NMR structural characterization of the N-terminal domain of the adenylyl cyclase-associated protein (CAP) from *Dictyostelium discoideum*

Chrystelle Mavoungou[a], Lars Israel[b], Till Rehm[a], Dorota Ksiazek[a], Marcin Krajewski[a], Grzegorz Popowicz[a], Angelika A. Noegel[c], Michael Schleicher[b] & Tad A. Holak[a]

[a]*Max Planck Institute for Biochemistry, 82152 Martinsried, Germany;* [b]*Adolf Butenandt Institute, Cell Biology, Ludwig Maximilians-University, 80336 Munich, Germany;* [c]*Institute for Biochemistry, University of Cologne, 50931 Cologne, Germany*

Abstract

Cyclase-associated proteins (CAPs) are highly conserved, ubiquitous actin binding proteins that are involved in microfilament reorganization. The N-termini of CAPs play a role in Ras signaling and bind adenylyl cyclase; the C-termini bind to G-actin. We report here the NMR characterization of the amino-terminal domain of CAP from *Dictyostelium discoideum* (CAP(1-226)). NMR data, including the steady state ^{1}H-^{15}N heteronuclear NOE experiments, indicate that the first 50 N-terminal residues are unstructured and that this highly flexible serine-rich fragment is followed by a stable, folded core starting at Ser 51. The NMR structure of the folded core is an α-helix bundle composed of six antiparallel helices, in a stark contrast to the recently determined CAP C-terminal domain structure, which is solely built by β-strands.

Introduction

Cyclase associated proteins (CAPs) are ubiquitous proteins in eukaryotes and exhibit multifunctional activities due to the existence of domains involved in actin binding, adenylyl cyclase association, SH3 binding and oligomerization. CAP was first isolated from *Saccharomyces cerevisiae* as a component of the adenylyl cyclase complex (Field et al., 1990), assuming the role of a bridging protein that links the nutritional response signaling and changes in the cytoskeleton (Hubberstey and Mottillo, 2002; Stevenson and Theurkauf, 2000; Gottwald et al., 1996).

The CAP protein of *Dictyostelium discoideum* is involved in the microfilament reorganization at anterior and posterior plasma membrane regions (Gottwald et al., 1996). While CAPs of other organisms are trifunctional, *D. discoideum* CAP is so far known as a bifunctional protein. The G-actin binding activity has been localized to the carboxy-terminal domain of the protein (residues 306-464), which is separated by a proline rich linker region of 39 residues from

the N-terminal domain encompassing residues 1-226 (Figure 1). The N-terminal domain of CAP localizes the protein to the membrane (Noegel et al., 1999).

A comparison of the amino-terminal domain of CAP proteins of 14 organisms revealed a conserved RLEXAXXRLE motif (Hubberstey and Mottillo, 2002). This highly conserved motif interacts with adenylyl cyclase in yeast and has been termed the 'CAP signature' motif. In the *D. discoideum* CAP this motif has been replaced by a RLD-RLE motif. It is yet unclear if this conservative change of one amino acid could affect the adenylyl cyclase binding activity of the *D. discoideum* CAP.

In the present report we describe the NMR characterization of the amino-terminal domain of CAP from *D. discoideum* CAP(1-226) and the determination of the 3D structure of the stable folded core of this domain. Recently the structure of the N-terminal domain of CAP (residues 51-226) was also solved by X-ray diffraction (Ksiazek et al., 2003).
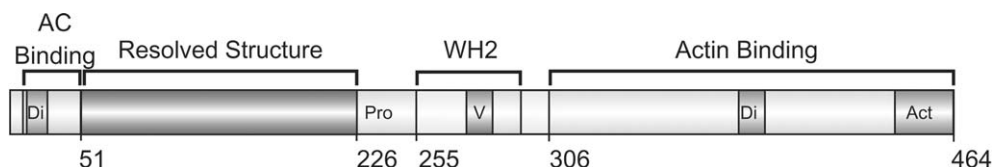
Figure 1. Domain structure of *D. discoideum* CAP (Gottwald et al., 1996; Hubberstey and Mottillo, 2002; Paunola et al., 2002). An adenylate cyclase binding domain (AC) and a dimerization domain (Di) are located at the amino terminus and are followed by the proline-rich region (Pro) and the WH2 domain (which includes a highly conserved verprolin homology region (V)). At the carboxy terminus is an actin binding domain (Act) and a second dimerization site (Di). The N-terminal domain consisting of residues 51-226 (CAP-N) used for our structure studies is highlighted.

Materials and methods

Sample preparation

The sample preparation of CAP was performed as previously described (Gottwald et al., 1996; Rehm et al., 2002a). The primary sequence of the folded N-terminal fragment (CAP-N) comprises residues Ser 51-Thr 226 (Figure 1). Uniformly ^{15}N and ^{15}N-^{13}C, as well as amino acid type selectively ^{15}N labeled samples, were prepared following standard procedures of Senn et al. (1987). The following samples were available at concentrations ranging from 0.5 to 1.2 mM at pH 7.3 containing 10% D_2O: Uniformly ^{15}N-, uniformly ^{15}N-^{13}C, selectively ^{15}N-Ala, ^{15}N-Phe, ^{15}N-Gly, ^{15}N-Ile, ^{15}N-Lys, ^{15}N-Leu, ^{15}N-Val, ^{15}N-Gly/^{15}N- Ser-labeled and unlabeled protein samples.

