Chapter 4

Spin-labels - a novel approach for the structure determination of multi-domain proteins

A structure determination of a protein the size of a pleckstrin double-domain construct (28-30kDa) is a very difficult task by conventional NMR methods, even though the structures and the assignments of all pleckstrin single domains are known. Residual dipolar couplings (RDCs) can be used to determine the relative orientation of two or more domains (Skrynnikov et al., 2000; Varadan et al., 2004), but they are insensitive to how close or far the domains are from each other. Since the inter-domain linkers between the domains of pleckstrin are about fifteen residues in length, the domains are not restricted to stay close to each other in a NMR structure calculation with RDCs alone. If the two domains do specifically interact with each other and possibly form a compound structure, a different method has to be employed to report on the "contact area" of both domains. Chemical shift perturbation analysis which compares the single domains with the double-domain constructs would provide this type of information, but unfortunately the data are inconclusive for N-PH_DEP (B. Simon, unpublished results); for DEP_C-PH, no change in chemical shift is observed on C-PH, implying that the two domains may not interact at all (chapter 3).

In the current chapter, the approach to employ spin-labels to obtain long-range structural information is applied to pleckstrin single- and double-domain constructs. The method has already been used successfully by others to calculate the structure of a relatively large globular protein (Battiste and Wagner, 2000). Here, it is employed to scan the surface of the multi-domain protein pleckstrin so as to define the regions where the domains are close in space. In this way, "contact information" is obtained and combined with RDC data for a NMR structure determination.

Spin-labels are usually attached to proteins via cysteine sidechains. Hence, introducing cysteines by site-directed mutagenesis offers the possibility to attach spin-labels at any chosen position. This strategy is known as site-directed spin-labelling (Hubbell et al., 1998). Until now, it has only been applied to small and/or intrinsically cysteine-free proteins because a cysteine-free protein is required as a starting point. The work presented in this chapter is the first instance that site directed spin-labelling is applied to a multi-domain protein with 7 native cysteine residues.

4.1 Results

4.1.1 Analysis of spin-labelled pleckstrin C-PH

When one or more spin-labels are attached to a protein via cysteine sidechains, the resonance lines of nuclei that are close in space to a spin-label are broadened and the intensity of their peaks is reduced (see section 1.2.4 "Theory"). This effect is called paramagnetic relaxation enhancement (PRE). The reduction in peak intensity can be measured by comparing a spectrum where the spin-label is active (PRE spectrum) with a reference spectrum where the spin-label is chemically inactivated, e.g. by reduction with ascorbic acid.

To get a first impression of PRE, all pleckstrin single and double domain constructs (N-PH, DEP, N-PH_DEP, DEP_C-PH and C-PH) are spin-labelled following a published protocol (Battiste and Wagner, 2000) with slight modifications in the purification procedure (cf. Chapter 6 "Materials and Methods"). The $^{1}H,^{15}N$-HSQC spectra of most constructs are so severely affected by the spin-labels that only few NMR signals are left – the majority of peaks have strongly increased linewidths. Most likely, this is the result of having several spin-labels attached to each molecule (see below). Only the spectrum of C-PH is notably different: about half of the peaks are unaffected by the spin-label, while the other half displays varying degrees of reduced peak intensity (Fig 4.1 panel A and B). When the residues with reduced peak intensity are mapped onto the structure of C-PH (Fig. 4.1C), they cluster significantly on one side of the domain. The first β-sheet of C-PH is strongly affected (β1, β2, and β3 strands), whereas the reverse side of the domain comprising mainly the second β-sheet (β5, β6 and β7 strands) is unaffected.

There are two cysteine residues on C-PH. Cys^{250} is located on the β1 strand, whereas Cys^{295} resides in the β4–β5 loop (yellow and white spheres, respectively, in Fig. 4.1C). Clearly, peaks with reduced intensity in the PRE spectrum are clustered in space around Cys^{250}, while the region of Cys^{295} is unaffected by PRE, including Cys^{295} itself. All of this strongly suggests that Cys^{250} is spin-labelled and Cys^{295} is not, in agreement with the structure of C-PH (Chapter 2): Cys^{250} is fully solvent exposed, while Cys^{295} is buried in the core of the domain. There are several NOEs between the H^{γ} of Cys^{295} and other core residues.

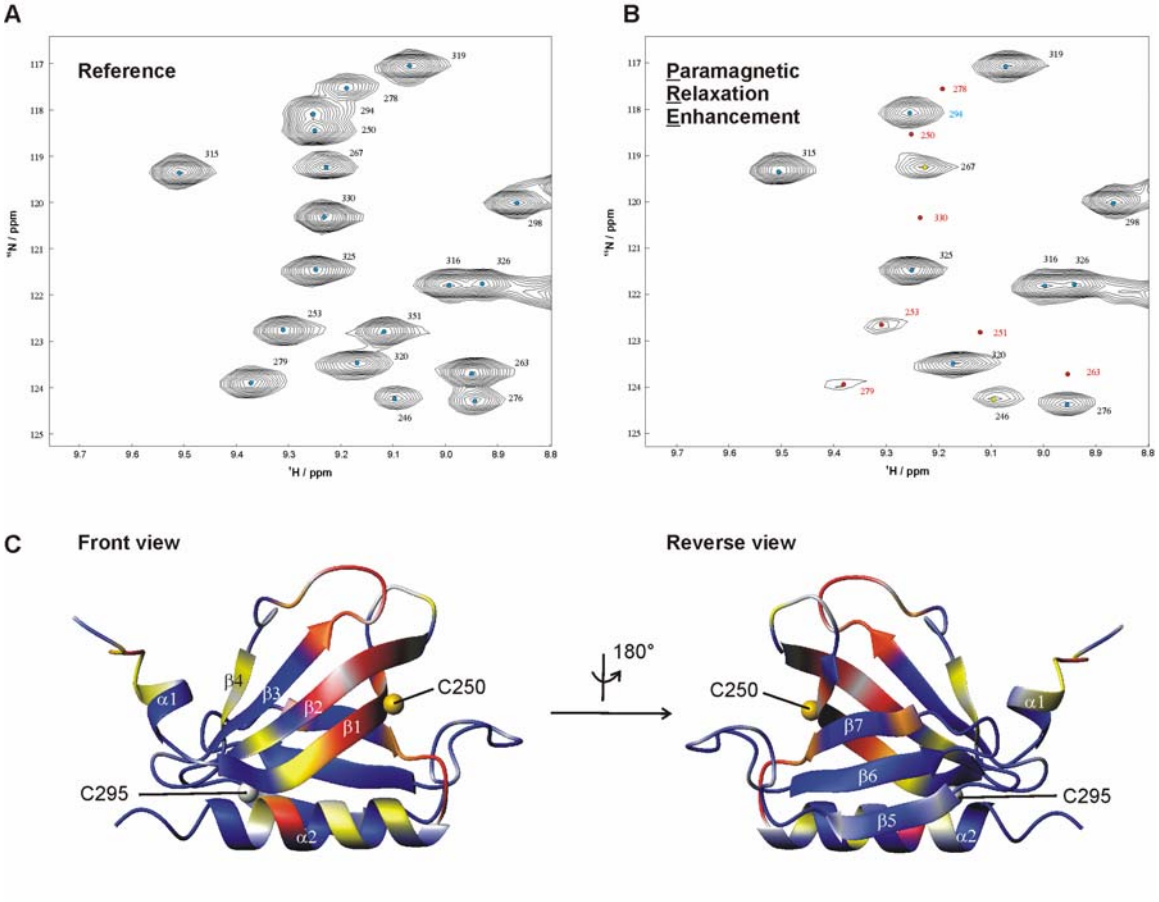

Figure 4.1: Effects of paramagnetic relaxation enhancement (PRE) caused by spin-labelling. A and B: *β-sheet area of a $^{1}H,^{15}N$-HSQC spectrum of C-PH. Panel **A** shows the reference spectrum, whereas panel **B** shows the spectrum with paramagnetic relaxation enhancement caused by the spin-label. Peak positions and assignments are indicated. In panel **B**, the dots representing the peak position are coloured according to whether the peak is unaffected (blue), mildly affected (yellow) or strongly affected (red). **C:** the same colour coding as in **B** is used to map the affected area onto the structure of C-PH. A front and a reverse view are presented.*

The reduction in signal intensity of a amide backbone proton (H^N) in a $^1H,^{15}N$-HSQC caused by a spin-label is correlated to the distance in space between H^N and the paramagnetic electron of the spin-label. The peak intensity ratio between PRE (with spin-label) and reference (with inactive spin-label) spectra, I_{PRE}/I_0 , is calculated for every H^N and plotted against the distance to S^γ of Cys^{250} (Fig 4.2). The distribution of distances against I_{PRE}/I_0 shows a clear correlation, although with substantial scatter.

