Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Structural Characterisation of the Catalytic Core of the Class III Adenylyl Cyclase Rv0386 from *Mycobacterium tuberculosis*

Christina Pancevac

Master of Science in Engineering Biology

Heidelberg, 2006

Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

Christina Pancevac,

Civilingenjörsexamen i teknisk biologi,

Master of Science in Engineering Biology,

born in Göteborg, Sweden

Oral-Examination:……………………………………………….

Structural Characterisation of the Catalytic Core of the Class III Adenylyl Cyclase Rv0386 from *Mycobacterium tuberculosis*

Referees: Prof. Dr. Irmgard Sinning

               Prof. Dr. Heiner Schirmer

TABLE OF CONTENTS

Acknowledgements

First, I would like to thank Prof. Dr. Irmi Sinning for giving me the opportunity to carry out my PhD thesis in her lab. It has been great to work in a very well equipped lab including the possibility to test crystals in the in-house X-ray facility.

I am indebted to Prof. Dr. Heiner Schirmer for being my second supervisor. I am deeply grateful for his constant interest and support.

My thanks also go to Prof. Dr. Werner Buselmaier and Prof. Dr. Thomas Rausch for being so kind to be in my thesis committee.

I would also like to express my gratitude to Dr. Ivo Tews for his supervision, constant support, continuous interest, constructive suggestions and always showing an interest in discussions. I will always remember our cake group meetings.

Many thanks go to Dr. Dmitri Svergun, EMBL-Hamburg for the invaluable help during the small angle X-ray scattering experiments, for processing and analysing the data. I am sincerely obliged to him for always being there for my questions and for a nice collaboration.

It has also been a pleasure to work with Dr. Lucila Castro, Universität Tübingen. We have had fruitful discussions and great collaboration. I thank Dr. Jürgen Linder and Prof. Dr. Joachim Schultz, Universität Tübingen for introducing the adenylyl cyclase topic into our lab.

I would like to express my gratitude to Jacek Mazurkiewicz and Dr. Karsten Rippe, Kirchhoff Institut für Physik, Universität Heidelberg for carrying out the analytical ultracentrifugation experiments. It has been a pleasure to work with you.

Thanks to Dr. Jens Pfannstiel for performing the ESI-MS experiment. His friendly and open character contributed to a positive environment.

My gratefulness goes to Dr. Kyriaki Galani for her critical remarks and proof reading of my PhD thesis. I enjoyed our cheerful get-together in my life outside the lab.

Many thanks go to Ken Rosendal, Badri Konkimalla, Dr. Ulrike Dürrwang, Eva-Maria Knapp, Felix Heise and Oliver Schlenker for a great working atmosphere, support and productive discussions. I have enjoyed working with you and would like to send my faithful gratitude for motivating me during difficult moments. I wish you all the best in the future.

Satu Honkala, Angela Ku, Karolina Lennerth and Gabriella Wastenson I would like to thank for their support and encouragement. Real friends are never to far away. Vänner för livet.

I would also like to thank my brother Daniel. Thank you for being there. It will be great watching my cute nephew William growing up.

Last but not least, I would like to thank Oliver and my parents for their enduring support, love, understanding and patience, in bad and good moments. My gratitude is from the bottom of my heart and as a thank you I dedicate this thesis to them. Without you this thesis would not have been written. Thank you for always being there. You mean everything to me.
Volim vas.

Abstract

Tuberculosis (TB) is a chronic disease, with one third of the worlds' population infected. Multi drug resistant *Mycobacterium tuberculosis* strains have evolved, which are unaffected by the commonly used antibiotics. If the disease is not treated it could spread uncontrollably and become a pandemy. It is therefore imperative to identify novel drug targets. There is a great interest in adenylyl cyclases (ACs) from *M. tuberculosis*, which generate the universal second messenger cAMP, mostly since regulatory processes involving cAMP are vital for the pathogen. Elevated levels of cAMP are observed in phagosomes during infection, which protects the bacteria from elimination. The high number of ACs in *M. tuberculosis* suggests that the pathogen has the ability to sense and respond to a number of extracellular and intracellular signals through cAMP signalling. These processes are of fundamental interest, but investigation of the involved enzymes may also directly lead into the drug discovery process and result in new strategies for treatment.

This study deals with the X-ray crystallographic structure determination of the catalytic core of cyclase homology domain (CHD) of the class III adenylyl cyclase Rv0386 from *Mycobacterium tuberculosis*. This CHD is exceptional as it can use both, ATP and GTP, as substrates. Class III CHDs have been shown to exist as dimers in a head-to-tail arrangement. In contrast, the structure presented here identifies a novel dimeric contact, which we define as head-to-head. In this state, the CHD can not form proper active sites and thus is inhibited. This is in line with the low catalytic activity observed *in vitro*. The interface contact of the Rv0386 CHD dimer is substantial and buries 1497 Å^2. To test whether the contact was generated during crystallisation, the in-solution properties of Rv0386 CHD were examined by small angle X-ray scattering, which confirmed the head-to-head interface.

To learn about the substrate recognition of the Rv0386 CHD, a model of the active state was created. It shows GTP/ATP binding by the non-canonical purine binding residues Asn106 and Gln57 made possible with 180° rotation of both amide side chains. A discussion of homologous AC structures reveals that in the commonly observed eight-stranded β-sheet Rv0386 lacks the conserved β-strand β5. In the head-to-tail arrangement of a catalytically active dimer, β5 contributes to the interface which suggests that β5 forms in Rv0386 when in the active conformation. Since β5 also carries catalytic residue Asn106, this region might be an activity switch operated on homo-dimerisation of the enzyme. To obtain the active conformation, one of the two CHDs of Rv0386 as observed in the crystal structure would have to rotate around 48.2° and translated by 16.8 Å.

Crystallisation trials to obtain the active state were carried out in the presence of substrates or inhibitors. However, the dimer contact of Rv0386 CHD to form the head-to-head dimer is clearly more stable than the head-to-tail dimer in the active state, even when attempting to occupy the substrate binding pockets. The cofactor manganese is required for Rv0386 CHD catalytic activity *in vitro*. In this work manganese additionally was identified to interfere with dimerisation of Rv0386 CHD, which was unexpected. In the presence of Mn^{2+}, Rv0386 CHD is monomeric as characterised by size exclusion chromatography, analytical ultracentrifugation and small angle X-ray scattering. Attempts were made to convert the monomeric species to a catalytic dimer by addition of substrate analogs. Initial crystals obtained would require further improvement.

Zusammenfassung

Tuberkulose (TB) ist eine chronische Krankheit, an der ein Drittel der Weltbevölkerung infiziert ist. Mehrfachresistente *Mycobacterium tuberculosis*-Stämme sind entstanden, die durch bisher existierende Antibiotika nicht bekämpft werden können. Wird die Krankheit nicht behandelt, könnte sie sich unkontrolliert ausbreiten und zur Pandemie anwachsen. Es ist daher dringend geboten, neue Methoden zur Bekämpfung von Tuberkulose zu finden. Das Interesse an mycobakteriellen Adenylat-Zyklasen (Adenylyl Cyclases, ACs), die den universellen *second messenger* cAMP herstellen, ist groß, insbesondere weil regulatorische Prozesse, an denen cAMP beteiligt ist, lebenswichtig für dieses Pathogen zu sein scheinen. Man findet erhöhte cAMP Konzentrationen während der Infektion in den Phagosomen, und dies scheint die Bakterien vor ihrer Vernichtung zu bewahren. Die große Zahl von ACn bei *M. tuberculosis* legt nahe, daß das Pathogen die Fähigkeit hat, extra- und intrazelluläre Signale wahrzunehmen und mittels cAMP darauf zu reagieren. Diese Vorgänge sind von fundamentalem Interesse, Studie der involvierten Enzyme könnte aber direkt zur Entwicklung von Medikamenten und damit zu neuen Behandlungsmethoden führen.

In dieser Studie wird die kristallographische Röntgenstruktur der katalytischen Domäne (cyclase homology domain, CHD) der Klasse III-Adenylatzyklase Rv0386 aus *Mycobacterium tuberculosis* untersucht. Rv0386 ist eine außergewöhnliche Adenylatzyklase, da sie sowohl ATP als auch GTP als Substrat verwenden kann. Bisher wurden Klasse III-ACs als gegenläufige Dimere (head-to-tail) gefunden. Im Gegensatz dazu zeigt die Struktur von Rv0386 CHD einem neuen, Kopf-Kopf homodimeren Zustand, in dem sich keine aktiven Zentren ausbilden. Dies erklärt auch die geringe Aktivität des Proteins *in vitro*. Die Kontaktfläche des Dimers ist mit 1497 Å^2 recht groß. Um diesen neuen Dimer hinsichtlich der Kristallisation abzuklären, wurden die Eigenschaften in Lösung mittels Kleinwinkelröntgenstreuung untersucht. Dabei konnte der Kopf-Kopf Dimer verfiziert werden.

Um die Substraterkennung zu studieren, wurde hier ein Modell des aktiven Zustands erstellt. Gezeigt werden die ATP- und GTP bindenden Aminosäuren Asn106 und Gln57, die beide Substrate erkennen können, wenn die Seitenketten jeweils um 180° gedreht werden. Eine Diskussion der homologen AC-Strukturen zeigt, dass im Gegensatz zum üblicherweise identifizierten achtsträngigen β-Faltblatt der Klasse III-ACs bei Rv0386 CHD β-Strang β5 fehlt. Bei den gegenläufigen Dimeren trägt β5 zum Interface bei, was vermuten läßt, daß β5 in Rv0386 beim Übergang vom inhibierten zum aktiven Zustand gebildet wird. Da β5 zusätzlich das katalytische Asn106 trägt, kann diese Region als Aktivitätsschalter während der Interaktion der Monomere im katalytischen Dimer angesehen werden. Um vom inaktiven Dimer zum katalytisch aktiven Dimer zu gelangen, müsste man eines der Monomeren um 48,2° drehen und 16,8 Å entlang der Rotationsachse verschieben.

Kristallisation wurde mit zusätzliche Substraten oder Inhibitoren versucht. Allerdings scheint der Kopf-Kopf Dimer von Rv0386 CHD so stabil zu sein, dass der gegenläufige Dimer des aktiven Zustands nicht gebildet wird, sogar nach dem Versuch, die Substrat-Bindungsstellen zu besetzen. Co-Faktor Mangan ist notwendig für die Katalyse von Rv0386 CHD *in vitro*. Unerwarteterweise wird in dieser Arbeit gezeigt, dass Mangan zusätzlich das Oligomerisierungsverhalten beeinflusst. In Gegenwart von Mn^{2+} ist Rv0386 CHD ein Monomer, wie über Gel-Chromatographie, analytische Ultrazentrifugation und Kleinwinkelstreuung nachgewiesen. Versuche, den Monomer über Zugabe von Substrat-Analoga in den Dimer zu überführen, führte zu Kristallen, die jedoch für die Analyse noch weiter verbessert werden müssten.

Publication and Conference Presentations

This thesis is included into the publication below:

Christina Pancevac, Dmitri I. Svergun, Irmgard Sinning, Ivo Tews. (2006) Regulatory principles of dimer formation in the catalytic domains of the adenylyl cyclase Rv0386 from *Mycobacterium tuberculosis.* (in preparation).

Poster presented at the following conferences:

Christina Pancevac, Dmitri I. Svergun, Irmgard Sinning, Ivo Tews. (2005) The mycobacterial adenylyl cyclase Rv0386 crystallised in a dimeric conformation suggestive of the autoinhibited state, Structural Biology of Molecular Recognition, Murnau, Germany, September 15th-17th.

Christina Pancevac, Dmitri I. Svergun, Irmgard Sinning, Ivo Tews. (2005) The mycobacterial adenylyl cyclase Rv0386 crystallised in a dimeric conformation suggestive of the autoinhibited state, SB-Net (Structural Biology Network), Tällberg, Sweden, June 17th-20th.

Christina Pancevac, Lucila I. Castro, Joachim Schultz, Irmgard Sinning, Jürgen Linder, Ivo Tews. (2004) The dimeric structure of the mycobacterial adenylyl cyclase Rv0386: a possible autoinhibited enzymatic state, EMBO Conference on Structures in Biology, Heidelberg, Germany, November, 10th-13th.

List of Figures

List of Tables

Abbreviations

A	adenosine analog
Å	Ångström (1×10^{-10} m)
AC	adenylate cyclase
Ad	adenine
AIDS	acquired immune deficiency syndrome
AmpR	ampicillin resistant
5´AMP	adenosine 5'monophosphate
ATP	adenosine triphosphate
AU	asymmetric unit
AUC	analytical ultracentrifugation
BCG	bacillus calmette-guerin vaccine
C1	cyclase homology domain 1 of mammalian AC
C2	cyclase homology domain 2 of mammalian AC
Ca^{2+}	calcium
cAM	calmodulin
cAMP	3´,5´-cyclic adenosine monophosphate
cGMP	3',5'-cyclic guanosine monophosphate
CHD	cyclase homology domain
C-His	C-terminal histag
CmR	Chloramphenicol resistant
DLS	dynamic light scattering
E	ground state
E*	substrate-free activated state
E**	substrate/product bound state
E. coli	*Escherichia coli*
EM	cryo electron microscopy
EPAC	exchange proteins directly activated by cAMP
ESI-MS	electrospray ionisation mass spectrometry
ESRF	European synchrotron radiation facility
Fc	calculated structure factor
Fo	experimentally observed structure factor
FOM	figure of merit
FP	structure factor of the native protein
FPH	structure factor of the derivatised protein
Gα	α subunit of heterotrimeric G protein
GAF	cGMP phosphodiesterases adenylyl cyclases and formate-hydrolase transcription activator
GAP	GTPase-activating protein
Gβ	β subunit of heterotrimeric G protein
GC	guanylate cyclase
GDP	guanosine diphosphate
Gγ	γ subunit of heterotrimeric G protein
Gi	inhibitory heterotrimeric G protein
5´GMP	guanosine 5'monophosphate
GPCR	G-protein coupled receptor
Gs	stimulatory heterotrimeric G protein
GTP	guanosine triphosphate

HAMP	histidine kinases, adenylyl cyclases, methyl-accepting chemotaxis proteins and phosphatases
HCO_3^-	bicarbonate
Hepes	4-(2-hydroxyethyl)piperazine-1-ethanesulphonic acid
HisK	histidine kinase domains
HIV	human immunodeficiency virus
HR	highest resolution bin
H.sapiens	*Homo sapiens*
HTH	helix-turn-helix DNA binding domain
IPTG	isopropyl-thio-β-D-galactopyranoside
I(q)	scattering intensity
KanR	kanamycin resistant
kD	kilodalton
Kd	dissociation constant
λ	wavelength
mAC	mammalian adenylyl cyclase
MAD	multiple anomalous dispersion
MDR	multidrug-resistant
Me	metal binding
MtAC	AC from *M.tuberculosis*
M. tuberculosis	*Mycobacterium tuberculosis*
Mw	molecular weight
N-His	N-terminal histag
NMR	nuclear magnetic resonance
OD_{600}	optical density at $\lambda = 600$ nm
ON	over night
Pα	α-phosphate
Pβ	β-phosphate
PAM	protein associated with Myc
PBS	phosphor-buffered saline
PCR	polymerase chain reaction
PDB	protein database
PDE	phosphodiesterase
PEG	polyethylene glycol
Pγ	γ-phosphate
PKA	protein kinase A
PKC	protein kinase C
PP2B	protein phosphatase 2B
PPi	pyrophosphate
PPIase	protein cis-trans peptidylprolyl isomerase
P(q)	form factor
PVDF	polyvinylidene difluoride
rAC	receptor-type membrane bound adenylyl cyclases
r.m.s.d	root mean square deviation
RGS	regulator of G protein signaling
Rv0386	AC of *M. tuberculosis*, open reading frame number 386
Rv1264	AC of *M. tuberculosis*, open reading frame number 1264
Rv1900c	AC of *M. tuberculosis*, open reading frame number 1900
sAC	soluble adenylate cyclase
SAD	single anomalous dispesion

SAXS	small angle X-ray scattering
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SeMet	L-selenomethionine
S(q)	structure factor
TB	tuberculosis
T.brucei	*Trypanosoma brucei*
6TM	domain with 6 transmembrane helices
TPR	tetratricopeptide repeat domain
Tr	transition state
TRIS	tris-(hydroxymethyl)-aminomethane
V_E	retention volume
V_{max}	maximum velocity at infinite substrate concentration

Abbreviations for Residues

1 letter code	3 letter code	Full name
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagines
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophane
Y	Tyr	Tyrosine
X		any residue
aa		amino acid

Tillägnat Oliver och mina föräldrar
Посвећено Оливеру и мојим родитељима
Dedicated to Oliver and my parents

Wer andere erkennt, ist gelehrt.
Wer sich selbst erkennt ist weise.
Wer andere besiegt, hat Muskelkraft.
Wer sich selbst besiegt, ist stark.
Wer zufrieden ist, ist reich.
Wer seine Mitte nicht verliert,
ist unüberwindlich.

Lao-Tse

Chapter 1

INTRODUCTION

1.1 Adenylyl Cyclases

1.1.a Intracellular Signals

In the late 1950's Sutherland and coworkers discovered the ubiquitous second messenger 3′,5′-cyclic adenosine monophosphate (cAMP), which lead to the essential milestone for the prototype of hormone signaling through second messengers (Rall and Sutherland, 1962). Sutherland's finding gave him the Nobel prise in 1971. The discovery that 3',5'-cyclic guanosine monophosphate (cGMP) is as well substantial in signal transduction came shortly after the cAMP finding. The enzymes adenylyl cyclase (AC, EC 4.6.1.1) catalyses the synthesis of cyclic AMP and pyrophosphate (PPi) from adenosine triphosphate (ATP) (Rall and Sutherland, 1962) and guanylyl cyclase (GC, EC 4.6.1.2) catalyses the synthesis of cyclic GMP and PPi from guanosine triphosphate (GTP). Both cAMP and cGMP are degraded by cyclic AMP/GMP phosphodiesterases (PDEs) to adenosine 5'monophosphate (5'AMP) and guanosine 5'monophosphate (5'GMP) respectively (Conti and Jin, 1999).

Intracellular receptors such as cyclic nucleotide-dependent protein kinases, PDEs, cyclic nucleotide-gated ion channels, Rap1 guanine nucleotide-exchange factor (EPACs), (de Rooij et al., 1998) cGMP phosphodiesterases adenylyl cyclases and formate-hydrolase transcription activator (GAF) domains, can be activated by eukaryotic cAMP and cGMP. The key regulators

of these signaling pathways are kinases and their activation by second messengers initiate a cascade of downstream events resulting in a cellular response. Protein kinase A (PKA) regulates, sugar, glycogen and lipid metabolism and protein kinase C (PKC) regulates cell division and proliferation. As the second messenger cAMP controls many important functions in the cell, the significance of ACs is highlighted.

1.1.b Adenylyl Cyclase Signaling Pathway in Mammals

Mammalian and bacterial adenylyl cyclases are regulated in different manners. So far, the regulation of prokaryotic ACs is not very well understood. In contrast, the regulation of mammalian adenylyl cyclases (mACs) is most studied and will be described in the following paragraphs. G-protein coupled receptors (GPCRs) are found in all eukaryotes and belong to the largest family of cell surface receptors. They are known to be integral membrane proteins that comprise seven transmembrane helices and can be activated by a range of ligands, such as light, calcium, amino acids, fatty acids, neurotransmitters (serotonin), hormones (dopamine, adrenalin), odorants, pheromones, small molecules and proteins. The same ligand can activate many different G-protein linked receptors.

Upon binding of an extracellular ligand to a GPCR, the receptor alters its conformation and activates the stimulatory heterotrimeric G protein (Gs) comprising subunits α, β and γ (Figure 1.1). In doing so, Gsα releases GDP replacing it with GTP. GTP binding causes the G protein to dissociate into two activated components, Gα and Gβγ. The concomitant conformational change of Gα allows the interaction with its target proteins. The Gβγ conformation is not altered, but the functional surface interacting with Gα is now accessible. Target proteins of the dissociated G proteins are enzymes or ion channels localised in the plasma membrane. The active Gα binds to and stimulates the enzyme adenylyl cyclase to convert ATP to the second messenger cAMP. The lifetime of the active Gα is short, its GTPase activity is stimulated when bound to adenylyl cyclase. GTP hydrolyses to GDP, inactivates Gα and AC, and Gα reassociates with Gβγ. In most cells cAMP is found at a concentration of approximately 10^{-7} M in the cytosol. Cyclic AMP activates mainly cyclic–AMP-dependent PKA. The binding of four cAMPs to the regulatory

subunits of PKA, leads to a conformational change dissociating the regulatory from the catalytic subunits in order to activate the catalytic subunits of the kinase (Alberts et al., 2002).

In addition, to stimulatory G protein (Gs) there is an inhibitory G protein (Gi). There are GPCRs, such as α2-adrenergic receptors that are coupled to adenylyl cyclase by an inhibitory G protein inhibiting the activity of the adenylyl cyclase. The Gi protein comprises the subunits β and γ, just as Gs, but Gi has a different α (α_i), which inhibits the adenylyl cyclase. The complex of βγ dissociated from Gi, acts inhibitory by binding to free α_s (Gs) available in the proximity.
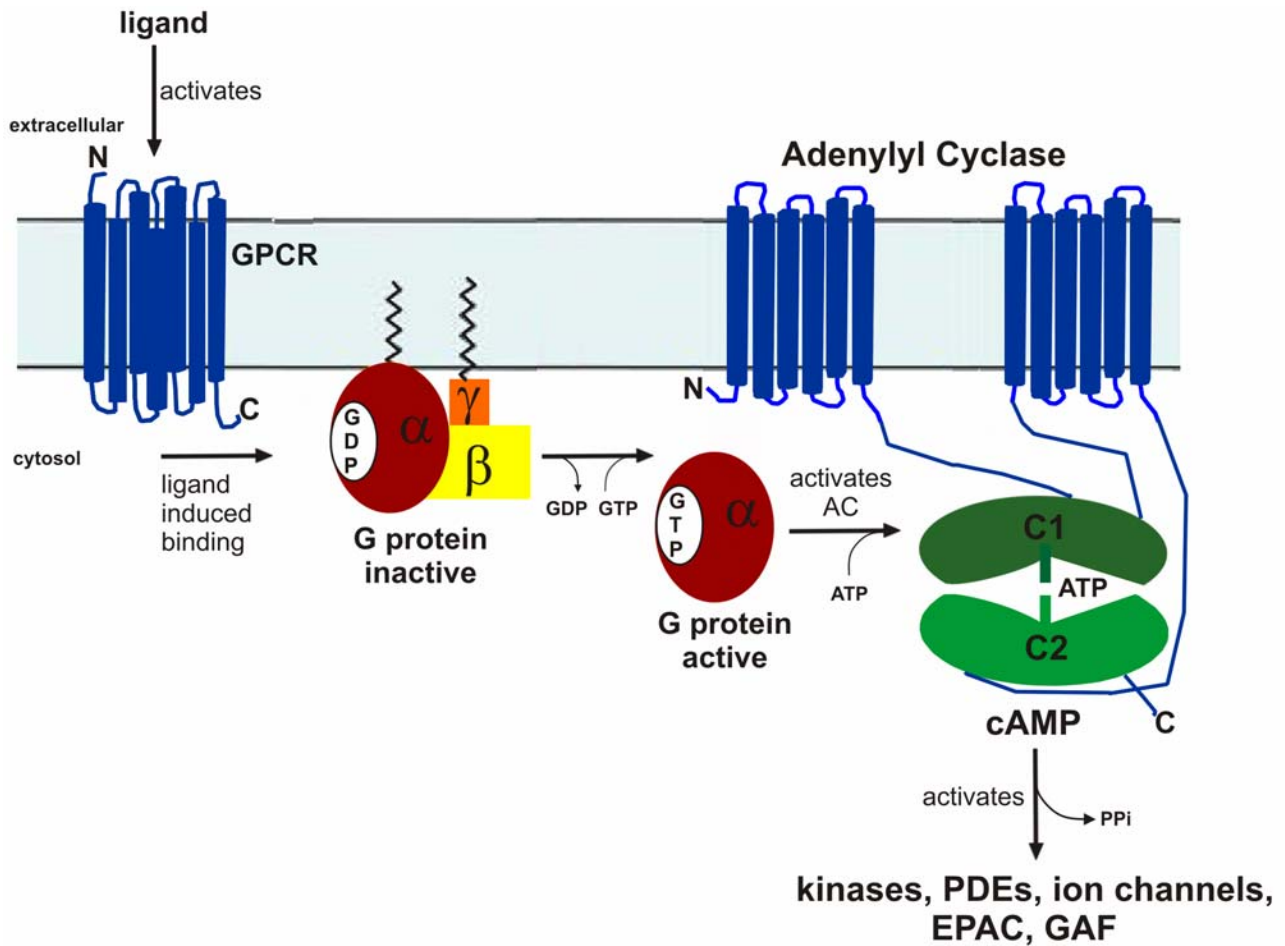
Figure 1.1: Adenylyl cyclase signaling pathway. An extracellular ligand binds to a G-protein coupled receptor (GPCR) which changes the conformation of the receptor and alters the conformation of the heterotrimeric G protein bound to the receptor. The G protein releases its bound GDP and replaces it with GTP, resulting in an activation and dissociation of Gα from Gβγ. The active Gα activates adenylyl cyclases, which in turn converts ATP to cAMP. The second messenger cAMP activates downstream proteins, such as cyclic AMP dependent protein kinases, phosphodiesterases, ions channels, exchange proteins directly activated by cAMP (EPAC) and cGMP phosphodiesterases adenylyl cyclases and formate-hydrolase transcription activator (GAF) proteins.

1.2 Classification of Adenylyl Cyclases

In eukaryotic and eubacterial organisms adenylyl cyclases play an important role in signal transduction (Barzu and Danchin, 1994; Sunahara et al., 1996). On the basis of the amino acid sequence adenylyl cyclases are classified into three groups (Danchin, 1993). Recently, these were expanded into six classes based on homology and structural organisation (Cotta et al., 1998; Sismeiro et al., 1998; Tellez-Sosa et al., 2002).

Class I adenylyl cyclases, "the enterobacterial class", is highly conserved in enterobacteria (Baker and Kelly, 2004) and in many gram-negative bacteria (Barzu and Danchin, 1994). Class II adenylyl cyclases, present in the pathogens *Bacillus anthracis* (Gram-positive bacterium causing anthrax) and *Bordetella pertussis* (Gram-negative bacterium causing whooping cough) is for its occurrence called "the toxic class". Class III adenylyl cyclases, the "universal class" are found in prokaryotes and eukaryotes. Class III are closely related to guanylyl cyclases (Linder and Schultz, 2003; Schultz et al., 1992) which suggest that guanylyl cyclases and adenylyl cyclases share a common ancestor. Class IV to VI are all found in microbes, but have not been studied extensively and contain only a single species each. The class IV AC is present in *Aeromonas hydrophila* (Sismeiro et al., 1998), class V adenylyl cyclase is found in *Prevotella ruminicola* (Cotta et al., 1998) and class VI AC in *Rhizobium etli* (Tellez-Sosa et al., 2002). ACs class I-III are sufficiently characterised and introduced in the following sections.

1.2.a AC Class I

The class I adenylyl cyclases are highly conserved. They occur in the human pathogens *Salmonella typhimurium*, *Yersinia pestis* and *Escherichia coli* and the Gram-negative bacteria *Aeromonas hydrophila* and *Vibrio cholerae* (Danchin, 1993).

Class I adenylyl cyclases are characteristically rich in cysteine and histidine residues and consisting of two domains (Roy et al., 1983). The catalytic domain is situated at the N-terminus and the glucose sensitive regulatory domain at the C-terminus. The enzyme is activated by phosphorylation by an important histidine, located in the regulatory domain (Danchin, 1993).

The adenylyl cyclase CyaA of *E. coli* is regulated by glucose. Low levels of glucose increases the synthesis of cAMP and high glucose levels inhibit the AC. High levels of cAMP contribute to transcriptional activation of the catabolite gene activator protein (CAP). CAP activates genes that enable *Escherichia coli* to use alternative carbon sources, when the favoured glucose is unavailable. The complex of cAMP and CAP binds to specific DNA sequences near target promotors and thereby initiates transcription of various genes, including the lac operon. In this way, the level of cAMP in the cell determines if the expression of the target gene is switched on or off (Alberts et al., 2002; Baker and Kelly, 2004). So far, structural information from class I adenylyl cyclases is lacking.

1.2.b AC Class II

Class II ACs are present in Gram-negative bacterium *Bordetella pertussis*, causing whooping cough and the Gram-positive bacterium *Bacillus anthracis* causing anthrax. The AC toxins share structural and functional characteristics and they are activated by the host protein Ca^{2+} binding calmodulin (CaM), which is not present in bacteria (Glaser et al., 1988; Kather and Aktories, 1983; Mock et al., 1988).

Class II adenylyl cyclases are extracellular toxins. The enzymes are secreted, penetrating mammalian cells, and subsequently use mammalian CaM for activation to synthesize high levels of cyclic AMP. The accumulation of cAMP in the host cells disrupts the normal regulation of the host's AC/PDE and intracellular signaling pathways (Ahuja et al., 2004; Barzu and Danchin, 1994). The primary target of the ACs is the immune effector cells; CD11b[+] dendritic cells which are the most potent antigen presenting, as well as CD8[+] CTL and CD4[+] T helper cells (Simsova et al., 2004). The ultimate goal of adenylyl cyclase toxins is to accumulate cAMP in the immune effector cells, which poisons the immune system and enables survival in the host. It is unknown if class II ACs also have an intracellular role in the pathogens (Ahuja et al., 2004; Barzu and Danchin, 1994).

Bacillus anthracis ACs comprise two domains, where the domain situated at the N-terminus is important for interacting with another bacterial toxin and the located domain at the C-terminus binds calmodulin (Ahuja et al., 2004). The crystal structures of the *Bacillus anthracis* anthrax oedema factor, a calmodulin activated AC has been determined with and without bound calmodulin (Drum et al., 2002). Oedema factor and mammalian ACs do not share structural homology. Additionally, oedema factor functions as monomer and mammalian ACs as dimers.

1.2.c AC Class III

Class III adenylyl cyclases are present in eukaryotic and prokaryotic species. The catalytic domain of class III ACs, also designated as the cyclase homology domain (CHD), is greatly diverging in their primary sequences (Barzu and Danchin, 1994; Sunahara et al., 1996) but shared by all class III ACs. The classification of class III was made based on Bayesian algorithms for multiple sequence alignment and database searching (McCue et al., 2000), suggesting class III adenylyl cyclases to be classified as cytoplasmic, receptor-type membrane bound and integral membrane proteins (Figure 1.2), as described below.

Domain Organisation of Class III AC

Two cyclase homology domains form one adenylyl cyclase with one or two active centres in the interface of the dimer. The active centre comprises six canonical catalytic residues involved in catalysis (Sunahara et al., 1998; Tesmer et al., 1999; Yan et al., 1997b). Soluble adenylyl cyclases (sACs) are mainly present in prokaryotes and form homodimers, with two symmetric catalytic centres including additional regulatory domains (Figure 1.2A). Receptor-type membrane bound adenylyl cyclases (rACs) do as well form homodimers with two symmetric catalytic sites and usually contain an additional extracellular domain (Figure 1.2B). Adenylyl cyclases arranged as integral membrane proteins are present in mammals and in some prokaryotes (Figure 1.2C) comprising 6 predicted transmembrane regions (6TM) followed by a cytoplasmic catalytic domain (C1, C2) in a dimeric arrangement; 6TM-C1-6TM-C2. The catalytic domains C1 and C2 are homologous, well conserved with the sequence identity of 25 %. Biochemical and X-ray structures revealed that mammalian catalytic domains function as heterodimers with one pseudosymmetric catalytic centre and catalysis takes place at the interface of the dimer (Tesmer et al., 1997; Tesmer et al., 1999; Whisnant et al., 1996; Zhang et al., 1997). The function of the

transmembrane regions is unknown. So far, the transmembrane regions are known to localise in the membrane and to coordinate the interaction of the catalytic domain. The catalytic core alone is capable of the catalysis and it is still unknown why the transmembrane region comprise 12 transmembrane helices, instead of two (Hurley, 1999; Tang and Hurley, 1998). Topological analogy to transporters has suggested a role in transmembrane transport, which has not been confirmed (Krupinski et al., 1989).
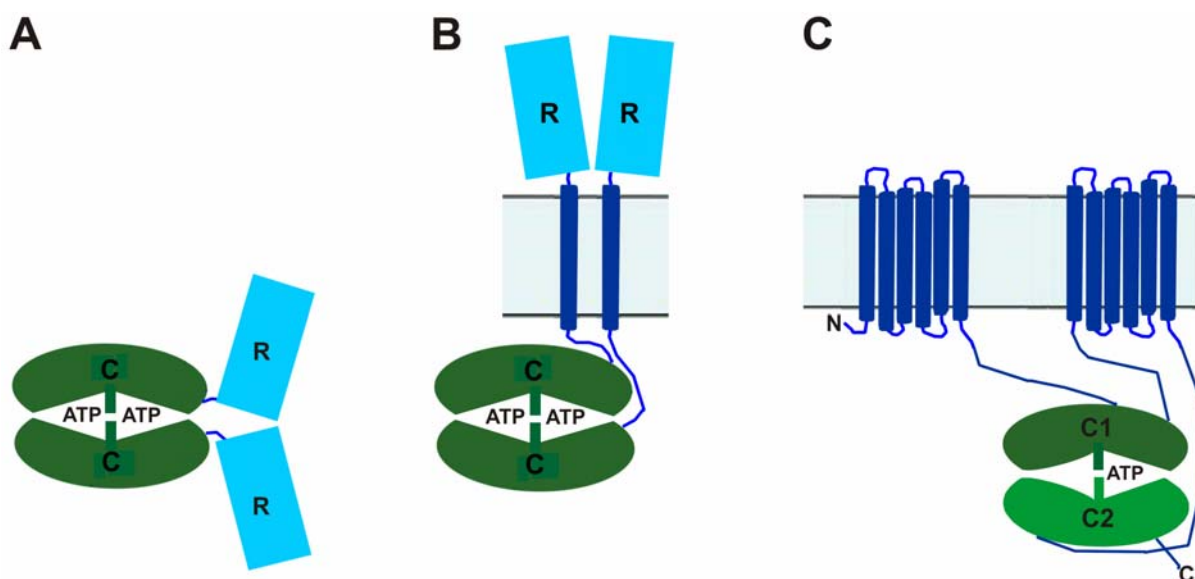
Figure 1.2: Classification of class III adenylyl cyclases. Cyclic homology domains (C, C1, C2) are visible in green, transmembrane regions in dark blue and in light blue the extracellular -and regulatory (R) domains are shown. A) Cytoplasmic homodimeric ACs with two catalytic centres and additional regulatory domains attached. This is the most occurring subclass in prokaryotic ACs. B) Receptor-type ACs are homodimers comprising two catalytic centres, an extracellular domain and a single transmembrane helix. C) Integral membrane proteins include 6TM-C1-6TM-C2. The catalytic core contains one pseudosymmetric catalytic centre.

AC Class IIIa-d

Recently, a classification of class III adenylyl cyclases into subgroups IIIa to IIId has been presented by (Linder and Schultz, 2003) based on primary sequence alignment of cyclase homology domains (Figure 1.3) (Barzu and Danchin, 1994; Linder and Schultz, 2003). Six canonical residues have been identified important for substrate binding (Sunahara et al., 1998; Tesmer et al., 1999; Yan et al., 1997b). The substrate-defining residues Lys and Asp are selecting ATP over GTP. Two conserved Asp coordinate two metal-cofactors (Mg^{2+} and Mn^{2+}), and Asn and Arg are involved in transition-state stabilisation. Additionally, the conserved Arg is binding to the γ-phosphate of the substrate. Another characteristic of class III ACs is the arm region, also

called dimerisation arm, situated in the interface of two cyclase homology domains, believed to be important for dimerisation of the catalytic core. Detailed descriptions of the differences in class IIIa-d are described below and for alignment see (Figure 1.3).

AC Class IIIa

Class IIIa adenylyl cyclases are pseudoheterodimeric in metazoans (Krupinski et al., 1989) and homodimeric in *Mycobacterium tuberculosis* (Guo et al., 2001) and the cyanobacterium *Spirulina platensis* (Kasahara et al., 2001). Distinctive signature motifs are the predicted (F/Y)XX(F/Y)D motif which seems to be important for interface formation (Tang et al., 1995; Tesmer et al., 1997) and the EKI<u>K</u> motif comprising the substrate defining lysine (underlined). Furthermore, class IIIa catalytic homology domains contain the conserved arm region (Tesmer et al., 1997) which is 14 residues in length.
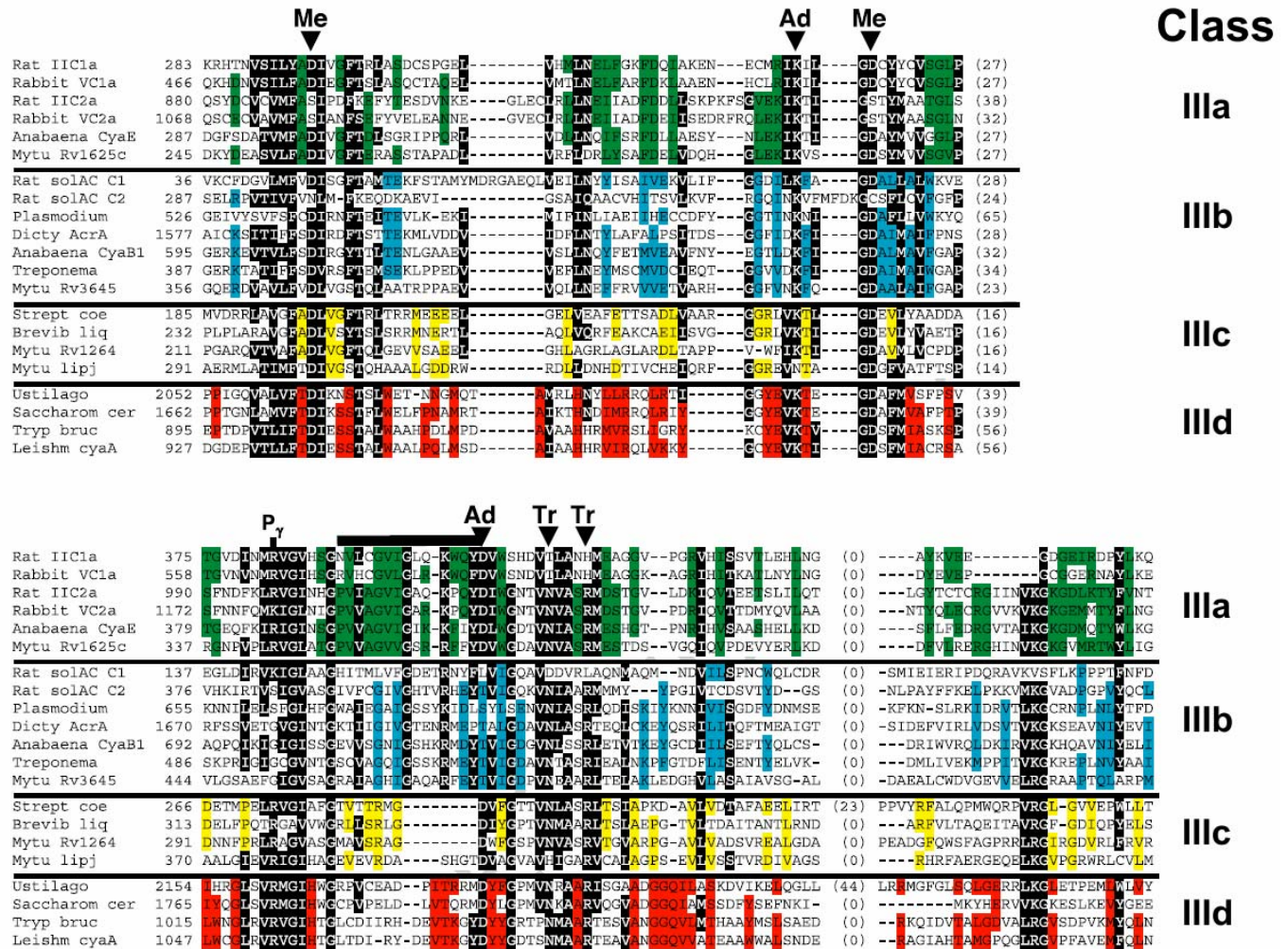
Figure 1.3: Sequence alignment of class III AC cyclase homology domains, sub-categorised in class IIIa-d. Residues similar in all cyclase homology domains have black background and conserved residues specific for its class are depicted in green for class IIIa, cyan for class IIIb, yellow for class IIIc and in red for class IIId. The six canonical residues are marked with Me: metal binding, Ad: adenine/purine binding, Tr: transition-state stabilising. The conserved Arg involved in γ-phosphate (Pγ) binding in class IIIa, IIIc and IIId is as well defined. The arm region is shown as a top bar, between the conserved Gly and the downstream adenine binding region. Image taken from (Linder and Schultz, 2003).

AC Class IIIb

Class IIIb ACs are present in Gram-negative bacteria, Gram-positive bacteria like *M. tuberculosis*, protozoa like *Plasmodium* and in mammals. The characteristics of class IIIb cyclase homology domain is a replacement of the substrate–defining aspartate with a threonine or serine (Kanacher et al., 2002). Additionally, a phosphate–binding arginine is frequently replaced by a serine or glycine and the arm region mentioned above comprises 15 residues (Linder and Schultz, 2003).

AC Class IIIc

Class IIIc adenylyl cyclases have been detected in Gram-negative and Gram-positive bacteria. The arm region is completely absent or shortened to only 7-11 residues (Linder et al., 2002; Linder and Schultz, 2003). Furthermore, several cyclase homology domains contain substitutions in the six canonical catalytic residues, suggestive of a number of different catalytic mechanisms (Linder and Schultz, 2003).

AC Class IIId

Class IIId ACs are present in trypanosomatids, protozoan and fungi (Naula et al., 2001). This class comprises a YEV<u>K</u>T motif which is a prominent signature sequence, neighbouring the substrate-defining lysine. The arm region usually contains 14 residues (Linder and Schultz, 2003). An additional feature of this class is the Δ-subdomain (Bieger and Essen, 2001), existing in the trypanosomal cyclase homology domain. Interestingly, fungal class IIId does not contain any Δ-subdomain. So far, the function of this subdomain is unknown.

1.3 Regulation of Class III Adenylyl Cyclases

1.3.a Adenylyl Cyclase Class III Enzyme Cycle

Mammalian class III adenylyl cyclases have ten different isoforms (AC I-AC X) (Sunahara and Taussig, 2002) and they are until now the best biophysically and biochemically characterised

ACs (Hurley, 1998; Sunahara et al., 1996; Tang and Gilman, 1992). The cytoplasmic catalytic core has mainly been studied, resulting in a model of the enzyme cycle (Figure 1.4) (Tang and Hurley, 1998).

E*:A,PPi

E*:PPi ←→ E*:cAMP,PPi

E ←→ E* ←→ E**:ATP ←→ E**:cAMP,PPi

E** ←→ E**:PPi

E**:A,PPi
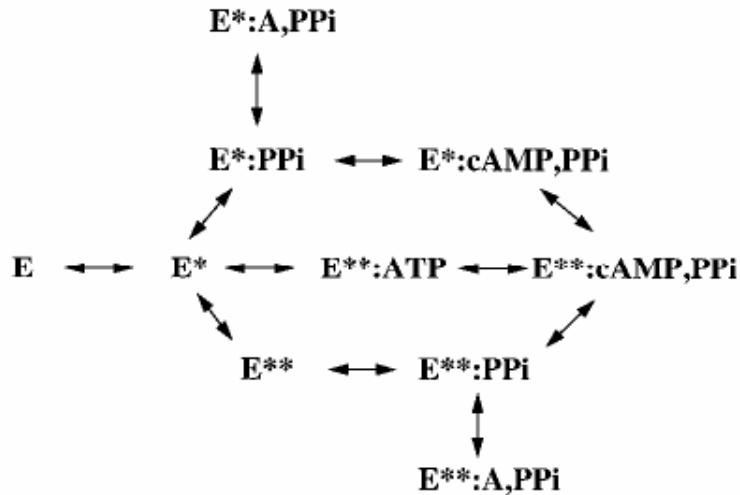
Figure 1.4: Model of the adenylyl cyclase catalysis cycle. The apo-form of the enzyme (E) converts to the semi-activated state (E*) by an activator like Gsα. E* changes to the fully activated state (E**) in the substrate (ATP) bound form. Subsequently catalysis takes place and forms the products cAMP and pyrophosphate (PPi). Adenosine analogs (A) are known as P-site inhibitors. Figure taken from (Tang and Hurley, 1998).