NMR spectroscopy

All NMR experiments were recorded at 300 K on Bruker DRX 600 and DMX 750 spectrometers equipped with triple resonance probeheads and pulsed-fields gradient units. The sequence specific resonance assignment was accomplished as reported previously by Rehm et al. (2002a) using a pair of HNCA and CBCA(CO)NH triple-resonance spectra, with the help of ^{15}N-HSQC and ^{13}C-HSQC spectra of the uniformly labeled samples of CAP-N, and also with the ^{15}N-HSQC spectra of the amino acid type selectively labeled samples. HNCO, 3D ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC spectra were also used for the assignment. All spectra were processed with the XWinNMR software package of Bruker and analyzed using the program Sparky (Goddard et al., 2000). The chemical shifts of CAP-N have been deposited in the BioMagResBank under the accession number 5393. The ^{1}H-^{15}N heteronuclear NOE experiment was recorded at pH 6.8 and 7.4 on the full length N-terminal CAP (226 residues) using a version of the experiment as described previously (Farrow et al., 1994; Mühlhahn et al., 1996; Renner et al., 2002). Saturation of the amide protons was carried out by applying 120° pulses prior to the experiment; the presaturation time was 2.5 s for all ^{1}H-^{15}N heteronuclear NOE experiments. 2048 data points and 256 increments were measured in the direct and indirect dimensions, respectively. NOE values were calculated by scaling ratios of peak heights in the NOE experiment with ^{1}H presaturation and the standard HSQC experiment obtained from the same sample. Recording of the NOE experiment without proton saturation using the same sample was not possible due to degradation and the low concentration of CAP(1-226). This simplified approach introduces an additional error of approximately 10–20% to the NOE values (Renner et al., 2002).

The assignment of the short construct could be transferred to the full length N-terminal CAP taking into account a few differences in chemical shifts of residues near amino acid 51 for the two constructs. Tentatively assigned peaks were verified using the HSQC spectra of the ^{15}N amino acid type selectively labeled samples and the NOESY-HSQC spectra, both recorded for the full length construct. Except for a few such residues, peaks that were not present in the HSQC spectra of the short construct were assumed to belong to the first 50 residue fragment. In general most of the resonances of the first 50 residues were either not present in the HSQC spectrum or clustered around the H^N chemical shift of 8.3 ppm (Figure 2A). No sequence specific assignment could be obtained for the visible resonances.

Input constraints and structure calculation

Interproton NOE distance constraints were generated from the integrated cross-peaks volumes from the 3D ^{15}N-NOESY-HSQC and 13C-NOESY-HSQC spectra
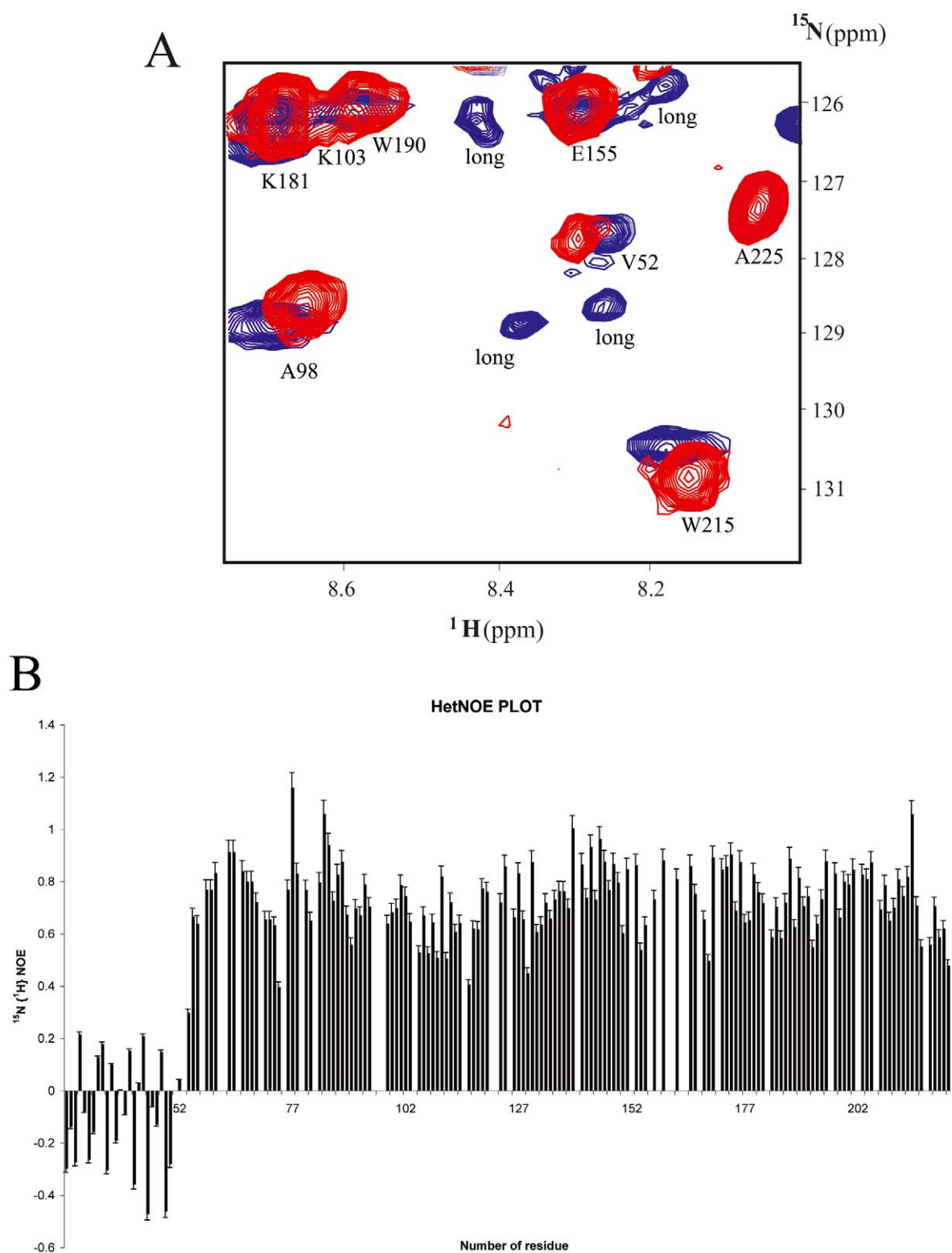
Figure 2. (A) Details of the HSQC spectra of CAP(1-226) (blue) and CAP-N (red). Peaks labelled 'long' are present only in the full-length construct, clustered around HN values of 8.3 ppm, show low or negative NOEs and have an appearance typical for an unstructured protein segment (Rehm et al., 2002b). (B) ^{1}H-^{15}N-heteronuclear NOE plot of the backbone amide protons of CAP(1-226). The NOE that could be measured for the first 50 residues are shown at the left-hand side of the graph. The NOEs of the assigned resonances start at Val 52.