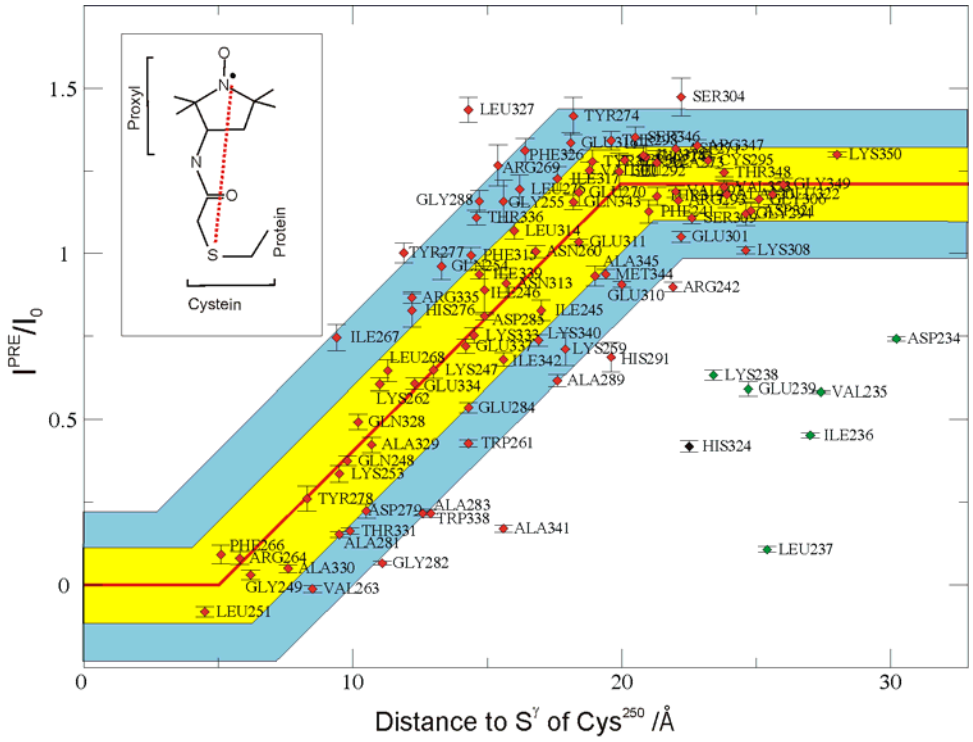

Figure 4.2: Relation between distance to the spin-label and paramagnetic relaxation enhancement (PRE). *All analysable residues of C-PH are plotted (diamonds) according to their PRE and distance to S^γ of Cys^{250}. PRE of a residue is calculated as peak intensity ratio between spectra with and without a paramagnetic spin-label (I_{PRE}/I_0). Error bars are calculated from the noise in the measurements. The red trendline is constructed manually to calibrate distance against PRE. The yellow area includes all points within ±2.5Å and ±0.1 I_{PRE}/I_0 of the red line. The blue area has twice the width (±5Å, ±0.2 I_{PRE}/I_0) of the yellow area.* ***Inset****: the chemical structure of the spin label and its attachment to cysteine sidechains is illustrated. The red dotted line is the distance between the cysteine S^γ to the unpaired electron (black dot) of the spin label.*

Surprisingly, the plateau at large distances from the spin-label is not at an I_{PRE}/I_0 of 1.0, but notably above (between 1.2 and 1.3). This may be caused by the effect of PRE on the longitudinal relaxation rate (R1), which accelerates the return to equilibrium magnetisation in the delay between the scans. Mainly residues from secondary structure

elements display the largest I_{PRE}/I_0 ratios (these residues also have the lowest R1 in the non-spin-labelled sample).

There are several sources of systematic errors that contribute to the deviation from a more ideal distribution, e.g. uncertainty in the position of the radical electron of the spin label (which is not located at the position of the cysteine's Sγ, inset Fig. 4.2) and ensemble and time averaging of the paramagnetic transverse relaxation enhancement (ΔR2) over r^{-6} (which is expected to strongly bias ΔR2 in favour of short distances, cf. section 1.2.4). For the purpose of obtaining distance restraints from I_{PRE}/I_0, which is not an exact approach, only the approximate magnitude of the errors has to be known so that appropriate upper and lower restraint boundaries can be chosen. In Fig. 4.2, a red trendline is added manually to correlate distance with PRE: I_{PRE}/I_0 is zero for 0-5Å, and beyond a distance of 20Å, I_{PRE}/I_0 is the trimmed mean value (mean value after discarding the upper 20% of all outliers) of all points in this area. Between 5 and 20Å, a straight line connects the two levels. The yellow area in Fig. 4.2C includes all points within ±2.5Å of the trendline and within ±0.1 I_{PRE}/I_0 in the flat parts of the curve. The blue area is constructed in the same way, only with twice the width in both dimensions (±5Å distance, ±0.2 I_{PRE}/I_0). About 50% of all residues are within the yellow area and more than 85% are in the blue area. Thus, it should be possible to use a trendline as shown in Fig. 4.2 as an empirical calibration curve for distance *versus* paramagnetic relaxation enhancement if error bounds of ±5Å are applied.

The outliers of the plot are mostly residues from the N-terminus of C-PH (aa 234-239, blue diamonds in Fig 4.2). As the N-terminus is flexible, its position is not well represented in any single structure. It samples multiple conformations and by consequence gives averaged, non-interpretable PRE values. Since averaging is dependent on the distance by r^{-6} (cf. section 1.2.4), I_{PRE}/I_0 of flexible residues is expected to be strongly biased towards low values, as observed for residues 234-239. This shows that flexible residues will have to be treated with care when analysing PRE. A cut-off criterion to eliminate flexible residues could be applied, e.g. heteronuclear NOE > 0.6.

4.1.2 Determination of the number of accessible cysteines in pleckstrin

The PRE data for C-PH are readily analysable and, in principle, could be used immediately as input for a structure calculation. All other pleckstrin constructs suffer from too strong PRE, most likely originating from more than one spin-label per molecule. The reaction with Ellmann's reagent (5-5'-Dithio-bis(2-nitrobenzoic acid), a cysteine specific reagent) is employed to estimate the number of accessible cysteines of pleckstrin. The formation of the thiolate product is measured at 412nm, and the number of accessible cysteines per molecule is calculated based on the extinction coefficient of the product. The results are shown in Table 4.1. Cys^{59} and Cys^{103} of N-PH both react with DTNB and must therefore be accessible. The interpretation of the data concerning the two or three cysteines of the DEP domain (Cys^{155}, Cys^{160} ±Cys^{226}) is not straightforward: one or two may be accessible; it seems that again the peculiarities of the DEP domain hamper a clear interpretation (cf. chapter 3). For the C-PH domain, it is already established by the analysis of PRE data that Cys^{250} is accessible and Cys^{295} is not.