Mammalian adenylyl cyclases are present in the ground state (E), substrate-free activated state (E*) and substrate/product bound state (E**) (Tesmer et al., 1997; Zhang et al., 1997). The conformational transition of the cyclic homology domain from E to E* arise by activators or is blocked by inhibitors. Mammalian cyclase homology domains are activated by $G_s\alpha$, $G\beta\gamma$ (interacting with $G_s\alpha$), forskolin, Ca^{2+} -calmodulin and inhibited by $G_i\alpha$, $G\beta\gamma$ (interacting with $G_i\alpha$) and free Ca^{2+}. The E* to E** transition occurs on substrate binding, which induces a conformational change required for catalysis to proceed. In the E** state ATP is converted to cAMP. The E* state is reformed after the release of the products cAMP and PPi (Tang and Hurley, 1998).

Mammalian cyclase homology domains bind to ATP with affinity in micromolar range. P-site inhibitors are known adenosine analogs, named according to the intact purine ring (Desaubry et

al., 1996b; Johnson et al., 1979; Johnson et al., 1989; Londos and Wolff, 1977; Wolff et al., 1981)and bound in their product release state (E**:PPi). P-site inhibitors and cAMP compete to bind to the P-site of adenylyl cyclases in the presence of PPi. Interestingly, the adenosine analogs are not competing for the substrate binding in the absence of PPi. The pyrophosphate enhances the binding of P-site inhibitors, which have nanomolar affinity to the adenylyl cyclase. In contrast, cAMP is binding with low affinity (usually about 20 mM) to the enzyme (Dessauer and Gilman, 1997). The enzyme cycle outlined is not applicable for prokaryotic adenylyl cyclases, hence the low affinity of P-site inhibitors for bacterial adenylyl cyclases (Kanacher et al., 2002; Sinha et al., 2005). This is suggestive of a different E** state or no E** state for the prokaryotic adenylyl cyclases.

1.3.b Regulation of mammalian Adenylyl Cyclases

The mACs are postulated to be important for learning, memory, olfaction, drug dependency, long-term potentiation, sperm capacitation and synaptic plasticity (Sunahara and Taussig, 2002). Some prokaryotic ACs are regulated by the same effectors as the mammalian ACs. For comparison, a detailed description of the mammalian adenylyl cyclase effectors follows. Mammalian adenylyl cyclases are regulated via modulators, such as the heterotrimeric G protein (Gαβγ), P-site inhibitors, forskolin, pyrophosphate, calcium, calcium-calmodulin, bicarbonate, other proteins and posttranslational modification (Table 1.1).

Stimulation by Gsα

The GTP-bound G-protein α-subunit $G_s\alpha$ is known to activate all mammalian adenylyl cyclases (E to E*), except the soluble AC type X (Sunahara and Taussig, 2002). The complex of $G_s\alpha$-GTP binds into the groove of the catalytic domain C2 including α2, α3 and the N-terminal catalytic domain C1 (Figure 1.5). The $G_s\alpha$ binding site is about 30 Å away from the catalytic site (Tesmer et al., 1997; Yan et al., 1997a; Zimmermann et al., 1998). The $G_s\alpha$ regions contributing to the binding is switch II, including α2 and the loop regions of α4/β6 and α3/β5. Switch II is highly mobile and is one of the regions that changes conformation upon binding to GTP (Berlot and Bourne, 1992; Lambright et al., 1994). It has been demonstrated that $G_s\alpha$ in the GTP-bound form activates the AC 10-fold compared to the GDP-bound form (Sunahara et al., 1997a). After the

GTP hydrolysis the dissociation of the $G_s\alpha$ from the AC takes place and $G_s\alpha$ reassembles with $G\beta\gamma$. Subsequently, the signalling of $G_s\alpha$ and $G\beta\gamma$ is abolished. A specific regulator of G protein signaling (RGS), PXI-RGS serves as GTPase-activating proteins (GAP) for $G_s\alpha$ (Zheng et al., 2001), and increases the low intrinsic hydrolysis rate. As a consequence, $G_s\alpha$ is inactivated. The GTPase activity of $G_s\alpha$ can weakly be increased by the adenylyl cyclase itself (Scholich et al., 1999).

Class	Regulator	*H. sapiens* ACs	Effects
G proteins	Gs_α	All except X	+
	$G\beta\gamma$	I	-
		II, IV, VII	+
	Gi_α	I, V, VI	-
	Go_α	I	-
	Gz_α	I, V	-
Small chemicals	Forskolin	All except IX, X	+
	Adenosine	All	-
Feedback inhibition	PKA, PKC	II, V, VI, VII	+/-
	Ca^{2+}/CaM	I, V, VI, VIII	+/-
	PP2B	IX	-

Table 1.1: Regulators of mammalian adenylyl cyclases. ACI-X are regulated by a number of effectors, either stimulating or inhibiting the activity. PKA = protein kinase A, PKC = protein kinase C, PP2B = protein phosphatase 2B. Taken from (Findeisen, 2005).

Inhibition by Giα

The members of the inhibitory G protein family, $G_i\alpha$ include $G_i\alpha_1$, $G_i\alpha_2$, $G_i\alpha_3$, $G_o\alpha$, $G_z\alpha$ and inhibit adenylyl cyclases (E state) (Sunahara and Taussig, 2002). $G_i\alpha$ does not compete with $G_s\alpha$ for the binding to the enzymes. Instead, structural modelling and mutagenesis experiments propose a binding site for $G_i\alpha$, located in the crevice formed by α2 and α3 of C1, symmetrically opposite of the AC compared to $G_s\alpha$, (Dessauer et al., 1998; Tesmer et al., 1997; Yan et al., 1997a). The $G_i\alpha$ binding site is close to the AC active site. A rotation of the C1 catalytic domain is proposed as the inhibitory mechanism (Hurley, 1999) resulting in a prevention of ATP binding.
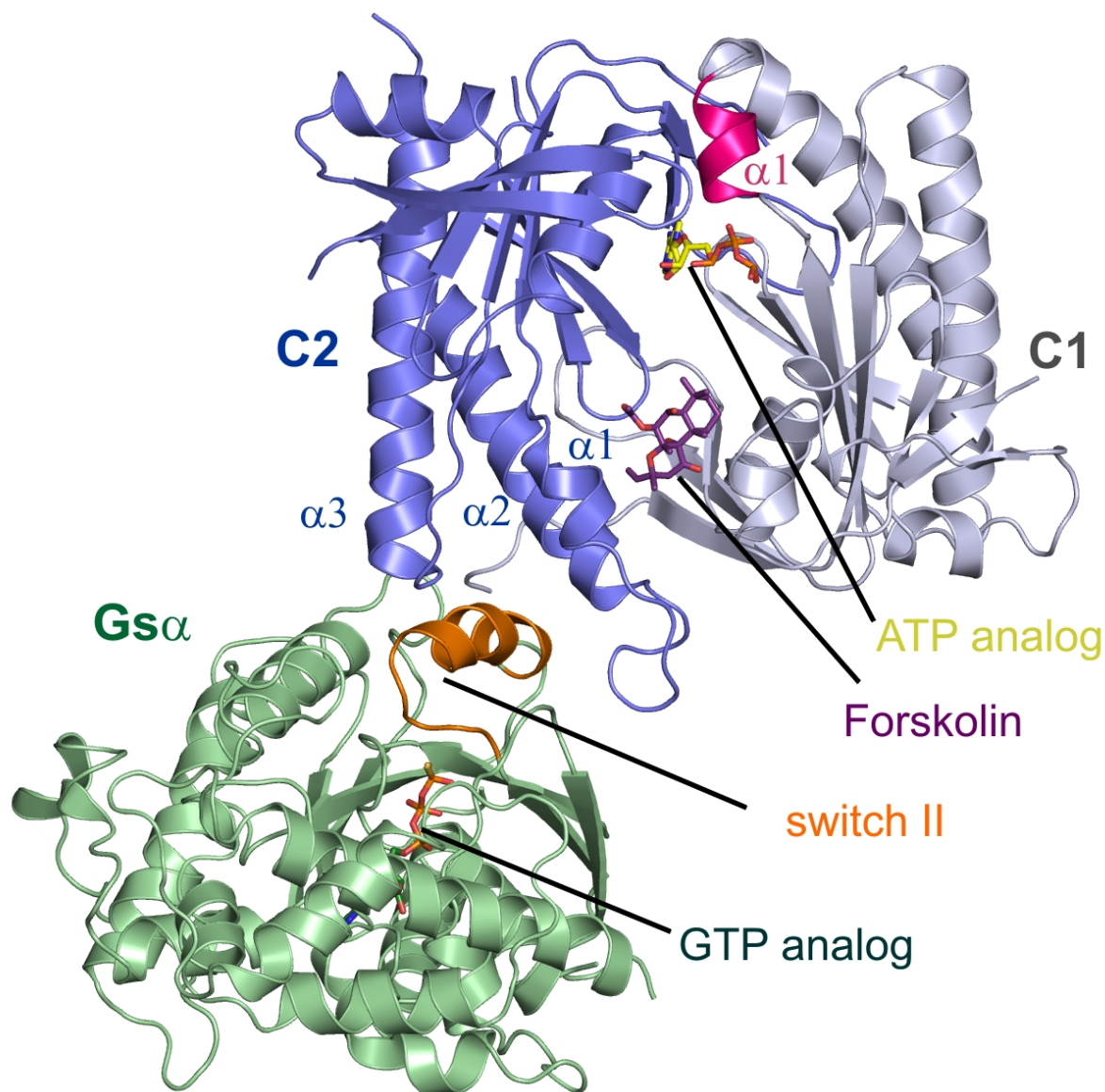
Figure 1.5: The catalytic core of mammalian AC (VC1 and IIC2) in complex with bound ATP analog, forskolin in the interface if the heterodimer and Gsα · GTP analog (PDB: 1CJK). C1, C2 and Gsα are rendered in blue, grey and green respectively. The ATP analog and the activator forskolin are colored in yellow and violet respectively. The α1-helix in C1moves upon substrate binding and is displayed in magenta. The switch II region (in orange) of the Gsα-GTP complex binds into the cavity of the C2 domain including α2, α3 and the N-terminal catalytic domain C1.

Regulation by Gβγ

The G protein Gβγ subunits are strong modulators of adenylyl cyclases. They can act stimulatory or inhibitory depending on the AC type (Gao and Gilman, 1991; Tang and Gilman, 1991). The Gβγ heterodimer has proven to be the most potent negative regulator of AC type I and AC type VIII (E state), thus overriding the effect of forskolin, Ca^{2+}-calmodulin and $G_s\alpha$ (Sternweis and Robishaw, 1984; Sunahara and Taussig, 2002). Shortly after the disassociation from $G_i\alpha$, Gβγ activates AC II, AC IV and AC VII, but only in presence of $G_s\alpha$ (E to E*), suggesting Gβγ enhances the capacity of $G_s\alpha$ to activate AC (Sunahara and Taussig, 2002). The Gβγ binding site of AC type II has been located with peptide competition studies (Chen et al., 1995). This region flanks the β3-α3 loop and the first two-thirds of the α3-helix (Zhang et al., 1997)is the proposed binding site for Gβγ close to the $G_s\alpha$ site, but the binding sites do not overlap (Hurley, 1999).

Regulation by Forskolin

The dipterine forskolin from the Ayurvedic herb Coleus forskohlii is a powerful, hydrophobic, *in vitro* activator of all mACs, except AC IX and AC X. The forskolin binding site is located on one moiety and the substrate binding catalytic site is located on the other moiety of the enzyme (Figure 1.5) (Tesmer et al., 1997). Forskolin binds in a hydrophobic crevice located in the interface of the cyclase homology domain (Liu et al., 1997; Sunahara and Taussig, 2002; Tesmer et al., 1997). Forskolin activates adenylyl cyclases (E to E*) by gluing the two catalytic domains C1 and C2 together and making hydrophobic and hydrogen bonding interactions (Zhang et al., 1997). The AC residues binding forskolin are conserved in all types, with minor differences in AC type IX (Hurley, 1999). Type IX is not activated by forskolin, since the forskolin binding residues Ser and Leu are altered. Site-directed mutagenesis of Ala to Ser and Tyr to Leu, results in an activation of AC type IX with forskolin (Yan et al., 1998).

In contrast to mammalian ACs, biochemical assays *in vitro* have shown that adenylyl cyclases from *M. tuberculosis* are not activated by forskolin (Castro, 2004; Linder), probably due to the substitution of the forskolin binding residue Ser to Asp in mycobacterial ACs.

Regulation by P-site inhibitors

A class of adenosine analogs, known as P-site inhibitors, inhibit the activity of adenylyl cyclases, without competing with ATP for binding (Dessauer and Gilman, 1997; Florio and Ross, 1983; Londos and Wolff, 1977). Most of the P-site inhibitors lack one or more hydroxyl groups with respect to the ribose. The most potent are polyphosphorylated at the 3'-position, lack a 2'-hydroxyl group (Desaubry and Johnson, 1998; Desaubry et al., 1996a; Desaubry et al., 1996b), and are analogs of cAMP (Desaubry et al., 1996a). P-site inhibitors inhibit the enzymes by binding to the conformation of the adenylyl cyclases that strongly resembles the posttransition state or product bound state (E**:PPi). The inhibitor binding is enhanced by the presence of pyrophosphate (Dessauer et al., 1999). It has been found that all AC isoforms, except the soluble AC X are inhibited by P-site inhibitors (Sunahara and Taussig, 2002). The soluble ACs from *M. tuberculosis* are not inhibited by P-site inhibitors. Canonical substrate binding residues vary in mammalian and mycobacterial ACs, which might give differences in potency of P-site inhibitors and possibly different catalytic mechanisms (Linder and Schultz, 2003). This could be important for drug development of mycobacterial ACs.

Regulation by Pyrophosphate

Adenylyl cyclase converts ATP to pyrophosphate and cyclic AMP. It was revealed that the rate-limiting step is the release of pyrophosphate (Dessauer and Gilman, 1997). Therefore, high concentrations of pyrophosphate have been used to force the enzyme into the product-bound conformation and hinder ATP binding. Additionally, compounds such as foscarnet or phosphonoformic acid mimic pyrophosphate and are as well inhibiting the activity of AC (Kudlacek et al., 2001).

Regulation by Ca^{2+}-calmodulin

Ca^{2+}-calmodulin stimulates activity of AC I, AC VIII and probably AC III (Cali et al., 1994; Choi et al., 1992; Krupinski et al., 1989). Some agents, elevating Ca^{2+} levels in the cells are not able to activate Ca^{2+} sensitive ACs. In contrast, voltage-gated Ca^{2+} channels stimulated by the change in the voltage across the membranes are effective in activating these AC isoforms (Fagan et al., 2000; Gu and Cooper, 2000). In conclusion, Ca^{2+}-sensitive ACs detect and respond to capacitative cation entry rather than global cytosolic cation concentrations (Gu and Cooper,

2000). Non-physiological concentrations of Ca^{2+} (mM) inhibit all ACs, but AC V and AC VI are inhibited by micromolar Ca^{2+} concentrations (Cooper et al., 1998).

Regulation by Posttranslational Modification

Posttranslational modification such as glycosylation, phosphorylation and S-nitrosylation are regulating the activity of ACs. Protein kinases phosphorylate ACs, mainly resulting in an inhibitory effect. The cyclic–AMP-dependent protein kinase A (PKA) inhibits AC V and AC VI through a negative feedback mechanism (Chen et al., 1997; Iwami et al., 1995).

Regulation by Bicarbonate and Ca^{2+}

The soluble mammalian AC type X (Buck et al., 1999) is not regulated by G proteins, hormones, P-site inhibitors or forskolin, but strongly activated by bicarbonate and Ca^{2+}. AC X comprises a catalytic domain at the N-terminus and a domain at the C-terminus with an unknown function (Sunahara and Taussig, 2002). Based on amino acid analysis, AC type X has the highest similarity to the ACs from cyanobacteria. Furthermore, the cyanobacterial ACs are regulated by bicarbonate (Chen et al., 2000) and it has been suggested that AC type X functions as a metabolic sensor, resembling cyanobacterial ACs (Zippin et al., 2001).

Regulation by Other Proteins

Recently, several proteins have been identified to interact directly with adenylyl cyclases. They were identified through yeast two-hybrid and copurification experiments using the catalytic core as baits. The biological significance of the adenylyl cyclase interaction partners has yet to be determined. The protein associated with Myc (PAM) is a very potent inhibitor of AC type I and V (Scholich et al., 2001) and the *E. coli* protein cis-trans peptidylprolyl isomerase (PPIase) SlyD is interacting and inhibiting AC type VII (Yan et al., 2001). The interaction of the RGS2 protein and AC III is of higher importance (Sinnarajah et al., 2001), since the RGS proteins are known to increase the GTPase activity of the heterotrimeric G proteins $G\alpha_i$ and $G\alpha_q$ (Ingi et al., 1998; Kehrl and Sinnarajah, 2002).

1.3.c Regulation of ACs from *M. tuberculosis*

There is a great interest in adenylyl cyclases from *Mycobacterium tuberculosis* (*Mt*ACs), mostly since regulatory processes involving cAMP seem important in this human pathogen as derived from the high number of predicted ACs, compared to other organisms of similar genome size. 15 putative ACs have been found in the *M. tuberculosis* genome of 4400 genes (Cole et al., 1998) compared to the *E. coli* genome of 4300 genes containing only one AC, and the more complex organism of *Homo sapiens* comprising only 10 ACs in a genome of 28 000 genes. All 15 putative ACs from *M. tuberculosis* belong to class III (Table 1.2), and one of the adenylyl cyclases is receptor-type membrane bound, five are integral membrane proteins and nine are cytoplasmic. So far, nine ACs from *M. tuberculosis* have shown to posses *in vitro* AC activity e.g Rv1318c, Rv1319c, Rv1320c, Rv3645, Rv2212, Rv1900c (Linder et al., 2004; Motaal, 2006; Sinha et al., 2005)Rv1625c, Rv1264 (Guo et al., 2001; Linder et al., 2002; Reddy et al., 2001) and Rv0386 presented in this thesis (Castro et al., 2005) (Table 1.2). The high number of ACs in *M. tuberculosis* suggests that the pathogen has the ability to sense and respond to extracellular and intracellular signals through cAMP and that cAMP signalling is of high importance in the pathogenic agent (McCue et al., 2000).

Regulation of Integral Membrane, Receptor-Type and Soluble *Mt*ACs

In contrast to eukaryotic ACs the prokaryotic ACs are not regulated via protein kinase C and G-protein coupled receptor (GPCR)- G-protein signal transduction pathways, (Cooper, 2003) since the GPCR-G-protein pathway is not existing in prokaryotes. Interestingly, adenylyl cyclases Rv2435c and Rv1625c from *M. tuberculosis* are classified as integral membrane proteins, possibly responding to extracellular signals and regulated by GTPases (G proteins). Sequence alignment of Rv1625c illustrates a similarity to eukaryotic ACs and suggests a horizontal gene transfer from eukaryotes, as a survival advantage for the pathogen. Receptor-type membrane bound ACs resemble the eukaryotic receptor guanylyl cyclases and are supposed to sense extracellular signals (McCue et al., 2000). From the genome sequences of prokaryotic soluble ACs, several regulatory domains directly attached to the catalytic domains have been identified. These regulatory domains and intracellular signals seem to regulate the catalytic domains of prokaryotic soluble ACs (Linder and Schultz, 2003). Additionally, the genome sequence of *M.*

tuberculosis H37Rv revealed eukaryotic–like protein kinases (Cole et al., 1998) which could be involved in intracellular signalling pathways mediated by cAMP.

Class IIIa	Class IIIb	Class IIIc
Rv1625c	Rv1318c	Rv0386
Rv2435c	Rv1319c	Rv891c
	Rv1320c	Rv1358
	Rv3645	Rv1359
		Rv1647
		Rv1264
		Rv1900c
		Rv2212
		Rv2488c

Table 1.2: Classification of predicted adenylyl cyclases from *M. tuberculosis*. AC activity could be shown *in vitro* for ACs highlighted in grey. All 15 ACs are categorised into class IIIa-c according to (Linder and Schultz, 2003)

All class III adenylyl cyclases from *M. tuberculosis* comprise a cyclase homology domain on either the C-terminus or N-terminus with additional one or more putative regulatory domains (Figure 1.6). These include a helix-turn-helix DNA binding domain (HTH), an α/β-hydrolase-fold, transmembrane helices, histidine kinase domains (HisK), HAMP domain, GAF domain, PAS domain etc (Linder and Schultz, 2003). For class III AC domains, not all regulatory processes are understood in full detail.

Catalytic centre

In contrast to mammalian heterodimeric ACs comprising one catalytic centre and a forskolin binding site, class III adenylyl cyclases from bacteria and protozoa form homodimeric ACs with two catalytic centres. The catalytic residues from both cyclase homology domains form the catalytic centres, located in the interface of the dimer. Class III cyclase homology domains comprise six conserved substrate binding residues. These residues are non-conserved in a number of class III adenylyl cyclases from prokaryotic organisms such as *M. tuberculosis* (Table 1.3) (Linder and Schultz, 2003). Most of the noncanonical cyclase homology domains still need to be examined in detail (Castro, 2004). However, ACs Rv1264, Rv2212 and Rv1647 from *M. tuberculosis* comprise the same canonical residues as previously described for mammalian ACs.
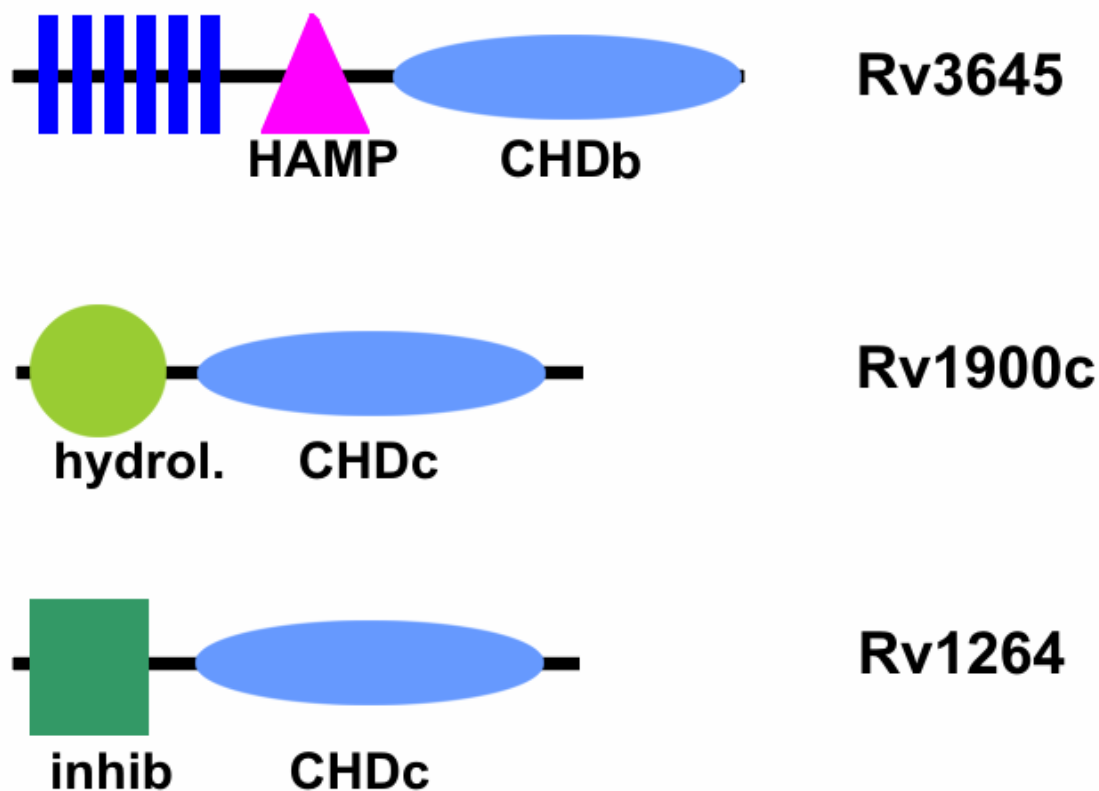
Figure 1.6: Schematic overview of examples of class III AC domain organisation. Represented are catalytic domains and a variety of regulatory domains from *M. tuberculosis*. The cyclase homology domains (CHD) are indicated by subscripts according to the subclassification IIIa-d. Dark blue bars correspond to predicted transmembrane helices, HAMP: tandemly arranged amphoteric α-helices present in histidine kinases, adenylyl cyclases, methyl-accepting chemotaxis proteins and phosphatases, hydrol: α/β-hydrolase-fold, inhib: autoinhibitory domain. Figure according to (Linder and Schultz, 2003).

Adenylyl Cyclases	Metal binding		Purine binding		Transition-state stabilisation	
Rv1625c	D	D	K	D	N	R
Rv2435c	D	N	R	D	S	Q
Rv1318c	D	D	K	T	N	R
Rv1319c	D	D	K	T	N	R
Rv1320c	D	D	K	T	N	R
Rv3645	D	D	K	T	N	R
Rv1647	D	D	K	D	N	R
Rv1264	D	D	K	D	N	R
Rv2212	D	D	K	D	N	R
Rv0386	D	D	Q	N	N	R
Rv891c	D	D	R	L	N	R
Rv1900c	D	D	N	D	H	R
Rv2488c	D	D	Q	N	D	R
Rv1358	D	S	Q	T	N	C
Rv1359	G	D	Q	I	N	E

Table 1.3: Overview of the six amino acids that form the catalytic centre of class III adenylyl cyclases from *M. tuberculosis*. Deviations of the conserved catalytic residues are coloured in cyan.

Regulation by Additional Domains

AC Rv3645 from *M. tuberculosis* comprises the cyclase homology domain at the C terminus in addition to two regulatory domains at the N terminus (Figure 1.6). Biochemical assays have shown that the HAMP domain fused to the cyclase homology domain increased the activity 21-fold. The HAMP domain is a direct modulator of the AC activity. This suggests that HAMP domains on their own have regulatory functions and in addition function as signal transducers (Linder et al., 2004). The numerous additional regulatory domains could increase the activity of low activity ACs from *M. tuberculosis in vitro*.

Regulation by Bicarbonate

Recently, it was demonstrated that Rv1319c class III AC containing the purine binding residues Lys and Thr (Table 1.3), was responsive to bicarbonate (Cann et al., 2003). Lys is believed to coordinate the bicarbonate in the catalytic cleft. In contrast, mycobacterial ACs like Rv1264 and Rv1625c having an Asp instead of Thr did not show any activation by bicarbonate. To date, all ACs regulated by HCO_3^- contain a Thr residue (Cann et al., 2003; Chen et al., 2000) and the unresponsive ones instead have Asp (Chen et al., 2000). This suggests that a large number of prokaryotic class III ACs, namely the ones with Thr at this position could be activated by HCO_3^-

(**Table 1.3**). Additionally, a novel mechanism is also strongly indicated. Bicarbonate is ubiquitous in the extracellular, intracellular milieu and is important in the prokaryotic and eukaryotic systems. However, it is unknown by which mechanism the organisms detect and respond to HCO_3^- (Cann et al., 2003).

Regulation by Mn^{2+}

The physiological ATP concentration in *Mycobacterium tuberculosis* is reported to be 1 mM (Reddy et al., 2001), but the intracellular concentrations of the cations Mn^{2+} and Mg^{2+} are unknown. Biochemical assays *in vitro* show that many prokaryotic cyclase homology domains require the cofactor Mn^{2+} for activity (Coudart-Cavalli et al., 1997; Kasahara et al., 2001). However, a few bacterial and mammalian ACs use Mg^{2+} as a cofactor, which is most likely the physiologically more important cation (Dessauer et al., 1997). It has been reported that ACs from *M. tuberculosis* require millimolar concentrations of Mn^{2+} for activity *in vitro* (Guo et al., 2001; Linder et al., 2004; Linder and Schultz, 2003; Shenoy et al., 2004). Mn^{2+} is possibly physiologically relevant in *M. tuberculosis* as the genome comprises a Mn^{2+} transporter that may be responsible for Mn^{2+} intracellular concentrations (Cole et al., 1998; Linder et al., 2004; Reddy et al., 2001). It has been reported that macrophages can import Mn^{2+} and thus provide elevated levels of Mn^{2+} for activation of resident mycobacterial ACs. The increased concentration of cAMP is secreted into the macrophages, resulting in inhibition of several phagocyte-related processes including phagosomes-lysosome fusion (Lowrie et al., 1975) and propagation of *M. tuberculosis* in the host macrophages.

1.4 Structural Characterisation of Class III ACs

1.4.a Structures of Mammalian ACs

Crystal Structures

The first X-ray structure of an adenylyl cyclase was published in 1997, containing the catalytic domain of rat type II AC C2 homodimer in complex with its potent activator forskolin (Zhang et al., 1997). Mammalian homodimers are only formed under *in vitro* conditions and are inactive, since the C1 domain is lacking and with it some of the catalytic residues essential for binding of ATP. The typical CHD is structurally described by a three-layer α/β sandwich: a central β-sheet with eight β-strands surrounded by five α-helices (Figure 1.7). Furthermore, the structure shows the two catalytic domains in a dimeric assembly, including forskolin in the interface of the dimer. The AC dimer is in a head-to-tail conformation, also called wreath-like.

Subsequently, the crystal structure of the activated Gsα·GTPγS (Sunahara et al., 1997b) in complex with AC rat IIC2 · canine VC1 · forskolin was determined (Tesmer et al., 1997). Additionally, crystal structures were determined of the apo form as well as of complexes with a range of ATP analogs, divalent metal ions and P-site inhibitors (Tesmer et al., 2000; Tesmer et al., 1997; Tesmer et al., 1999).

Active Site

The C1 and C2 domains form the catalytic core of mammalian ACs. The active site is located in a deep cleft at the interface between the C1 and C2 domains, and its asymmetry is essential for the function of ACs. The interface occupies the active site at one end and the binding sites of forskolin and Gα_s at the other (Figure 1.5). Forskolin and Gα_s · GTP appear to be responsible for the orientation of the C1 and C2 domains and thereby the formation of the activated state. Upon binding of Gα_s and forskolin the C1 domain is rotated 7° around the C2 domain, believed to position the active site for catalysis (Tesmer et al., 1997; Tesmer et al., 1999; Zhang et al., 1997). The active site (Figure 1.8) includes the α4, α5 helices of the C2 domain and β1, α1, β2-β3 turn of the C1 domain. The catalytic residues are conserved in mammalian AC isoforms. Lys938 and

Asp1018 are binding to the adenine ring and correspond to Glu and Cys respectively in guanylyl cyclases. With site-directed mutagenesis the substrate specificity for ACs can be changed from ATP to GTP (Sunahara et al., 1998), although GTP cannot be converted to cGMP. Due to a contact of the main chain carbonyl of Ile1019 to adenine–N6 which disfavours guanine. In contrast, the conversion of GC to AC has been demonstrated, with mutagenesis leading to successful conversion of ATP to cAMP (Sunahara et al., 1998; Tucker et al., 1998). The residues coordinating the binding of the ribose and triphosphate of the substrate are conserved in all ACs and GCs. Cyclase activity is severely impaired when these are substituted to non-conserved residues. The α-, β-, and γ-phosphate of the nucleotide are coordinated by Arg1029, Lys1065 in AC type IIC2 and Arg484 in AC type VC1. The residues Asp440 and Asp396 in AC type VC1 are highly conserved and assist in coordinating the phosphates by coupling to Mg^{2+} and Mn^{2+} that stabilise the α-phosphate during catalysis (Tesmer et al., 1999).
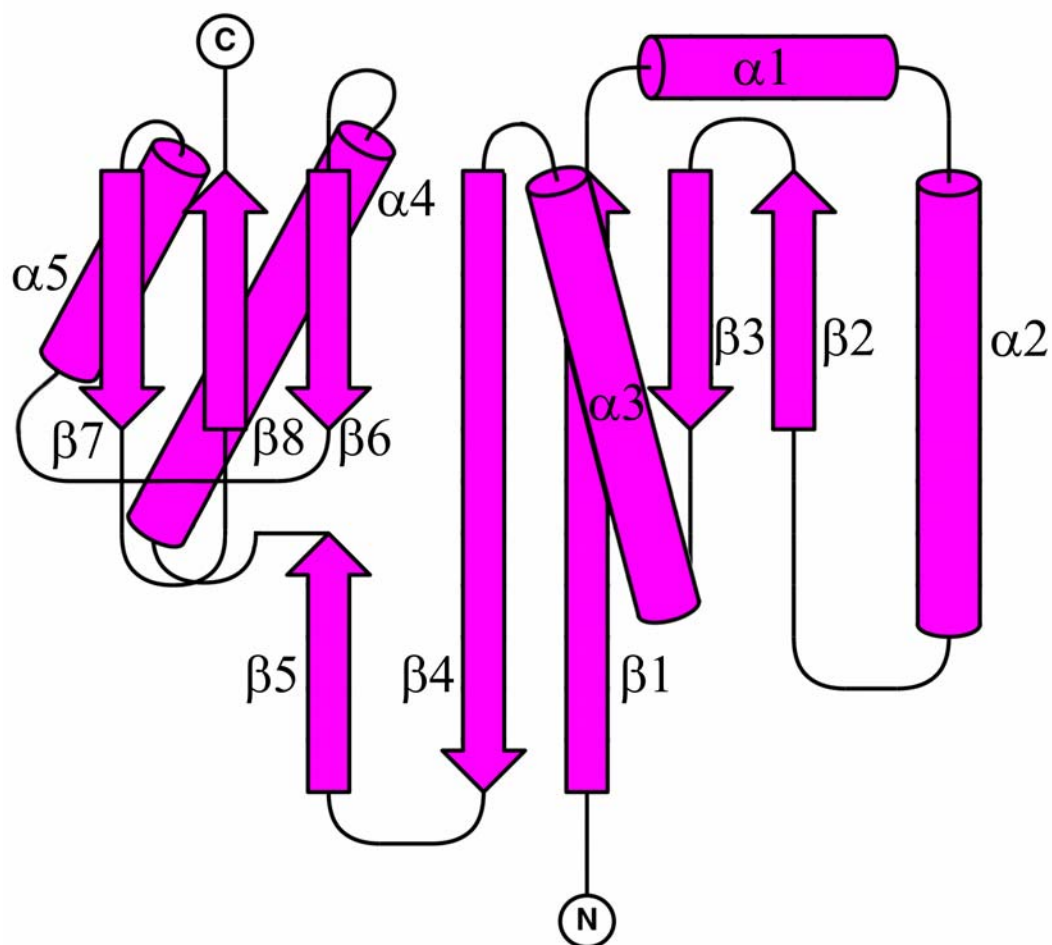
Figure 1.7: Topology of the adenylyl cyclase homology domain, consisting of a central β-sheet enclosed by five α-helices. This picture is based on the PDB entry code: 1AB8.
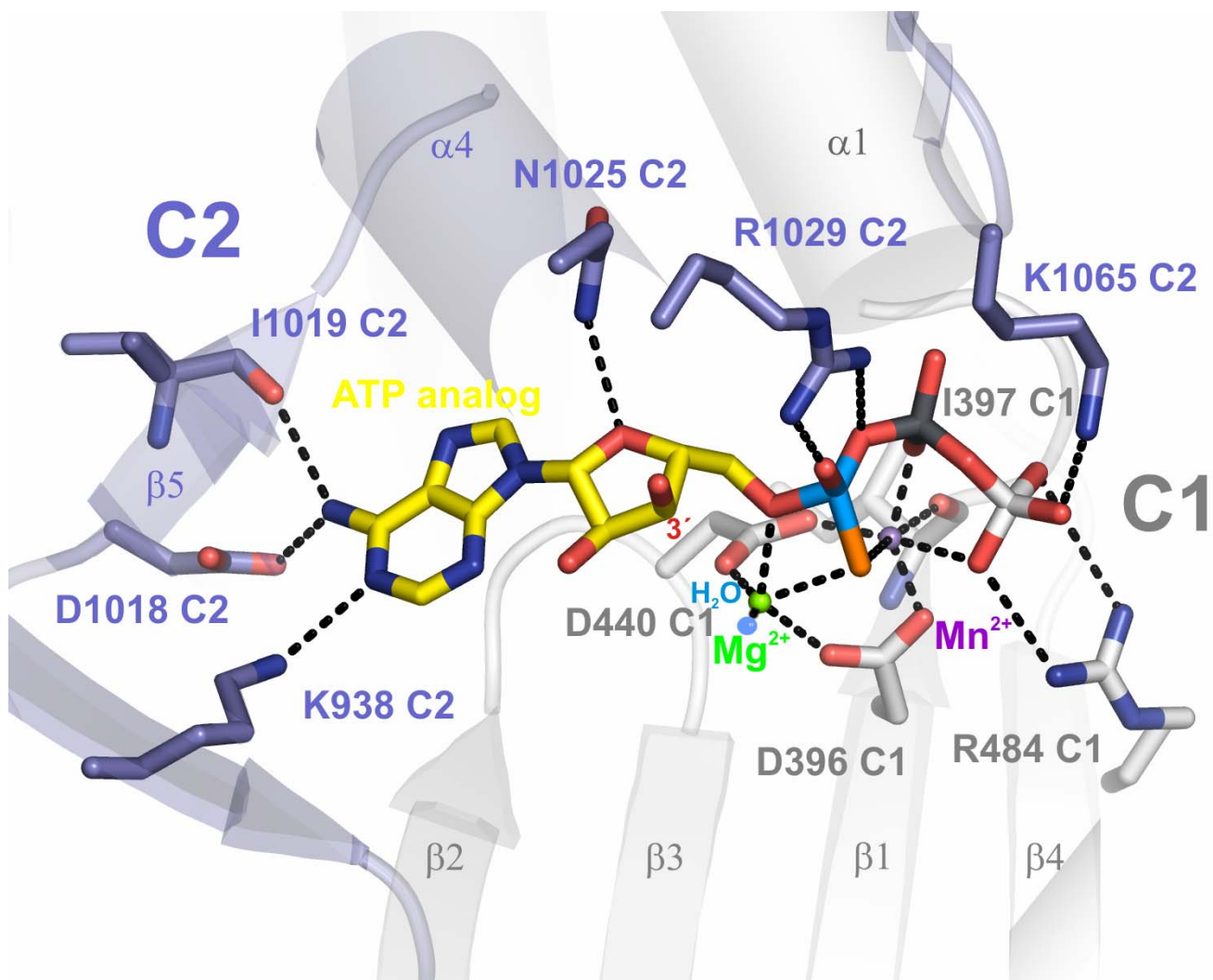
Figure 1.8: The nucleotide binding site in mammalian AC. An ATP analog rendered in yellow is located in the interface of canine V C1 (grey) and rat II C2 (slate) (PDB:1CJK). Lys938, Asp1018 and main chain of Ile1019 belong to C2 and are specific adenine binding residues. Asn1025 C2 and Arg1029 C2 stabilise the transition state and the two divalent metals (spheres in green; Mg^{2+} and violet; Mn^{2+}) are coordinated by Asp440 C1, Asp396 C1, main chain of Ile397 C1 and water (sphere in blue). Additionally, Lys1065 C2 and Arg484 C1 bind and stabilise the γ-phosphate. The 3´ hydroxyl of ribose is depicted in red and named 3´.

Catalytic Mechanism

From the first heterodimeric AC crystal structure ($G\alpha_s \cdot$ VC1 $\cdot$ IIC2 $\cdot$ 2'-deoxy-3'-AMP $\cdot$ PPi) the catalytic mechanism for the synthesis of cAMP was not obvious. Therefore, Sprang and co-workers decided to determine four crystal structures of the AC catalytic core with two different ATP analogs and various divalent metal ions (Tesmer et al., 1999).

The active site of ACs resembles the active site of DNA polymerases, and it was therefore initiative to propose the same two-metal ion mechanism, based on the assumption of development from a common ancestor. The core of three out of four families of DNA polymerases contain a palm domain composed of the $\beta\alpha\beta\beta\alpha\beta$ motif. This motif is present in the AC C1 and C2 domains (Artymiuk et al., 1997; Steitz et al., 1993; Steitz et al., 1994). Based on the various activated AC structures and a comparison with DNA polymerases, a model of the catalytic mechanism of adenylyl cyclases was established.

A hydrogen bond is expected to be established between Asn1025 and the O4' oxygen of ATP (Figure 1.8). The nucleophilicity of the 3' hydroxyl ATP is increased by Mg^{2+}, acting as a Lewis acid. The distance between Mg^{2+} and the 3' hydroxyl of ribose is 2 Å. Then, the 3' hydroxyl group is positioned for an in-line attack on the α-phosphate. This reaction generates cAMP and PPi. In Figure 1.8 the 3′hydroxyl of ribose is positioned away from Mg^{2+}, but in the nucleotide bound state the 3′hydroxyl would be positioned towards Mg^{2+} and result in catalysis. Upon substrate binding, both ACs and DNA polymerases induce a transition from an open to closed conformation (Doublie et al., 1999; Tesmer and Sprang, 1998). The transition causes the breakdown of the fingers on the palm domain of DNA polymerases. For ACs, it has been noted that the α1 helix shifts 2 Å into the active site, upon substrate binding. A number of loops around ATP are as well broken down in ACs, to form the active site. The break down of these loops brings basic residues into the active site which stabilise the triphosphate of the substrate, the pentavalent phosphate transition state of the reaction, and the pyrophosphate leaving group. It has been suggested that $G_s\alpha$ induces the break down of loops around ATP to form a proper active site. $G_i\alpha$ should on the other hand hinder the break down and stabilises thereby an open, inactive conformation (Tesmer et al., 1999).

1.4.b AC Structures from *Trypanosoma brucei*

Cyclase homology domains from *Trypanosoma brucei* have been structurally determined in a monomeric state. Since the AC catalytic centres are situated at the interface of two monomers, *T. brucei* ACs cannot perform catalysis and therefore are in an inactive state. The structure has a high structural similarity to other CHDs, with the exception of an inserted Δ-subdomain positioned close to the active site. The Δ-subdomain comprises 36 residues, is inserted in the α3-β4 loop and might be an allosterically regulatory protein binding site. It is proposed that activation of these monomeric ACs occurs with a ligand-induced dimerisation, similar to membrane bound guanylyl cyclases (Bieger and Essen, 2001).

1.5 AC Rv0386 from *M. tuberculosis*

1.5.a Overview and Domain Organisation

In 1998 Cole and co-workers published the whole genome of the aetiological agent, *Mycobacterium tuberculosis* H37Rv. Five genes were annotated as putative ACs including Rv1625c, Rv1318c, Rv1319c, Rv1320c and Rv1264. The name of the gene Rv0386 was for the first time noted in this genome project and assigned as a transcriptional regulator (Cole et al., 1998). Later, 15 putative ACs were detected with computational methods, among those Rv0386 (McCue et al., 2000). Rv0386 is characterised as a cytoplasmic class IIIc AC, with shortened dimerisation arm in the cyclase homology domain and acting as a guanylyl cyclase converting GTP to cGMP (Castro et al., 2005).

In (Gamieldien et al., 2002) Rv0386 is placed between 19 other genes from *M. tuberculosis* that may have been acquired by horizontal gene transfer from eukaryotes. This possibly had survival advantages for the ancient mycobacterium, before it developed into an eukaryotic pathogen (McCue et al., 2000).

Rv0386 comprises four domains with the molecular weight of 117 kD. The AC CHD domain is found at the N-terminus of the protein, followed by an AAA-ATPase domain, a tetratricopeptide repeat (TPR) domain and a luxR helix-turn-helix (HTH) DNA binding

domain at the C-terminus (Figure 1.9). Interestingly, the putative AC genes Rv1358 and Rv2488 from *M. tuberculosis* show an identical domain organisation as Rv0386.
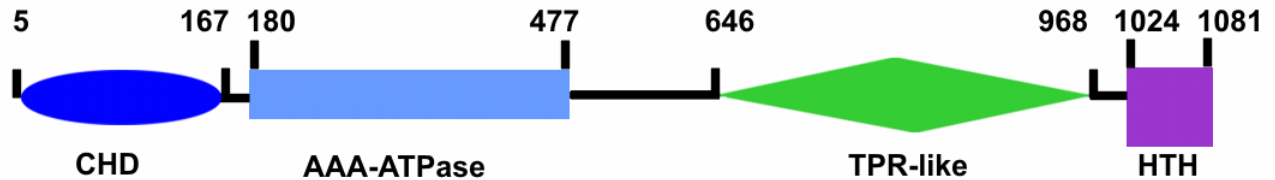
Figure 1.9: Domain organisation of Rv0386. The cyclase homology domain (CHD) in dark blue, the AAA-ATPase domain in light blue, the TPR-like domain in green and the luxR helix-turn-helix (HTH) DNA binding domain in violet.

The AAA-ATPase domain is responsible for ATP binding and hydrolysis and has 40 % similarity to several prokaryotic transcriptional regulators, like AfsR from *Streptomyces coelicolor* (Horinouchi et al., 1990). The TPR domain functions as protein-protein recognition module in e.g. intracellular signalling and transcription control (Cliff et al., 2005). The HTH DNA binding domain is involved in transcriptional regulation and present in most prokaryotic and some eukaryotic transcription factors (Aravind et al., 2005). The regulatory domains of AC Rv0386 are suggested to participate in transcription (Castro, 2004). The presence of the helix-turn-helix DNA binding domain strongly proposes that the adenylyl cyclase Rv0386 is connected to a transcriptional effector.

