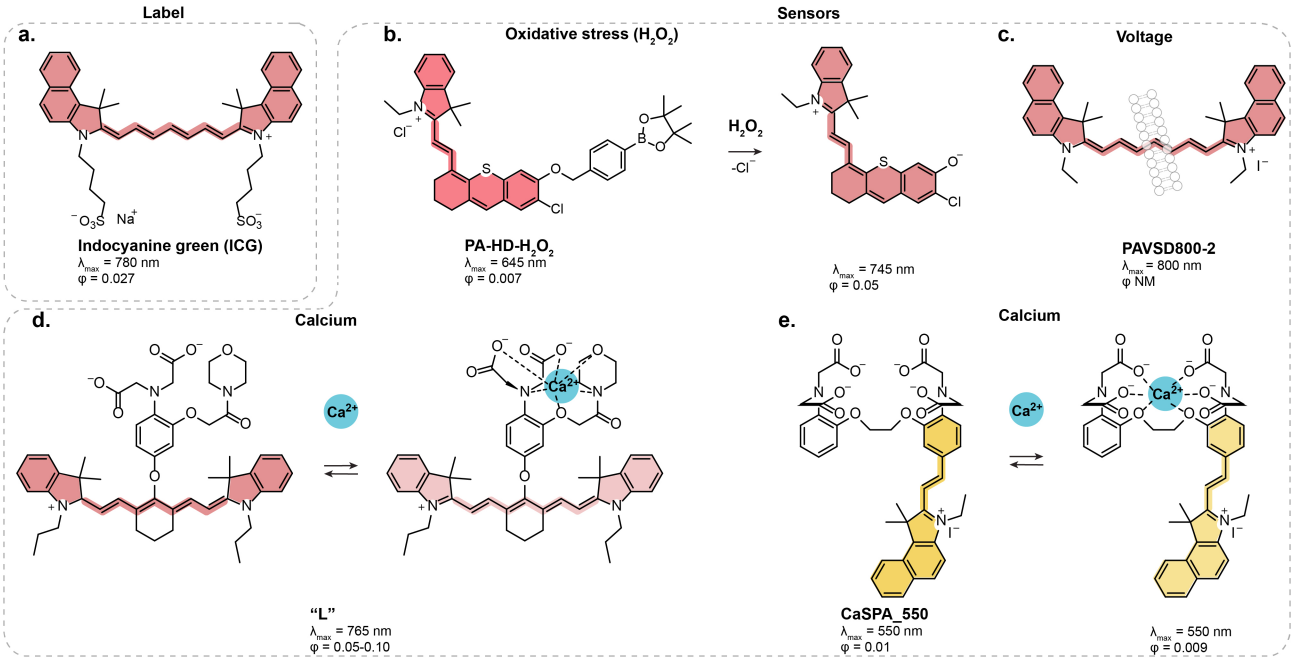


Figure 7: Cyanine-based photoacoustic labels and sensors. Label: (**a**) Indocyanine green (**ICG**).[64] Sensors: (**b**) Hemicyanine hydrogen peroxide sensor, **PA-HD-H$_2$O$_2$**.[56] (**c**) Voltage indicator **PAVSD800-2**.[59] (**d**) Calcium sensor, "**L**".[62] (**e**) Calcium sensor **CaSPA_550.**[63]

BODIPYs

BODIPY dyes have also been regularly used in small-molecule acoustogenic probes because of their easy functionalisation, tuneable absorption into the NIR and biocompatibility. Motivated by detection of copper (II) in Alzheimer's disease models, activatable Cu^{2+} probes, **APC-1** and **APC-2**, were developed using an aza-BODIPY chromophore with 2-picolinic ester moieties, which are selectively cleaved by Cu^{2+}-mediated hydrolysis (Figure 8a).[65] These two probes show ratiometric turn-ons up to 101-fold, but were too hydrophobic for *in vivo* use. PA brain imaging with BODIPY probes has been scarce, but recently a series of donor-acceptor *meso*-trifluoromethyl BODIPY probes were developed, which can cross the blood-brain barrier.[66-68] Functionalisation with a Cu^{2+}-selective chelator, N^1,N^2-*bis*(2-(pyridine-2-yl)ethyl)benzene-1,2-diamine, made a reversible Cu^{2+} probe which has a hypsochromic shift upon binding of the metal ion and an additional free radical anion NIR absorption peak (Figure 8b).[66] *In vivo* imaging of Parkinson's disease mouse models, revealed elevated Cu^{2+} levels in the corpus callosum, hippocampus and thalamus as the disease progresses, and identified expression of DMT1 transporters as the primary cause. The ratiometric property was advantageous to resolve specific signal from the background contrast of haemoglobin.[69] Modification of the responsive group enabled adaptation to other stimuli. Using catechol, oxidizable by superoxide to benzoquinone with a red-shift in absorption, which can in turn be reduced by glutathione, provided a reporter of oxidative stress.[67] Similarly, a NO-responsive derivative was also designed, and all three types of probe used to screen the therapeutic effect of natural polyphenols.[68] These probes are useful for imaging brain pathophysiology, but there is to date no BODIPY-based PA probe able to report on neuronal activity.

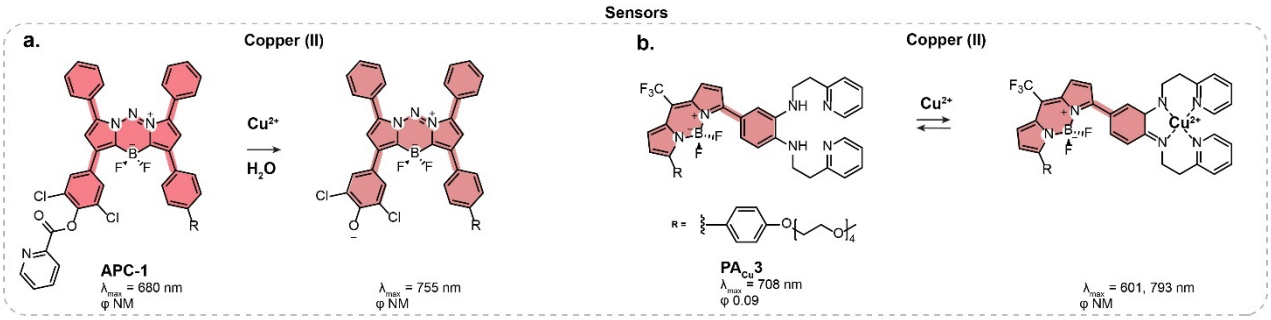


Figure 8: BODIPY-based photoacoustic sensors. (**a**) Activatable copper (II) sensor, **APC-1**.[65] (**b**) BBB permeable copper (II) sensor, **PA$_{Cu}$3**.[66]

Xanthenes

By contrast, the xanthene family of dyes are responsible for one of the most widely used fluorescent synthetic calcium indicators for reporting neuronal activity, **Oregon Green BAPTA**.[70, 71] Xanthenes usually have high quantum yields, but quenching mechanisms can be exploited to generate sufficient PA signal.[72] For example, removal of the bridging O-atom from fluorescein gives phenolphthalein which has >2-fold greater PA signal due to its rotational flexibility.[73] Although classic xanthenes like fluorescein and rhodamine absorb in the visible wavelength range, NIR wavelengths can be attained. Substitution of the bridging xanthene O atom for electron-deficient groups such as ketone results in absorbance >850 nm.[74] Cyclisation with julolidine auxochromes reduced potential nucleophilic attack on the bridging ketone in aqueous solution, resulting in 10-fold greater PA signal for **KeJuR** than **ICG** in a defibrinated sheep's blood tissue phantom (Figure 9a).[75] A donor-acceptor-donor design, **XanthCR880**, with the xanthene acceptor core conjugated to thienylpiperidine, to strengthen the push-pull system, enabled PA signal detection 4 cm deep through swine tissue (Figure 9b).[76, 77] The ease of synthetic functionalisation of xanthenes also makes them amenable to sensor design, however these PA probes have been applied to cancer biology or other diseases rather than neuroimaging. For *in vivo* PA measurements of tumour pH, NIR dyes SNARF-5F and benzo[a]phenoxazine were repurposed into nanoprobes.[78-80] Modification of the thienylpiperidine probe including incorporation of *o*-phenylenediamine and encapsulation with a reference dye, generated a ratiometric nitric oxide sensor, **rNP-NO**.[81] To extend conjugation and hence red-shift absorption, rhodamines can be combined with hemicyanines in the same scaffold. **NRh** displays Cu^{2+}-sensitive absorption and could detect Cu^{2+} in mice *in vivo* after subcutaneous administration in nanomicelles.[82] With a similar structure, a dual modal probe for hypoxia was developed, with a nitroreductase-sensitive moiety attached to the rhodamine amino group, enabling imaging of drug-induced liver hypoxia in a mouse model.[83]

Furthermore, there is an intrinsic feature of rhodamines, the environmentally-sensitive open-closed equilibrium, which will be discussed in more detail in section I.3.3, that can be exploited to make activatable probes and sensors. A silicon rhodamine dark quencher, via a combination of photoinduced electron transfer (PeT) and twisted intramolecular charge transfer (TICT) mechanisms, **PA-MMSiNQ**, possesses a mercaptomethyl group at the 2-position on the pendant phenyl ring (Figure 9c).[84, 85] The probe rests predominantly in the spirocyclic, closed, non-absorbent form until selective oxidation by HOCl locks the dye in the open, PA active form. Proof-of-principle was demonstrated in the mouse subcutis with elevated PA signal upon HOCl injection. An analogous cell permeable phosphinate-ester based rhodamine probe for HOCl, **SNR$_{700}$-HOCl**, has also subsequently been developed.[86] Another strategy to red-shift rhodamines is to introduce asymmetry and extend the π system. To report on the tumour biomarker human NAD(P)H: quinone oxidoreductase isozyme 1 (hNQO1), a rhodol with a masked phenolic hydroxyl group was designed, **Rhodol-PA** (Figure 9d).[87] hNQO1-catalysed elimination of the masking group shifts the equilibrium strongly towards the open form providing a turn-on in PA and NIR fluorescence signal. The high contrast mechanism of this probe was demonstrated in mice after intratumoural and tail vein injections, showing an average 8- and 7-fold enhancement in PA signal for tumours overexpressing hNQO1. Recently, a similar extended rhodol, **GX-5-CO**, but with NIR-II absorption, enabled detection of carbon monoxide in the vasculature of hypertensive mice.[88] Nevertheless, as

for the other synthetic dye families, labelling specificity remains lacking, and these elegant designs for xanthene-based PA sensors are yet to be employed to other, neuronal activity-related stimuli such as calcium.

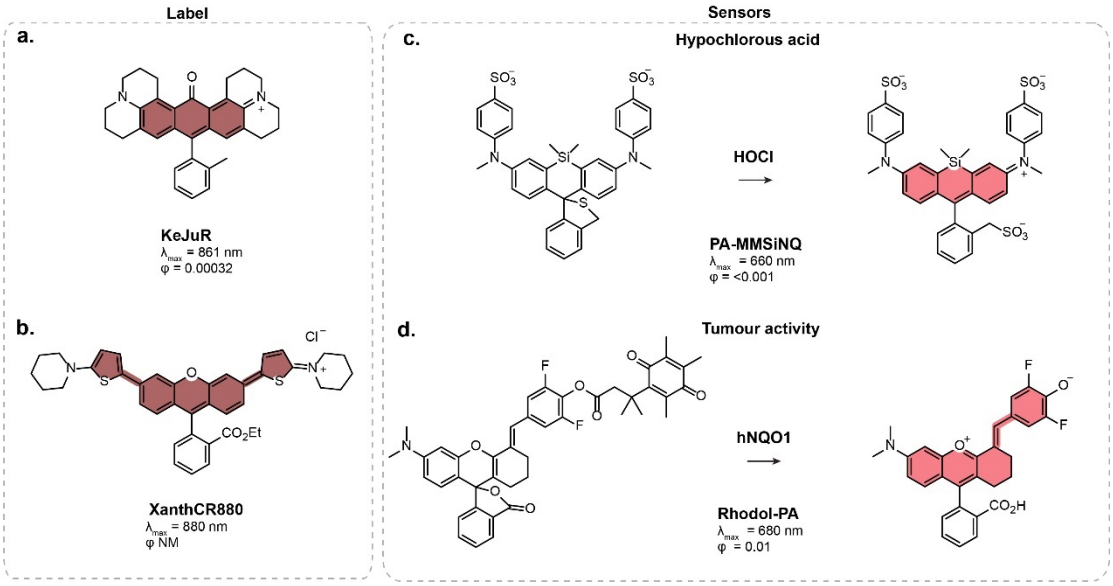


Figure 9: Xanthene photoacoustic labels and sensors. Labels: (**a**) Ketone rhodamine **KeJuR**,[75] (**b**) NIR absorbing **XanthCR880**.[76] Sensors: (**c**) hypochlorous acid activatable sensor, **PA-MMSiNQ**.[84] (**d**) hNQO1 activatable probe, **Rhodol-PA**.[87]

Miscellaneous dyes

Finally, a few more dyes from other families are worth mentioning because of their PA properties. **Methylene blue** (Figure 10a) is a common reference PA dye, which is also FDA-approved and has been involved in numerous studies, for example, in the identification of sentinel lymph nodes in rats[89] and the development of nanoparticles for photothermal therapy.[90] The strong PA signal arises from significant intersystem crossing, due to its bridging sulphur heteroatom, and subsequent nonradiative relaxation from the triplet state. In addition, the only other synthetic calcium indicators used for PAI, are the heteroatom-containing, chromogenic **arsenazo III** and **chlorophosphonazo III** (Figure 10b). **Arsenazo III** shows a nonlinear increase in absorbance and PA signal with calcium concentration in tissue phantoms, and Arsenazo-treated cells had a higher PA signal than background and showed fluctuations over 30 s of PAM imaging.[91] However, there are several limitations including cell impermeability, concentration-dependent signal and its well-known toxicity from nonspecific protein binding and ROS generation.[92, 93] **Chlorophosphonazo III** is a cell-permeable analogue with lower toxicity which could be used to monitor, via OR-PAM, the increase in intracellular calcium in response to thapsigargin treatment in cell culture.[94] Unfortunately, the sensitivity remains very low with only a 16-20% increase in signal in response to 25 µM of calcium.

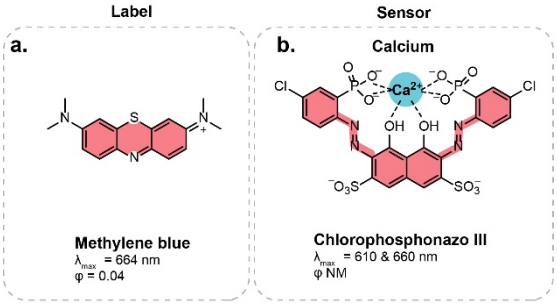


Figure 10: Additional small-molecule photoacoustic labels and sensors. Label: (**a**) PA reference dye, methylene blue.[95] Sensor: (**b**) calcium indicator, **chlorophosphonazo III**.[96]

I.3.2.3 Genetically-encoded reporters

The second major class of reporter for PAI consists of genetically-encoded chromophores. They have the intrinsic advantage of labelling specificity, as controlled expression from the promoter of the gene of interest enables targeting to specific cell populations or even organelles.[40] These proteins can also be widely expressed across whole tissues and organs, including the brain, and their stable long-term expression facilitates longitudinal imaging. The majority of protein-based PA contrast agents, including enzymatic reporters and fluorescent proteins, have been repurposed from other imaging modalities and applications, primarily fluorescence microscopy. Notwithstanding, strongly absorbent proteins have been engineered specifically to function as PA reporters, such as non-fluorescent chromoproteins and bacterial phytochromes.

Enzymatic reporters

Enzymatically synthesised pigments were the first genetically-encoded reporters for PAI, providing strong photobleaching-resistant signal. Only low expression levels are required as each enzyme catalyses many reactions resulting in accumulation of an absorbing molecular product to high concentration, and hence strong signal. β-galactosidase, encoded by lacZ, cleaves the colourless substrate X-gal into galactose and a side product which is oxidised into 5,5'-dibromo-4,4'-dichloroindigo, a strong absorber in the 600-700 nm range. This was used in gliosarcoma detection but delivery of X-gal is required and can cause skin irritation.[97, 98] Avoiding potential administration issues, the fully genetic label tyrosinase catalyses the rate-limiting step in the biosynthetic pathway of melanin pigments (Figure 11) so has also been exploited to generate strong, broad spectrum PA contrast for imaging tumour xenografts in mice for example.[99, 100, 101] However, high levels of melanin precursors can be cytotoxic if not sequestered by melanosomes, specialised organelles present in pigment cells.[102] This can be avoided through inducible gene expression or expression of auto-assembling nano-compartments which can sequester melanin.[100, 103] Despite these limitations, these enzymes can act as transcriptional reporters of neuronal activity through control with promoters of immediate early genes such as c-Fos, but their visualisation has never been attempted with PAI.[104] Although unsuitable for imaging of neural activity in terms of its dynamics and resolution, a melanin-based biosensor has been developed, involving relocalisation of melanosomes in response to GPCR signalling.[105] *Xenopus lavis* cells were transplanted into the zebrafish brain, and aggregation of melanosomes could be detected with PA on stimulation by a hormone, melatonin, or by a GPCR agonist when a DREADD was expressed.

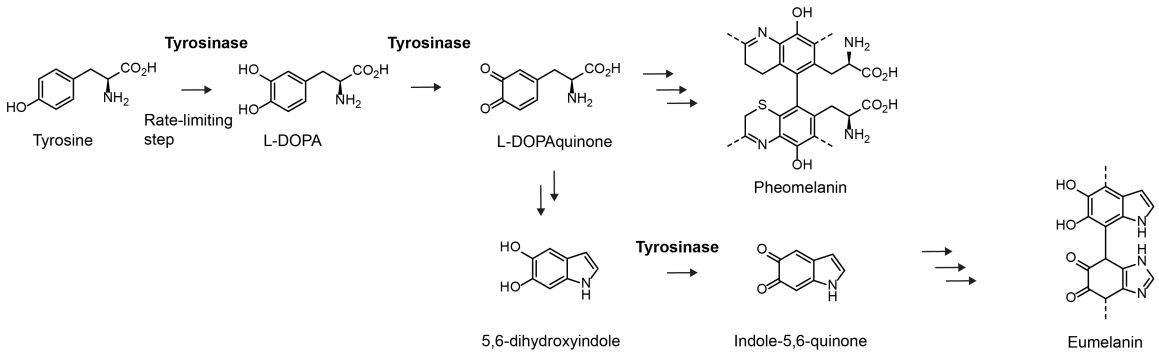


Figure 11: Enzymatic reporters for PAI. Biosynthetic pathway for the synthesis of melanin pigments.

GFP-derived chromoproteins and fluorescent proteins

As for synthetic dyes and enzymatically-produced pigments, ideal protein-based PA contrast agents possess high extinction coefficients and minimal fluorescence. The green fluorescent protein (GFP) family also contains non-fluorescent members called chromoproteins ($\Phi < 0.001$) which are an effective mode of genetically-encoded PA contrast. Laufer *et al.* characterised a series of fluorescent proteins and three chromoproteins (cjBlue[106] (Figure 12a), aeCP597[107] and E2 Crimson NF) with photoacoustic spectroscopy.[108] Normalising the

photoacoustic spectra by the maximum extinction coefficient provides a measure of photoacoustic generation efficiency, which were more than twice greater for the chromoproteins, likely through a combination of reduced radiative relaxation and ground state depopulation due to the faster relaxation times. These nonfluorescent mutants also showed greater photostability, losing 10% of signal compared to 30-50% for the fluorescent proteins under the same conditions, with the shorter relaxation times minimising oxidative bleaching. Directed evolution of chromoproteins Ultramarine[109] and cjBlue was performed to improve the PA signal via a combination of absorption and PA colony screening, which resulted in up to 4.3-fold enhancement in PA signal.[110] As a proof-of-principle, one of the new variants, tdUltramarine2, maintained higher PA signal *in vivo* than its parent protein, when expressed in *E. coli* cells injected into the ear of a rat. Fusing Ultramarine variants with EGFP led to a prototype Förster/ fluorescence resonance energy transfer (FRET) biosensor for protease activity (Figure 12b), with a decrease in PA signal from the acceptor after incubation with trypsin. However, chromoprotein-based calcium indicators are yet to be developed, as these proteins still need to be optimised for expression in mammalian cells, and so fluorescent proteins have been predominantly used.

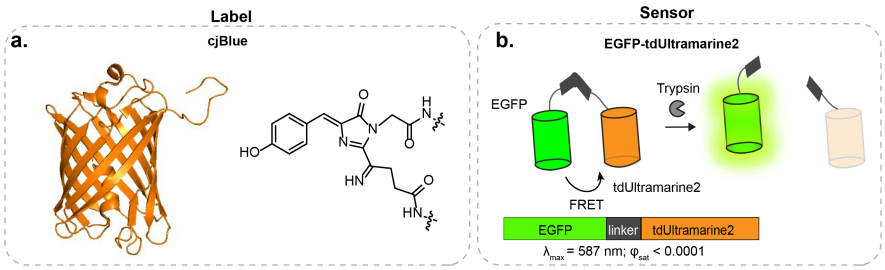


Figure 12: Chromoproteins as photoacoustic reporters. (**a**) Crystal structure and chromophore of cjBlue (PDB: 2IB5).[106] (**b**) Structure of the protease sensor EGFP-tdUltramarine2.[110]

Although counterintuitive due to their high quantum yields, fluorescent proteins can be repurposed into PA contrast agents. This was first demonstrated by expression of EGFP or mCherry in *Drosophila* and zebrafish, with the transgenic tissues, including the brain, visible at several millimetres depth.[32] In fluorescence microscopy, genetically-encoded calcium indicators (GECIs) are the most commonly used tools to monitor neural activity because of their high sensitivity, brightness and fast kinetics, thanks to a huge protein engineering effort over the past 20 years, with the GCaMP series setting the standard.[111-114] These sensors consist of enhanced GFP (EGFP) fused to calmodulin (the calcium-binding domain) and a calmodulin-binding peptide (Figure 13a, b). The intensiometric calcium response stems primarily from a significant absorption change, which enables photoacoustic detection, despite the large fluorescence increase (Figure 13c). Calcium binding to calmodulin instigates a conformational change involving association with the calcium-binding peptide, which alters the environment surrounding the EGFP chromophore. This chromophore exists in an equilibrium between protonated fluorescent and non-protonated fluorescent states, due to the tyrosine's phenolate exposure to the solvent through openings in the β-barrel protein structure (Figure 13d). In the calcium-bound form of the sensor, the deprotonated, brighter fluorescent state is stabilised via the altered hydrogen bond network, thus increasing absorption and fluorescence.[115] These sensors have enabled the first functional photoacoustic calcium imaging in zebrafish and mice. Using GCaMP5G, calcium activity in immobilised and freely swimming zebrafish larvae was visualised.[116] In response to chemical neurostimulation with pentylenetetrazole (PTZ), an increase in PA signal was observed in the spinal cord with a maximum turn-on of 1.8 over 0.1-0.3 s just before movement. GCaMP5G could also resolve calcium dynamics in 0.5 mm^3 voxels after PTZ-stimulation in an isolated adult zebrafish brain, which could not be followed by fluorescence below superficial brain regions due to scattering. The haemoglobin-free brains of *Drosophila* expressing GCaMP5G in cholinergic neurons were also imaged *in vivo* via PAM, with and without the melanin-maintaining cuticle.[117] PA signals recorded in the antennal lobe in response to odour stimulation showed ~14% and ~23% increases respectively and depth-dependent kinetics.

Scaling up to mice, neural activity from GCaMP6f-expressing brain slices stimulated by potassium perfusion could be detected by PAT with a PA signal increase up to 15%, even after covering with a 1.5 mm layer of brain tissue.[117] Gottschalk et *al.* first imaged an *ex vivo* bloodless brain expressing GCaMP6f with a hybrid planar

fluorescence and PAT set-up for volumetric imaging across the whole mouse brain.[31] Injecting PTZ into the frontal cortex stimulated large increases in PA signal up to 150%, which initiated around the injection site before propagating into the other hemisphere, and these waves of calcium activity were validated by the correlated fluorescence response. Simultaneous injection of tetrodotoxin, a sodium channel blocker, inhibited the response. They were also able to record calcium dynamics *in vivo* across the brains of GCaMP6f expressing mice. Pulses of hind paw electrical stimulation induced responses up to 4% $\Delta PA/PA_0$, which were temporally and spatially correlated with the fluorescence signals, but resolvable deeper in the brain, highlighting the depth advantage of this modality. Equivalently treated GCaMP6s-expressing mice, showed similar peak times but slower decay of the PA signals, as observed in the fluorescence responses. PAM was also used to image neural activity *in vivo*, focused on an area with minimal vasculature, and local PA responses were detected in response to the stimulation but could not be concretely distinguished from the slower haemodynamic changes.[117]

During *in vivo* imaging, the strong PA signal from endogenous chromophores, particularly haemoglobin, is a principal issue. One method to improve the optical contrast is to use differential imaging with photoswitchable proteins which modulate their signals at different wavelengths. A reversibly switchable derivative of GCaMP5G was engineered, rsGCaMP1.1, displaying photochromism and a low quantum yield ($\Phi = 0.016$), whilst an affinity variant, rsGCaMP1.4-ER, can record calcium dynamics in the endoplasmic reticulum (ER).[118] The *cis-trans* isomerization responsible for the photoswitching between 'on' and 'off' states is dependent on calcium-binding, and induced by illumination at 488 and 405 nm (Figure 13e), so enables application in super-resolution fluorescence microscopy as well as PAI. Relative calcium concentrations could be determined from fitting of the photoswitching kinetics from tubes containing purified rsGCaMP with different calcium concentrations. With the tubes subcutaneously implanted in mice the two lowest concentrations (< 65 nM) were not resolvable due to the poor tissue penetration of blue light. Implanted HeLa cells expressing rsGCaMP1.1 in the back of a mouse treated with ionomycin/Ca^{2+} could be just about distinguished from resting cells. Whilst being a promising design for novel reversibly switchable GECIs and presenting a low quantum yield, these sensors cannot resolve the other problem of GFP-derived indicators: the poor light penetration of green excitation wavelengths. PA imaging of GCaMP6f-expressing mice revealed that, at the peak absorption wavelength of GCaMP6f (488 nm), contrast can be provided only down to ~1.5 mm, whereas at 650 nm, the whole brain, ~7 mm deep is visible, which highlights the need for far-red/NIR calcium indicators.[31]

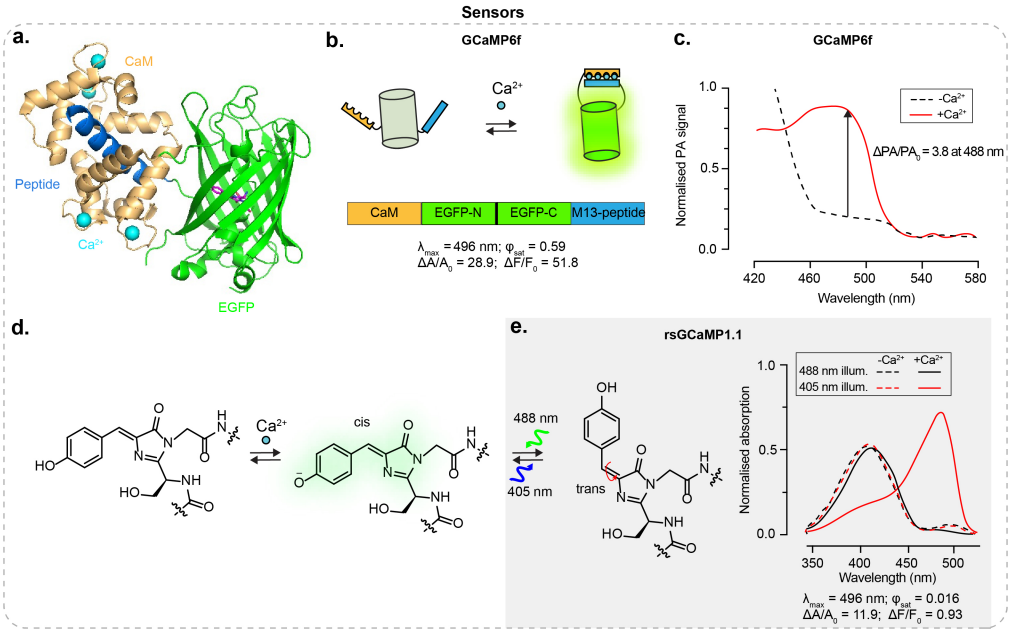


Figure 13: GFP-based genetically-encoded calcium indicators. (**a**) GCaMP6m crystal structure (PDB: 3WLD);[119] CaM: calmodulin. (**b**) Modular structure of GCaMP proteins and photophysical properties of GCaMP6f.[112, 119] (**c**) Photoacoustic spectrum of purified GCaMP6f; adapted from ref.[31] (**d**) The equilibrium of the EGFP chromophore between protonated and unprotonated forms. (**e**) The photoswitching mechanism, absorption spectrum and photophysical properties of rsGCaMP1.1; adapted from ref.[118]

NIR phytochrome-derived fluorescent proteins

To achieve NIR absorption wavelengths requires alternative chromophore chemistry, achieved by biliverdin-binding fluorescent proteins derived from bacteriophytochromes (BphPs). These photosensory receptor proteins have three conserved protein domains (PAS, GAF, PHY) joined via α-helices with the biliverdin IXα (BV) chromophore housed in a pocket of the GAF domain and covalently bound to the cysteine residue of the PAS domain. Biliverdin, a linear tetrapyrrole chromophore (Figure 14), is an essential cofactor produced through haem metabolism, so endogenously present in mammalian cells. Although primarily designed for fluorescence imaging, their low quantum yields and NIR excitation wavelengths make them promising for PA application. The first BphP used for PAI, iRFP, (ε = 105000 $M^{-1}{\cdot}cm^{-1}$; λ_{max} = 692 nm; Φ = 0.059) was expressed in a mouse xenograft mammary gland cancer model, incorporated endogenous BV, and produced resolved images down to 4 mm depth.[120, 121] Development of novel phytochromes for deep-tissue imaging[122] enabled visualisation of iRFP670- and iRFP720-expressing tumours and vasculature even deeper to 8 mm, whilst maintaining tens of micrometre resolution.[123] Phytochromes are actually by nature photoswitchable with the BV chromophore undergoing a *cis-trans* isomerisation between red-absorbing (Pr) and far-red (Pfr) or near-red (Pnr) absorbing states (Figure 14), and the non-switchable reporters described above have been engineered to stabilise one of these states. Photoswitchable phytochromes have been exploited for differential PAI such as BphP1, which photoconverts from the Pfr to the Pr state on illumination with 730-790 nm, while 630 nm light induces the reverse transformation.[124] BphP1-expressing cells in the kidney could be imaged at ~8 mm depth by PAT with contrast-to-noise ratio ~20 after image subtraction, and tumour metastases in the liver could be longitudinally monitored for 1 month. Showcasing the versatility of PA, the reversible photoswitching facilitates super-resolution PAM in an analogous manner to RESOLFT,[125] demonstrated in fixed bacteria and U87 cells. There are several other examples of NIR photoswitchable proteins for PAI,[126-128] however these have not yet been applied to calcium or neuroimaging.

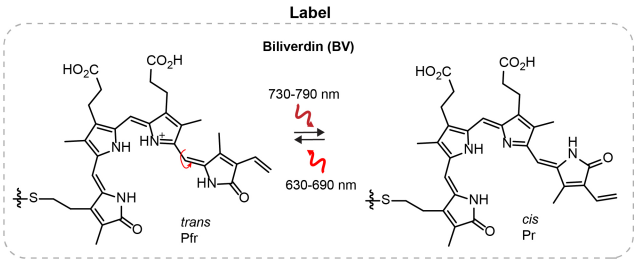


Figure 14: NIR bacterial phytochromes. The photoswitching mechanism of the biliverdin chromophore; adapted from ref.[40]

There have been significant protein engineering efforts to develop NIR calcium sensors for fluorescence imaging analogous to the diverse range of green-shifted variants, but these have however so far, only resulted in sensors with much smaller sensitivity and limited photoacoustic applicability. NIR-GECO1, was prepared by insertion of a calmodulin-peptide domain into the monomeric, biliverdin-dependent protein mIFP (Figure 15a).[129, 130] Although not characterised by PAI, the NIR absorption (λ_{max} = 678 nm) and low quantum yield (Φ_{apo} = 0.063, Φ_{sat} = 0.019) suggest that this protein could be used with this modality. However, the calcium-dependent decrease in absorption is small ($\Delta A/A_0$ = -0.68; Figure 15b), the photostability is relatively poor for a fluorescent protein, which is a limitation of the mIFP family,[131] and the BV binding efficiency is also low. *In vivo* expression of NIR-GECO1 in the sensorimotor cortex of mice resulted in only a 0.3% fluorescence decrease upon paw electrical stimulation. Two second-generation variants (NIR-GECO2 and NIR-GECO2G) developed via directed evolution yielded only minor improvements in brightness and sensitivity, with very similar photophysical properties.[132] The most significant improvement was a >2-fold increase in calcium-binding affinity which could explain the higher sensitivity in cultured neurons and acute brain slices. These new variants facilitated *in vivo* calcium imaging in *C. elegans* under osmotic and optogenetic stimulation, and spontaneous neural activity of *Xenopus laevis.* Recently, NIR-GECO2G has been used *in vivo* in the mouse cortex with marked improvement over NIR-GECO1 ($\Delta F/F_0$ ~-3) and brightness could be improved 2.5-fold by using Blvra$^{-/-}$ mice overexpressing BV.[133, 134] These mice were also imaged with PAT, and calcium dynamics could be detected with the aid of spectral unmixing from the haemoglobin signals.[135] However, rapid photobleaching limited recording to only a 90 s time window. A third generation, the NIR-GECO3 series, recently replaced mIFP with the brighter miRFP680 chromophore but did not result in a significant improvement in performance with fluorescence imaging in cultured cells and mice *in vivo.*[136] Also the calcium-dependent absorption change is reduced compared to the previous generations, which would decrease photoacoustic sensitivity.

Similar designs for intensiometric NIR phytochrome-based calcium indicators have been further investigated. GAF-CaMP2, also has a CaM-M13 calcium-binding domain inserted into a fluorescent protein, GAF-FP. This single-domain phytochrome should aid *in vivo* maturation rate, expression level and delivery but it has only been characterised *in vitro* so far.[137] The absorption and fluorescence spectra are slightly blue-shifted compared to NIR-GECO1 but optimised for commonly used 633/640 nm laser lines, is positive-going and has very similar quantum yields (Φ_{apo} = 0.018, Φ_{sat} = 0.048). However, function was only demonstrated in mammalian cell culture, with exogenous biliverdin added, as a ratiometric fusion with sfGFP (Figure 15c), giving a dynamic range of $\Delta R/R_0$ = 169%. Moreover, the absorbance turn-on is not sufficiently large ($\Delta A/A_0$ = 0.75) for an excellent photoacoustic calcium sensor.

By contrast, the other major approach to GECIs is ratiometric, involving two fluorescent proteins, which show Ca^{2+}-dependent FRET due to changes in distance or orientation between the two fluorophores. iGECI, joins two phytochrome-based NIR fluorescent proteins, miRFP670 and miRFP720 via calmodulin, the M13 peptide and short, flexible linkers (Figure 15d).[138] In the calcium-saturated state, FRET is favoured, resulting in a 600% increase in the fluorescence ratio. iGECI showed superior performance to NIR-GECO1 in neurons with higher signal-to-noise ratio (SNR), sensitivity and on-kinetics, but slower decay, and enabled functional imaging in acute brain slices to greater depths than was possible for GCaMP6s. *In vivo* fluorescence microscopy of the

mouse cortex determined a 3% $\Delta F/F_0$ in the miRFP670 donor channel during hindlimb electrical stimulation and up to 23% with intraperitoneal injection of clozapine *N*-oxide. The sensor was used for *in vivo* two-photon imaging in the motor visual cortex, and no photoacoustic characterisation of iGECI was reported. iGECI is likely unsuited for this modality with the decrease in miRFP670 fluorescence not sufficient for PA modulation as the absorbance remains constant.

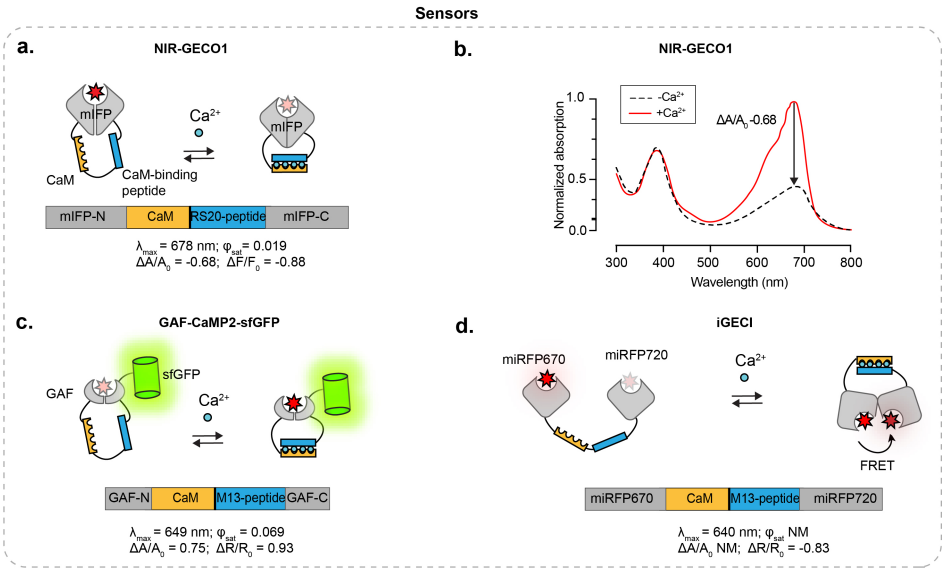


Figure 15: NIR phytochrome-based calcium indicators. (**a**) Structure of the calcium sensor NIR-GECO1.[129] (**b**) Absorption spectrum of NIR-GECO1, adapted from ref.[129] (**c**) Structure of the intensiometric calcium sensor GAF-CaMP2-sfGFP.[137] (**d**) Structure of the FRET-based calcium sensor iGECI.[138]

In summary, the development of photoacoustic sensors for neural activity is still incipient, and there is a huge disparity between the large number of neural activity indicators developed for fluorescence microscopy, and the critical absence of robust sensors adapted to PAI. To date, there are no neurotransmitter indicators, two examples of voltage sensitive dyes, and a handful of calcium sensors. Currently, the only potential options for PAI of mouse brains at far-red wavelengths are the nonspecific cyanine-based sensors from Westmeyer *et al.* (Figure 7d, e), and the NIR-GECO phytochromes from Campbell *et al.* (Figure 15a, b) which also show limited dynamic range, and suffer from poor photostability.

I.3.3 Hybrid reporters and biosensors

The limitations of the two main reporter classes described previously (synthetic and genetically-encoded), particularly the lack of targetability for small-molecules, and suboptimal fluorescent protein performance in the NIR, are conserved in the fluorescence microscopy field. This has led to the development of chemigenetic approaches, combining bright, tuneable synthetic fluorophores with self-labelling proteins (SLPs). These proteins recognise and bind to a small-molecule ligand, which can be incorporated into the structure of common fluorophores and delivered to biological systems expressing the protein. The most commonly used is HaloTag,[139, 140] and I wrote a review on the diverse applications of this platform for bioimaging:[141] Cook, A.; Walterspiel, F.; Deo, C. HaloTag-Based Reporters for Fluorescence Imaging and Biosensing, *Chembiochem* **2023**, *24* (12), e202300022. DOI: 10.1002/cbic.202300022.

I.3.3.1 The HaloTag self-labelling protein

The HaloTag self-labelling protein is derived from a 33 kDa bacterial enzyme, *Rhodococcus* dehalogenase DhaA, and binds irreversibly to a chloroalkane ligand (Figure 16a). The catalytic function is inhibited due to

mutations of key catalytic site residues in the 15 Å deep substrate binding tunnel, especially of the histidine base (H272), preventing release of the covalent ester intermediate (Figure 16b). The high specificity of HaloTag for bright fluorophore ligands makes it highly attractive for a broad range of imaging applications.[142] Crucially, dyes with finely tuneable optical and biochemical properties can be used, providing small-molecule flexibility which can, in principle, enable the expansion of this approach to PAI. Amongst the variety of self-labelling tags, HaloTag has the fastest labelling kinetics with specific chloroalkane ligands, which in general are more cell permeant and bioavailable than ligands for alternative tags. Moreover, HaloTag can achieve higher levels of fluorescence turn-on with widely used rhodamine fluorophores compared to other scaffolds.

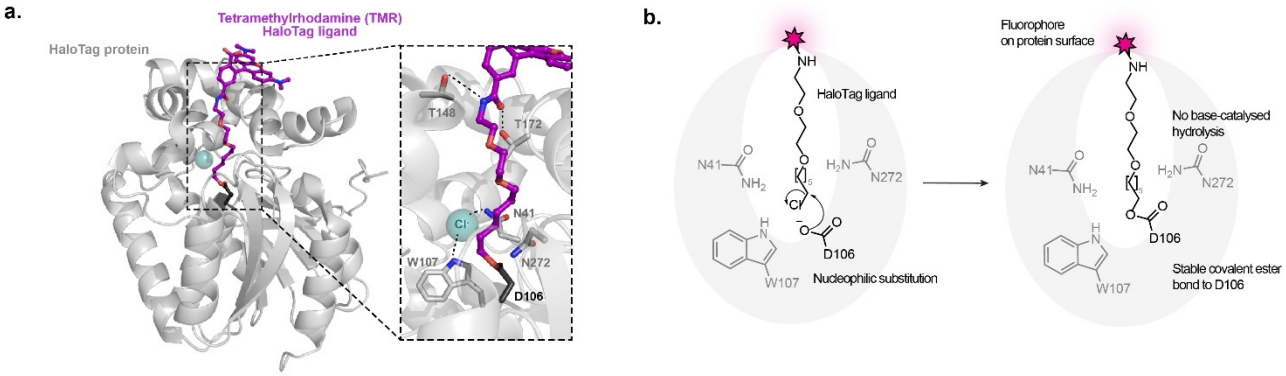


Figure 16: HaloTag protein, a self-labelling protein for chloroalkane ligands. (**a**) Crystal structure of HaloTag (PDB: 6Y7A)[143] bound to the tetramethylrhodamine (**TMR**) HaloTag ligand, with zoomed in inset of the binding tunnel. (**b**) Mechanism of covalent binding to the HaloTag ligand; adapted from ref.[141]

Rhodamines, from the xanthene family of dyes, have been applied extensively as fluorophore partners for SLPs, and HaloTag in particular, showing high brightness and photostability, cell permeability and ease of functionalisation (Figure 17a).[144] These dyes exist in equilibrium between a closed, non-absorbent lactone form and an open, highly absorbent and fluorescent zwitterionic form. This has led to the design of fluorogenic probes, which are predominantly closed in solution but display a large increase in absorption and fluorescence upon binding to HaloTag, due to a shift in equilibrium towards the open form, as a result of tight interactions between the dye and residues on the protein surface (Figure 16a, 17b). This equilibrium can be quantified and finely altered by rational chemical modifications, yielding ligands with enhanced fluorogenicity, across the visible spectrum, exemplified by **JF635-HTL**.[145] Fluorogenicity is highly advantageous for imaging: avoiding additional washing procedures due to the minimal background signal, and facilitating the design of biosensors with an absorption/ fluorescence modulation mechanism.

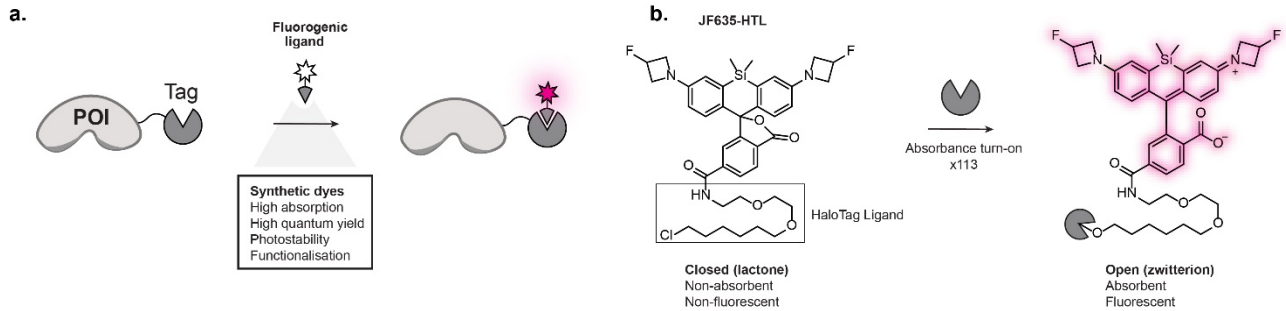


Figure 17: Fluorogenicity of rhodamines with HaloTag. (**a**) Schematic of labelling HaloTag, fused to a protein of interest (POI), with highly performant synthetic dyes. (**b**) The open-closed equilibrium of rhodamines, with **JF635-HTL** as an example.[145]

I.3.3.2 HaloTag-based biosensors

The HaloTag chemigenetic platform is amenable for the development of such biosensors[141] for a range of stimuli, including calcium and voltage, so is able to monitor neural activity. These sensors can be grouped into two categories, with either synthetic or protein sensing domains, but both exploit the key chemigenetic advantage of labelling specificity with bright fluorophores.

HaloTag-targeted biosensors (synthetic sensing domain)

This class uses fully synthetic biosensors, consisting of fluorophores functionalised with specific binding motifs. These are targeted to HaloTag fusion proteins, facilitating specific subcellular localisation, and HaloTag binding can also amplify the sensitivity of ligands exhibiting fluorogenicity. The main advantage is the use of optimised synthetic binding domains with high specificity for their target analyte, and greater diversity than protein-sensing domains. However, the downside is that the fluorophores tend to be bulky with the additional sensing domains, which can restrict cell permeability and bioavailability.

The most common design for HaloTag-targeted calcium sensors uses the same mechanism as the classic fluorescent, and also the new PA, synthetic calcium indicators.[13, 62, 63] PeT from the BAPTA (or equivalent) moiety to the rhodamine core quenches fluorescence but calcium-binding requires coordination of the nitrogen lone pairs, disrupting the electron transfer and so increasing fluorescence. One of the first examples was RhoCa-Halo, which could be excited in the orange channel, enabling multiplexing with green sensors but presented relatively low brightness and sensitivity.[146] To improve these key properties, the optimal position of the HaloTag ligand on another orange-emitting rhodamine derivative, **JF549**, was investigated until settling on the azetidine rings.[147] The resultant sensor, **JF549-BAPTA** could record single action potentials in cultured neurons. Substitution of the bridging O atom in the rhodamine core for silicon, generated the first fluorogenic far-red sensor, **JF646-BAPTA**, albeit with weak fluorogenicity ($\Delta F/F_0 = 2.4$ upon binding to HaloTag when saturated with Ca^{2+}), combining PeT quenching with the modulation of the open-closed equilibrium due to HaloTag binding (Figure 18a). Another synthetic strategy for engineering fluorogenic rhodamine ligands replaces the spirocyclic 3-carboxylic acid with electron-deficient amides (MaP dyes).[148] Attachment of a BAPTA or MOBHA calcium chelator onto this sulphonamide led to the MaPCa series, including the far-red **MaPCa-656** with >100-fold fluorescence turn-on upon HaloTag binding and $\Delta F/F_0 \geq 6$ in response to calcium (Figure 18b).[149] The choice of calcium-binding moiety resulted in low and high affinity variants, which were used to monitor calcium dynamics in rat hippocampal neurons in the ER and cytosol respectively. In addition, labelling H-Luc, a fusion of HaloTag with a luciferase, generates far-red bioluminescent Ca^{2+} indicators. Bioluminescence resonance energy transfer (BRET) occurs between the luciferase and the MaPCa dyes, resulting in up to 6.5-fold ratiometric changes.[149] The Ca^{2+}-dependent change is manifested primarily in fluorescence rather than absorption for all of these calcium sensors and they have high quantum yields ($\Phi_{sat} > 0.4$), so would likely not be suitable for photoacoustic application.

The approach to develop HaloTag-based voltage reporters required incorporation of different functional groups into the rhodamine structure, although also employs a PeT mechanism. Rhodamines are connected to an aniline donor via a lipophilic molecular wire to mediate membrane potential-dependent PeT, as the quenching is inhibited on membrane depolarisation, resulting in an increase in fluorescence. This indicator is dim and red-shifted so would generate significant PA signal but the sensitivity would depend on the PeT quenching effect on absorption. Incorporation of the HaloTag ligand and an extended, flexible linker via the spirocyclic amide mediates specific labelling of neuronal membranes for activity recordings in brain slices and *Drosophila* (Figure 18c).[150-152] Nevertheless, these voltage sensors and the calcium indicators discussed in the previous paragraph, would likely struggle to monitor neural activity in living rodents because of limited bioavailability of these complicated ligands.

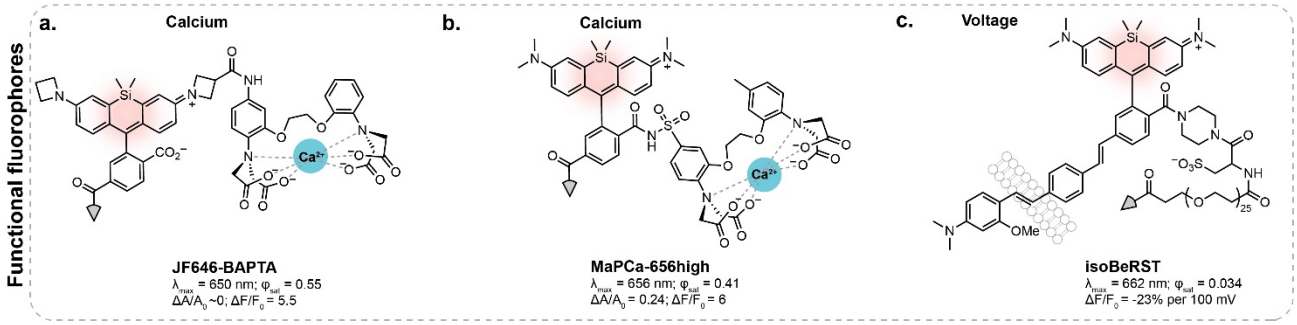


Figure 18: HaloTag-targeted biosensors. Calcium sensors: (**a**) **JF646-BAPTA**[147] and (**b**) **MaPCa-656high.**[149] Voltage indicator: (**c**) **isoBeRST**.[152] Adapted from ref.[141]

Chemigenetic biosensors (genetically-encoded sensing domain)

Alternatively, the sensing mechanism can originate from the protein scaffold via functionalisation of HaloTag with protein-based sensing domains. Both ratiometric and intensiometric designs of biosensors are possible. The former involves combination of HaloTag with a fluorescent protein or luciferase to form a FRET or BRET pair. Conformational changes of the protein in response to the stimuli alter the FRET/BRET efficiency with the HaloTag-bound fluorophore acceptor, hence modulating the fluorescence ratio in the donor and acceptor channels. By contrast, intensiometric HaloTag biosensors are usually labelled with environmentally-sensitive fluorophores, which exhibit fluorogenicity. With these platforms the protein's structural alterations modify the dye-protein interactions and hence the photophysical properties of the bound dye. The key advantage of these biosensors with genetically-encoded sensing domains is the spectral flexibility afforded by simply changing the dye ligand. In addition, these ligands are simpler, only consisting of the fluorophore, which aids bioavailability and delivery to living animals.

The development of these sensors requires extensive protein engineering. For example, heavy mutagenesis of the protein-protein interface in a HaloTag-EGFP fusion resulted in almost quantitative FRET efficiency even with non-spectrally matched fluorophores like **SiR** and **JF669**.[153] Copying the design of the GECI yellow cameleon 3.6, the calcium-binding protein calmodulin and the M13 binding peptide were inserted between the proteins to generate a ratiometric response to calcium, increasing FRET efficiency with calcium concentration (Figure 19a). The flexibility of the chemigenetic modular design enabled preparation of colour variants, by substituting EGFP for other fluorescent proteins or using different fluorophores as the HaloTag-bound acceptor. The best performing variant, ChemoG-CaM-SiR, was able to detect single action potentials in cultured neurons.

As mentioned above, insertion of a luciferase, NanoLuc, into HaloTag to make H-Luc enables BRET and red-shifting of emission wavelengths over 200 nm from the native substrate furimazine (460 nm).[154, 155] HaloTag-NanoLuc-mediated bioluminescence can also be readout for electrical activity. Q-BOLT uses a two-acceptor system with BRET favoured to the HaloTag-bound **TMR-HTL** ligand, rather than to the membrane-localised voltage-sensitive dye dipicrylamine, upon membrane depolarisation, and vice versa.[156] However, this system is complicated by the delivery of two small-molecules, **TMR-HTL** and dipicrylamine, and cannot compete with the outstanding performance of Voltron. This voltage indicator is FRET-based, consisting of Ace2, a microbial rhodopsin voltage-sensing domain, fused to HaloTag.[157] Voltron functions via a membrane depolarisation-dependent rise in FRET between the opsin retinal cofactor and the bound **JF525-HTL** rhodamine dye (Figure 19b). The brightness and high photostability of these dyes make Voltron one of the most performant voltage indicators *in vivo* to date, facilitating extended neuronal recordings in the mouse visual cortex.[157, 158] Such ratiometric sensors have less potential for PAI, as like iGECI, the fluorescent donor's absorbance and hence PA signal is not modulated, whilst the bright, fluorescent HaloTag ligand acceptor would generate minimal signal.

Another limitation of these ratiometric sensors is that the FRET partners have blue/green emission wavelengths, and the red-shifted fluorescent protein alternatives are quite dim, and so it is difficult to use far-red synthetic dyes because of poor FRET efficiency (with the ChemoX system being an exception). Therefore, intensiometric HaloTag-based biosensors have the advantage of using bright, far-red fluorogenic ligands in conjunction with protein scaffolds which undergo sufficient conformational changes to exploit their environmental sensitivity. For example, in an analogous design to the GCaMP GECIs, two chemigenetic Ca^{2+} indicators, HaloCaMP1a and HaloCaMP1b, were engineered by fusing a circularly permutated HaloTag protein to calmodulin and a calmodulin-binding peptide, from myosin light chain kinase (MLCK) or calmodulin-dependent kinase kinase (CKK; Figure 19c), respectively.[143] The sensor is completed by labelling with **JF635-HTL** which becomes more open in the Ca^{2+}-bound state, inducing an increase in absorbance and fluorescence. Varying the small-molecule dye creates a series of derivatives with tuneable wavelengths, sensitivities, brightness and affinities for calcium. Single action potentials could be recorded in cultured neurons, where HaloCaMP outperformed the far-red GECI NIR-GECO1 in terms of sensitivity, brightness, rise and decay times. rHCaMP is another scaffold with CaM-M13 inserted into HaloTag, but functions as a ratiometric sensor with a unique colour-shifting coumarin-benzopyrilium ligand.[148] With the same 'plug-and-play' strategy, replacing GFP in the voltage sensors ASAP and ArcLight by fusion of HaloTag variants to voltage-sensitive domains, led to bright, far-red chemigenetic versions (Figure 19d).[143] These systems rely predominantly on absorption changes as the rhodamine ligands shift to the open form, and so would show a modulation in PA signal too, but their quantum yields need to be reduced to enhance signal. Proof of function of these sensors *in vivo* is yet to be demonstrated because of limited dye bioavailability. However, recently, using a BBB-permeable ligand, **JF669-HTL**, a novel Ca^{2+} indicator, WHaloCaMP, was developed (Figure 19e).[159] The dye is only very weakly fluorogenic, and the Ca^{2+}-dependent fluorescence changes, 7-fold *in vitro*, result from PeT quenching via a tryptophan residue incorporated into the HaloTag protein surface. WHaloCaMP was used for *in vivo* neuronal calcium imaging in flies, zebrafish and mice but unfortunately the small absorption modulation hinders PAI application.

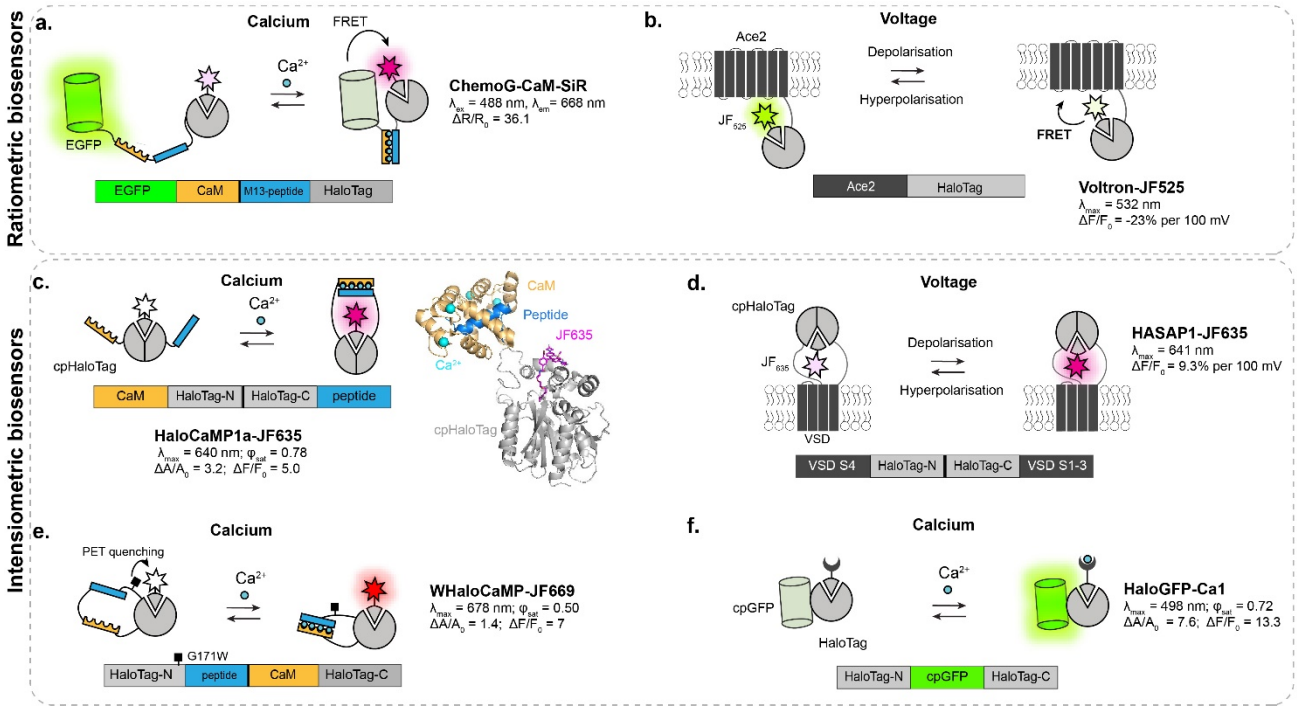


Figure 19: Ratiometric and intensiometric chemigenetic biosensors for neuroimaging. Ratiometric biosensors: (**a**) ChemoG-CaM, labelled with **SiR-HTL**, for calcium,[153] (**b**) Voltron, labelled with **JF525-HTL**, for voltage.[157] Intensiometric biosensors: (**c**) HaloCaMP1a, labelled with **JF635-HTL**, for calcium, plus the crystal structure of HaloCaMP1b bound to **JF635-HTL** (PDB: 6U2M);[143] (**d**) HASAP1, labelled with **JF635-HTL**, for voltage;[143] (**e**) WHaloCaMP, labelled with **JF669-HTL**, for calcium;[159] (**f**) HaloGFP-Ca1, using a synthetic BAPTA binding domain and protein chromophore.[160] Adapted from ref.[141]

Another strategy maintains the use of a synthetic sensing domain by targeting, for example, BAPTA, to a cpGFP-HaloTag protein via incorporation of HTL into the binding motif. With **BAPTA-HTL** bound to cpGFP-HaloTag, calcium-binding influences the GFP chromophore's protonation state (Figure 19f).[160] This approach enables exploitation of the wider diversity of synthetic binding motifs but for calcium and voltage, there are already sensitive protein domains, such as calmodulin. As for the GCaMP sensors, there is a significant absorption change for HaloGFP-Ca1 which could be exploited for PAI application, but the limitation of poorly penetrating green wavelengths remains and the extinction coefficient is also quite low (24000 $M^{-1} \cdot cm^{-1}$).

These hybrid systems provide solutions to the deficiencies in the properties of both small-molecule and genetically-encoded protein sensors. The use of synthetic chromophores provides easy access to far-red wavelengths whilst maintaining a specific, universal targeting method via the HaloTag protein. However, they are currently unsuitable for PAI with their high brightness, and so adaptation is required to develop novel chemigenetic PA probes.

I.4 Objectives and outline

I.4.1 Objectives

To continue to improve our understanding of the brain requires a range of imaging techniques to visualise neuronal activity across spatial scales. Photoacoustic imaging is a relatively novel mesoscopic modality able to fill a niche amongst current methods, which are already performant at the neuronal and whole brain scale, through its unique combination of key parameters: molecular sensitivity, penetration depth, speed and resolution. However, this introduction has hopefully highlighted the lack of suitable reporters of neural activity for this modality, lagging behind those for fluorescent microscopy, which has prevented PAI from becoming a widely used tool in the neuroscience field. Despite the slower kinetics relative to the electrical signals, calcium imaging is the predominant method for monitoring neural activity thanks to the high sensitivity of calcium indicators, and the time gap to electrophysiology continues to be narrowed with the development of faster sensors. Therefore, to achieve the ultimate goal of this project, whole brain photoacoustic neuroimaging in mice, we needed to first develop novel calcium sensors optimised for this modality. Performant PA calcium sensors would enable recording of neural circuit activity and connectivity across brain regions throughout the several millimetre depths of the mouse brain at high resolution, in the range of tens of micrometres, providing a correlative imaging tool between the micro- and macroscales.

I.4.2 Design

Novel photoacoustic calcium sensors must meet certain criteria, which, as discussed previously in this chapter, are not currently met by existing platforms, in order to provide sufficient contrast and sensitivity at imaging depths down to the bottom of the mouse brain. Strong absorption at far-red/NIR wavelengths, negligible fluorescence quantum yield and high photostability are essential to generate high PA signal intensity, whilst being capable of localising specifically to particular cells or subcellular regions, with the primary target being the neuronal cytosol for detection of activity-dependent calcium fluctuations.[23] To function as calcium sensors, they must have a sensitive and selective modulation mechanism in response to calcium. Moreover, these indicators need fast kinetics to capture the calcium dynamics elicited by neurons firing, following the rapid electrical signal rise (<10 ms) and decay times (~50−70 ms) for action potentials, and respond at physiological calcium concentrations (~100−200 nM).[161, 162] Currently, the wide diversity of hybrid biosensor designs is revolutionising molecular imaging by combining the excellent photophysical properties of small-molecule dyes with the labelling specificity afforded by genetically-encoded proteins. Consequently, to meet the stringent requirements for a photoacoustic calcium indicator, we selected a chemigenetic approach: exploiting the specificity of HaloTag-based protein scaffolds for neuronal labelling, which are highly amenable to sensor design, with synthetic ligands engineered for strong far-red absorption and PA signal.

The rapid labelling kinetics and specificity of HaloTag, as well as *in vivo* applicability and flexibility to protein engineering, makes it the standout chemigenetic platform. Rhodamine ligands are the most performant binding partner with HaloTag and exhibit this environmentally sensitive equilibrium between a non-absorbing lactone form and a highly absorbing zwitterionic form, which can be exploited for turn-on probes and sensors (see section I.3.3.2). By designing quenched, far-red/NIR absorbing rhodamine derivatives, also possessing this open-closed equilibrium, we could generate "acoustogenic" ligands, adapting the chemigenetic approach to the photoacoustic modality (Figure 20). These ligands would remain nonfluorescent in both open and closed forms, but would show significant increases in absorbance and hence PA signal upon binding to the HaloTag protein scaffold, enabling application as performant photoacoustic labels.

Concerning the calcium indicator design, due to the requirement for the low fluorescence and the origin of the photoacoustic effect,[163] we reasoned that sensors functioning via changes in absorption can achieve higher sensitivity than quantum yield-modulated sensors. The photoacoustic signal generated by a molecular reporter can be approximated as:

$$PA = \Gamma \cdot F \cdot \mu_a(1 - \Phi)$$

Equation 1: Photoacoustic signal (PA) generation by a reporter. Γ: Grüneisen parameter (which quantifies the thermodynamic efficiency of the heat to pressure conversion), F: optical fluence (the light intensity per unit area), μ_a: absorption coefficient, Φ: fluorescence quantum yield.[163]

Therefore, the PA response of a biosensor to calcium can be determined by:

$$\frac{\Delta PA}{PA_0} = \frac{\mu_a^+(1-\Phi^+) - \mu_a^-(1-\Phi^-)}{\mu_a^-(1-\Phi^-)} = \frac{\mu_a^+(1-\Phi^+)}{\mu_a^-(1-\Phi^-)} - 1$$

Equation 2: Photoacoustic signal (PA) response from a calcium sensor. μ_a^+: absorption coefficient in the calcium-saturated state, μ_a^-: absorption coefficient in the calcium-free state, Φ^+: fluorescence quantum yield in the calcium-saturated state, Φ^-: fluorescence quantum yield in the calcium-free state.

Comparing theoretical sensor properties where only either quantum yield or extinction coefficient are modulated is revealing (Table 1). A sensor with low Φ = 0.1 which reduces by a factor of 2 in the calcium-bound state, results in only a 5% increase in PA signal (**A**), whilst a two-fold rise in extinction coefficient doubles PA signal (**B**). Stronger quenching from this low starting value cannot achieve greater than an 11% increase in PA signal (**C**), so to achieve sufficient sensitivity the calcium-free state must be bright and subject to a larger decrease (>10-fold) in Φ on calcium-binding (**D**), or the modulation mechanism should be dominated by absorption. This is supported by the performance comparison of reversibly switchable proteins in PAI, where photochromic proteins, presenting absorption changes, demonstrated far greater sensitivity than so-called "QY changers".[164]

Table 1: Predicted effects of extinction coefficient and quantum yield changes on photoacoustic signal and sensitivity of theoretical calcium sensors.

	Φ_{-Ca}	ε_{-Ca}	Φ_{+Ca}	ε_{+Ca}	PA_{-Ca}	PA_{+Ca}	$\Delta PA/PA_0$
A	0.1	100000	0.05	100000	90000	95000	0.0555
B	0.05	50000	0.05	100000	47500	95000	1.00
C	0.1	100000	0.001	100000	90000	99900	0.110
D	0.525	100000	0.05	100000	47500	95000	1.00

Therefore, mechanisms such as PeT, which the majority of synthetic Ca^{2+} sensors rely on,[13] and FRET, are in principle less advantageous for photoacoustic probes. PeT can still be employed in PA sensors when inducing a sufficient absorption change, along with the fluorescence modulation, for example in **CaSPA**[63] and "**L**"[62] (Figure 7d, e). However, the sensitivity of these indicators is low and so we must consider other systems where absorption changes predominate, exemplified by the GCaMP series. At far-red wavelengths, the chemigenetic HaloTag-based calcium indicator scaffolds, HaloCaMP[143] and rHCaMP[148] appear the most promising for readaptation to PAI (Figure 19). These sensors are principally absorption modulated via the opening of the

rhodamine open-closed equilibrium and are tuneable to the choice of dye ligand. Therefore, the usual fluorogenic rhodamine ligands could be substituted with synthetically quenched, acoustogenic ligands. Conformational changes in HaloCaMP induced by calcium binding influence the surrounding environment of the bound ligand, opening the equilibrium to provide strong absorption, and thus, PA turn-on (Figure 20).

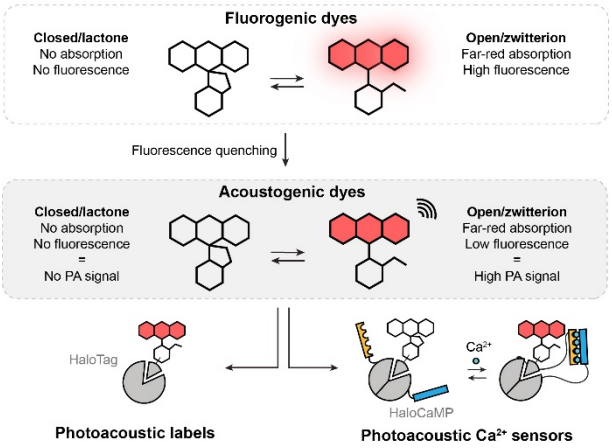


Figure 20: General concept for the design of chemigenetic photoacoustic labels and calcium sensors dependent on the open-closed equilibrium of acoustogenic dyes engineered from their fluorogenic analogues. Adapted from ref.[165]

I.4.3 Outline

In this thesis, I describe the first-generation of chemigenetic probes and calcium sensors developed for photoacoustic imaging, which has the potential to resolve the current lack of robust reporters for this modality and lead to the establishment of PAI as a permanent fixture in the neuroimaging field. The experimental pipeline (Figure 21) involves four key stages, which will be discussed in the following chapters. First, I designed and synthesised the acoustogenic rhodamine derivatives, which form the synthetic component of our labels and calcium sensors. These were then characterised *in vitro* for their photophysical properties, including PA performance, bound to HaloTag-based protein scaffolds. Subsequently, I attempted to rationally optimise the dye chromophore structure to improve acoustogenicity and began to undergo protein engineering of the HaloCaMP scaffold. The fourth stage involves evaluation of the *in vivo* labelling efficiency of the acoustogenic ligands and PAI of mouse brain tissues.

In Chapter II, I describe the design of our PA calcium sensors combining novel acoustogenic ligands with HaloTag-based protein scaffolds. The synthesis of a library of these quenched ligands will be discussed, before spectroscopic characterisation in absorption and PA when bound to their HaloTag binding partners. The resulting calcium sensors showed superior performance to current indicators such as NIR-GECO1 with sensitivity up to 7.5-fold. Chapter III details our attempts to optimise these first-generation sensors, focusing on the key properties of acoustogenicity, calcium affinity and bioavailability. I designed four HaloCaMP affinity variants, however further modifications are required to improve their overall performance. In addition, I explored methods for evaluation of protein mutants in future engineering endeavours and have begun to establish colony screening for chemigenetic PA calcium sensors. Finally, Chapter IV shows the application of our labels and sensors in systems of gradually increasing complexity up to *in vivo* brain labelling in mice. Our acoustogenic probes demonstrate strong performance in tissue-mimicking phantoms, *ex vivo* labelled mouse brain slices and can label neurons *in vivo*, but require optimisation of bioavailability for consistent labelling efficiency.

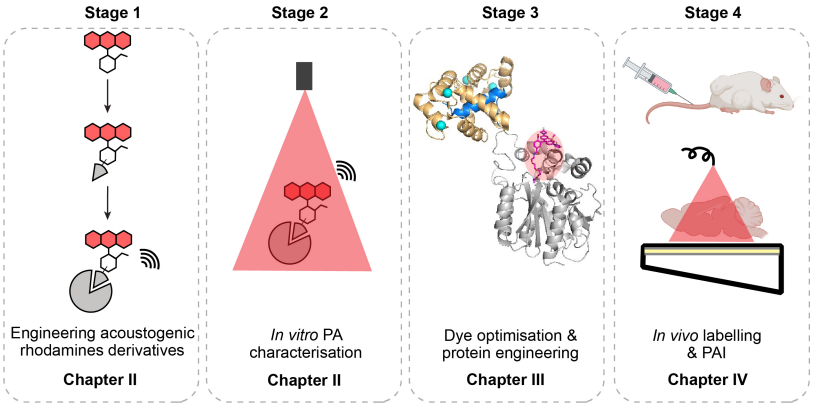


Figure 21: Pipeline for the development of the first chemigenetic photoacoustic probes and calcium sensors.

Chapter II: Design, engineering and *in vitro* characterisation of chemigenetic photoacoustic probes and calcium sensors

The work discussed in this chapter was conducted in collaboration with Nikita Kaydanov (Prevedel group, EMBL), Dr. Begoña Ugarte-Uribe (Deo group, EMBL), and with the support of the Protein Expression and Purification Core Facility (PEPCF, EMBL), and Dr. Dieter Schollmeyer (Johannes Gutenberg Universität Mainz). I performed the chemical syntheses and full chemical characterisations, the preparation of dye solutions, and all absorption and fluorescence spectroscopy measurements whilst Nikita developed the custom spectrometer and conducted the photoacoustic spectroscopy of the probes. Begoña and the PEPCF helped purify the HaloTag-based proteins for dye characterisation. Dr. Schollmeyer collected the X-ray diffraction data and solved the crystal structure of the small-molecule dye **S21**.

The results presented in this chapter were in part published in the Journal of the American Chemical Society (JACS):[165] Cook, A.; Kaydanov, N.; Ugarte-Uribe, B.; Boffi, J. C.; Kamm, G. B.; Prevedel, R.; Deo, C.; Chemigenetic Far-Red Labels and Ca^{2+} Indicators Optimized for Photoacoustic Imaging, *J Am Chem Soc,* **2024**, *146* (34), 23963-23971. DOI: 10.1021/jacs.4c07080. I adapted several figures from our publication for use in this chapter, and integrated additional unpublished data.

II.1 Design, synthesis, and characterisation of acoustogenic dye scaffolds

The emerging mesoscopic modality of photoacoustic imaging has very promising applications in neuroscience, but is currently limited by an absence of optimised reporters. To resolve this deficiency, we set out to develop novel "acoustogenic" probes and calcium sensors, using the recently established chemigenetic approach, previously never applied to PAI.

The first step was the design and synthesis of suitable dye scaffolds, compatible for functionalisation and conjugation to HaloTag, which would meet the key photophysical requirements of: strong far-red/NIR absorption to access the optical window, minimising background absorption from haemoglobin in particular; negligible Φ to maximise the photoacoustic generation efficiency; and a chromogenic open-closed equilibrium, providing high absorbance turn-ons upon binding to the HaloTag protein partner. First, I synthesised and evaluated the free dyes, before selecting the most promising candidates for incorporation of the HaloTag ligand (HTL) to quantify the acoustogenicity in the presence of the HaloTag protein. This is an established strategy for the development of fluorogenic ligands, with the free dye screening saving additional synthetic steps.[145, 166] Spectroscopic characterisation of the photophysical properties of the free dyes was subsequently performed, directing rational chemical modifications for optimisation of the open-closed equilibrium in particular. The absorption maxima and extinction coefficients of the free dyes were measured in aqueous buffer at physiological pH, and in MeCN/H_2O mixtures to evaluate their open-closed equilibrium and estimate their chromogenicity. To assess this property, we defined dyes with $200 < \varepsilon < 15000$ $M^{-1}\cdot cm^{-1}$ as chromogenic, based on the extensive work on fluorogenic rhodamine derivatives, particularly by the Lavis and Johnsson groups.[145, 148, 166] A low extinction coefficient of the dye ligand in solution, where the closed form predominates, is needed to achieve the highest absorption, and hence PA signal, turn-on upon HaloTag binding. This does not guarantee chromogenicity however, as opening of the dye on the protein depends strongly on the dye scaffold and the dye-protein interactions which are hard to predict, and usually dyes do not attain their fully open state bound to the protein. However, if the free dye is "too closed" in solution with no detectable absorption peak, corresponding to $\varepsilon < 200$ $M^{-1}\cdot cm^{-1}$, the interaction with HaloTag is usually insufficient to result in substantial turn-on. Three major scaffolds were investigated, based on silicon rhodamines, pyrrole-xanthene and Malachite Green lactone derivatives.

II.1.2 Silicon rhodamines

II.1.2.1 Synthesis

Silicon rhodamines (or Si-rhodamines) are a well-established family of dyes with high extinction coefficients at far-red (~10^5 $M^{-1}\cdot cm^{-1}$) absorption wavelengths and have been used to make several highly fluorogenic dyes, such as **JF635** and **MaP700** with >100-fold absorbance turn-ons,[145, 148] due to an open-closed equilibrium which

lies predominantly closed in solution. Consequently, I made Si-rhodamines with *N*-aryl (**1-4**) or *N*-methylpiperazine (**6**) auxochromes (Figure S1), which function as quenching groups likely through a combination of PeT and TICT mechanisms.[85, 167] Additionally, I synthesised an unsymmetrical Si-rhodamine (**5**) with a single azetidine ring to investigate whether a single *N*-aryl group is sufficient for quenching. Last year, dim NIR-absorbing dihydroquinoline-fused Si-rhodamines were described and therefore also included in our library (**7**, **8**).[168] Such red-shifted rhodamine scaffolds are often too closed to be chromogenic, so substitution of fluorines on the pendant phenyl ring was also investigated as this modification shifts the equilibrium towards the open form.[166, 169] These were synthesised via previously reported synthetic procedures from Si-fluorescein ditriflates precursors[169] before functionalisation using palladium-catalysed Buchwald-Hartwig cross-couplings with amines (Figure 22).[167] The synthesis was generally successful, despite lower yields for the *N*-aryl cross coupling with the tetrafluorinated ditriflate precursor. The dihydroquinoline-fused ring compounds were synthesised using the recently published lactol condensation method.[168]

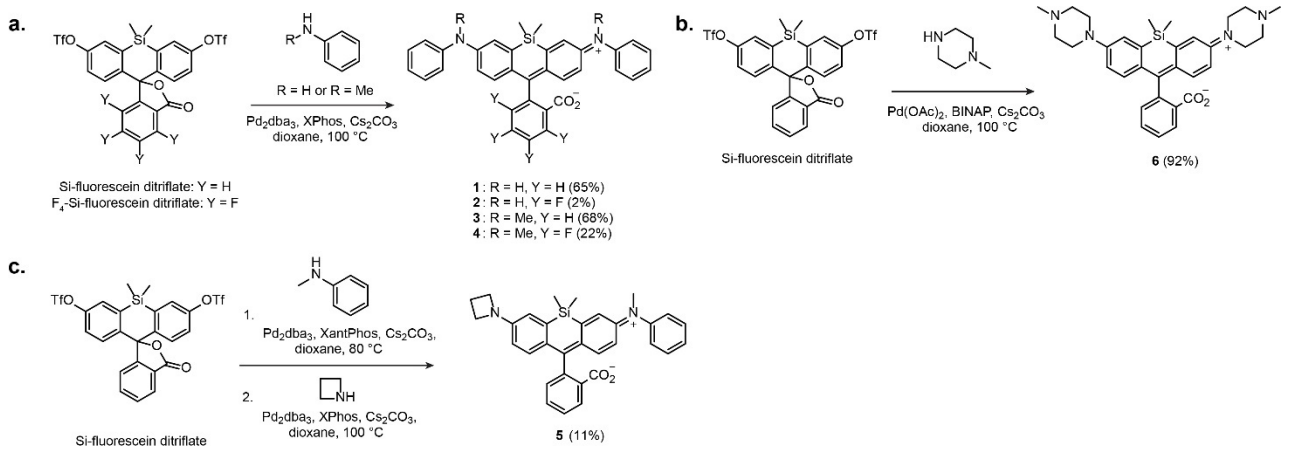


Figure 22: Synthesis of silicon rhodamine free dyes 1-6. Buchwald-Hartwig cross-coupling of Si-fluorescein ditriflate precursors with (**a**) aniline or *N*-methylaniline, (**b**) *N*-methylpiperazine, (**c**) *N*-methylaniline and then azetidine.

II.1.2.2 Spectroscopic characterisation

The nonfluorinated dyes **1** and **3** possess a too closed equilibrium with absorption below the detection limit (ε < 200 $M^{-1}{\cdot}cm^{-1}$, Figure 23a, S3). However, their fluorinated equivalents **2**, **4** as well as **5**, **6** showed a detectable absorption signal (200 < ε < 2000 $M^{-1}{\cdot}cm^{-1}$), which is an indicator of potential chromogenicity, when bound to the HaloTag protein (Figure 23a, S3).[166] This was supported by the measurements in MeCN/H_2O mixtures,[170, 171] with **1** and **3** remaining undetectable whilst very small absorbance peaks (ε > 200 $M^{-1}{\cdot}cm^{-1}$) were visible for **2**, **4**-**6** (Figure 23b). MeCN/H_2O was chosen instead of the standard dioxane/H_2O combination because the ligands appeared to show improved solubility, enabling differentiation between completely closed and chromogenic dyes. Because of the low absorption, relatively high dye concentrations are required (5 µM) to differentiate between chromogenic and completely closed dyes, which then results in aggregation of these hydrophobic dyes in more aqueous solutions. For **8**, I measured a substantially higher extinction coefficient than reported, measured at 1.25 µM rather than 5 µM, suggesting that this dye is actually fully open (Figure 23a, S3). Indeed, the MeCN/H_2O titration at 5 µM demonstrated that **8** is highly susceptible to aggregation in aqueous solution with a peak at ε = 98000 $M^{-1}{\cdot}cm^{-1}$ in 30% MeCN/H_2O (dielectric constant = 65.6) before a rapid decrease with increasing H_2O content, which would explain this discrepancy (Figure 23b). The nonfluorinated analogue **7** also aggregates but is still chromogenic, whilst the chromogenic dyes **2** and **4** show this aggregation behaviour too but at ≥ 80% water. In terms of wavelength, all of the compounds showed desirable absorption in the far-red region (λ_{max} > 640 nm; Figure 23a, S3), confirmed by absorption measurements in acidic conditions to fully open the dyes (2,2,2-trifluoroethanol (TFE) containing 0.1% v/v trifluoroacetic acid (TFA); Figure S4).[145]

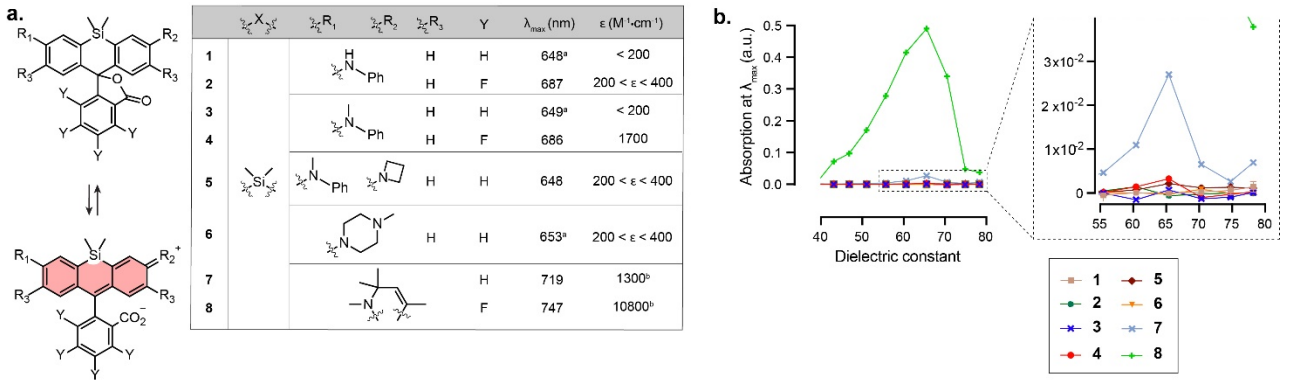


Figure 23: Si-rhodamines as acoustogenic dyes. (**a**) General structure and open-closed equilibrium of the synthesised dyes **1-8** with absorption properties; all measurements were made in 10 mM HEPES, pH = 7.4 except for [a] in 2,2,2-trifluoroethanol (TFE) containing 0.1% trifluoroacetic acid (TFA); [b]apparent value due to aggregation. Values are the mean of 3 replicates. (**b**) Absorption at λ_{max} of free dyes **1-8** in MeCN/H_2O mixtures.[171, 172] All measurements were performed at 5 µM dye concentration, values are the mean of 3 replicates. Adapted from ref.[165]

II.1.3 Pyrrole-xanthene derivatives

II.1.3.1 Synthesis

The second family selected consist of pyrrole-functionalised O-xanthenes, which provide a new strategy to bring xanthenes into the NIR region by extending the electronic conjugation system.[173] These have low fluorescence quantum yields, again likely through a combination of PeT and/or TICT mechanims, and their high potential for photoacoustic imaging was demonstrated with a thienylpiperidine analogue (Figure 9b).[76] The target compounds **9-12** were obtained by Suzuki coupling of the fluorescein ditriflate precursors[167] with the corresponding commercially available boronic esters (Figure 24, S1).

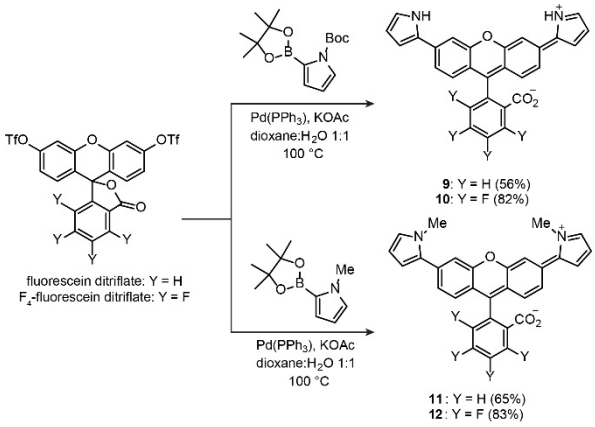
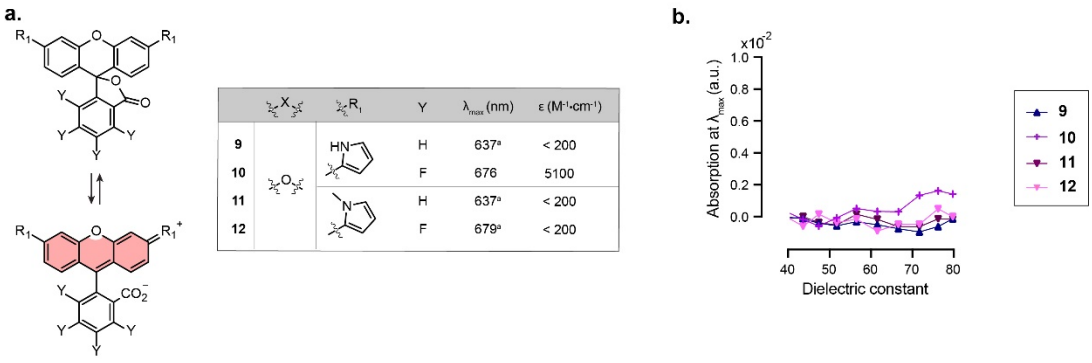


Figure 24: Synthesis of pyrrole-xanthene derivates 9-12. Suzuki coupling of boronic esters with fluorescein ditriflate precursors.

II.1.3.2 Spectroscopic characterisation

As expected, these dyes showed λ_{max} > 640 nm, however the nonfluorinated dyes **9, 11** and fluorinated dye **12** were largely closed (ε < 200 $M^{-1}{\cdot}cm^{-1}$, Figure 25, S3) in aqueous buffer. Dye **10** showed potential chromogenicity (ε = 5100 $M^{-1}{\cdot}cm^{-1}$) with low absorbance compared to the fully open state (~10^5 $M^{-1}{\cdot}cm^{-1}$).[173] This dye also possesses a broad absorption spectrum, possibly due to hydrogen bonding of the pyrrole NH with H_2O (Figure S3).



	X	R1	Y	λmax (nm)	ε (M-1·cm-1)
9	O	HN	H	637[a]	< 200
10	O	HN	F	676	5100
11	O	N	H	637[a]	< 200
12	O	N	F	679[a]	< 200

Figure 25: Pyrrole-xanthene derivatives as acoustogenic dyes. (**a**) General structure and open-closed equilibrium of the synthesised dyes **9-12** with absorption properties; all measurements were made in 10 mM HEPES, pH = 7.4 except for [a]in TFE containing 0.1% TFA. Values are the mean of 3 replicates. (**b**) Absorption at λ_{max} of free dyes **9-12** in MeCN/H_2O mixtures.[171, 172] All measurements were performed at 5 µM dye concentration, values are the mean of 3 replicates. Adapted from ref.[165]

II.1.4 Malachite Green lactone derivatives

II.1.4.1 Synthesis

Completing our library of acoustogenic dyes, are triarylmethane lactones, derivatives of the dyestuff Malachite Green with the *o*-carboxylic acid group on the pendant phenyl ring to form the open-closed equilibrium. Absence of the bridging atom in the xanthene system provides far-red absorption and crucially minimal fluorescence because of the rotational freedom of the phenyl rings.[174, 175] A simple one-step Friedel-Crafts acylation of *N,N*-dimethylaniline with (tetrafluoro)phthalic anhydride afforded the original nonfluorinated and fluorinated compounds **13** and **14** respectively (Figure 26). **14** has a slightly lower extinction coefficient and a similar structure to the pioneering fluorogenic dye **SiR**, amenable for considerable structural modification and so derivatives were designed and synthesised with the aim of lowering the absorbance, by marginally closing the equilibrium, to improve the chromogenicity.

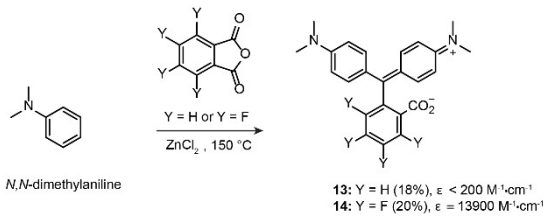


Figure 26: Synthesis of Malachite Green lactone derivatives 13, 14. Friedel-Crafts acylation of *N,N*-dimethylaniline with (tetrafluoro)phthalic anhydride.

N-auxochrome modification

Given the success of replacing the *N,N*-dimethyl groups in **SiR** with azetidines to make **JF646** and **JF635**, which increased the absorbance turn-on upon binding to HaloTag, from 7 to 21 and 113-fold respectively,[145, 176] I attempted to synthesise the equivalent Malachite Green lactone structures. However, this proved a far greater synthetic challenge than expected (Figure 27). Palladium-catalysed cross-coupling of the tetrafluorinated ditriflate precursor with substituted azetidines (Route A) resulted in competing nucleophilic aromatic substitution (S_NAr) of fluorines on the bottom ring. This S_NAr reaction is too fast occurring in preference to cross-coupling under milder conditions and occurs even in the absence of catalyst and ligand at room

temperature. Protecting the 6-position, which is most susceptible to substitution,[166] could not prevent S_NAr, likely in the 4-position. To reduce the electrophilicity of the bottom ring, I synthesised a tetrachlorinated ditriflate precursor,[177] which I planned to convert back to the fluorinated compound after the cross-coupling using a halide exchange (Halex) reaction process (Route B).[178] However, the cross-coupling conditions continued to facilitate the S_NAr process, preventing isolation of the desired product. The direct Friedel-Crafts acylation of 1-phenylazetidine also failed, likely due to decomposition of the azetidine under the harsh acidic conditions and high temperature (Route C). Another strategy was devised to use 1-(4-bromo/iodophenyl)-azetidine in a metal-halogen exchange with organolithium or Grignard reagents before addition to the anhydride (Route D). Both Claire Deo and I attempted multiple procedures using various organometallic reagents, with and without an additional transmetallation step to a magnesium or lanthanide species. Although some minor product could be formed (isolated yield < 10%), additional substitution predominantly occurred on the bottom ring. The next approach was to lithiate the methyl or *t*-Bu ester of tetrafluorobenzoic acid, with the acidity of the *o*-hydrogen being sufficient for exchange due to the strong electron-withdrawing effect of the fluorines, before addition to the corresponding diphenylketone (Route E). Again, the yield was too low, so I hoped to improve the exchange by iodinating the substrate but the protected ester product could not be formed. Fortunately, I received help from Dr. Jon Grimm (Lavis lab, Janelia Research Campus), with their now published lactol condensation method,[168] which resulted in yields of ~75% for the different azetidine Malachite Green lactones, **15, 16**, **17, 18** (Figure 28). I found that the reaction is extremely sensitive to impurities and requires recrystallisation of the acid starting material and pure benzophenones for good yields.

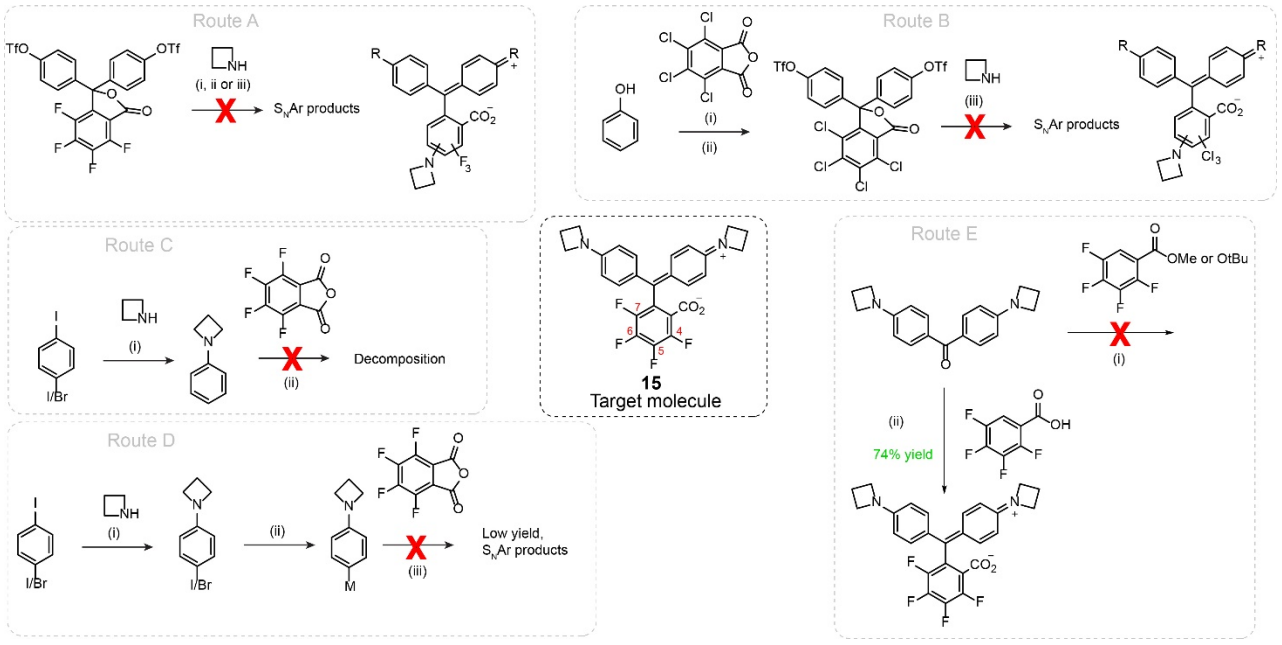


Figure 27: Synthetic routes attempted for preparation of 15. Route A: Palladium-catalysed Buchwald cross-coupling with azetidine; (i)[167] $Pd_2(dba)_3$, XPhos, Cs_2CO_3, dioxane, 100 °C; (ii)[179] $Pd_2(dba)_3$, XPhos, K_3PO_4, THF, 80°C; (iii) dioxane, rt. **Route B**: synthesis of tetrachlorinated ditriflate precursor, and subsequent cross coupling, followed by the Halex process; (i)[177] $MeSO_3H$, 90°C; (ii)[167] Tf_2O, pyridine, CH_2Cl_2; (iii)[167] $Pd_2(dba)_3$, XPhos, Cs_2CO_3, dioxane, 100°C. **Route C**: Friedel-Crafts acylation of 1-phenylazetidine with tetrafluorophthalic anhydride; (i)[169] azetidine, CuI, K_3PO_4, 1-butanol, ethylene glycol; (ii)[180] $ZnCl_2$, 150°C. **Route D**: metal-halogen exchange, transmetallation to magnesium or lanthanum, and electrophile addition to tetrafluorophthalic anhydride; (i)[169] azetidine, CuI, K_3PO_4, 1-butanol, ethylene glycol; (ii)[169, 181, 182] *t*-BuLi, MgI_2 or $(n\text{-Bu})_2$-*i*-PrMgLi alone, -78°C; *t*-BuLi, -78°C then $MgBr_2.OEt_2$, -10°C or $LaCl_3.2LiCl$, -78°C; all in THF; (iii) tetrafluorophthalic anhydride, THF, -78°C–-10°C. **Route E**: lactol condensation with tetrafluorobenzoic acid derivatives; (i) methyl/ *t*-butyl-2,3,4,5-tetrafluorobenzoate, *n*-BuLi, THF, -78°C; (ii) tetrafluorobenzoic acid, *n*-BuLi, THF, -78°C.[168]

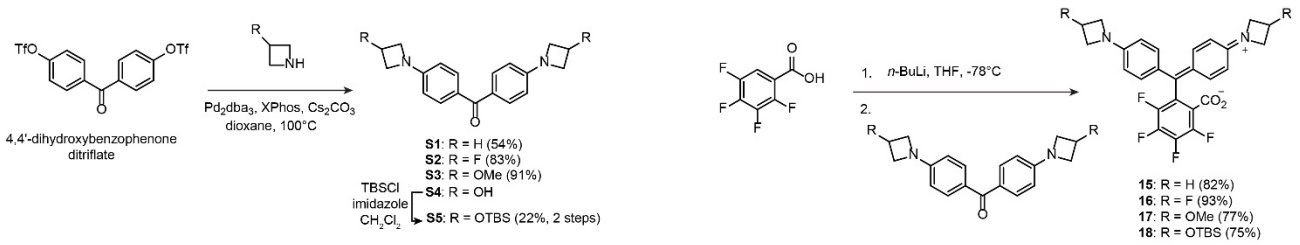


Figure 28: Synthesis of Malachite Green lactone derivatives 15, 16, 17, 18 with azetidine groups. Cross-coupling[167] of azetidine derivatives with the diphenylketone triflate precursor before the recently developed lactol condensation method with tetrafluorobenzoic acid.[168]

Xanthene core and bottom ring modification

Another strategy to close the equilibrium involved modifying the xanthene core directly, through Friedel-Crafts acylation of aniline derivatives with tetrafluorophthalic anhydride (Figure 29). I incorporated fluorines at the 2', 7' positions due to their *meta*-directing effect on the electrophilic aromatic substitution reaction, and methyl groups at the 1', 8' positions because of their *ortho*-selectivity (Figure 29a). Unsurprisingly, the deactivating effect of the fluorines resulted in very low yields but enough for characterisation. Moving to the lower half of the scaffold, another approach, used in the MaP dye series,[148, 183] replaces the 3-carboxylic acid with electron-deficient amides, which have greater propensity for forming the closed, spirolactam form. As this has quite a large effect on the equilibrium, I chose the most electron-withdrawing substituent, which closes the equilibrium the least, cyanobenzenesulphonamide, to couple to **14** (Figure 29b). To protect against potential S_NAr with the sulphonamide, I also synthesised the 6-*t*-Bu ester derivative (**S22**) but nevertheless the reaction failed. The electron-withdrawing nature of the tetrafluorinated ring appears to strongly deactivate the carboxylic acid for amide coupling. In terms of the equilibrium, the tetrafluorination strongly opens the dye so I was interested to see the effect of incomplete fluorination. When synthesising the HaloTag ligand of **14**, I fortuitously extracted the trifluorinated molecule **21** as a by-product (Figure 29c). With these Malachite Green lactone derivatives in hand, their absorption properties could then be evaluated.

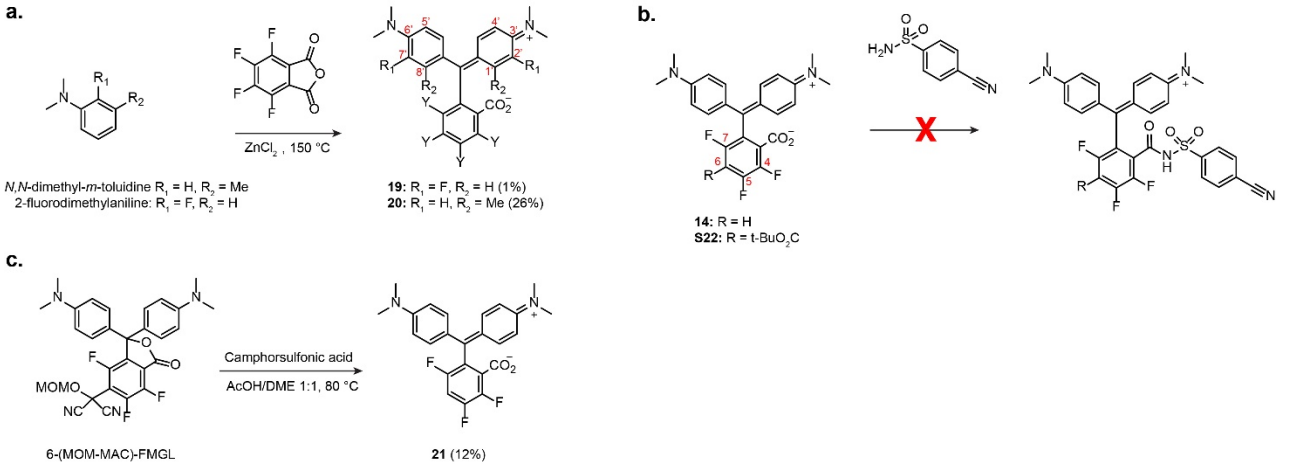
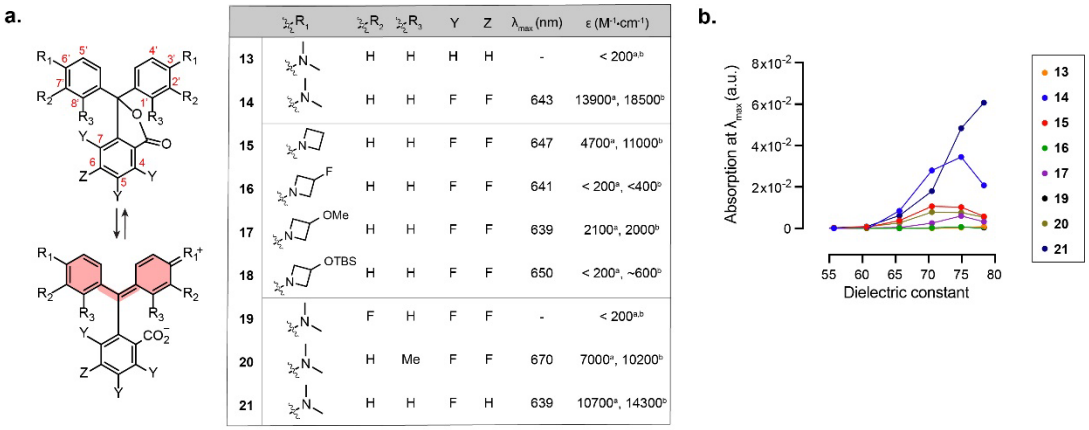


Figure 29: Synthesis of Malachite Green lactone derivatives 19-21 and attempted MaP synthesis. (**a**) Friedel-Crafts acylation of aniline derivatives with tetrafluorophthalic anhydride to form **19** and **20**. (**b**) Activation of the *o*-carboxylic acid on the fluorinated ring could not enable coupling to the sulphonamide. (**c**) Decarboxylation of MOM-MAC derivative forms trifluorinated compound **21**.

II.1.4.2 Spectroscopic characterisation

The original Malachite Green lactone was too closed but fluorination successfully brought this scaffold into the chromogenic range with ε = 13900 $M^{-1}·cm^{-1}$ for **14** (Figure 30, S3). However, I was unable to fully open this scaffold in either acidic or basic conditions to determine the maximum absorption (Figure S4). This prevented the use of $K_{L\text{-}Z}$ to quantify the open-closed equilibrium and compare with the other families hence the choice of the D_{50} method with MeCN/H_2O.[145, 170] However, quantitative D_{50} values could not be determined because either the absorbance of such predominantly closed, chromogenic dyes was too low or aggregation occurred, as observed for the Si-rhodamines.

The azetidine-functionalised molecules presented a similar trend in extinction coefficients to their Si-rhodamine equivalents. Incorporation of the azetidine rings closes the dye and the presence of electron-withdrawing substituents in the 3-position on the azetidine results in further closure. However, the Malachite Green scaffold is generally more closed than Si-rhodamines, which means the perfectly tuned **JF635** with 3-fluoroazetidine, becomes too closed with dye **16**. The extinction coefficients of dyes **15** (4700 $M^{-1}·cm^{-1}$) and **17** (2100 $M^{-1}·cm^{-1}$), predict improved chromogenicity compared to the parent compound **14** (Figure 30a, S3). Fluorination of the xanthene core strongly closed the equilibrium so chromogenicity was lost in **19** (ε < 200 $M^{-1}·cm^{-1}$, Figure 30a, S3). Methylation, which has not been used to tune the open-closed equilibrium of dyes before, halved ε for **20**, successfully closing the equilibrium, whilst providing a desirable 30 nm bathochromic shift (Figure 30a, S3). There is precedent for this modification resulting in small wavelength shifts, in both directions depending on a combination of steric and electronic factors, for di- and tri-arylmethane dyes.[184] The partial defluorination, another novel strategy, was successful in reducing absorbance (ε = 10700 $M^{-1}·cm^{-1}$ for **21**). The MeCN/H_2O measurements confirmed the trend in extinction coefficients and the chromogenicity candidates via small absorption peaks (**14, 15, 17, 20, 21**) but also highlighted aggregation at ≥ 90% water (Figure 30b). The extinction coefficients for this family were also measured in buffer with 10 mM SDS additive to resolve aggregation, which confirmed the measured trends in equilibrium, without significantly opening up the dyes (Figure 30a).



	R_1	R_2	R_3	Y	Z	λ_{max} (nm)	ε ($M^{-1}·cm^{-1}$)
13		H	H	H	H	-	< 200[a,b]
14		H	H	F	F	643	13900[a], 18500[b]
15		H	H	F	F	647	4700[a], 11000[b]
16	F	H	H	F	F	641	< 200[a], <400[b]
17	OMe	H	H	F	F	639	2100[a], 2000[b]
18	OTBS	H	H	F	F	650	< 200[a], ~600[b]
19		F	H	F	F	-	< 200[a,b]
20		H	Me	F	F	670	7000[a], 10200[b]
21		H	H	F	H	639	10700[a], 14300[b]

Figure 30: Malachite Green lactones as acoustogenic dyes. (**a**) General structure and open-closed equilibrium of the synthesised Malachite Green lactone dyes. (**b**) Absorption properties; all measurements were made in 10 mM HEPES, pH = 7.4. Extinction coefficients were measured in [a] 10 mM HEPES, pH = 7.4 and [b] 10 mM HEPES, pH = 7.4 with 10 mM SDS additive. Values are the mean of 3 replicates. (**c**) Absorption at λ_{max} of free dyes **13-21** in MeCN/H_2O mixtures.[171, 172] All measurements were performed at 5 µM dye concentration, values are the mean of 3 replicates. Adapted from ref.[165]

After characterisation of the absorption properties of this dye library, I had a shortlist of acoustogenic dyes, selected for their minimal quantum yields and far-red extinction coefficients in the chromogenic range (200 < ε < 15000 $M^{-1}·cm^{-1}$), for synthesis of the HaloTag ligand and evaluation of their labelling kinetics and acoustogenicity with the HaloTag protein. These promising candidates include the fluorinated *N*-aryl Si-

rhodamines **2** and **4**, the unsymmetrical Si-rhodamine **5**, as well as nonfluorinated **6** and **7** with *N*-methylpiperazine and dihydroquiniline auxochromes respectively, the fluorinated pyrrole-xanthene **10**, and the Malachite Green lactones **14**, **15**, **17**, **20**, and **21**.

II.2 Synthesis and characterisation of HaloTag ligands as photoacoustic labels

II.2.1 Synthesis

The candidate acoustogenic dyes ($200 < \varepsilon < 15000$ $M^{-1}\cdot cm^{-1}$) were synthesised with the HaloTag ligand attached in the 6-position on the pendant phenyl ring. For the nonfluorinated ligands (**5-**, **6-**, **9-**, **11-HTL**), established synthetic routes[185] were used to make the 6-*t*-Bu ester precursor before deprotection and amide coupling to form the HTL (Figure 31, S5). However, the fluorinated compounds were more challenging.

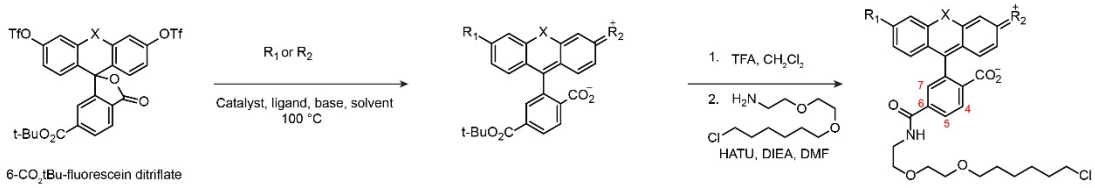


Figure 31: General synthesis of nonfluorinated Si-rhodamine and pyrrole-xanthene HaloTag ligands. Cross-coupling of amines or boronic esters with 6-CO_2-*t-Bu*-Si-fluorescein ditriflate precursors, followed by acid deprotection and amide coupling to form the final HTL products.

The synthesis of the HaloTag ligands of the fluorinated compounds required a different approach. Masked acyl cyanide (MAC) chemistry is a recently reported method for rhodamine functionalisation which enables regioselective substitution of a single fluorine atom on the tetrafluorinated bottom ring of rhodamines for a carboxyl group using 2-(methoxymethoxy)malononitrile as an umpolung acyl anion equivalent.[166, 186] The intermediate can be converted to the HTL amide (Figure 32, Route A) through TFA removal of MOM, forming a gem-dicyano alcohol. Direct amide coupling installs the HTL, in a base-mediated cyanide elimination, via the reactive acyl cyanide intermediate. In my hands however, the acidic conditions used for removal of the MOM group did not provide the expected compound, and attempts to form the HTL using this method were unsuccessful, for unknown reasons. Therefore, I tried several alternative routes to circumvent this issue with limited success (Figure S6). However, then I synthesised alternative MAC reagents bearing different protecting groups, and assessed their regioselectivity with tetrafluorinated substrates (Figure 32, Route B). From the three reagents attempted (tetrahydropyran (THP)-MAC, *t*-butyldimethylsilyl (TBS)-MAC and ethyl ether (EE)-MAC), THP-MAC had the highest yield in preparation, and showed an identical reactivity pattern to MOM-MAC, with S_NAr principally on the 6-position. This route enabled synthesis of the remaining fluorinated ligands **10-**, **15-**, **16-**, **17-**, **20-**, **22-**, **23-HTL** (Figure 33, S7).

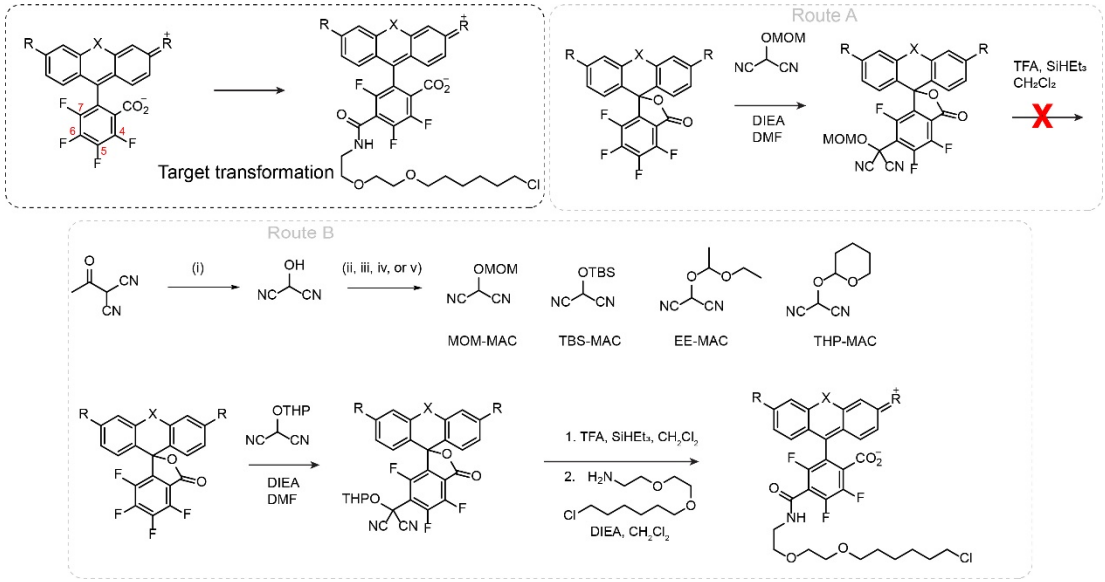


Figure 32: Synthesis of fluorinated rhodamine HaloTag ligand derivates. Route A: the established MAC chemistry substitution, deprotection and amide coupling; (i) MOM-MAC, DIEA, DMF; (ii) TFA, $SiHEt_3$, CH_2Cl_2 then HaloTag ligand, DIEA, CH_2Cl_2.[166] **Route B**: synthesis and substitution with alternative MAC reagents; (i) peracetic acid, acetic acid, H_2O; (ii)[187] dimethoxymethane, phosphorus pentoxide, CH_2Cl_2; (iii)[187] *t*-butyldimethylsilyl chloride, imidazole, DMF; (iv)[188] ethylvinyl ether, *p*-toluenesulfonic acid, toluene; (v)[188] 3,4-dihydro-2H-pyran, *p*-toluenesulfonic acid, toluene.

The remaining difluorinated compound, **23-HTL**, is the HTL equivalent of **21**, and was synthesised via a new method which I developed. Starting from **14**, I synthesised the 6-methyl ester precursor using the THP-MAC chemistry via coupling with methanol after the deprotection, in place of HTL-NH_2 (Figure 33). Alternatively, from the MOM-MAC intermediate, deprotection with H_2SO_4 formed the carboxylic acid, which could be esterified with *N*,*N*-dimethylformamide di-*tert*-butyl acetal to form the *t*-Bu ester (Figure S7f). Addition of $NaBH_4$ mediated regioselective hydrogenation of the fluorine in the 4-position as the major isolated product (~90%), with the 5-position being the minor product. The regioselectivity of this reaction has been determined by X-ray crystallography (see section VI.2). The crystal structure of the 6-methyl ester compound **S21** revealed a 9:1 split in electron density of the fluorine across the carbons at positions 4 and 5 providing definitive evidence of the regioselectivity. I also tested the defluorination directly on **14-HTL** which was less selective and seemed to progress to the monofluorinated and nonfluorinated species. This indicates that the nature of the substitution on the 6-position has a large effect on the regioselectivity, which could be worth exploring to expand defluorination as a tuning method for the open-closed equilibrium of rhodamine derivatives.

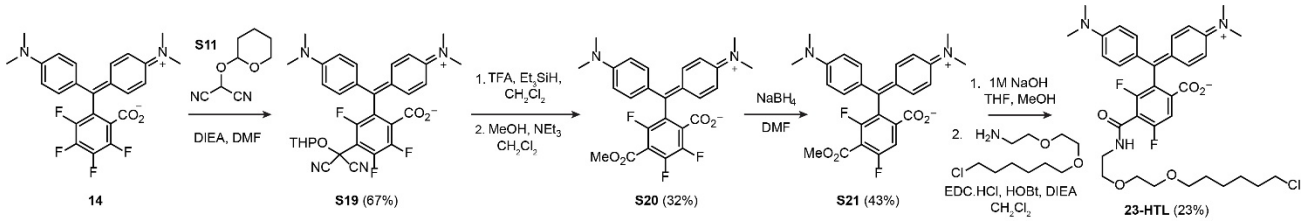


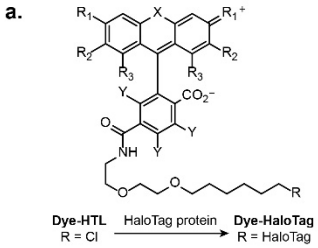
Figure 33: Synthesis of defluorinated Malachite Green lactone HaloTag ligand 23-HTL. Regioselective substitution with THP-MAC and the resulting intermediate is deprotected and coupled with methanol to form the methyl ester. Regioselective defluorination with $NaBH_4$, and subsequent deprotection and amide coupling forms the difluorinated HTL product.

With the synthesis of this library of acoustogenic HaloTag ligands (Figure S2) complete, I started to characterise their key photophysical properties for application as photoacoustic labels and calcium sensors.

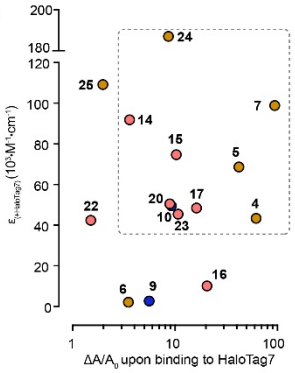
II.2.2 Absorption characterisation

To characterise the labelling kinetics and chromogenicity with HaloTag protein, the absorption of these ligands was measured with and without the protein and monitored over time. Every compound showed chromogenicity to differing degrees, with absorption increasing upon HaloTag binding, due to the shift in the equilibrium towards the open form (Figure 34, S8). This validated our approach to select chromogenic free dyes based on their absorbance within the range $200 < \varepsilon < 15000$ $M^{-1}{\cdot}cm^{-1}$, with lower absorbance increasing turn-on with HaloTag above a ~200 $M^{-1}{\cdot}cm^{-1}$ threshold (Figure S9). To assess binding rates, we used a faster, semi-quantitative approach by assuming pseudo-first order kinetics and determining a rate constant by fitting an exponential model to the increase in absorption over time.

a.



b.



c.

Compound	Protein	λ_{max} (nm)	λ_{em} (nm)	ε ($M^{-1}{\cdot}cm^{-1}$)	Φ	$\Delta A/A_0$	$\Delta PA/PA_0$	k / 10^{-3} s^{-1}
4-HTL	+HaloTag7	683	705	44300	< 0.01	62	-	0.19
5-HTL	+HaloTag7	654	679	68800	0.01	42	-	4.9 (fast), 0.81 (slow)
6-HTL	+HaloTag7	648	658	3000	0.11	3.5	-	9.3
24-HTL	+HaloTag7	687	701	185000	0.26	9.5	-	12
25-HTL	+HaloTag7	676	691	109000	0.37	2.0	-	10
7-HTL	+HaloTag7	726	749	99000	0.13	94	-	0.85
9-HTL	+HaloTag7	655	710	3600	< 0.01	5.6	-	6.3
10-HTL	+HaloTag7	695	742	50300	< 0.01	9.2	6.9	1.4
	+HaloTag9	700	-	86300	-	18	8.6	-
14-HTL	+HaloTag7	646	666	92400	0.001	3.6	2.8	7.6
	+HaloTag9	648	-	114000	-	4.6	2.9	-
15-HTL	+HaloTag7	648	662	75100	< 0.01	10	8.1	8.9
	+HaloTag9	650	-	81600	-	11	-	-
16-HTL	+HaloTag7	638	650	11400	< 0.01	21	-	10
17-HTL	+HaloTag7	643	661	48800	< 0.01	16	12	12
	+HaloTag9	645	-	53200	-	17	-	-
20-HTL	+HaloTag7	672	713	50700	< 0.01	8.9	5.7	33
	+HaloTag9	670	-	76800	-	13	-	-
22-HTL	+HaloTag7	644	658	43100	< 0.01	1.5	1.0	5.0
23-HTL	+HaloTag7	640	656	45900	< 0.01	11	5.7	11
	+HaloTag9	641	-	50500	-	13	-	-

Figure 34: Absorption properties of acoustogenic ligands with HaloTag protein. (**a**) General structure of synthesised HaloTag ligands and formation of the ligand-HaloTag adduct. (**b**) Plot of extinction coefficient of the HaloTag-bound ligand against absorbance turn-on upon binding for the HaloTag ligands. "HTL" was removed from the labels for clarity. Yellow: Si-rhodamines, blue: pyrrole-xanthenes, red: Malachite Green derivatives. The dashed box highlights compounds with the desired combination of high chromogenicity and HaloTag-bound absorption. (**c**) Table of properties of the HaloTag ligands bound to HaloTag protein. HaloTag7 is the HaloTag protein variant used unless otherwise specified. All measurements were performed in 10 mM HEPES, pH 7.4, containing 0.1 $mg{\cdot}mL^{-1}$ of CHAPS. Values are the mean of 3 replicates.

II.2.2.1 Silicon rhodamines

In the Si-rhodamine family, **4-HTL** displayed a high absorbance turn-on ($\Delta A/A_0 = 62$), however, the binding kinetics were very slow, requiring >24 h to reach the maximum absorption (Figure 35a, b). This is too slow for labelling *in vivo*, as the dye would likely be cleared out of the bloodstream before saturating the HaloTag protein target. Clashes of the bulky phenyl rings on the protein surface are likely responsible, so I hypothesised that the unsymmetrical ligand **5-HTL**, may bind faster. Indeed, **5-HTL** has faster binding kinetics, whilst maintaining the negligible fluorescence (Φ ~0.01) and high chromogenicity ($\Delta A/A_0 = 42$), but still remains too slow (Figure 35a, b). In order to confirm that it is definitely the binding which is slow and not slow opening of the equilibrium on the protein, I performed a competition assay for **4-HTL** and **5-HTL** with the fully open dye **JF549-HTL**, adding the competing ligand after 50% of the maximum absorption had been reached. The plateau in absorption at λ_{max} for the Si-rhodamines confirmed that the slow kinetics observed are due to dye-ligand binding (Figure

35c). Piperazine ring-based quenching groups increased the labelling speed but **6-HTL** possesses a too closed equilibrium, so the HaloTag-bound absorption remains low (Figure 34, S8). The fluorinated piperazine and morpholino substituted Si-rhodamines (Figure S2) similarly displayed rapid labelling kinetics and chromogenicity, with **24-HTL** standing out with an exceptionally high extinction coefficient for a rhodamine (Figure 34, 35, S8). Unfortunately, the PeT quenching from these groups is not sufficient for PA application, with high quantum yields when conjugated to HaloTag ($\Phi > 0.10$). Finally, the dihydroquinoline derivative **7-HTL**, showed the highest chromogenicity of the series ($\Delta A/A_0 = 94$; Figure 34, 35a, S8). However, the bulky fused ring structures also slowed the binding kinetics, to a similar range to the unsymmetrical **5-HTL**, although this ligand can label mammalian cells in 30 minutes of incubation (Figure 35b).[168] Nevertheless, this compound, as for the piperazine and morpholino Si-rhodamines, presents higher fluorescence when bound to HaloTag ($\Phi = 0.13$).

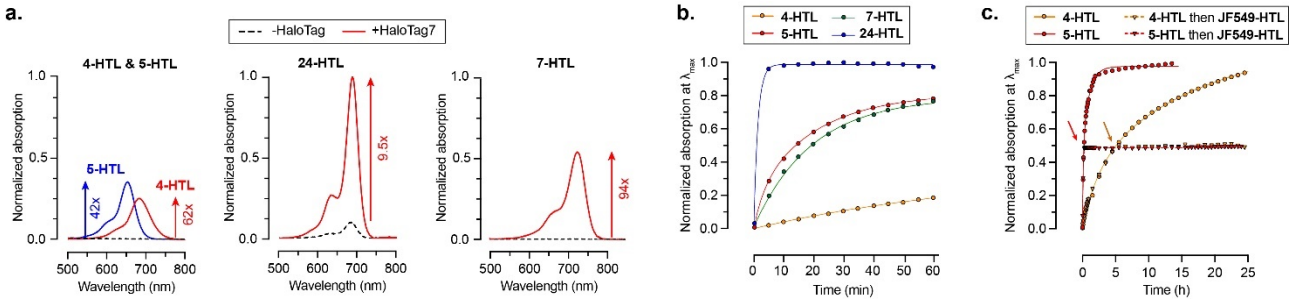


Figure 35: Chromogenicity and labelling kinetics of selected Si-rhodamine HTLs with HaloTag7 protein. (**a**) Normalised absorption spectra of ligands **4-HTL**, **5-HTL**, **24-HTL** and **7-HTL** ligands in the absence or presence of HaloTag7 protein; spectra normalised to the maximum absorption of **24-HTL**. (**b**) Binding kinetics of the same ligands to HaloTag7, with absorption normalised to each ligand's maximum. (**c**) Labelling kinetics of slow binding ligands **4-HTL** and **5-HTL**, including a binding competition assay with **JF549-HTL**, which was added in excess after absorption had reached ~50% of its maximum for each sample; the absorption at λ_{max} plateaus, proving that the slow kinetics are due to slow ligand binding. All measurements were performed in 10 mM HEPES, pH 7.4, containing 0.1 mg·mL^{-1} of CHAPS with 5 µM of HaloTag ligand and 7.5 µM of HaloTag7 protein.

II.2.2.2 Pyrrole-xanthene derivatives

For the pyrrole-xanthenes, the heteroaromatic rings were accommodated by the HaloTag binding pocket, resulting in fast labelling (Figure 34, 36b). **10-HTL** was the sole chromogenic candidate and displayed the expected absorbance turn-on ($\Delta A/A_0 = 9.2$), with the characteristic broad spectrum (Figure 36a). As the position of equilibrium was unknown for these compounds and the synthesis was straightforward, I had already made **9-HTL** and **11-HTL** before the free dye characterisation. When bound to HaloTag, **9-HTL** showed a small absorption increase but with a low extinction coefficient, and **11-HTL** remained completely closed with undetectable absorption (Figure S8).

II.2.2.3 Malachite Green lactone derivatives

The Malachite Green lactones proved to be excellent binding partners for HaloTag as well, with fast binding kinetics, reaching absorption maximum in < 5 minutes in our conditions, and showing absorbance turn-ons up to 21-fold (Figure 34, 36, S8). Crucially, their low quantum yields ($\Phi < 0.01$) were maintained bound to HaloTag, suggesting that they still retain conformational freedom in the HaloTag binding pocket, in contrast to fluorogen activating proteins, which induce a fluorescence turn-on for Malachite Green.[189] The quantum yield of **14**-HaloTag7 was measured more precisely using the relative method with a reference dye (Oxazine 1), to give a value of Φ ~0.001. The turn-on of the different dyes follows the trend observed for the equilibrium, with

the more closed dyes displaying lower protein-free absorption and higher chromogenicity (Figure 34). However, there is also a compensatory decrease in the HaloTag-bound absorption, particularly evident for the most closed ligands of the series, such as the fluorinated **16-HTL**, which prevents even greater turn-ons (Figure S8). In an attempt to reduce the compromise in signal observed with the more closed ligands, and increase the photoacoustic turn-on upon HaloTag binding, I evaluated a selection of our library with HaloTag9, a "brighter" variant which can shift particular rhodamine derivatives further towards the open form and hence increase extinction coefficients.[190] The azetidine-based ligands were largely unaffected by the two-point mutations introduced in this protein scaffold. However, there was a significant increase in absorption (up to 73%) for **10-HTL**, **14-HTL** and **20-HTL** (Figure 36a, S8). This suggests that with tailored protein engineering, involving site-saturated mutagenesis of the dye binding cavity and linkers, greater enhancement of absorption can be achieved with these ligands.

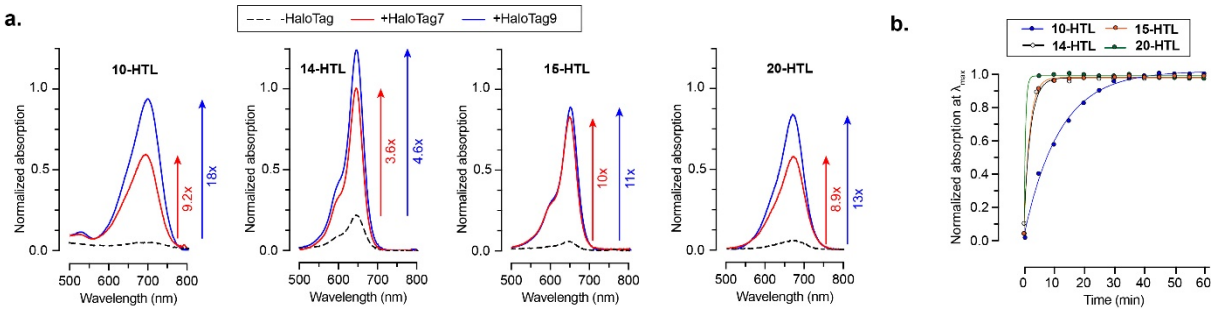


Figure 36: Chromogenicity and labelling kinetics of selected pyrrole-xanthene and Malachite Green lactone HTLs with HaloTag protein. (**a**) Normalised absorption spectra of ligands **10-, 14-, 15-** and **20-HTL** ligands in the absence or presence of HaloTag7 or HaloTag9 protein. Spectra normalised to the maximum absorption of **14-HTL** bound to HaloTag7. (**b**) Binding kinetics of the same ligands to HaloTag7 normalised to each ligand's maximum.