Table 4.1: Estimated number of accessible cysteine in pleckstrin

Protein		**Number of cysteines / molecule**		
	aa	*DTNB*	*Total*	*Accessible*
N-PH	1-103	2.7	2	2/2
N-PH C102A	1-103	0.8	1	1/1
N-PH_DEP	1-243	2.9	5	3/5
DEP	122-221	1.5	2	1/2 or 2/2
DEP_C-PH	122-350	2.6	3	2/4 or 3/4
C-PH	234-350	N.D. [1]	2	1/2 (spin label)

[1] N.D. = not determined

4.1.3 Elimination of native cysteines

According to the numbers presented in Table 4.1, all constructs of pleckstrin except C-PH (and N-PH C102A) contain more than one accessible cysteine. Thus, more than one spin-label can be attached to all other constructs, which leads to the undesired excessive

PRE seen for N-PH, DEP, N-PH_DEP and DEP_C-PH constructs. To avoid this, it is necessary to generate constructs with only a single accessible cysteine by mutagenesis.
In total there are 7 cysteines in pleckstrin out of which 4-6 may be accessible. It is recommended to mutate cysteine to either serine or alanine, according to whether it is in a more polar or apolar environment, respectively (Bordo and Argos, 1991). Following these guidelines, all cysteines except Cys^{295} are mutated one at a time. The resulting constructs with a single mutation (C59A, C102A, C155A, C160A, C226S, C250A) are screened for solubility and correct folding (HSQC spectra). All mutants are soluble and have a HSQC spectrum that is very similar to the wildtype. In the next step, cysteine mutations are accumulated until double-domain constructs are available with only one native accessible cysteine left for spin-labelling. All single-native-cysteine mutants yield folded protein. For clarity, single-native-cysteine mutants are named after the position that can be spin-labelled. For example, "DEP_C-PH C226" is the construct in which all accessible cysteines except Cys^{226} are mutated (C155A/C160A/C250A). In the following, only this shorthand notation will be used.

Table 4.2: Single native cysteine spin-labelled samples.

Protein	*Spin-label*	*PRE observed*	*Observation*
N-PH_DEP	C59	Yes	Good quality spectra. PRE consistent.
	C102	Yes	PRE in consistent.
	C155	No	-
	C160	Yes	Good quality spectra.
	C226	Yes	Spectrum changes strongly. PRE and reference spectrum can hardly be compared.
DEP_C-PH	C155	No	-
	C160	Yes	Good quality spectra. PRE consistent.
	C226	Yes	Spectrum changes strongly. PRE and reference spectrum can hardly be compared.
	C250	Yes	Good quality spectra. PRE consistent.

All single-native-cysteine mutants are spin-labelled and PRE data is collected and analysed. The results of the analysis are briefly summarised in Table 4.2. The following conclusions can be made: a) Cys^{155} in DEP is not accessible, while Cys^{160} and Cys^{226} are;

b) spin-labels in inter-domain linker regions give rise to very broad and unspecific PRE, which is probably caused by the conformational freedom of the linker; c) for all the others, PRE correlates quite well with distances within the domain the spin-label is attached to; d) no cross-domain PRE (i.e. reduction in peak intensity for residues that are not in the domain the spin-label is attached to) is observed for any of the non-linker spin-labelled proteins. Therefore, the structural information to be gained from these data cannot be used to help define domain-domain orientations in pleckstrin.

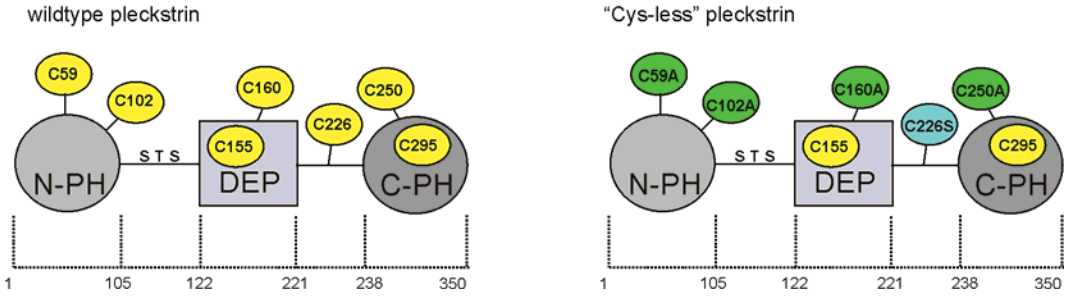

Figure 4.3: Summary of native cysteine mutagenesis. *All cysteines in wildtype pleckstrin are illustrated as yellow spheres. Inaccessible cysteines are positioned inside the shape representing the domain. In the "Cys-less" variant of pleckstrin, all accessible cysteines are mutated to alanine (green) or serine (blue).*

4.1.4 Site-directed spin-labelling

The aim of making spin-labelled pleckstrin double-domain constructs is to obtain structural restraints between two domains. Unfortunately, none of the single-native-cysteine mutants supply the desired cross-domain PRE information. Therefore a new set of double-domain constructs which contain a selectively introduced cysteine in the "Cys-less" pleckstrin (Fig. 4.3) background is produced for site directed spin-labelling. These mutants, hereinafter referred to as "SDSL" mutants – for Site-Directed Spin-Labelling, are again named after the position that is spin-labelled (e.g. N-PH_DEP "C75" is C59A/C102A/C160A/C226S/K75C).

The residue that is replaced with cysteine by site-directed mutagenesis is usually charged or polar and fully solvent accessible so as to secure good reactivity with the spin-labelling reagent and minimal perturbation of the protein's structure. Cysteines for site-directed spin-labelling are predominantly introduced at "strategic" positions, i.e. at positions that

are likely to cause cross-domain PRE or that are biologically relevant. In the case of N-PH_DEP, chemical shift comparison with the single N-PH and DEP domains provides some initial hints where the two domains might interact (B. Simon, unpublished results). Accordingly, cysteine residues are preferentially introduced close to the putative contact area. For DEP_C-PH, no such putative interaction surface is defined. Hence, cysteine mutations are placed more randomly, but with a higher density close to the functional or putatively functional areas, i.e. the phosphoinositide binding site of C-PH (cf. chapter 2) and the β1–β2 "finger" of DEP, where the only known loss-of-function mutation of a DEP domain is located (Boutros et al., 1998). This strategy is aimed at utilizing all prior knowledge that is available, be it structural, biochemical or mutational data.

SDSL mutants are produced in batches of ten at a time. Procedures for cloning, solubility screening, expression, purification and spin-labelling are streamlined and optimised to allow processing of several samples in parallel. Mutagenesis and screening procedures were developed by G. Stier and will be described in detail elsewhere. Description of the optimised protocols for protein purification, spin-labelling and sample preparation can be found in chapter 6 ("Materials and Methods"). The measurement time required for one spin-labelled double-domain construct is between 3 and 12 hours per HSQC on a Bruker 600MHz spectrometer with a cryoprobe; since every sample has to be measured a second time after reducing the spin-label with ascorbic acid (reference spectrum) the total spectrometer time for 10 strategic surface mutants amounts to about 5 days.

The assignments of the wildtype protein are transferred to the "Cys-less" variant by means of a 3D-^{15}N-NOESY. The HSQC spectra of SDSL mutants only show minor differences to the "Cys-less" variant because there is only one additional mutation (and not 4 or 5 compared to the wildtype).

4.1.5 Cross-domain PRE by spin-labels on N-PH_DEP

40 SDSL mutants of N-PH_DEP are cloned, 30 of them are expressed and 21 are spin-labelled and analysed (see Appendix B for a full list). About 10% of all constructs fail at the level of cloning, 20% do not express any protein and another 20% yield too low amounts of protein. The success rate is about 50%, which is very satisfactory. 21 different SDSL mutants are more than enough to scan the surface of a protein like

pleckstrin N-PH_DEP because of the large radius (10-25Å) of paramagnetic relaxation enhancement caused by the spin-label.

