Journal of Biomolecular NMR, **24:** 89–102, 2002.
KLUWER/ESCOM

Solution structure of the coiled-coil trimerization domain from lung surfactant protein D

Helena Kovacs[c,d], Sean I. O'Donoghue[c,e], Hans-Jürgen Hoppe[b], David Comfort[a,g], Kenneth B. M. Reid[b], Iain D. Campbell[a] & Michael Nilges[c,f,*]
[a]*Department of Biochemistry and* [b]*MRC Immunochemistry Unit, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.;* [c]*European Molecular Biology Laboratory, D-69012 Heidelberg, Germany. Present addresses:* [d]*Bruker BioSpin AG, Industriestrasse 26, CH-8117 Fällanden, Switzerland;* [e]*Lion Bioscience AG, Waldhoferstr. 98, 69123 Heidelberg, Germany;* [f]*Unité de Bio-informatique structurale, Institut Pasteur, 25-28 rue du docteur Roux, F-75015 Paris, France;* [g]*Department of Chemistry and Biochemistry, UCLA-DOE Laboratory of Structural Biology, University of California Los Angeles, Los Angeles, CA 90095, U.S.A.*

Received 8 May 2002; Accepted 7 August 2002

Key words: ambiguous distance restraints, coiled coil, lung surfactant protein, NMR-spectroscopy, trimer

Abstract

Surfactant protein D (SP-D) is one of four known protein components of the pulmonary surfactant lining the lung alveoli. It is involved in immune and allergic responses. SP-D occurs as a tetramer of trimers. Trimerization is thought to be initiated by a coiled coil domain. We have determined the solution structure of a 64-residue peptide encompassing the coiled coil domain of human SP-D. As predicted, the domain forms a triple-helical parallel coiled coil. As with all symmetric oligomers, the structure calculation was complicated by the symmetry degeneracy in the NMR spectra. We used the symmetry-ADR (ambiguous distance restraint) structure calculation method to solve the structure. The results demonstrate that the leucine zipper region of SP-D is an autonomously folded domain. The structure is very similar to the independently determined X-ray crystal structure, differing mainly at a single residue, Tyr248. This residue is completely symmetric in the solution structure, and markedly asymmetric in the crystalline phase. This difference may be functionally important, as it affects the orientation of the antigenic surface presented by SP-D.

Abbreviations: ADR – ambiguous distance restraint; CRD – carbohydrate recognition domain; HSQC – heteronuclear single quantum correlation; MBL – mannose binding lectin; NMR – nuclear magnetic resonance; NOE – nuclear Overhauser effect; NOESY NOE – spectroscopy; SP-D – surfactant protein D; TOCSY – total correlation spectroscopy.

Introduction

Pulmonary surfactant reduces surface tension in the air-liquid interface of the lung alveoli, thereby preventing alveolar collapse at the end of expiration. The surfactant contains mainly phospholipids. In addition, four polypeptides are present in minor amounts, called surfactant proteins A through D (SP-A, SP-B, SP-C and SP-D; Johansson et al., 1994). SP-D participates in immune and allergic responses (Weaver and Whitsett, 1991) and has also been implicated in the reorganization or turnover of pulmonary surfactant (Kishore et al. 1996).

Electron microscopy studies showed the SP-D protein to occur as a tetramer of trimers. The trimers are rod-shaped, and the four rods are arranged into a quaternary cross shape (Thiel and Reid, 1989). Each trimer rod consists of three identical polypeptide chains associated in parallel. The polypeptide chain has four distinct structural domains: an N-terminal

*To whom correspondence should be addressed. E-mail: nilges@pasteur.fr

segment involved in oligomerization through interchain disulfide bridges; a collagen-like triple-helical domain; a 'neck' domain with typical coiled coil heptad pattern; and a C-terminal C-type lectin or carbohydrate recognition domain (CRD), which binds to distinct carbohydrates on the surface of pathogens (Miyamura et al., 1994). SP-D belongs to a class of modular fluid-phase proteins called collectins (Hoppe and Reid, 1994), which all share the same domain organization (Holmskov et al., 1994).