The regulatory function has not been tested *in vitro* due to folding problems of the holoenzymes, including Rv0386. How the AC domain is regulated via the additional domains is something that needs to be elucidated in the future.

1.6 Tuberculosis

1.6.a History and Epidemiology

Tuberculosis (TB) is accounted for about 1.7 million deaths in 2004 (World Health Organisation Tuberculosis Fact Sheet; http://www.who.int/topics/tuberculosis/en/, #168). This life threatening disease is claiming more lives than acquired immune deficiency syndrome (AIDS) and malaria combined. The fear is that TB will kill 40 million people within the next 20 years, if a vaccine is not found (Carrington, 1999). The disease is caused by the pathogen *Mycobacterium tuberculosis*. The pathogen is latent for a couple of years and the risk of developing TB is highest up to two years after infection. AIDS and human immunodeficiency virus (HIV) increase the risk more than 100 times for developing the active state of TB. One third of the world's population has latent TB, but only 5-10 % develops the active disease and can transmit the bacteria.

Tuberculosis is also known as wasting disease, consumption and the white plague (Cole et al., 1998). Until the mid of the 19th century, it was believed that TB was hereditary. The finding that TB is contagious came in 1865, and was discovered by the French surgeon Jean-Antoine Villemin. In 1882 the German scientist Robert Koch discovered the gram positive bacteria causing the disease, subsequently receiving the Nobel Prize in physiology or medicine in 1905 for his discovery. But, it would take until 1943 for the first drug, streptomycin to be identified, by the American scientist Selman Waksman. Within the next ten years two further drugs against *M. tuberculosis* were discovered. This lead to an increased rate of people cured from TB and a decrease in mortality rates. However, in the 1980s it was noticed that the rate of TB infected people was rising. TB is a life threatening disease if not treated.

Today, the highest TB mortality regions are in Africa. This is correlated with the high rate of HIV, which increases the risk of infection and death by TB. South-East Asia also has high TB mortality rates. The former Soviet Union countries had a rise in TB in the beginning of the 1990s, but after 2001 it declined {http://www.who.int/mediacentre/factsheets/fs104/en/index.html, #168}.

1.6.b Treatment

Tuberculosis was until the 1940s untreatable and people with TB were sent to sanatoriums, special rest homes for treatment. Later, it was discovered that antibiotics such as isoniazid, rifampicin, pyrazinamid, ethambutol and streptomycin were very efficient in curing TB. In many countries, a vaccine program included the Bacillus Calmette-Guerin (BCG) vaccine, with 80% protection of TB. Tuberculosis is usually treated with a standard six months therapy, despite of HIV-infection using a combination cocktail of four drugs. Rifampicin and isoniazid are the most potent antibiotics. The standard therapy is two months medication of isoniazid, rifampicin, pyrazinamid and ethambutol with additional four months of isoniazid and rifampicin. Streptomycin is only used in case of antibiotic resistance or toxic/negative side effects using the other antibiotics. During the therapy, sputum is weekly controlled for presence of acid-fast staining bacilli by microscopy and cultural identification of *M. tuberculosis* in media. If the sputum cultures are positive after two months treatment, the therapy is extended to duration of nine months or more (Lange et al., 2005).

It is essential that the patient makes a complete therapy, with all four antibiotics taken regularly, once per day for a period of at least six months. If not, the risk of developing multidrug resistant tuberculosis is increased. Recently, it was discovered that the former Soviet Union countries have new multidrug-resistant (MDR) *M. tuberculosis* strains resistant to the most powerful anti-TB drugs i.e. isoniazid and rifampicin. More than 50 million people worldwide are infected with MDR strains and 15% of all TB cases are caused by MDR strains of *M. tuberculosis* (Kaufmann, 2002). Drug resistant TB can be treated with chemotherapy, but this is an expensive therapy. Therefore, the development of new drugs is of high importance.

1.6.c Pathogenesis

Tuberculosis is transferred from person to person through air. Only persons with active TB can spread the disease, but not all exposed to *M. tuberculosis* become infected. There are three factors that determine if TB is transmitted: 1) How contagious the TB patient is. 2) The length of the exposure to *M. tuberculosis*. 3) The environment in which the exposure took place.

The main port of the entry of *M. tuberculosis* and the importance of disease manifestation is represented by the lung. The pathogen is inhaled into the lungs and travel to the alveolar (small air sacs of the lung) macrophages, where it is engulfed but not eliminated. In the draining lymph nodes specific T cells are stimulated, which induce the bacterial control in small granulomatous lesions of the lung. Here, *M. tuberculosis* is not completely eradicated (Figure 1.10). Instead, a balance between the host defence and the bacterial persistence develops, that might be lifelong. This represents the latent state of TB and does not have to develop into the active state of the disease (Kaufmann, 2001; Manabe and Bishai, 2000). Drugs cannot eliminate the pathogen in the latent state, but can decrease the risk of developing the active state (http://www.infektionsbiologie.ch). Once the active state of TB develops and remains untreated, 50% of the patients have a fatal outcome (Kaufmann, 2001; Manabe and Bishai, 2000). The active state of the disease develops from the latent state when the hosts' immune system is weakened due to HIV infection, immune suppressive medication through cancer, diabetes, rheumatoid arthritis, malnourishment or great age. A person with the active disease infects up to 15 persons per year (Kaufmann, 2000).

1.6.d Intracellular Lifestyle of *M. tuberculosis*

Macrophages are used as a residence for *M. tuberculosis* which has two important implications for the survival of the pathogen (Schaible et al., 1999). Firstly, cholesterol is used by the bacteria as a docking site, promoting receptor-ligand interactions (Gatfield and Pieters, 2000). The primary interaction of the surface receptors influences the fate of the pathogen in the macrophages. The interaction with toll-like receptors and immunoglobulin receptors (FcRs) stimulate host-defence mechanisms. In contrast, interaction with complement receptors promotes the survival of the pathogen (Armstrong and Hart, 1975; Brightbill et al., 1999; Schorey et al., 1997; Thoma-Uszynski et al., 2001). The second implication is the survival mechanism of *M. tuberculosis* within the phagosome. Other microbes are eliminated within phagosomes harsh environment (Schaible et al., 1999) but *M. tuberculosis* survives, as it is able to arrest the maturation of the phagosome at an early stage to prevent fusion of the phagosome with the lysosomes (Figure 1.11) (Armstrong and Hart, 1975; Ferrari et al., 1999; Harth and Horwitz, 1999; Russell et al., 1996).

It has been noted that the macrophage ingestion of *M. tuberculosis* results in an increase in cAMP levels and the fusion of lysosomes with the bacterium-containing phagosomes was impaired. The explanation is that the bacterium is defending itself from destruction by producing cAMP in the phagosomes (Lowrie et al., 1979; Lowrie et al., 1975). Recently, AC Rv1264 has been found to be active at pH 5.5 and inhibited at pH 7.5 *in vitro*. A suggested explanation is that Rv1264 is a pH sensor important for *M. tuberculosis* survival in the macrophages (Tews et al., 2005).
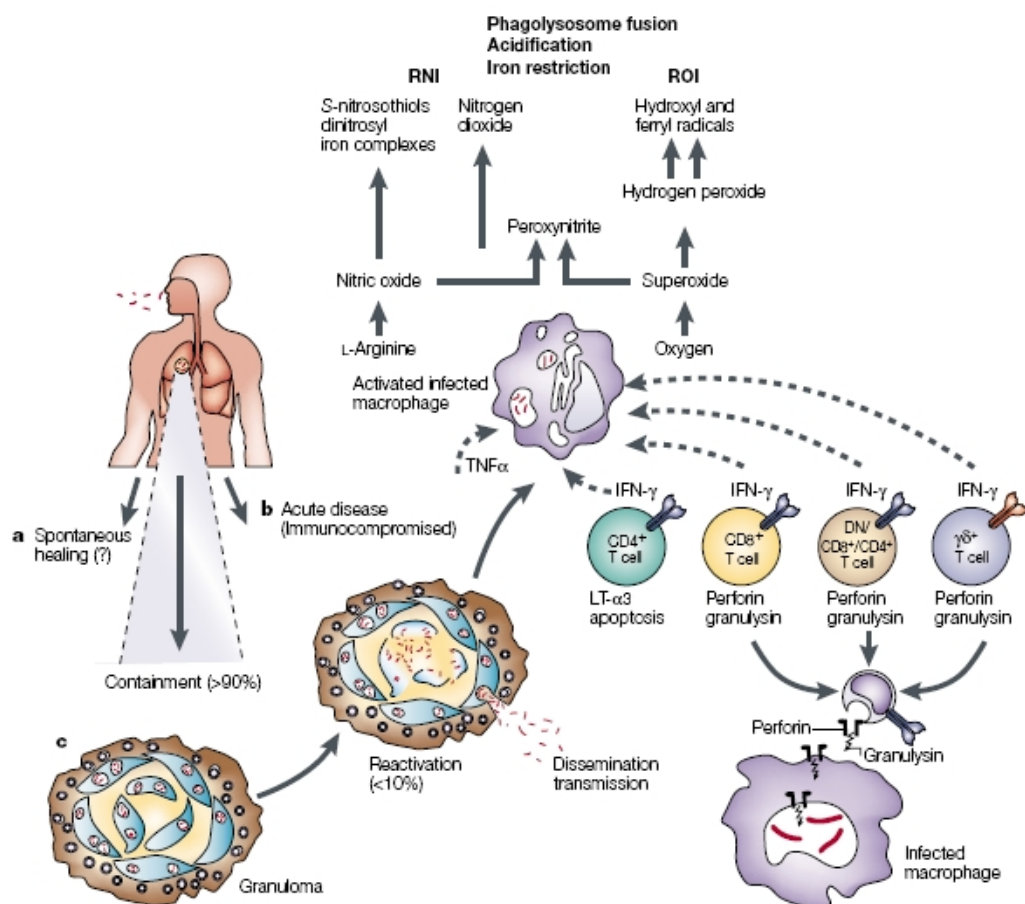
Figure 1.10: Infection of tuberculosis and host defence. **A)** Abortive infection of *M. tuberculosis* resulting in spontaneous healing. **B)** TB develops directly after infection. **C)** In most cases the pathogen is contained in the granulomas in a latent state, and later <10% of the infected, develop the active state of TB as a result of reactivation. Effector T cells including CD4[+], CD8[+], γδ[+] and macrophages participate in the control of TB. T cells produce the important macrophage activators: tumour-necrosis factor-α (TNF-α) and interferon-γ (IFN-γ). This results in phagosomal maturation and production of reactive oxygen intermediates (ROI), reactive nitrogen intermediates (RNI), lymphotoxin-α3 (LT-α3) involved in antimicrobial defence. Recently, it was shown that granulysin and perforin produced from T cells, directly kills *M. tuberculosis*. Picture taken from (Kaufmann, 2001)
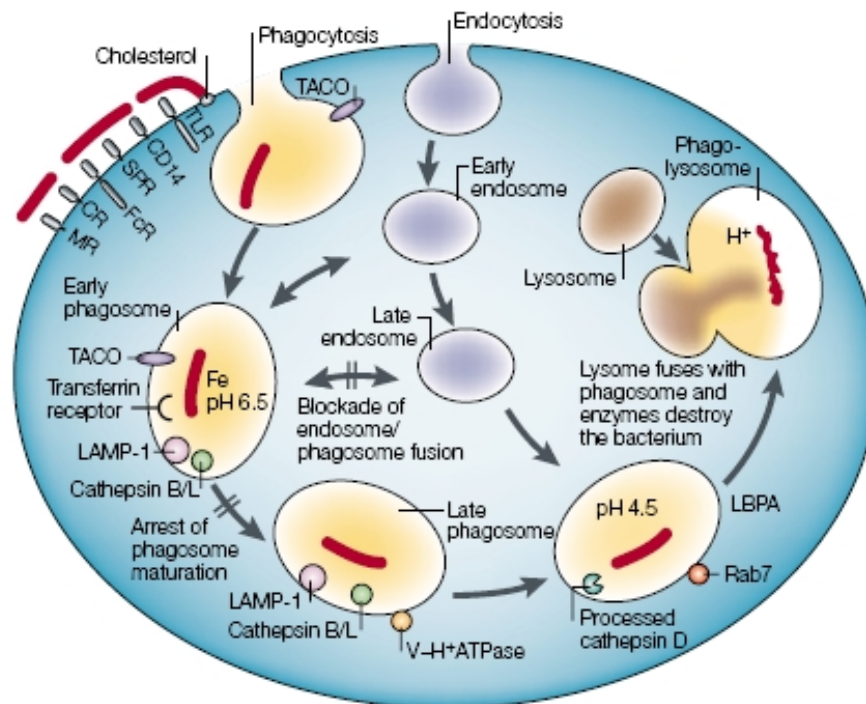
Figure 1.11: *Mycobacterium tuberculosis* intracellular lifestyle. This picture illustrates *M. tuberculosis* influence of the phagosome and endosome maturation process. At the early encounter between *M. tuberculosis* and macrophages a number of surface receptors are involved. The docking site between the pathogen and surface receptors is facilitated by cholesterol. Subsequently, the bacteria are engulfed and end up in the phagosome. The phagosome maturation is arrested at an early stage, caused by TACO. The phagosomal acidification and the fusion with the endosomal pathway are hindered by *M. tuberculosis*. In some cases, the phagosomes mature to form phagolysosomes. This occurs in activated macrophages, especially after stimulation by INF-γ. Then, the phagosomal and endosomal pathways mature until the last stage is reached where the phagosome and endosome merge with the lysosome. MR, mannose receptor; CR, complement receptor; FcR, receptor for the constant fragment of immunoglobulin; SPR, surfactant protein receptor; TLR, Toll-like receptor; TACO, tryptophane, aspartate-containing coat protein; LAMP-1, lysosomal-associated membrane protein 1; V–H+ATPase, vacuolar ATP-dependent proton pump; Rab7, member of the small GTPase family;LBPA, lysobiphosphatic acid. Taken from (Kaufmann, 2001).

1.7 Goal of this Study

15 adenylyl cyclases from *M. tuberculosis* are of high importance for survival of the pathogen in the human host. Biophysical and structural approach of mycobacterial ACs would result in a better understanding of this microbe causing the life threatening disease of tuberculosis and could consequently lead to new treatment. In the beginning of these studies the only available X-ray structures of AC class III cyclase homology domains were from mammals and *Trypanosoma brucei*. Rv0386 from *M. tuberculosis* is a multidomain protein, comprising an AC catalytic domain at the N-terminus, followed by an AAA-ATPase domain, TPR domain and a HTH DNA binding domain at the C-terminus. The dissimilarity of the substrate binding residues of Rv0386 compared to other ACs, and the promiscuous role choosing ATP and GTP as a substrate made this protein of high interest. Subsequently, this work was focussed on the class III cyclase homology domain of AC Rv0386 from *M. tuberculosis*.

The goals of this study were:
- Determination of the X-ray structure of the catalytic core of Rv0386.
- Examination of the modes of substrate binding and the catalytic mechanism.
- Comparison of the crystal structures of the catalytic domain of Rv0386 with homologous ACs.
- Validation of a newly found interface, (which became available during this thesis) with analytical ultracentrifugation, small angle X-ray scattering and site-directed mutagenesis.
- Understanding the regulation of AC Rv0386.
- Establishing buffer conditions to define the mode of oligomerisation for probing of different conformations of cyclase homology domain Rv0386.
- Characterisation and crystallisation of the CHD Rv0386 including the cofactor manganese and substrate analogs, hence this divalent cation probably improves the activity of the protein.

Chapter 2

MATERIALS AND METHODS

2.1 Cloning of the Catalytic Domain of Rv0386

2.1.a   N-His CHD of Rv0386

The catalytic domain of Rv0386 (residue 1-175) from *Mycobacterium tuberculosis* (GenBank Accession Number BX842573) is cloned between the 5´BamHI and 3´HindIII cleavage sites in the vector pQE30 (Qiagen) adding a N-terminal Met-Arg-Gly-Ser-His$_6$-Gly-Ser tag. The amino acid sequence of this construct is shown below (Figure 2.1).

```
                        1            11           21           31
MRGSHHHHHHGS  MSKLLPRG TV  TLLLADVEGS   TWLWETHPDD  MGAAVARLDK
41            51           61           71           81
AVSGVIAA HD   GVRPVEQG EG  DSFVLAFACA   SDAVAAALDL  QRARLAPIRL
91            101          111          121          131
RIGVHTGE VA   LRDEGNYA GP  TINRTARLRD   LAHGGQTV LS GVTESLVIDR
141           151          161          171
LPDKAWLV DL   GTHALRDLSR   PERVMQLC HP  ELRID*
```

Figure 2.1: The amino acid sequence of the N-His CHD of Rv0386 construct. The pQE30 vector and the residues of the catalytic domain of Rv0386 are depicted in black and blue respectively. The star is assigned for a stopcodon.

2.1.b C-His CHD of Rv0386

The catalytic domain (1-175) is cloned into the 5´BamHI and 3´BglII restriction sites of vector pQE60 (Qiagen) containing a N-terminal elongation of Met-Gly-Gly-Ser and a C-terminal Arg-Ser-His$_6$ tag (Figure 2.2 and Figure 2.3). Both constructs (C-His CHD of Rv0386 and N-His CHD of Rv0386) were made by Lucila I. Castro (Castro, 2004). These two constructs were available through collaboration with the group of Prof. J.E Schultz, Universität Tübingen, Germany.

```
                          BamHI
                         ~~~~~~~
             M  G  G  S   M  S  K  L  L  P  R  G  T  V  T  L  L  L  A  D  V  E  G  S  T  W  L  W  E  T  H  P  D  D ·
            ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      1     ATGGGAGGAT CCATGAGCAA GTTGCTGCCA CGGGGCACAG TGACATTGCT GTTGGCCGAC GTCGAGGGTT CCACCTGGCT GTGGGAGACC CATCCAGACG
            TACCCTCCTA GGTACTCGTT CAACGACGGT GCCCCGTGTC ACTGTAACGA CAACCGGCTG CAGCTCCCAA GGTGGACCGA CACCCTCTGG GTAGGTCTGC

            · M  G  A   A  V  A  R  L  D  K  A  V  S  G  V  I  A  A  H  D  G  V  R  P  V  E  Q  G  E  G  D  S  F ·
            ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    101     ACATGGGTGC TGCCGTGGCG CGCCTCGACA AAGCCGTGTC TGGTGTGATT GCCGCCCATG ACGGCGTACG CCCAGTCGAG CAGGGTGAGG GTGATAGCTT
            TGTACCCACG ACGGCACCGC GCGGAGCTGT TTCGGCACAG ACCACACTAA CGGCGGGTAC TGCCGCATGC GGGTCAGCTC GTCCCACTCC CACTATCGAA

            · V  L  A   F  A  C  A  S  D  A  V  A  A  A  L  D  L  Q  R  A  R  L  A  P  I  R  L  R  I  G  V  H  T
            ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    201     TGTCCTCGCG TTCGCCTGCG CGTCGGATGC CGTGGCCGCC GCGTTGGACT TGCAGCGAGC GCGGCTCGCA CCGATCCGGT TGCGCATAGG CGTGCACACC
            ACAGGAGCGC AAGCGGACGC GCAGCCTACG GCACCGGCGG CGCAACCTGA ACGTCGCTCG CGCCGAGCGT GGCTAGGCCA ACGCGTATCC GCACGTGTGG

             G  E  V  A   L  R  D  E  G  N  Y  A  G  P  T  I  N  R  T  A  R  L  R  D  L  A  H  G  G  Q  T  V  L  S ·
            ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    301     GGGGAGGTCG CGCTCCGCGA CGAAGGCAAC TATGCCGGTC CGACCATCAA CCGGACCGCG CGCCTGCGTG ACTTGGCGCA TGGGGGCCAG ACGGTGCTCT
            CCCCTCCAGC GCGAGGCGCT GCTTCCGTTG ATACGGCCAG GCTGGTAGTT GGCCTGGCGC GCGGACGCAC TGAACCGCGT ACCCCCGGTC TGCCACGAGA

            · G  V  T   E  S  L  V  I  D  R  L  P  D  K  A  W  L  V  D  L  G  T  H  A  L  R  D  L  S  R  P  E  R ·
            ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    401     CGGGCGTGAC CGAAAGCCTG GTCATCGATC GCCTCCCGGA CAAAGCATGG CTGGTTGACC TGGGGACGCA CGCGCTGCGG GATCTGTCGC GTCCGGAGCG
            GCCCGCACTG GCTTTCGGAC CAGTAGCTAG CGGAGGGCCT GTTTCGTACC GACCAACTGG ACCCCTGCGT GCGCGACGCC CTAGACAGCG CAGGCCTCGC
                                                                       BglII
                                                                      ~~~~~~~
            · V  M  Q   L  C  H  P  E  L  R  I  D  R  S  H  H  H  H  H  H  *
            ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    501     GGTAATGCAG CTGTGTCATC CCGAATTGCG TATCGATAGA TCTCATCACC ATCACCATCA CTAA
            CCATTACGTC GACACAGTAG GGCTTAACGC ATAGCTATCT AGAGTAGTGG TAGTGGTAGT GATT
```

Figure 2.2: Nucleotide and translated protein sequence of the C-His CHD of Rv0386 construct are colored in blue. The vector sequence is depicted in black. The catalytic domain of Rv0386 has been cloned into the pQE60 vector between the BamHI and BglII restriction sites. The figure has been created in Vector NTI.

*Bam*HI (123)

C-His CHD Rv0386

Ampicillin

*Bgl*II (654)

pQE60/C-His CHD Rv0386
3957 bp
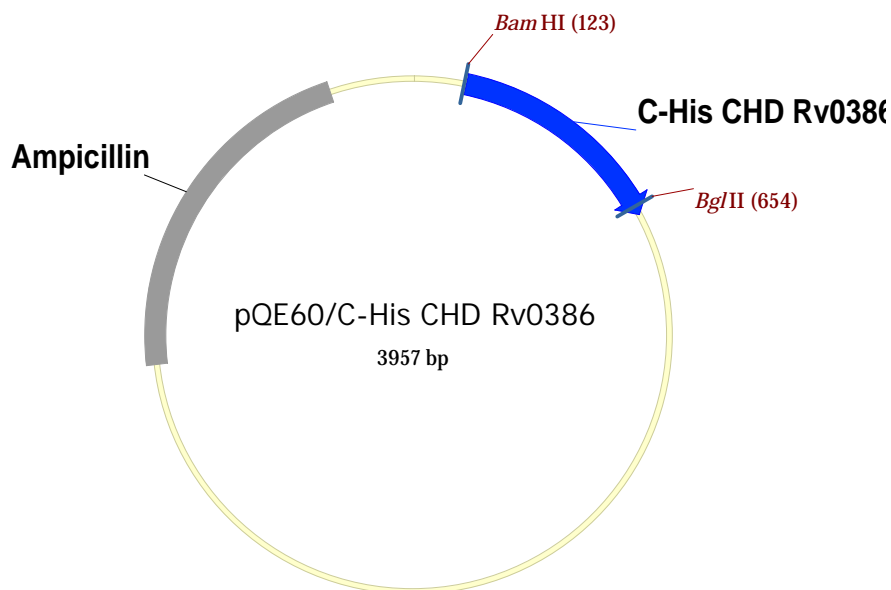
Figure 2.3: Overview of the circular plasmid map of the construct pQE60/C-His CHD Rv0386 residue 1-175. Vector NTI has been used for making the picture.

2.1.c Rv0386 CHD$_{(1-168)}$ C168S

In order to identify a construct suitable for crystallisation, the sequence of the catalytic domain of Rv0386 was aligned against the sequence of the crystallised catalytic domain of Rv1900c, both from *Mycobacterium tuberculosis*. The sequence of the catalytic domain of Rv1900c represents the resolved structure and therefore the construct of the catalytic domain of Rv0386 was designed accordingly. Mutation Cys168Ser was introduced to avoid possible formation of non native disulphide bridges.

The shorter construct containing residue 1-168 and the point mutation Cys168Ser was generated by the QuickChange XL site-directed mutagenesis kit (Stratagene, Germany). The DNA of the holoenzyme of Rv0386$_{(1-1085)}$ in pQE30 (kindly provided by Lucila I. Castro, Tübingen Universität, Germany) was used as a template and primers carrying the desired mutation (Table 2.1 and 2.3) were used for the PCR. 18 cycles of PCR were performed with the Pfu Turbo DNA Polymerase at the annealing temperature 60 °C. The elongation temperature was carried out at 68 °C for 14 minutes in order to amplify the ca 7 kb DNA template (Table 2.2). As a negative control a sample of all components added for the PCR was mixed, but was not used for PCR.

5 µl	10 x reaction buffer
10 ng	dsDNA template
1 µl (125 ng)	Oligonucleotide primer #1
1 µl (125 ng)	Oligonucleotide primer #2
1 µl	dNTP mix
3 µl	Quicksolution
Double destilled H_2O up to a final volume of 50 µl	
1 µl	Pfu Turbo DNA Polymerase (2.5 U/µl)

Table 2.1: Polymerase chain reaction mixture for one sample.

Denaturation		95 °C	1 min
18 cycles	Denaturation	95 °C	50 sec
	Primer annealing	60 °C	50 sec
	Extension	68 °C	14 min (2 min/kb of plasmid length)
Fill up		68 °C	7 min
		4 °C	∞

Table 2.2: The polymerase chain reaction program.

All samples (including the control) were digested with 1 µl Dpn I for 2 h at 37 °C in order to digest non-mutated DNA. 5µl of the Dpn I treated DNA was used for the transformation in XL1-Blue Supercompetent *E.coli* cells. If the negative control does not have any colonies, then all non-mutated DNA is digested and the colonies on the other transformation plates are positives, containing the desired construct. The construct was confirmed by DNA sequencing (MWG, Germany).

2.1.d Rv0386 CHD$_{(1-172)}$ and Rv0386 CHD$_{(1-172)}$ Cys168Ser

The QuickChange XL site-directed mutagenesis kit (Stratagene, Germany) was used for making the constructs containing residue 1-172 with and without the point mutation

Cys168Ser. The holoenzyme Rv0386/pQE30 was used as a template for the PCR, together with the primers (Table 2.3). The exact procedure of making these constructs has been described in 2.1.c. The correct sequence of the constructs was confirmed by DNA sequencing (MWG, Germany).

2.1.e Interface Mutants: V132A and E98A

Two mutants on the dimerisation interface of the CHD of Rv0386 were made. V132A $CHD_{(1-175)}$ of Rv0386 and E98A $CHD_{(1-175)}$ of Rv0386, both in the pQE30 vector have been constructed and sequenced by Dr. Jürgen U. Linder, (Prof. J.E Schultz group, Tübingen Universität, Germany).

Construct	Forward primer	Reverse primer
Rv0386 $CHD_{(1-168)}$C168S	5´ g cgg gta atg cag ctg **agt tga** ccc gaa ttg cgt atc g 3´	5´ c gat acg caa ttc ggg **tca act** cag ctg cat tac ccg c 3´
Rv0386 $CHD_{(1-172)}$C168S	5´ g cgg gta atg cag ctg **agt** cat ccc gaa **tga** cgt atc gat ttc ccg ccg 3´	5´cgg cgg gaa atc gat acg **tca** ttc ggg atg **act** cag ctg cat tac ccg c 3´
Rv0386 $CHD_{(1-172)}$	5´ g cag ctg tgt cat ccc gaa **tga** cgt atc gat ttc ccg ccg 3´	5´cgg cgg gaa atc gat acg **tca** ttc ggg atg aca cag ctg c 3´

Table 2.3: Primers used for cloning shorter constructs of the catalytic domain of Rv0386. Mutations and stop codons are depicted in bold.

2.2 Expression and Purification of the CHD of Rv0386

2.2.a Overexpression

The construct C-His CHD of Rv0386 was transformed into *Escherichia coli* strain BL21(DE3) (pREP4) (AmpR and KanR). Single cell streak out was performed on a LB agar plate containing Ampicillin and Kanamycin. A pre-culture was set up with Luria Broth (LB) medium containing 100 µg/ml Amp and 30 µg/ml Kan and a single colony from the single colony streak out, grown at 37 °C ON. The pre-culture was inoculated into 10-20 L of LB medium in the presence of 50 µg/ml Carbenicillin and 30 µg/ml Kanamycin. Cells were grown at 37 °C until an OD_{600} = 0.4, then the temperature was decreased to 23 °C and the cells were grown to OD_{600} = 0.6-0.8. Prior to induction a glycerol stock was made containing 750 µl cells and 250 µl glycerol, stored at -80 °C. The cells were induced with 100 µM isopropyl-β-D-thiogalactosid (IPTG), at 23 °C for 4-16 h.

The cells were harvested by centrifugation at 5300 g (Sorvall RC12BP, rotor: Sorvall H 12000), for 20 minutes, at 4 °C. The pellet was resuspended in cold PBS (phosphor-buffered saline: 150 mM NaCl, 10 mM Na/K phosphate pH 7.4) and transferred in 50 ml Falcon tubes, re-pelleted by centrifugation at 6500 g (Heraeus Multifuge 35-R) at 4 °C, for 60 min, flash frozen in liquid nitrogen and stored at -80 °C.

The construct C-His CHD of Rv0386 has also been overexpressed in *Escherichia coli* strain B834(DE3) with LB medium, identical to the description in 2.2.a.

2.2.b Over expression of Rv0386 CHD$_{(1-172)}$, Rv0386 CHD$_{(1-172)}$ Cys168Ser and Rv0386 CHD$_{(1-168)}$ Cys168Ser

Over expression was carried out with the constructs of Rv0386 CHD$_{(1-172)}$, Rv0386 CHD$_{(1-172)}$ Cys168Ser and Rv0386 CHD$_{(1-168)}$ Cys168Ser. The constructs were transformed into the *Escherichia coli* strains Rosetta(DE3) pLysS (AmpR and CmR) and BL21(DE3) (pREP4) (AmpR and KanR).

Pre-cultures were prepared with LB medium containing 100 µg/ml Amp and 30 µg/ml Kan for BL21(DE3) (pREP4) cells and 100 µg/ml Amp and 50 µg/ml Chloramphenicol for

Rosetta(DE3) pLysS cells. Additionally, a single colony was inoculated and grown at 37 °C ON. The pre-culture was inoculated into LB medium in the presence of antibiotics. Cells were grown at 37 °C until an OD_{600} = 0.4, then the temperature was decreased to 23 °C and 16 °C and the cells were grown to OD_{600} = 0.6-0.8. A glycerol stock was made prior to induction, containing 250 µl glycerol and 750 µl cells, stored at -80 °C. The cells were induced with 60 µM isopropyl-β-D-thiogalactosid (IPTG), at 23 °C for 4 h or 16 °C, ON.

Afterwards, the cells were harvested by centrifugation at 5300 g (Sorvall RC12BP, rotor: Sorvall H 12000), for 20 minutes, at 4 °C. Cold PBS (phosphor-buffered saline: 150 mM NaCl, 10 mM Na/K phosphate pH 7.4) was used for resuspending the pellet. The suspention was transferred to 50 ml Falcon tubes, centrifugated at 6500 g (Heraeus Multifuge 35-R) at 4 °C, for 60 min, frozen in liquid nitrogen and stored at -80 °C.

2.2.c Purification of C-His CHD of Rv0386

10-15 grams of cells were thawn on ice and suspended in 100-150 ml lysis buffer (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 20 mM imidazole, 5% glycerol, 5mM β-mercaptoethanol) supplemented with one tablet complete protease inhibitor (EDTA free, Boehringer Ingelheim) per 50 ml lysis buffer. The pellet was lysed on ice with a Sonicator (Branson sonifier 250) at 50 % duty cycle, 25 sec pulse, 25 sec pause in total 20 minutes. Cellular debris was separated by ultracentrifugation (Sorvall Discovery 100, rotor: Ti647.5) at 120 000 g for 40 minutes at 4 °C. The protein was filtered over a 0.22 µm sterile PVDF filter (Millipore) and loaded on a Chelating Sepharose column (Pharmacia) loaded with Ni^{2+} ions, pre equilibrated with Buffer A (50 mM Tris-HCl pH 8.0, 10 mM imidazole, 10 mM NaCl, 5% glycerol, 2 mM $MgCl_2$, 5mM β-mercaptoethanol). The protein was washed to baseline with 40 mM, 60 mM and 80 mM imidazole B buffer (50 mM Tris-HCl pH 8.0, 10 mM NaCl, 500 mM imidazole, 5% glycerol, 2 mM $MgCl_2$, 5 mM β-mercaptoethanol). The protein eluted over a linear gradient, with an increasing imidazole concentration of 80 to 500 mM. The fractions of the catalytic domain of Rv0386 were collected, pooled and subjected to Size Exclusion Chromatography on Superdex 75 26/60 (Pharmacia Biotech) at 4 °C. The gel filtration was carried out in 20 mM Tris-HCl pH 8.5, 10 mM NaCl, 5 % glycerol and 2mM β-ME. The protein containing fractions were analysed by SDS-PAGE followed by Comassie brilliant blue staining. The fractions containing the dimer were collected, pooled and concentrated to 10-20 mg/ml using Amicon Ultra centrifugal devices (Millipore) with a cut-off of 10 kD. An aliquot of the

protein was incubated with 1 mM of the potent inhibitor $\alpha_1\beta CH_2ATP$ at 4 °C, 16-19 h. The protein was ultracentrifuged at 100 000 g (Optima Max , Beckman Coulter) for 45 minutes at 4 °C, to remove aggregates before crystallisation.

2.2.d Purification of C-His CHD of Rv0386 as Monomer

C-His CHD of Rv0386 was transformed in BL21(DE3) (pREP4) or B834(DE3) *E.coli* strains. The over-expression was performed as described in section 2.2.a. The cell disruption and nickel affinity purification was carried out according to section 2.2.c. The protein was dialysed in 20 mM Tris-HCl pH 8.5, 500 mM NaCl, 10 % glycerol, 2 mM β-ME, 0.5 mM $MnCl_2$ ON in 4 °C and run on a gelfiltration column Superdex 75 26/60 (Pharmacia Biotech) in a buffer containing the same components as the dialysis buffer. The protein in the monomeric peak was collected and concentrated using Amicon Ultra centrifugal devices (Millipore) with a cut-off of 10 kD to ca 1 mg/ml. The protein was incubated with 1 mM $\alpha_1\beta CH_2ATP$ at 4 °C, ON. Prior to crystallisation the protein was ultracentrifuged at 100 000 g (Optima Max, Beckman Coulter) for 45 minutes at 4 °C, in order to remove aggregates.

2.2.e Purification of N-His CHD of Rv0386

N-His CHD of Rv0386 was transformed in BL21(DE3) (pREP4) or BL21(DE3) *E.coli* strains. The over-expression was carried out according to section 2.2.a.

10 grams of cells corresponding to 4 L of cell culture were thawn on ice and suspended in 200 ml lysis buffer (50 mM Tris-HCl pH 8.5, 300 mM NaCl, 10% glycerol, 2 mM β-ME, 1.0 mM $MnCl_2$) supplemented with one tablet complete protease inhibitor (EDTA free, Boehringer Ingelheim) per 50 ml lysis buffer. The cells were disrupted on ice with a Sonicator (Branson sonifier 250) at 50 % duty cycle, 25 sec pulse, 25 sec pause in total 40 minutes. Cellular debris was separated by ultracentrifugation (Sorvall Discovery 100, rotor: Ti647.5) at 120 000 g for 40 minutes at 4 °C. The protein was purified over a Chelating Sepharose column (Pharmacia) loaded with Ni^{2+} ions, pre equilibrated with Buffer A (50 mM Tris-HCl pH 8.5, 10 mM imidazole, 10 mM NaCl, 10% glycerol, 2 mM β-ME, 1.0 mM $MnCl_2$). The protein was washed to baseline with 50 mM imidazole B buffer (50 mM Tris-HCl pH 8.5, 10 mM NaCl, 500 mM imidazole, 10% glycerol, 2 mM β-ME, 1.0 mM $MnCl_2$). The protein eluted over a linear gradient of 10 to 100 % corresponding to 50 to 500 mM imidazole. The fractions

of the catalytic domain of Rv0386 were collected, pooled and concentrated. One part of the protein was used for crystallisation, while the other was polished on Superdex 75 26/60 (Pharmacia Biotech) at 4 °C. The gel filtration was carried out in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 % glycerol, 2 mM β-ME, 1.0 mM MnCl$_2$. The protein containing fractions were analysed by SDS-PAGE followed by Comassie brilliant blue staining. The fractions containing the dimer were collected, pooled and concentrated to 10 mg/ml using Amicon Ultra centrifugal devices (Millipore) with a cut-off of 10 kD. Aggregates were removed by ultracentrifugation at 100 000 g (Optima Max, Beckman Coulter) for 45 minutes at 4 °C. Drops were set up for crystallisation, by the hanging drop method.

2.2.f   Purification of Rv0386 CHD$_{(1\text{-}172)}$ Cys168Ser and Rv0386 CHD$_{(1\text{-}168)}$ Cys168Ser

Four grams of cells were thawn on ice and suspended in 50 ml lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 0.02% MTG) added with one tablet complete protease inhibitor (EDTA free, Boehringer Ingelheim) per 50 ml lysis buffer. Disruption of the cells was accomplished on ice with a Sonicator (Branson sonifier 250) at 50 % duty cycle, 25 sec pulse, 25 sec pause in total 20 minutes. Ultracentrifugation (Sorvall Discovery 100, rotor: Ti647.5) at 120 000 g for 40 minutes at 4 °C was carried out to separate the cellular debris. The protein was loaded on a dripping column comprising 1 ml Chelating Sepharose (Pharmacia) resin loaded with Ni^{2+} ions, pre equilibrated with 5 column volumes of lysis buffer. The protein was washed with 50 mM imidazole B buffer (buffer A: 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol, 0.02% MTG, buffer B: 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole, 10% glycerol, 0.02% MTG). The protein was step eluted with 150 mM and 500 mM imidazole B buffer. The catalytic domain of Rv0386 was collected in eight fractions, one ml each, pooled and concentrated. An aliquot was dialysed at 4 °C and used for crystallisation, while the remainder was purified further on a Superdex 75 26/60 (Pharmacia Biotech) at 4 °C in 50 mM Tris-HCl pH 8.5, 500 mM NaCl, 10 % glycerol and 2mM β-mercaptoethanol. The protein containing fractions were analysed by SDS-PAGE followed by Comassie brilliant blue staining. The fractions containing the protein were collected, pooled and concentrated to 5-10 mg/ml with a cut-off of 10 kD using Amicon Ultra centrifugal devices (Millipore). Aggregates were removed by ultracentrifugation at 100 000 g

(Optima Max, Beckman Coulter) for 45 minutes at 4 °C. The protein was used for crystallisation.

2.2.g Production of Selenomethionine Substituted Protein

The C-His CHD of Rv0386 construct was transformed into methionine requiring auxotroph *Escherichia coli* strain B834(DE3) and plated out on a LB agar plate containing Ampicillin. Pre-culture: a single colony was spread on a LB agar plate containing Amp, incubated in 37 °C ON. The overgrown agar plate was harvested and used for inoculation of 1 L minimal media (2xM9, 0.4 % glucose, 19 amino acids (-Met) at 40 µg/ml, Seleno-L-methionine at 40 µg/ml and vitamins: riboflavin, niacinamide, pyridoxine monohydrochloride, thiamine at 1µg/ml) (Table 2.4-2.6) (Ramakrishnan et al., 1993) with 50 µg/ml Carbenicillin. The culture was grown for 4h, at 37 °C until the OD_{600} reached a value of 0.7. As a control to see if the plasmid was still inside the cells, a small sample of the culture was spread on a LB-agar plate containing Amp and left in 37 °C, ON. The cells were induced with 100 µM IPTG, for 5h at 23 °C. The cells were harvested and stored in -80 °C, as described in section 2.2.a. The buffers used for the protein purification were extensively degassed, to avoid oxidation of the SeMet protein. Purification was performed according to section 2.2.c.

20 x M9	
10 g	NH_4Cl
30 g	KH_2PO_4
68 g	Na_2HPO_4 (or 128g Na_2HPO_4 x 7 H_2O)
Add water to make 500 ml and autoclave.	

Table 2.4: The components of the 20 x M9 media.

Amino acid mix I: all amino acids except Met, Phe, Trp, Tyr. Concentration of each amino acid: 4mg/ml. Filter sterilise and store at -20 °C.
Amino acid mix II: Phe, Trp, Tyr. Concentration of each amino acid: 4mg/ml. The solution is adjusted to pH 8.0 with NaOH to improve solubility before bringing it up to volume. The mixture could be turbid. Filter sterilise and store at +20 °C.
Vitamins: riboflavin (= lactoflavin, vit B2), niacinamide (= nicotinamide, vit PP), thiamine (= vit B1), pyridoxine monohydrochloride (= vit B6). Concentration of each: 1mg/ml. Make a vitamin mix of niacinamide, thiamine and pyridoxine monohydrochloride. Riboflavin does not dissolve completely. Filter sterilise the vitamin mix and riboflavin and store at -20 °C.

Table 2.5: The mixtures, concentrations and storage temperatures of the vitamins, amino acid mix I, amino acid mix II.

1L of minimal media supplemented with SeMet	
2 ml	1 M $MgSO_4$
100 ml	20 x M9
2 ml	$FeSO_4$ x 7 H_2O at 12.5 mg/ml
10 ml	40% Glucose
10 ml	Amino acid mix I at 4mg/ml (all residues except Met, Phe, Trp, Tyr)
10 ml	Amino acid mix II at 4mg/ml includes Phe, Trp, Tyr
1 ml	Vitamins at 1 mg/ml includes riboflavin, niacinamide, thiamine, pyridoxine monohydrochloride
4 ml	Seleno-L-Methionine at 10 mg/ml
Add water to make 1L and make sure that pH is about 7.4	
Autoclaved components: 40% Glucose, 1M $MgSO_4$, 20 x M9 The other solutions were filter sterilised (vitamins, aa mix I, aa mix II, SeMet)	

Table 2.6: Minimal media supplemented with SeMet. Full description of the media can be found on (http://alf1.mrc-lmb.cam.ac.uk/~ramak/madms/segrowth.html)

2.3 Activity Studies of the Catalytic Domain of Rv0386

The activity of the AC has been determined in a volume of 100 µl at 30 °C for 20 minutes (Salomon et al., 1974). The catalytic domain of Rv0386 was incubated with 50 mM 3-(N-morpholino)-propanesulfonic acid pH 7.5, 22 % glycerol, 5 mM $MnCl_2$, 850 µM [α-^{32}P]ATP and 2 mM [2,8-^{3}H]cAMP. The kinetic analysis was conducted from 10 µM to 2.3 mM ATP. From the Lineweaver-Burk plot the kinetic constants have been derived (Lineweaver and Burk, 1934). Standard errors of the mean of the multiple assays were usually <10 %. In order to determine the GC activity, the guanine nucleotides were used instead of the adenine nucleotides(Klumpp and Schultz, 1982). The activity assays have been performed by Lucila I. Castro in the laboratory of J.E Schultz (Tübingen Universität, Germany) (Castro, 2004).

2.4 Biophysical Characterisation

2.4.a Dynamic Light Scattering

In order to determine if the catalytic domain of Rv0386 from *Mycobacterium tuberculosis* is monodisperse or polydisperse, the protein was measured in different concentrations, buffers and temperatures in a 12 µl quartz cuvette in a Protein Solutions DynaPro temperature controlled Microsampler.

2.4.b Mass Spectrometry

Electrospray ionisation mass spectrometry (ESI-MS) was performed by Dr. Jens Pfannstiel, member of the group of Prof. Johannes Lechner, Biochemie-Zentrum Heidelberg, Universität Heidelberg.

2.4.c Analytical Ultracentrifugation

In order to determine shape and oligomeric state of the catalytic domain of Rv0386 in solution analytical ultracentrifugation (AUC) was performed. Since the reducing agent β-mercaptoethanol absorbs, the protein was purified according to section 2.2.c and 2.2.d and then dialysed into the buffer containing 20 mM Tris-HCl pH 8.5, 10 mM NaCl, 5 % glycerol and 0.5 mM DTT at 4 °C ON. 5% glycerol is the maximal amount compatible with AUC. Additionaly, the Mn^{2+} purified sample was dialysed into 50 mM Tris-HCl pH 8.5, 500 mM NaCl, 5 % glycerol, 0.5 mM DTT and 50 mM Tris-HCl pH 8.5, 10 mM NaCl, 5 % glycerol, 0.5 mM DTT, 0.5 mM $MgCl_2$. Protein concentration was measured at Abs_{280} and the protein concentration used for the experiment was 1 mg/ml. The volume used for each protein sample was 500 µl, with dialysis buffer as reference. The rotor A N60 Ti contained three cells and each cell had two compartments, one for the buffer and the other for the protein. Sedimentation velocity run was conducted in a Beckman Coulter Optima XLA analytical ultracentrifuge, 200 000 g, 2.5 h, at 20 °C. The software SEDNTERP (Laue et al., 1992) was used for calculating the partial volume, Hydropro (Garcia De La Torre et al., 2000) for determining the theoretical s-values and Sedfit (Schuck, 2000) for calculating the actual s-values. The analytical ultracentrifugation run was performed by Jacek Mazurkiewicz, in the group of PD Karsten Rippe, Kirchhoff Institut für Physik, Universität Heidelberg.