Given the hydrophobicity of these scaffolds, highlighted by the absorption measurements with the free dyes (Figure 30b), aggregation also occurred with the HaloTag ligands. When measuring these photophysical properties, I noticed a concentration-dependence on the absorption of the free ligand and hence on the $\Delta A/A_0$ values, with $\Delta A/A_0$ increasing at higher concentrations. This aggregation behaviour made measurement of an accurate extinction coefficient for the free ligands difficult. I selected 1.25 µM as an appropriate concentration for reporting the absorption properties, as **14-HTL** remained in the linear range whilst providing sufficient absorption above the sensitivity limit of the spectrometer (Figure 37a). Screening different detergents (CHAPS, Pluronic F127, Tween20, SDS) confirmed CHAPS as the most suitable for helping to maintain solubility without significantly affecting the dye equilibrium.[191] Running a time course of **14-HTL** in solution overnight, revealed slow aggregation, decreasing absorption by 43% after 17 h (Figure 37b). In comparison, **JF635-HTL** absorption reduced by 31%, highlighting that this is not unusual for such chromogenic dyes. Structural modifications can reduce this propensity for aggregation. **22-HTL**, with hydrophilic hydroxyl groups, maintained linearity between absorption and concentration until 4 µM (Figure 37c), however unfortunately the absorbance turn-on upon HaloTag binding was lower than for the other ligands (Figure 34, S8). More hydrophilic derivatives could be investigated with the polar groups further away from the xanthene core to minimise the impact on the equilibrium.

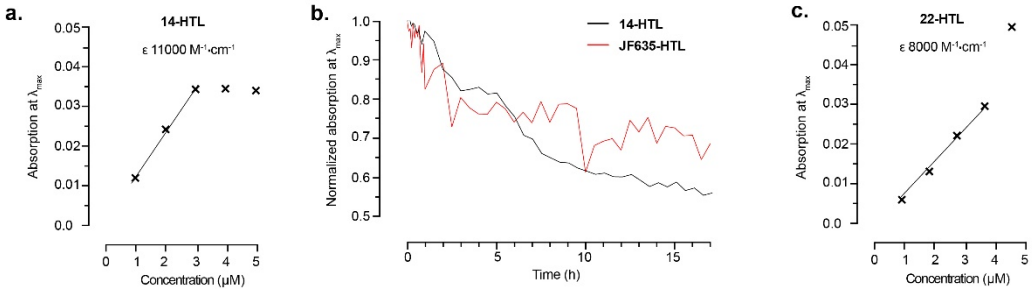


Figure 37: Investigating the aggregation of acoustogenic HaloTag ligands. (**a**) Absorption at 646 nm of **14-HTL** at 1–5 μM. (**b**) Normalised absorption time course of **14-HTL** and **JF635-HTL** at 1.25 μM. (**c**) Absorption at 644 nm of **22-HTL** at 1–5 μM. All measurements were performed in 10 mM HEPES, pH 7.4, containing 0.1 mg·mL^{-1} of CHAPS, in the absence of HaloTag7.

II.2.3 Photoacoustic characterisation

Next, we set out to measure the photoacoustic properties of these novel probes. In this screening pipeline, absorption is a suitable proxy for photoacoustic signal because I selected dyes with minimal quantum yields, but their photoacoustic properties need to be characterised, for which we used a multimodal spectrometer developed by Nikita Kaydanov (Prevedel group). Solutions of my probes were added to a glass cuvette which was submerged in a water tank and irradiated by a pulsed laser, from which acoustic signals were detected by a focused transducer (Figure 38a). The spectral features and peak turn-on values ($\Delta PA/PA_0$) with HaloTag protein were very similar for all of the ligands, validating the use of absorption as an initial approximation for photoacoustic signal (Figure 34, S9). For example, **14-HTL** bound to HaloTag7, showed identical peaks in absorption and PA, with $\Delta PA/PA_0 = 2.8$ (Figure 38b). In general, there was a consistent trend of increasing $\Delta PA/PA_0$ values with $\Delta A/A_0$ (Figure 38c), although the former was marginally lower, perhaps because of very minor increases in fluorescence between the free dye and protein-bound form due to conformational restriction. HaloTag9 bound ligands **10-HTL** and **14-HTL** also showed the expected higher PA signal thanks to their stronger absorption (Figure 34, S10).

High photostability is also a requirement for effective PA probes. Nikita adapted his spectrometer set-up into a well plate configuration so an entire well could be illuminated for homogenous photobleaching. Pulsed laser illumination for 1 h at peak wavelength, resulted in minimal bleaching of the PA signal of **14**-HaloTag7, compared to bleaching by 60-70% for Cy5, a spectrally matched common cyanine fluorophore, and the far-red fluorescent protein, mIFP (Figure 38d, S11).[129, 192] This further validates our choice of rhodamine derivatives as chromophores for PAI, and highlights the photostability advantage of using synthetic chromophores over their fluorescent protein counterparts.

In addition, to validate the proof-of-principle of using synthetically quenched rhodamine derivatives, we evaluated the PA signal of **14-HTL** against the bright, spectrally similar, fluorogenic **JF635-HTL** (Figure 38e).[145] When bound to HaloTag7 both **JF635-HTL** and **14-HTL** have comparable extinction coefficients, and therefore absorbance. However, the photoacoustic signal was 17-fold higher for **14**-HaloTag7 because of its minimal quantum yield (750-fold smaller than **JF635**-HaloTag7, $\Phi = 0.75$). This demonstrates that for strong photoacoustic signal from these rhodamine-based chromophores, severe fluorescence quenching is a necessity, which is achieved by our series of acoustogenic ligands, to function as photoacoustic labels.

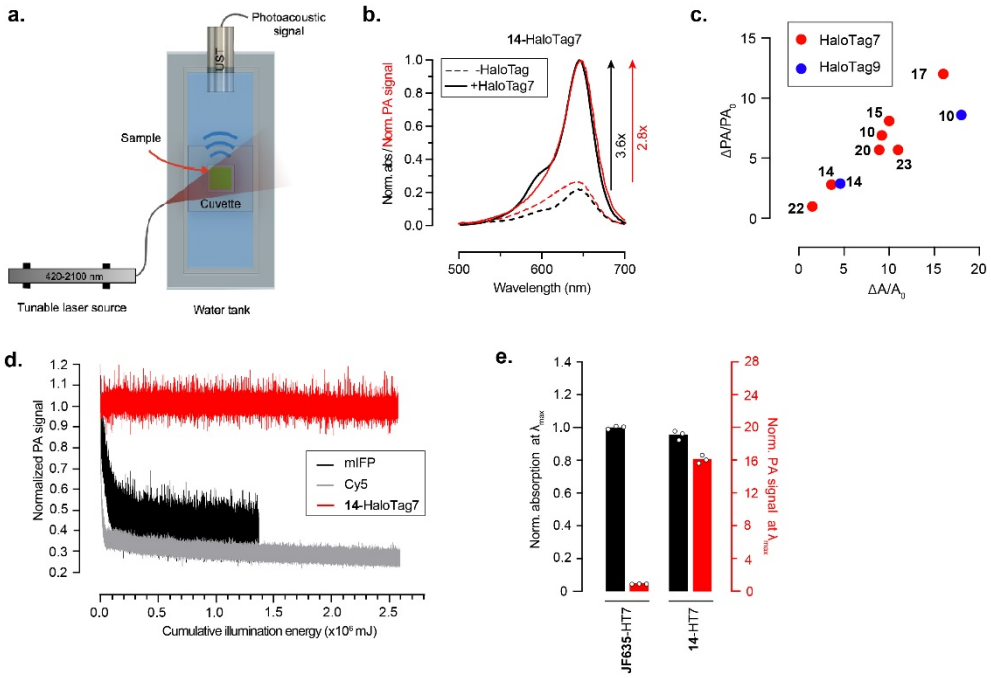


Figure 38: Photoacoustic characterisation of HaloTag conjugates. (**a**) Schematic of the custom-built photoacoustic spectrometer (from above): dye solution in a cuvette is placed inside a water bath for acoustic coupling; the sample is excited with a tuneable, pulsed laser and the PA signal generated recorded with an ultrasound transducer (UST). (**b**) Overlaid absorption and PA spectra of **14-HTL** with (solid lines) and without HaloTag7 protein (dashed lines). (**c**) Plot of $\Delta PA/PA_0$ vs $\Delta A/A_0$ for chosen HaloTag ligands in combination with HaloTag7 protein. "HTL" was removed from the labels for clarity. (**d**) Normalised PA signal of absorbance-matched samples, during irradiation at $\lambda_{PA} = \lambda_{max}$ for 1 h at 100 Hz pulse repetition rate, plotted against cumulative illumination energy. (**e**) Normalised absorption at λ_{max} (black) and normalised PA signal at $\lambda_{PA} = \lambda_{max}$ (red) for **JF635-HTL** and **14-HTL** bound to HaloTag7 (HT7); values were normalised to **JF635-HTL**. All measurements were performed in 10 mM HEPES, pH 7.4, containing 0.1 mg·mL^{-1} of CHAPS at a dye ligand concentration of 1.25 µM and protein concentration of 1.9 µM, in triplicate.

II.3 Design and characterisation of chemigenetic photoacoustic calcium sensors

With this selection of acoustogenic ligands, which can function as photoacoustic labels with HaloTag, the next stage was to apply them in our design for chemigenetic calcium sensors. The HaloTag-based HaloCaMP[143] and rHCaMP[148] proteins when combined with rhodamine derivatives can confer calcium-dependent absorption changes. When the calmodulin domain of these sensors binds calcium, the protein adopts a different conformation, which forms more favourable interactions with the open form of bound ligands, resulting in a shift in the equilibrium and a resultant increase in absorption (Figure 19c). By labelling these scaffolds with our acoustogenic ligands, we hoped to achieve absorption and photoacoustic turn-ons upon calcium-binding. To determine calcium sensitivity, I labelled the protein scaffolds with my acoustogenic ligands, and measured absorption and PA signal in their calcium-saturated and calcium-free states using Ca^{2+} or ethylene glycol-bis(β-aminoethyl ether)-*N,N,N′,N′*-tetraacetic acid (EGTA)-containing buffers. EGTA is a calcium chelator so sequesters any free calcium in solution. All of the ligands tested showed calcium-dependent responses with all three protein variants but the different properties highlight the tunability afforded by simply changing the dye ligand (Figure 39, S12, S13).

II.3.1 Calcium sensitivity

The most closed ligands selected after the evaluation on HaloTag (**17-HTL** and **23-HTL**) showed low extinction coefficients with both HaloCaMPs, even in the Ca^{2+}-saturated conformation (Figure 39a, S12). Despite their

higher quantum yields, I also tested the Si-rhodamines **7-HTL** and **24-HTL** with the HaloCaMP variants in absorption (Figure 39a, S12). **24-HTL** was calcium-insensitive, with the dye in the fully open form independent of calcium state with both HaloCaMP1a and HaloCaMP1b. **7-HTL** displayed only a small negative response with HaloCaMP1a. With the third protein scaffold, rHCaMP, all of our ligands performed less well with low extinction coefficients and small negative changes in absorbance upon calcium-binding, so we did not characterise their photoacoustic properties (Figure 39a, S12). Such a different response is not surprising given this sensor was engineered in parallel with a specific coumarin-benzopyrylium fluorophore,[193] so the dye-protein interactions are less influenced by the calcium-dependent conformational change. This type of customised protein engineering is crucial to improving the sensitivity and performance of these first-generation absorption-based calcium sensors.

a.

Compound	protein	λ_{max} (nm)	ε_{sat} ($M^{-1}\cdot cm^{-1}$)	$\Delta A/A_0$	$\Delta PA/PA_0$	K_d
24-HTL	HaloCaMP1a	687	169000	0.1	-	-
	HaloCaMP1b	689	168000	0.1	-	-
7-HTL	HaloCaMP1a	728	13900	-0.5	-	-
	HaloCaMP1b	725	64200	0.1	-	-
	rHCaMP	726	17600	-0.7	-	-
10-HTL	HaloCaMP1a	700	23600	1.4	1.8	6.9 mM
	HaloCaMP1b	704	36000	2.1	1.3	16 nM
	rHCaMP	693	5400	-0.5	-	-
14-HTL	HaloCaMP1a	646	69100	0.7	1.1	300 mM; 18 mM
	HaloCaMP1b	645	29300	3.3	3.1	17 nM
	rHCaMP	645	17000	-0.4	-	-
15-HTL	HaloCaMP1a	649	30500	0.8	1.5	4.8 mM
	HaloCaMP1b	650	15200	2.4	3.2	10 nM
	rHCaMP	648	7500	-0.3	-	-
17-HTL	HaloCaMP1a	643	13200	4.1	-	-
	HaloCaMP1b	642	4100	1.5	-	-
	rHCaMP	645	5300	-0.4	-	-
20-HTL	HaloCaMP1a	674	20900	1.0	2.0	980 µM; 10 mM
	HaloCaMP1b	675	15300	16	7.5	33 nM
	rHCaMP	670	5900	0.01	-	-
23-HTL	HaloCaMP1a	640	14700	0.9	-	-
	HaloCaMP1b	638	7300	0.2	-	-
	rHCaMP	640	6100	-0.4	-	-
NIR-GECO1[a]		678[a]	20000[a]	-0.68[a]	-0.5	215 nM[a]
GCaMP6f[b]		496[b]	61900[b]	28.9[b]	3.8[c]	375 nM[b]
CASPA_550[d]		550[d]	77745[d]	~-0.5[d]	~-0.5[d]	~4 µM[b]
"L"[e]		765[e]	194000[e]	~-0.5[e]	~-0.5[e]	11.3 µM[c]

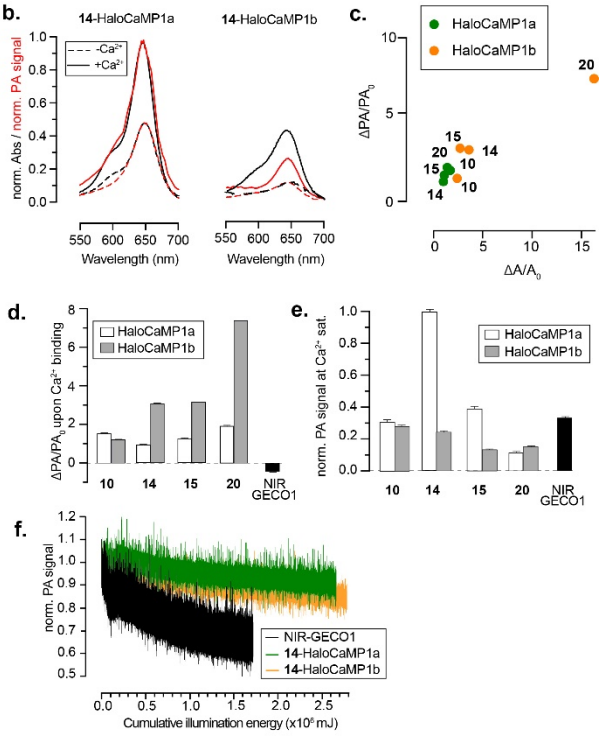


Figure 39: Characterisation of novel photoacoustic chemigenetic calcium sensors. (**a**) Photophysical properties of PA calcium sensors. All measurements were performed in 30 mM MOPS, 100 mM KCl, pH 7.2, containing 0.1 mg·mL^{-1} of CHAPS, to which either 10 mM EGTA (for the calcium-free state) or excess $CaCl_2$ was added (for the calcium-saturated state, i.e. 500 mM $CaCl_2$ for HaloCaMP1a or 10 mM $CaCl_2$ for HaloCaMP1b and rHCaMP). Values are the mean of 3 replicates. [a]Data from ref.[129] [b]Data from ref.[112] [c]Data from ref.[31] [d]Data from ref.[63] [e]Data from ref.[62] (**b**) Normalised absorption (black) and PA (red) spectra of **14-HTL** bound to HaloCaMP1a or HaloCaMP1b in the Ca^{2+}-free (dashed lines) and Ca^{2+}-saturated (solid lines) states; normalised to the Ca^{2+}-saturated state of **14**-HaloCaMP1a. (**c**) $\Delta PA/PA_0$ vs $\Delta A/A_0$ for selected calcium sensors upon binding Ca^{2+}; "HTL" was omitted from the labels for clarity. (**d**) $\Delta PA/PA_0$ upon binding calcium, measured at $\lambda_{PA} = \lambda_{max}$, for HaloCaMP1a and HaloCaMP1b labelled with **10-HTL**, **14-HTL**, **15-HTL** or **20-HTL** compared to NIR-GECO1. (**e**) Normalised PA signal in the Ca^{2+}-saturated state for the same sensors as in (**d**). (**f**) Photoacoustic signal of absorbance-matched samples of calcium sensors in the Ca^{2+}-saturated state during irradiation at $\lambda_{PA} = \lambda_{max}$ for 1 h at 100 Hz pulse repetition rate; normalised PA signal plotted against cumulative illumination energy.

Overall, four ligands (**10-HTL**, **14-HTL**, **15-HTL** and **20-HTL**) showed the best performance with the HaloCaMP variants in terms of PA signal intensity and calcium sensitivity (Figure 39a-d, S12, S13). In general, these ligands showed higher signal with HaloCaMP1a, and turn-ons in absorption and PA signal of 0.7-1.4 and 1.1-2.0 respectively. By contrast, the conjugates with HaloCaMP1b were more sensitive, displaying $\Delta A/A_0$ between 2.1-16 and $\Delta PA/PA_0$ between 1.3-7.5, with **20-HTL** resulting in the largest responses. As expected, the correlation in absorption and photoacoustic spectra and between $\Delta A/A_0$ and $\Delta PA/PA_0$ was conserved (Figure 39c). The trend observed across the different dyes follows the observation made with the Janelia Fluor ligands

developed alongside the HaloCaMP variants.[143] Dyes are more open in both Ca^{2+}-free and Ca^{2+}-saturated states bound to HaloCaMP1a, resulting in higher signal, whereas they remain more closed on HaloCaMP1b, but calcium-binding results in higher turn-ons. The extinction coefficients are lower for HaloCaMP-bound ligands than with HaloTag, highlighting that they do not reach the fully open state, which would require further protein engineering. Nevertheless, in comparison to the dim, far-red calcium sensor NIR-GECO1,[129] all of these 8 sensors are more sensitive, and are positive-going, which becomes even more advantageous in more complex imaging conditions with high background signal (Figure 39d). The existing synthetic PA calcium sensors also show lower sensitivity (Figure 39a).[62, 63] Whilst GCaMP6f shows larger dynamic range due to several generations of protein engineering to optimise the absorbance turn-on, the high quantum yield and blue excitation wavelength limit PA application to transparent samples and shallow depths.[31] The intensity of the Ca^{2+}-saturated **14**-HaloCaMP1a is also 3-fold greater than NIR-GECO1, and was the standout with >2-fold greater intensity than any other sensor combination (Figure 39e).

II.3.2 Photostability

In the same manner as for the labels, we investigated the photostability of these calcium sensors. Illumination for 1 h at λ_{max} reduced the photoacoustic signal of the Ca^{2+}-saturated **14**-HaloCaMP1a and **14**-HaloCaMP1b by only 10% and 15% respectively, whereas the far-red calcium sensor, NIR-GECO1, decreased by 35% (Figure 39f, S14). The relative decay of NIR-GECO1 and its parent chromophore mIFP match the reported values in fluorescence, validating our photostability measurements in PA spectroscopy.[129, 192] The spectroscopic measurements of PA signal only required very short exposure times and so our chemigenetic sensors should show even greater performance than in the spectroscopic measurements, when applied to biological systems with extended imaging periods, thanks to their superior photostability.

II.3.3 Calcium affinity

For the sensors to function at physiological neuronal calcium concentrations, they must have calcium affinities in the range experienced by activated neurons. The cytosolic calcium concentration in resting neurons is around 50–100 nM, which can rise over 1000-fold during an action potential.[9, 194] The affinity can be quantified by the dissociation constant, K_d, representing the concentration of calcium at which 50% of the maximum sensor response is achieved. As I have established absorbance as a suitable approximation for PA signal for these quenched ligands, I determined calcium affinities through CaEGTA titrations, measuring absorption as the readout. Surprisingly, the HaloCaMP1a sensors displayed low millimolar affinity, contrasting with the nanomolar range of the HaloCaMP1a-bound Si-rhodamine JF ligands (Figure 39a).[143] **14**- and **20**-HaloCaMP1a show biphasic behaviour (Figure 40a) with a higher affinity micromolar step (K_{d1} = 304 and 980 µM, respectively), then a weaker millimolar phase (K_{d2} = 18 and 10 mM, respectively). We can hypothesise that this behaviour results from independent association of the Ca^{2+} ions to the C-terminal (higher affinity) and N-terminal (lower affinity) domains of calmodulin, as was established for CaM-M13 pairs.[53] The first fluorescent protein-based calcium sensor, Cameleon-1 also shows this biphasic response along with the more recent FGCaMP, both of which use a CaM-M13 combination, like HaloCaMP1a.[143, 195, 196] Nonetheless, this does not explain the dye-dependent nature of this response, as **10**-HaloCaMP1a and **15**-HaloCaMP1a both present a single millimolar phase (Figure 40a). Different dye-protein interactions result in a conformation that facilitates cooperativity of calcium binding to the two terminal domains,[197] or alternatively the first higher affinity binding phase occurs, but does not translate into a response in absorption. A potential application of such millimolar sensors could be the monitoring of interstitial calcium dynamics at the intercellular or tissue level using PAM or PAT respectively, providing complementary information to the cytosolic activity reported by high affinity sensors.[198] Furthermore, the responses in the micromolar range are ideal for imaging calcium fluxes in organelles such as the ER.[199-201] Protein targets involved in the secretory pathway for neuronal growth and

maintenance, for example ERGIC53 and Rab51b, could also be tracked, with variations in intensity from calcium fluctuations as they move between compartments.[198, 202] However, for recording of cytosolic neuronal activity, these are currently unsuitable and will require protein engineering to increase their affinity.

By contrast, combining the same ligands with HaloCaMP1b results in high affinity sensors with nanomolar K_d values (10–33 nM; Figure 39a, 40b) in the same range as the original HaloCaMP sensors,[143] and jGCaMP7s and jGCaMP7b which can be used for a variety of *in vivo* neuroimaging applications from monitoring activity in neurites and neuropils to whole neuronal populations.[113] The strong affinity means that the sensor could become saturated at low numbers of action potentials so would also benefit from tuning of affinity to ~100−200 nM to maximise the observable dynamic range in physiological conditions.

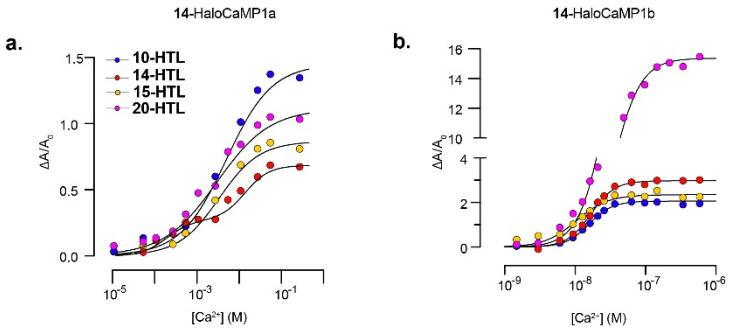


Figure 40: Ca^{2+} affinity of acoustogenic calcium sensors. Ca^{2+} titrations of (**a**) HaloCaMP1a and (**b**) HaloCaMP1b labelled with **10-HTL**, **14-HTL**, **15-HTL** or **20-HTL**; measurements were performed in duplicate, in 30 mM MOPS, 100 mM KCl, pH 7.2 via titration of two solutions containing 10 mM K_2EGTA or 10 mM CaEGTA.

II.3.4 Calcium binding and unbinding kinetics

Action potentials and the resultant spikes in cytoplasmic calcium concentration are rapid, transient events which occur in the order of milliseconds.[162, 194] Therefore, calcium sensors must be capable of fast responses to calcium to accurately report on neural activity, although those with protein-sensing domains are intrinsically limited by the timescale of the conformational changes.[203] To measure both the calcium binding and unbinding kinetics of our novel sensors using a stopped-flow spectrophotometer, we recorded absorbance changes upon mixing solutions containing Ca^{2+} or the EGTA chelator, and compared against existing indicators. The slower decay on calcium dissociation from the sensor is usually the limiting factor for resolution of action potentials, and indeed **14**-HaloCaMP1b presented quite slow unbinding, ~10-fold slower than **JF635**-HaloCaMP1b and NIR-GECO1 (Figure 41a, S15). Both of these sensors can reliably be used for functional imaging in cultured neurons so **14**-HaloCaMP1b is already close to the required range.[129, 143] The very strong affinity of **14**-HaloCaMP1b likely contributes to the slower decay, so subtle affinity tuning would also improve the dissociation speed. Interestingly, the unbinding kinetics of **JF635**-HaloCaMP1b in absorbance were faster than in fluorescence (k_{off} = 0.43 s^{-1}).[143] Both **JF635**-HaloCaMP1a and NIR-GECO1 by contrast showed very similar rates to the reported values so these differences should be explored further to determine if this is a probe specific, or more general effect (Figure S15a, b).

14-HaloCaMP1b and **JF635**-HaloCaMP1b both displayed fast calcium binding, although our acoustogenic calcium sensor was 3-fold slower (Figure 41b, S15c). Importantly though, **14**-HaloCaMP1b is in a similar range to the GCaMP6 and jGCaMP7 series, which have been widely used for *in vivo* brain recordings, so should be performant when applied to measuring transients in neurons.[113]

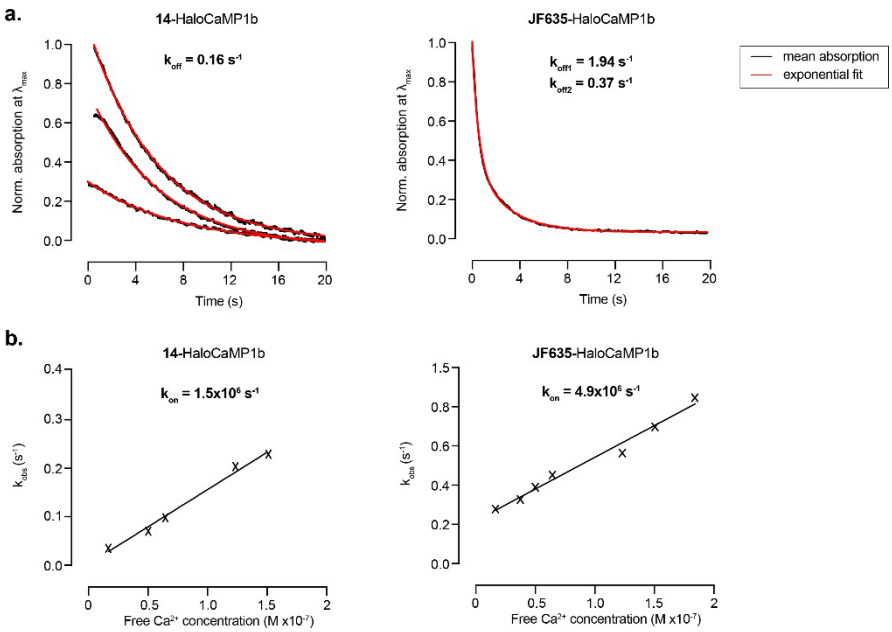


Figure 41: Ca^{2+} binding and unbinding kinetics of acoustogenic calcium sensors. (**a**) Calcium dissociation kinetics of **14**-HaloCaMP1b and **JF635**-HaloCaMP1b (black curves); sensors in 30 mM MOPS, 100 mM KCl, 100 µM $CaCl_2$, pH 7.2 were mixed with high EGTA-containing buffer (30 mM MOPS, 100 mM KCl, 10 mM EGTA, pH 7.2), and k_{off} determined by fitting with exponential functions (red curves); 3 concentrations of sensors for **14**-HaloCaMP1b, 1 concentration for **JF635**-HaloCaMP1b; plots shown are the mean of at least 4 runs. (**b**) Calcium binding kinetics of **14**-HaloCaMP1b and **JF635**-HaloCaMP1b; determination of k_{on} from k_{obs} values for each calcium concentration from Figure S15c.

II.3.5 Calcium selectivity and pH stability

To complete the sensor characterisation, I used **14-HTL** as a representative ligand to evaluate the calcium selectivity and pH stability of these sensors. Another divalent cation, Mg^{2+}, can be present at millimolar concentrations in the cytoplasm and compete for calcium-binding sites.[204] **14**-HaloCaMP1a and **14**-HaloCaMP1b showed minimal responses to Mg^{2+} and the presence of Mg^{2+} did not affect sensor function on subsequent addition of Ca^{2+} (Figure 42a, b).

Maintaining response function across the physiological pH range of the neuronal cytosol is also a key parameter, as the resting pH can vary between neurons (e.g. 6.3−7.7 for hippocampal neurons), and fluctuations occur during electrical activity.[205] The function of **14**-HaloCaMP1a and **14**-HaloCaMP1b was conserved in this relevant cytosolic range (Figure 42c, d). There was a loss of sensitivity below pH 6, however, the behaviour when the proteins were labelled with **JF635-HTL** was very similar, which had not previously been characterised (Figure 42c, d). **JF635**-HaloCaMP1a/b performs well in cultured neurons,[143] suggesting that this pH sensitivity at lower pH does not impede function.

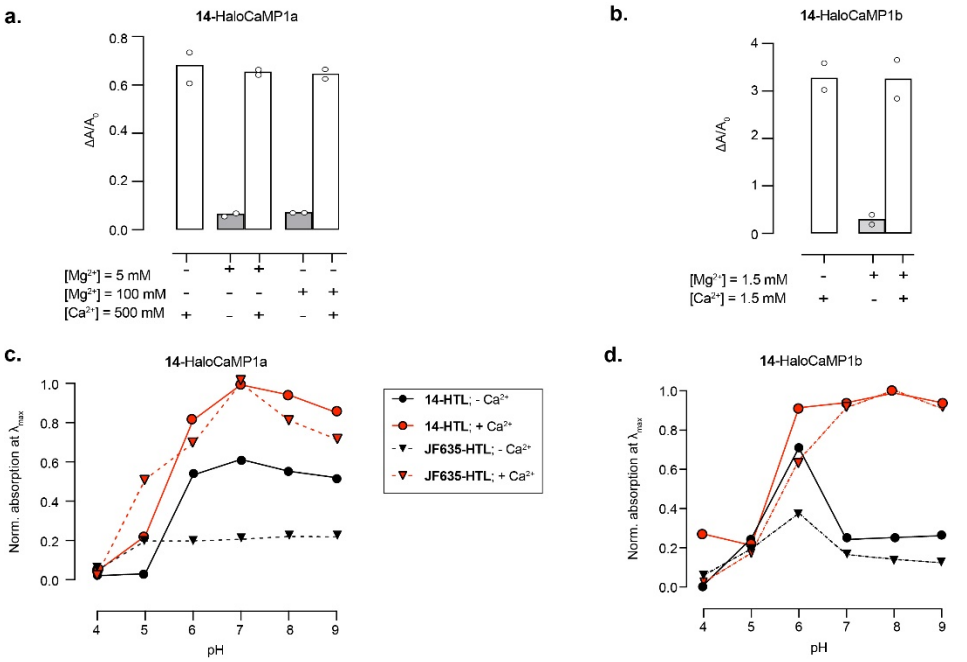


Figure 42: Ca^{2+} selectivity and pH stability of acoustogenic calcium sensors. (**a-b**) Selectivity of sensors (**a**) **14**-HaloCaMP1a and (**b**) **14**-HaloCaMP1b for Ca^{2+} over Mg^{2+}; measurements were performed in 30 mM MOPS, 100 mM KCl, pH 7.2 containing EGTA (10 µM for HaloCaMP1a, 10 mM for HaloCaMP1b), in duplicate. All measurements were performed at 1.25 µM dye and 1.9 µM protein. (**c-d**) pH sensitivity of calcium sensors: (**c**) **14**-HaloCaMP1a and (**d**) **14**-HaloCaMP1b; absorbance at λ_{max} measured in citrate, phosphate and Tris buffers in the absence (black) or presence (red) of calcium, and compared with **JF635**-labelled sensors; measurements were performed in duplicate.

In summary, combining the four acoustogenic ligands **10-, 14-, 15**- and **20-HTL** together with the two HaloCaMP variants, gives eight calcium sensors optimised for PAI. These span across the full physiological Ca^{2+} concentration range with affinities from nanomolar to millimolar, provide strong photoacoustic signal intensities and have superior sensitivities to current indicators such as NIR-GECO1 and CaSPA.

II.4 Conclusion and perspectives

This chapter has described the design and development of the first chemigenetic labels and calcium sensors for photoacoustic imaging. First, I synthesised a library of >20 dyes across three families, Si-rhodamines, pyrrole-xanthene derivatives and Malachite Green lactones. These dyes were systematically evaluated for the key photophysical properties of high far-red/NIR absorption, minimal fluorescence quantum yield and chromogenicity, exploiting the open-closed equilibrium of rhodamine derivatives. I introduced rational chemical modifications in order to improve these properties before assessment of the chromogenicity and labelling kinetics upon binding of the corresponding HaloTag ligands to the HaloTag protein. During this screening, Malachite Green lactones were used in PAI for the first time, generating strong photoacoustic signal. The modifications included two new synthetic strategies for tuning the open-closed equilibrium, nominally 1',8'-methylation and mono-hydrodefluorination. The best performing ligands showed rapid labelling kinetics and up to a 12-fold increase in photoacoustic signal. Investigation of the calcium sensitivity of these ligands, with the absorbance-modulated calcium sensor scaffolds, HaloCaMP1a and HaloCaMP1b, resulted in 8 photoacoustic calcium indicators with a range of affinities and sensitivities. **These are the first positive-going, far-red/NIR calcium indicators designed for PAI, and outperform existing sensors for this modality in spectroscopic measurements with up to 7.5-fold photoacoustic turn-on.**

Whilst this screening pipeline for the development of chromogenic dyes is well established, further method development would accelerate the process, providing easier access to optimally tuned dyes. First, the two methods (D_{50} and $K_{L\text{-}Z}$)[170,145] for quantification of the open-closed equilibrium of these environmentally-sensitive dyes have limitations. Both approaches are unable to distinguish between the highly chromogenic and completely closed dyes below a certain threshold (D_{50} ~75, $K_{L\text{-}Z}$ <0.0001). Moreover, the low absorption of these predominantly closed dyes necessitates high concentrations, which can lead to aggregation, preventing quantification.[206, 207] Addition of detergents such as SDS can resolve aggregates,[206] but these can influence the equilibrium and have scaffold-dependent effects. To remove the issue of low signal from the closed form, perhaps an NMR titration method could be developed where NMR spectra are recorded in solvent mixtures of different polarity and the equilibrium determined based on relative differences in chemical shift between the open and closed forms.[171]

Furthermore, on the synthetic side, this work highlighted the challenge in the synthesis of fluorinated dyes, in particular **15**. The susceptibility of the tetrafluorinated ring to S_NAr by amines prevented the convenient late-stage derivatisation using palladium-catalysed cross-coupling.[167] The current routes for these polyfluorinated dyes instead install the auxochromes early in the synthesis before addition of the bottom ring.[166, 168] This impedes simple derivatisation of dyes near the end of the synthesis, to evaluate the effect of different nitrogen substituents, for example, on the photophysical properties (as was possible for nonfluorinated rhodamines).[145, 167, 176] Therefore, there is an existing requirement for development of new synthetic methods for these fluorinated scaffolds enabling late-stage derivatisation.

As an aside, I have developed novel strategies for tuning photophysical properties, which demand further exploration on rhodamines used for fluorescence imaging. Selective removal of a single fluorine atom could improve the turn-on of modestly chromogenic dyes, such as the bioavailable **JF669**,[169, 208] by slightly closing the equilibrium (Figure 43a). Adjusting the reaction conditions and using different ester or amide groups on the 6-position could enable removal of a second fluorine atom for larger equilibrium shifts of more open dyes, making them more cell permeable.[166] Methylation of the xanthene scaffold could also enhance the fluorogenicity for dyes such as **JF646**[176] or **MaP555**[148], and their corresponding targetable Ca^{2+} indicators **JF646-BAPTA**[147] and the **MaPCa** dyes[149], which still have room for further closure.

Whilst the initial properties of our sensors are promising, there is still room for improvement and so we will continue to explore these dye scaffolds and also the HaloCaMP protein. The dynamic range of our calcium indicators are good with $\Delta PA/PA_0$ up to 7.5, but still far smaller than current GCaMP sensors, the neuroscience field's gold standard tool. Additional synthetic modifications could help identify more chromogenic structures, minimising free ligand signal, but with a reduced trade-off in the maximum PA signal bound to HaloTag. For example, 3,3-dimethylazetidine or pyrrolidine substituents could be worth exploring (Figure 43b). However, a crucial parallel step to exploit the full equilibrium range of these scaffolds is to work on the HaloTag-based protein partner. As demonstrated by the improvement in photoacoustic turn-on for dyes **10-HTL** and **14-HTL** with HaloTag9, targeted mutations can enhance photophysical properties. The HaloCaMP proteins are existing calcium sensors, originally engineered with the structurally different Si-rhodamine **JF635-HTL**. Tailored protein engineering of the HaloCaMP scaffold with a Malachite Green lactone or pyrrole-xanthene dye would optimise the dye-protein interactions, maximising the environmental change experienced by the bound ligand upon the calcium-induced conformational change, and so increase dynamic range. Modifications to HaloCaMP can also improve the other important sensor property which needs tuning for optimised performance in neurons, calcium-binding affinity. In the next chapter, I discuss my attempts to improve these first-generation calcium sensors with regard to their calcium affinity, sensitivity, and bioavailability.

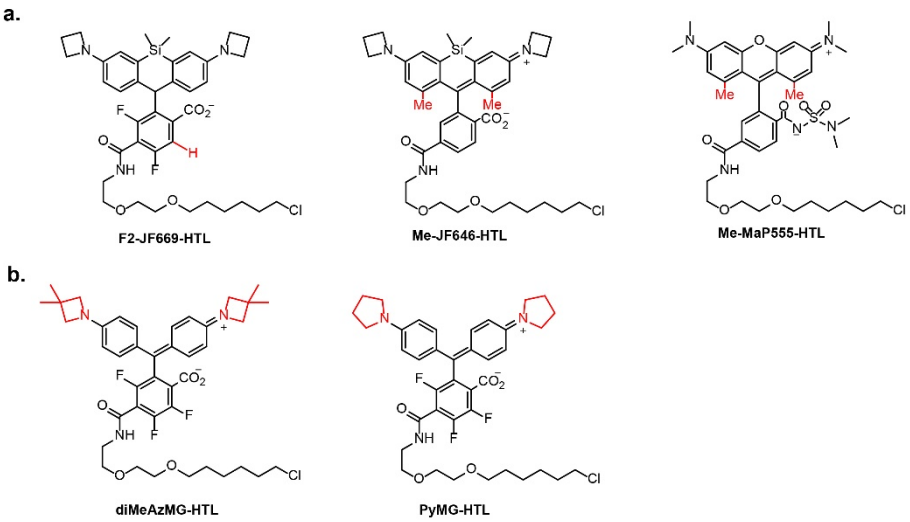


Figure 43: Future dye structures to explore. (**a**) New tuning methods of the open-closed equilibrium applied to fluorescent rhodamine derivatives: defluorinated **JF669-HTL, "F2-JF669-HTL"**; methylated xanthene scaffolds of **JF646-HTL** and **MaP555-HTL, "Me-JF646-HTL"** and **"Me-MaP555-HTL"**, with the modifications highlighted in red. (**b**) Potential acoustogenic Malachite Green lactone ligands with 3,3-dimethylazetidine and pyrrolidine auxochromes, with the modifications highlighted in red.

Chapter III: Optimising biosensors – protein engineering and development of a high throughput PA screening assay

The work discussed in this chapter was conducted in collaboration with Dr. Claire Deo and Dr. Begoña Ugarte-Uribe (Deo group, EMBL), Nikita Kaydanov (Prevedel group, EMBL) and Dr. Kenana Al-Adem and Dr. Simon Göllner (Stiel group, Helmholtz Zentrum Münich). Claire performed the ADPA well plate assay, Begoña helped me with the cloning of the NanoLuc constructs and Nikita performed the photoacoustic spectroscopy in lysates. Kenana and Simon performed the labelling and photoacoustic microscopy of bacterial colonies on their set-up.

III.1 Objective

The development of novel probes and calcium sensors is an iterative process, involving several stages continuously trying to improve existing properties and diversify their biological applications. For example, the original version of the most widely used fluorescent calcium sensors, the GCaMP series, suffered from low brightness, limited dynamic range, slow kinetics and was susceptible to photobleaching and pH changes.[209] Two pairs of mutations were consecutively introduced to enhance maturation efficiency and brightness, resulting in GCaMP2, 200 times brighter than GCaMP1, which could monitor calcium dynamics of a mouse embryonic heart *in vivo*.[210, 211] It was not until GCaMP3, 8 years later, with higher protein stability, larger dynamic range and strong calcium affinity, achieved via mutations close to the EGFP chromophore at the CaM/M13 domain interface and the EF hands (calcium-binding motifs) of CaM, that sufficient sensitivity for neuroimaging was achieved.[111] Several GCaMP5 variants were subsequently created with improved brightness and sensitivity to more reliably detect single action potentials *in vivo*, but still were inferior in kinetics and sensitivity to the best synthetic dyes, such as Oregon-Green BAPTA.[212] Screening variants in neurons, especially targeting the cpEGFP-CaM and CaM/M13 interfaces for mutagenesis, led to a big jump in performance with the development of GCaMP6f, GCaMP6s and GCaMP6m, optimised for rapid kinetics, high sensitivity or a compromise between the two.[112] The jGCaMP7 series was more specifically tailored to particular imaging requirements, for example tracking of whole neuronal populations or visualising thin neuropil or neurite processes.[113] The most recent iteration, the jGCaMP8 sensors, improved calcium-binding kinetics, without a trade-off in brightness or sensitivity, by exchanging the calcium binding peptide.[114]

The GCaMP story clearly demonstrates the need for continued modifications and sensor tuning to improve performance. Now we have our first-generation of photoacoustic calcium sensors, we must follow a similar path of stepwise improvements to make them suitable for whole-brain *in vivo* neuroimaging. One difference is that these chemigenetic reporters have two components which must be modified in parallel: the synthetic, acoustogenic ligands, and the calcium-sensitive HaloCaMP protein scaffold. The ligands developed in Chapter II were combined with already established protein scaffolds: HaloTag, HaloCaMPs and rHCaMP.[139, 140, 143, 193] These were originally engineered using very specific fluorophores: **TMR**, **JF635**, and a benzocoumarin respectively, which have different structures to our series of PA dyes. Optimising the specific dye-protein interactions via protein engineering is required to improve the performance of our sensors, particularly to increase acoustogenicity to at least match the turn-ons achieved by their fluorogenic counterparts, increasing the dynamic range of the calcium sensors and tuning their calcium affinity to the physiological range of neurons (sections III.2, III.3). However, as a relatively new modality, method development for high-throughput screening of protein mutants is necessary and our exploratory attempts will be discussed (section III.4). Moreover, selection of the acoustogenic ligand for re-engineering of the HaloCaMP scaffold must also consider bioavailability, as this dye will be used for subsequent *in vivo* experiments. This is not easy to predict for environmentally-sensitive dye ligands without *in vivo* evaluation, but I have attempted to assess their potential for blood-brain barrier (BBB) permeation using *in vitro* assays (section III.5).

III.2 Combining tryptophan quenching with acoustogenic ligands

To begin our attempts to enhance the acoustogenicity of our ligands upon binding to HaloTag, I first explored the effect of tryptophan-based quenching from the HaloTag protein surface on our synthetically quenched acoustogenic ligands. Tryptophan point mutations close to the dye binding site have been introduced in dim HaloTag variants G171W and P174W to modulate the brightness and fluorescence lifetime of bound

fluorophores.[190, 213] The former mutant has also been used in the WHaloCaMP sensor to provide the proximity-induced, calcium-dependent turn-on mechanism (Figure 19e).[159] In our PA probe platform, we hoped for a synergistic effect, with the tryptophan's additional quenching increasing the PA signal of the bound ligands (Figure 44a). Therefore, we measured the absorption and PA signal of **14-HTL** and **JF635-HTL** as a comparison, bound to the G171W and P174W HaloTag variants in our custom-built spectrometer (see section II.2). The tryptophan-based quenching enhanced the PA signal of **JF635**-HaloTag7 but still lagged behind **14**-HaloTag7 with 7.2- and 2.2-fold weaker signal for **JF635**-G171W and **JF635**-P174W respectively (Figure 44b). Unfortunately combining these tryptophan-mutated variants with **14-HTL** made no further improvement in signal, because of the already negligible fluorescence, and the reduced absorption with G171W. However, at least our approach of using synthetic quenching mechanisms over tryptophan-based quenching from the protein surface is validated, given the superior signal intensity of **14**-HaloTag7.

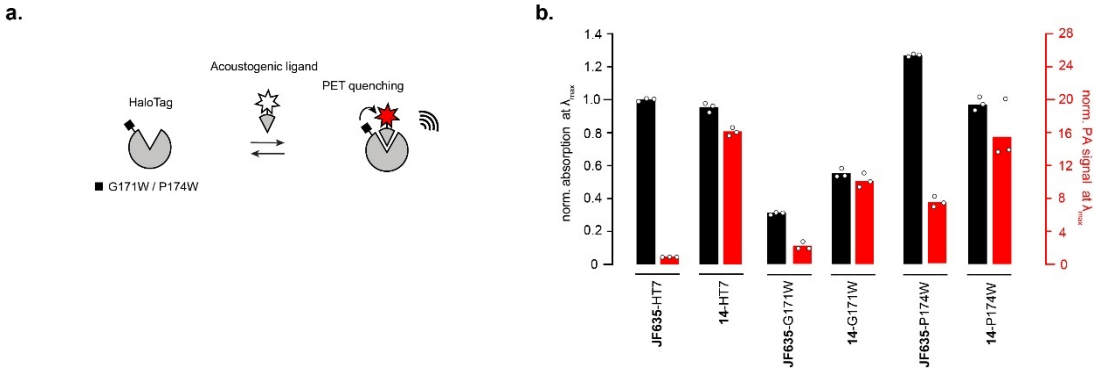
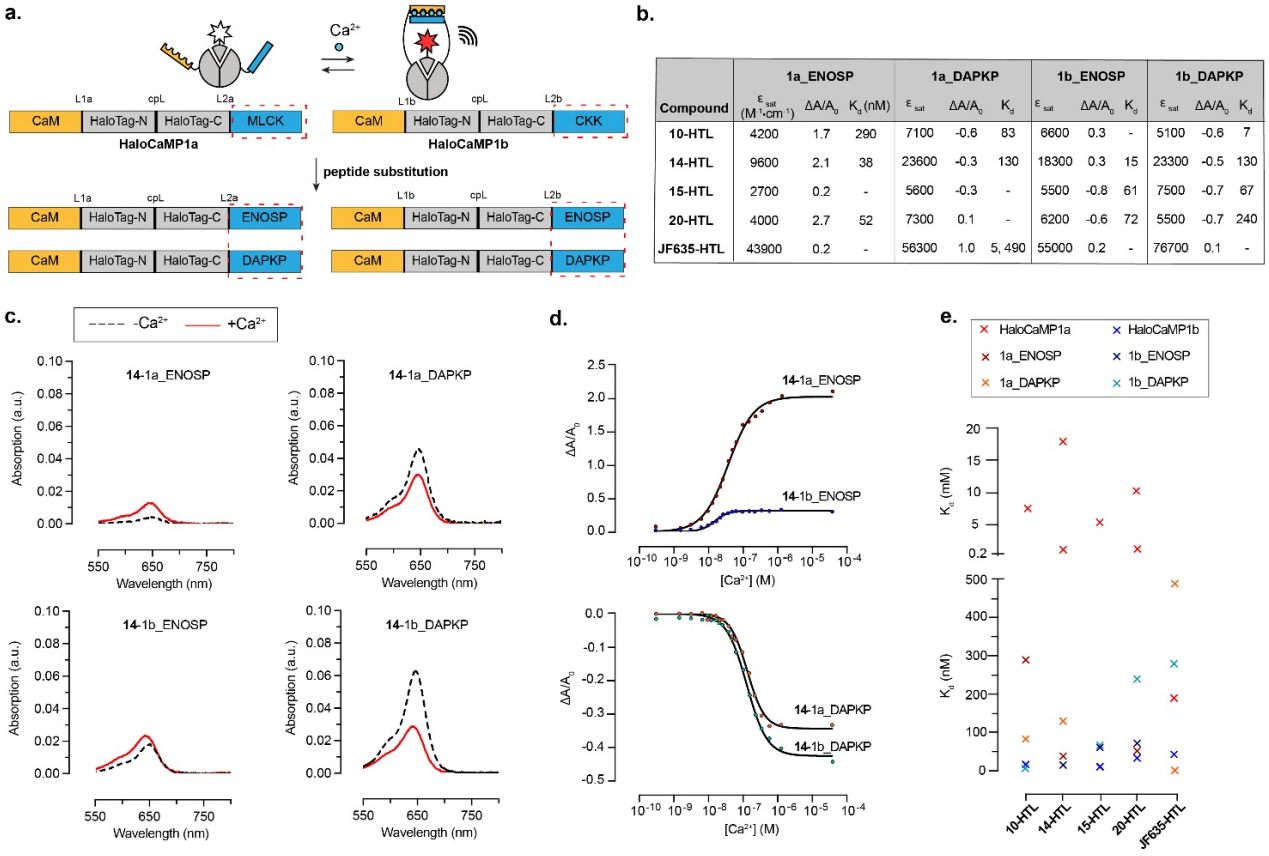


Figure 44: Tryptophan quenching from HaloTag variants with acoustogenic ligands. (**a**) Combining acoustogenic ligands with tryptophan-mutated HaloTag variants might enhance PA signal. (**b**) Normalised absorption at λ_{max} (black) and normalised photoacoustic signal at $\lambda_{PA} = \lambda_{max}$ (red) for **JF635-HTL** and **14-HTL** bound to HaloTag7 (HT7), or variants G171W and P174W. Values were normalised to **JF635**-HaloTag7. All measurements were performed in 10 mM HEPES, pH 7.4, containing 0.1 mg·mL^{-1} CHAPS at a dye ligand concentration of 1.25 µM and protein concentration of 1.9 µM, in triplicate.

III.3 HaloCaMP affinity variants

As discussed in Chapter II, the calcium-binding affinity of our photoacoustic calcium sensors requires optimisation to improve performance in living cells. Matching the K_d to the physiological range of neurons (~100–200 nM) would enable the maximum dynamic range of the sensor to be exploited during recording of neural activity.[214] Therefore, the HaloCaMP1a sensors require a large decrease in K_d from their current millimolar levels, whilst the HaloCaMP1b indicators need a subtle increase. To modify affinity in sensors with calmodulin-based calcium-binding domains, targeted mutations can be introduced into either the calmodulin or peptide sequences to alter the strength of their associative interactions, or the whole peptide can be exchanged.[114, 159, 215-217] Other key photophysical properties, such as sensitivity and response kinetics, are also dependent on calcium binding, so can be affected by these modifications. The complex series of interactions involved in the calcium-induced conformational changes make it impossible to predict *in situ* the effects of a particular modification but these can be estimated from previous engineering attempts. During the development of jGCaMP8, an extensive screen of GCaMP6s constructs with different calcium-binding peptides was performed in order to identify variants with improved kinetics without compromising brightness or sensitivity.[114] From this initial screen, they identified two peptides, endothelial nitric oxide synthase (ENOSP) and death-associated protein kinase 1 (DAPKP), which resulted in variants with the best combination of photophysical properties (K_d, $\Delta F/F_0$, k_{off}, k_{on}, Hill coefficient). These had contrasting effects on affinity with the ENOSP substitution reducing affinity of GCaMP6s significantly more (144 to 1062 nM) than DAPKP (205 nM). Given the structural and functional similarity of HaloCaMP to GCaMP6s (Figure 13, 19c), I decided to

perform the peptide substitution of both HaloCaMP1a and HaloCaMP1b with these two peptide sequences to assess their effects on affinity (Figure 45a).



Compound	1a_ENOSP			1a_DAPKP			1b_ENOSP			1b_DAPKP		
	ε_{sat} ($M^{-1}\cdot cm^{-1}$)	$\Delta A/A_0$	K_d (nM)	ε_{sat}	$\Delta A/A_0$	K_d	ε_{sat}	$\Delta A/A_0$	K_d	ε_{sat}	$\Delta A/A_0$	K_d
10-HTL	4200	1.7	290	7100	-0.6	83	6600	0.3	-	5100	-0.6	7
14-HTL	9600	2.1	38	23600	-0.3	130	18300	0.3	15	23300	-0.5	130
15-HTL	2700	0.2	-	5600	-0.3	-	5500	-0.8	61	7500	-0.7	67
20-HTL	4000	2.7	52	7300	0.1	-	6200	-0.6	72	5500	-0.7	240
JF635-HTL	43900	0.2	-	56300	1.0	5, 490	55000	0.2	-	76700	0.1	-

Figure 45: Design and characterisation of HaloCaMP affinity variants. (**a**) Approach of peptide substitution in HaloCaMP1a and HaloCaMP1b to generate affinity variants. (**b**) Photophysical properties of the HaloCaMP affinity variants with **10-, 14-, 15-, 20-** and **JF635-HTL**. (**c**) Absorption spectra of **14-HTL** with the four affinity variants. (**d**) Ca^{2+} titrations of HaloCaMP variants labelled with **14-HTL**. (**e**) Plot of K_d values for the HaloCaMP variants labelled with **10-**, **14-**, **15-**, **20-** and **JF635-HTL**. All measurements were performed in 30 mM MOPS, 100 mM KCl, pH 7.2, containing 0.1 mg·mL^{-1} of CHAPS, to which either 10 mM EGTA (for the calcium-free state) or 10 mM $CaCl_2$ was added (for the calcium-saturated state); values are the mean of 3 replicates, for the titrations 2 replicates; all measurements were performed at 1.25 µM dye and 1.9 µM protein.

I cloned via Gibson assembly four new HaloCaMP variants (1a and 1b with DAPKP or ENOSP peptide, Figure S16), which I then transformed into *E. coli* for expression and purification. The dynamic ranges, extinction coefficients and calcium affinities were subsequently measured after conjugation with the four principal acoustogenic ligands developed in Chapter II (**10-**, **14-**, **15-** and **20-HTL**) and **JF635-HTL** (Figure 45, S17, S18). Regarding affinity, the peptide substitution was largely successful with the millimolar K_ds of the HaloCaMP1a sensors replaced by nanomolar values (Figure 45b, e). In general, the DAPKP peptide adjusted the affinity into the desired ~100−200 nM range, whereas the sensors with ENOSP presented slightly higher affinities. The effect of the peptide substitution on calcium binding and unbinding kinetics would also be interesting to explore to see if the rate enhancement is conserved from the jGCaMP8 sensors. Unfortunately, despite the affinity improvement, sensitivity was compromised for the majority of the sensors, with the exception of the 1a_ENOSP variant for **10-**, **14-** and **20-HTL** increasing up to 2.7-fold turn-on (Figure 45, S17). Interestingly, several of the dye-HaloCaMP variant combinations switched to a negative response particularly with the DAPKP peptide, with higher signal in the absence of calcium (Figure 45c, S17). Additionally, the dyes were significantly more closed on these variants than with the parent proteins, including **JF635-HTL**, revealed by the low extinction coefficients, which would currently limit PA signal intensity (Figure 45b). The biphasic behaviour of dyes **14-** and **20-HTL** with HaloCaMP1a was abolished, however **JF635**-1a_DAPKP and **10-**

1b_DAPKP showed independent association of calcium to the two domains with an unusual switch between negative and positive responses (Figure 45d, S18).

Therefore, more extensive protein engineering is required to improve the collective photophysical properties of these sensors. The variation in properties between the sensors highlights the specificity of the dye-protein interactions and the need for a tailored engineering approach using a selected dye throughout the screening process. However, for such an approach involving likely thousands of mutants, a high-throughput assay for screening and identifying the most performant mutants must be developed.

III.4 Development of a high-throughput screening assay for PA biosensors

III.4.1 Absorption measurements in bacterial lysates

For HaloTag-based fluorescent biosensors, we have an established lysate screening assay in the lab which ideally would be directly adapted to my photoacoustic sensors using an alternative readout (Figure 46a). Site-saturation mutagenesis using a degenerated NNS codon would lead to libraries of HaloCaMP variants which could be expressed in *E. coli* as EGFP fusions, and the fluorescent colonies picked and grown in 96-well plates to cover all possible amino acids in the mutated position. The cells would be lysed, and protein concentration quantified via EGFP fluorescence using a plate reader. After labelling with the selected dye-ligand for tailored protein engineering, the lysate would be split into two and Ca^{2+} or EGTA-containing buffers added. PA signal (or equivalent) would be measured in each state, from which $\Delta PA/PA_0$ could be calculated. However, unlike for fluorescence, there is no commercially available photoacoustic plate reader. Therefore, I examined the sensitivity of absorption plate readers to confirm whether absorbance could be used as a readout, an approach which was successful for screening of dye candidates in Chapter II. With purified HaloTag protein bound to **14-HTL**, I observed a linear relationship between absorbance and concentration to ~250 nM and to ~500 nM with the lower absorbing HaloCaMP variants (Figure 46b). This could be sufficient for such an assay if protein expression is good as we usually attain 500 nM – 2 µM protein in these wells, although at least 1.5-fold excess is required to ensure all dye is bound.

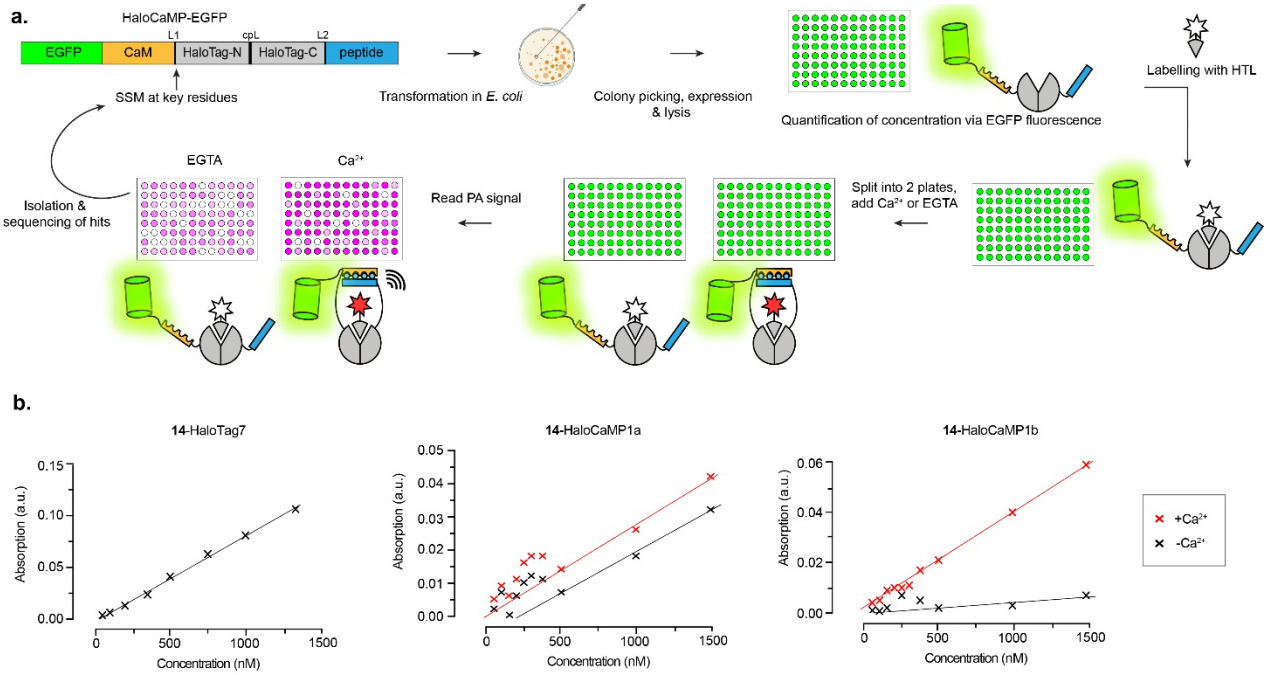


Figure 46: Design of a high-throughput screening assay, using absorbance as a readout. (**a**) Ideal pipeline for high-throughput screening of mutant variants for our PA sensors. (**b**) Plots of absorption versus concentration for purified HaloTag7 (in 10 mM HEPES, pH 7.4, containing 0.1 mg·mL^{-1} CHAPS), HaloCaMP1a and HaloCaMP1b (in 100 mM Tris-HCl, pH 7.5, 300 mM NaCl, 0.1 mg·mL^{-1} CHAPS) labelled with **14-HTL** measured using an absorbance plate reader.

To test whether these measurements could be extended to screening in bacterial lysates, I labelled the *E. coli* lysate after expression of the parent calcium sensors HaloCaMP1a and HaloCaMP1b, split the lysate into two 96-well plates, mixed with Ca^{2+} or EGTA buffers and measured absorbance. Unfortunately, very variable results were observed due to a high background signal from the lysate (OD ~ 0.1) which was similar to or greater than the signal from the sensor. Therefore, even minor fluctuations in this background could obscure the calcium-dependent signal changes. I attempted to clarify the lysate through filtering or by using histidine-tag-functionalised magnetic beads, which enabled a crude affinity chromatography purification in each well, but no major improvements were observed (Figure 47a). In general, no linear absorption-concentration relationship was found and whilst in some wells the expected turn-ons were measured, for HaloCaMP1a in particular, the large well-to-well variations prevent use of this method for screening mutant variants. To identify if PA signal was also strongly altered by lysate composition, we measured the absorption and PA signal of **14-HTL** labelled HaloCaMP1a lysates in a cuvette with Nikita's multimodal spectrometer (Figure 47b). The variation in the non-zero baseline for absorbance reinforced the problem with the plate reader measurements, shifting the peak absorption and hence the turn-on values, whereas the PA spectrum was relatively unaffected. This suggests that PA could resolve this problem but we have yet to try our PA spectroscopy set-up in the well plate reader configuration (as used for the photostability measurements in II.2.3, II.3.2) for bacterial lysate measurements.

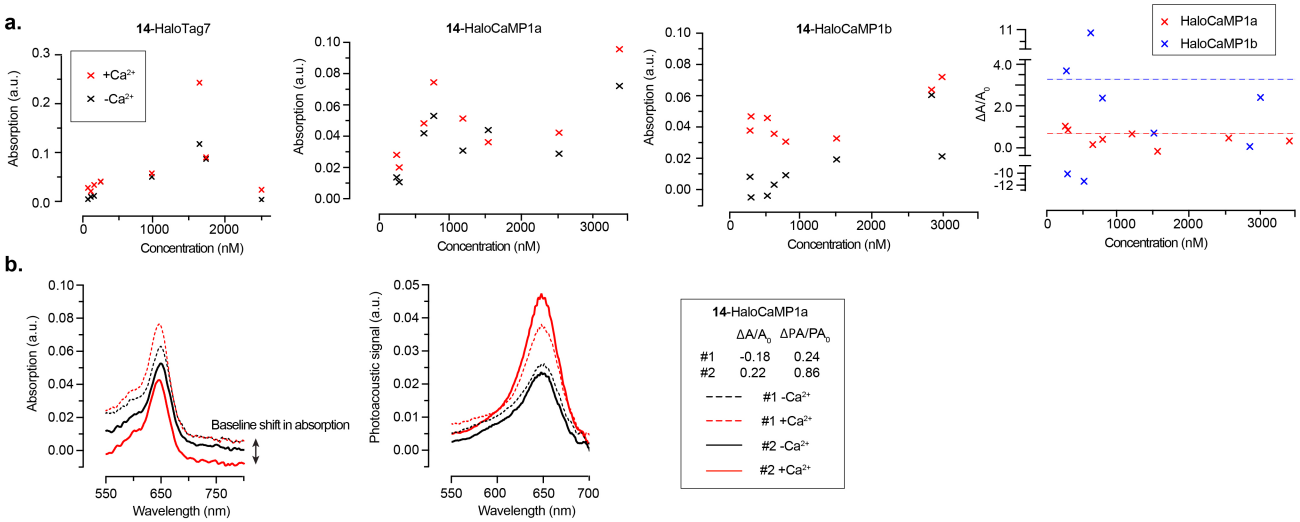


Figure 47: Trials for absorption screening of HaloCaMP1a/b in bacterial lysates. (**a**) Plots of absorption versus concentration for *E. coli* lysates (in 100 mM Tris-HCl, pH 7.5, 300 mM NaCl, 0.1 mg mL^{-1} lysozyme, 1 mM PMSF, DNase) after expression of HaloTag7 (far-left), HaloCaMP1a (middle-left) and HaloCaMP1b (middle-right), and labelling with **14-HTL**, plus the absorbance turn-ons for each pair of measurements (far-right); dotted lines indicate turn-on of purified **14**-HaloCaMP proteins. (**b**) Absorption and photoacoustic spectra of HaloCaMP1a lysates (in #1 100 mM Tris-HCl, pH 7.5, 300 mM NaCl, #2 100 mM Tris-HCl, pH 7.5, 100 mM NaCl), and labelled with **14-HTL**.

III.4.2 Exploiting the photosensitizing properties of Malachite Green as a readout

As an alternative to the measurement of absorption or photoacoustic signal from lysates, we considered whether we could generate a fluorescent readout from the change in PA signal, which could be measured with high sensitivity using commercial 96-well plate readers. The majority of our sensors are based on the Malachite Green dyestuff, which is not only a fluorogenic dye but can also function as a photosensitiser. Photosensitisers can generate singlet oxygen species upon light absorption resulting in local biological effects such as protein inactivation (CALI)[218] or cell death for photodynamic therapy.[219] An iodine-substituted Malachite Green derivative has been shown to present a high singlet oxygen quantum yield when bound to a fluorogen activating protein (FAP dL5**).[189] The iodine heteroatoms enhance intersystem crossing to the long-lived triplet state

which can react with molecular oxygen to form singlet oxygen ($^1O_2$; Figure 48a).[220,174] The unsubstituted Malachite Green displayed 40% of this $^1O_2$-generating capacity, therefore, we wondered whether our fluorinated Malachite Green lactone structure would exhibit a detectable degree of photosensitising ability. Singlet oxygen generation can be determined by the bleaching rate of a fluorescent scavenger such as anthracene-9,10-dipropionic acid (ADPA),[221] providing a fluorescence readout, which could be combined with our assay design from Figure 46a. ADPA could be added into the wells and the bleaching monitored during irradiation of the HaloCaMP complexes labelled with **14-HTL**. Faster bleaching would indicate that the HaloCaMP mutant shifts the dye more towards the open form, corresponding to stronger PA signal (Figure 48b). A greater difference in ADPA bleaching rate between the calcium-bound and calcium-free forms would then signify larger $\Delta PA/PA_0$.

To test the photosensitising capability of our acoustogenic ligands, Claire performed a plate reader assay of **14-HTL** with ADPA. Free **14-HTL** or pre-labelled **14**-HaloTag7 were added to wells at different concentrations (1-10 µM) with ADPA (50 µM) and continuously illuminated with an LED at 650 nm. The fluorescence of ADPA was measured at intervals throughout the illumination period. No fluorescence decay was observed independent of concentration with no difference between the negative controls lacking **14-HTL** (Figure 48c). Therefore, these Malachite Green lactones bound to HaloTag are not sufficiently strong photosensitisers to use this property as a readout in the screening assay. Similar iodination might be required to increase the photosensitising strength but this would strongly alter the position of equilibrium of the ligands and so affect their behaviour on HaloCaMP binding. Alternatively, the conformational restriction of Malachite Green afforded by the FAP binding may also be necessary to promote singlet oxygen generation, along with fluorescence, which does not seem to occur when bound to HaloTag given the negligible fluorescence of our ligands. On a positive note, the failure of this assay suggests Malachite Green lactones likely do not exert phototoxic effects on illumination in cells, although a neuronal cell viability assay should be performed to confirm this.

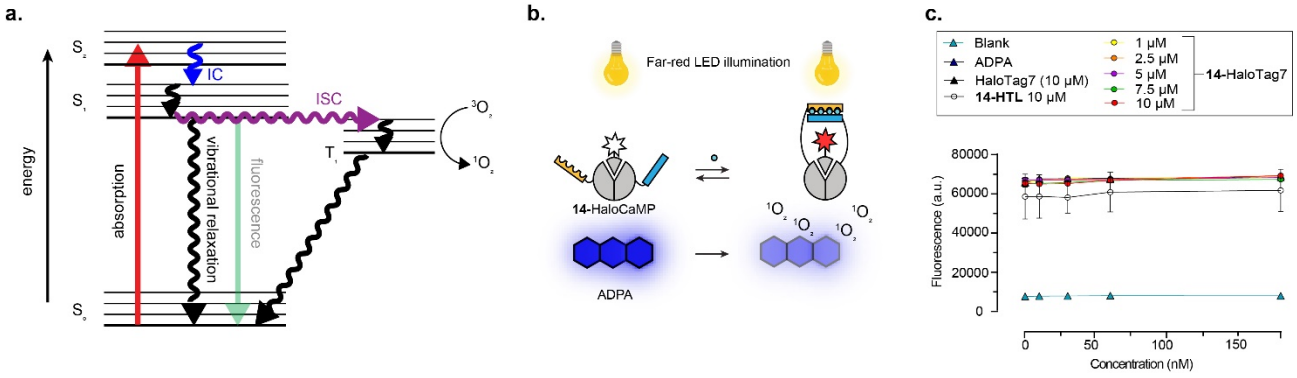
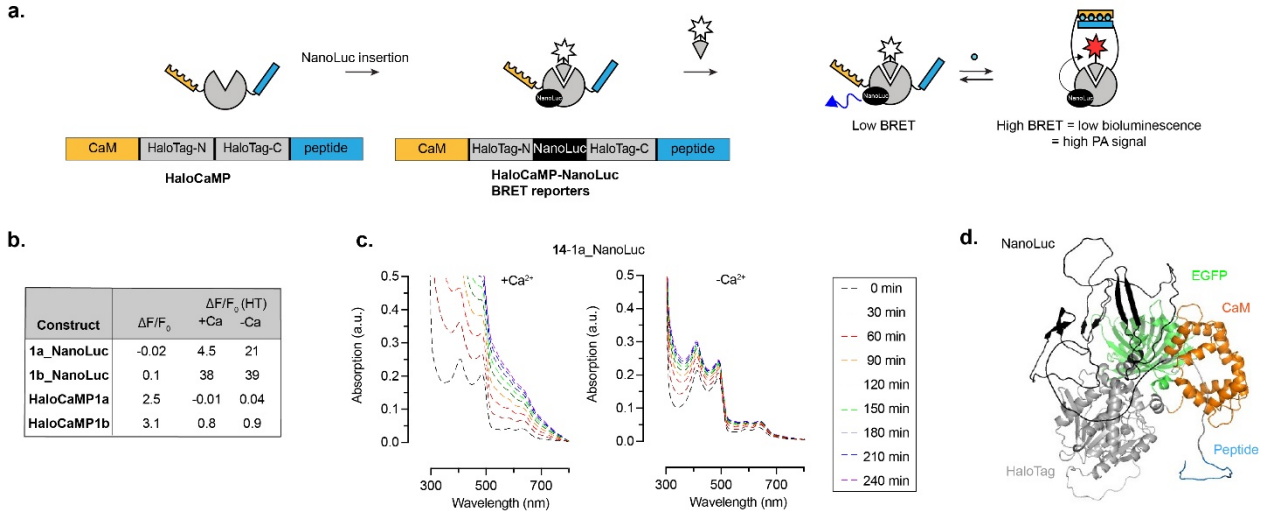


Figure 48: Investigating the photosensitising ability of our Malachite Green lactones. (**a**) Jablonski diagram outlining the process of reactive oxygen species generation upon illumination of a photosensitiser; IC: internal conversion; ISC: intersystem crossing; T_1: first triplet state. (**b**) Mechanism of the proposed ADPA assay with HaloCaMP protein variants. (**c**) Photobleaching ADPA assay using **14-HTL** with and without HaloTag protein at different concentrations (1-10 µM), λ_{ex} = 374 nm, λ_{em} = 390 nm, LED illumination at 650 nm.

III.4.3 Development of a HaloCaMP BRET sensor

Another option to generate an emission readout from our acoustic biosensors, is to form a FRET or BRET pair by combining HaloCaMP with a fluorescent protein or luciferase donor respectively. HaloCaMP labelled with an acoustogenic ligand would act as the acceptor and so the energy transfer's efficiency would be determined by the absorption of the bound ligand, which is in turn dependent on the open-closed equilibrium. HaloCaMP mutants capable of opening up the dye more, would show higher FRET/BRET efficiency and hence lower fluorescence or bioluminescence signal from the donor channel. Designing such a negative reporter first requires

sufficient spectral overlap of the donor's emission with the dye's excitation peak. Additionally, the orientation and close proximity of the partners are also essential and the recent ChemoCaM design (Figure 19a) shows high FRET efficiency is possible with rhodamines of similar far-red wavelengths to our ligands.[153] However, this required extensive mutagenesis of the HaloTag-EGFP interface in the region where HaloCaMP has the calcium-responsive domains, so designing a HaloCaMP-EGFP FRET reporter would necessitate full engineering from scratch. The BRET reporter, H-Luc, which also achieves sufficient BRET efficiency with the far-red **SiR** ligand, has NanoLuc inserted into HaloTag in a different position.[154] Therefore, insertion of NanoLuc into HaloCaMP to make a calcium-dependent BRET reporter seemed more feasible (Figure 49a). The assay from Figure 46 would then be adapted with expression of HaloCaMP-NanoLuc variants, labelling with the selected acoustogenic ligand and the NanoLuc substrate, then bioluminescence measurement after addition of Ca^{2+}/EGTA buffers. Similarly to the photosensitiser design, a decrease in the bioluminescence signal in the calcium-saturated state would indicate a greater PA response.



Construct	$\Delta F/F_0$	$\Delta F/F_0$ (HT) +Ca	$\Delta F/F_0$ (HT) -Ca
1a_NanoLuc	-0.02	4.5	21
1b_NanoLuc	0.1	38	39
HaloCaMP1a	2.5	-0.01	0.04
HaloCaMP1b	3.1	0.8	0.9

Figure 49: Development of a HaloCaMP-NanoLuc BRET reporter. (**a**) The design of HaloCaMP-NanoLuc BRET reporters and their mechanism of action. (**b**) Table of fluorescence turn-on values for 1a_NanoLuc, 1b_NanoLuc, HaloCaMP1a, and HaloCaMP1b lysates labelled with **JF635-HTL**; $\Delta F/F_0$: the fluorescence turn-on between the calcium-saturated and calcium-free states, $\Delta F/F_0$ (HT): fluorescence turn-on after addition of excess HaloTag7 for either the calcium-saturated or calcium-free states. (**c**) Absorption spectrum of 1a_NanoLuc labelled with **14-HTL** at 2.5 µM in 30 mM MOPS pH 7.2, 100 mM KCl, 500 mM $CaCl_2$ or 10 mM EGTA. (**d**) Representative example of predicted Colabfold structure of 1a_NanoLuc.

A Gibson assembly cloning strategy afforded the two constructs, 1a_NanoLuc and 1b_NanoLuc, with NanoLuc inserted into position 154 of the cpHaloTag protein domain of HaloCaMP1a and HaloCaMP1b respectively (Figure 49a, S19). Before examining BRET efficiency, the first test was to determine if HaloTag ligand binding and calcium sensitivity was retained after insertion of the NanoLuc domain. The two HaloCaMP-NanoLuc chimeras were expressed in *E. coli* and the corresponding lysates were labelled with **JF635-HTL** then separated into well plates with Ca^{2+}/EGTA buffers. $\Delta F/F_0$ was determined before and after incubation with excess purified HaloTag protein to determine the bound fraction of dye. 1a_NanoLuc and 1b_NanoLuc lysates showed negligible calcium sensitivity (Figure 49b $\Delta F/F_0$) and a large increase in fluorescence upon addition of excess HaloTag compared to HaloCaMP1a and 1b (Figure 49b $\Delta F/F_0$ (HT)). This suggests that HaloTag ligand binding to these constructs is impaired, as all unbound dye was captured by the additional HaloTag. I also measured the absorption spectra of purified 1a_NanoLuc after addition of **14-HTL** (Figure 49c). Strong aggregation of the protein occurs in both high and low calcium conditions, and there is an unusual dual peak from the EGFP fusion, which highlights a protein instability issue. The loss of function is perhaps not surprising with the high steric demand to maintain the complete folding of all of the closely packed domains. The predicted Colabfold[222, 223] structure of these proteins supports these results, with the most probable structures often containing completely or partially unfolded NanoLuc and/or calmodulin-binding peptide domains, and a split HaloTag protein (Figure

49d, S20). Protein sequences of alternative insertion positions in HaloTag (141-G or 142) and different cpNanoLucs, based on the best performing H-Luc variants,[154] showed similar unfolded protein structure predictions.

In summary, our attempts to make a HaloCaMP-based BRET reporter to provide a luminescence readout for screening of protein mutants were unsuccessful. We tried to simply replace HaloTag in the H-Luc bioluminescent reporter with the calcium-sensitive HaloCaMP but the resulting protein lost HaloTag ligand binding function. Design of the desired reporter, or with a fluorescent protein partner, is likely possible but would require extensive protein engineering work. Moreover, the same dye ligand would be used in the development of the reporter and in the engineering of the calcium sensor, so this strategy would not be generalisable if we want to explore different acoustogenic dye scaffolds in the future. Therefore, we decided to consider alternative screening methods.

III.4.4 Colony screening

High-throughput screening of protein mutants is not solely restricted to a 96-well plate format. Another method which has been established for optimising calcium sensors by protein engineering directly screens bacterial colonies using fluorescence and photoacoustic microscopy.[224, 225] A plate of colonies is imaged by PAM (Figure 50a) and quantified, before and after adding high calcium buffer, providing a universally applicable method, independent of the dye or protein scaffold. This has the added advantage of not requiring colony picking, growing and lysis of clones to identify initial hits, and is thus faster compared to screening in lysates. However, these examples use fully genetically-encoded sensors, so do not require dye labelling. Therefore, to explore whether we could use a similar approach, I grew and labelled *E. coli* colonies expressing HaloTag7-EGFP with **JF635-HTL** or **14-HTL** by gentle application of dye solution onto the Petri dish. These plates were imaged on a fluorescence gel imager, which enables whole plate imaging at EGFP and far-red wavelengths (Figure 50b). Quantifying the fluorescence signal showed an increase in the far-red/EGFP ratio, providing evidence of colony labelling for both dyes without the need to break open the cells, although the fluorescence signal for the plate labelled with **14-HTL** was lower, because of its lower quantum yield (Figure 50b, c). However, the labelling was inhomogenous, probably due to variable dye diffusion across the plate.

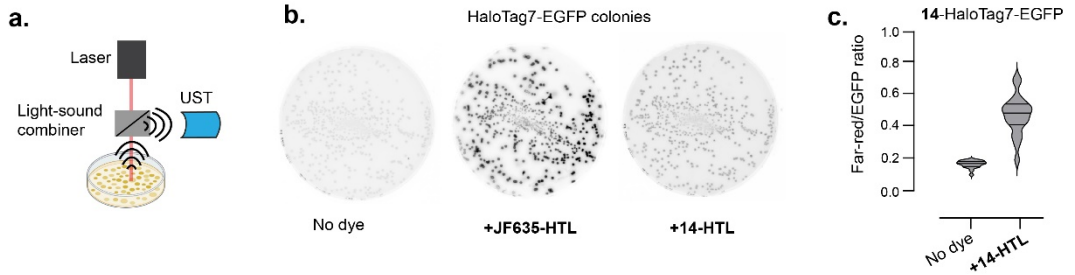


Figure 50: Initial tests for bacterial colony screening using photoacoustic microscopy. (**a**) Schematic of PAM set-up for imaging bacterial colonies. (**b**) Representative fluorescence images of HaloTag7-EGFP expressing colonies labelled with **JF635-HTL** or **14-HTL**. (**c**) Quantification of a plate of HaloTag7-EGFP colonies before and after labelling with **14-HTL**.

To explore this screening approach further, we started a collaboration with the Stiel lab (Helmholtz Zentrum Münich), experts in protein reporters for PAI, who have also developed a prototype PAM set-up for colony screening. Kenana optimised the dye labelling by covering the colonies with a layer of low gelling point agarose containing **14-HTL**, which was left to set overnight. PAM detected consistent, strong PA signal from labelled HaloTag7-EGFP colonies, indicating homogenous labelling, which could be normalised for their differences in size by PA signal from EGFP (Figure 51a, b). Labelling HaloCaMP1a-EGFP colonies with the same method was successful resulting in ~33% less signal, after normalisation, in the calcium-saturated state than **14-**

HaloTag7, consistent with our spectroscopic measurements on the purified proteins (Figure 51b). The colonies were saturated with calcium simply by adding a Ca^{2+}-rich buffer, also containing ionomycin and polylysine to aid permeation, on top of the agarose gel and incubated overnight. To measure both the calcium-bound and calcium-free states in one plate, after measurement of the calcium-bound signal, the buffer was removed and replaced with EGTA-rich buffer, then incubated overnight before re-imaging. With HaloCaMP1a-EGFP colonies we observed a ~30% decrease in photoacoustic signal in the Ca^{2+}-free form, consistent with the $\Delta A/A_0$ of 0.7 for the purified protein (Figure 51c). This difference was larger than the very small increase in the normalised PA signal at 646 nm due to a decrease in the EGFP signal over time from unlabelled HaloCaMP1a-EGFP colonies (Figure S21). We should be able to perform the same procedure for HaloCaMP1b colonies, although we may have to measure the Ca^{2+}-free form first because these high affinity sensors will bind any residual calcium. The time for incubation with Ca^{2+} could also be reduced to minimise changes in the PA signal between measurements, from EGFP in particular. Therefore, we hope that we have sufficient imaging sensitivity to differentiate performance of mutants, enabling us to screen topological variants of HaloCaMP and perform site-saturation mutagenesis of linkers and residues in close proximity to the dye binding site.

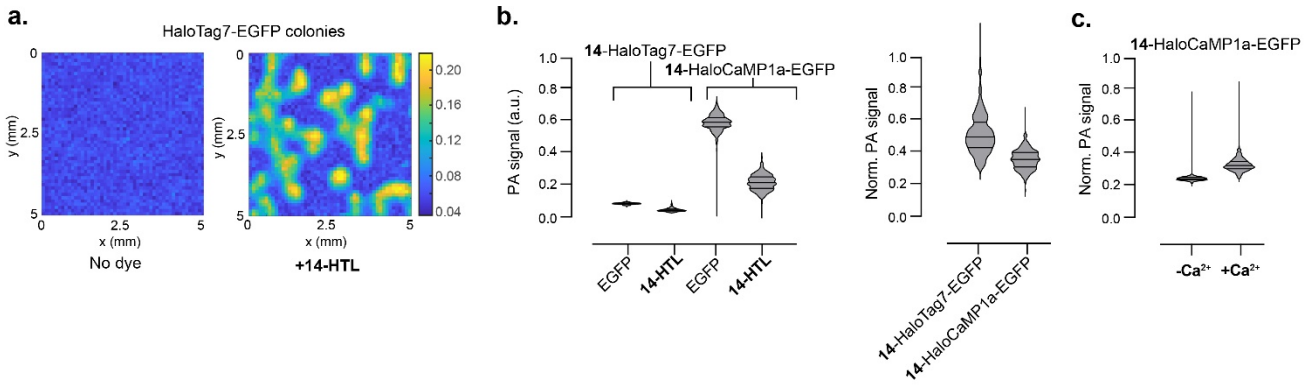


Figure 51: Photoacoustic microscopy of HaloTag7-EGFP and HaloCaMP1a colonies labelled with 14-HTL. (**a**) PAM images of HaloTag7-EGFP colonies with and without **14-HTL**, λ_{PA} = 660 nm. (**b**) PA signal from HaloTag7-EGFP and HaloCaMP1a-EGFP (in the Ca^{2+}-saturated state) colonies labelled with **14-HTL** (left); normalised signal at 646 nm from the same colonies (right); normalisation is performed by dividing the PA signal at 646 nm from the dye by the EGFP signal at 495 nm. (**c**) Normalised PA signal distribution from HaloCaMP1a-EGFP labelled with **14-HTL** in the Ca^{2+}-saturated and Ca^{2+}-free states. Data obtained from Kenana (Stiel Lab).

III.5 Evaluating dye bioavailability

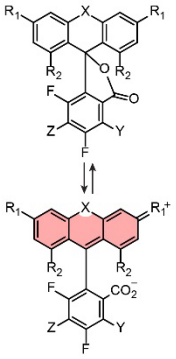
Before we begin extensive screening of HaloCaMP protein mutants, we must carefully select the acoustogenic ligand which we would like to engineer the protein against to optimise the resultant properties. A crucial parameter for application of our PA calcium sensors in the mouse brain, which we have not yet considered, is dye bioavailability. The HaloCaMP protein will be expressed in neurons but requires labelling with the acoustogenic dye ligand. Ideally, the dye would be delivered into the bloodstream, transported and able to cross the blood-brain barrier (BBB) facilitating systemic labelling for whole-brain photoacoustic calcium imaging.

Given the importance of small-molecule delivery to the brain in drug discovery, there has been extensive research into the key properties which favour permeation of the BBB. The renowned Lipinski's "rule of five" was determined by computational analysis of a library of compounds which had passed to Phase II clinical trials and provide an indication of the likelihood of a molecule crossing the BBB by passive diffusion, without active transport by a membrane protein.[226] Compounds with molecular weight < 500 Da are more likely to diffuse passively across the BBB, able to occupy temporary gaps in the phospholipid bilayer of cell membranes. All rhodamine-based HaloTag ligands are comfortably over this threshold, but systemic brain labelling is possible

with such ligands.[208] Membrane barriers are hydrophobic so molecules must be sufficiently lipophilic to be able to cross them via diffusion. The octanol-water partition coefficient, logP, describes this key parameter, originally recommended to be < 5 but this includes the vast majority of small-molecule drugs and further studies have identified 1.5−2.5 as the ideal range.[227] Similarly regarding a molecule's hydrophilicity, possessing fewer hydrogen bond donors and acceptors is advantageous (< 5 donors, < 10 acceptors), and minimising the polar surface area. Furthermore, a net neutral or slightly positive electrostatic charge avoids repulsion from phospholipids and reducing molecular flexibility (< 5 rotatable single bonds) helps orientation into the organised membrane structure and decreases bulkiness.[228]

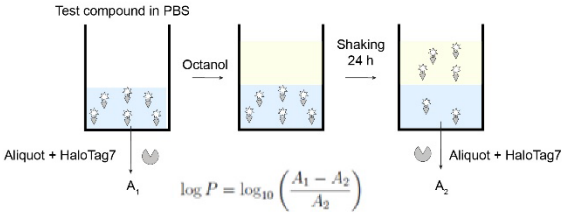
To investigate these properties in our series of acoustogenic ligands, I performed *in silico* prediction of their chemical properties using existing tools,[229, 230] to determine candidates with potentially superior bioavailability. A complication for these ligands is the open-closed equilibrium, which is influenced by the surrounding environment, and although the majority of molecules remain closed in solution and also when crossing a hydrophobic membrane, there will be a proportion of the open form too, so I calculated the properties for both (Figure 52a). The compounds all have similar masses around 700 Da so this cannot be a distinguishing factor. Molar refractivity describes the size and polarizability of a molecule and these ligands again exceed the limit.[231] The molecular flexibility is also high but only due to the HaloTag ligand, which contributes the majority of these rotatable bonds, even for the molecular rotor-based Malachite Green lactones. Moreover, the polar surface area and the number of hydrogen bond donors and acceptors are all within the recommended ranges. However, these molecules are likely too lipophilic, with predicted logP values for the predominant closed form > 5, aside from **22-HTL**, and even the open form of the dyes do not fall in the desired range for BBB-permeable molecules. To validate these predictions, I attempted to measure the logP values experimentally using the shake-flask method.[232] I had to adapt this method for our acoustogenic ligands because of their low absorption, so before and after partitioning between octanol and buffer, an aliquot was incubated with excess purified HaloTag7 protein (Figure 52b). Due to the high lipophilicity of these compounds, almost all of the ligands partitioned into the octanol phase and so the absorbance signal from the PBS layer was very low with no clear peak (Figure 52c). Therefore, the calculated logP value should be considered as the lower end of an approximation with logP > 3.8 for **14-HTL** and > 2.5 for **22-HTL**, which is consistent with the *in silico* predictions.

a.



Compound	X	R1	R2	Y	Molecular weight (Da)	H-bond donors	H-bond acceptors	Refractivity (m^3 mol^{-1})	TPSA ($Å^2$)	RBN	logP
Target					< 500	< 5	< 10	40 - 130	< 140	< 5	1.5 - 2.5
10-HTL	O	HN	H	F	734	3 (3)	6 (6)	187 (187)	115 (127)	15 (16)	8.2 (3.2)
14-HTL	H, H	N	H	F	676	1 (1)	7 (6)	176 (176)	80 (94)	17 (17)	6.3 (4.7)
15-HTL	H, H	N	H	F	700	1 (1)	7 (6)	181 (181)	80 (94)	17 (17)	6.5 (5.0)
17-HTL	H, H	N-OMe	H	F	760	1 (1)	9 (8)	194 (194)	99 (112)	19 (19)	5.8 (4.3)
20-HTL	H, H	N	Me	F	704	1 (1)	7 (6)	186 (185)	80 (94)	17 (17)	6.9 (5.4)
22-HTL	H, H	N-OH	H	F	732	3 (3)	9 (8)	184 (184)	121 (134)	17 (17)	4.5 (3.0)
23-HTL	H, H	N	H	H	658	1 (1)	7 (6)	176 (176)	80 (94)	17 (17)	6.1 (4.6)

b.



c.

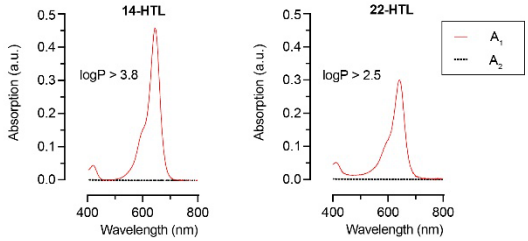


Figure 52: Drug-like properties of acoustogenic HaloTag ligands. (**a**) Table of predicted chemical properties of our acoustogenic ligands, and their chemical structures, using the Atomistica and Molinspiration tools,[229, 230] with the value for the closed form followed by the value for the open form in brackets: TPSA (topological polar surface area), RBN (number of rotatable bonds). Z = C(O)-HTL for the general chemical structure, full structures in Figure S2. (**b**) Method for determining the partition coefficient, logP. (**c**) Absorption spectra of **14-HTL** and **22-HTL** bound to HaloTag7 during the logP determination.

There are a plethora of *in vitro* BBB models suitable for screening dye permeability of which the most commonly used is the transwell model with an endothelial cell monolayer (e.g. Caco-2).[233] Whilst these cell-based models are a more accurate representation of the BBB with endothelial cells joined by tight junctions restricting passive transport, and active efflux transporters, their establishment and maintenance is costly and time-consuming. The parallel artificial membrane permeability assay (PAMPA), has also been used in the pharmaceutical industry for a lower cost, high-throughput drug screening alternative.[234] This method consists of a filter membrane coated with a brain lipid solution, which separates two aqueous layers, the donor solution of the compound to be tested and an acceptor solution. Incubation for >18 h allows sufficiently permeable compounds to cross the membrane and reach an equilibrium. The concentration of compound in the acceptor well can be determined via UV/Vis absorption using a plate reader and compared to the expected absorption if all of the compound had permeated the membrane.

For concentration determination of our ligands, HaloTag7 protein was again required. Initially I added the HaloTag7 protein to the acceptor solution after the incubation period but the signal was undetectable for the chromogenic dyes **JF635-HTL** and **14-HTL**. This highlights the same problem from the logP determination that these assays are designed for smaller, non-environmentally sensitive drug compounds. Increasing DMSO concentration or addition of a detergent (Pluronic F127) did not aid permeation, although the predominantly open fluorophore **JF549-HTL** was able to cross the membrane. However, addition of HaloTag7 protein directly to the acceptor well forced the equilibrium, encouraging more dye to cross the membrane, which could be then detected via absorption measurements (Figure 53a). Whilst this modification prevents comparison of the permeability rates with known BBB-permeable compounds, the values of our acoustogenic ligands could be compared with JF ligands that have already been used *in vivo*. A series of JF-HaloTag ligands were screened for bioavailability in HaloTag-expressing mice via a pulse-chase assay, where the first ligand was delivered by retroorbital injection before perfusion with a spectrally separable ligand 12-18 h afterwards.[208] The trend in the fraction of labelled protein in these experiments is conserved with the absorption ratio measured in my PAMPA assay, validating this method to screen dye bioavailability (Figure 53b).

Screening our best performing acoustogenic ligands revealed that the hydroxylated **22-HTL** was the most permeable, matching the absorbance ratio of the bioavailable dyes **JF552-HTL, JF669-HTL** and additionally **JF549-HTL** (Figure 53c). Plotting the extinction coefficient of the corresponding free dyes against absorption highlights **22-HTL** as an outlier, with the greater permeability facilitated by the OH groups rather than opening of the equilibrium. The three Malachite Green lactones **14-, 17**- and **23-HTL** presented greater absorption ratios than **JF585-HTL**, which has also been demonstrated to label neurons in the mouse brain via tail vein and intraperitoneal injections,[145] suggesting they could also achieve brain labelling via intravenous (IV) injections to some extent. **JF635-HTL** has been reported to have poor bioavailability[159] *in vivo* so the even less permeable ligands **10-, 15-** and **20-HTL**, which also had the highest predicted logP values, are unlikely to be able to achieve significant labelling. Given the positive impact of the OH groups, this is perhaps surprising for the pyrrole-xanthene with dual NH groups but the high logP value for the closed form suggests that the additional heterocyclic rings outweigh the hydrophilicity provided by the NH.

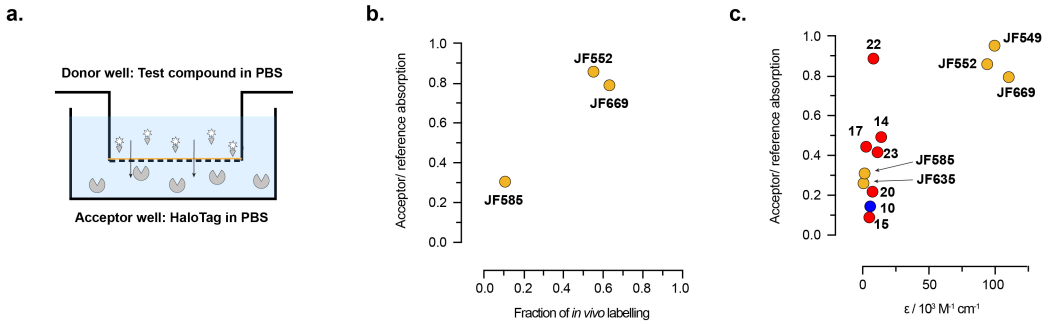


Figure 53: Screening the bioavailability of acoustogenic HaloTag ligands. (**a**) Principle of the PAMPA BBB assay with HaloTag ligands. (**b**) Comparison of the absorption ratio of Janelia Fluor ligands in the PAMPA BBB assay with the fraction of *in vivo* brain labelling achieved in mice from ref.[208] "HTL" has been omitted from the labels for clarity. (**c**) Plot of absorption ratio of acoustogenic and Janelia Fluor HaloTag ligands against extinction coefficient of the corresponding free dye.

III.6 Conclusion and future directions

In Chapter III, I have described the next steps towards a second-generation of acoustogenic calcium indicators, aiming to improve the properties of those developed in Chapter II. I designed, cloned and characterised four HaloCaMP affinity variants swapping the calmodulin-binding peptides in the original HaloCaMP1a and HaloCaMP1b proteins for ENOSP and DAPKP peptides. Characterisation of their absorption properties when labelled with our best performing acoustogenic ligands (**10-, 14-, 15**- and **20-HTL**) revealed improved affinities, mostly in the 50−300 nM range. However, these sensors compromised on dynamic range and PA signal, with very low extinction coefficients, therefore require further engineering to improve the collective properties. Such targeted engineering needs to screen large numbers of protein mutants and so I have pursued several approaches to develop a suitable high-throughput method. Using absorption as a readout from bacterial lysates was not sensitive enough due to the high background signal. Alternative approaches to engineer a fluorescence signal by following the bleaching of ADPA or developing a NanoLuc-based BRET reporter also did not work. For the former, the Malachite Green ligands were not sufficiently strong photosensitisers, whilst for the latter, ligand binding was impaired to the HaloTag domain. However, through our collaboration with the Stiel group, we are establishing a method for PAM colony screening of our calcium sensors, which we hope will be robust and sensitive to facilitate the reliable differentiation of the performance of individual mutants. Finally, I explored the physicochemical properties of the acoustogenic ligands through *in silico* predictions, experimental logP calculations and an artificial membrane permeability assay, and we can use these results to design more bioavailable ligands.

The varied performance of the affinity variants with the four different dye ligands highlights the importance of the specific dye-protein interactions to sensor function. Once a protein has been engineered in conjunction with a particular dye, then it is very difficult to match this performance with another ligand. Therefore, an extensive protein engineering approach of the HaloCaMP scaffold with one of our acoustogenic ligands needs to be started. Whilst the choice of dye is not yet clear-cut, especially if the most acoustogenic ligands indeed show limited bioavailability, which is further investigated in Chapter IV, we can already consider the approach on the protein side. The best option might be to have several starting points as we do not know which property will be easier to optimise: the absorption/ PA intensity, dynamic range or affinity. Overall, HaloCaMP1b exhibits the greatest photoacoustic turn-on with the majority of ligands so should be the principal starting scaffold. Additionally, I would suggest engineering of 1b_DAPKP because the affinity is tuned in the optimal range for 3 out of the 4 ligands, and shows reasonable sensitivity, albeit negative. The direction of the response could be reversed with targeted mutations. The HaloTag scaffold has been extensively explored for sensor design and the same positions are consistently targeted, especially 143 and 154, which result in the most performant sensors.[141] Therefore, the focus should be on optimisation of the linkers via site-saturation mutagenesis because

of their importance in conferring the conformational transformations and their influence on the protein's stability and surrounding packing of interfaces (Figure 54a).[119, 235] Moreover, subtle modifications should be introduced around the ligand binding site to promote favourable interactions with these novel Malachite Green lactone scaffolds, that would sit differently in the binding pocket compared to Si-rhodamines (Figure 54a). Resolution of a crystal structure of one of these dye ligands bound to the protein could help guide the design to identify key interaction residues as in the development of HaloCaMP with **JF635-HTL**.[143] However, the ligands often adopt multiple possible conformations, and especially with the molecular flexibility of the Malachite Green structure, rational mutations might not be possible, so an unguided approach could also be followed.

The next stage is to perform an initial site-directed mutagenesis test of the HaloCaMP protein scaffold using our promising colony screening method (Figure 54b). Mutations could be introduced into one of the linker regions, colonies grown, labelled with **14-HTL** and PAM performed before and after Ca^{2+} addition. Three colonies could be picked, with high, medium and low $\Delta PA/PA_0$, sequenced and purified. Absorption and photoacoustic characterisation of these purified proteins would then confirm whether the relative performance measured from their respective colonies is consistent with the protein variants' properties. Perhaps there may be too much variability, even after correction for expression via the EGFP signal, and so multiple plates with the same colonies may have to be grown, imaged and averaged to provide a reliable quantification. Our preliminary lysate measurements also suggested that PA signal may be less susceptible than absorption to background scattering so alternatively we could continue the development of the PA plate reader. The sole major adaptation required would be an automated stage to shift the plate from well-to-well, keeping the alignment of the laser and transducer. We measured the detection limit of 0.0012 OD at 650 nm using a reference dye ($NiCl_2$), which would correspond to a minimum ~300 nM concentration of **14**-HaloCaMP1b, very much within the achievable protein concentration range in our lysate assay (Figure 46a).

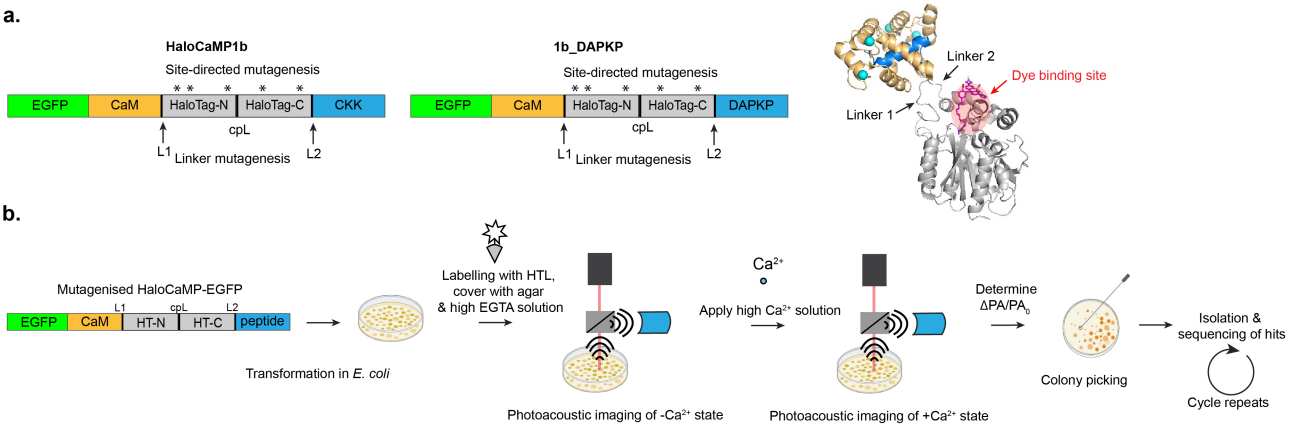


Figure 54: Protein engineering of the HaloCaMP scaffold for acoustogenic calcium sensors using colony screening. (**a**) Starting HaloCaMP constructs for protein engineering and target areas for mutagenesis highlighted on the HaloCaMP1b crystal structure (PDB 6U2M),[143] linkers (black arrows) and around the dye binding site (red area). (**b**) Pipeline for using PAM as the method for screening colonies of HaloCaMP mutants.

Predicting bioavailability and BBB permeability is very difficult and even more so when the compounds are environmentally sensitive. Nevertheless, the data presented in this chapter suggest that bioavailability of almost all of these compounds will be limited. The high logP values indicate that these dyes are too lipophilic and the addition of the hydroxyl groups to **22-HTL** drastically increased permeability in the PAMPA-BBB assay. Unfortunately, this modification provided a trade-off in sensitivity ($\Delta PA/PA_0 = 1.0$ with HaloTag7). Therefore, the design of novel acoustogenic ligands should focus on slightly increasing hydrophilicity whilst minimising the impact on the open-closed equilibrium. Displacement of OH by one methylene group away from the xanthene core may achieve this, or hydroxyl-functionalised 5- or 6-membered rings, which have small, additional closing effects on the equilibrium relative to azetidines (Figure 55a).[176] The incorporation of PEG

groups, widely used to improve small-molecule bioavailability, is another option also via the auxochromes to not impede HaloTag binding (Figure 55a). To reduce molecular weight, unsymmetrical compounds with only one PEG group could be made. I would additionally explore the effect of aliphatic heterocyclic groups, as in **24-**, **25-HTL** from Chapter II (Figure 55a). These groups increase the polarity, tune the equilibrium,[167] whilst perhaps providing enhancement in HaloTag-bound absorption, and quenching is still achieved via the free rotation of the Malachite Green lactone scaffold. For the pyrrole-xanthene family, analogous hydroxylated and PEGylated compounds could be tested with potentially positional effects (Figure 55b). Although substitution of the NH for O, S or Se would likely result in too closed dyes, *N*-modification could also be investigated (Figure 55b).

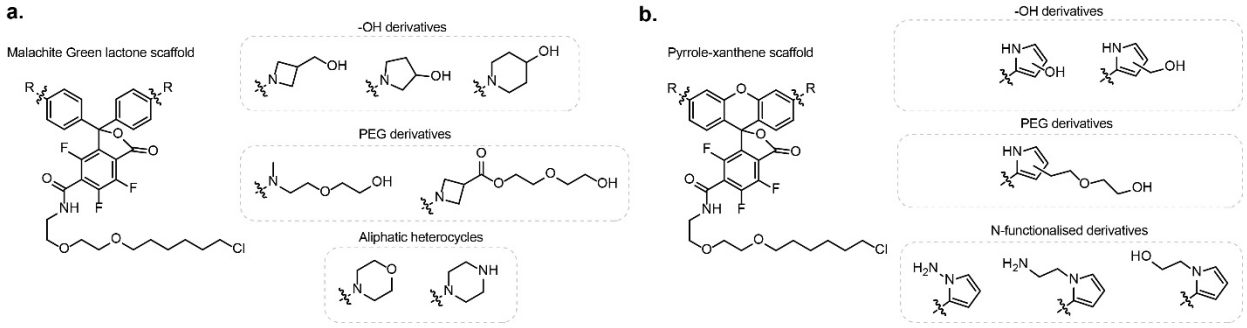


Figure 55: Modified acoustogenic ligand structures to improve bioavailability. (**a**) Malachite Green lactone derivatives, grouped into hydroxylated, PEGylated and aliphatic heterocyclic classes. (**b**) Pyrrole-xanthene derivatives, grouped into hydroxylated, PEGylated and *N*-functionalised classes.

Experimental validation of cell permeability and bioavailability in living systems is required of any new derivatives and also our current ligands. During the development of our calcium sensors, we have already started to assess these key properties. In Chapter IV, I shall discuss our cell labelling, phantom imaging, *ex vivo* and *in vivo* mouse labelling experiments with the first-generation of our chemigenetic photoacoustic calcium sensors. These results inform the next direction of our PA probe and sensor development as we incrementally improve our PA calcium sensors for *in vivo* imaging of neural activity.

Chapter IV: Towards photoacoustic neuroimaging in mice

The work presented in this chapter was conducted in collaboration with Nikita Kaydanov, Juan Boffi and Gretel Kamm (Prevedel group, EMBL), and Begoña Ugarte-Uribe (Deo group, EMBL). Juan and Gretel performed the AAV and dye injections, plus the sacrifice of the mice, brain perfusion and slicing. Nikita performed the photoacoustic tomography and fluorescence microscopy of our mouse brain samples, and also helped with the sacrifice, perfusion and slicing. Begoña prepared and imaged the U2OS cells and helped me to culture the spheroids.

Part of this work was also published in the aforementioned publication on my project: Cook, A.; Kaydanov, N.; Ugarte-Uribe, B.; Boffi, J. C.; Kamm, G. B.; Prevedel, R.; Deo, C.; Chemigenetic Far-Red Labels and Ca^{2+} Indicators Optimized for Photoacoustic Imaging, *J Am Chem Soc,* **2024**, *146* (34), 23963-23971. DOI: 10.1021/jacs.4c07080.[165] I adapted several figures from our publication for use in this chapter, and integrated additional unpublished data.

IV.1 Objective

The ultimate goal of this ambitious project is to achieve brain-wide PA calcium imaging in a living mouse. As we have seen from the previous chapter optimisation attempts to improve the first-generation sensors are ongoing, as their performance, whilst promising, is not yet sufficient for visualising activity of neuronal populations *in vivo*. Nevertheless, all novel sensors require several validation and optimisation steps to lay the groundwork for such experiments in the future. Following the example of GCaMP sensor development again, we set out to assess the performance of our new probes and sensors in an analogous series of experiments building up to *in vivo* mouse neuroimaging (Figure 56). First, the calcium indicators are characterised in purified protein solution, and we took this a step further by embedment in tissue-mimicking phantoms, to simulate scattering conditions in a realistic imaging set-up. Next, we aimed to assess our PA sensors' performance in different cultured cellular systems, a parallel to the evaluation of GCaMP performance in rat hippocampal neurons. Subsequently, transgenic mouse brain slices can be imaged.[111] Finally, our shared final goal with the GCaMPs is to express the calcium sensors in the mouse brain, and monitor transient calcium dynamics in response to stimuli.[112]

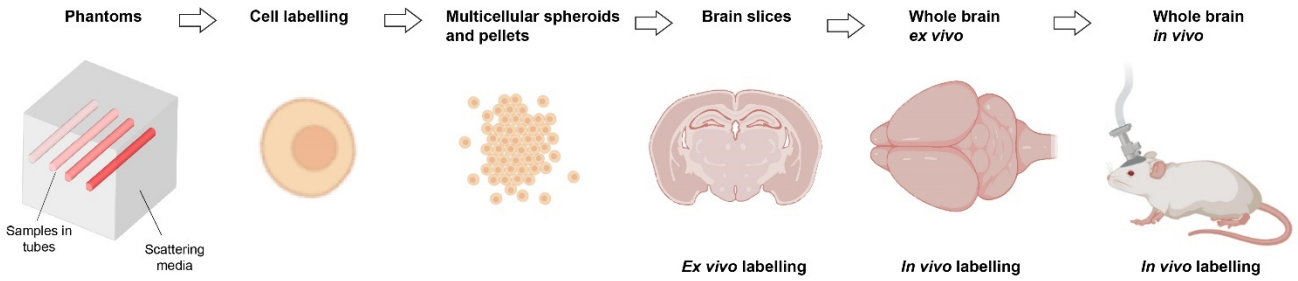


Figure 56: Pipeline of the establishment of photoacoustic neuroimaging in mice with our novel chemigenetic labels and sensors. Created with Biorender.

Several challenges were encountered during these steps, due to the use of a custom PAT set-up, and the additional complexity of dye labelling of the HaloTag-based protein scaffolds. We had to establish the PAT set-up (section IV.2) to provide an initial assessment into whether our probes can generate detectable photoacoustic signal in scattering conditions (section IV.3). Moreover, the cell permeability of the acoustogenic ligands needed to be evaluated, via labelling of cytosolically expressed HaloTag proteins (section IV.4). Once cellular labelling had been confirmed, we could image clusters of cells via PAT and HaloTag-expressing mouse brain slices after *ex vivo* labelling (section IV.5). Finally, it was important to explore dye delivery methods and investigate the *in vivo* labelling efficiency of the acoustogenic ligands (section IV.6).

IV.2 All-optical photoacoustic tomography

Given the size of the mouse brain (~13.2 x 11.4 x 8.0 mm),[236] PAI across the entirety of this organ requires tomography, as discussed in Chapter I, thanks to its sufficient field of view and penetration depth below 1 cm. Nikita Kaydanov continued the work of Jakub Czuchnowski (Prevedel group, EMBL),[237, 238] to establish an all-optical Fabry-Pérot-based photoacoustic tomographic set-up for biological imaging. As for the spectroscopy measurements, a tuneable pulsed laser excites the sample, but the beam passes through a transparent Fabry-Pérot interferometer, which functions as the ultrasound sensor instead of commonly used piezoelectric transducers, before illuminating the sample (Figure 57).[239] The interferometer consists of two mirrors separated by a polymer film, whose thickness is varied in response to the acoustic waves emitted by the sample (Figure 57 inset). An interrogation laser beam at infrared wavelengths, which the mirrors reflect, is raster scanned across the surface of the interferometer and from the variation at each point, due to the changes in reflectivity of the mirrors, the 2D spatiotemporal distribution of photoacoustic waves can be inferred. Reconstruction algorithms, for example, back projection or time-reversal, generate a 3D PA image from the collected data. Such all-optical set-ups have several advantages, over those which use ultrasound detectors, for *in vivo* imaging.[240] Higher spatial resolution can be attained as well as greater sensitivity at higher resolution, and these sensors possess a broader detection bandwith. In addition, the Fabry-Pérot is optically transparent, providing flexibility in the orientation of the set-up, including an inverted configuration, allowing direct positioning onto the sample, and possible combination with a fluorescence microscope for multi-modal imaging. This could facilitate validation of calcium dynamics reported by our PA calcium sensors, with a fluorescent GECI, for example. There is additionally more space for imaging a freely moving animal because these interferometers are smaller than bulky transducer cups, ideal for neuroimaging purposes. Finally, fast acquisition times can be achieved, which was the previous limitation of the all-optical method, by using high pulse repetition frequency excitation lasers and multiplexing the laser beams for sensor readout.[241]

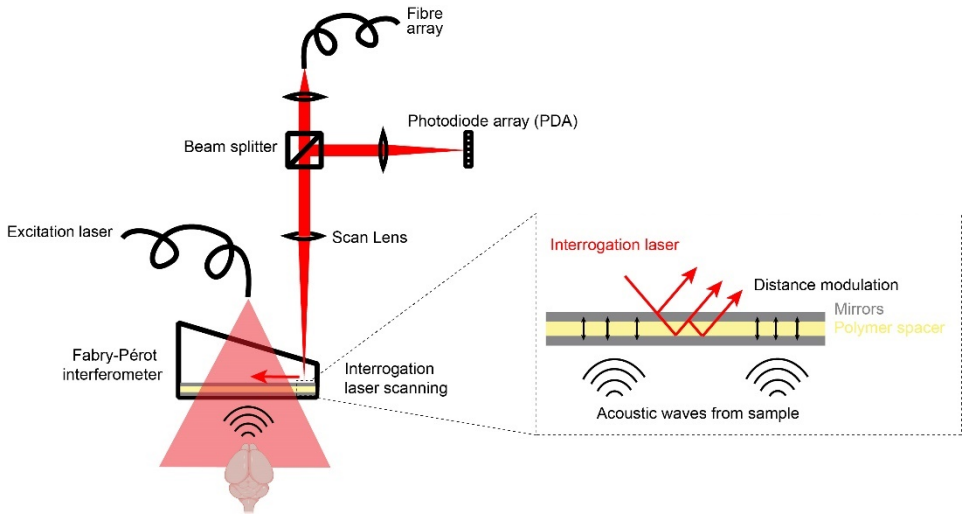


Figure 57: The all-optical Fabry-Pérot based photoacoustic tomographic set-up. Overall schematic of the set-up. Inset: principle of the Fabry-Pérot interferometer.

IV.3 Photoacoustic tomography in tissue-mimicking phantoms

With our series of calcium indicators in hand, fully characterised by spectroscopy, the next major question was whether they provide sufficient PA signal to be visualised in realistic imaging conditions, using PAT. This can be initially tested in tissue-mimicking phantoms, the simplest of which consists of dye solutions in tubes embedded in scattering media. However, initial imaging of readily available silicon tubing was unsuccessful with weak signal and imaging artefacts. The tube material can have a large impact on the PA signal,[242] so we

tried to make tubeless phantoms using gel wax. Gel wax has the advantage of durability, longevity, and greater transparency compared to commonly used agar-based phantoms, plus possesses similar acoustic attenuation and speed of sound properties to tissue.[243] However, removal of the glass capillaries after setting resulted in rupturing of the gel and leakage of added dyes into the phantom (Figure S22). Therefore, returning to the tubes, we assessed the signal of a far-red reference PA dye, methylene blue,[50] across different tube materials (Figure S23) and settled on the use of 1 mm diameter polyethylene (PE) tubes fixed at different depths, using a 3D printed holder, in a coupling medium of water (non-scattering), 10% or 60% v/v milk/H_2O mixtures, of which the latter has scattering coefficients representative of the mouse brain (Figure 58a).[244-246] Solutions of purified protein labelled with acoustogenic ligand were loaded into the tubes and placed at the desired depth away from the Fabry-Pérot sensor.

To evaluate the PA signal intensity of our probes against existing reporters, we compared the depth-dependence of the signal of **14**-HaloTag7 against mIFP,[130] the chromophore component of NIR-GECO1.[129] **14**-HaloTag7 could be detected down to 1.2 cm in the 60% milk phantom, sufficient to cover the depth of the whole brain, and deeper than the 0.8 mm detection limit for mIFP in both 10% and 60% milk phantoms (Figure 58b, S24). Side-by-side in adjacent tubes, we detected ~6-fold greater signal from **14**-HaloTag7 (Figure 58c). Varying concentration of **14**-HaloTag7 at 5 mm depth, which would be deep in the mouse brain, revealed the sensitivity limit of ~1.5 µM (Figure 58d, e). This was determined by the intersection of the noise floor and the linear fit of the signal intensities of **14**-HaloTag7, as a function of concentration, between 5−50 µM. AAV-induced calcium indicator protein expression levels, such as GCaMP3, in neurons, are usually in the tens of micromolar range, which suggests this is easily achievable in terms of the protein tag, with sufficient dye delivery being required for labelling.[247, 248]

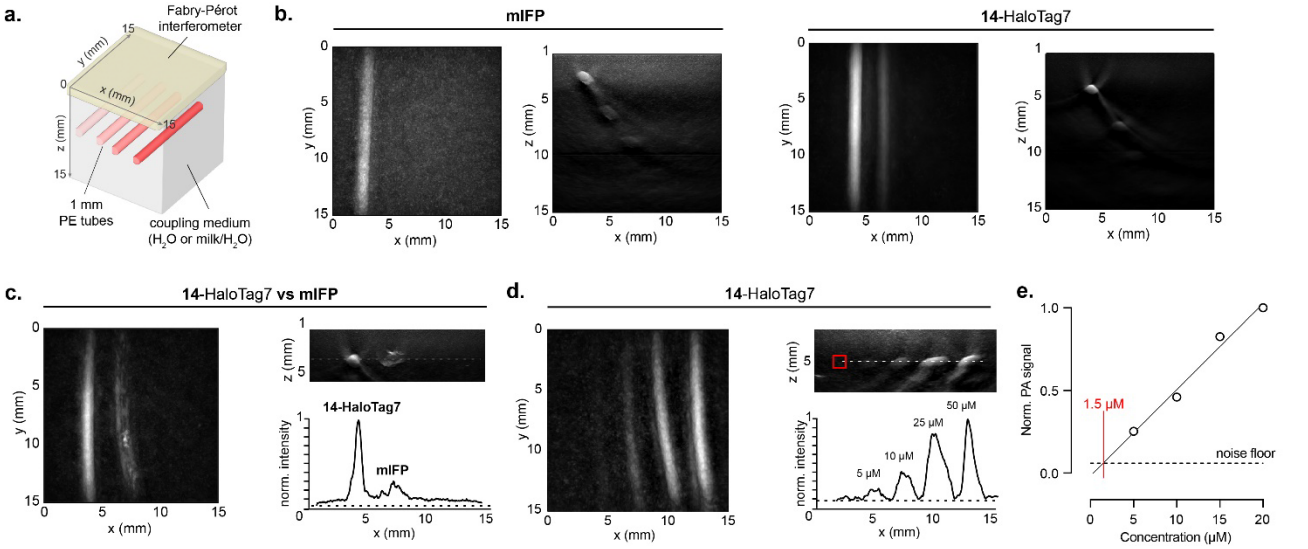


Figure 58: PAT imaging of solutions of our chemigenetic labels in tissue-mimicking phantoms. (**a**) Schematic representation of the tissue-mimicking phantom set-up showing relative position of the polyethylene (PE) tubes and the Fabry-Pérot interferometer; tubes were filled with a 50 µM protein solution, and immobilised in a coupling medium of 60% v/v milk/H_2O. Representative PAT images of: (**b**) mIFP and **14**-HaloTag7 at different depths, representative of n = 2 images; (**c**) side-by-side comparison of **14**-HaloTag7 (left tube) and mIFP (right tube) at ~3.5 mm depth; (**d**) **14**-HaloTag7 at 5, 10, 25 or 50 µM, at ~5 mm depth, representative of n = 2 images. (**e**) Plot of detection limit determination of **14**-HaloTag7 in the 60% v/v milk/H_2O scattering conditions: linear fitting of the maximum intensity values of the tubes against concentration, and the detection limit was determined by the intersection with the noise floor. The noise floor was estimated as a mean of the background plus standard deviation of the background in 1 mm^3 shown with the red ROI in panel (**d**). For images (**b-d**): PAT was performed at $\lambda_{PA} = \lambda_{max}$ for each probe; the (xy) images are maximum intensity projections (MIPs) and (xz) images are MIPs of the reconstructed side view; all images are displayed to their respective minimum and maximum intensities; the Fabry-Pérot interferometer is located on the z = 0 plane; intensity measurement was performed using a line profile (250 µm thick) on the (xz) MIPs, and quantification displayed below the side view.

For the calcium sensors, we imaged the Ca^{2+}-saturated and Ca^{2+}-free states in adjacent tubes to determine calcium sensitivity (Figure 59a, S25). The $\Delta PA/PA_0$ values of **14**-HaloCaMP1a, **14**-HaloCaMP1b and NIR-GECO1 were retained from spectroscopy, with larger positive PA responses for our two sensors. Imaging the Ca^{2+}-saturated state of these three reporters side-by-side resulted in >2-fold higher intensity for **14**-HaloCaMP1a, and similar signal for **14**-HaloCaMP1b and NIR-GECO1 (Figure 59b, S25). Perhaps the greater photostability of **14**-HaloCaMP1b is responsible for closing the signal intensity gap with NIR-GECO1 from the spectroscopy measurements, due to the >40-fold longer exposure time in the PAT imaging.

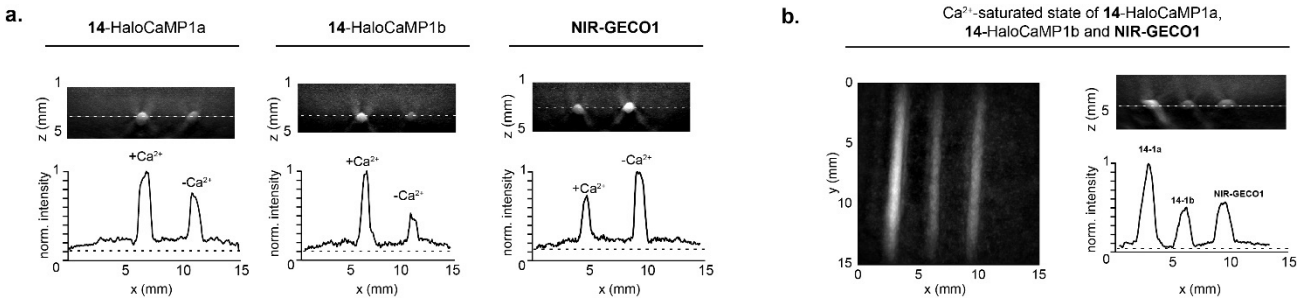


Figure 59: PAT imaging of solutions of our chemigenetic calcium sensors in tissue-mimicking phantoms. Representative PAT images of: (**a**) **14**-HaloCaMP1a, **14**-HaloCaMP1b or NIR-GECO1 in the calcium-saturated and the calcium-free states, at ~3.5 mm depth; images are representative of n = 3 replicates; (**b**) the calcium-bound states of **14**-HaloCaMP1a **(14-1a**, left tubes), **14**-HaloCaMP1b **(14-1b**, middle tubes) and NIR-GECO1 (right tubes); images are representative of n = 2 replicates. For images (**a-b**): PAT was performed at $\lambda_{PA} = \lambda_{max}$ for each probe; the (xy) images are maximum intensity projections (MIPs) and (xz) images are MIPs of the reconstructed side view; all images are displayed to their respective minimum and maximum intensities; the Fabry-Pérot interferometer is located on the z = 0 plane; intensity measurement was performed using a line profile (250 µm thick) on the (xz) MIPs, and quantification displayed below the side view.

Overall, these results were encouraging for visualisation of our photoacoustic labels and calcium sensors in brain tissue, as we could detect them several millimetres deep in scattering conditions representative of the mouse brain, and observe calcium-dependent PA signal differences with the labelled HaloCaMP sensors. Therefore, the proof-of-principle of our approach of engineering acoustogenicity with quenched rhodamines derivatives and HaloTag-based protein scaffolds was validated, and we proceeded with imaging of biological samples.

IV.4 Cellular labelling and imaging with acoustogenic ligands

IV.4.1 Cell permeability

Before performing PAT imaging of complex biological systems, it was essential to confirm the cell permeability and labelling efficiency of our acoustogenic ligands in living cells. U2OS cells expressing HaloTag7-EGFP in the cytosol were subjected to a pulse-chase assay, involving initial incubation with an acoustogenic ligand (**10-**, **14-**, **15-**, or **22-HTL**), washing, followed by addition of the bright, cell permeable competing ligand **JF549-HTL**.[145] As expected, the positive control showed bright fluorescence in the JF549 channel, co-localizing with EGFP expression (Figure 60a). Pre-labelling with the acoustogenic ligands **10-HTL**, **14-HTL** and **15-HTL** resulted in very minimal fluorescence in the JF549 channel, with a JF549/EGFP signal ratio of ~5% relative to the positive control, suggesting that these compounds are cell permeable, and were able to bind to ~95% of the cytosolically expressed HaloTag protein (Figure 60a, b). **22-HTL** showed slightly lower labelling ~85%, which is in contrast to its high predicted permeability from the bioavailability assay, and might be due to differences in labelling kinetics or sample quality (Figure 60b).