using the program Sparky (Goddard et al., 2000). The NOEs were divided into three groups (strong, medium, and weak), and given upper distance bounds of 4.2, 5.2 and 6.0 Å, respectively, based on the volumes of characteristic sequential and medium-range NOEs for residues within ordered structural elements. The 2D ^{1}H NOESY spectra ($\tau_m = 120$ ms, $\tau_m = 40$ ms and $\tau_m = 20$ ms) were also used, mostly for the NOEs between amide protons; the 2D NOESY in D_2O ($\tau_m = 120$ ms) was used for the aromatic side chain NOEs. The peak volumes obtained from the 2D spectra were translated into distance constraints by grouping into the following distance ranges: 1.8-2.8 Å, 1.8-3.5 Å and 1.8-5.5 Å, using secondary structure elements for calibration. These groups of distance ranges correspond respectively to strong, medium and weak NOEs. Pseudo-atom corrections were added for distances that involved aromatic ring protons and non-stereospecifically assigned methyl and methylene protons, according to the method of Wüthrich (1986).

A 3D HNHA experiment was used to extract the values of $^3J_{HN\alpha}$ – coupling constants to determine ϕ angles. Additional inputs were provided by predicted ϕ and ψ torsion angles obtained from the program TALOS (Cornilescu et al., 1999). This torsion angle prediction uses the strong dependence of chemical shifts of backbone atoms to the local conformation in proteins. The minimum ranges used for all torsion angles constraints were $\pm 20°$.

1500 NOE distance constraints and 358 torsion angles restraints were used for the structure calculation in the program CNS_solve 1.1 (Brünger, 1995). An initial ensemble of 30 structures was generated in the Cartesian coordinate space using standard protocols for simulated annealing (Nilges et al., 1988, Weber et al., 2000). A stepwise refinement protocol was performed using the Powell energy minimization algorithms to obtain the lowest possible energy for all the structures (Powell, 1977). A final selection of the CAP-N structures was based on the low total energy criteria and the converged structure ensemble was then analyzed with the software program Suppose (Smith, 2002) for the rmsd calculations (Table 1). The superposition of the final structures <SA> was carried out using the software program MOLMOL (Koradi et al., 1996) and Swiss-PdbViewer 3.7.

Table 1. Parameters characterizing the structure determination of CAP-N in solution at pH 7 and 300 K

Parameters	<SA>	
RMS deviations from idealized geometry		
Bond lengths (Å)	0.0046 ± 0.000	
Angles (deg)	0.6191 ± 0.031	
Impropers (deg)	0.5332 ± 0.021	
Energies (kcal.mol^{-1})		
E_{all}	340.4	
E_{bond}	3.3	
RMS deviations of NOE violations		
Number of violations ≥ 0.2 Å	21.9	± 4.150
Number of violations ≥ 0.5 Å	0.21	± 0.043
RMS deviations from experimental constraints (Å)		
NOE class all (1500)	0.062	± 0.043

<SA> represents the ensemble of 18 structures. All parameters are calculated using the program CNS_solve 1.1.

Results and discussion

^{1}H-NMR spectra and sequence analysis of CAP(1-226) and CAP(51-226)

The full-length N-terminal construct of the *D. discoideum* CAP, encompassing residues 1-226, was unstable in that the proton 1D NMR spectra revealed the presence of cleaved peptide fragments in the sample after a few days, despite of the presence of a cocktail of protease inhibitors (Figure 3). Sequence analysis, mass spectrometry and a careful examination of the ^{1}H-^{15}N spectra of the amino acid type selectively labeled samples, indicated that the primary cleavage site was located between Ala 50 and Ser 51 (Figure 4). A subsequently cloned CAP-N (CAP(51-226)) was not further degraded, even after several months. ^{15}N-HSQC spectra of the full length construct show peaks which are mostly clusterred around 8.3 ppm ^{1}H chemical shift (Figure 2A). The linewidth is noticeable smaller for these residues and no NOESY crosspeaks could be detected in the NOESY-HSQC spectrum. The heteronuclear-NOE spectrum gave NOE values much lower than the average for residues 51-226, indicating increased flexibillity of this part (Figure 2B). Taken together the first 50 residues seem to be in a random-coil conformation. The spectra of CAP(1-226) were recorded with the fresh sample and ^{1}H-1D spectra were recorded before and after longer measurements (HSQC, HET-NOE) to monitor cleavage products.
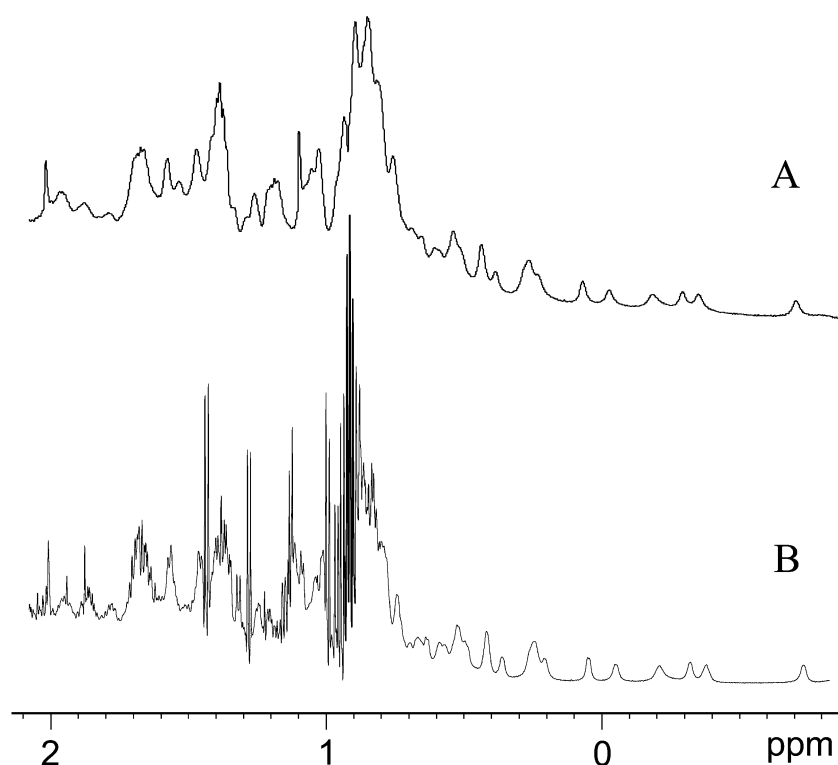
Figure 3. The aliphatic region of 1D proton spectra of CAP(1-226). (A) CAP(1-226) freshly prepared. (B) CAP(1-226) after 7 days at room temperature. Short poly-peptides give rise to sharp signals around 1 ppm.