2.5 Crystallisation, Data Collection and Structure Determination

2.5.a Crystallisation

Formation of macromolecular crystals are influenced by many variables, such as homogeneity of the protein, protein purity, concentration of the protein and precipitant, temperature, pH, vibrations, age of protein, source, and presence of ligands. In the most common methods of growing protein crystals, pure and homogeneous protein is dissolved in an aqueous buffer containing a precipitant (e.g. ammonium sulfate or polyethylene glycol) at a concentration where the protein is still soluble and does not precipitate. The precipitating conditions are then reached by controlled evaporation of water allowing for crystal growth.

 In the hanging drop vapour diffusion method the protein/precipitant solution is equilibrated in a closed container with a larger reservoir, where the precipitant concentration is in the appropriate range to produce crystals. Vapor diffusion in this closed system results in net transfer of water from the protein solution to the reservoir, until the precipitant concentration is the same in both solutions. The volume of the protein droplet is much smaller than the reservoir. The final concentration of the precipitant in the protein solution is approximately the same as in the reservoir. The net transfer of water terminates at system equilibrium and the protein solution is sustained at the highest precipitant concentration. Amorphous precipitation usually occurs when protein-and precipitant concentrations are too high. In case of no precipitation, protein-and precipitant concentrations would have to be adjusted.

In order to find a proper crystallisation condition, different commercial sparse-matrix and grid screens were tested. 500 µl of precipitant solution is filled into a crystallisation plate containing 24 wells (Molecular Dimensions Limited). 1.5 µl of precipitant is mixed with 1.5 µl of protein solution on a siliconised cover slide (Hampton Research). The cover slide containing the drop hanging up-side down was sealed with grease (hanging drop vapour diffusion method). For initial crystallisation conditions crystallisation screens Wizard I and II, Cryo I and II (Emerald Biostructures) MembFac, Crystal Screen I and II, Index Screen, Peg/Ion Grid Screen (Hampton Research) were used. Conditions containing crystals were optimised by individually devised screens around the initial condition by altering buffer, precipitant, and additive concentrations and identities, as well as altering the temperature of crystallisation (19 °C and 4°C) and protein concentration. Crystals grew ON, but the biggest and best diffracting crystals took from a few days to 2 weeks to grow. Rod shaped crystals of the catalytic domain of Rv0386 grew in hanging drops using vapour diffusion at 19 °C by mixing 1.5 µl protein solution (10-20 mg/ml) with 1.5 µl of 0.1 M Hepes pH 7.5 and 0.7 M sodium acetate. The selenomethionine crystals were grown in order to perform a single anomalous dispersion experiment (SAD). The SeMet crystals grew as thick rods. 1.5 µl Se-Met protein (10-20 mg/ml) was crystallised in 1.5 µl of 5% (w/v) PEG2000, 0.1 M Hepes pH 7.5 and 0.2 M KCl. After 2-4 days native and Se-Met crystals grew to a size of 30 x 30 x 250 µm. Rv0386 catalytic domain crystallised in the tetragonal space group P4$_1$22. The first non-diffracting crystals of Rv0386 CHD were originally grown by Lucila I. Castro (Universität Tübingen, Germany) (Castro, 2004).

2.5.b Crystal Freezing and Data Collection

For most crystals the cryo-solution, i.e. a solution which avoids the formation of ice upon freezing, is the same as the crystallisation conditions with an additional 20-25 % (w/v) glycerol, polyethylenglycol or ethylenglycol (Garman and Schneider, 1997). A protein crystal is moved to a cryo-solution for a few seconds. Thereafter the crystal is transferred into liquid nitrogen using a nylon loop of size 0.05 mm to 0.5 mm (depending on the size of the crystal) and stored frozen at 100 °K until data collection. At 100 °K crystals are exposed to X-rays, because at this temperature radiation damage is lower than at room temperature, conformational disorder and thermal vibrations are reduced.

In the year 1895 X-rays were discovered by Röntgen. X-rays are electromagnetic radiation of wavelength 0.1-100 Å and can be described by three parameters, wavelength, amplitude and phase. X-rays are not affected by magnetic fields and cannot be focused by optical lenses. The images are therefore derived indirectly by recording the diffraction pattern of X-rays. The wavelength of X-rays emitted by the X-ray source and the wavelength of all recorded X-rays are the same, which means that the diffraction is inelastic and that the wavelength is known. Diffraction of a single molecule cannot be detected with present technology, since X-rays are scattered very weakly. Usually, diffraction from multiple proteins is erased by each other but in proteins that form a crystal diffraction is enhanced and can be detected if it fulfils a criterion known as Bragg's law. Bragg´s law of diffraction was established in the year of 1912 by W.H Bragg and W.L Bragg and it demonstrates the conditions that produce diffraction. The first structure was determined in 1912 from NaCl. The 17 kD sized protein Myoglobin was the first protein structure, solved by Max Perutz in 1958.

$$\text{Bragg´s law: } 2\, d_{hkl}\, \sin\theta = n\, \lambda$$

d_{hkl}: spacing between two lattice planes

hkl: miller indices, identifies equivalent, parallel planes

θ: diffraction angle

n: an integer

λ: wavelength of x-ray (copper target emits at 1.5418Å, CuK_{α})

X-rays are produced by two methods. The first method is that rotating anodes or sealed tubes generate X-rays by accelerating electrons at high voltage against a metal target (most common is copper or molybdenum) and the second method is to use synchrotron radiation, where the electrons emit energy in the curved motion and the energy is emitted as X-rays. Protein crystals are exposed to a strong X-ray source, resulting in a diffraction pattern (reflections) that is recorded by a X-ray detector. In order to test diffraction of crystals and heavy-atom derivatives an in-house Bruker Nonius X-ray generator was used with a MAR345 imaging system (Marresearch). The crystals were diffracting to ca 8-10 Å in-house. All data used for structure determination was recorded on CCD image plates (ADSC) at the European synchrotron radiation facility (ESRF) in Grenoble, France. The beamlines used for data collection were ID14-4 (native) and ID14-1 (SeMet).

2.5.c Data Processing

Autoindexing can be done in two ways. One is to autoindex the collected data by entering the unit cell in the program and all possible orientations of this cell are searched until a match is found. In the second, the reciprocal space is searched for vectors, three non-coplanar groups are searched for, and the shortest reciprocal space distance (the largest in real space) is taken to be the cell edges. The correct cell is usually the one that has the longest axes (in real space) and highest symmetry.

Crystals are constructed from unit cells, repeated by translation in all three directions. The unit cell edges are assigned a, b, c and the angles between them are α, β and γ. The unit cell edges a, b, c in real space corresponds to the coordinate system h, k, l (Miller indices) in reciprocal space, respectively, giving rise to the reflection spot (h k l). The program DENZO (Otwinowski and Minor, 1997) determinates parameters from the pattern of recorded diffraction spots and allocates the spots to lattice planes: called indexing. The intensities of the spots by integration and profile fitting is measured by DENZO. Scalepack (Otwinowski and Minor, 1997) was used for scaling images from different images, since intensities on different images vary in strength due to the thickness of the crystal in the beam or due to crystal decay. Reflections that diverge above a threshold from the mean are assumed to be incorrectly measured and can be rejected by Scaling programs. The Rsym value is an important parameter of the accuracy of a data set. In Rsym the differences between symmetry

related spots that in the ideal case would be identical in intensity is compared. Data with Rsym < 0.05 are good data, Rsym < 0.10 probably usable data, Rsym > 0.20 are doubtful data. Two datasets scaled together with a Rsym of 0.57 do not have any relationship to each other. Rsym of 0.35 can be reached with random data scaled in bins with multiple scale factors. The reflections are calculated as:

$$Rsym = \frac{\sum h \sum i |I(h) - I(h)i|}{\sum h \sum i I(h)i}$$

I(h) is the mean intensity after scaling

I(h)i is the intensity of a recorded diffraction spot

2.5.d Molecular Replacement

The phase of each reflection is the only parameter not known, so far. There are several possibilities to determine this. A homologous structure with sequence identity of ca 50 %, can be used as a model in molecular replacement. With this high sequence identity it should be straightforward, since the structures are most likely very similar. In molecular replacement, firstly introduced by Michael Rossmann (Rossmann, 1972), the phases from structure factors of a related, known protein are used as an initial guess of the unknown structure phases. The known protein is called the phasing model and it is fitted into the unit cell of the unknown structure. By calculating a Patterson function in a random P1 cell and then superimposing them (cross rotation matrix) the orientation of the two molecules are determined. Thereafter, the translation function in the real cell is determined. This method assumes as that the phases of the phasing model are the same as that of the unknown protein. R-factor calculation and correlation functions estimate the quality of the solutions. Molecular replacement failed for CHD of Rv0386. Monomeric, dimeric and truncated models from different homologous proteins were used, unsuccessfully. The advantage of molecular replacement is that it is the fastest method of solving a structure, from a single native data set when a homologous model is available. However, initial phases are biased towards the phasing model, making it difficult to align the correct structure.

2.5.e Single Anomalous Dispersion (SAD)

Another possibility to determine phases is to take advantage of the heavy atom´s (selenium) capacity to absorb X-rays of specific wavelength. Due to this absorption, breakdown of Friedel´s law occurs and the reflections h k l and -h -k –l (Bijvoet pairs) do not have equal intensity anymore (Figure 2.4). This effect is called anomalous dispersion or anomalous scattering. Elements absorb and emit X-rays, and at wavelengths just below their emission wavelength the absorption drops sharply. This abrupt change in absorption at a specific wavelength is called the absorption edge. Anomalous scattering occurs when the X-ray wavelength is near the heavy atom´s (selenium) absorption edge, the heavy atom absorbs a fraction of the radiation and is reemitted with altered phase. The highest signal of anomalous scattering is obtained when the wavelength of the heavy atom´s absorption edge is used for data collection. Carbon, oxygen and nitrogen, do not contribute to anomalous scattering, since the absorption edge for these elements are not near the wavelength of the X-rays used in crystallography.

Differences in intensity are usually very small, so data are collected with a high redundancy to improve the signal to noise ratio and allow more accurate estimates of the intensities. It is necessary to use one crystal where the Friedel pairs are collected under similar conditions. Data for SeMet substituted CHD of Rv0386 crystals were collected at ID 14-1 at ESRF, Grenoble, France which is a beamline with a fixed wavelength at $\lambda = 0.934$ Å. Prior to the data collection, a wavelength scan was performed of SeMet substituted crystals at ID 14-4 and revealed a peak wavelength of $\lambda = 0.9793$ Å for selenium. Therefore the wavelength at ID 14-1 was sub optimal for the anomalous signal. The structure of the catalytic domain of Rv0386 was solved using single anomalous dispersion (SAD). SAD data can be collected at synchrotron beamlines whose wavelength cannot be altered (ESRF, ID 14-1, ID 14-2 etc). Multiple anomalous dispersion (MAD) on the other hand, uses the anomalous dispersion at multiple wavelengths and therefore data has to be collected at synchrotron beamlines whose wavelength can be altered. MAD was not tried, since the method SAD was successful. The substructure solution of anomalous scatterers are determined, by an anomalous Patterson map that calculates the position of the anomalous scatterers. The substructure of the Se-Met derivatised CHD of Rv0386 was found with either of the programs SHAKE 'n' BAKE (SnB) (Miller et al., 1994) or SOLVE (Terwilliger and Berendzen, 1999). Phasing was carried out with the program SHARP (de la Fortelle and Bricogne, 1997), and phases were improved by solvent flattening with Solomon (Abrahams and Leslie, 1996). Solvent flattening improves

the signal to noise ratio in electron density maps. Usually 30-70 % of the crystal volume is occupied by protein. The rest is comprised of solvent, that adds random noise to the diffraction. With the program Solomon a clear electron density map was obtained.

By replacing the sulphur atoms to selenium atoms in methionines, centres are created that are more electron dense in the SeMet substituted protein compared to the native protein. The reflection intensities are therefore altered. The difference between the structure factor of the native protein (FP) and the derivatised protein crystal (FPH) is defined as FH. The substructure solutions of heavy atom centres can be defined from Patterson maps, and phases are determined (Figure 2.5) for structure determination.
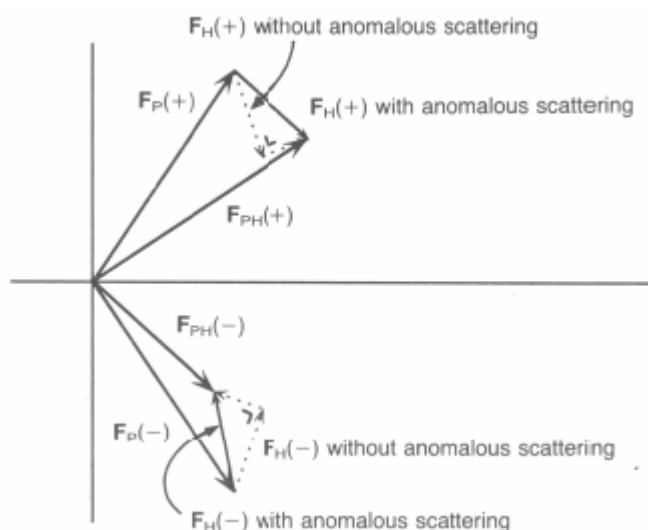

Figure 2.4: Breakdown of Friedel's law. Friedel's law states that the amplitude of FP (+) and FP (-) are the same. The amplitude of FP (+) and FP (-) is equal for the native protein but the phase is reversed. The protein with an anomalous scatterer has both a different phase and amplitude between FPH (+) and FPH (-). Figure taken from (Drenth, 1994).
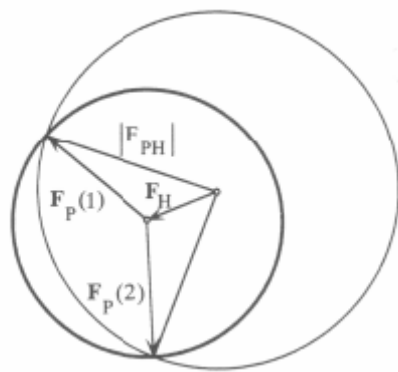
Figure 2.5: Phase determination using isomorphous replacement. The structure factor of the native protein is FP. When the substructure of the heavy atoms has been determined, phase and magnitude of FH can be calculated. With FPH measured experimentally, there are two possible phase values for FP, FP1 and FP2. Which of these is correct can be determined by a second derivative, which intersects with either FP1 or FP2. Figure taken from (Drenth, 1994).

2.5.f Refinement

The electron density of the unit cell can be calculated and represented in a map. The electron density is the primary experimental result, in which a protein model is built. The following formula is used to calculate the electron density at any point within the unit cell. This is the Fourier series that gives the first electron density map:

Fourier synthesis:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l w_{hkl} |F(hkl)_{obs}| e^{-2\pi i(hx+ky+lz-\alpha'(hkl)calc)}$$

ρ (x y z) is the value of electron density at position x, y, z

F (h k l) $_{obs}$ is the structure factor of reflection h k l from the native intensities of the new protein

W $_{hkl}$ is the weight factor, which is close to 1.0 if α ´ (h k l) is among the most reliable phases

α ´ (h k l) $_{calc}$ is the phase of reflection h k l from the anomalous dispersion or phasing model

The maps were visualised by Program O (Jones et al., 1991). The Cα chain of Rv0386 catalytic domain was traced by using a poly-alanine model. In order to make the protein

model to refine, the start with a poly-alanine model was very useful. If the first protein model is derived from molecular replacement, the protein model will contain errors because the protein could be in another conformation. In order to improve the protein model, manual building is done in O. R-factor is a reliability factor that is used for estimating the error of the protein model. R-factor is the difference between the experimentally derived structure factors and the calculated structure factors that would be obtained if the protein model was correct.

$$R-factor = \frac{\sum \left| F_o(h) - F_c(h) \right|}{\sum \left| F_o(h) \right|}$$

Fo experimentally observed structure factor of reflection h
Fc calculated structure factor of reflection h

A completely random model would have a R-factor about 55 %. R-factor of refined structures are between 20 and 30 percent, depending on the resolution of the diffraction. In Table 3.4 the refined structure of the CHD of Rv0386 is presented. Higher resolution structures tend to have lower R-factor values. The R-factor alone is influenced by the changes made and therefore not free of model bias, i.e. it can lead to over-refinement. Approximately 5% of reflections are left out from refinement and therefore independent of any manipulation. A Rfree value (Brunger, 1993) is calculated from these reflections, using the same formula as for calculating the R-factor. Typically the R-factor should be about the resolution × 10 and the Rfree should be a few % higher. At a resultion of 1.6 Å a good structure may have R-factor = 17 % and Rfree = 20 %. If the difference between R-factor and Rfree is ca 10 % there is a risk that the protein model has been over-refined. If the R-factor falls during refinement and Rfree do not decrease, then there is a risk that the model is wrong. It is important that both Rfree and Rfactor decline.

Refinement programs can refine the atomic coordinates of atomic positions as well as their B-factors that represent the level of dynamic disorder of individual atoms. Iterative model building and refinement using the program REFMAC 5 (Murshudov et al., 1999) was done. As the model improves, the restraints can be relaxed. Waters were placed by the program ARP/wARP (Perrakis et al., 1999). In the final refinement a few cycles of TLS (Winn et al., 2001) was used. The quality of the refined model was evaluated using WHAT IF (Vriend, 1990) and PROCHECK (Laskowski et al., 1993). These programs validate the quality of the

Ramachandran plot, as well as errors like atoms that are unlikely near to each other, improbable hydrogen bond arrangements. The Ramachandran plot compares the mainchain angles of proteins (ψ and φ) with those from known structures at high resolution and on this basis defines allowed, generously allowed and disallowed regions. The program SFCHECK (Vaguine et al., 1999) checks the quality of the structure-factor data and compares it to the structure factos deposited in the Nucleic Acid Database (Berman et al., 1992).

2.5.g Data Analysis and Interpretation

The residues of the protein model into secondary structure elements were assiged with the program DSSP(Kabsch and Sander, 1983). PYMOL (DeLano, 2002) has been used in drawing all figures of the protein structure. A sequence alignment was generated using the program CLUSTALX (Thompson et al., 1997). The topology of the cyclase homology domain was designed with TopDraw (Bond, 2003). DynDom (Hayward and Berendsen, 1998) generated the domain movements between the inhibited and modeled active state. Corel Photo-Paint 12 was utilised for photo editing purposes.

2.6 Small Angle X-ray Scattering (SAXS)

2.6.a Introduction to SAXS

In the end of the 19th century, Röntgen discovered radiation with wavelength much smaller than that of visible light. Because of their unknown nature, Röntgen named this high-energy radiation "X-rays". Soon after this discovery, von Laue and his associates discovered that crystals scatter X-rays in distinct patterns. It was quickly recognized that these patterns give direct insight into the structure of the materials that caused the scattering. Since these early discoveries, many technical advances made X-ray scattering one of the most powerful characterization tools available for both homogeneous and heterogeneous materials. Today,

scattering from X-rays, neutrons and light is used by scientists in many different disciplines to study an extensive range of materials ranging from polymers to proteins.

The method small angle X-ray scattering (SAXS) makes it possible to study individual proteins and large complexes, under nearly physiological conditions and in solution. It is applicable to study the shape of low resolution three-dimensional particles, oligomeric mixtures, complex formation, quaternary structure and kinetics of assembly/dissociation or folding/unfolding processes. Analysis of structural changes upon variation in ionic strength, pH, temperature etc. is also eligible. SAXS can be assigned for molecules with the molecular weight of 5 kDa to 100 MDa in a resolution from 15 to 10 000 Å. SAXS is the only successful method to obtain the shape of macromolecules with molecular masses of a few hundred kDa in their native condition, hence too small for cryo electron microscopy (EM) and too large for nuclear magnetic resonance (NMR). The disadvantage with SAXS is that it cannot (easily) be applied to study viscous samples.

Monodisperse samples (>90 % is desired) are required for structure analysis, i.e determining shape and quaternary structure. The monodispersity of the sample can be determined by native gel filtration, analytical ultracentrifugation (AUC), dynamic light scattering (DLS) etc. For a complete study 2 to 5 mg of protein is required. A concentration range from 1 to 20 mg/ml (1, 3, 5, 10, 15, 20 mg/ml) in a sample volume of ca 100 µl per measurement is necessary. In order to decrease the aggregation/damage during the data collection the reducing agent DTT (2mM) was added to the sample and the buffer before the experiment. If the sample is aggregated, the scattering data will be difficult or even impossible to interpret. The corresponding buffer is measured before and after each sample of the macromolecular solution. Buffers containing high amounts of salts and glycerol (>0.5 M) should be avoided, since the signal of the sample with low protein concentrations is difficult to detect because of the limited contrast. The cell is cleaned and refilled manually after each measurement. The buffers should be in sufficient amount (> 10 ml) to clean the cell and dilute the samples. The SAXS measurement was performed at the X33 SAXS-beamline, EMBL, Hamburg, Germany. The collection, processing and the analysis of the SAXS data was done with the help of Dr. Dmitri Sverguns group.

2.6.b Theory of SAXS

Small angle X-ray scattering (SAXS) uses a small diffraction angle when a sample in solution is irradiated with X-rays and the resulting scattering pattern is measured (Figure 2.6). A low resolution structure is thereafter determined from the observed reflections on the detector. Scattering patterns are caused by the interference of secondary waves that are emitted from various structures (electrons for X-rays) when irradiated. Scattering of X-rays is caused by differences in electron density. Unlike an electron micrograph, small angle X-ray scattering patterns do not give morphological information directly. The result of a SAXS experiment gives rise to a scattering curve and it is essentially the intensity of the Fourier transform of the electron density and must be interpreted in order to determine morphology. One fundamental problem with any scattering experiment is that two different morphologies can, in theory, give identical scattering patterns. Generally, one cannot reconstruct the exact microstructure uniquely from a SAXS pattern because in a scattering experiment only the scattered radiation intensity can be measured and all phase information is lost. Therefore, one cannot be absolutely sure that a scattering pattern is due to a particular morphology. However, something usually is known about the system in question, so that it is often (but not always) reasonable to assume that if a particular model is shown to fit the scattering pattern, then the model is a correct description of the morphology. Nevertheless, many different approaches exist to extract morphological information from a SAXS pattern. Interpreting the SAXS data can be a very difficult task. Models can be fitted to give information about the radius of gyration, shape of the particles, surface-to-volume ratio of the scattering objects, crystallisation of polymer materials.
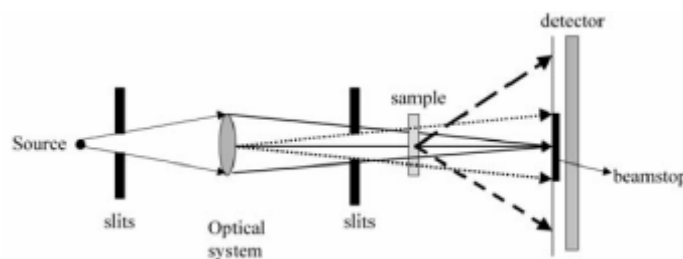

Figure 2.6: Instrumentation of SAXS. Picure taken from (Koch et al., 2003).

Scattering of a protein in solution allows the determination of the radius of gyration of the macromolecule, which yields the shape, and the correlation, which depends on the distance between clusters. The scattering sample can give various information, i.e the decline of the scattering curve and from features in the curve such as peaks that correspond to a correlation function between the particles. During the SAXS measurement there are three components that are measured, i.e the scattering of the sample (signal that we require), background scattering (water) and noise i.e. X-rays scattered from secondary objects like slits.

$$I(q) \propto \frac{d\Sigma(q)}{d\Omega} = \frac{N}{V} V^2_{particle} (\rho_1 - \rho_2)^2 P(q) S(q)$$

$I(q)$: intensity of scattering

N :   number of scatterers in the volume V of sample

$V_{particle}$: volume of the individual scattering item

ρ: density of the particle (ρ_1) or the matrix (ρ_2)

$P(q)$: form factor

$S(q)$: structure factor

In a diluted system the clusters in solution scatter independently, and thereby there are no interparticle effects. Due to this, the sum of independent particle scattering is equal to the total scattering intensity I(q). Since there are no interactions between particles, there are no correlation effects and therefore the structure factor S(q) has no affect on the scattering.

$$I(q) \propto \frac{d\Sigma(q)}{d\Omega} = \frac{N}{V} V^2_{particle} (\rho_1 - \rho_2)^2 P(q)$$

The size and shape about the scatterers can be calculated by the form factor P(q).

Monodisperse Spheres

It can be difficult or even impossible to interpret the SAXS data if the sample (scattering spheres) is not monodisperse. In Figure 2.7 measured data of silica spheres are shown in orange and the size and shape can be determined from the form factor $P(q)$ in blue.
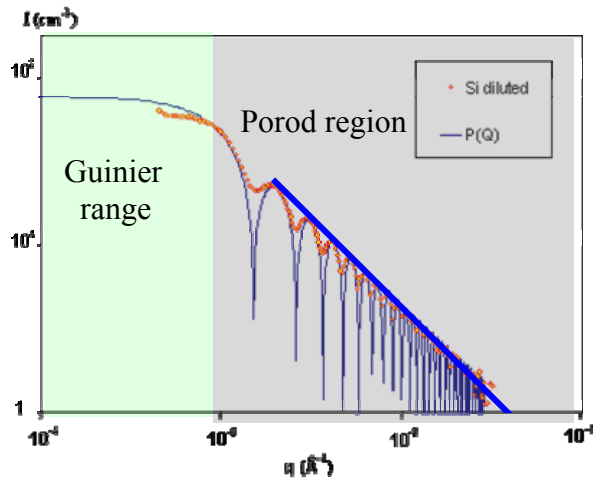
Figure 2.7: From the scattering curve the form factor P(Q) of silica spheres is determined.

Particles Interacting

With increased density for the silica spheres, the correlation effect becomes important, since it changes the scattering features (especially at low q values). The interparticle structure can be determined by interpreting the correlation effect, appearing as additional peaks in the scattering (Figure 2.8).
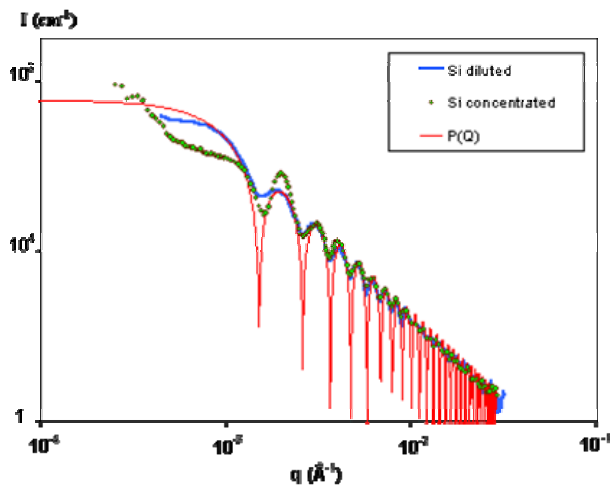
Figure 2.8: Experimental data on interacting charged Silica spheres. The intensity scattering signal I(q) is devided into two terms: the structure factor S(Q) and the form factor P(Q).

Form factors for Spheres

Below is the formula for the form factor P(q) for a spherical scatterer of density contrast ρ_0 and radius R. The expected scattering intensity can be calculated and through the Guinier and the Porod approximations the scattering curve can be fully investigated.

$$P(q) = \left(V\rho_0 \frac{3}{(qR)^2} \left[\frac{\sin qR}{qR} - \cos qR \right] \right)^2$$

Guinier Approximation

From the text above we know that the intensity of the scattering is proportional to the structure factor and the form factor.

$$I(q) \propto P(q)S(q)$$

In this case there are no particles interacting (diluted system) and therefore the structure factor does not have any effect on the scattering, and thus S(q)=1.

The Guinier region is marked in the scattering curve in Figure 2.7. The Guinier region is referred to a very low angle range of the scattering plot, where q < 1/R. In the Guinier area a plotted graph of ln I(q) against q^2 should give a straight line.

$$I(q) = V^2 \rho_0^2 e^{-\frac{q^2 R^2}{5}}$$

$$\ln I(q) = \frac{-q^2 R^2}{5} + c$$

The size of the particles (assuming spherical scatterers) can be determined from the gradient of the line. The equation above is applicable on spherical particles, but the equation can be adopted to fit an arbitrary shaped particle, by using the radius of gyration, which is the centre of electron mass. In conclusion, from the Guinier region of the scattering curve both shape and size can be extracted.

Porod´s law

The Porod region (Figure 2.7) is an area of the scattering curve with higher q and it is applied on the surface properties of the particle.

$$I(q) \propto P(q) = \left(V\rho_0 \frac{3}{(qR)^2} \left[\frac{\sin qR}{qR} - \cos qR \right] \right)^2$$

In the Porod region, qR>>1, when the $\cos^2(qR)$ = 1/2 and sin (qR)/qR→0, the equation is simplified to:

$$I(q) \approx V^2 \rho_0{}^2 \frac{9\langle \cos^2 qR \rangle}{(qR)^4} \approx V^2 \rho_0{}^2 \frac{9}{2q^4 R^4}$$

Towards larger angles there is a decrease in the scattering intensity, which follows a q^{-4} dependence. An exception is though, rough or complex surfaces of the scattering particles that will not have a q^{-4} dependence at higher angles. The surface area and the unit volume are assigned to the V^2/R^4 factor, which is a measure of the particle smoothness.

The experimental data is fitted by using the form factor for spherical scattering particles.
In order to fit the data correctly, the absolute scattering intensity must be known which is related to the concentration of the scatterers. This is a difficult parameter to quantify but is essential in order to establish the true scattering strength of the particle.

2.6.c Programs used for Processing and Analysis of SAXS Data

PRIMUS

A program suite where the main program PRIMUS (Konarev et al., 2003) used for processing of the small angle scattering data, manipulates and analyses the data. PRIMUS performs averaging, subtraction of background and merging of the SAXS data measured in different angular ranges. The buffer is extracted and the data is extrapolated to zero concentration for selected data files with different sample concentrations (no zero sample concentration). The

radius of gyration, R_g (for rod-type, flat and globular particles) from the Guinier plot, zero intensity, Porod's volume and molecular weight are evaluated. PRIMUS can via pop-up menus analyse characteristic functions by indirect Fourier transformation. It can also perform peak analysis for partially ordered systems and shape approximations in terms of three-parametric geometrical bodies can be found.

DAMMIN

DAMMIN (Svergun, 1999) is used in order to restore *ab initio* low resolution shape. The program also restores internal structure of biological macromolecules in solution from isotropic scattering. DAMMIN fills densely packed particles (small spheres or dummy atoms) into a sphere of diameter equal to the maximun particle size. Each dummy atom should be assigned either to solvent or particle. Next step is the minimization, by which the program searches for compact dummy atoms that fits the SAXS data (Figure 2.9). Simulated annealing (SA) is a minimization method where the temperature is high and the changes are almost random, thereafter it is cooled down. In the end a configuration with nearly minimum energy is reached. Simulated annealing is used in order to find a configuration that fits the data while minimizing the interfacial area (Figure 2.10). Application of the method is to determine the shape of several proteins from experimental x-ray scattering data.
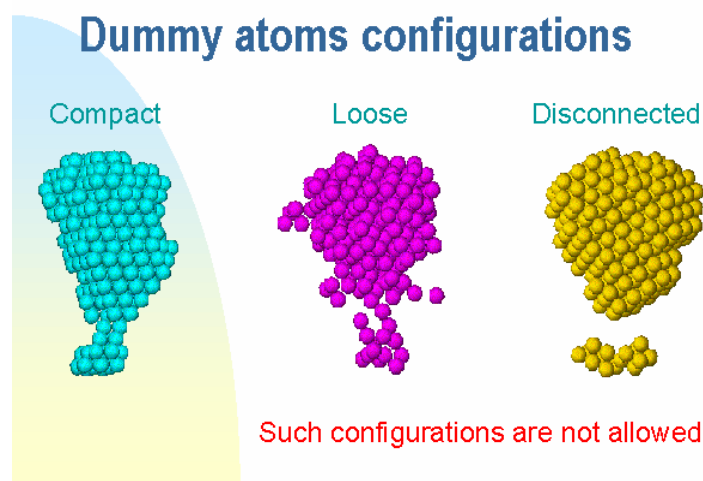

Figure 2.9: An example of dummy atom configurations after minimization
in the range of compact to disconnected arrangement. Figure taken from
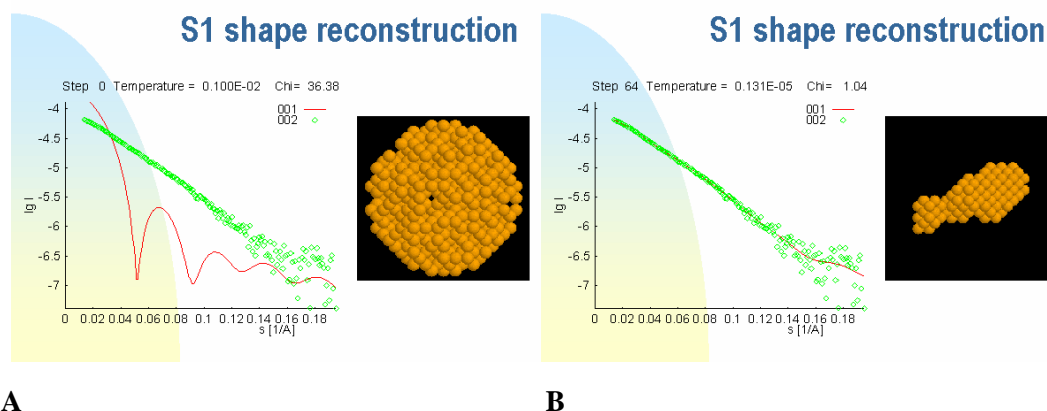(http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin/sld004.htm)

Figure 2.10: The pictures present the shape determination of the s1 subfragment of the myosin head. The graph in **A**) presents the shape reconstruction in the beginning when the fit to the experimental data is bad ($\chi = 36.38$) and the graph in **B**) shows a good fit ($\chi = 1.04$) to the experimental data. The dummy atoms in orange are shown to be very well fitted in B) Figure taken from (http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin/sld006.htm) and (http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin/sld024.htm)

MASSHA

MASSHA (Konarev et al., 2001) is applied for rigid-body refinement and 3D rendering. The program displays and manipulates low resolution models and high resolution atomic structures, which are portrayed as ensembles of beads. The program allows interactive and automated rigid body refinement of the quaternary structure of macromolecular complexes against solution scattering data. Alignment of structural models of different resolution and origin are performed automatically. The low-or high resolution models are refined to fit the experimental solution scattering data.

GLOBSYMM

GLOBSYMM (Petoukhov and Svergun, 2005) is applied for rigid body modelling of symmetric oligomers. The program searches for the quaternary structure of a symmetric oligomer built from monomers. From the high-resolution model the partial scattering amplitudes of the monomer $A_{lm}(s)$ are computed by CRYSOL (Svergun et al., 1995). The monomer is centered at the origin in a reference orientation. The monomer is rotated by the Euler angles (α, β, γ) and thereafter it is substituted by the vector r = (r, ψ, φ), creating a symmetric oligomer. The partial scattering amplitudes of the rotated and shifted monomer $B_{lm}(s)=B_{lm}(s, \alpha, \beta, \gamma, r)$ are analytically expressed via the amplitudes $A_{lm}(s)$ and the elements of the finite rotation matrix. The scattering from the entire oligomer is:

$$I_{calc}(s,r,\alpha,\beta,\gamma) = 2\pi^2(nk)^2 \sum_{1=0}^{\infty} \sum_{m=1}^{1} \left| B_{lm}(s) \right|^2$$

where the summation runs only over m equal to zero or multiples of n and all imaginary parts vanish in the case of k=2. The scattering intensity from the symmetric oligomer can be calculated from the equation above, thereby it is possible to search for the positional and rotational parameters to minimize discrepancy between the experimental data and the calculated curve.

In order to prevent oligomeric structures with steric clashes and those with too loose contacts between the monomers, a crude dummy atom structure of the monomer is produced. This lead to a fast evaluation of the parameters of contacts and the overlap between monomers, thereby the less favorable structures were left out on in the course of minimization.

OLIGOMER

OLIGOMER (Konarev et al., 2003) determines the molecular weights, radii of gyration and volume fractions for each of the mixtures of proteins with known scattering intensities. The program uses experimental scattering curves from a mixture of proteins as input data and determines the form-factors for each component of these mixtures. *I(s)* is the experimental scattering intensity from a mixture of different proteins (components), and mentioned as contributions w_i of each protein $I_i(s)$ to resulting scattering:

$$I(s) = \sum \left(w_i \times I_i(s) \right)$$

w_i is the volume fraction and $I_i(s)$ is the scattering intensity, from the *i*-th component.

If the intensities from the proteins are known, OLIGOMER uses the algorithm of unconstrained or nonnegative least-squares to minimize the deviation between the experimental and calculated scattering curves. The program estimates molecular weights of the proteins (components) in order to determine the radius of gyration and volume fractions

from experimental or model scattering curves of components of the mixture. Therefore, it is important that the proper molecular weight (calculated from the primary sequence) of each protein (component) is entered into the program, otherwise the values of the volume fractions will be incorrect.

Chapter 3

RESULTS

In the beginning of this thesis, no structural information was available on adenylyl cyclases (AC) from *M. tuberculosis*. AC Rv0386 from *M. tuberculosis* is a multidomain protein, comprising a catalytic AC domain at the N-terminus, followed by an AAA-ATPase domain, TPR domain and a helix-turn-helix DNA binding domain at the C-terminus. The catalytic core of AC Rv0386 converts ATP to cAMP and GTP to cGMP, which is extraordinary. Structural studies were applied to increase our understanding of the molecular mechanism of the catalytic core of Rv0386. Several constructs were designed, for overexpression, purification and biophysical characterisation, presented in the first section of this chapter. Crystallisation with native and selenomethionine substituted protein and structure determination with SAD is described in the second part. Small angle X-ray scattering, analytical ultracentrifugation and mutagenesis was applied to verify a unique inhibited conformation of the X-ray structure, presented in the successive section. Additionally, crystals of the catalytic core in the probably active state were included.

3.1 Purification and Biophysical Characterisation of the Catalytic Domain of Rv0386

3.1.a Purification of C-His CHD of Rv0386

The cyclase homology domain of Rv0386 comprising an uncleavable C-terminal histidine tag was expressed in the *Escherichia coli* strains BL21(DE3), BL21(DE3)pREP4 and B834(DE3) as described in section 2.2.a. The highest yield of over expressed protein was obtained with B834(DE3) (not shown). The catalytic domain was purified in two steps, using first affinity chromatography and then size exclusion chromatography. The histidine tag was well exposed and the protein bound with high affinity to the nickel column. After the washing step of 80 mM imidazole, the protein was eluted over a linear gradient (80-500 mM imidazole) with the elution peak at 125 mM imidazole in a pure form, without visible Coomassie stained contaminations (Figure 3.1). The second purification step, size exclusion chromatography was introduced to remove aggregation. Molecular weights given from gel filtration runs are derived from a calibration run with protein standards (Sigma). For Superdex 75 10/30, the aggregated protein gives a peak at the void volume (8.08 ml) and the catalytic domain elutes as a dimer at retention volume $V_E = 10.48$ ml, corresponding to a molecular weight of 45 kD (Figure 3.2). Using a Superdex 75 26/60, the aggregated protein elutes at 126 ml (void volume) and the CHD at the retention volume $V_E = 174$ ml, corresponding to 45 kD. These values are compatible with the calculated molecular weight of 40.4 kD of the dimer of the catalytic domain of Rv0386. The protein yield after gel filtration corresponds to 1 mg per liter bacterial culture, which is sufficient for crystallisation.
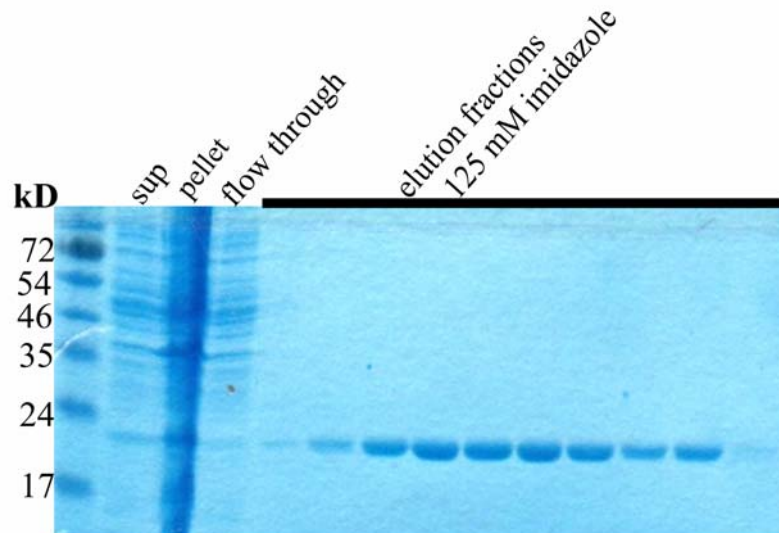
Figure 3.1: Affinity purification of C-His CHD of Rv0386. The 12% Coomassie stained SDS-PAGE shows a protein band at 20 kD, which corresponds to the expected molecular weight of 20 kD of the histidine tagged catalytic domain of Rv0386.
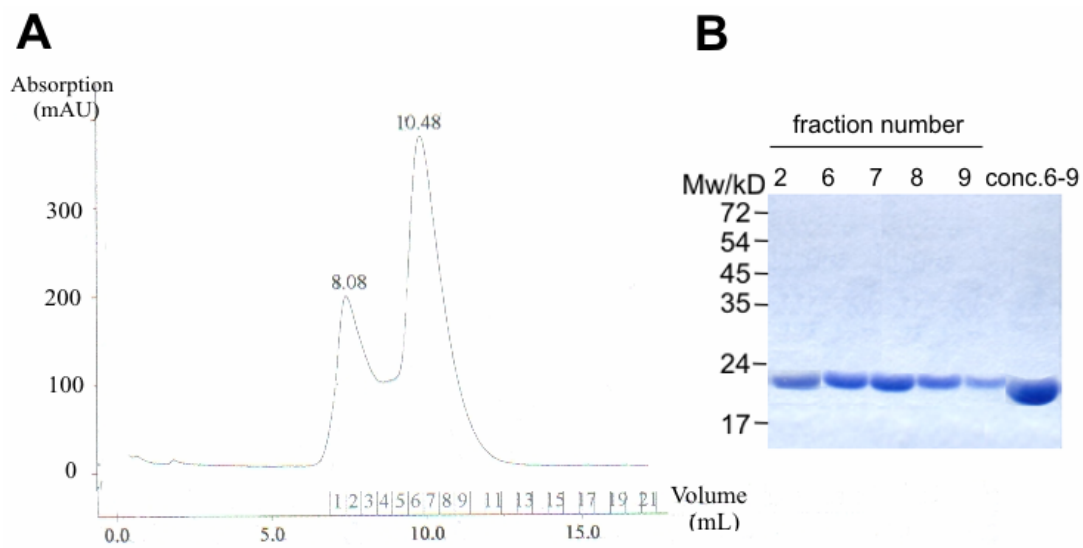
Figure 3.2: Size exclusion chromatography chromatogram of C-His CHD Rv0386. **A)** The main peak at retention volume, V_E = 10.48 ml, corresponds to a molecular weight of 45 kD (calculated Mw of the CHD dimer = 40.4 kD). The aggregation peak is visible at the void volume of 8.08 ml. **B)** Samples from the elution peaks in A are visible on the Coomassie stained SDS-PAGE. C-His CHD Rv0386 is observed as a band at a molecular weight of 20 kD.