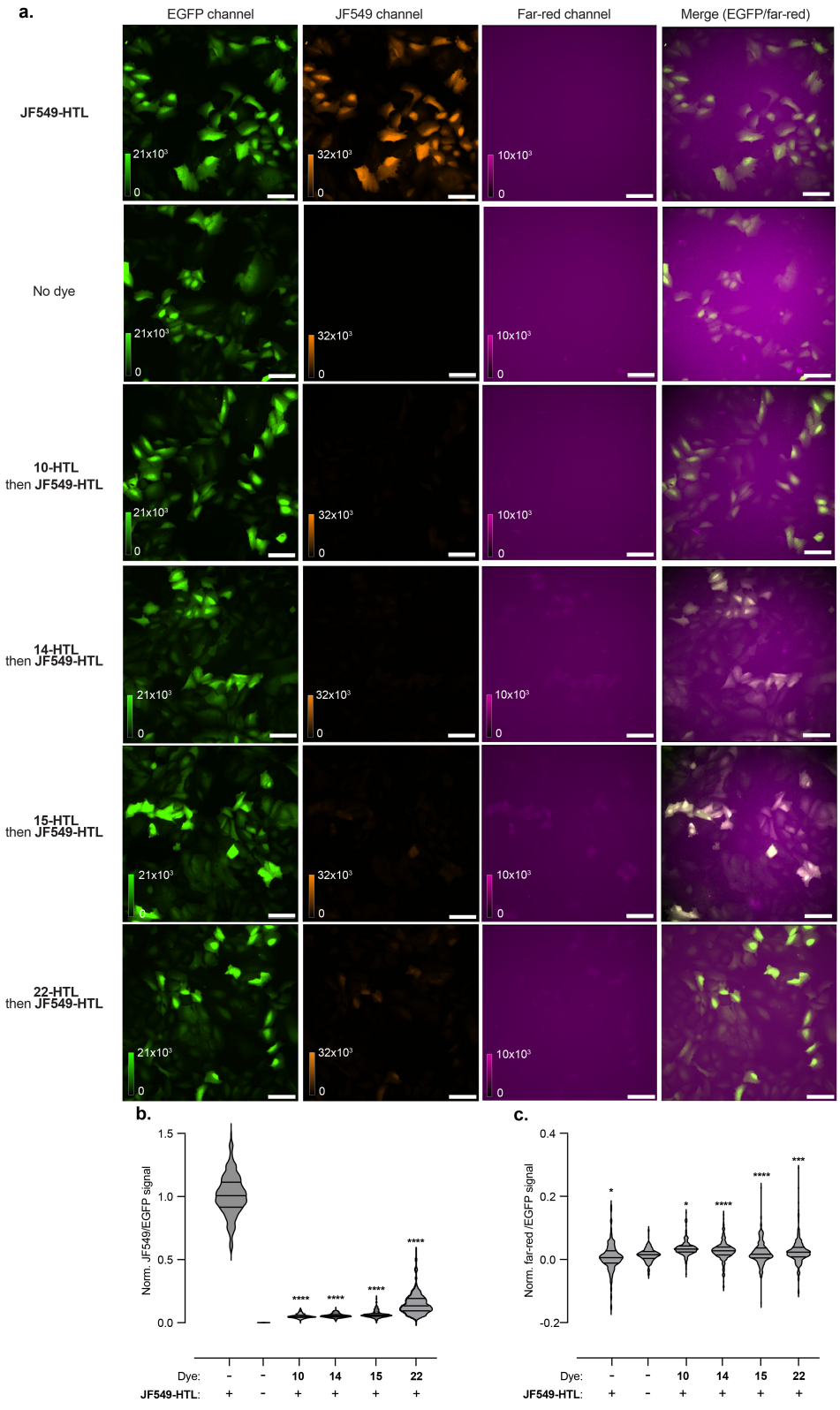


Figure 60: Evaluating cell permeability of the acoustogenic ligands. (**a**) Representative widefield fluorescence images of live U2OS cells expressing HaloTag7-EGFP labelled with acoustogenic dye ligands (1 µM), followed by **JF549-HTL** (0.5 µM) to assess labelling efficiency. From left to right: fluorescence in EGFP, JF549 and far-red channels; overlay of the EGFP and far-red channels; representative images from 2 wells for each condition, 4 FOVs per well; scale bars: 100 µm. (**b**) Normalised fluorescence ratio of the JF549/EGFP channels. "HTL" was omitted from the labels for clarity.. (**c**) Normalised fluorescence ratio of the far-red/EGFP channels; N = 200-300 cells for each condition. Statistical test: unpaired t-test vs the unlabelled control value ****p < 0.0001, ***p < 0.0005, *p < 0.02.

In the far-red channel, very weak fluorescence signal was detected due to the low quantum yields of the dyes, and the long exposure times result in high background, so colocalization of the EGFP and far-red channels are unfortunately not so informative in terms of labelling specificity (Figure 60a, c). Nevertheless, quantification reveals a significantly greater far-red/EGFP fluorescence ratio for **14-**, **15-** and **22-HTL** than the unlabelled control (Figure 60c). **22-HTL** has a higher p-value, either because of reduced labelling or lower brightness, given the 2-fold lower extinction coefficient when HaloTag-bound compared to **14-HTL**. The further red-shifted excitation and emission spectra of **10-HTL** were suboptimal for the filter cube used so the far-red fluorescence is minimal, as for **JF549-HTL**.

IV.4.2 PAT imaging of multicellular spheroids and pellets

These 2D cultured cells unfortunately cannot be imaged with PAT to further confirm labelling with our acoustogenic ligands. Instead, a 3D sample is required, to provide sufficient signal. Therefore, Begoña and I attempted to form spheroids using the same U2OS cells, expressing HaloTag7-EGFP.[249] Using pre-coated specialised well plates enabled formation of spheroids, but these proved too fragile, breaking apart during manipulation and transfer to an imaging plate. Attempts to grow them directly in the imaging plates or chambers failed to form spheroids. Alternatively, I prepared *E. coli* pellets, expressing HaloTag7 which were labelled in suspension with **14-HTL** (Figure 61a) before mounting onto the Fabry–Pérot for imaging. The concentration of **14-HTL** in the pellets could be quantified in the range 50−80 µM, with labelling clearly visible by the intense blue colour of the labelled pellets. However, consistent mounting and imaging of the fragile pellets proved a challenge. Embedding in a 2% agar gel, revealed strong signal from the labelled pellets in comparison to the unlabelled cells, but with high variability because of dispersion of the pellets due to pipetting into water, the acoustic coupling medium (Figure 61b). Sandwiching of pellets in low-melting point agarose, even after flash-freezing in liquid nitrogen and/or freeze-drying, also failed to provide robust measurements due to their fragility (Figure 61c). Adding the pellets into the same tubes as for the original phantoms also did not provide a strong signal, perhaps because of air gaps within the cell suspension, disrupting the acoustic coupling (Figure 61d). Direct loading onto the surface of the Fabry-Pérot provided a strong signal but the dispersion of the pellet on the surface was variable (Figure 61e).

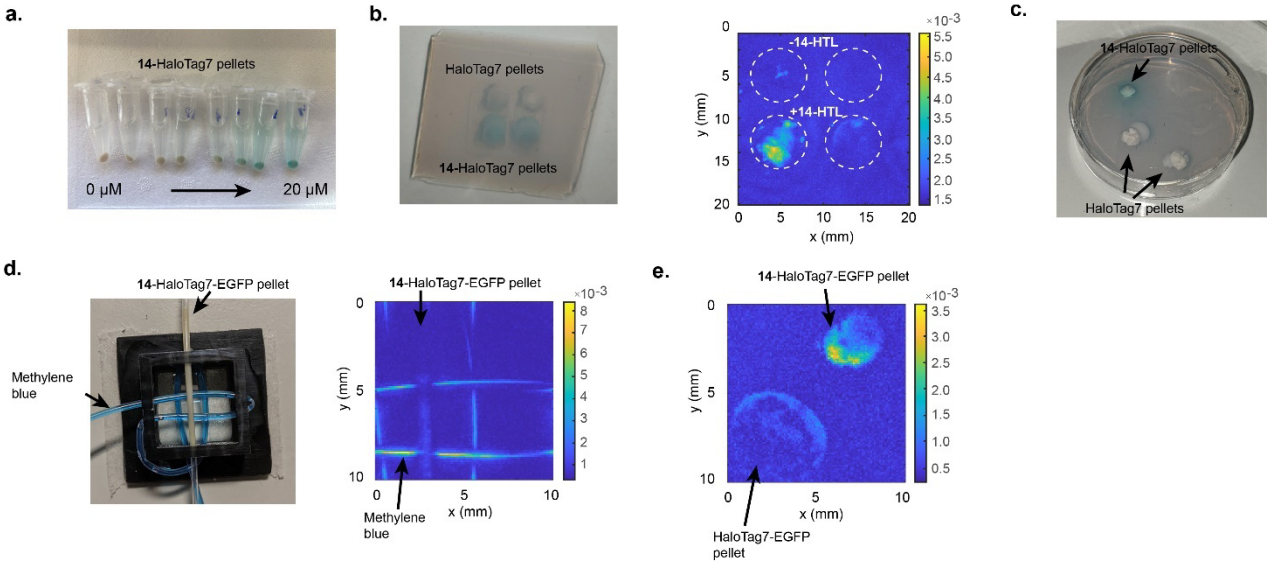


Figure 61: Photoacoustic tomography of HaloTag *E. coli* pellets labelled with 14-HTL. (**a**) Image of HaloTag7 pellets labelled with **14-HTL** at increasing concentrations in the suspension up to 20 µM. (**b**) Agar well phantom loaded with HaloTag7 pellets, two unlabelled and two labelled with **14-HTL**, photo of phantom (left), MIP (xy) of PAT image (right). (**c**) Gel wax "sandwich" phantom of two unlabelled, and one **14-HTL** labelled HaloTag7 pellet. (**d**) PE tube loaded with a **14-HTL** labelled HaloTag7-EGFP pellet, above a tube filled with methylene blue solution (0.27 mM in H_2O): photo of phantom (left), MIP (xy) of PAT image (right). (**e**) MIP (xy) of PAT image of **14-HTL** labelled and unlabelled HaloTag7-EGFP pellets loaded directly onto the Fabry-Pérot interferometer. PAT was performed at λ_{PA} = 646 nm.

IV.5 Brain slice labelling

Rather than persisting with optimisation of mounting these bacterial pellets, as we had demonstrated cellular labelling with our acoustogenic dyes, and generation of strong PA signal in tissue-mimicking conditions with the phantoms, we decided to move forward and attempt to label mouse brain tissues. To deliver the HaloTag anchor, mice were stereotaxically injected into the righthand hippocampus with an AAV1-HaloTag7-EGFP virus, controlled by the synapsin promoter. After 4 weeks to allow integration of the viral DNA and neuronal-specific expression of HaloTag7-EGFP, the mice were sacrificed, brains perfused with PBS, removed and sliced in coronal sections (~300 µM thick). The slices were labelled with dye in PBS and washed before evaluation by fluorescence microscopy and PAT. Using fluorescence microscopy, before labelling only the EGFP fluorescent signal from HaloTag7-EGFP expressing neurons in the hippocampus could be observed (Figure 62a). Specific labelling of this hippocampal region with **14-HTL** or **15-HTL**, where HaloTag7 was being expressed by the neurons, was evident by the correlation of the dim, far-red fluorescence from the dye with the EGFP fluorescence (Figure 62a, b). PAT was then performed on the same slices, and the distinctive shape of the labelled hippocampi was clearly visible in the images of slices labelled with **14-HTL** or **15-HTL** due to the strong PA signal (Figure 62a, b). Slices with minimal HaloTag7-EGFP expression showed no specific labelling with the acoustogenic ligand, evidenced by negligible far-red fluorescence and PA signals (Figure 62c). Quantification of the far-red/EGFP fluorescence ratio in the hippocampus of brain slices labelled with these two dyes showed a nonsignificant difference so relative labelling efficiency could not be deduced (Figure 62d).

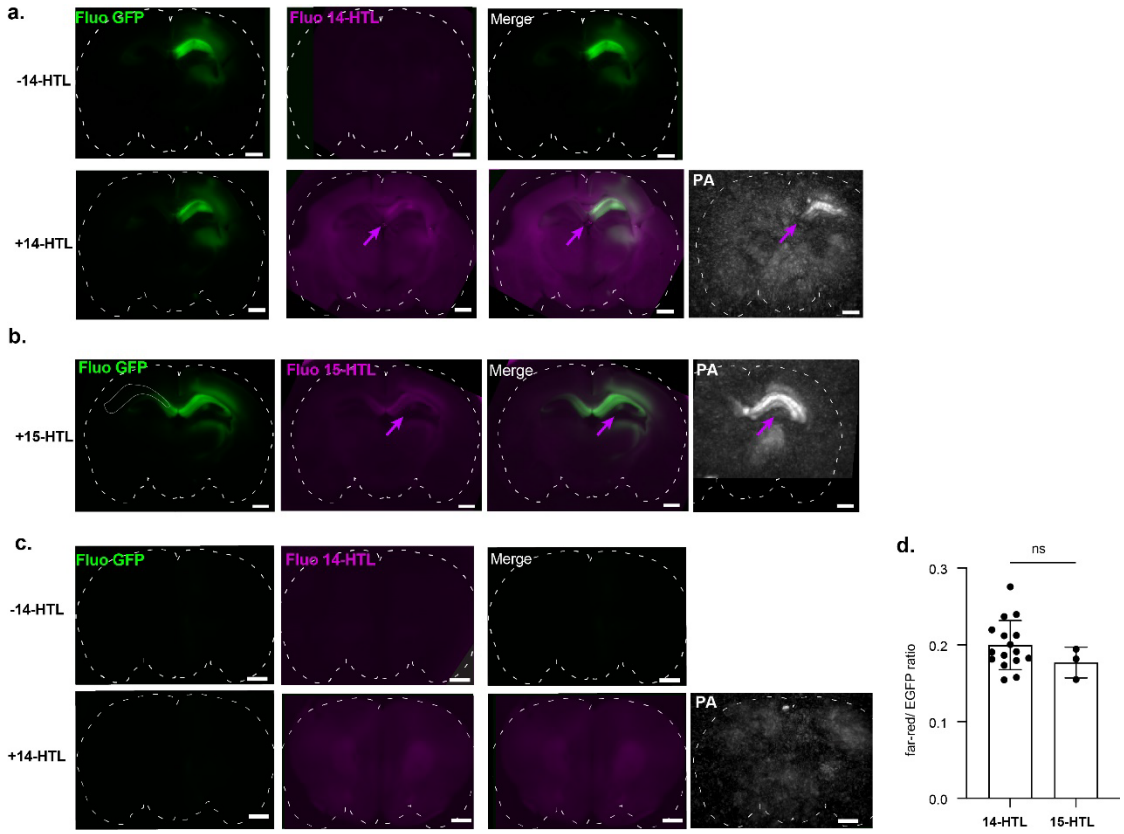


Figure 62: Fluorescence microscopy and photoacoustic tomography of mouse coronal brain slices labelled *ex vivo* with 14-HTL and 15-HTL. (**a-c**) From left to right: fluorescence images of EGFP channel; far-red channel; overlay of EGFP and far-red channels; PAT MIPs (~2 mm thickness) at $\lambda_{PA} = \lambda_{max}$; scale bars: 1 mm. (**a**) Slices with high hippocampal HaloTag7-EGFP expression before and after bath labelling with **14-HTL**, magenta arrows highlight signals from the bound dye ligand, images are representative of n = 16 slices, from n = 3 mice. (**b**) Slices with high hippocampal HaloTag7-EGFP expression before and after bath labelling with **15-HTL**, images are representative of n = 3 slices, from n = 1 mouse. (**c**) Slices with low hippocampal HaloTag7-EGFP expression before and after bath labelling with **14-HTL**, images are representative of n = 4 slices, from n = 3 mice. (**d**) Far-red/EGFP fluorescence ratio for brain slices labelled with **14-HTL** and **15-HTL**. Bars show mean and standard deviation. Statistical test: unpaired t-test, ns: non-significant.

Additionally, I examined **10-HTL** and **20-HTL** for labelling brain slices, but their bathochromic wavelengths lay outside of the far-red filter cube of the widefield fluorescence microscope. Confocal imaging had to be used which resulted in inhomogeneities because the curled slices, after labelling at 37°C, were not completely in the focal plane (Figure S26). Nevertheless, imaging at different depths focused on the hippocampus robustly supported the specific labelling for **10-HTL** in particular (Figure 63a, b). The labelling specificity of **20-HTL** was confirmed by the correlation of the PA signal with the widefield EGFP fluorescence image (Figure 63c). However, the signal from the **20-HTL** labelled brain slices was weaker than with **14-HTL** or **15-HTL**, and **10-HTL** was undetectable. This was due to a combination of the lower absorption of these dyes HaloTag conjugates (Figure 34), and the lower laser power at red wavelengths (~60% for **20-HTL** at 670 nm, ~40% for **10-HTL** at 695 nm, relative to the power at 646 nm for **14-HTL**). Consequently, with the current implementation of our PAT set-up and first-generation of PA chemigenetic probes, **14-HTL** or **15-HTL** are the best performing acoustogenic ligands for labelling brain tissue.

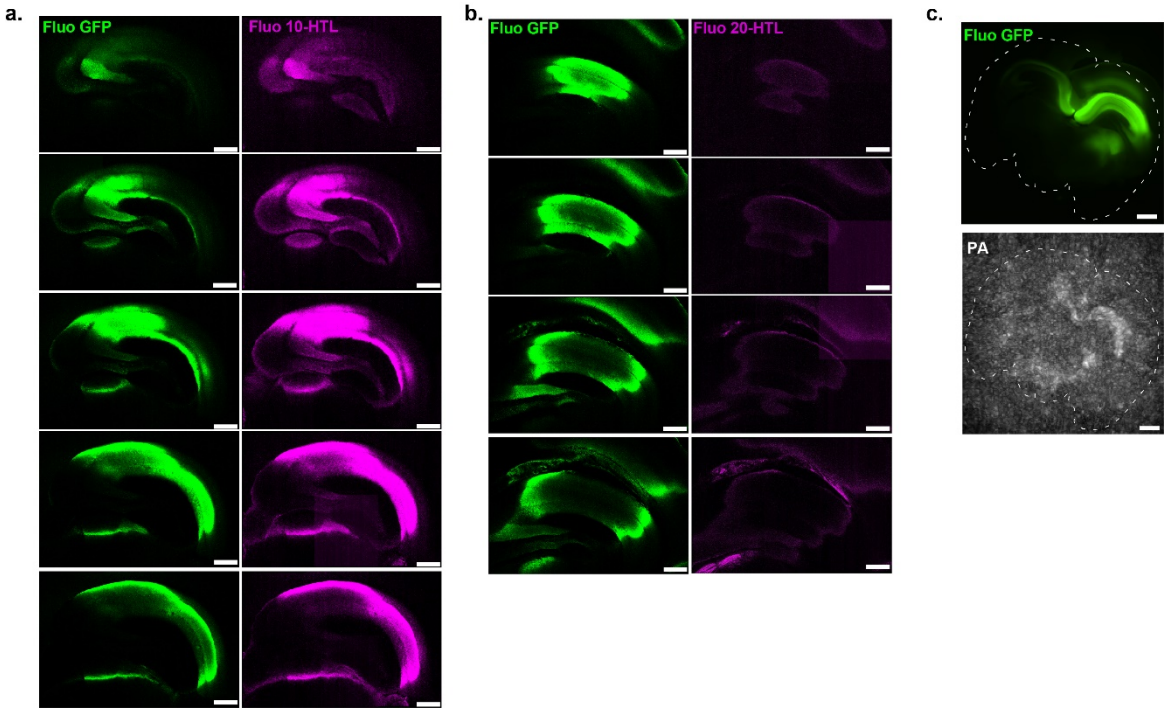


Figure 63: Fluorescence microscopy and photoacoustic tomography of mouse coronal brain slices labelled *ex vivo* with 10-HTL and 20-HTL. (**a-b**) Confocal microscopy z-stacks of the hippocampus after labelling with (**a**) **10-HTL** or (**b**) **20-HTL**, EGFP channel (left), far-red channel (right), scale bar: 500 µm. (**c**) Widefield fluorescence image (EGFP channel) and PAT image of a slice after labelling with **20-HTL**; λ_{PA} = 670 nm, scale bar: 1 mm.

IV.6 *In vivo* labelling

The following major experiment was to determine the labelling efficiency of our acoustogenic ligands *in vivo,* which we could again evaluate via a combination of widefield fluorescence imaging and PAT of the brain *ex vivo*. The labelling efficiency depends not only on the dye, but also the delivery method and so we first wanted to deduce whether significant labelling could be achieved via intravenous (IV) delivery, and if not then we could circumvent the blood-brain barrier via stereotaxic injection to the brain. The same AAV1-HaloTag7-EGFP used for the brain slices was injected into the cortex, hippocampus and thalamus of 4 mice, before **15-HTL** was delivered via either a tail vein (IV) or intracerebroventricular (ICV) injection. This preliminary dye delivery method assessment was performed before I ran the bioavailability assay and, at this stage, we thought **15-HTL** would be the most suitable choice because of its lower PA signal from the unbound dye, and hence superior acoustogenicity to **14-HTL**, providing a higher ceiling for improvement during subsequent engineering stages. The mice were sacrificed 24 h after dye injection, brains perfused with PBS/4% PFA and sliced 100 µM thick.

Widefield fluorescence microscopy revealed strong neuronally localised HaloTag7-EGFP expression, especially in the hippocampus (Figure 64a, S27). Gratifyingly, we also observed some weak far-red signal in individual neurons after ICV injection, which colocalises with the HaloTag7-EGFP expression and supports specific labelling with **15-HTL** (Figure 64a-c). The labelling was not homogenous though, evidenced by the large variation between neurons in each ICV labelled mouse, for example, slices #3 and #4 have stronger colocalization than #1 and #2 (Figure 64). Moreover, it was difficult to conclusively determine labelling efficiency, due to the very low brightness of the dyes, which could not be validated via PAT as no signal was detected from these thin slices. Nonetheless, quantification of the fluorescence signal from individual neurons strongly suggests that labelling only occurred for the ICV injections, with a 3.2-fold greater average far-red/EGFP signal (Figure 64d, e).

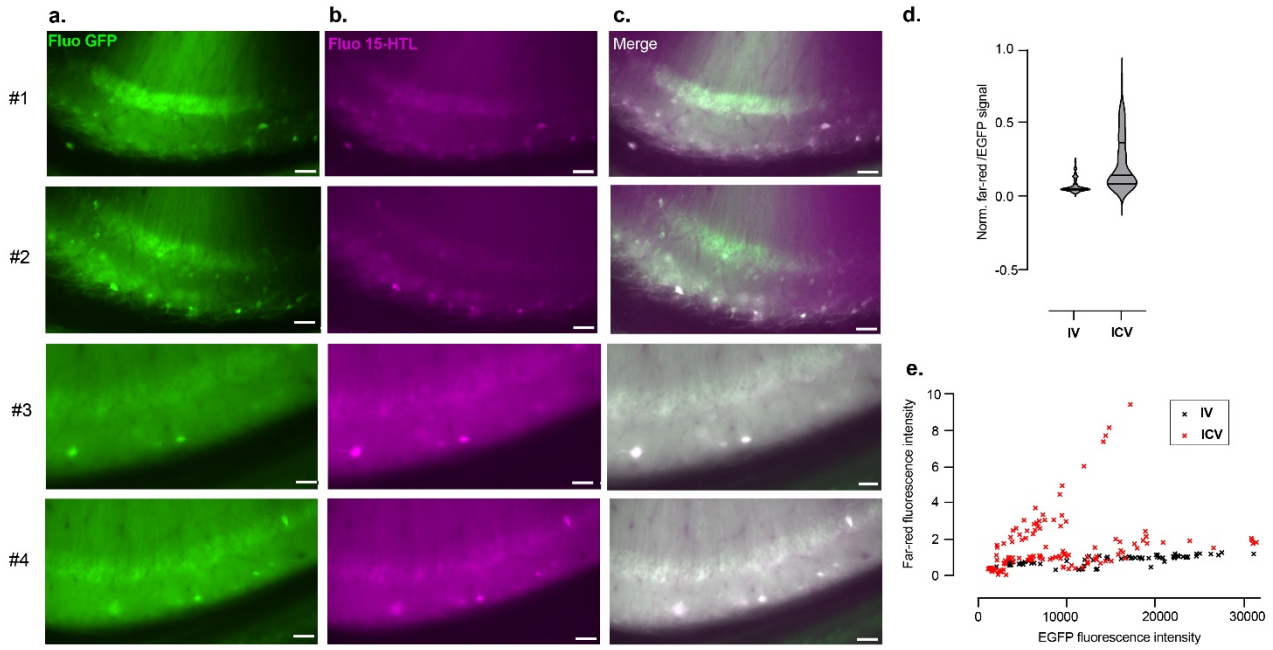


Figure 64: Evaluating neuronal labelling efficiency after delivery of 15-HTL via tail vein (IV) or intracerebroventricular (ICV) injection to HaloTag7-EGFP expressing mice. (**a-c**) Representative widefield fluorescence images of 4 coronal slices of a mouse brain expressing HaloTag7-EGFP and labelled with **15-HTL** delivered through intracerebroventricular injection *in vivo*; (**a**) EGFP channel; (**b**) far-red channel; (**c**) overlay of the EGFP and far-red channels; scale bars: 50 µM. (**d**) Normalised fluorescence ratio of the far-red/EGFP channels for mice labelled via IV or ICV injection; N = 32-89 neurons for each mouse. (**e**) Plot of far-red against EGFP fluorescence intensity for individual neurons in brain slices labelled either by IV or ICV with **15-HTL**.

Therefore, to increase the likelihood of sufficient neuronal labelling for imaging by avoiding the BBB, we continued to deliver our acoustogenic ligands via ICV injection. Given the undetectable PA signal in the initial brain slices, we focused on using our ligand with the strongest HaloTag-bound PA signal, **14-HTL**, for whole *ex vivo* brain PAT imaging. The control brain (without viral or dye injection) showed no photoacoustic signal, aside from the strong signal from the melanin-containing olfactory bulbs (Figure 65a, S28a). However, brains labelled with **14-HTL** delivered via ICV into the left lateral ventricle, in the opposite hemisphere to where the virus was injected, displayed variable results, which demonstrate the potential PA application of these ligands but underline current limitations in bioavailability and dye delivery. One brain displayed strong, specific PA signal from the right hemisphere in the region corresponding to the viral injection site, which was absent in the unlabelled brain (mouse #1, Figure 65b, S28b). To confirm that the signal arose from the HaloTag-bound ligand, the brain was fixed, sliced into ~300 µm slices and imaged via fluorescence microscopy. Using the mouse brain atlas,[250] the corresponding slices were paired with their coronal section in the PAT reconstruction. The PAT signal conclusively correlates strongly with the far-red fluorescence, which is predominantly localised to the hippocampal region, where the bright EGFP fluorescence indicates high HaloTag7 expression (Figure 65c, d,

S29). However, negligible labelling occurred in the thalamus despite high HaloTag7-EGFP expression, suggesting that the bioavailability of the dye needs to be improved (Figure 65c, d, g, S29). The importance of this parameter was emphasised as we were unable to replicate this strong degree of *in vivo* labelling in the following two mice (mice #2, #3). Slicing and fluorescence microscopy confirmed labelling, albeit with ~6-fold lower far-red/EGFP fluorescence intensities compared to the strongly labelled mouse (Figure 65e-g). PAT imaging of the whole brains and the partially labelled slices alone could not detect any signal, confirming the limited labelling. We also tried delivering the dye into the ventricle in the same hemisphere where the virus was injected to reduce diffusion distance to the target HaloTag7-EGFP expressing neurons, but no improvement was observed. Therefore, whilst the resounding success of neuronal labelling and PAT imaging in one mouse was very promising, the variable labelling efficiency highlights that these acoustogenic ligands are not currently optimal for *in vivo* delivery, and perhaps ICV injection is also not the best distribution method for systemic brain labelling.

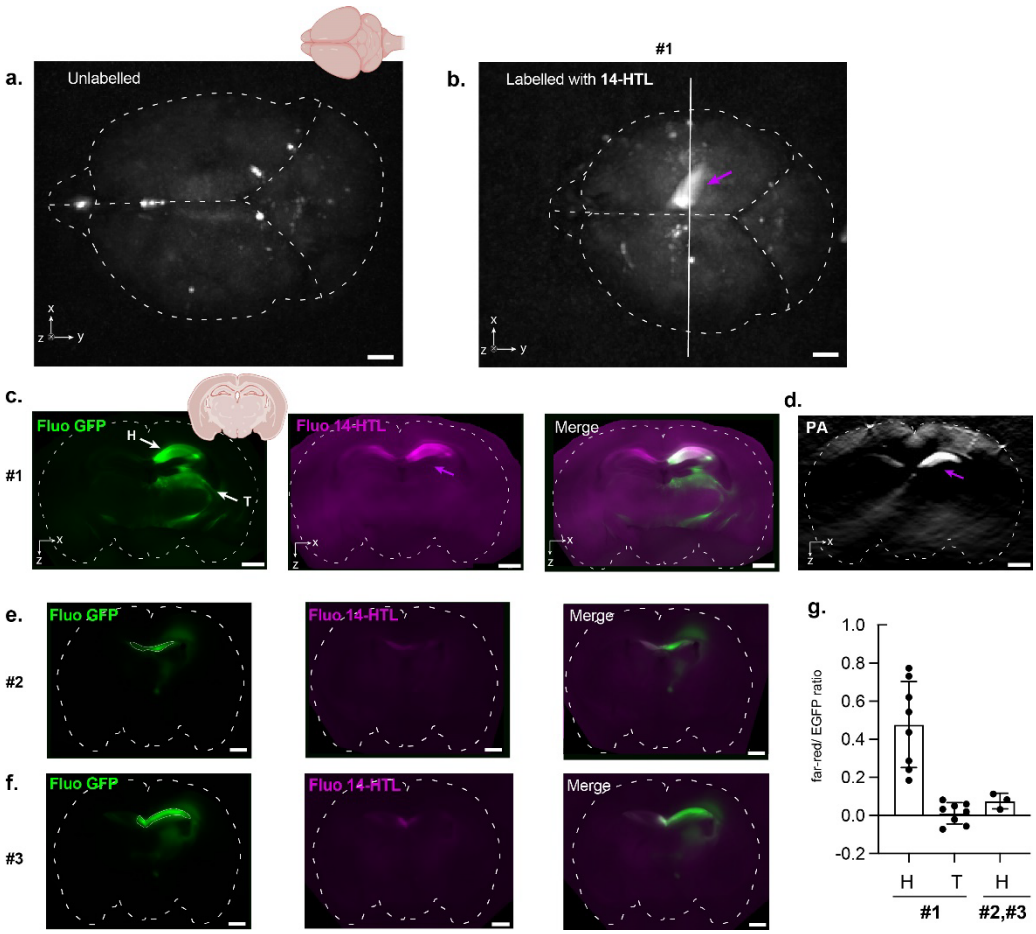


Figure 65: *In vivo* labelling of a mouse brain with 14-HTL via intracerebroventricular injection (ICV) and visualisation *ex vivo* with photoacoustic tomography. Photoacoustic tomography of a whole *ex vivo* mouse brain expressing HaloTag7-EGFP in neurons, labelled with **14-HTL** delivered by ICV *in vivo*, and fluorescence images of coronal slices. (**a**, **b**) PAT maximum intensity axial projection (λ_{PA} = 646 nm) of the entire mouse brain between 2.5 and 5.75 mm from the surface of the Fabry-Pérot interferometer to exclude strong endogenous signal from the olfactory bulbs: (**a**) control, unlabelled mouse and (**b**) mouse labelled with **14-HTL**; the magenta arrow indicates signal from **14**-HaloTag7, the white vertical line indicates the position of the coronal slice. (**c**) Widefield fluorescence images of a selected coronal slice, from top to bottom: EGFP channel with white arrows indicating the hippocampus (H) and thalamus (T) regions; far-red channel; and the merge of the two channels; images are from one of three mice injected with **14-HTL**. (**d**) Corresponding coronal slice from the PAT volume. (**e-f**) Widefield fluorescence images of coronal slices of mice brains showing partial labelling with **14-HTL**, from top to bottom: EGFP channel; far-red channel; and the merge of the two channels. (**g**) Far-red/EGFP fluorescence ratio for hippocampus (H) and thalamus (T) regions of mouse #1 (**c**) vs the hippocampus (H) of mice #2 and #3 (**e, f**).

IV.7 Conclusion and future perspectives

In this final results chapter, slowly moving up the biological complexity scale towards living mice, I have demonstrated the potential of our chemigenetic photoacoustic probes and calcium sensors for imaging of brain tissues. Under scattering conditions within tissue-mimicking phantoms, a representative acoustogenic ligand, **14-HTL**, bound to HaloTag-based protein scaffolds, generates strong PA signal enabling detection at depths surpassing those of the mouse brain, and retains calcium sensitivity with the HaloCaMP proteins. The concentration of probe needed for detection is also in the low micromolar range, which is achievable with AAV-mediated delivery of protein-based calcium sensors and so requires only optimisation of the dye delivery process for efficient labelling. Our series of acoustogenic ligands are cell-permeable and label cytosolically expressed HaloTag protein efficiently in mammalian cells. They can also permeate and label neurons within mouse brain slices, and produce sufficient PA signal for detection in these slices. Specifically labelled hippocampal regions can additionally be visualised within a whole mouse brain *ex vivo*. However, the *in vivo* labelling efficiency was variable, likely due to a combination of limited dye bioavailability and possible inhomogeneities in distribution of the cerebrospinal fluid from the lateral ventricles where the dye was delivered. Whilst these labels and sensors show great potential for *in vivo* application, further improvements are necessary before robust *in vivo* labelling can be achieved.

The inconsistent *in vivo* labelling results unfortunately confirm the predictions of limited bioavailability from the artificial membrane assay (see section III.5). The initial labelling efficiency comparison between IV and ICV delivery methods was not ideal because first we unwittingly selected one of our least bioavailable dyes, **15-HTL**, which was unlikely to be able to cross the BBB. Moreover, at this stage we only had fluorescence microscopy to evaluate labelling, which was not conclusive due to the sparse labelling and low brightness of the dye. Given the similar performance of **14-HTL** compared to **JF585-HTL** in the PAMPA-BBB assay, this dye might be capable of at least some neuronal labelling after IV delivery, which would have been more informative for our selection of dye delivery method. However, it is worth noting that in general the variability of dye labelling in such *in vivo* experiments is generally high,[208] and there are few examples of delivery of rhodamine derivatives and/or HaloTag ligands by stereotaxic brain injections.[251, 252] Furthermore, there is limited knowledge on dye distribution by this method, as the mechanisms of solute transport in the brain are not fully understood, but involve a combination of random, passive diffusion and convective flow via the recently characterised glymphatic system.[253, 254] These factors could contribute to an additional level of variability between individual mice.

Less invasive IV methods for dye delivery are preferable to minimise perturbation and so synthesis of modified acoustogenic ligands for improved bioavailability and BBB permeability, such as the structures proposed in Chapter III (Figure 55), should be explored to improve this key parameter. These new derivatives must be first evaluated for their absorption and photoacoustic properties *in vitro,* in spectroscopic and phantom measurements, plus assessment of their cell permeability and predicted bioavailability in the artificial membrane assay. With suitable acoustogenic candidates available, we could then evaluate their *in vivo* labelling efficiency via IV delivery to HaloTag-EGFP expressing mice (Figure 66a). Perfusion with an orthogonal fluorescent dye after 24 h would saturate the HaloTag protein, allowing quantification of labelling, as in cell culture (Figure 60).[208] PAT imaging of these slices would also be important to validate the best performing dye candidates.

Our immediate next steps, however, are to establish functional photoacoustic imaging with our current acoustogenic calcium sensors in acute brain slices. Using analogous AAV1s containing the two HaloCaMP constructs, the calcium scaffold can be expressed in neurons before confirming labelling *ex vivo* with our acoustogenic ligands via fluorescence microscopy and PAT, as for the HaloTag anchor (Figure 62). Given the current detection limit for the PAT set-up and the lower PA signal of dye-HaloCaMP conjugates compared to HaloTag, **14-HTL** should be used for labelling. Then neuronal function will be temporarily maintained using an artificial cerebrospinal fluid (ACSF)-supplied perfusion system and neurons excited using electrical or chemical stimulation during PAT imaging. This would demonstrate the first example of monitoring calcium dynamics in neurons with chemigenetic photoacoustic calcium sensors. Following this, we could attempt to

image a stimulated mouse brain *ex vivo,* taking crucial steps on the path towards brain-wide *in vivo* neuroimaging.

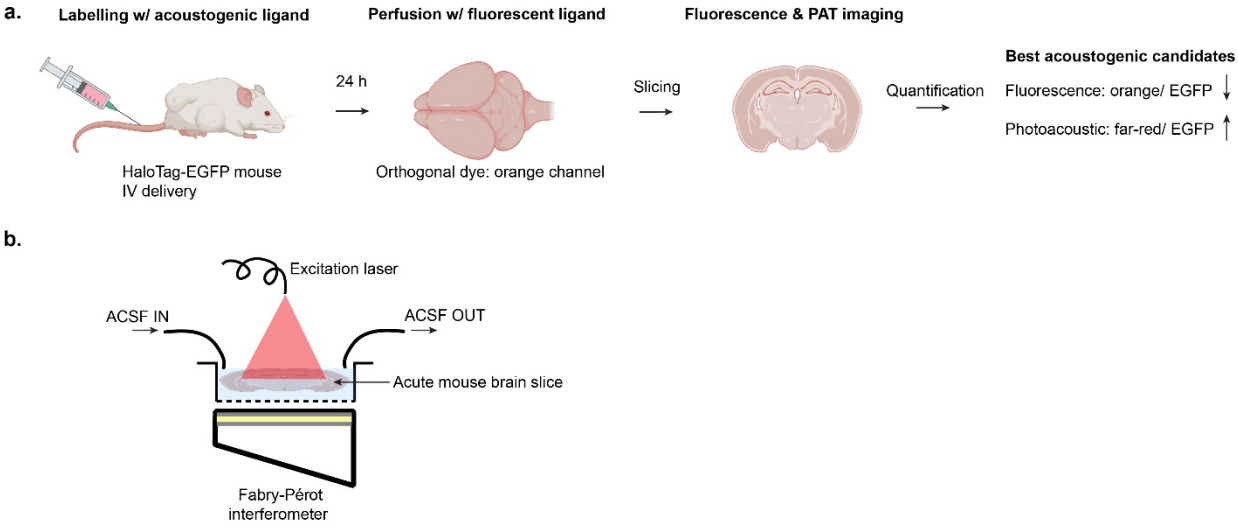


Figure 66: Future experiments with our chemigenetic probes and calcium sensors. (**a**) Schematic for assessment of *in vivo* labelling efficiency of acoustogenic dye ligands. (**b**) Set-up for PAT imaging of acute mouse brain slices.

Chapter V: General discussion

Obtaining a deep understanding of brain-wide neural activity requires a variety of highly optimised molecular imaging tools. The GCaMP series of GECIs is the perfect example of one of these tools, the result of over 20 years of protein engineering to yield a diverse selection of finely tuned sensors for different applications in neurons. However, sensor development has focused primarily on fluorescence microscopy which, despite all the technological advancements, is fundamentally depth limited by the scattering of light. Therefore, to fully understand the integrated function of the entirety of the brain our attention must turn to other imaging modalities. Photoacoustic imaging, which uses the unique combination of light and sound, can provide functional imaging at the mesoscale with high spatiotemporal resolution, for capturing widespread dynamics between neural circuits and different brain regions. In this work, we have developed a novel series of photoacoustic probes and calcium sensors, designed specifically for this modality, in order to resolve the deficiency in existing reporters. Indeed, our calcium indicators outperform currently available calcium indicators in crucial parameters such as dynamic range and photostability *in vitro*, and can be used to specifically label brain tissues, generating strong photoacoustic signal.

To develop these probes, we leveraged the chemigenetic approach for sensors, combining the advantages of tuneable, far-red synthetic chromophores with the labelling specificity of the self-labelling protein, HaloTag. The development pipeline involved first the design, synthesis and evaluation of acoustogenic rhodamine-based ligands which could be combined with HaloTag and the existing calcium-sensitive HaloCaMP scaffolds. These probes were then imaged in tissue-mimicking phantoms, to demonstrate their calcium sensitivity and strong PA signal in scattering conditions, and subsequently in brain tissue, showcasing their labelling specificity. However, the limited bioavailability revealed by *in vivo* delivery of these acoustogenic ligands must be improved for future applications in living mice.

In general, the major limitation for chemigenetic probe development is indeed the delivery of the dye-ligand to the specifically expressed protein target, evidenced by the few examples of *in vivo* demonstration of these tools.[157, 159] To achieve our goal of whole-brain neuroimaging, our acoustogenic ligands must be reliably delivered *in vivo* and achieve systemic brain labelling. Now we have demonstrated the proof-of-principle of developing chemigenetic PA probes, we should adapt the approach for the design of future PA sensors by prioritising bioavailability, and not select ligands primarily on sensitivity. Chemigenetic sensors for *in vivo* applications using other modalities should also follow this approach, exemplified by the development of WHaloCaMP.[159] At an early stage in screening acoustogenic dye candidates, the most bioavailable ligands should be identified: first by using *in vitro* screening methods like the artificial membrane assay or cell culture techniques, and then *in vivo* using the aforementioned pulse-chase labelling experiments for example.[208] Subsequently, the protein scaffold should be engineered in conjunction with the selected probe to enhance the sensing mechanism. For our PA calcium sensors, the most chromogenic of the bioavailable ligands would be selected for tailored engineering of the HaloCaMP scaffold to improve the dynamic range via modulation of the specific dye-protein interactions.

Evaluation of the bioavailability of more ligands would help to understand the necessary chemical properties for bioavailable chromogenic ligands and guide future design. The work by the Lavis group on rhodamine derivatives seemed to suggest that modestly chromogenic dyes were the most cell and tissue permeable, due to a greater tendency to adopt the closed, lipophilic form to cross phospholipid membranes.[166] By contrast, the poor bioavailability of **JF635** indicated a potential trade-off in return for high chromogenicity. However, the position of equilibrium is clearly not the only factor, demonstrated by the differing $K_{L\text{-}Z}$ of the most bioavailable JF ligands, and even some very chromogenic dyes can also cross the BBB.[208, 255] There are likely functional groups, potentially OH in the case of the Malachite Green lactone acoustogenic ligands, which when present can improve this key property. These groups need to be identified for other dye scaffolds which could be combined with HaloCaMP as we diversify our acoustogenic ligands. Therefore, synthetic modifications and *in vivo* screening are essential to facilitate the next generation of chemigenetic sensors for application in living animals.

Regarding these alternative scaffolds, the dye flexibility of the chemigenetic system enables diversification to other families, which should also be explored for PA probe development. Cyanines are of particular interest because of their high extinction coefficients, up to ~200000 $M^{-1}{\cdot}cm^{-1}$, approximately two-fold greater than our current acoustogenic dyes, and provide a further red-shift (λ_{max} > 700 nm) away from the competing haemoglobin absorption. Moreover, the lower toxicity of these compounds is highlighted by the FDA- and EMA-approved **ICG** and participation of **IR800CW** in clinical studies. Recently, rhodamine-like open-closed equilibriums have been established in cyanines through intramolecular cyclisation with carboxylic acid or amide groups resulting in chromogenicity with either HaloTag or SNAP-tag self-labelling proteins, with up to 24- and 11-fold absorbance turn-ons respectively.[171, 207] For PA application in our chemigenetic platform, quenched derivatives using both chromogenic strategies, would need to be designed, due to their high quantum yields (Φ = 0.1−0.5), which maintain high photostability (Figure 67). Quenching groups facilitating intramolecular charge transfer could be incorporated into the structure, such as NO_2 or tetrazine, via the indoleninium moieties.[256] Substituents on the heptamethine chain can also have a large effect on quantum yield, extinction coefficient and photostability.[257] Alternatively, the fluorescence of other cyanine scaffolds is lower, for example the bioavailable **ICG** (Φ = 0.027)[64], which could be modified to introduce the open-closed equilibrium, via either method, and to facilitate cell permeability, the sulfonic acid groups could be masked with esters labile to intracellular esterases in mammalian cells.[258]

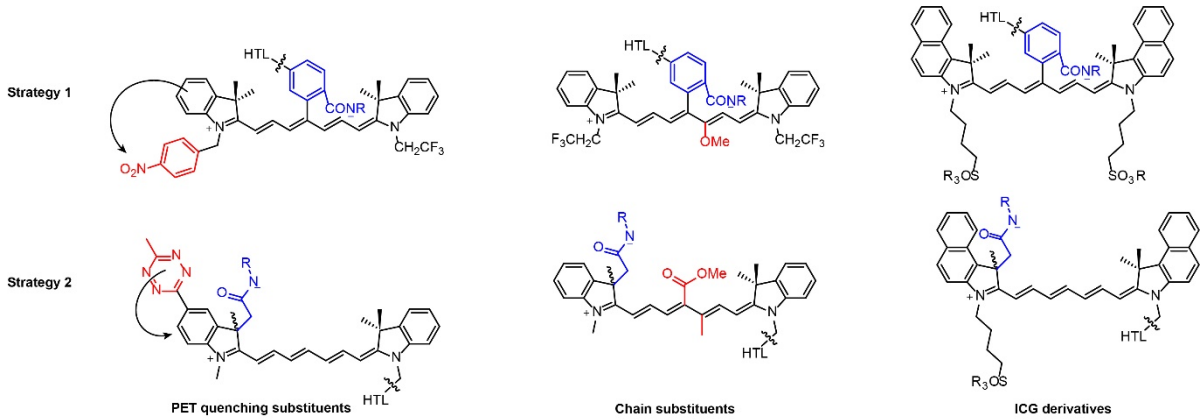


Figure 67: Engineering acoustogenic cyanines for PAI. Potential acoustogenic dye structures; moieties responsible for the open-closed equilibrium are highlighted in blue, from two recently published methods;[171, 207] and groups for quenching in red.

In addition, I believe we have laid the groundwork to accelerate future PA chemigenetic-based probe development. Nikita's multimodal spectrometer is ideal for characterisation of crucial photophysical properties of dye candidates in solution. We have adapted screening of bacterial colonies using PAM for our sensors, which will enable evaluation of protein mutants labelled with acoustogenic dyes. Alternatively, the well plate configuration of our photoacoustic spectroscopy set-up, with automation, could provide a complementary method for screening of bacterial lysates. Consequently, once a series of bioavailable acoustogenic ligands has been synthesised, we can improve their calcium-sensing properties with HaloCaMP, including dynamic range and calcium affinity, via protein engineering. Moreover, we are already establishing an acute brain slice platform for PAI with our acoustogenic calcium sensors, for the first photoacoustic recording of neuronal calcium fluctuations using chemigenetic tools. This would enable evaluation of the performance of new sensors when recording neuronal activity. We hope that this work is just the beginning of chemigenetic photoacoustic probe development, bringing this powerful modality to the forefront of a "loud" future for neuroimaging.

Chapter VI: Supplementary Information

VI.1 Supplementary Figures and Tables

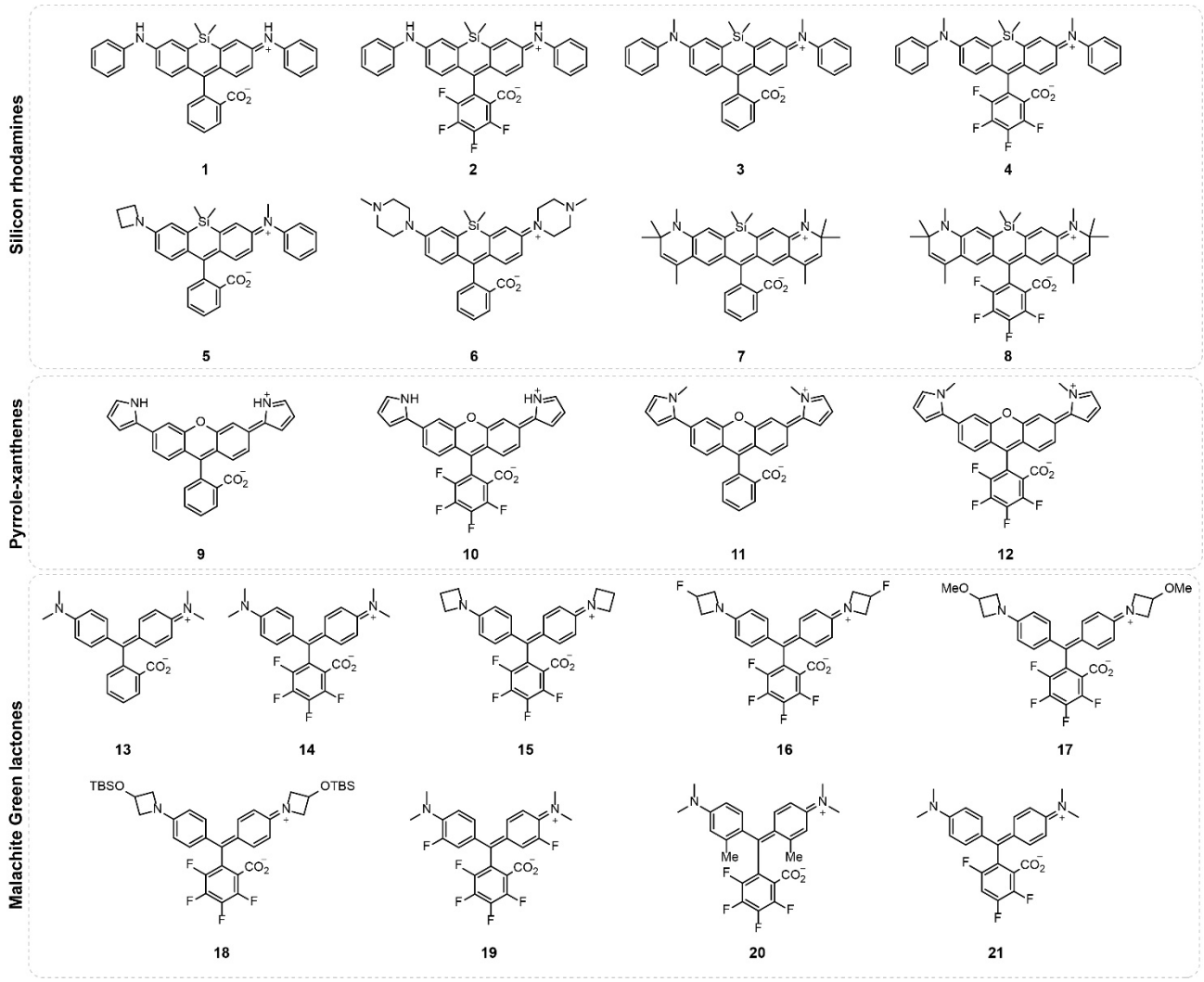


Figure S1: **Chemical structures of the free dyes synthesised and characterised.**

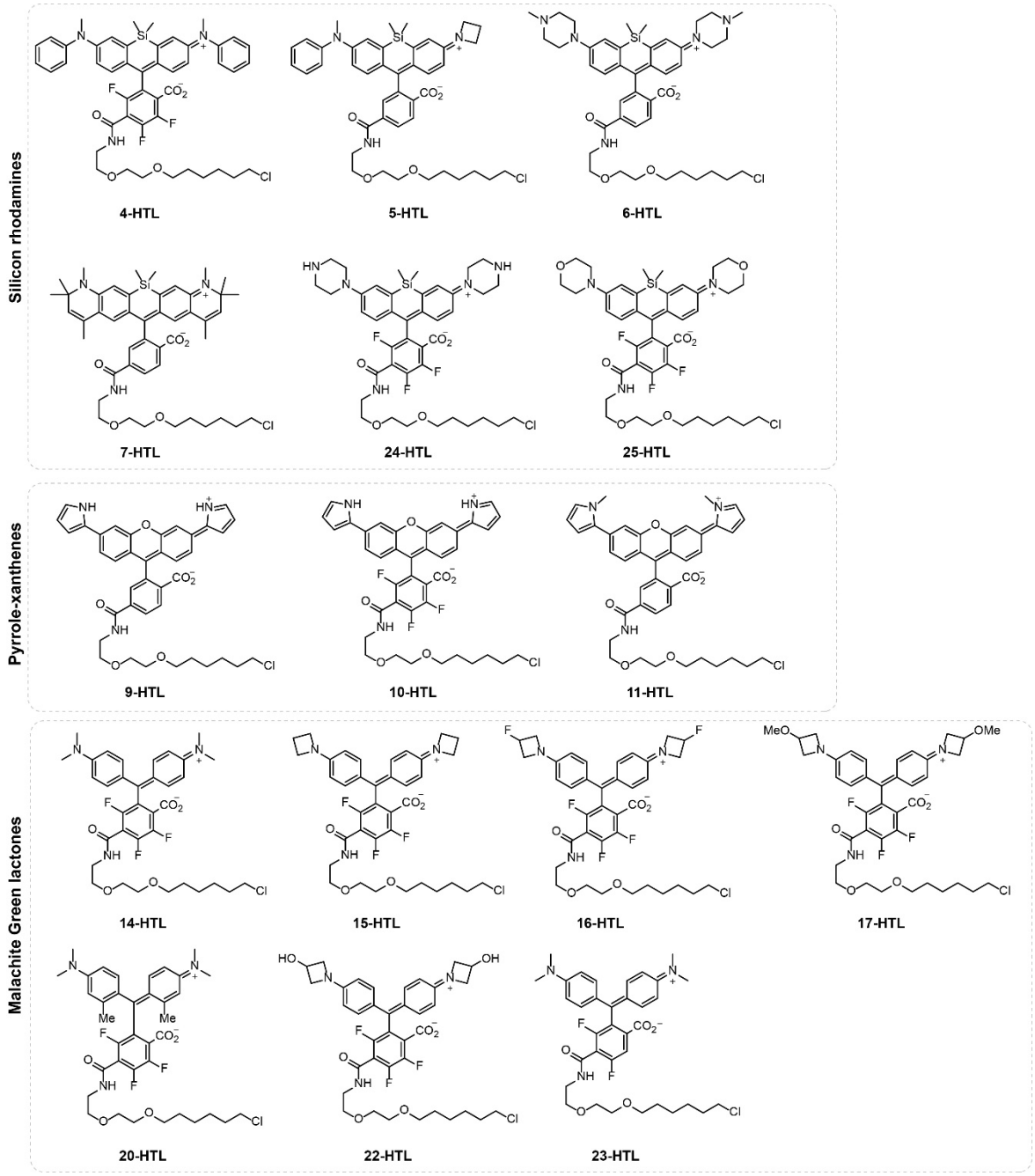


Figure S2: Chemical structures of the HaloTag ligands synthesised and characterised. In addition to the synthesised ligands I received an aliquot of **7-HTL**, **24-HTL** and **25-HTL** from the Lavis lab for testing.

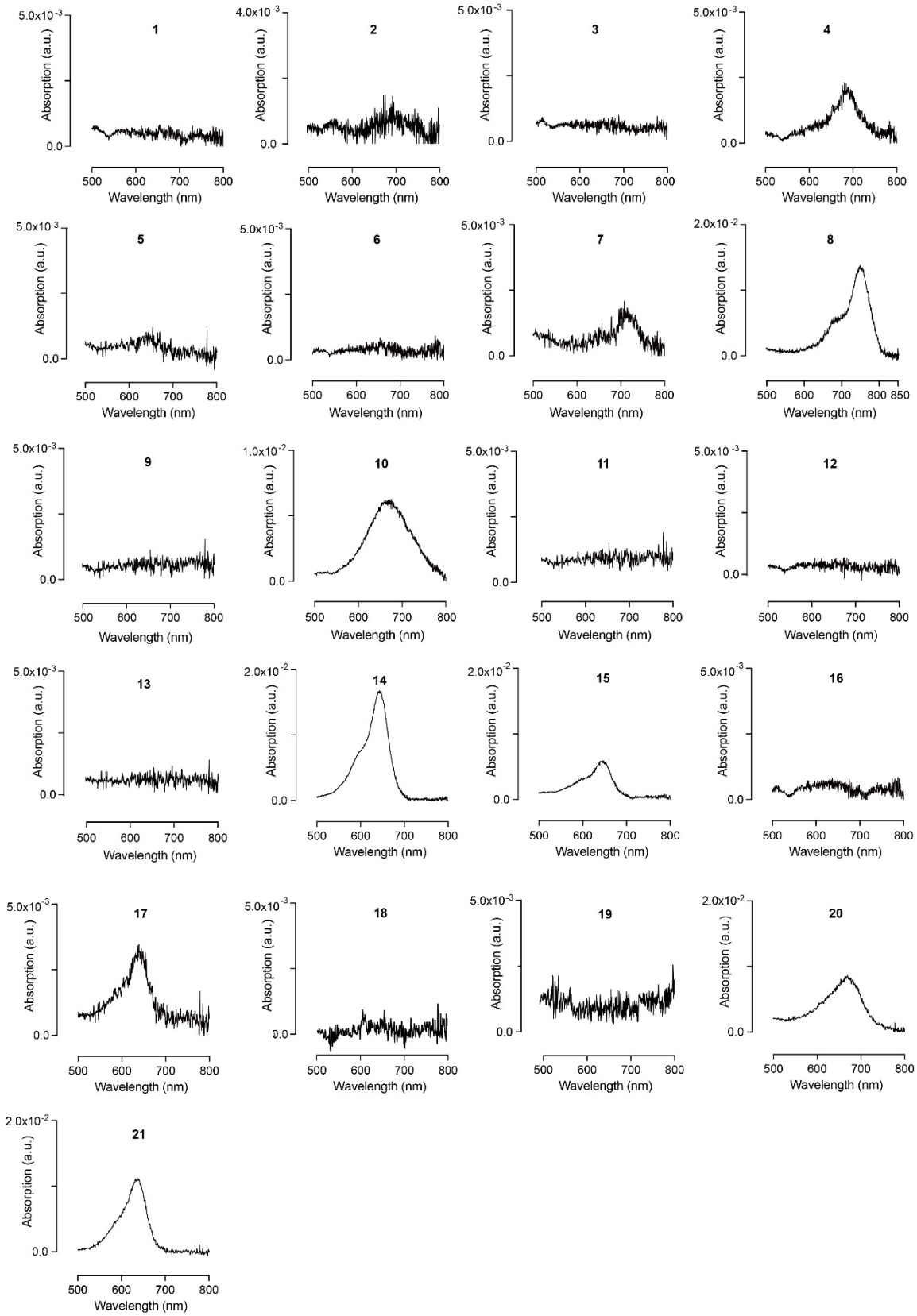


Figure S3: Absorption spectra of free dyes 1-21. All measurements were performed in 10 mM HEPES, pH 7.4 at 1.25 µM dye concentration.

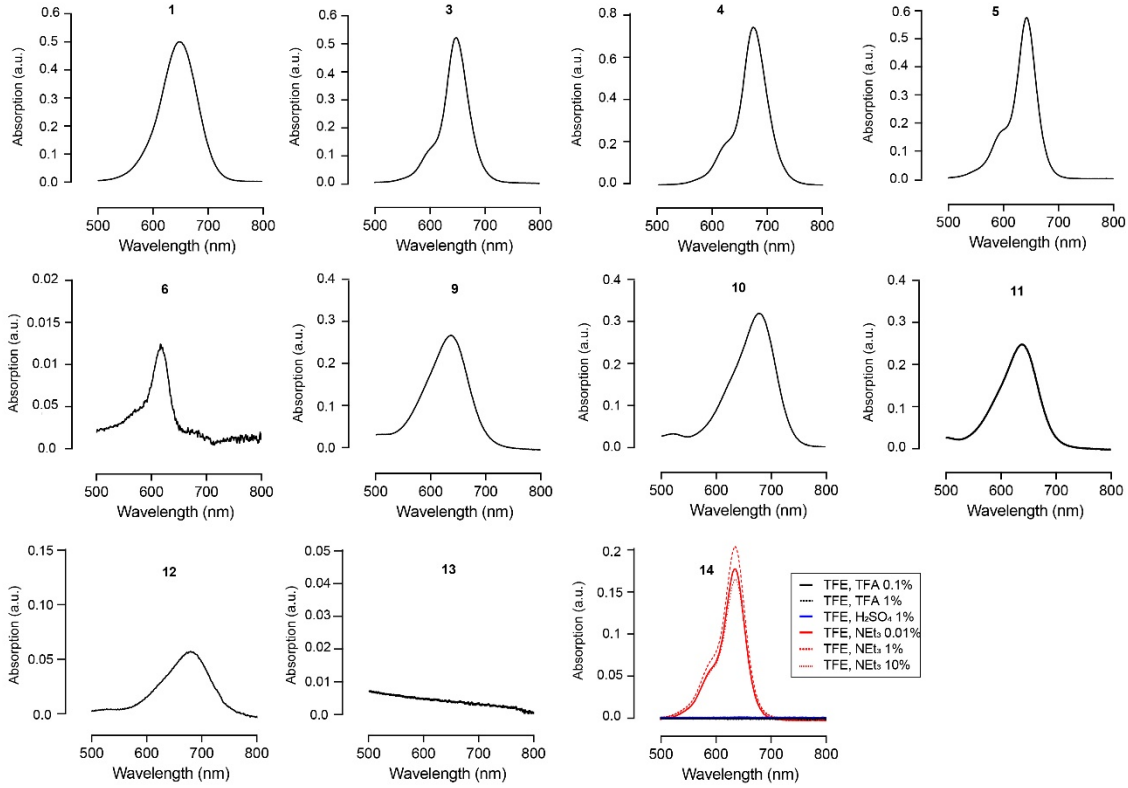


Figure S4: Absorption spectra of free dyes 1-14 in acidic conditions to elicit dye opening. All measurements were performed in 2,2,2-trifluoroethanol (TFE) containing 0.1% TFA at 5 µM dye concentration, except for **14** which was also measured in TFE with 1% TFA or 1% H_2SO_4, 0.01%, 1% or 10% NEt_3.

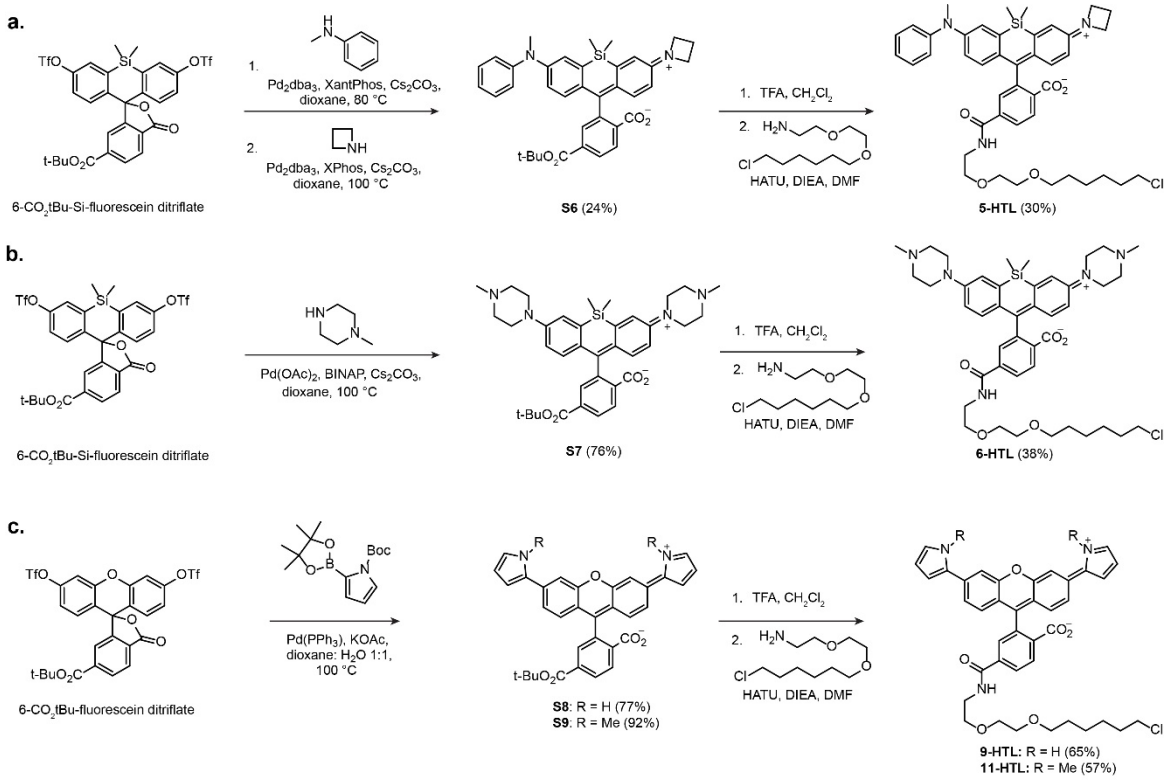


Figure S5: Synthesis of nonfluorinated HaloTag ligands 5-, 6-, 9- and 11-HTL. (**a, b**) Cross-coupling of (**a**) *N*-methylaniline and azetidine, (**b**) *N*-methylpiperazine with 6-CO_2*t-Bu*-Si-fluorescein ditriflate precursors, followed by acid deprotection and amide coupling to form **5-HTL** and **6-HTL** respectively. (**c**) Suzuki coupling of 6-CO_2*t-Bu*-fluorescein ditriflate precursor with a pyrrole boronic ester followed by acid deprotection and amide coupling to form **9-HTL** or **11-HTL**.

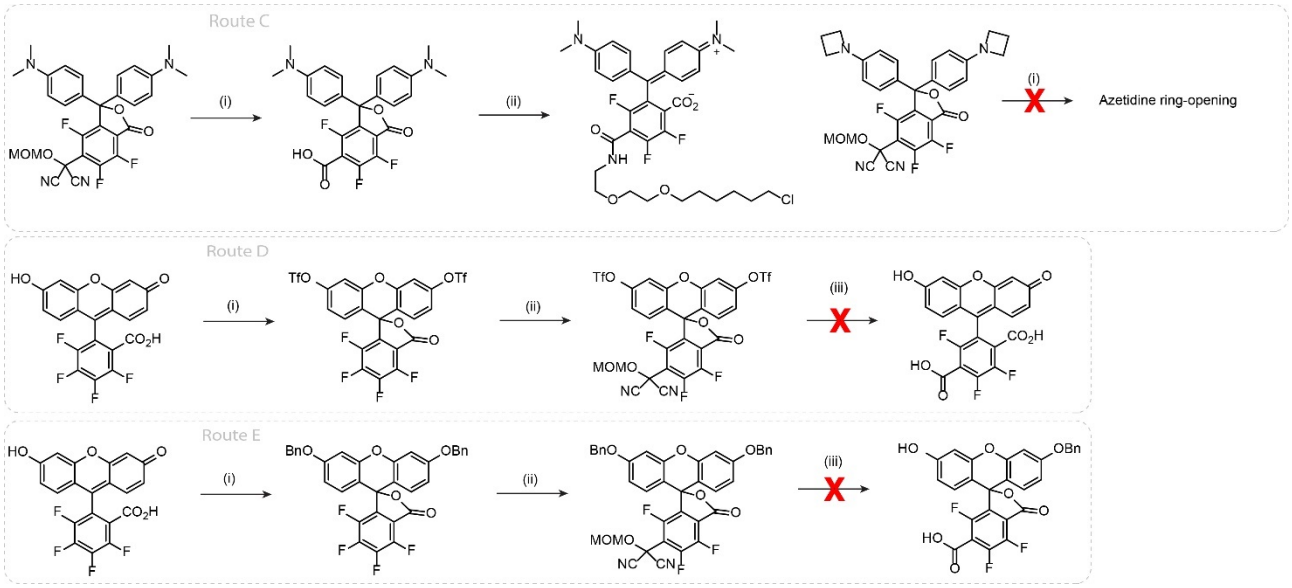


Figure S6: Other synthetic routes attempted in the preparation of HaloTag ligands from tetrafluorinated rhodamine derivatives. Route C: alternative acid deprotection of MOM-MAC Malachite Green lactone intermediates; (i)[180] 1M H_2SO_4, THF, 60°C; (ii)[259] HTL, EDC.HCl, HOBt, DIEA, CH_2Cl_2. This was successful for the synthesis of **14-HTL**, however it was not suitable for all compounds: the acidic conditions induced ring-opening of the azetidines,[260, 261] whilst the pyrrole-xanthene scaffold decomposed. **Route D:** To synthesise the pyrrole-xanthene HTLs I tried to perform the MAC chemistry before incorporation of the auxochromes, however triflate hydrolysis resulted in low yielding reactions; (i)[167] Tf_2O, pyridine, CH_2Cl_2; (ii)[166] MOM-MAC, DIEA, DMF; (iii)[180] 1M H_2SO_4, THF, 60°C. **Route E**: protection of fluorescein OH groups with orthogonal benzyl ether groups, was a better option, but the harsh acidic conditions for the MOM-MAC deprotection resulted in partial deprotection of the benzyl ethers, making this already long synthetic route even lower yielding; (i)[262] benzyl bromide, K_2CO_3, acetone; (ii)[166] MOM-MAC, DIEA, DMF; (iii)[180] 1M H_2SO_4, THF, 60°C.

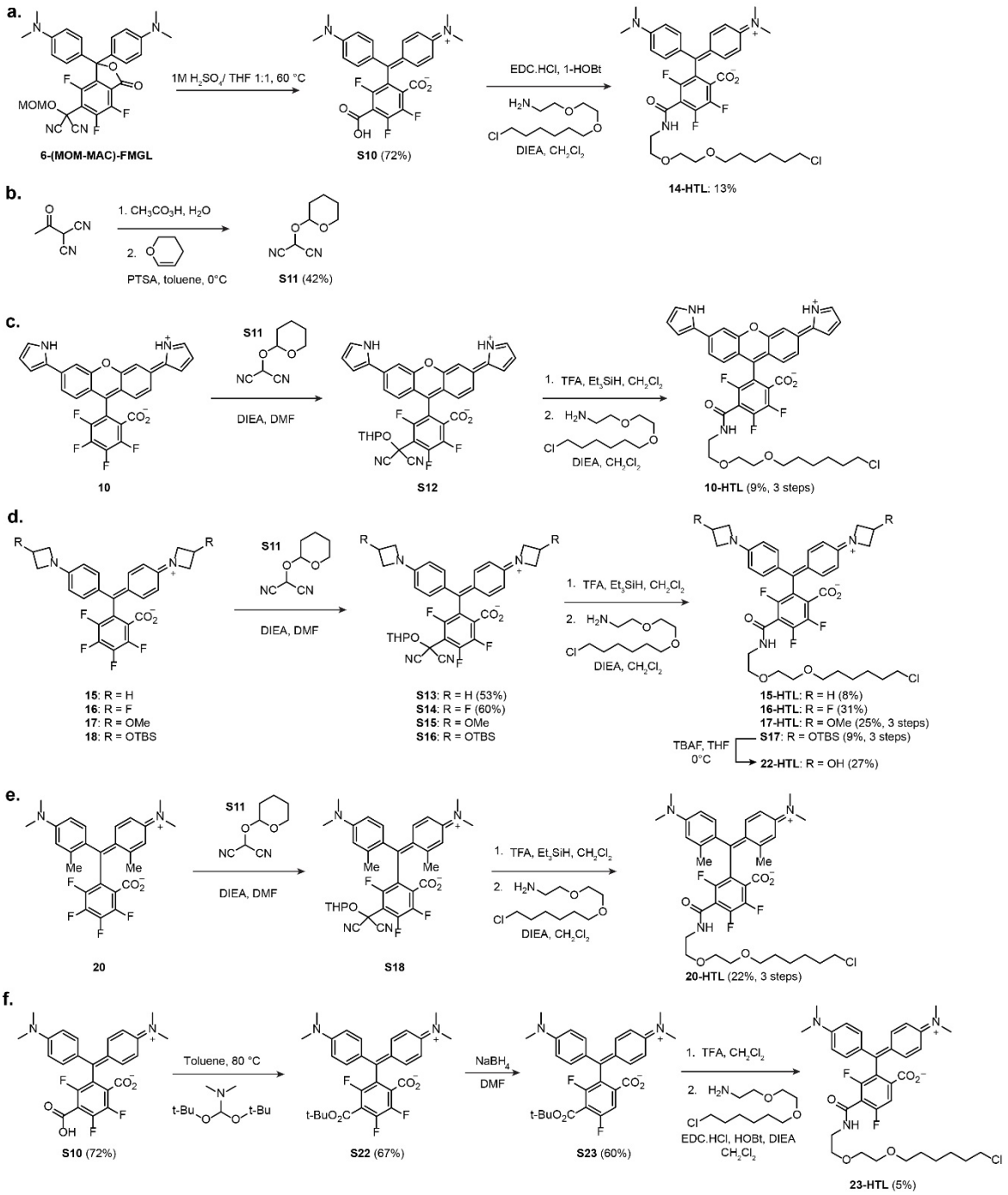


Figure S7: Synthesis of fluorinated HaloTag ligands 10-, **15-**, **16-**, **17-**, **20-**, **22-**, and **23-HTL**. (**a**) Synthesis of **14-HTL** from the MOM-MAC intermediate via deprotection in H_2SO_4 and subsequent amide coupling. (**b**) Synthesis of THP-MAC from malononitrile. (**c-e**) Regioselective substitution of THP-MAC in the 6-position followed by acid deprotection and amide coupling with HaloTag ligand. (**c**) Synthesis of **10-HTL** using the alternative MAC route. (**d**) Synthesis of azetidine-substituted Malachite Green lactone HaloTag ligands **15-, 16-, 17-** and **22-HTL** using the alternative MAC route. (**e**) Synthesis of **20-HTL** using the alternative MAC route. (**f**) Alternative synthesis of **23-HTL** via the MOM-MAC intermediate, which is hydrolysed to the carboxylic acid and esterified with *N,N*-dimethylformamide dimethyl acetal to form the *t*-Bu ester; regioselective defluorination with $NaBH_4$, and subsequent deprotection and amide coupling forms the difluorinated HTL product.

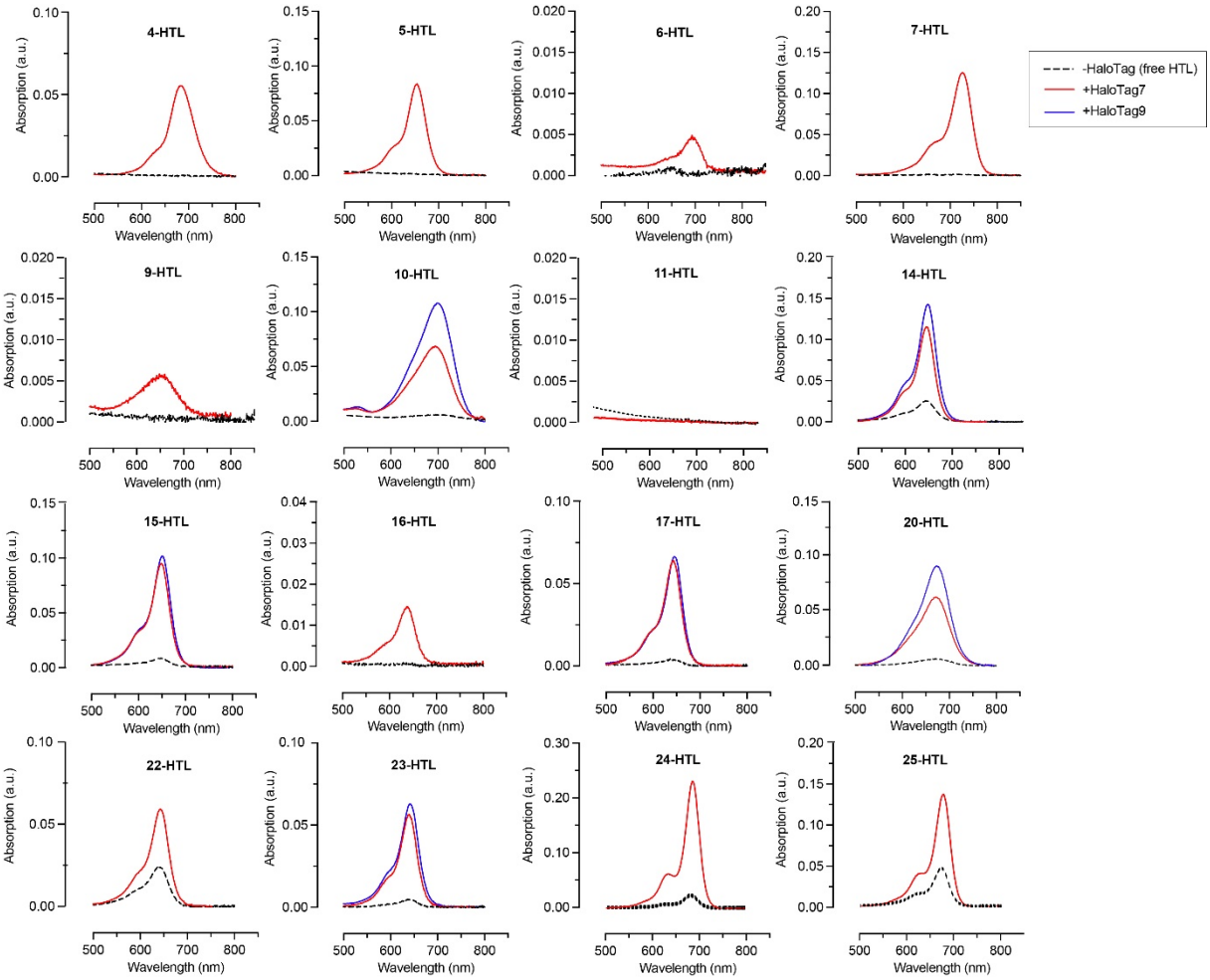


Figure S8: Absorption spectra of all HaloTag ligands in the absence of protein (dashed black lines), bound to HaloTag7 (red lines) or bound to HaloTag9 (blue lines). All measurements were performed in 10 mM HEPES, pH 7.4, containing 0.1 mg·mL^{-1} of CHAPS at a dye ligand concentration of 1.25 µM and protein concentration of 1.9 µM, in triplicate.

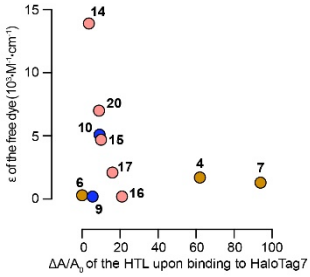


Figure S9: **Using free dye absorption as a predictor of chromogenicity.** Extinction coefficients of the free dyes (see Figures 23, 25, 30) vs $\Delta A/A_0$ for the corresponding HaloTag ligands upon binding to HaloTag protein (see Figure 34).

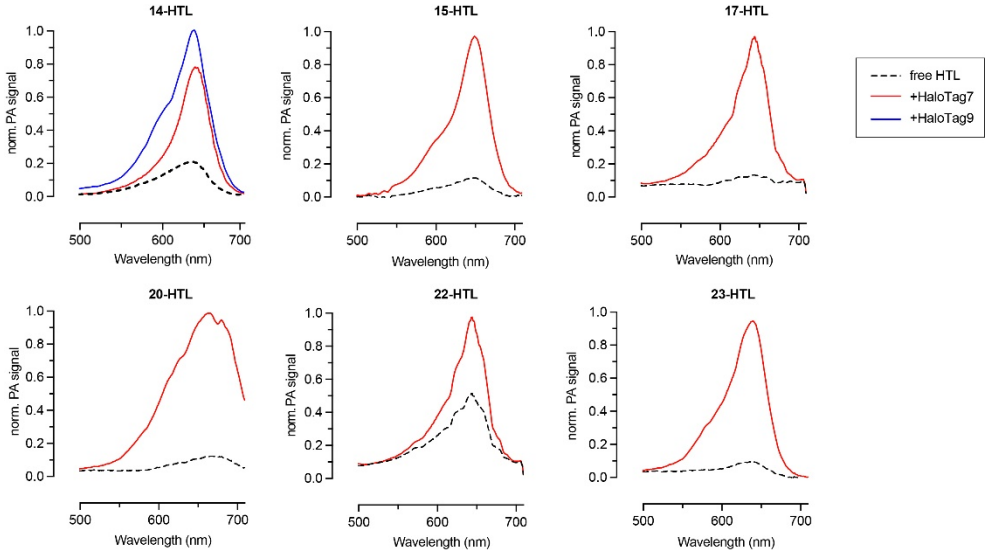


Figure S10: Normalised photoacoustic spectra of selected HaloTag ligands in the absence of protein (dashed black lines) or bound to HaloTag7 (red lines) or HaloTag9 (blue lines). All measurements were performed in 10 mM HEPES, pH 7.4, containing 0.1 mg·mL^{-1} of CHAPS at a dye ligand concentration of 1.25 µM and protein concentration of 1.9 µM, in triplicate.

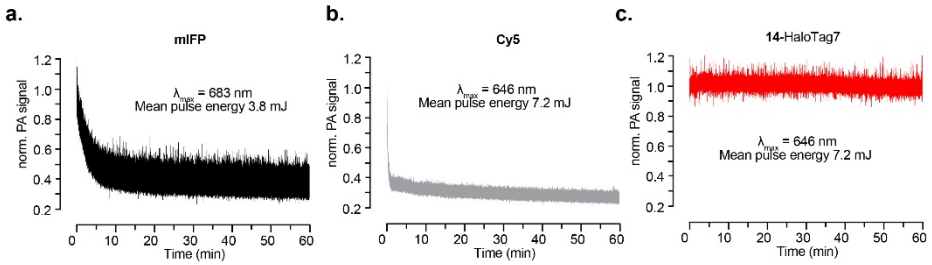
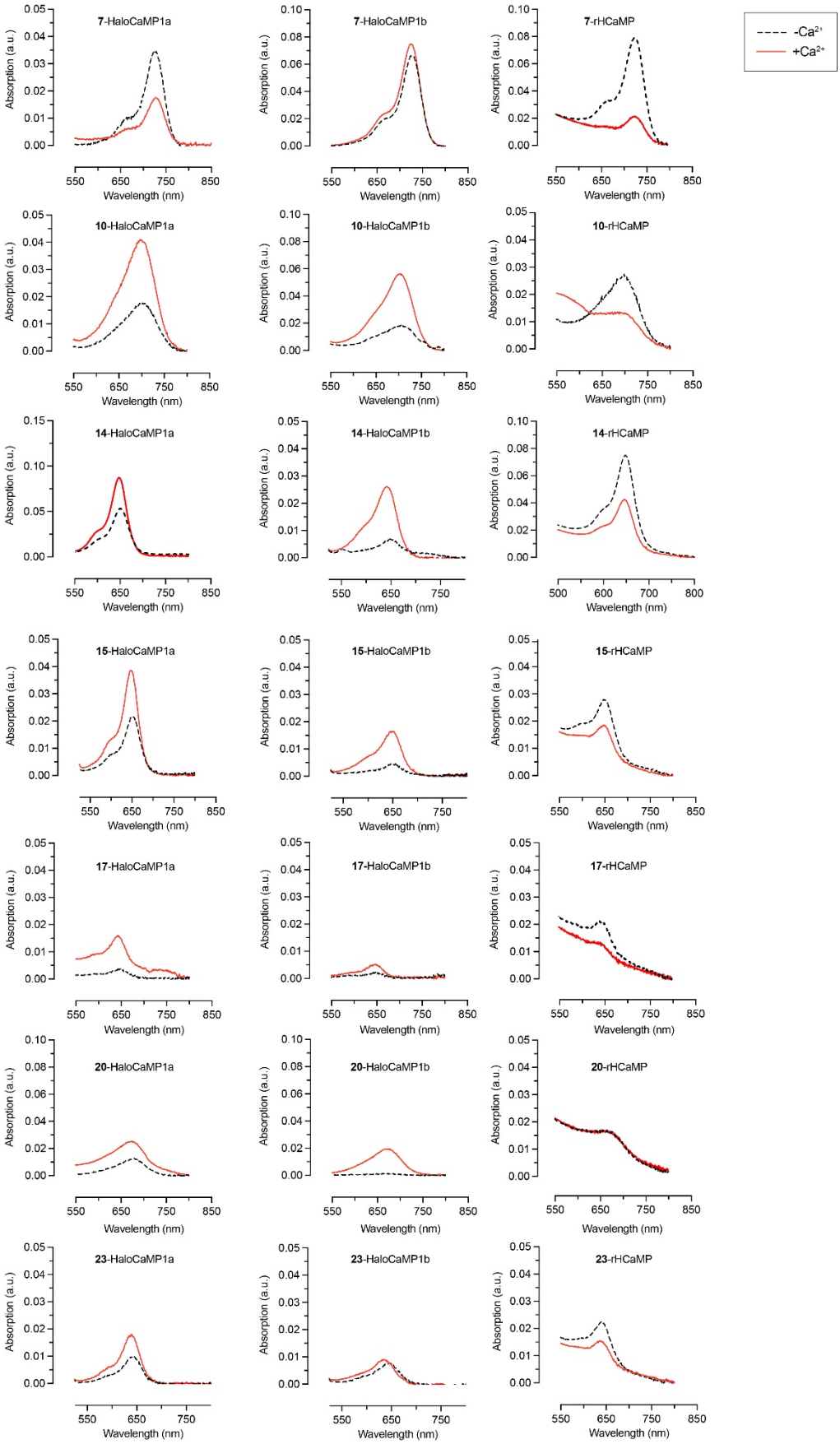


Figure S11: Photostability of photoacoustic labels. Individual photobleaching curves from Figure 38d at different λ_{PA}: (**a**) mIFP (683 nm), (**b**) Cy5 (646 nm), (**c**) **14**-HaloTag7 (646 nm). Normalised photoacoustic signal plotted against time.


7-HaloCaMP1a
7-HaloCaMP1b
7-rHCaMP
10-HaloCaMP1a
10-HaloCaMP1b
10-rHCaMP
14-HaloCaMP1a
14-HaloCaMP1b
14-rHCaMP
15-HaloCaMP1a
15-HaloCaMP1b
15-rHCaMP
17-HaloCaMP1a
17-HaloCaMP1b
17-rHCaMP
20-HaloCaMP1a
20-HaloCaMP1b
20-rHCaMP
23-HaloCaMP1a
23-HaloCaMP1b
23-rHCaMP
Absorption (a.u.)
Wavelength (nm)
-Ca2+
+Ca2+

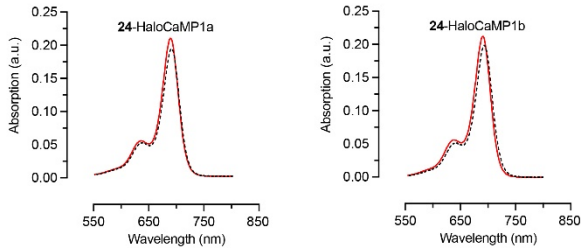


Figure S12: Absorption spectra of the chemigenetic calcium sensors in the calcium-free (dashed black lines) and calcium-bound states (solid red lines). All measurements were performed in 30 mM MOPS, 100 mM KCl, pH 7.2, containing 0.1 mg·mL^{-1} CHAPS to which either 10 mM EGTA (for the calcium-free state) or excess $CaCl_2$ (for the calcium-saturated state, i.e. 500 mM $CaCl_2$ for HaloCaMP1a or 10 mM $CaCl_2$ for HaloCaMP1b and rHCaMP) was added. For HaloCaMP1a and HaloCaMP1b, solutions were prepared with 1.25 µM HaloTag ligand and 1.9 µM protein. For rHCaMP, solutions were prepared with 2.5 µM HaloTag ligand and 3.8 µM protein**.** Measurements were performed in triplicate.

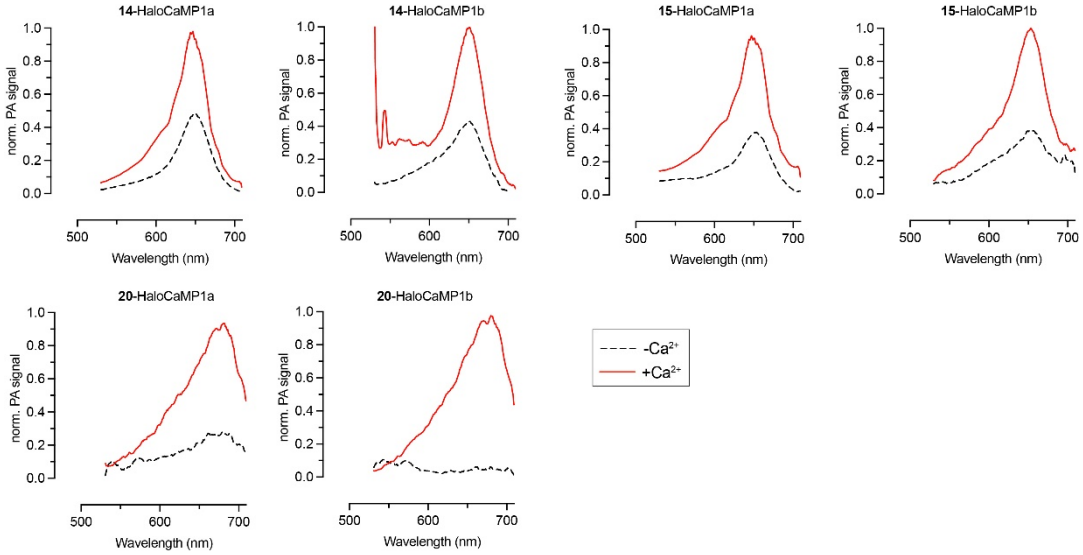


Figure S13: Photoacoustic spectra of selected chemigenetic calcium sensors in the calcium-free (dashed black lines) and calcium-saturated states (solid red lines). All measurements were performed in 30 mM MOPS, 100 mM KCl, pH 7.2, containing 0.1 mg·mL^{-1} of CHAPS to which 10 mM EGTA was added (for the calcium-free state), or excess $CaCl_2$ was added (for the calcium-saturated state, 500 mM $CaCl_2$ for HaloCaMP1a or 10 mM $CaCl_2$ for HaloCaMP1b). Each spectrum was normalised to the maximum PA signal at λ_{PA} for the Ca^{2+}-bound state. Solutions were prepared with 1.25 µM HaloTag ligand and 1.9 µM protein. Measurements were performed in triplicate.

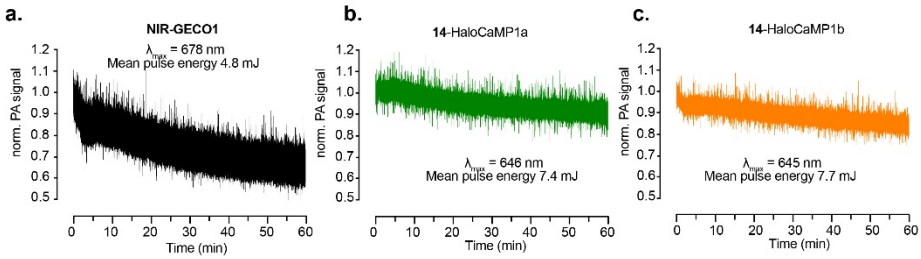


Figure S14: Photostability of photoacoustic calcium sensors. Individual photobleaching curves from Figure 39f at different λ_{PA}: (**a**) NIR-GECO1 (λ_{PA} = 678 nm), (**b**) **14**-HaloCaMP1a (λ_{PA} = 646 nm), (**c**) **14**-HaloCaMP1b (λ_{PA} = 645 nm). Normalised photoacoustic signal plotted against time.

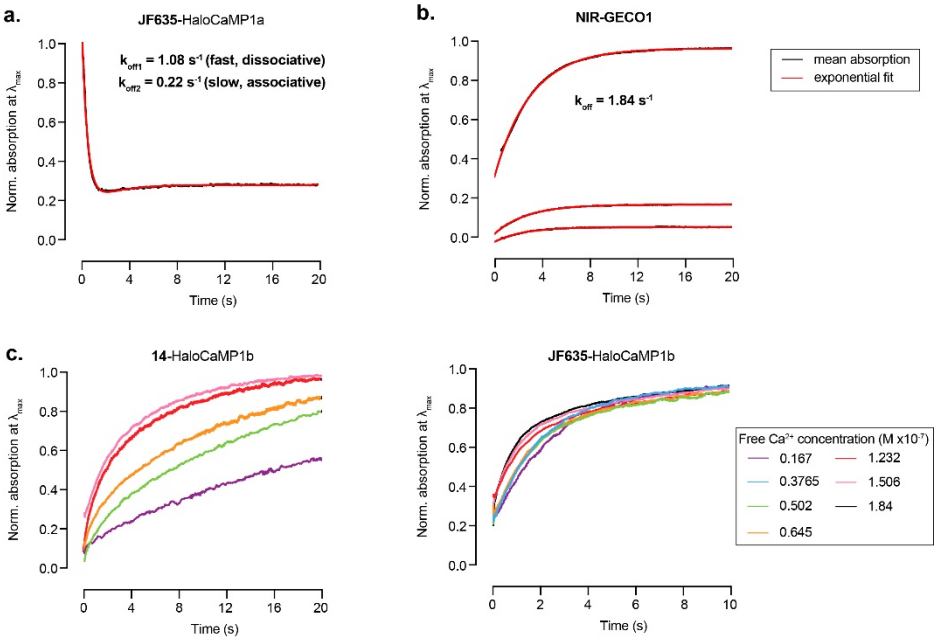


Figure S15: Ca^{2+} binding and unbinding kinetics of acoustogenic calcium sensors. (**a**) Calcium dissociation kinetics of **JF635**-HaloCaMP1a, (**b**) NIR-GECO1. Sensors in 30 mM MOPS, 100 mM KCl, 100 µM $CaCl_2$ or 10 µM for NIR-GECO1, pH 7.2 were mixed with high EGTA-containing buffer (30 mM MOPS, 100 mM KCl, 10 mM EGTA, pH 7.2), and k_{off} determined by fitting of average absorption at λ_{max} (black curves) with exponential functions (red curves); 1 concentration for **JF635**-HaloCaMP1a, 3 concentrations for NIR-GECO1. (**c**) Calcium binding kinetics of **14**-HaloCaMP1b and **JF635**-HaloCaMP1b; sensors in 30 mM MOPS, 100 mM KCl, 10 mM K_2EGTA, pH 7.2 were mixed with CaEGTA-containing buffer (30 mM MOPS, 100 mM KCl, 10 mM CaEGTA, pH 7.2). Plots shown are the mean of at least 4 runs.

(a)

MHHHHHHHHHHGSRKKTFKEVATAVKIIAMLMGPEFAAETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPL
TEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARL
AKSLPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISGGGTGGSGGTGGSGGTGGSMAEIGTGFPFDPHYV
EVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFM
DAFIEALGLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWLEFARDQLTEEQIAEFKEAFSL
FDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVF
DKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAKLEISGGGSGMVSKGEE
LFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMA
DKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAA
GITLGMDELYK

(b)

MHHHHHHHHHHGSRKKWKQSVRLISLCQRLSRPEFAAETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTE
VEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAK
SLPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISGGGTGGSGGTGGSGGTGGSMAEIGTGFPFDPHYVEV
LGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDA
FIEALGLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWLEFARDQLTEEQIAEFKEAFSLFD
KDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVFD
KDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAKLEISGGGSGMVSKGEEL
FTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMA
DKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAA
GITLGMDELYK

(c)

MHHHHHHHHHHGSRKKTFKEVATAVKIIAMLMGPKETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEV
EMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKS
LPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISGGGTGGSGGTGGSGGTGGSMAEIGTGFPFDPHYVEVL
GERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAPKHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFI
EALGLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPFARDQLTEEQIAEFKEAFSLFDKD
GDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVFDKD
GNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAKISGGGSGMVSKGEELFTGV
VPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFK
SAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQK
NGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYK

(d)

MHHHHHHHHHHGSRKKWKQSVRLISLCQRLSRPKETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEVE
MDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSL
PNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISGGGTGGSGGTGGSGGTGGSMAEIGTGFPFDPHYVEVLG
ERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAPKHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIE
ALGLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPFARDQLTEEQIAEFKEAFSLFDKDG
DGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVFDKDG
NGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAKISGGGSGMVSKGEELFTGVVP
ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSA
MPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDE
LYK

Annotations: purple – start codon; red – His-tag; brown – linker; blue – peptide; grey – cpHaloTag; orange – calmodulin; green – EGFP.

Figure S16: Amino acid sequences of HaloCaMP affinity variants annotated with sequence features. (**a**) 1a_ENOSP, (**b**) 1a_DAPKP, (**c**) 1b_ENOSP, (**d**) 1b_DAPKP.

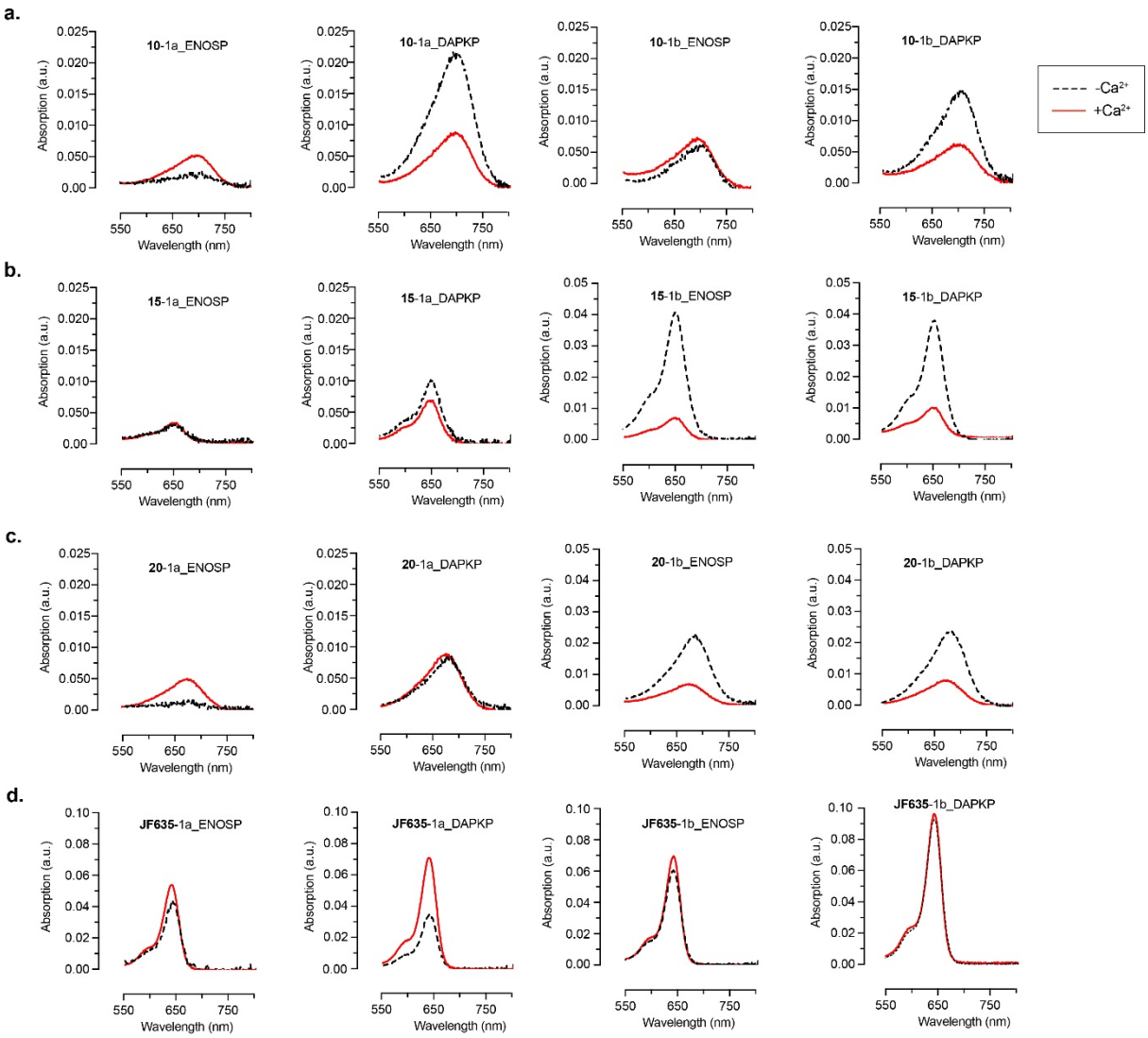


Figure S17: Absorption spectra of HaloCaMP variants labelled with 10-, 15-, 20- and JF635-HTL. (**a**) **10-HTL**, (**b**) **15-HTL**, (**c**) **20-HTL**, (**d**) **JF635-HTL**. All measurements were performed in 30 mM MOPS, 100 mM KCl, pH 7.2, containing 0.1 mg·mL^{-1} of CHAPS, to which either 10 mM EGTA (for the calcium-free state) or 10 mM $CaCl_2$ was added (for the calcium-saturated state); values are the mean of 3 replicates; all measurements were performed at 1.25 µM dye and 1.9 µM protein.

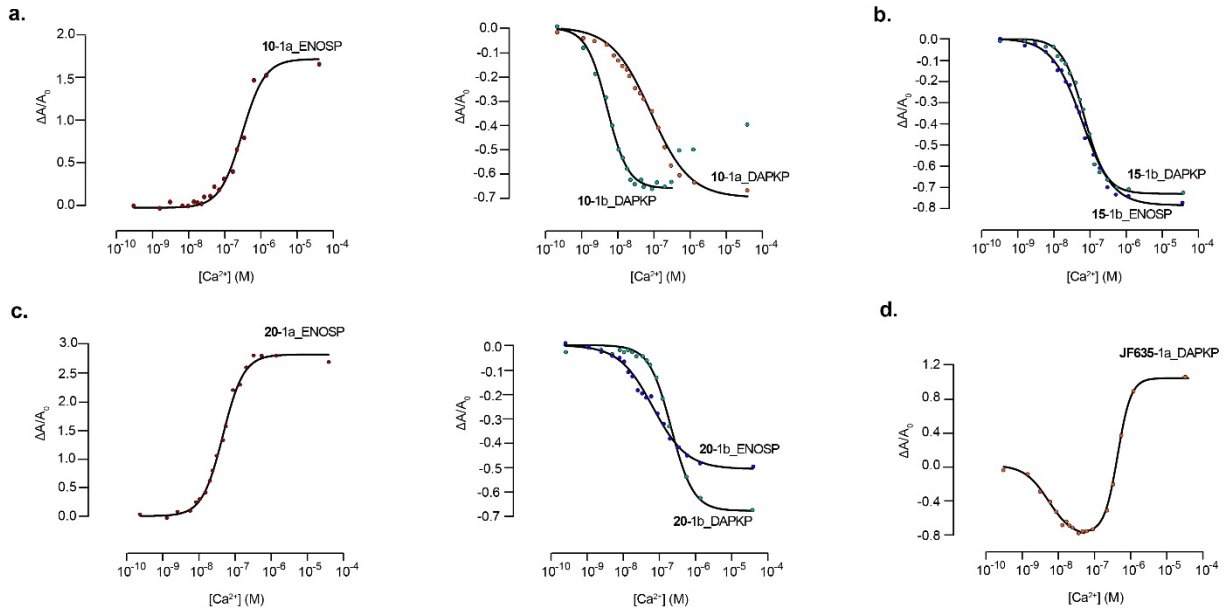


Figure S18: Ca^{2+} titrations of HaloCaMP variants labelled with 10-, 15-, 20- and JF635-HTL. (**a**) **10-HTL**, (**b**) **15-HTL**, (**c**) **20-HTL**, (**d**) **JF635-HTL**. All measurements were performed in 30 mM MOPS, 100 mM KCl, pH 7.2, containing 0.1 mg·mL^{-1} of CHAPS, to which either 10 mM EGTA (for the calcium-free state) or 10 mM $CaCl_2$ was added (for the calcium-saturated state); values are the mean of 2 replicates.

(**a**)

MLQNELALKLAGLDINKTGGSHHHHHHGSARRKWQKTGHAVRAIGRLSSPEFAAETFQAFRTSGDQMGQIEKI
FKVVYPVDDHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLF
RVTINGVTGWRLCERILAGGTGGSGGTGGSMVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTPI
QRIVLSGENGLKIDIHVIIPYEGLDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPN
ELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLI
GSEIARWLSTLEISGGGTGGSGGTGGSGGTGGSMAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGN
PTSSYVWRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALGFHW
AKRNPERVKGIAFMEFIRPIPTWDEWLEFARDQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTE
AELQDMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEKLTD
EEVDEMIREADIDGDGQVNYEEFVQMMTAKLEISGGGSGMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEG
EGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNY
KTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD
HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

(**b**)

MLQNELALKLAGLDINKTGGSHHHHHHGVRVIPRLDTLILVKAMGHRKRFGNPFRPKETFQAFRTSGDQMGQI
EKIFKVVYPVDDHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSL
LFRVTINGVTGWRLCERILAGGTGGSGGTGGSMVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSV
TPIQRIVLSGENGLKIDIHVIIPYEGLDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRF
PNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNP
DLIGSEIARWLSTLEISGGGTGGSGGTGGSGGTGGSMAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFL
HGNPTSSYVWRNIIPHVAPKHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALG
FHWAKRNPERVKGIAFMEFIRPIPTWDEWPFARDQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNP
TEAELQDMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEKL
TDEEVDEMIREADIDGDGQVNYEEFVQMMTAKISGGGSGMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEG
EGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNY
KTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD
HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

Annotations: purple – start codon; yellow – nuclear export signal (NES); red –His-tag; brown – linker; blue – peptide; grey – cpHaloTag; black – cpNanoLuc; orange – calmodulin; green – EGFP.

Figure S19: Amino acid sequences of HaloCaMP-NanoLuc BRET reporter proteins annotated with sequence features. (**a**) 1a_NanoLuc, (**b**) 1b_NanoLuc.

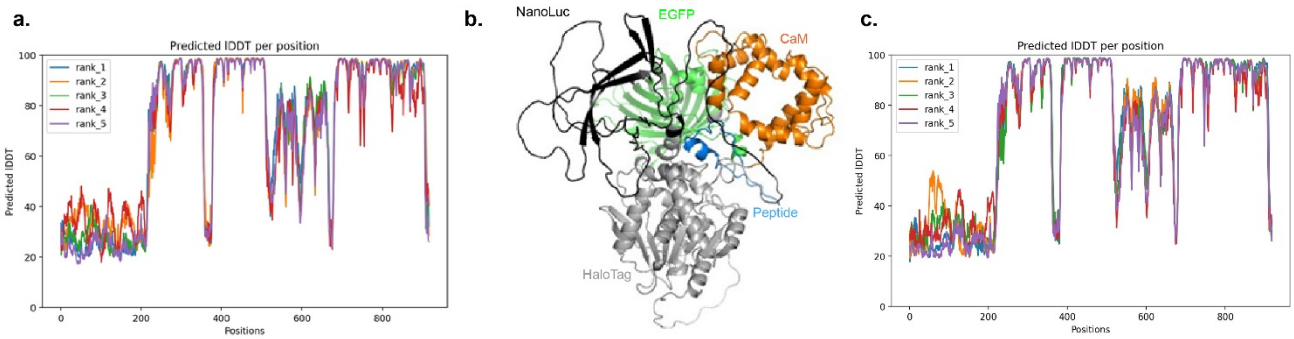


Figure S20: Predicted protein structures of HaloCaMP-NanoLuc BRET reporter proteins with ColabFold.[222, 223] (**a**) Predicted IDDT per position for 1a_NanoLuc. (**b**) Representative example of predicted Colabfold structure of 1b_NanoLuc. (**c**) Predicted IDDT per position for 1b_NanoLuc.

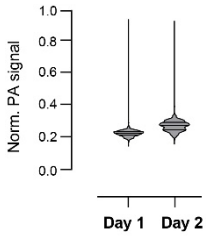


Figure S21: Control for the change in the normalised PA signal of HaloCaMP1a bacterial colonies over time. Unlabelled HaloCaMP1a colonies in the presence of MOPS buffer (30 mM MOPS, 100 mM KCl, pH 7.2). PA signal measured at 495 nm and 646 nm, normalised by EGFP signal at 495 nm. Data obtained by Kenana (Stiel lab).

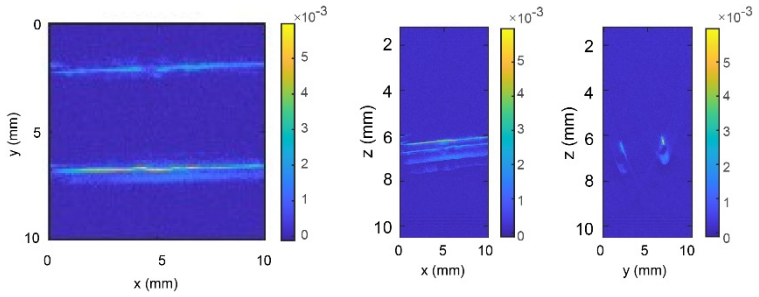


Figure S22: PAT imaging of solutions of methylene blue in tubeless gel wax phantoms. The phantoms were filled with a high concentration of methylene blue solution (0.27 mM) in water then submerged in a coupling medium of H_2O. $\lambda_{PA} = \lambda_{max}$ = 646 nm for **14**-HaloTag7. The (xy) images are maximum intensity projections (MIPs) and (xz and yz) images are MIPs of the reconstructed side view.

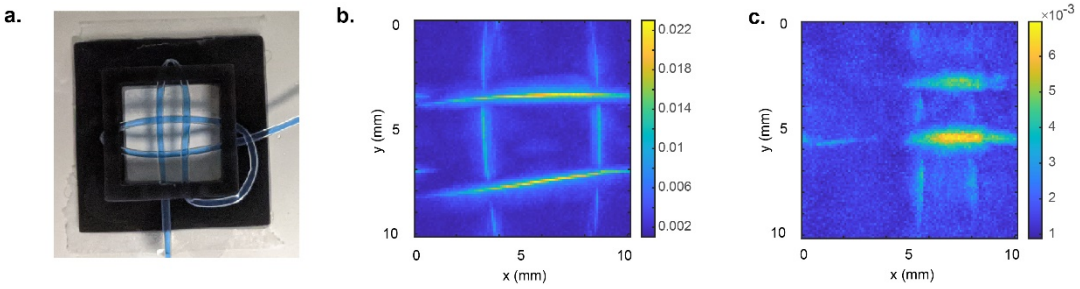


Figure S23: PAT imaging of solutions of methylene blue in tubes. Tubes were filled with a high concentration of methylene blue solution (0.27 mM) in water, and immobilised in a coupling medium of H_2O. (**a**) Photo of methylene blue-filled tube in the phantom set-up; (**b**) polyethylene tube (ID 0.86 mm, OD 1.27 mm, Hugo Sachs Elektronik); (**c**) Teflon tube (ID 1.07 mm, OD 1.47 mm, Alpha Wire). λ_{PA} = 646 nm = λ_{max} for **14**-HaloTag7. The silicon tubing resulted in minimal detectable signal.

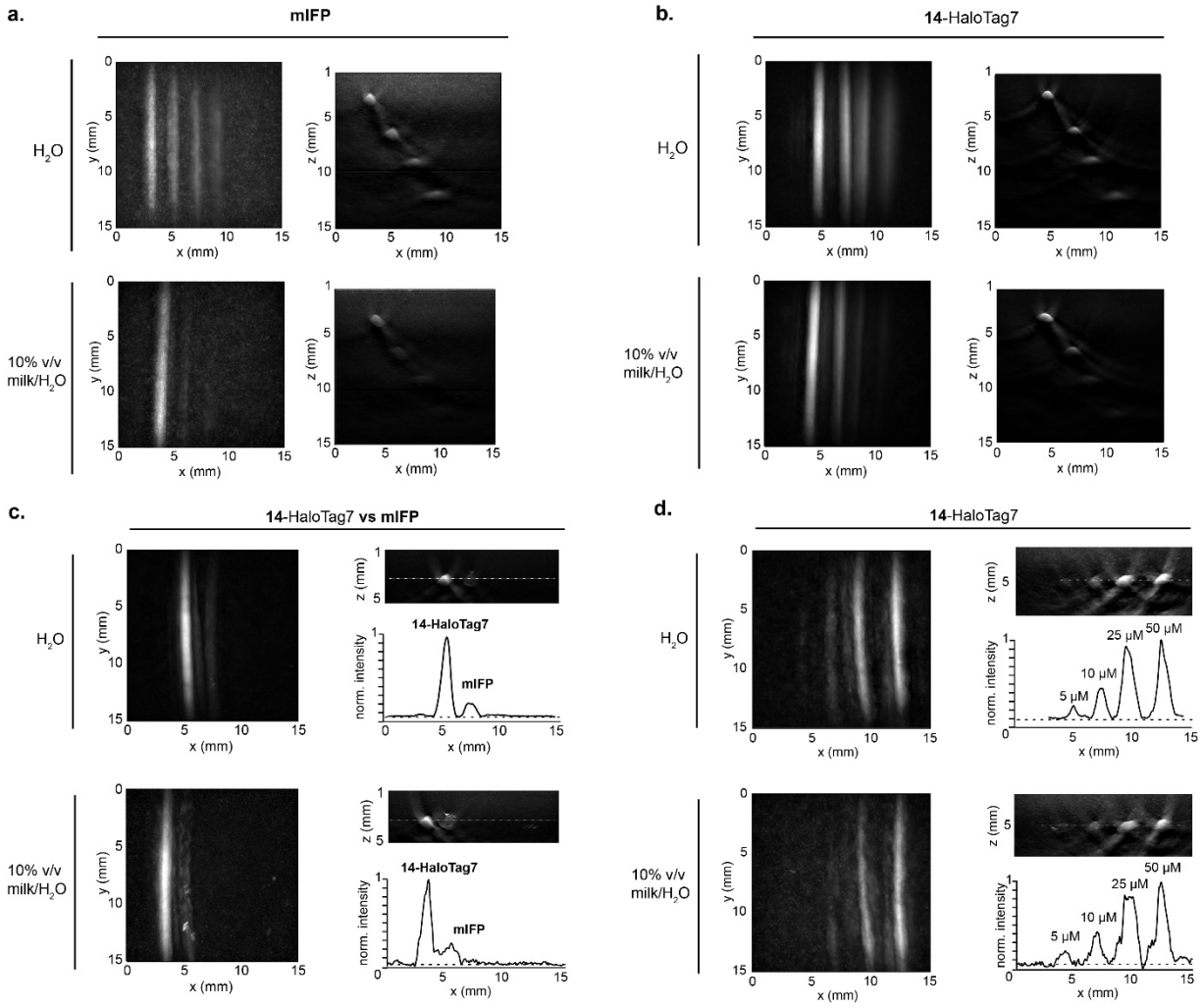


Figure S24: PAT imaging of solutions of our chemigenetic probes in tissue-mimicking phantoms. Tubes were filled with a 50 µM protein solution, and immobilised in a coupling medium of H_2O or 10% v/v milk/H_2O. Representative photoacoustic tomography images of tissue-mimicking phantoms of: (**a**) mIFP at different depths, representative of n = 2 images; (**b**) **14**-HaloTag7 at different depths, representative of n = 2 images; (**c**) side-by-side comparison of **14**-HaloTag7 (left tube) and mIFP (right tube) at ~3.5 mm depth; (**d**) **14**-HaloTag at 5, 10, 25 or 50 µM, at ~5 mm depth, representative of n = 2 images.

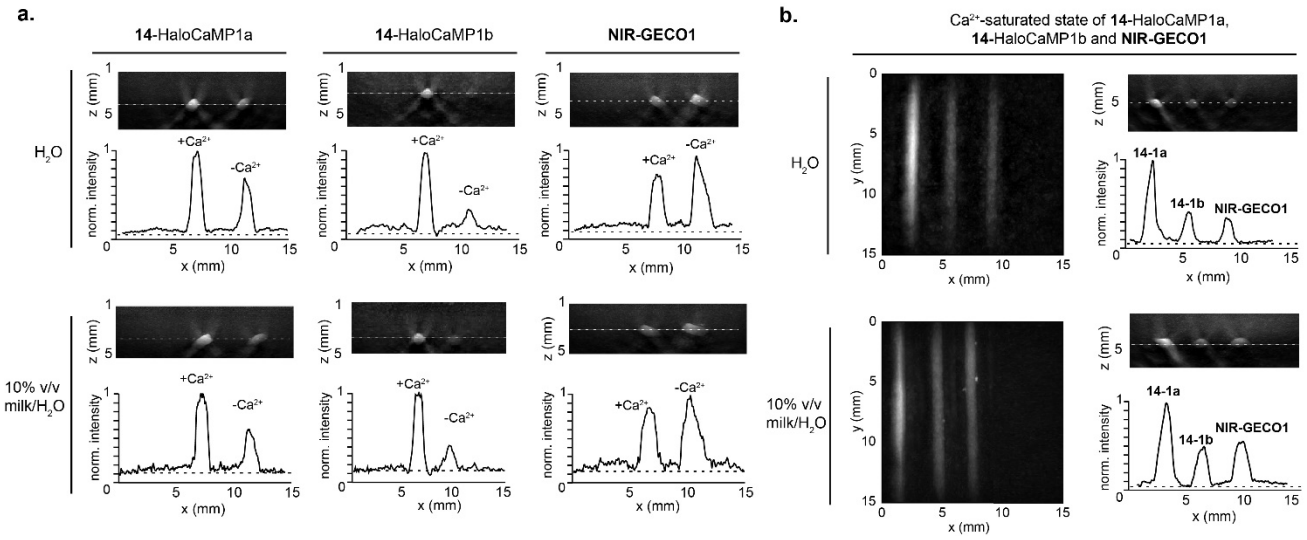


Figure S25: PAT imaging of solutions of our chemigenetic calcium sensors in tissue-mimicking phantoms. Tubes were filled with a 50 µM protein solution, and immobilised in a coupling medium of H_2O or 10% v/v milk/H_2O; Representative photoacoustic tomography images of tissue-mimicking phantoms of: (**a**) **14**-HaloCaMP1a, **14**-HaloCaMP1b or NIR-GECO1 in the calcium-saturated and the calcium-free states, at ~3.5 mm depth; images are representative of n = 3 replicates; (**b**) the calcium-saturated states of **14**-HaloCaMP1a **(14-1a**, left tubes), **14**-HaloCaMP1b **(14-1b**, middle tubes) and NIR-GECO1 (right tubes).

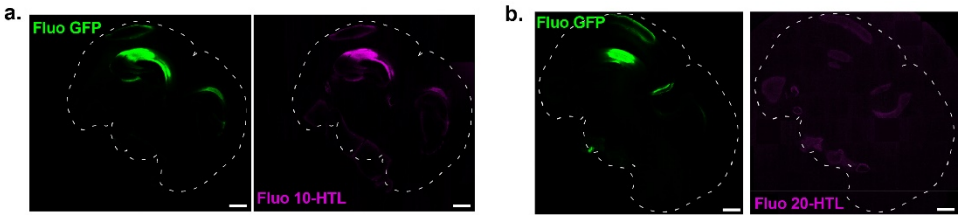


Figure S26: Fluorescence microscopy of mouse coronal brain slices labelled *ex vivo* with 10- and 20-HTL. (**a-b**) Confocal microscopy tile scans of the hippocampus after labelling with (**a**) **10-HTL** or (**b**) **20-HTL**, EGFP channel (left), far-red channel (right). Images are representative of n = 4 slices, from n = 1 mouse. Scale bars: 1 mm.

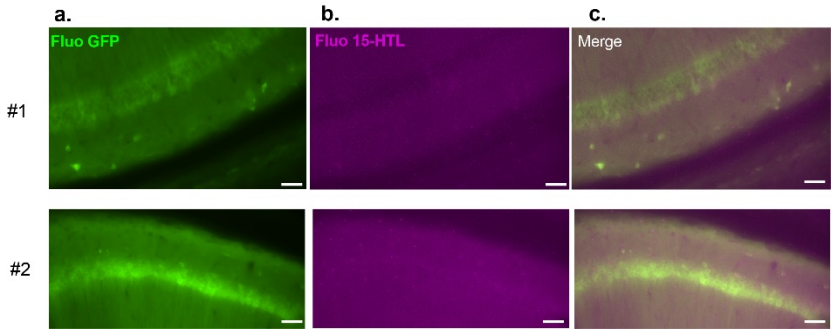


Figure S27: Fluorescence microscopy of mouse coronal brain slices labelled *in vivo* via tail injection of 15-HTL. (**a-c**) Representative widefield fluorescence images of 2 coronal slices of a mouse brain expressing HaloTag7-EGFP and labelled with **15-HTL** delivered via tail vein injection *in vivo*; (**a**) EGFP channel; (**b**) far-red channel; (**c**) overlay of the EGFP and far-red channels; scale bars: 50 µM.

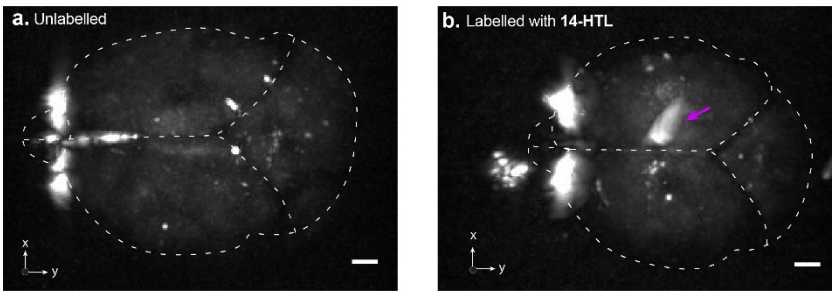


Figure S28: Photoacoustic tomography of whole *ex vivo* mouse brain expressing HaloTag7-EGFP in neurons including the olfactory bulbs. (**a-b**) PAT maximum intensity axial projections of the entire brain (same brains from Figure 65a, b) including strong endogenous signal from the olfactory bulbs for: (**a**) unlabelled brain; (**b**) labelled with **14-HTL** via intracerebroventricular injection *in vivo*. Scale bars: 1 mm.

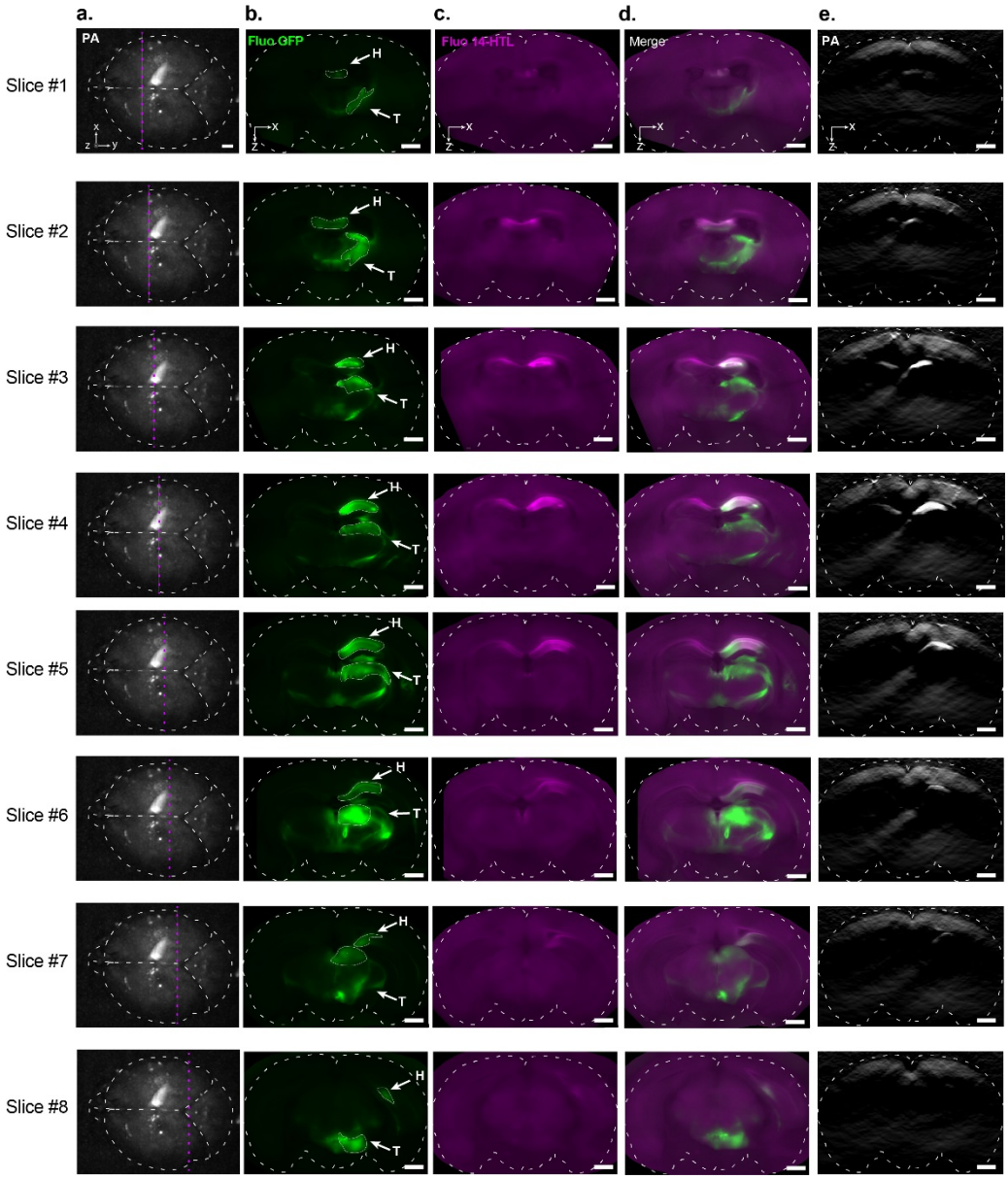
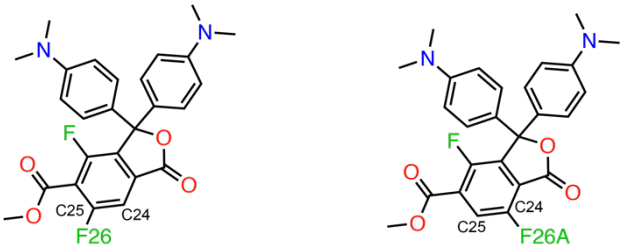


Figure S29: Photoacoustic tomography of a whole *ex vivo* mouse brain expressing HaloTag7-EGFP in neurons, labelled with 14-HTL delivered by intracerebroventricular injection *in vivo*, and fluorescence images of coronal slices. Each row corresponds to a coronal slice of the brain volume. (**a**) PAT maximum intensity axial projection of the entire brain between 2.5 and 5.75 mm from the surface of the Fabry-Pérot interferometer to exclude strong endogenous signal from the olfactory bulbs; the magenta dashed line indicates the position of the coronal slice (λ_{PA} = 646 nm). (**b**) Widefield fluorescence image in the EGFP channel, outlined areas indicate the hippocampus (H) and thalamus (T) ROIs showing expression of HaloTag which are used for quantification (Figure 65g). (**c**) Widefield fluorescence image in the far-red channel. (**d**) Overlay of the EGFP and **14-HTL** channels. (**e**) Corresponding coronal slice from the PAT volume. Scale bars: 1 mm. These slices are from one of three mice injected with **14-HTL** (mouse #1).

VI.2 X-Ray Crystallography

To confirm the position of the two fluorine atoms on compound **23-HTL**, we solved a crystal structure of the isolated intermediate **S21**. Suitable crystals were grown by vapour diffusion of cyclohexane in a solution of MeOH/CH_2Cl_2, and X-Ray diffraction data was collected by Dr. Dieter Schollmeyer at the Johannes Gutenberg Universität Mainz.

X-Ray diffraction revealed the structure of **S21** (CCDC Deposition Number 2336812), the major product present at 90%, bearing the hydrogen at the 4-position (C25 in the atom numbering scheme, Figures C1, C2). In addition, the crystal revealed the presence of another regioisomer **S21** present at 10% proportion, and bearing the hydrogen at the 5-position. The two regioisomers could not be fully separated at that stage, and the 90:10 mixture was used for the synthesis of the HaloTag ligand **23-HTL**. This compound could be purified as a single isomer, as shown by the ^{1}H NMR spectrum and LCMS trace, and the NMR characteristics (chemical shifts and coupling constants for the fluorines and the hydrogen on the bottom ring) are in good accordance with that of compound **S21**, confirming the chemical structure proposed for **23-HTL**.



S21 (major - 90%) **S21'** (minor - 10%)

Figure C1: Chemical structures and proportions of the isomers observed in the single crystal of compound **S21**.

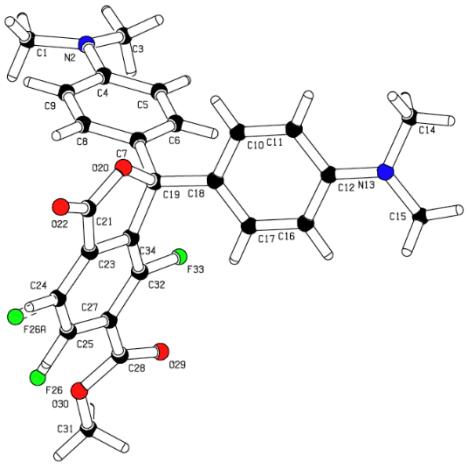


Figure C2: Molecular structure with atom numbering scheme. The crystal was composed of 90% of compound bearing fluorine at C25 (= F26), and 10% of compound bearing fluorine at C24 (= F26A).