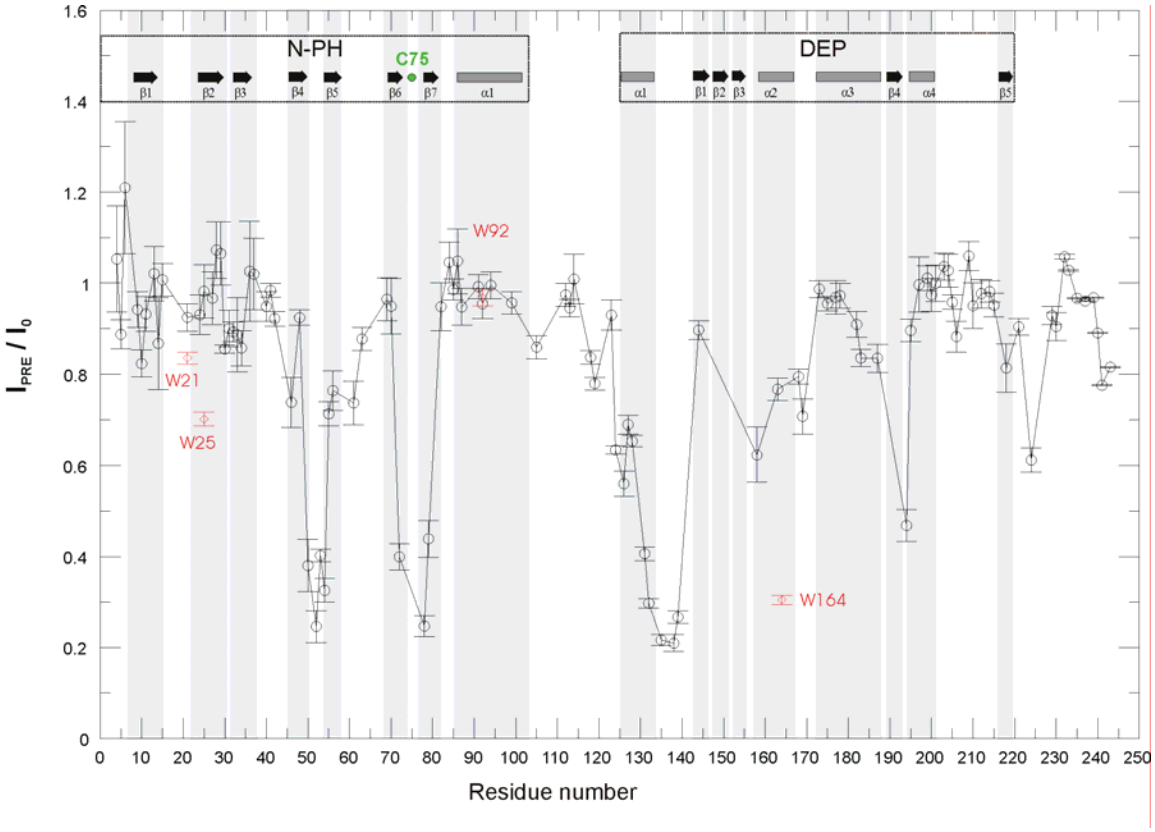

Figure 4.4: PRE of the spin-label attached to C75 of N-PH_DEP. *I_{PRE}/I_0 of backbone amide groups is plotted against the sequence position in N-PH_DEP. Tryptophane sidechain $H^{\varepsilon 1}/N^{\varepsilon 1}$ are coloured in red. The secondary structure elements of N-PH and DEP are shaded in grey.*

Analysis of the 21 spin-labelled mutants reveals that N-PH and DEP are closely associated in the N-PH_DEP double-domain construct. 5 different SDSL mutants cause strong cross-domain paramagnetic relaxation enhancement. One such construct, N-PH_DEP C75, is shown in Fig. 4.4. Clearly, there are regions on both N-PH and DEP that have decreased I_{PRE}/I_0 . On N-PH, low I_{PRE}/I_0 values are seen for the β4–β5 loop and for strands β4, β5, β6 and β7. On DEP, it is mostly helix α1, the loop after α1 and the sidechain Hε1/Nε1 of Trp^{164} that have very low I_{PRE}/I_0. Also parts of helix α2 of DEP have a significantly lower I_{PRE}/I_0 than the rest of the molecule. When the I_{PRE}/I_0 ratio is colour-coded and mapped onto the N-PH and DEP structures (Fig. 4.5) it can be seen that the residues with a reduced peak intensity ratio cluster structurally on both domains.

Therefore, the data are self-consistent and demonstrate that N-PH and DEP interact with each other. Clearly, PRE on DEP is localised on the C-terminal part of α1 and the surrounding area.