3.1.b Purification of N-His CHD of Rv0386

The protein purification was carried out in two steps using affinity chromatography and gel filtration. The N-terminal his tagged CHD of Rv0386 bound with high affinity to the nickel resin and was eluted with a gradient of 50 to 500 mM imidazole. The protein eluted in two peaks, the first between 150 and 250 mM imidazole and second between 250 and 500 mM imidazole. The second elution peak gave rise to protein that precipitated easily at low salt concentrations and at higher protein concentrations. The eluted catalytic domain was in a pure form as shown on the SDS-PAGE (Figure 3.3). The protein yield is 18 mg per liter bacterial culture. Size exclusion chromatography was performed to analyse the oligomeric state and to separate aggregated protein. The catalytic domain has a tendency to precipitate at salt concentrations lower than 300 mM, and it was best to keep a minimal glycerol concentration of 10 %. Gel filtration over Superdex 75 26/60 visualises the oligomerisation state of the N-terminal histidine tagged catalytic domain of Rv0386 (Figure 3.4). The aggregated protein eluted at the void volume 111 ml and the dimeric catalytic domain has a retention volume of V_E=174 ml, corresponding to the molecular weight of 40 kD which is close to the theoretical molecular weight of 41.4 kD. Tetrameric, octameric and higher oligomeric species eluted at V_E=150 ml corresponding to a molecular weight of 80 kD, 141 ml to 160 kD and 124 ml to 250 kD respectively.
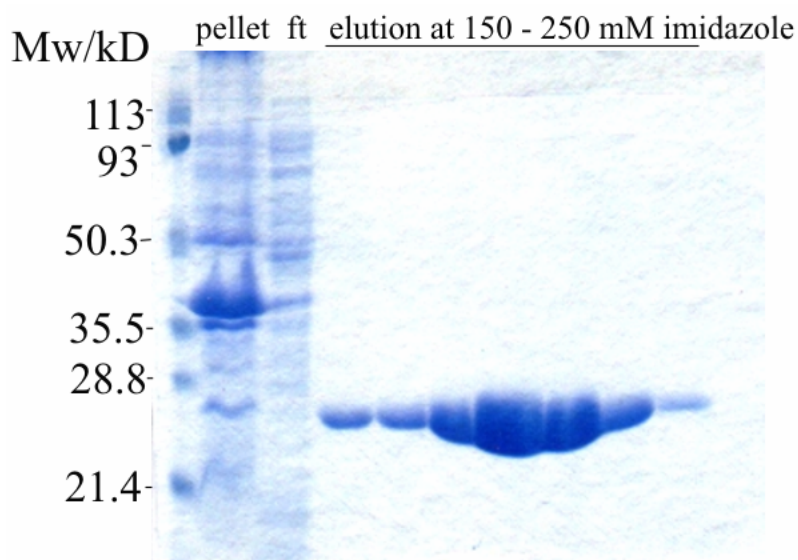

Figure 3.3: Affinity chromatography of the N-His CHD of Rv0386. The protein was analysed by SDS-PAGE and stained with Coomassie Brilliant Blue. Ft= flow through.
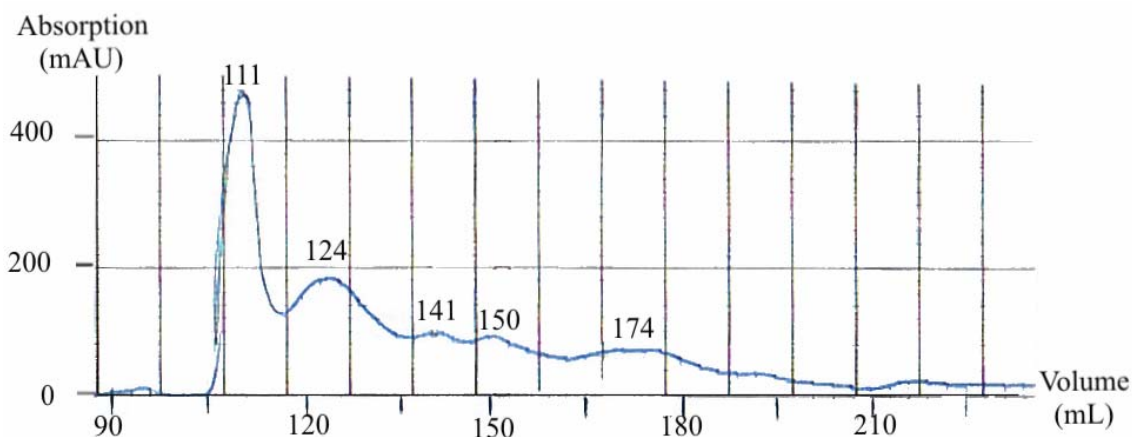
Figure 3.4: Size exclusion chromatography of the N-His CHD of Rv0386. The protein oligomerises and the dimeric peak is visible at retention volume, V_E = 174 ml, corresponding to the Mw of 40 kD, which is close to the expected Mw of 41.4 kD.

3.1.c Purification of SeMet Substituted Rv0386 CHD

Over expression, purification and crystallisation of selenomethionine substituted C-His CHD Rv0386 was carried out in order to obtain phase information for a SAD experiment. In most cases, the over expression of selenomethionine substituted protein gives a very low yield, due to the slow growth in minimal media. In contrast, over expression of SeMet substituted C-His CHD Rv0386 was very successful, producing high amount of protein in short time. The bacterial cells grown in SeMet substituted minimal media behaved as during native over expression. The cells reached an optical density of 0.7, within 4 hours. During the purification of the SeMet substituted C-His CHD of Rv0386 the buffers were intensely degassed, in order to decline the risk of oxidation of the SeMet protein. During affinity chromatography the protein eluted at 125 mM imidazole in an almost pure form. The catalytic domain is observed on the Coomassie stained SDS-PAGE with a molecular weight of approximately 20 kD (Figure 3.5). The size exclusion chromatography separates the aggregated protein from the dimeric form (Figure 3.6). The former elutes at the void volume 121 ml and the latter at retention volume V_E= 167 ml, corresponding to a dimer of Mw of 40 kD, compatible to the expected molecular weight of 40.4 kD. The protein was analysed on a Coomassie stained SDS-PAGE, where a single band of the catalytic domain can be observed. The protein yield after size exclusion chromatography is 5 mg per liter bacterial culture.
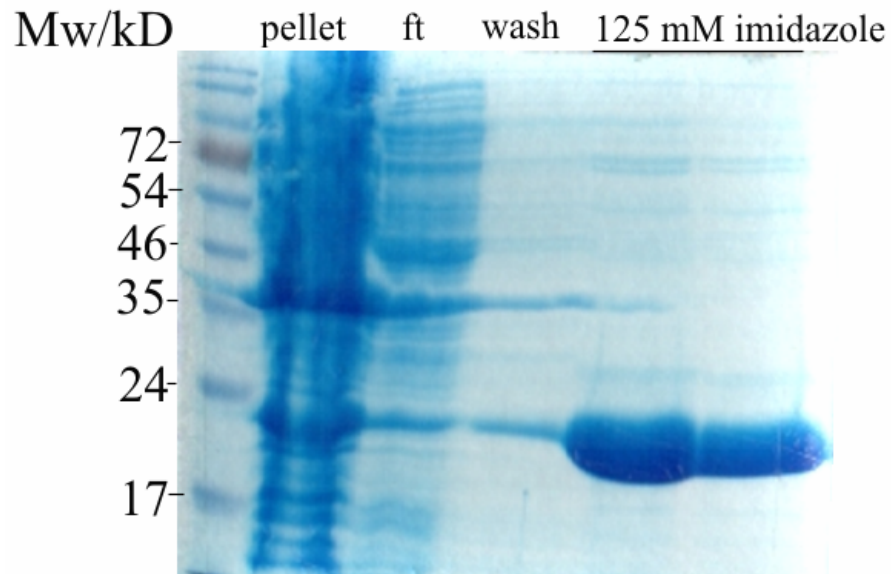
Figure 3.5: Affinity chromatography of SeMet substituted C-His CHD of Rv0386. The protein is eluted in an almost pure form and in high amount observed on the Coomassie stained SDS-PAGE.
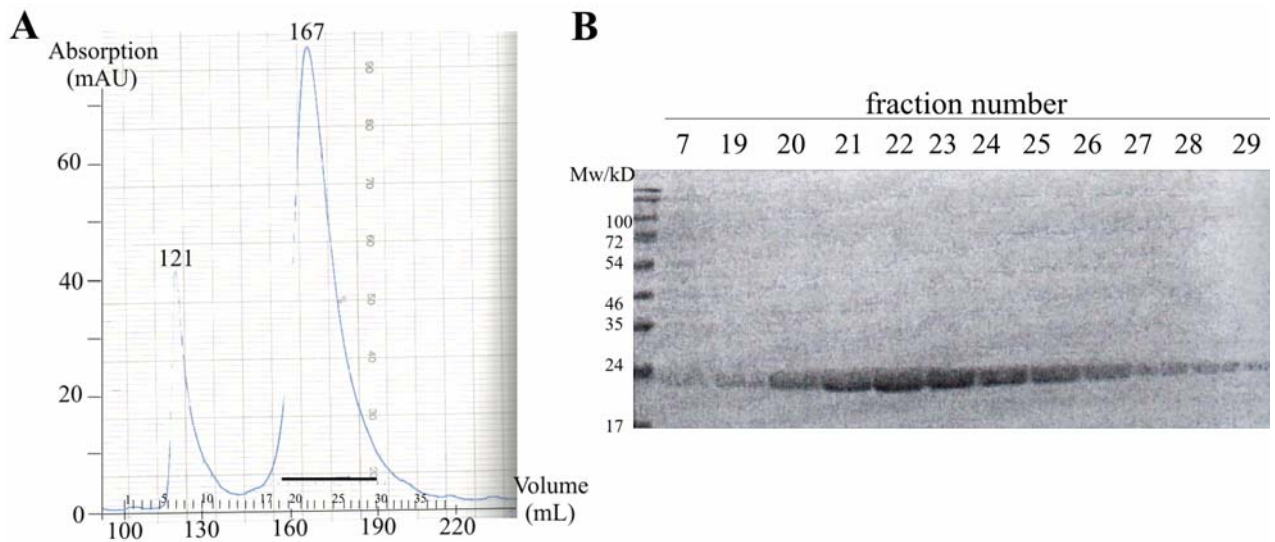
Figure 3.6: Size exclusion chromatography of SeMet substituted C-His CHD Rv0386. **A)** The main peak is at retention volume V_E=167 ml, corresponding to a dimer of a molecular weight of 40 kD, compatible to the expected value of 40.4 kD. Aggregated protein elutes at void volume 121 ml. **B)** SeMet substituted protein is in a pure state after gel filtration as observed on the Coomassie stained SDS-PAGE.

3.1.d Activity Studies

AC Rv0386 is a low activity protein compared to other ACs from *M. tuberculosis*. Exceptionally, the enzyme converts ATP and GTP to cAMP and cGMP, respectively. The catalytic domain of Rv0386 exhibits 20 % guanylyl cyclase side-activity (Castro, 2004; Castro et al., 2005).

The V_{max} values of the Rv0386 C-His CHD and the Rv0386 holoenzyme are 5.6 and 3 nmol cAMP·mg^{-1}·min^{-1}, respectively (in the presence of 850 µM ATP). In contrast, the V_{max} of cyclase homology domains of Rv1900c and Rv1264 is 180-fold higher compared to the Rv0386 C-His CHD. The Hill coefficient of 0.9 for C-His CHD of Rv0386 indicates no cooperativity. The low activity could possibly be stimulated by the other regulatory domains of Rv0386 and by an unknown external effector (Castro, 2004; Castro et al., 2005).

Oligomerisation of the Catalytic AC Domain of Rv0386

During the protein dependence measurement of N-His CHD Rv0386, it was noted that protein specific activity increased 5-fold at protein concentrations between 46.5 to 465 nM. In contrast, higher protein concentrations resulted in a slight decrease of the specific activity. This suggests a rapid construction of a homodimer with formation of oligomers induced by increasing protein concentrations (Figure 3.7). The catalytic domain of Rv0386 oligomerises in solution forming monomers, dimers, trimers, tetramers and hexamers as shown by SDS-PAGE followed by western blot analysis. The sample of N-His CHD of Rv0386 applied for the western blot experiment had AC and GC activities of 3.7 and 1.1 nmol cAMP·mg^{-1}·min^{-1}, respectively. The dissociation constant (Kd) was determined to 140 nM, indicative of a high affinity of the catalytic domains for each other. The Kd of C-His CHD was determined to 1.5 µM, which indicates a slower oligomerisation process of this construct compared to N-His CHD (Castro, 2004).
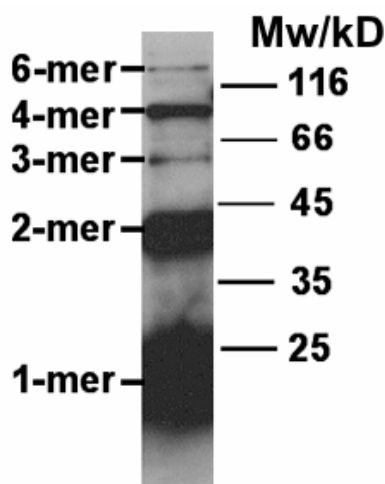
Figure 3.7: Western blot analysis of the oligomeric state of the cyclic homology domain of Rv0386. The primary antibody applied is a specific antibody for the catalytic domain of Rv0386. Picture taken from (Castro, 2004).

Activity studies of the catalytic core of Rv0386 detected low activity for all constructs (Table 3.1). N-His CHD and C-His CHD have a V_{max} of 4.8 and 3.9 nmol cAMP·mg^{-1}·min^{-1}, respectively. The C-His CHD has 81% AC activity of the N-His CHD. In order to analyse whether a shorter C-terminus influences the activity, constructs with truncated C-terminus were expressed for kinetic measurements. Measured AC activity for Rv0386 CHD$_{(1-172)}$ was 30% compared to N-His CHD. The mutations are not substituting the substrate binding residues and should therefore not influence the activity. No AC and GC activity could be detected for Rv0386 CHD$_{(1-172)}$ Cys168Ser and Rv0386 CHD$_{(1-168)}$ C168S indicating that the C-terminus is important for the activation.

No activation by forskolin and no inhibition by P-site inhibitors was observed, which is in agreement with soluble ACs from *M. tuberculosis* (Castro, 2004; Castro et al., ; Linder, ; Sinha et al., 2005). The divalent cofactors Mn^{2+} and Mg^{2+} were tested *in vitro* and it was found that Mn^{2+} improves the activity of the AC Rv0386 (Castro, 2004; Castro et al., 2005) which is in agreement with all other ACs tested from *M. tuberculosis* (Guo et al., 2001; Linder et al., 2002).

Protein	µM in assay	AC activity nmol/mg/min
C-His CHD Rv0386	0.2	2.1
	0.5	3.2
	1.4	3.5
N-His CHD Rv0386	0.2	2.7
	0.5	4.0
	1.4	4.3
Rv0386 CHD$_{(1-172)}$	0.2	0.00
	0.5	1.10
	1.4	1.40
Rv0386 CHD$_{(1-172)}$ Cys168Ser	0.2	0.00
	0.5	0.00
	1.4	0.30
Rv0386 CHD$_{(1-168)}$ C168S	0.2	0.00
	0.5	0.00
	1.4	0.20

Table 3.1: AC activities of AC constructs under identical conditions. Assay performed with MOPS buffer pH 7.5, 5 mM Mn^{2+}, 850 µM ATP, for 20 min at 30°C. Detection limit was 0.15 nmol/mg/min.

3.1.e Purification of Rv0386 C-His CHD as Monomer

The cofactor Mn^{2+} is important for the activity of the AC Rv0386 as shown by (Castro et al., 2005). Therefore, the metal ion Mn^{2+} was included in the purification and crystallisation of C-His CHD Rv0386. Mn^{2+} was added after the affinity chromatography step to the dialysis buffer, prior to size exclusion chromatography. When $MnCl_2$ is added, the protein has a tendency to aggregate and precipitate. To prevent aggregation 10% glycerol and 500 mM NaCl was added which kept the protein stable. Surprisingly, when ≥0.5 mM $MnCl_2$ is added to the gel filtration buffer, the catalytic domain is observed as a monomer. In comparison, magnesium chloride ($MgCl_2$) does not affect the catalytic domain of Rv0386 in the same way and the protein stays dimeric. The protein behaves identical with $MgCl_2$ or without any addition of cofactor.

Affinity chromatography over a nickel column gives an almost pure protein, as evaluated by SDS-PAGE (Figure 3.8). The molecular weight of 24 kD corresponds to the theoretical molecular weight of 20 kD of the catalytic domain of Rv0386. The gel filtration column Superdex 75 10/30, gives a peak at the void volume (8.20 ml) of aggregated protein and the

catalytic domain elutes as a monomer at retention volume $V_E = 11.77$ ml, corresponding to a molecular weight of 22 kD (Figure 3.9). Aggregated protein elutes at the void volume 120 ml of Superdex 75 26/60 and the CHD at the retention volume $V_E = 195$ ml, corresponding to 22 kD. This value is compatible to the theoretical molecular weight of 20 kD of the monomer of the CHD of Rv0386. After gel filtration the protein yield was 2.5 mg per liter bacterial cell culture. In summary, the oligomeric state of the cyclase homology domain was changed upon addition of the cofactor Mn^{2+} which was unexpected.
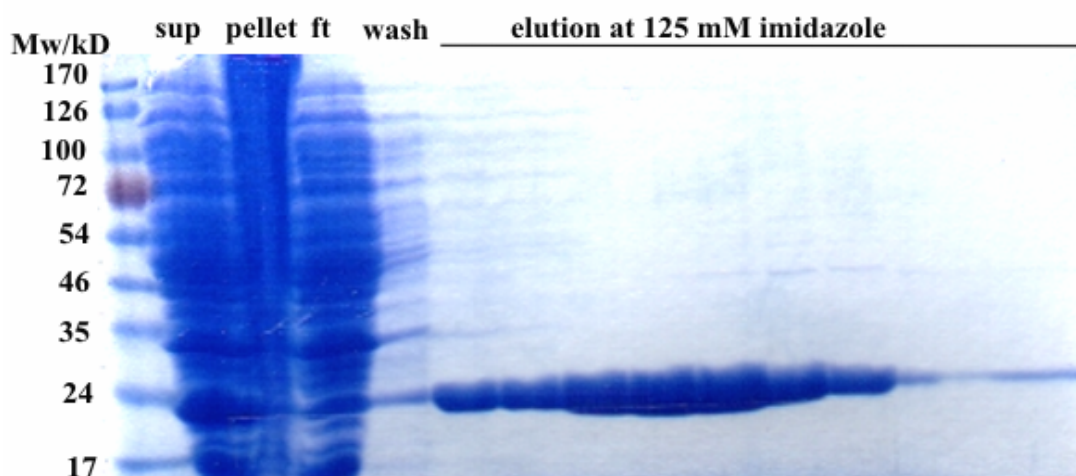
Figure 3.8: Affinity chromatography of C-His CHD of Rv0386. The catalytic domain elutes from the nickel column in an almost pure form. The molecular weight at the SDS-PAGE gives a strong band at 24 kD, which corresponds to the expected Mw of 20 kD for the catalytic domain of Rv0386. sup = supernatant, ft = flow through.
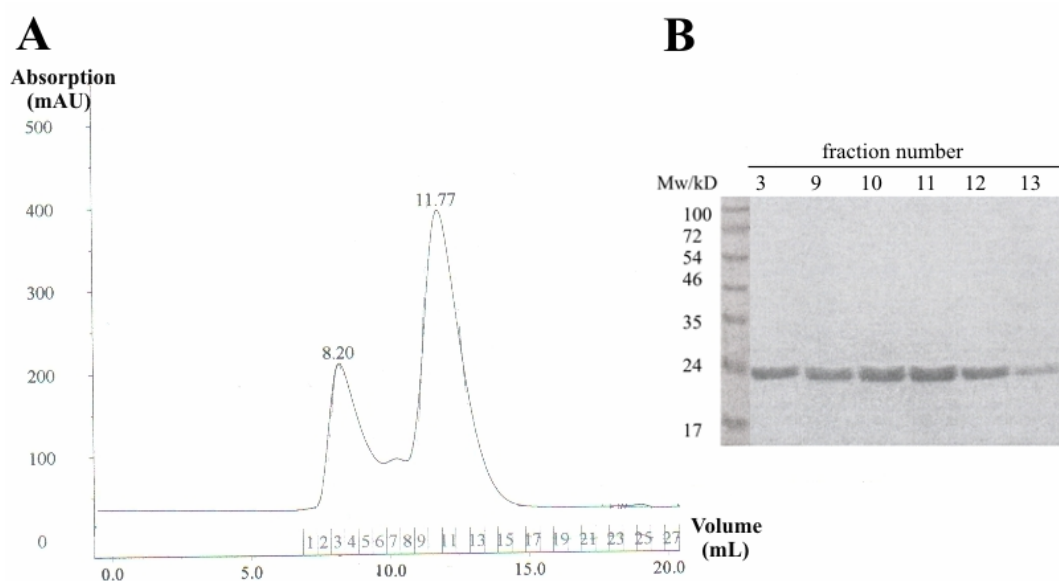
Figure 3.9: Size exclusion chromatography of C-His CHD of Rv0386 eluted as monomer. **A)** The aggregation peak is visible at the void volume of 8.2 ml of Superdex 75 10/30. The main peak at retention volume, $V_E = 11.77$ ml corresponds to a molecular weight of 22 kD, agreeing with an expected monomer of Mw of 20 kD. **B)** Samples were analysed on Coomassie stained SDS-PAGE.

3.1.f Degradation Test of CHD Rv0386

The N-His CHD of Rv0386 was left for 5 days in 4°C. This test was carried out to analyse the protein stability after leaving the protein in the refrigerator a couple of days, prior to crystallisation. After 5 days the catalytic domain has formed degradation bands as observed on the Coomassie stained SDS-PAGE in (Figure 3.10). Thus, the protein is not suitable for crystallisation after a few days of storing at 4°C.
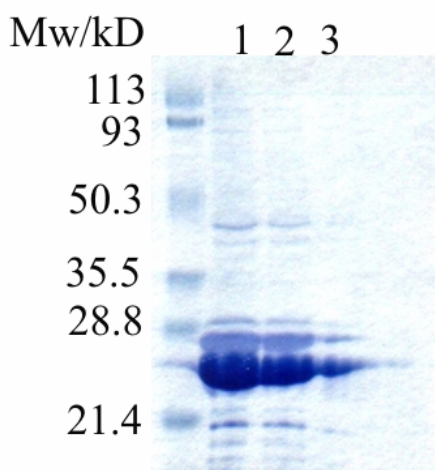
Figure 3.10: Degradation test of the CHD of Rv0386. The protein is analysed on the SDS-PAGE after 5 days in 4 °C. Several degradation bands and minor contaminations are detected. Lane 1: 40 µg, Lane 2: 20 µg, Lane 3: 10 µg protein.

3.1.g Mass Spectrometry

In order to determine if the cyclase homology domain of Rv0386 is fully intact, electrospray ionisation mass spectrometry (ESI-MS) was chosen. This technique is very sensitive and accurate and the main species of the catalytic domain of Rv0386 was determined to the molecular weight of 20121 ± 0.02 Da (Figure 3.11), which is identical to the expected value of the demethionised CHD Rv0386. The minor species comprising the molecular weight of 40 242 ± 0.02 Da, corresponds to the dimeric demethionised CHD Rv0386 with the Mw of 40 242 Da.
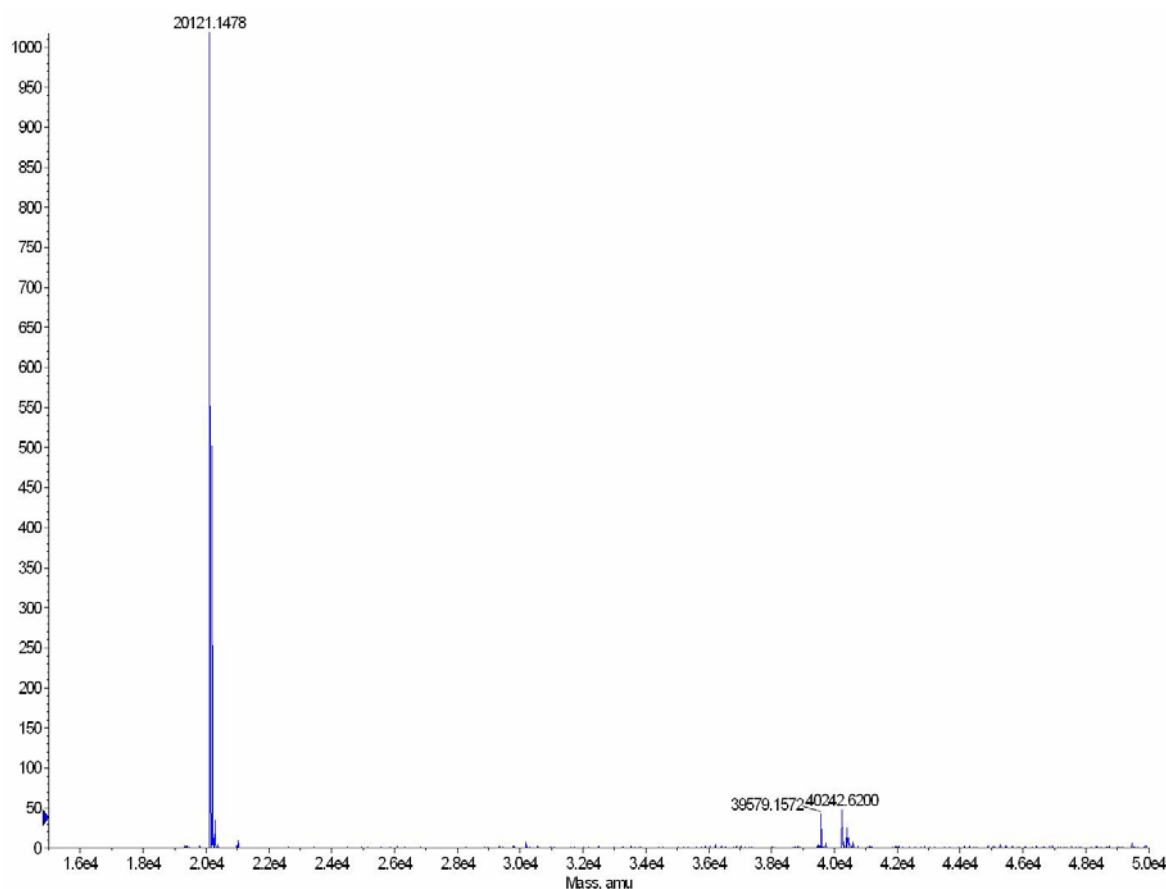
Figure 3.11: Electrospray ionisation mass spectrometry of the catalytic domain of Rv0386. The main peak at molecular weight of 20 121 Da correspond to the theoretical value of the demethionised catalytic domain. The minor peak at 40242 Da, matches the expected molecular weight of the dimeric demethionised catalytic domain.

3.1.h Dynamic Light Scattering

Dynamic light scattering was carried out as an entry test for crystallisation. For a poly dispersity of <20-25% the protein is regarded as mono disperse and has a good chance to crystallise due to homogeneity and a single species in solution. The C-His CHD of Rv0386 was purified with affinity- and size exclusion chromatography and the measurement was carried out at protein concentration of 1mg/ml, in buffer comprising 50 mM Tris pH 8.5, 500 mM NaCl, 2 mM BME, 10% glycerol, at 12 °C. Prior to the measurement aggregation was removed by ultra centrifugation at 100 000 g for 45 to 60 min at 4°C. The protein is mono disperse with a poly dispersity of 20.4 % and a molecular weight of 47 kD (Figure 3.12),

corresponding to a dimer with the theoretical molecular weight of 40 kD. The cyclase homology domain of Rv0386 is homogeneous and suitable for crystallisation.
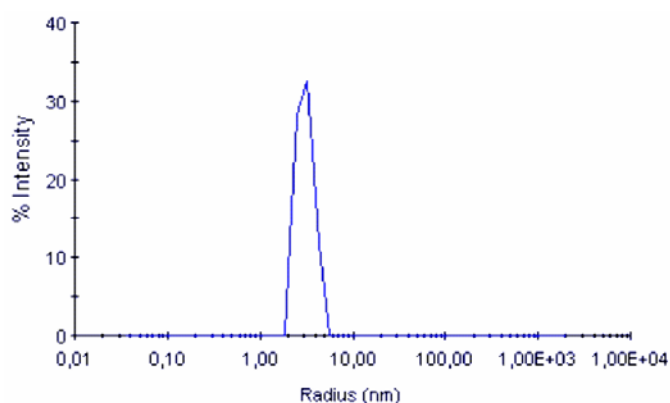
R(nm)	PD(%)	Mw(kD)	Mass(%)	Baseline	SOS-error	Intensity (counts/s)
3.1	20.4	47	100	1.000	210	79 000

Figure 3.12: Dynamic light scattering of the catalytic core of Rv0386. The protein is mono disperse with a molecular weight of 34 kD, corresponding to a dimer, compatible to the expected Mw of 40 kD. R=hydrodynamic radius, PD=poly dispersity, Mw=molecular weight, SOS=sum of squares.

3.2 Structure Determination of the Catalytic Core of Rv0386

3.2.a Crystallisation of the Catalytic Domain of Rv0386

The catalytic domain of Rv0386 was crystallised in several crystal forms with the hanging drop vapour diffusion method, with protein comprising the N-terminal or C-terminal histidine tag (N-His CHD or C-His CHD). There were a number of different morphologies observed (Table 3.2). Generally, crystals were grown within 24 h to a couple of days. The crystals were tested at the synchrotron radiation source (ESRF, Grenoble, France) for diffraction quality. None of these initial crystals diffracted. Further screening of different morphology crystals,

resulted in a breakthrough. Small, tiny rod-like crystals of the C-His CHD Rv0386 grown in the condition 0.5 M Sodium Acetate, 0.1 M imidazole pH 6.5 were tested at ESRF, Grenoble with the diffraction of 8 Å (Table 3.2).
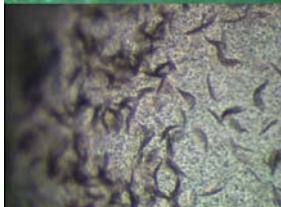
Construct	Condition	Diffraction in ESRF, Grenoble	Crystals
N His CHD Rv0386	10 % PEG 6000 2 M NaCl 9% Glycerol (Castro, 2004)	No diffraction	
N His CHD Rv0386	10 %PEG 8000 0.1 M CHES pH 9.5 0.2 M NaCl	No diffraction	
C His CHD Rv0386	15% PEG 4000 0.1 M Tris HCl pH 8.5 0.2 M Sodium Acetate	No diffraction	
C His CHD Rv0386	0.5 M Sodium Acetate 0.1 M imidazole pH 6.5 (Castro, 2004)	Diffraction to 8 Å	

Table 3.2: Crystals of N-His and C-His CHD Rv0386 grown in different conditions.

Fine screening with other buffers but at identical pH, lyotropic and chaotropic precipitants of the Hofmeister series, most commonly used precipitants, additives and inhibitors were applied to improve the crystals. Additionally, the protein purification was modified where a size exclusion chromatography step was included. Furthermore, the complete protein purification procedure was carried out more rapidly. Within one day, the cells were broken up, affinity- and size exclusion chromatography were carried out, including incubation of concentrated protein with ATP analogue. The following day crystallisation trials were set up. Crystals increased in thickness and size in several conditions. Protein solution of 1.5 µl (10-20 mg/ml) was mixed with 1.5 µl crystal condition of 0.1 M Hepes pH 7.5 and 0.7 M sodium acetate. After 2-4 days at 19 °C the rod shaped crystals grew to a size of 30 x 30 x 250 µm^3 (Figure

3.13) and a complete native dataset was collected at ESRF, beam line ID 14-4 to a resolution of 2.95 Å.
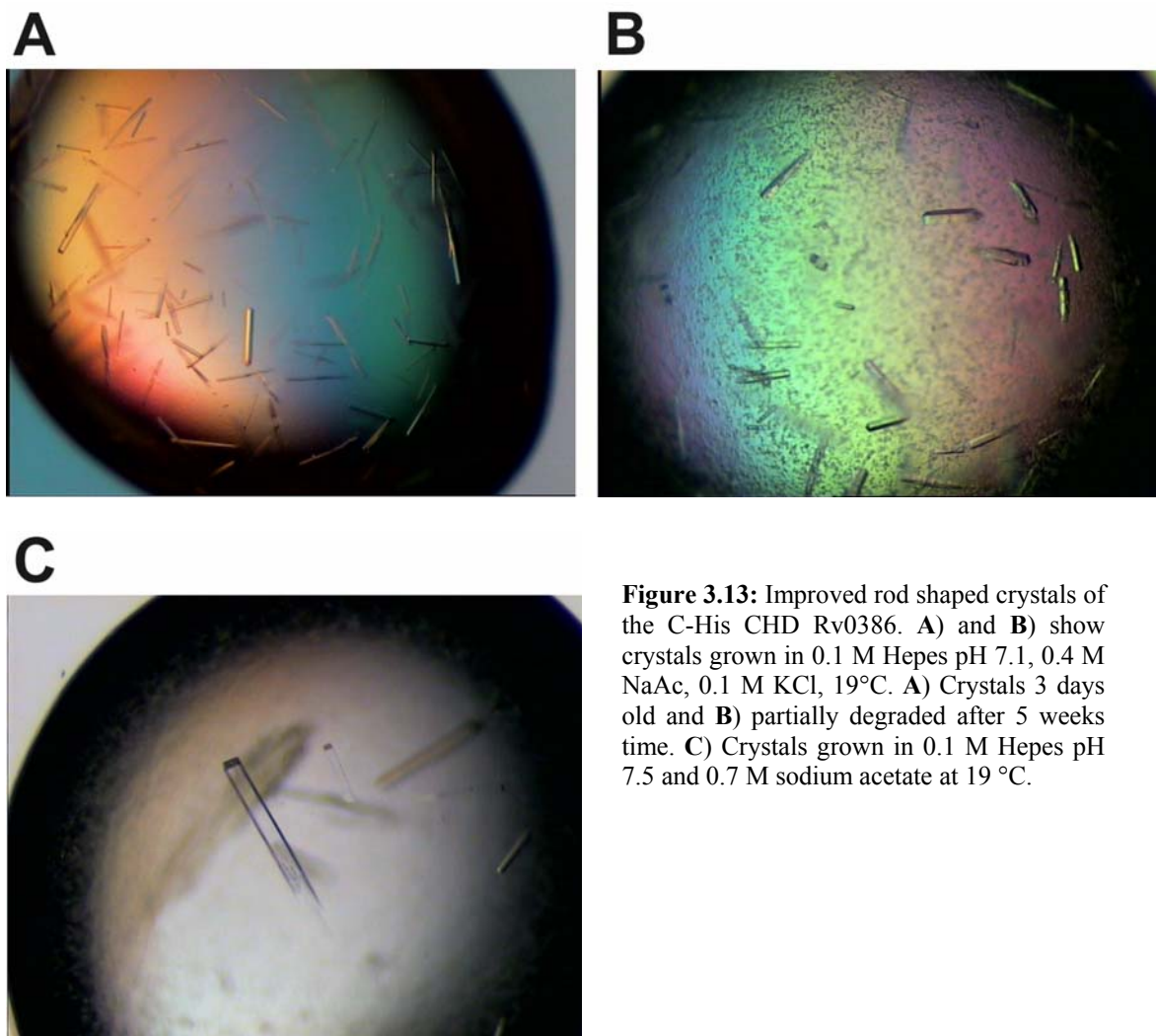
Figure 3.13: Improved rod shaped crystals of the C-His CHD Rv0386. **A)** and **B)** show crystals grown in 0.1 M Hepes pH 7.1, 0.4 M NaAc, 0.1 M KCl, 19°C. **A)** Crystals 3 days old and **B)** partially degraded after 5 weeks time. **C)** Crystals grown in 0.1 M Hepes pH 7.5 and 0.7 M sodium acetate at 19 °C.

3.2.b   Crystallisation of the Se-Met Modified Catalytic Domain of Rv0386

In order to obtain phase information for a SAD experiment selenomethionine substituted C-His CHD was crystallised. The selenomethionine modified protein behaved comparable to the native protein and crystallised in similar conditions with the hanging drop vapour diffusion method. 1.5 µl of protein solution (14 mg/ml) mixed with 1.5 µl of 0.6 M NaAc, 0.1 M Hepes pH 7.25, set up for crystallisation at 16 °C resulted in rod shaped crystals. Initial data collected were insufficient to determine the structure, due to the limited size of the crystals obtained. Larger crystals were grown from PEG2000. Protein solution of 1.5 µl (17-20 mg/ml) was mixed with 1.5 µl of 5% (w/v) PEG2000, 0.1 M Hepes pH 7.5 and 0.2 M KCl and set up for crystallisation at 19 °C. After 2-4 days the Se-Met substituted crystals grew to a size of 30 x 30 x 250 µm^3 with sharp and distinct edges of rod shaped crystals (Figure 3.14).
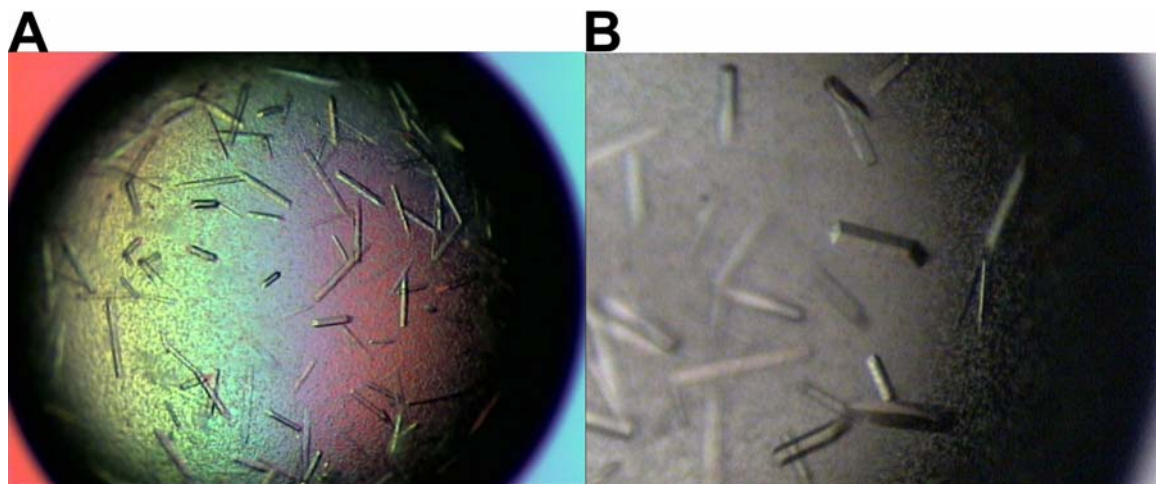
Figure 3.14: Improved Se-Met modified crystals. **A)** Thicker crystals grown in the condition of 0.6 M NaAc, 0.1 M Hepes pH 7.25 at 16 °C. **B)** Seleno methionine containing crystals with sharp edges grown in 5% (w/v) PEG2000, 0.1 M Hepes pH 7.5 and 0.2 M KCl at 19°C.

3.2.c Structure Determination of the Catalytic Domain of Rv0386

The complete dataset of the PEG2000 selenomethionine modified crystals were collected at ESRF, beam line ID 14-1 (Grenoble). The crystals were collected in one sweep, over the oscillation range of 180 °. Statistics of the data collection are given in Table 3.3. The crystals belong to the space group P4$_1$22 with unit cell dimensions of a* = b* = 82.90 Å and c* = 275.18 Å. The asymmetric unit contains 4 molecules. Measured reflections were integrated and scaled using Denzo and Scalepack (Otwinowski and Minor, 1997). I tried to solve the structure of the catalytic core of Rv0386 with molecular replacement, employing monomeric, dimeric, and truncated models from a number of homologous models. The sequence identity was typically as low as 14%. The structure could not be solved with molecular replacement, possibly due to low sequence identity and slight differences in orientation of the α-helices. The crystal structure of CHD Rv0386 was finally solved with the SAD phasing technique. 8 out of 12 selenium sites were identified in the four chains resulting in a map with the figure of merit (FOM) of 0.31 with the program SOLVE (Terwilliger and Berendzen, 1999). In all four chains the Met1 is removed due to post-translational modification, which explains the lacking selenium sites. Phasing was carried out with the program SHARP (de la Fortelle and Bricogne, 1997), and phases were improved by solvent flattening with Solomon (Abrahams and Leslie, 1996). The Rv0386 catalytic domain was initially built in O (Jones et al., 1991) as a poly-alanine model, which was necessary to start the refinement. Later on, as the phases

improved and R-factor converged, the side-chains were added. Iterative model building and refinement using the programs O (Jones et al., 1991) and Refmac 5 (Murshudov et al., 1999) lead to the current model comprising four chains with a few residues at the N- and C-terminus omitted but without gaps in the protein chain. The final map had a figure of merit of 0.8. Refinement statistics gave an R_{free} of 30.6 % and an R_{factor} of 22.1 % (Table 3.4). The quality of the final refined model was checked by PROCHECK (Figure 3.15), WHAT IF and SFCHECK (Figure 3.16). The model is shown in a well refined state.

Data collection

	Native	SeMet
Unit cell	a* = b *= 82.9 Å c* = 275.2 Å $\alpha = \beta = \gamma = 90°$	a* = b* = 82.9 Å c* = 275.2 Å $\alpha = \beta = \gamma = 90°$
No of chains per AU	4	4
Space group	$P4_122$	$P4_122$
Resolution (Å)	25 – 2.95	25 - 3.10
Mosaicity	0.23	0.31
Unique reflections	21091	18397
Redundancy	12.8	19
Completeness (%)	99.9	99.8
R_{merge}*(%)	10.6	6.9
I / σ	18.7	20
HR shell in Å	3.06 - 2.95	3.15 - 3.10
R_{merge}* (HR) in %	47.9	26.6
I / σ (HR)	4.5	6.3

Table 3.3: Data collection statistics of the native and selenomethionine modified crystals of the catalytic domain of Rv0386. The former crystals grown in NaAc diffracted to 2.95 Å and the latter grown in PEG2000 diffracted to 3.1 Å.

*$R_{merge} = \Sigma_{hkl} \Sigma_i | I_i (hkl)- <I(hkl)> | / \Sigma_{hkl} \Sigma_i I_i (hkl)$; where HR corresponds to the highest resolution shell.

Refinement statistics

No. of amino acids	A: 3-173 B: 3-173 C: 4-183 D: 5-175
No. of protein atoms	5259
No. of water atoms	287
RMSD bond (Å)	0.016
RMSD angle (°)	1.788
R_{free} [‡] (%)	30.6
R_{factor} [§] (%)	22.1
Ramachandran (%)	
Residues in most favoured regions	84.0
Residues in additional allowed regions	15.7
Residues in generously allowed regions	0.3
Residues in disallowed regions	0.0

Table 3.4: Refinement statistics of the catalytic domain of Rv0386.

[§]$R_{factor} = \Sigma \mid F_o - F_c \mid / \Sigma F_o$, where Fo and Fc are observed and calculated structure factors respectively.

[‡]R_{free} is calculated with 5 % of the data.

A)

Residues in most favoured regions [A,B,L]	500	84.3%	
Residues in additional allowed regions [a,b,l,p]	91	15.3%	
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.3%	
Residues in disallowed regions	0	0.0%	
	----	-------	
Number of non-glycine and non-proline residues	593	100.0%	
Number of end-residues (excl. Gly and Pro)	8		
Number of glycine residues (shown as triangles)	60		
Number of proline residues	32		

Total number of residues	693		

B Main-chain parameters

Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Comparison values Band width	No. of band widths from mean	
a. %-tage residues in A, B, L	593	84.3	66.5	10.0	1.8	BETTER
b. Omega angle st dev	684	8.7	6.0	3.0	0.9	Inside
c. Bad contacts / 100 residues	7	1.0	20.0	10.0	-1.9	BETTER
d. Zeta angle st dev	633	2.2	3.1	1.6	-0.6	Inside
e. H-bond energy st dev	407	0.7	1.0	0.2	-1.6	BETTER
f. Overall G-factor	693	-0.2	-0.8	0.3	1.9	BETTER

C Side-chain parameters

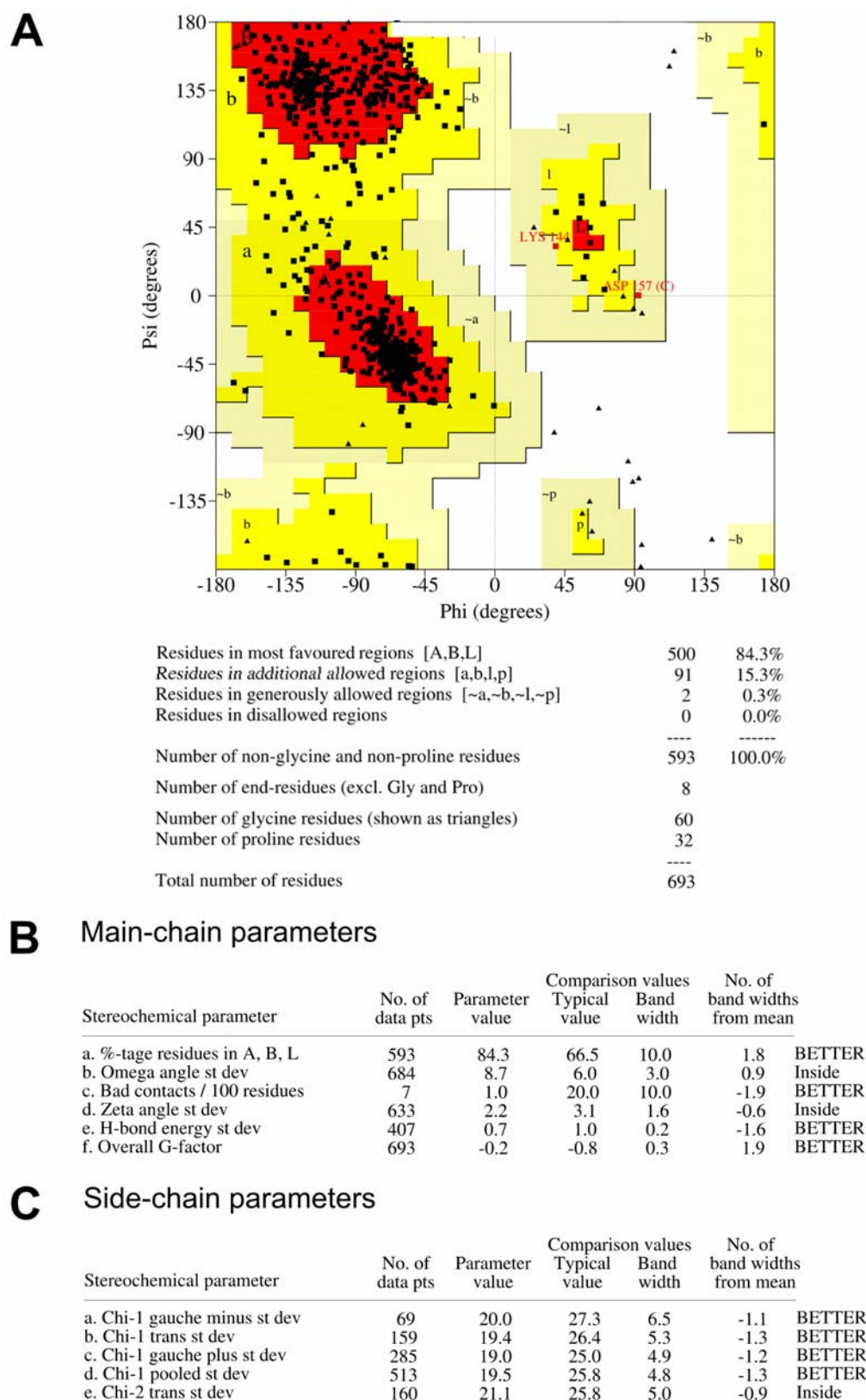| Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Comparison values Band width	No. of band widths from mean	
a. Chi-1 gauche minus st dev	69	20.0	27.3	6.5	-1.1	BETTER
b. Chi-1 trans st dev	159	19.4	26.4	5.3	-1.3	BETTER
c. Chi-1 gauche plus st dev	285	19.0	25.0	4.9	-1.2	BETTER
d. Chi-1 pooled st dev	513	19.5	25.8	4.8	-1.3	BETTER
e. Chi-2 trans st dev	160	21.1	25.8	5.0	-0.9	Inside

Figure 3.15: Stereochemical parameters of the final model. **A)** Ramachandran plot of the catalytic core of Rv0386 including the **B**) main-chain parameters and **C**) side-chain parameters, generated by PROCHECK.