Table C1: Crystal data and structure refinement

Identification code	f2mgome	
Empirical formula	$C_{26}H_{24}F_2N_2O_4$	
moiety formula	$C_{26}H_{24}F_2N_2O_4$	
Formula weight	466.47	
Temperature	120(2) K	
Wavelength, radiation type	1.54178Å, CuKα	
Diffractometer	STOE STADIVARI	
Crystal system	Monoclinic	
Space group name, number	P 2_1/c, (14)	
Unit cell dimensions	a = 19.0356(5) Å	α = 90°
	b = 6.76850(10) Å	β = 118.528(2)°
	c = 19.9025(5) Å	γ = 90°
Volume	2252.94(10) Å^3	
Number of reflections	15982	
and range used for lattice parameters	4.45° <=θ<= 68.46°	
Z	4	
Density (calculated)	1.375 Mg/m^3	
Absorption coefficient	0.872 mm^{-1}	
Absorption correction	Integration	
Max. and min. transmission	0.9773 and 0.9080	
F(000)	976	
Crystal size, colour and form	0.030 x 0.060 x 0.120 mm^3, colorless needle	
Theta range for data collection	4.449 to 68.725°.	
Index ranges	-22<=h<=22, -7<=k<=8, -24<=l<=23	
Number of reflections:		
collected	24162	
independent	4119 [R(int) = 0.0426]	
observed [I>2sigma(I)]	2870	
Completeness to theta = 67.7°	99.5 %	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	4119 / 1 / 321	
Goodness-of-fit on F^2	0.977	
Final R indices [I>2sigma(I)]	R1 = 0.0421, wR2 = 0.1052	
R indices (all data)	R1 = 0.0638, wR2 = 0.1126	
Largest diff. peak and hole	0.365 and -0.275 eÅ^{-3}	
Remark	position of F26 split at C24 (10%) and C25 (90%)	

VI.3 Materials and Methods for Chapter II

VI.3.1 General Experimental Information

VI.3.1.1 Synthesis

Compounds **7**, **8**, **4-HTL**, **7-HTL, 8-HTL**, **24-HTL and 25-HTL** were synthesised as previously reported[168] or gifted by the Lavis group (Janelia Research Campus, HHMI).

Commercial reagents were obtained from reputable suppliers (e.g. Merck, TCI, BLDpharm) and used as received. All solvents used for chemical reactions were of anhydrous grade, purchased in septum-sealed bottles stored under an inert atmosphere. Reactions under an inert atmosphere were sealed with septa and purged under vacuum/argon on a Schlenk line. Reactions were performed in round-bottomed flasks or septum-sealed vials.

Reactions were monitored by thin layer chromatography (TLC) on precoated aluminium plates (silica gel 60 F254, 200 µm thickness) or by LCMS (Agilent 1260 Infinity II; ZORBAX SB-C18 18 µm 80 Å, 2.1x50 mm column, 5 to 20 µL injection, 5–95% CH_3CN/H_2O gradient with constant 0.1% v/v HCO_2H additive, 8 to 10 min run, 0.5 mL/min flow, ESI positive ion mode, detection at 254 nm). TLC plates were visualised by UV illumination (254 nm) or developed with stains (e.g. cerium ammonium molybdate or Seebach stain). Compounds were purified by flash chromatography on an automated purification system (Biotage Isolera One) using pre-packed silica cartridges (Biotage Sfär Duo, 60 Å pores, 60 µm particles size) or by preparative HPLC (Agilent 1260 Infinity II, Phenomenex Gemini NX 21.2x150 mm, 5 µm C18 column). High-resolution mass spectrometry was performed by the Metabolomics Core Facility at EMBL Heidelberg.

NMR spectra were recorded on a 400 MHz spectrometer (Bruker 400 UltraShield) in deuterated solvents. All spectra were recorded at 298 K. ^{1}H and ^{13}C chemical shifts (δ) were referenced to residual solvent peaks, and ^{19}F chemical shifts (δ) were referenced to $CFCl_3$. Data for ^{1}H NMR spectra are reported as chemical shift (δ in ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentuplet, dd = doublet of doublets, ddd= doublet of doublets of doublets, dt = doublet of triplets, tt = triplet of triplets, dtt = doublet of triplets of triplets, m = multiplet, br s = broad signal), coupling constant (*J* in Hz), integration. Data for ^{13}C NMR spectra are reported by chemical shift (δ in ppm) with hydrogen substitution (C, CH, CH_2, CH_3) information obtained from DEPT spectra. Data for ^{19}F NMR are reported by chemical shift (δ in ppm) and coupling constant (*J* in Hz). Data was processed using Mnova from Mestrelab. The ^{13}C NMR spectra are not reported for compounds containing fluorinated aryl rings due to complex couplings. Determination of sample concentration for UV-Vis and fluorescence spectroscopy was performed by ^{1}H NMR of samples in DMSO-d_6 containing 5 mM DMF for relative integration.

VI.3.1.2 Cloning, Protein Expression and Purification

pET51b-HaloTag7 (Addgene, 167266), pBR322-HaloTag9 (Addgene, 169324) and pET51b-rHCaMP (Addgene, 187106) plasmids were received as gifts from Kai Johnsson (MPI for Medical Research, Heidelberg). pRSET-HaloCaMP1a-EGFP (Addgene, 138327) and pRSET-HaloCaMP1b-EGFP (Addgene, 138328) were received as gifts from Eric Schreiter (Janelia Research Campus, HHMI). The HaloCaMP1a and HaloCaMP1b inserts were cloned via Gibson assembly into the pET51b backbone for controlled protein overexpression. Sanger sequencing (Eurofins Genomics) was used to verify DNA sequences in the obtained plasmid constructs. The pDuEx2-NIRGECO1 plasmid (Addgene, 113680) was used for NIR-GECO1 insert amplification followed by Gibson assembly into the pET51b backbone.

HaloTag7, HaloCaMP1a and HaloCaMP1b proteins were expressed and purified as previously described.[143] In summary, pET51b plasmids containing either HaloTag7, HaloCaMP1a-EGFP or HaloCaMP1b-EGFP, were transformed into T7 Express Competent *E. coli* cells (NEB). Cell colonies were grown in LB medium

containing ampicillin (100 μg·mL^{-1}) overnight at 37°C at 200 r.p.m. The starting culture was diluted 100-fold in LB containing ampicillin (100 μg·mL^{-1}) and incubated at 37°C, 200 r.p.m until the OD reached 0.4-0.6. Protein expression was then induced by IPTG (1 mM) and the culture incubated at 18°C, 200 r.p.m for 14-20 h. After centrifugation, cell pellets were lysed [50 mM Tris-HCl (tris(hydroxymethyl)aminomethane hydrochloride), pH 8.0, 300 mM NaCl, 5 mM imidazole, 5 mM β-mercaptoethanol, protease inhibitor (cOmplete, Roche, 25x stock), 0.01 mg·mL^{-1} DNase, 5 mM $MgCl_2$, 5 mg·mL^{-1} n-octyl-β-D-glucopyranoside] in a microfluidizer and the lysate was clarified by centrifugation. Purification from lysate was achieved by nickel-affinity chromatography using HisTrap^{TM} HP column [(GE Healthcare), equilibration buffer 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10-20 mM imidazole, 5 mM β-mercaptoethanol; elution buffer 50 mM Tris-HCl pH 8.0, 100mM NaCl, 50-250 mM imidazole (stepwise gradient), 5 mM β-mercaptoethanol]. Selected protein fractions were then pooled, followed by size exclusion chromatography using a Superdex 200 16/600 column (GE Healthcare) at a flow rate of 0.5 mL·min^{-1} in 50 mM Tris-HCl, 100 mM NaCl, pH 7.4. Protein concentration was estimated using the extinction coefficient of EGFP ($\varepsilon_{488\ nm}$) = 55900 M^{-1}·cm^{-1}) or using the extinction coefficient at 280 nm calculated from the protein sequence (Expasy).

HaloTag9 and rHCaMP were expressed as described above. Purification from lysate was achieved by gravity nickel-affinity chromatography using PureCube Ni-NTA Agarose (Cube Biotech), and the same buffers as above except the elution buffer (50 mM Tris-HCl pH 8.0, 100mM NaCl, 500 mM imidazole), followed by PD-10 desalting column (Cytiva) for buffer exchange to 50 mM Tris-HCl, 100 mM NaCl, pH 7.4. Protein concentration was estimated using the extinction coefficient at 280 nm calculated from the protein sequence (Expasy).

NIR-GECO1 and mIFP (Addgene 54620) were expressed and purified as for HaloTag9 and rHCaMP, except with an additional incubation step with the biliverdin cofactor. For this, 100 mL of cell lysates were incubated on ice with 300 μL of biliverdin (20 mg·mL^{-1} stock solution in DMSO), for 1 h on ice before loading onto the Ni-NTA Agarose column.

VI.3.2 UV-Vis, Fluorescence and Photoacoustic Spectroscopy

VI.3.2.1 UV–Vis and Fluorescence Spectroscopy

All measurements were performed at room temperature (23±2°C). Compounds were prepared as stock solutions at 1 mM in DMSO which were diluted in solvents and buffers to a final DMSO concentration not exceeding 1% v/v. Spectroscopy in buffer solutions was performed using 1-cm path length polystyrene cuvettes (Ratiolab GmbH). With organic solvents 1-cm path length quartz cuvettes were used (Hellma Suprasil Quartz). Absorption spectra were recorded on a Cary Model 60 spectrophotometer (Agilent). Fluorescence spectra were recorded on a JASCO spectrofluorometer (FP-8500). Data was analysed and graphs were plotted using Prism (GraphPad). Absorption spectra, maximum absorption wavelength (λ_{max}), extinction coefficient (ε) and maximum emission wavelength (λ_{em}) were measured in triplicate under the following conditions, reported values for ε are the mean of 3 replicates. Spectra were recorded at 1.25 μM dye concentration, unless otherwise stated. For the free dyes, measurements were made in 10 mM HEPES (4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid) pH 7.4. For the Malachite Green lactones, ε were also measured with 10 mM SDS (sodium dodecyl sulphate) additive. For the corresponding HaloTag ligands, 0.1 mg·ml^{-1} CHAPS (3-((3-cholamidopropyl)dimethylammonio)-1-propanesulfonate) was added to the buffer. For measurements in the presence of proteins, the dye ligand was incubated with 1.5 equivalent (eq) of purified protein for at least 3 h at room temperature. For assessment of labelling kinetics, HaloTag ligand (to a final concentration of 5 μM) was added to a solution of HaloTag protein (7.5 μM) and absorption spectra measured at 5 min intervals for the first hour and subsequently 30 min intervals. In the competition assay, **JF549-HTL** (to a final concentration of 10 μM) was added after absorption had reached approximately 50% of its maximum value. An exponential model

was used to fit the binding curves assuming pseudo-first order association kinetics with either one (equation 3) or two (equation 4) association phases.

$$y = a_0 + a_1(1 - e^{-b_1 t})$$

Equation 3: Exponential fit for a single association phase. The rate constants correspond to the value of b_1.

$$y = a_0 + a_1(1 - e^{-b_1 t}) + a_2(1 - e^{-b_2 t})$$

Equation 4: Exponential fit for two association phases. The rate constants correspond to the value of b_1 and b_2.

To further quantify the open-closed equilibrium, we measured absorption spectra of dyes at 5 µM in MeCN/H_2O mixtures,[2] and in 2,2,2-trifluoroethanol (TFE) containing 0.1% v/v trifluoroacetic acid (TFA).[11] The MeCN/H_2O solvent mixture was found to be more reliable in quantifying the equilibrium of our ligands than the commonly used dioxane/water titration, probably because of improved solubility in acetonitrile. The Malachite Green scaffold did not fully open in either acidic or basic conditions. To try to determine the maximum absorption of the Malachite Green scaffold, 5 µM solutions of **14** were prepared in acidic (TFE with TFA or H_2SO_4 0.1%-1%) or basic conditions (TFE, MeOH or EtOH with NEt_3, pyridine or DIPEA 0.01%-10%) and absorbance spectra were measured.

Fluorescence quantum yields (Φ) were measured using an absolute quantum yield measurement system (Quantaurus, C11347, Hamamatsu). Measurements were carried out using dilute samples (A<0.1), self-absorption corrections were performed using the Quantaurus software. The fluorescence quantum yield of compound **14-HTL** bound to HaloTag7 was calculated precisely using the relative method with Oxazine 1 (Φ = 0.11 in EtOH) as a standard.[12]

VI.3.2.2 Photoacoustic Spectroscopy

A custom-built multimodal spectroscopy set-up was used to measure photoacoustic spectra (Figure 38a). The set-up consists of a water tank holding a standard 1-cm glass cuvette containing the sample. The system utilizes a 100 Hz pulsed tuneable laser source (EVO I OPO, Innolas Laser GmbH) to excite photoacoustic signals, which are subsequently detected using a focused water-immersion transducer operating at a frequency of 0.5 MHz (V301, Olympus IMS). The acquired ultrasound signals were recorded via a data acquisition (DAQ) card with a sampling rate of 125 MSa/s (ATS9440, Alazar Technologies Inc.) and processed with a custom MATLAB code. The peak-to-peak value of the ultrasound waveform was used as a metric of photoacoustic signal. The photoacoustic spectra were measured by sweeping through the wavelength range of the excitation laser (420-709 nm) and acquiring photoacoustic signals at each wavelength. Each spectrum was corrected for changes in laser pulse energy with wavelength, and the background photoacoustic signal from a cuvette with corresponding buffer solution was subtracted to remove any signal from the glass cuvette. Single wavelength measurements were recorded at the maximum absorption wavelength and averaged for 1000 pulses.

For photostability measurements, the photoacoustic spectroscopy set-up was adapted into a well plate reader configuration. The laser light was adjusted to illuminate the entire well and photobleach the sample homogenously. Single wavelength photobleaching were recorded at the maximum absorption wavelength for 1 h at 100 Hz pulse repetition rate with average power ranging from 100 to 500 mW depending on the wavelength. The photoacoustic amplitudes were corrected for power fluctuations between pulses and normalised to initial values. The photobleaching data was then plotted against the cumulative deposited energy to compare bleaching rates of samples at different wavelengths where the laser power is different. 1 mL samples were prepared and absorption matched to ~0.065 cm^{-1} using the Cary spectrometer. Absorption was measured again after irradiation for validation.

VI.3.2.3 Calcium Titrations and K_d Determination

For the high affinity sensors, calcium titrations were performed using a commercial EGTA (ethylene glycol-bis(β-aminoethyl ether)-*N,N,N′,N′*-tetraacetic acid) /Ca·EGTA buffer system (Invitrogen), to which 0.1 mg·ml^{-1} CHAPS was added, following the associated protocol. For the low affinity sensors, buffer solutions (30 mM MOPS (3-(*N*-morpholino)propanesulfonic acid), 100 mM KCl, pH 7.2) containing various calcium concentrations (0 µM, 20 µM, 100 µM, 200 µM, 500 µM, 1 mM, 2 mM, 5 mM, 10 mM, 20 mM, 50 mM, 100 mM, 500 mM) were prepared.

For calcium titrations of the HaloTag ligands in the presence of HaloCaMP protein, the dye was incubated with 1.5 eq of purified HaloCaMP protein for 3 h at room temperature. All the calcium titrations were performed in duplicate. K_d values were obtained by fitting the curve of the absorption at λ_{max} as a function of free [Ca^{2+}] with the Hill equation.

$$y = \frac{y_{max} \cdot x^h}{K_d^h + x^h}$$

Equation 5: The Hill equation. h: Hill coefficient, K_d: dissociation constant.

VI.3.2.4 Calcium binding and unbinding kinetics

14-HTL or **JF635-HTL** were incubated with purified HaloCaMP1a or 1b (1.5 eq) at room temperature overnight. For calcium unbinding, the labelled protein samples or NIR-GECO1 were diluted to 5 - 50 µM in 30 mM MOPS, 100 mM KCl, 100 µM $CaCl_2$ (10 µM $CaCl_2$ for NIR-GECO1). A stopped flow device (SFM-3000, BioLogic) was used to rapidly mix the calcium sensors with a solution of 30 mM MOPS, 100 mM KCl, 10 mM EGTA, pH 7.2 in a 1:1 ratio. For calcium binding, the labelled protein samples were diluted to 5 - 30 µM in 30 mM MOPS, 100 mM KCl, 10 mM EGTA. The same stopped-flow device was used to rapidly mix the calcium sensors with a solution of 30 mM MOPS, 100 mM KCl, 10 mM Ca·EGTA, pH 7.2 in different ratios keeping the concentration of the sensor constant but altering free calcium concentration between 1.67×10^{-8} M and 4.52×10^{-7} M.

For both binding and unbinding, the samples were excited using a Xe lamp at λ_{max} for each sensor, and absorbance detected by a photomultiplier tube. Changes in absorption over time were fit to the exponential models described in Equations 3 and 4. Measurements with **14**-HaloCaMP1b were made with protein prepared from three independent dye labelling reactions, while for the published probes **JF635**-HaloCaMP1a and **JF635**-HaloCaMP1b only a single dye labelling reaction was used.

VI.3.2.5 Calcium Selectivity

14-HTL (60 µM) was incubated with purified HaloCaMP1a and HaloCaMP1b (90 µM, 1.5 eq) at room temperature for 3 h. For HaloCaMP1a, the labelled protein was diluted to 1.25 µM in 30 mM MOPS pH 7.2, 100 mM KCl, 10 µM EGTA. $CaCl_2$ or $MgCl_2$ were added to final concentrations of either 500 mM $CaCl_2$, 5 mM $MgCl_2$ or 100 mM $MgCl_2$. $CaCl_2$ was then added to the $MgCl_2$ containing solutions to reach 500 mM Ca^{2+} concentration. Absorption was measured at each step. For HaloCaMP1b, the labelled protein was diluted to 1.25 µM in 30 mM MOPS pH 7.2, 100 mM KCl, 10 mM EGTA. $CaCl_2$ or $MgCl_2$ were added to a concentration of 1.5 mM (sufficient free Ca^{2+} concentration to reach the calcium-saturated state of the sensor). To the $MgCl_2$-containing solution, 1.5 mM $CaCl_2$ was subsequently added. All measurements were performed in duplicate.

VI.3.2.6 pH Stability

Purified HaloCaMP1a and HaloCaMP1b (90 µM, 1.5 eq) and **14-HTL** or **JF635-HTL** (60 µM) were incubated at room temperature for 3 h. The labelled protein was diluted to 1.25 µM, into the following buffer systems: citrate (pH = 4.0, 5.0), phosphate (pH = 6.0, 7.0), Tris-HCl (pH = 8.0, 9.0), containing 10 mM buffer, 150 mM NaCl and 0.1 $mg \cdot mL^{-1}$ CHAPS. 10 mM EGTA was added for the calcium-free state, or excess $CaCl_2$ was added for the calcium-bound state (i.e. 500 mM for HaloCaMP1a or 10 mM for HaloCaMP1b). Absorbance spectra were recorded, all measurements were performed in duplicate.

VI.4 Materials and Methods for Chapter III

VI.4.1 Cloning, Protein Expression and Purification

ENOSP or DAPKP peptide blocks (from Integrated DNA Technologies) were cloned via Gibson assembly into the original HaloCaMP1a and HaloCaMP1b plasmids with pET51b backbone for controlled protein overexpression, with an EGFP C-terminal fusion. The HaloCaMP-NanoLuc BRET reporters' (1a_NanoLuc, 1b_NanoLuc) inserts (from Integrated DNA Technologies) were cloned via Gibson assembly into the same pET51b backbone. Sanger sequencing (Eurofins Genomics) was used to verify DNA sequences in the obtained plasmid constructs.

The HaloCaMP affinity variants and HaloCaMP-NanoLuc reporters were expressed and lysed as described for the HaloTag-based proteins in Chapter II (see section VI.3.1.2). For the HaloCaMP affinity variants, purification from lysate was achieved by gravity nickel-affinity chromatography using PureCube Ni-NTA Agarose [(Cube Biotech), equilibration buffer 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10-20 mM imidazole, 5 mM β-mercaptoethanol; elution buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 500 mM imidazole)], followed by PD-10 desalting column (Cytiva) for buffer exchange to 50 mM Tris-HCl, 100 mM NaCl, pH 7.4. For purification of the HaloCaMP-NanoLuc proteins, the cleared lysate was loaded overnight at 4°C onto a HisTrap^{TM} HP column (GE Healthcare), and the same elution procedure followed as described in VI.3.1.2. The buffer was exchanged to 50 mM Tris-HCl, 100 mM NaCl, pH 7.4 using a PD-10 desalting column (Cytiva). Protein concentration was estimated for all constructs using the extinction coefficient of EGFP ($\varepsilon_{488\ nm}$ = 55900 $M^{-1}\cdot cm^{-1}$).

VI.4.2 UV–Vis and Fluorescence Spectroscopy

Measurements were performed as described for Chapter II (see section VI.3.2.1). Spectra were recorded at 1.25 µM dye concentration, unless otherwise stated. For measurements in the presence of proteins, the dye ligand was incubated with 1.5 equivalent (eq) of purified protein for at least 3 h at room temperature. Reported values for ε_{sat} with the HaloCaMP affinity variants are the mean of 3 replicates, and were measured in 30 mM MOPS pH 7.2, 100 mM KCl, 10 mM $CaCl_2$. For labelling of 1a_NanoLuc with **14-HTL**, the absorption spectra were measured every 30 mins for at least 4 h to follow binding/ aggregation.

VI.4.3 Calcium Titrations and K_d Determination

Calcium titrations were performed as described in Chapter II (see section VI.3.2.3), using a commercial EGTA (ethylene glycol-bis(β-aminoethyl ether)-*N,N,N′,N′*-tetraacetic acid)/Ca·EGTA buffer system (Invitrogen) to which 0.1 $mg\cdot ml^{-1}$ CHAPS was added, following the associated protocol.

The HaloTag ligands were incubated with 1.5 eq of purified HaloCaMP protein variants for 3 h at room temperature. All the calcium titrations were performed in duplicate. K_d values were obtained by fitting the curve of the absorption at λ_{max} as a function of free Ca^{2+} concentration with the Hill equation (equation 5).

VI.4.4 Plate readers

Absorption plate reader measurements were performed with a BioTek Epoch2 microplate spectrophotometer. Absorption values were measured at λ_{max}, corrected for pathlength via absorption at 977/ 900 nm and absorption from blank buffer-only wells subtracted. For the purified protein measurements, **14-HTL** was incubated with HaloTag for 3 h (20 µM dye, 1.5 eq protein, in 10 mM HEPES, pH 7.4, CHAPS 0.1 mg mL^{-1}) and then diluted to final concentrations of 0.05, 0.1, 0.25, 0.35, 0.5, 0.75, 1, 2 µM in triplicate in a 96-well plate with 150 µL volume in each well. For the HaloCaMP proteins, **14-HTL** was incubated with HaloCaMP1a or 1b for 3 h (10 µM dye, 1.5 eq protein, in 100 mM Tris-HCl, pH 7.5, 300 mM NaCl, CHAPS 0.1 mg mL^{-1}) and then diluted to

final concentrations of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.75, 1, 2, 3 μM in a 96-well plate with 150 μL volume in each well. Each concentration of protein was diluted with either 30 mM MOPS pH 7.2, 100 mM KCl, 10 mM $CaCl_2$, or 30 mM MOPS pH 7.2, 100 mM KCl, 10 mM EGTA.

Fluorescence plate reader measurements were performed with a TECAN Infinite M1000 Pro fluorescent plate reader. EGFP protein concentration could be determined from a fluorescence calibration curve performed by Begoña (λ_{ex} = 488 nm, λ_{em} = 507 nm). For measurement of **JF635-HTL** fluorescence the following parameters were used: λ_{ex} = 640 nm, λ_{em} = 660 nm, excitation and emission bandwidths of 5 nm.

VI.4.5 Absorption lysate screening assay

pRSET plasmids containing either HaloTag7-EGFP, HaloCaMP1a-EGFP or HaloCaMP1b-EGFP, were transformed into T7 Express Competent *E. coli* cells (NEB). Colonies were grown overnight at 37°C on LB agar plates containing ampicillin. Green fluorescent colonies were picked and inoculated 24- or 96-deep well plates with LB medium containing ampicillin (100 $\mu g \cdot mL^{-1}$), and incubated at 30°C, 300 r.p.m for 72 h. The cultures were pelleted by centrifugation (4000 r.p.m, 30 min) and the pellets were resuspended in lysis buffer (100 mM Tris-HCl pH 7.5, 300 mM NaCl, 0.1 mg mL^{-1} lysozyme, 1 mM PMSF, DNase; 2 mL for 24-well plates or 500 μL for 96-well plates). The cells were lysed via freeze-thawing cycles (5x) and incubated overnight at 30°C. The supernatant was isolated via centrifugation (4000 r.p.m, 30 min) and a 150 μL aliquot from each well transferred to a transparent 96-well plate (Thermo Scientific™ Nunc™ MicroWell™ 96-Well Sterile Microplate). Protein concentration was determined via EGFP fluorescence using the fluorescent plate reader.

Ni-NTA magnetic beads (Cube Biotech) were equilibrated by mixing with lysis buffer (500 μL buffer per 40 μL of 25% magnetic bead suspension) before separation and buffer removal using a magnetic plate. The cleared lysate was added to fresh 24- or 96- deep well plates with equilibrated magnetic beads and incubated for 1 h at 4°C on a tilt shaker. **14-HTL** (13 μM stock in DMSO, 2 eq) was added to half of these wells for the incubation period (pre-labelling). The beads were separated and supernatant removed using a magnetic plate. 500 μL/2 mL (for 24-, 96-well plates respectively) wash buffer (100 mM Tris-HCl pH 7.4, 300 mM NaCl) was added to each well and mixed by vortexing then the beads were allowed to separate on the magnetic stand and supernatant removed. 150/300 μL elution buffer (100 mM Tris-HCl pH 7.4, 300 mM NaCl) was added and the supernatant transferred to a transparent 96-well plate. Protein concentration was determined via EGFP fluorescence using the fluorescent plate reader. The remaining unlabelled wells were incubated with **14-HTL** (50 μM stock in DMSO, 0.5 eq, 1.5 – 10 μL) at 4°C with shaking for 1 h (post-labelling). Since the protein concentration in the 24-well plates was higher, these were used for the calcium assay. The eluate from each well was split into two and either high Ca^{2+} (100 mM MOPS pH 7.2, 100 mM KCl, 100 mM $CaCl_2$) or EGTA (100 mM MOPS pH 7.2, 100 mM KCl, 100 mM EGTA) buffer was added (7.5 μL volume, total 150 μL per well). The absorbance at 488 and 645 nm was measured for the calcium-saturated and calcium-free states using the plate reader.

Measurements were also performed on the cleared lysate without the magnetic bead purification or after filtration of the lysate through 0.2 μm filters.

VI.4.6 Photoacoustic spectroscopy of HaloCaMP1a lysates

HaloCaMP1a lysates were prepared as above in lysis buffer except with either 100 mM NaCl or 300 mM NaCl. **14-HTL** (25 μM stock in DMSO, 0.5 eq) was incubated with the lysates for 3 h. 150 μL of high Ca^{2+} (100 mM MOPS pH 7.2, 100 mM KCl, 100 mM $CaCl_2$) or EGTA (100 mM MOPS pH 7.2, 100 mM KCl, 100 mM EGTA) buffer was added to each 3 mL. The photoacoustic spectra were measured on the custom-built set-up outlined in Chapter II (see section VI.3.2.2).

VI.4.7 ADPA Photosensitiser assay

Solutions of either pre-labelled **14**-HaloTag7 (1, 2.5, 5, 7.5, 10 µM), **14-HTL** (10 µM) or HaloTag7 (10 µM) were mixed with ADPA (50 µM) in 10 mM HEPES pH 7.4, in a transparent 96-well plate. The plate was illuminated with a custom-built LED box for 1 hour at 650 nm, then stored in the fridge overnight, and illuminated for 2 hours the next morning. The fluorescence of ADPA was measured using the TECAN plate reader (λ_{ex} = 374 nm, λ_{em} = 390 nm). The values were normalised to the values before overnight incubation.

VI.4.8 NanoLuc binding assay

Lysate from *E. coli* expressing HaloCaMP1a, HaloCaMP1b, 1a_NanoLuc and 1b_NanoLuc was prepared as previously described. 150 µL of each cleared lysate in triplicate was added to a transparent 96-well plate for protein concentration determination using EGFP fluorescence. 1 mL of each lysate was labelled with **JF635-HTL** (0.5 eq, 13 µM stock in DMSO) for 2 h. 135 µL of each labelled lysate was added to 6 wells in the 96-well plate and 15 µL of either high Ca^{2+} (100 mM MOPS pH 7.2, 100 mM KCl, 3.5 M $CaCl_2$ for HaloCaMP1a proteins or 100 mM for HaloCaMP1b) or EGTA (100 mM MOPS pH 7.2, 100 mM KCl, 100 mM EGTA) buffer was added. Purified HaloTag7 protein (1 eq, 126 µM in 50 mM Tris-HCl pH 7.4, 100 mM NaCl) was added to each well and incubated for 45 min. The fluorescence of **JF635-HTL** was measured before and after HaloTag addition using the TECAN plate reader and the $\Delta F/F_0$ values were calculated.

VI.4.9 Protein structure prediction

For 3D protein structure prediction, protein sequences were inputted into ColabFold version 1.5.2 (https://github.com/sokrypton/ColabFold) installed through EasyBuild (https://docs.easybuild.io/version-specific/supported-software/c/ColabFold/). Specifically, we used the colabfold_batch script with the following options (values that differ from the default are marked with a *): model-type = "alphafold2_ptm", stop-at-score = 100, num-recycle = None, recycle-early-stop-tolerance = None, num-ensemble = 1, num-seeds = 1, random-seed = 0, num-models = 5, recompile-padding = 10, host-url = "https ://api.colabfold.com", msa-mode = "mmseqs2_uniref_env", *amber = True, num-relax = 0, *templates = True, custom-template-path = None, rank = "auto", pair-mode = "unpaired_paired", sort-queries-by = "length", save-single-representations = False, save-pair-representations = False, use-dropout = False, max-seq = None, max-extra-seq = None, max-msa = None, disable-cluster-profile = False, zip = False, *use-gpu-relax = True, save-all = False, save-recycles = False, overwrite-existing-results = False, disable-unified-memory = False. For each sequence, five models were generated and ranked by their pIDDT scores. The predicted structures were visualised and analysed using PyMOL.

VI.4.10 Fluorescence colony screening

pET51b-HaloTag7-EGFP plasmid was transformed into T7 Express Competent *E. coli* cells (NEB). Colonies were grown overnight at 37°C on LB agar plates inoculated with ampicillin. The plates were imaged on a fluorescent gel imager (GE Typhoon FLA 9000) using EGFP and Cy5 filter wavelengths. 1 mL of **JF635-HTL** or **14-HTL** (1 µM in HEPES 10 mM, pH 7.4, CHAPS 0.1 $mg \cdot mL^{-1}$) was carefully added to the plate using a pipette and dispersed by tilting. After incubation for 1 h at room temperature, the plates were re-imaged.

VI.4.11 Photoacoustic colony screening

pET51b plasmids containing HaloTag7-EGFP and HaloCaMP1a-EGFP were transformed into BL21(DE3) *E. coli* cells (NEB). Colonies were grown overnight at 24°C on LB agar plates inoculated with ampicillin for 48 h.

For the HaloTag7-EGFP colonies, **14-HTL** was added to a final concentration of 5 µM (0.4% DMSO) in a SeaPrep Agarose solution (10 mM HEPES, pH 7.4) on top of the colonies and incubated overnight at 4°C.

Photoacoustic microscopy imaging was performed at 495 nm and 646 nm to record the signal from EGFP and **14-**HaloTag7 respectively.

For the HaloCaMP1a-EGFP colonies, the SeaPrep Agarose was prepared in 30 mM MOPS pH 7.2, 100 mM KCl, with addition of 50 $\mu g \cdot ml^{-1}$ polylysine, 50 $\mu g \cdot ml^{-1}$ ionomycin and **14-HTL** (at a final concentration of 5 µM). After 1 hour of gel formation, the high Ca^{2+}-buffer containing (30 mM MOPS pH 7.2, 100 mM KCl, 500 mM $CaCl_2$) was added on the SeaPrep layer and the plate was incubated overnight in the fridge at 4°C. The following day, the PA signal of the Ca^{2+}-bound state using the PAM set-up was measured at 495 nm and 646 nm for the EGFP and the acoustogenic dye signals, respectively. After the measurement, the buffer was aspirated and replaced with high EGTA buffer (30 mM MOPS pH 7.2, 100 mM KCl, 500 mM EGTA), left overnight in the fridge at 4°C and photoacoustic imaging performed on the set-up as before. A control plate of HaloCaMP1a was measured where no Ca^{2+} or EGTA buffers were added (30 mM MOPS pH 7.2, 100 mM KCl).

VI.4.12 *In silico* prediction of physicochemical properties

Molecules were drawn in ChemDraw v19.1 and inputted into the following software tools: the ADME calculator from Atomistica[229] and the molecular properties calculator from Molinspiration Cheminformatics,[230] for prediction of physicochemical properties. Atomistica predicts the number of H bond donors and acceptors, logP, molar refractivity and topological polar surface area (TPSA). Molinspiration Cheminformatics estimates number of H bond donors and acceptors, logP, TPSA, molecular volume and the number of rotatable bonds.

VI.4.13 logP determination

The shake-flask method was used to determine the partition coefficient, logP. To pre-saturate the two phases, a 1:1 mixture of n-octanol and PBS (pH 7.4) was shaken for 24 h, and then centrifuged (30 min, 2500 r.p.m) to separate the phases. **14-HTL** or **22-HTL** was added to an aliquot of the pre-saturated PBS phase to a final concentration of 5 µM (0.05% DMSO) and incubated with purified HaloTag7 protein (1.5 eq) for 3 h. The absorption spectrum was measured in the Cary spectrometer with A_1, the value at λ_{max}. **14-HTL** or **22-HTL** was added to a second aliquot of pre-saturated PBS which was mixed with an equal volume of pre-saturated n-octanol and shaken at 37°C for 24 h. Centrifugation (30 min, 2500 r.p.m) separated the two phases. The PBS layer was then incubated with purified HaloTag7 protein (1.5 eq) for 3 h and the absorption spectrum recorded, with A_2 the value at λ_{max}. LogP was calculated using the following equation:

$$\log P = \log_{10}\left(\frac{A_1 - A_2}{A_2}\right)$$

Equation 6: Determination of the partition coefficient. P: partition coefficient. A_1: absorbance of dye in pre-saturated PBS. A_2: absorbance of dye in the PBS layer after mixing with n-octanol.

VI.4.14 PAMPA-BBB assay

The *in vitro* parallel artificial membrane permeability assay was performed using the BioAssaySystems kit. Solutions of the test compounds were prepared in PBS. 5 µL of suspended porcine brain lipids were used to coat the well membrane. The acceptor plate was filled with 300 µL PBS (pH 7.2). 200 µL of test compounds were added to their respective donor wells and then laid over the acceptor plate. The plate was incubated at 37°C for 22 h.

The first tests used **JF635-HTL**, **JF549-HTL** and **14-HTL** at 100 or 250 µM concentration in PBS, with the acceptor solution of only PBS. After incubation, an aliquot of the acceptor solution (e.g. 7.5 µL) was incubated with purified HaloTag7 protein (1.5 eq) and absorption at λ_{max} subsequently measured using the Epoch 2 plate

reader. Only **JF549-HTL** was detectable. Increasing DMSO concentration (up to 40%) and/or Pluronic F127 (up to 20%) did not reach a detectable level for **JF635-HTL** or **14-HTL**. Consequently, for detection of all screened HaloTag ligands, the test compounds were prepared at 25 μM (2.5% DMSO) in PBS and the acceptor well contained purified HaloTag (15 μM, 1.5 eq). If the compound can fully permeabilize the membrane, the final dye concentration will be 10 μM. The resulting acceptor solution could be directly transferred to a 96-well plate for absorbance measurements.

A reference solution was prepared by incubation of 10 μM HaloTag ligand with HaloTag7 protein (15 μM, 1.5 eq) for at least 3 h and the absorption at λ_{max} was recorded. For quantification, the absorption of the blank control (only PBS) was subtracted from the absorption for the reference and experimental solutions. The relative permeability was determined by the ratio of absorption of the experimental solutions with the corresponding reference solution.

VI.5 Materials and Methods for Chapter IV

VI.5.1 Photoacoustic Tomography

A custom-built photoacoustic tomography set-up based on a transparent Fabry-Pérot ultrasound sensor was used for PA imaging. Briefly, the set-up is based on a 100 Hz pulsed tunable laser source (EVO I OPO, Innolas Laser GmbH) that excites the photoacoustic signal in a 420-709 nm range. The ultrasound waves are then detected by a home-made Fabry-Pérot sensor[237] interrogated with a 1520-1620 nm CW laser (Venturi TLB-8800, Newport). The Fabry-Pérot sensor was raster-scanned with a 15x15 mm^2 FOV and 50 µm step resulting in a 301x301 ultrasound sampling points. The photoacoustic image was then reconstructed with either a back-projection or time-reversal algorithm from a k-Wave toolbox for MATLAB.[263]

VI.5.2 Phantoms

Tubes. For the tissue-mimicking phantoms, protein solutions at 50 µM were loaded in PE tubing (ID 0.86 mm, OD 1.27 mm, Hugo Sachs Elektronik), secured in a 3D-printed holder and fixed in a petri dish. They were submerged in coupling media of water or milk-water mixtures (10% or 60% whole cow's milk). The photoacoustic data was acquired at respective excitation wavelengths and the resulting image was formed with a k-Wave reconstruction algorithm and then exported to ImageJ, where the maximum intensity projections of a 3D image were extracted for xy, xz and yz views. The maximum intensity projections were then analysed with a line profile to quantify the values from each sample. The detection limit in the scattering phantoms for **14**-HaloTag7 was determined by linear fitting the signal intensities of **14**-HaloTag7 in the 5-20 µM range, and intersection with the noise floor. The noise floor was estimated as a mean of the background plus standard deviation of the background in 1 mm^3 (Figure 58d).

Three different tube materials were initially tested: PE (ID 0.86 mm, OD 1.27 mm, Hugo Sachs Elektronik), Teflon (ID 1.07 mm, OD 1.47 mm, Alpha Wire), silicon (ID 0.4 mm, OD 1.0 mm, VWR). The tubes were filled with methylene blue (0.27 mM in H_2O), fixed in a 3D printed holder and imaged on the PAT set-up with H_2O coupling media.

Gel wax. Gel wax was melted in a beaker heated to ~100 °C in an oil bath, transferred to a round-bottomed flask and degassed on a Schlenk line whilst heating ~130 °C. Glass capillaries were fixed across a Petri dish and the degassed gel wax was slowly poured until the capillaries were submerged to a depth of 1-2 mm. The gel was allowed to set and then the capillaries carefully removed. Lubricants were used to coat the capillaries but did not prevent rupture of the channel on capillary removal.

VI.5.3 Fluorescence microscopy in cultured cells

Widefield imaging was performed at the Advanced Light Microscopy Facility (EMBL Heidelberg), on a Nikon Ti-E microscope equipped with a Spectra X light engine (Lumencore) with a 20x objective (CFI P-Apo 20x Lambda/0.75/1,00) and imaged onto a scientific complementary metal–oxide–semiconductor camera (pco.edge 4.2 CL). A quad bandpass filter cube was used to image EGFP (excitation 485/20, emission 525/50), **JF549-HTL** (excitation 531/40, emission 593/40) and far-red ligands (excitation 650/13, emission 692/40).

VI.5.4 Cell culture

U2OS cells stably expressing HaloTag7–EGFP fusion protein (gift from the Johnsson group, MPI for Medical Research, Heidelberg) were cultured in Dulbecco's modified Eagle medium [(DMEM, high glucose (4.5 $g \cdot L^{-1}$)

with phenol red (Gibco), supplemented with 10% v/v fetal bovine serum (ThermoFisher Scientific), penicillin (100 units·mL^{-1}), streptomycin (100 µg·mL^{-1}; Gibco), hygromycin B (100 µg·mL^{-1}; Sigma Aldrich), 2 mM L-Glutamine (Gibco) and 1 mM sodium pyruvate (Gibco)], and maintained at 37°C in a humidified 5% v/v CO_2 environment. All labelled cells were imaged 18–24 h post-plating.

U2OS HaloTag7-EGFP cells were seeded in a chambered coverslip with 8 individual wells (Ibidi, 80806). The next day cells were incubated with 1 µM HaloTag ligand (**10-HTL**, **14-HTL**, **15-HTL**, **22-HTL** or **JF549-HTL,** the positive control) for 2 h at 37°C, washed three times with imaging buffer [DMEM low glucose (1 g·L^{-1}) without phenol red, supplemented with 10% v/v fetal bovine serum, 2 mM L-Glutamine (Gibco) and 1 mM sodium pyruvate]. Following this, 0.5 µM of **JF549-HTL** was added to the medium, and the cells were incubated for 30 min more at 37°C. The cells were subsequently washed twice with PBS, incubated for 15 min at 37°C and then washed three times with imaging buffer before widefield fluorescence imaging as described above. For quantification, ROIs were manually drawn around all visible cells in the field of view in the EGFP channel. The fluorescence intensity for each ROI was measured in the different channels. The average background intensity was subtracted from the mean intensity value in each channel, and the fluorescence ratios of 549/EGFP and far-red/EGFP were calculated.

VI.5.5 Spheroids

U2OS HaloTag7-EGFP cells were cultured in Dulbecco's modified Eagle medium [(DMEM, high glucose (4.5 g·L^{-1}) with phenol red (Gibco), supplemented with 10% v/v fetal bovine serum (ThermoFisher Scientific), penicillin (100 units·mL^{-1}), streptomycin (100 µg·mL^{-1}; Gibco), hygromycin B (100 µg·mL^{-1}; Sigma Aldrich), 2 mM L-Glutamine (Gibco) and 1 mM sodium pyruvate (Gibco)], and maintained at 37°C in a humidified 5% v/v CO_2 environment. These cells were seeded into faCellitate BIOFLOAT^{TM} U-bottomed 96-well plates at densities in the range 3125–50000 viable cells per well. The spheroid cultures were monitored using a light microscope. For imaging, transfer of the spheroids was attempted to 96-well glass-bottomed plates (Zell-Kontakt GmbH).

Alternatively, 100 µL faCellitate BIOFLOAT^{TM} FLEX coating solution was added to flat-bottomed 8-well chambers (Ibidi, 80806). After 3 minutes, the liquid was aspirated and then the plate allowed to dry for 30 minutes before seeding with U2OS HaloTag7-EGFP cells as above.

In a final coating strategy, 96-well U-bottomed plates or Ibidi chambers were coated with 200 µL hyaluronic acid solution (Sigma Aldrich) at 1, 2.5 or 5 mg·mL^{-1}. The plates were dried at 37°C for 3-4 days before seeding with U2OS HaloTag7-EGFP cells as above.[264]

VI.5.6 Preparation of bacterial pellets

pET51b plasmids containing HaloTag7 were transformed into T7 Express Competent *E. coli* cells (NEB) via electroporation. Cell colonies were grown in LB medium containing ampicillin (100 µg·mL^{-1}) overnight at 37°C at 200 r.p.m. The starting culture was diluted 100-fold in LB containing ampicillin (100 µg·mL^{-1}) and incubated at 37°C, 200 r.p.m until the OD reached 0.4-0.6. Protein expression was then induced by IPTG (1 mM) and the culture incubated at 18°C, 200 r.p.m for 14-20 h. After centrifugation (15 min at 3500 r.p.m), pellets (from 5-10 mL of culture) were resuspended in 1 mL PBS. **14-HTL** (1 mM in DMSO) was added to give a final concentration of 20 µM and the suspension incubated for at least 3 h, shaking at 850 r.p.m at room temperature. A second round of centrifugation (5 min at 7500 r.p.m) and subsequent removal of supernatant provided the labelled pellets.

For quantification of saturation, the volume of the suspension after labelling and of the supernatant was measured. The supernatant was diluted into 1 mL PBS and passed through a 0.2 µM filter to remove any bacteria. The filtered supernatant was incubated with excess HaloTag protein (20 µM) at room temperature for 4 h before the absorption spectrum was recorded. From the extinction coefficient of HaloTag-bound **14-HTL**, the concentration of dye in the supernatant could be determined, and using the following equation, the concentration of dye inside the pelleted cells could be deduced:

$$c_{tot} \cdot v_{tot} = c_s \cdot v_s + c_p \cdot v_p$$

Equation 7: C_{tot}: total **14-HTL** concentration; V_{tot}: total volume; C_S: **14-HTL** concentration in the supernatant; V_S: volume of supernatant; C_P: **14-HTL** concentration in the pellet; V_P: volume of pellet.

VI.5.7 Methods for mounting the pellets onto the Fabry-Pérot

Agar wells. A 2% agar gel was allowed to set on a 3D printed mould with 3 mm diameter cylinders to form wells. After removal of the mould, the gel was stuck to the bottom of a Petri dish, mounted onto the PAT set-up and submerged in H_2O for acoustic coupling. The pellets were carefully pipetted into the wells.

Embedding in low melting point agarose. 2% low melting point agarose gel (UltraPure™ Invitrogen) was allowed to slowly set in a Petri dish and labelled pellets (at room temperature, frozen or freeze-dried) were placed on top just before the gel completely set. A second layer of agarose was applied, after cooling to close to the setting point, onto the embedded pellet. The Petri dish was mounted onto the PAT set-up and submerged in water for acoustic coupling.

Tubes. The pellets were taken up with a 1 mL syringe and needle, air bubbles removed, and slowly injected into the PE tubing (ID 0.86 mm, OD 1.27 mm, Hugo Sachs Elektronik). The tubes were fixed by a 3D printed tube holder, mounted onto the PAT set-up and submerged in H_2O for acoustic coupling. For comparison the pellet containing tube was positioned over a tube with methylene blue (0.27 mM in H_2O).

Direct imaging. The pellets were taken up with a 1 mL syringe and needle, air bubbles removed, and slowly applied onto the surface of the Fabry-Pérot interferometer.

VI.5.8 General mouse *in vivo* methods

This work followed the European Communities Council Directive (2010/63/EU) to minimise animal pain and discomfort. All procedures described in this paper were approved by EMBL's committee for animal welfare and institutional animal care and use, under license 22-004_HD_RP. Experiments were performed on 12-week-old male C57BL/6 mice from the EMBL Heidelberg core colonies or Charles River Laboratories. During the course of the study, mice were housed in groups of 1–5 in makrolon type 2L or 3H cages, in ventilated racks at room temperature and 50% humidity while kept in a 12/12 h light/dark cycle. Food and water were available *ad libitum*.
For neuronal-specific expression of HaloTag7-EGFP, AAV1-synapsin1-NES-HaloTag7-EGFP (6.0×10^{13} viral genomes per mL) was used. For stereotaxic injection of AAVs, mice were anesthetised with isoflurane (Baxter) vapour mixed with O_2 (5% for induction and 1–1.5% for maintenance). Eye ointment was applied (Bepanthen, Bayer) and Xyolcain 1% (AstraZeneca) was subcutaneously injected in the scalp for preincisional local anaesthesia. During anaesthesia, body temperature was monitored and maintained with an RWD ThermoStar system. A small incision was produced in the scalp to expose the dorsal cranium and, using the bregma as a landmark, a small burr hole centred at ML 1.3 mm, AP -2 mm was made with a dental drill (Microtorque, Harvard Apparatus). AAV injections were performed with an RWD 68803 stereotaxic fame equipped with a 68055 mouse adapter in the hippocampus and thalamus on the left hemisphere at depths -1.2 mm and 3 mm

respectively, using a syringe at a rate of ~4 µL·h^{-1}. Approximatively 300 nL were injected per spot. Mice were sutured, meloxican administered subcutaneously (Metacan, Boehringer Ingelheim) dosed at 1 µg·g^{-1} for pain relief and single housed after the viral injection surgery and had a recovery period of at least 4 weeks before further experiments for HaloTag7-EGFP expression.

VI.5.9 Fluorescence Microscopy and Photoacoustic Imaging of Brain Slices

The animals were sacrificed with CO_2 and the brains were transcardially perfused with ~35 mL of PBS without Ca^{2+} and Mg^{2+} containing 4% paraformaldehyde to remove blood from tissues. The brains were fixated by immersion in PBS with 4% paraformaldehyde for 72 h at 4ºC. The fixed brain was washed three times with PBS and then sliced (200-300 µm thick) with a vibratome (Leica VT1200) in PBS without Ca^{2+} and Mg^{2+}. The slices were imaged via widefield fluorescence imaging (as described above) before and after dye labelling with **14-HTL** and **15-HTL** with the 4x objective in transmission, GFP and far-red channels.

Slices labelled with **10-HTL** or **20-HTL** were imaged via confocal microscopy using a Leica Stellaris 8 microscope equipped with a Spectra X light engine (Lumencore) with a 10x objective (HC PL FLUOTAR10x / NA 0.30 Dry) and imaged onto a 4x HyD S detector. A white light laser was used to image **10-HTL** (excitation 690 nm, emission 710-800 nm) and **20-HTL** (excitation 660 nm, emission 680-800 nm). The whole brain slices were imaged via tile scanning and then a z-stack image of a tile focused on the hippocampus (57 µm intervals).

For dye labelling of brain slices, the slices were soaked in a bath of PBS containing 10 µM of **10-HTL**, **14-HTL, 15-HTL** or **20-HTL** for 2 h at 37°C, rotating gently on a tilt shaker to ensure homogenous labelling (n = 16 slices for **14-HTL**, n = 3 slices for **10-HTL**, **15-HTL,** and **20-HTL**). The slices were washed with PBS without Ca^{2+} and Mg^{2+} (3 x 20 min) at 37°C before photoacoustic imaging.

For PAT imaging the brain slices were fixed in low-gelling temperature agarose (A9045-10G, Sigma) in a petri dish. The sample was submerged in PBS without Ca^{2+} and Mg^{2+} for acoustic coupling and placed in the PAT set-up for imaging. The widefield fluorescence image of the agarose-fixed slice was taken in the GFP channel for co-registering the PAT image with the original fluorescence images.

For quantification of fluorescence intensity from brain slices, ROIs were manually drawn around the hippocampus or thalamus regions with visible EGFP expression (one ROI per slice and brain region, see Figure 62d). The fluorescence intensity for each ROI was measured in the different channels. The average background intensity was subtracted from the mean intensity value in each channel, and the fluorescence ratios of far-red/EGFP were calculated. Each data point corresponds to one brain slice, measured as one ROI.

VI.5.10 *In vivo* dye delivery

For intracerebroventricular injections, a 15 mM solution of **14-HTL** or **15-HTL** in DMSO was prepared. 2 µL of this solution was added to 2 µL of Pluronic F127 in DMSO (20% w/w), and 16 µL of sterile saline was added to give a final concentration of 1.5 mM of dye ligand. 2 µL (3 nmol) of this dye ligand solution was administered via intracerebroventricular injection (n = 3 mice for **14-HTL** and n = 2 mice for **15-HTL**).

For tail vein injections, a 3 mM solution of **15-HTL** in DMSO was prepared. 150 µL was added to 50 µL of Pluronic F127 in DMSO (20% w/w), and 300 µL of sterile saline was added to give a final concentration of 1.5 mM of dye ligand. 100 µL (150 nmol) of this dye ligand solution was administered via tail vein injection (n = 2 mice for **15-HTL**).

24 hours after dye injection, animals were sacrificed with CO_2. The **15-HTL** injected brains were transcardially perfused with ~35 mL of PBS without Ca^{2+} and Mg^{2+} containing 4% paraformaldehyde. They were directly sliced 100 µm thick with a vibratome (Leica VT1200) in PBS without Ca^{2+} and Mg^{2+}. The slices were transferred

to a slide and a cover slip added before fluorescence microscopy. The slices were imaged via widefield fluorescence imaging (as described above) with the 10x objective in transmission, GFP and far-red channels.

After sacrificing, the **14-HTL** labelled brains were transcardially perfused with ~35 mL of PBS without Ca^{2+} and Mg^{2+} to remove blood from tissues. The perfused brain was then acutely dissected fresh in ice cold PBS without Ca^{2+} and Mg^{2+}. Immediately after dissection, the **14-HTL** injected brains were imaged in our PAT set-up by gluing to the bottom of a petri dish and submersion in PBS without Ca^{2+} and Mg^{2+} for acoustic coupling. After imaging, the brain was retrieved from the PAT set-up, fixated by immersion in PBS with 4% paraformaldehyde for 72 h at 4ºC, and directly sliced 300 µm thick with a vibratome (Leica VT1200) in PBS without Ca^{2+} and Mg^{2+}. Based on the mouse brain atlas,[15] and using morphological landmarks, including the splitting of coronal hemispheres, each slice was mapped to their corresponding coronal section in the PAT volume.

For quantification of fluorescence intensity from individual neurons within the brain slices, ROIs were manually drawn around the cytosol of neurons with visible EGFP expression. The fluorescence intensity for each ROI was measured in the different channels. The average background intensity was subtracted from the mean intensity value in each channel, and the fluorescence ratios of far-red/EGFP were calculated.

Chapter VII: References

(1) Bressler, S. L.; Menon, V.; Large-scale brain networks in cognition: emerging methods and principles, *Trends Cogn Sci,* **2010**, *14* (6), 277-290.
(2) Dasgupta, D.; Warner, T. P. A.; Erskine, A.; Schaefer, A. T.; Coupling of Mouse Olfactory Bulb Projection Neurons to Fluctuating Odor Pulses, *J Neurosci,* **2022**, *42* (21), 4278-4296.
(3) Blackshaw, S.; Scholpp, S.; Placzek, M.; Ingraham, H.; Simerly, R.; Shimogori, T.; Molecular pathways controlling development of thalamus and hypothalamus: from neural specification to circuit formation, *J Neurosci,* **2010**, *30* (45), 14925-14930.
(4) Zemla, R.; Basu, J.; Hippocampal function in rodents, *Curr Opin Neurobiol,* **2017**, *43*, 187-197.
(5) McKay, B. E.; Turner, R. W.; Physiological and morphological development of the rat cerebellar Purkinje cell, *J Physiol,* **2005**, *567* (3), 829-850.
(6) Angeles Fernandez-Gil, M.; Palacios-Bote, R.; Leo-Barahona, M.; Mora-Encinas, J. P.; Anatomy of the brainstem: a gaze into the stem of life, *Semin Ultrasound CT MR,* **2010**, *31* (3), 196-219.
(7) Tau, G. Z.; Peterson, B. S.; Normal development of brain circuits, *Neuropsychopharmacology,* **2010**, *35* (1), 147-168.
(8) Deo, C.; Lavis, L. D.; Synthetic and genetically encoded fluorescent neural activity indicators, *Curr Opin Neurobiol,* **2018**, *50*, 101-108.
(9) Berridge, M. J.; Lipp, P.; Bootman, M. D.; The Versatility and Universality of Calcium Signalling *Nature Reviews Molecular Cell Biology,* **2000**, *1*, 11-21.
(10) Grienberger, C.; Konnerth, A.; Imaging calcium in neurons, *Neuron,* **2012**, *73* (5), 862-885.
(11) Kakiuchi, S.; Yamazaki, R.; Calcium dependent phosphodiesterase activity and its activating factor (PAF) from brain: Studies on cyclic 3′,5′-nucleotide phosphodiesterase (III), *Biochem Biophys Res Commun,* **1970**, *41* (5), 1104-1110.
(12) Cheung, W. Y.; Cyclic 3',5'-Nucleotide Phosphodiesterase, *Biochem Biophys Res Commun,* **1970**, *38* (3), 533-538.
(13) Grynkiewicz, G.; Poenie, M.; Tsien, R. Y.; A new generation of Ca^{2+} indicators with greatly improved fluorescence properties, *Journal of Biological Chemistry,* **1985**, *260* (6), 3440-3450.
(14) Sabatini, B. L.; Tian, L.; Imaging Neurotransmitter and Neuromodulator Dynamics In Vivo with Genetically Encoded Indicators, *Neuron,* **2020**, *108* (1), 17-32.
(15) Minatohara, K.; Akiyoshi, M.; Okuno, H.; Role of Immediate-Early Genes in Synaptic Plasticity and Neuronal Ensembles Underlying the Memory Trace, *Front Mol Neurosci,* **2015**, *8*, 78.
(16) Phillips, A. A.; Chan, F. H.; Zheng, M. M.; Krassioukov, A. V.; Ainslie, P. N.; Neurovascular coupling in humans: Physiology, methodological advances and clinical implications, *J Cereb Blood Flow Metab,* **2016**, *36* (4), 647-664.
(17) Glover, G. H.; Overview of functional magnetic resonance imaging, *Neurosurg Clin N Am,* **2011**, *22* (2), 133-139.
(18) Bean, B. P.; The action potential in mammalian central neurons, *Nat Rev Neurosci,* **2007**, *8* (6), 451-465.
(19) Engelman, H. S.; MacDermott, A. B.; Presynaptic ionotropic receptors and control of transmitter release, *Nat Rev Neurosci,* **2004**, *5* (2), 135-145.
(20) Ametamey, S. M.; Honer, M.; Schubiger, P. A.; Molecular Imaging with PET, *Chem Rev,* **2008**, *108*, 1501-1516.
(21) Bell, A. G.; On the Production and Reproduction of Sound by Light, *American Journal of Science,* **1880**, *s3-20* (118), 305-324.
(22) Dean-Ben, X. L.; Gottschalk, S.; Mc Larney, B.; Shoham, S.; Razansky, D.; Advanced optoacoustic methods for multiscale imaging of in vivo dynamics, *Chem Soc Rev,* **2017**, *46* (8), 2158-2198.
(23) Weber, J.; Beard, P. C.; Bohndiek, S. E.; Contrast agents for molecular photoacoustic imaging, *Nat Methods,* **2016**, *13* (8), 639-650.
(24) Attia, A. B. E.; Balasundaram, G.; Moothanchery, M.; Dinish, U. S.; Bi, R.; Ntziachristos, V.; Olivo, M.; A review of clinical photoacoustic imaging: Current and future trends, *Photoacoustics,* **2019**, *16*, 100144.
(25) Wang, L. V.; Yao, J.; A practical guide to photoacoustic tomography in the life sciences, *Nat Methods,* **2016**, *13* (8), 627-638.
(26) Hu, S.; Maslov, K.; Tsytsarev, V.; Wang, L. V.; Functional transcranial brain imaging by optical-resolution photoacoustic microscopy, *J Biomed Opt,* **2009**, *14* (4), 040503.
(27) Kim, J.; Kim, J. Y.; Jeon, S.; Baik, J. W.; Cho, S. H.; Kim, C.; Super-resolution localization photoacoustic microscopy using intrinsic red blood cells as contrast absorbers, *Light Sci Appl,* **2019**, *8*, 103.
(28) Park, S.; Lee, C.; Kim, J.; Kim, C.; Acoustic resolution photoacoustic microscopy, *Biomedical Engineering Letters,* **2014**, *4* (3), 213-222.
(29) Stein, E. W.; Maslov, K.; Wang, L. V.; Noninvasive, in vivo imaging of the mouse brain using photoacoustic microscopy, *J Appl Phys,* **2009**, *105* (10), 102027.
(30) Xia, J.; Yao, J.; Wang, L. V.; Photoacoustic tomography: principles and advances, *Electromagn Waves (Camb),* **2014**, *147*, 1-22.

(31) Gottschalk, S.; Degtyaruk, O.; Mc Larney, B.; Rebling, J.; Hutter, M. A.; Dean-Ben, X. L.; Shoham, S.; Razansky, D.; Rapid volumetric optoacoustic imaging of neural dynamics across the mouse brain, *Nat Biomed Eng,* **2019**, *3* (5), 392-401.
(32) Razansky, D.; Distel, M.; Vinegoni, C.; Ma, R.; Perrimon, N.; Köster, R. W.; Ntziachristos, V.; Multispectral opto-acoustic tomography of deep-seated fluorescent proteins in vivo, *Nature Photonics,* **2009**, *3* (7), 412-417.
(33) Steinberg, I.; Huland, D. M.; Vermesh, O.; Frostig, H. E.; Tummers, W. S.; Gambhir, S. S.; Photoacoustic clinical imaging, *Photoacoustics,* **2019**, *14*, 77-98.
(34) Andreev, V. G.; Karabutov, A. A.; Oraevsky, A. A.; Detection of Ultrawide-Band Ultrasound Pulses in Optoacoustic Tomography, *IEEE Trans Ultrason Ferroelectr Freq Control,* **2003**, *50* (10), 1383-1390.
(35) Ntziachristos, V.; Going deeper than microscopy: the optical imaging frontier in biology, *Nat Methods,* **2010**, *7* (8), 603-614.
(36) Stuker, F.; Ripoll, J.; Rudin, M.; Fluorescence molecular tomography: principles and potential for pharmaceutical research, *Pharmaceutics,* **2011**, *3* (2), 229-274.
(37) Deffieux, T.; Demene, C.; Pernot, M.; Tanter, M.; Functional ultrasound neuroimaging: a review of the preclinical and clinical state of the art, *Curr Opin Neurobiol,* **2018**, *50*, 128-135.
(38) Jin, Z.; Lakshmanan, A.; Zhang, R.; Tran, T. A.; Rabut, C.; Dutka, P.; Duan, M.; Hurt, R. C.; Malounda, D.; Yao, Y.; et al.; Ultrasonic reporters of calcium for deep tissue imaging of cellular signals, *bioRxiv,* **2023**.
(39) Razansky, D.; Klohs, J.; Ni, R.; Multi-scale optoacoustic molecular imaging of brain diseases, *Eur J Nucl Med Mol Imaging,* **2021**, *48*, 4152–4170.
(40) Brunker, J.; Yao, J.; Laufer, J.; Bohndiek, S. E.; Photoacoustic imaging using genetically encoded reporters: a review, *J Biomed Opt,* **2017**, *22* (7), 070901.
(41) Yao, J.; Wang, L.; Yang, J. M.; Maslov, K. I.; Wong, T. T.; Li, L.; Huang, C. H.; Zou, J.; Wang, L. V.; High-speed label-free functional photoacoustic microscopy of mouse brain in action, *Nat Methods,* **2015**, *12* (5), 407-410.
(42) Li, L.; Xia, J.; Li, G.; Garcia-Uribe, A.; Sheng, Q.; Anastasio, M. A.; Wang, L. V.; Label-free photoacoustic tomography of whole mouse brain structures ex vivo, *Neurophotonics,* **2016**, *3* (3), 035001.
(43) Olefir, I.; Ghazaryan, A.; Yang, H.; Malekzadeh-Najafabadi, J.; Glasl, S.; Symvoulidis, P.; O'Leary, V. B.; Sergiadis, G.; Ntziachristos, V.; Ovsepian, S. V.; Spatial and Spectral Mapping and Decomposition of Neural Dynamics and Organization of the Mouse Brain with Multispectral Optoacoustic Tomography, *Cell Rep,* **2019**, *26* (10), 2833-2846
(44) Shi, J.; Wong, T. T. W.; He, Y.; Li, L.; Zhang, R.; Yung, C. S.; Hwang, J.; Maslov, K.; Wang, L. V.; High-resolution, high-contrast mid-infrared imaging of fresh biological samples with ultraviolet-localized photoacoustic microscopy, *Nat Photonics,* **2019**, *13*, 609-615.
(45) Burton, N. C.; Patel, M.; Morscher, S.; Driessen, W. H.; Claussen, J.; Beziere, N.; Jetzfellner, T.; Taruttis, A.; Razansky, D.; Bednar, B.; et al.; Multispectral opto-acoustic tomography (MSOT) of the brain and glioblastoma characterization, *Neuroimage,* **2013**, *65*, 522-528.
(46) Li, L.; Zhu, L.; Ma, C.; Lin, L.; Yao, J.; Wang, L.; Maslov, K.; Zhang, R.; Chen, W.; Shi, J.; et al.; Single-impulse Panoramic Photoacoustic Computed Tomography of Small-animal Whole-body Dynamics at High Spatiotemporal Resolution, *Nat Biomed Eng,* **2017**, *1* (5), 0071.
(47) Wang, X.; Pang, Y.; Ku, G.; Xie, X.; Stoica, G.; Wang, L. V.; Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain, *Nature Biotechnology,* **2003**, *21* (7), 803-806.
(48) Zhang, P.; Li, L.; Lin, L.; Hu, P.; Shi, J.; He, Y.; Zhu, L.; Zhou, Y.; Wang, L. V.; High-resolution deep functional imaging of the whole mouse brain by photoacoustic computed tomography in vivo, *J Biophotonics,* **2018**, *11* (1), e201700024.
(49) Chen, Z.; Zhou, Q.; Dean-Ben, X. L.; Gezginer, I.; Ni, R.; Reiss, M.; Shoham, S.; Razansky, D.; Multimodal Noninvasive Functional Neurophotonic Imaging of Murine Brain-Wide Sensory Responses, *Adv Sci (Weinh),* **2022**, *9* (24), e2105588.
(50) Borg, R. E.; Rochford, J.; Molecular Photoacoustic Contrast Agents: Design Principles & Applications, *Photochem Photobiol,* **2018**, *94* (6), 1175-1209.
(51) Kim, G.; Huang, S. W.; Day, K. C.; O'Donnell, M.; Agayan, R. R.; Day, M. A.; Kopelman, R.; Ashkenazi, S.; Indocyanine-green-embedded PEBBLEs as a contrast agent for photoacoustic imaging, *J Biomed Opt,* **2007**, *12* (4), 044020.
(52) Wang, X.; Ku, G.; Wegiel, M. A.; Bornhop, D. J.; Stoica, G.; Wang, L. V.; Noninvasive photoacoustic angiography of animal brains in vivo with near-infrared light and an optical contrast agent, *Opt Lett,* **2004**, *29* (7), 730-732.
(53) Laufer, J.; Zhang, E.; Beard, P.; Evaluation of Absorbing Chromophores Used in Tissue Phantoms for Quantitative Photoacoustic Spectroscopy and Imaging, *IEEE J. Quantum Electron* **2010**, *16* (3), 600-607.

(54) Liu, C.; Chen, J.; Zhu, Y.; Gong, X.; Zheng, R.; Chen, N.; Chen, D.; Yan, H.; Zhang, P.; Zheng, H.; et al.; Highly Sensitive MoS_2-Indocyanine Green Hybrid for Photoacoustic Imaging of Orthotopic Brain Glioma at Deep Site, *Nanomicro Lett,* **2018**, *10* (3), 48.
(55) Zhu, M.; Sheng, Z.; Jia, Y.; Hu, D.; Liu, X.; Xia, X.; Liu, C.; Wang, P.; Wang, X.; Zheng, H.; Indocyanine Green-holo-Transferrin Nanoassemblies for Tumor-Targeted Dual-Modal Imaging and Photothermal Therapy of Glioma, *ACS Appl Mater Interfaces,* **2017**, *9* (45), 39249-39258.
(56) Gardner, S. H.; Brady, C. J.; Keeton, C.; Yadav, A. K.; Mallojjala, S. C.; Lucero, M. Y.; Su, S.; Yu, Z.; Hirschi, J. S.; Mirica, L. M.; et al.; A General Approach to Convert Hemicyanine Dyes into Highly Optimized Photoacoustic Scaffolds for Analyte Sensing, *Angew Chem Int Ed Engl,* **2021**, *60* (34), 18860-18866.
(57) Zhang, S.; Chen, H.; Wang, L.; Qin, X.; Jiang, B. P.; Ji, S. C.; Shen, X. C.; Liang, H.; A General Approach to Design Dual Ratiometric Fluorescent and Photoacoustic Probes for Quantitatively Visualizing Tumor Hypoxia Levels In Vivo, *Angew Chem Int Ed Engl,* **2022**, *61* (7), e202107076.
(58) Huang, W. J.; Zhang, X.; Chen, W. W.; Role of oxidative stress in Alzheimer's disease, *Biomed Rep,* **2016**, *4* (5), 519-522.
(59) Zhang, H. K.; Yan, P.; Kang, J.; Abou, D. S.; Le, H. N.; Jha, A. K.; Thorek, D. L.; Kang, J. U.; Rahmim, A.; Wong, D. F.; et al.; Listening to membrane potential: photoacoustic voltage-sensitive dye recording, *J Biomed Opt,* **2017**, *22* (4), 45006.
(60) Kang, J.; Zhang, H. K.; Kadam, S. D.; Fedorko, J.; Valentine, H.; Malla, A. P.; Yan, P.; Harraz, M. M.; Kang, J. U.; Rahmim, A.; et al.; Transcranial Recording of Electrophysiological Neural Activity in the Rodent Brain in vivo Using Functional Photoacoustic Imaging of Near-Infrared Voltage-Sensitive Dye, *Front Neurosci,* **2019**, *13*, 579.
(61) Rao, B.; Zhang, R.; Li, L.; Shao, J. Y.; Wang, L. V.; Photoacoustic imaging of voltage responses beyond the optical diffusion limit, *Sci Rep,* **2017**, *7* (1), 2560.
(62) Mishra, A.; Jiang, Y.; Roberts, S.; Ntziachristos, V.; Westmeyer, G. G.; Near-Infrared Photoacoustic Imaging Probe Responsive to Calcium, *Anal Chem,* **2016**, *88* (22), 10785–10789.
(63) Roberts, S.; Seeger, M.; Jiang, Y.; Mishra, A.; Sigmund, F.; Stelzl, A.; Lauri, A.; Symvoulidis, P.; Rolbieski, H.; Preller, M.; et al.; Calcium Sensor for Photoacoustic Imaging, *J Am Chem Soc,* **2018**, *140* (8), 2718-2721.
(64) Philip, R. P., A.; Bäumler, W.; Szeimies, R. M.; Abels, C.; Absorption and fluorescence spectroscopic investigation of indocyanine green, *Journal of Photochemistry and Photobiology A: Chemistry,* **1996**, *96* (1-3), 137-148.
(65) Li, H.; Zhang, P.; Smaga, L. P.; Hoffman, R. A.; Chan, J.; Photoacoustic Probes for Ratiometric Imaging of Copper(II), *J Am Chem Soc,* **2015**, *137* (50), 15628-15631.
(66) Jiang, Z.; Zhang, C.; Wang, X.; Ling, Z.; Chen, Y.; Guo, Z.; Liu, Z.; A Small-Molecule Ratiometric Photoacoustic Probe for the High-Spatiotemporal-Resolution Imaging of Copper(II) Dynamics in the Mouse Brain, *Angew Chem Int Ed Engl,* **2024**, *63* (13), e202318340.
(67) Cui, Y.; Wang, X.; Jiang, Z.; Zhang, C.; Liang, Z.; Chen, Y.; Liu, Z.; Guo, Z.; A Photoacoustic Probe with Blood-Brain Barrier Crossing Ability for Imaging Oxidative Stress Dynamics in the Mouse Brain, *Angew Chem Int Ed Engl,* **2023**, *62* (9), e202214505.
(68) Jiang, Z.; Liang, Z.; Cui, Y.; Zhang, C.; Wang, J.; Wang, H.; Wang, T.; Chen, Y.; He, W.; Liu, Z.; et al.; Blood-Brain Barrier Permeable Photoacoustic Probe for High-Resolution Imaging of Nitric Oxide in the Living Mouse Brain, *J Am Chem Soc,* **2023**, *145* (14), 7952-7961.
(69) Wang, S.; Sheng, Z.; Yang, Z.; Hu, D.; Long, X.; Feng, G.; Liu, Y.; Yuan, Z.; Zhang, J.; Zheng, H.; et al.; Activatable Small-Molecule Photoacoustic Probes that Cross the Blood-Brain Barrier for Visualization of Copper(II) in Mice with Alzheimer's Disease, *Angew Chem Int Ed Engl,* **2019**, *58* (36), 12415-12419.
(70) Tsien, R. Y.; New calcium indicators and buffers with high selectivity against magnesium and protons: design, synthesis, and properties of prototype structures, *Biochemistry,* **1980**, *19* (11), 2396-2404.
(71) Minta, A.; Kao, J. P. Y.; Tsien, R. Y.; Fluorescent indicators for cytosolic calcium based on rhodamine and fluorescein chromophores, *Journal of Biological Chemistry,* **1989**, *264* (14), 8171-8178.
(72) Brondsted, F.; Stains, C.; Xanthene-Based Dyes for Photoacoustic Imaging and Their Use as Analyte-Responsive Probes, *Chemistry,* **2024**, *30* (37), e202400598.
(73) Boguta, A., Wrobel, D.; Fluorescein and Phenolphthalein—Correlation of Fluorescence and Photoelectric Properties, *Journal of Fluorescence,* **2001**, *11* (2), 129-137.
(74) Daly, H. C.; Matikonda, S. S.; Steffens, H. C.; Ruehle, B.; Resch-Genger, U.; Ivanic, J.; Schnermann, M. J.; Ketone Incorporation Extends the Emission Properties of the Xanthene Scaffold Beyond 1000 nm, *Photochem Photobiol,* **2021**, *98* (2), 325-333.
(75) Zhou, X.; Fang, Y.; Wimalasiri, V.; Stains, C. I.; Miller, E. W.; A long-wavelength xanthene dye for photoacoustic imaging, *Chem Commun (Camb),* **2022**, *58*, 11941-11944.
(76) Rathnamalala, C. S. L.; Pino, N. W.; Herring, B. S.; Hooper, M.; Gwaltney, S. R.; Chan, J.; Scott, C. N.; Thienylpiperidine Donor NIR Xanthene-Based Dye for Photoacoustic Imaging, *Org Lett,* **2021**, *23* (19), 7640-7644.

(77) Chi, W.; Qi, Q.; Lee, R.; Xu, Z.; Liu, X.; A Unified Push–Pull Model for Understanding the Ring-Opening Mechanism of Rhodamine Dyes, *The Journal of Physical Chemistry C,* **2020**, *124* (6), 3793-3801.
(78) Jo, J.; Lee, C. H.; Kopelman, R.; Wang, X.; In vivo quantitative imaging of tumor pH by nanosonophore assisted multispectral photoacoustic imaging, *Nat Commun,* **2017**, *8* (1), 471.
(79) Chen, Q.; Liu, X.; Chen, J.; Zeng, J.; Cheng, Z.; Liu, Z.; A Self-Assembled Albumin-Based Nanoprobe for In Vivo Ratiometric Photoacoustic pH Imaging, *Adv Mater,* **2015**, *27* (43), 6820-6827.
(80) Horvath, T. D.; Kim, G.; Kopelman, R.; Ashkenazi, S.; Ratiometric photoacoustic sensing of pH using a "sonophore", *Analyst,* **2008**, *133* (6), 747-749.
(81) Rathnamalala, C. S. L.; Hernandez, S.; Lucero, M. Y.; Swartchick, C. B.; Kalam Shaik, A.; Hammer, N. I.; East, A. K.; Gwaltney, S. R.; Chan, J.; Scott, C. N.; Xanthene-Based Nitric Oxide-Responsive Nanosensor for Photoacoustic Imaging in the SWIR Window, *Angew Chem Int Ed Engl,* **2023**, *62* (13), e202214855.
(82) Wang, S.; Yu, G.; Ma, Y.; Yang, Z.; Liu, Y.; Wang, J.; Chen, X.; Ratiometric Photoacoustic Nanoprobe for Bioimaging of Cu^{2+}, *ACS Appl Mater Interfaces,* **2019**, *11* (2), 1917-1923.
(83) Fan, X.; Ren, T.; Yang, W.; Zhang, X.; Yuan, L.; Activatable photoacoustic/fluorescent dual-modal probe for monitoring of drug-induced liver hypoxia in vivo, *Chem Commun (Camb),* **2021**, *57*, 8644-8647.
(84) Ikeno, T.; Hanaoka, K.; Iwaki, S.; Myochin, T.; Murayama, Y.; Ohde, H.; Komatsu, T.; Ueno, T.; Nagano, T.; Urano, Y.; Design and Synthesis of an Activatable Photoacoustic Probe for Hypochlorous Acid, *Anal Chem,* **2019**, *91* (14), 9086-9092.
(85) Myochin, T.; Hanaoka, K.; Iwaki, S.; Ueno, T.; Komatsu, T.; Terai, T.; Nagano, T.; Urano, Y.; Development of a series of near-infrared dark quenchers based on Si-rhodamines and their application to fluorescent probes, *J Am Chem Soc,* **2015**, *137* (14), 4759-4765.
(86) Brondsted, F.; Fang, Y.; Li, L.; Zhou, X.; Grant, S.; Stains, C. I.; Single Atom Stabilization of Phosphinate Ester-Containing Rhodamines Yields Cell Permeable Probes for Turn-On Photoacoustic Imaging, *Chemistry,* **2024**, *30* (1), e202303038.
(87) Liu, F.; Shi, X.; Liu, X.; Wang, F.; Yi, H. B.; Jiang, J. H.; Engineering an NIR rhodol derivative with spirocyclic ring-opening activation for high-contrast photoacoustic imaging, *Chem Sci,* **2019**, *10* (40), 9257-9264.
(88) Li, J.; Wang, J.; Xu, L.; Chi, H.; Liang, X.; Yoon, J.; Lin, W.; A Class of Activatable NIR-II Photoacoustic Dyes for High-Contrast Bioimaging, *Angew Chem Int Ed Engl,* **2024**, *63* (2), e202312632.
(89) Song, K. H.; Stein, E. W.; Margenthaler, J. A.; Wang, L. V.; Noninvasive photoacoustic identification of sentinel lymph nodes containing methylene blue in vivo in a rat model, *J Biomed Opt,* **2008**, *13* (5), 054033.
(90) Vy Phan, T. T.; Bharathiraja, S.; Nguyen, V. T.; Moorthy, M. S.; Manivasagan, P.; Lee, K. D.; Oh, J.; Polypyrrole–methylene blue nanoparticles as a single multifunctional nanoplatform for near-infrared photo-induced therapy and photoacoustic imaging, *RSC Advances,* **2017**, *7* (56), 35027-35037.
(91) Dana, N.; Fowler, R. A.; Allen, A.; Zoldan, J.; Suggs, L.; Emelianov, S.; In vitro photoacoustic sensing of calcium dynamics with Arsenazo III, *Laser Phys Lett,* **2016**, *13*, 075603.
(92) Beeler, T. J. S., A.; Martonosi, A.; The binding of arsenazo III to cell components, *Biochimica et Biophysica Acta,* **1980**, *629* (2), 317-327.
(93) Docampo, R.; Moreno, S. N.; Mason, R. P.; Generation of free radical metabolites and superoxide anion by the calcium indicators arsenazo III, antipyrylazo III, and murexide in rat liver microsomes, *Journal of Biological Chemistry,* **1983**, *258* (24), 14920-14925.
(94) Liu, W. W.; Chen, S. H.; Li, P. C.; Functional photoacoustic calcium imaging using chlorophosphonazo III in a 3D tumor cell culture, *Biomed Opt Express,* **2021**, *12* (2), 1154-1166.
(95) Tardivo, J. P.; Del Giglio, A.; de Oliveira, C. S.; Gabrielli, D. S.; Junqueira, H. C.; Tada, D. B.; Severino, D.; de Fatima Turchiello, R.; Baptista, M. S.; Methylene blue in photodynamic therapy: From basic mechanisms to clinical applications, *Photodiagnosis Photodyn Ther,* **2005**, *2* (3), 175-191.
(96) Ferguson, J. W.; Richard, J. J.; O'Laughlin, J. W.; Banks, C. V.; Simultaneous Spectrophotometric Determination of Calcium and Magnesium with Chlorophosphonazo III, *Analytical Chemistry,* **1964**, *36* (4), 796-799.
(97) Li, L.; Zemp, R. J.; Lungu, G.; Stoica, G.; Wang, L. V.; Photoacoustic imaging of lacZ gene expression in vivo, *J Biomed Opt,* **2007**, *12* (2), 020504.
(98) Cai, X.; Li, L.; Krumholz, A.; Guo, Z.; Erpelding, T. N.; Zhang, C.; Zhang, Y.; Xia, Y.; Wang, L. V.; Multi-scale molecular photoacoustic tomography of gene expression, *PLoS One,* **2012**, *7* (8), e43999.
(99) Paproski, R. J.; Forbrich, A. E.; Wachowicz, K.; Hitt, M. M.; Zemp, R. J.; Tyrosinase as a dual reporter gene for both photoacoustic and magnetic resonance imaging, *Biomedical Optics Express,* **2011**, *2* (4), 771–780.
(100) Paproski, R. J.; Heinmiller, A.; Wachowicz, K.; Zemp, R. J.; Multi-wavelength photoacoustic imaging of inducible tyrosinase reporter gene expression in xenograft tumors, *Sci Rep,* **2014**, *4*, 5329.
(101) Krumholz, A.; Vanvickle-Chavez, S. J.; Yao, J.; Fleming, T. P.; Gillanders, W. E.; Wang, L. V.; Photoacoustic microscopy of tyrosinase reporter gene in vivo, *J Biomed Opt,* **2011**, *16* (8), 080503.
(102) Raposo, G.; Marks, M. S.; Melanosomes--dark organelles enlighten endosomal membrane transport, *Nat Rev Mol Cell Biol,* **2007**, *8* (10), 786-797.

(103) Sigmund, F.; Massner, C.; Erdmann, P.; Stelzl, A.; Rolbieski, H.; Desai, M.; Bricault, S.; Worner, T. P.; Snijder, J.; Geerlof, A.; et al.; Bacterial encapsulins as orthogonal compartments for mammalian cell engineering, *Nat Commun,* **2018**, *9* (1), 1990.
(104) Reijmers, L. G., Perkins, B. L., Matsuo, N., Mayford, M.; Localization of a Stable Neural Correlate of Associative Memory, *Science,* **2007**, *317* (5842), 1230-1233.
(105) Lauri, A.; Soliman, D.; Omar, M.; Stelzl, A.; Ntziachristos, V.; Westmeyer, G. G.; Whole-Cell Photoacoustic Sensor Based on Pigment Relocalization, *ACS Sens,* **2019**, *4* (3), 603-612.
(106) Chan, M. C.; Karasawa, S.; Mizuno, H.; Bosanac, I.; Ho, D.; Prive, G. G.; Miyawaki, A.; Ikura, M.; Structural characterization of a blue chromoprotein and its yellow mutant from the sea anemone Cnidopus japonicus, *J Biol Chem,* **2006**, *281* (49), 37813-37819.
(107) Shkrob, M. A.; Yanushevich, Y. G.; Chudakov, D. M.; Gurskaya, N. G.; Labas, Y. A.; Poponov, S. Y.; Mudrik, N. N.; Lukyanov, S.; Lukyanov, K. A.; Far-red fluorescent proteins evolved from a blue chromoprotein from Actinia equina, *Biochem J,* **2005**, *392*, 649-654.
(108) Laufer, J.; Jathoul, A.; Pule, M.; Beard, P.; In vitro characterization of genetically expressed absorbing proteins using photoacoustic spectroscopy, *Biomed Opt Express,* **2013**, *4* (11), 2477-2490.
(109) Pettikiriarachchi, A.; Gong, L.; Perugini, M. A.; Devenish, R. J.; Prescott, M.; Ultramarine, a chromoprotein acceptor for Forster resonance energy transfer, *PLoS One,* **2012**, *7* (7), e41028.
(110) Li, Y.; Forbrich, A.; Wu, J.; Shao, P.; Campbell, R. E.; Zemp, R.; Engineering Dark Chromoprotein Reporters for Photoacoustic Microscopy and FRET Imaging, *Sci Rep,* **2016**, *6*, 22129.
(111) Tian, L.; Hires, S. A.; Mao, T.; Huber, D.; Chiappe, M. E.; Chalasani, S. H.; Petreanu, L.; Akerboom, J.; McKinney, S. A.; Schreiter, E. R.; et al.; Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators, *Nature Methods,* **2009**, *6* (12), 875-881.
(112) Chen, T. W.; Wardill, T. J.; Sun, Y.; Pulver, S. R.; Renninger, S. L.; Baohan, A.; Schreiter, E. R.; Kerr, R. A.; Orger, M. B.; Jayaraman, V.; et al.; Ultrasensitive fluorescent proteins for imaging neuronal activity, *Nature,* **2013**, *499* (7458), 295-300.
(113) Dana, H.; Sun, Y.; Mohar, B.; Hulse, B. K.; Kerlin, A. M.; Hasseman, J. P.; Tsegaye, G.; Tsang, A.; Wong, A.; Patel, R.; et al.; High-performance calcium sensors for imaging activity in neuronal populations and microcompartments, *Nat Methods,* **2019**, *16* (7), 649-657.
(114) Zhang, Y.; Rozsa, M.; Liang, Y.; Bushey, D.; Wei, Z.; Zheng, J.; Reep, D.; Broussard, G. J.; Tsang, A.; Tsegaye, G.; et al.; Fast and sensitive GCaMP calcium indicators for imaging neural populations, *Nature,* **2023**, *615* (7954), 884-891.
(115) Akerboom, J.; Rivera, J. D.; Guilbe, M. M.; Malave, E. C.; Hernandez, H. H.; Tian, L.; Hires, S. A.; Marvin, J. S.; Looger, L. L.; Schreiter, E. R.; Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design, *J Biol Chem,* **2009**, *284* (10), 6455-6464.
(116) Dean-Ben, X. L.; Sela, G.; Lauri, A.; Kneipp, M.; Ntziachristos, V.; Westmeyer, G. G.; Shoham, S.; Razansky, D.; Functional optoacoustic neuro-tomography for scalable whole-brain monitoring of calcium indicators, *Light Sci Appl,* **2016**, *5* (12), e16201.
(117) Zhang, R.; Li, L. S.; Rao, B.; Rong, H.; Sun, M. Y.; Yao, J.; Chen, R.; Zhou, Q.; Mennerick, S.; Raman, B.; et al.; Multiscale photoacoustic tomography of neural activities with GCaMP calcium indicators, *J Biomed Opt,* **2022**, *27* (9), 096004.
(118) Mishra, K.; Fuenzalida-Werner, J. P.; Pennacchietti, F.; Janowski, R.; Chmyrov, A.; Huang, Y.; Zakian, C.; Klemm, U.; Testa, I.; Niessing, D.; et al.; Genetically encoded photo-switchable molecular sensors for optoacoustic and super-resolution imaging, *Nat Biotechnol,* **2022**, *40*, 598-605.
(119) Ding, J.; Luo, A. F.; Hu, L.; Wang, D.; Shao, F.; Structural basis of the ultrasensitive calcium indicator GCaMP6, *Sci China Life Sci,* **2014**, *57* (3), 269-274.
(120) Filonov, G. S.; Piatkevich, K. D.; Ting, L. M.; Zhang, J.; Kim, K.; Verkhusha, V. V.; Bright and stable near-infrared fluorescent protein for in vivo imaging, *Nat Biotechnol,* **2011**, *29* (8), 757-761.
(121) Filonov, G. S.; Krumholz, A.; Xia, J.; Yao, J.; Wang, L. V.; Verkhusha, V. V.; Deep-tissue photoacoustic tomography of a genetically encoded near-infrared fluorescent probe, *Angew Chem Int Ed Engl,* **2012**, *51* (6), 1448-1451.
(122) Shcherbakova, D. M.; Verkhusha, V. V.; Near-infrared fluorescent proteins for multicolor in vivo imaging, *Nat Methods,* **2013**, *10* (8), 751-754.
(123) Krumholz, A.; Shcherbakova, D. M.; Xia, J.; Wang, L. V.; Verkhusha, V. V.; Multicontrast photoacoustic in vivo imaging using near-infrared fluorescent proteins, *Sci Rep,* **2014**, *4*, 3939.
(124) Yao, J.; Kaberniuk, A. A.; Li, L.; Shcherbakova, D. M.; Zhang, R.; Wang, L.; Li, G.; Verkhusha, V. V.; Wang, L. V.; Multiscale photoacoustic tomography using reversibly switchable bacterial phytochrome as a near-infrared photochromic probe, *Nat Methods,* **2016**, *13* (1), 67-73.
(125) Godin, A. G.; Lounis, B.; Cognet, L.; Super-resolution microscopy approaches for live cell imaging, *Biophys J,* **2014**, *107* (8), 1777-1784.

(126) Li, L.; Shemetov, A. A.; Baloban, M.; Hu, P.; Zhu, L.; Shcherbakova, D. M.; Zhang, R.; Shi, J.; Yao, J.; Wang, L. V.; et al.; Small near-infrared photochromic protein for photoacoustic multi-contrast imaging and detection of protein interactions in vivo, *Nat Commun,* **2018**, *9* (1), 2734.
(127) Chee, R. K. W.; Li, Y.; Zhang, W.; Campbell, R. E.; Zemp, R. J.; In vivo photoacoustic difference-spectra imaging of bacteria using photoswitchable chromoproteins, *J Biomed Opt,* **2018**, *23* (10), 1-11.
(128) Märk, J.; Dortay, H.; Wagener, A.; Zhang, E.; Buchmann, J.; Grötzinger, C.; Friedrich, T.; Laufer, J.; Dual-wavelength 3D photoacoustic imaging of mammalian cells using a photoswitchable phytochrome reporter protein, *Communications Physics,* **2018**, *1*, 3.
(129) Qian, Y.; Piatkevich, K. D.; Mc Larney, B.; Abdelfattah, A. S.; Mehta, S.; Murdock, M. H.; Gottschalk, S.; Molina, R. S.; Zhang, W.; Chen, Y.; et al.; A genetically encoded near-infrared fluorescent calcium ion indicator, *Nat Methods,* **2019**, *16* (2), 171-174.
(130) Yu, D.; Baird, M. A.; Allen, J. R.; Howe, E. S.; Klassen, M. P.; Reade, A.; Makhijani, K.; Song, Y.; Liu, S.; Murthy, Z.; et al.; A naturally monomeric infrared fluorescent protein for protein labeling in vivo, *Nat Methods,* **2015**, *12* (8), 763-765.
(131) Matlashov, M. E.; Shcherbakova, D. M.; Alvelid, J.; Baloban, M.; Pennacchietti, F.; Shemetov, A. A.; Testa, I.; Verkhusha, V. V.; A set of monomeric near-infrared fluorescent proteins for multicolor imaging across scales, *Nat Commun,* **2020**, *11* (1), 239.
(132) Qian, Y.; Cosio, D. M. O.; Piatkevich, K. D.; Aufmkolk, S.; Su, W. C.; Celiker, O. T.; Schohl, A.; Murdock, M. H.; Aggarwal, A.; Chang, Y. F.; et al.; Improved genetically encoded near-infrared fluorescent calcium ion indicators for in vivo imaging, *PLoS Biol,* **2020**, *18* (11), e3000965.
(133) Kobachi, K., Kuno, S., Sato, S., Sumiyama, K., Matsuda, M., Terai, K.; Biliverdin Reductase-A Deficiency Brighten and Sensitize Biliverdin-binding Chromoproteins, *Cell Structure and Function,* **2020**, *45* (2), 131-141.
(134) Shaykevich, S.; Little, J. P.; Campbell, R. E.; Shoham, S., In vivo characterization of the near-infrared genetically encoded calcium indicator NIR-GECO2G. In *SPIE BiOS*, San Francisco, 2023; Vol. 12365, p 1236509. DOI: 10.1117/12.2650973.
(135) Shaykevich, S.; Chan, R. W.; Chan, K. C.; Campbell, R. E.; Razansky, D.; Shoham, S., Towards multimodal functional imaging of brain dynamics using hybrid near-infrared optoacoustic and fluorescence imaging. In *SPIE BiOS*, San Francisco, 2024; Vol. PC12828, p PC128280C. DOI: 10.1117/12.3003298.
(136) Chai, F.; Fujii, H.; Le, G. N. T.; Lin, C.; Ota, K.; Lin, K. M.; Pham, L. M. T.; Zou, P.; Drobizhev, M.; Nasu, Y.; et al.; Development of an miRFP680-Based Fluorescent Calcium Ion Biosensor Using End-Optimized Transposons, *ACS Sens,* **2024**, *9* (6), 3394-3402.
(137) Subach, O. M.; Barykina, N. V.; Anokhin, K. V.; Piatkevich, K. D.; Subach, F. V.; Near-Infrared Genetically Encoded Positive Calcium Indicator Based on GAF-FP Bacterial Phytochrome, *Int J Mol Sci,* **2019**, *20* (14), 3488.
(138) Shemetov, A. A.; Monakhov, M. V.; Zhang, Q.; Canton-Josh, J. E.; Kumar, M.; Chen, M.; Matlashov, M. E.; Li, X.; Yang, W.; Nie, L.; et al.; A near-infrared genetically encoded calcium indicator for in vivo imaging, *Nat Biotechnol,* **2021**, *39*, 368-377.
(139) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Friedman Ohana, R.; Urh, M.; et al.; HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis, *ACS Chem Biol,* **2008**, *3* (6), 373-382.
(140) Encell, L. P.; Friedman Ohana, R.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Wood, M. G.; Los, G. V.; McDougall, M. G.; Zimprich, C.; Karassina, N.; et al.; Development of a Dehalogenase-Based Protein Fusion Tag Capable of Rapid, Selective and Covalent Attachment to Customizable Ligands, *Current Chemical Genomics,* **2012**, *6*, 55-71.
(141) Cook, A.; Walterspiel, F.; Deo, C.; HaloTag-Based Reporters for Fluorescence Imaging and Biosensing, *Chembiochem,* **2023**, *24* (12), e202300022.
(142) Hoelzel, C. A.; Zhang, X.; Visualizing and Manipulating Biological Processes by Using HaloTag and SNAP-Tag Technologies, *Chembiochem,* **2020**, *21* (14), 1935-1946.
(143) Deo, C.; Abdelfattah, A. S.; Bhargava, H. K.; Berro, A. J.; Falco, N.; Farrants, H.; Moeyaert, B.; Chupanova, M.; Lavis, L. D.; Schreiter, E. R.; The HaloTag as a general scaffold for far-red tunable chemigenetic indicators, *Nat Chem Biol,* **2021**, *17*, 718-723.
(144) Lavis, L. D.; Teaching Old Dyes New Tricks: Biological Probes Built from Fluoresceins and Rhodamines, *Annu Rev Biochem,* **2017**, *86*, 825-843.
(145) Grimm, J. B.; Muthusamy, A. K.; Liang, Y.; Brown, T. A.; Lemon, W. C.; Patel, R.; Lu, R.; Macklin, J. J.; Keller, P. J.; Ji, N.; et al.; A general method to fine-tune fluorophores for live-cell and in vivo imaging, *Nat Methods,* **2017**, *14* (10), 987-994.
(146) Best, M.; Porth, I.; Hauke, S.; Braun, F.; Herten, D. P.; Wombacher, R.; Protein-specific localization of a rhodamine-based calcium-sensor in living cells, *Org Biomol Chem,* **2016**, *14* (24), 5606-5611.
(147) Deo, C.; Sheu, S. H.; Seo, J.; Clapham, D. E.; Lavis, L. D.; Isomeric Tuning Yields Bright and Targetable Red Ca^{2+} Indicators, *J Am Chem Soc,* **2019**, *141* (35), 13734-13738.

(148) Wang, L.; Tran, M.; D'Este, E.; Roberti, J.; Koch, B.; Xue, L.; Johnsson, K.; A general strategy to develop cell permeable and fluorogenic probes for multicolour nanoscopy, *Nat Chem,* **2020**, *12* (2), 165-172.
(149) Mertes, N.; Busch, M.; Huppertz, M. C.; Hacker, C. N.; Wilhelm, J.; Gurth, C. M.; Kuhn, S.; Hiblot, J.; Koch, B.; Johnsson, K.; Fluorescent and Bioluminescent Calcium Indicators with Tuneable Colors and Affinities, *J Am Chem Soc,* **2022**, *144* (15), 6928-6935.
(150) Deal, P. E.; Liu, P.; Al-Abdullatif, S. H.; Muller, V. R.; Shamardani, K.; Adesnik, H.; Miller, E. W.; Covalently Tethered Rhodamine Voltage Reporters for High Speed Functional Imaging in Brain Tissue, *J Am Chem Soc,* **2020**, *142* (1), 614-622.
(151) Kirk, M. J.; Benlian, B. R.; Han, Y.; Gold, A.; Ravi, A.; Deal, P. E.; Molina, R. S.; Drobizhev, M.; Dickman, D.; Scott, K.; et al.; Voltage Imaging in Drosophila Using a Hybrid Chemical-Genetic Rhodamine Voltage Reporter, *Front Neurosci,* **2021**, *15*, 754027.
(152) Ortiz, G.; Liu, P.; Deal, P. E.; Nensel, A. K.; Martinez, K. N.; Shamardani, K.; Adesnik, H.; Miller, E. W.; A silicon-rhodamine chemical-genetic hybrid for far red voltage imaging from defined neurons in brain slice, *RSC Chem Biol,* **2021**, *2* (6), 1594-1599.
(153) Hellweg, L.; Edenhofer, A.; Barck, L.; Huppertz, M. C.; Frei, M. S.; Tarnawski, M.; Bergner, A.; Koch, B.; Johnsson, K.; Hiblot, J.; A general method for the development of multicolor biosensors with large dynamic ranges, *Nat Chem Biol,* **2023**, *19*, 1147–1157.
(154) Hiblot, J.; Yu, Q.; Sabbadini, M. D. B.; Reymond, L.; Xue, L.; Schena, A.; Sallin, O.; Hill, N.; Griss, R.; Johnsson, K.; Luciferases with Tunable Emission Wavelengths, *Angew Chem Int Ed Engl,* **2017**, *56* (46), 14556-14560.
(155) Thirukkumaran, O. M.; Wang, C.; Asouzu, N. J.; Fron, E.; Rocha, S.; Hofkens, J.; Lavis, L. D.; Mizuno, H.; Improved HaloTag Ligand Enables BRET Imaging With NanoLuc, *Front Chem,* **2019**, *7*, 938.
(156) Benlian, B. R.; Klier, P. E. Z.; Martinez, K. N.; Schwinn, M. K.; Kirkland, T. A.; Miller, E. W.; Small Molecule-Protein Hybrid for Voltage Imaging via Quenching of Bioluminescence, *ACS Sens,* **2021**, *6* (5), 1857-1863.
(157) Abdelfattah, A. S.; Kawashima, T.; Singh, A.; Novak, O.; Liu, H.; Shuai, Y.; Huang, Y.-C.; Campagnola, L.; Seeman, S. C.; Yu, J.; et al.; Bright and photostable chemigenetic indicators for extended in vivo voltage imaging, *Science,* **2019**, *364*, 699-704.
(158) Abdelfattah, A. S.; Valenti, R.; Zheng, J.; Wong, A.; Team, G. P.; Podgorski, K.; Koyama, M.; Kim, D. S.; Schreiter, E. R.; A general approach to engineer positive-going eFRET voltage indicators, *Nat Commun,* **2020**, *11* (1), 3444.
(159) Farrants, H.; Shuai, Y.; Lemon, W. C.; Hernandez, C. M.; Zhang, D.; Yang, S.; Patel, R.; Qiao, G.; Frei, M. S.; Grimm, J. B.; et al.; A modular chemigenetic calcium indicator for multiplexed in vivo functional imaging, *Nat Methods,* **2024**.
(160) Zhu, W.; Takeuchi, S.; Imai, S.; Terada, T.; Ueda, T.; Nasu, Y.; Terai, T.; Campbell, R. E.; Chemigenetic indicators based on synthetic chelators and green fluorescent protein, *Nat Chem Biol,* **2023**, *19* (1), 38-44.
(161) Oheim, M.; van 't Hoff, M.; Feltz, A.; Zamaleeva, A.; Mallet, J. M.; Collot, M.; New red-fluorescent calcium indicators for optogenetics, photoactivation and multi-color imaging, *Biochim Biophys Acta,* **2014**, *1843* (10), 2284-2306.
(162) Lin, M. Z.; Schnitzer, M. J.; Genetically encoded indicators of neuronal activity, *Nat Neurosci,* **2016**, *19* (9), 1142-1153.
(163) Yao, J.; Wang, L. V.; Sensitivity of photoacoustic microscopy, *Photoacoustics,* **2014**, *2* (2), 87-101.
(164) Vetschera, P.; Mishra, K.; Fuenzalida-Werner, J. P.; Chmyrov, A.; Ntziachristos, V.; Stiel, A. C.; Characterization of Reversibly Switchable Fluorescent Proteins in Optoacoustic Imaging, *Anal Chem,* **2018**, *90* (17), 10527-10535.
(165) Cook, A.; Kaydanov, N.; Ugarte-Uribe, B.; Boffi, J. C.; Kamm, G. B.; Prevedel, R.; Deo, C.; Chemigenetic Far-Red Labels and Ca^{2+} Indicators Optimized for Photoacoustic Imaging, *J Am Chem Soc,* **2024**, *146* (34), 23963-23971.
(166) Grimm, J. B.; Tkachuk, A. N.; Xie, L.; Choi, H.; Mohar, B.; Falco, N.; Schaefer, K.; Patel, R.; Zheng, Q.; Liu, Z.; et al.; A general method to optimize and functionalize red-shifted rhodamine dyes, *Nat Methods,* **2020**, *17* (8), 815-821.
(167) Grimm, J. B.; Lavis, L. D.; Synthesis of Rhodamines from Fluoresceins Using Pd-Catalyzed C-N Cross-Coupling, *Organic Letters,* **2011**, *13* (24), 6354–6357.
(168) Grimm, J. B.; Tkachuk, A. N.; Patel, R.; Hennigan, S. T.; Gutu, A.; Dong, P.; Gandin, V.; Osowski, A. M.; Holland, K. L.; Liu, Z. J.; et al.; Optimized Red-Absorbing Dyes for Imaging and Sensing, *J Am Chem Soc,* **2023**, *145* (42), 23000-23013.
(169) Grimm, J. B.; Brown, T. A.; Tkachuk, A. N.; Lavis, L. D.; General Synthetic Method for Si-Fluoresceins and Si-Rhodamines, *ACS Cent Sci,* **2017**, *3* (9), 975-985.
(170) Lukinavicius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Correa, I. R., Jr.; Luo, Z. G.; et al.; A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins, *Nat Chem,* **2013**, *5* (2), 132-139.

(171) Usama, S. M.; Marker, S. C.; Li, D. H.; Caldwell, D. R.; Stroet, M.; Patel, N. L.; Tebo, A. G.; Hernot, S.; Kalen, J. D.; Schnermann, M.; Method To Diversify Cyanine Chromophore Functionality Enables Improved Biomolecule Tracking and Intracellular Imaging, *J Am Chem Soc,* **2023**, *145* (27), 14647-14659.
(172) Gagliardi, L. G., Castells, C. B., Rafols, C., Roses, M., Bosch, E.; Static Dielectric Constants of Acetonitrile/Water Mixtures at Different Temperatures and Debye−Hückel *A* and a_0B Parameters for Activity Coefficients, *J. Chem. Eng. Data,* **2007**, *52*, 1103-1107.
(173) Rajapaksha, I.; Chang, H.; Xiong, Y.; Marder, S.; Gwaltney, S. R.; Scott, C. N.; New Design Strategy Toward NIR I Xanthene-Based Dyes, *J Org Chem,* **2020**, *85* (19), 12108-12116.
(174) Deng, Z.; Li, L.; Jia, H.; Li, N. F.; He, J.; Li, M. D.; Phillips, D. L.; Li, Y.; Insights into the Photodynamics of Fluorescence Emission and Singlet Oxygen Generation of Fluorogen Activating Protein-Malachite Green Systems, *Chemistry,* **2023**, *29* (16), e202203684.
(175) Luo, Y.; Zhou, L.; Du, L.; Xie, Y.; Lou, X. Y.; Cai, L.; Tang, B. Z.; Gong, P.; Zhang, P.; Malachite green: a long-buried water-soluble AIEgen with near-infrared fluorescence for living cell nucleus staining, *Chem Commun (Camb),* **2024**, *60* (11), 1452-1455.
(176) Grimm, J. B.; English, B. P.; Chen, J.; Slaughter, J. P.; Zhang, Z.; Revyakin, A.; Patel, R.; Macklin, J. J.; Normanno, D.; Singer, R. H.; et al.; A general method to improve fluorophores for live-cell and single-molecule microscopy, *Nat Methods,* **2015**, *12* (3), 244-250.
(177) Sabnis, R. W.; A facile synthesis of phthalein indicator dyes, *Tetrahedron Letters,* **2009**, *50* (46), 6261-6263.
(178) Froese, R. D.; Whiteker, G. T.; Peterson, T. H.; Arriola, D. J.; Renga, J. M.; Shearer, J. W.; Computational and Experimental Studies of Regioselective S_NAr Halide Exchange (Halex) Reactions of Pentachloropyridine, *J Org Chem,* **2016**, *81* (22), 10672-10682.
(179) Taeufer, T.; Pospech, J.; Palladium-Catalyzed Synthesis of N,N-Dimethylanilines via Buchwald-Hartwig Amination of (Hetero)aryl Triflates, *J Org Chem,* **2020**, *85* (11), 7097-7111.
(180) Lavis, L. D.; Grimm, J. B.; Red-Shifted Fluorophores, *The United States Patent and Trademark Office,* **2021**, *US 2021/0171490 Al*.
(181) Butkevich, A. N.; Modular Synthetic Approach to Silicon-Rhodamine Homologues and Analogues via Bis-aryllanthanum Reagents, *Org Lett,* **2021**, *23* (7), 2604-2609.
(182) Kneisel, F. F. M., Y.; Knapp, K. M.; Zipse, H.; Knochel, P.; Stereoselective cyclizations mediated by functionalized organomagnesium reagents and catalyzed by cobalt or copper salts, *Tetrahedron Lett,* **2002**, *43*, 4875-4879.
(183) Lardon, N.; Wang, L.; Tschanz, A.; Hoess, P.; Tran, M.; D'Este, E.; Ries, J.; Johnsson, K.; Systematic Tuning of Rhodamine Spirocyclization for Super-resolution Microscopy, *J Am Chem Soc,* **2021**, *143* (36), 14592-14600.
(184) Hallas, G. P., K. N.; Waring, D. R.; Humpston, J. R.; Jones, A. M.; Steric effects in di- and tri-arylmethane dyes. Part 13. Electronic absorption spectra of derivatives of Crystal Violet, Malachite Green, and Michler's Hydrol Blue exhibiting simultaneous central and terminal steric distortion, *J. Chem. Soc. Perkin Trans. 2,* **1977**, *4*, 450-456.
(185) Butkevich, A. N.; Belov, V. N.; Kolmakov, K.; Sokolov, V. V.; Shojaei, H.; Sidenstein, S. C.; Kamin, D.; Matthias, J.; Vlijm, R.; Engelhardt, J.; et al.; Hydroxylated Fluorescent Dyes for Live-Cell Labeling: Synthesis, Spectra and Super-Resolution STED, *Chemistry,* **2017**, *23* (50), 12114-12119.
(186) Nemoto, H.; Kubota, Y.; Yamamoto, Y.; Development of a New Acyl Anion Equivalent for the Preparation of Masked Activated Esters and Their Use To Prepare a Dipeptide, *J Org Chem,* **1990**, *55* (15), 4515-4516.
(187) Yang, K. S.; Nibbs, A. E.; Turkmen, Y. E.; Rawal, V. H.; Squaramide-catalyzed enantioselective Michael addition of masked acyl cyanides to substituted enones, *J Am Chem Soc,* **2013**, *135* (43), 16050-16053.
(188) Nemoto, H.; Li, X.; Ma, R.; Suzuki, I.; Shibuya, M.; A three-step preparation of MAC, *Tetrahedron Lett,* **2003**, *44*, 73-75.
(189) Szent-Gyorgyi, C.; Schmidt, B. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; et al.; Fluorogen-activating single-chain antibodies for imaging cell surface proteins, *Nat Biotechnol,* **2008**, *26* (2), 235-240.
(190) Frei, M. S.; Tarnawski, M.; Roberti, M. J.; Koch, B.; Hiblot, J.; Johnsson, K.; Engineered HaloTag variants for fluorescence lifetime multiplexing, *Nat Methods,* **2021**, *19*, 65-70.
(191) Zheng, Q.; Ayala, A. X.; Chung, I.; Weigel, A. V.; Ranjan, A.; Falco, N.; Grimm, J. B.; Tkachuk, A. N.; Wu, C.; Lippincott-Schwartz, J.; et al.; Rational Design of Fluorogenic and Spontaneously Blinking Labels for Super-Resolution Imaging, *ACS Cent Sci,* **2019**, *5* (9), 1602-1613.
(192) Shcherbakova, D. M.; Baloban, M.; Emelyanov, A. V.; Brenowitz, M.; Guo, P.; Verkhusha, V. V.; Bright monomeric near-infrared fluorescent proteins as tags and biosensors for multiscale imaging, *Nat Commun,* **2016**, *7*, 12405.
(193) Wang, L.; Hiblot, J.; Popp, C.; Xue, L.; Johnsson, K.; Environmentally Sensitive Color-Shifting Fluorophores for Bioimaging, *Angew Chem Int Ed Engl,* **2020**, *59* (49), 21880-21884.

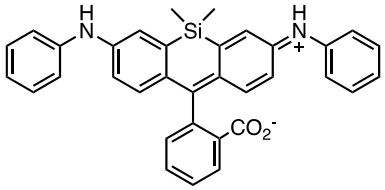
(194) Maravall, M.; Mainen, Z. F.; Sabatini, B. L.; Svoboda, K.; Estimating Intracellular Calcium Concentrations and Buffering without Wavelength Ratioing, *Biophys J,* **2000**, *78*, 2655–2667.
(195) Miyawaki, A.; Llopis, J.; Heim, R.; McCaffery, J. M.; Adams, J. A.; Ikura, M.; Tsien, R. Y.; Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin, *Nature,* **1997**, *388* (6645), 882-887.
(196) Barykina, N. V.; Subach, O. M.; Piatkevich, K. D.; Jung, E. E.; Malyshev, A. Y.; Smirnov, I. V.; Bogorodskiy, A. O.; Borshchevskiy, V. I.; Varizhuk, A. M.; Pozmogova, G. E.; et al.; Green fluorescent genetically encoded calcium indicator based on calmodulin/M13-peptide from fungi, *PLoS One,* **2017**, *12* (8), e0183757.
(197) Linse S.; Helmersson, A. F., S.; Calcium binding to calmodulin and its globular domains, *The Journal of Biological Chemistry,* **1991**, *266* (13), 8050-8054.
(198) Valiente-Gabioud, A. A.; Garteizgogeascoa Suner, I.; Idziak, A.; Fabritius, A.; Basquin, J.; Angibaud, J.; Nagerl, U. V.; Singh, S. P.; Griesbeck, O.; Fluorescent sensors for imaging of interstitial calcium, *Nat Commun,* **2023**, *14* (1), 6220.
(199) Suzuki, J.; Kanemaru, K.; Ishii, K.; Ohkura, M.; Okubo, Y.; Iino, M.; Imaging intraorganellar Ca^{2+} at subcellular resolution using CEPIA, *Nat Commun,* **2014**, *5*, 4153.
(200) Henderson, M. J.; Baldwin, H. A.; Werley, C. A.; Boccardo, S.; Whitaker, L. R.; Yan, X.; Holt, G. T.; Schreiter, E. R.; Looger, L. L.; Cohen, A. E.; et al.; A Low Affinity GCaMP3 Variant (GCaMPer) for Imaging the Endoplasmic Reticulum Calcium Store, *PLoS One,* **2015**, *10* (10), e0139273.
(201) Deng, X.; Yao, X. Q.; Berglund, K.; Dong, B.; Ouedraogo, D.; Ghane, M. A.; Zhuo, Y.; McBean, C.; Wei, Z. Z.; Gozem, S.; et al.; Tuning Protein Dynamics to Sense Rapid Endoplasmic-Reticulum Calcium Dynamics, *Angew Chem Int Ed Engl,* **2021**, *60* (43), 23289-23298.
(202) Kennedy, M. J.; Hanus, C.; Architecture and Dynamics of the Neuronal Secretory Network, *Annu Rev Cell Dev Biol,* **2019**, *35*, 543-566.
(203) Tay, L. H.; Griesbeck, O.; Yue, D. T.; Live-cell transforms between Ca^{2+} transients and FRET responses for a troponin-C-based Ca^{2+} sensor, *Biophys J,* **2007**, *93* (11), 4031-4040.
(204) Brocard, J. B.; Rajdev, S.; Reynolds, I. J.; Glutamate-induced increases in intracellular free Mg^{2+} in cultured cortical neurons, *Neuron,* **1993**, *11*, 751-757.
(205) Chesler, M.; Regulation and Modulation of pH in the Brain, *Physiological Reviews,* **2003**, *83* (4), 1183-1221.
(206) Bucevičius, J.; Kostiuk, G.; Gerasimaitė, R.; Gilat, T.; Lukinavičius, G.; Enhancing the biocompatibility of rhodamine fluorescent probes by a neighbouring group effect, *Chemical Science,* **2020**, *11* (28), 7313-7323.
(207) Martin, A.; Rivera-Fuentes, P.; A general strategy to develop fluorogenic polymethine dyes for bioimaging, *Nat Chem,* **2024**, *16* (1), 28-35.
(208) Mohar, B.; Grimm, J. B.; Patel, R.; Brown, T. A.; Tillberg, P. W.; Lavis, L. D.; Spruston, N.; Svoboda, K.; Brain-wide measurement of protein turnover with high spatial and temporal resolution *bioRxiv,* **2023**.
(209) Nakai, J.; Ohkura, M.; Imoto, K.; A high signal-to-noise Ca^{2+} probe composed of a single green fluorescent protein, *Nature Biotechnology,* **2001**, *19*, 137-141.
(210) Ohkura, M.; Matsuzaki, M.; Kasai, H.; Imoto, K.; Nakai, J.; Genetically Encoded Bright Ca^{2+} Probe Applicable for Dynamic Ca^{2+} Imaging of Dendritic Spines, *Anal Chem,* **2005**, *77*, 5861-5869.
(211) Tallini, Y. N.; Ohkura, M.; Choi, B.-R.; Ji, G.; Imoto, K.; Doran, R.; Lee, J.; Plan, P.; Wilson, J.; Xin, H.-B.; et al.; Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca^{2+} indicator GCaMP2, *Proc Natl Acad Sci U S A,* **2006**, *112* (6), 4753-4758.
(212) Akerboom, J.; Chen, T. W.; Wardill, T. J.; Tian, L.; Marvin, J. S.; Mutlu, S.; Calderon, N. C.; Esposti, F.; Borghuis, B. G.; Sun, X. R.; et al.; Optimization of a GCaMP calcium indicator for neural activity imaging, *J Neurosci,* **2012**, *32* (40), 13819-13840.
(213) Frei, M. S.; Koch, B.; Hiblot, J.; Johnsson, K.; Live-Cell Fluorescence Lifetime Multiplexing Using Synthetic Fluorescent Probes, *ACS Chem Biol,* **2022**, *17* (6), 1321-1327.
(214) Inoue, M.; Genetically encoded calcium indicators to probe complex brain circuit dynamics in vivo, *Neurosci Res,* **2021**, *169*, 2-8.
(215) Inoue, M.; Takeuchi, A.; Horigane, S.-i.; Ohkura, M.; Gengyo-Ando, K.; Fujii, H.; Kamijo, S.; Takemoto-Kimura, S.; Kano, M.; Nakai, M.; et al.; Rational design of a high-affinity, fast, red calcium indicator R-CaMP2, *Nat Methods,* **2015**, *12*, 64-70.
(216) Helassa, N.; Podor, B.; Fine, A.; Torok, K.; Design and mechanistic insight into ultrafast calcium indicators for monitoring intracellular calcium dynamics, *Sci Rep,* **2016**, *6*, 38276.
(217) Shen, Y.; Dana, H.; Abdelfattah, A. S.; Patel, R.; Shea, J.; Molina, R. S.; Rawal, B.; Rancic, V.; Chang, Y. F.; Wu, L.; et al.; A genetically encoded Ca^{2+} indicator based on circularly permutated sea anemone red fluorescent protein eqFP578, *BMC Biol,* **2018**, *16* (1), 9.
(218) Bulina, M. E.; Lukyanov, K. A.; Britanova, O. V.; Onichtchouk, D.; Lukyanov, S.; Chudakov, D. M.; Chromophore-assisted light inactivation (CALI) using the phototoxic fluorescent protein KillerRed, *Nat Protoc,* **2006**, *1* (2), 947-953.

(219) Abrahamse, H.; Hamblin, Michael R.; New photosensitizers for photodynamic therapy, *Biochemical Journal,* **2016**, *473* (4), 347-364.
(220) He, J.; Wang, Y.; Missinato, M. A.; Onuoha, E.; Perkins, L. A.; Watkins, S. C.; St Croix, C. M.; Tsang, M.; Bruchez, M. P.; A genetically targetable near-infrared photosensitizer, *Nat Methods,* **2016**, *13* (3), 263-268.
(221) Lindig, B. A.; Rodgers, M. A. J.; Schaap, A. P.; Determination of the Lifetime of Singlet Oxygen in D_20 Using 9,10-Anthracenedipropionic Acid, a Water-Soluble Probe, *J Am Chem Soc,* **1980**, *102* (17), 5590-5593.
(222) Mirdita, M.; Schutze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M.; ColabFold: making protein folding accessible to all, *Nat Methods,* **2022**, *19* (6), 679-682.
(223) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al.; Highly accurate protein structure prediction with AlphaFold, *Nature,* **2021**, *596* (7873), 583-589.
(224) Thestrup, T.; Litzlbauer, J.; Bartholomaus, I.; Mues, M.; Russo, L.; Dana, H.; Kovalchuk, Y.; Liang, Y.; Kalamakis, G.; Laukat, Y.; et al.; Optimized ratiometric calcium sensors for functional in vivo imaging of neurons and T lymphocytes, *Nat Methods,* **2014**, *11* (2), 175-182.
(225) Hofmann, U. A. T.; Fabritius, A.; Rebling, J.; Estrada, H.; Dean-Ben, X. L.; Griesbeck, O.; Razansky, D.; High-Throughput Platform for Optoacoustic Probing of Genetically Encoded Calcium Ion Indicators, *iScience,* **2019**, *22*, 400-408.
(226) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J.; Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Advanced Drug Delivery Reviews,* **2001**, *46* (1-3), 3-26.
(227) Misra, A.; Ganesh, S.; Shahiwala, A.; Shah, S. P.; Drug delivery to the central nervous system: a review, *Journal of Pharmaceutical Sciences,* **2003**, *6* (2), 252-273.
(228) Mikitsh, J. L.; Chacko, A. M.; Pathways for small molecule delivery to the central nervous system across the blood-brain barrier, *Perspect Medicin Chem,* **2014**, *6*, 11-24.
(229) Armakovic, S.; Armakovic, S. J.; Atomistica.online – web application for generating input files for ORCA molecular modelling package made with the Anvil platform, *Molecular Simulation,* **2023**, *49* (1), 117-123.
(230) *Molinspiration Cheminformatics free web services*. Molinspiration Cheminformatics, https://molinspiration.com/cgi/properties (accessed 26/08/2024).
(231) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J.; A Knowledge-Based Approach in Designing Combinatorial or Medicinal Chemistry Libraries for Drug Discovery. 1. A Qualitative and Quantitative Characterization of Known Drug Databases, *J Comb Chem,* **1999**, *1*, 55-68.
(232) Andrés, A.; Rosés, M.; Ràfols, C.; Bosch, E.; Espinosa, S.; Segarra, V.; Huerta, J. M.; Setup and validation of shake-flask procedures for the determination of partition coefficients (logD) from low drug amounts, *European Journal of Pharmaceutical Sciences,* **2015**, *76*, 181-191.
(233) Helms, H. C.; Abbott, N. J.; Burek, M.; Cecchelli, R.; Couraud, P. O.; Deli, M. A.; Forster, C.; Galla, H. J.; Romero, I. A.; Shusta, E. V.; et al.; In vitro models of the blood-brain barrier: An overview of commonly used brain endothelial cell culture models and guidelines for their use, *J Cereb Blood Flow Metab,* **2016**, *36* (5), 862-890.
(234) Di, L.; Kerns, E. H.; Fan, K.; McConnell, O. J.; Carter, G. T.; High throughput artificial membrane permeability assay for blood-brain barrier, *Eur J Med Chem,* **2003**, *38* (3), 223-232.
(235) Wang, Q.; Shui, B.; Kotlikoff, M. I.; Sondermann, H.; Structural basis for calcium sensing by GCaMP2, *Structure,* **2008**, *16* (12), 1817-1827.
(236) Wang, Q.; Ding, S. L.; Li, Y.; Royall, J.; Feng, D.; Lesnar, P.; Graddis, N.; Naeemi, M.; Facer, B.; Ho, A.; et al.; The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas, *Cell,* **2020**, *181* (4), 936-953.
(237) Czuchnowski, J.; Prevedel, R.; Adaptive optics enhanced sensitivity in Fabry-Perot based photoacoustic tomography, *Photoacoustics,* **2021**, *23*, 100276.
(238) Czuchnowski, J., Establishing a high-sensitivity photoacoustic tomography system for applications in life sciences. Ruprecht-Karls-Universität Heidelberg, Heidelberg, 2021.
(239) Zhang, E.; Laufer, J.; Beard, P.; Backward-mode multiwavelength photoacoustic scanner using a planar Fabry–Perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues, *Applied Optics,* **2008**, *47* (4), 561-577.
(240) Wissmeyer, G.; Pleitez, M. A.; Rosenthal, A.; Ntziachristos, V.; Looking at sound: optoacoustics with all-optical ultrasound detection, *Light Sci Appl,* **2018**, *7*, 53.
(241) Huynh, N. T.; Zhang, E.; Francies, O.; Kuklis, F.; Allen, T.; Zhu, J.; Abeyakoon, O.; Lucka, F.; Betcke, M.; Jaros, J.; et al.; A fast all-optical 3D photoacoustic scanner for clinical vascular imaging, *Nat Biomed Eng,* **2024**.
(242) Arconada-Alvarez, S. J.; Lemaster, J. E.; Wang, J.; Jokerst, J. V.; The development and characterization of a novel yet simple 3D printed tool to facilitate phantom imaging of photoacoustic contrast agents, *Photoacoustics,* **2017**, *5*, 17-24.

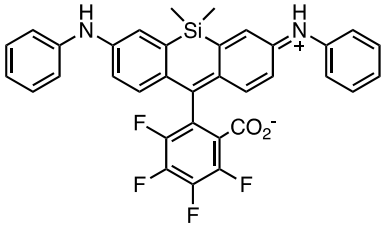
(243) Maneas, E.; Xia, W.; Ogunlade, O.; Fonseca, M.; Nikitichev, D. I.; David, A. L.; West, S. J.; Ourselin, S.; Hebden, J. C.; Vercauteren, T.; et al.; Gel wax-based tissue-mimicking phantoms for multispectral photoacoustic imaging, *Biomed Opt Express,* **2018**, *9* (3), 1151-1163.
(244) Lai, P.; Xu, X.; Wang, L. V.; Dependence of optical scattering from Intralipid in gelatin-gel based tissue-mimicking phantoms on mixing temperature and time, *J Biomed Opt,* **2014**, *19* (3), 35002.
(245) Aernouts, B.; Van Beers, R.; Watte, R.; Huybrechts, T.; Lammertyn, J.; Saeys, W.; Visible and near-infrared bulk optical properties of raw milk, *J Dairy Sci,* **2015**, *98* (10), 6727-6738.
(246) Genina, E. A.; Bashkatov, A. N.; Tuchina, D. K.; Dyachenko Timoshina, P. A.; Navolokin, N.; Shirokov, A.; Khorovodov, A.; Terskov, A.; Klimova, M.; Mamedova, A.; et al.; Optical properties of brain tissues at the different stages of glioma development in rats: pilot study, *Biomed Opt Express,* **2019**, *10* (10), 5182-5197.
(247) Huber, D.; Gutnisky, D. A.; S., P.; O'Connor, D. H.; Wiegert, J. S.; Tian, L.; Oertner, T. G.; Looger, L. L.; Svoboda, K.; Multiple dynamic representations in the motor cortex during sensorimotor learning, *Nature,* **2012**, *484* (7395), 473–478.
(248) Wallace, D. J.; Meyer zum Alten Borgloh, S.; Astori, S.; Yang, Y.; Bausen, M.; Kugler, S.; Palmer, A. E.; Tsien, R. Y.; Sprengel, R.; Kerr, J. N.; et al.; Single-spike detection in vitro and in vivo with a genetic Ca^{2+} sensor, *Nat Methods,* **2008**, *5* (9), 797-804.
(249) Costa, E. C.; de Melo-Diogo, D.; Moreira, A. F.; Carvalho, M. P.; Correia, I. J.; Spheroids Formation on Non-Adhesive Surfaces by Liquid Overlay Technique: Considerations and Practical Approaches, *Biotechnol J,* **2018**, *13* (1), 1700417.
(250) Paxinos, G., Franklin, K. B. J., *The Mouse Brain in Stereotaxic Coordinates*; Academic Press, 2001.
(251) Masch, J. M.; Steffens, H.; Fischer, J.; Engelhardt, J.; Hubrich, J.; Keller-Findeisen, J.; D'Este, E.; Urban, N. T.; Grant, S. G. N.; Sahl, S. J.; et al.; Robust nanoscopy of a synaptic protein in living mice by organic-fluorophore labeling, *Proc Natl Acad Sci U S A,* **2018**, *115* (34), E8047-E8056.
(252) Shields, B. C.; Yan, H.; Lim, S. S. X.; Burwell, S. C.; Fleming, E. A.; Cammarata, C. M.; Kahuno, E. W.; Vagadia, P. P.; Loughran, M. H.; Zhiquan, L.; et al.; Thousandfold Cell-Specific Pharmacology of Neurotransmission, *bioRxiv,* **2022**.
(253) Iliff, J. J.; Wang, M.; Zeppenfeld, D. M.; Venkataraman, A.; Plog, B. A.; Liao, Y.; Deane, R.; Nedergaard, M.; Cerebral arterial pulsation drives paravascular CSF-interstitial fluid exchange in the murine brain, *J Neurosci,* **2013**, *33* (46), 18190-18199.
(254) Mestre, H.; Hablitz, L. M.; Xavier, A. L.; Feng, W.; Zou, W.; Pu, T.; Monai, H.; Murlidharan, G.; Castellanos Rivera, R. M.; Simon, M. J.; et al.; Aquaporin-4-dependent glymphatic solute transport in the rodent brain, *Elife,* **2018**, *7*, e40070.
(255) Wu, W.; Yan, K.; He, Z.; Zhang, L.; Dong, Y.; Wu, B.; Liu, H.; Wang, S.; Zhang, F.; 2X-Rhodamine: A Bright and Fluorogenic Scaffold for Developing Near-Infrared Chemigenetic Indicators, *J Am Chem Soc,* **2024**, *146* (16), 11570-11576.
(256) Niu, H.; Liu, J.; O'Connor, H. M.; Gunnlaugsson, T.; James, T. D.; Zhang, H.; Photoinduced electron transfer (PeT) based fluorescent probes for cellular imaging and disease therapy, *Chem Soc Rev,* **2023**, *52* (7), 2322-2357.
(257) Stackova, L.; Muchova, E.; Russo, M.; Slavicek, P.; Stacko, P.; Klan, P.; Deciphering the Structure-Property Relations in Substituted Heptamethine Cyanines, *J Org Chem,* **2020**, *85* (15), 9776-9790.
(258) Choi, A.; Miller, S. C.; Reductively-labile sulfonate ester protecting groups that are rapidly cleaved by physiological glutathione, *Organic & Biomolecular Chemistry,* **2017**, *15* (6), 1346-1349.
(259) St John-Campbell, S.; White, A. J. P.; Bull, J. A.; Single operation palladium catalysed C(sp(3))-H functionalisation of tertiary aldehydes: investigations into transient imine directing groups, *Chem Sci,* **2017**, *8* (7), 4840-4847.
(260) Akiyama, T. D., K.; Fuchibe, K.; Cu(I)-Catalyzed Enantioselective [2 + 2] Cycloaddition of 1-Methoxyallenylsilane with α-Imino Ester: Chiral Synthesis of α,β-Unsaturated Acylsilanes, *Organic Letters,* **2003**, *5* (20), 3691-3693.
(261) Hara, D. M., T.; Kazuta, Y.; Norimine, Y.; Amino, H.; Suganuma, M., Fujiyama, S.; Yamashita, K.; Okada, M.; Nishikawa, Y.; Iwanaga, S.; Fluorescent Dye and Use Thereof, *The United States Patent and Trademark Office,* **2021**, *US 2021/0255191 A1*.
(262) Bering, L.; Jeyakumar, K.; Antonchick, A. P.; Metal-Free C-O Bond Functionalization: Catalytic Intramolecular and Intermolecular Benzylation of Arenes, *Org Lett,* **2018**, *20* (13), 3911-3914.
(263) Treeby, B. E.; Cox, B. T.; k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, *J Biomed Opt,* **2010**, *15* (2), 021314.
(264) Carvalho, M. P.; Costa, E. C.; Correia, I. J.; Assembly of breast cancer heterotypic spheroids on hyaluronic acid coated surfaces, *Biotechnol Prog,* **2017**, *33* (5), 1346-1357.