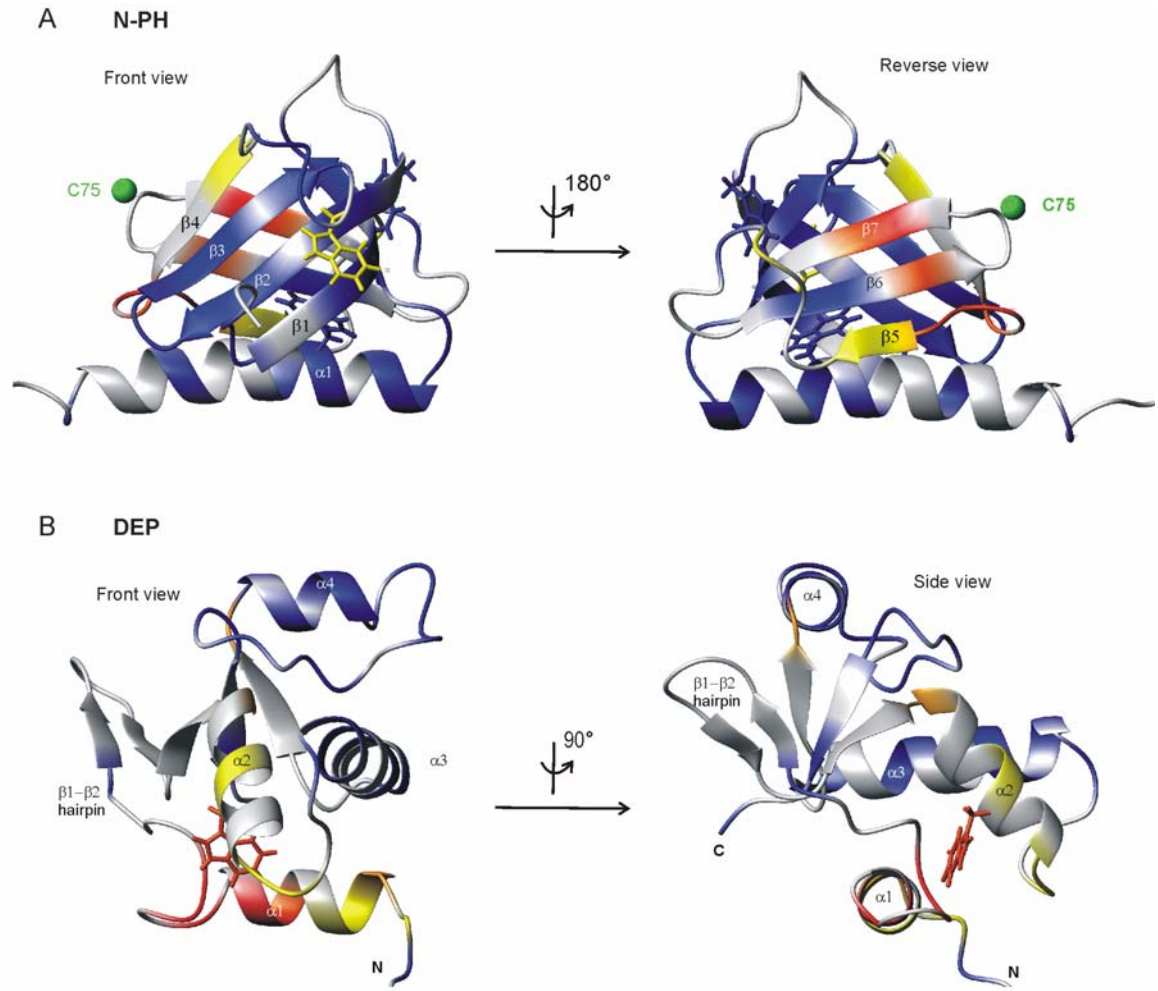

Figure 4.5: PRE of the spin-label attached to C75 of N-PH_DEP mapped onto the structures of N-PH and DEP. *Decrease in peak intensity caused by the spin-label is colour coded: blue (no effect), yellow (slight decrease in I_{PRE}/I_0), orange (strong decrease in I_{PRE}/I_0) and red (very strong decrease in I_{PRE}/I_0). Grey: no data. The ribbon representation of N-PH is shown in panel* ***A****, as front and reverse view. In panel* ***B****, the DEP structure is shown in front and side view. Tryptophane sidechains are also shown and coloured.*

The relationship between known distance and I_{PRE}/I_0 ratio (only for the N-PH domain, Fig. 4.6) is very good. There are much fewer outliers and the I_{PRE}/I_0 ratio at distances larger than 20Å is almost exactly 1.0. A trendline for the PRE-distance relationship is drawn in red in Fig. 4.6. More than 80 % of all residues are within the yellow zone

(±2.5Å, ±0.1 I_{PRE}/I_0 from the red line). Thus, N-PH_DEP spin-labelled on Cys^{75} produces excellent results: PRE is consistent on both domains (Fig. 4.5) and the relationship between distance and PRE is very good (Fig. 4.6).

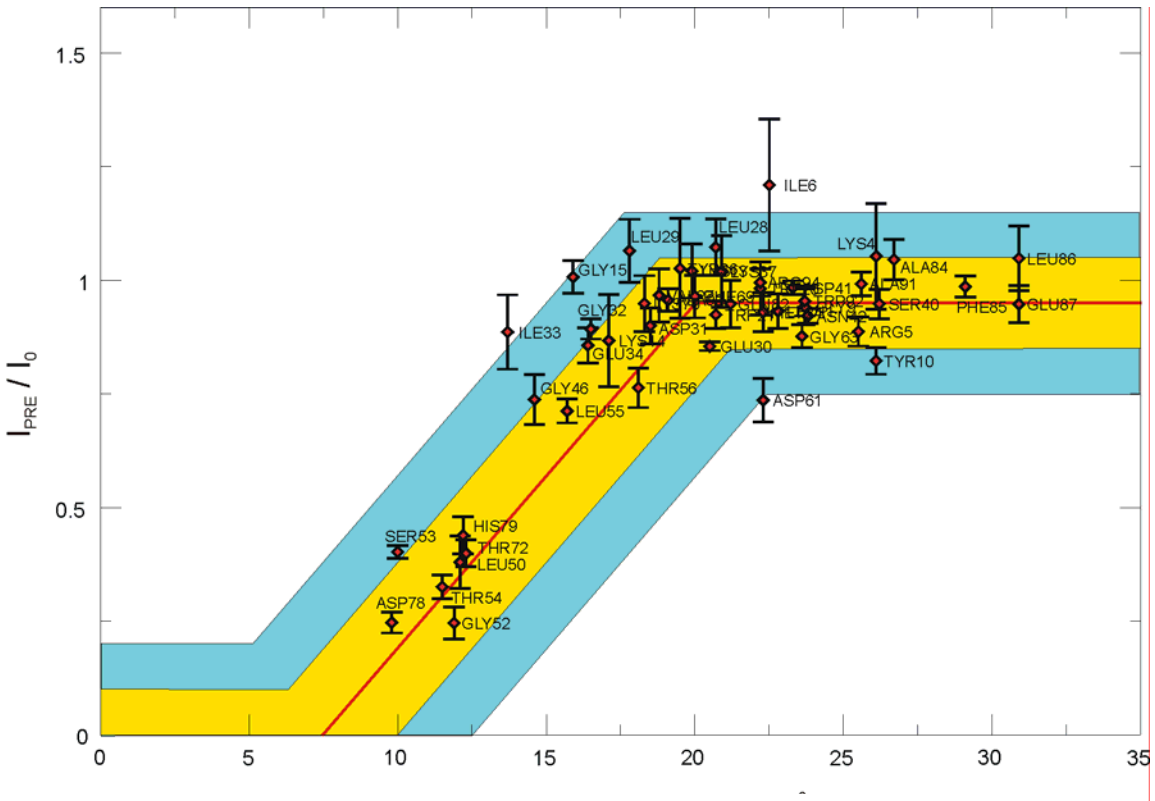

Figure 4.6: Relation between distance to the spin-label and paramagnetic relaxation enhancement (PRE) of N-PH_DEP C75 (N-PN only). *All residues of N-PH are plotted (diamonds) according to their PRE and distance to S^{γ} of Cys^{75}. The red trendline is constructed manually to calibrate distance against I_{PRE}/I_0 .The yellow area includes all points within ±2.5Å and ±0.1 I_{PRE}/I_0 of the red line. The blue area has twice the width (±5Å, ±0.2 I_{PRE}/I_0)* of the yellow area.

4.1.6 Spin-label scanning of DEP_C-PH

45 cysteine mutants for site-directed spin-labelling of DEP_C-PH are cloned, 21 are expressed, spin-labelled and measured and 19 of them produce interpretable PRE data. The success rate (19 out of 45 = 42%) is somewhat lower than for N-PH_DEP. All 19 positions used for site-directed spin-labelling are highlighted on the structures of DEP and C-PH in Fig. 4.7 and are listed in Appendix B.

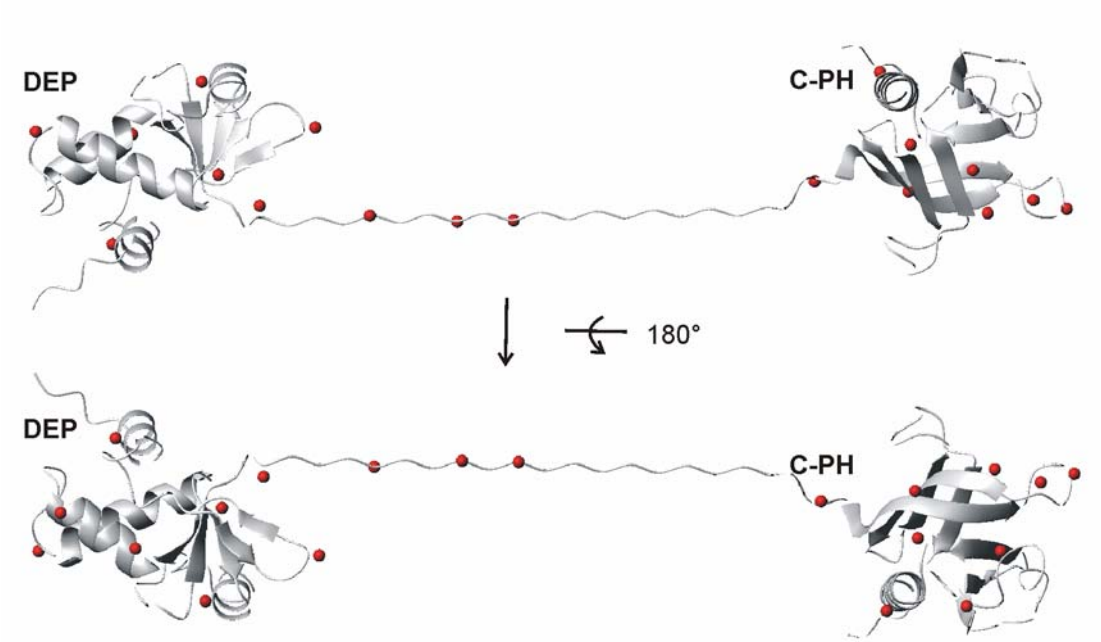

Fig. 4.7: Surface scanning of DEP_C-PH. *All positions to which spin-labels are attached for surface scanning of DEP_C-PH are shown as red spheres on a model where the linker is fully extended.*