Crystal

Cell parameters:

a: 82.90 Å	b: 82.90 Å	c: 275.18 Å
α: 90.00°	β: 90.00°	γ: 90.00°

Space group: P 41 2 2

Model

5547 atoms (288 water molecules)

Number of chains:	5
Volume not occupied by model:	47.0 %
<B> (for atomic model):	28.8 Å ^{2}
σ(B):	8.74 Å ^{2}
Matthews coefficient:	3.04
Corresponding solvent % :	59.27

Refinement

Program:	REFMAC 5.2.0019
Nominal resolution range:	25.00 – 2.95 Å
Reported R–factor:	0.221
Number of reflections used:	20003
Reported Rfree:	0.31
Sigma cut–off:	N.A.

Structure Factors

Input

Nominal resolution range:	24.91 – 2.95 Å
Reflections in file:	21091
Unique reflections above 0:	21091
above 1σ:	20679
above 3σ:	18595

SFCHECK

Nominal resolution range:	24.91 – 2.95 Å

(max. from input data, min. from author)

Used reflections:	21091
Completeness:	99.9 %
R_stand(F) = <σ(F)>/<F> :	0.040

Anisotropic distribution of Structure Factors
    ratio of eigen values: 1.0000 1.0000 0.9628

B_overall (by Patterson):	46.4 Å ^{2}
Optical resolution:	2.14 Å
Estimated minimal error:	0.024 Å

Model vs. Structure Factors

R–factor for all reflections:	0.298
Correlation factor:	0.829
R–factor:	0.302
for F > 2.0σ	
nom. resolution range:	25.00 – 2.95Å
reflections used:	20677
Rfree:	0.371
Nfree:	1063
R–factor without free–refl.:	0.299
Non free–reflections:	19614
<u> (error in coords by Luzzati plot):	0.438 Å
Estimated maximal error:	0.174 Å
DPI:	0.463 Å

Scaling

Scale:	0.417
Bdiff:	1.64

Anisothermal Scaling (Beta):
  6.2105 6.2105 5.6782 0.0000 0.0000 0.0000

Solvent correction – Ks,Bs: 0.690 250.255

Figure 3.16: The quality of the structure factor data of the catalytic domain of Rv0386 verified by SFCHECK.

3.2.d Overall Crystal Structure

We obtained tetragonal crystals that diffracted to 2.95 Å resolution. The structure of the cyclic homology domain of Rv0386 (Figure 3.17 A, B, C) was solved by SAD. The asymmetric unit contains four molecules (A, B, C, D) forming two homodimers (AB and CD) (Figure 3.18). Each homodimer comprises two monomers oriented in a head-to-head arrangement, where the catalytic residues do not form active sites, thus this conformation is assigned here as the inhibited state of CHD Rv0386. The refined model (Table 3.4) of the structure included complete chains for residues 3-173 for chains A and B, residues 4-183 for chain C and amino acid 5-175 for chain D. A flexible loop region ranging from residue 99-109 is partially disordered in all four chains. There are two further extensive contacts between monomers that belong to different head-to-head dimers. Chain C contains an entirely ordered C-terminal his_6-tag (residues 178-183) and it binds between symmetry related molecule A (A') and chain B where it occupies part of the conserved active site of molecule A'. Further, residues 150-173 in chain A are exchanged towards the residues in the related symmetry molecule A' (Figure 3.19).

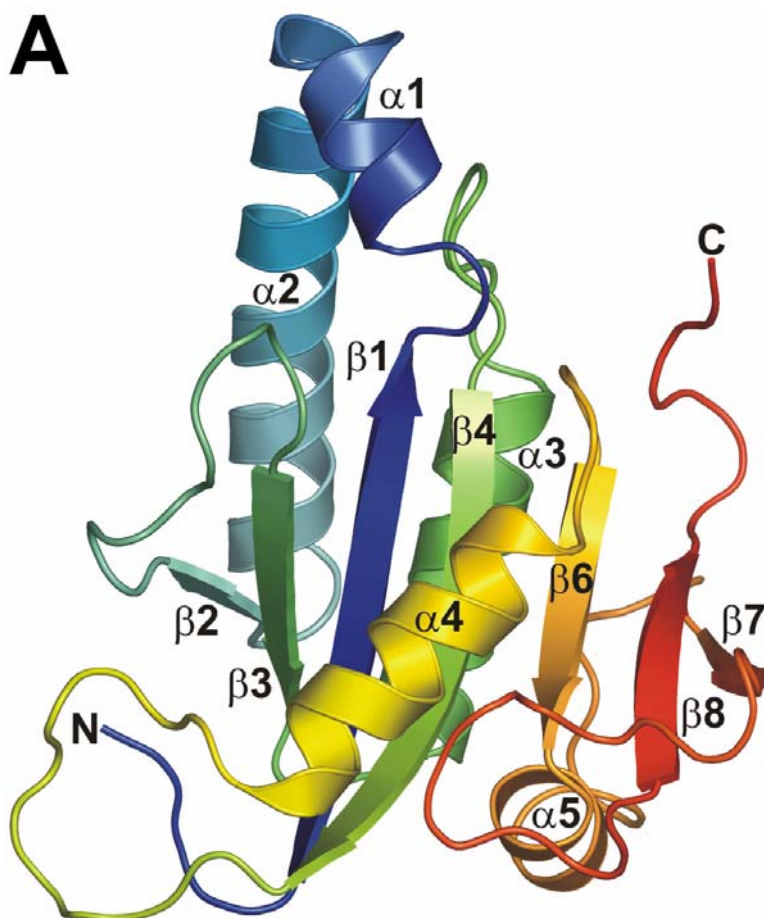**Figure 3.17:** Overview of the X-ray structure of the catalytic core of Rv0386. **A)** Overall structure of the CHD of Rv0386. The catalytic domain of Rv0386 as monomer in rainbow, ranging from the N-terminus in blue to the C-terminus in red. The secondary structure elements are labelled. All figures are created by PyMOL (DeLano, W.L., The PyMOL Molecular graphics System (2002), http://www.pymol.org).
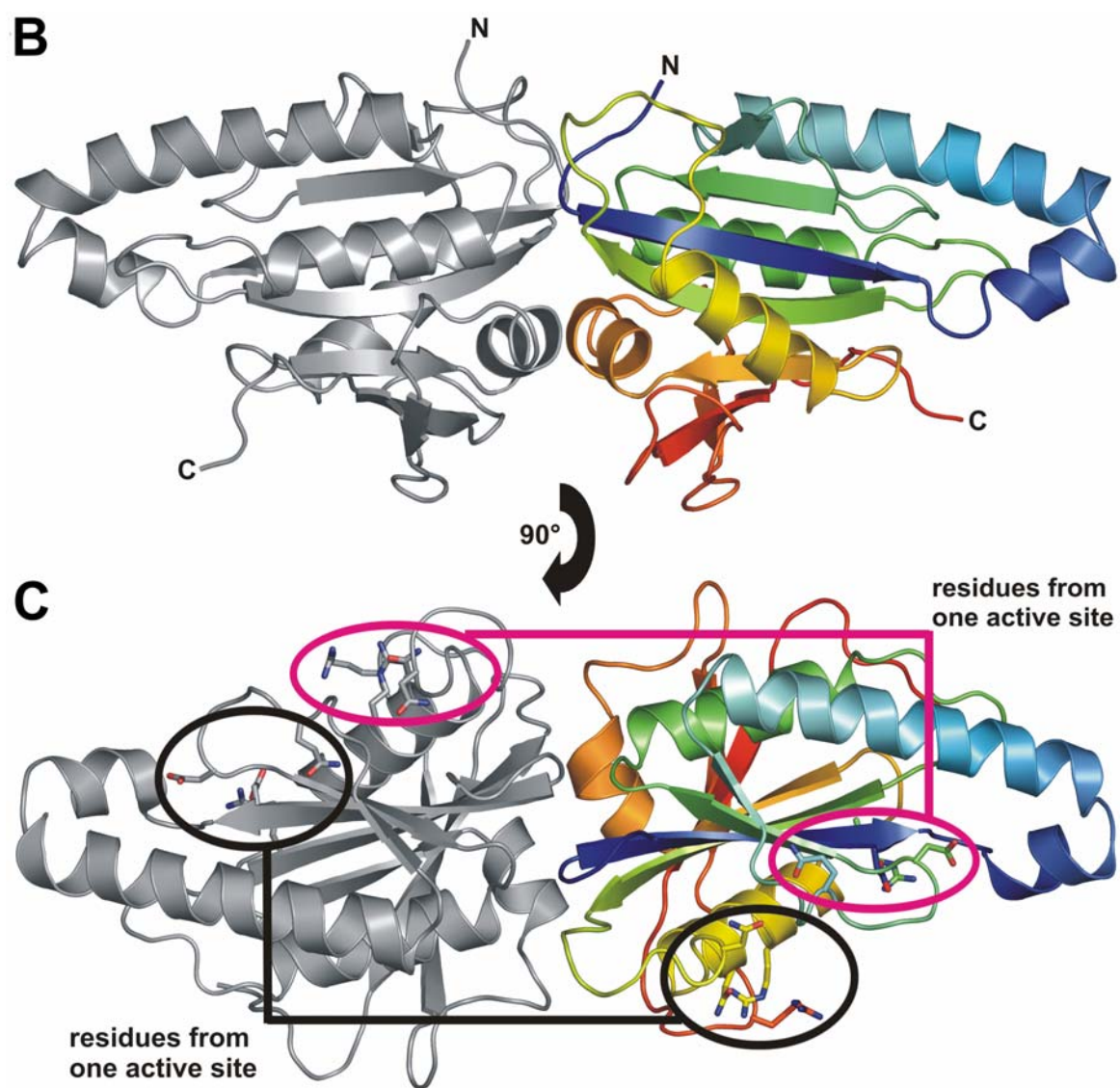
Figure 3.17: B) Inhibited dimer in the head-to-head conformation. The elongated dimer in cartoon is rendered in grey and rainbow. **C)** 90 degrees rotated compared to B). The two catalytic sites with corresponding residues are coloured in magenta and black. The catalytic residues do not form active sites and they are shown in sticks where nitrogen, blue; oxygen, red.

The monomers in the asymmetric unit are very similar. The root mean square deviation (r.m.s.d) between corresponding Cα atoms, between A and B chain is 0.85 (unit) with 141 Cα atoms, between chain C and B = 0.93 with 167 Cα atoms, between chain D and
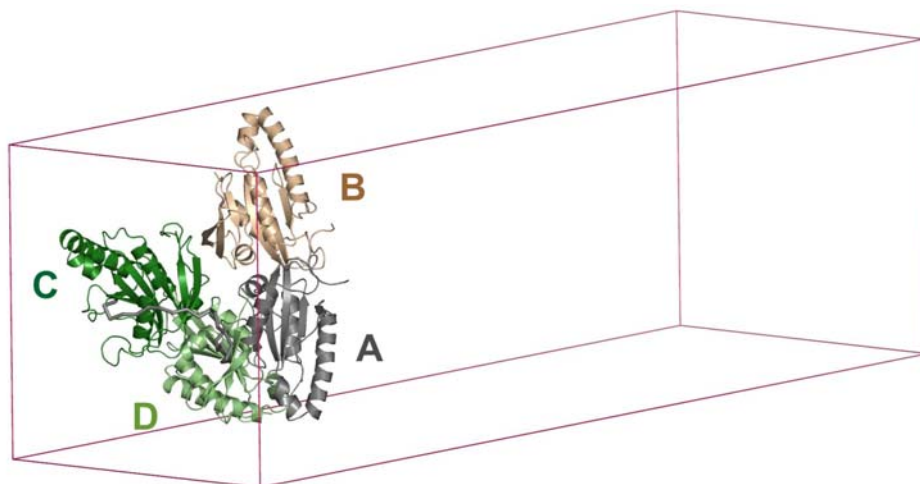
Figure 3.18: Four molecules in the asymmetric unit. Homodimers AB and CD in cartoon are rendered in grey-wheat and dark green-light green respectively. The edges of the asymmetric unit are depicted in magenta.

Figure 3.19: Extended region of monomer A. Molecules AB in grey-wheat and the symmetry related molecule A' in cyan are shown in cartoon. The extended part of the A chain and A' (residue 150-173) are coloured in black and deepteal respectively.

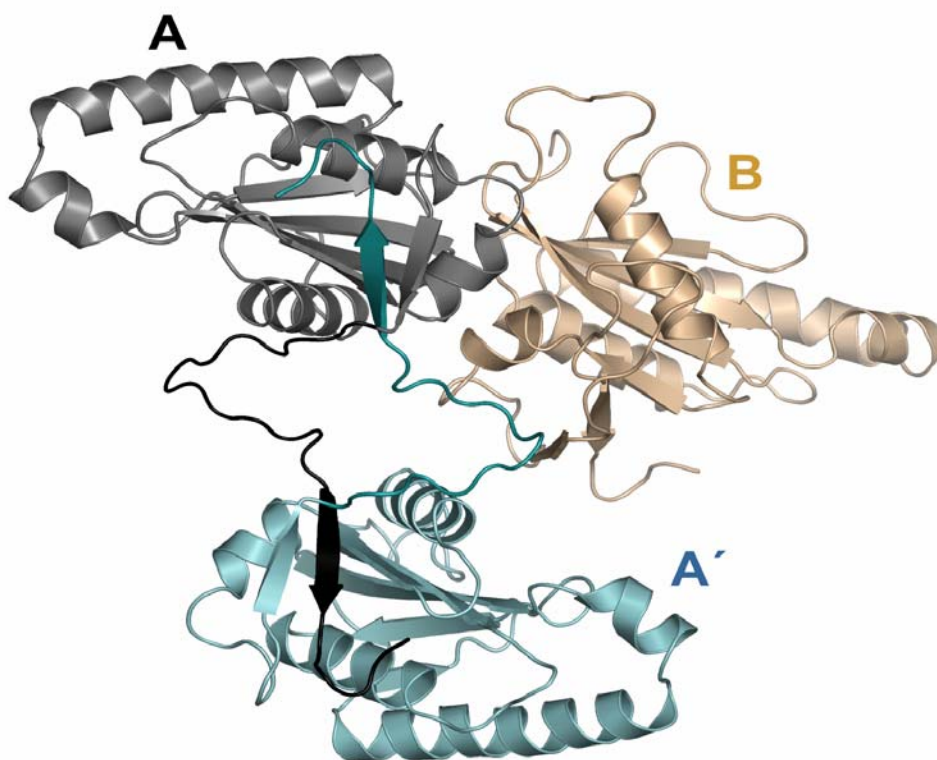B = 0.95 with 166 Cα atoms. The differences between the molecules are found in flexible loops and terminal residues. The flexible loops in all molecules are in the areas of residues 54-59, 99-109 and 149-162. The first eight N-terminal amino acids and the C-terminal amino acids starting from residue 172 do not superimpose. The dimension of the catalytic domain of Rv0386 is 31 x 40 x 42 Å^3. The extended regions at the C-terminus, residue C178-C183 and A150-A173 extend about 30 Å from the main protein bodies.

The structure consists of a three-layer α/β sandwich. The central part consists of a mixed β-sheet, surrounded by 4 α-helices. The β-sheet contains seven β strands in the order 2-3-1-4-6-8-7. The β strands 2-3-1-4 and 6-8-7 are antiparallel, while β4 and β6 are parallel. The α-helices 1-2-3 are situated on one side of the antiparallel 2-3-1-4 β-sheet and on the opposite side α4 is located. In contrast to the other α-helices, α5 is 90 ° rotated towards the central β-2-3-1-4-6 sheet. The secondary structure elements are different in the four chains. The β$_2$ (residues 50 to 52) in chain B and β$_7$β$_8$ (amino acids 146-152; 163-168 respectively) in chain A are not formed. The latter secondary structure elements are altered to a coiled loop, buried into the active site of the A' (described above) and thereby stabilising the formation of the crystal.

3.2.e Ordered Hexahistidine Tag

The complete hexahistidine tag on the C-terminus (amino acids 178-183) can be traced in electron density for chain C and it binds between chain B and A' (Figure 3.20). The His$_6$-tag is bound in a groove on the surface of molecule A and B and enters hydrophilic and hydrophobic contacts. Indeed, the histidine tag is binding in the active site of molecule A'. It is in close proximity to the secondary structure elements α4, β3, β1, β4 from A' and β7, β8 belonging to chain B (to distinguish between the chains the subscript letters A, A', B, C, D are used). The catalytic residues Gln57$_{A'}$ and Asn106$_{A'}$ (Castro et al., 2005) are in vicinity to the histidine-tag, but neither of them make contacts to the hexahistidines. The residue Glu59$_{A'}$ from loop β2-β3 hydrogen bond with His181$_C$. The main-chain carbonyl oxygen of His179$_C$ and the nitrogen of the main-chain amide of Glu59$_{A'}$ do also hydrogen bond. Arg119$_{A'}$ from α4 forms a hydrogen bond to the main chain carbonyl oxygen of His178$_C$. Between His183$_C$ and the main-chain carbonyl oxygen of Lys144$_B$ from loop α5-β7 a hydrogen bond is

established. Hydrophobic interactions are identified between the carbon-β of His183$_C$ and the carbon-γ of Lys144$_B$. The carbon-β of Ala108$_{A'}$ from loop β4-α4 and the carbon-δ$_1$ of Ile112$_{A'}$ from α4 form a hydrophobic interaction with the carbon of the terminal His183$_C$ and the carbon-β of His182$_C$ respectively. In the centre of the histidine-tag aromatic-ring stacking between Trp146$_B$-His179$_C$ is observed.

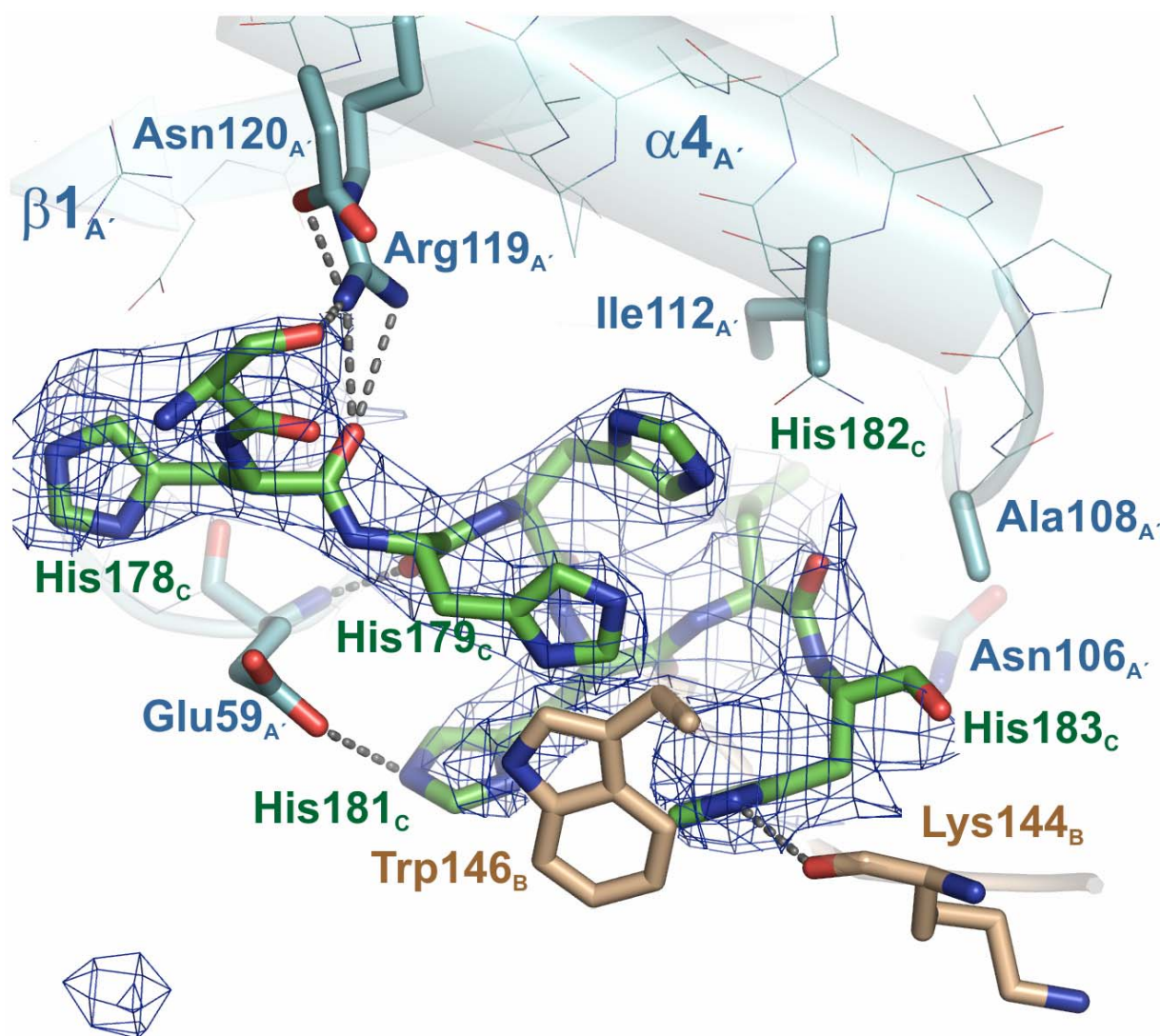**Figure 3.20:** Ordered hexa-histidine tag. The complete histidine tag from chain C situated between chain B and A' in the electron density 2F$_o$ – F$_c$ at the 1 δ level. His$_6$-tag and residues interacting with the hexa histidines are visible in sticks, with nitrogen, blue; oxygen, red. Hydrogen bonds are showed as dashed lines in grey. Chain B, A' and the histidine tag are rendered in wheat, cyan and dark green respectively.

3.2.f Inhibited Dimer

The homodimer of the catalytic domain of Rv0386 buries 1497 $Å^2$ of solvent-accessible surface area per monomer (Brunger et al., 1998). The homodimer forms a head-to head contact, where the catalytic residues do not form active sites. The catalytic residues in Rv0386 are not in the interface of the dimer, but 50 Å apart (Figure 3.17 C). The monomers are arranged to each other in a non-crystallographic two-fold axis. The interface between the homodimers AB and CD are identical. The secondary structure elements forming the interface are $α5_{AB}$, $β4_{AB}$ and $α3_{AB}$, with hydrophobic and hydrophilic interactions. In the centre of the interface the hydrophobic region is formed between $Leu136_B$ and $Val132_A$, $Leu136_A$ and $Val132_B$. In the periphery, two hydrogen bonds are formed between $Ser71_A$ and $Glu98_B$, $Ser71_B$ and $Glu98_A$. All interacting residues are facing each other across the two-fold axis (Figure 3.21). The extensive interactions in the interface of the two monomers, suggest that our structure exists in solution. Several attempts were made of co-crystallising the catalytic domain of Rv0386 with ATP non-hydrolysable analogues to force the active form. We observe this construct of the AC only in this inhibited dimeric state.
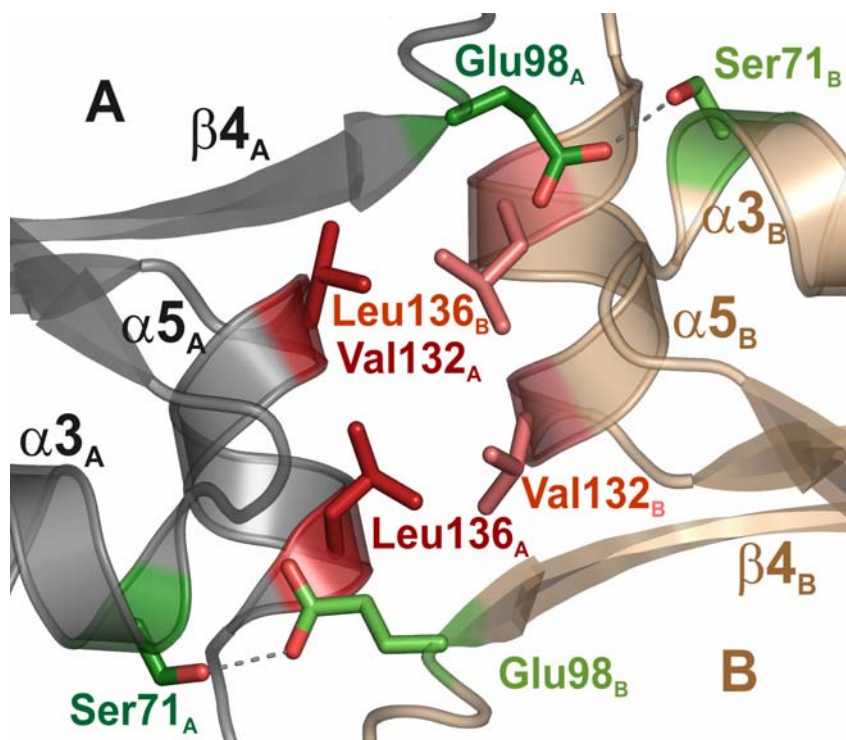
Figure 3.21: Interface of the catalytic domain of Rv0386. The monomers A and B in a head-to-head conformation are in cartoon and rendered in grey and wheat respectively. Polar and hydrophobic interactions are identified in the interface and shown in dark green-light green and red-pink respectively. Hydrogen bonds are formed between $Glu98_A$ and $Ser71_B$, $Glu98_B$ and $Ser71_A$. Hydrophobic interactions are found between $Val132_A$ and $Leu136_B$, $Val132_B$ and $Leu136_A$.

3.3 Analysis of the Shape of CHD Rv0386 in Solution

3.3.a Small Angle X-ray Scattering

In contrast to other crystal structures from homologous proteins, our crystal structure shows an elongated shape with a new dimeric form of the protein, which we postulate to be a novel regulatory state of adenylyl cyclases. Small angle x-ray scattering (SAXS) is a sensitive technique to detect macromolecular shape in solution and was therefore applied to the catalytic domain of Rv0386. SAXS data were collected in the protein concentration range of 1-20 mg/ml. For the evaluation of the SAXS data a number of models were designed, and these were tested for a good fit to the experimental data. The models included the X-ray structure, symmetry related molecules and the modeled active state of the catalytic core of Rv0386 in a number of conformations and oligomeric states. The modeled active state of CHD Rv0386 was generated by superimposing the X-ray structure of CHD Rv0386 on CHD Rv1900c open and closed state (Sinha et al., 2005). The oligomeric contact of the models was based on the interface from head-to-tail, head-to-head CHD dimers and the binding region of CHDs to Gsα. The dimeric model based on the head-to-head contact and the tetrameric model on head-to-head and head-to-tail contacts fit very well to the experimental data (Figure 3.24). Examples of models which did not give a good fit to the small angle X-ray scattering data are presented in Figure 3.22 and 3.23 and based on head-to-tail dimers, symmetry related molecules in oligomeric states. Table 3.5 summarizes the overall parameters calculated from the scattering data: molecular mass M_r, radius of gyration R_g, maximum particle size D_{max} and the volume. At 1 mg/ml aggregation is observed, but at higher concentrations the solutions were monodisperse, which is an essential requirement for the analysis.

The protein is in a concentration-dependent equilibrium oligomeric mixture. At protein concentration 1mg/ml the protein shows the elongated shape of an inhibited head-to-head dimer (Figure 3.24). The model yields a good fit with the discrepancy of $\chi = 1.79$ (Table 3.5). The systematic deviations may be explained by the presence of higher oligomers, i.e. tetramers. Models of open and closed head-to-tail dimers of CHD Rv1900c did not give a good fit ($\chi = 2.26$ and $\chi = 2.42$ respectively) to the experimental data. At concentration 4.4 mg/ml a tetramer is observed (Figure 3.24). The tetrameric model was created by

superimposing the elongated catalytic domain of Rv0386 on the open form of Rv1900c CHD (PDB entry code: 1YBT) (r.m.s.d of 1.60 Å for 144 Cα-atoms).
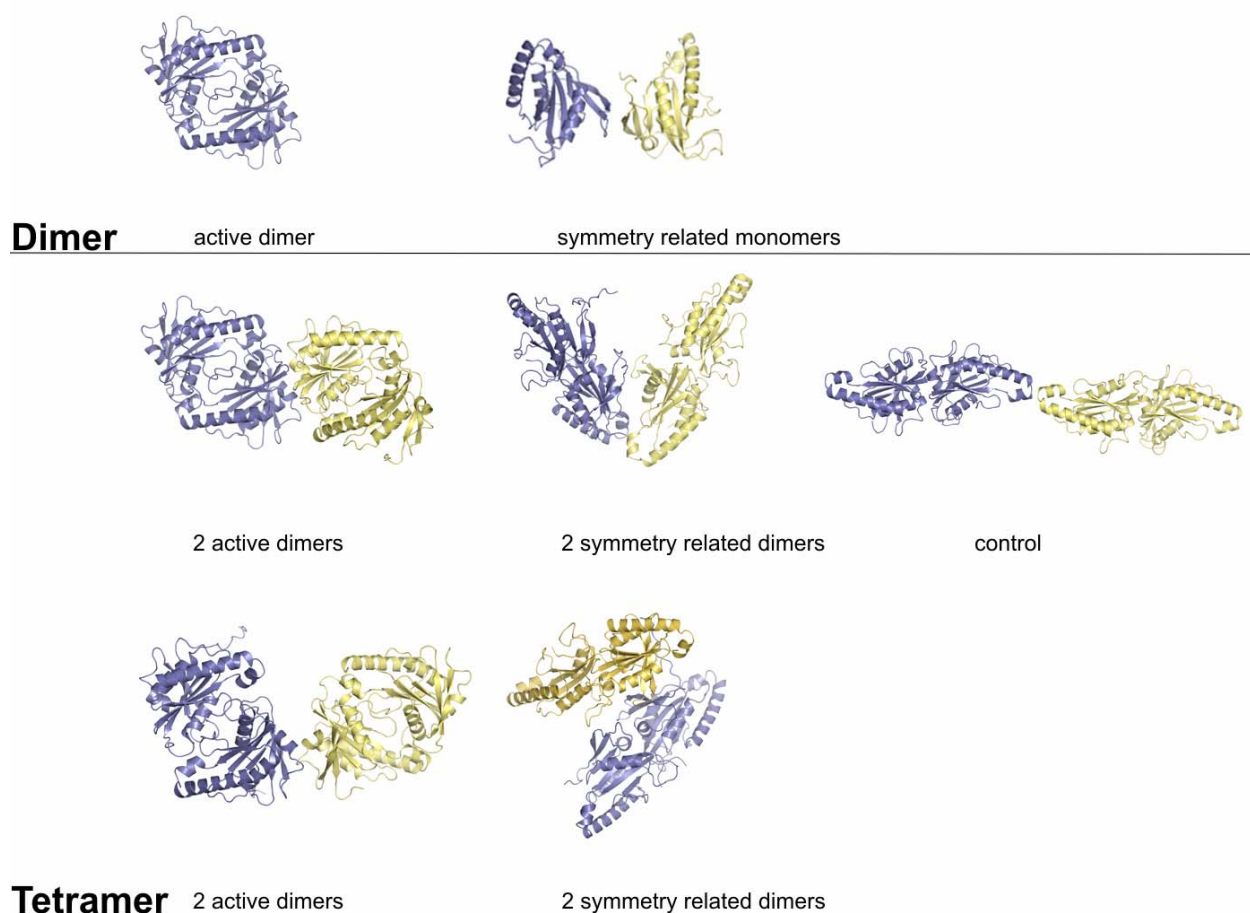
Figure 3.22: Small angle X-ray scattering models of CHD Rv0386. Dimers are depicted in the upper panel and tetramers in the lower panel. The models are designed from inhibited-and modelled active states and symmetry related molecules in the crystal. None of the models above gave a good fit to the experimental data.
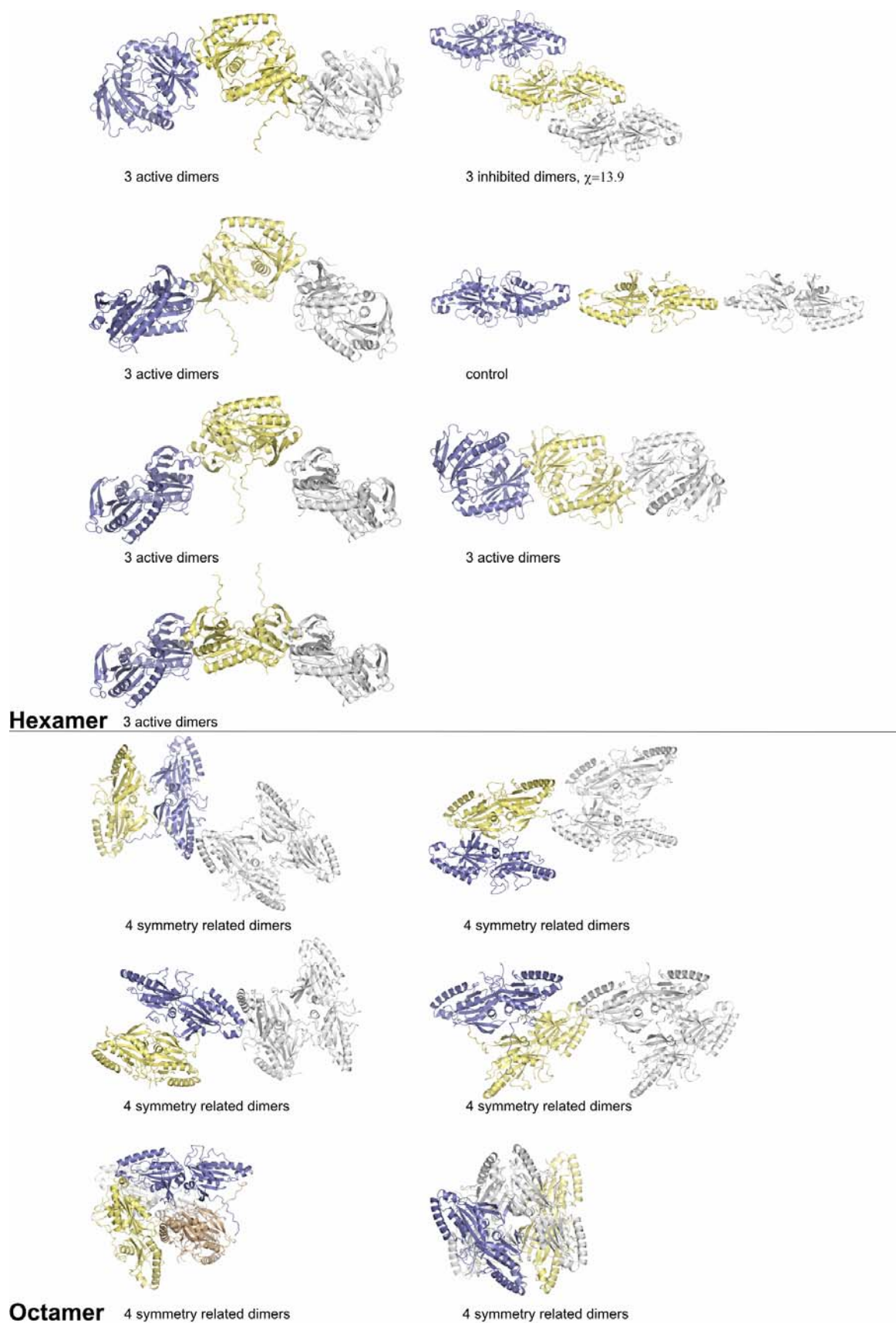
Figure 3.23: Small angle X-ray scattering models of higher oligomers. Inhibited, symmetry related molecules and modelled active state of CHD Rv0386 in various oligomeric conformations. None of the models presented above did give a good fit to the small angle X-ray scattering data. The best model of a hexamer at protein concentration 8 mg/ml gave a χ of 13.9, depicted in the most upper right panel.

SAXS Data

Sample	M_{rseq}[a] (kD)	M_{rexp}[a] (kD)	D_{max}[b] (nm)	Volume[c] (nm^3)	R_g[d] (nm)	χ[e]	Model
Rv0386$_{CHD}$ c= 1mg/ml	18.5	18 ± 2	6 ± 0.5	45 ± 5	18.9 ± 0.5	0.83	Monomer
Rv0386$_{CHD}$ c= 1mg/ml	37	45 ± 4.5	10 ± 1.0	88 ± 8.8	29.66 ± 1.5	1.79	Dimer
Rv0386$_{CHD}$ c= 4.4 mg/ml	74	70 ± 7.0	11 ± 1.1	120 ± 12	32.1 ± 1.6	1.03	Tetramer

Table 3.5: Small angle X-ray scattering data of the catalytic domain of Rv0386.

[a] $M_{r\ seq}$ and $M_{r\ exp}$ are molecular masses predicted from the primary sequence and calculated from SAXS scattering data, respectively.

[b] D_{max} are experimental maximum diameters computed with the program GNOM (Svergun, 1992).

[c] Volume is the Porod volume of hydrated particle.

[d] R_g experimental radius of gyration calculated with GNOM (Svergun, 1992).

[e] χ Chi-square between experimental and theoretical curve.

Two head-to-head dimers forming two catalytic sites in the centre of the open tetramer fit best the experimental SAXS data (discrepancies provided by high resolution models, $\chi = 1.03$) see Table 3.5. At this concentration protein is nearly tetrameric, possibly with 5% of dimers (suggested by OLIGOMER). Head-to-tail models forming tetramers gave an impaired fit. At protein concentrations of 8 mg/ml the suggested species is that of a hexamer, octamer or a mixture of the two (data not shown). The best model of the octamer was constructed by rigid body modelling by GLOBSYMM (Petoukhov and Svergun, 2005) from two tetramers (described above), in an elongated shape ($\chi = 3.8$) forming head-to-tail contacts however, the model shows steric clashes. The best model of a hexamer (Figure 3.23) gives a bad fit to the experimental data ($\chi = 13.9$). The model was created from three head-to-head dimers forming head-to-tail contacts and four catalytic centres.

Manganese treated CHD Rv0386 was observed as monomer by SAXS (Table 3.5). The fit of one catalytic monomer of the X.-ray structure CHD Rv0386 to the experimental SAXS data

gave a very good fit of $\chi = 0.83$. This result is in agreement with observed size exclusion chromatography and analytical ultracentrifugation experiments.

The SAXS data verifies the elongated shape observed for CHD Rv0386 dimer in the crystal structure. At protein concentrations of 1 mg/ml Rv0386 catalytic domain is observed as a head-to-head dimer possibly within a mixture of tetramers. At 4 mg/ml a tetrameric species is observed, composing two active sites from two head-to-head dimers. The tetramer is most likely only formed *in vitro*, in contrast to the head-to-head dimer which might also be a state of the protein *in vivo*. The oligomerisation in SAXS correlates the results described earlier on size exclusion chromatography and activity assays *in vitro*.
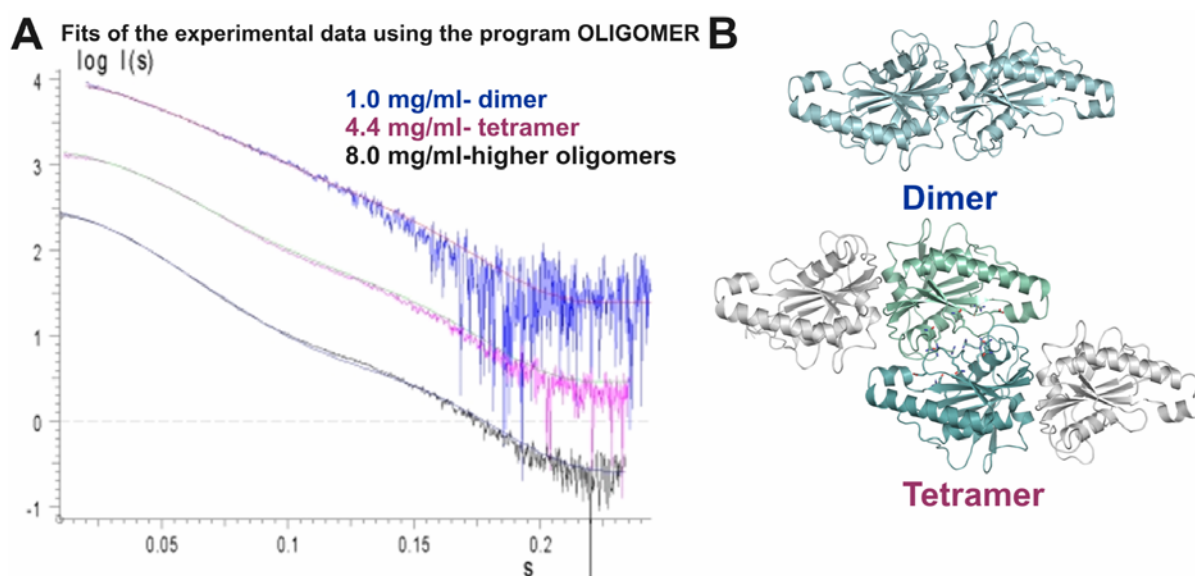
Figure 3.24: X-ray scattering patterns and models. **A)** Experimental scattering patterns from mycobacterial adenylyl cyclase Rv0386 solutions at concentrations 1, 4.4 and 8.0 mg/ml (blue, magenta and black lines) along with the fits computed by a mixture of dimers and tetramers (red), from open tetramers (green) and from tentative octamers (dark blue). The plot displays the logarithm of the scattering intensity as a function of momentum transfer $s = 4\pi \, sin(\theta)/\lambda$ where 2θ is the scattering angle and $\lambda = 1.5$ Å is the X-ray wavelength. The curves are displaced along the abscissa axis for better visualization. The figure has been prepared with Sigmaplot **B)** Upper panel: The model of the dimer in solution, corresponding to the inhibited dimer as observed in the crystal structure. The dimer is in cartoon and cyan. Lower panel: A model of a tetramer is formed upon interaction of two head-to-head dimers. The interaction between the two elongated dimers is such that active sites are formed by the two central monomers (greencyan-dark cyan).

3.3.b Mutagenesis of the Interface of the Head-to-head Rv0386

The crystal structure of the catalytic core of Rv0386 was determined in a unique inhibited dimeric conformation. In order to analyse the interface of the catalytic core of Rv0386 mutation analysis was carried out. The mutant Val132Ala behaves as the wild type concerning adenylyl cyclase/guanylyl cyclase activity. In contrast, the activity of the Glu98Ala mutant is only 15% of the wild type. The former mutation did not have any effect. The bad protein expression of the latter mutation points to a folding problem. The mutation E98A most likely disrupts the interface of the head-to-head dimer and consequently is in a monomeric conformation. Monomers easily aggregate and precipitate without stabilising agents such as glycerol and high salt. Subsequently, the unspecific aggregation might have been the cause of the decreased activity.