Chapter VIII: Appendix

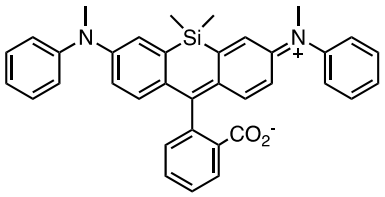
VIII.1 Synthetic Procedures and Characterisations



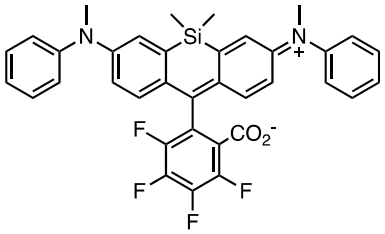
(1): Si-fluorescein ditriflate (40.1 mg, 0.063 mmol, 1 eq), Pd_2dba_3 (5.7 mg, 0.006 mmol, 0.1 eq), XPhos (9.0 mg, 0.019 mmol, 0.3 eq) and Cs_2CO_3 (57.1 mg, 0.175 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (1 mL) and subsequently aniline (17 µL, 0.150 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 3 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with CH_2Cl_2 and the solvent was evaporated under reduced pressure. Purification was performed by silica gel column chromatography (5% MeOH/CH_2Cl_2) followed by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were partially concentrated to remove CH_3CN then lyophilised overnight to afford the title compound as a blue-green solid (26.0 mg, 65%, TFA salt). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.20 (br s, 2H), 8.09 (d, J = 7.6 Hz, 1H), 7.71 (t, J = 7.6 Hz, 1H), 7.60 (t, J = 7.6 Hz, 1H), 7.38 – 7.28 (m, 7H), 7.15 (d, J = 7.7 Hz, 4H), 7.07 (t, J = 7.7 Hz, 2H), 6.91 – 6.83 (m, 4H), 0.58 (s, 3H), 0.47 (s, 3H). ^{13}C NMR ($CDCl_3$, 101 MHz) δ 170.7 (C), 153.6 (C), 143.0 (C), 142.3 (C), 137.8 (C), 136.0 (C), 134.0 (CH), 129.6 (CH), 129.2 (CH), 128.6 (CH), 126.9 (C), 126.2 (CH), 124.9 (CH), 122.1 (CH), 121.9 (CH), 118.6 (CH), 117.8 (CH), 77.4 (C), 0.3 (Si-CH_3), -1.6 (Si-CH_3). Analytical HPLC: t_R = 5.4 min, 98% purity (5–95% CH_3CN /H_2O, gradient with constant 0.1% v/v formic acid additive, 10 min run, 0.5 mL/min flow, detection at 254 nm). HRMS (ESI) calcd for $C_{34}H_{29}N_2O_2Si$ $[M+H]^+$ 525.1993, found 525.2001.



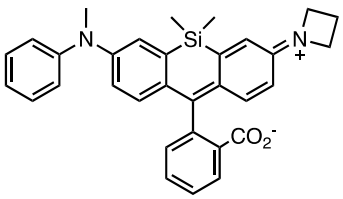
(2): F_4-Si-fluorescein ditriflate (42.6 mg, 0.060 mmol, 1 eq), Pd_2dba_3 (5.5 mg, 0.006 mmol, 0.1 eq), XPhos (8.6 mg, 0.018 mmol, 0.3 eq) and Cs_2CO_3 (54.8 mg, 0.168 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (1 mL) and subsequently aniline (13 µL, 0.144 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 19 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc and the solvent was evaporated under reduced pressure. Purification was performed by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–15% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (30–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were partially concentrated to remove CH_3CN then lyophilised overnight to afford the title compound as a blue solid (0.89 mg, 2%, TFA salt). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 7.71 (s, 2H), 7.51 (d, J = 2.6 Hz, 2H), 7.32 – 7.25 (m, 4H), 7.22 – 7.16 (m, 4H), 7.07 (dd, J = 8.7, 2.6 Hz, 2H), 6.99 (dd, J = 8.6, 1.6 Hz, 2H), 6.93 (tt, J = 7.3, 1.2 Hz, 2H), 0.59 (s, 3H), 0.54 (s, 3H). ^{19}F NMR (376 MHz, $(CD_3)_2CO$) δ -139.83 – -139.97 (m, 1F), -141.76 – -141.91 (m, 1F), -146.03 – -146.18 (m, 1F), -154.05 – -154.20 (m, 1F). Analytical HPLC: t_R = 4.8 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% v/v formic acid additive, 8 min run, 0.5 mL/min flow, detection at 254 nm). HRMS (ESI) calcd for $C_{34}H_{24}F_4N_2O_2Si$ $[M+H]^+$ 597.1616, found 597.1600.



(3): Si-fluorescein ditriflate (38.5 mg, 0.060 mmol, 1 eq), Pd_2dba_3 (5.5 mg, 0.006 mmol, 0.1 eq), XPhos (8.6 mg, 0.018 mmol, 0.3 eq) and Cs_2CO_3 (54.9 mg, 0.169 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (1 mL) and subsequently *N*-methylaniline (16 µL, 0.144 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 3 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with CH_2Cl_2 and the solvent was evaporated under reduced pressure. Purification by silica gel column chromatography (10% EtOAc/cyclohexane) afforded the title compound as a blue-green solid (22.5 mg, 68%). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.98 (d, *J* = 7.6 Hz, 1H), 7.68 (t, *J* = 7.6 Hz, 1H), 7.57 (t, *J* = 7.6 Hz, 1H), 7.38 (d, *J* = 7.6 Hz, 1H), 7.30 (t, *J* = 7.7 Hz, 4H), 7.25 (d, *J* = 2.8 Hz, 2H), 7.09 (d, *J* = 7.7 Hz, 4H), 7.04 (t, *J* = 7.7 Hz, 2H), 6.81 (d, *J* = 8.8 Hz, 2H), 6.77 (dd, *J* = 8.8, 2.8 Hz, 2H), 3.33 (s, 6H), 0.57 (s, 3H), 0.54 (s, 3H). ^{13}C NMR ($(CD_3)_2CO$, 101 MHz) δ 170.4 (C), 155.0 (C), 149.3 (C), 149.1 (C), 137.4 (C), 136.5 (C), 135.1 (CH), 130.3 (CH), 130.1 (CH), 128.5 (CH), 127.1 (C), 126.3 (CH), 125.5 (CH), 123.8 (CH), 123.5 (CH), 123.4 (CH), 119.9 (CH), 91.3 (C), 40.3 (N-CH_3), 0.2 (Si-CH_3), -1.5 (Si-CH_3). Analytical HPLC: t_R = 4.7 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% v/v formic acid additive, 10 min run, 0.5 mL/min flow, detection at 254 nm). HRMS (ESI) calcd for $C_{36}H_{33}N_2O_2Si$ $[M+H]^+$ 553.2306, found 553.2314.

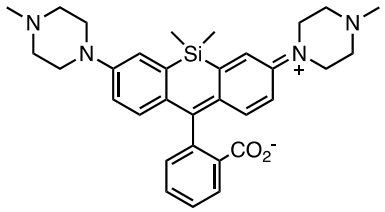


(4): F_4-Si-fluorescein ditriflate (33.5 mg, 0.047 mmol, 1 eq), Pd_2dba_3 (4.3 mg, 0.005 mmol, 0.1 eq), XPhos (6.8 mg, 0.014 mmol, 0.3 eq) and Cs_2CO_3 (43.1 mg, 0.132 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (1 mL) and subsequently N-methylaniline (12 µL, 0.113 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 3 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with CH_2Cl_2 and the solvent was evaporated under reduced pressure. Purification was performed by silica gel column chromatography (0–10% EtOAc/cyclohexane) followed by reverse phase HPLC (30–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were partially concentrated to remove CH_3CN then lyophilised overnight to afford the title compound as a blue-green solid (7.7 mg, 22%, TFA salt). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.39 (t, *J* = 7.9 Hz, 4H), 7.22 – 7.15 (m, 6H), 7.14 (d, *J* = 2.8 Hz, 2H), 6.86 (d, *J* = 8.9 Hz, 2H), 6.74 (dd, *J* = 8.9, 2.8 Hz, 2H), 3.42 (s, 6H), 0.47 (s, 3H), 0.45 (s, 3H). ^{19}F NMR ($CDCl_3$, 376 MHz) δ -137.5 – -137.7 (m, 1F), -138.6 (td, *J* = 19.6, 3.6 Hz, 1F), -145.7 – -146.0 (m, 1F), -151.6 (td, *J* = 19.6 Hz, 3.6 Hz, 1F). HRMS (ESI) calcd for $C_{36}H_{29}F_4N_2O_2Si$ $[M+H]^+$ 625.1929, found 625.1920.

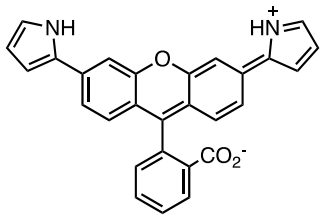


(5): Si-fluorescein ditriflate (100 mg, 0.157 mmol, 1 eq), $Pd_2(dba)_3$ (14.4 mg, 0.016 mmol, 0.1 eq), XantPhos (27.3 mg, 0.048 mmol, 0.3 eq) and Cs_2CO_3 (102.4 mg, 0.314 mmol, 2.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (2 mL) and subsequently *N*-methylaniline (17.0 µL, 0.157 mmol, 1 eq) were added then the reaction mixture was stirred at 80°C for 2 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc and the solvent was evaporated under reduced pressure. Partial purification

was performed by silica gel column chromatography (cyclohexane:CH_2Cl_2:toluene (2:2:1)) and afforded the intermediate as a yellow-green solid. The intermediate showed moderate stability and was directly engaged in the next step.
The intermediate (~40 mg, 0.067 mmol), Pd_2dba_3 (6.1 mg, 0.007 mmol, 0.1 eq), XPhos (9.5 mg, 0.020 mmol, 0.3 eq) and Cs_2CO_3 (60.7 mg, 0.186 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (1 mL) and subsequently azetidine (11 µL, 0.186 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 3 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc and the solvent was evaporated under reduced pressure. The crude material was purified by silica gel column chromatography (10–15% TBME/cyclohexane, linear gradient) followed by reverse phase HPLC (10–50% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined and lyophilised to afford the title compound as a blue-green solid (10.7 mg, 11% yield over 2 steps, TFA salt). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.99 (d, *J* = 7.6 Hz, 1H), 7.69 (td, *J* = 7.5, 1.1 Hz, 1H), 7.58 (td, *J* = 7.5, 1.1 Hz, 1H), 7.36 – 7.28 (m, 3H), 7.19 (dd, *J* = 14.7, 2.7 Hz, 2H), 7.14 – 7.03 (m, 3H), 6.97 (d, *J* = 8.8 Hz, 1H), 6.82 (d, *J* = 8.8 Hz, 1H), 6.72 – 6.77 (m, 2H), 4.24 (t, *J* = 7.7 Hz, 4H), 3.33 (s, 3H), 2.54 (p, *J* = 7.7 Hz, 2H), 0.592 (s, 3H), 0.587 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.2 (C), 152.8 (C), 148.6 (C), 148.1 (C), 146.3 (C), 139.3 (C), 139.2 (C), 137.1 (C), 134.4 (C), 134.0 (CH), 129.6 (CH), 129.4 (CH), 129.0 (CH), 128.7 (CH), 127.0 (C), 126.5 (CH), 125.0 (CH), 123.6 (CH), 123.2 (CH), 122.3 (CH), 120.7 (CH), 118.8 (CH), 116.6 (CH), 77.4 (C), 54.8 (CH_2), 40.2 (CH_3), 16.8 (CH_2), 0.3 (Si-CH_3), -1.7 (Si-CH_3). Analytical HPLC: t_R = 5.2 min, 95% purity, (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{32}H_{31}N_2O_2Si$ $[M+H]^+$ 503.2149, found 503.2145.

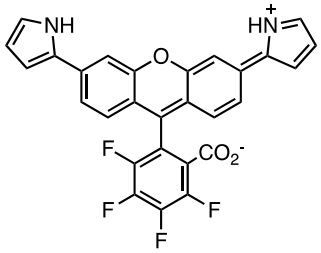


(6): Si-fluorescein ditriflate (41.0 mg, 0.048 mmol, 1 eq), $Pd(OAc)_2$ (2.2 mg, 0.010 mmol, 0.2 eq), BINAP (9.1 mg, 0.015 mmol, 0.3 eq) and Cs_2CO_3 (44.4 mg, 0.136 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. Toluene (1 mL) and subsequently 1-methylpiperazine (13 µL, 0.117 mmol, 2.4 eq) were added and the reaction mixture was stirred at 100°C for 48 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with MeOH and the solvent was evaporated under reduced pressure. Purification by silica gel column chromatography (10% MeOH/CH_2Cl_2) afforded the title compound as a pale green solid (23.8 mg, 92%). ^{1}H NMR (400 MHz, CD_3CN) δ 7.93 (d, *J* = 7.6 Hz, 1H), 7.71 (td, *J* = 7.5, 1.2 Hz, 1H), 7.61 (td, *J* = 7.5, 0.9 Hz, 1H), 7.28 – 7.32 (m, 2H), 7.22 (d, *J* = 7.7 Hz, 1H), 6.83 – 6.89 (m, 4H), 3.82 (br s, 4H), 3.49 (br s, 4H), 3.12 (m, 8H), 2.80 (s, 6H), 0.65 (s, 3H), 0.56 (s, 3H). ^{13}C NMR (CD_3CN, 101 MHz) δ 171.2 (CO), 155.6 (C), 149.9 (C), 137.5 (C), 136.6 (C), 135.7 (CH), 130.3 (CH), 129.0 (CH), 126.6 (CH), 126.4 (C), 125.0 (CH), 121.7 (CH), 118.5 (CH), 91.1 (C), 53.7 (CH_2), 46.5 (CH_2), 43.5 (CH_3), 0.0 (Si-CH_3), -1.1 (Si-CH_3). Analytical HPLC: t_R = 2.5 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{32}H_{39}N_4O_2Si$ $[M+H]^+$ 539.2837, found 539.2825.

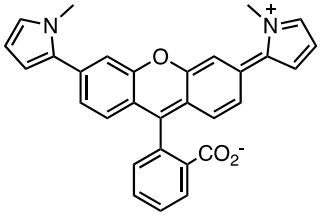


(9): Fluorescein ditriflate (92.4 mg, 0.155 mmol, 1 eq), *N*-Boc-pyrrole-2-boronic acid pinacol ester (136 mg, 0.465 mmol, 3.0 eq), $Pd(PPh_3)_4$ (10.7 mg, 0.009 mmol, 0.06 eq) and KOAc (91.2 mg, 0.930 mmol, 6.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. A mixture of 1,4-dioxane/H_2O (2/1, 2.85 mL) was added and the reaction mixture was stirred at 100°C for 23 h. After cooling to room temperature, the reaction mixture was extracted with $CHCl_3$. The organic layers were combined, washed with brine, dried over Na_2SO_4 and the solvent was evaporated under

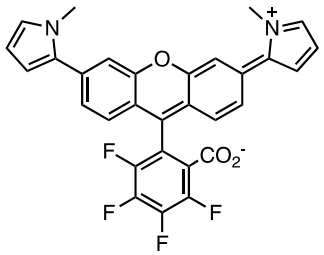
reduced pressure. Silica gel column chromatography (Biotage Sfär Duo 5 g, 0–30% CH_2Cl_2/cyclohexane, linear gradient) afforded the title compound as a blue solid (37.7 mg, 56%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.87 (s, 2H), 8.06 – 8.00 (m, 1H), 7.69 – 7.58 (m, 2H), 7.31 (d, *J* = 1.8 Hz, 2H), 7.16 – 7.07 (m, 3H), 6.87 – 6.91 (m, 2H), 6.73 (d, *J* = 8.3 Hz, 2H), 6.58 – 6.52 (m, 2H), 6.28 (q, *J* = 2.8 Hz, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.1 (CO), 153.1 (C), 151.8 (C), 135.6 (C), 135.4 (CH), 130.7 (C), 130.0 (CH), 128.6 (CH), 126.4 (C), 125.4 (CH), 124.2 (CH), 120.3 (CH), 119.6 (CH), 116.0 (C), 111.5 (CH), 110.5 (CH), 107.7 (CH), 75.2 (C). Analytical HPLC: t_R = 4.3 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{28}H_{19}N_2O_3$ $[M+H]^+$ 431.1390, found 431.1388.



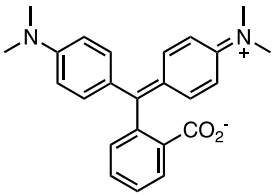
(10): F_4-Fluorescein ditriflate (30.8 mg, 0.046 mmol, 1 eq), *N*-Boc-pyrrole-2-boronic acid pinacol ester (40.6 mg, 0.139 mmol, 3.0 eq), $Pd(PPh_3)_4$ (3.2 mg, 0.003 mmol, 0.06 eq) and KOAc (27.2 mg, 0.277 mmol, 6.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. A mixture of 1,4-dioxane/H_2O (2/1, 1.05 mL) was added and the reaction mixture was stirred at 100°C for 46 h. After cooling to room temperature, the reaction mixture was extracted with $CHCl_3$. The organic layers were combined, washed with brine, dried over Na_2SO_4 and the solvent was evaporated under reduced pressure. Silica gel column chromatography (30% EtOAc/cyclohexane) afforded the title compound as a blue solid (15.0 mg, 65%). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 10.72 (br s, 2H), 7.62 (d, *J* = 1.8 Hz, 2H), 7.48 (dd, *J* = 8.4, 1.8 Hz, 2H), 7.22 (d, *J* = 8.3 Hz, 2H), 6.96 (td, *J* = 2.7, 1.4 Hz, 2H), 6.70 – 6.73 (m, 2H), 6.20 – 6.24 (m, *J* = 3.5, 2.4 Hz, 2H). ^{19}F NMR (376 MHz, $(CD_3)_2CO$) δ -139.98 (td, *J* = 19.6, 9.3 Hz, 1F), -143.84 (td, *J* = 19.3, 4.3 Hz, 1F), -144.45 (td, *J* = 18.8, 9.0 Hz, 1F), -152.37 (td, *J* = 18.6, 4.3 Hz, 1F). Analytical HPLC: t_R = 4.9 min, 98% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 10 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{28}H_{15}F_4N_2O_3$ $[M+H]^+$ 503.1013, found 503.1017.



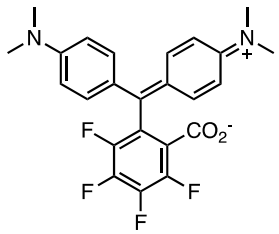
(11): Fluorescein ditriflate (77.7 mg, 0.130 mmol, 1 eq), 1-methyl-2-pyrroleboronic acid pinacol ester (81.0 mg, 0.391 mmol, 3.0 eq), $Pd(PPh_3)_4$ (9.0 mg, 0.007 mmol, 0.06 eq) and KOAc (76.7 mg, 0.781 mmol, 6.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. A mixture of 1,4-dioxane/H_2O (2/1, 2.4 mL) was added and the reaction mixture was stirred at 100°C for 23 h. After cooling to room temperature, the reaction mixture was extracted with CH_2Cl_2. The organic layers were combined, washed with brine, dried over Na_2SO_4 and the solvent was evaporated under reduced pressure. Silica gel column chromatography (Biotage Sfär Duo 5 g, 0–20% EtOAc/cyclohexane, linear gradient) afforded the title compound as a grey solid (49.1 mg, 82%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.07 (d, *J* = 7.5 Hz, 1H), 7.75 – 7.63 (m, 2H), 7.35 (d, *J* = 1.7 Hz, 2H), 7.28 (d, *J* = 7.6 Hz, 1H), 7.10 (dd, *J* = 8.2, 1.7 Hz, 2H), 6.85 (d, *J* = 8.3 Hz, 2H), 6.75 (t, *J* = 2.3 Hz, 2H), 6.29 – 6.33 (m, 2H), 6.22 (t, *J* = 3.1 Hz, 2H), 3.73 (s, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.5 (CO), 153.2 (C), 151.4 (C), 136.1 (C), 135.3 (CH), 133.2 (C), 130.1 (CH), 128.2 (CH), 126.7 (C), 125.4 (CH), 125.0 (CH), 124.2 (CH), 124.1 (CH), 117.0 (C), 116.4 (CH), 109.9 (CH), 108.3 (CH), 82.7 (C), 35.5 (CH_3). Analytical HPLC: t_R = 4.7 min, 98% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{30}H_{23}N_2O_3$ $[M+H]^+$ 459.1703, found 459.1701.



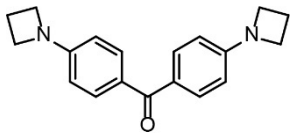
(12): F_4-Fluorescein ditriflate (32.4 mg, 0.046 mmol, 1 eq), 1-methyl-2-pyrroleboronic acid pinacol ester (40.6 mg, 0.139 mmol, 3.0 eq), $Pd(PPh_3)_4$ (3.2 mg, 0.003 mmol, 0.06 eq) and KOAc (27.2 mg, 0.277 mmol, 6.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. A mixture of 1,4-dioxane/H_2O (2/1, 1.05 mL) was added and the reaction mixture was stirred at 100°C for 22 h. After cooling to room temperature, the reaction mixture was extracted with CH_2Cl_2. The organic layers were combined, washed with brine, dried over Na_2SO_4 and the solvent was evaporated under reduced pressure. Silica gel column chromatography (15% EtOAc/cyclohexane) afforded the title compound as an off-white solid (21.4 mg, 88%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$, 5 mM DMF) δ 7.51 (t, J = 1.1 Hz, 2H), 7.31 (d, J = 1.1 Hz, 4H), 6.93 (dd, J = 2.7, 1.8 Hz, 2H), 6.38 (dd, J = 3.7, 1.8 Hz, 2H), 6.11 (dd, J = 3.7, 2.6 Hz, 2H), 3.75 (s, 6H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -137.72 – -138.11 (m, 1F), -141.36 – -141.74 (m, 2F), -149.48 – -149.81 (m, 1F). Analytical HPLC: t_R = 5.4 min, 98% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 10 min run, 0.5 mL/ min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{30}H_{19}F_4N_2O_3$ $[M+H]^+$ 531.1326, found 531.1331.



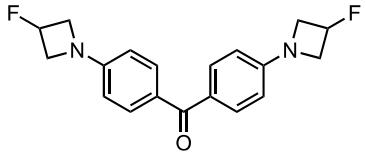
(13): *N,N*-dimethylaniline (1.03 mL, 8.10 mmol, 2 eq), phthalic anhydride (600 mg, 4.05 mmol, 1 eq) and $ZnCl_2$ (1.14 mg, 8.10 mmol, 2 eq) were loaded into a sealed vial and stirred vigorously at 150°C for 5 h. The reaction mixture was cooled to room temperature then dissolved progressively in organic solvents (CH_3CN and EtOAc). Aqueous HCl (1 M, 10 mL) was added and the organic solvents evaporated under reduced pressure. The aqueous phase was neutralised to pH = 5-6 with NaOH (1M), which formed a green precipitate. The precipitate was dissolved with EtOAc. The layers were separated and the aqueous layer was extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Trituration with cyclohexane to remove excess *N,N*-dimethylaniline followed by purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–30% EtOAc/cyclohexane, linear gradient) afforded the title compound as a yellow solid (271 mg, 18%). For characterisation by NMR and UV/Vis spectroscopy, purification of an analytical fraction was performed by reverse phase HPLC (10–50% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, and lyophilised to afford a fraction of the pure compound as a TFA salt. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.94 (d, J = 7.6 Hz, 1H), 7.75 (t, J = 7.6, 1H), 7.61 (t, J = 7.6 Hz, 1H), 7.55 (d, J = 7.6 Hz, 1H), 7.41 – 7.35 (m, 8H), 3.13 (s, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.3 (C), 150.8 (C), 145.1 (C), 139.1 (C), 134.9 (CH), 130.2 (CH), 129.1 (CH), 126.5 (CH), 125.3 (C), 124.1 (CH), 119.0 (CH), 90.5 (C), 45.0 (CH_3). Analytical HPLC: t_R = 4.2 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{24}H_{21}N_2O_2$ $[M+H]^+$ 373.1911, found 373.1907.



(14): *N,N*-dimethylaniline (2.00 g, 16.53 mmol, 2 eq), tetrafluorophthalic anhydride (1.82 g, 8.27 mmol, 1 eq) and $ZnCl_2$ (2.25 g, 8.27 mmol, 2 eq) were loaded into a sealed vial and stirred vigorously at 150°C for 4 h. The reaction mixture was cooled to room temperature then dissolved progressively in organic solvents (CH_3CN and EtOAc). Aqueous HCl (1 M, 50 mL) was added and the organic solvents evaporated under reduced pressure. The aqueous layer was neutralised to pH = 5-6 with NaOH (1M), which formed a green precipitate. The precipitate was dissolved with EtOAc. The layers were separated and the aqueous layer was extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 50 g, 0-30% EtOAc/cyclohexane) afforded the title compound as a blue-green solid (746 mg, 20%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.15 (d, J = 8.6 Hz, 4H), 6.65 (d, J = 8.6 Hz, 4H), 2.96 (s, 12H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -137.87 (t, J = 20.2 Hz, 1F), -138.82 – -139.06 (m, 1F), -143.33 – -143.59 (m, 1F), -151.88 – -152.18 (m, 1F). Analytical HPLC: t_R = 4.7 min, 98% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{24}H_{21}F_4N_2O_2$ $[M+H]^+$ 445.1534, found 445.1530.

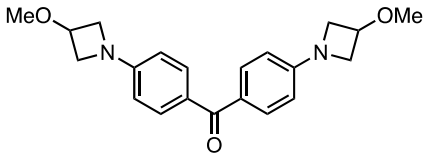


(S1): 4,4'-Dihydroxybenzophenone ditriflate (1.00 g, 2.10 mmol, 1 eq), $Pd_2(dba)_3$ (193 mg, 0.210 mmol, 0.1 eq), XPhos (301 mg, 0.631 mmol, 0.3 eq) and Cs_2CO_3 (1.92 g, 5.89 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (8 mL) and subsequently azetidine (340 µL, 5.05 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 5 h. After cooling to room temperature, the mixture was filtered through Celite, washed with EtOAc, and the solvent evaporated under reduced pressure. Purification by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–30% EtOAc/cyclohexane, linear gradient) followed by recrystallisation from MeOH afforded the title compound as a yellow solid (333 mg, 54%). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 7.62 (d, J = 8.6 Hz, 4H), 6.43 (d, J = 8.6 Hz, 4H), 3.98 (t, J = 7.3 Hz, 8H), 2.42 (p, J = 7.3 Hz, 4H). ^{13}C NMR (101 MHz, $(CD_3)_2CO$) δ 193.4 (C), 155.0 (C), 132.3 (CH), 127.8 (C), 110.3 (CH), 52.4 (CH_2), 17.2 (CH_2). Analytical HPLC: t_R = 4.1 min, 97% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{19}H_{21}N_2O$ $[M+H]^+$ 293.1648, found 293.1643.

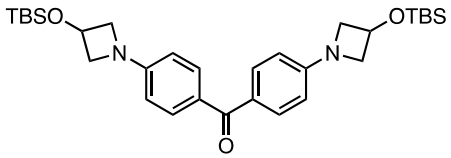


(S2): 4,4'-Dihydroxybenzophenone ditriflate (1.50 g, 3.13 mmol, 1 eq), $Pd_2(dba)_3$ (287 mg, 0.313 mmol, 0.1 eq), XPhos (448 mg, 0.939 mmol, 0.3 eq), Cs_2CO_3 (5.30 g, 16.28 mmol, 5.2 eq) and 3-fluoroazetidine hydrochloride (838 mg, 7.52 mmol, 2.4 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (10 mL) was added and the reaction mixture was stirred at 100°C for 4 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc and the solvent evaporated under reduced pressure. Purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–30% EtOAc/cyclohexane, linear gradient) followed by recrystallisation from MeOH afforded the title compound as a pale-yellow solid (833 mg, 81%). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 7.65 (d, J = 8.7 Hz, 4H), 6.54 (d, J = 8.7 Hz, 4H), 5.56 (dtt, $^2J_{HF}$ = 57.3 Hz, J = 8.8, 3.3 Hz, 2H), 4.39 – 4.29 (m, 4H), 4.11 – 4.00 (m, 4H). ^{13}C NMR (101 MHz, $(CD_3)_2CO$) δ 193.5 (C), 154.3 (d, $^4J_{CF}$ = 1.5 Hz, C), 132.4 (CH), 128.6 (C), 119.8 (C), 116.5 (C), 111.2 (CH), 84.1 (d, $^1J_{CF}$ = 203.0 Hz, CFH), 59.9 (d, $^2J_{CF}$ = 24.2 Hz, CH_2). ^{19}F NMR (376 MHz, $(CD_3)_2CO$) δ -179.83

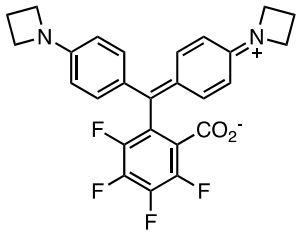
– -179.86 (m, 2F). Analytical HPLC: t_R = 4.3 min, 95% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{19}H_{19}N_2O_2$ $[M+H]^+$ 329.1460, found 329.1454.



(S3): 4,4'-Dihydroxybenzophenone ditriflate (1.01 g, 2.12 mmol, 1 eq), $Pd_2(dba)_3$ (0.212 g, 0.212 mmol, 0.1 eq), XPhos (0.303 g, 0.636 mmol, 0.3 eq), Cs_2CO_3 (3.59 g, 11.02 mmol, 5.2 eq) and 3-methoxyazetidine hydrochloride (0.628 g, 5.085 mmol, 2.4 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (8 mL) was added then the reaction mixture was stirred at 100°C for 4 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc, and the solvent evaporated under reduced pressure. Purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–50% EtOAc/cyclohexane, linear gradient) afforded the title compound as a yellow solid (679 mg, 91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (d, J = 8.5 Hz, 4H), 6.43 (d, J = 8.5 Hz, 4H), 4.40 – 4.34 (m, 2H), 4.09 – 4.03 (m, 4H), 3.85 – 3.80 (m, 4H), 3.35 (s, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.3 (CO), 153.5 (C), 132.2 (CH), 127.7 (C), 110.2 (CH), 70.0 (CH), 58.6 (CH_2), 56.3 (CH_3). Analytical HPLC: t_R = 3.8 min, 97% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{21}H_{25}N_2O_3$ $[M+H]^+$ 353.1860, found 353.1853.

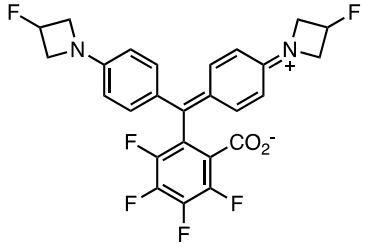


(S5): 4,4'-Dihydroxybenzophenone ditriflate (1.09 g, 2.29 mmol, 1 eq), $Pd_2(dba)_3$ (209 mg, 0.229 mmol, 0.1 eq), XPhos (327 mg, 0.686 mmol, 0.3 eq), Cs_2CO_3 (4.32 g, 13.3 mmol, 6 eq) and 3-hydroxyazetidine hydrochloride (642 mg, 6.86 mmol, 3 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (10 mL) was added then the reaction mixture was stirred at 100°C for 4 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc and MeOH, and the solvents evaporated under reduced pressure. Partial purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–50% EtOAc/cyclohexane, linear gradient) afforded the hydroxy intermediate **S4** as an orange-red solid. The intermediate was dissolved in CH_2Cl_2 (80 mL), cooled to 0°C, and imidazole (1.47 g, 21.6 mmol, 6 eq) followed by *tert*-butyldimethylsilyl chloride (1.63 g, 10.8 mmol, 3 eq) were added. The reaction mixture was stirred at room temperature under argon for 14 h, after which it was washed with H_2O and brine. The organic layer was dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness. Purification by silica gel chromatography (Biotage Sfär Duo 25 g, 0–15% EtOAc/cyclohexane, linear gradient) followed by recrystallisation from MeOH afforded the title compound as a pale-yellow solid (268 mg, 21% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.70 (d, J = 8.5 Hz, 4H), 6.43 (d, J = 8.5 Hz, 4H), 4.82 – 4.74 (m, 2H), 4.27 – 4.18 (m, 4H), 3.79 – 3.74 (m, 4H), 0.90 (s, 18H), 0.09 (s, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.3 (CO), 153.6 (C), 132.2 (CH), 127.7 (C), 110.5 (CH), 62.4 (CH_2), 61.9 (CH_2), 25.9 (CH_3), 18.1 (C), -4.8 (CH_3). HRMS (ESI) calcd for $C_{31}H_{49}N_2O_3Si_2$ $[M+H]^+$ 553.3276, found 553.3267.

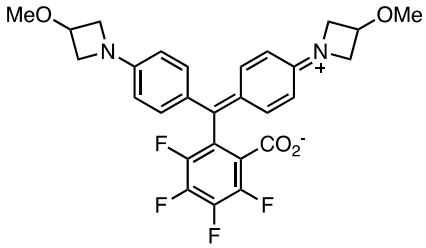


(15): A solution of 2,3,4,5-tetrafluorobenzoic acid (2.01 g, 10.3 mmol, 10 eq) in anhydrous THF (20 mL) was cooled to -78°C under argon. *n*-Butyllithium (2.5 M in hexanes, 8.27 mL, 20.6 mmol, 20 eq) was added, and the reaction was stirred at -78°C for 3 h. **S1** (302 mg, 1.03 mmol, 1 eq) in anhydrous THF (20 mL) was added, the reaction was allowed to

warm to room temperature and stirred for 16 h. The reaction mixture was subsequently diluted with saturated NH_4Cl and H_2O and extracted with EtOAc (×3). The combined organic layers were washed with saturated $NaHCO_3$ and brine, dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness. Purification by silica gel chromatography (Biotage Sfär Duo 25 g, 0–30% TBME/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (360 mg, 74%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$, 5 mM DMF) δ 7.03 (d, J = 8.4 Hz, 4H), 6.42 – 6.37 (d, J = 8.4 Hz, 4H), 3.81 (t, J = 7.2 Hz, 8H), 2.30 (p, J = 7.2 Hz, 4H). ^{19}F NMR (376 MHz, $(CD_3)_2CO$) δ -138.94 (t, J = 20.0 Hz, 1F), -140.60 (td, J = 19.3, 8.8 Hz, 1F), -144.83 (ddd, J = 20.9, 18.0, 8.8 Hz, 1F), -153.52 (ddd, J = 20.4, 17.9, 4.5 Hz, 1F). Analytical HPLC: t_R = 4.7 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{26}H_{21}F_4N_2O_2$ $[M+H]^+$ 469.1534, found 469.1525.

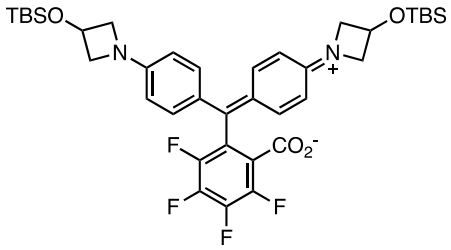


(16): A solution of 2,3,4,5-tetrafluorobenzoic acid (1.21 g, 6.24 mmol, 10 eq) in anhydrous THF (15 mL) was cooled to -78°C under argon. *N*-Butyllithium (2.5 M in hexanes, 5.00 mL, 12.5 mmol, 20 eq) was added, and the reaction was stirred at -78°C for 3 h. A solution of **S2** (205 mg, 0.624 mmol, 1 eq) in anhydrous THF (15 mL) was added, and the reaction was allowed to warm to room temperature and stirred for 18 h. The reaction mixture was subsequently diluted with saturated NH_4Cl and water and then extracted with EtOAc (×3). The combined organic layers were washed with saturated $NaHCO_3$ and brine, dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness. Purification by silica gel chromatography (Biotage Sfär Duo 25 g, 0–30% TBME/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (232 mg, 74%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$, 5 mM DMF) δ 7.07 (d, J = 8.5 Hz, 4H), 6.48 (d, J = 8.5 Hz, 4H), 5.48 (dtt, $^2J_{HF}$ = 57.5 Hz, J = 8.8, 3.0 Hz, 2H), 4.23 – 4.09 (m, 4H), 3.95 – 3.85 (m, 4H). ^{19}F NMR (376 MHz, $(CD_3)_2CO$) δ -138.83 – -139.07 (m, 1F), -139.52 – -139.76 (m, 1F), -142.56 – -151.87 (m, 1F), -151.51 – -151.81 (m, 1F), -178.49 – -178.93 (m, 2F). Analytical HPLC: t_R = 4.4 min, 97% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{26}H_{19}F_4N_2O_2$ $[M+H]^+$ 505.1345, found 505.1338.

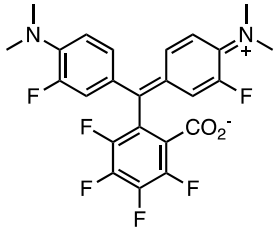


(17): A solution of 2,3,4,5-tetrafluorobenzoic acid (3.45 g, 17.8 mmol, 10 eq) in anhydrous THF (42 mL) was cooled to -78°C under argon. *n*-Butyllithium (2.5 M in hexanes, 14.2 mL, 35.5 mmol, 20 eq) was added, and the reaction was stirred at -78°C for 3 h. A solution of **S3** (636 mg, 1.78 mmol, 1 eq) in anhydrous THF (35 mL) was added, the reaction was allowed to warm to room temperature and stirred for 18 h. The reaction mixture was subsequently diluted with saturated NH_4Cl and water and then extracted with EtOAc (3x). The combined organic layers were washed with saturated $NaHCO_3$ and brine, dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness. Purification by silica gel chromatography (Biotage Sfär Duo 25 g, 0–30% EtOAc/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (656 mg, 77%). For full characterisation by NMR and UV/Vis spectroscopy an analytical fraction of the product was further purified by reverse phase HPLC (10–100% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness to afford a pure sample. ^{1}H NMR (400 MHz, CD_3CN) δ 7.12 – 7.04 (m, 4H), 6.46 – 6.38 (m, 4H), 4.30 (tt, J = 6.1, 4.2 Hz, 2H), 4.07 – 4.02 (m), 3.68 – 3.63 (m, 4H), 3.27 (s, 6H). ^{19}F NMR (376 MHz, $(CD_3)_2SO$) δ -138.89 – -139.20 (m, 1F), -139.83 – -140.08 (m, 1F), -142.91 – -143.44 (m, 1F), -151.96 – -152.41 (m, 1F). Analytical HPLC: t_R = 4.5 min, >99 % purity (5–95% CH_3CN/H_2O, gradient

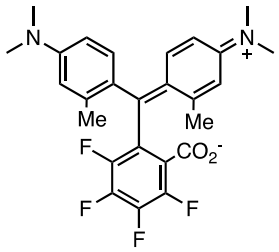
with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{28}H_{25}F_4N_2O_4$ $[M+H]^+$ 529.1745, found 529.1735.



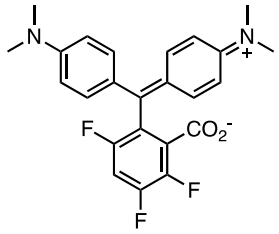
(18): A solution of 2,3,4,5-tetrafluorobenzoic acid (941 mg, 4.85 mmol, 10 eq) in anhydrous THF (10 mL) was cooled to -78°C under argon. *n*-Butyllithium (2.5 M in hexanes, 3.88 mL, 9.70 mmol, 20 eq) was added, and the reaction was stirred at -78°C for 3 h. A solution of **S4** (268 mg, 0.485 mmol, 1 eq) in anhydrous THF (20 mL) was added, the reaction was allowed to warm to room temperature and stirred for 17 h. The reaction mixture was subsequently diluted with saturated NH_4Cl and water and then extracted with EtOAc (3x). The combined organic layers were washed with saturated $NaHCO_3$ and brine, dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness. Purification by silica gel chromatography (Biotage Sfär Duo 25 g, 0–15% EtOAc/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (155 mg, 44%). For full characterisation by NMR and UV/Vis spectroscopy an analytical fraction of the product was further purified by column chromatography (Biotage Sfär Duo 5 g, 0–15% EtOAc/cyclohexane, linear gradient) to afford a pure sample. ^{1}H NMR (400 MHz, $(CD_3)_2SO$, 5 mM DMF) δ 7.04 (d, J = 8.4 Hz, 4H), 6.43 (d, J = 8.7 Hz, 4H), 4.77 (p, J = 5.5 Hz, 2H), 4.13 (t, J = 7.1 Hz, 4H), 3.52 (dd, J = 8.0, 4.7 Hz, 4H), 0.86 (s, 18H), 0.06 (s, 12H). ^{19}F NMR (376 MHz, $(CD_3)_2SO$) δ -138.85 – -139.05 (m, 1F), -139.69 – -139.87 (m, 1F), -142.79 – -143.00 (m, 1F), -151.81 – -151.97 (m, 1F). Analytical HPLC: t_R = 7.7 min, 97% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/ min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{38}H_{49}F_4N_2O_4Si_2$ $[M+H]^+$ 729.3162, found 729.3139.



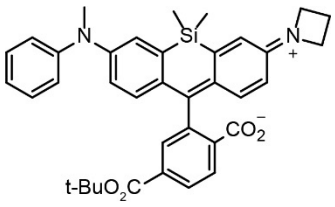
(19): 2-fluorodimethylaniline (510 mg, 3.67 mmol, 2 eq), tetrafluorophthalic anhydride (400 mg, 1.84 mmol, 1 eq) and $ZnCl_2$ (500 mg, 3.67 mmol, 2 eq) were loaded into a sealed vial then stirred vigorously at 150°C for 4 h. The reaction mixture was cooled to room temperature then dissolved progressively in organic solvents (CH_3CN and EtOAc). Aqueous HCl (1 M, 12 mL) was added and the organic solvents were evaporated under reduced pressure. The aqueous phase was neutralised to pH = 5-6 with NaOH (1M), which formed a green precipitate. The precipitate was dissolved with EtOAc. The layers were separated and the aqueous layer was extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Trituration with cyclohexane removed some excess 2-fluorodimethylaniline, and purification was performed by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–15% EtOAc/cyclohexane, linear gradient), followed by reverse phase HPLC (10–50% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, and lyophilised to afford the title compound as a white solid (5.4 mg, 1%, TFA salt). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.01 – 6.89 (m, 4H), 6.82 (t, J = 8.9 Hz, 2H), 2.89 (d, J = 1.1 Hz, 12H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -118.96 (m, 2F), -136.03 – -136.23 (m, 1F), -136.55 – -136.71 (m, 1F), -140.30 – -140.49 (m, 1F), -147.32 (td, J = 39.0, 4.9 Hz, 1F). Analytical HPLC: t_R = 4.8 min, 97% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{24}H_{22}F_3N_2O_2$ $[M+H]^+$ 481.1345, found 481.1340.



(20): *N,N*-dimethyl-*m*-toluidine (1.33 mL, 9.13 mmol, 2 eq), tetrafluorophthalic anhydride (1.00 g, 4.57 mmol, 1 eq) and $ZnCl_2$ (1.25 g, 9.13 mmol, 2 eq) were loaded into a sealed vial then stirred vigorously at 150°C for 3.5 h. The reaction mixture was cooled to room temperature then dissolved progressively in organic solvents (CH_3CN and EtOAc). Aqueous HCl (1 M, 20 mL) was added and the organic solvents evaporated under reduced pressure. The aqueous phase was neutralised to pH = 5-6 with NaOH (1M), which formed a green precipitate. The precipitate was dissolved with EtOAc, the layers were separated and the aqueous layer was extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Trituration with cyclohexane removed residual *N,N*-dimethyl-*m*-toluidine and purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0-30% EtOAc/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (476 mg, 22%). For full characterisation by NMR and UV/Vis spectroscopy, purification of an analytical fraction was performed by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, and lyophilised to afford the title compound as a pale green solid (73.3 mg, 3%, TFA salt). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.80 (dd, *J* = 8.8, 2.8 Hz, 2H), 6.53 (s, 2H), 6.42 (dd, *J* = 8.8, 2.8 Hz, 2H), 2.95 (d, *J* = 1.2 Hz, 12H), 2.11 (s, 6H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -137.22 – -137.95 (td, *J* = 20.0, 8.4 Hz, 1F), -139.28 – -139.68 (td, *J* = 20.0, 8.4 Hz, 1F), -143.49 – -143.96 (m, 1F), -151.44 (td, *J* = 19.5, 4.3 Hz, 1F). Analytical HPLC: t_R = 4.8 min, 98% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{26}H_{25}F_4N_2O_2$ $[M+H]^+$ 473.1847, found 473.1841.

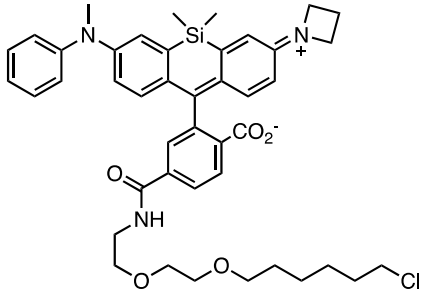


(21): Compound **21** was isolated as a by-product when attempting to synthesize the HaloTag ligand, **14-HTL**. **6-(MOM-MAC)-FMGL[8]** (223 mg, 0.431 mmol, 1 eq), and camphorsulfonic acid (110 mg, 0.475 mmol, 1.1 eq) were loaded into a vial which was sealed and evacuated/backfilled with argon. AcOH/DME (1/1, 6.4 mL) was added and the reaction stirred at 80°C for 68 h. The solvent was evaporated under reduced pressure, and the aqueous phase was neutralised with saturated $NaHCO_3$ before the reaction mixture was extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. To the crude intermediate, a solution of HaloTag(O2) amine (as the TFA salt, 229 mg, 0.678 mmol, 3 eq), DIEA (394 µL, 2.26 mmol, 10 eq) and HATU (258 mg, 0.678 mmol, 3 eq) in DMF (3.5 mL) were added. The reaction mixture was stirred for 18 h at room temperature and then evaporated to dryness. Purification was performed by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–20% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (20–80% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$, extracted with EtOAc (3x), dried over Na_2SO_4, filtered and evaporated to dryness to afford the title compound as a blue-green solid (20.6 mg, 11%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 – 7.31 (m, 4H), 7.29 (d, *J* = 5.2 Hz, 1H), 7.20 – 7.15 (m, 4H), 3.11 (s, 12H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -113.82 (d, *J* = 19.5 Hz, 1F), -129.4 – -129.3 (m, 1F), -142.13 (t, *J* = 20.9 Hz, 1F). Analytical HPLC: t_R = 4.5 min, 96% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{24}H_{22}F_3N_2O_2$ $[M+H]^+$ 427.1628, found 427.1624.



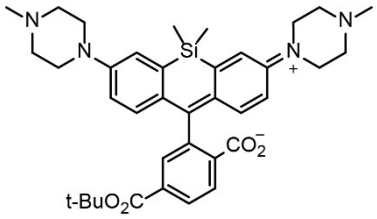
(S6): 6-CO_2*t*-Bu-fluorescein ditriflate (202 mg, 0.273 mmol, 1 eq), $Pd_2(dba)_3$ (25.0 mg, 0.027 mmol, 0.1 eq), XantPhos (47.5 mg, 0.082 mmol, 0.3 eq) and Cs_2CO_3 (178 mg, 0.547 mmol, 2.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (4 mL) and subsequently *N*-methylaniline (29.6 µL, 0.273 mmol, 1 eq) were added then the reaction mixture was stirred at 80°C for 2 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc and the solvent was evaporated under reduced pressure. Purification by silica gel column chromatography (Biotage Sfar Duo 10 g, cyclohexane:CH_2Cl_2: toluene (10:0:0–5:3.5:1.5), linear gradient) afforded the intermediate as a pale green solid. The intermediate showed moderate stability and was directly engaged in the next step.

This intermediate (~63 mg, 0.091 mmol), $Pd_2(dba)_3$ (8.3 mg, 0.009 mmol, 0.1 eq), XPhos (12.9 mg, 0.027 mmol, 0.3 eq) and Cs_2CO_3 (82.6 mg, 0.254 mmol, 2.8 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane (2 mL) and subsequently azetidine (14.7 µL, 0.217 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 3.5 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with EtOAc and the solvent was evaporated under reduced pressure. Purification by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–20% EtOAc/cyclohexane, linear gradient) afforded the title compound as a blue solid (35 mg, 21% over 2 steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.13 (dd, *J* = 8.1, 1.3 Hz, 1H), 7.97 (d, *J* = 8.0 Hz, 1H), 7.87 (s, 1H), 7.33 – 7.27 (m, 2H), 7.23 (d, *J* = 2.6 Hz, 1H), 7.12 – 7.07 (m, 2H), 7.04 (t, *J* = 7.3 Hz, 1H), 6.88 – 6.82 (m, 2H), 6.79 (dd, *J* = 8.9, 2.7 Hz, 1H), 6.71 (br s, 1H), 6.35 (br s, 1H), 3.93 (t, *J* = 7.4 Hz, 4H), 3.33 (s, 3H), 2.39 (q, *J* = 7.2 Hz, 2H), 1.56 (s, 9H), 0.62 (s, 3H), 0.56 (s, 3H). Analytical HPLC: t_R = 5.6 min, 95% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{37}H_{39}N_2O_4Si$ $[M+H]^+$ 603.2674, found 603.2664.

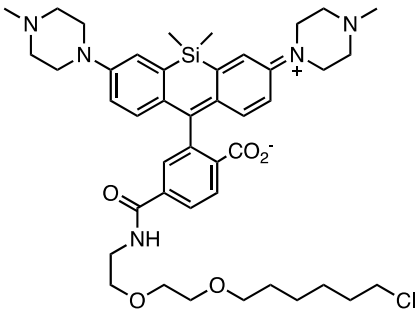


(5-HTL): S6 (56.3 mg, 0.093 mmol) was dissolved in CH_2Cl_2 (1.95 mL). TFA (0.39 mL) was added and the reaction was stirred at room temperature for 6 h. Toluene (2 mL) was added, the reaction mixture evaporated to dryness and then azeotroped with MeOH (3x) to give the crude carboxylic acid as a blue solid. This intermediate, HaloTag(O2) amine (TFA salt, 47.3 mg, 0.140 mmol, 1.5 eq) and HATU (53.3 mg, 0.140 mmol, 1.5 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. DMF (2 mL) and subsequently DIEA (81 µL, 0.467 mmol, 5 eq) were added. The reaction mixture was stirred at room temperature for 24 h. The solvent was evaporated under reduced pressure before purification by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–100% TBME/cyclohexane, linear gradient) followed by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were partially concentrated to remove CH_3CN then lyophilised overnight to afford the title compound as a blue solid (24.1 mg, 30%, TFA salt). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 (d, *J* = 7.9 Hz, 1H), 7.90 (dd, *J* = 8.0, 1.4 Hz, 1H), 7.74 – 7.72 (m, 1H), 7.30 (t, *J* = 7.9 Hz, 2H), 7.23 (d, *J* = 2.4 Hz, 1H), 7.10 (d, *J* = 7.6 Hz, 2H), 7.04 (t, *J* = 7.3 Hz, 1H), 6.83 – 6.74 (m, 4H), 6.72 (d, *J* = 2.7 Hz, 1H), 6.33 (dd, *J* = 8.7, 2.6 Hz, 1H), 3.93 (t, *J* = 7.3 Hz, 4H), 3.68–3.60 (m, 6H), 3.55 (dd, *J* = 5.8, 3.1 Hz, 2H), 3.50 (t, *J* = 6.6 Hz, 2H), 3.40 (t, *J* = 6.7 Hz, 2H), 3.33 (s, 3H), 2.39 (p, *J* = 7.3 Hz, 2H), 1.76 – 1.67 (m, 2H), 1.52 (p, *J* = 6.9 Hz, 2H), 1.43 – 1.35 (m, 2H), 1.34 – 1.26 (m, 2H), 0.60 (s, 3H), 0.55 (s, 3H). ^{13}C NMR (101 MHz, $(CD_3)_2CO$) δ 170.0 (CO), 166.1 (CO), 155.8 (C), 152.1 (C), 149.2 (C), 149.1 (C), 141.3 (C), 137.4

(C), 136.8 (C), 136.1 (C), 132.9 (C), 130.3 (CH), 129.0 (C), 128.9 (CH), 128.6 (CH), 128.5 (CH), 126.3 (CH), 124.1 (CH), 124.0 (CH), 123.8 (CH), 123.0 (CH), 119.8 (CH), 116.4 (CH), 113.4 (CH), 91.8 (C), 71.5 (CH_2), 70.9 (CH_2), 70.8 (CH_2), 70.0 (CH_2), 53.8 (CH_2), 45.8 (CH_2), 40.6 (CH_2), 40.3 (N-CH_3), 33.3 (CH_2), 30.3 (CH_2), 27.3 (CH_2), 26.1 (CH_2), 17.4 (CH_2), 0.2 (Si-CH_3), -1.2 (Si-CH_3). Analytical HPLC: t_R = 3.7 min, 97% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/ min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{43}H_{51}ClN_3O_5Si$ $[M+H]^+$ 752.3281, found 752.3268.

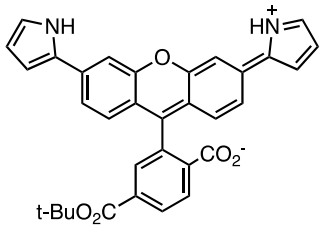


(S7): 6-CO_2*t*-Bu-Si-fluorescein ditriflate (88.1 mg, 0.119 mmol, 1 eq), $Pd(OAc)_2$ (5.4 mg, 0.024 mmol, 0.2 eq), BINAP (22.3 mg, 0.036 mmol, 0.3 eq) and Cs_2CO_3 (109 mg, 0.334 mmol, 2.8 eq) were added to a vial. The vial was sealed and evacuated/backfilled with argon. Toluene (1 mL) and subsequently *N*-methylpiperazine (32 µL, 0.286 mmol, 2.4 eq) were added then the reaction mixture was stirred at 100°C for 48 h. After cooling to room temperature, the reaction mixture was filtered through Celite, washed with MeOH and the solvent was evaporated under reduced pressure. Purification by silica gel column chromatography (10% MeOH/CH_2Cl_2) afforded the title compound as a pale blue solid (57.7 mg, 76%). For NMR characterisation, an analytical fraction of the product was further purified by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness to afford a pure sample. ^{1}H NMR (400 MHz, CD_3OD) δ 8.13 (dd, *J* = 8.1, 1.3 Hz, 1H), 7.99 (d, *J* = 8.0 Hz, 1H), 7.69 (t, *J* = 1.0 Hz, 1H), 7.29 (t, *J* = 1.6 Hz, 2H), 6.90 – 6.83 (m, 4H), 3.26 (t, *J* = 5.1 Hz, 8H), 2.68 (t, *J* = 5.1 Hz, 8H), 2.40 (s, 6H), 0.68 (s, 3H), 0.58 (s, 3H). ^{13}C NMR ($(CD_3)_2CO$, 101 MHz) δ 170.3 (CO), 164.6 (CO), 157.0 (C), 150.9 (C), 138.3 (C), 136.3 (C), 135.1 (C), 130.7 (CH), 128.9 (C), 128.3 (CH), 126.5 (CH), 125.0 (CH), 120.9 (CH), 118.0 (CH), 90.9 (C), 82.9 (C), 55.1 (CH_2), 48.1 (CH_2) 45.4 (CH_3), 28.1 (CH_3), -0.2 (d, Si-CH_3). Analytical HPLC: t_R = 2.9 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/ min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{37}H_{47}N_4O_4Si$ $[M+H]^+$ 639.3361, found 639.3347.

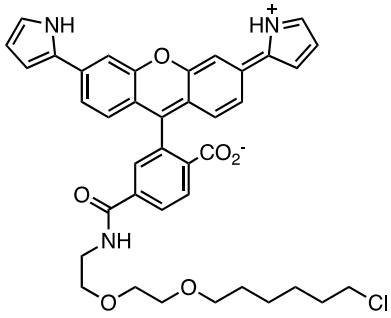


(6-HTL): S7 (57.7 mg, 0.090 mmol) was dissolved in CH_2Cl_2 (2 mL) and TFA (0.4 mL) was added. The reaction was stirred at room temperature for 4 h. Toluene (2 mL) was added, the reaction mixture evaporated to dryness and then azeotroped with MeOH (3x) to give the resulting acid as a blue solid. This acid intermediate (53.9 mg, 0.077 mmol), HaloTag(O2) amine (TFA salt, 39.2 mg, 0.116 mmol, 1.5 eq) and HATU (44.1 mg, 0.116 mmol, 1.5 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. DMF (1 mL) and subsequently DIEA (67 µL, 0.387 mmol, 5 eq) were added then the reaction mixture was stirred at room temperature for 4.5 h. The reaction mixture was evaporated to dryness before purification by silica gel column chromatography (10% MeOH/CH_2Cl_2) followed by reverse phase HPLC (10–50% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were partially concentrated to remove CH_3CN then lyophilised overnight to afford the title compound as a blue solid (26.8 mg, 38%, TFA salt). ^{1}H NMR ($(CD_3)_2CO$, 400 MHz) δ 8.12 (dd, *J* = 7.95, 1.40 Hz, 1H), 8.01 (dd, *J* = 8.0, 2.0 Hz, 1H), 7.78 (s, 1H), 7.48 (d, *J* = 2.8 Hz, 2H), 7.00 – 6.94 (m, 2H), 6.89 (dd, *J* = 8.8, 1.8 Hz, 2H), 3.70 – 3.16 (m, 16H), 3.13 – 3.03 (m,

4H), 2.92 (s, 6H), 1.80 – 1.66 (m, 2H), 1.49 – 1.25 (m, 8H), 0.72 (s, 3H), 0.58 (s, 3H). ^{13}C NMR (CD_3CN, 101 MHz) δ 170.5 (CO), 166.8 (CO), 155.7 (C), 150.0 (C), 141.8 (C), 137.6 (C), 136.1 (C), 129.2 (CH), 129.1 (CH), 128.5 (C), 126.9 (CH), 123.8 (CH), 121.8 (CH), 118.5 (CH), 91.4 (C), 71.6 (CH_2), 70.9 (CH_2), 70.8 (CH_2), 69.9 (CH_2), 53.7 (CH_2), 46.5 (CH_2), 46.2 (CH_2), 43.5 (CH_3), 40.7 (CH_2), 33.3 (CH_2), 30.2 (CH_2), 27.3 (CH_2), 26.1 (CH_2), 0.1 (Si-CH_3), -1.00 (Si-CH_3). Analytical HPLC: t_R = 2.4 min, 95% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 10 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{43}H_{59}ClN_5O_5Si$ $[M+H]^+$ 788.3969, found 788.3953.

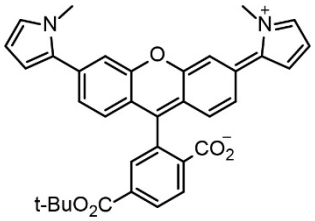


(S8): 6-CO_2t-Bu-fluorescein ditriflate (79.5 mg, 0.114 mmol, 1 eq), 1-*N*-Boc-pyrrole-2-boronic acid pinacol ester (100 mg, 0.342 mmol, 3.0 eq), $Pd(PPh_3)_4$ (7.9 mg, 0.069 mmol, 0.06 eq) and KOAc (67.2 mg, 0.685 mmol, 6.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane/H_2O (2/1, 1.2 mL) was added and the reaction mixture was stirred at 100°C for 68 h. After cooling to room temperature, the reaction mixture was extracted with $CHCl_3$. The organic layers were combined, washed with brine, dried over Na_2SO_4 and the solvent was evaporated under reduced pressure. Silica gel column chromatography (30% EtOAc/cyclohexane) afforded the title compound as a blue solid (46.7 mg, 77%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.78 (br s, 2H), 8.26 (d, *J* = 8.0 Hz, 1H), 8.09 (d, *J* = 8.0 Hz, 1H), 7.75 (s, 1H), 7.35 – 7.32 (m, 2H), 7.13 (dd, *J* = 8.3, 1.8 Hz, 2H), 6.92 – 6.88 (m, 2H), 6.73 (d, *J* = 8.3 Hz, 2H), 6.61 – 6.54 (m, 2H), 6.32 – 6.28 (m, 2H), 1.53 (s, 9H). ^{13}C NMR (101 MHz, $(CD_3)_2CO$) δ 172.3 (CO), 164.7 (CO), 154.1 (C), 152.7 (C), 146.1 (C), 139.3 (C), 137.1 (C), 131.9 (CH), 131.0 (C), 129.5 (CH), 126.0 (CH), 125.5 (CH), 121.2 (CH), 120.4 (CH), 116.3 (C), 111.8 (CH), 110.6 (CH), 108.4 (CH), 83.3 (C), 82.9 (C), 28.1 (CH_3). Analytical HPLC: t_R = 5.2 min, 98% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 10 min run, 0.5 mL/ min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{33}H_{27}N_2O_5$ $[M+H]^+$ 531.1914, found 531.1919.

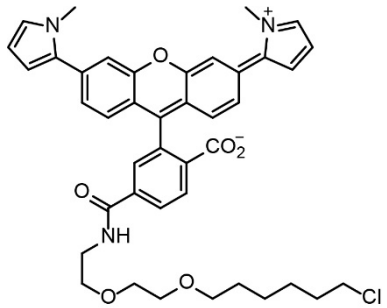


(9-HTL): S8 (44.2 mg, 0.083 mmol) was dissolved in CH_2Cl_2 (2 mL), and TFA (0.4 mL) was added. The reaction was stirred at room temperature for 6 h. Toluene (2 mL) was added, the reaction mixture evaporated to dryness and then azeotroped with MeOH (3x) to give the crude carboxylic acid as a purple solid. This acid intermediate (34.9 mg, 0.059 mmol), HaloTag(O2) amine (TFA salt, 28.8 mg, 0.089 mmol, 1.5 eq) and HATU (33.8 mg, 0.089 mmol, 1.5 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. DMF (1 mL) and subsequently DIEA (52 µL, 0.297 mmol, 5 eq) were added then the mixture was stirred at room temperature for 3 h. The reaction mixture was evaporated to dryness and purification by silica gel column chromatography (20% EtOAc/cyclohexane) afforded the title compound as a blue solid (38.3 mg, 68%). ^{1}H NMR ($(CD_3)_2CO$, 400 MHz) δ 10.75 (br s, 2H), 8.25 (dd, *J* = 8.0, 1.4 Hz, 1H), 8.10 (d, *J* = 8.0 Hz, 2H), 7.80 (s, 1H), 7.61 (d, *J* = 1.8 Hz, 2H), 7.41 (dd, *J* = 8.4, 1.8 Hz, 2H), 6.96 – 6.92 (m, 2H), 6.86 (d, *J* = 8.4 Hz, 2H), 6.70 – 6.66 (m, 2H), 6.22 – 6.18 (m, 2H), 3.56 – 3.51 (m, 4H), 3.51 – 3.45 (m, 4H), 3.43 – 3.38 (m, 2H), 3.30 (t, *J* = 6.5 Hz, 2H), 1.72 – 1.63 (m, 2H), 1.46 – 1.22 (m, 6H). ^{13}C NMR ($(CD_3)_2CO$, 101 MHz) δ 168.9 (CO), 166.0 (CO), 154.2 (C), 152.6 (C), 142.3 (C), 137.0 (C), 130.9 (C), 130.4 (CH), 129.5 (CH), 129.3 (C), 125.8 (CH), 123.3 (CH), 121.2 (CH), 120.3 (CH), 116.3 (C), 111.7 (CH), 110.6 (CH), 108.3 (CH), 83.2 (C), 71.4 (CH_2), 70.8 (CH_2), 70.6

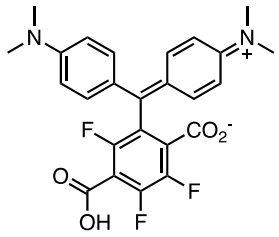
(CH_2), 70.0 (CH_2), 46.7 (CH_2), 40.5 (CH_2), 33.3 (CH_2), 30.2 (CH_2), 27.3 (CH_2), 26.1 (CH_2). Analytical HPLC: t_R = 4.9 min, 99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 10 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{39}H_{39}ClN_3O_6$ $[M+H]^+$ 680.2522, found 680.2515.



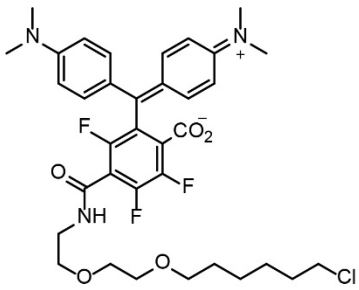
(S9): 6-CO_2*t*-Bu-fluorescein ditriflate (77.3 mg, 0.111 mmol, 1 eq), 1-methyl-2-pyrroleboronic acid pinacol ester (69.0 mg, 0.333 mmol, 3.0 eq), $Pd(PPh_3)_4$ (7.7 mg, 0.067 mmol, 0.06 eq) and KOAc (65.4 mg, 0.666 mmol, 6.0 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. 1,4-Dioxane/H_2O (2/1, 1.2 mL) was added and the reaction mixture was stirred at 100°C for 18 h. After cooling to room temperature, the reaction mixture was extracted with $CHCl_3$. The organic layers were combined, washed with brine, dried over Na_2SO_4 and the solvent was evaporated under reduced pressure. Silica gel column chromatography (15% EtOAc/cyclohexane) afforded the title compound as a blue solid (57.1 mg, 92%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.25 (d, *J* = 8.0, 1H), 8.10 (d, *J* = 8.0 Hz, 1H), 7.85 (s, 1H), 7.36 (d, *J* = 1.7 Hz, 2H), 7.11 (dd, *J* = 8.2, 1.7 Hz, 2H), 6.82 (d, *J* = 8.2 Hz, 2H), 6.76 (t, *J* = 1.9 Hz, 2H), 6.30 – 6.34 (m, 2H), 6.23 – 6.21 (m, 2H), 3.74 (s, 6H), 1.57 (s, 9H). ^{13}C NMR ($CDCl_3$, 101 MHz) δ 168.7 (C), 164.3 (C), 153.2 (C), 151.3 (C), 138.6 (C), 136.3 (C), 133.1 (C), 131.1 (CH), 129.7 (C), 128.2 (CH), 125.3 (CH), 125.2 (CH), 125.0 (CH), 124.1 (CH), 116.5 (CH), 116.3 (C), 110.0 (CH), 108.3 (CH), 83.0 (C), 82.8 (C), 35.5 (CH_3), 28.2 (CH_3). HRMS (ESI) calcd for $C_{35}H_{31}N_2O_5$ $[M+H]^+$ 559.2227, found 559.2228.



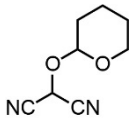
(11-HTL): S9 (55.1 mg, 0.10 mmol) was dissolved in CH_2Cl_2 (2 mL), and TFA (0.4 mL) was added. The reaction was stirred at room temperature for 7 h. Toluene (2 mL) was added, the reaction mixture evaporated to dryness and then azeotroped with MeOH (3x) to give the crude carboxylic acid as a dark blue solid. This acid intermediate (35.2 mg, 0.057 mmol), HaloTag(O2) amine (TFA salt, 28.8 mg, 0.085 mmol, 1.5 eq) and HATU (32.5 mg, 0.085 mmol, 1.5 eq) were loaded into a vial. The vial was sealed and evacuated/backfilled with argon. DMF (1 mL) and subsequently DIEA (50 µL, 0.285 mmol, 5 eq) were added then the mixture was stirred at room temperature for 3 h. The reaction mixture was evaporated to dryness and purification by silica gel column chromatography (30% EtOAc/cyclohexane) followed by reverse phase HPLC (10–50% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were partially concentrated to remove CH_3CN then lyophilised overnight to afford the title compound as a blue solid (26.4 mg, 57%, TFA salt). ^{1}H NMR ($(CD_3)_2CO$, 400 MHz) δ 8.27 (dd, *J* = 8.0, 1.4 Hz, 1H), 8.15 – 8.07 (m, 2H), 7.85 (s, 1H), 7.44 (d, *J* = 1.8 Hz, 2H), 7.24 (dd, *J* = 8.3, 1.8 Hz, 2H), 6.95 (d, *J* = 8.3 Hz, 2H), 6.86 – 6.81 (m, 2H), 6.34 – 6.30 (m, 2H), 6.14 – 6.09 (m, 2H), 3.77 (s, 6H), 3.57 – 3.52 (m, 4H), 3.52 – 3.46 (m, 4H), 3.46 – 3.41 (m, 2H), 3.35 – 3.29 (m, 2H), 1.74 – 1.65 (m, 2H), 1.48 – 1.24 (m, 6H). ^{13}C NMR ($(CD_3)_2CO$, 101 MHz) δ 168.8 (C), 165.9 (C), 154.0 (C), 152.0 (C), 142.4 (C), 137.3 (C), 133.2 (C), 130.4 (CH), 129.3 (C), 129.2 (CH), 126.2 (CH), 125.9 (CH), 124.7 (CH), 123.5 (CH), 117.4 (C), 116.4 (CH), 110.8 (CH), 108.7 (CH), 82.9 (C), 71.4 (CH_2), 70.8 (CH_2), 70.7 (CH_2), 70.0 (CH_2), 45.7 (CH_2), 40.7 (CH_2), 40.5 (CH_2), 35.7 (CH_3), 33.3 (CH_2), 27.3 (CH_2), 26.1 (CH_2). Analytical HPLC: t_R = 5.3 min, >99.0% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 10 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{41}H_{43}ClN_3O_6$ $[M+H]^+$ 708.2824, found 708.2823.



(S10): 6-(MOM-MAC)-FMGL[168] (600 mg, 1.09 mmol, 1 eq) in a mixture of THF (60 mL) and H_2SO_4 (1 M in H_2O, 60 mL) was stirred at 60°C for 64 h. The solvent was evaporated under reduced pressure and the aqueous solution was neutralised with saturated $NaHCO_3$ before extraction with 15% *i*-PrOH in $CHCl_3$ (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–15% MeOH/CH_2Cl_2 with constant 1% AcOH additive, linear gradient) afforded the title compound as a blue-green solid (367 mg, 72%). An analytical fraction of the product was further purified by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined and lyophilised to afford the analytically pure compound as a TFA salt. ^{1}H NMR (400 MHz, CD_3OD) δ 7.19 (d, J = 8.5 Hz, 4H), 6.91 (d, J = 8.9 Hz, 4H), 3.03 (s, 12H).^{19}F NMR (376 MHz, CD_3OD) δ -116.20 (d, J = 21.8 Hz, 1F), -133.12 (d, J = 20.4 Hz, 1F), -142.98 (t, J = 21.3 Hz, 1F). Analytical HPLC: t_R = 3.3 min, >99.0% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{25}H_{22}F_3N_2O_4$ $[M+H]^+$ 471.1526, found 471.1518.

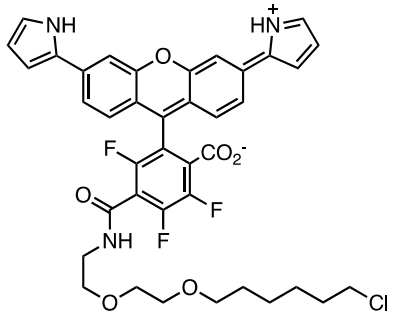


14-HTL: **S10** (0.164 g, 0.348 mmol, 1 eq), HaloTag(O2) amine (TFA salt, 1.174 g, 3.48 mmol, 10 eq), EDC.HCl (334 mg, 1.74 mmol, 5 eq), and 1-hydroxybenzotriazole hydrate (0.141 g, 1.04 mmol, 3 eq) were loaded into a sealed vial, which was evacuated/ backfilled under argon. CH_2Cl_2 (3 mL) and then DIEA (424 µl, 2.44 mmol, 7 eq) were added and the reaction stirred at room temperature for 15 h. Saturated $NaHCO_3$ was added, and the mixture was extracted with CH_2Cl_2 (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Purification was performed via silica gel column chromatography (Biotage Sfar Duo 5 g, 0–50% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness to afford the title compound as a blue-green solid (31.8 mg, 13%). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 8.16 (s, 1H, NH), 7.13 (d, J = 8.8 Hz, 4H), 6.70 (d, J = 8.8 Hz, 4H), 3.62 (t, J = 5.4 Hz, 2H), 3.59 – 3.53 (m, 6H), 3.52 – 3.47 (m, 2H), 3.38 (t, J = 6.4 Hz, 2H), 2.97 (s, 12H), 1.74 (p, J = 6.9 Hz, 2H), 1.54 – 1.46 (m, 2H), 1.39 – 1.25 (m, 4H). ^{19}F NMR (376 MHz, $(CD_3)_2CO$) δ -118.14 –-118.33 (m, 1F), -135.48 – -135.78 (m, 1F), -144.13 – -144.42 (m, 1F). Analytical HPLC: t_R = 4.67 min, >99.0% purity (5–95% MeCN/ H_2O, linear gradient, with constant 0.1% formic acid additive, 11 min run, 0.5 mL/min flow, ESI, positive ion mode, UV detection at 254 nm). HRMS (ESI) calcd for $C_{35}H_{42}ClF_3N_3O_5$ $[M+H]^+$ 676.2760, found 676.2748.



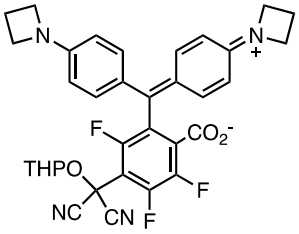
(S11): Acetylmalononitrile (1.39 g, 12.9 mmol, 1 eq) was dissolved in H_2O (30 mL), and a peracetic acid solution (9.6% in AcOH, 30 mL) was added. The reaction was stirred at room temperature for 3 h. The solvent was evaporated under reduced pressure. The crude off-white intermediate was resuspended in toluene (33 mL) and cooled to 0°C. *p*-

Toluenesulfonic acid (246 mg, 1.29 mmol, 0.1 eq) and 3,4-dihydro-2H-pyran (1.77 mL, 19.4 mmol, 1.5 eq) were added, and the mixture was stirred for 2 h at 0°C. The reaction was diluted with EtOAc and the organic layer was washed with H_2O, saturated $NaHCO_3$ and brine. The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–50% CH_2Cl_2/cyclohexane, linear gradient) afforded the title compound as a white solid (844 mg, 40%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.38 (s, 1H), 5.01 (t, J = 2.6 Hz, 1H), 3.79 (td, J = 10.8, 2.8 Hz, 1H), 3.74 – 3.65 (m, 1H), 1.87 – 1.51 (m, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 111.01 (CN), 110.97 (CN), 98.6 (CH), 62.7 (CH_2), 51.8 (CH), 29.1 (CH_2), 24.7 (CH_2), 17.8 (CH_2). HRMS (ESI) calcd for $C_8H_9N_2O_2$ $[M-H]^-$ 165.0670, found 165.0672.



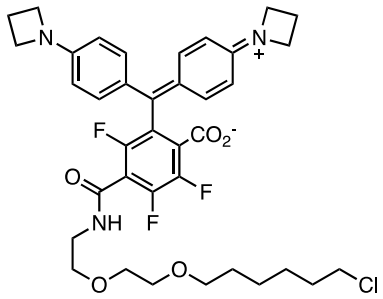
(10-HTL): 10 (93.3 mg, 0.186 mmol, 1 eq) and **S11** (28.6 mg, 0.186 mmol, 1 eq) were dissolved in DMF (4.1 mL) under argon, and DIEA (66.3 µL, 0.372 mmol, 2 eq) was added. The reaction was stirred at room temperature for 3 h, then evaporated to dryness. Purification by silica gel chromatography (Biotage Sfär Duo 10 g, 0–50% TBME/cyclohexane, linear gradient) afforded compound **S12** as a dark blue solid. This intermediate showed limited stability and was directly engaged in the next step.

This intermediate **S12** (~106 mg, 0.164 mmol) was dissolved in CH_2Cl_2 (10.9 mL). Triethylsilane (2.18 mL) and then trifluoroacetic acid (1.09 mL) were added and the reaction was stirred at room temperature for 3 h. Toluene (5 mL) was added, and the reaction mixture was evaporated to dryness. The residue was added to a solution of HaloTag(O2) amine (TFA salt, 110 mg, 0.333 mmol, 2 eq) and DIEA (292 µL, 1.64 mmol, 10 eq) in CH_2Cl_2 (2 mL), and the reaction was stirred at room temperature for 16 h. The solvent was evaporated under reduced pressure and purification was performed by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–60% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (20–80% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, neutralised with saturated $NaHCO_3$, and extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4 and evaporated to dryness to afford the title compound as a blue-green solid (12.9 mg, 9% over 3 steps). ^{1}H NMR (400 MHz, CD_3CN) δ 9.94 (s, 2H), 7.52 (d, J = 1.8 Hz, 2H), 7.43 (s, 1H), 7.39 (dd, J = 8.3, 1.8 Hz, 2H), 7.12 (d, J = 8.3 Hz, 2H), 6.92 (q, J = 2.3 Hz, 2H), 6.67 (q, J = 2.6 Hz, 2H), 6.23 (q, J = 2.7 Hz, 2H), 3.55 – 3.36 (m, 8H), 3.27 (t, J = 6.5 Hz, 2H), 1.68 (p, J = 6.8 Hz, 2H), 1.44 – 1.18 (m, 8H). ^{19}F NMR (376 MHz, CD_3CN) δ -123.09 (d, J = 21.4 Hz, 1F), -133.62 (d, J = 21.5 Hz, 1F), -142.86 (t, J = 21.6 Hz, 1F). Analytical HPLC: t_R = 4.5 min, >99% purity (5–95% CH_3CN/H_2O, linear gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{39}H_{36}ClF_3N_3O_6$ $[M+H]^+$ 734.2239, found 734.2239.

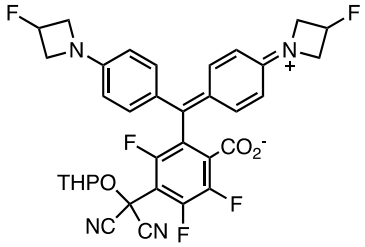


(S13): 15 (208 mg, 0.444 mmol, 1 eq) and **S11** (68.5 mg, 0.444 mmol, 1 eq) were combined in DMF (9.9 mL) under argon, and DIEA (159 µL, 0.818 mmol, 2 eq) was added. The reaction was stirred at room temperature for 3 h, then evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–30% EtOAc/cyclohexane) afforded the title compound as a blue-green solid (145 mg, 53%). For characterisation by NMR spectroscopy, an analytical fraction of the product was further purified by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–25%

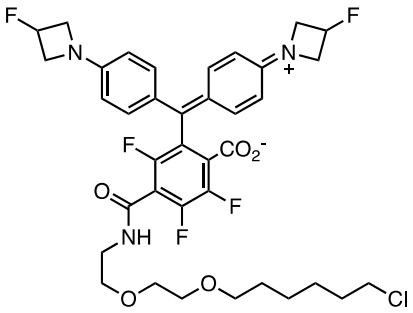
EtOAc/cyclohexane). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 7.09 – 7.14 (m, 4H), 6.38 – 6.43 (m, 4H), 4.70 (t, *J* = 3.0 Hz, 1H), 3.88 (t, *J* = 7.2 Hz, 8H), 3.67 – 3.75 (m, 1H), 3.50 – 3.57 (m, 1H), 2.36 (p, *J* = 7.3 Hz, 4H), 1.90 – 1.81 (m, 2H), 1.75 – 1.50 (m, 4H). ^{19}F NMR (376 MHz, $(CD_3)_2CO$) δ -112.62 (d, *J* = 21.8 Hz, 1F), -129.26 (d, *J* = 19.8 Hz, 1F), -141.66 (t, *J* = 20.8 Hz, 1F). HRMS (ESI) calcd for $C_{34}H_{30}F_3N_4O_4$ $[M+H]^+$ 615.2214, found 615.2196.



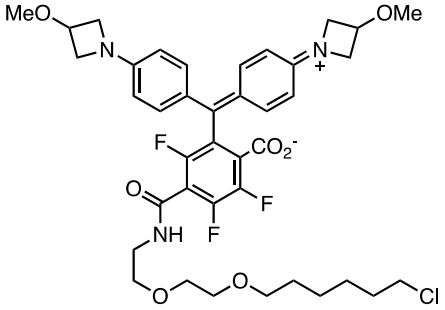
(15-HTL): S13 (114 mg, 0.186 mmol) was dissolved in CH_2Cl_2 (12.4 mL) then triethylsilane (1.24 mL) and trifluoroacetic acid (2.5 mL) were added. The reaction was stirred at room temperature for 2 h. Toluene (5 mL) was added, and the reaction mixture was evaporated to dryness. The residue was added to a solution of HaloTag(O2) amine (TFA salt, 126 mg, 0.373 mmol, 2 eq) and DIEA (333 µL, 1.86 mmol, 10 eq) in CH_2Cl_2 (3 mL), and the reaction was stirred at room temperature for 15 h. The solvent was evaporated under reduced pressure and purification was performed by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–60% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (20–80% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, neutralised with saturated $NaHCO_3$, and extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4 and evaporated to dryness to afford the title compound as a blue-green solid (8.9 mg, 7%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$), 5 mM DMF) δ 9.14 (t, *J* = 5.6 Hz, 1H), 7.02 (d, *J* = 8.3 Hz, 4H), 6.39 (d, *J* = 8.3 Hz, 4H), 3.81 (t, *J* = 7.2 Hz, 8H), 3.60 (t, *J* = 6.6 Hz, 2H), 3.51 (m, 4H), 3.46 (m, 2H), 3.40 (q, *J* = 5.6 Hz, 2H), 2.29 (q, *J* = 7.3 Hz, 4H), 1.68 (p, *J* = 6.8 Hz, 2H), 1.46 (p, *J* = 6.7 Hz, 2H), 1.31 (m, 6H). ^{19}F NMR (376 MHz, CD_3CN) δ -118.58 (d, *J* = 21.8 Hz, 1F), -135.14 (d, *J* = 21.3 Hz, 1F), -143.24 (t, *J* = 21.6 Hz, 1F). Analytical HPLC: t_R = 4.8 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{37}H_{42}ClF_3N_3O_5$ $[M+H]^+$ 700.2760, found 700.2743.



(S14): 16 (102 mg, 0.202 mmol, 1 eq) and **S11** (31.2 mg, 0.202 mmol, 1 eq) were combined in DMF (4.5 mL) under argon, and DIEA (72 µL, 0.404 mmol, 2 eq) was added. The reaction was stirred at room temperature for 3.5 h, then evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–30% EtOAc/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (79.2 mg, 60%). For characterisation by NMR spectroscopy, an analytical fraction of the product was further purified by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–60% CH_2Cl_2/cyclohexane followed by Biotage Sfär Duo 5 g, 0–25% EtOAc/cyclohexane, linear gradient). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 7.16 (d, *J* = 8.7, 4H), 6.56 – 6.47 (m, 4H), 5.50 (dtt, $^2J_{HF}$ = 57.3 Hz, *J* = 8.8, 3.0 Hz, 2H), 5.47 (t, *J* = 3.1 Hz, 1H), 4.30 – 4.19 (m, 4H), 4.01 – 3.88 (m, 4H), 3.75 – 3.67 (m, 1H), 3.57 – 3.50 (m, 1H), 1.90 – 1.78 (m, 2H), 1.75 – 1.51 (m, 4H). ^{19}F NMR (376 MHz, $CD_3)_2CO$) δ -112.64 (d, *J* = 21.8 Hz), -129.01 (d, *J* = 20.4 Hz), -141.46 (t, *J* = 20.8 Hz), -179.60 – -179.80 (m, 2F). HRMS (ESI) calcd for $C_{34}H_{28}F_5N_4O_4$ $[M+H]^+$ 651.2025, found 651.2009.



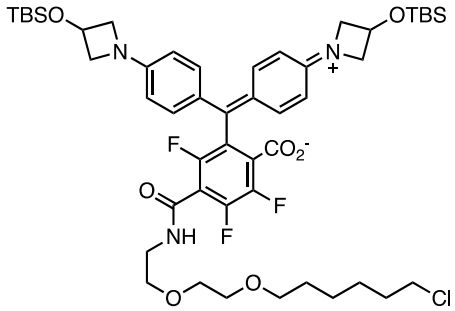
(16-HTL): S14 (16.6 mg, 0.026 mmol) was dissolved in CH_2Cl_2 (3.4 mL), triethylsilane (0.34 mL) and trifluoroacetic acid (0.68 mL) were added. The reaction was stirred at room temperature for 1 h. Toluene (2 mL) was added, and the reaction mixture was evaporated to dryness. The residue was added to a solution of HaloTag(O2) amine (TFA salt, 17.2 mg, 0.051 mmol, 2 eq) and DIEA (46 µL, 0.255 mmol, 10 eq) in CH_2Cl_2 (1.5 mL), and the reaction was stirred at room temperature for 16 h. The solvent was evaporated under reduced pressure and purification was performed by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–60% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, neutralised with saturated $NaHCO_3$, and extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4 and evaporated to dryness to afford the title compound as a blue-green solid (5.8 mg, 30%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$, 5 mM DMF) δ 9.18 (t, *J* = 5.8 Hz, 1H), 7.06 (d, *J* = 8.3 Hz, 4H), 6.47 (d, *J* = 8.3 Hz, 4H), 5.47 (m, 2H), 4.16 (m, 4H), 3.89 (ddd, *J* = 24.2, 9.5, 3.1 Hz, 4H), 3.55 – 3.27 (m, 12H), 1.67 (p, *J* = 6.9 Hz, 2H), 1.46 (p, *J* = 6.8 Hz, 2H), 1.25 (m, *J* = 11.9 Hz, 2H), 0.85 (t, *J* = 6.6 Hz, 2H). ^{19}F NMR (376 MHz, CD_3CN) δ -118.51 (dd, *J* = 22.2 Hz, 2.1 Hz, 1F), -134.89 (dd, *J* = 21.1 Hz, 2.5 Hz, 1F), -143.01 (t, *J* = 21.8 Hz, 1F), -179.95 – -179.98 (m, 2F). Analytical HPLC: t_R = 4.5 min, 98% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{37}H_{40}ClF_5N_3O_5$ $[M+H]^+$ 736.2571, found 736.2554.



(17-HTL): **17** (656 mg, 1.24 mmol, 1 eq) and **S11** (191 mg, 1.24 mmol, 1 eq) were combined in DMF (27.5 mL) under argon, and DIEA (443 µL, 2.48 mmol, 2 eq) was added. The reaction was stirred at room temperature for 4 h, then evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–30% EtOAc/cyclohexane, linear gradient) afforded the compound **S15** as a blue-green solid. This intermediate showed limited stability and was directly engaged in the next step.

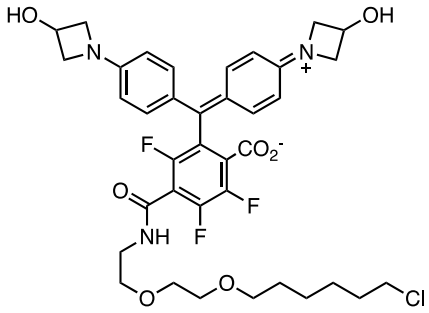
S15 (~455 mg, 0.675 mmol) was taken up in CH_2Cl_2 (45 mL), triethylsilane (4.50 mL) and trifluoroacetic acid (9 mL) were added. The reaction was stirred at room temperature for 2 h. Toluene (5 mL) was added, and the reaction mixture was evaporated to dryness. The residue was added to a solution of HaloTag(O2) amine (TFA salt, 455 mg, 1.35 mmol, 2 eq) and DIEA (1.21 mL, 6.75 mmol, 10 eq) in CH_2Cl_2 (30 mL), and the reaction was stirred at room temperature for 15 h. The solvent was evaporated under reduced pressure and purification was performed by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–50% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (10–100% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, neutralised with saturated $NaHCO_3$, and extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4 and evaporated to dryness to afford the title compound as a blue-green solid (243 mg, 26% over 3 steps). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 8.36 (br s, 1H), 7.11 (d, *J* = 8.7 Hz, 4H), 6.44 (d, *J* = 8.7 Hz, 4H), 4.37 – 4.31 (m, 2H), 4.12 – 4.08 (m, 4H), 3.70 – 3.65 (m, 4H), 3.65 – 3.60 (m, 2H), 3.60 - 3.54 (m, 6H), 3.51 – 3.48 (m, 2H), 3.39 (t, *J* = 6.5 Hz, 2H), 3.28 (s, 6H), 1.74 (p, *J* = 5.6 Hz, 2H), 1.50 (p, *J* = 6.8 Hz, 2H), 1.45 – 1.28 (m, 4H). ^{19}F NMR (376 MHz, $(CD_3)_2SO$) δ -118.77 (d, *J* = 21.8 Hz, 1F), -134.76 – -134.98 (d, *J* = 23.2 Hz, 1F), -143.10 (t, *J* = 22.0 Hz, 1F). Analytical HPLC: t_R = 4.5 min, 98.5

% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{39}H_{46}ClF_3N_3O_7$ $[M+H]^+$ 760.2971, found 760.2950.

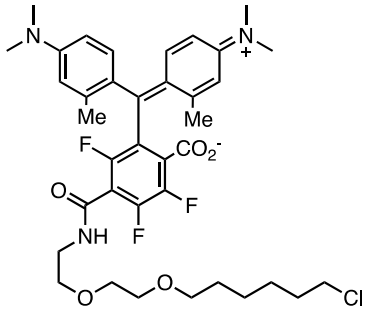


(S17): 18 (159 mg, 0.288 mmol, 1 eq) and **S11** (44.4 mg, 0.288 mmol, 1 eq) were combined in DMF (6 mL) under argon, and DIEA (103 µL, 0.576 mmol, 2 eq) was added. The reaction was stirred at room temperature for 3 h, then evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–15% EtOAc/cyclohexane, linear gradient) afforded compound **S16** as a green solid. This intermediate showed limited stability and was directly engaged in the next step.

S16 (~85 mg, 0.098 mmol) was taken up in CH_2Cl_2 (6.5 mL), triethylsilane (650 µL) and trifluoroacetic acid (1.3 mL) were added. The reaction was stirred at room temperature for 4 h. Toluene (2 mL) was added, and the reaction mixture was evaporated to dryness. The residue was added to a solution of HaloTag(O2) amine (TFA salt, 65.9 mg, 0.195 mmol, 2 eq) and DIEA (175 µL, 0.977 mmol, 10 eq) in CH_2Cl_2 (2 mL), and the reaction was stirred at room temperature for 15 h. The solvent was evaporated under reduced pressure and purification was performed by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–30% EtOAc/cyclohexane, linear gradient) to afford the title compound as a green solid (23.6 mg, 9% over 3 steps). ^{1}H NMR (400 MHz, CD_3CN) δ 7.35 (t, J = 5.5 Hz, 1H), 7.07 (d, J = 8.3 Hz, 4H), 6.41 (d, J = 8.3 Hz, 4H), 4.76 (tt, J = 6.2, 4.7 Hz, 2H), 4.16 – 4.10 (m, 4H), 3.60 – 3.45 (m, 12H), 3.41 (t, J = 6.5 Hz, 2H), 3.34 (t, J = 6.5 Hz, 2H), 1.79 – 1.66 (m, 2H), 1.54 (p, J = 6.9 Hz, 2H), 1.48 – 1.24 (m, 4H), 0.89 (s, 18H), 0.08 (s, 12H). ^{19}F NMR (376 MHz, CD_3CN) δ -118.45 (d, J = 21.8 Hz, 1F), -135.11 (d, J = 21.1 Hz, 1F), -143.25 (t, J = 21.7 Hz, 1F). Analytical HPLC: t_R = 7.5 min, 95% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{49}H_{70}F_3N_3O_7Si_2$ $[M+H]^+$ 960.4387, found 960.4364.

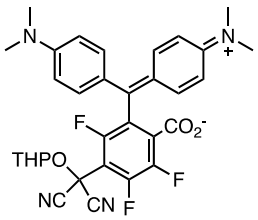


(22-HTL): S17 (6.9 mg, 0.007 mmol, 1 eq) in THF (1 mL) was cooled to 0°C and a solution of tetrabutylammonium fluoride (36 mM in THF, 1 mL, 5 eq) was added. The reaction was stirred at 0°C for 1 h. Saturated NH_4Cl was added, and the mixture was extracted with EtOAc (3x). The combined organic layers were washed with brine, dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness. Purification was performed by silica gel chromatography (Biotage Sfär Duo 5 g, 0–5% $MeOH/CH_2Cl_2$, linear gradient) followed by reverse phase HPLC (10–100% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, neutralised with saturated $NaHCO_3$, and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4 and evaporated to dryness to afford the title compound as a blue solid (1.4 mg, 27%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$) δ 9.15 (t, J = 5.6 Hz, 1H), 7.02 (d, J = 8.5 Hz, 4H), 6.61 (br s, 2H), 6.41 (d, J = 8.5 Hz, 4H), 5.67 – 5.58 (m, 2H), 4.59 – 4.51 (m, 4H), 4.07 (t, J = 7.2 Hz, 4H), 3.62 – 3.43 (m, 12H), 1.67 (p, J = 7.0 Hz, 2H), 1.47 (p, J = 7.0 Hz, 2H), 1.39 – 1.10 (m, 4H). ^{19}F NMR (376 MHz, $(CD_3)_2SO$) δ -118.79 (d, J = 20.9 Hz), -134.87 (d, J = 23.5 Hz), -143.05 – -143.29 (m, 1F). Analytical HPLC: t_R = 3.9 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{37}H_{42}ClF_3N_3O_7$ $[M+H]^+$ 732.2658, found 732.2637.

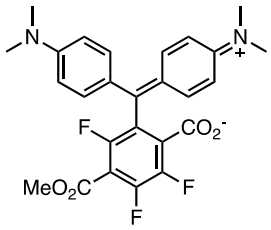


(20-HTL): 20 (59.7 mg, 0.126 mmol, 1 eq) and **S11** (19.5 mg, 0.126 mmol, 1 eq) were combined in DMF (3 mL) under argon, and DIEA (45 µL, 0.252 mmol, 2 eq) was added. The reaction was stirred at room temperature for 3 h, then evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–20% EtOAc/cyclohexane, linear gradient) afforded compound **S18** as a blue-green solid. This intermediate showed limited stability and was directly engaged in the next step.

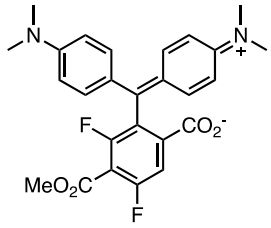
This intermediate **S18** (~57 mg, 0.093 mmol) was dissolved in CH_2Cl_2 (6.2 mL). Triethylsilane (620 µL) and then trifluoroacetic acid (1.24 mL) were added and the reaction was stirred at room temperature for 3 h. Toluene (5 mL) was added, and the reaction mixture was evaporated to dryness. The residue was added to a solution of HaloTag(O2) amine (TFA salt, 62.6 mg, 0.186 mmol, 2 eq) and DIEA (165 µL, 0.932 mmol, 10 eq) in CH_2Cl_2 (2 mL), and the reaction was stirred at room temperature for 16 h. The solvent was evaporated under reduced pressure and purification was performed by silica gel column chromatography (Biotage Sfär Duo 5 g, 0–50% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (10–95% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were combined, neutralised with saturated $NaHCO_3$, and extracted with EtOAc (3x). The organic layers were combined, dried over Na_2SO_4 and evaporated to dryness to afford the title compound as a blue-green solid (20.1 mg, 23% over 3 steps). ^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 8.18 – 8.25 (m, 1H), 6.83 (d, J = 8.2 Hz, 2H), 6.61 (s, 2H), 6.46 (dd, J = 8.8, 2.8 Hz, 2H), 3.64 – 3.60 (m, 2H), 3.60 – 3.52 (m, 6H), 3.52 – 3.47 (m, 2H), 3.39 (t, J = 6.5 Hz, 2H), 2.94 (s, 12H), 2.07 (s, 6H), 1.78 – 1.69 (m, 2H), 1.55 – 1.47 (m, 2H), 1.47 – 1.26 (m, 4H). ^{19}F NMR (376 MHz, CD_3CN) δ -118.61 (d, J = 21.4 Hz, 1F), -135.00 (dd, J = 21.1, 2.8 Hz, 1F), -143.75 (t, J = 21.7 Hz, 1F). Analytical HPLC: t_R = 4.8 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{37}H_{46}ClF_3N_3O_5$ $[M+H]^+$ 704.3073, found 704.3052.



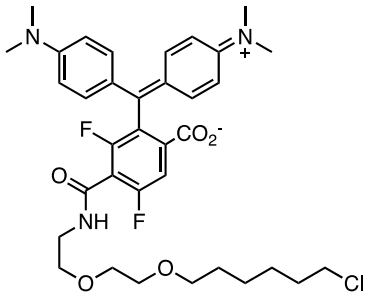
S19: **14** (253 mg, 0.570 mmol, 1 eq) and **S11** (87.8 mg, 0.570 mmol, 1 eq) were combined in DMF (12.6 mL) under argon, and DIEA (203 µL, 1.139 mmol, 2 eq) was added. The reaction was stirred at room temperature for 4 h, then evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–30% EtOAc/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (225 mg, 67%). For characterisation by NMR spectroscopy, an analytical fraction of the product was further purified by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–25% EtOAc/cyclohexane, linear gradient). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.14 (d, J = 8.4 Hz, 4H), 6.69 (d, J = 8.1 Hz, 4H), 5.52 (m, 1H), 3.90 – 3.81 (m, 1H), 3.68 (d, J = 13.3 Hz, 1H), 2.97 (s, 12H), 1.90 – 1.53 (m, 6H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -112.41 (d, J = 22.5 Hz, 1F), -127.44 (d, J = 20.1 Hz, 1F), -139.61 (t, J = 21.5 Hz, 1F). HRMS (ESI) calcd for $C_{32}H_{30}F_3N_4O_4$ $[M+H]^+$ 591.2214, found 591.2211.



S20: S19 (616 mg, 1.04 mmol) was taken up in CH_2Cl_2 (69 mL), triethylsilane (7 mL) and trifluoroacetic acid (14 mL) were added. The reaction was stirred at room temperature for 4 h. Toluene (10 mL) was added, and the reaction mixture was evaporated to dryness. To the residue was added a solution of methanol (1.06 mL, 26.1 mmol, 25 eq) and triethylamine (1.45 mL, 10.4 mmol, 10 eq) in CH_2Cl_2 (12 mL), and the reaction was stirred at room temperature for 15 h. The solvent was evaporated under reduced pressure and purification was performed by silica gel column chromatography (Biotage Sfär Duo 25 g, 0–20% EtOAc/cyclohexane, linear gradient) to afford the title compound as a blue-green solid (161 mg, 32%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.14 (d, J = 8.4 Hz, 4H), 6.65 (d, J = 8.4 Hz, 4H), 3.97 (s, 3H), 2.96 (s, 12H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -115.33 (d, J = 22.6 Hz), -131.62 (d, J = 17.2 Hz, 1F), -141.97 (t, J = 20.0 Hz, 1F). Analytical HPLC: t_R = 5.2 min, 97.3% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{26}H_{24}F_3N_2O_4$ $[M+H]^+$ 485.1683, found 485.1680.



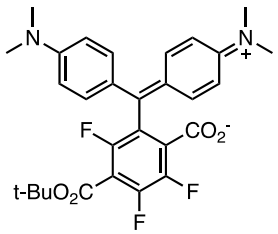
S21: S20 (211 mg, 0.437 mmol, 1 eq) was loaded in a vial which was sealed and evacuated/backfilled with argon. DMF (3 mL) was added, followed by dropwise addition of a solution of $NaBH_4$ (82.6 mg, 2.18 mmol, 5 eq) in DMF (1.5 mL). The reaction was stirred at room temperature for 4 h and carefully quenched with H_2O. The mixture was extracted with CH_2Cl_2 (3x), the organic layers were combined, washed with brine, dried over Na_2SO_4, filtered and evaporated to dryness. Purification was performed by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–15% EtOAc/cyclohexane, linear gradient) to afford the title compound as a blue-green solid (87.5 mg, 43%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.50 (d, J = 7.1 Hz, 1H), 7.15 (d, J = 8.6 Hz, 4H), 6.65 (d, J = 8.6 Hz, 4H), 3.96 (s, 3H), 2.96 (s, 12H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -107.66 (s, 1F), -109.96 (d, J = 4.7 Hz, 1F). Analytical HPLC: t_R = 5.0 min (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). Purity was estimated at 90% based on crystallography. HRMS (ESI) calcd for $C_{26}H_{24}F_2N_2O_4$ $[M+H]^+$ 467.1777, found 467.1776.



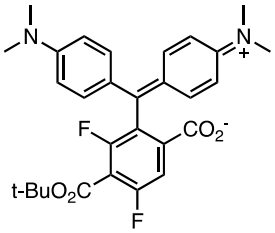
(23-HTL): S21 (117 mg, 0.250 mmol, 1 eq) was dissolved in 1:1 THF: MeOH (22 mL) under argon. 1M NaOH (1 mL, 4 eq) was added dropwise and the reaction stirred at room temperature for 24 h. The reaction was neutralised (1M HCl), diluted in H_2O and the solvents evaporated. The mixture was extracted with *i*-PrOH/$CHCl_3$ 15% (3x), before the organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. This crude intermediate, HaloTag(O2) amine (TFA salt, 843 mg, 2.50 mmol, 10 eq), EDC.HCl (359 mg, 1.87 mmol, 7.5 eq) and HOBt (152 mg, 1.13 mmol, 4.5 eq) were loaded into a sealed vial, which was evacuated/backfilled under argon. CH_2Cl_2 (5 mL) and then DIEA (0.46 mL, 2.6 mmol, 10 eq) were added and the reaction stirred at room temperature for 22 h. Saturated $NaHCO_3$ was added, and the mixture was extracted with CH_2Cl_2 (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to

dryness. Purification was performed by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–50% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (20–80% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness to afford the title compound as a blue-green solid (39 mg, 23%). ^{1}H NMR (400 MHz, $(CD_3)_2SO$, 5 mM DMF) δ 9.09 (t, *J* = 5.5 Hz, 1H), 7.83 (d, *J* = 6.9 Hz, 1H), 7.02 (d, *J* = 8.7 Hz, 4H), 6.70 (d, J = 8.7 Hz, 4H), 3.59 (t, *J* = 6.6 Hz, 2H), 3.54 – 3.49 (m, 4H), 3.48 – 3.44 (m, 2H), 3.40 (q, *J* = 5.6 Hz, 2H), 2.89 (d, *J* = 5.2 Hz, 12H), 1.67 (p, *J* = 6.7 Hz, 2H), 1.45 (p, *J* = 6.8 Hz, 2H), 1.39 – 1.21 (m, 6H). ^{19}F NMR (376 MHz, $(CD_3)_2SO$) δ -110.15 (d, *J* = 6.2 Hz, 1F), -113.84 (d, *J* = 6.1 Hz, 1F). Analytical HPLC: t_R = 4.6 min, ≥99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{35}H_{43}ClF_2N_3O_5$ $[M+H]^+$ 658.2854, found 658.2845.

Alternative method: **S23** (226 mg, 0.444 mmol, 1 eq) was dissolved in CH_2Cl_2 (9.3 mL) under argon. TFA (1.9 mL) was added dropwise and the reaction stirred at room temperature for 4 h. Toluene (2 mL) was added, the reaction mixture evaporated to dryness and then azeotroped with MeOH (3x). This crude intermediate, HaloTag(O2) amine (TFA salt, 1.55 g, 4.66 mmol, 10 eq), EDC.HCl (661 mg, 3.45 mmol, 7.5 eq) and HOBt (280 mg, 2.07 mmol, 4.5 eq) were loaded into a sealed vial, which was evacuated/backfilled under argon. CH_2Cl_2 (5.4 mL) and then DIEA (0.842 mL, 4.83 mmol, 10.5 eq) were added and the reaction stirred at room temperature for 22 h. Saturated $NaHCO_3$ was added, and the mixture was extracted with CH_2Cl_2 (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Purification was performed by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–60% EtOAc/cyclohexane, linear gradient) followed by reverse phase HPLC (20–80% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness before a second column (Biotage Sfär Duo 5 g, 0–55% EtOAc/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (13 mg, 5%).

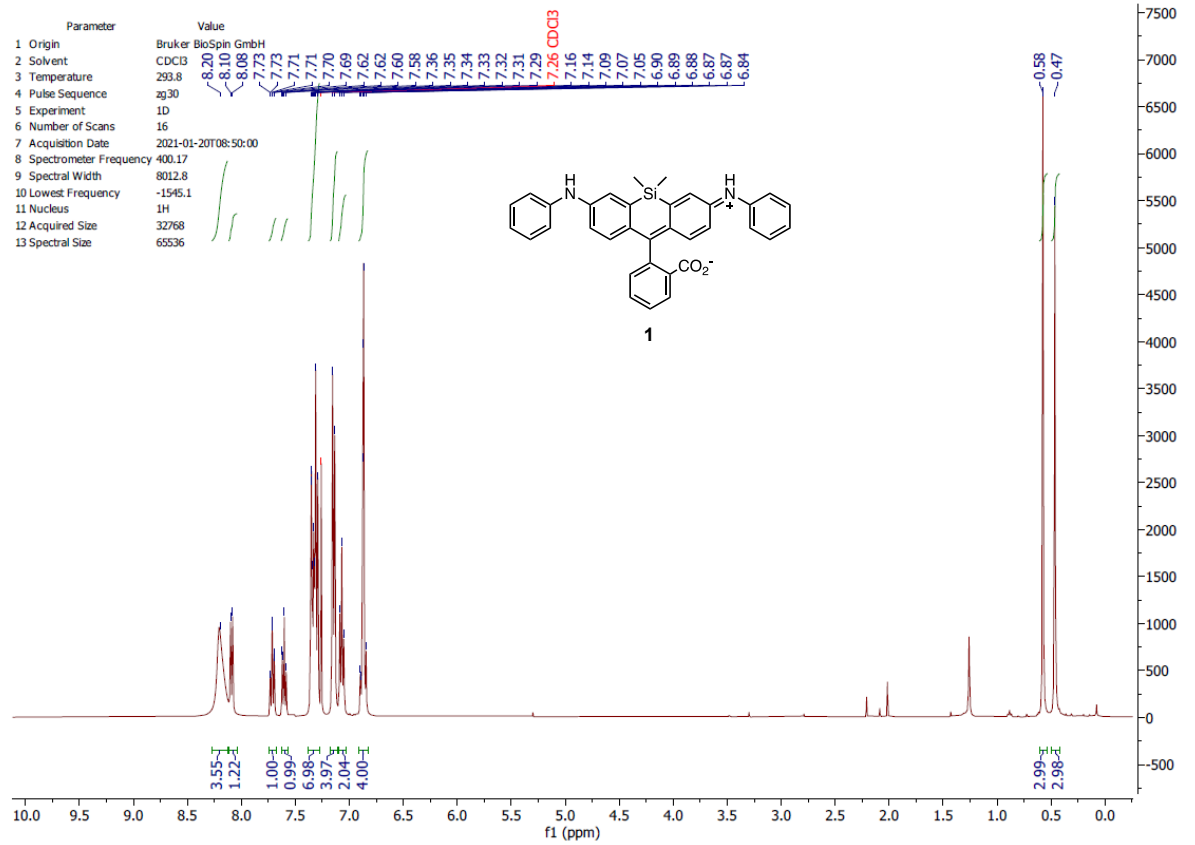


S22: S10 (102 mg, 0.217 mmol, 1 eq) was added to a vial which was sealed and evacuated/backfilled with argon. Toluene (0.5 mL) was added and the mixture heated to 80°C. *N,N*-dimethylformamide di-*tert*-butyl acetal (0.312 mL, 1.30 mmol, 6 eq) was added dropwise. After 15 min, a further 6 eq of *N,N*-dimethylformamide di-*tert*-butyl acetal were then added and the mixture stirred for an additional 30 min to complete the reaction. The mixture was cooled to room temperature, H_2O and saturated $NaHCO_3$ were added, and the aqueous layer was extracted with CH_2Cl_2 (3x). The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–60% TBME/cyclohexane, linear gradient) afforded the title compound as a blue-green solid (76.5 mg, 67%). An analytical fraction of the product was further purified by reverse phase HPLC (20–80% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness to afford a pure sample. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 (d, *J* = 8.8 Hz, 4H), 7.29 (d, *J* = 8.8 Hz, 4H), 3.14 (s, 12H), 1.60 (s, 9H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -116.96 (d, J = 21.9 Hz, 1F), -130.54 (d, *J* = 21.1 Hz, 1F), -140.24 (t, *J* = 21.7 Hz, 1F). Analytical HPLC: t_R = 5.0 min, 94% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{29}H_{30}F_3N_2O_4$ $[M+H]^+$ 527.2152, found 527.2148.

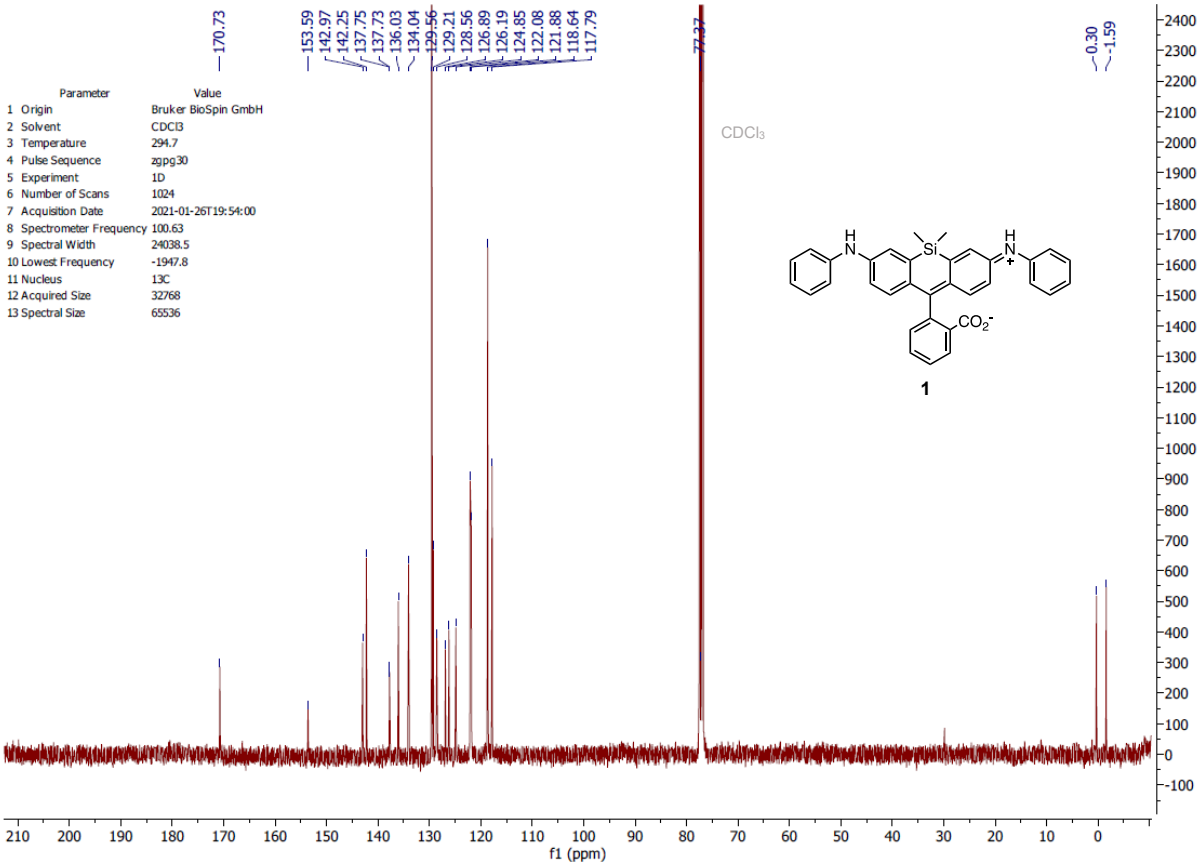


S23: S22 (52.0 mg, 0.099 mmol, 1 eq) was loaded in a vial which was sealed and evacuated/backfilled with argon. DMF (1 mL) was added, followed by dropwise addition of a solution of $NaBH_4$ (18.7 mg, 0.494 mmol, 5 eq) in DMF (0.5 mL). The reaction was stirred at room temperature for 4 h and carefully quenched with H_2O. The mixture was extracted with CH_2Cl_2 (3x), the organic layers were combined, washed with brine, dried over Na_2SO_4, filtered and evaporated to dryness. Purification was performed by silica gel column chromatography (Biotage Sfär Duo 10 g, 0–30% EtOAc/cyclohexane, linear gradient) to afford the title compound as a blue-green solid (30.3 mg, 60%). An analytical fraction of the product was further purified by reverse phase HPLC (30–70% CH_3CN/H_2O, linear gradient with constant 0.1% v/v TFA additive). The pooled HPLC product fractions were neutralised with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were dried over Na_2SO_4, filtered and evaporated to dryness to afford a pure sample. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 (d, J = 6.9 Hz, 1H), 7.17 (d, J = 8.5 Hz, 4H), 6.70 (br s, 4H), 2.97 (s, 12H), 1.57 (s, 9H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -111.89 – -111.97 (s, 1F), -116.95 (d, J = 25.1 Hz, 1F). Analytical HPLC: t_R = 5.0 min, >99% purity (5–95% CH_3CN/H_2O, gradient with constant 0.1% formic acid additive, 8 min run, 0.5 mL/min flow, UV detection at 254 nm). HRMS (ESI) calcd for $C_{29}H_{31}F_3N_2O_4$ $[M+H]^+$ 509.2246, found 509.2241.