The resonances of residues close to residue 51 are not at the exact same position in the spectra of the short and long constructs, indicating that the unstructured flexible N-terminal part is still attached to the well-structured core in the latter.

CAP exhibits the domain organization of all CAP-homologues (Gerst et al., 1992): The amino-terminal domain is separated from the carboxy-terminal domain by a proline-rich linker domain (Figure 1). Hubberstey and Mottillo (2002) reviewed the functional domains, using yeast CAP as an example, as follows: The amino-terminal domain interacts directly with adenylyl cyclase. The carboxy-terminal domain contains one dimerization site and is responsible for actin binding. The poly-proline linker domain carries

a SH3 recognition sequence. In *D. discoideum* CAP this latter domain is not conserved.

Our NMR characterization of the amino-terminal 226 residues of *D. discoideum* CAP revealed that the first 50 residues are unstructured and very flexible (Figure 2). This segment was removed for the structure determination reported here. Interestingly in *S. cerevisiae* CAPs the adenylyl cyclase binding domain, as well as the dimerization site, are found within these first 50 residues. This domain is not well-conserved, apart from a RLEXAXXRLE motive that is thought to be responsible for a coiled-coil interaction with adenylyl cyclase (Nishida et al., 1998). Our data corroborate a hypothesis whereby the un-
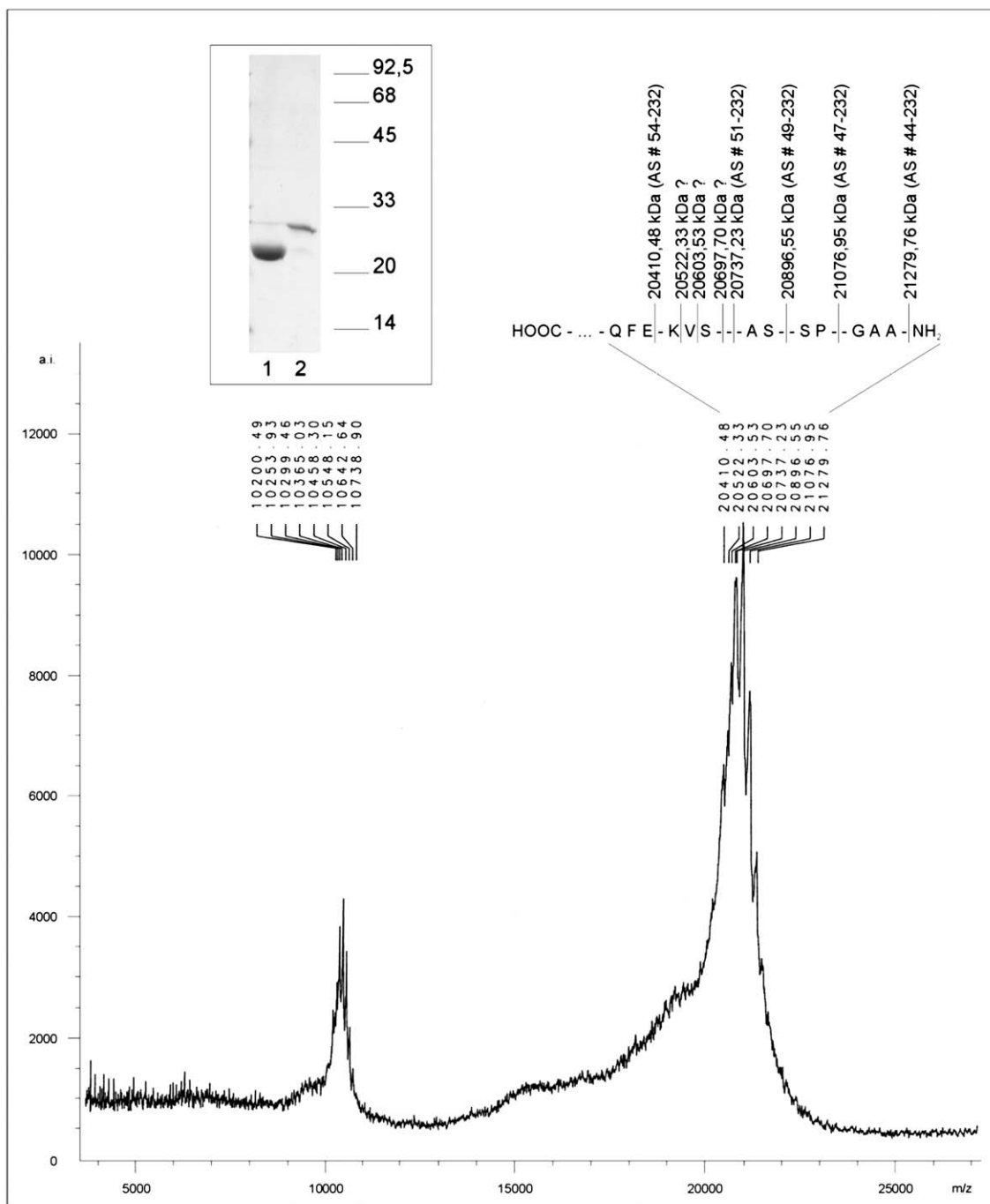
Figure 4. Mass spectrum of the amino terminal domain of CAP (1-226 residues plus a 6 residue His-tag) after 7 days at room temperature. The mass spectrum displays proteolytic degradation of CAP(1-226), despite of the presence of a cocktail of protease inhibitors. Masses of the cleaved peptides could be identified between 20410 and 21279 for residues 54-232, 51-232, 49-232, 47-232 and 44-232. The inset shows the SDS gel of the purified proteolytic product (lane 1) of CAP(1-226) (lane 2), corresponding molecular weight of 20 kDa and 25 kDa, respectively.
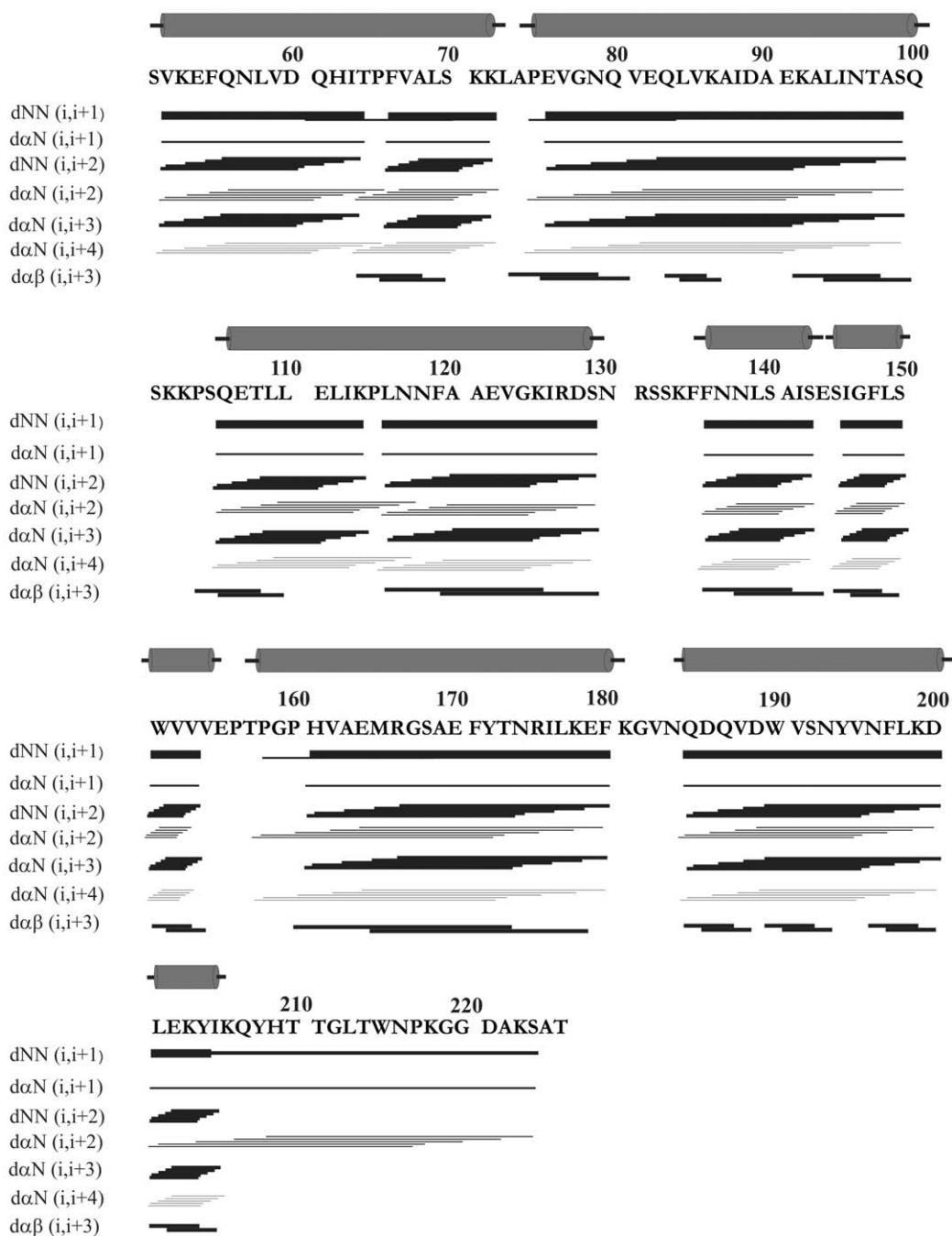
Figure 5. Summary of the short range connectivities ($|i - j| < 5$) involving the H^N, H^α and H^β protons. The thickness of the lines reflects the intensities of the cross-peaks observed in the spectra. The intensities of the signals corresponding to H^α (i) $-HN(i+1)$ and H^α (i) $-H^N(i+3)$ were nearly the same in all the spectra with few exceptions. The stretches of the H^N(i)-H^N(i+1) NOEs together with the presence of H^α (i) $-H^N$(i+3, 4) and H^α (i)-H^β (i+3) indicates the presence of helices extending from 52-73, 75-100, 107-128, 136-153, 158-180 and 185-208.