Apart from SP-D, two collectin proteins were studied at atomic resolution: the mannose binding lectins (MBL) in human (PDB code 1hup, Sheriff et al. 1994) and in rat (1rtm, Weis and Drickamer, 1994). Fragments that include the neck and CRD domain, but exclude the collagen-like domain, form trimers. The trimer interface is dominated by the neck domain, which forms a triple-helical parallel coiled coil. It has been postulated that trimerization in the collectins is initiated in the coiled coil, which then nucleates the formation of the collagen triple helix (Hoppe et al., 1994). The collectin coiled coil region would then act as an autonomous trimerization domain, analogous to the leucine zipper dimerization domain (Landschulz et al., 1988).

Another possible role for the coiled coil domain is to determine the relative orientation of the three CRD domains, thus influencing antigen recognition and binding. The CRD orientation varies within the collectin super family: in human MBP (1hup), the CRDs are separated by 45 Å; in rat MBP (1rtm), the separation is 53 Å. Sheriff et al. (1994) proposed that these variations might arise from differences both in the coiled coil inter-helical spacing, and in the interface between the CRDs and the coiled coil domain. This interface involves mostly hydrophobic residues in the C-terminal part of the coiled coil domain.

In the present study, we extend a previous NMR study of a 64-residue fragment encompassing the coiled coil domain and seven repeats of the adjacent collagen domain of the human SP-D (Hoppe et al., 1994). We had previously verified that this fragment forms a stable trimer, within pH 3.0–8.5 and below 55 °C. Judging from the NMR spectra, the collagen-like domain is not folded. When this domain is protolytically removed, the remaining coiled coil domain (plus seven residues from the CRD domain) is still able to trimerize after heat denaturation. In contrast, the isolated coiled coil domain does not trimerize in rat MBL, but needs the CRD domain for the formation of stable trimers (Weis and Drickamer, 1994). This difference is probably due to MBL having only three complete repeats in the coiled coil domain, while SP-D has a fourth repeat. Because these repeats consist of seven residues they are called heptads. The fourth heptad repeat of SP-D has a somewhat unusual composition, with aromatic residues at the *a* and *d* positions (Phe and Trp, respectively). Figure 1 shows schematically the helical wheel of the SP-D neck peptide sequence and the packing of the three helices into a coiled coil trimer.

We report here the complete solution structure of the coiled coil trimer of SP-D, and compare it to the independently determined X-ray crystal structure of a fragment comprising the coiled coil and CRD domains from SP-D (PDB accession code 1bo8; Hakansson et al., 1999). The determination of the NMR structure was considerably complicated by the symmetry of the coiled coil. Symmetric oligomers always present a special difficulty for the NMR structure determination because of the symmetry degeneracy in the NMR spectra, which leads to ambiguity in the derived distance data. Whereas the experimental approach to address this problem by asymmetric labelling and X-filtered NMR-experiments (Folmer et al., 1997) is sufficient to resolve the ambiguity in symmetric homodimers, it alone is no longer sufficient for trimers or higher order oligomers. For these, more elaborate experimental schemes have to be employed (Jasanoff, 1998).

In the absence of additional experimental information to alleviate the symmetry ambiguity one has to resort to manual assignment or an appropriate calculation method. To date, the only generally applicable calculation approach allowing for flexibility of the monomers incorporates the ambiguous NOEs into the structure calculation in the form of ambiguous distance restraints (ADRs) and uses special restraints to ensure a symmetric arrangement of the monomers (Nilges, 1993; O'Donoghue et al., 1996). This 'symmetry-ADR method' has been successfully used in the determination of several symmetric oligomers, including a hexamer (O'Donoghue et al., 2000). We used the symmetry-ADR method successfully to the determination of the solution structure of the SP-D trimer.