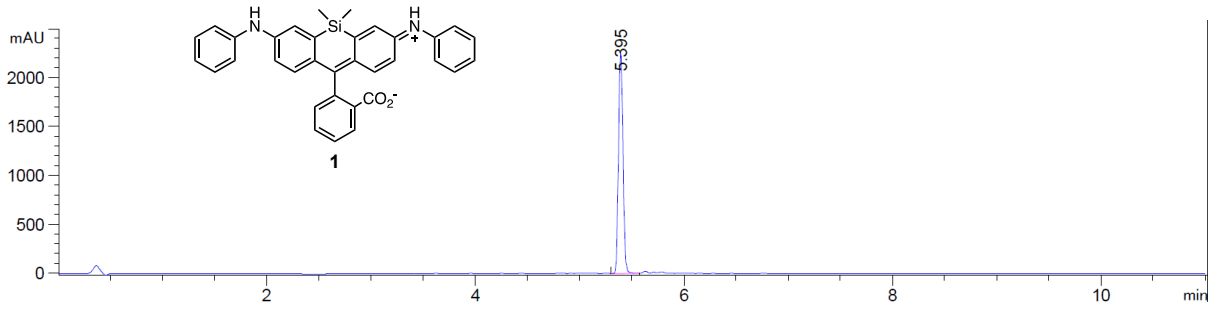
VIII.2 NMR spectra and HPLC traces

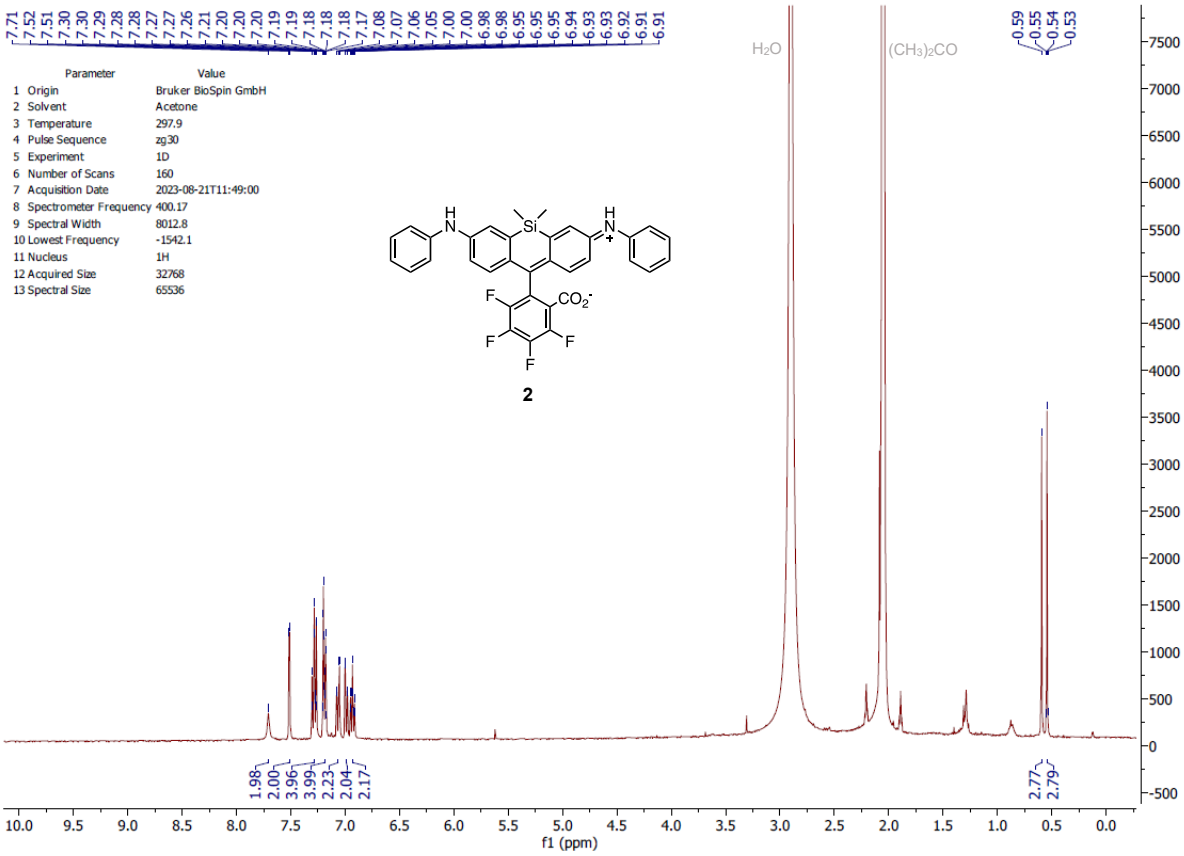


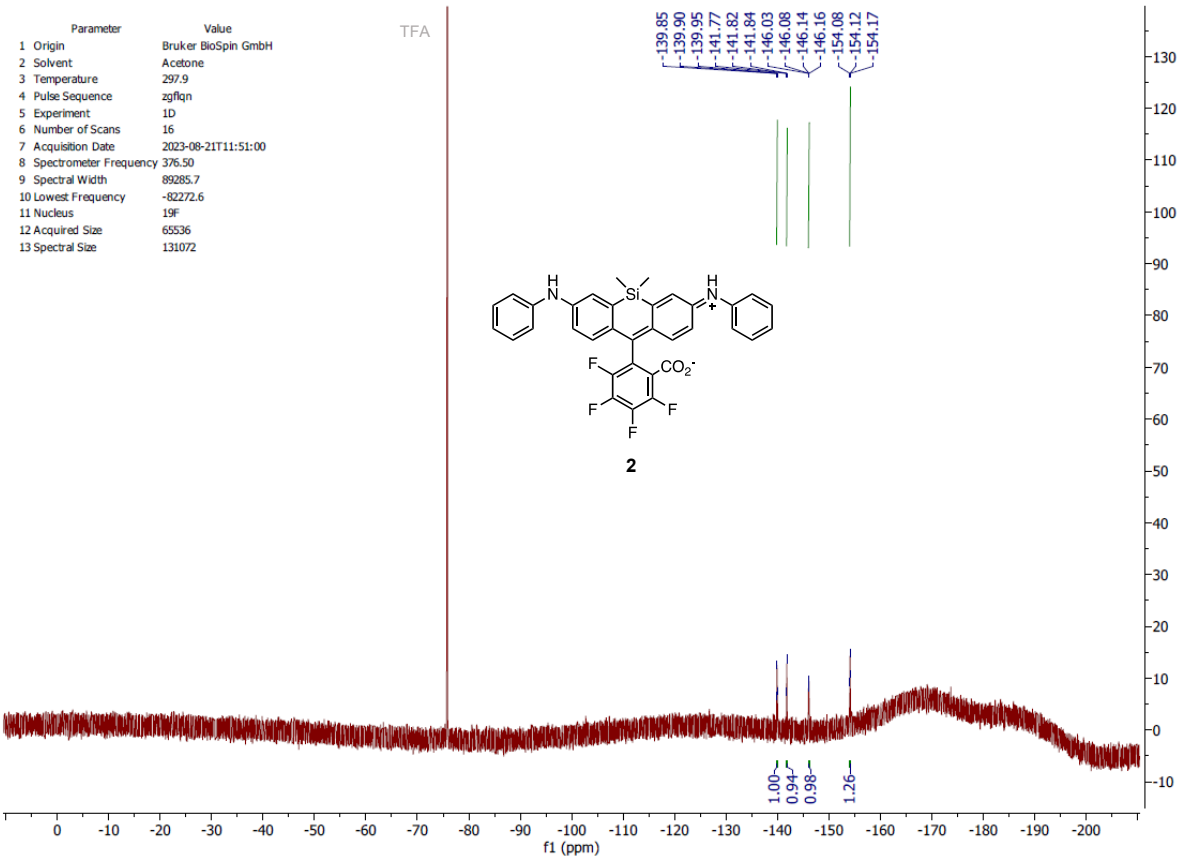
1

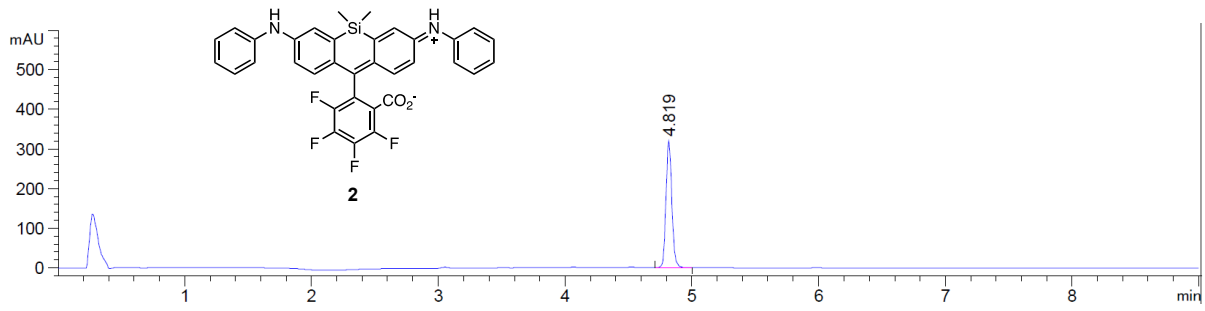


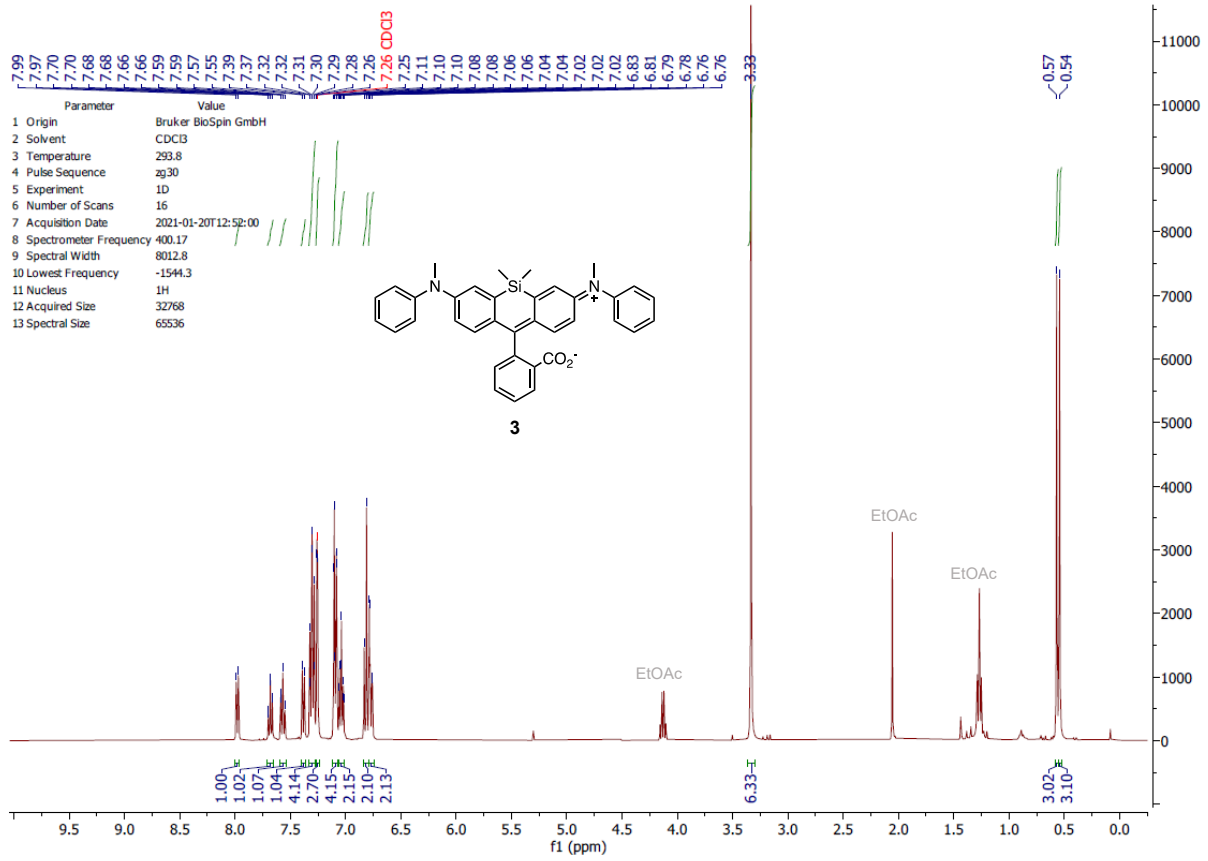
1

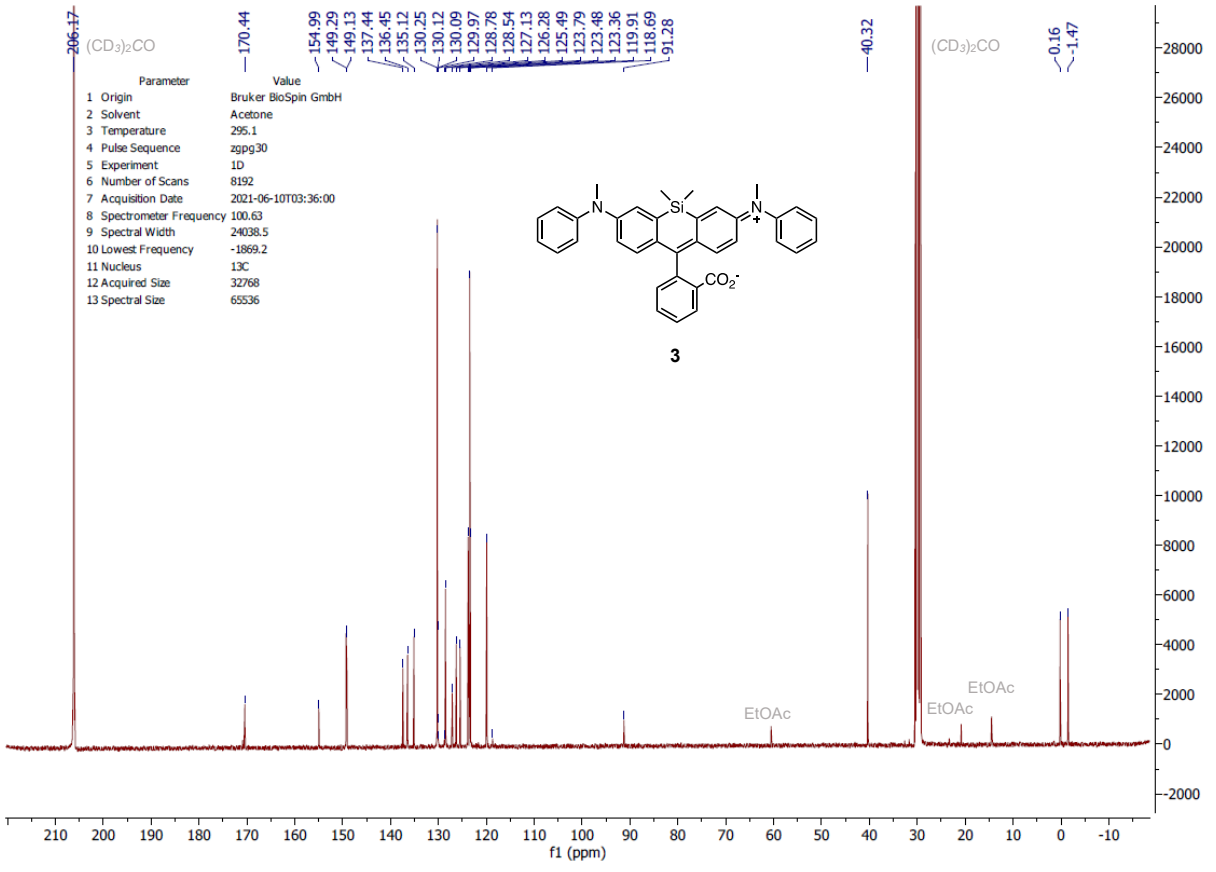

mAU
2000
1500
1000
500
0
2
4
6
8
10
min
5.395
1

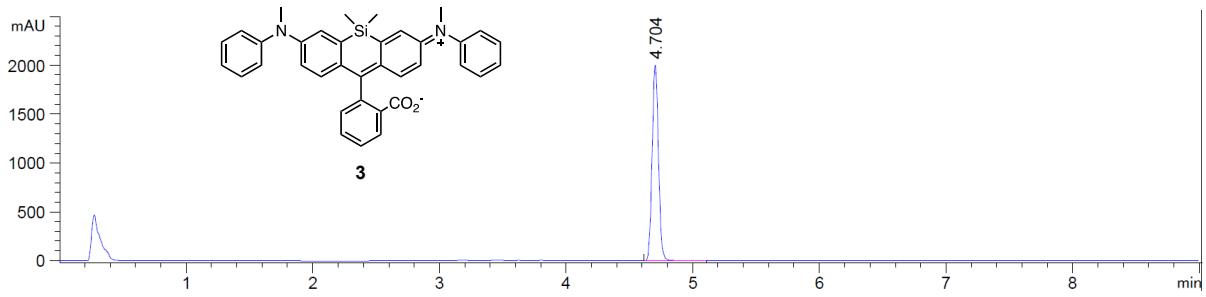

7.71 7.52 7.51 7.30 7.30 7.29 7.28 7.28 7.27 7.27 7.26 7.21 7.20 7.20 7.20 7.19 7.19 7.18 7.18 7.18 7.17 7.08 7.07 7.06 7.05 7.00 7.00 6.98 6.98 6.95 6.95 6.95 6.94 6.93 6.93 6.92 6.91 6.91
H2O
(CH3)2CO
0.59 0.55 0.54 0.53
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 297.9
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 160
7 Acquisition Date 2023-08-21T11:49:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.1
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
2
1.98 2.00 3.96 3.99 2.23 2.04 2.17
2.77 2.79
f1 (ppm)


TFA
-139.85 -139.90 -139.95 -141.77 -141.82 -141.84 -146.03 -146.08 -146.14 -146.16 -154.08 -154.12 -154.17
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 297.9
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-21T11:51:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82272.6
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
2
1.00 0.94 0.98 1.26
f1 (ppm)

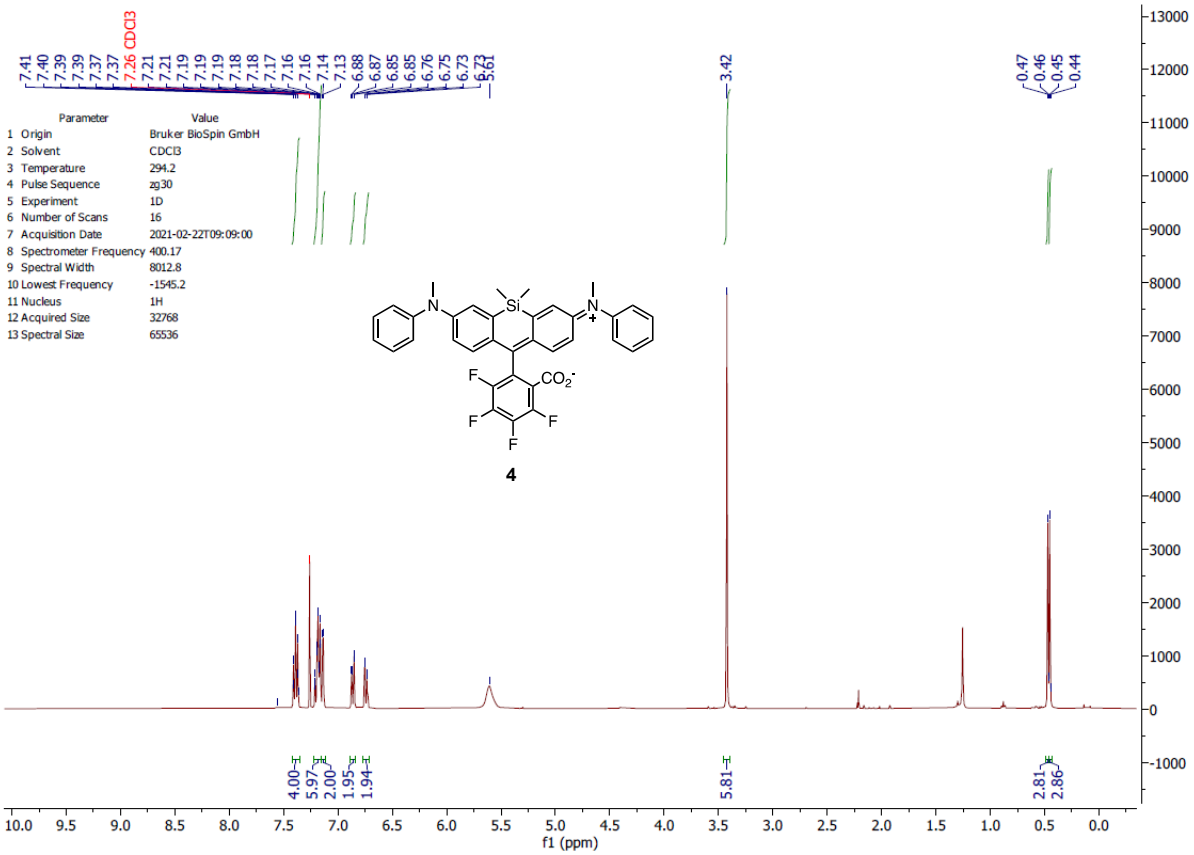

mAU
500
400
300
200
100
0
1
2
3
4
5
6
7
8
min
4.819
F
F
F
F
CO2
Si
H
N
H
N
2


7.99 7.97 7.70 7.70 7.68 7.68 7.66 7.66 7.59 7.59 7.57 7.55 7.39 7.37 7.32 7.32 7.31 7.30 7.29 7.28 7.26 CDCl3 7.25 7.11 7.10 7.10 7.08 7.08 7.06 7.06 7.04 7.04 7.02 7.02 7.02 6.83 6.81 6.79 6.78 6.76 6.76
3.33
0.57 0.54
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-01-20T12:52:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1544.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
3
EtOAc
EtOAc
EtOAc
1.00 1.02 1.07 1.04 4.14 2.70 4.15 2.15 2.10 2.13
6.33
3.02 3.10
f1 (ppm)

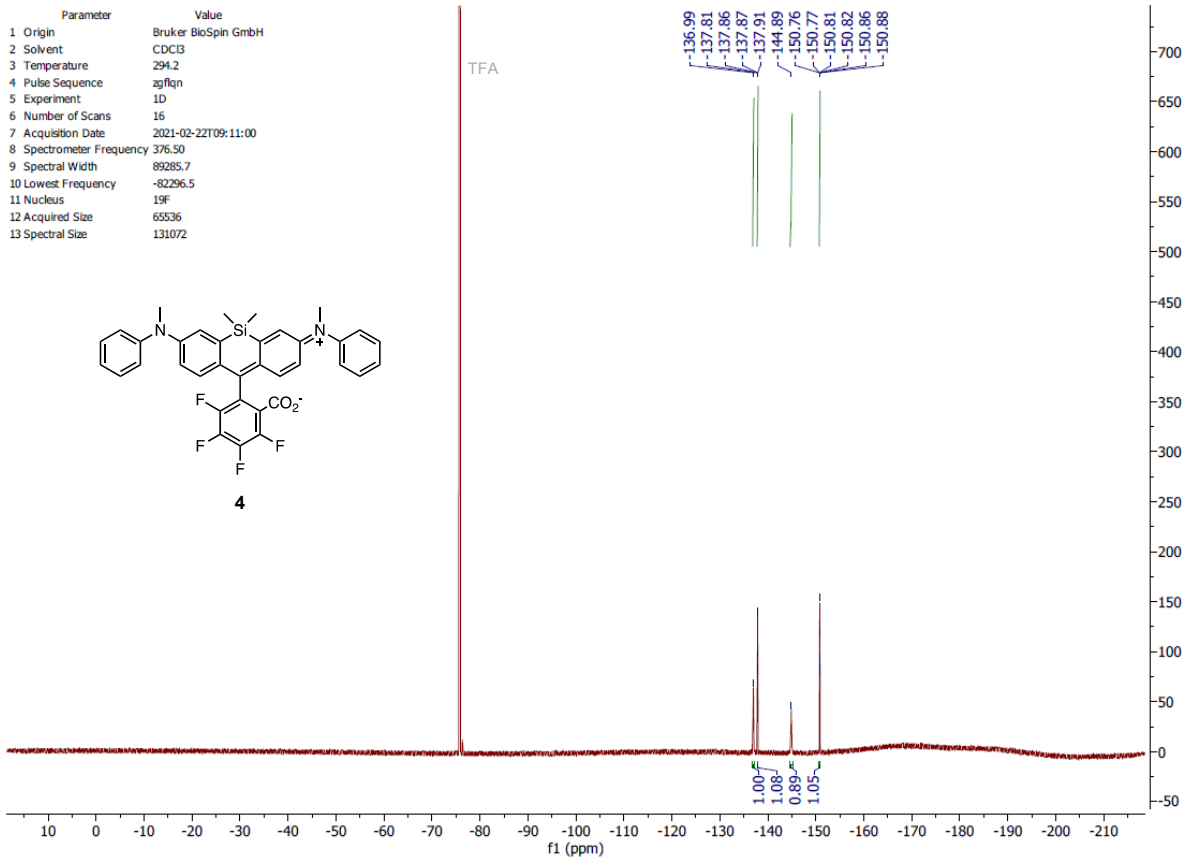

206.17
(CD3)2CO
170.44
154.99 149.29 149.13 137.44 136.45 135.12 130.25 130.12 130.09 129.97 128.78 128.54 127.13 126.28 125.49 123.79 123.48 123.36 119.91 118.69 91.28
40.32
(CD3)2CO
0.16 -1.47
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 295.1
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 8192
7 Acquisition Date 2021-06-10T03:36:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1869.2
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
3
EtOAc
EtOAc
EtOAc
f1 (ppm)



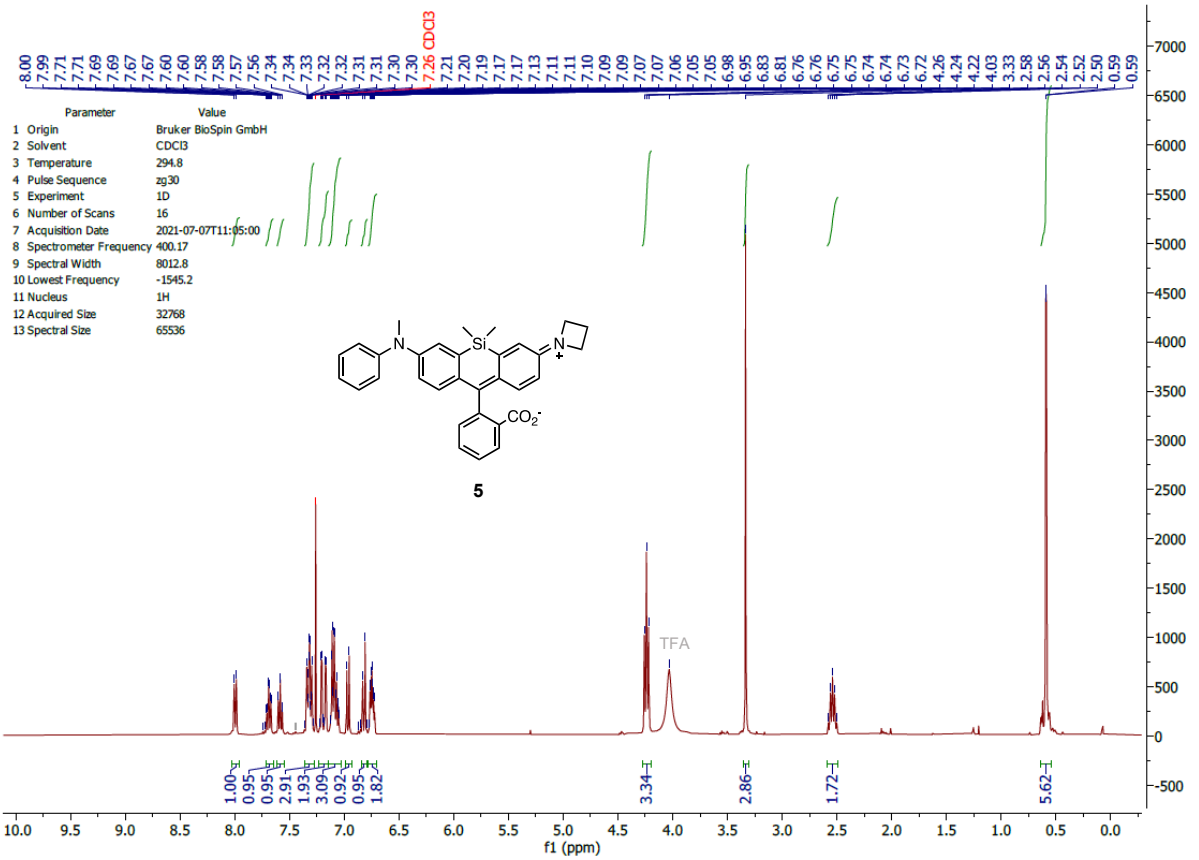
3


7.41 7.40 7.39 7.39 7.37 7.37 7.26 CDCl3 7.21 7.21 7.19 7.19 7.19 7.18 7.18 7.17 7.16 7.16 7.14 7.13 6.88 6.87 6.85 6.85 6.76 6.75 6.73 6.73 6.61
3.42
0.47 0.46 0.45 0.44
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.2
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-02-22T09:09:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.2
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
4.00 5.97 2.00 1.95 1.94
5.81
2.81 2.86
f1 (ppm)

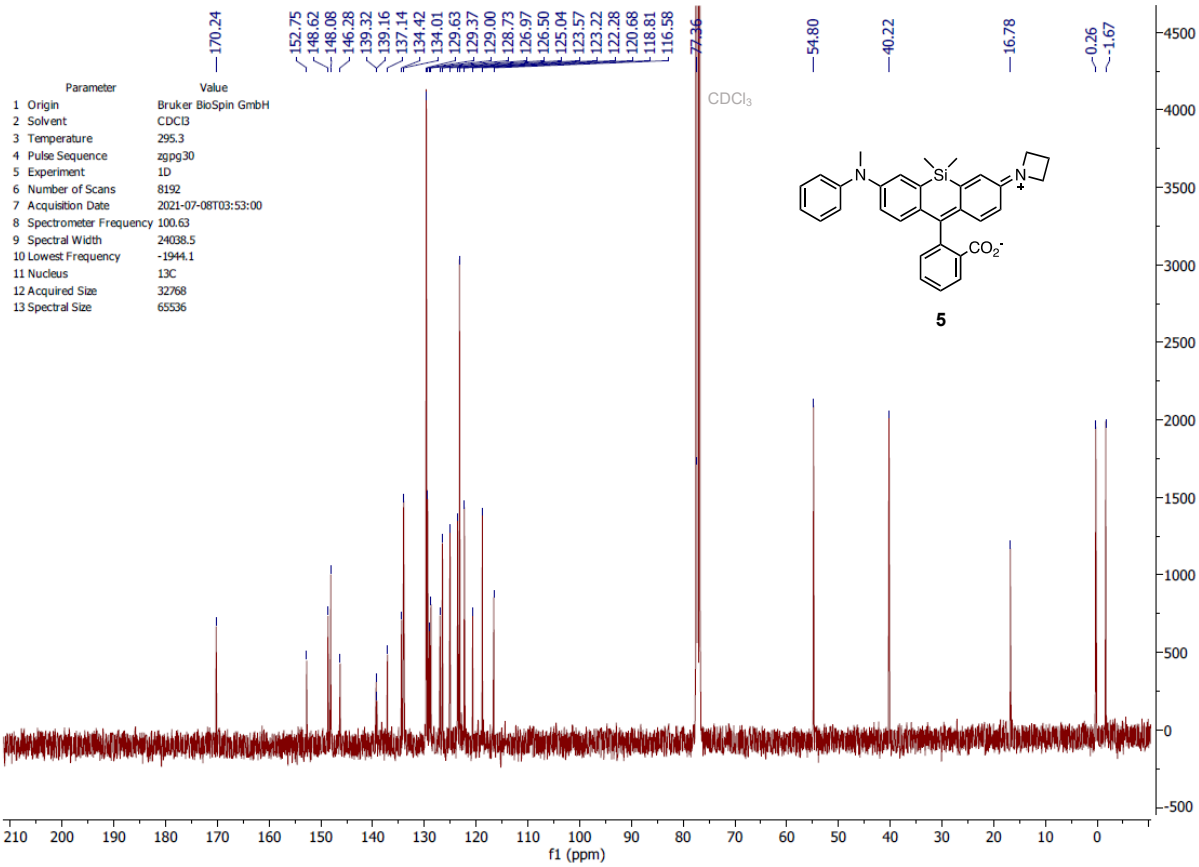
4


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.2
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-02-22T09:11:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82296.5
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
-136.99 -137.81 -137.86 -137.87 -137.91 -144.89 -150.76 -150.77 -150.81 -150.82 -150.86 -150.88
1.00 1.08 0.89 1.05
f1 (ppm)

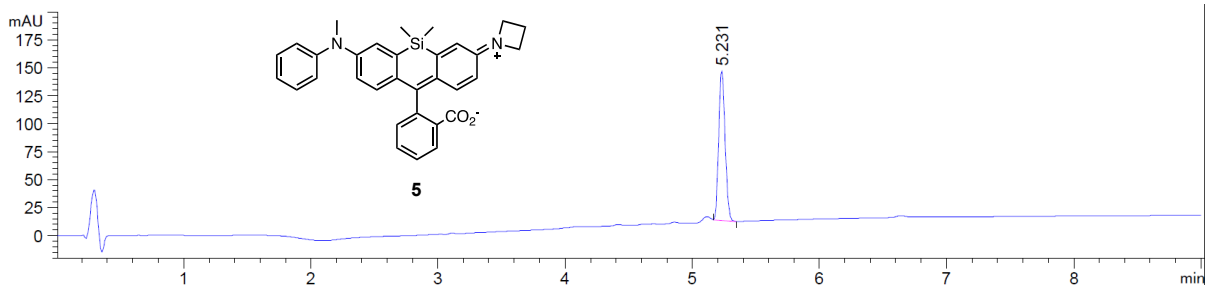
4


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-07-07T11:05:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.2
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
CDCl3
TFA
f1 (ppm)

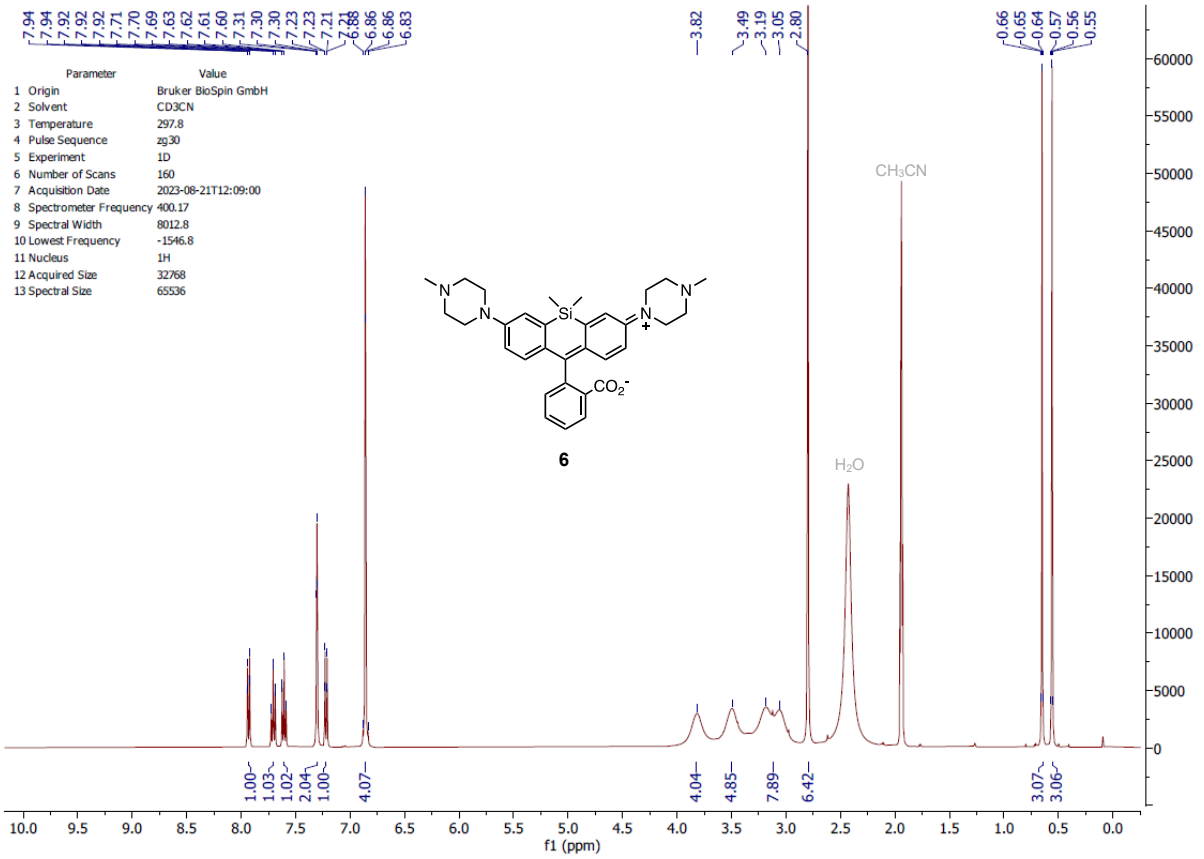
5

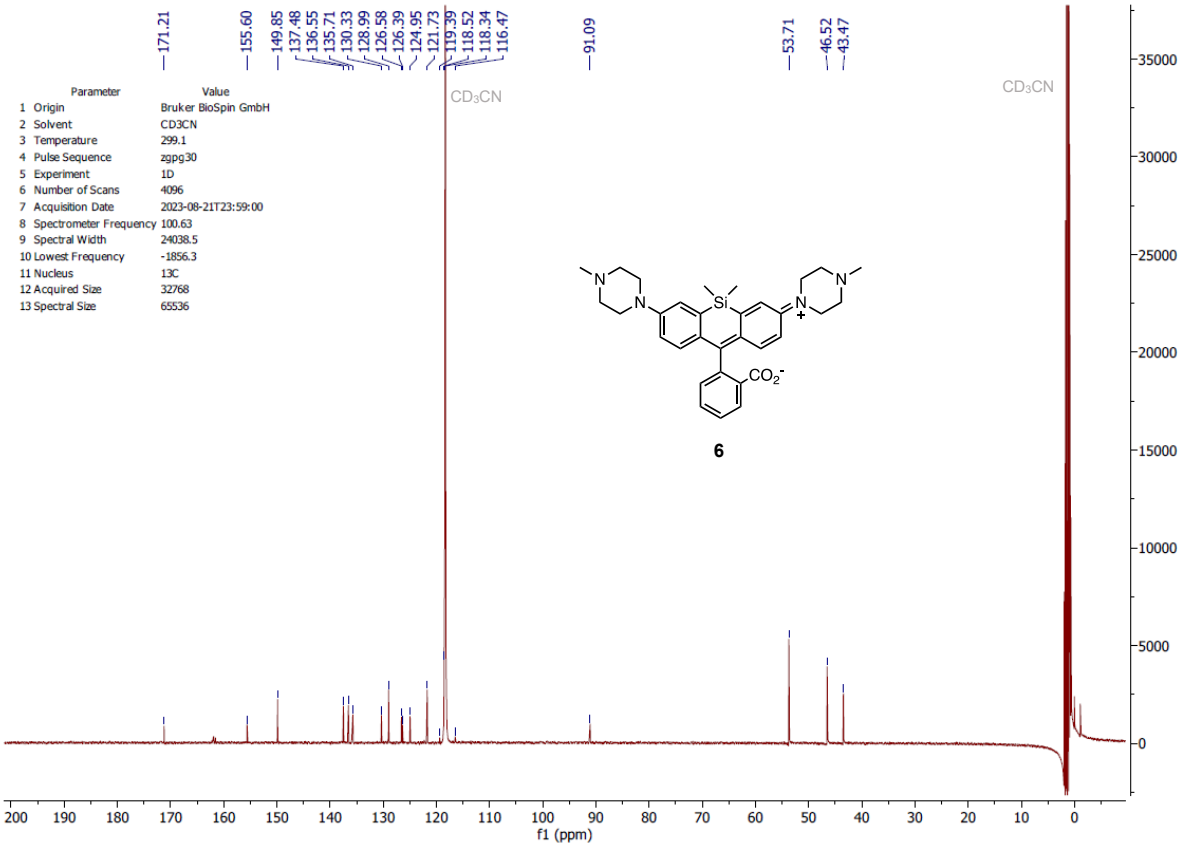

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 295.3
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 8192
7 Acquisition Date 2021-07-08T03:53:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1944.1
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
CDCl3
f1 (ppm)

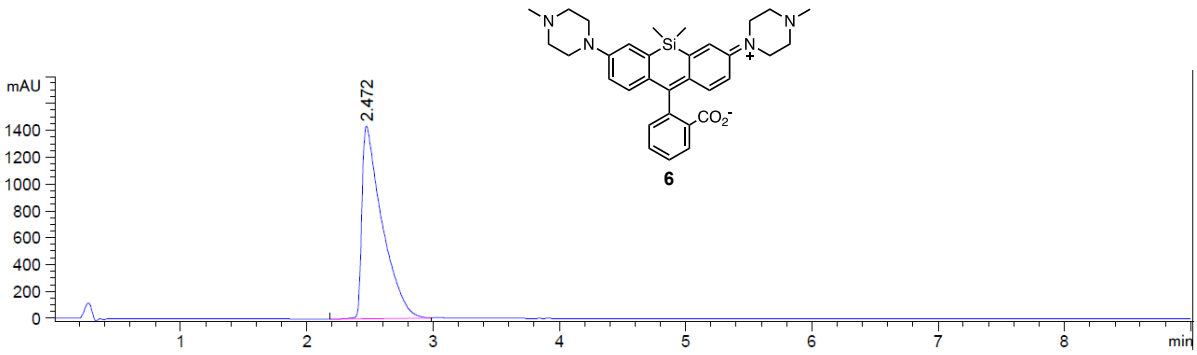
5



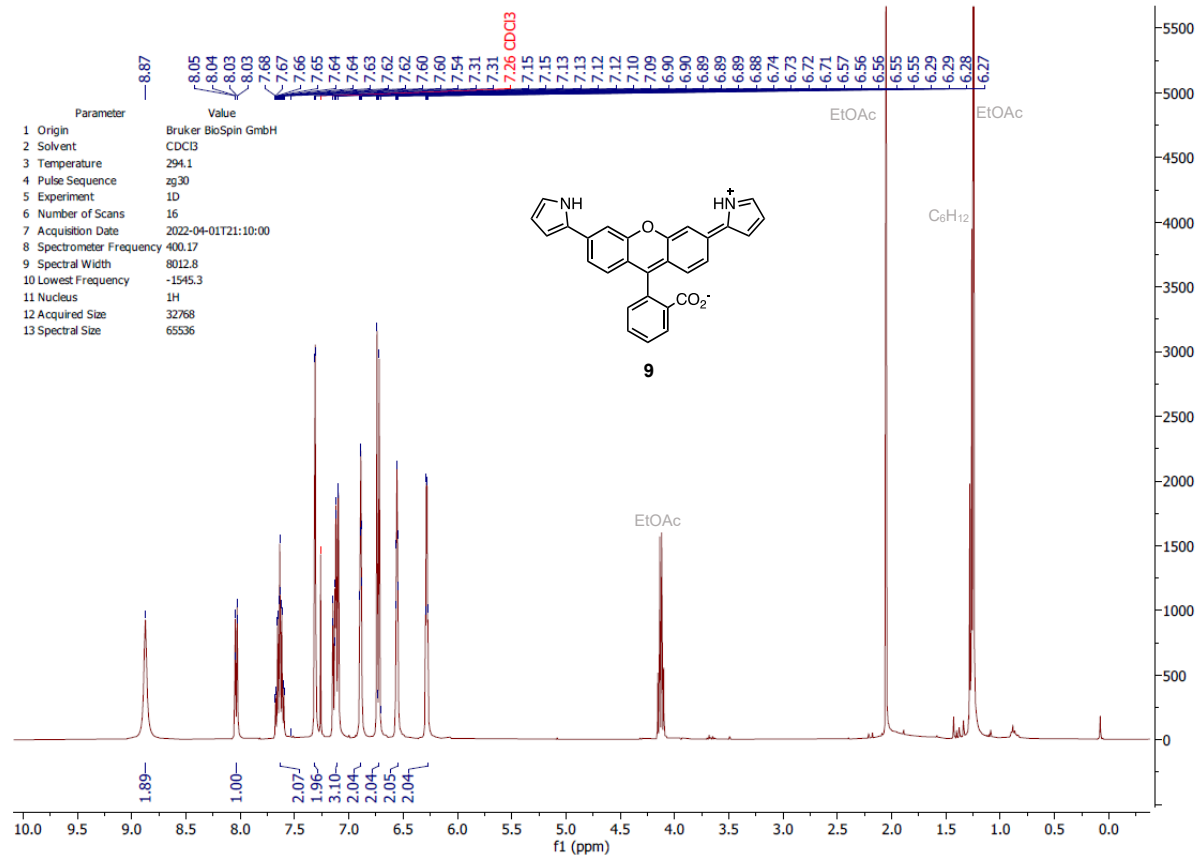
5

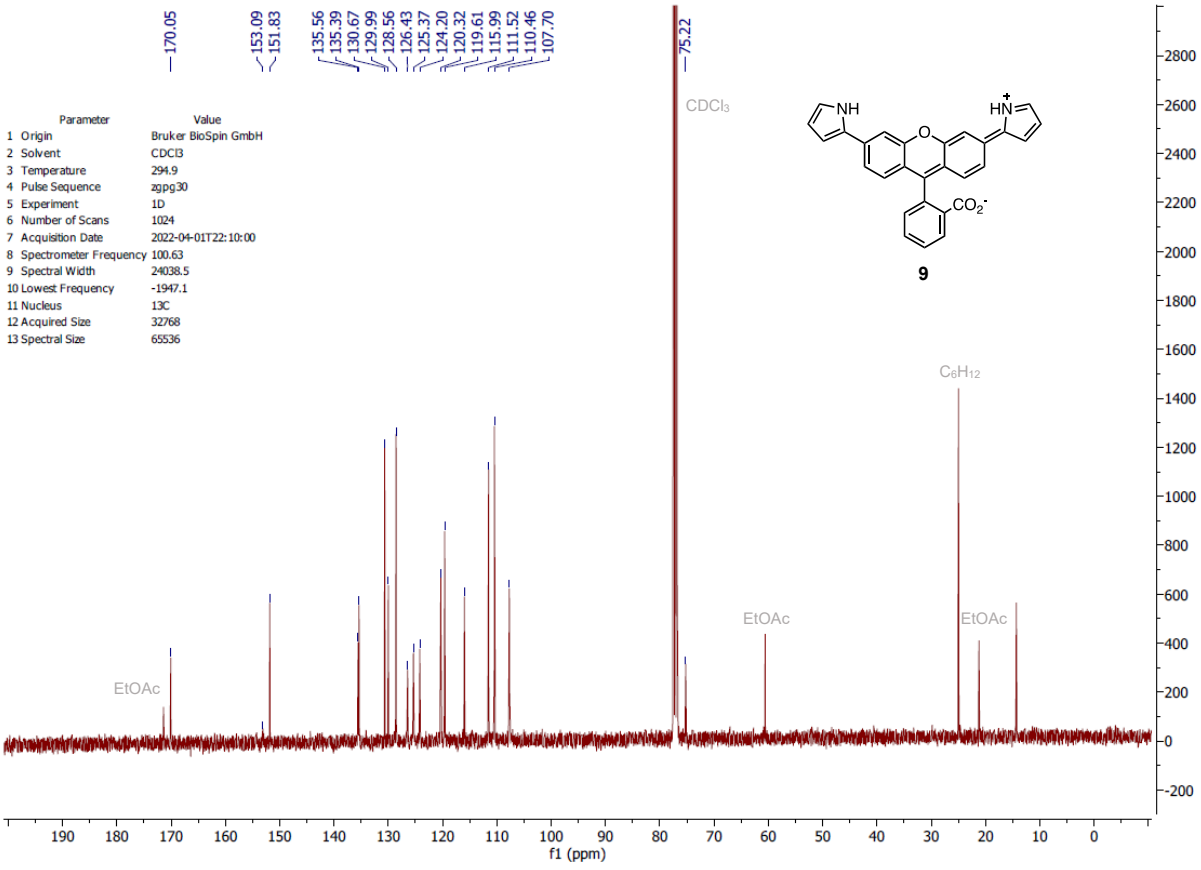

7.94
7.94
7.92
7.92
7.92
7.71
7.70
7.69
7.63
7.62
7.61
7.60
7.31
7.30
7.30
7.23
7.23
7.21
6.88
6.86
6.86
6.83
3.82
3.49
3.19
3.05
2.80
0.66
0.65
0.64
0.57
0.56
0.55
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 297.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 160
7 Acquisition Date 2023-08-21T12:09:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1546.8
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
CH3CN
H2O
6
1.00
1.03
1.02
2.04
1.00
4.07
4.04
4.85
7.89
6.42
3.07
3.06
f1 (ppm)

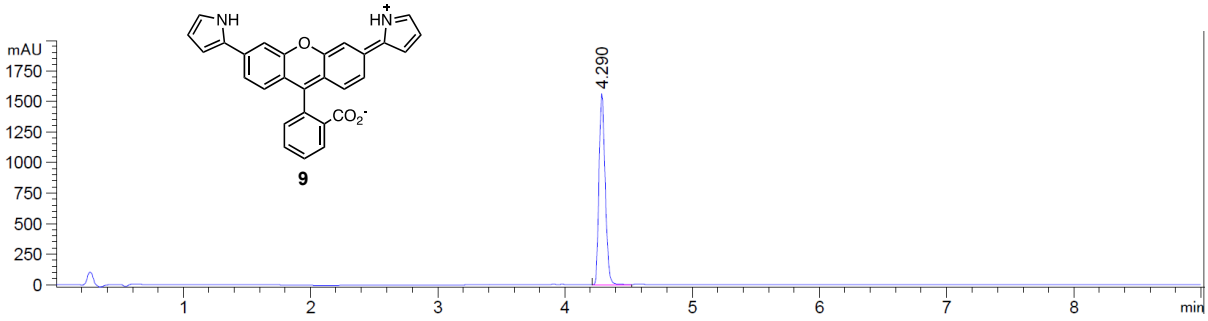

171.21
155.60
149.85
137.48
136.55
135.71
130.33
128.99
126.58
126.39
124.95
121.73
119.39
118.52
118.34
116.47
91.09
53.71
46.52
43.47
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 299.1
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 4096
7 Acquisition Date 2023-08-21T23:59:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1856.3
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
CD3CN
CD3CN
6
f1 (ppm)

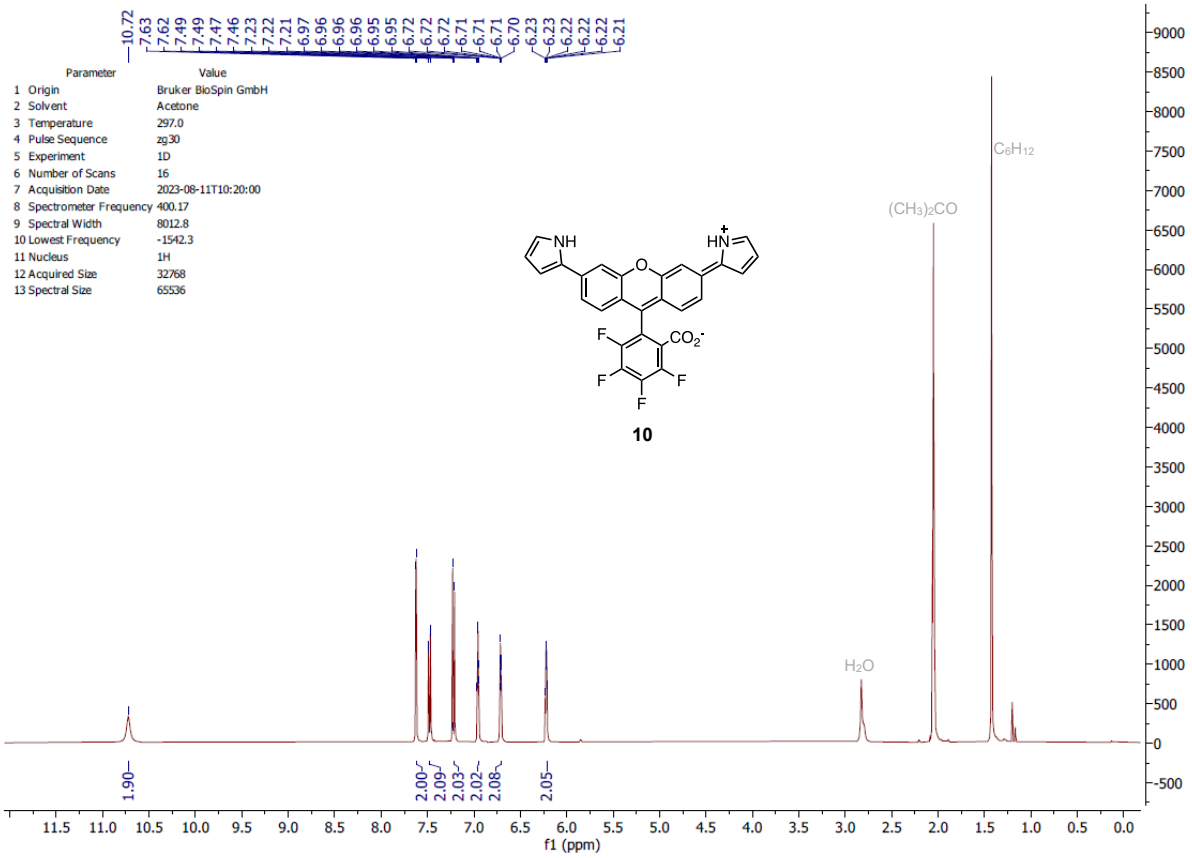

mAU
1400
1200
1000
800
600
400
200
0
2.472
1
2
3
4
5
6
7
8
min
N
N
Si
N
CO2-

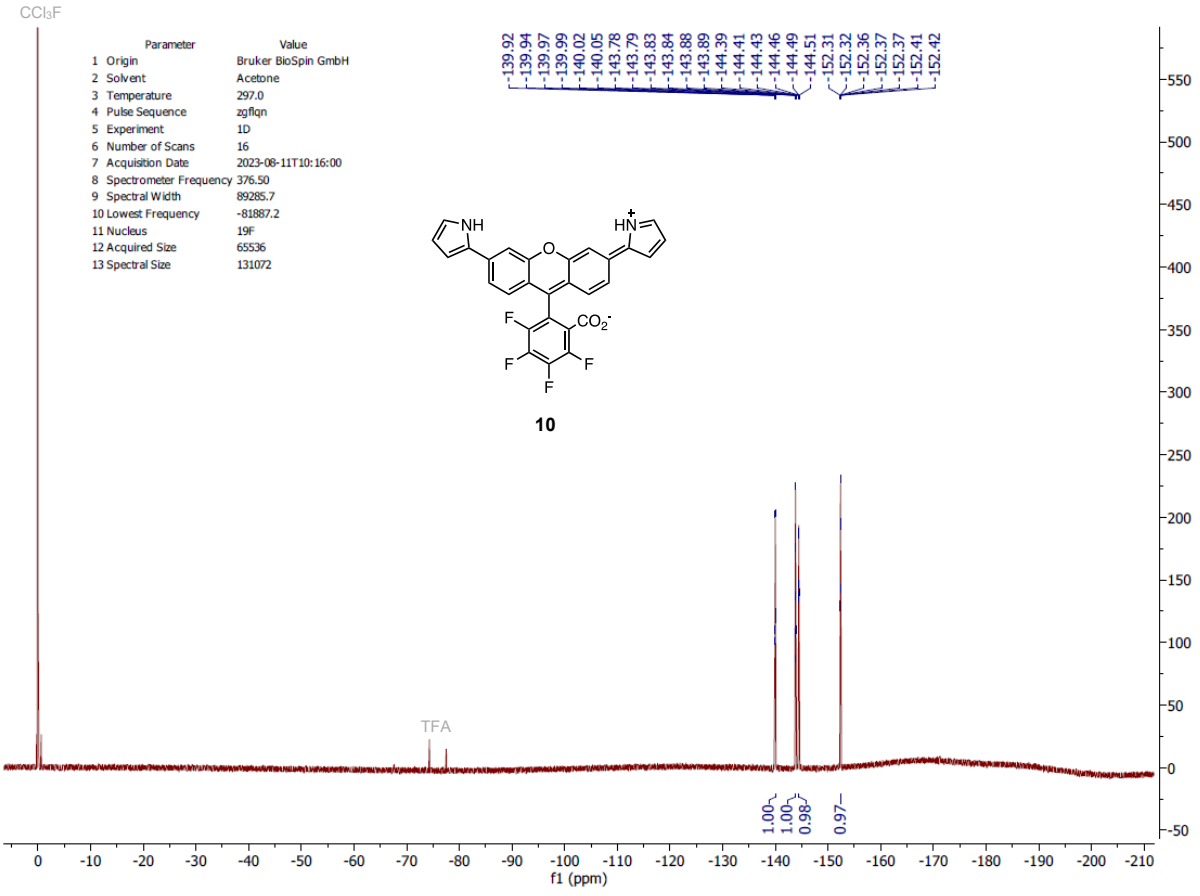
6

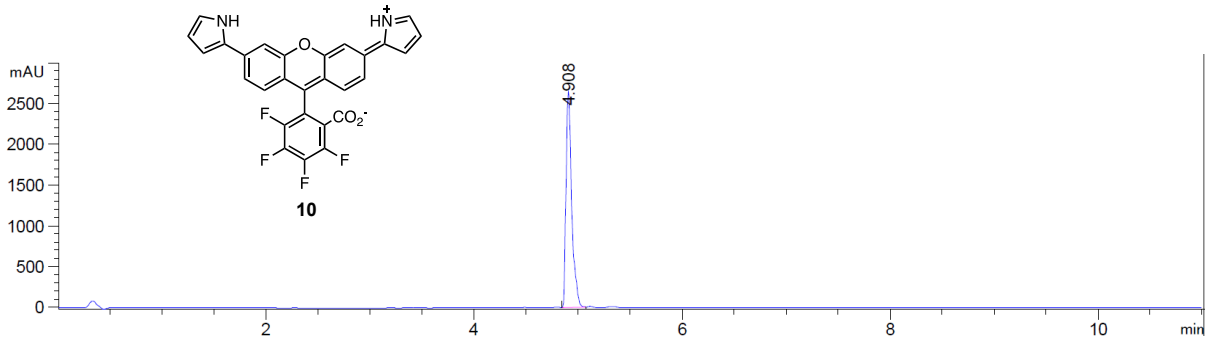

8.87
8.05
8.04
8.03
8.03
7.68
7.67
7.66
7.65
7.64
7.64
7.63
7.62
7.62
7.60
7.60
7.54
7.31
7.31
7.26 CDCl3
7.15
7.15
7.13
7.13
7.12
7.12
7.10
7.09
6.90
6.90
6.89
6.89
6.89
6.88
6.74
6.73
6.72
6.71
6.57
6.56
6.56
6.55
6.55
6.29
6.29
6.28
6.27
EtOAc
EtOAc
C6H12
EtOAc
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.1
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-04-01T21:10:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
NH
HN
O
CO2-
9
1.89
1.00
2.07
1.96
3.10
2.04
2.04
2.05
2.04
f1 (ppm)

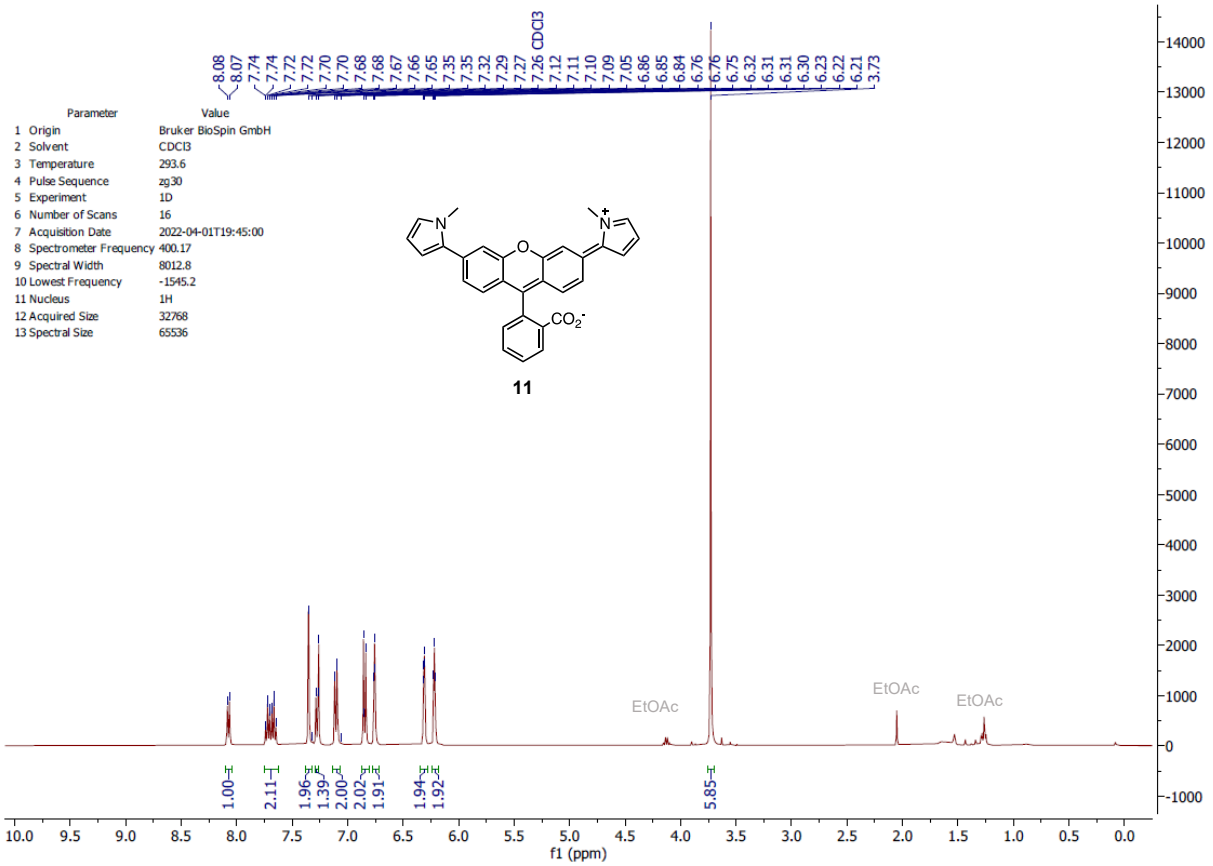

170.05
153.09
151.83
135.56
135.39
130.67
129.99
128.56
126.43
125.37
124.20
120.32
119.61
115.99
111.52
110.46
107.70
75.22
CDCl3
EtOAc
EtOAc
EtOAc
C6H12
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.9
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 1024
7 Acquisition Date 2022-04-01T22:10:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1947.1
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
NH
HN
O
CO2-
9
f1 (ppm)

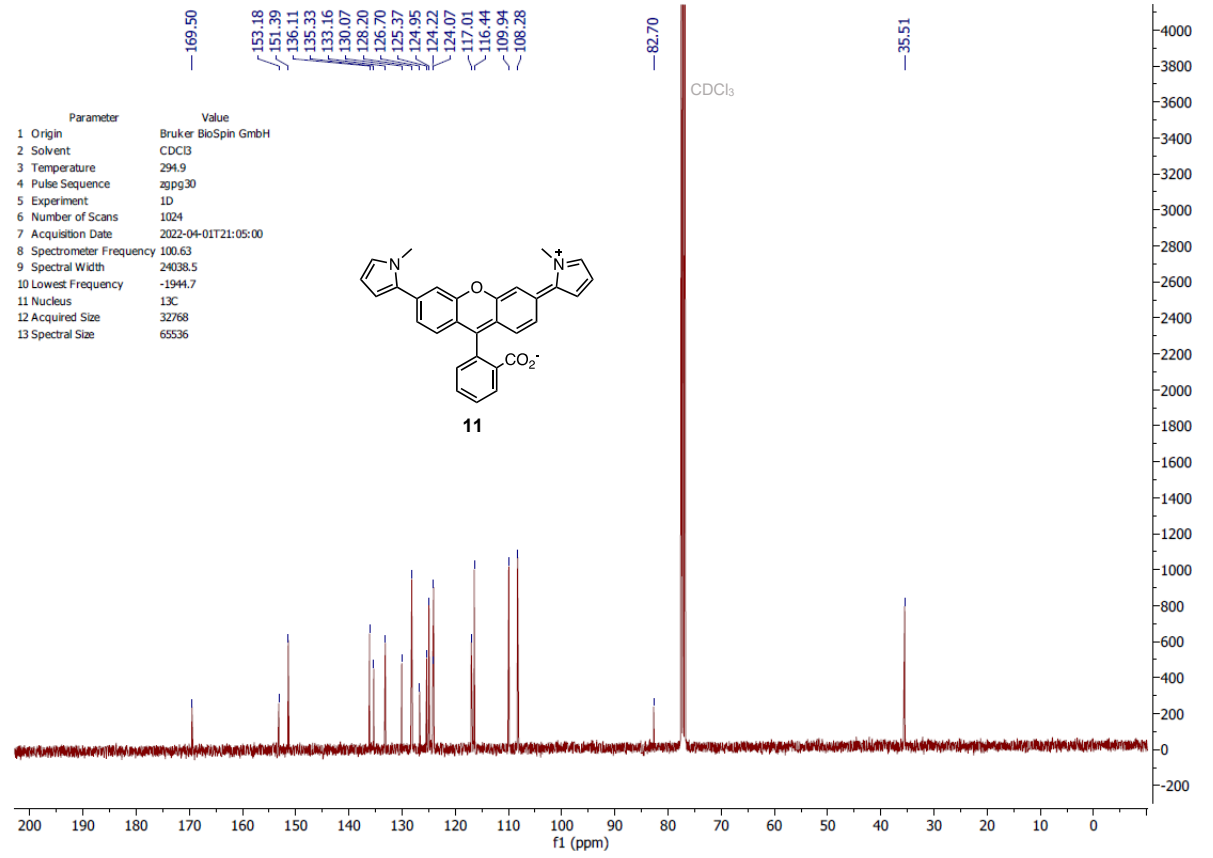

NH
HN
O
CO2-
9
mAU
1750
1500
1250
1000
750
500
250
0
4.290
1
2
3
4
5
6
7
8
min

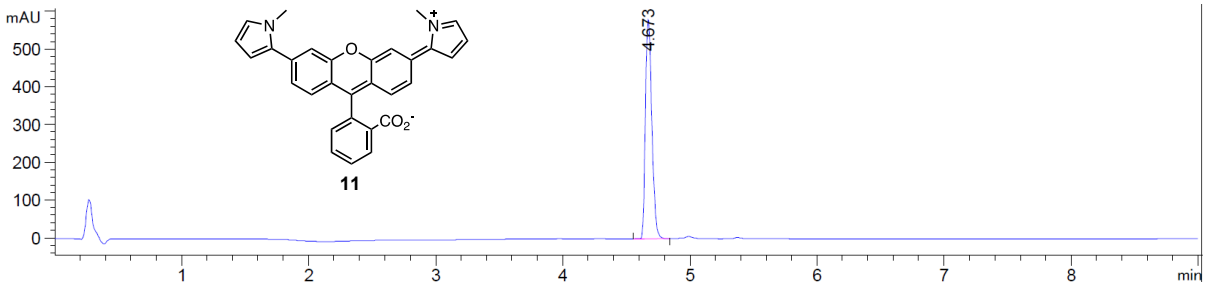

10.72
7.63
7.62
7.49
7.49
7.47
7.46
7.23
7.22
7.21
6.97
6.96
6.96
6.95
6.95
6.72
6.72
6.71
6.71
6.70
6.23
6.22
6.22
6.21
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 297.0
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-11T10:20:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
NH
HN
O
F
CO2-
F
F
F
10
C6H12
(CH3)2CO
H2O
1.90
2.00
2.09
2.03
2.02
2.08
2.05
f1 (ppm)


CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 297.0
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-11T10:16:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81887.2
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
-139.92
-139.94
-139.97
-139.99
-140.02
-140.05
-143.78
-143.79
-143.83
-143.84
-143.88
-143.89
-144.39
-144.41
-144.43
-144.46
-144.49
-144.51
-152.31
-152.32
-152.36
-152.37
-152.37
-152.41
-152.42
NH
HN
O
F
CO2-
F
F
F
10
TFA
1.00
1.00
0.98
0.97
f1 (ppm)

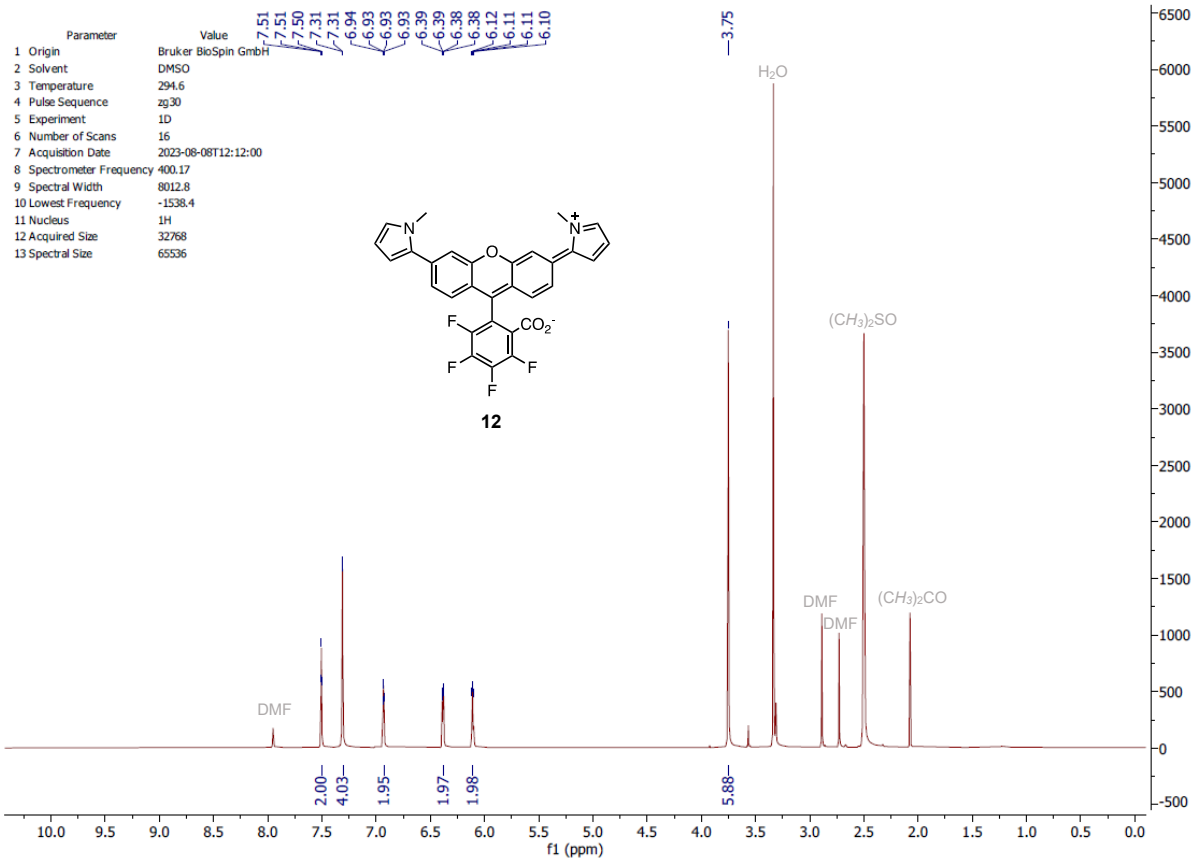

NH
HN+
O
F
CO2-
F
F
F
10
mAU
2500
2000
1500
1000
500
0
4.908
2
4
6
8
10
min


8.08 8.07 7.74 7.74 7.72 7.72 7.70 7.70 7.68 7.68 7.67 7.66 7.65 7.35 7.35 7.32 7.29 7.27 7.26 CDCl3 7.12 7.11 7.10 7.09 7.05 6.86 6.85 6.84 6.76 6.76 6.75 6.32 6.31 6.31 6.30 6.23 6.22 6.21 3.73
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.6
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-04-01T19:45:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.2
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
11
EtOAc
EtOAc
EtOAc
1.00 2.11 1.96 1.39 2.00 2.02 1.91 1.94 1.92 5.85
f1 (ppm)

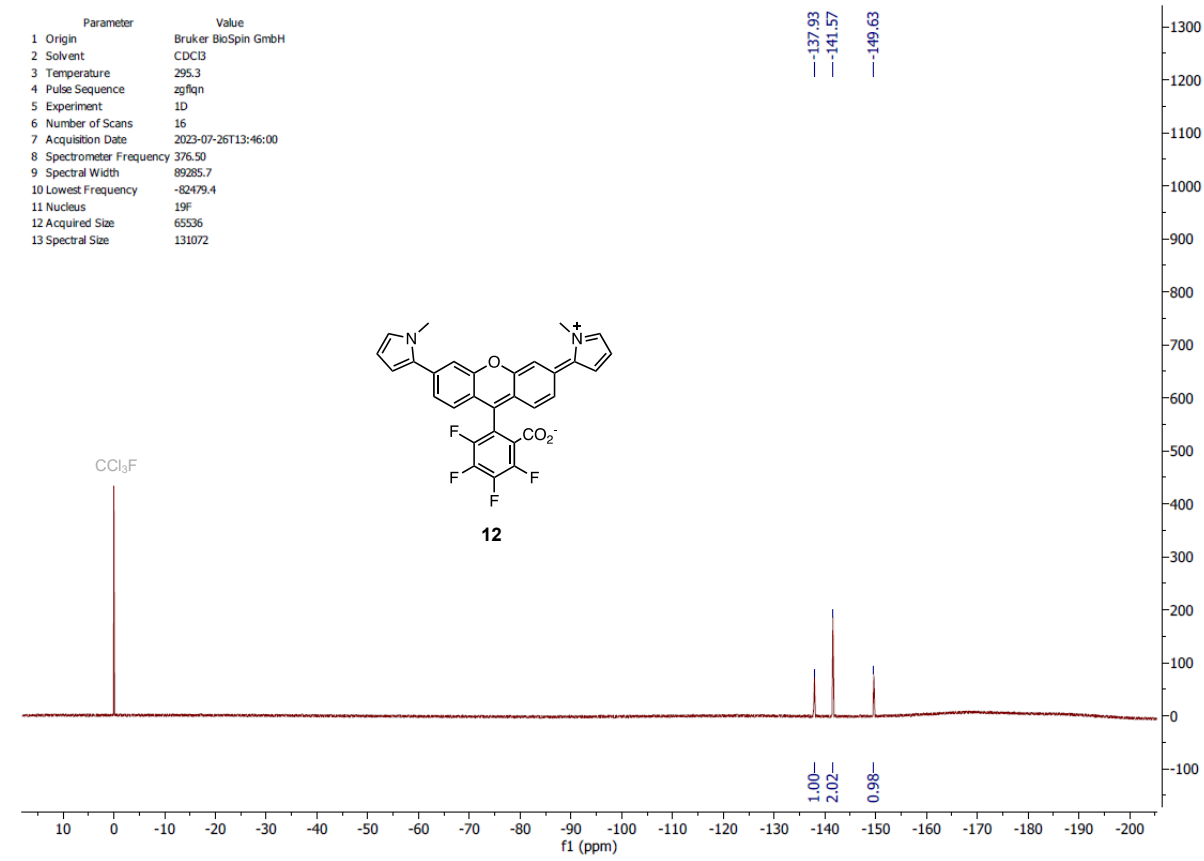

169.50 153.18 151.39 136.11 135.33 133.16 130.07 128.20 126.70 125.37 124.95 124.22 124.07 117.01 116.44 109.94 108.28 82.70 CDCl3 35.51
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.9
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 1024
7 Acquisition Date 2022-04-01T21:05:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1944.7
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
11
f1 (ppm)



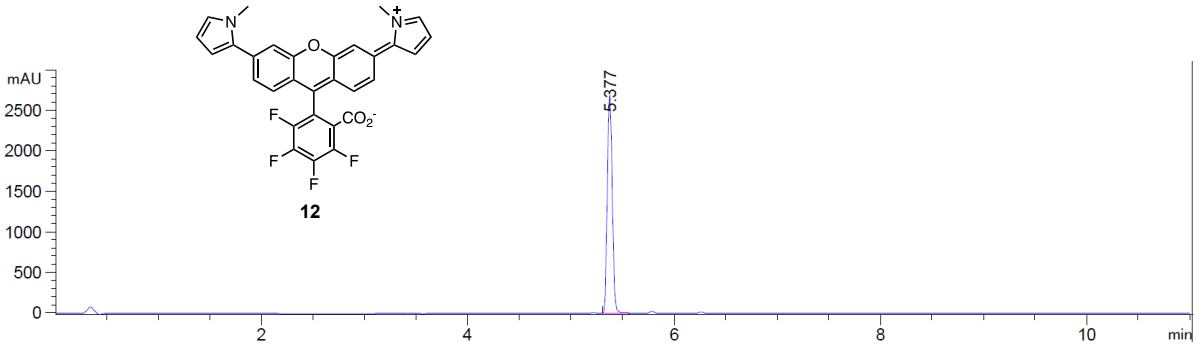
11

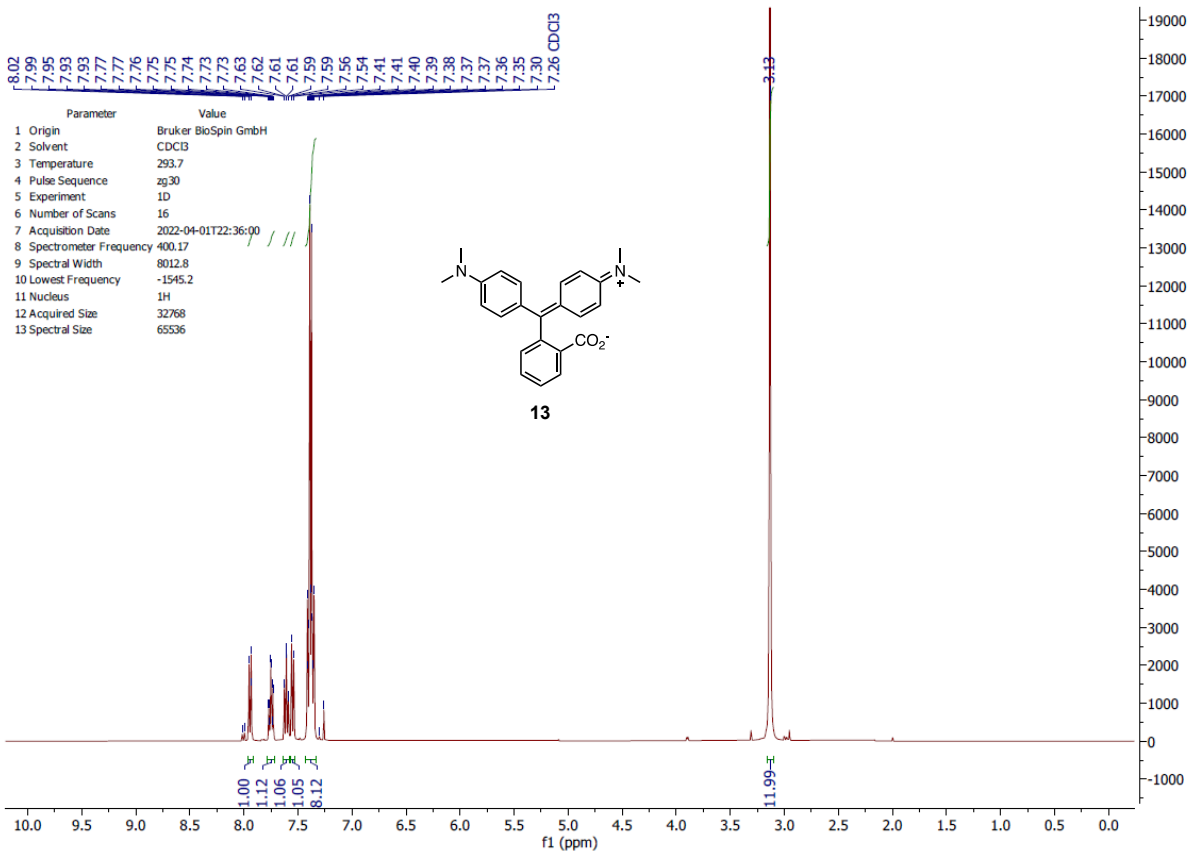

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 294.6
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-08T12:12:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.4
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
7.51 7.51 7.50 7.31 7.31 6.94 6.93 6.93 6.93 6.39 6.39 6.38 6.38 6.12 6.11 6.11 6.10
3.75
H2O
(CH3)2SO
DMF
DMF
(CH3)2CO
DMF
2.00 4.03 1.95 1.97 1.98 5.88
f1 (ppm)

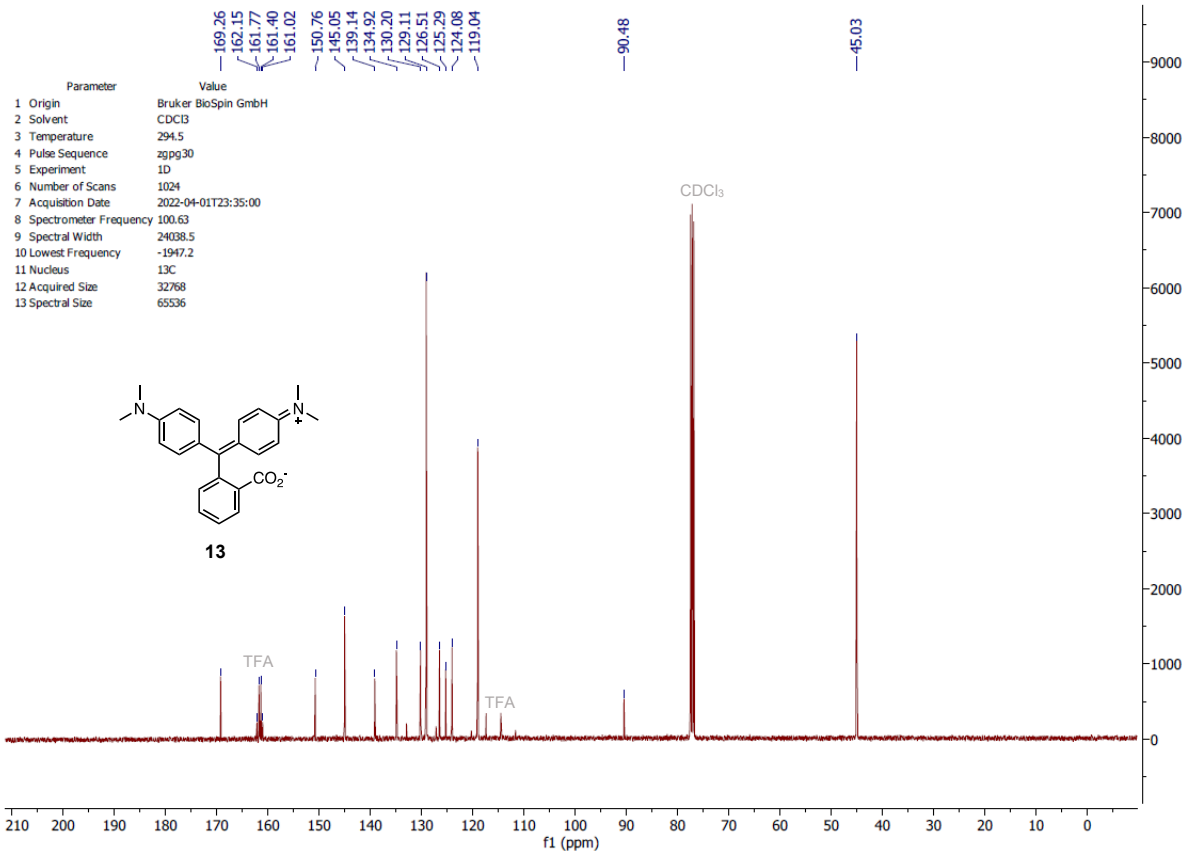
12

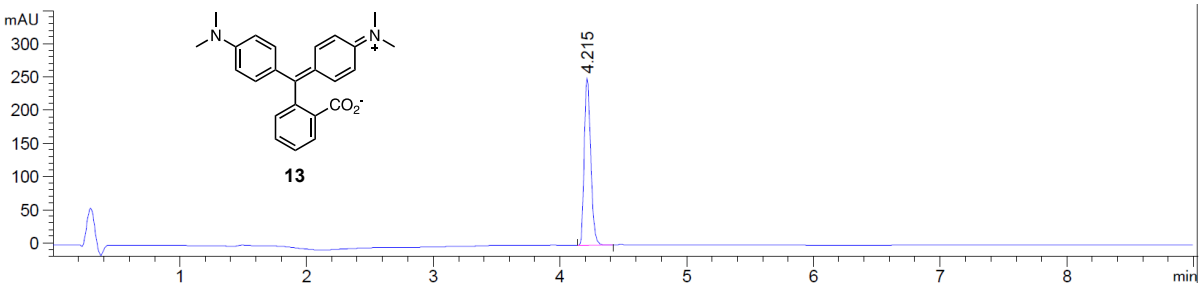

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 295.3
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-07-26T13:46:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82479.4
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
-137.93 -141.57 -149.63
CCl3F
1.00 2.02 0.98
f1 (ppm)

12

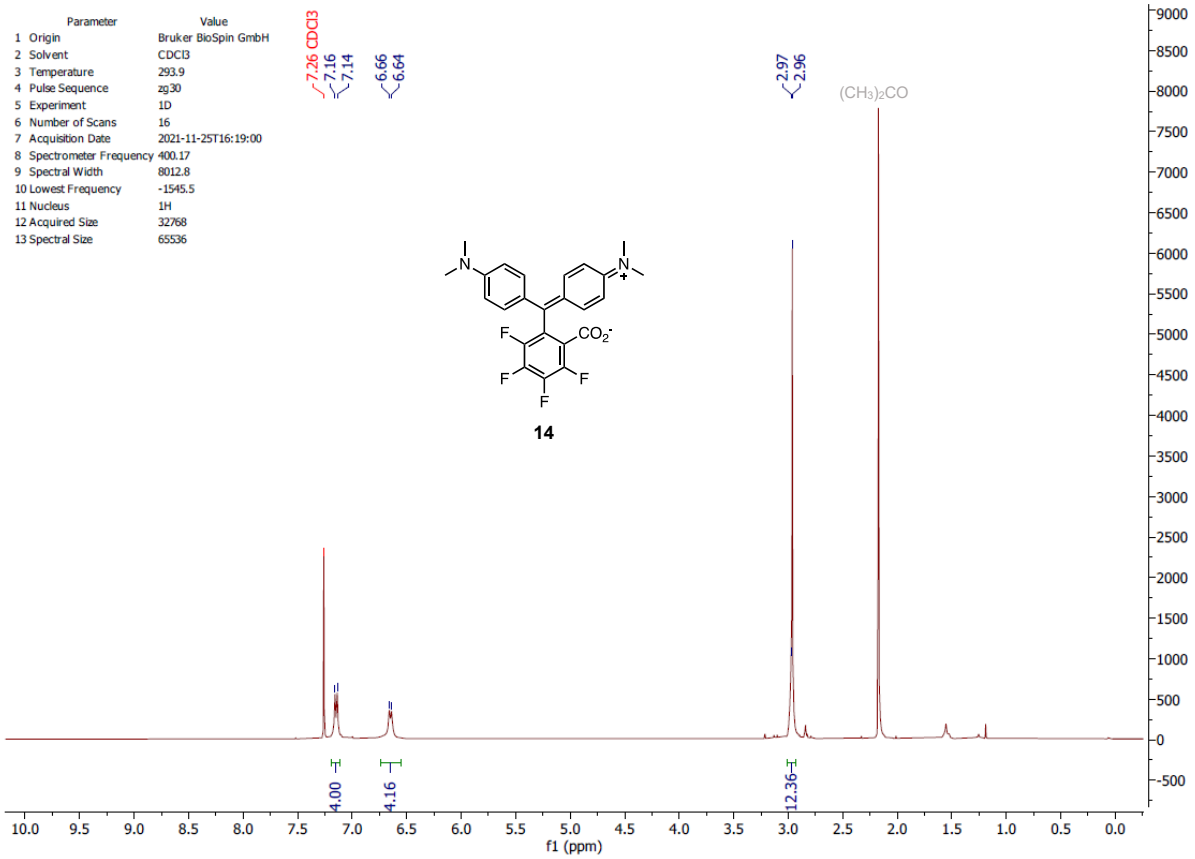

mAU
2500
2000
1500
1000
500
0
2
4
6
8
10
min
5.377
F
F
F
F
CO2-
12


8.02 7.99 7.95 7.93 7.93 7.77 7.77 7.76 7.75 7.75 7.74 7.73 7.73 7.63 7.62 7.61 7.61 7.59 7.59 7.56 7.54 7.41 7.41 7.40 7.39 7.38 7.37 7.37 7.36 7.35 7.30 7.26 CDCl3
3.13
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.7
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-04-01T22:36:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.2
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
13
1.00 1.12 1.06 1.05 8.12 11.99
f1 (ppm)

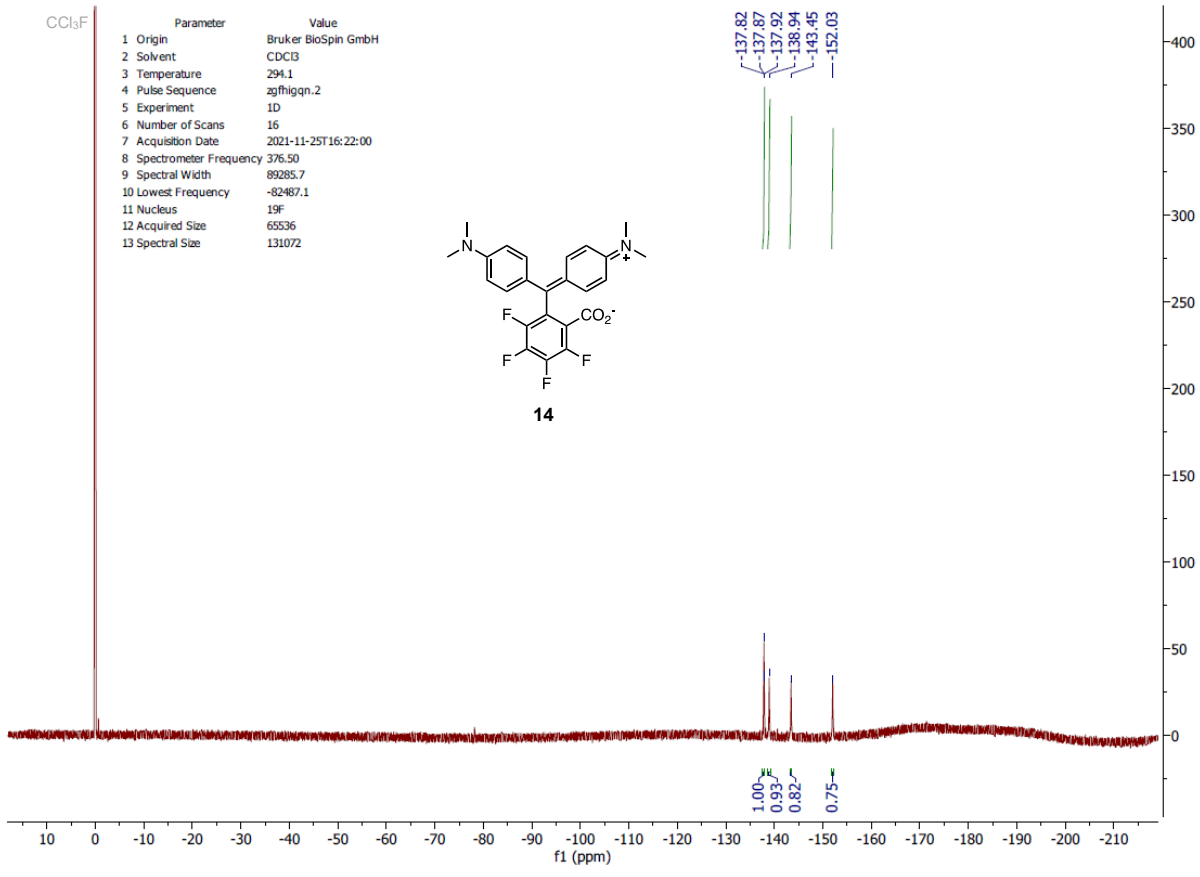

169.26 162.15 161.77 161.40 161.02 150.76 145.05 139.14 134.92 130.20 129.11 126.51 125.29 124.08 119.04
90.48
45.03
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.5
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 1024
7 Acquisition Date 2022-04-01T23:35:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1947.2
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
13
CDCl3
TFA
TFA
f1 (ppm)



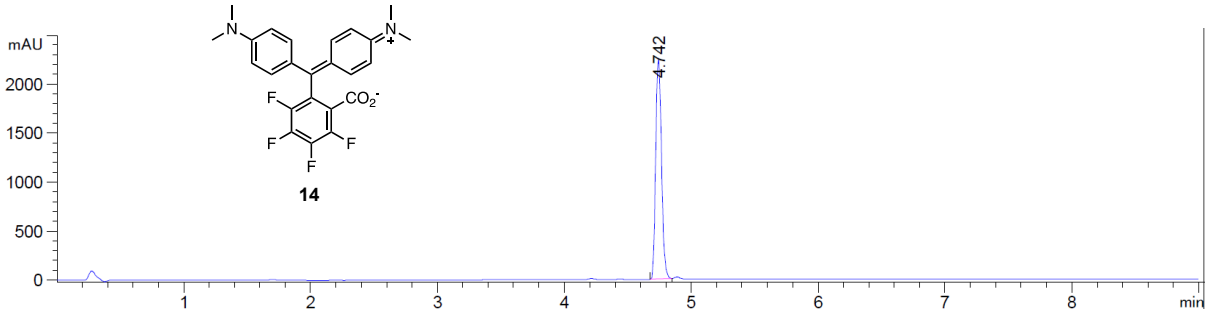
13


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.9
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-11-25T16:19:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
7.26 CDCl3
7.16
7.14
6.66
6.64
2.97
2.96
(CH3)2CO
4.00
4.16
12.36
f1 (ppm)

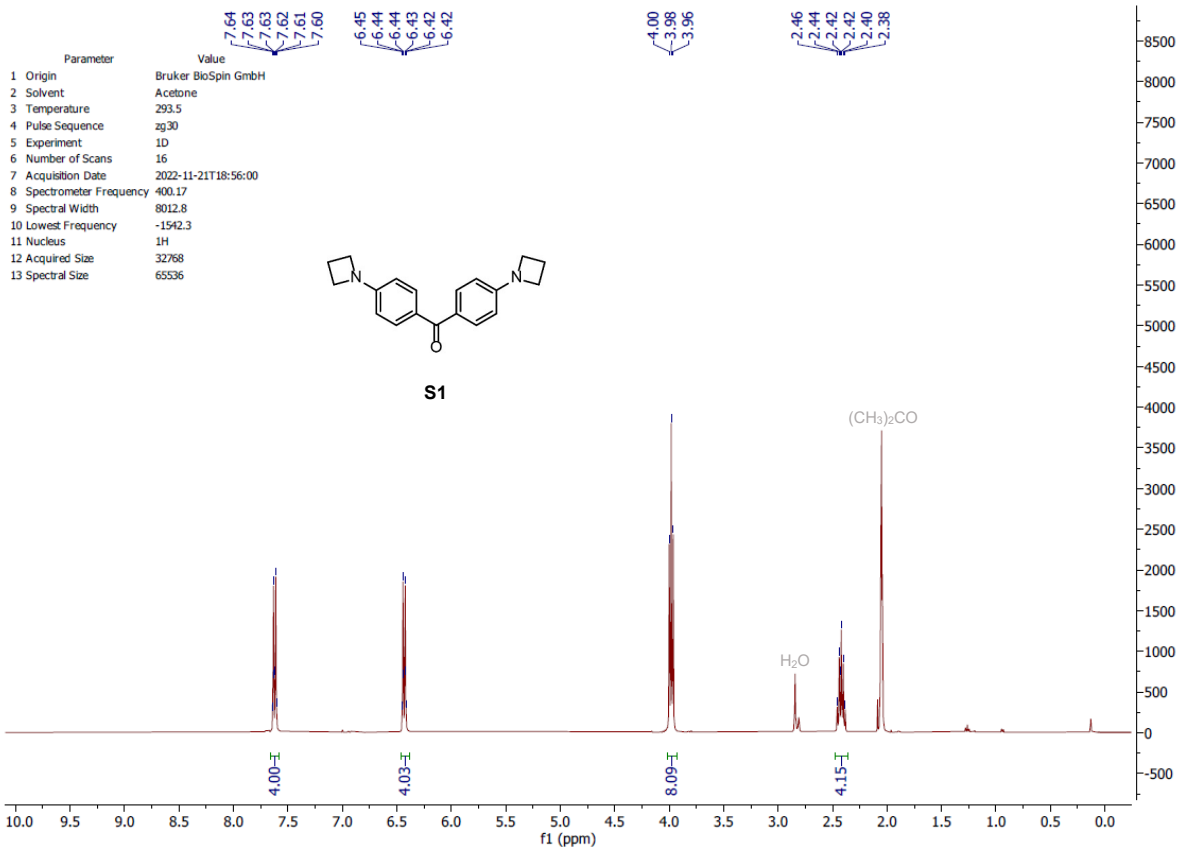
14

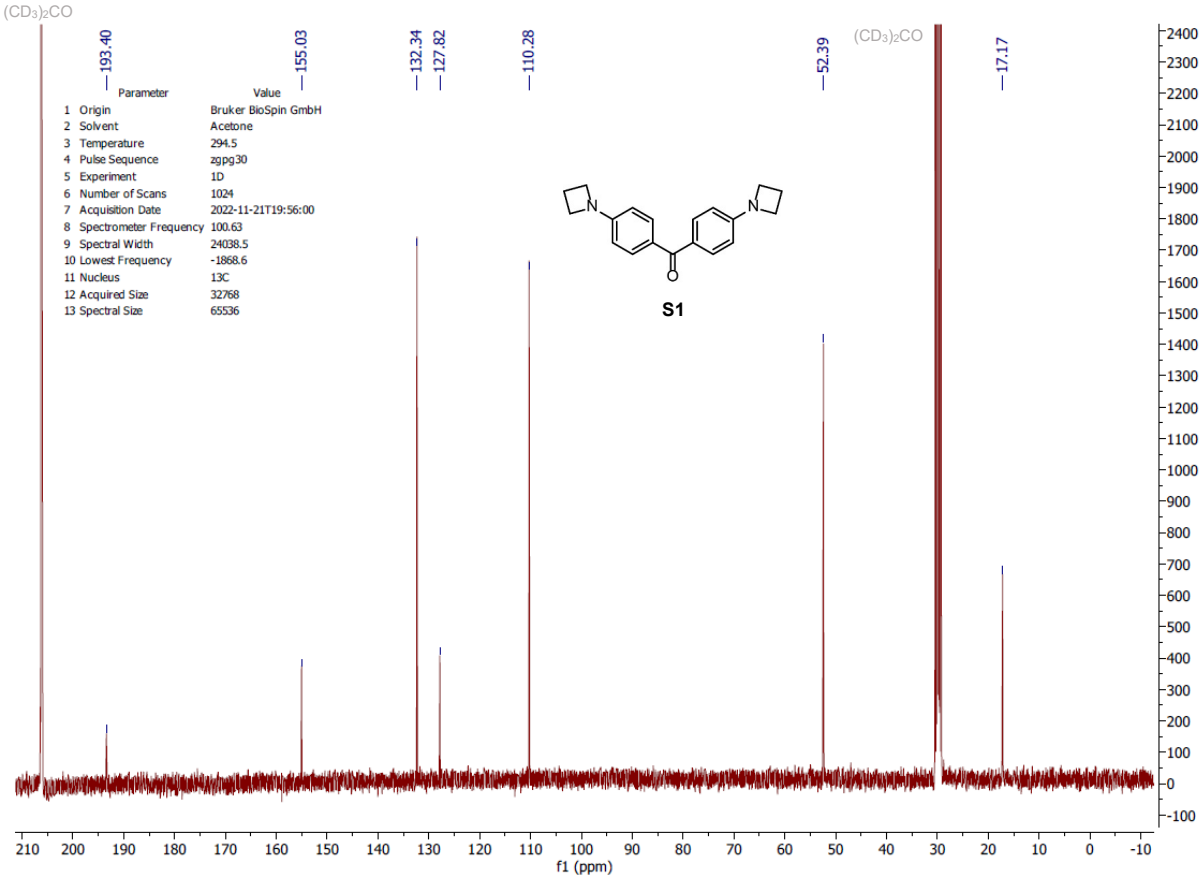

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.1
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-11-25T16:22:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82487.1
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
-137.82
-137.87
-137.92
-138.94
-143.45
-152.03
1.00
0.93
0.82
0.75
f1 (ppm)

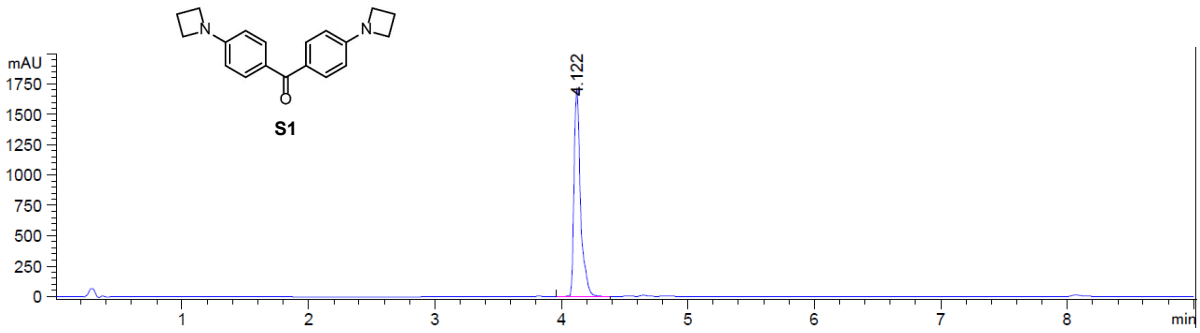
14

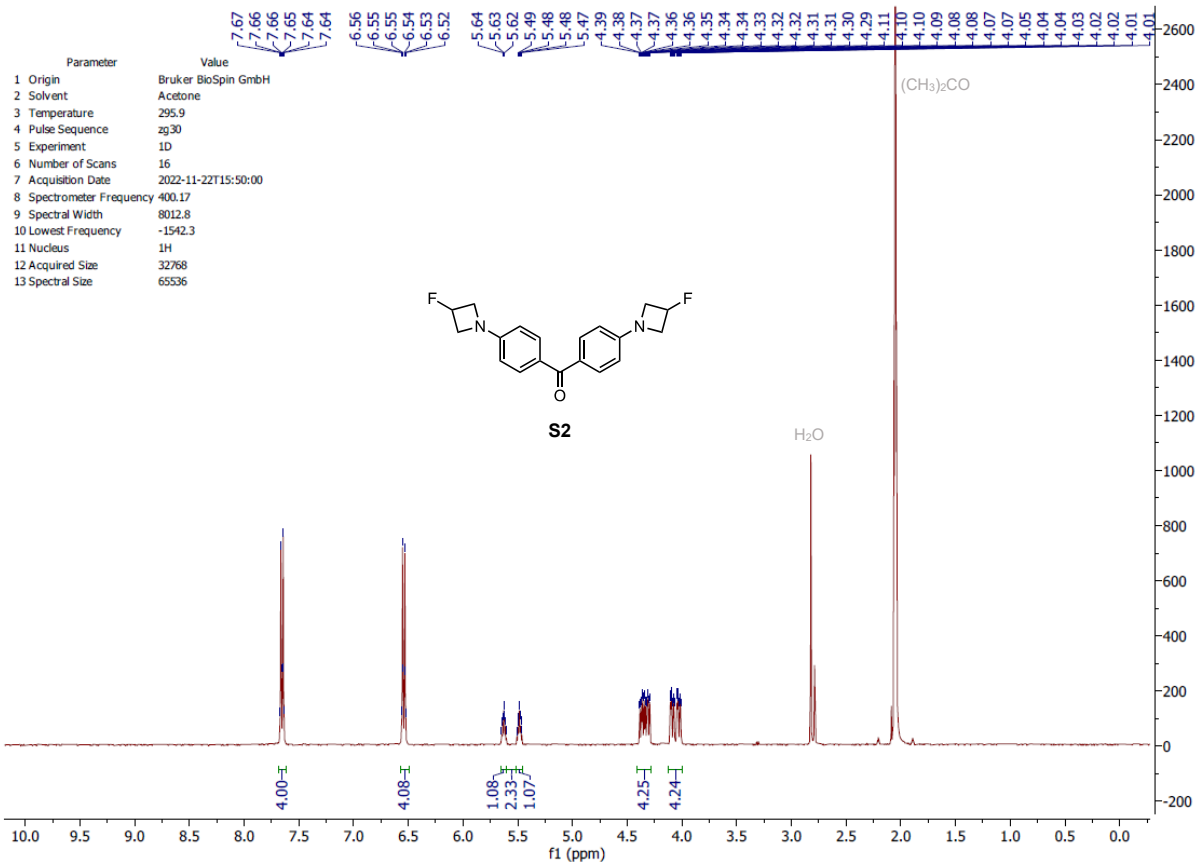


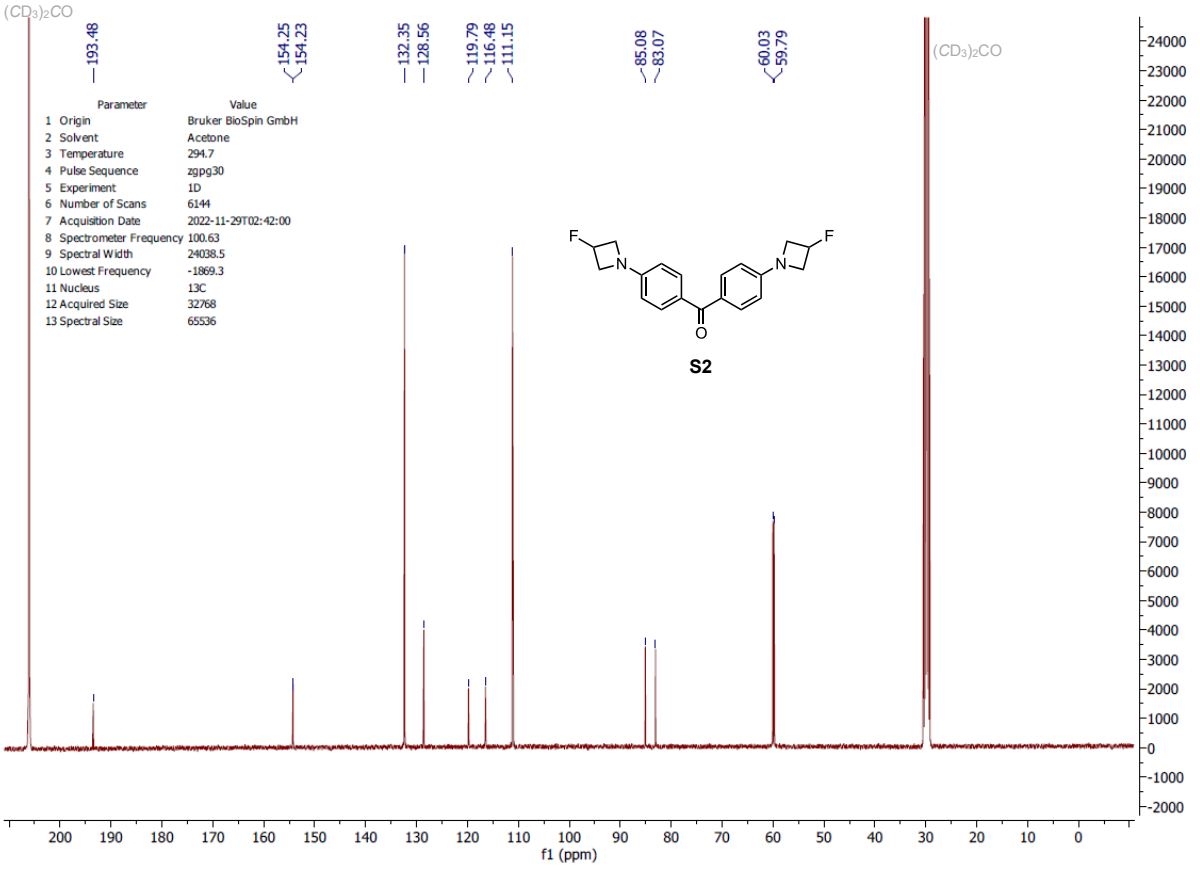
14

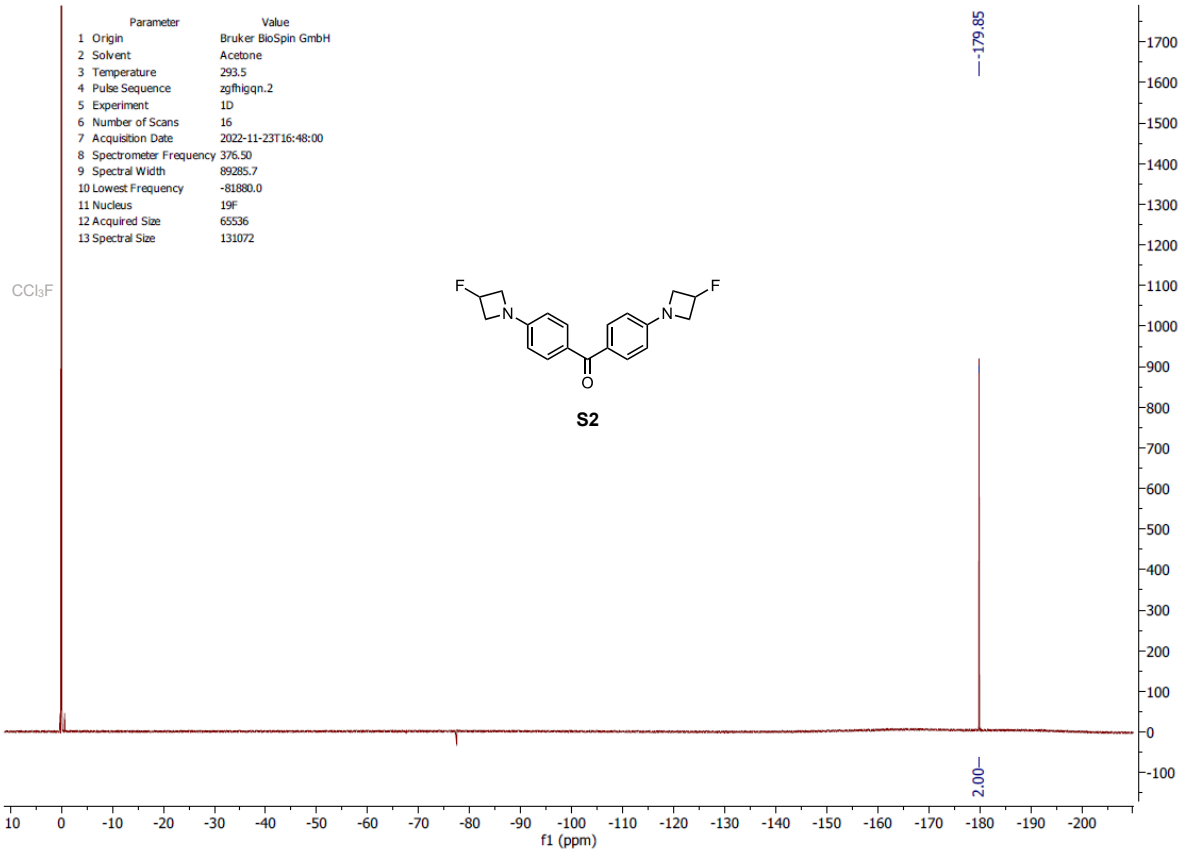

7.64
7.63
7.63
7.62
7.61
7.60
6.45
6.44
6.44
6.43
6.42
6.42
4.00
3.98
3.96
2.46
2.44
2.42
2.42
2.40
2.38
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 293.5
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-11-21T18:56:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
S1
(CH3)2CO
H2O
4.00
4.03
8.09
4.15
f1 (ppm)


(CD3)2CO
193.40
155.03
132.34
127.82
110.28
52.39
(CD3)2CO
17.17
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 294.5
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 1024
7 Acquisition Date 2022-11-21T19:56:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1868.6
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
S1
f1 (ppm)

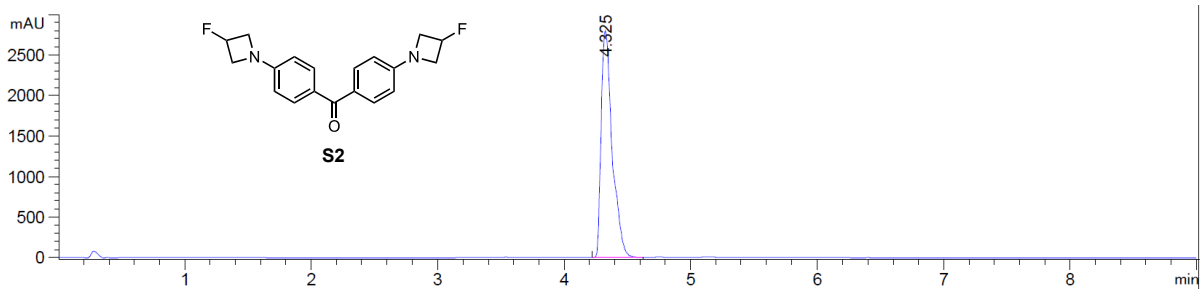

S1
mAU
1750
1500
1250
1000
750
500
250
0
4.122
1
2
3
4
5
6
7
8
min


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 295.9
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-11-22T15:50:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
(CH3)2CO
H2O
S2
f1 (ppm)

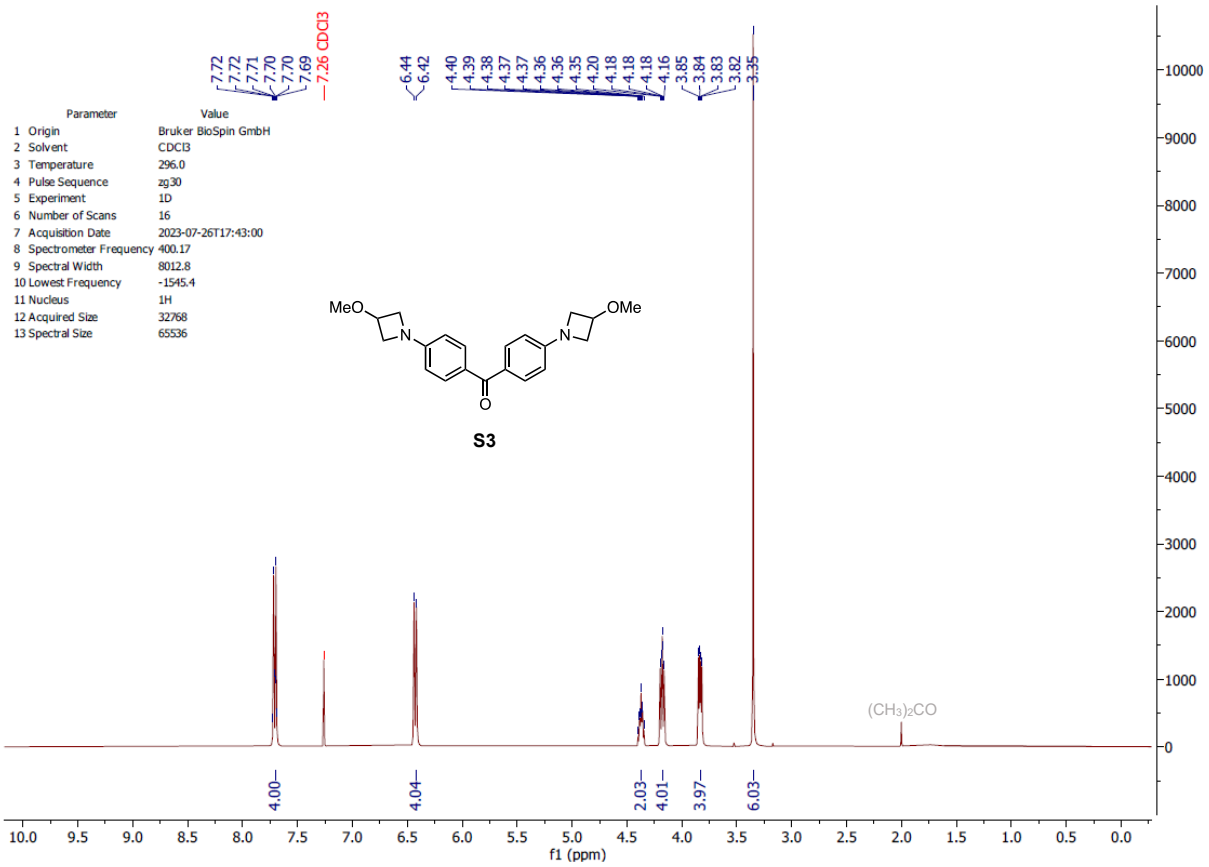

(CD3)2CO
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 294.7
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 6144
7 Acquisition Date 2022-11-29T02:42:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1869.3
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
S2
f1 (ppm)

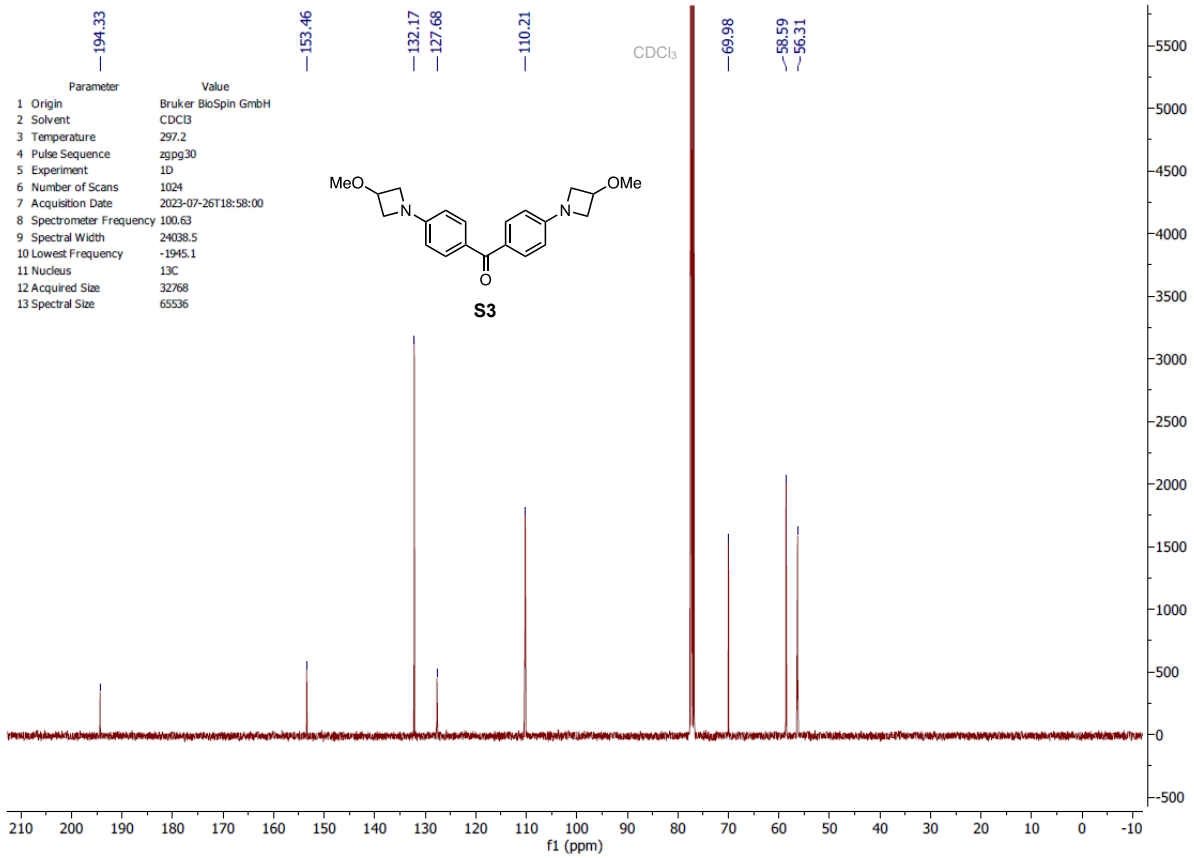

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 293.5
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-11-23T16:48:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81880.0
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
-179.85
2.00
f1 (ppm)

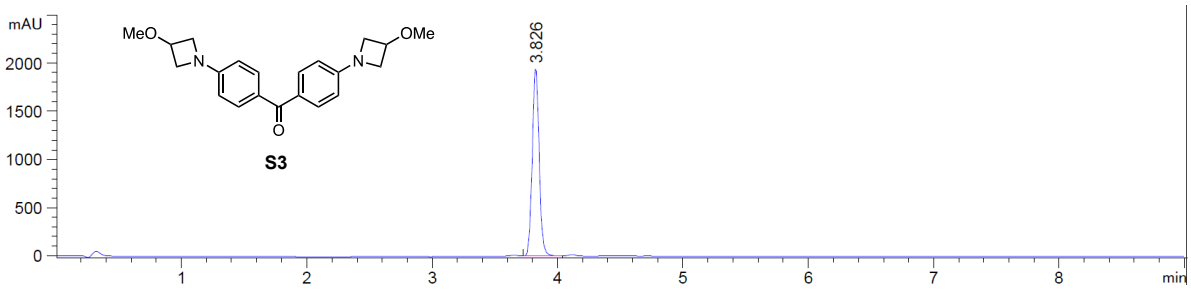
S2


mAU
4.325
min

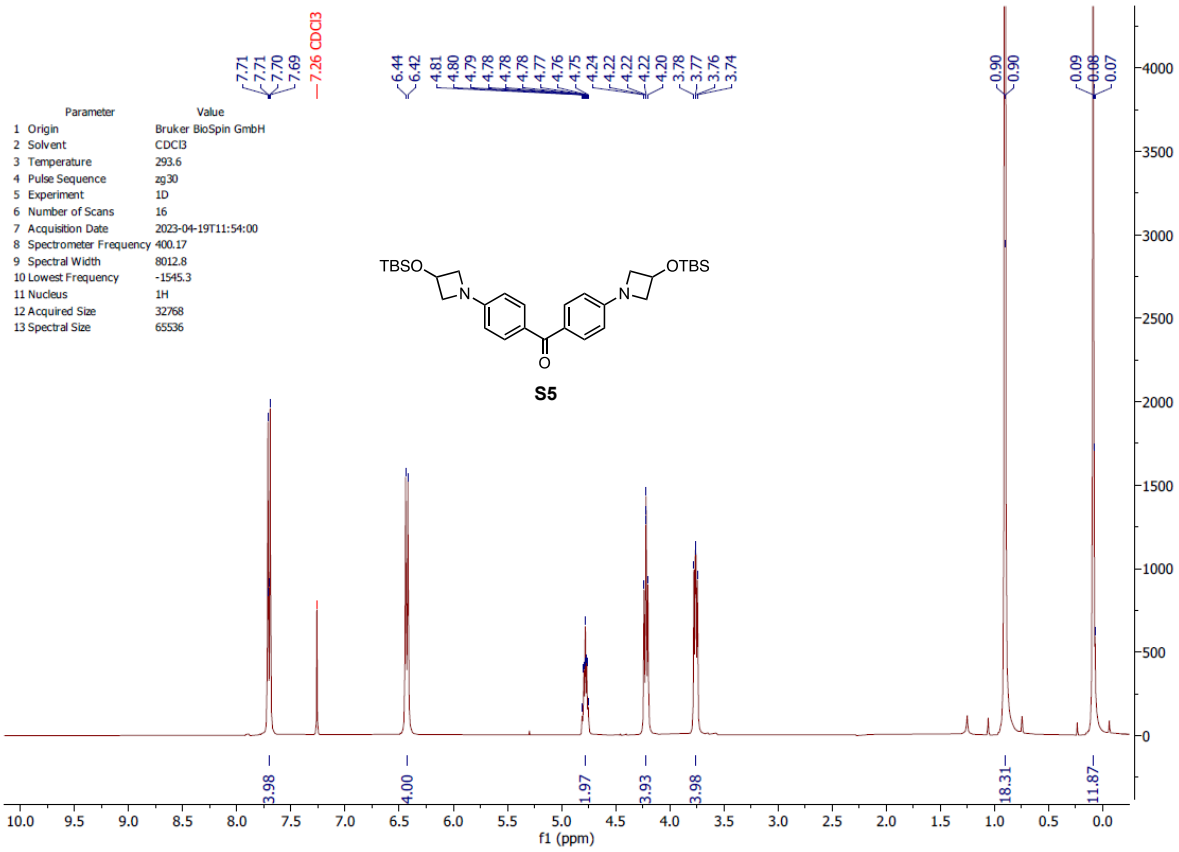
S2

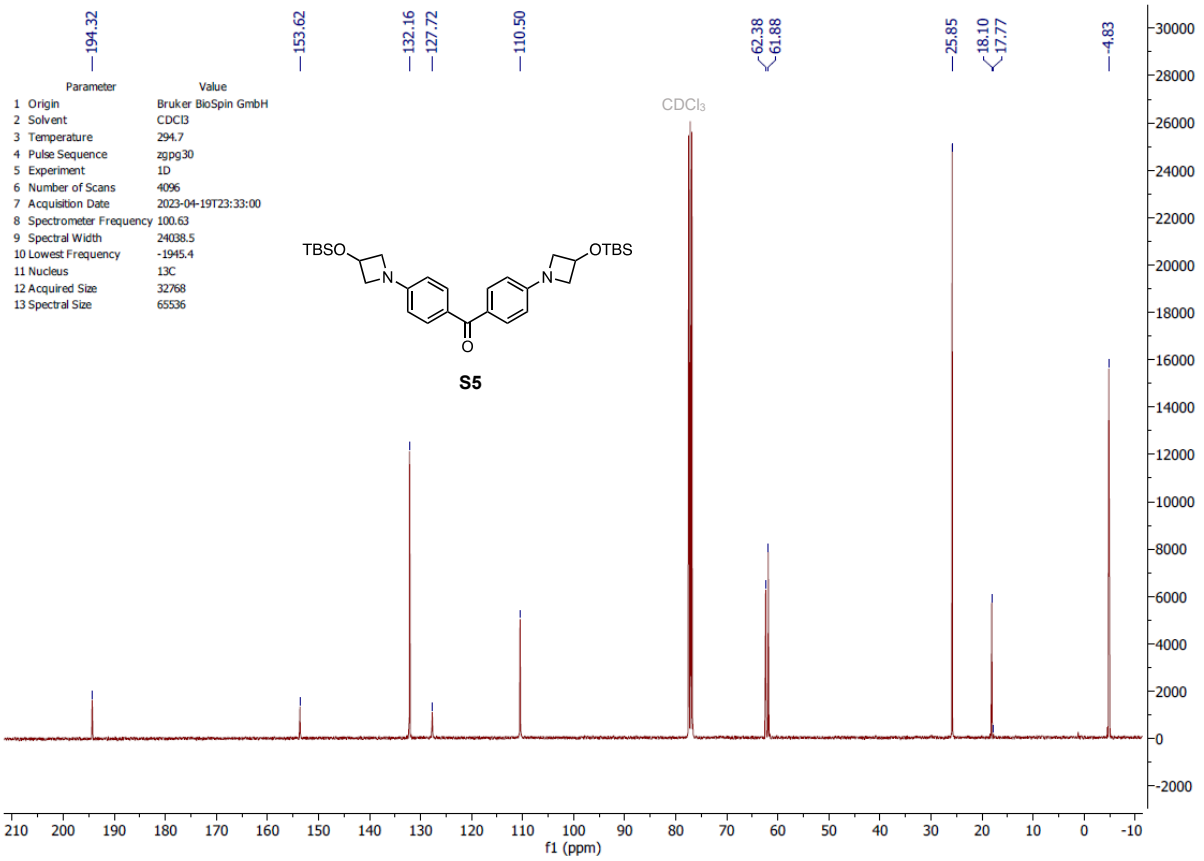

7.72
7.72
7.71
7.70
7.70
7.69
7.26 CDCl3
6.44
6.42
4.40
4.39
4.38
4.37
4.37
4.36
4.36
4.35
4.20
4.18
4.18
4.18
4.16
3.85
3.84
3.83
3.82
3.35
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 296.0
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-07-26T17:43:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.4
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
MeO
OMe
N
N
O
S3
(CH3)2CO
4.00
4.04
2.03
4.01
3.97
6.03
f1 (ppm)

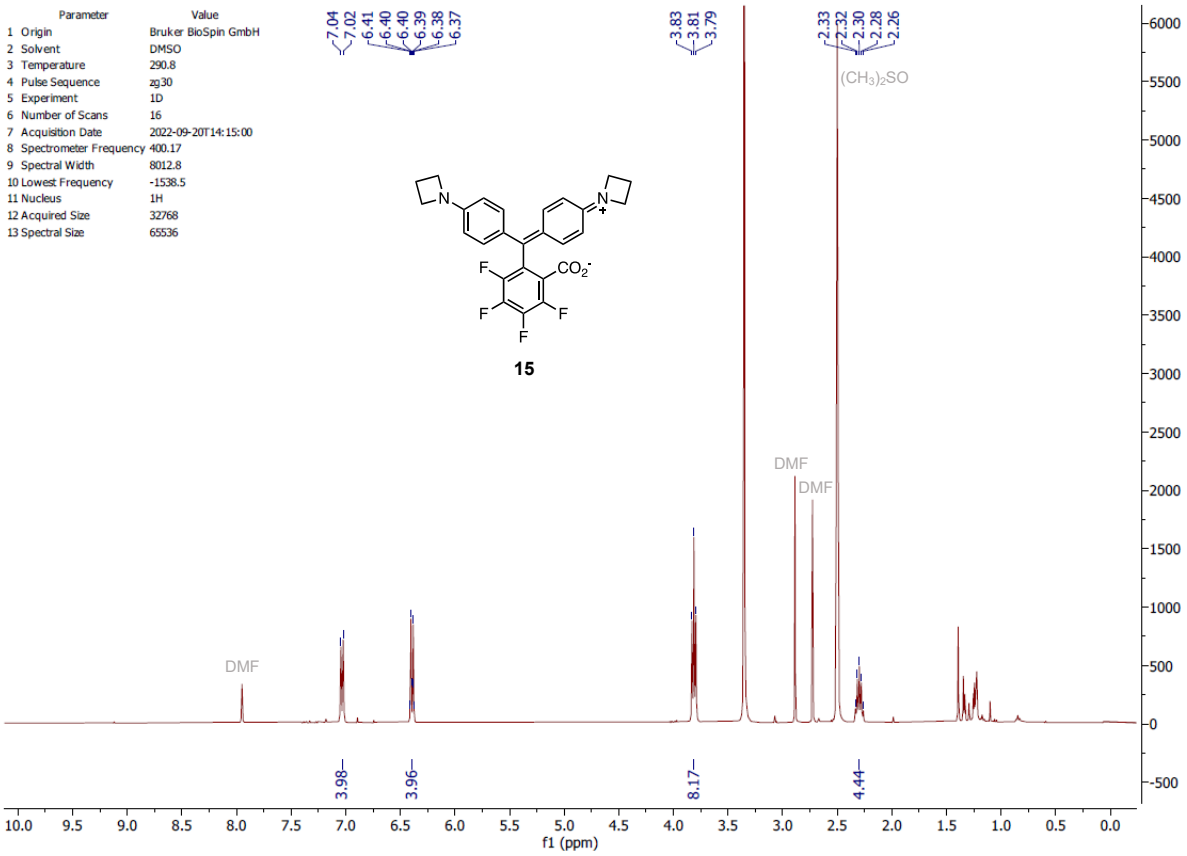

194.33
153.46
132.17
127.68
110.21
CDCl3
69.98
58.59
56.31
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 297.2
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 1024
7 Acquisition Date 2023-07-26T18:58:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1945.1
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
MeO
OMe
N
N
O
S3
f1 (ppm)



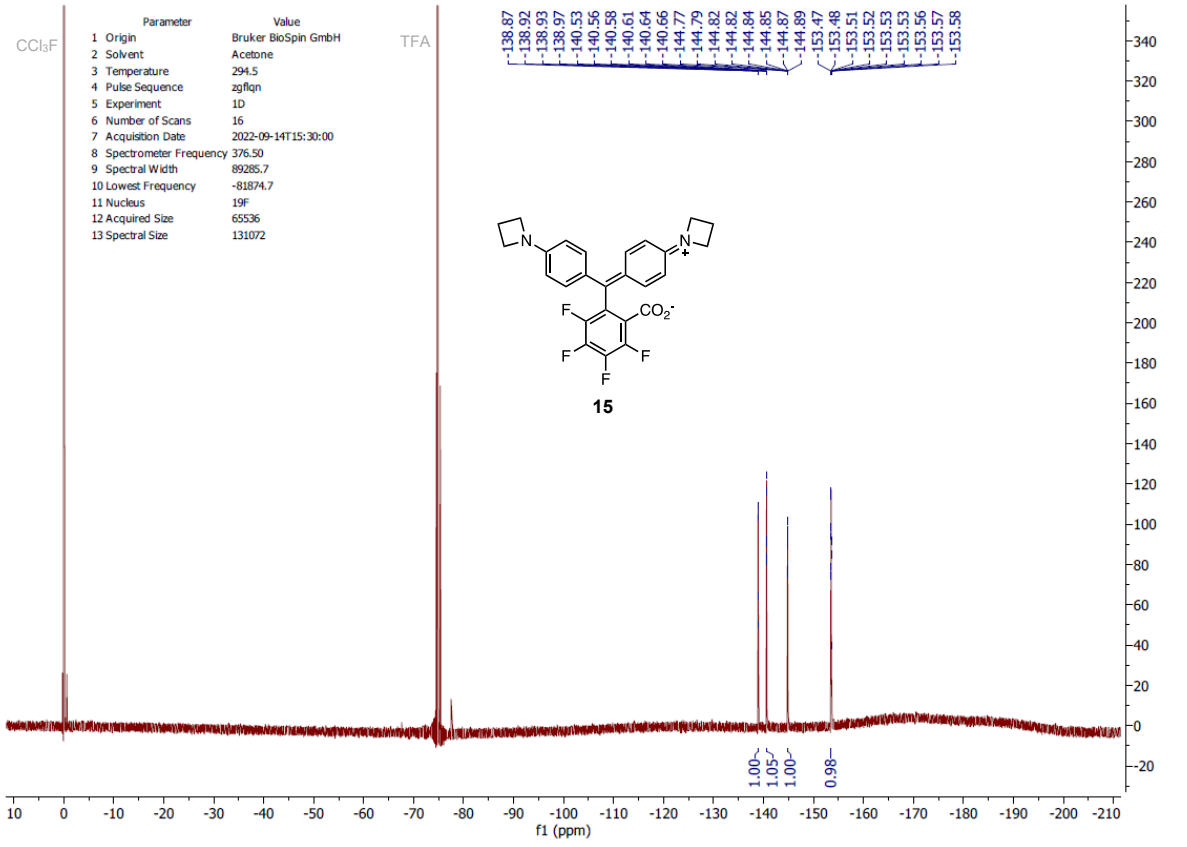
S3


7.71
7.71
7.70
7.69
7.26 CDCl3
6.44
6.42
4.81
4.80
4.79
4.78
4.78
4.78
4.77
4.76
4.75
4.24
4.22
4.22
4.22
4.20
3.78
3.77
3.76
3.74
0.90
0.90
0.09
0.08
0.07
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.6
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-04-19T11:54:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
TBSO
OTBS
N
N
O
S5
3.98
4.00
1.97
3.93
3.98
18.31
11.87
f1 (ppm)

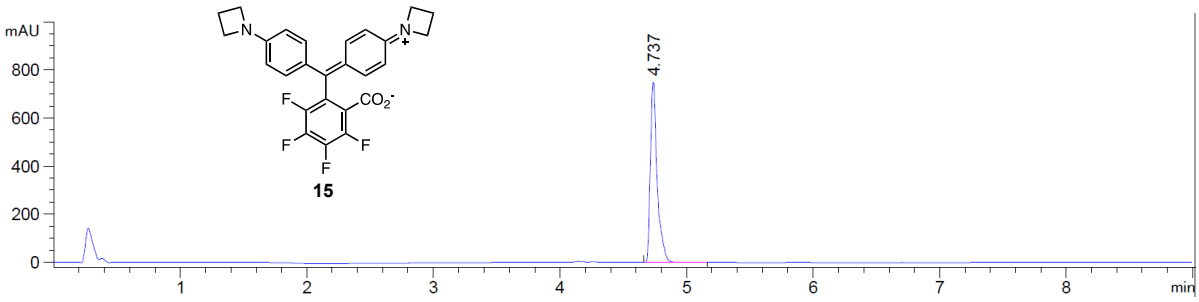

194.32
153.62
132.16
127.72
110.50
62.38
61.88
25.85
18.10
17.77
-4.83
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.7
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 4096
7 Acquisition Date 2023-04-19T23:33:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1945.4
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
CDCl3
TBSO
OTBS
N
N
O
S5
f1 (ppm)


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 290.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-09-20T14:15:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
7.04
7.02
6.41
6.40
6.40
6.39
6.38
6.37
3.83
3.81
3.79
2.33
2.32
2.30
2.28
2.26
(CH3)2SO
DMF
DMF
DMF
3.98
3.96
8.17
4.44
f1 (ppm)

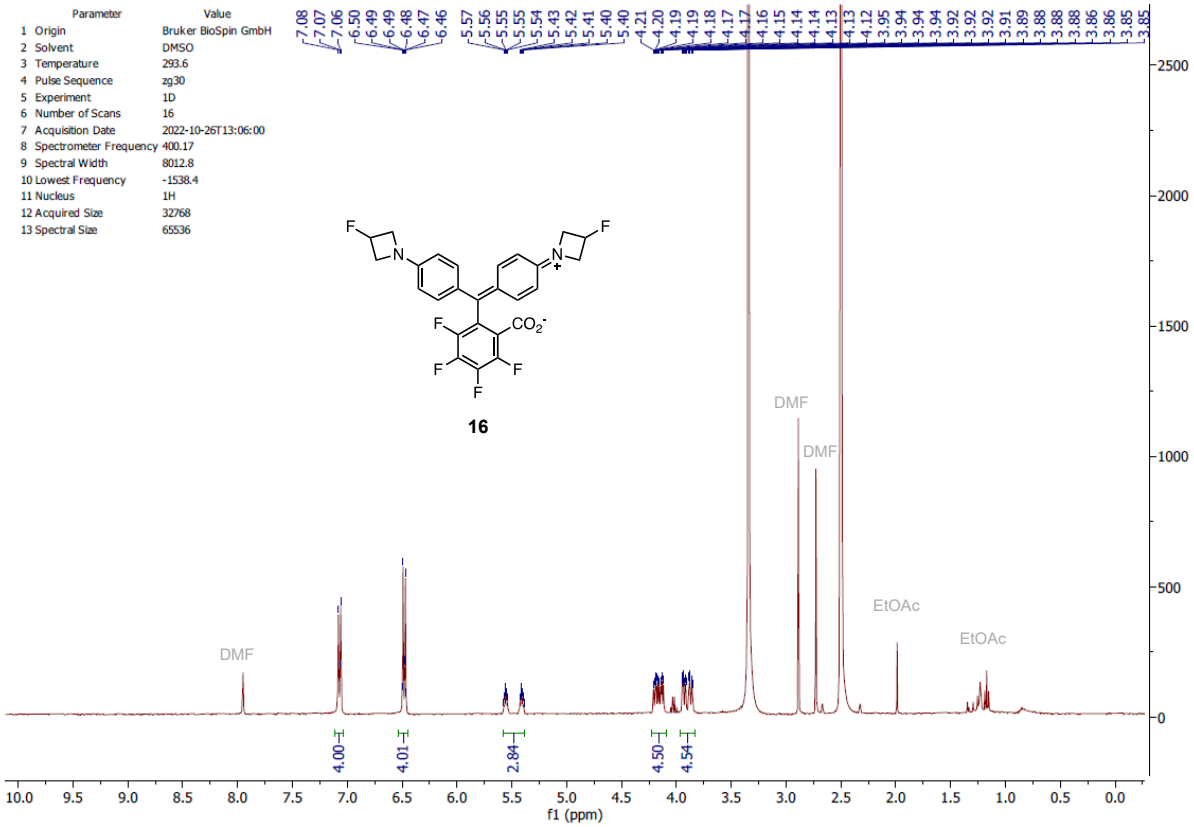
15


CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 294.5
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-09-14T15:30:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81874.7
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
138.87
138.92
138.93
138.97
140.53
140.56
140.58
140.61
140.64
140.66
144.77
144.79
144.82
144.82
144.84
144.85
144.87
144.89
153.47
153.48
153.51
153.52
153.53
153.53
153.53
153.56
153.57
153.58
1.00
1.05
1.00
0.98
f1 (ppm)