In conclusion, interpretation of Glu98 as an interface residue of the catalytic domain of Rv0386 in the inhibited state is correlated with the mutagenesis data presented. The interface is important for stability of the protein. In the modeled active state Glu98 is not located in the interface and would not bind the substrate. Therefore, it is unlikely that E98A mutant would have a direct effect on the activity, if the CHD of Rv0386 would have formed a head-to-tail dimer, i.e the active state.

3.3.c Analytical Ultracentrifugation

Sedimentation velocity experiments were applied to analyse the elongated shape and to characterise the oligomeric state of the catalytic domain of Rv0386 in solution. The velocity data were analysed by the Lamm equation for a dimer and monomer model and the continuous size distribution function C(M). C(M) supplies a distribution of weight averaged molecular weights or sedimentation coefficients that fit best to the data. For the protein dialysed in 20 mM Tris-HCl pH 8.5, 10 mM NaCl, 5 % glycerol and 0.5 mM DTT a dimer was expected. The buffer density and viscosity were calculated with the program Sednterp to 1.01433 kg/m^3 and 1.179x10^{-2} Ns/m^2 respectively. Additionally, the program Sedfit calculated the sedimentation coefficient, S to 2.5, corresponding to the molecular weight of 42900 Da, compatible to the theoretical Mw of 40 400 Da, i.e a dimer (Figure 3.25).

Figure 3.25: Sedimentation velocity of the catalytic domain of Rv0386. **A)** Data generated from a sedimentation velocity experiment at 200 000g and absorbance at 280 nm. The lines are the best fit to the data. **B)** Residuals from the fit in A. **C)** C(M) function corresponding to the best fit to the data in A. The actual S value calculated from Sedfit is 2.5, corresponding to the molecular weight of 42900 Da, similar to the theoretical Mw of 40 400 Da, i.e a dimer.

For the Mn^{2+} treated protein dialysed in 50 mM Tris-HCl pH 8.5, 500 mM NaCl, 5 % glycerol, 0.5 mM DTT a monomer was expected. The buffer density and viscosity were calculated with the program Sednterp to 1.03429 kg/m^3 and 1.2262x10^{-2} Ns/m^2 respectively. The program Sedfit determined the sedimentation coefficient, S to 1.68, corresponding to a molecular weight of 23 kD, consistent with the theoretical Mw of 20.2 kD, corresponding to a monomer (Figure 3.26).
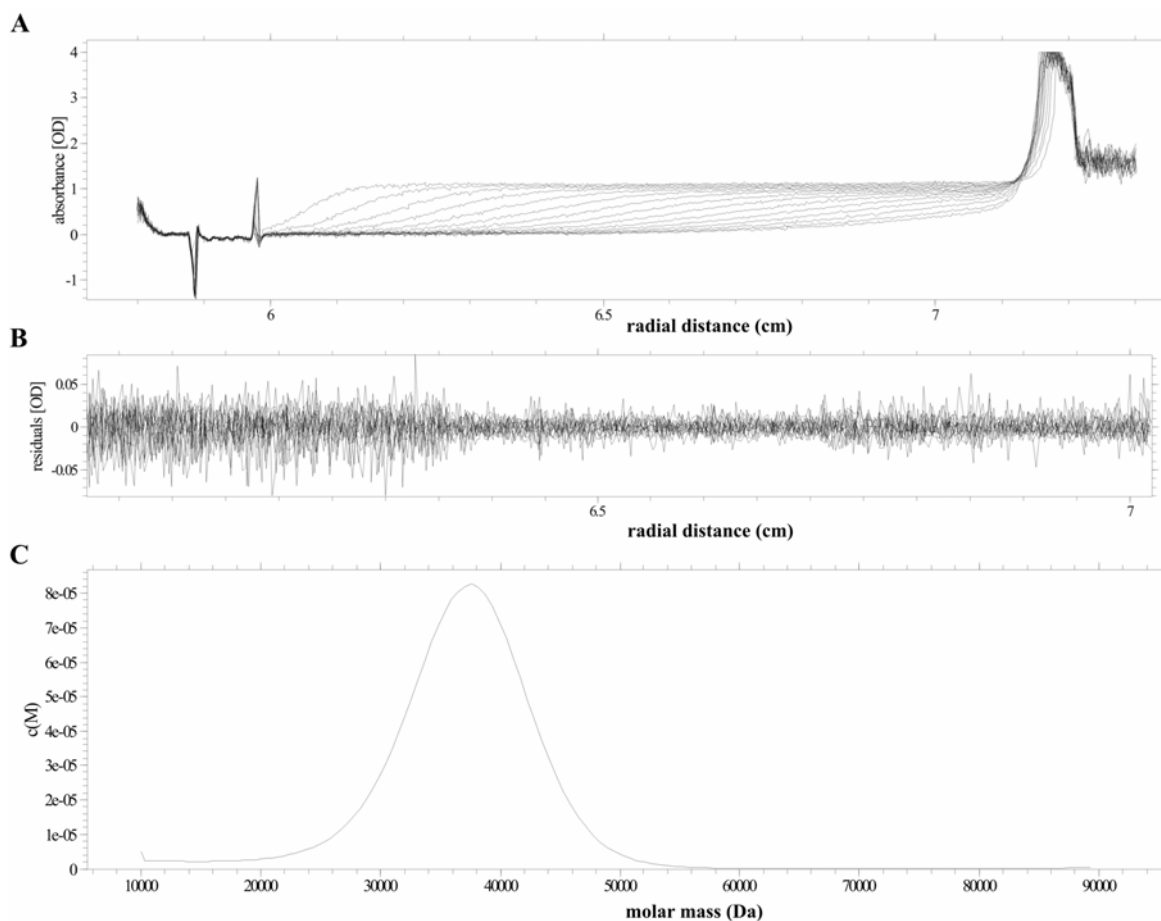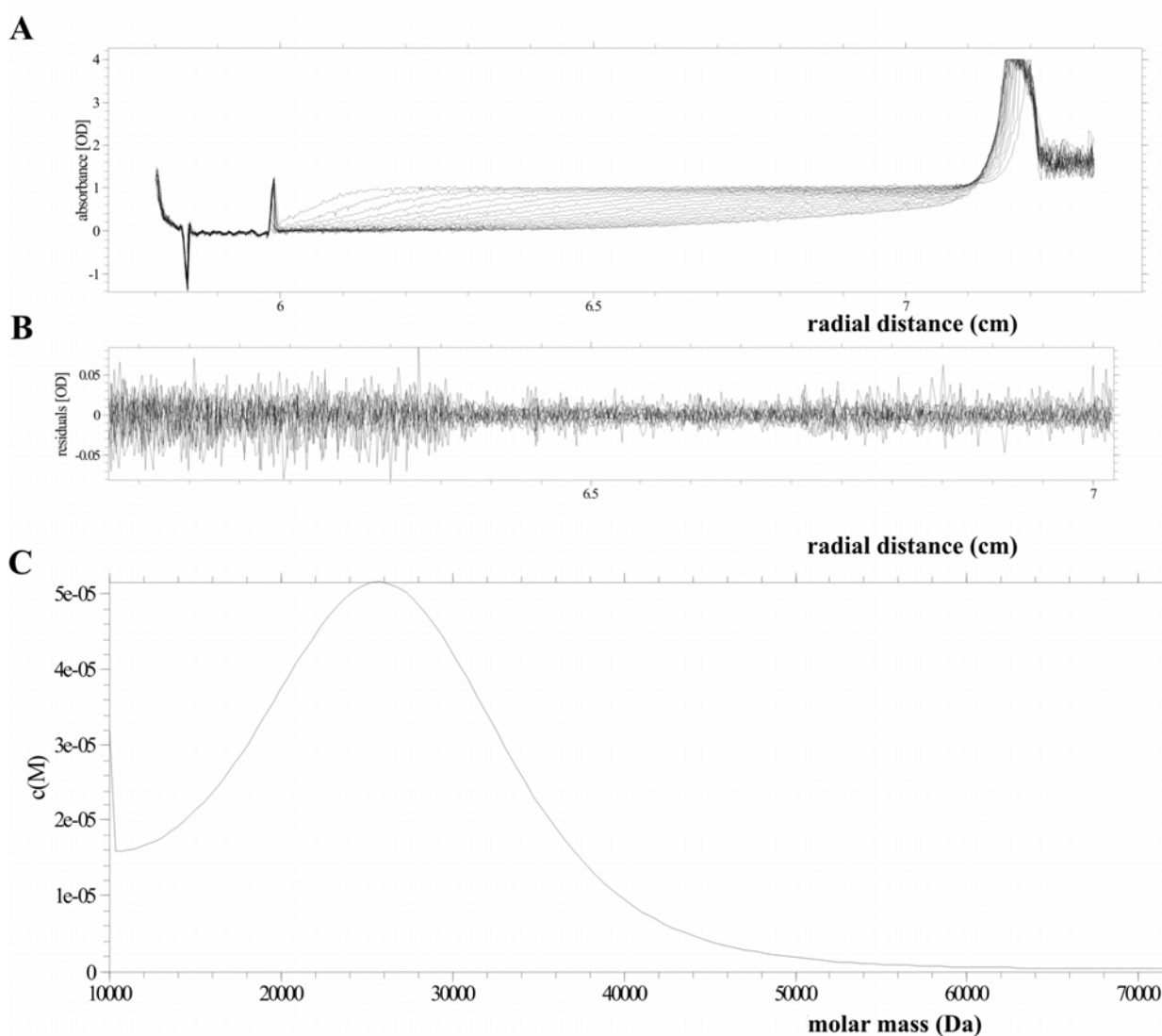
Figure 3.26: Sedimentation velocity of the CHD of Rv0386. **A)** Data produced from a sedimentation velocity experiment, run at 200 000g and absorbance at 280 nm. The lines are the best fit to the data. **B)** Residuals from the fit above. **C)** C(M) function corresponding to the best fit to the data in A. The actual S value calculated from Sedfit is 1.68, matching the molecular weight of 23 kDa, consistent to the theoretical Mw of 20.2 kDa, i.e monomer.

The inhibited head-to-head crystal structure, the modeled active catalytic core of Rv0386 and the monomer were employed to calculate the theoretical S values with the program Hydropro. A monomer with and without the histidine tag contribute to the theoretical sedimentation coefficient, S of 1.50 and 1.56, consistent with the experimental S value of 1.68. The inhibited CHD dimer of the X-ray structure of Rv0386 result in the theoretical S value of 2.23 and the theoretical S value 2.47 for the modeled active state was compared to the experimental S of 2.5. Since the differences in S values for the inhibited X-ray structure and the modeled active

structure of CHD Rv0386 were small, the inhibited conformation could not be verified in solution by analytical ultracentrifugation. However, the dimeric and monomeric state of the catalytic core of Rv0386 analysed by size exclusion chromatography was confirmed by sedimentation velocity experiments.

3.4 Crystallisation of the Monomer or Active Dimer of the Catalytic Core of Rv0386

The catalytic domain of Rv0386 purified as a monomer and incubated with the non-hydrolysable ATP analogue $\alpha_1\beta CH_2ATP$, gave rise to crystals in new conditions. The crystals had different morphologies as seen in (Figure 3.27). Small crystals were grown and further optimisation would have been required to improve the quality and size for higher resolution data. Due to time restrictions, the crystals were not improved.
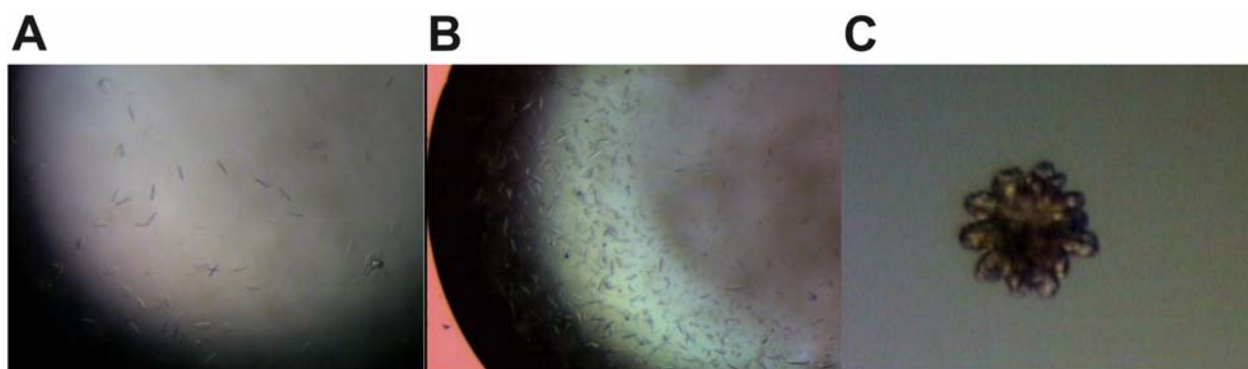
Figure 3.27: Crystals of the monomer or active state of the catalytic core of Rv0386. **A)** Crystals grown in 15 % PEG4000, 0.1 M BTP pH 8.0, 0.2 M NaCl and in 16.5 % PEG4000, 0.1 M CHES pH 9.8, 0.2 M NaCl at 12°C. **B)** Crystals were grown in 5 % PEG3350, 0.1 M Hepes pH 7.5, 0.2 M KAc at 12°C. **C)** Crystals grown in 28% PEG400, 0.1M Hepes· Na pH 7.5, 0.2M CaCl$_2$ at 12°C.

Chapter 4

DISCUSSION

Each year 3000 000 people are killed by the pathogen *Mycobacterium tuberculosis.* One third of the world's population is infected, but only 10% develop the disease tuberculosis. Signalling pathways used by *M. tuberculosis* are investigated in order to develop new methods for elimination of the disease. In this thesis, the catalytic domain of the adenylyl cyclase Rv0386 has been studied, which is one of in total 15 class III ACs from *M. tuberculosis*. The AC Rv0386 is unusual in the sense that it uses both ATP and GTP as substrate, converting it into cAMP and cGMP respectively. We have determined the X-ray structure of the catalytic core of AC Rv0386 from *Mycobacterium tuberculosis* in a new elongated dimeric state. Small angle X-ray scattering confirmed the shape of the protein in solution and the oligomerisation effect by the cofactor Mn^{2+} gives an explanation of the low activity of the CHD Rv0386. The low activity AC Rv0386 could possibly be stimulated by regulatory domains or by an unknown effector. The suggested model of the active dimer is likely to reflect the native state of the activated catalytic core of AC Rv0386. An activation mechanism of the cyclase homology domain of Rv0386 is proposed.

4.1 Inhibited CHD of Rv0386 from *M. tuberculosis*

4.1.a X-ray Structure of the Inhibited Catalytic Core of Rv0386

Non diffracting crystals grown in several conditions were successfully improved and the X-ray structure of the inhibited CHD Rv0386 was determined with SAD, to a resolution of 2.95 Å. The CHD is in a new elongated dimeric fashion, forming a head-to-head dimer. The active CHDs from *M. tuberculosis* have the catalytic residues in the interface of the homodimer, contributing to the formation of two active sites. In contrast, the CHD of Rv0386 comprises catalytic residues that are not in the interface of the dimer, but 50 Å away (Figure 3.17C). With disrupted catalytic sites, the substrate cannot bind and catalysis cannot take place. Therefore, the crystal structure of the catalytic domain of Rv0386 is in an inhibited state.

The homodimer of the catalytic domain of Rv0386 buries 1497 $Å^2$ of solvent-accessible surface area per monomer (Brunger et al., 1998) and contains hydrophobic and hydrophilic interactions (3.17B,C). This extensive interface makes it reasonable to believe that the structure is existing in solution. To confirm this, small angle X-ray scattering was applied. High resolution models of the head-to-head and head-to-tail open and closed catalytic domains of Rv0386 were used for SAXS. The crystal structure of the CHD Rv0386 gave the best fit to the experimental data (Figure 3.24, Table 3.5), which confirms that the head-to-head CHD Rv0386 is present in solution and in an inhibited state.

Co-crystallisation of the CHD Rv0386 with ATP non-hydrolysable analogues to force the active form was conducted. We observe this construct of the AC only in this inhibited dimeric state. High affinity inhibitors are not available for the mycobacterial class III soluble adenylyl cyclases: P-site inhibitors that bind to the mammalian G-protein dependent adenylyl cyclases with nanomolar affinity have only low affinity. This can possibly be explained by differences in the ability to bind substrate. The catalytic residues in class III ACs are indicated in the sequence alignment in Figure 4.1.
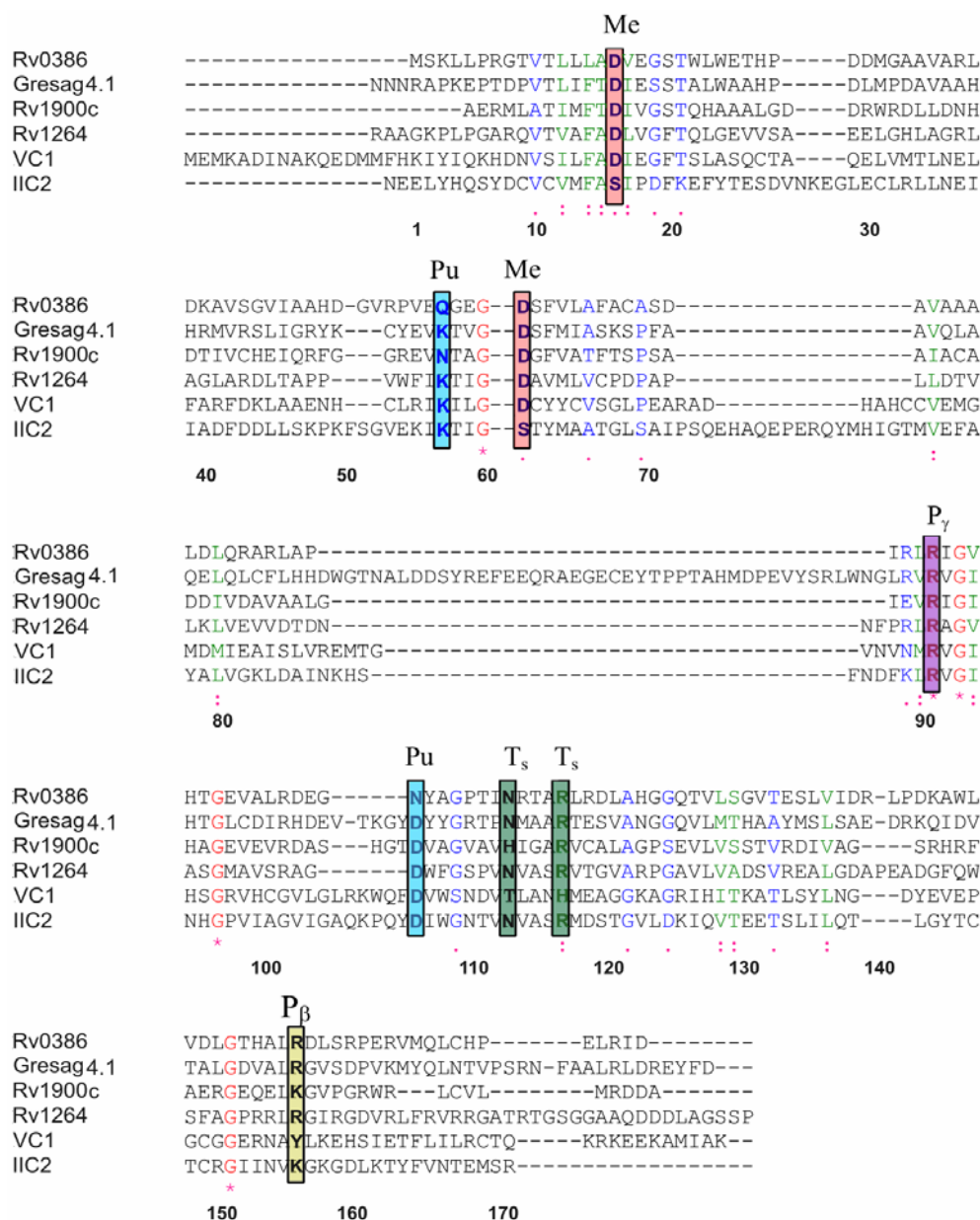
Figure 4.1: Sequence alignment of homologous CHDs. The sequence of *M. tuberculosis* Rv0386 (homodimer) on top, followed by *Trypanosoma brucei* Gresag 4.1(monomer), *M. tuberculosis* Rv1900c (homodimer), *M. tuberculosis* Rv1264, canine VC1 and rat IIC2 (heterodimer). Amino acid numbering (below of each block) is according to the AC Rv0386 sequence. Eight residues involved in catalysis are in coloured boxes. Pink: metal-binding (Me) (Rv0386: D16 and D61), cyan: purine-binding (Pu) (Rv0386: Q57 and N106), violet: γ-phosphate binding (Pγ) (Rv0386: R91), yellow: β-phosphate binding (Pβ) (Rv0386: R156), green: transition state stabilising (Ts) (Rv0386: N113 and R117).

4.1.b Low Activity and the Effect of Mn^{2+} in Oligomerisation

The catalytic domain of Rv0386 is a low activity protein, displaying a Vmax of 5.6 nmol cAMP·mg-1·min-1 at 850 µM ATP, which is 180 fold lower compared to high activity CHDs Rv1900c and Rv1264 from *Mycobacterium tuberculosis* (Castro, 2004; Castro et al., 2005). The cofactor Mn^{2+} gives rise to higher activity *in vitro* compared to the divalent cofactor Mg^{2+} (Castro, 2004). In this work, it was observed by size exclusion chromatography that addition of Mn^{2+} changes the state of CHD Rv0386, from dimeric to monomeric. This is in agreement with analytical ultracentrifugation and small angle X-ray scattering data. The monomeric species precipitates easily in 5% glycerol and low salt (10 mM NaCl). In contrast, the Rv0386 CHD head-to-head dimers stay in solution in 10 mM NaCl and 5% glycerol over three days.

From the SAXS experiment and the crystal structure head-to-head dimers can be observed (Figure 3.17C, Figure 3.24). Binding of Mn^{2+} to the catalytic residues in the active centre might induce a conformational change in the CHD of Rv0386 destabilising the head-to-head homodimer interface and generating the monomers. This is observed as aggregation and precipitation in size exclusion chromatography, activity studies, analytical ultracentrifugation and small angle X-ray scattering. The monomers are unstable in solution, due to the exposure of hydrophobic surfaces and absence of interaction partners. For stabilisation, a part of the monomers form inhibited head-to-head dimers, higher head-to-head oligomers and active head-to-tail dimers. The Kd is determined to 140 nM (Castro, 2004), indicative of a high affinity of the catalytic domains for each other. Since the head-to-tail dimers are in minority, the catalysis takes place in less extent. This is in agreement with the low activity of CHD Rv0386 (Castro, 2004; Castro et al., 2005). In contrast to Mn^{2+}, cofactor Mg^{2+} does not contribute to any change in oligomeric state in AC Rv0386, observed by size exclusion chromatography. Indicating that Mn^{2+} is essential for inducing the inhibited to active state of CHD of Rv0386. An external activator and addition of regulatory domains would be a prerequisite for confirmation of the suggested regulation of AC Rv0386.

The low activity could possibly be stimulated by the other regulatory domains and by an unknown activator (Castro, 2004; Castro et al., 2005). An example of this is the CHD of Rv3645 which can be stimulated 21-fold by the HAMP domain (Linder et al., 2004). To test this for Rv0386 the CHD has been cloned with its neighbouring AAA-ATPase domain (Castro, 2004) but the protein was observed in inclusion bodies in recombinant expression. The individual domains of Rv0386, i.e AAA-ATPase, tetratricopeptide repeat (TPR) domain,

luxR helix-turn-helix DNA binding domain as well as the full length protein were also observed in inclusion bodies (Castro, 2004). Further cloning and expression is required. Expression with more constructs including fusion proteins, chaperons, weaker promoters, stepwise increase/decrease of 1-3 residues, lower expression temperatures, low IPTG concentrations, glucose-lactose expression, media with more or less nutrients, are a commonly used strategy to increase solubility of the single domains as well as for the full length Rv0386. Furthermore, identification and co-expression of the activator with the additional CHD would be of advantage. Another possibility is to design a linker between the CHDs, thus creating a head-to-tail CHD of Rv0386.

4.1.c Putative Activator

The activator and its binding region of Rv0386 is not known. A putative modulator could bind to α2, α3 of one catalytic monomer including the N-terminal region of the other catalytic monomer, in the active state of Rv0386. This is in agreement with the mammalian AC in complex with $G_s\alpha$ (Tesmer et al., 1997).

It is suggested that prokaryotic class III ACs comprising the purine binding residues Thr and Lys are activated by bicarbonate *in vitro* (Cann et al., 2003; Chen et al., 2000). These ACs include conserved substrate binding residues, except for the purine binding Thr (**Table 1.3**). Other ACs with non conserved substrate binding residues might also be regulated by an external activator. This could also be true for non conserved Asn106 and Gln57 in AC Rv0386. It is possible that the activator stabilises the residues responsible for substrate selectivity, thereby allowing efficient nucleotide binding and increasing the activity. The bicarbonate might mimic an *in vivo* external regulator or regulatory domains. Another example is forskolin involved in gluing the mammalian cyclic homology domains together *in vitro*, most likely mimicking an unknown activator *in vivo*. Most of the mycobacterial CHDs have low intrinsic activity (Castro et al., 2005) and might need an activator for enhanced activity.

4.1.d Crystals of Rv0386 CHD in the Active State?

The catalytic domain of Rv0386 was also produced in a monomeric state, observed by size exclusion chromatography, analytical ultracentrifugation, activity studies and small angle X-ray scattering. The monomer is unstable and precipitates easily, unless high amount of salt and glycerol is added. Prior to crystallisation non hydrolysable substrate analogues were added to the protein, to induce the active state of the catalytic domains as homodimer (head-to-tail). This could happen, since the substrate analogue stabilises the interface of the head-to-tail dimer. The analogue would stabilise the active dimer conformation. Crystals were grown in new conditions and crystal forms, possibly related to a conformation of the enzyme as monomer or a dimer with bound substrate analogue. The head-to-tail Rv1900c (Sinha et al., 2005) and Rv1264 (Tews et al., 2005) have been structurally determined with bound substrate analogue and apoform respectively. The latter comprises a SO_4^{2-} in the interface from the crystallisation buffer, stabilising the interface of the active dimer. Due to time restrictions, the crystals of Rv0386 were not improved.

4.2 Comparison with Homologous Structures

4.2.a Similarity of the Catalytic Domain of Rv0386 with Related Proteins

The catalytic domain of Rv0386 is structurally highly related to all class III CHDs of ACs. The sequence identities of the catalytic domain of Rv0386 compared to the mammalian domains is only 14 % with canine type VC1 and 14 % with rat type IIC2. Superimposing the CHD of Rv0386 on the type IIC2 domain from rat gives a r.m.s.d of 1.85 Å for 115 Cα-atoms and on type VC1 domain from canine gives a r.m.s.d of 1.57 Å for 133 Cα-positions (PDB entry code: 1CJK). The superimposition of CHD of Rv0386 on Rv1900c in the open conformation (r.m.s.d of 1.60 Å for 144 Cα-atoms) (PDB entry code: 1YBT) and on Rv1264 active and inhibited conformation give a r.m.s.d of 1.66 Å for 124 Cα-atoms and 1.60 Å for 120 Cα- atoms respectively.

The topology of the catalytic domain of Rv0386 shows a divergence compared to the topology of mammalian, trypanosomal and mycobacterial catalytic domains (Table 4.1). The

catalytic domain of Rv0386 is similar to other mycobacterial catalytic domains, lacking the Δ-subdomain of the GRESAG proteins from *Trypanosoma brucei* (Bieger and Essen, 2001). The dimerisation arm, (Linder et al., 2002; Linder and Schultz, 2003; Tesmer et al., 1997) situated between β_4-β_5 in all CHDs, contains a 6-residues gap in Rv0386 compared to the mammalian, trypanosomal and Rv1900c ACs. The dimerisation arm of Rv0386 has the similar length as the CHD of Rv1264. Unusually, the β_5 is only missing in Rv0386 (Figure 4.2**)** consistent with DSSP assignment, but present in all other catalytic domains. The shorter β_4 and longer α4 contributes to that the highly conserved β-hairpin motif (β_4-β_5) is missing in Rv0386 catalytic domain.

The 6-amino acids gap and the missing ß5 make the dimer interface more unstable and could drive the protein into a head-to-head conformation. In the X-ray structure of Rv0386 the dimerisation arm is in the vicinity of the dimer interface, but it is not part of the dimer interface.

Interestingly, residues in vicinity to the missing dimerisation arm and the regions involved in catalysis have high B-factors, indicative of disorder or high flexibility in Rv0386. The His$_6$-tag and the elongated region from the A chain do not have high B-factors, since they are stabilised by the surrounding B chain and A' and molecule A' respectively.

M. tuberculosis Rv0386	$\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3\beta_4$ $\alpha_4\beta_6\alpha_5\beta_7\beta_8$
Rat type II C_2 (homodimer)	$\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3\beta_4\beta_5\alpha_4\beta_6\alpha_5\beta_7\beta_8$
Canine type VC1	$\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3\beta_4\beta_5\alpha_4\beta_6\alpha_5\beta_7\alpha_6\alpha_7\beta_8$
M. tuberculosis Rv1900	$\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3\beta_4\beta_5\alpha_4\beta_6\alpha_5\beta_7\beta_8$
M. tuberculosis Rv1264 (active)	$\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3\beta_4\beta_5\alpha_4\beta_6\alpha_5\beta_7\beta_8$
M. tuberculosis Rv1264 (inhibited)	β_1 $\alpha_2\beta_2\beta_3\alpha_3\beta_4\beta_5\alpha_4\beta_6\alpha_5\beta_7\beta_8$
T. Brucei GRESAG	$\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3\alpha_{3a}\alpha_{3b}\beta_4\beta_5\beta_6\alpha_4\beta_7\alpha_5\alpha_6\beta_8\beta_9$

Table 4.1: Topology of the catalytic domain of class III adenylyl cyclases, according to DSSP.
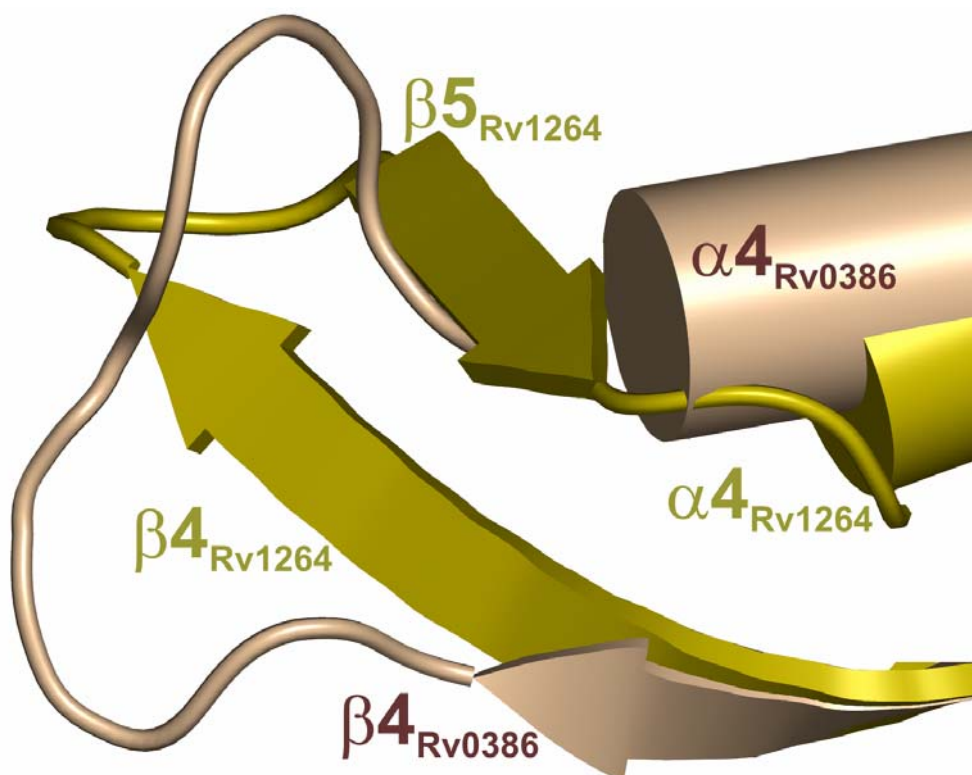
Figure 4.2: Missing dimerisation arm of Rv0386 compared to Rv1264. The CHD of Rv0386 and Rv1264 are in cartoon and rendered in wheat and olive respectively.

4.3 Expected Active Conformation

4.3.a Model of the Active State of Rv0386 in Comparison with the Mammalian Heterodimer

When the CHD of Rv0386 was superimposed on the AMPCPP-bound CHD dimer of Rv1900c (PDB entry code1YBU) (rmsd of 1.75 Å for 142 Cα-atoms) clashes were observed between Asn113A and Thr21B, Thr21A and Asn113B, Gly58B and Gly58A. No clashes were observed when superimposing the catalytic domain of Rv0386 on the mammalian AC type IIC2 domain from rat and on type VC1 domain from canine (r.m.s.d of 1.85 Å for 115 Cα-atoms and r.m.s.d of 1.57 Å for 133 Cα-positions respectively, PDB: 1CJK) and therefore, the overlay was done with the mammalian AC. From this superposition a model of the active site of Rv0386 in complex with P-site inhibitor ATPαS-Rp (Figure 4.3B) and GTP (Figure 4.3C) was created. Nucleotides ATP and GTP are accepted as substrate of the enzyme (Castro

et al., 2005), and the model shows how the two residues Gln57$_B$ (loop β$_2$-β$_3$) and Asn106$_B$ (loop β$_4$- α$_4$) take up their function in substrate recognition, by an amide switch. Nitrogen and oxygen of the Gln57$_B$ amide hydrogen bond to adenine N1 and N6 respectively.

Amide nitrogen of Asn106B hydrogen bonds to the amide oxygen of Gln57$_B$. Asn106$_B$ and Gln57$_B$ correspond respectively to Lys938 and Asp1018 from mammalian IIC2 (Figure 4.3). Single or double mutation of either Gln57$_B$ or Asn106$_B$ to the usual substrate-specifying lysine-aspartate pair; Q57K/N106D abolished AC and GC activity. Mutation Gln57Glu neither showed any detectable activity. Interestingly, Asn106Ser had 36% of wild type AC activity and no GC activity. This confirms that the AC and GC activity are dependent on the Gln57 and Asn106 pair. Asn106 can act as a hydrogen-bond acceptor via its carbonyl oxygen from the amide and as a hydrogen-bond donor via the nitrogen in the amide. Aspartate and glutamate can only act as hydrogen-bond acceptors in contrast to serine that can serve as a hydrogen-bond donor. Therefore, this supports that Rv0386 uses a novel mode for substrate selection (Castro et al., 2005). In Rv1900c CHD the non conserved substrate-specifying residues Asn342-Asp395 (Sinha et al., 2005) differ to the usual lysine-aspartate couple. The Asn342-Asp395 pair does not hydrogen bond to the AMPCPP adenine moiety and is therefore nonessential for catalysis.

Asp61$_A$ (loop β2-β3) and Asp16$_A$ (β1) could coordinate the Mn^{2+} as shown for Asp440 and Asp396 respectively of VC1 (Tesmer and Sprang, 1998; Tesmer et al., 1999). Manganese is an essential co-factor required for the activity of the catalytic domain of Rv0386 and other ACs from *Mycobacterium tuberculosis* (Castro, 2004; Guo et al., 2001; Linder et al., 2004; Linder and Schultz, 2003; Shenoy et al., 2004). Manganese may very well be physiologically relevant, since a putative Mn^{2+} transporter is found in the genome of *M. tuberculosis*, possibly sustaining high concentrations of intracellular Mn^{2+} (Cole et al., 1998; Linder et al., 2004; Reddy et al., 2001). Many ACs from bacteria prefer the cofactor Mn^{2+} over Mg^{2+}. A few bacterial CHDs and the mammalian ACs use Mg^{2+} instead (Dessauer and Gilman, 1997). Asn113$_B$ (α4) is equivalent to Asn1025$_B$ of type IIC2, and likely binds to the ATP ribose oxygen orienting the substrate during catalysis (Tesmer et al., 1997; Yan et al., 1997). Arg117$_B$ (α4) potentially binds to the α-phosphate of ATPαS-Rp, as the corresponding Arg1029 of IIC2. Arg91$_A$ (β4) is equivalent to Arg484 of VC1, binding to the γ-phosphate of the ATPαS-Rp. Arg156$_B$ (loop β7- β8) corresponds to Lys1065 of IIC2 and Rv0386 would in the active state potentially move towards the γ-phosphate of ATPαS-Rp and bind there. With

the exception of the purine binding residues Asn106 and Gln57, the conserved substrate binding residues in Rv0386 depicted in the sequence alignment (Figure 4.1) and the modeled active Rv0386 (Figure 4.3) suggest the catalytic mechanism of Rv0386 may be similar to that in most of the other ACs.
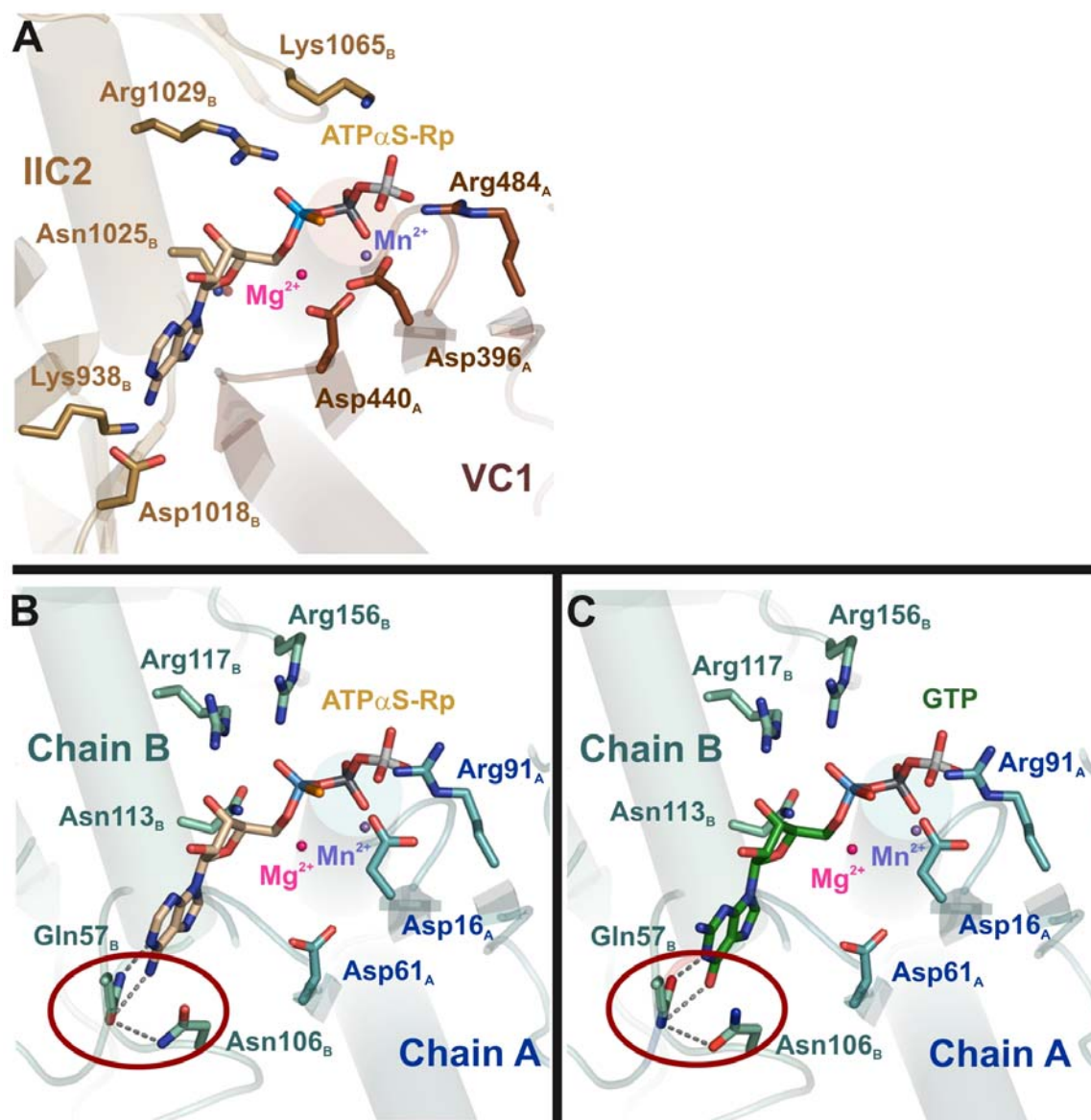

Figure 4.3: Model of the catalytic site of Rv0386 in comparison with the active site of the mammalian AC (PDB code: 1CJK). The catalytic domains of the AC Rv0386 were superimposed onto the active state of the mammalian heterodimeric AC with ATPαS-Rp. Rv0386 forms a regular dimer without clashes with two active sites. Residues playing a role in substrate binding and catalysis, ATPαS-Rp and GTP are shown in sticks with nitrogen, blue; oxygen, red; sulphur, orange; α-phosphate, cyan; β-phosphate, black; γ-phosphate, grey. Mg^{2+} and Mn^{2+} are highlighted in pink and violet respectively. Substrate specifying residues Gln57 and Asn106 binding ATP analogue (B left panel) and GTP (C right panel) by a possible amide switch are encircled in ruby. **A)** A view of the structure of the G-protein activated hetero-dimeric canine AC type V C1 and rat AC type II C2 active site in complex with P-site inhibitor ATPαS-Rp. VC1 (chain A) and IIC2 (chain B) is coloured in brown and beige respectively. ATPαS-Rp is depicted in wheat. **B)** Model of Rv0386 CHD in the active site modeled after the mammalian AC. View of the catalytic site of Rv0386 in complex with ATPαS-Rp. The rendered monomers chain B and chain A are coloured in cyangreen and darkcyan respectively. **C)** Active site of Rv0386 in complex with GTP as a model. GTP is rendered in dark green and the monomers are described in B.

The modelled active state of CHD Rv0386 buries 1361 $\mathring{A}^2$ of solvent-accessible surface area per monomer. In comparison, the mammalian AC type VC1 domain from canine and type IIC2 domain from rat is contributing to the interface of 3800 $\mathring{A}^2$. The dimerisation arm (Linder et al., 2002; Linder and Schultz, 2003; Tesmer et al., 1997) of the modeled active state of Rv0386 excludes 6 amino acids (Figure 4.4B) relative to the mammalian heterodimer (Figure 4.4A). The missing dimerisation arm and shorter loops are important reasons of the difference in buried interface.

Modelled active Rv0386 comprises diminished interactions in the interface compared to homologous ACs. The missing β_5, 6-residues gap in the dimerisation arm, shorter loops and declined interactions in interface of the active dimer could make it more demanding forming the head-to-tail dimer, driving the formation to the head-to-head dimer. Upon, stimulation by the other regulatory domains and by an unknown activator *in vivo*, Rv0386 could arrange in the active state (Castro et al., 2005). A potent non-hydrolysable analogue is important as well, hence stabilising the interaction between the monomers in the active conformation. This has been shown *in vitro* by, class III AC crystal structures containing a non-hydrolysable ATP analogue (Sinha et al., 2005; Tesmer et al., 1997) or a sulphate (Tews et al., 2005), which has contributed to the stability and formation of the active state.
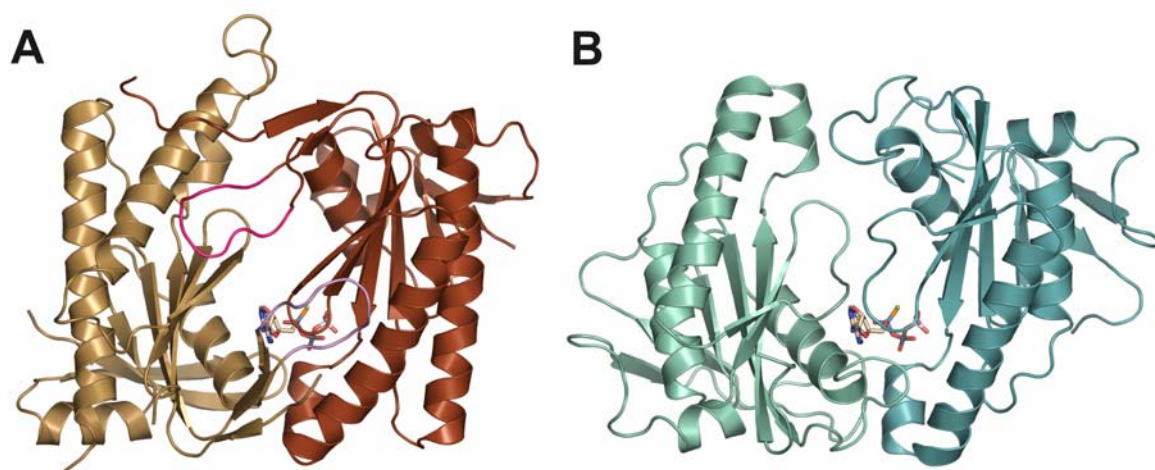

Figure 4.4: Modelled active CHD of Rv0386 compared to the rat AC type II C2 AC and canine AC type V C1 (PDB code: 1CJK). The ATPαS-Rp as described in Figure 4.3. **A)** Mammalian heterodimer C2 in beige, C1 in brown rendered in cartoon. Dimerisation arm high lighted in magenta and violet. **B)** Rv0386 homodimer in cartoon rendered in greencyan and darkcyan.