folded N-terminal 50-residue fragment would become structured only when bound to adenylyl cyclase.

NMR structure of CAP-N

An overview of the observed sequential, medium range NOEs for CAP-N is shown in Figure 5. The sequential and the medium range backbone NOEs were used for the determination of the secondary structure of the *D. discoideum* CAP protein. The NOEs, chemical shifts, and J couplings show the presence of a fully helical structure (Figure 5).

The three-dimensional structure of CAP-N indeed consist of six antiparallel helices, (Figures 6 and 7), each of them containing at least 10 to 20 amino acids. The helices are arranged into a six-helix bundle, which is connected in the complete protein to the C-terminal domain through a proline rich linker. In detail, the folded N-terminal domain consists of six helices in the regions extending from 52-73, 75-100, 107-128, 136-153, 158-180 and 185-208. In helix α1 one turn is distorted and ϕ and ψ angles for the two amino acids 61 and 62 do not fit the ideal helix conformation (ϕ, ψ : $-91.2°$, $-70.1°$; $-118.5°$, $1.2°$, respectively, in the minimized averaged (SA$_m$ structure). The pairwise backbone atomic rms difference of the structure ensemble is 1.4 ± 0.3 Å to the mean structure, excluding residues 100-105 and 213-226 (Figure 6).

The structure of the C-terminal domain of *S. cerevisae* CAP has been solved recently (Roswarski et al., to be published) (PDB ID: 1K4Z). In contrast to our N-terminal domain structure, the C-terminus of CAP is built solely by parallel β-strands that form a right-handed β-helix of six turns. The β-helix itself forms a homodimer with two β-structures arranged antiparallel to each other. It is interesting to note that the cyclase and actin binding sites are located in the whole protein on positions that are structurally independent from each other.