In addition to the problems posed by symmetry, coiled coils present a further challenge for NMR, since the critical interface residues are usually of the same type (i.e., all leucines or isoleucines), and hence difficult to assign due to poor chemical shift dispersion. Coiled coil structures solved by NMR include the

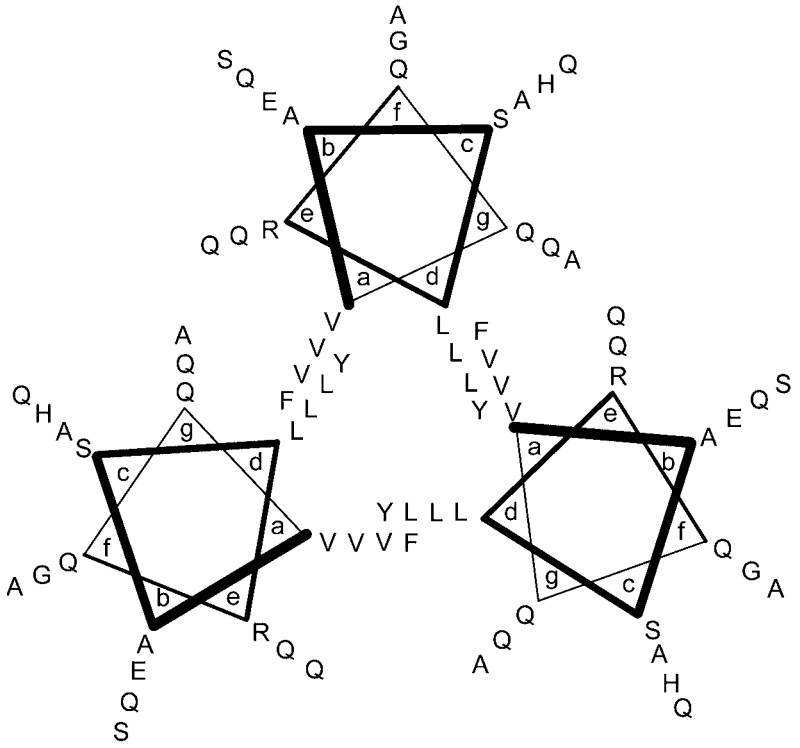

Figure 1. Projection of residues V224-Y248 from each of the three SPD neck peptide sequences onto a helical wheel, viewing down the helical axis starting at the N-terminal end. Every seven residues constitute a structural 7-residue repeat called a heptad. The individual positions of the seven residues are denoted by letters *a* to *f*. The hydrophobic interface residues in positions *a* and *d* are high-lighted in magenta.

symmetric dimer of the Jun leucine zipper domain (O'Donoghue et al., 1993, 1996; 1jun, Junius et al., 1996); two symmetric coiled coil trimers by means of X-filtered NMR-experiments, gp41 (2ezo, Caffrey et al., 1997) and the cartilage matrix protein (1aq5, Dames et al., 1998); the tetramerization domain of the Mnt repressor, which consists of two right-handed coiled coils (1qey, Nooren et al., 1999) and the GAL4 dimerization domain (1hbw, Hidalgo et al., 2001). There are also the NMR-studies of the solution structure of a C-terminal coiled-coil domain from bovine IF(1), which is an antiparallel coiled coil with histidine residues in the dimer interface (1hf9, Gordon-Smith et al., 2001) and the spectrin repeat, which is an asymmetric antiparallel triple-helical coiled coil (1aj3, Pascual et al., 1997). Several coiled coils have been determined at atomic resolution by X-ray crystallography. These include several triple-helical coiled coils (1qfu, Wilson et al., 1981; 1cos, Lovejoy et al., 1993; 1gcm, Harbury et al., 1993, 1994; 1rtm, Weis and Drickamer, 1994; 1hup, Sheriff et al., 1994; 1mof, Fass et al., 1996; 1swi, 1eij, Gonzalez et al., 1996a, b; 1aik, Chan et al., 1997; 1coi, Ogihara et al., 1997).

Methods

Protein sample

The 64-residue polypeptide GSPGLKGDKGIPGD-KGAKGESGLPD - VASLRQQVEALQGQVQHLQA AFSQYKK - VELPNGGIPHRD was expressed and purified, both ^{15}N-labelled and unlabelled, as described previously (Hoppe et al. 1994). In addition to the coiled coil domain, comprising residues V224-K250 (separated above by dashes), the polypeptide contained seven N-terminal G-X-Y collagen triplets (G202-P222) plus D223, and a C-terminal continuation of seven residues (V251-G257) of the sequence encoding the CRD domain. The vector also contained an additional G-S-P triplet in the N-terminus and five residues from the pGex polylinker in its C-terminus. The concentration of the unlabelled NMR sample was 2.0 mM and the concentration of the ^{15}N-labelled sample was 2.8 mM. The samples were dissolved in 90% H_2O/10% D_2O at pH 4.4.