In general, DEP_C-PH SDSL mutants are more difficult to handle and to analyse than N-PH_DEP SDSL mutants. Several SDSL DEP_C-PH mutants tend to aggregate in standard NMR buffer (100mM NaCl) and have to be measured in 200-500mM NaCl. The aggregation tendencies also appear to influence the PRE data. A considerable number of residues experiences large PRE irrespective of the position of the spin-label in the protein, i.e. reduced peak intensity of the same signal is seen for many different spin-labelled SDSL mutants without any correlation to the position of the spin-label. This unspecific PRE could be caused by transient association of two spin-labelled molecules, where the spin-label from one molecule induces PRE on the other. Battiste and Wagner also report an incidence of unspecific PRE involving a tryptophane sidechain (Battiste and Wagner, 2000). Due to the hydrophobic properties of the spin-label's proxyl group, the authors suggest an intermolecular hydrophobic mechanism directly involving the spin-label. In the case of DEP_C-PH, aggregation tendencies are observed already in the absence of the spin-label (Chapter 3). Mutagenesis and spin-labelling involving the replacement of a hydrophilic sidechain with the cysteinyl proxyl derivative appear to enhance these tendencies in some cases. Thus, both intrinsic and extrinsic (mutation plus

spin-label) factors may induce intermolecular association, resulting in intermolecular on top of the desired intramolecular paramagnetic relaxation enhancement.

Due to the presence of unspecific – most likely intermolecular – PRE, the interpretation of DEP_C-PH data is quite difficult. Careful analysis of all available data does not reveal any clear cross-domain PRE as for N-PH_DEP. Yet, DEP_C-PH is expected to yield the same strong cross-domain PRE as N-PH_DEP because the chemical shift changes are also much smaller. Some weaker cross-domain PRE is observed for some spin-labelled mutants of DEP_C-PH, but the data are not consistent enough to confidently exclude the possibility of an intermolecular effect. Currently, the data do not support a specific association of DEP and C-PH in DEP_C-PH. Thus, the two domains may not interact *in cis*, in agreement with relaxation and chemical shift analysis (Chapter 3). The effects leading to the large extent of unspecific PRE in DEP_C-PH are still under investigation.

4.1.7 Preliminary structure of N-PH_DEP

PRE data from six N-PH_DEP SDSL constructs are used in a structure calculation (C1, C4, C74, C75, C132 and C199). Each spin-label dataset is normalised such that the plateau peak intensity ratio level is 1.0. Then, four classes of residues are defined according to their peak intensity ratio:

(i) $I_{PRE}/I_0 \geq 0.85$

(ii) $0.85 > I_{PRE}/I_0 \geq 0.60$

(iii) $0.60 > I_{PRE}/I_0 \geq 0.45$

(iv) $0.45 > I_{PRE}/I_0$,

Distance restraints between the H^N of each residue and the Sγ of the SDSL mutant's cysteine are defined for all classes based on the distance correlations shown in Fig. 4.2 and 4.6. Class (ii)-(iv) have upper limits between 10 and 20Å (Table 4.3). For unaffected residues (class (i)) only a lower distance limit is defined (15Å), resulting in a "repulsive restraint". The spin-label restraints are summarised in Table 4.3:

Table 4.3: Spin-label restraints for the structure calculation of N-PH_DEP:

Class	Distance	N_{total} [1]	$N_{X\text{-}domain}$ [2]	
i	>15Å	492	265	"Unaffected"
ii	8-20Å	47	23	"Mildly affected"
iii	5-15Å	34	11	"Strongly affected"
iv	1.8-10Å	108	32	"Bleached"

[1] Total number of spin-label restraints

[2] Number of cross-domain restraints

Since there is no precedent for generating distance restraints directly from intensity ratios, the optimal distance boundaries have to be determined by trial and error. The values presented in Table 4.3 do not represent optimal, but sufficiently usable boundaries. The upper distance boundaries are set "conservatively" i.e. quite large, with the intention to avoid over-refinement with respect to the PRE restraint. Since the number of restraints is large, the domains are "pulled together" with a sufficiently strong force.

A total of 681 spin-label restraints are used in the structure calculation, most of which are repulsive (class i). The number of "attractive" restraints (class ii-iv) which define a distance between the N-PH and DEP is 66. In addition, 149 H^N-N RDC restraints and 376 TALOS restraints are included in the structure calculation. To ensure that the domain structures of N-PH and DEP are kept intact, non-crystallographic symmetry (NCS) restraints are employed. The CNS protocols of the structure calculation are developed in collaboration with B. Simon and will be described in detail elsewhere.

The best 20 out of 100 calculated structures all have a very similar appearance. The orientation between N-PH and DEP domain is well defined, and the secondary structure elements of both domains superimpose satisfactorily (data not shown). A ribbon representation of the preliminary structure of N-PH_DEP is shown in Fig. 4.8. The spin-labels used in the calculation are shown as large spheres and labelled with their residue number. The class (iv) distance restraints (residues that are most strongly affected by the spin-label) are illustrated as straight lines between the spin-label and the H^N of the affected residue.