Comparison to the X-ray structure

The overall folds of the structures solved by NMR and X-ray crystallography are very similar (Figure 8). The RMS deviation between an averaged NMR structure and the X-ray model is 1.79 Å for all backbone heavy atoms (excluding the C-terminal residues 209-226) and even lower (1.6 Å) for helices only. The structures differ mostly in loop regions that lack well-defined secondary structure elements. The number of NOEs assigned to these regions is lower than that for

the helical regions and larger rms differences are observed in the ensemble of the NMR structures in these parts. Therefore it is not possible to ascertain whether the difference is due to flexibility of these fragments or lack of the NMR constraining data (or both). On the other hand, loop positions in the X-ray structure can be restricted by crystal packing. For CAP-N this indeed seems to be the case as all loops and the C-terminus are in contacts with molecules from neighboring cell units.

The α-helix between residues 53 to 72 (α1) is almost identical in both structures, so are helixes α3 and α6. The beginning of α2 (between residues 76-99) is located in a more external position in the NMR model, while its end is buried deeper into the center of the molecule compared with the X-ray structure. Helix α5 is bent in the NMR model, near residue 170, shifting amino acids 158-169 to a more external position, while the X-ray model shows a straight helix.

The major difference in the structures pertains to helix α4. The NMR structure shows a continuous α-helix built by residues 136-153, while in the crystallographic model the helix ends at residue 143 and residues 144-158 form a long linker between the helices α4 and α5. CAP-N crystallized both as a dimer and a monomer from the same drop (30% PEG8000, 0.2 M MgCl$_2$, 10 mM β-mercaptoethanol and 0.1 