4.4 Modelled Active Interface

4.4.a CHD Interface of Modelled Active State of Rv0386 Compared to Active Rv1264

The catalytic domain of Rv0386 was superimposed on the active apo-form CHD of Rv1264 (PDB entry code1Y11) (rmsd of 1.66 Å for 124 Cα-atoms). Molecule D and B of Rv0386 were overlaid on active Rv1264 B and A respectively. Rv1264 was chosen for comparison, since the dimerisation arm as well is missing and it is crystallised in the active and inhibited state. The modeled active state of the catalytic domain of Rv0386 embodies an interface of 1361 $Å^2$ compared to the catalytic domain of Rv1264 that forms an interface of 1900 $Å^2$. The interface of the modelled active state of Rv0386 CHD contains hydrophobic and polar interactions. The interactions in the interface are described as follows. Nitrogen of the main-chain amide $Leu158_D$ (loop β_7- β_8) hydrogen bonds to the side-chain of $Glu25_B$ ($\alpha1$). Polar contacts are also established between the side-chains of $Thr21_B$ ($\alpha1$), and $Arg117_D$ ($\alpha4$) and the nitrogen side-chain of $Asn113_D$ ($\alpha4$). Side-chains of $Asp120_D$ ($\alpha4$), $Arg119_D$ ($\alpha4$) and $Glu59_B$ (loop β_2-β_3) form polar contacts. Main-chain carbonyl oxygen of $Glu56_B$ (loop β_2-β_3) hydrogen bond to nitrogen of the main-chain amide of $Glu59_D$ (loop β_2-β_3). The side–chain of $Glu104_B$ (β_4- α_4) is as well hydrogen bonding to the nitrogen of the amide $Gly32_D$ ($\alpha2$). A hydrophobic interaction is established between $Trp24_B$ ($\alpha1$) and $Ala108_D$ (β_4- α_4).

The model of the catalytic domain of Rv0386 demonstrates that Rv0386 can arrange in the active state. Minor clashes are observed in the model between $Gly58_B$ and $Gly58_D$ (loop β_2-β_3), and between side-chains of $Asn113_B$ (α_4) and $Thr21_D$ (α_1). In the genuine active state, loop β_2-β_3 could be flexible and possible interactions can be established between Asn/Thr.

4.5 Transition of the Inhibited to the Expected Active State

4.5.a Regulation of CHD of Rv0386

In this study the X-ray structure of the catalytic domain of the adenylyl cyclase Rv0386 has been determined. For the first time an AC is established in an inhibited head-to-head fashion. The exact alignment of CHDs in the active state is important for the regulation of adenylyl cyclases, as the catalytic dimer recognises residues from both chains to form proper catalytic centres. The movement required to convert the inhibited to the modelled active conformation of CHD Rv0386 was investigated using the program DynDom {Hayward, 1998 #93; http://www.sys.uea.ac.uk/dyndom}. The inhibited and the active states were superimposed, generating one fixed molecule for the two states. DynDom calculated the domain motion i.e the rotation/translation vector to convert one conformation into the other.

The catalytic residues in the inhibited state of CHD Rv0386 are 50 Å apart. For the active state to be formed, the catalytic monomers are rotated with 48.2° and translated with 16.8 Å along the rotation axis, (Figure 4.5) creating two catalytic sites in the interface between the monomers with catalytic residues 25 Å apart. A schematic drawing visualises this effect (Figure 4.6).

Other crystal structures of adenylyl cyclases also show domain movements. The active mammalian heterodimer of the C1, C2 catalytic domain bound to an activating Gsα protein and ATP analogue (Tesmer et al., 1997) has been compared to the C2 homodimer (Zhang et al., 1997) resulting in a rotation of 7°, which has been suggested to be the activation mechanism. The significance of this comparison is unclear, due to two different constructs employed. One is an inactive homodimer with no catalytic centres and the other an active heterodimer with two catalytic centres. Still, the crystal structure of the mammalian CHD including the transmembrane helices is missing and the complete AC enzyme could possibly generate a larger movement with additional membrane spanning segments and Giα protein. The open to closed conformation of CHD Rv1900c binding AMPCPP, results in a rotation of 16.6° and translation of 11.4 Å along the rotation axis (Sinha et al., 2005). The additional N terminal α/β-hydrolase domain of Rv1900c has not been crystallised together with the

catalytic domain and therefore one cannot derive its regulatory foundation. It is not clear if Rv1900c also might form an inhibited state just as Rv0386 and Rv1264. The only known inhibited structure so far, has been from holoenzyme Rv1264 (Tews et al., 2005). For the active state of Rv1264 to be formed from the inhibited state, the αN10-switch helix unforms, α1-switch forms and the catalytic monomers are rotated 55° with a translation of 6 Å along the rotation axis. These sterical changes causes 40 fold activation. The regulatory domain is barely changed and interacts mainly with the catalytic domain. Rv0386 in the active state is likewise expected to be stabilised by interaction of the catalytic domain with other regulatory domains.
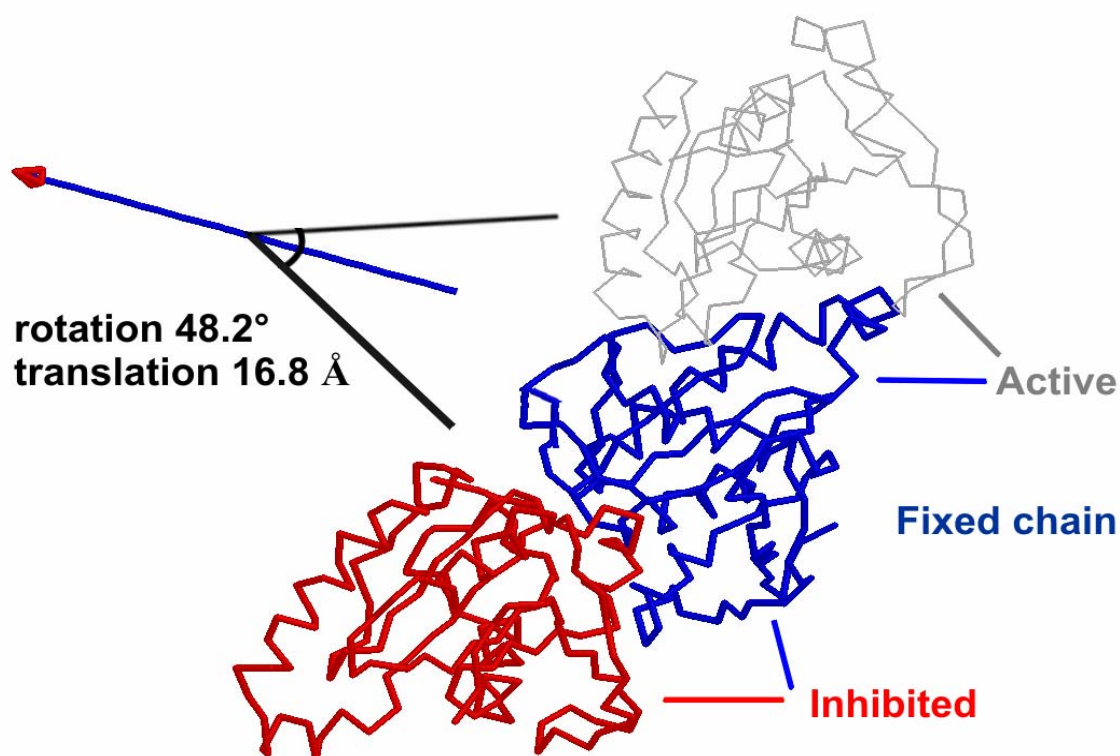
Figure 4.5: Movement of the inhibited to the modelled active conformation of CHD Rv0386. The two protein states in ribbon, and superimposed where the fixed chain and the screw axis is coloured in blue. Inhibited catalytic domain of Rv0386 rotates 48.2 ° and 16.8 Å translation along rotation axis creating the active state. The inhibited head-to-head dimer in red, blue and the modelled active head-to-tail dimer in blue, white. Picture was generated by DynDom (Hayward and Berendsen, 1998)and Rasmol (Sayle and Milner-White, 1995).

Figure 4.6: Schematic drawing of the model of Rv0386 CHD activation. Left panel: inhibited state with disassembled active sites in grey and beige. Middle panel: movement of the monomers with rotation of 48.2° and translation of 16.8 Å along the rotation axis. Monomers coloured in light green. Right panel: active state in dark blue and aqua marine with two ATPs bound in the active sites.

4.5.b Putative β5-switch

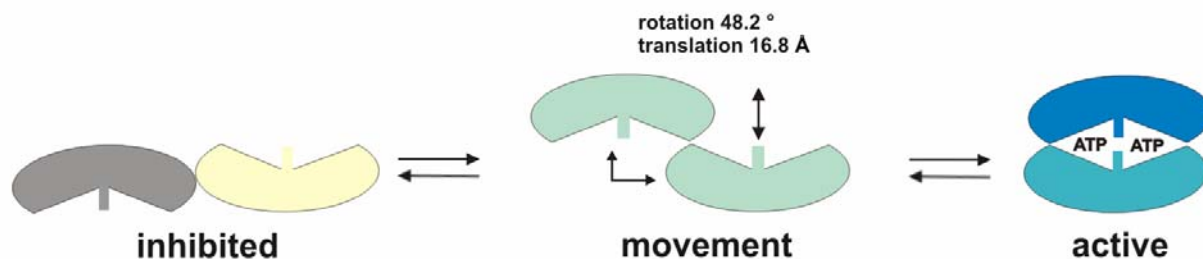Comparison of the catalytic domain of Rv0386 with homologous CHDs revealed a difference in the secondary structure elements (Table 4.1). In other class III CHD structures from *M. tuberculosis, T. Brucei* and mammals, β5 is present. So far, the Rv0386 CHD structure is the only one with absent β5. Instead, a random coiled loop is observed. This region comprises the non conserved catalytic residue Asn106, which binds to the purine of the substrate in the modeled active state. This is in agreement with the purine binding residues D1018 and I1019 from mammalian AC as well located at β5 (Figure 1.8). In the inhibited crystal structure of CHD Rv0386 this catalytic residue is not oriented in the same direction as corresponding residue in other CHD crystal structures. In the electron density Asn106 has two conformations, which shows its flexibility. The active conformation of the protein would contain Asn106 in a different orientation and binding the substrate.

The structures of the holoenzyme Rv1264 show the relationship between the catalytic and regulatory domains in the inhibited and active state (Tews et al., 2005). The α1-switch binds to the β/γ-phosphates in the mammalian ACs and it is expected to do that as well for Rv1264. The inhibited state of Rv1264 contains an unformed α1-switch and in the active state it is forming an α-helix. As a comparison, the inhibited head-to-head CHD Rv0386 has an unformed β5. For the active state of Rv0386 to be created it is possible that β5 is formed. This region could also improve the stability of the active dimer, increasing the affinity of the substrate and the buried surface. It is possible that unformed β5 in homologous structures

could result in an inhibited head-to-head conformation. Other AC structures need to be determined, to confirm this hypothesis.

4.6 Conclusions

This work presents the crystal and solution structure of the catalytic core of AC Rv0386. The enzyme was determined by X-ray in an inhibited state forming the unique head-to-head homodimer. The extensive interface buries 1497 Å^2 of solvent-accessible surface area per monomer and is in agreement with the solution structure determined by SAXS where this confirmation could be confirmed. The two active sites are 50 Å apart and cannot bind and convert ATP and GTP to cAMP and cGMP respectively. This study suggests that Mn^{2+} or a putative external activator make the transition of the inhibited to form the proper active state of the catalytic core of Rv0386. This rearrangement requires the monomers to rotate by 48.2° and translate by 16.8 Å along the rotation axis, creating two catalytic sites in the interface between the monomers with catalytic residues 25 Å apart.

A difference in secondary structure elements of our head-to-head dimer compared to general head-to-tail ACs, reveals a shorter β-sheet i.e absence of β5 in CHD Rv0386, highly conserved in all class III ACs. This β5 is a part of the modeled active interface responsible for the purine binding residue Asn106. This study suggests that β5 is formed in the active state of Rv0386. The region might improve the stability of the active dimer, increasing the affinity of the substrate. It would also increase the buried surface of 1361 Å. It is possible that not formed β5 in homologous structures could result in an inhibited head-to-head conformation.

The binding mode of ATP and GTP has been proposed from the modeled active state of Rv0386. The non conserved purine binding residues Asn106 and Gln57 are able to bind both substrates, possibly by 180° rotation of both amide side chains.

Bibliography

Abrahams, J.P. and Leslie, A.G. (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr*, **52**, 30-42.

Ahuja, N., Kumar, P. and Bhatnagar, R. (2004) The adenylate cyclase toxins. *Crit Rev Microbiol*, **30**, 187-196.

Alberts, B., Johnson, R.A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002) *Molecular Biology of the Cell*. Garland Science, New York.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M. and Iyer, L.M. (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev*, **29**, 231-262.

Armstrong, J.A. and Hart, P.D. (1975) Phagosome-lysosome interactions in cultured macrophages infected with virulent tubercle bacilli. Reversal of the usual nonfusion pattern and observations on bacterial survival. *J Exp Med*, **142**, 1-16.

Artymiuk, P.J., Poirrette, A.R., Rice, D.W. and Willett, P. (1997) A polymerase I palm in adenylyl cyclase? *Nature*, **388**, 33-34.

Baker, D.A. and Kelly, J.M. (2004) Structure, function and evolution of microbial adenylyl and guanylyl cyclases. *Mol Microbiol*, **52**, 1229-1242.

Barzu, O. and Danchin, A. (1994) Adenylyl cyclases: a heterogeneous class of ATP-utilizing enzymes. *Prog Nucleic Acid Res Mol Biol*, **49**, 241-283.

Berlot, C.H. and Bourne, H.R. (1992) Identification of effector-activating residues of Gs alpha. *Cell*, **68**, 911-922.

Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.H., Srinivasan, A.R. and Schneider, B. (1992) The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys J*, **63**, 751-759.

Bieger, B. and Essen, L.O. (2001) Structural analysis of adenylate cyclases from Trypanosoma brucei in their monomeric state. *Embo J*, **20**, 433-445.

Bond, C.S. (2003) TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics*, **19**, 311-312.

Brightbill, H.D., Libraty, D.H., Krutzik, S.R., Yang, R.B., Belisle, J.T., Bleharski, J.R., Maitland, M., Norgard, M.V., Plevy, S.E., Smale, S.T., Brennan, P.J., Bloom, B.R., Godowski, P.J. and Modlin, R.L. (1999) Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors. *Science*, **285**, 732-736.

Brunger, A.T. (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr D Biol Crystallogr*, **49**, 24-36.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, **54**, 905-921.

Buck, J., Sinclair, M.L., Schapal, L., Cann, M.J. and Levin, L.R. (1999) Cytosolic adenylyl cyclase defines a unique signaling molecule in mammals. *Proc Natl Acad Sci U S A*, **96**, 79-84.

Cali, J.J., Zwaagstra, J.C., Mons, N., Cooper, D.M. and Krupinski, J. (1994) Type VIII adenylyl cyclase. A Ca2+/calmodulin-stimulated enzyme expressed in discrete regions of rat brain. *J Biol Chem*, **269**, 12190-12195.

Cann, M.J., Hammer, A., Zhou, J. and Kanacher, T. (2003) A defined subset of adenylyl cyclases is regulated by bicarbonate ion. *J Biol Chem*, **278**, 35033-35038.

_____

Carrington, D. (1999) Hope for new TB treatment.
    http://news.bbc.co.uk/1/hi/sci/tech/specials/sheffield_99/448281.stm.
Castro, L.I. (2004) Mycobacterial adenylyl cyclases Rv1625c and Rv0386:orthodox vs.
    unorthodox catalysis. *Pharmazeutisches Institut, Pharmazeutische Biochemie.*
    Eberhard-Karls-Universität Tübingen, Germany, Tübingen, pp. 1-152.
Castro, L.I., Hermsen, C., Schultz, J.E. and Linder, J.U. (2005) Adenylyl cyclase Rv0386
    from Mycobacterium tuberculosis H37Rv uses a novel mode for substrate selection.
    Febs J, **272**, 3085-3092.
Chen, J., DeVivo, M., Dingus, J., Harry, A., Li, J., Sui, J., Carty, D.J., Blank, J.L., Exton,
    J.H., Stoffel, R.H. and et al. (1995) A region of adenylyl cyclase 2 critical for
    regulation by G protein beta gamma subunits. *Science*, **268**, 1166-1169.
Chen, Y., Cann, M.J., Litvin, T.N., Iourgenko, V., Sinclair, M.L., Levin, L.R. and Buck, J.
    (2000) Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor.
    Science, **289**, 625-628.
Chen, Y., Harry, A., Li, J., Smit, M.J., Bai, X., Magnusson, R., Pieroni, J.P., Weng, G. and
    Iyengar, R. (1997) Adenylyl cyclase 6 is selectively regulated by protein kinase A
    phosphorylation in a region involved in Galphas stimulation. *Proc Natl Acad Sci U S
    A*, **94**, 14100-14104.
Choi, E.J., Xia, Z. and Storm, D.R. (1992) Stimulation of the type III olfactory adenylyl
    cyclase by calcium and calmodulin. *Biochemistry*, **31**, 6492-6498.
Cliff, M.J., Williams, M.A., Brooke-Smith, J., Barford, D. and Ladbury, J.E. (2005)
    Molecular recognition via coupled folding and binding in a TPR domain. *J Mol Biol*,
    346, 717-732.
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V.,
    Eiglmeier, K., Gas, S., Barry, C.E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown,
    D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,
    Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Barrell, B.G. and et al. (1998)
    Deciphering the biology of Mycobacterium tuberculosis from the complete genome
    sequence. *Nature*, **393**, 537-544.
Conti, M. and Jin, S.L. (1999) The molecular biology of cyclic nucleotide phosphodiesterases.
    Prog Nucleic Acid Res Mol Biol, **63**, 1-38.
Cooper, D.M. (2003) Regulation and organization of adenylyl cyclases and cAMP. *Biochem
    J*, **375**, 517-529.
Cooper, D.M., Karpen, J.W., Fagan, K.A. and Mons, N.E. (1998) Ca(2+)-sensitive adenylyl
    cyclases. *Adv Second Messenger Phosphoprotein Res*, **32**, 23-51.
Cotta, M.A., Whitehead, T.R. and Wheeler, M.B. (1998) Identification of a novel adenylate
    cyclase in the ruminal anaerobe, Prevotella ruminicola D31d. *FEMS Microbiol Lett*,
    164, 257-260.
Coudart-Cavalli, M.P., Sismeiro, O. and Danchin, A. (1997) Bifunctional structure of two
    adenylyl cyclases from the myxobacterium Stigmatella aurantiaca. *Biochimie*, **79**, 757-
    767.
Danchin, A. (1993) Phylogeny of adenylyl cyclases. *Adv Second Messenger Phosphoprotein
    Res*, **27**, 109-162.
de la Fortelle, E. and Bricogne, G. (1997) *Methods in Enzymology*.
de Rooij, J., Zwartkruis, F.J., Verheijen, M.H., Cool, R.H., Nijman, S.M., Wittinghofer, A.
    and Bos, J.L. (1998) Epac is a Rap1 guanine-nucleotide-exchange factor directly
    activated by cyclic AMP. *Nature*, **396**, 474-477.
DeLano, W.L. (2002) The PYMOL Molecular Graphics System. *DeLano Scientific, San
    Carlos, CA, USA.*
Desaubry, L. and Johnson, R.A. (1998) Adenine nucleoside 3'-tetraphosphates are novel and
    potent inhibitors of adenylyl cyclases. *J Biol Chem*, **273**, 24972-24977.

Desaubry, L., Shoshani, I. and Johnson, R.A. (1996a) 2',5'-Dideoxyadenosine 3'-polyphosphates are potent inhibitors of adenylyl cyclases. *J Biol Chem*, **271**, 2380-2382.

Desaubry, L., Shoshani, I. and Johnson, R.A. (1996b) Inhibition of adenylyl cyclase by a family of newly synthesized adenine nucleoside 3'-polyphosphates. *J Biol Chem*, **271**, 14028-14034.

Dessauer, C.W. and Gilman, A.G. (1997) The catalytic mechanism of mammalian adenylyl cyclase. Equilibrium binding and kinetic analysis of P-site inhibition. *J Biol Chem*, **272**, 27787-27795.

Dessauer, C.W., Scully, T.T. and Gilman, A.G. (1997) Interactions of forskolin and ATP with the cytosolic domains of mammalian adenylyl cyclase. *J Biol Chem*, **272**, 22272-22277.

Dessauer, C.W., Tesmer, J.J., Sprang, S.R. and Gilman, A.G. (1998) Identification of a Gialpha binding site on type V adenylyl cyclase. *J Biol Chem*, **273**, 25831-25839.

Dessauer, C.W., Tesmer, J.J., Sprang, S.R. and Gilman, A.G. (1999) The interactions of adenylate cyclases with P-site inhibitors. *Trends Pharmacol Sci*, **20**, 205-210.

Doublie, S., Sawaya, M.R. and Ellenberger, T. (1999) An open and closed case for all polymerases. *Structure*, **7**, R31-35.

Drenth, J. (1994) *Principles of Protein X-ray crystallography*. Springer-Verlag, New York.

Drum, C.L., Yan, S.Z., Bard, J., Shen, Y.Q., Lu, D., Soelaiman, S., Grabarek, Z., Bohm, A. and Tang, W.J. (2002) Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature*, **415**, 396-402.

Fagan, K.A., Graf, R.A., Tolman, S., Schaack, J. and Cooper, D.M. (2000) Regulation of a Ca2+-sensitive adenylyl cyclase in an excitable cell. Role of voltage-gated versus capacitative Ca2+ entry. *J Biol Chem*, **275**, 40187-40194.

Ferrari, G., Langen, H., Naito, M. and Pieters, J. (1999) A coat protein on phagosomes involved in the intracellular survival of mycobacteria. *Cell*, **97**, 435-447.

Findeisen, F. (2005) Structural characterisation of signalling molecules: The Drosophila melanogaster GPCR Frizzled and the Mycobacterium tuberculosis adenylyl cyclase Rv1264. Rupertus Carolus University of Heidelberg, Germany, Heidelberg, pp. 1-187.

Florio, V.A. and Ross, E.M. (1983) Regulation of the catalytic component of adenylate cyclase. Potentiative interaction of stimulatory ligands and 2',5'-dideoxyadenosine. *Mol Pharmacol*, **24**, 195-202.

Gamieldien, J., Ptitsyn, A. and Hide, W. (2002) Eukaryotic genes in Mycobacterium tuberculosis could have a role in pathogenesis and immunomodulation. *Trends Genet*, **18**, 5-8.

Gao, B.N. and Gilman, A.G. (1991) Cloning and expression of a widely distributed (type IV) adenylyl cyclase. *Proc Natl Acad Sci U S A*, **88**, 10178-10182.

Garcia De La Torre, J., Huertas, M.L. and Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J*, **78**, 719-730.

Garman, E. and Schneider, T.R. (1997) Macromolecular cryocrystallography. *J. Appl. Cryst.*, **30**, 211-237.

Gatfield, J. and Pieters, J. (2000) Essential role for cholesterol in entry of mycobacteria into macrophages. *Science*, **288**, 1647-1650.

Glaser, P., Ladant, D., Sezer, O., Pichot, F., Ullmann, A. and Danchin, A. (1988) The calmodulin-sensitive adenylate cyclase of Bordetella pertussis: cloning and expression in Escherichia coli. *Mol Microbiol*, **2**, 19-30.

Gu, C. and Cooper, D.M. (2000) Ca(2+), Sr(2+), and Ba(2+) identify distinct regulatory sites on adenylyl cyclase (AC) types VI and VIII and consolidate the apposition of

capacitative cation entry channels and Ca(2+)-sensitive ACs. *J Biol Chem*, **275**, 6980-6986.

Guo, Y.L., Seebacher, T., Kurz, U., Linder, J.U. and Schultz, J.E. (2001) Adenylyl cyclase Rv1625c of Mycobacterium tuberculosis: a progenitor of mammalian adenylyl cyclases. *Embo J*, **20**, 3667-3675.

Harth, G. and Horwitz, M.A. (1999) An inhibitor of exported Mycobacterium tuberculosis glutamine synthetase selectively blocks the growth of pathogenic mycobacteria in axenic culture and in human monocytes: extracellular proteins as potential novel drug targets. *J Exp Med*, **189**, 1425-1436.

Hayward, S. and Berendsen, H.J. (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins*, **30**, 144-154.

Horinouchi, S., Kito, M., Nishiyama, M., Furuya, K., Hong, S.K., Miyake, K. and Beppu, T. (1990) Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in Streptomyces coelicolor A3(2). *Gene*, **95**, 49-56.

http://alf1.mrc-lmb.cam.ac.uk/~ramak/madms/segrowth.html.

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin/sld004.htm.

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin/sld006.htm.

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin/sld024.htm.

http://www.infektionsbiologie.ch. Mycobacterium tuberculosis.

Hurley, J.H. (1998) The adenylyl and guanylyl cyclase superfamily. *Curr Opin Struct Biol*, **8**, 770-777.

Hurley, J.H. (1999) Structure, mechanism, and regulation of mammalian adenylyl cyclase. *J Biol Chem*, **274**, 7599-7602.

Ingi, T., Krumins, A.M., Chidiac, P., Brothers, G.M., Chung, S., Snow, B.E., Barnes, C.A., Lanahan, A.A., Siderovski, D.P., Ross, E.M., Gilman, A.G. and Worley, P.F. (1998) Dynamic regulation of RGS2 suggests a novel mechanism in G-protein signaling and neuronal plasticity. *J Neurosci*, **18**, 7178-7188.

Iwami, G., Kawabe, J., Ebina, T., Cannon, P.J., Homcy, C.J. and Ishikawa, Y. (1995) Regulation of adenylyl cyclase by protein kinase A. *J Biol Chem*, **270**, 12481-12484.

Johnson, R.A., Saur, W. and Jakobs, K.H. (1979) Effects of prostaglandin E1 and adenosine on metal and metal-ATP kinetics of platelet adenylate cyclase. *J Biol Chem*, **254**, 1094-1101.

Johnson, R.A., Yeung, S.M., Stubner, D., Bushfield, M. and Shoshani, I. (1989) Cation and structural requirements for P site-mediated inhibition of adenylate cyclase. *Mol Pharmacol*, **35**, 681-688.

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, **47 (Pt 2)**, 110-119.

Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-2637.

Kanacher, T., Schultz, A., Linder, J.U. and Schultz, J.E. (2002) A GAF-domain-regulated adenylyl cyclase from Anabaena is a self-activating cAMP switch. *Embo J*, **21**, 3672-3680.

Kasahara, M., Unno, T., Yashiro, K. and Ohmori, M. (2001) CyaG, a novel cyanobacterial adenylyl cyclase and a possible ancestor of mammalian guanylyl cyclases. *J Biol Chem*, **276**, 10564-10569.

Kather, H. and Aktories, K. (1983) [The cAMP system and bacterial toxins]. *Klin Wochenschr*, **61**, 1109-1114.

Kaufmann, S.H. (2000) Is the development of a new tuberculosis vaccine possible? *Nat Med*, **6**, 955-960.

Kaufmann, S.H. (2001) How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol*, **1**, 20-30.

Kaufmann, S.H. (2002) Protection against tuberculosis: cytokines, T cells, and macrophages. *Ann Rheum Dis*, **61 Suppl 2**, ii54-58.

Kehrl, J.H. and Sinnarajah, S. (2002) RGS2: a multifunctional regulator of G-protein signaling. *Int J Biochem Cell Biol*, **34**, 432-438.

Klumpp, S. and Schultz, J.E. (1982) Characterization of a Ca2+-dependent guanylate cyclase in the excitable ciliary membrane from Paramecium. *Eur J Biochem*, **124**, 317-324.

Koch, M.H., Vachette, P. and Svergun, D.I. (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys*, **36**, 147-227.

Konarev, P.V., Petoukov, M.V. and Svergun, D.I. (2001) MASSHA - a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. *Journal of Applied Crystallography*, **34**, 527-532.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. and Svergun, D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, **36**, 1277-1282.

Krupinski, J., Coussen, F., Bakalyar, H.A., Tang, W.J., Feinstein, P.G., Orth, K., Slaughter, C., Reed, R.R. and Gilman, A.G. (1989) Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure. *Science*, **244**, 1558-1564.

Kudlacek, O., Mitterauer, T., Nanoff, C., Hohenegger, M., Tang, W.J., Freissmuth, M. and Kleuss, C. (2001) Inhibition of adenylyl and guanylyl cyclase isoforms by the antiviral drug foscarnet. *J Biol Chem*, **276**, 3010-3016.

Lambright, D.G., Noel, J.P., Hamm, H.E. and Sigler, P.B. (1994) Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. *Nature*, **369**, 621-628.

Lange, C., Schieferstein, C. and Toossi, Z. (2005) Opportunistic Infections; Tuberculosis. In Hoffmann, C., Rockstroh, J.K. and Kamps, B.S. (eds.), *HIV Medicine 2005*. Steinhäuser Verlag, pp. 388-400.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the strereochemical quality of protein structures. *Journal of Applied Crystallography*, **26**, 283-291.

Laue, T.M., Shah, B.D., Ridgeway, T.M. and Pelletier, S.L. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Royal Society of Chemistry, Cambridge.

Linder, J.U.

Linder, J.U., Hammer, A. and Schultz, J.E. (2004) The effect of HAMP domains on class IIIb adenylyl cyclases from Mycobacterium tuberculosis. *Eur J Biochem*, **271**, 2446-2451.

Linder, J.U., Schultz, A. and Schultz, J.E. (2002) Adenylyl cyclase Rv1264 from Mycobacterium tuberculosis has an autoinhibitory N-terminal domain. *J Biol Chem*, **277**, 15271-15276.

Linder, J.U. and Schultz, J.E. (2003) The class III adenylyl cyclases: multi-purpose signalling modules. *Cell Signal*, **15**, 1081-1089.

Lineweaver, H. and Burk, D. (1934) The determination of enzyme dissociation constants. *J. Am. Chem. Soc.*, **56**, 658-666.

Liu, Y., Ruoho, A.E., Rao, V.D. and Hurley, J.H. (1997) Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling and mutational analysis. *Proc Natl Acad Sci U S A*, **94**, 13414-13419.

Londos, C. and Wolff, J. (1977) Two distinct adenosine-sensitive sites on adenylate cyclase. *Proc Natl Acad Sci U S A*, **74**, 5482-5486.

_____

Lowrie, D.B., Aber, V.R. and Jackett, P.S. (1979) Phagosome-lysosome fusion and cyclic adenosine 3':5'-monophosphate in macrophages infected with Mycobacterium microti, Mycobacterium bovis BCG or Mycobacterium lepraemurium. *J Gen Microbiol*, **110**, 431-441.

Lowrie, D.B., Jackett, P.S. and Ratcliffe, N.A. (1975) Mycobacterium microti may protect itself from intracellular destruction by releasing cyclic AMP into phagosomes. *Nature*, **254**, 600-602.

Manabe, Y.C. and Bishai, W.R. (2000) Latent Mycobacterium tuberculosis-persistence, patience, and winning by waiting. *Nat Med*, **6**, 1327-1329.

McCue, L.A., McDonough, K.A. and Lawrence, C.E. (2000) Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in Mycobacterium tuberculosis. *Genome Res*, **10**, 204-219.

Miller, R., Gallo, S.M., Khalak, H.G. and Weeks, C.M. (1994) SnB: crystal structure determination via shake-and-bake. *Journal of Applied Crystallography*, **27**, 613-621.

Mock, M., Labruyere, E., Glaser, P., Danchin, A. and Ullmann, A. (1988) Cloning and expression of the calmodulin-sensitive Bacillus anthracis adenylate cyclase in Escherichia coli. *Gene*, **64**, 277-284.

Motaal. (2006) Mycobacterium tuberculosis adenylyl cyclases

Rv2435c and Rv2212c. *Fakultät für Chemie und Pharmazie*. Eberhard-Karls-Universität Tübingen, Tübingen, pp. 1-133.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. and Dodson, E.J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 1)**, 247-255.

Naula, C., Schaub, R., Leech, V., Melville, S. and Seebeck, T. (2001) Spontaneous dimerization and leucine-zipper induced activation of the recombinant catalytic domain of a new adenylyl cyclase of Trypanosoma brucei, GRESAG4.4B. *Mol Biochem Parasitol*, **112**, 19-28.

Otwinowski, Z. and Minor, W. (1997) *Methods in Enzymology,*

Macromolecular Crystallography, part A. Academic Press, San Diego.

Perrakis, A., Morris, R. and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol*, **6**, 458-463.

Petoukhov, M.V. and Svergun, D.I. (2005) Global rigid body modelling of macromolecular complexes against small-angle scattering data. *Biophys J*.

Rall, T.W. and Sutherland, E.W. (1962) Adenyl cyclase. II. The enzymatically catalyzed formation of adenosine 3',5'-phosphate and inorganic pyrophosphate from adenosine triphosphate. *J Biol Chem*, **237**, 1228-1232.

Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. and Sweet, R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, **362**, 219-223.

Reddy, S.K., Kamireddi, M., Dhanireddy, K., Young, L., Davis, A. and Reddy, P.T. (2001) Eukaryotic-like adenylyl cyclases in Mycobacterium tuberculosis H37Rv: cloning and characterization. *J Biol Chem*, **276**, 35141-35149.

Rossmann, M.G. (1972) *The molecular replacement method.* Gordon & Breach, NewYork.

Roy, A., Danchin, A., Joseph, E. and Ullmann, A. (1983) Two functional domains in adenylate cyclase of Escherichia coli. *J Mol Biol*, **165**, 197-202.

Russell, D.G., Dant, J. and Sturgill-Koszycki, S. (1996) Mycobacterium avium- and Mycobacterium tuberculosis-containing vacuoles are dynamic, fusion-competent vesicles that are accessible to glycosphingolipids from the host cell plasmalemma. *J Immunol*, **156**, 4764-4773.

Salomon, Y., Londos, C. and Rodbell, M. (1974) A highly sensitive adenylate cyclase assay. *Anal Biochem*, **58**, 541-548.

Sayle, R.A. and Milner-White, E.J. (1995) RASMOL: biomolecular graphics for all. *Trends Biochem Sci*, **20**, 374.

Schaible, U.E., Collins, H.L. and Kaufmann, S.H. (1999) Confrontation between intracellular bacteria and the immune system. *Adv Immunol*, **71**, 267-377.

Scholich, K., Mullenix, J.B., Wittpoth, C., Poppleton, H.M., Pierre, S.C., Lindorfer, M.A., Garrison, J.C. and Patel, T.B. (1999) Facilitation of signal onset and termination by adenylyl cyclase. *Science*, **283**, 1328-1331.

Scholich, K., Pierre, S. and Patel, T.B. (2001) Protein associated with Myc (PAM) is a potent inhibitor of adenylyl cyclases. *J Biol Chem*, **276**, 47583-47589.

Schorey, J.S., Carroll, M.C. and Brown, E.J. (1997) A macrophage invasion mechanism of pathogenic mycobacteria. *Science*, **277**, 1091-1093.

Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J*, **78**, 1606-1619.

Schultz, J.E., Klumpp, S., Benz, R., Schurhoff-Goeters, W.J. and Schmid, A. (1992) Regulation of adenylyl cyclase from Paramecium by an intrinsic potassium conductance. *Science*, **255**, 600-603.

Shenoy, A.R., Sreenath, N.P., Mahalingam, M. and Visweswariah, S.S. (2004) Characterization of phylogenetically distant members of the adenylyl cyclase family from mycobacteria: Rv1647 from M. tuberculosis and its ortholog ML1399 from M. leprae. *Biochem J*.

Simsova, M., Sebo, P. and Leclerc, C. (2004) The adenylate cyclase toxin from Bordetella pertussis--a novel promising vehicle for antigen delivery to dendritic cells. *Int J Med Microbiol*, **293**, 571-576.

Sinha, S.C., Wetterer, M., Sprang, S.R., Schultz, J.E. and Linder, J.U. (2005) Origin of asymmetry in adenylyl cyclases: structures of Mycobacterium tuberculosis Rv1900c. *Embo J*, **24**, 663-673.

Sinnarajah, S., Dessauer, C.W., Srikumar, D., Chen, J., Yuen, J., Yilma, S., Dennis, J.C., Morrison, E.E., Vodyanoy, V. and Kehrl, J.H. (2001) RGS2 regulates signal transduction in olfactory neurons by attenuating activation of adenylyl cyclase III. *Nature*, **409**, 1051-1055.

Sismeiro, O., Trotot, P., Biville, F., Vivares, C. and Danchin, A. (1998) Aeromonas hydrophila adenylyl cyclase 2: a new class of adenylyl cyclases with thermophilic properties and sequence similarities to proteins from hyperthermophilic archaebacteria. *J Bacteriol*, **180**, 3339-3344.

Steitz, T.A., Smerdon, S., Jager, J., Wang, J., Kohlstaedt, L.A., Friedman, J.M., Beese, L.S. and Rice, P.A. (1993) Two DNA polymerases: HIV reverse transcriptase and the Klenow fragment of Escherichia coli DNA polymerase I. *Cold Spring Harb Symp Quant Biol*, **58**, 495-504.

Steitz, T.A., Smerdon, S.J., Jager, J. and Joyce, C.M. (1994) A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science*, **266**, 2022-2025.

Sternweis, P.C. and Robishaw, J.D. (1984) Isolation of two proteins with high affinity for guanine nucleotides from membranes of bovine brain. *J Biol Chem*, **259**, 13806-13813.

Sunahara, R.K., Beuve, A., Tesmer, J.J., Sprang, S.R., Garbers, D.L. and Gilman, A.G. (1998) Exchange of substrate and inhibitor specificities between adenylyl and guanylyl cyclases. *J Biol Chem*, **273**, 16332-16338.

Sunahara, R.K., Dessauer, C.W. and Gilman, A.G. (1996) Complexity and diversity of mammalian adenylyl cyclases. *Annu Rev Pharmacol Toxicol*, **36**, 461-480.

Sunahara, R.K., Dessauer, C.W., Whisnant, R.E., Kleuss, C. and Gilman, A.G. (1997a) Interaction of Gsalpha with the cytosolic domains of mammalian adenylyl cyclase. *J Biol Chem*, **272**, 22265-22271.

Sunahara, R.K. and Taussig, R. (2002) Isoforms of mammalian adenylyl cyclase: multiplicities of signaling. *Mol Interv*, **2**, 168-184.

Sunahara, R.K., Tesmer, J.J., Gilman, A.G. and Sprang, S.R. (1997b) Crystal structure of the adenylyl cyclase activator Gsalpha. *Science*, **278**, 1943-1947.

Svergun, D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.*, **25**, 495-503.

Svergun, D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J*, **76**, 2879-2886.

Svergun, D.I., Barberato, C. and Koch, M.H.J. (1995) CRYSOL-a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography*, **28**, 768-773.

Tang, W.J. and Gilman, A.G. (1991) Type-specific regulation of adenylyl cyclase by G protein beta gamma subunits. *Science*, **254**, 1500-1503.

Tang, W.J. and Gilman, A.G. (1992) Adenylyl cyclases. *Cell*, **70**, 869-872.

Tang, W.J. and Hurley, J.H. (1998) Catalytic mechanism and regulation of mammalian adenylyl cyclases. *Mol Pharmacol*, **54**, 231-240.

Tang, W.J., Stanzel, M. and Gilman, A.G. (1995) Truncation and alanine-scanning mutants of type I adenylyl cyclase. *Biochemistry*, **34**, 14563-14572.

Tellez-Sosa, J., Soberon, N., Vega-Segura, A., Torres-Marquez, M.E. and Cevallos, M.A. (2002) The Rhizobium etli cyaC product: characterization of a novel adenylate cyclase class. *J Bacteriol*, **184**, 3560-3568.

Terwilliger, T.C. and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr*, **55**, 849-861.

Tesmer, J.J., Dessauer, C.W., Sunahara, R.K., Murray, L.D., Johnson, R.A., Gilman, A.G. and Sprang, S.R. (2000) Molecular basis for P-site inhibition of adenylyl cyclase. *Biochemistry*, **39**, 14464-14471.

Tesmer, J.J. and Sprang, S.R. (1998) The structure, catalytic mechanism and regulation of adenylyl cyclase. *Curr Opin Struct Biol*, **8**, 713-719.

Tesmer, J.J., Sunahara, R.K., Gilman, A.G. and Sprang, S.R. (1997) Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS. *Science*, **278**, 1907-1916.

Tesmer, J.J., Sunahara, R.K., Johnson, R.A., Gosselin, G., Gilman, A.G. and Sprang, S.R. (1999) Two-metal-Ion catalysis in adenylyl cyclase. *Science*, **285**, 756-760.

Tews, I., Findeisen, F., Sinning, I., Schultz, A., Schultz, J.E. and Linder, J.U. (2005) The structure of a pH sensing mycobacterial adenylyl cyclase holoenzyme. *Science*.

Thoma-Uszynski, S., Stenger, S., Takeuchi, O., Ochoa, M.T., Engele, M., Sieling, P.A., Barnes, P.F., Rollinghoff, M., Bolcskei, P.L., Wagner, M., Akira, S., Norgard, M.V., Belisle, J.T., Godowski, P.J., Bloom, B.R. and Modlin, R.L. (2001) Induction of direct antimicrobial activity through mammalian toll-like receptors. *Science*, **291**, 1544-1547.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, **25**, 4876-4882.

Tucker, C.L., Hurley, J.H., Miller, T.R. and Hurley, J.B. (1998) Two amino acid substitutions convert a guanylyl cyclase, RetGC-1, into an adenylyl cyclase. *Proc Natl Acad Sci U S A*, **95**, 5993-5997.

Vaguine, A.A., Richelle, J. and Wodak, S.J. (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 1)**, 191-205.

Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. *J Mol Graph*, **8**, 52-56, 29.

Whisnant, R.E., Gilman, A.G. and Dessauer, C.W. (1996) Interaction of the two cytosolic domains of mammalian adenylyl cyclase. *Proc Natl Acad Sci U S A*, **93**, 6621-6625.

Winn, M.D., Isupov, M.N. and Murshudov, G.N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **57**, 122-133.

Wolff, J., Londos, C. and Cooper, D.M. (1981) Adenosine receptors and the regulation of adenylate cyclase. *Adv Cyclic Nucleotide Res*, **14**, 199-214.

Yan, S.Z., Beeler, J.A., Chen, Y., Shelton, R.K. and Tang, W.J. (2001) The regulation of type 7 adenylyl cyclase by its C1b region and Escherichia coli peptidylprolyl isomerase, SlyD. *J Biol Chem*, **276**, 8500-8506.

Yan, S.Z., Huang, Z.H., Andrews, R.K. and Tang, W.J. (1998) Conversion of forskolin-insensitive to forskolin-sensitive (mouse-type IX) adenylyl cyclase. *Mol Pharmacol*, **53**, 182-187.

Yan, S.Z., Huang, Z.H., Rao, V.D., Hurley, J.H. and Tang, W.J. (1997a) Three discrete regions of mammalian adenylyl cyclase form a site for Gsalpha activation. *J Biol Chem*, **272**, 18849-18854.

Yan, S.Z., Huang, Z.H., Shaw, R.S. and Tang, W.J. (1997b) The conserved asparagine and arginine are essential for catalysis of mammalian adenylyl cyclase. *J Biol Chem*, **272**, 12342-12349.

Zhang, G., Liu, Y., Ruoho, A.E. and Hurley, J.H. (1997) Structure of the adenylyl cyclase catalytic core. *Nature*, **386**, 247-253.

Zheng, B., Ma, Y.C., Ostrom, R.S., Lavoie, C., Gill, G.N., Insel, P.A., Huang, X.Y. and Farquhar, M.G. (2001) RGS-PX1, a GAP for GalphaS and sorting nexin in vesicular trafficking. *Science*, **294**, 1939-1942.

Zimmermann, G., Zhou, D. and Taussig, R. (1998) Genetic selection of mammalian adenylyl cyclases insensitive to stimulation by Gsalpha. *J Biol Chem*, **273**, 6968-6975.

Zippin, J.H., Levin, L.R. and Buck, J. (2001) CO(2)/HCO(3)(-)-responsive soluble adenylyl cyclase as a putative metabolic sensor. *Trends Endocrinol Metab*, **12**, 366-370.