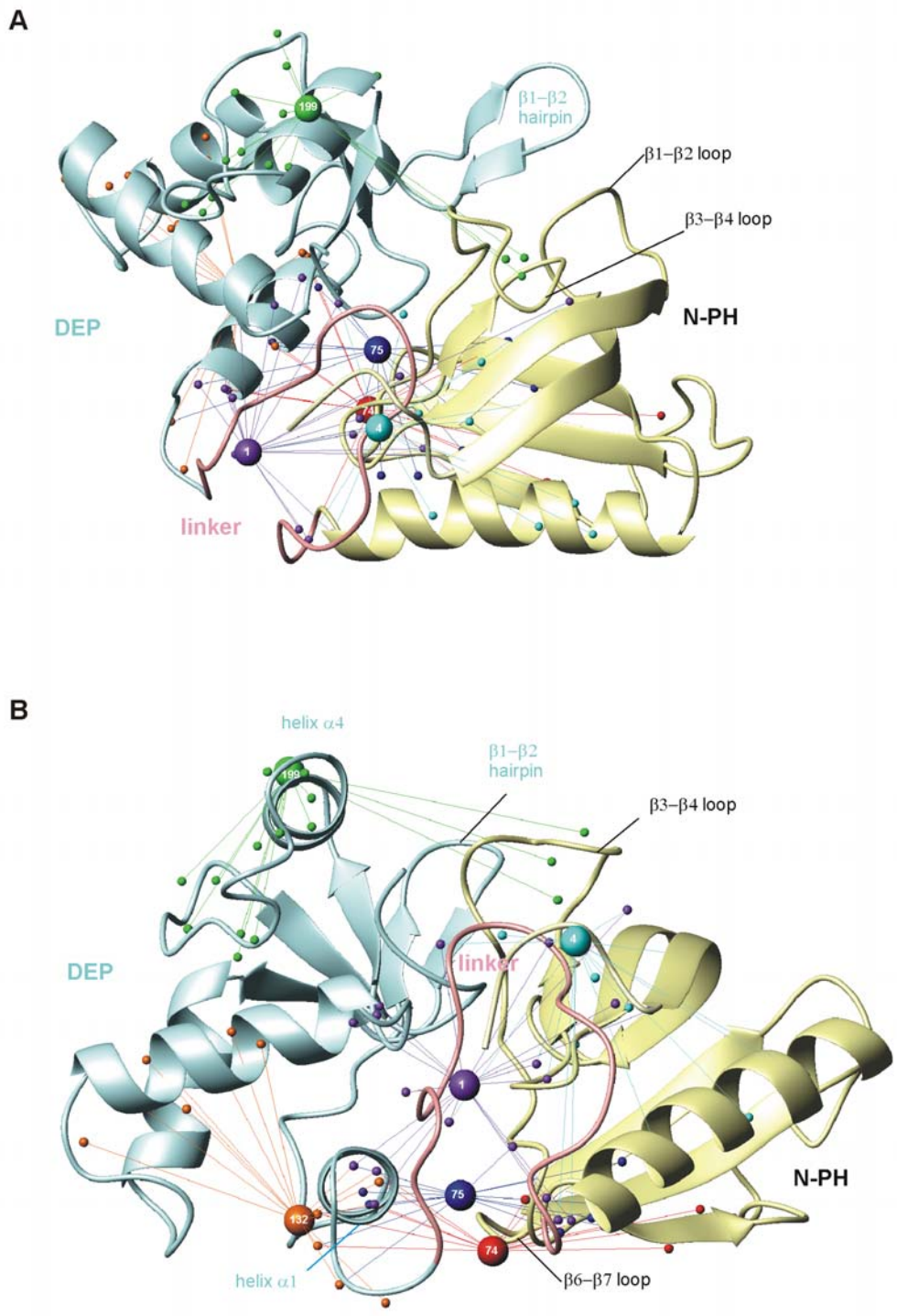

Fig. 4.8: Preliminary structure of N-PH_DEP*. The domain orientation is defined by restraints from six spin-labels which are shown as large spheres and labelled with their residue number. The H^N atoms that show the strongest PRE are represented as small spheres. The distance restraints for these atoms are represented as straight lines between the spin-label and the H^N. The ribbon representation of the structure is coloured in the following way: N-PH gold, N-PH_DEP linker pink and DEP light blue.*

The inter-domain linker (pink in Fig. 4.8) is not sufficiently restrained in the calculation, and the conformation shown is not representative. Nonetheless, it is interesting to note that it is located close to the interface between DEP and N-PH in most structures. The C-terminus of N-PH and the N-terminus of DEP are quite close in space, therefore the inter-domain linker adopts a conformation of an extended loop. The β1−β2 hairpin of DEP is near the phosphoinositide binding site of N-PH (β1−β2 loop, β3−β4 loop, Fig. 4.8A). This finding still has to be confirmed because there are no spin-label restraints for this region (the β1−β2 loop of N-PH is not assigned and the β1−β2 hairpin of DEP only has one "marker" residue with an isolated peak in the HSQC).

The interface between the domains is shown from a different angle in Fig. 4.8B. Spin-labels on the β5−β6 loop of N-PH (C74, C75) define the area where DEP and N-PH come closest together. Helix α1 of the DEP domain is at the centre of the interface between the domains. Also, the N-terminus of N-PH (spin-labels at C1 and C4) appears to be sandwiched between N-PH and DEP. Thus, the interface between the domains is quite extensive and may include parts of the inter-domain linker and the N-terminus of N-PH.

4.2 Discussion

The earliest attempts to obtain structural information from spin-labelled proteins date back to the 1970s (Kosen, 1989). Since then, the advent of isotopic labelling techniques and heteronuclear NMR has facilitated data analysis extremely, allowing the precise measurement of paramagnetically enhanced relaxation rates. Spin-labels are often employed to obtain long-range structural information that cannot be obtained otherwise by NMR. Examples include evidence of residual structure in unfolded proteins (Gillespie and Shortle, 1997; Lindorff-Larsen et al., 2004; Teilum et al., 2002), investigations of global structure in medium sized proteins (Battiste and Wagner, 2000; Gaponenko et al., 2000) and mapping of inter-molecular binding sites of medium size/large molecules (Liepinsh et al., 2001; Park et al., 2003).

All previous spin-label studies have employed intrinsically cysteine-free or single-cysteine proteins because only mono-derivatised proteins yield accurate data. The example of the pleckstrin protein described in this chapter shows that the limitation posed by the number of native cysteines no longer applies: even a relatively large number of native cysteines can be replaced by site-directed mutagenesis. This is a significant step for expanding the utility of spin-labels in the NMR field which is made possible by applying efficient molecular biology methods. A further advancement is the application of spin-labels to a multi-domain protein which has not been attempted before. As demonstrated for N-PH_DEP, the interpretation of PRE data from a double-domain construct is relatively straight-forward if the individual domain structures are known: PRE can be simply calibrated against known distances, in the same way as NOEs. This strategy simplifies the analysis of PRE data significantly because only peak intensities of 2D spectra have to be measured. The approach allows rapid extraction of structural restraints and their implementation in a structure calculation.

The derivation of PRE restraints from peak intensity ratios presented in this chapter is intended to allow the implementation of PRE restraints in a similar manner as NOE restraints. The NOE is by far the most widely used structural restraint that can be derived from NMR spectra. Despite several sources of systematic errors, NOE restraints are very robust. Since both NOE and PRE restraints depend on distance by r^{-6} it should be possible to derive PRE distance restraints the same way as for the NOE, i.e. by calibration against known distances plus generous error bounds. It is yet too early to assess or predict whether NOE-like calibration of PRE restraints is as reliable as NOE calibration, but the preliminary structure determination of pleckstrin N-PH_DEP is very promising.

The three-domain protein pleckstrin with its prototypic PH domains is an example of an "old" protein (first described in 1979, (Lyons and Atherton, 1979)) that has resisted all attempts to be crystallised. All domain structures have been solved individually. The site directed spin-labelling approach is applied to the two double-domain constructs of pleckstrin in order to obtain structural restraints between the domains and to define the position of the linker relative to the domains. Spin-labels are preferentially positioned in areas that are expected to lie close to the domain-domain interface. N-PH_DEP and DEP_C-PH give contrasting results: a large number of cross-domain PRE restraints can

be extracted for the former, but none for the latter, although a similar number of differently spin-labelled constructs are analysed. Clearly, N-PH-DEP and DEP_C-PH represent two different classes of two-domain proteins: one where the domains are very close and the other where the domains are relatively far apart. In fact, the N-PH_DEP double-domain protein is an example where the two domains adopt a "closed" conformation forming a substantial domain-domain interface.

The results of the spin-labels studies of N-PH_DEP and DEP_C-PH can be combined into a working model for the pleckstrin molecule. The C-terminal PH domain of DEP_C-PH appears to be independent of the DEP domain and could therefore function as an autonomous response element to PtdIns(3,4)P_2. This conclusion is in agreement the work on DEP_C-PH presented in the previous chapter, but also with published *in vivo* data (Ma and Abrams, 1999; Ma et al., 1997; Roll et al., 2000) which can be interpreted in the sense that C-PH has no effect when PtdIns(3,4)P_2 production is not stimulated. Conversely, N-PH appears to be the essential factor for morphological changes (membrane ruffling etc.) in all transfection experiments that requires PKC phosphorylation for its activity (Ma and Abrams, 1999; Ma et al., 1997; Roll et al., 2000); since N-PH can induce these transformations on its own, but not in a unphosphorylatable variant of PH1_DEP, the DEP domain may have an inhibitory role. Clearly, the closed conformation of the N-PH_DEP structure (Fig. 4.8) supports the "auto-inhibition" hypothesis. The basic residues at the "open side" of N-PH (β1−β2, β3−β4 loop) which have been shown to be required for most effects of N-PH are close to the DEP domain, but do not appear to be completely blocked (Fig. 4.8). Nonetheless, the mere proximity of the DEP domain may be sufficient to restrict access to the functional loops of N-PH. The auto-inhibition model of unphosphorylated pleckstrin is shown in the left panel of Fig. 4.9. Phosphorylation of the linker between N-PH and DEP could free N-PH from its inhibition by DEP, probably by strongly changing the electrostatic environment near the N-PH-DEP interface. The activated pleckstrin molecule would then feature an "open" N-PH domain that drives cytoskeletal rearrangements and the PtdIns(3,4)P_2 sensor C-PH (Fig. 4.9, right panel). The role of the DEP domain after PKC phosphorylation remains to be determined.