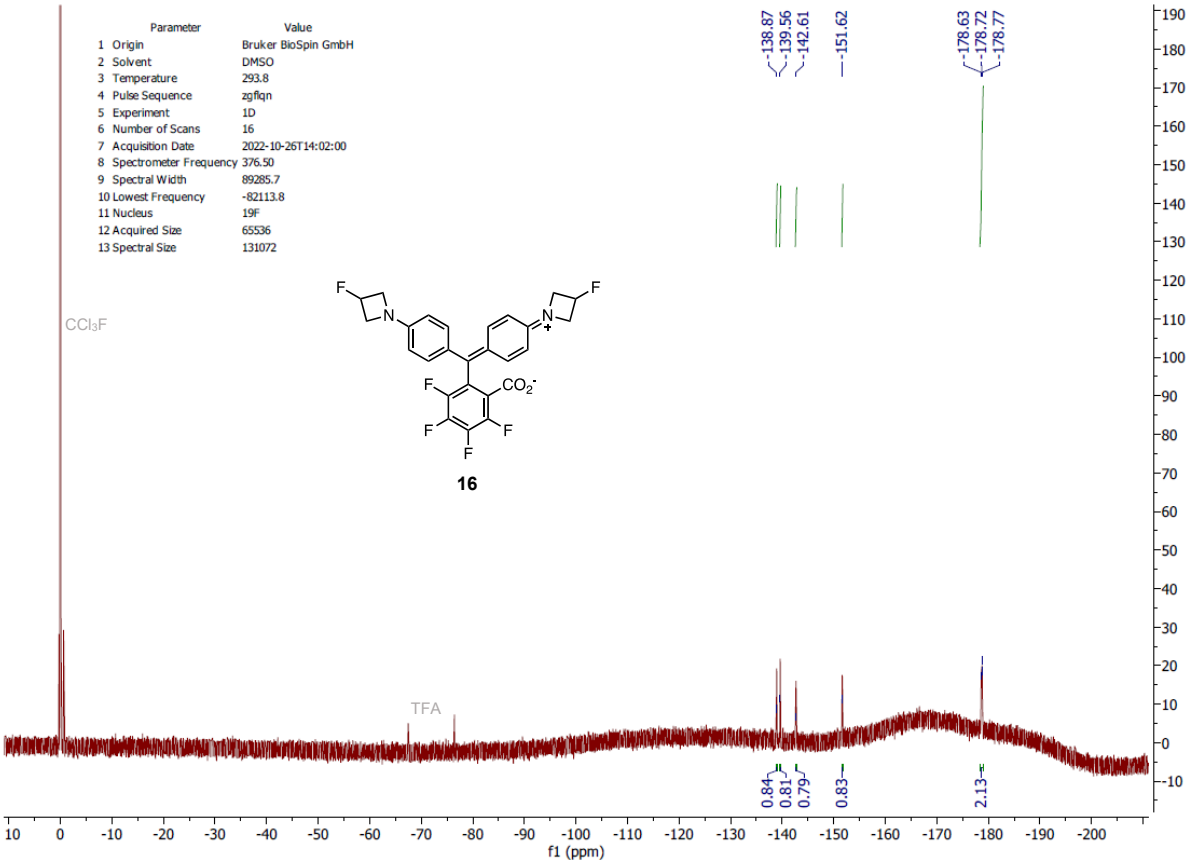
15



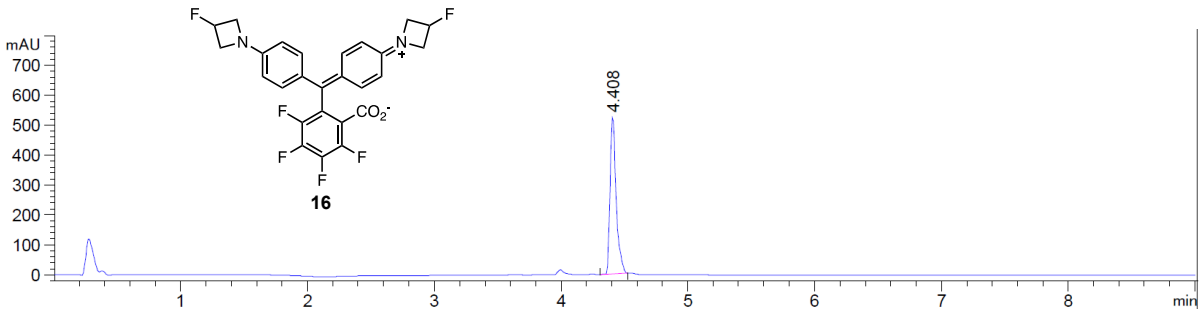
15


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 293.6
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-10-26T13:06:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.4
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
DMF
DMF
DMF
EtOAc
EtOAc
4.00
4.01
2.84
4.50
4.54
f1 (ppm)

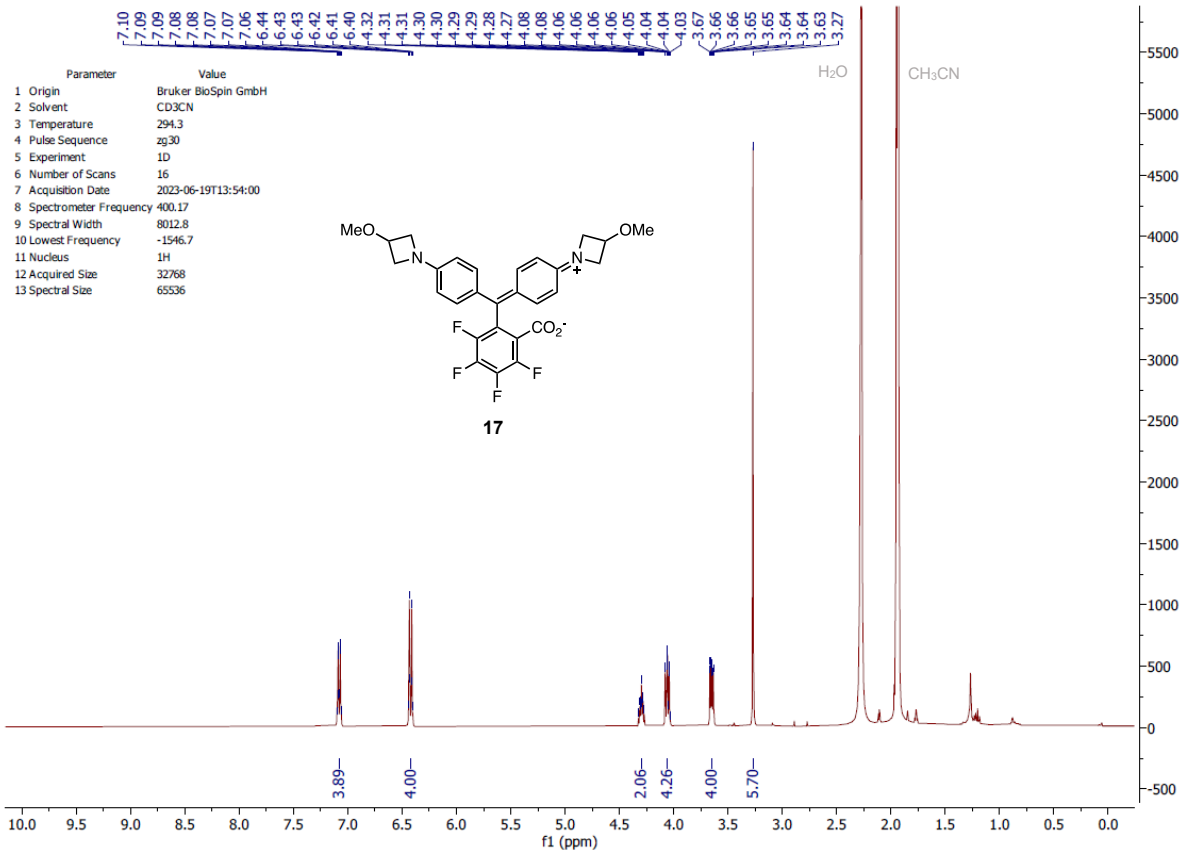
16


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 293.8
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-10-26T14:02:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82113.8
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
TFA
0.84
0.81
0.79
0.83
2.13
f1 (ppm)

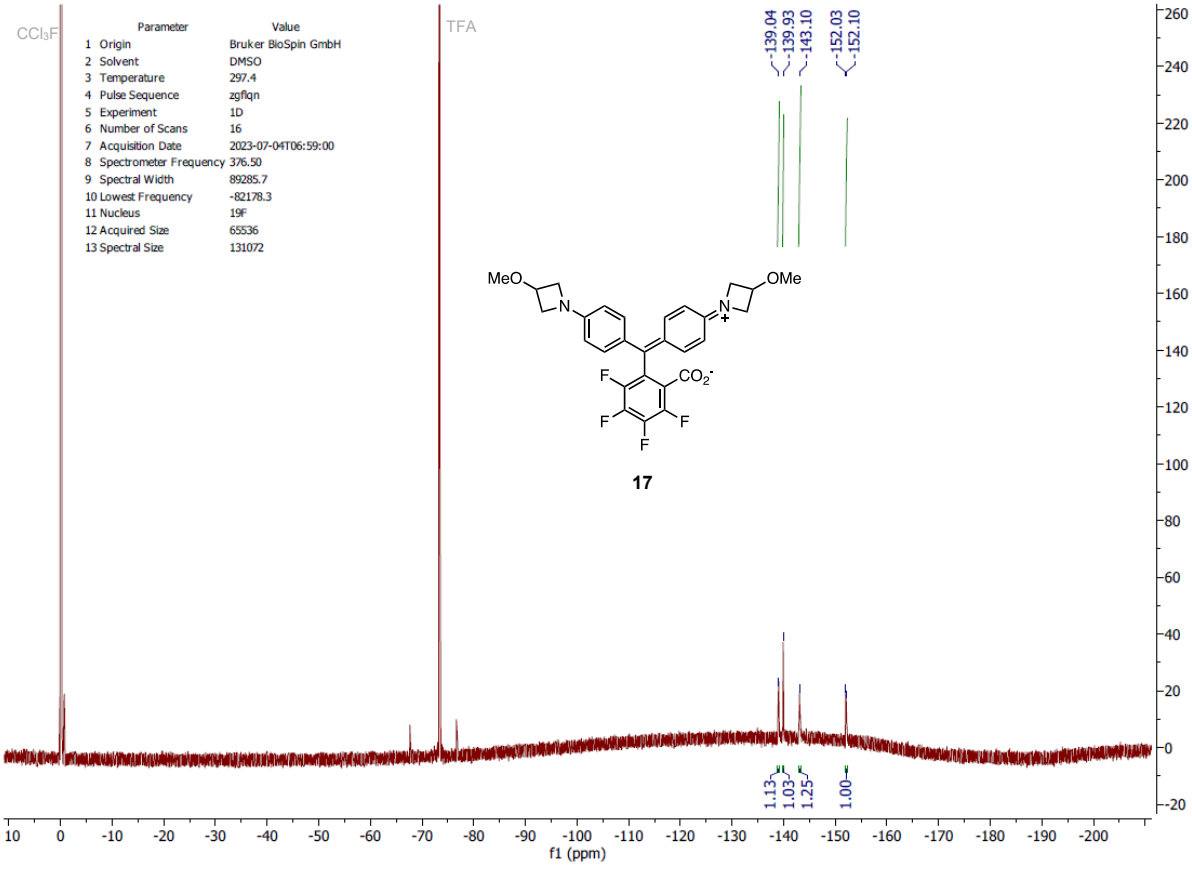
16


mAU
700
600
500
400
300
200
100
0
1
2
3
4
5
6
7
8
min
4.408
F
N
F
N+
F
CO2-
F
F
F

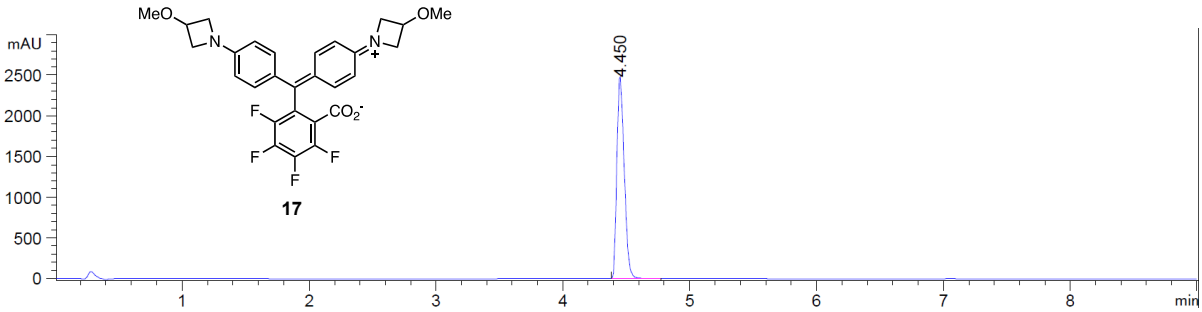
16

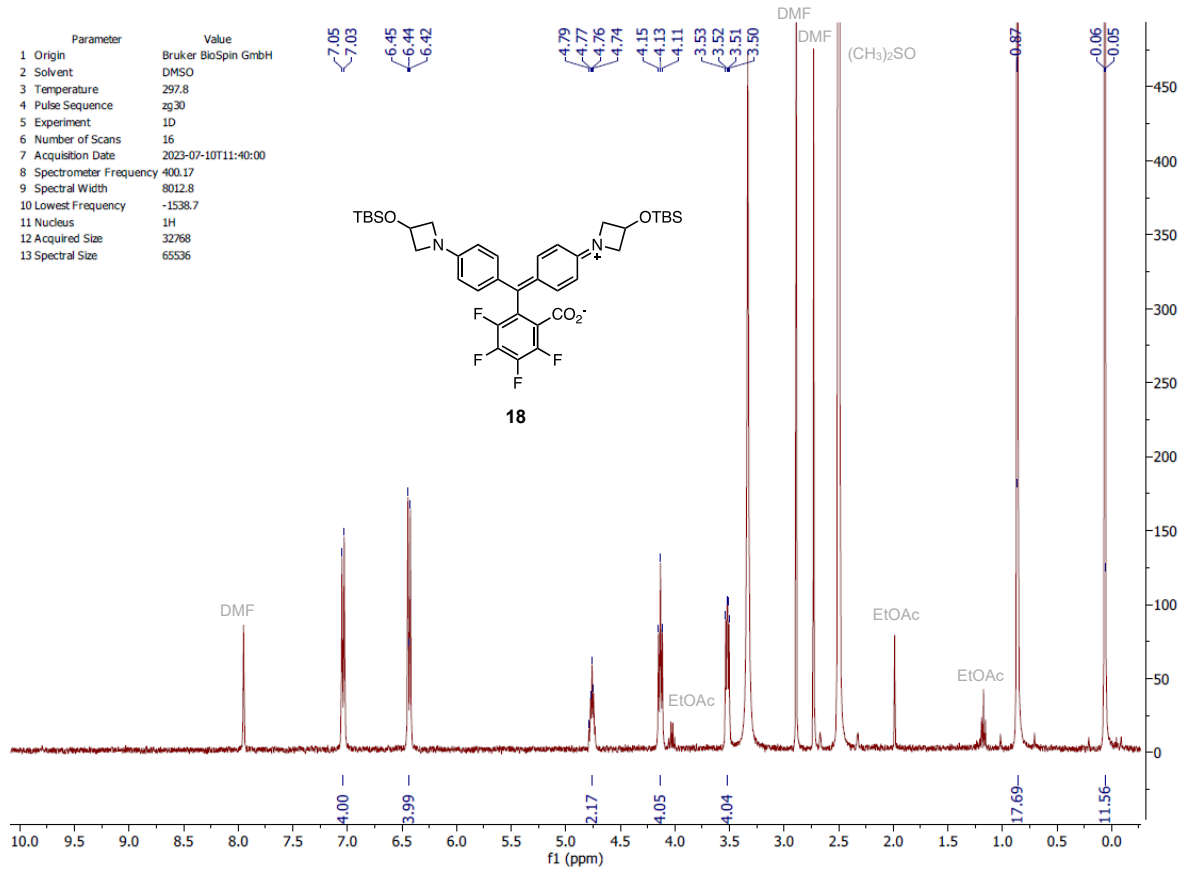

7.10 7.09 7.09 7.08 7.08 7.07 7.07 7.06 6.44 6.43 6.42 6.41 6.40 4.32 4.31 4.31 4.30 4.30 4.29 4.29 4.28 4.27 4.08 4.08 4.06 4.06 4.06 4.06 4.05 4.04 4.04 4.03 3.67 3.66 3.66 3.65 3.65 3.64 3.64 3.63 3.27
H2O
CH3CN
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 294.3
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-06-19T13:54:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1546.7
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
MeO
OMe
N
N+
F
CO2-
F
F
F
3.89 4.00 2.06 4.26 4.00 5.70
f1 (ppm)

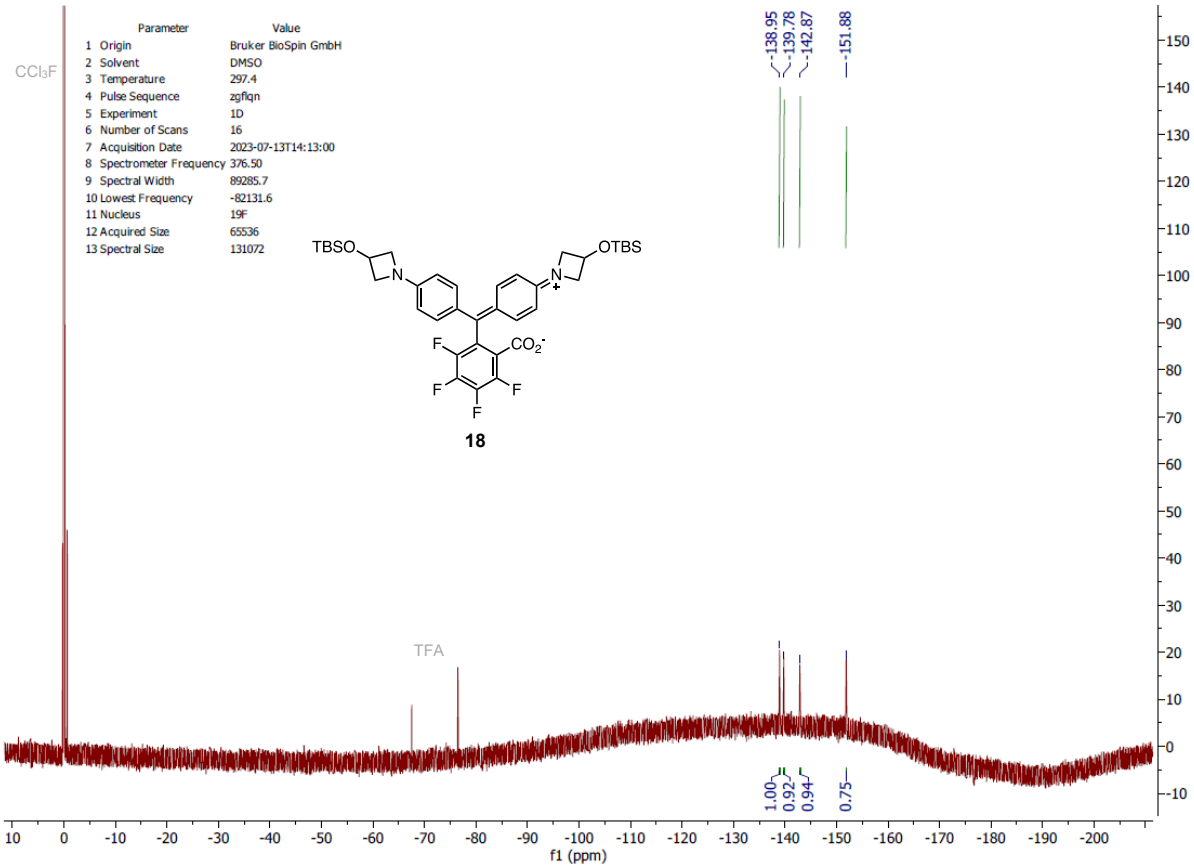
17

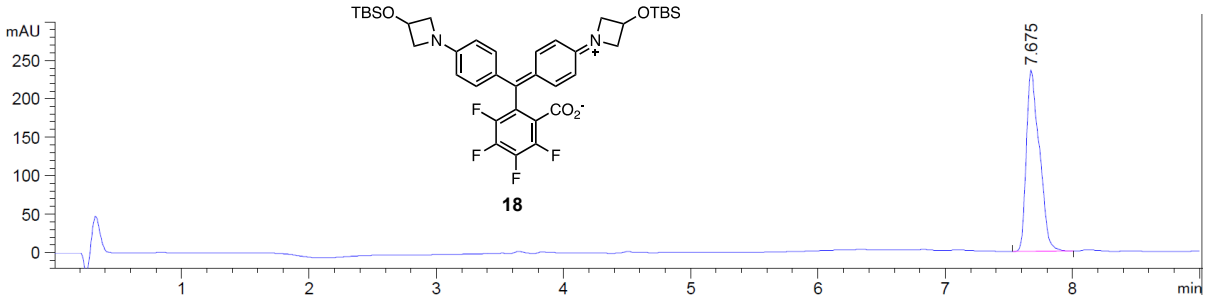

CCl3F
TFA
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 297.4
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-07-04T06:59:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82178.3
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
-139.04 -139.93 -143.10 -152.03 -152.10
MeO
OMe
N
N+
F
CO2-
F
F
F
1.13 1.03 1.25 1.00
f1 (ppm)

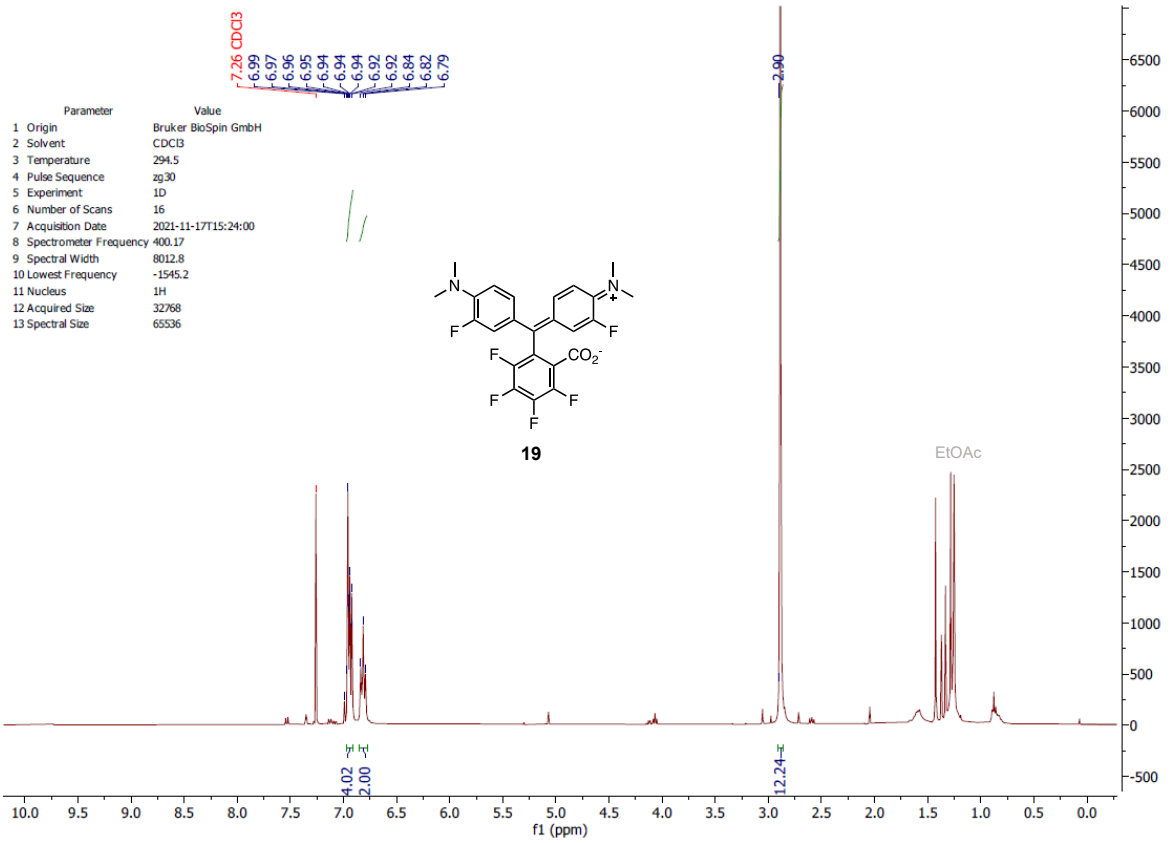
17

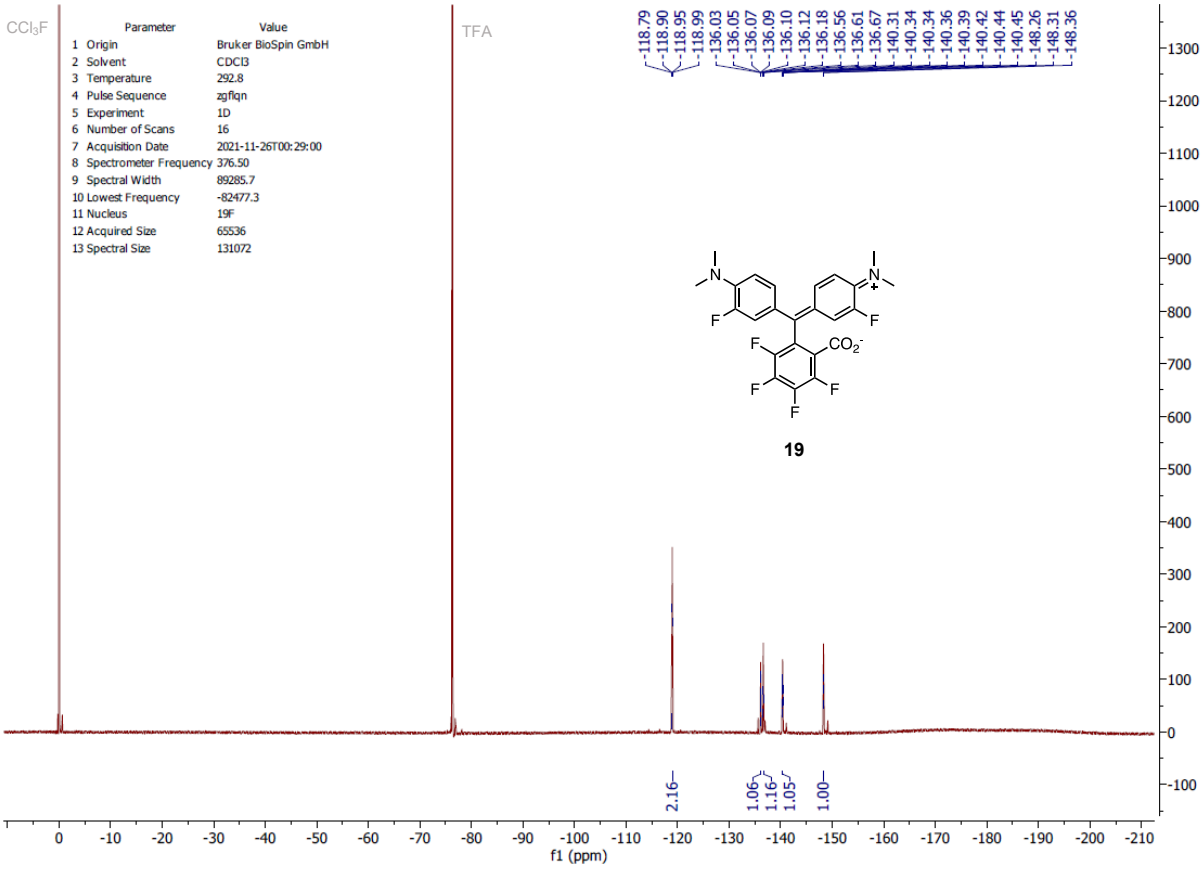

MeO
OMe
N
N+
F
F
F
F
CO2-
17
mAU
2500
2000
1500
1000
500
0
4.450
1
2
3
4
5
6
7
8
min

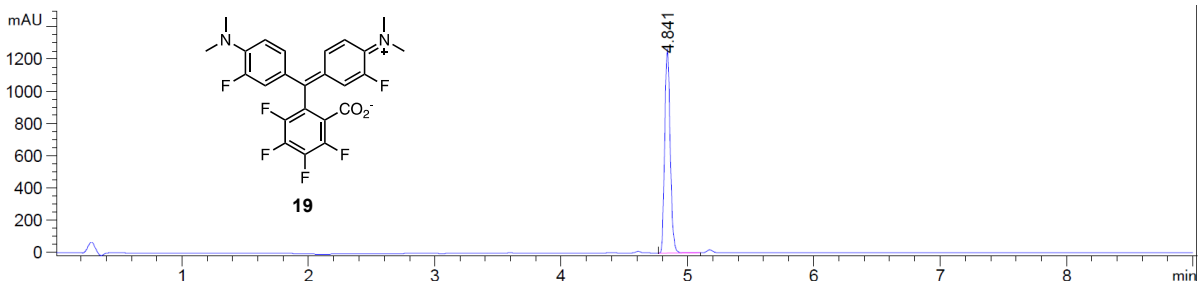

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 297.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-07-10T11:40:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.7
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
7.05
7.03
6.45
6.44
6.42
4.79
4.77
4.76
4.74
4.15
4.13
4.11
3.53
3.52
3.51
3.50
DMF
DMF
(CH3)2SO
0.87
0.06
0.05
TBSO
OTBS
F
CO2-
F
F
F
18
DMF
EtOAc
EtOAc
EtOAc
4.00
3.99
2.17
4.05
4.04
17.69
11.56
f1 (ppm)


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 297.4
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-07-13T14:13:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82131.6
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
-138.95
-139.78
-142.87
-151.88
TBSO
OTBS
F
CO2-
F
F
F
18
TFA
1.00
0.92
0.94
0.75
f1 (ppm)

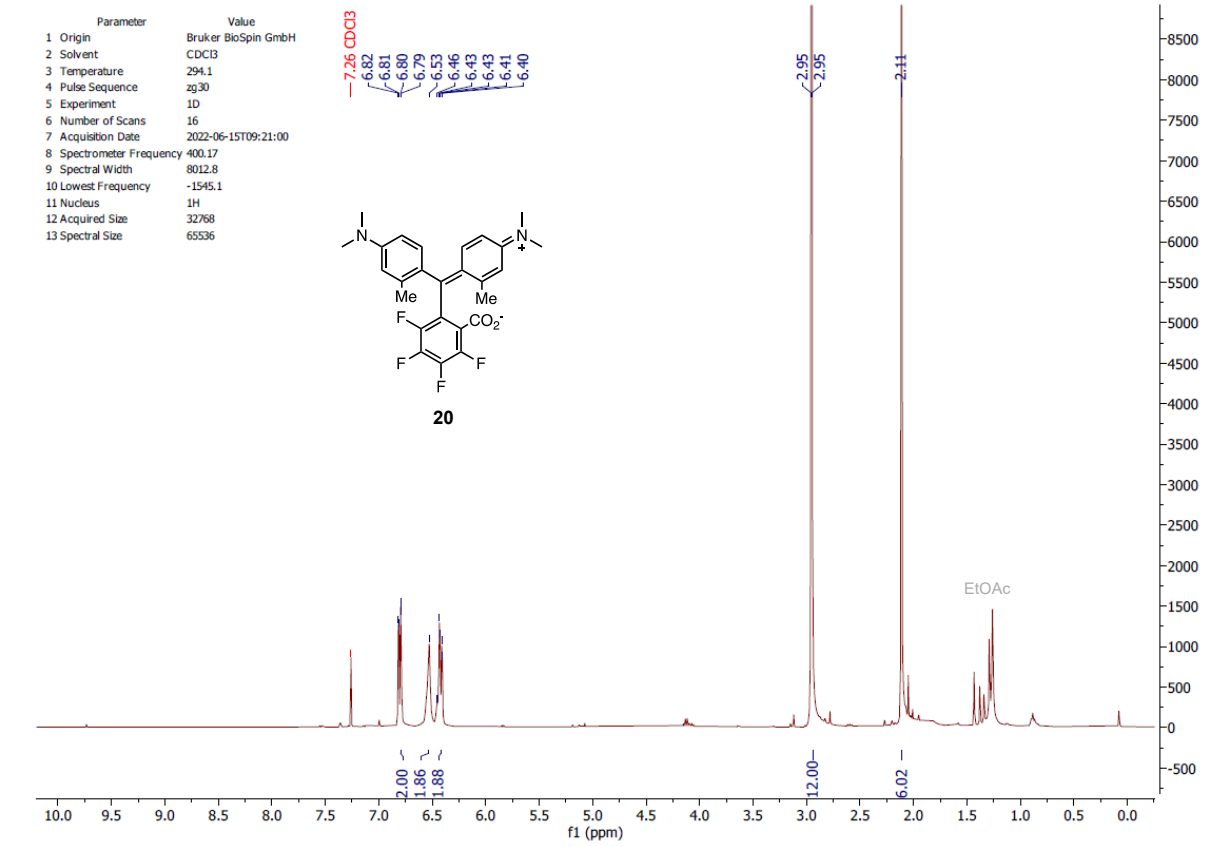

TBSO
OTBS
N
N+
F
CO2-
F
F
F
18
mAU
250
200
150
100
50
0
7.675
1
2
3
4
5
6
7
8
min


7.26 CDCl3
6.99
6.97
6.96
6.95
6.94
6.94
6.94
6.92
6.92
6.84
6.82
6.79
2.90
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.5
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-11-17T15:24:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.2
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
19
EtOAc
4.02
2.00
12.24
f1 (ppm)

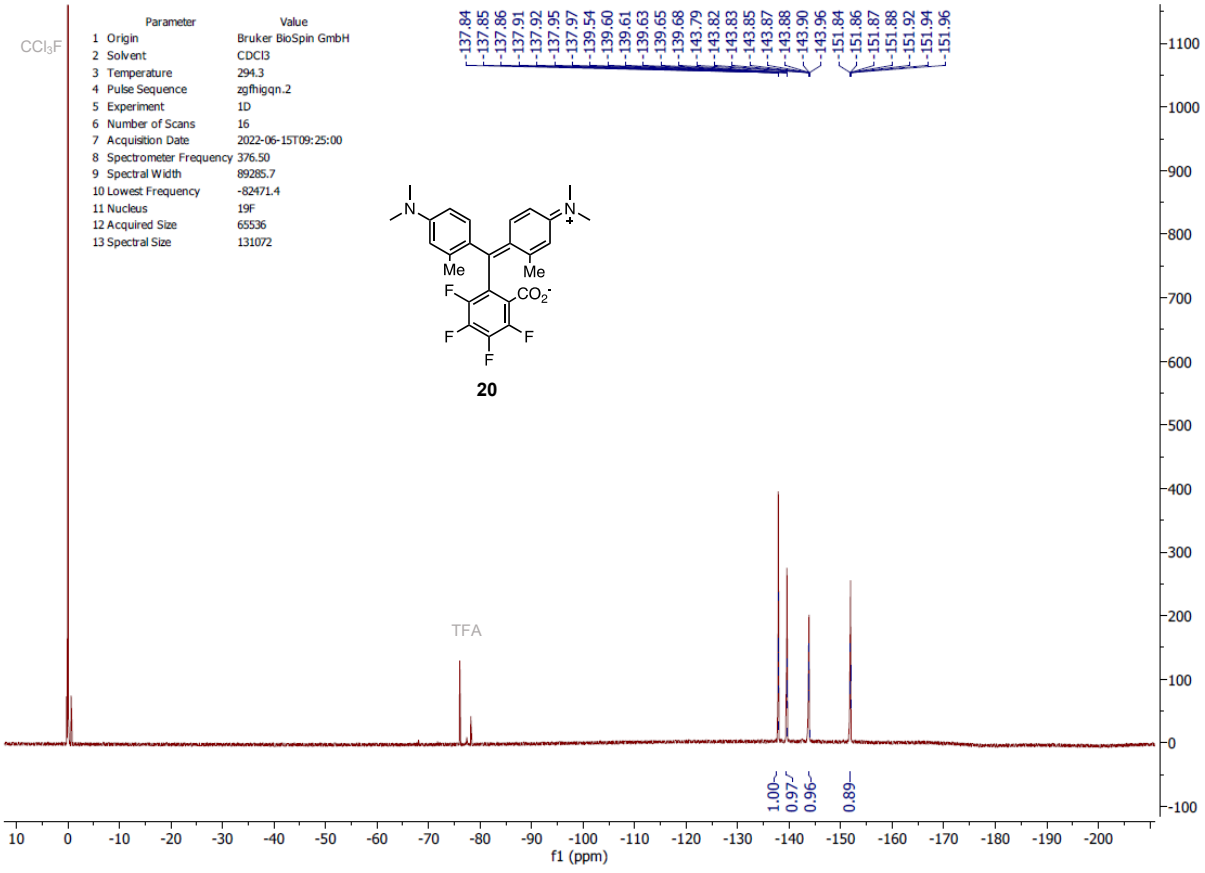

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 292.8
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-11-26T00:29:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82477.3
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
-118.79
-118.90
-118.95
-118.99
-136.03
-136.05
-136.07
-136.09
-136.10
-136.12
-136.18
-136.56
-136.61
-136.67
-140.31
-140.34
-140.34
-140.36
-140.39
-140.42
-140.44
-140.45
-148.26
-148.31
-148.36
19
2.16
1.06
1.16
1.05
1.00
f1 (ppm)



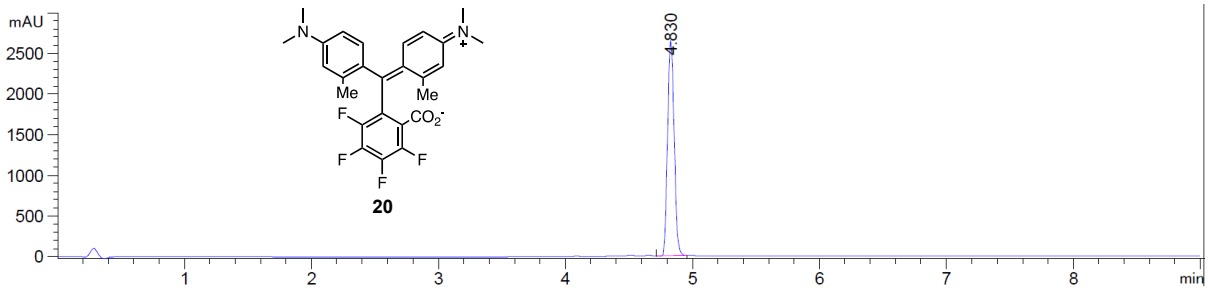
19

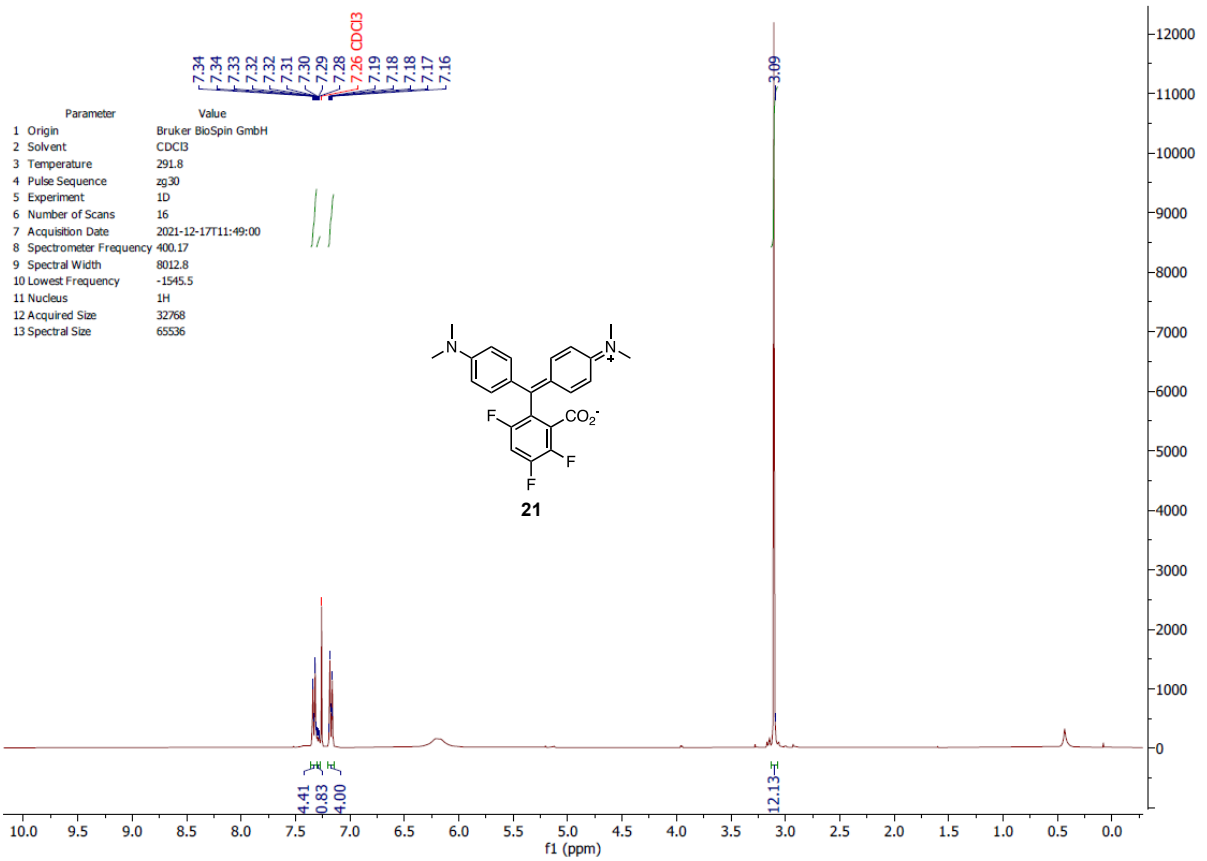

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.1
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-06-15T09:21:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.1
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
7.26 CDCl3
6.82
6.81
6.80
6.79
6.53
6.46
6.43
6.43
6.41
6.40
2.95
2.95
2.11
Me
Me
F
CO2-
F
F
F
EtOAc
2.00
1.86
1.88
12.00
6.02
f1 (ppm)

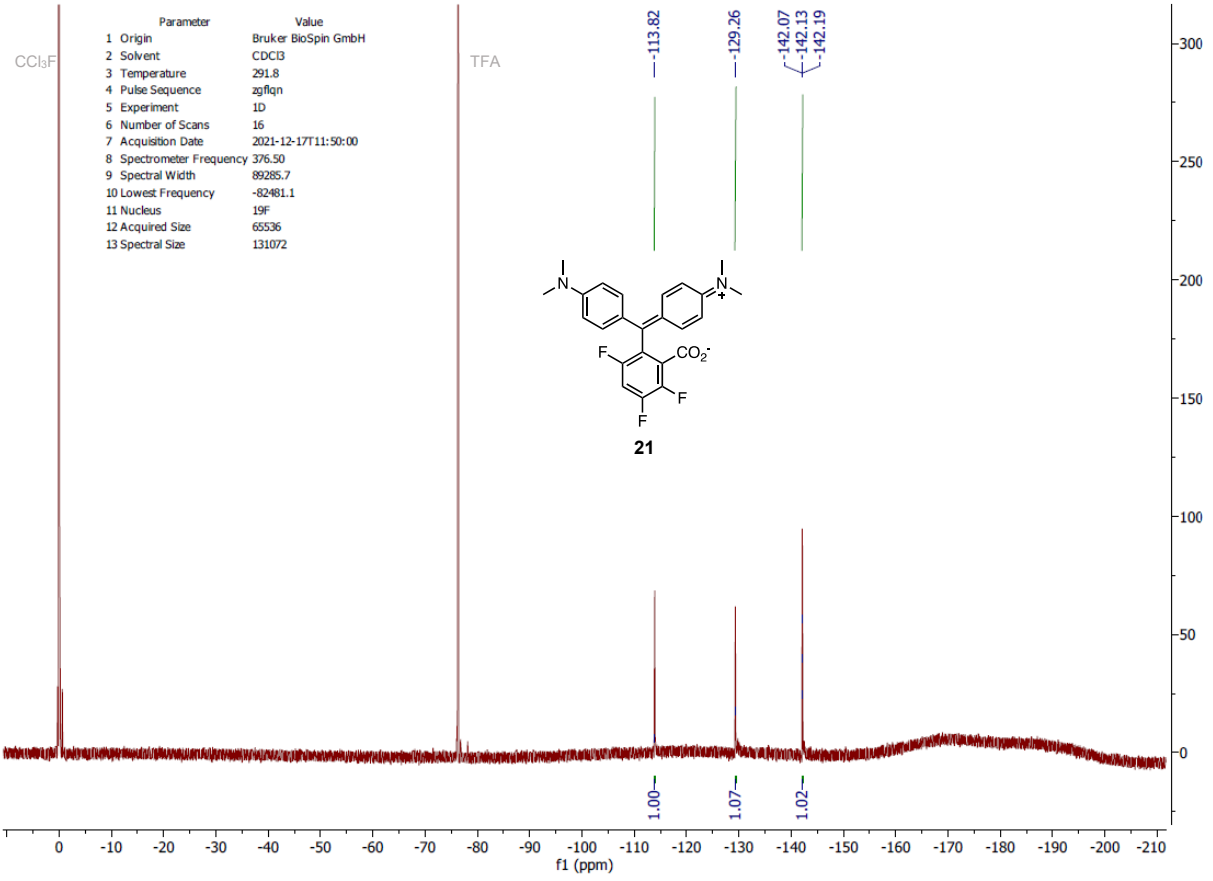
20

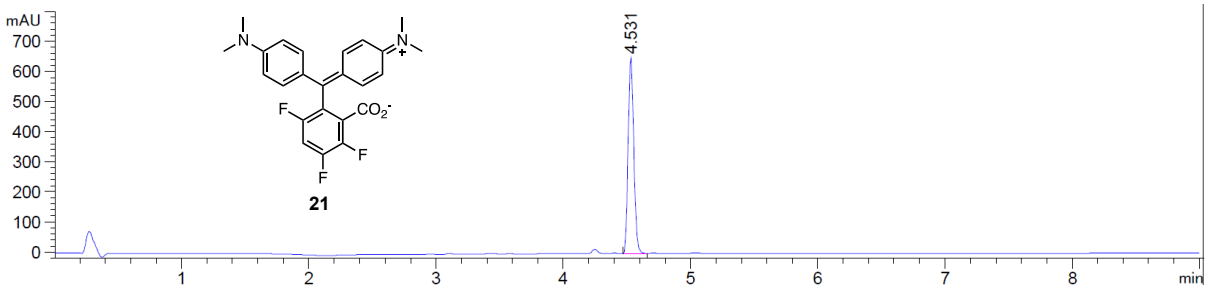

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.3
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-06-15T09:25:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82471.4
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
137.84
137.85
137.86
137.91
137.92
137.95
137.97
139.54
139.60
139.61
139.63
139.65
139.68
143.79
143.82
143.83
143.85
143.87
143.88
143.90
143.96
151.84
151.86
151.87
151.88
151.92
151.94
151.96
Me
Me
F
CO2-
F
F
F
TFA
1.00
0.97
0.96
0.89
f1 (ppm)

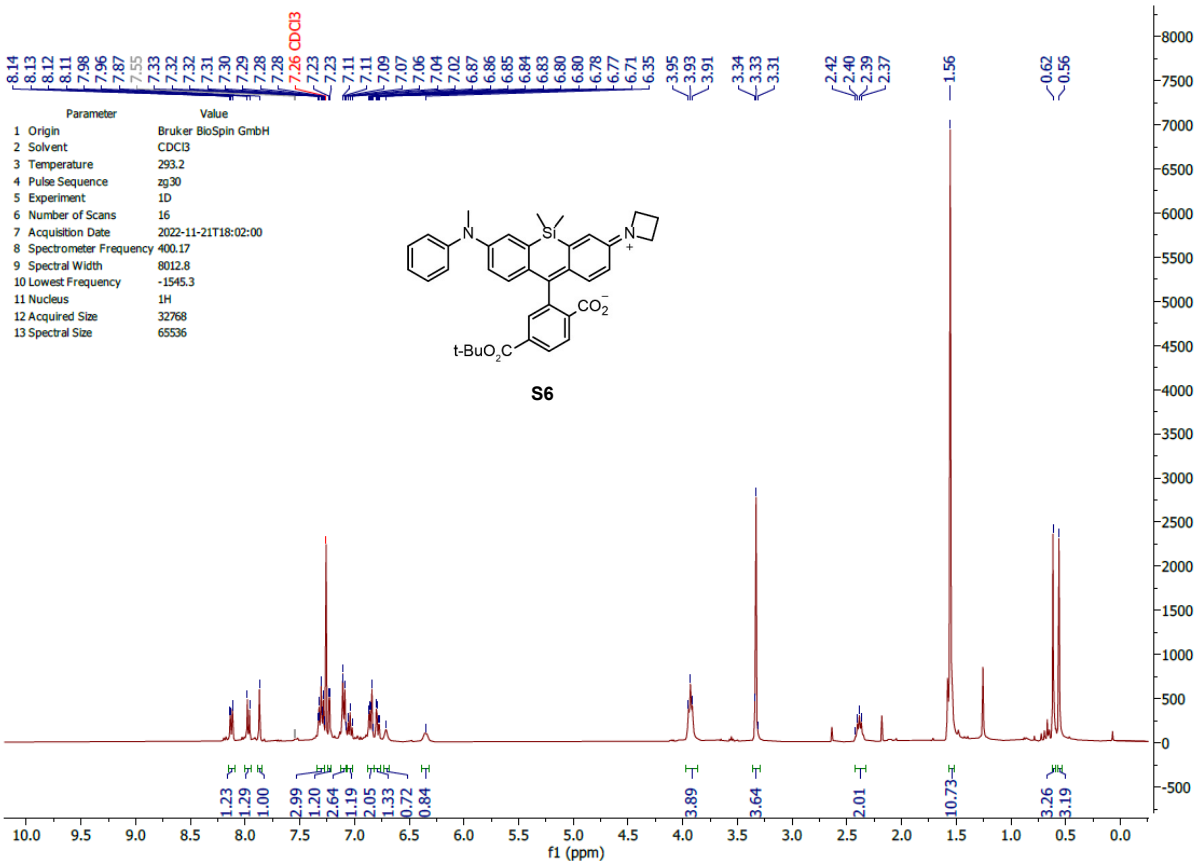
20

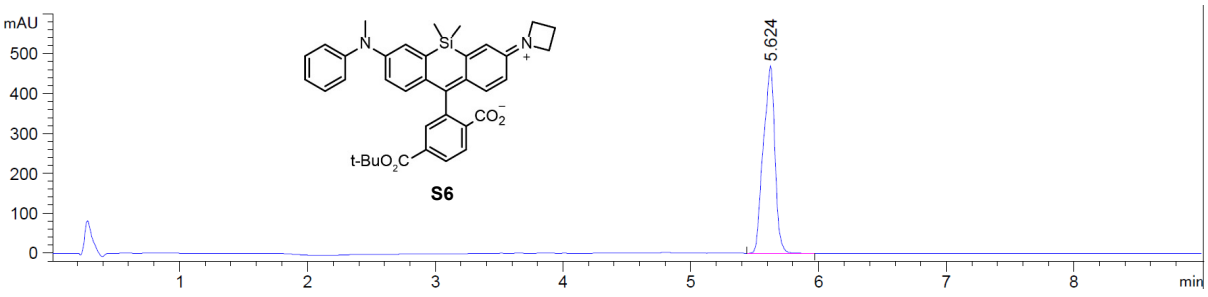

mAU
2500
2000
1500
1000
500
0
1
2
3
4
5
6
7
8
min
4.830
Me
Me
F
F
F
F
CO2-
20

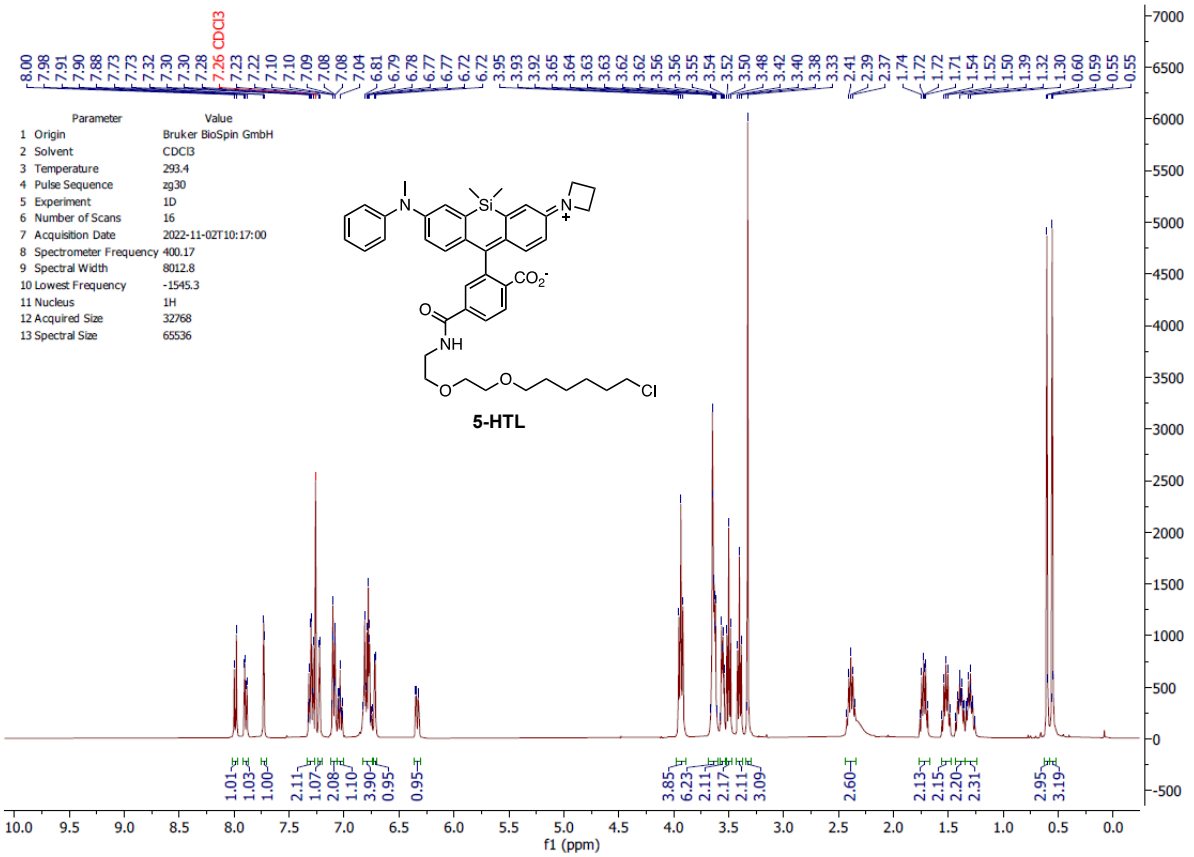

7.34
7.34
7.33
7.32
7.32
7.31
7.30
7.29
7.28
CDCl3
7.26
7.19
7.18
7.18
7.17
7.16
3.09
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 291.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-12-17T11:49:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
21
4.41
0.83
4.00
12.13
f1 (ppm)

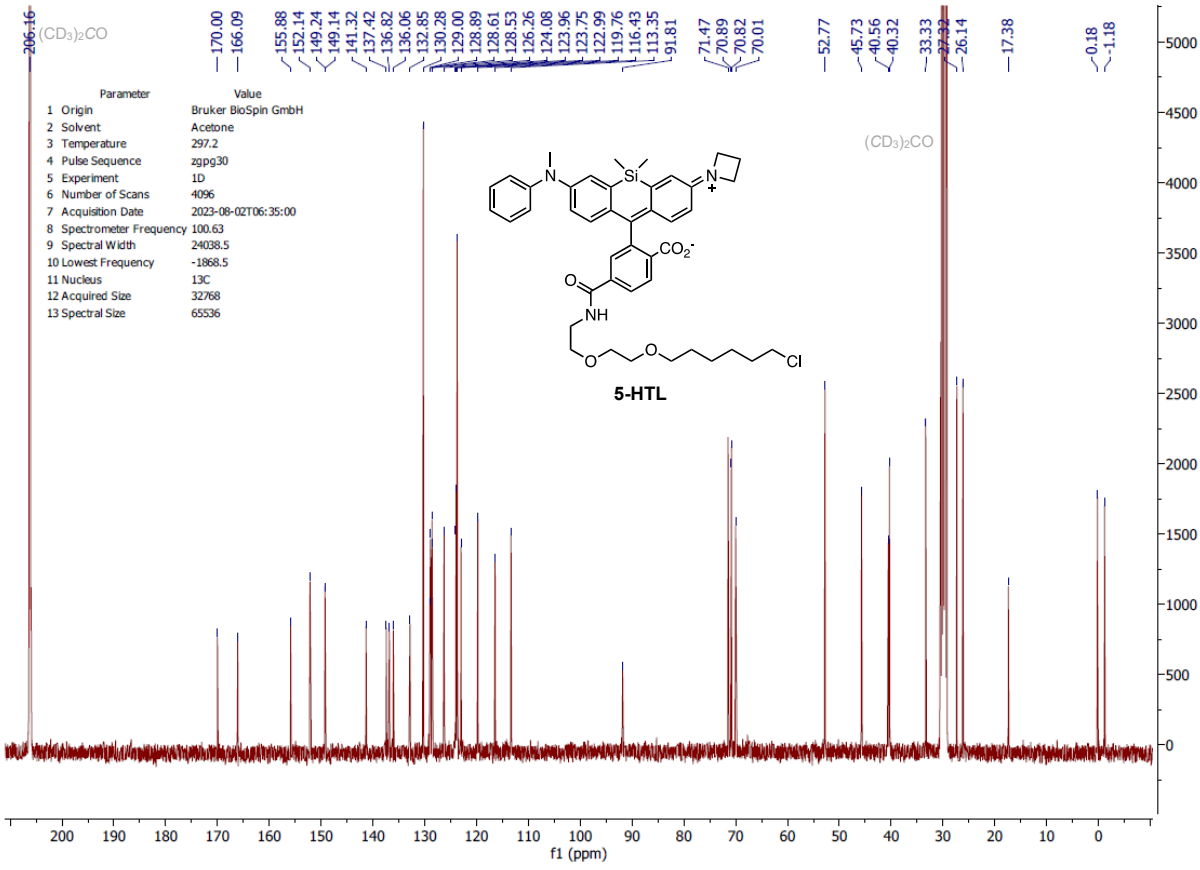

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 291.8
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-12-17T11:50:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82481.1
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
-113.82
-129.26
-142.07
-142.13
-142.19
21
1.00
1.07
1.02
f1 (ppm)

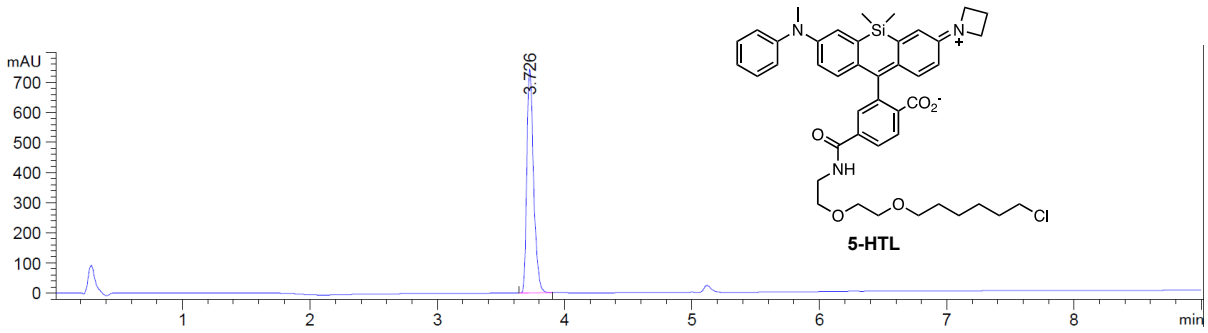

mAU
700
600
500
400
300
200
100
0
1
2
3
4
5
6
7
8
min
4.531
F
CO2-
F
F
21


8.14 8.13 8.12 8.11 7.98 7.96 7.87 7.55 7.33 7.32 7.32 7.31 7.30 7.29 7.28 7.28 7.26 CDCl3 7.23 7.23 7.11 7.11 7.09 7.07 7.06 7.04 7.02 6.87 6.86 6.85 6.84 6.83 6.80 6.80 6.78 6.77 6.71 6.35 3.95 3.93 3.91 3.34 3.33 3.31 2.42 2.40 2.39 2.37 1.56 0.62 0.56
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.2
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-11-21T18:02:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
t-BuO2C
CO2−
S6
1.23 1.29 1.00 2.99 1.20 2.64 1.19 2.05 1.33 0.72 0.84 3.89 3.64 2.01 10.73 3.26 3.19
f1 (ppm)

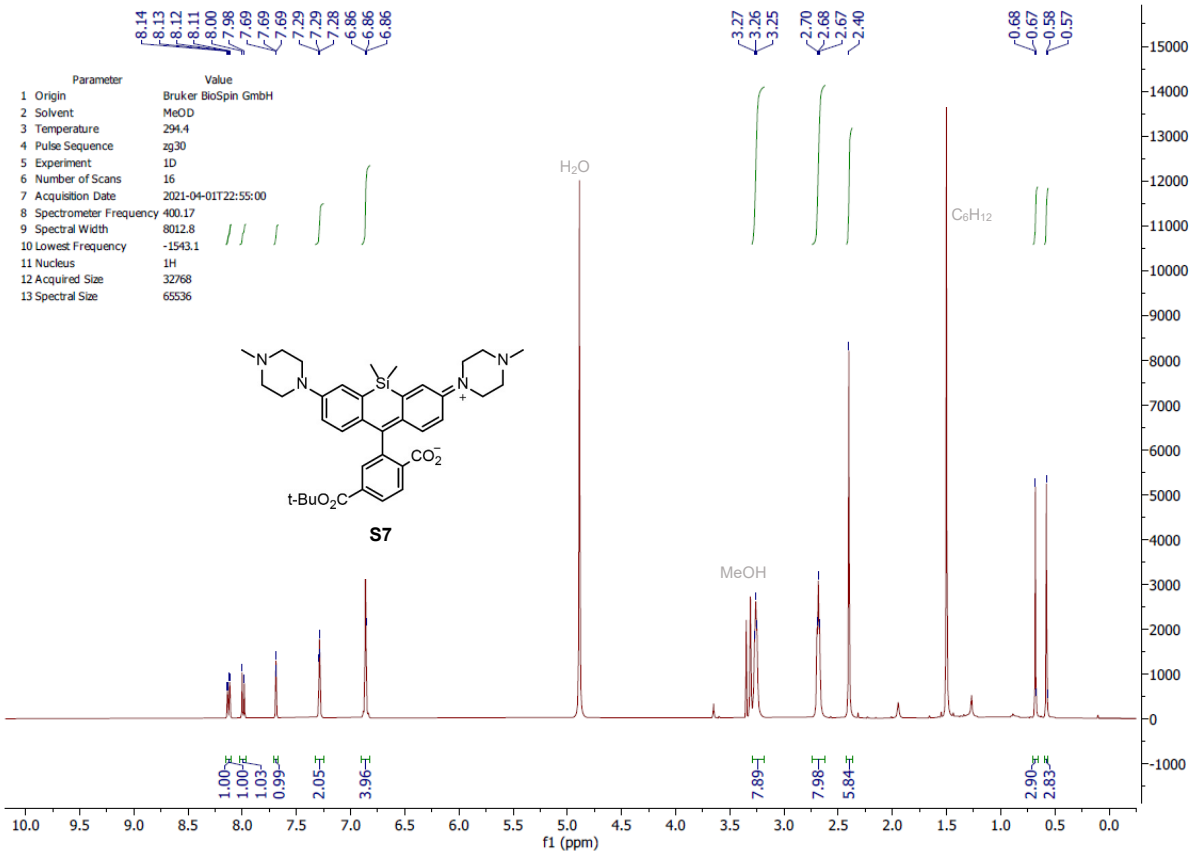

mAU
5.624
t-BuO2C
CO2−
S6
min

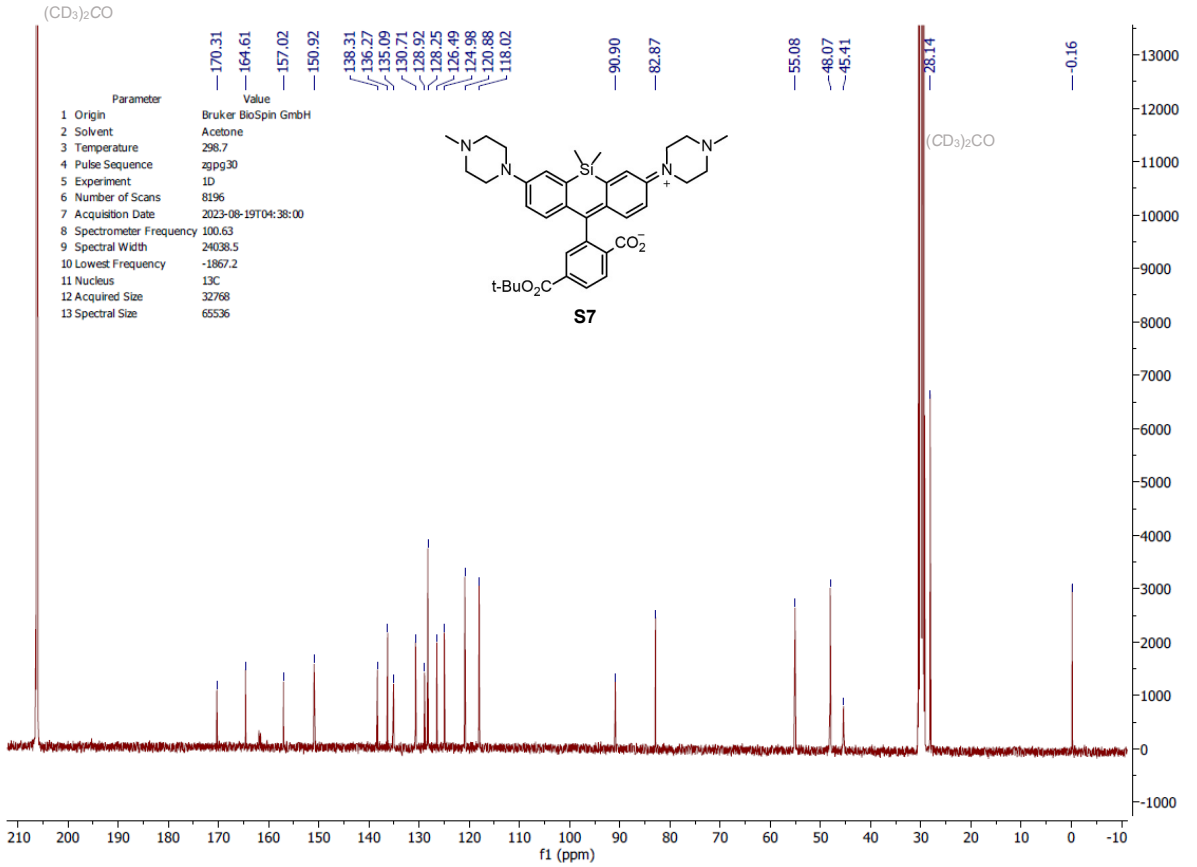

CDCl3
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.4
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-11-02T10:17:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
5-HTL
f1 (ppm)

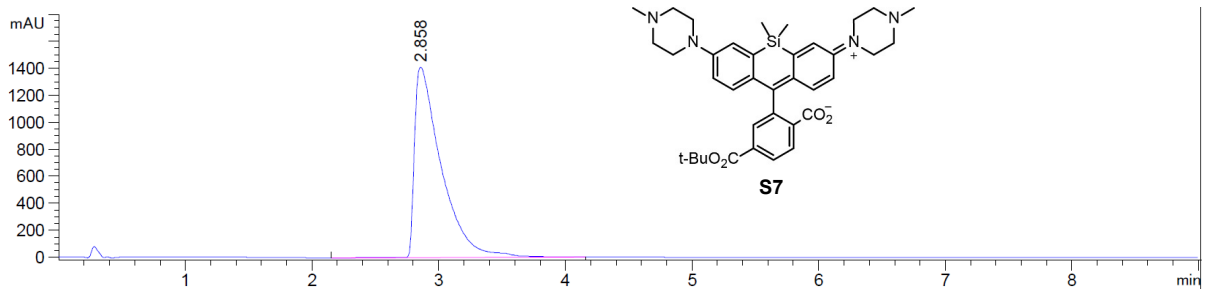

(CD3)2CO
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 297.2
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 4096
7 Acquisition Date 2023-08-02T06:35:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1868.5
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
5-HTL
f1 (ppm)


mAU
700
600
500
400
300
200
100
0
3.726
1
2
3
4
5
6
7
8
min
Si
N
N
CO2
O
NH
O
O
Cl

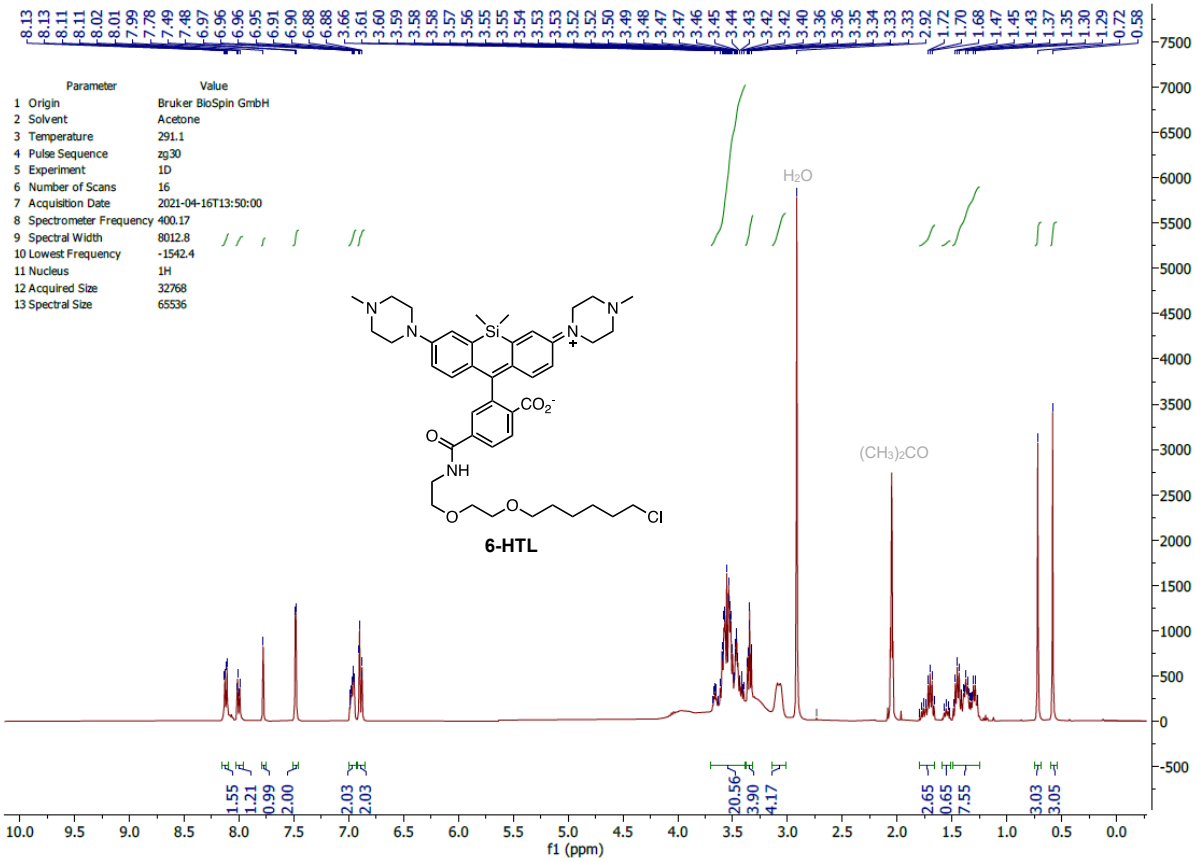
5-HTL

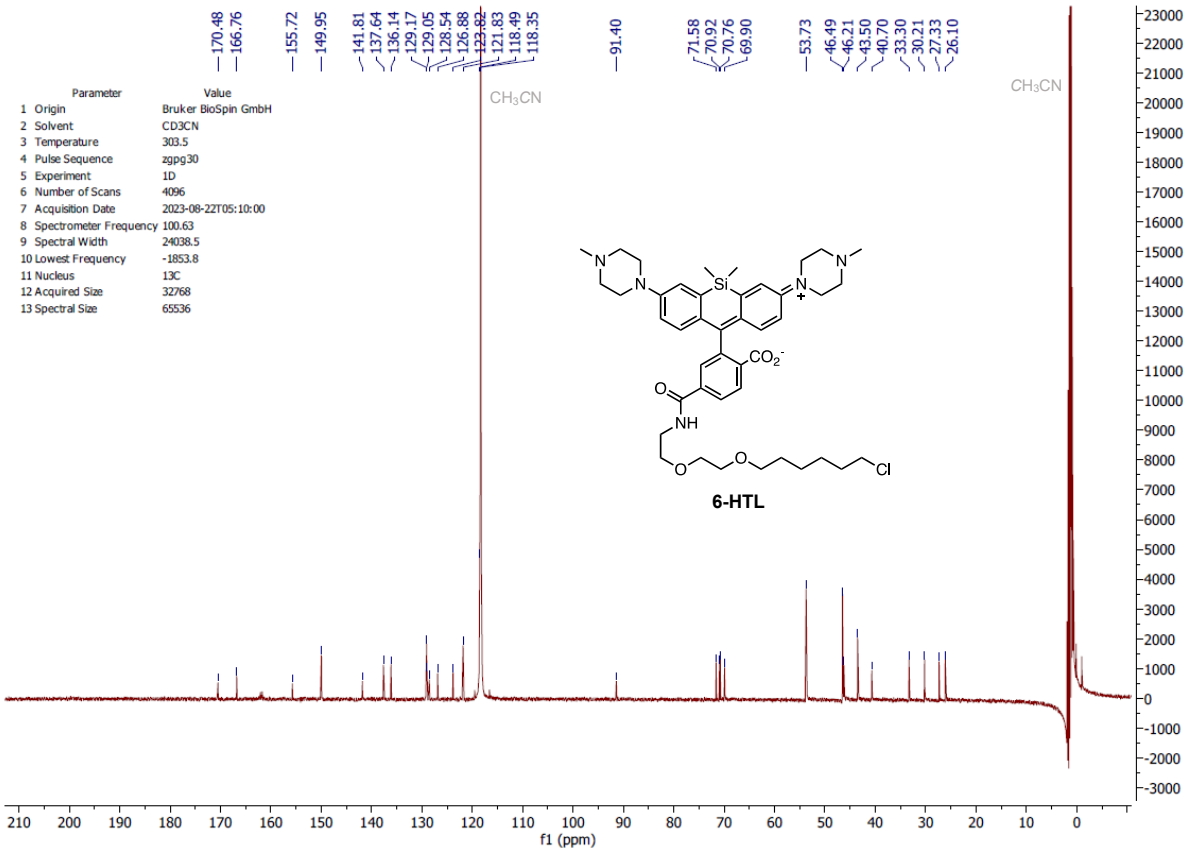

8.14
8.13
8.12
8.11
8.00
7.98
7.69
7.69
7.69
7.29
7.29
7.28
6.86
6.86
6.86
3.27
3.26
3.25
2.70
2.68
2.67
2.40
0.68
0.67
0.58
0.57
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent MeOD
3 Temperature 294.4
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-04-01T22:55:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1543.1
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
C6H12
MeOH
S7
1.00
1.00
1.03
0.99
2.05
3.96
7.89
7.98
5.84
2.90
2.83
f1 (ppm)

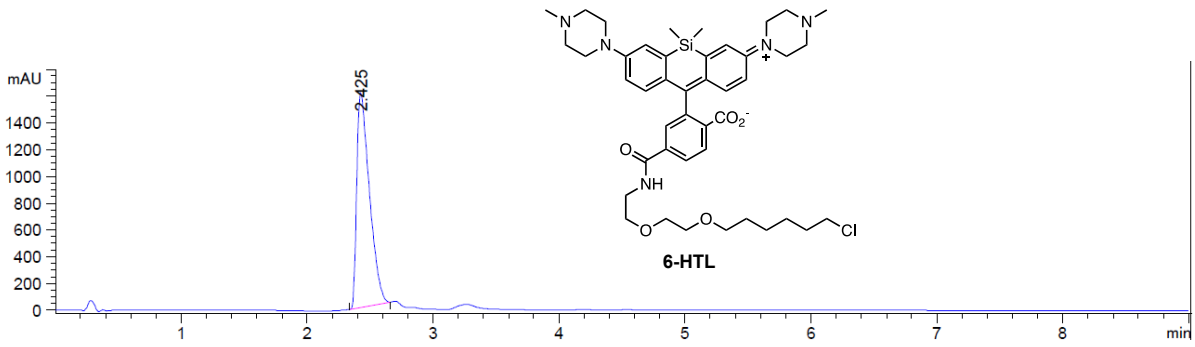

(CD3)2CO
170.31
164.61
157.02
150.92
138.31
136.27
135.09
130.71
128.92
128.25
126.49
124.98
120.88
118.02
90.90
82.87
55.08
48.07
45.41
28.14
-0.16
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 298.7
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 8196
7 Acquisition Date 2023-08-19T04:38:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1867.2
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
S7
(CD3)2CO
f1 (ppm)

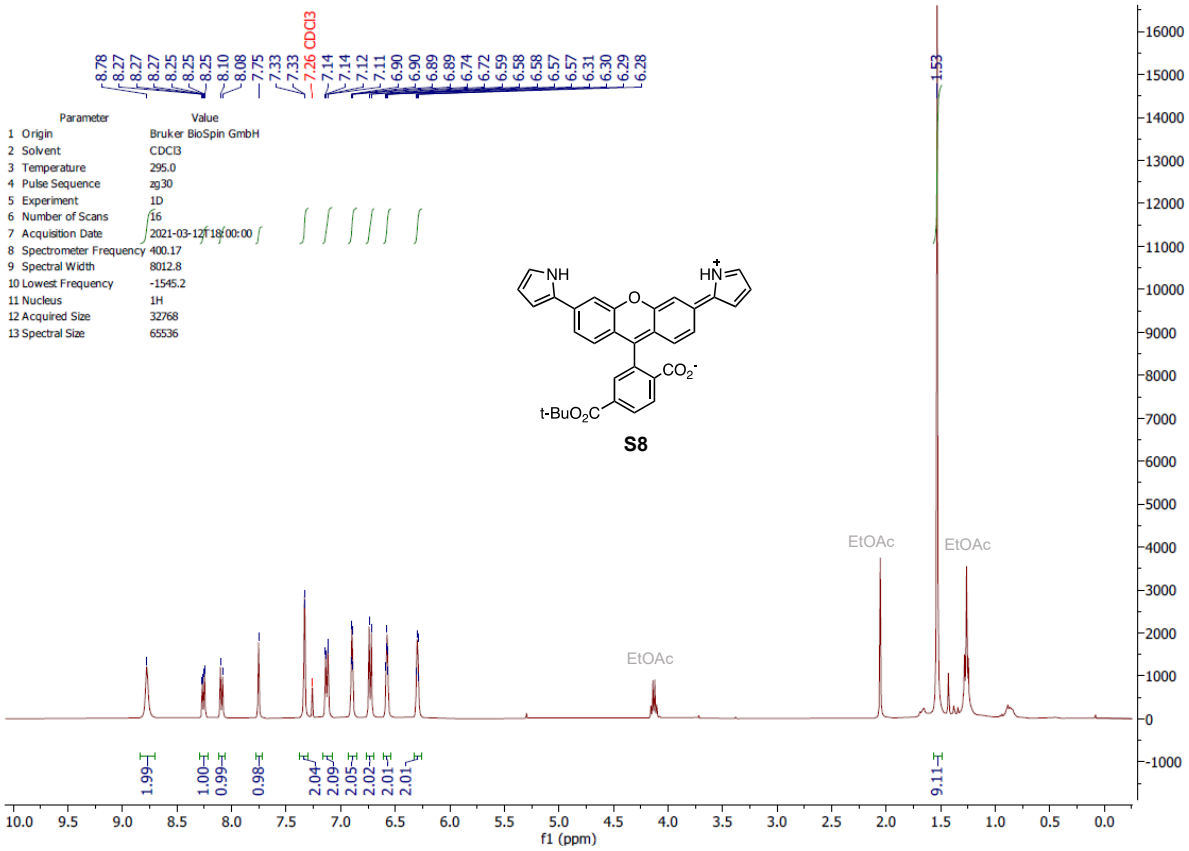

mAU
1400
1200
1000
800
600
400
200
0
2.858
1
2
3
4
5
6
7
8
min
N
N
Si
N
+
N
CO2−
t-BuO2C

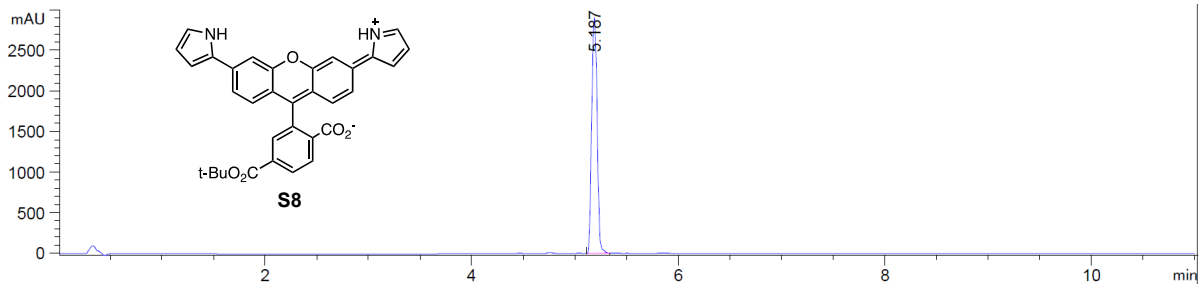
S7

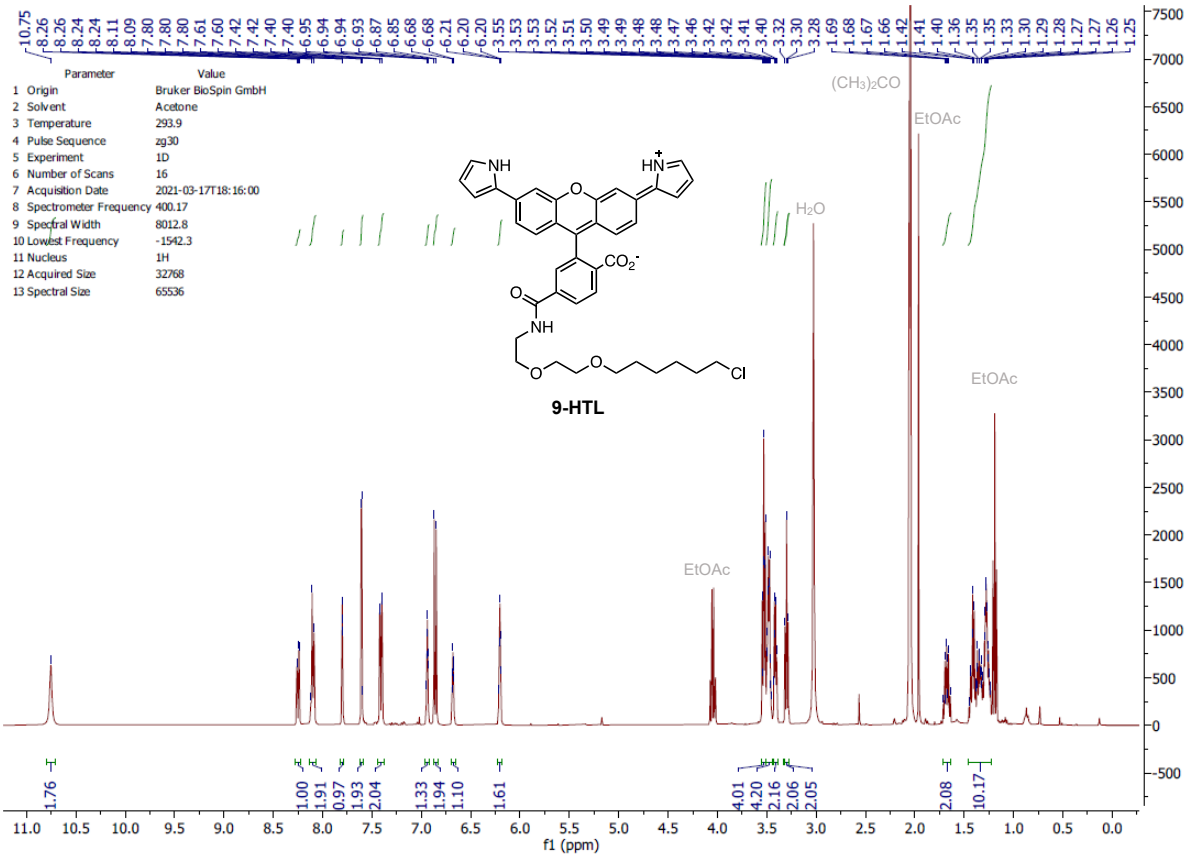

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 291.1
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-04-16T13:50:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.4
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
(CH3)2CO
6-HTL
f1 (ppm)

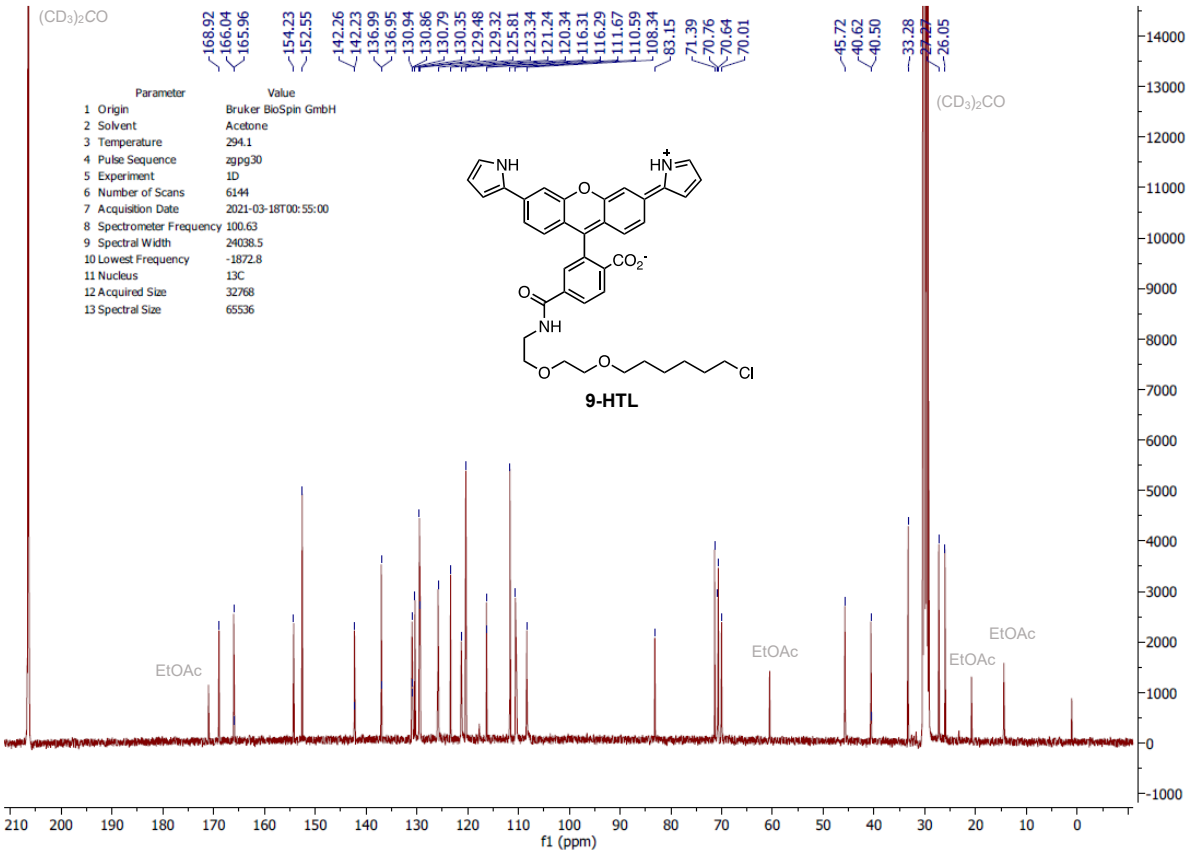

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 303.5
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 4096
7 Acquisition Date 2023-08-22T05:10:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1853.8
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
CH3CN
CH3CN
6-HTL
f1 (ppm)

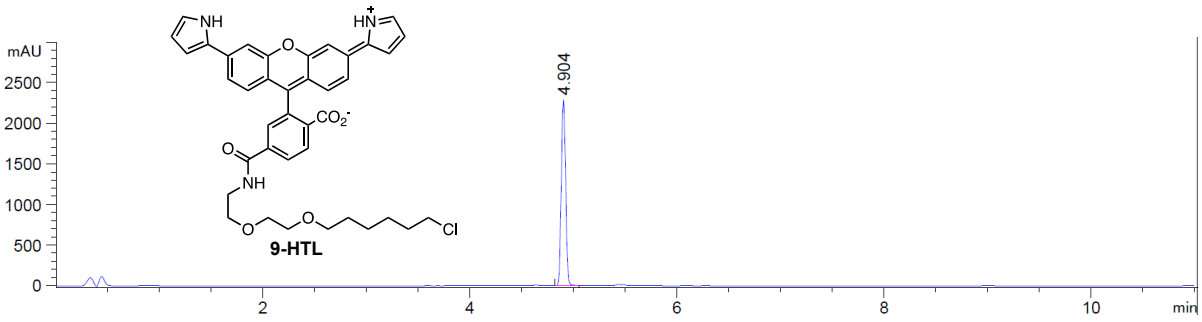

mAU
1400
1200
1000
800
600
400
200
0
2.425
1
2
3
4
5
6
7
8
min
Si
N
O
NH
Cl
6-HTL

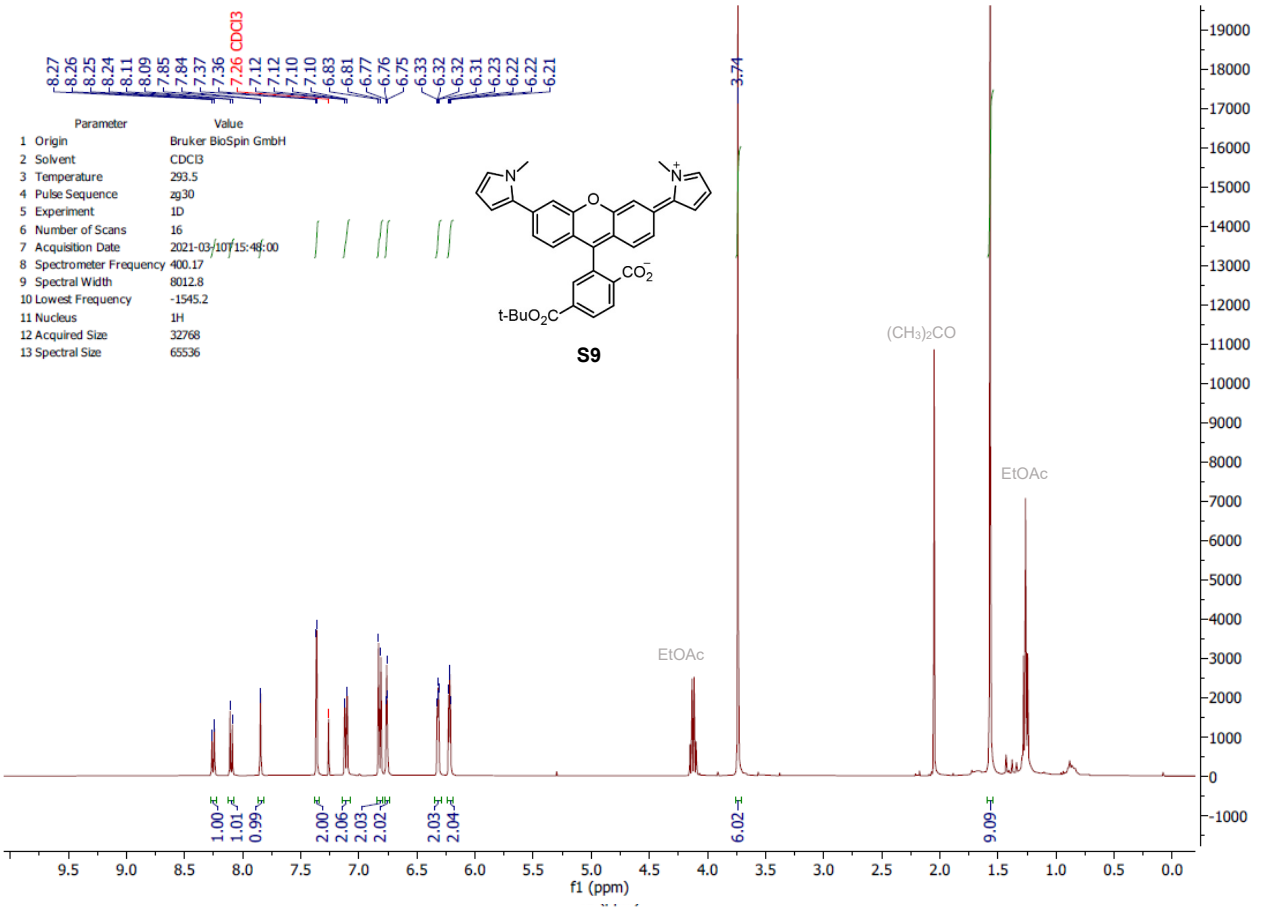

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 295.0
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-03-12T18:00:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.2
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
CDCl3
EtOAc
EtOAc
EtOAc
f1 (ppm)
S8

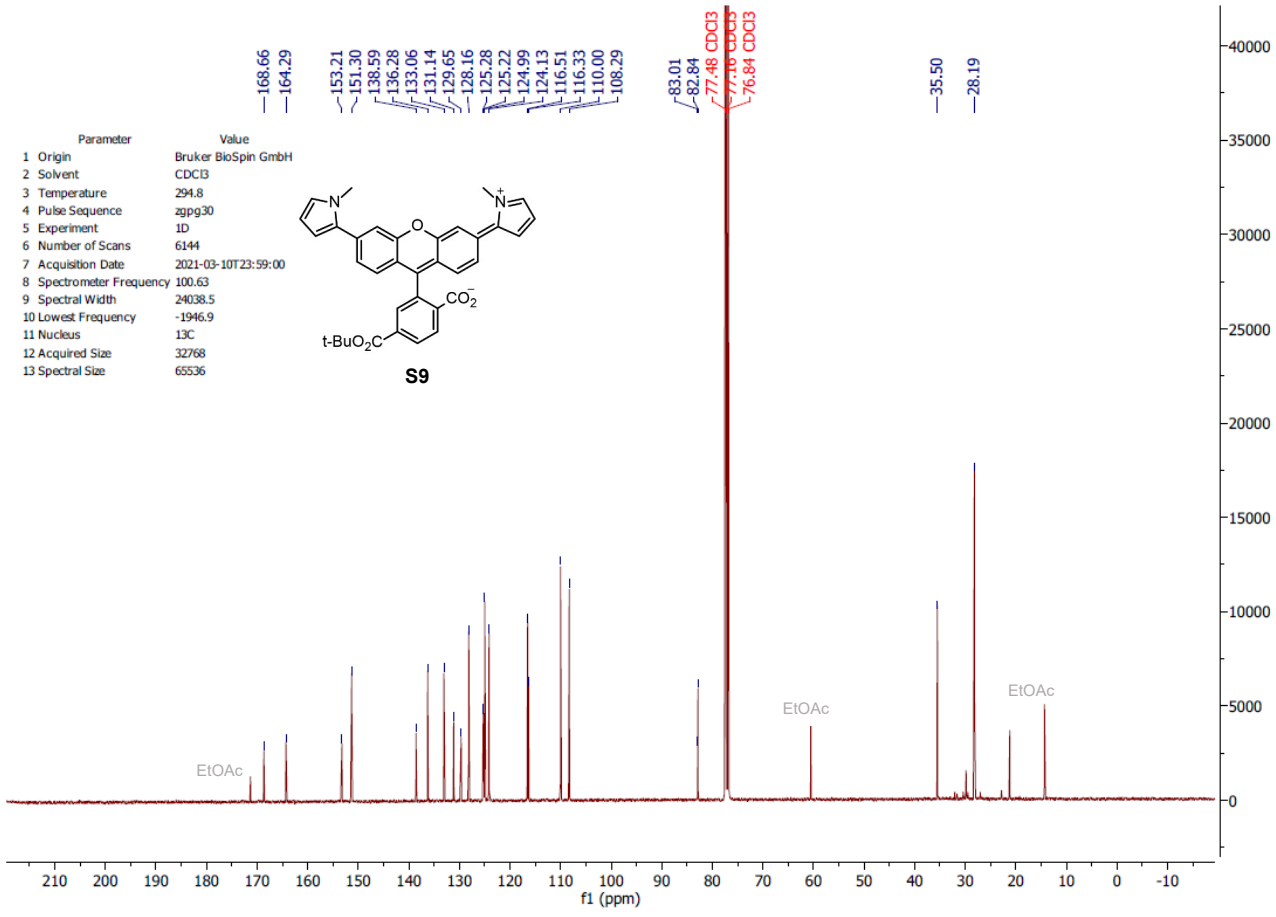

mAU
5.187
min
S8

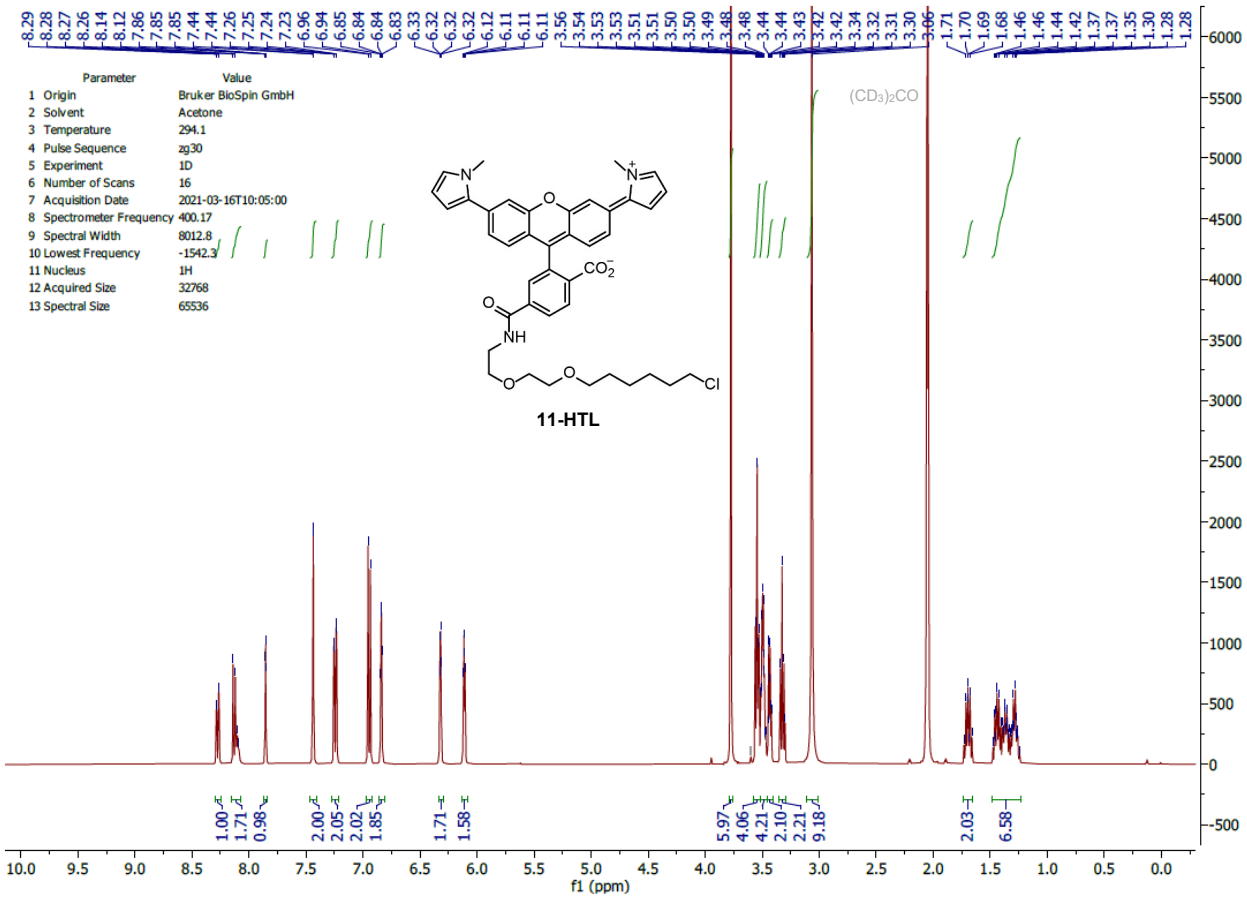

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 293.9
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-03-17T18:16:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
(CH3)2CO
EtOAc
H2O
EtOAc
EtOAc
9-HTL
f1 (ppm)

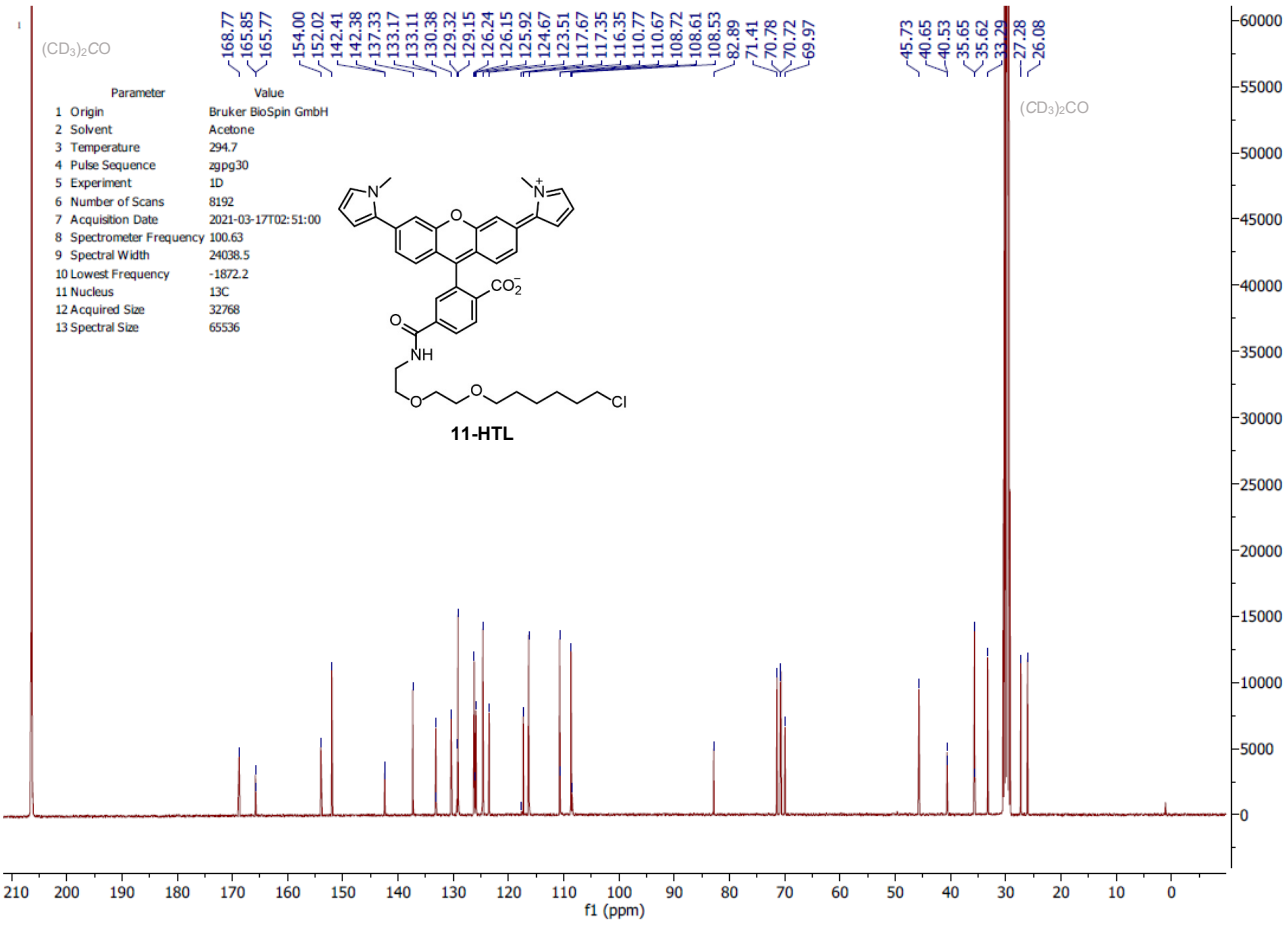

(CD3)2CO
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 294.1
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 6144
7 Acquisition Date 2021-03-18T00:55:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1872.8
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
(CD3)2CO
EtOAc
EtOAc
EtOAc
EtOAc
9-HTL
f1 (ppm)

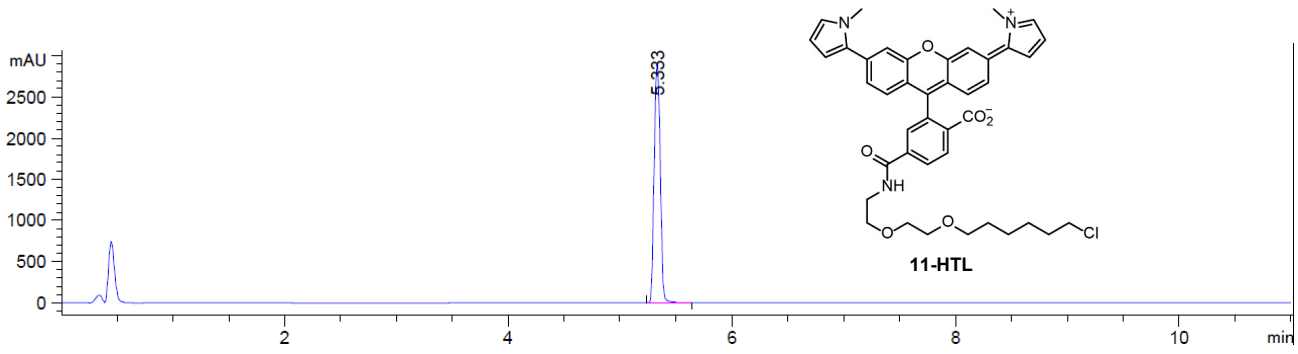

NH
HN
O
CO2-
O
NH
O
O
Cl
9-HTL
mAU
2500
2000
1500
1000
500
0
4.904
2
4
6
8
10
min

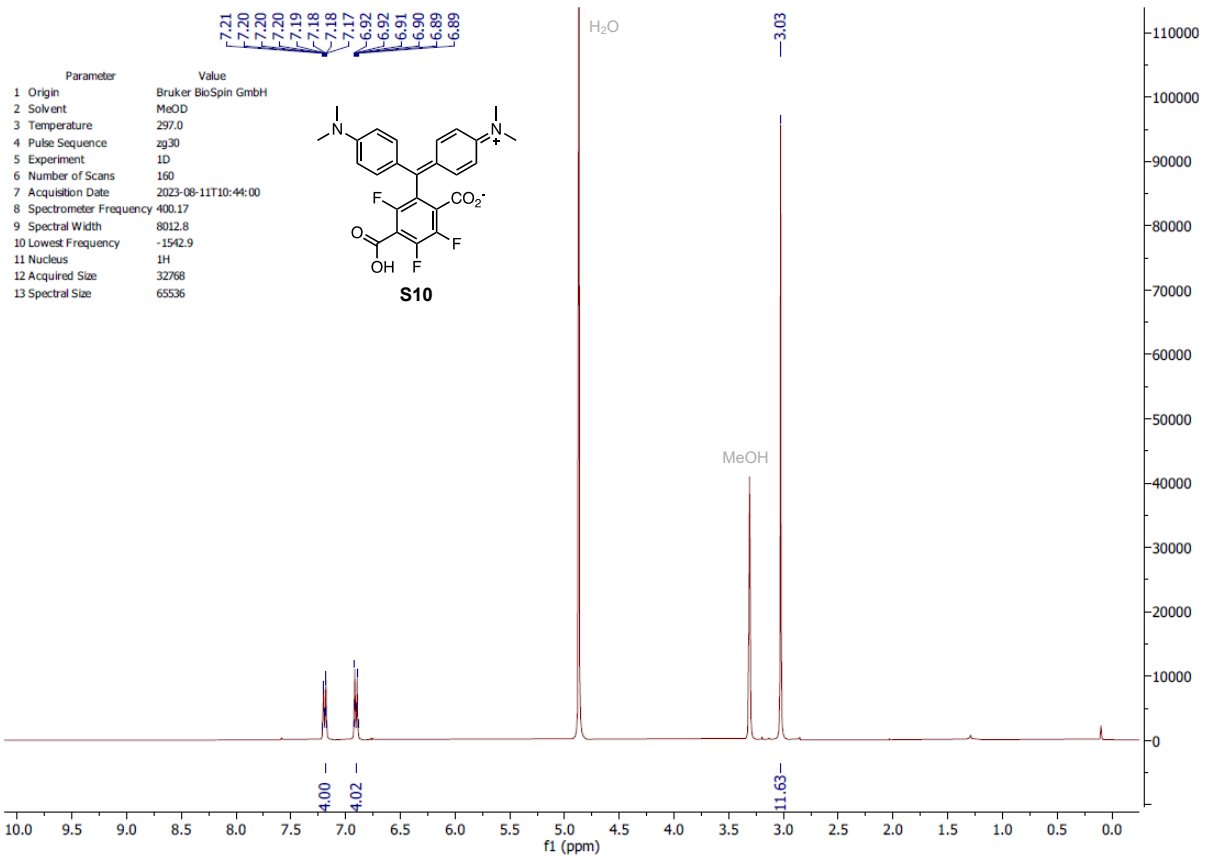

8.27
8.26
8.25
8.24
8.11
8.09
7.85
7.84
7.37
7.36
7.26 CDCl3
7.12
7.10
7.10
6.83
6.81
6.77
6.76
6.75
6.33
6.32
6.32
6.31
6.23
6.22
6.22
6.21
3.74
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.5
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-03-10T15:48:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.2
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
S9
(CH3)2CO
EtOAc
EtOAc
1.00
1.01
0.99
2.00
2.06
2.03
2.02
2.03
2.04
6.02
9.09
f1 (ppm)

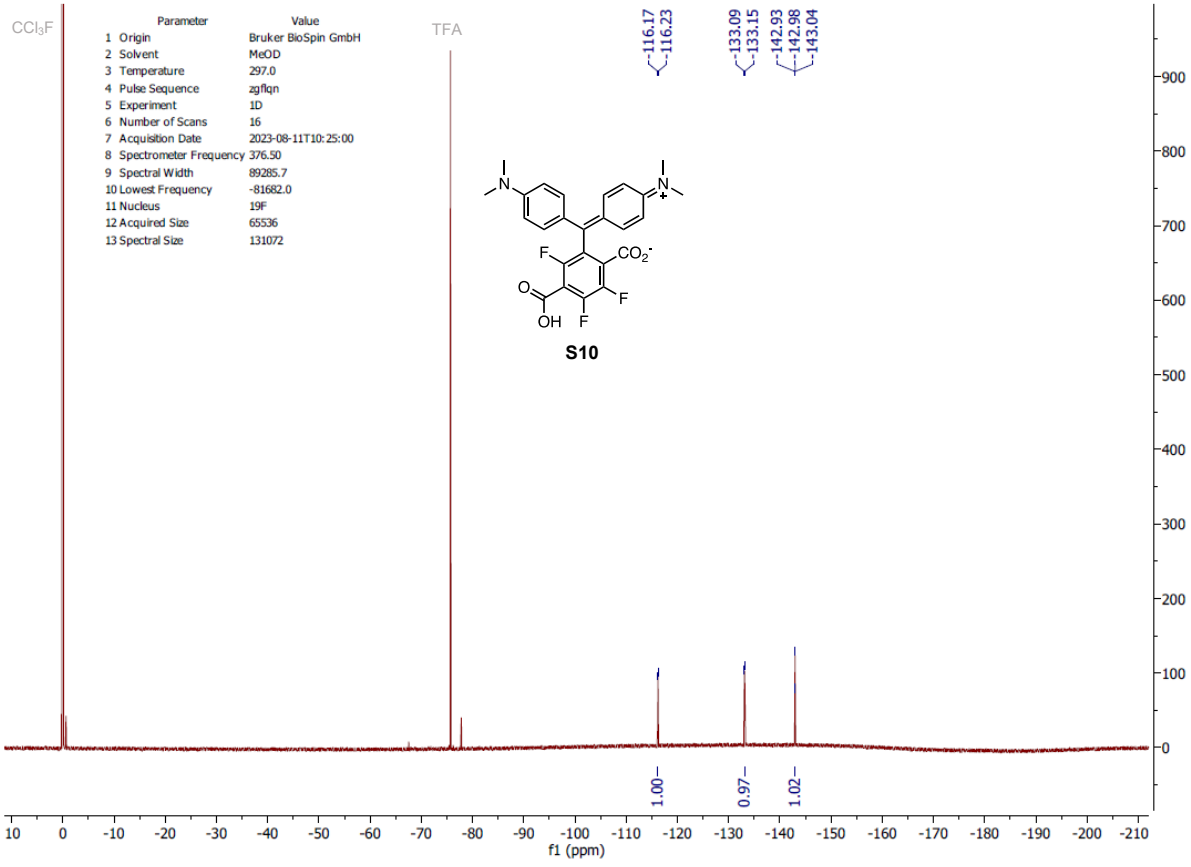

168.66
164.29
153.21
151.30
138.59
136.28
133.06
131.14
129.65
128.16
125.28
125.22
124.99
124.13
116.51
116.33
110.00
108.29
83.01
82.84
77.48 CDCl3
77.16 CDCl3
76.84 CDCl3
35.50
28.19
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.8
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 6144
7 Acquisition Date 2021-03-10T23:59:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1946.9
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
S9
EtOAc
EtOAc
EtOAc
f1 (ppm)

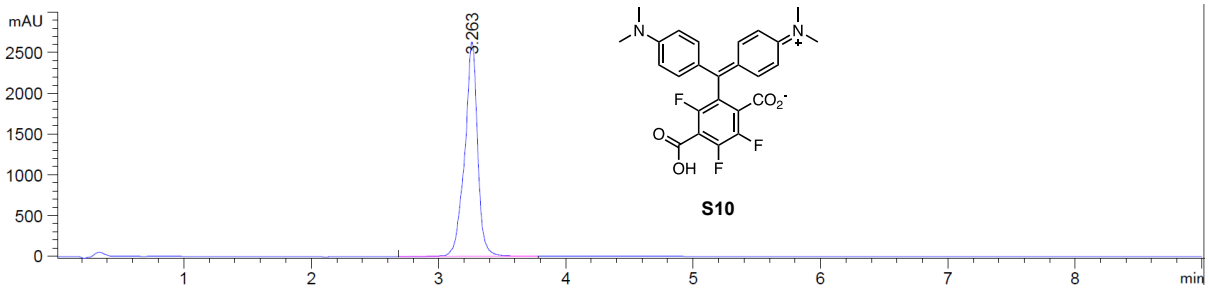

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 294.1
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2021-03-16T10:05:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
(CD3)2CO
11-HTL
f1 (ppm)


(CD3)2CO
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 294.7
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 8192
7 Acquisition Date 2021-03-17T02:51:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1872.2
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
11-HTL
f1 (ppm)

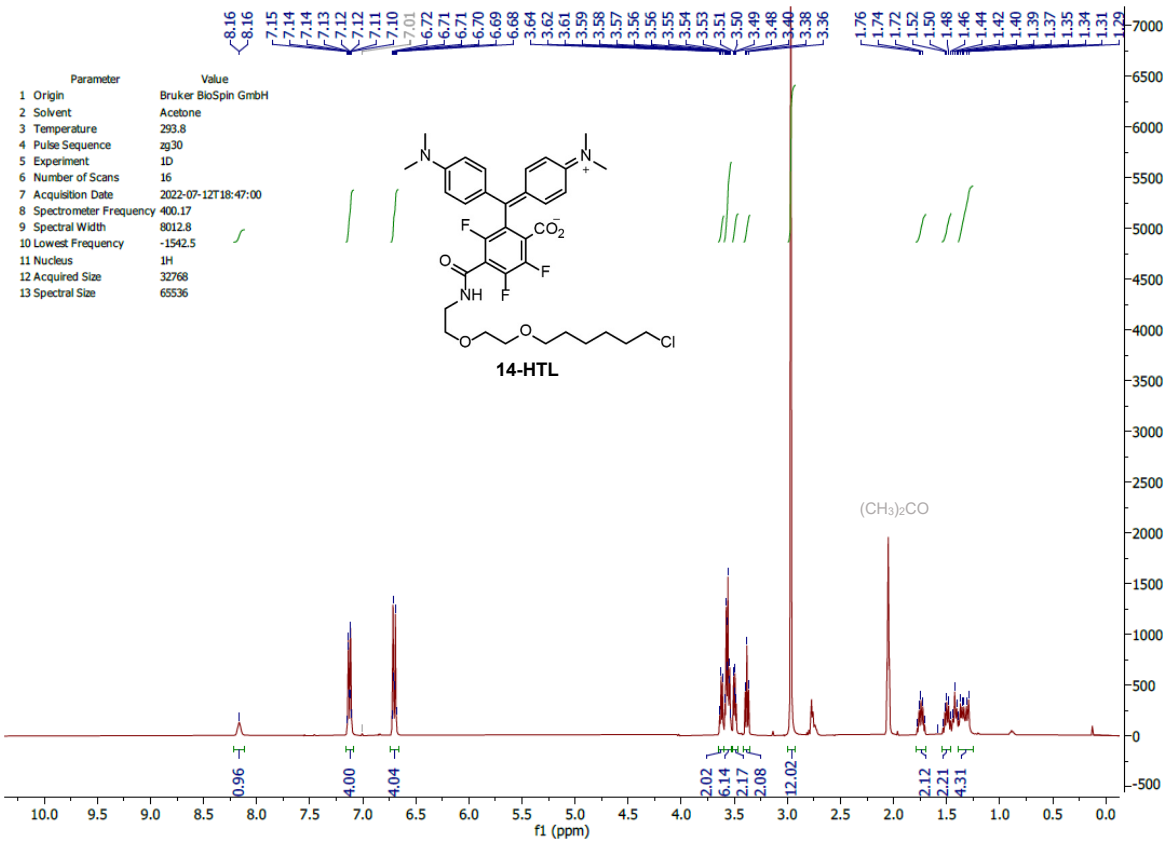

mAU
2500
2000
1500
1000
500
0
5.333
2
4
6
8
10
min
11-HTL

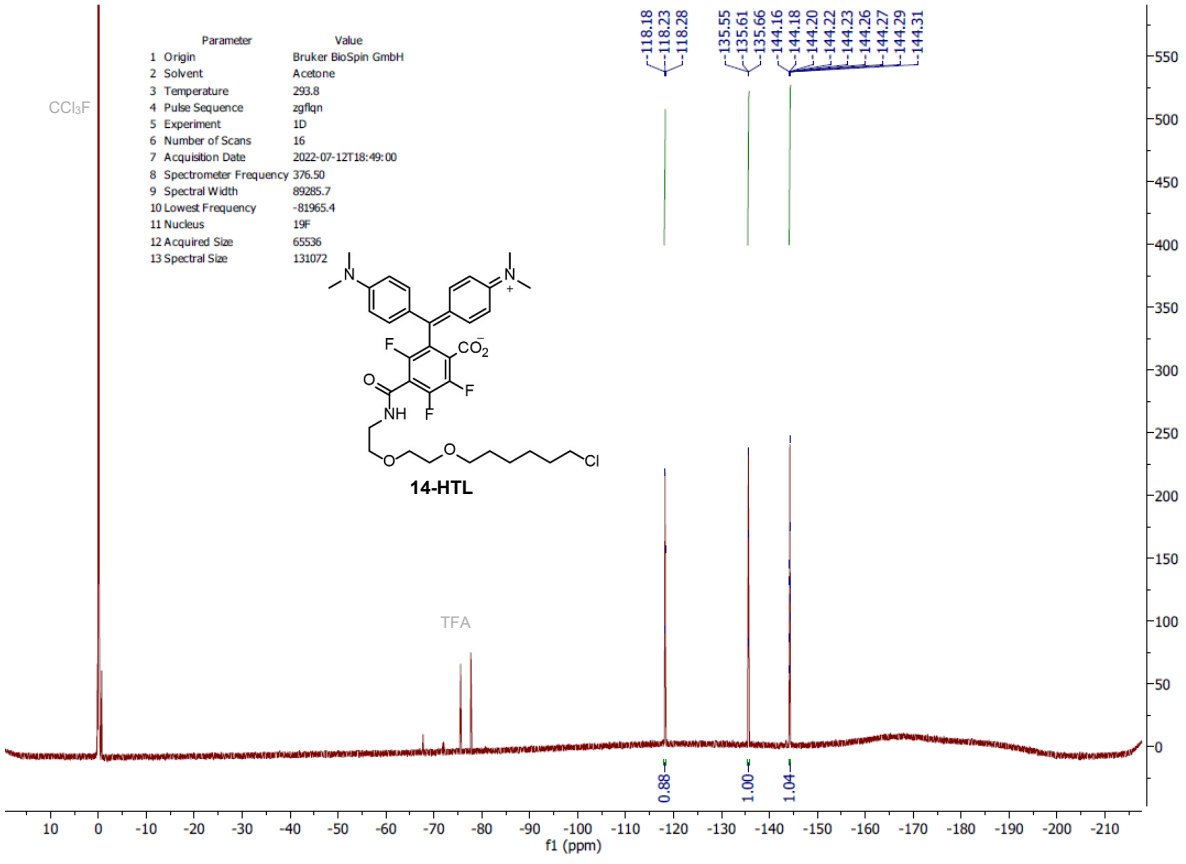

7.21
7.20
7.20
7.20
7.19
7.18
7.17
6.92
6.92
6.91
6.90
6.89
6.89
H2O
3.03
MeOH
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent MeOD
3 Temperature 297.0
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 160
7 Acquisition Date 2023-08-11T10:44:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.9
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
S10
4.00
4.02
11.63
f1 (ppm)

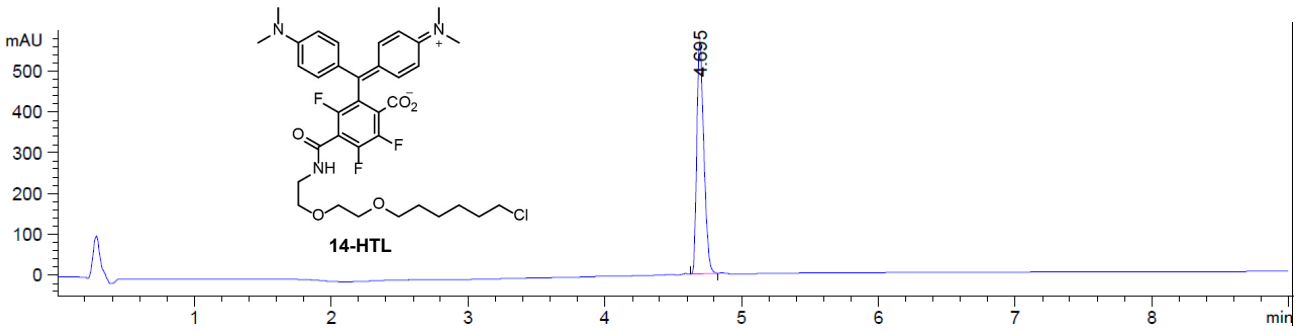

CCl3F
TFA
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent MeOD
3 Temperature 297.0
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-11T10:25:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81682.0
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
116.17
116.23
133.09
133.15
142.93
142.98
143.04
S10
1.00
0.97
1.02
f1 (ppm)