M MES, pH 6.1; Ksiazek et al., 2003). The final X-ray structure of the dimer comprises CAP-N and a magnesium atom, and was solved to 1.4 Å resolution. The monomer crystals diffracted to 1.7 Å. Comparison of the monomer structure with that of the dimer showed that they were essentially identical. For the dimer, Arg 127, Asp 128 (end of α3) and Glu 144 from each monomer are at hydrogen bonding distances to water molecules, which surround the magnesium ion at an average distance of 2.1 Å. The interaction of Mg with Glu 144 could then explain why helix α4 is broken at this residue in the X-ray model of the dimer. No Mg has been seen in the X-ray model of the monomer structure. However, Mg still could be present at Glu 144 since the metal might not be recognized at the 1.7 Å resolution for the monomer crystals and lower occupancy of Mg would further weaken chances of its detection. NMR titration of CAP-N with Mg^{2+} did not show any dimerization, which could be detected by changes in NMR linewidth or by induced chemical shifts. We propose that the NMR structure is closer to the native structure as crystals of both dimer and monomer were grown at a non-physiological concentration of MgCl$_2$ (0.2 M).
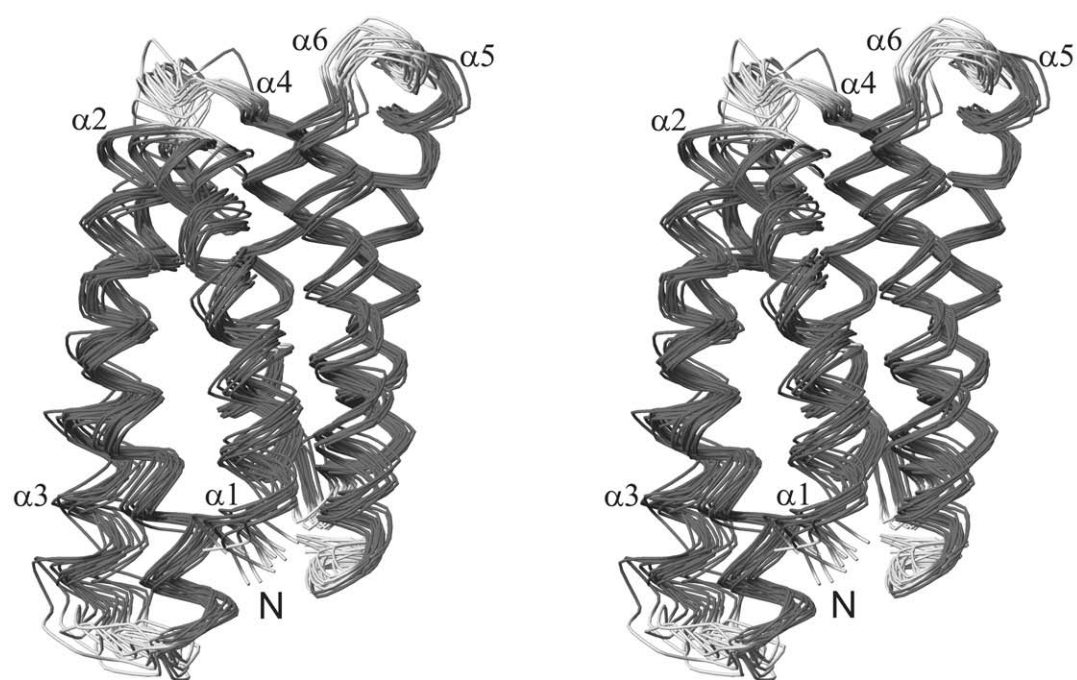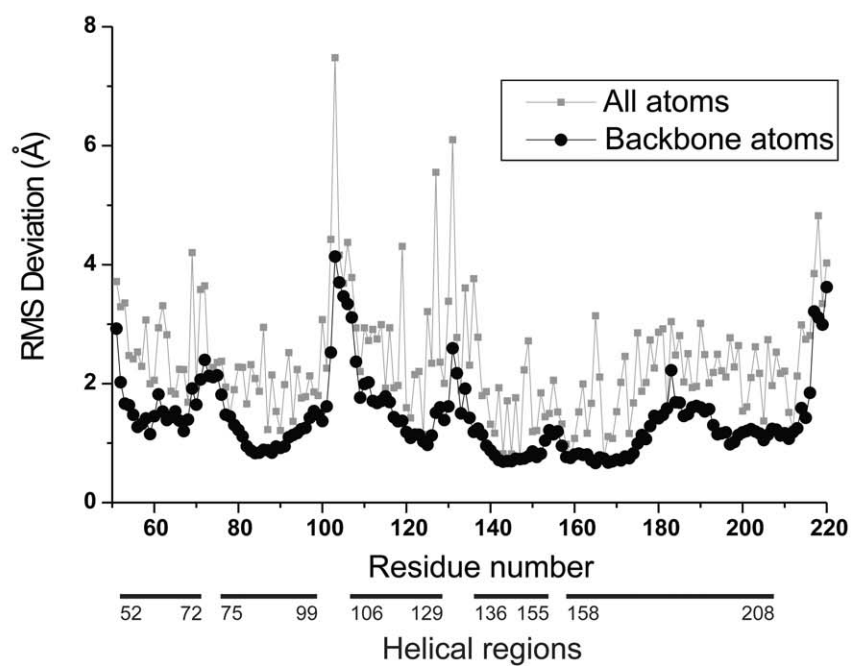
Figure 6. (A) Stereoview of the backbone atoms (N, C^α, C′ and O) of all residues for the family of 18 structures of CAP best fit to N, C^α and C′ atoms of the regions with regular secondary structure (52-73, 75-100, 107-128, 136-153, 158-180 and 185-208). (B) Atomic rms distributions for the heavy backbone atoms and for all heavy atoms of the <SA> to the minimized averaged structure (SA_m), best fit in all the helical regions.
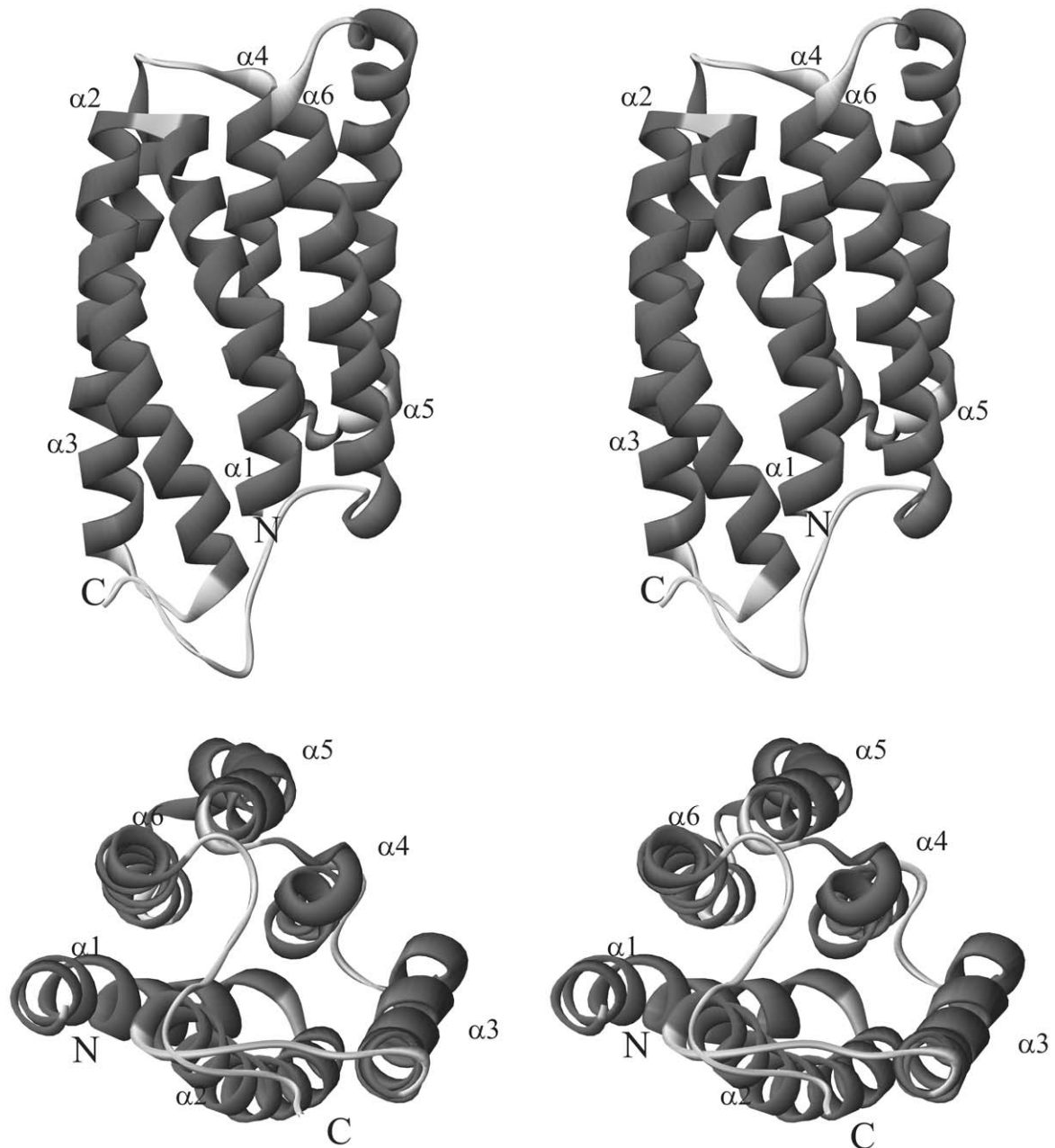
Figure 7. Ribbon plot of the minimized averaged (SA$_m$) CAP-N structure from two different views.