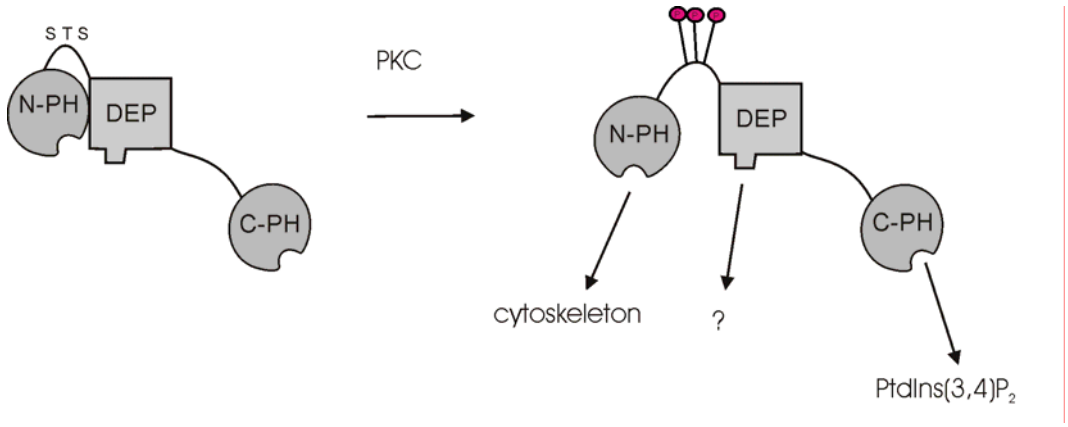

Fig. 4.9: Auto-inhibition model of pleckstrin. *The results of the spin-label studies are combined with results from previous chapters. PKC phosphorylation of pleckstrin opens the conformation of N-PH_DEP which causes reorganisation of the cytoskeleton by N-PH. C-PH is always accessible, but only comes into action after PKC phosphorylation when its ligand PtdIns(3,4)P_2 is produced.*

Only one structure is currently available of which the domain composition resembles pleckstrin: the Epac2 regulatory domain (Rehmann et al., 2003). It consists of a DEP domain sitting on top of two cyclic adenosine monophosphate (cAMP) binding domains which are in close contact. Evidently, a different situation is encountered as for pleckstrin: the two sensor domains interact with each other, but none is closely associated with the central DEP domain. The mechanism of activation is therefore also quite different: ligand binding disrupts the association of the sensor domains and results in transmission of the regulatory stimulus (Rehmann et al., 2003). Due to the striking similarity of the domain arrangement, one is forced to consider a similar possibility for pleckstrin. Although a strong coupling between the PH domains of pleckstrin is unlikely since no cooperativity or inhibition is observed in the full-length molecule by protein lipid overlay assays (Chapter 3), a more subtle N-PH-C-PH interaction or regulation can currently not be ruled out. This intriguing possibility awaits experimental validation in the full-length molecule.

In the current chapter, the feasibility of a de-novo structure determination of a two-domain protein by site-directed spin-labelling is demonstrated. All native accessible cysteines of pleckstrin are replaced and a large number of new cysteines are introduced for spin-labelling and PRE measurements. In the N-PH_DEP double-domain construct, the domains have a large interface that gives rise to cross-domain PRE by an

appropriately positioned spin-label. For DEP_C-PH, the domains appear to be far apart. A preliminary structure calculation of N-PH_DEP applying PRE and RDC restraints provides a first model of the interaction between N-PH, DEP and the inter-domain linker. This structure of N-PH_DEP shows a close association of DEP and N-PH and proposes an auto-inhibition model of unphosphorylated pleckstrin. It also constitutes an important milestone on the road to the structural characterisation of full-length pleckstrin.

4.3 References

Battiste, J. L., and Wagner, G. (2000). Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. Biochemistry *39*, 5355-5365.

Bordo, D., and Argos, P. (1991). Suggestions for "safe" residue substitutions in site-directed mutagenesis. J Mol Biol *217*, 721-729.

Boutros, M., Paricio, N., Strutt, D. I., and Mlodzik, M. (1998). Dishevelled activates JNK and discriminates between JNK pathways in planar polarity and wingless signaling. Cell *94*, 109-118.

Gaponenko, V., Howarth, J. W., Columbus, L., Gasmi-Seabrook, G., Yuan, J., Hubbell, W. L., and Rosevear, P. R. (2000). Protein global fold determination using site-directed spin and isotope labeling. Protein Sci *9*, 302-309.

Gillespie, J. R., and Shortle, D. (1997). Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. J Mol Biol *268*, 170-184.

Hubbell, W. L., Gross, A., Langen, R., and Lietzow, M. A. (1998). Recent advances in site-directed spin labeling of proteins. Curr Opin Struct Biol *8*, 649-656.

Kosen, P. A. (1989). Spin labeling of proteins. Methods Enzymol *177*, 86-121.

Liepinsh, E., Baryshev, M., Sharipo, A., Ingelman-Sundberg, M., Otting, G., and Mkrtchian, S. (2001). Thioredoxin fold as homodimerization module in the putative chaperone ERp29: NMR structures of the domains and experimental model of the 51 kDa dimer. Structure (Camb) *9*, 457-471.

Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C. M., Poulsen, F. M., and Vendruscolo, M. (2004). Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme a binding protein. J Am Chem Soc *126*, 3291-3299.

Lyons, R. M., and Atherton, R. M. (1979). Characterization of a platelet protein phosphorylated during the thrombin-induced release reaction. Biochemistry *18*, 544-552.

Ma, A. D., and Abrams, C. S. (1999). Pleckstrin induces cytoskeletal reorganization via a Rac-dependent pathway. J Biol Chem *274*, 28730-28735.

Ma, A. D., Brass, L. F., and Abrams, C. S. (1997). Pleckstrin associates with plasma membranes and induces the formation of membrane projections: requirements for phosphorylation and the NH2-terminal PH domain. J Cell Biol *136*, 1071-1079.

Park, S., Caffrey, M. S., Johnson, M. E., and Fung, L. W. (2003). Solution structural studies on human erythrocyte alpha-spectrin tetramerization site. J Biol Chem *278*, 21837-21844.

Rehmann, H., Prakash, B., Wolf, E., Rueppel, A., De Rooij, J., Bos, J. L., and Wittinghofer, A. (2003). Structure and regulation of the cAMP-binding domains of Epac2. Nat Struct Biol *10*, 26-32.

Roll, R. L., Bauman, E. M., Bennett, J. S., and Abrams, C. S. (2000). Phosphorylated pleckstrin induces cell spreading via an integrin-dependent pathway. J Cell Biol *150*, 1461-1466.

Skrynnikov, N. R., Goto, N. K., Yang, D., Choy, W. Y., Tolman, J. R., Mueller, G. A., and Kay, L. E. (2000). Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin. J Mol Biol *295*, 1265-1273.

Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002). Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling. J Mol Biol *324*, 349-357.

Varadan, R., Assfalg, M., Haririnia, A., Raasi, S., Pickart, C., and Fushman, D. (2004). Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. J Biol Chem *279*, 7055-7063.