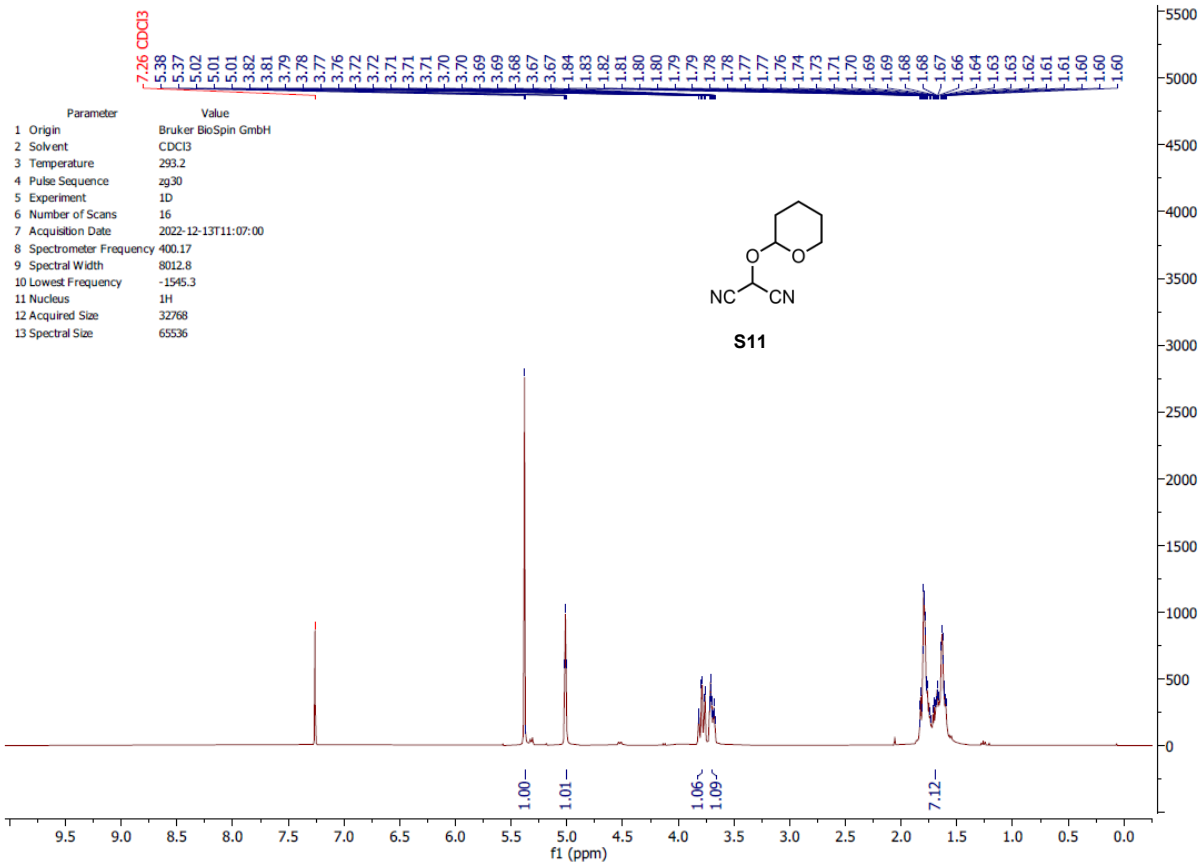


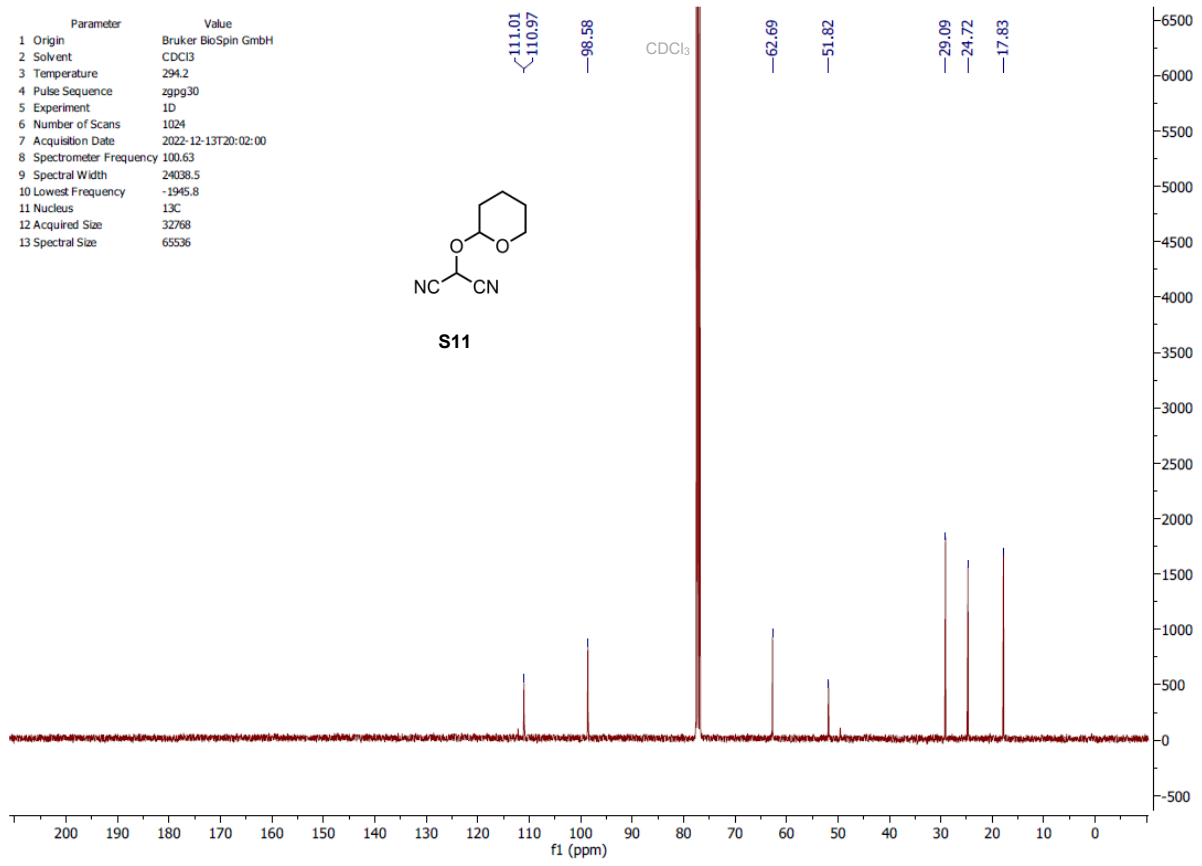
S10

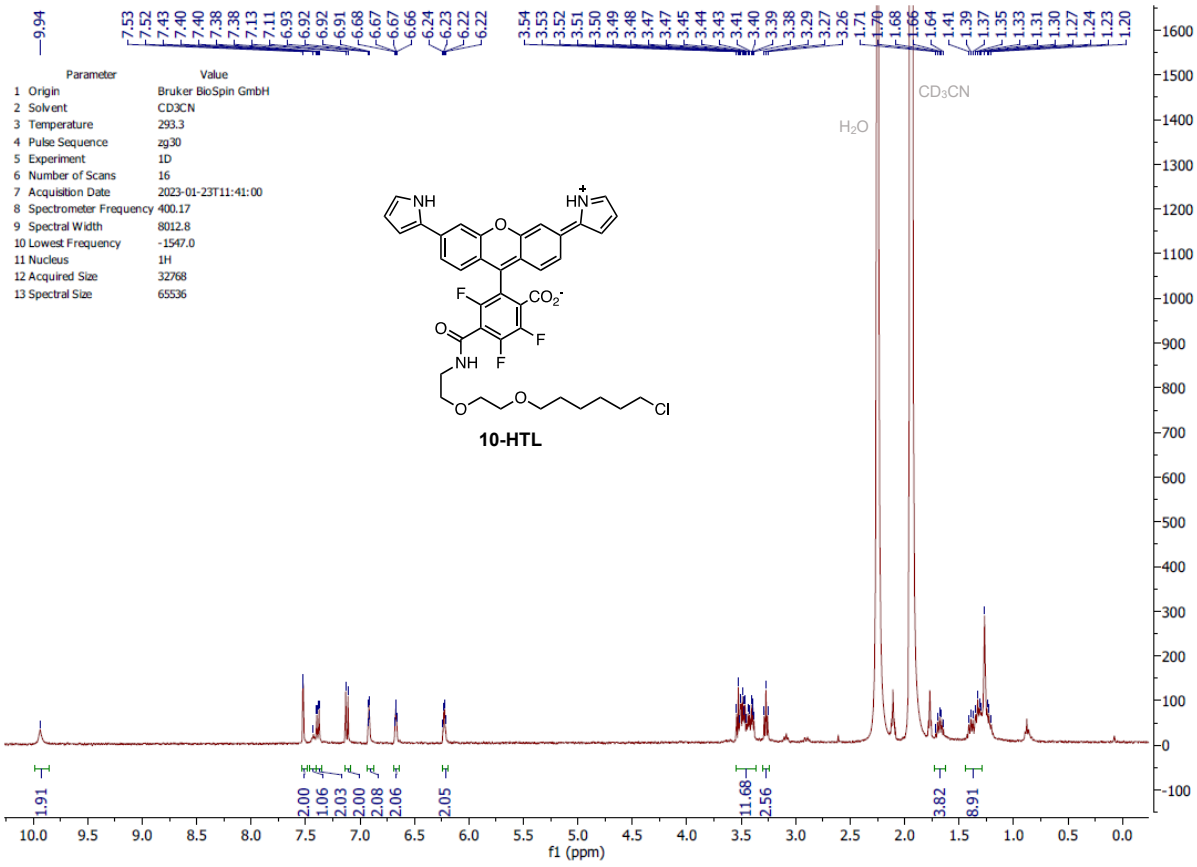

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 293.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-07-12T18:47:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
14-HTL
(CH3)2CO
f1 (ppm)

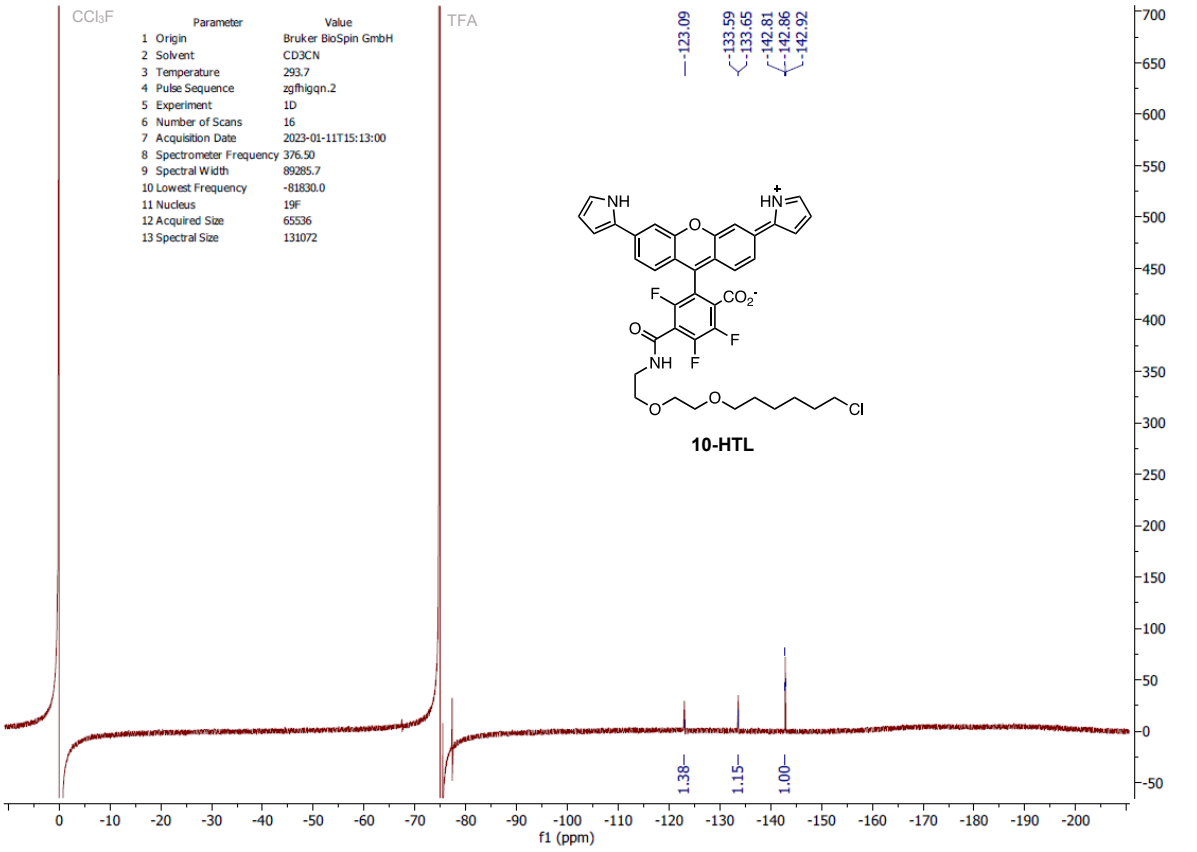

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 293.8
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-07-12T18:49:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81965.4
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
14-HTL
TFA
f1 (ppm)

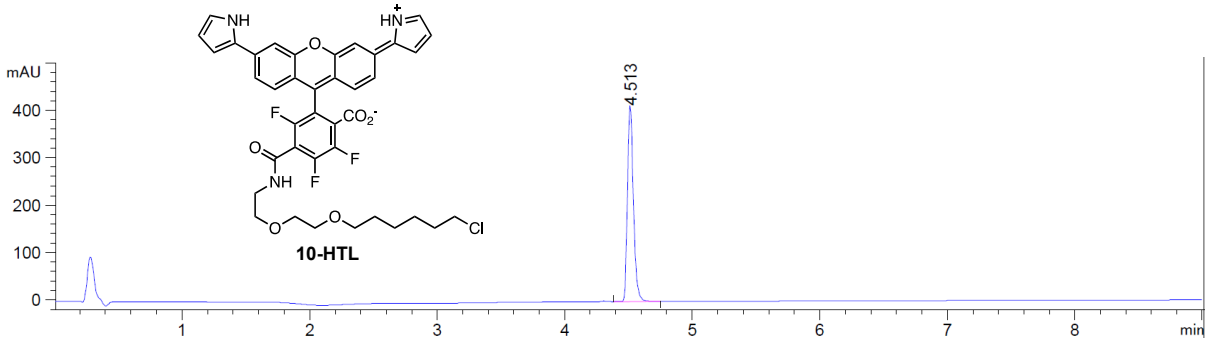

mAU
500
400
300
200
100
0
4.695
14-HTL
1
2
3
4
5
6
7
8
min

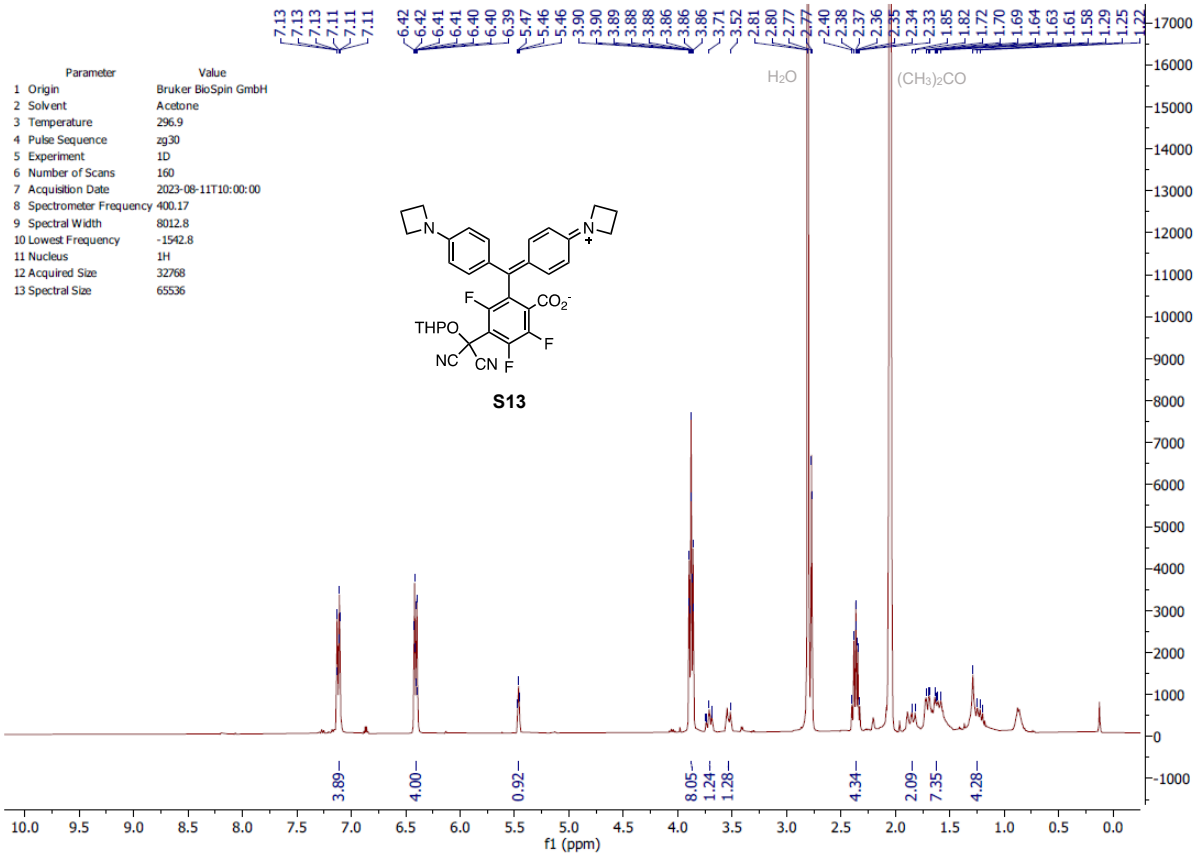

7.26 CDCl3
5.38 5.37 5.02 5.01 5.01 3.82 3.81 3.79 3.78 3.77 3.76 3.72 3.72 3.71 3.71 3.71 3.70 3.70 3.69 3.69 3.68 3.67 3.67 1.84 1.83 1.82 1.81 1.80 1.80 1.79 1.79 1.78 1.78 1.77 1.77 1.76 1.74 1.73 1.71 1.70 1.69 1.69 1.68 1.68 1.67 1.66 1.64 1.63 1.63 1.62 1.61 1.61 1.60 1.60 1.60
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 293.2
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-12-13T11:07:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
O
O
NC CN
S11
1.00 1.01 1.06 1.09 7.12
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)
5500 5000 4500 4000 3500 3000 2500 2000 1500 1000 500 0

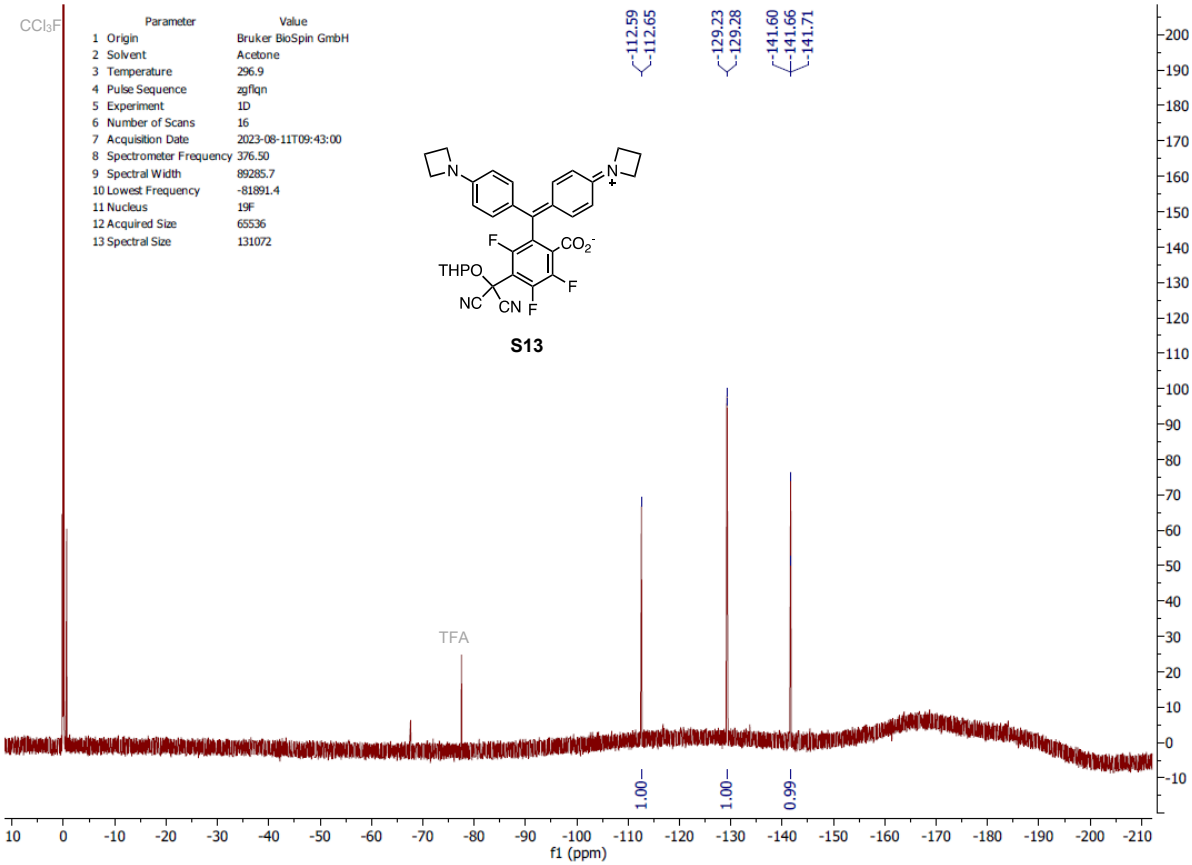

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.2
4 Pulse Sequence zgpg30
5 Experiment 1D
6 Number of Scans 1024
7 Acquisition Date 2022-12-13T20:02:00
8 Spectrometer Frequency 100.63
9 Spectral Width 24038.5
10 Lowest Frequency -1945.8
11 Nucleus 13C
12 Acquired Size 32768
13 Spectral Size 65536
111.01 110.97 98.58 CDCl3 62.69 51.82 29.09 24.72 17.83
O
O
NC CN
S11
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)
6500 6000 5500 5000 4500 4000 3500 3000 2500 2000 1500 1000 500 0 -500

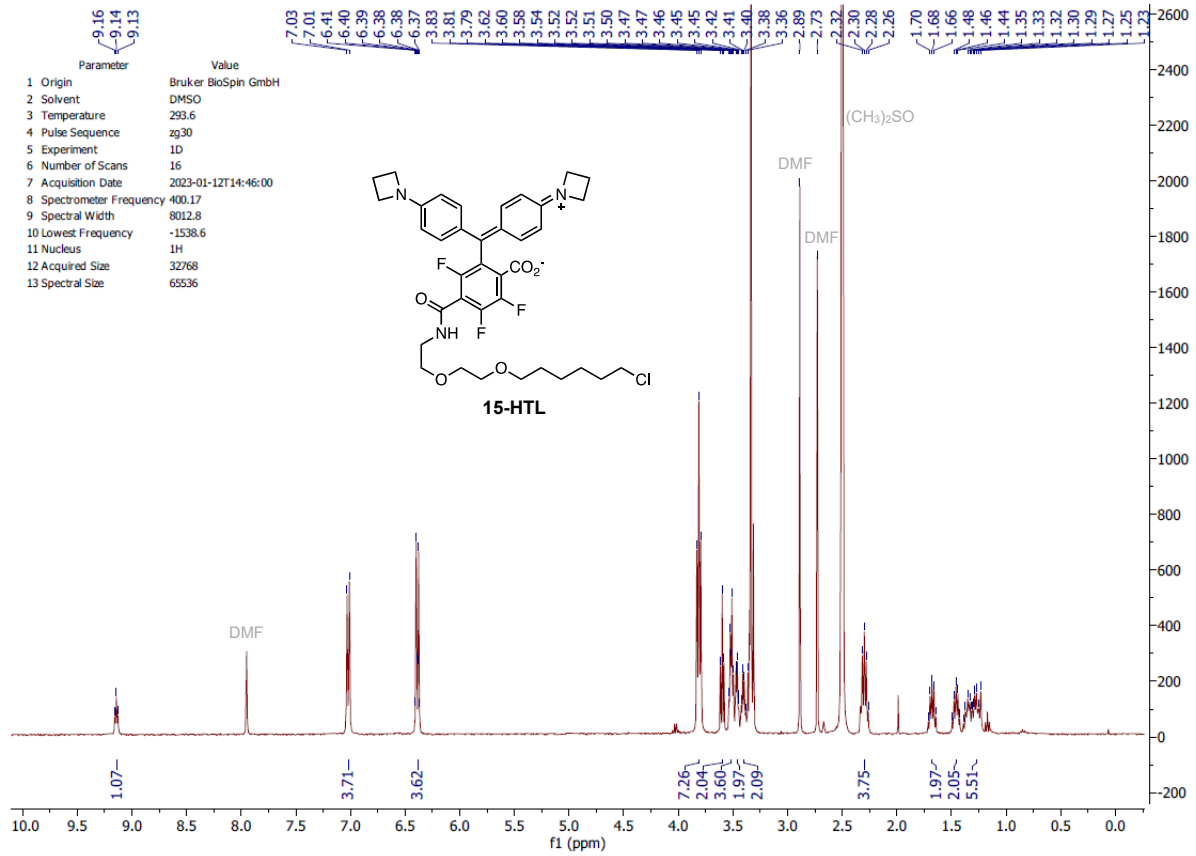

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 293.3
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-01-23T11:41:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1547.0
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
CD3CN
10-HTL
f1 (ppm)

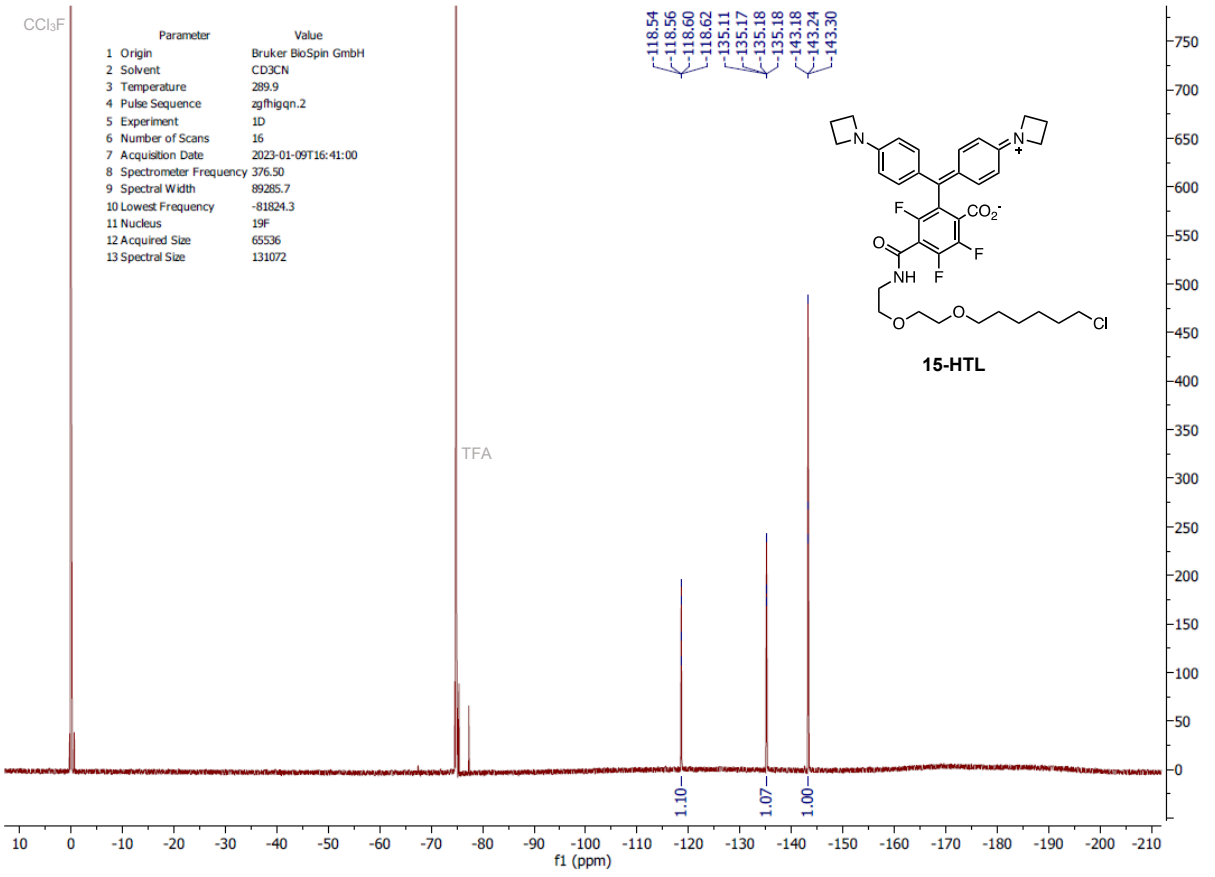

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 293.7
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-01-11T15:13:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81830.0
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
10-HTL
f1 (ppm)

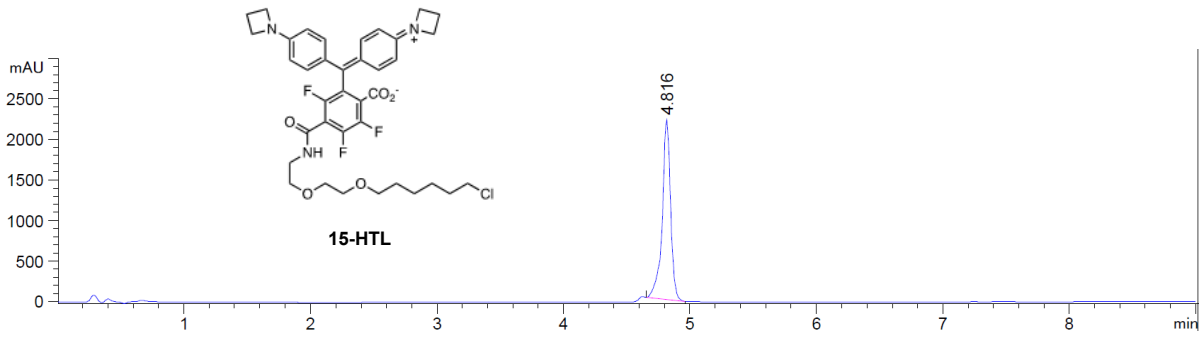

NH
HN
O
F
CO2
O
NH
F
F
O
O
Cl
10-HTL
mAU
400
300
200
100
0
4.513
1
2
3
4
5
6
7
8
min


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 296.9
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 160
7 Acquisition Date 2023-08-11T10:00:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.8
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
(CH3)2CO
S13
f1 (ppm)

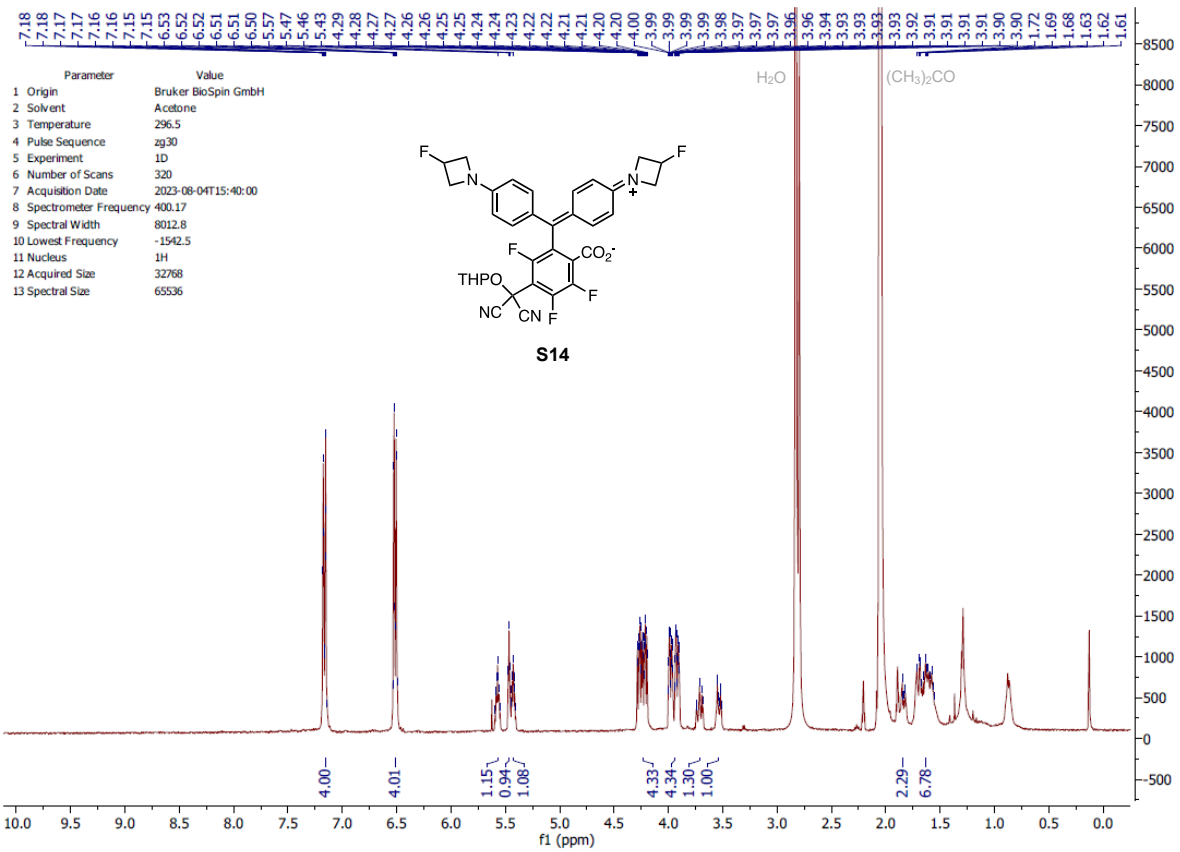

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 296.9
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-11T09:43:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81891.4
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
S13
f1 (ppm)


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 293.6
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-01-12T14:46:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.6
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
DMF
DMF
DMF
(CH3)2SO
15-HTL
f1 (ppm)

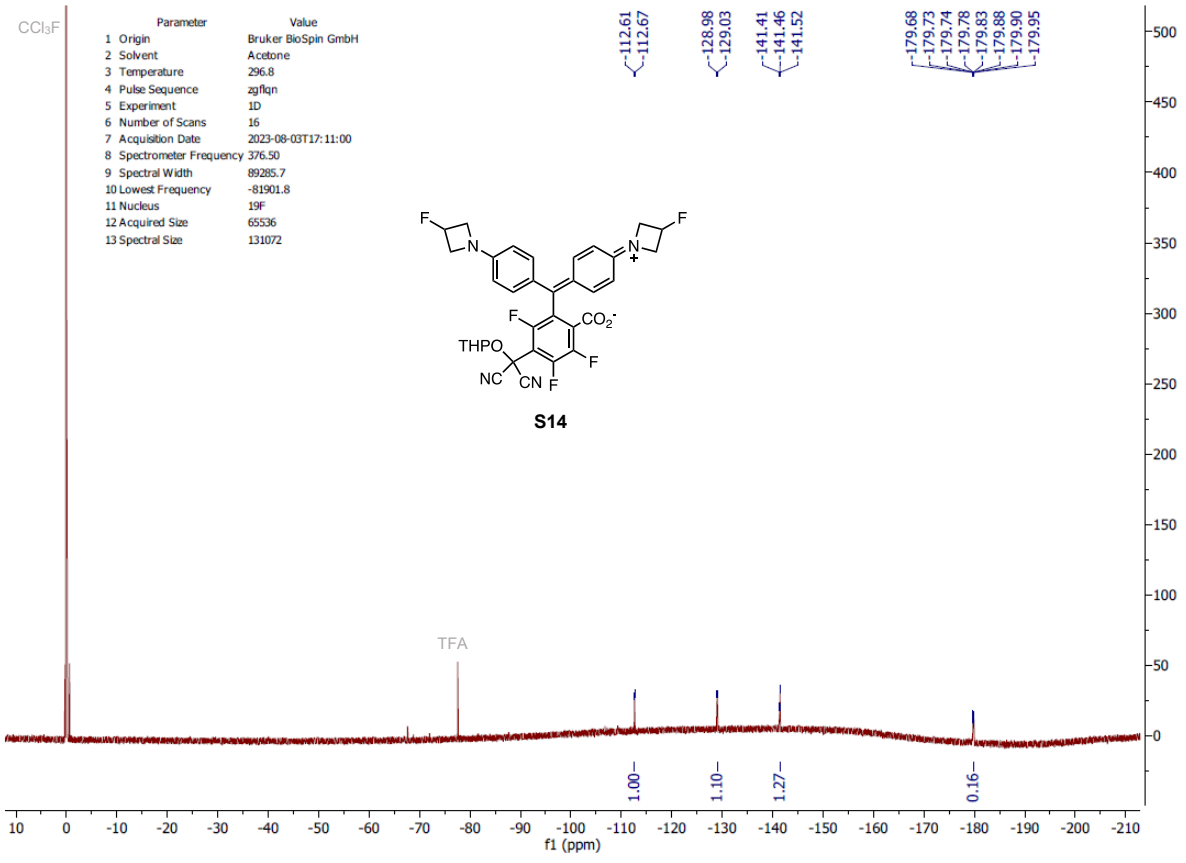

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 289.9
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-01-09T16:41:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81824.3
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
15-HTL
f1 (ppm)


mAU
2500
2000
1500
1000
500
0
1
2
3
4
5
6
7
8
min
4.816
F
CO2-
O
NH
F
F
O
O
Cl

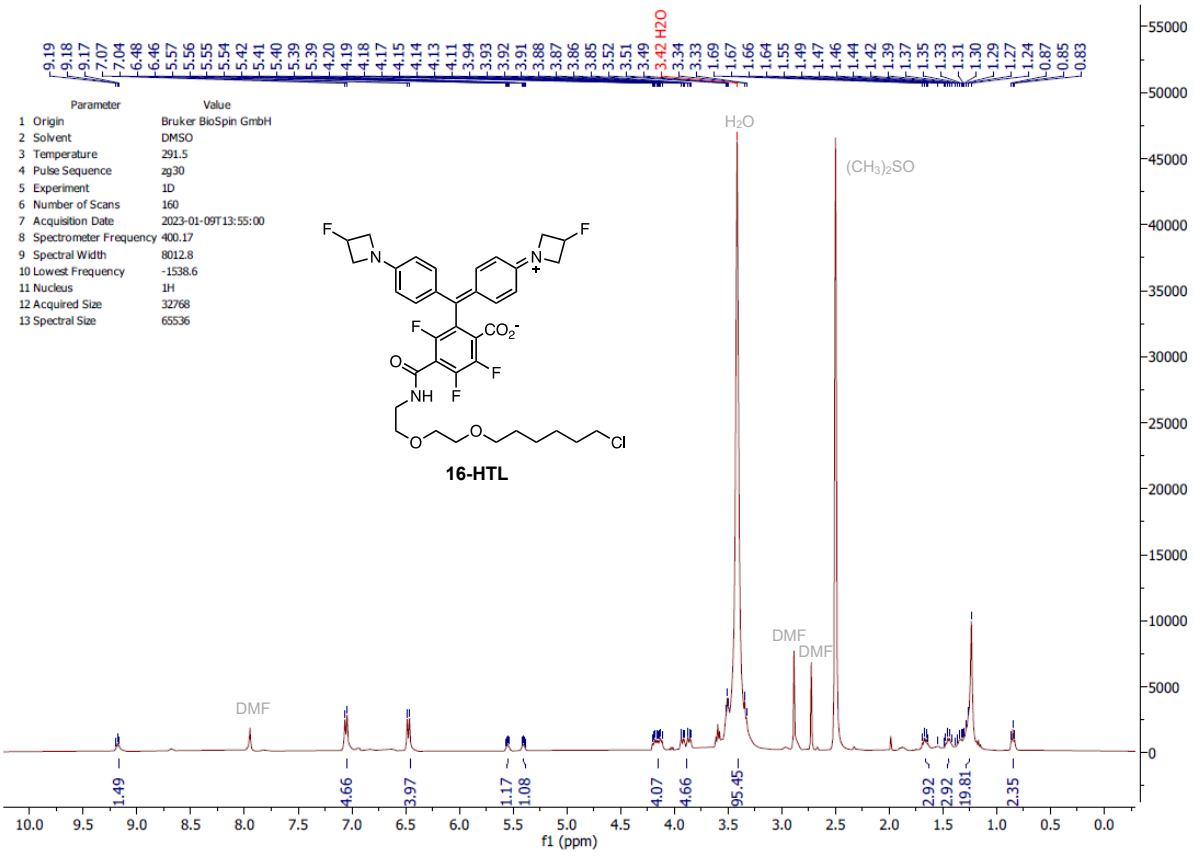
15-HTL

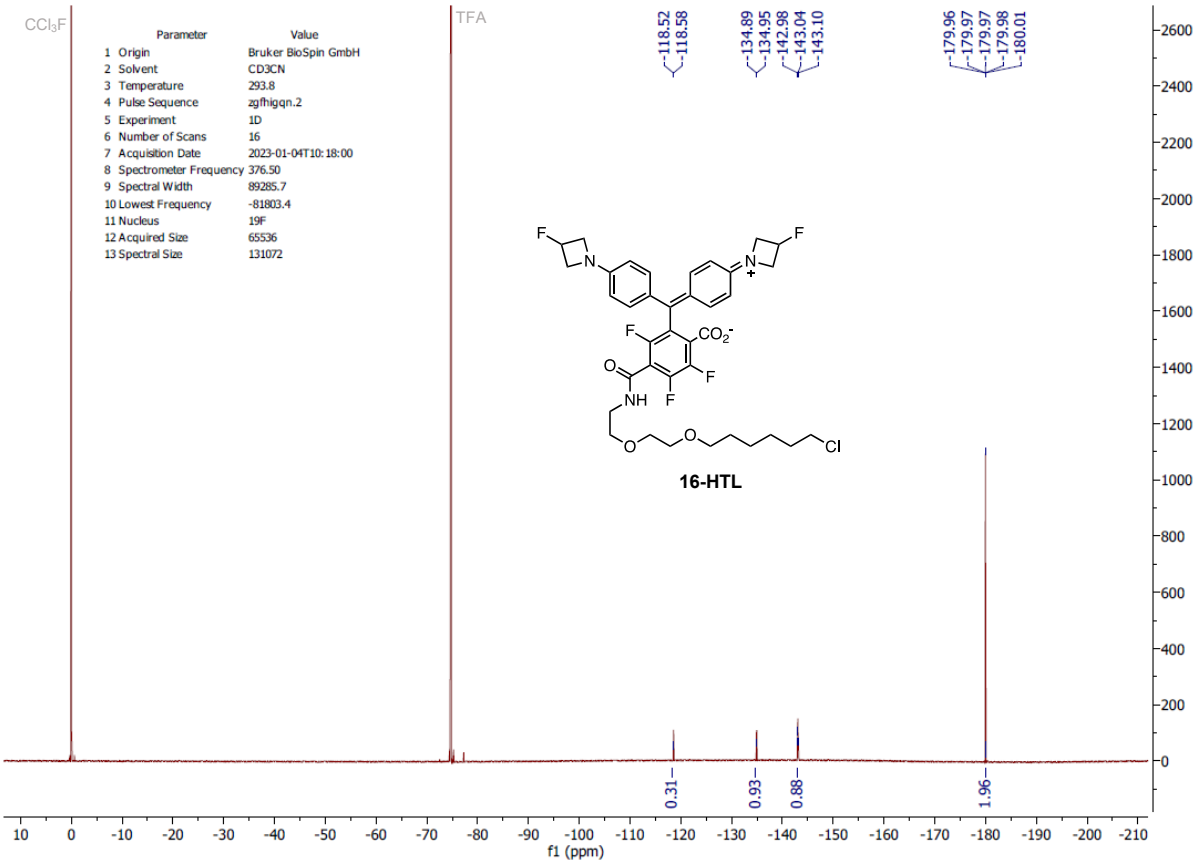

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 296.5
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 320
7 Acquisition Date 2023-08-04T15:40:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
(CH3)2CO
4.00
4.01
1.15
0.94
1.08
4.33
4.34
1.30
1.00
2.29
6.78
f1 (ppm)

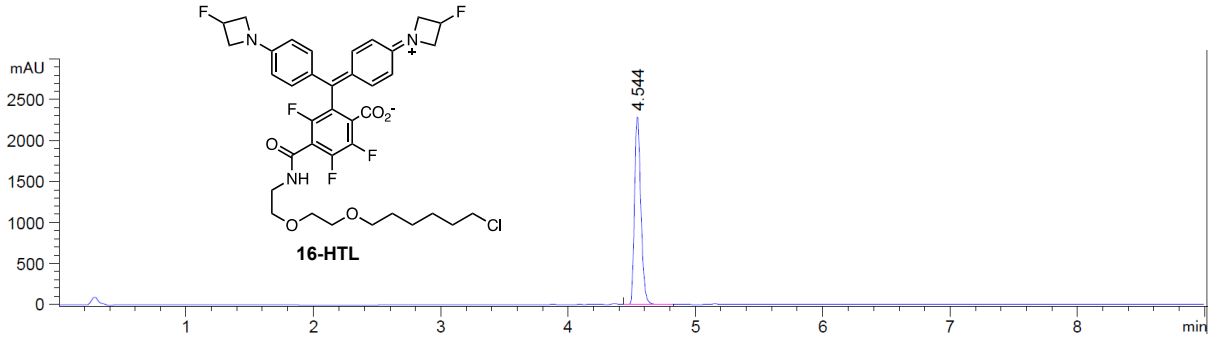
S14

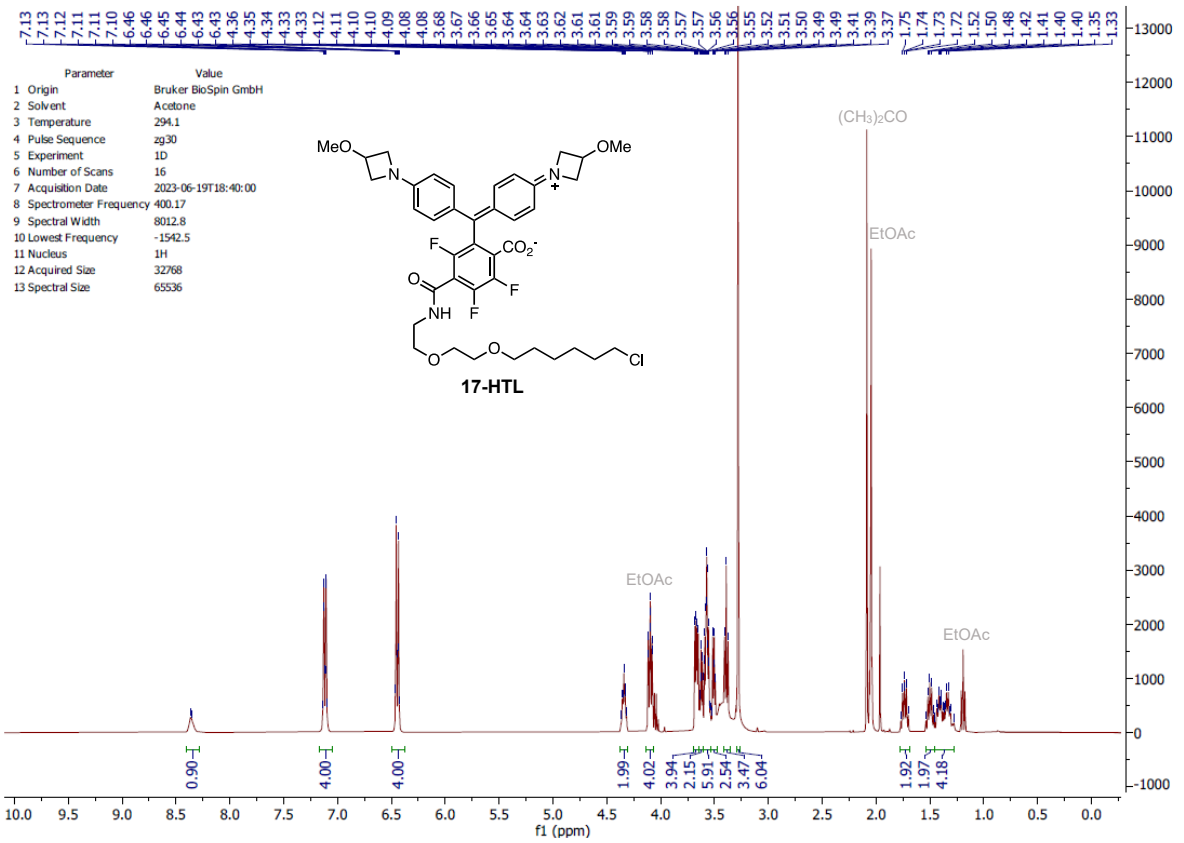

CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 296.8
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-03T17:11:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81901.8
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
1.00
1.10
1.27
0.16
f1 (ppm)

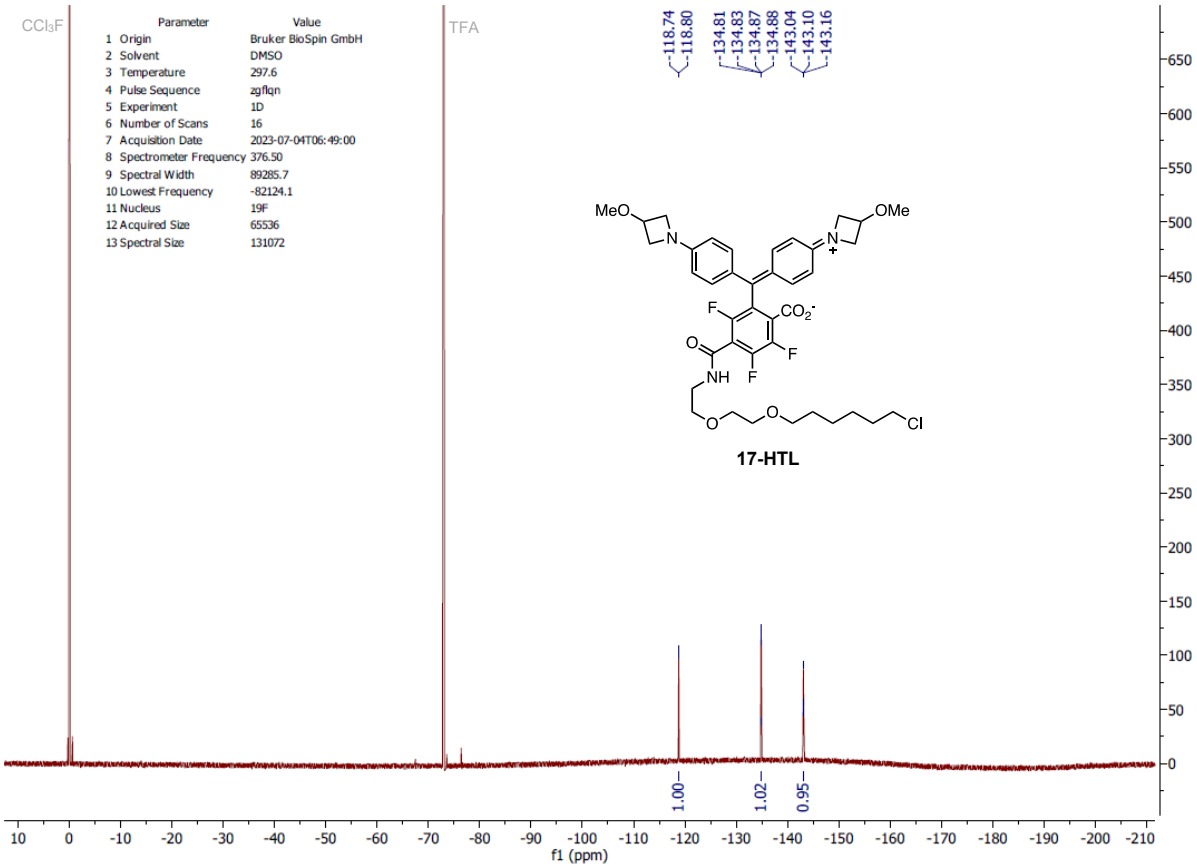
S14

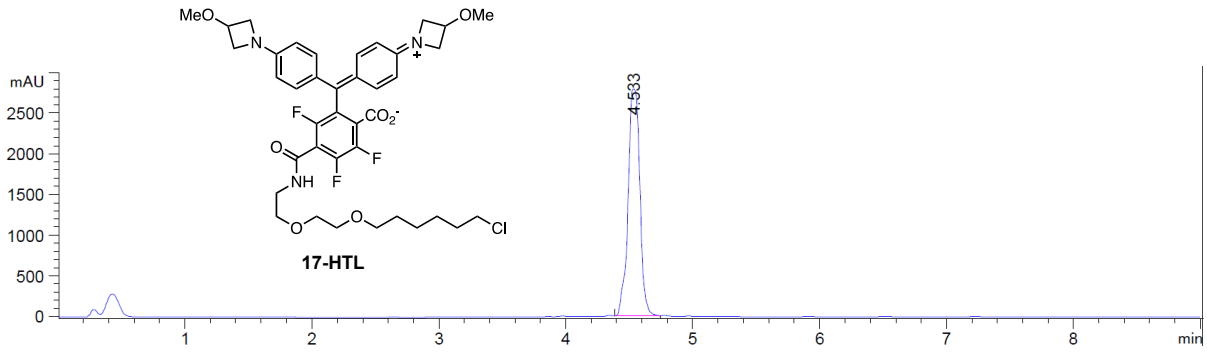

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 291.5
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 160
7 Acquisition Date 2023-01-09T13:55:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.6
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
(CH3)2SO
DMF
DMF
DMF
16-HTL
f1 (ppm)

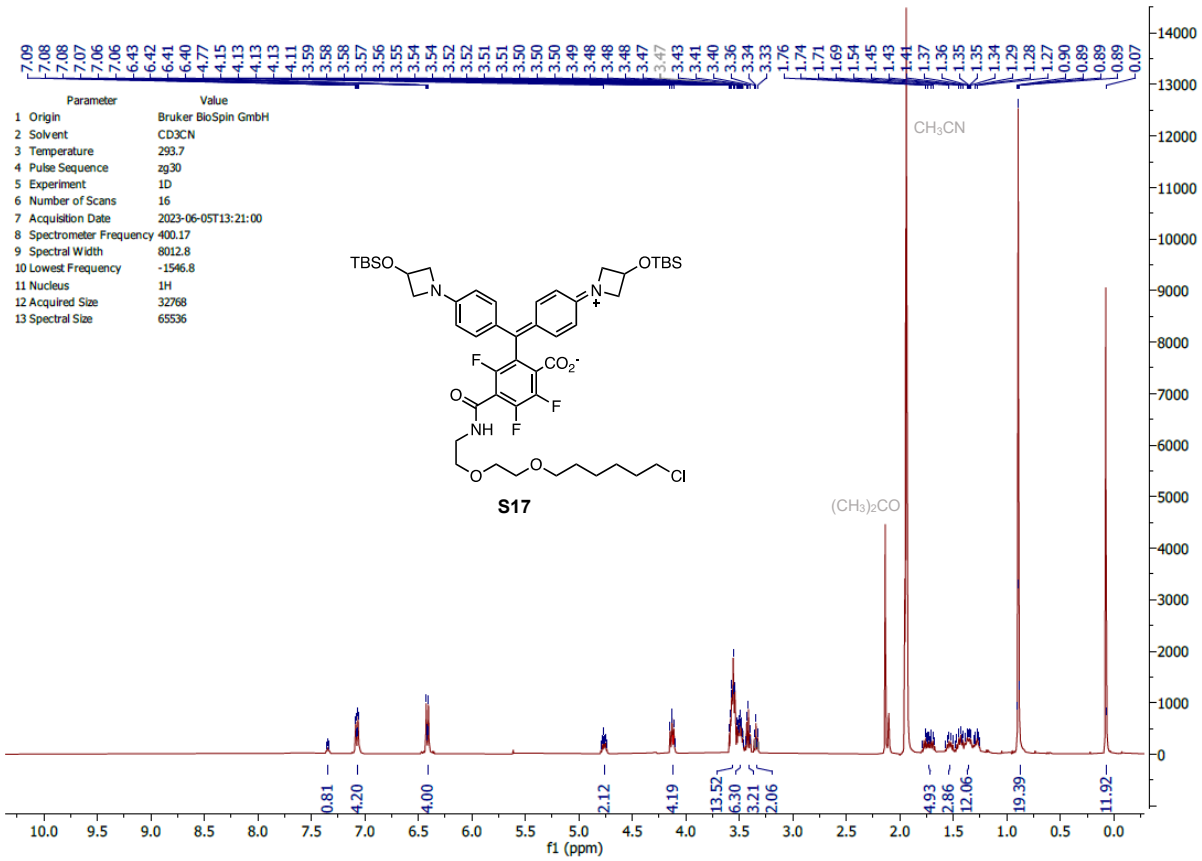

CCl3F
TFA
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 293.8
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-01-04T10:18:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81803.4
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
16-HTL
f1 (ppm)

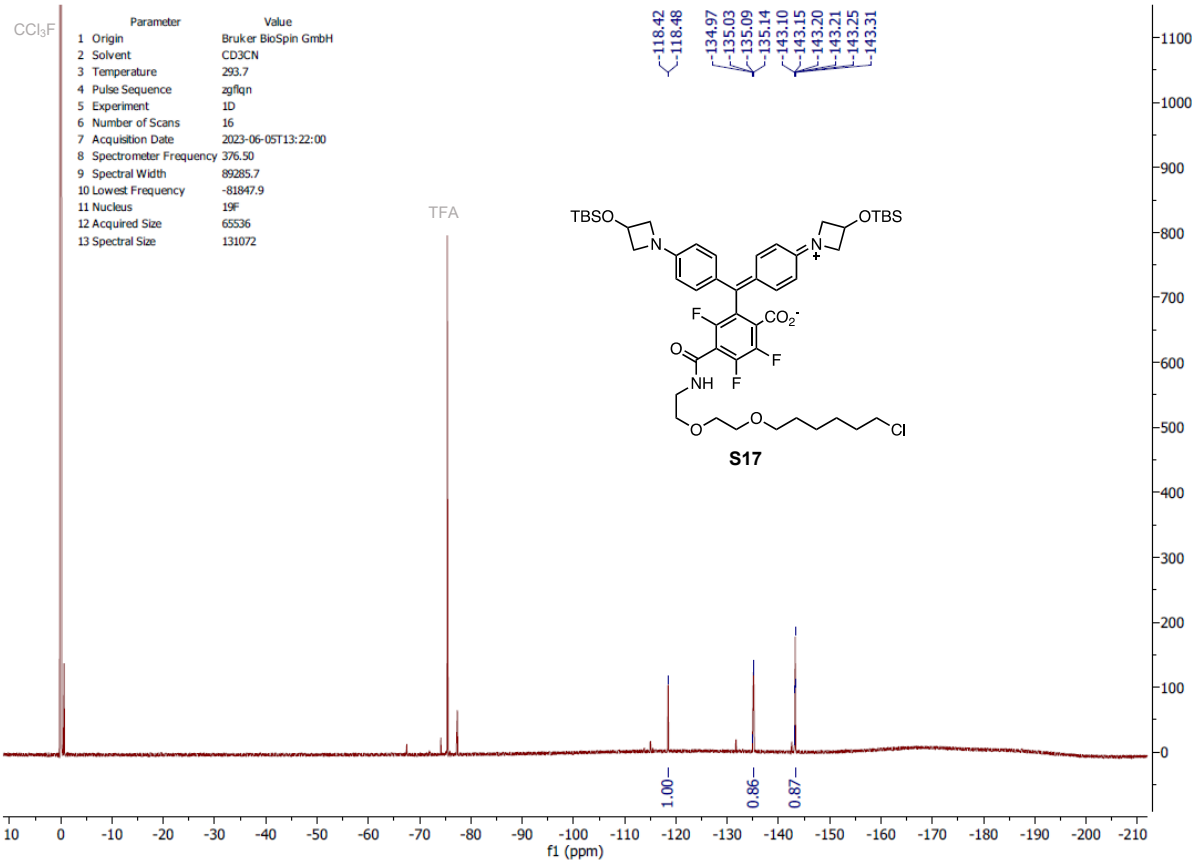

mAU
2500
2000
1500
1000
500
0
1
2
3
4
5
6
7
8
min
4.544
16-HTL

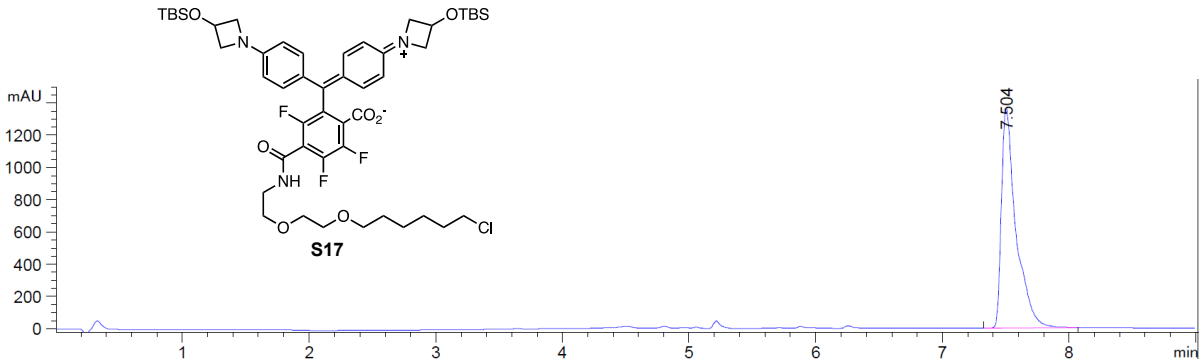

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 294.1
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-06-19T18:40:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
(CH3)2CO
EtOAc
EtOAc
EtOAc
17-HTL
f1 (ppm)

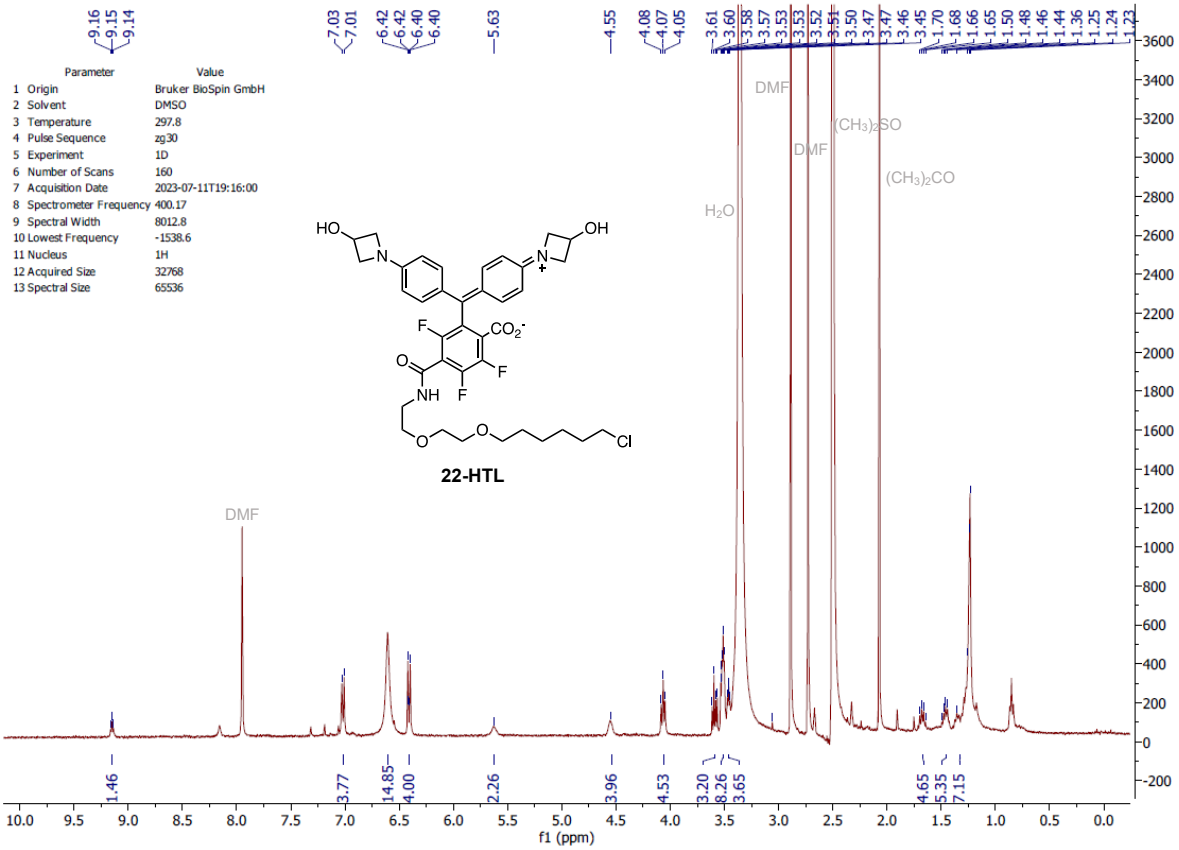

CCl3F
TFA
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 297.6
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-07-04T06:49:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82124.1
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
17-HTL
f1 (ppm)


MeO
OMe
F
CO2-
O
NH
F
F
O
O
Cl
17-HTL
4.533
mAU
2500
2000
1500
1000
500
0
1
2
3
4
5
6
7
8
min

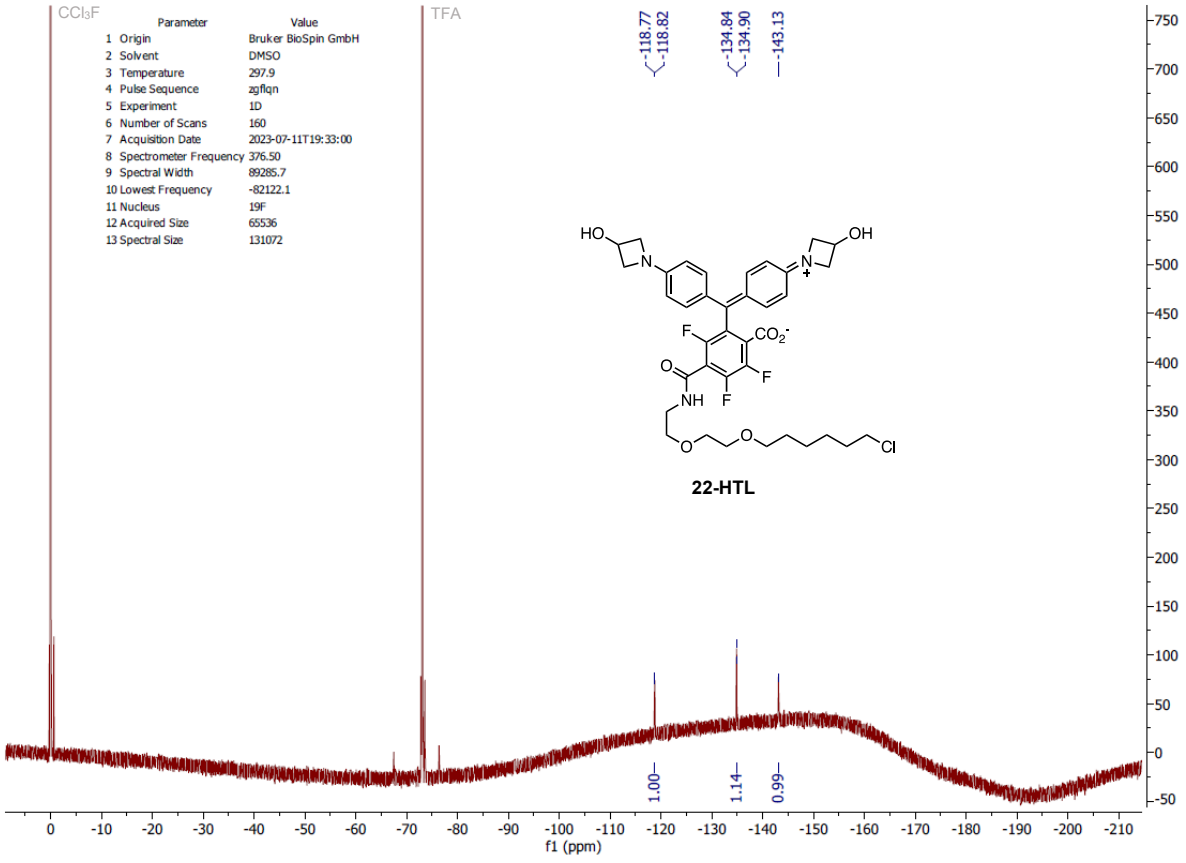

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 293.7
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-06-05T13:21:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1546.8
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
CH3CN
(CH3)2CO
TBSO
OTBS
S17
f1 (ppm)


CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 293.7
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-06-05T13:22:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81847.9
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
TBSO
OTBS
S17
f1 (ppm)

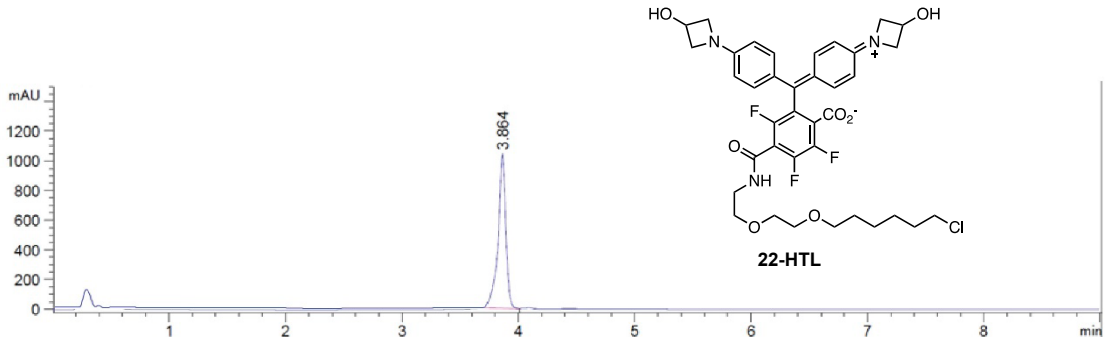

TBSO
OTBS
F
CO2-
O
NH
F
F
O
O
Cl
S17
mAU
1200
1000
800
600
400
200
0
7.504
1
2
3
4
5
6
7
8
min


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 297.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 160
7 Acquisition Date 2023-07-11T19:16:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.6
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
DMF
H2O
DMF
(CH3)2SO
DMF
(CH3)2CO
f1 (ppm)

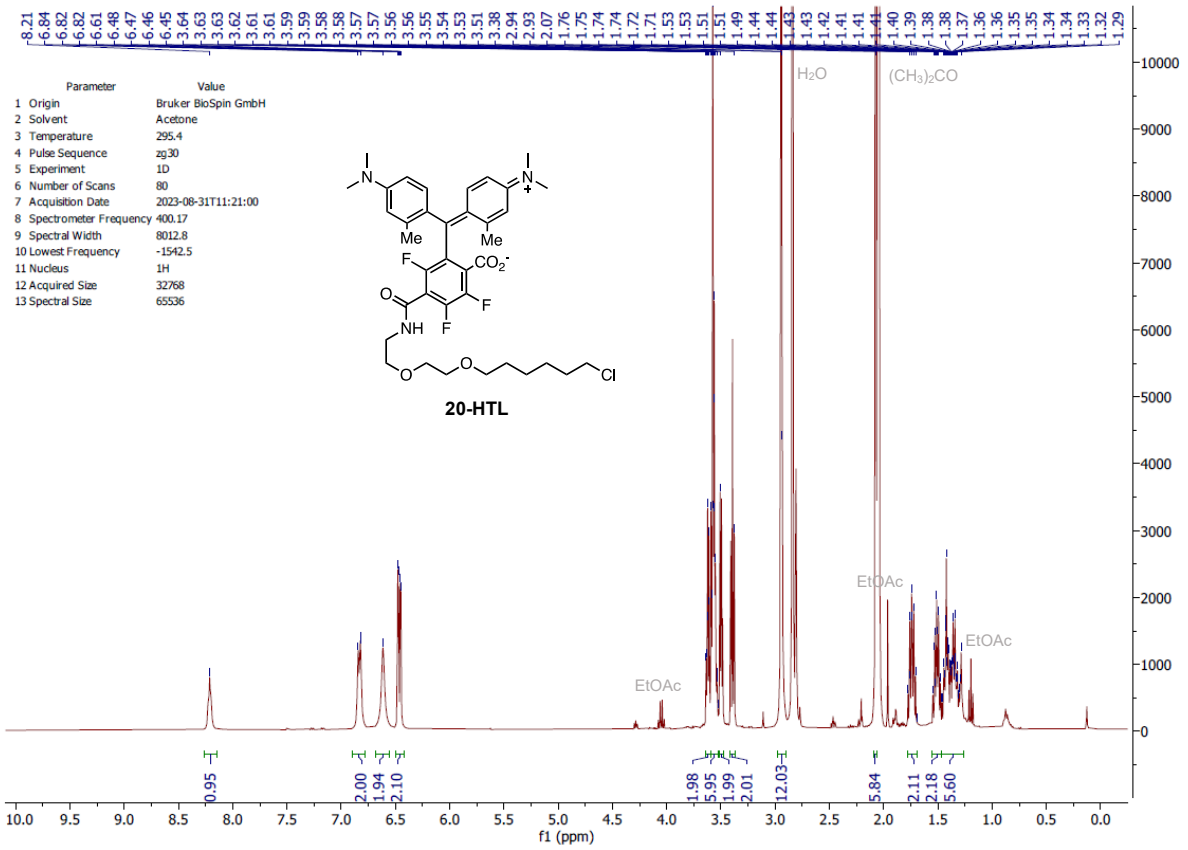
22-HTL

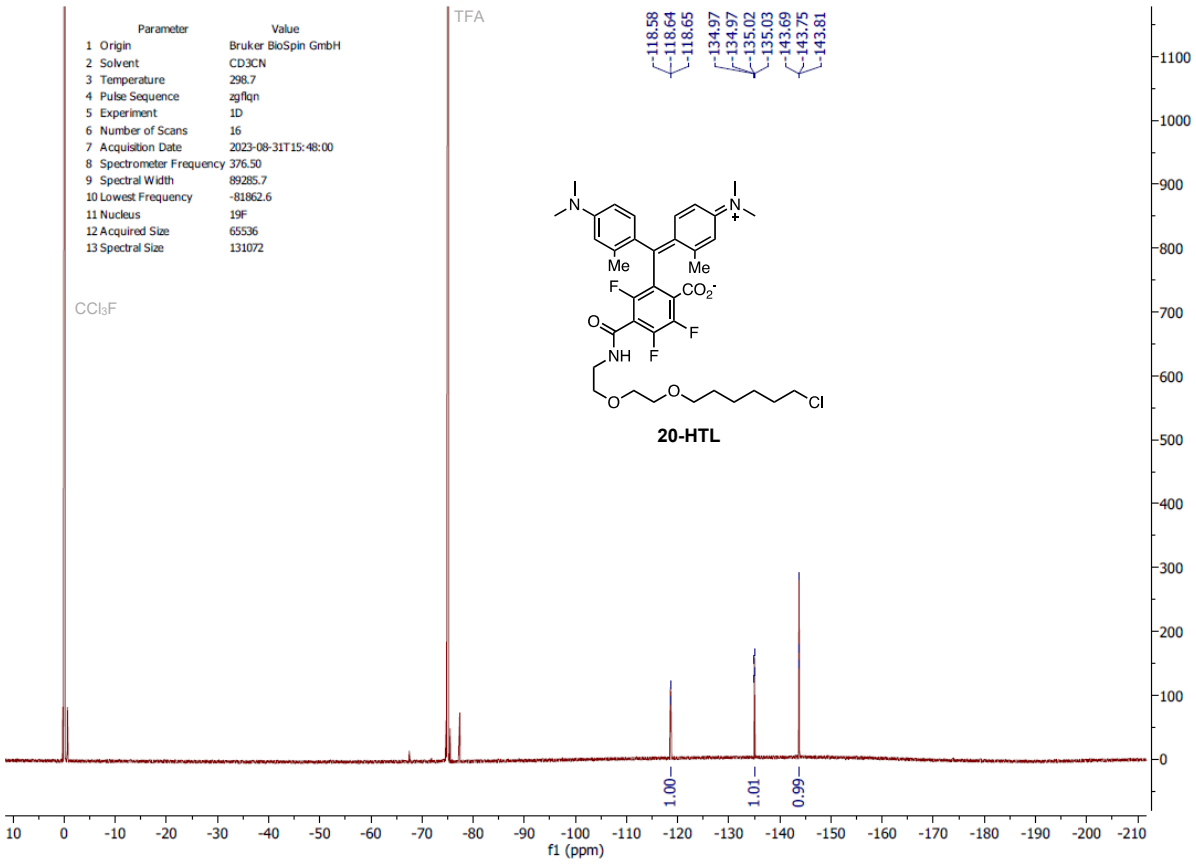

CCl3F
TFA
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 297.9
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 160
7 Acquisition Date 2023-07-11T19:33:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82122.1
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
f1 (ppm)

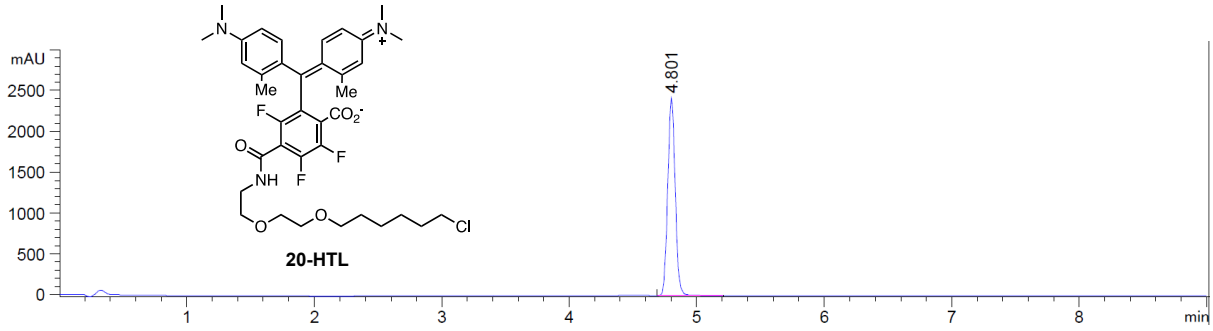
22-HTL

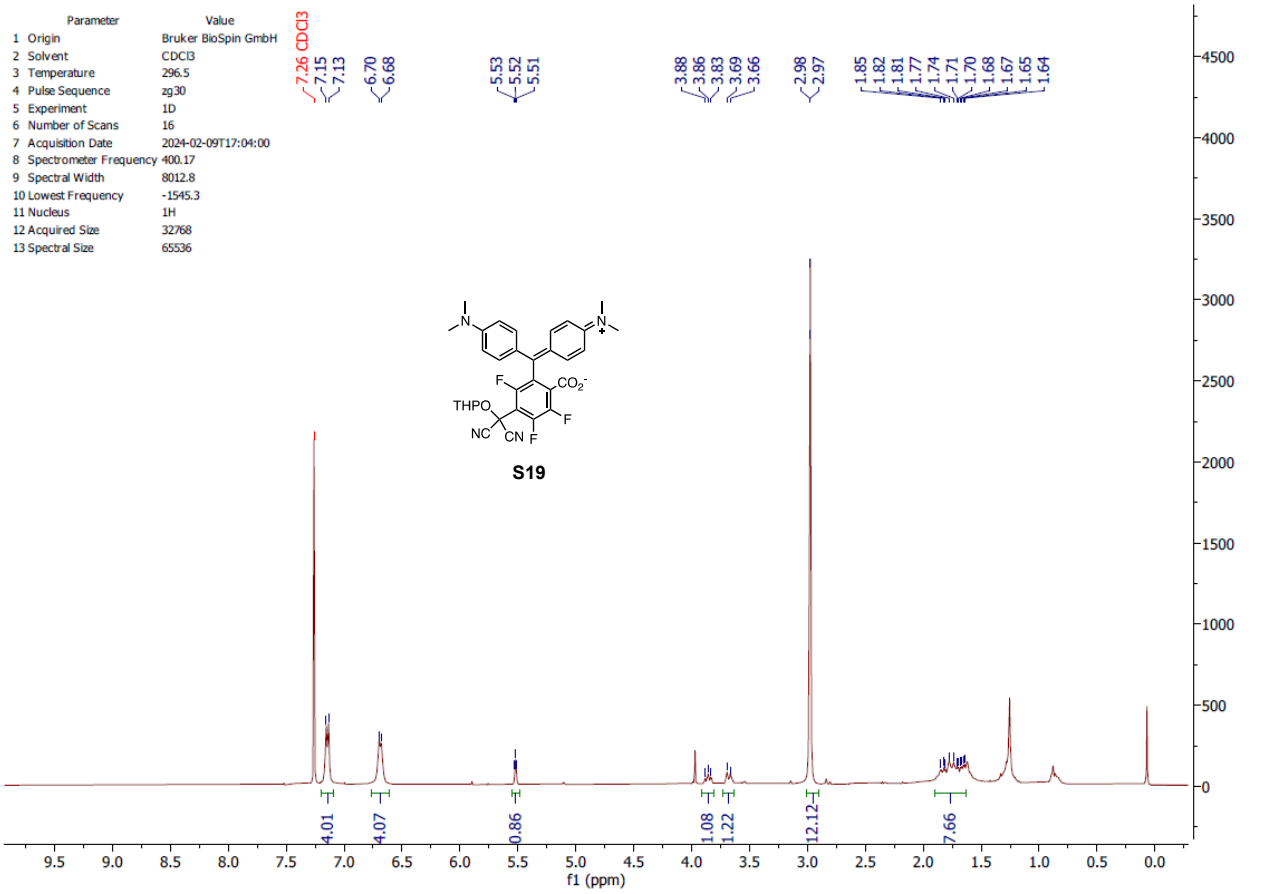

mAU
1200
1000
800
600
400
200
0
3.864
1
2
3
4
5
6
7
8
min
HO
OH
N
N+
F
CO2-
O
NH
F
F
O
O
Cl

22-HTL

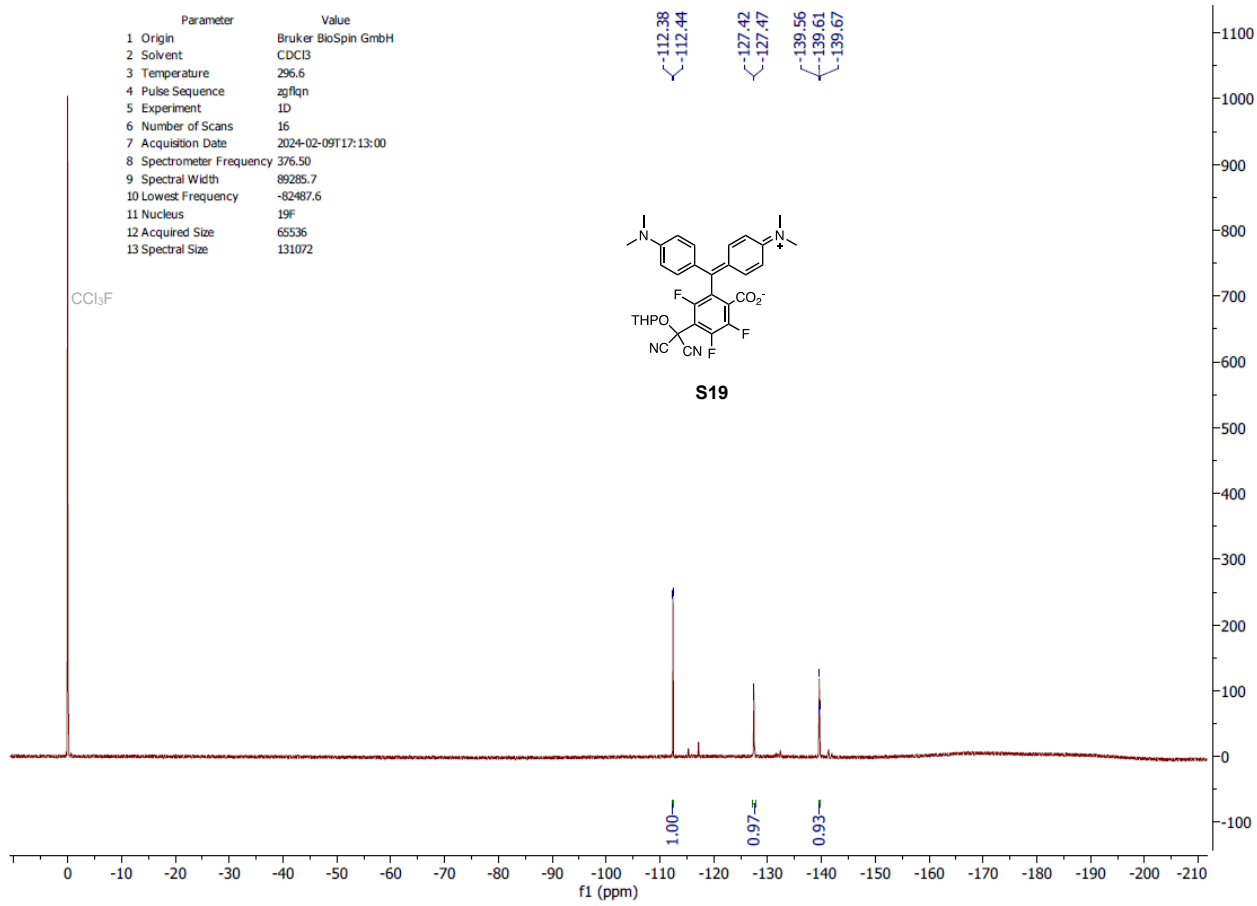

Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent Acetone
3 Temperature 295.4
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 80
7 Acquisition Date 2023-08-31T11:21:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1542.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
(CH3)2CO
EtOAc
EtOAc
EtOAc
20-HTL
f1 (ppm)


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CD3CN
3 Temperature 298.7
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2023-08-31T15:48:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -81862.6
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
TFA
20-HTL
f1 (ppm)

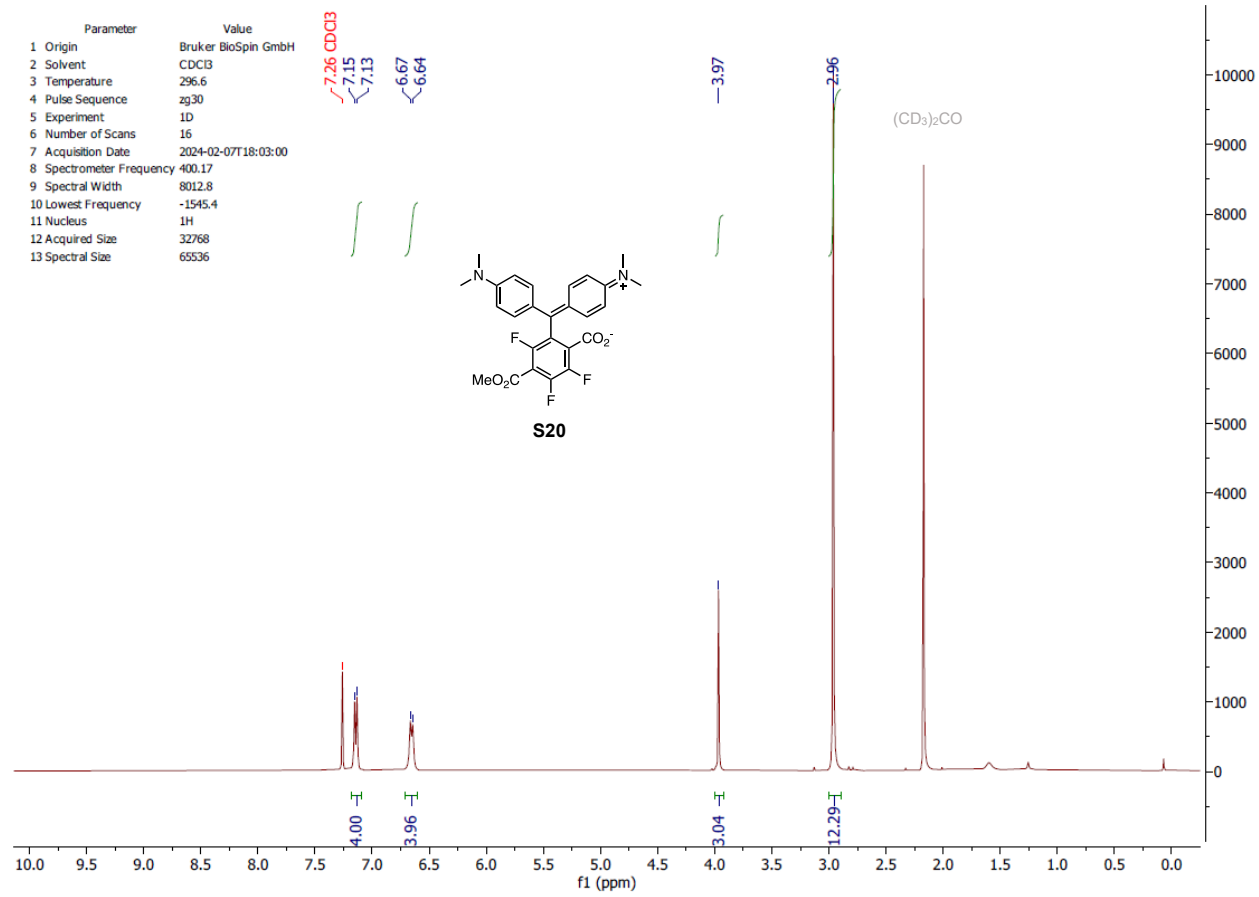

mAU
2500
2000
1500
1000
500
0
1
2
3
4
5
6
7
8
min
4.801
Me
Me
F
CO_2^-
F
O
NH
F
O
O
Cl
20-HTL


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 296.5
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2024-02-09T17:04:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.3
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
7.26 CDCl3
7.15
7.13
6.70
6.68
5.53
5.52
5.51
3.88
3.86
3.83
3.69
3.66
2.98
2.97
1.85
1.82
1.81
1.77
1.74
1.71
1.70
1.68
1.67
1.65
1.64
4.01
4.07
0.86
1.08
1.22
12.12
7.66
f1 (ppm)

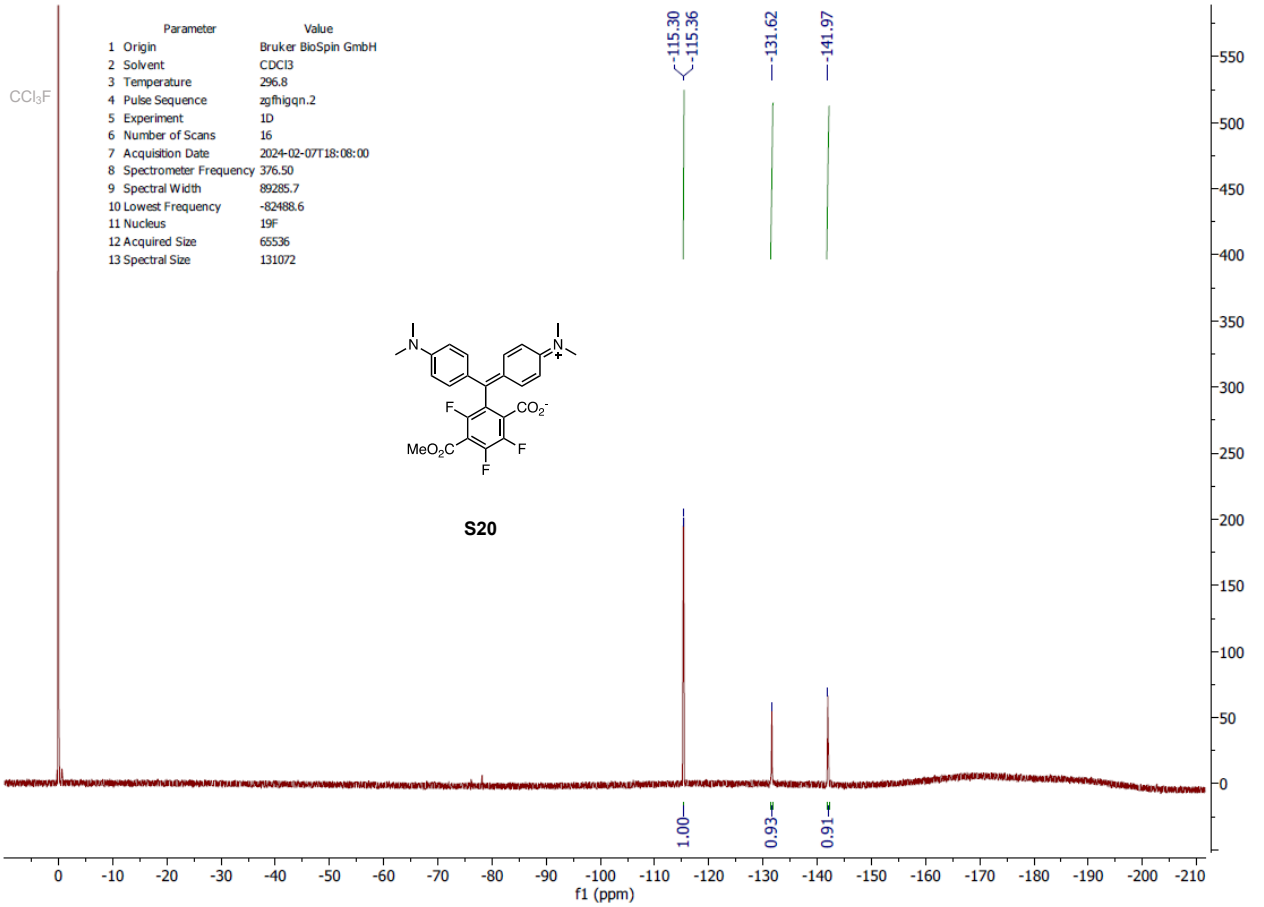
S19


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 296.6
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2024-02-09T17:13:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82487.6
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
-112.38
-112.44
-127.42
-127.47
-139.56
-139.61
-139.67
1.00
0.97
0.93
f1 (ppm)

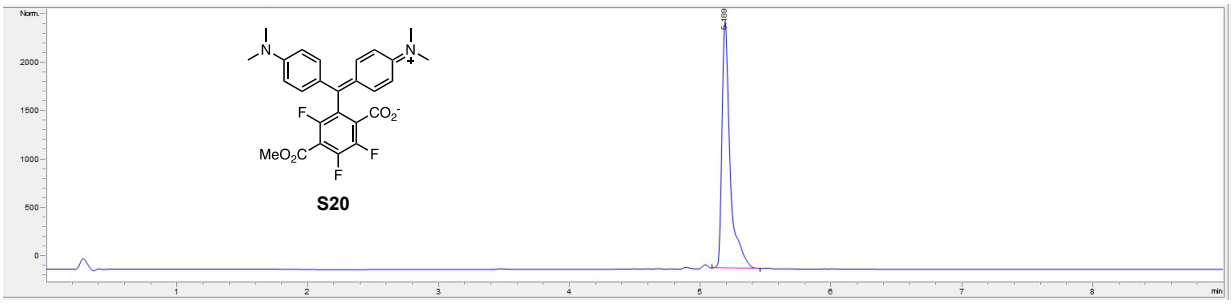
S19


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 296.6
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2024-02-07T18:03:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.4
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
7.26 CDCl3
7.15
7.13
6.67
6.64
3.97
2.96
(CD3)2CO
4.00
3.96
3.04
12.29
f1 (ppm)

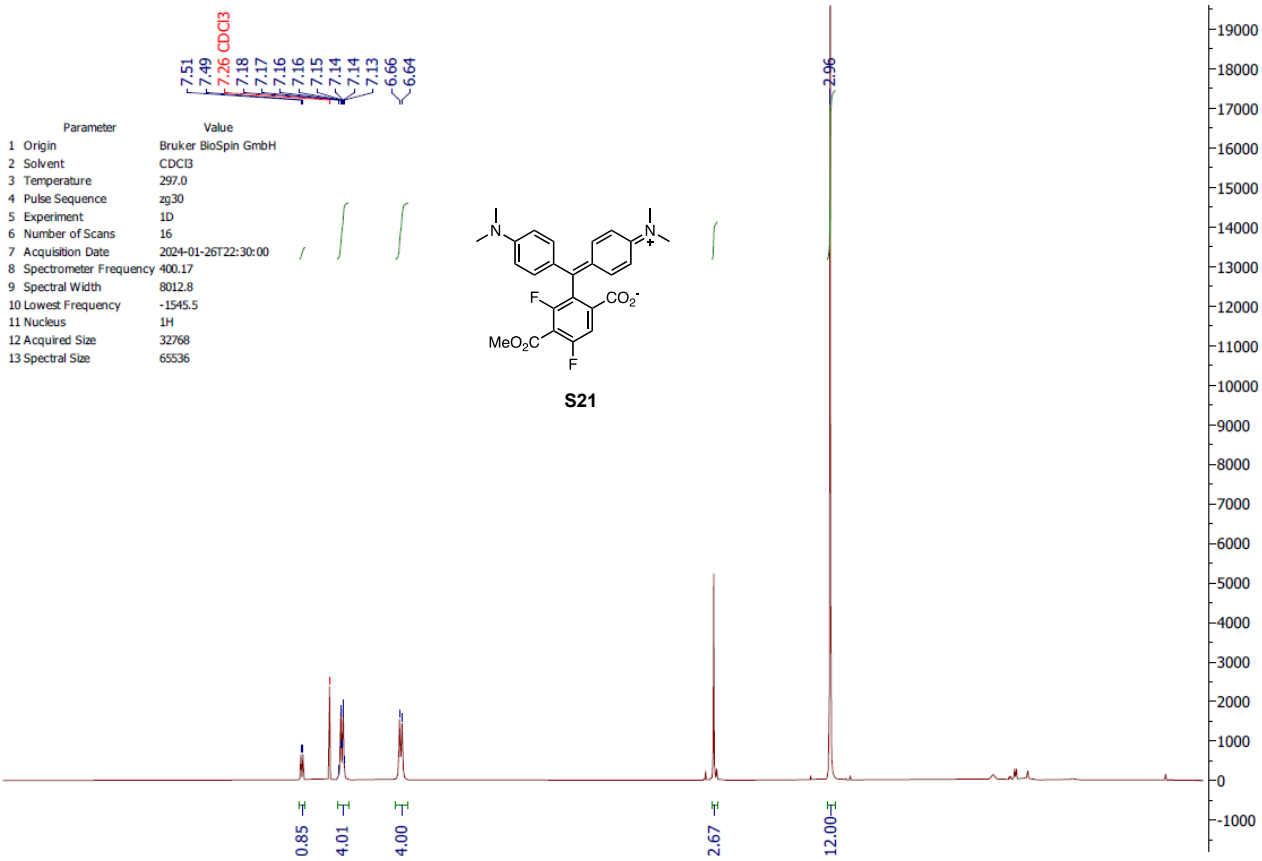
S20


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 296.8
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2024-02-07T18:08:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82488.6
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
-115.30
-115.36
-131.62
-141.97
1.00
0.93
0.91
f1 (ppm)

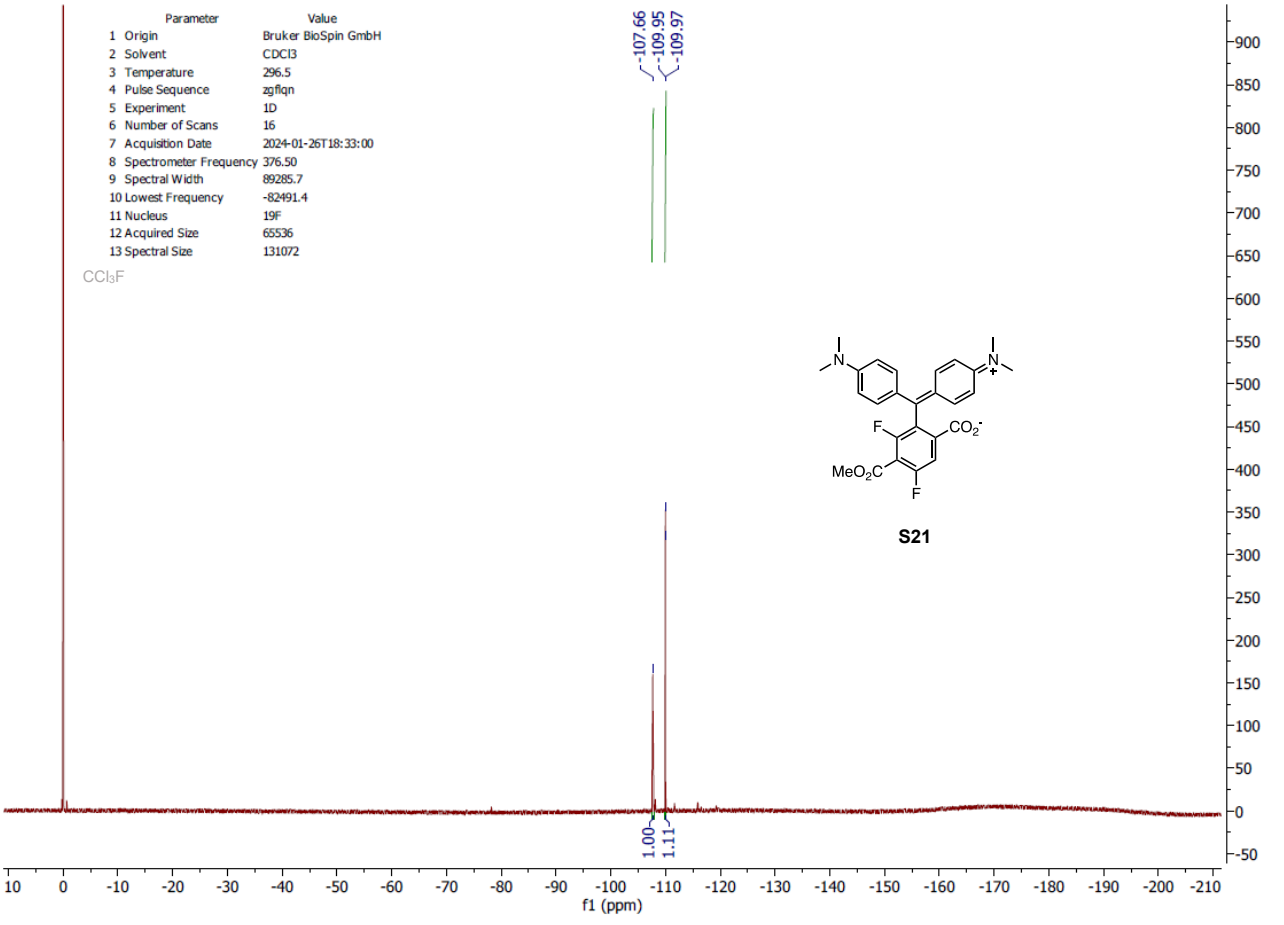
S20



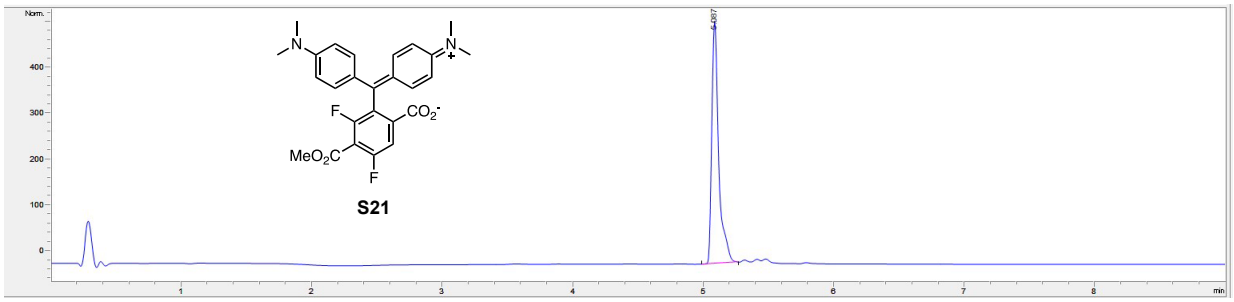
S20


7.51
7.49
7.26 CDCl3
7.18
7.17
7.16
7.16
7.15
7.14
7.14
7.13
6.66
6.64
2.96
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 297.0
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2024-01-26T22:30:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.5
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
0.85
4.01
4.00
2.67
12.00

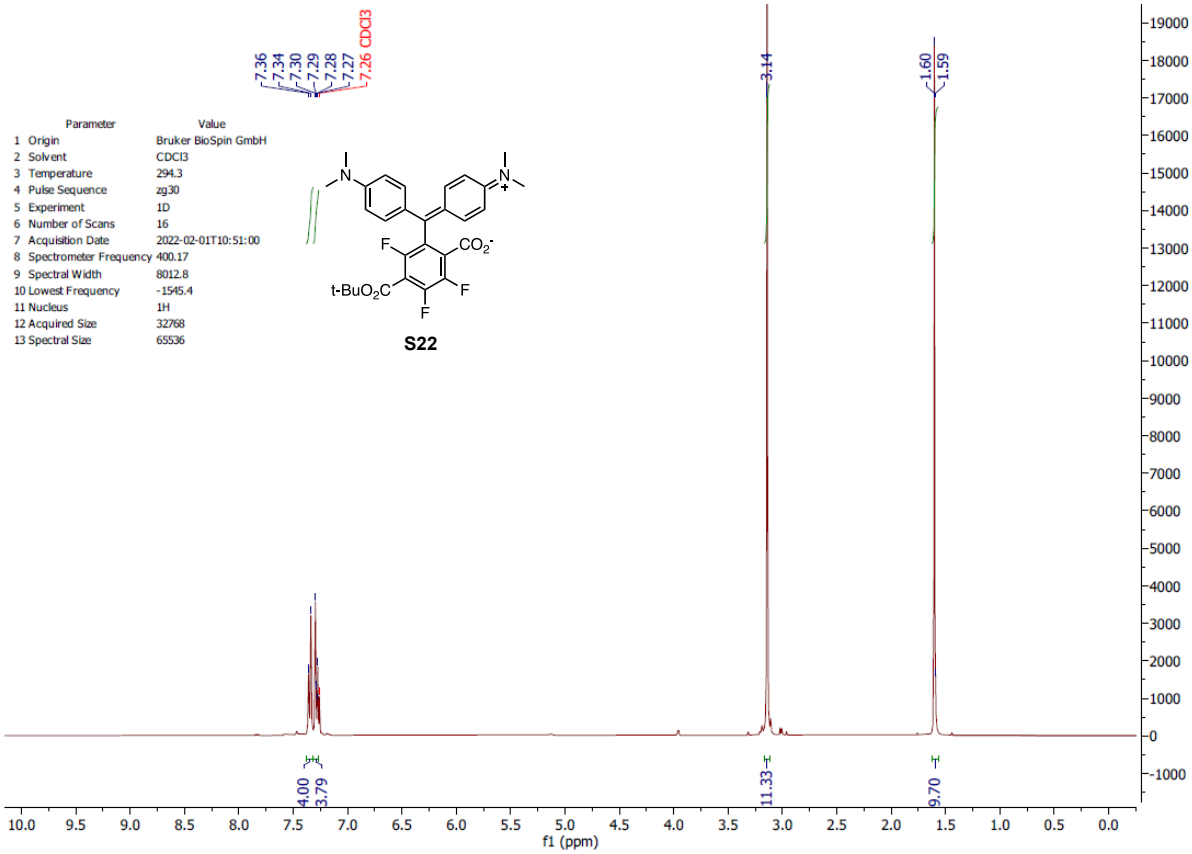
S21

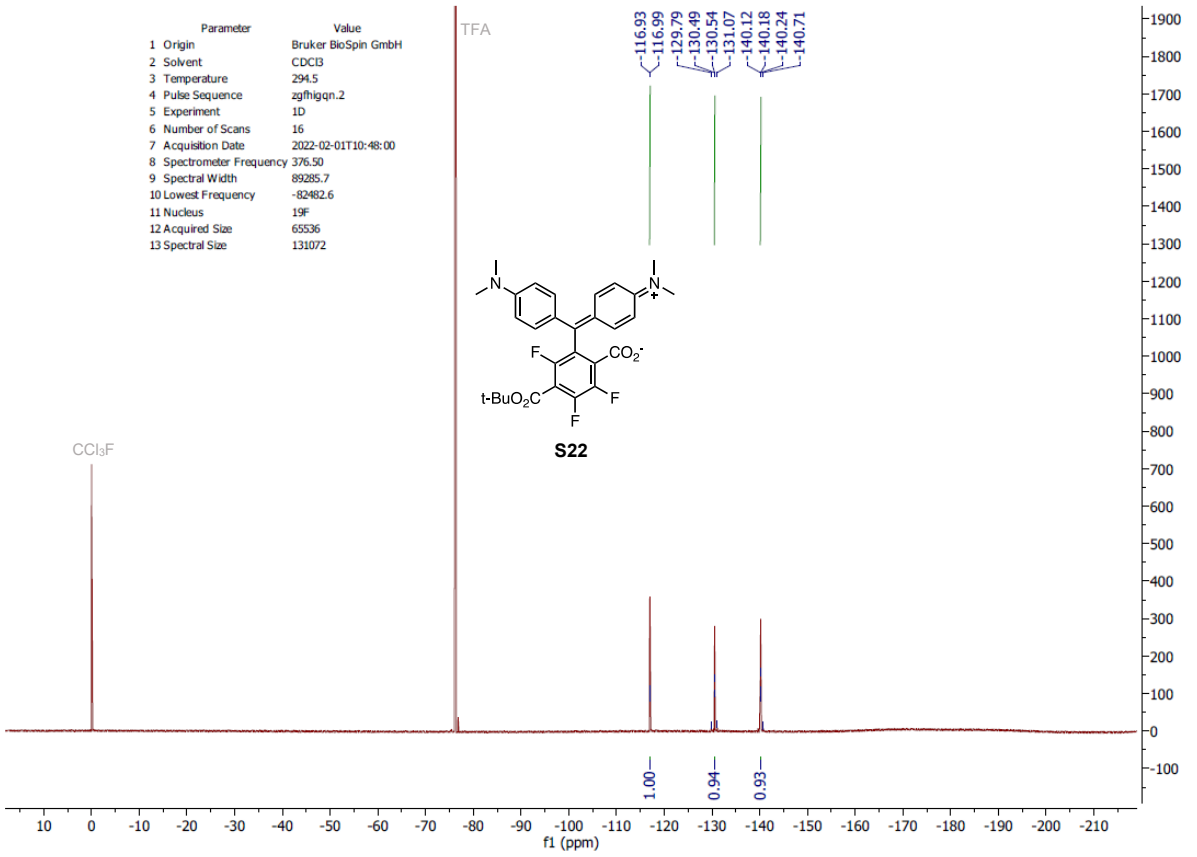

-107.66
-109.95
-109.97
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 296.5
4 Pulse Sequence zgflqn
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2024-01-26T18:33:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82491.4
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
CCl3F
1.00
1.11
f1 (ppm)

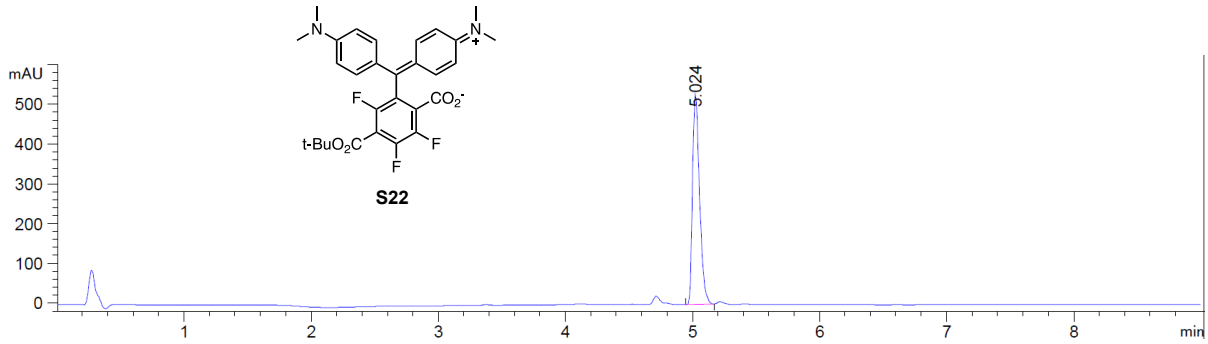
S21



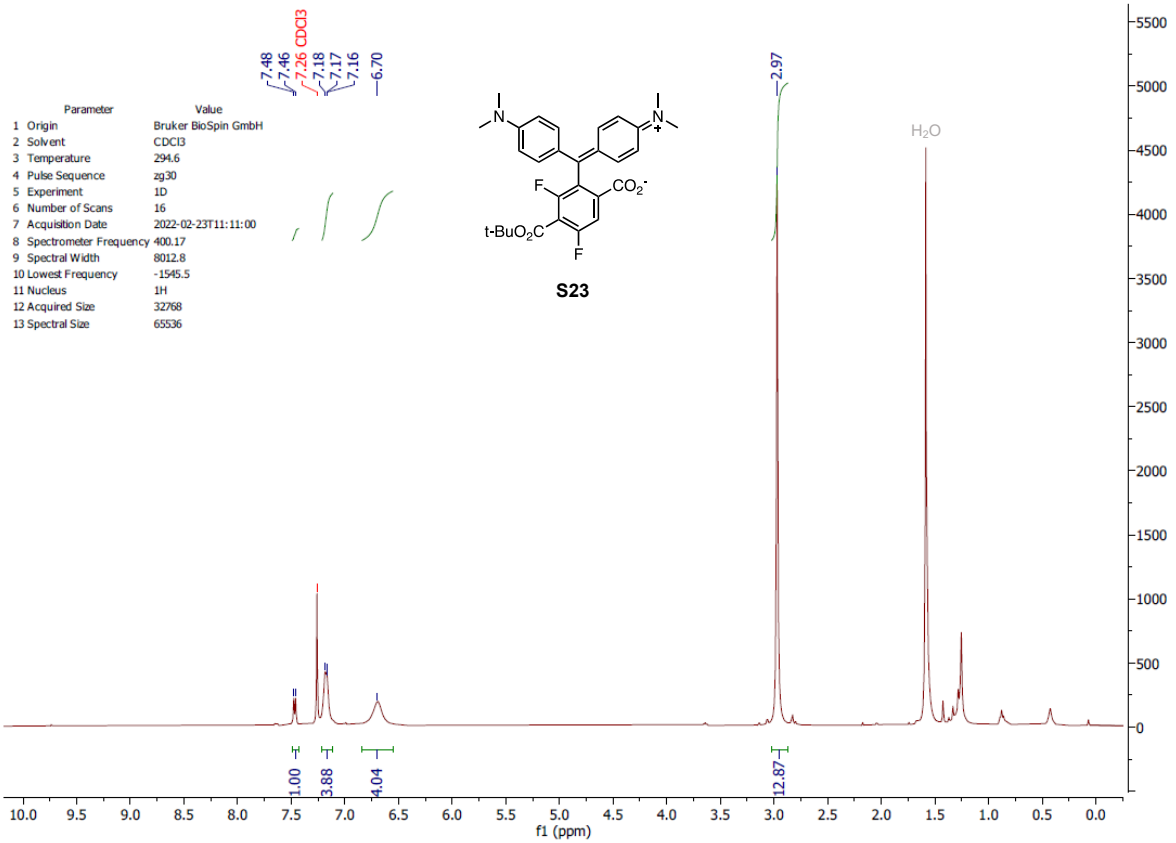
S21


7.36
7.34
7.30
7.29
7.28
7.27
7.26 CDCl3
3.14
1.60
1.59
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.3
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-02-01T10:51:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1545.4
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
S22
4.00
3.79
11.33
9.70
f1 (ppm)


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent CDCl3
3 Temperature 294.5
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-02-01T10:48:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82482.6
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
CCl3F
-116.93
-116.99
-129.79
-130.49
-130.54
-131.07
-140.12
-140.18
-140.24
-140.71
S22
1.00
0.94
0.93
f1 (ppm)

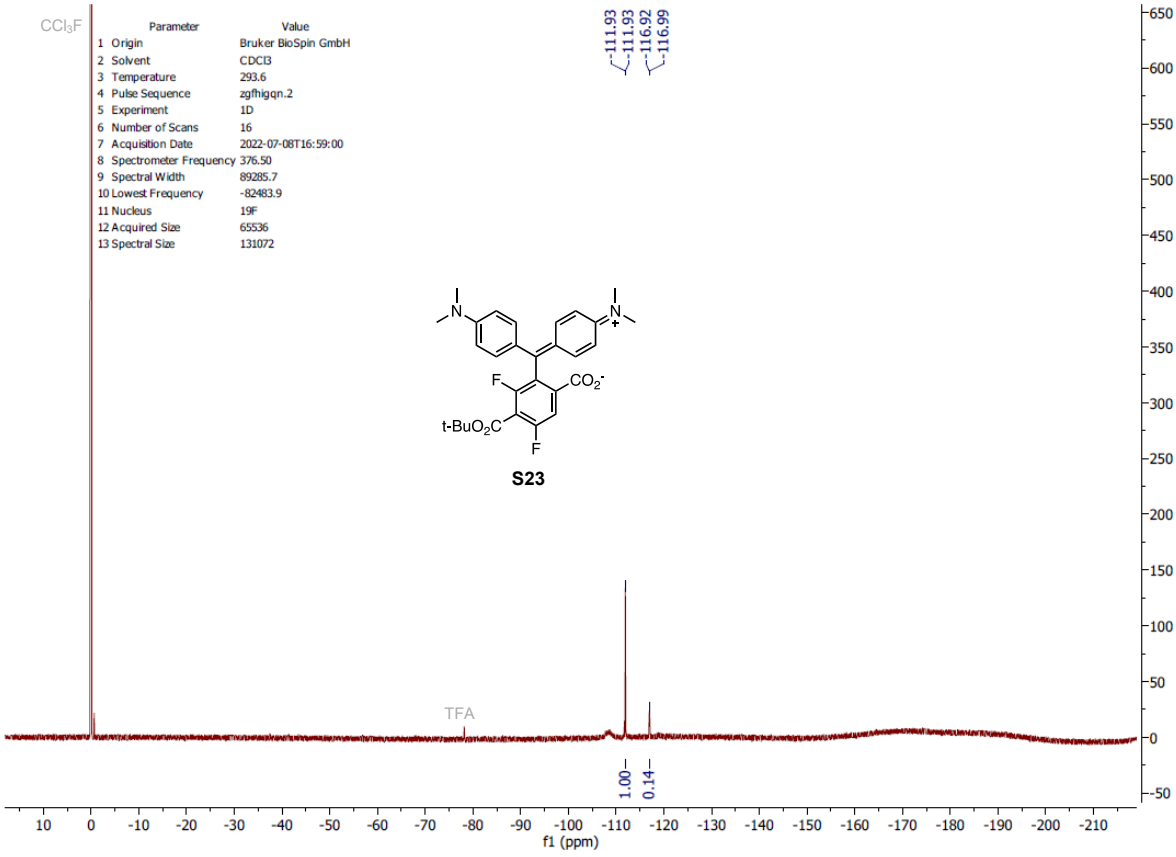


S22



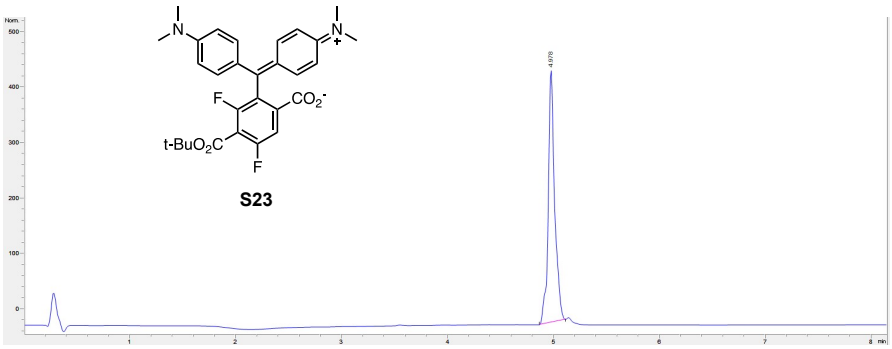
Parameter	Value
1 Origin	Bruker BioSpin GmbH
2 Solvent	CDCl3
3 Temperature	294.6
4 Pulse Sequence	zg30
5 Experiment	1D
6 Number of Scans	16
7 Acquisition Date	2022-02-23T11:11:00
8 Spectrometer Frequency	400.17
9 Spectral Width	8012.8
10 Lowest Frequency	-1545.5
11 Nucleus	1H
12 Acquired Size	32768
13 Spectral Size	65536

S23

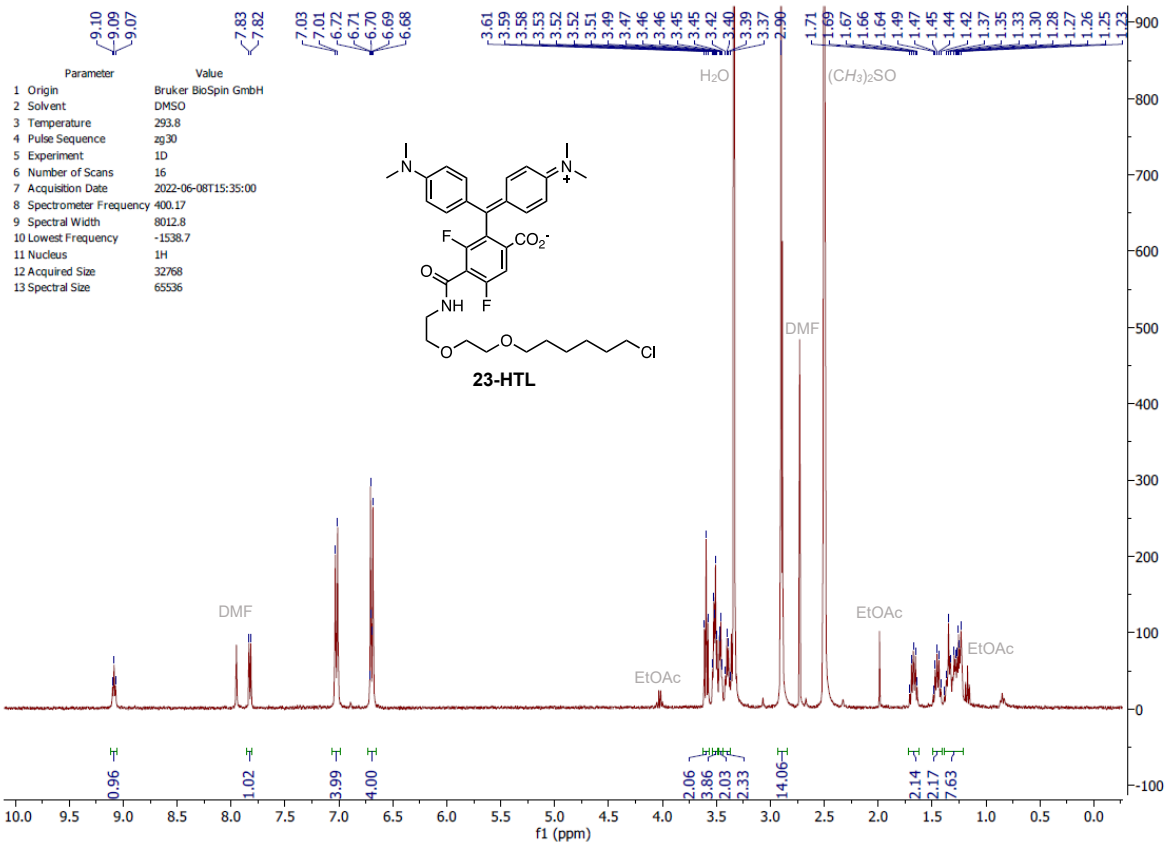


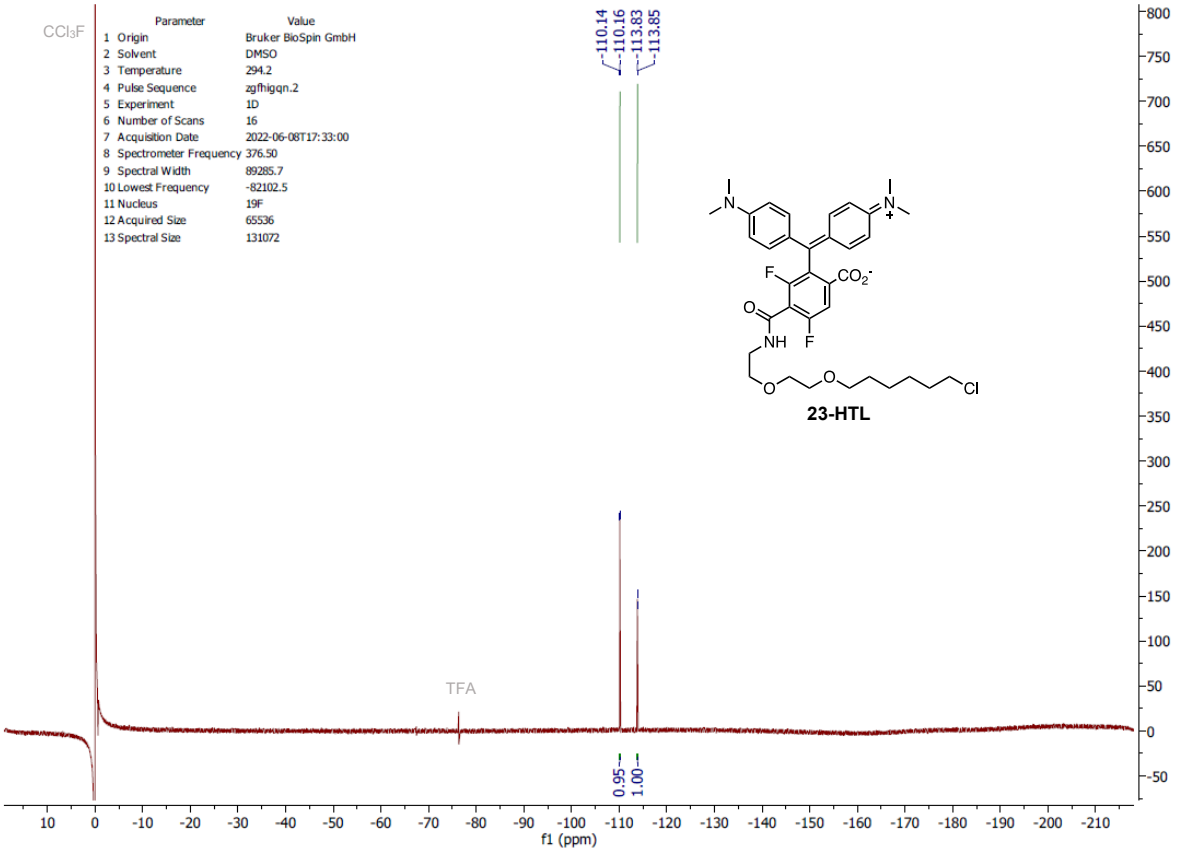
Parameter	Value
1 Origin	Bruker BioSpin GmbH
2 Solvent	CDCl3
3 Temperature	293.6
4 Pulse Sequence	zgfhigqn.2
5 Experiment	1D
6 Number of Scans	16
7 Acquisition Date	2022-07-08T16:59:00
8 Spectrometer Frequency	376.50
9 Spectral Width	89285.7
10 Lowest Frequency	-82483.9
11 Nucleus	19F
12 Acquired Size	65536
13 Spectral Size	131072

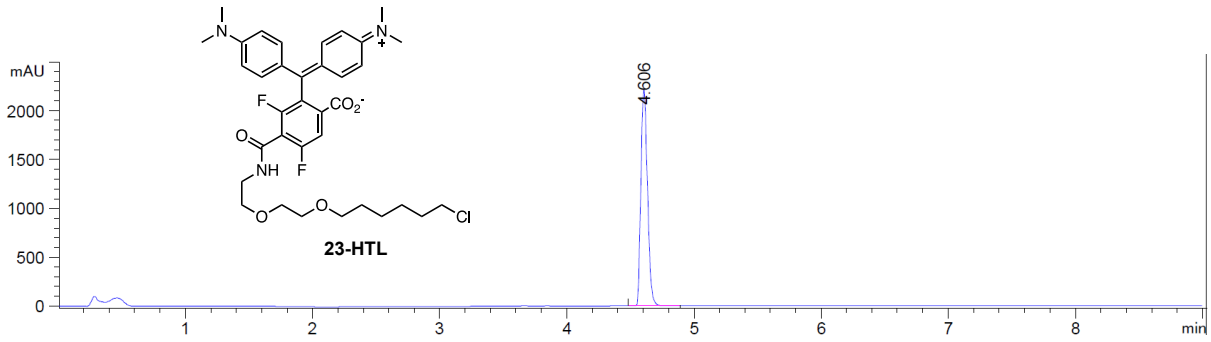
S23



S23


Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 293.8
4 Pulse Sequence zg30
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-06-08T15:35:00
8 Spectrometer Frequency 400.17
9 Spectral Width 8012.8
10 Lowest Frequency -1538.7
11 Nucleus 1H
12 Acquired Size 32768
13 Spectral Size 65536
H2O
(CH3)2SO
DMF
DMF
EtOAc
EtOAc
EtOAc
23-HTL
f1 (ppm)


CCl3F
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Solvent DMSO
3 Temperature 294.2
4 Pulse Sequence zgfhigqn.2
5 Experiment 1D
6 Number of Scans 16
7 Acquisition Date 2022-06-08T17:33:00
8 Spectrometer Frequency 376.50
9 Spectral Width 89285.7
10 Lowest Frequency -82102.5
11 Nucleus 19F
12 Acquired Size 65536
13 Spectral Size 131072
TFA
23-HTL
f1 (ppm)



23-HTL