Conclusions

Our NMR investigation of the 226 amino acid N-terminal domain of the CAP protein from *D. discoideum* revealed an unstructured 50 amino acid segment at the N-terminus. The remaining 176 residues form a stable, globular structure. The folded N-terminal core is exclusively composed of α-helices connected by irregular loops of 5-12 residues; the helices are arranged into a six-helix bundle.
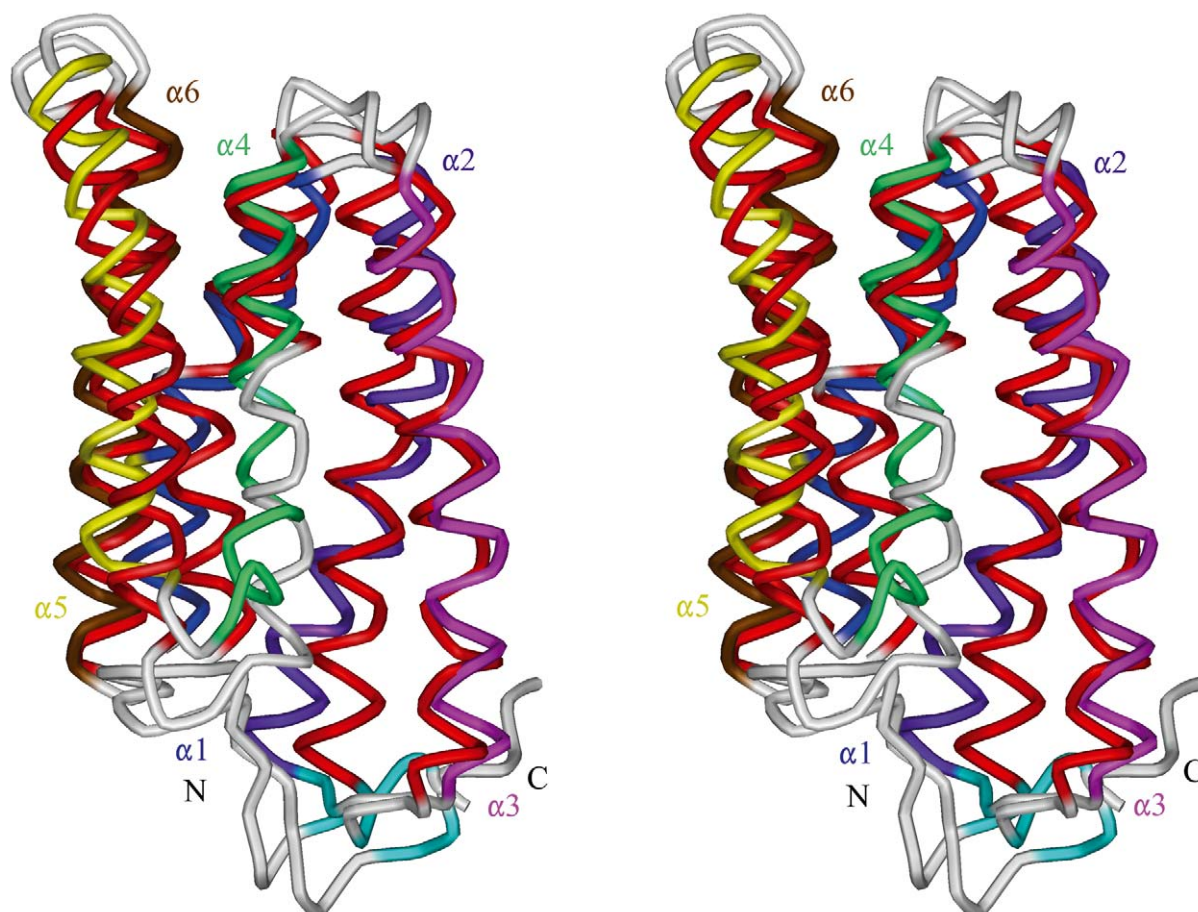
Acknowledgement

Figure 8. Stereoview of the C$^{\alpha}$-backbone of the X-ray structure (helices in red) superimposed on the minimized averaged NMR structure (the six helices are shown each in different colour).

References

Brünger, A.T. (1995) CNS_solve 1.1. *A System for X-Ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Cornilescu, G., Delaglio, F. and Bax, A. (1999) *J. Biomol. NMR*, **13**, 289–302.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Foreman- Kay, J.D. and Kay, L.E. (1994) *Biochemistry*, **33**, 5984–6003.

Field, J., Vojtek, A., Ballester, R., Bolger, G., Colicelli, J., Ferguson, K., Gerst, J., Kataoka, T., Powers, S., Riggs, M., Rodgers, L., Wieland, I., Wheland, B. and Wigler, M. (1990) *Cell*, **61**, 319–327.

Gerst, J.E., Rodgers, L., Riggs, M. and Wigler, M. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 4338–4342.

Goddard, T.D. and Kneller, D.G. (2000) '*SPARKY 3*', University of California, San Francisco.

Gottwald, U., Brokamp, R., Karakesisoglou, I., Schleicher, M. and Noegel A.A. (1996) *Mol. Biol. Cell*, **7**, 261–272.

Hubberstey, A.V. and Mottillo E.P. (2002) *FASEB J.*, **16**, 487–499.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 52–55.

Ksiazek, D., Brandstetter, H., Israel, L., Bourenkov, G.P., Katchalova, G., Janssen, K.-P., Bartunik, H. D., Noegel, A.A., Schleicher, M. and Holak, T.A. (2003) *Structure*, **11**, 1–20.

Mühlhahn, P., Bernhagen, J., Czisch, M., Georgescu, J., Renner, C., Ross, A., Bucala, R. and Holak, T.A. (1996) *Protein Sci.*, **5**, 2095–2103.

Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **239**, 129–136.

Nishida, Y., Shima, F., Sen, H., Tanaka, Y., Yanagihara, C., Yamawaki-Kataoka, Y., Kariya, K., and Kataoka, T. (1998) *J. Biol. Chem.*, **273**, 28019–28024.

Noegel, A.A., Rivero, F., Albrecht, R., Janssen, K.P., Koehler, J., Parent, C.A. and Schleicher, M. (1999) *J. Cell Sci.*, **112**, 3195–3203.

Powell, M. J. D. (1977) *Math. Program.*, **12**, 241–254.

Rehm, T., Mavoungou, Ch., Israel, L., Schleicher, M. and Holak, T.A. (2002a) *J. Biomol. NMR*, **23**, 337–338.

Rehm, T., Huber, R. and Holak, T.A. (2002b) *Structure*, **10**, 1613–1618.

Renner, C., Schleicher, M., Moroder, L. and Holak T.A. (2002) *J. Biomol. NMR*, **23**, 23–33.

Senn, H., Eugster, A., Otting, G., Suter, F. and Wüthrich, K. (1987) *Eur. Biophys. J.*, **14**, 301–316.

Stevenson, V. A. and Theurkauf, W. E. (2000) *Curr. Biol.*, **10**, R695–697.

Weber, T., Baumgartner, R., Renner, C., Marahiel, M.A. and Holak, T.A. (2000) *Structure*, **8**, 407–418.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.