NMR experiments

The spectra were recorded in Oxford on home-assembled 600 MHz and 750 MHz spectrometers interfaced to a GE Omega computer. The sample temperature was 30 °C, and the spectra were referenced to the water frequency at 4.70 ppm. Two-dimensional homo-nuclear spectra were recorded at 600 and 750 MHz ^{1}H fields with a spectral width of 13 ppm in both dimensions, and with typically 2048 complex points in the acquisition dimension and 1024 increments in the indirect dimension. The mixing time was 150 ms in the homo-nuclear NOESY experiments and 45 ms in the homo-nuclear TOCSY experiments. The three-dimensional experiments were recorded on the 600 MHz spectrometer. The ^{15}N sweep width was 30.68 ppm and the ^{1}H sweep width was 13.23 ppm. The ^{15}N carrier frequency was set to 121.25 ppm and $128\times32\times512$ complex points were collected. The mixing time was 150 ms in the 3D-NOESY-HMQC experiments and 30 ms in the 3D-TOCSY-HMQC. We used the software packages Felix 2.3 (Molecular Simulations Inc.) for spectral processing and XEASY (Bartels et al., 1995) for spectral analysis.

Experimental restraints

Proton-proton distance restraints (denoted by d) were calibrated against the NOE signal intensity as strong ($d < 3.0$ Å) medium ($3.0 < d < 4.0$ Å) or weak ($4.0 < d < 6.0$ Å). Unfortunately, we did not have the resources to perform mixed labelling or X-filtered experiments. Instead, we were able to assign certain backbone-backbone and backbone-H$^{\beta}$ NOEs as unambiguously intra-monomer, by analysing inter-proton distances in modelled and known coiled coil trimers (see Results). These restraints were used to calculate initial monomer structures, which served to identify characteristic α-helical hydrogen bonds. We used these hydrogen bonds in the subsequent calculations of the trimer as ambiguous distance restraints (ADRs), allowing for an acceptor-O at positions $i + 3$ or $i + 4$. The remaining ambiguous NOEs were used in the trimer calculations as symmetry-ADRs, i.e., including the intra-monomer and two different inter-monomer possibilities. There are two different ways to define these ADRs, either by using one restraint for each NOE including all intra- and all inter-monomer contributions to the NOE, or by using three separate restraints for each NOE, each including one intra-monomer and two inter-monomer contributions; we adopted the latter method.

Monomer structure calculations

All structure calculations for this work were performed with X-PLOR 3.851 (Brünger, 1993) using 'molecular dynamical simulated annealing' protocols similar to those described by Nilges et al. (1988). The monomer structures were calculated by adapting the 'refine_gentle.inp' protocol from the X-PLOR manual (Brünger, 1993), starting from chains with randomized ϕ and ψ angles. Prochiral pairs were treated with SUM averaging (Nilges, 1993), which makes pseudo atom corrections unnecessary. Structures were accepted if they showed no violations larger than 0.5 Å of distance restraints or 5 degrees of torsion angle restraints, respectively.

Trimer structure calculations

Each trimer calculation started from a different low-energy monomer structure. Two identical copies of the monomer were made, and the three monomers were placed in a parallel, symmetric orientation by rotation around one of the coordinate axes by $+/-120$ degrees. This already defined the symmetry axis of the trimer, but the interface between monomers varied randomly between calculations. The trimer structure was then calculated with the symmetry-ADR method essentially as described (Nilges, 1993; O'Donoghue et al., 1996; O'Donoghue and Nilges, 1999). In this method, the overall symmetry of the molecule is maintained with two symmetry restraints: a so-called non-crystallographic symmetry restraint, minimizing the structural difference between monomers (called 'NCS restraint' in X-PLOR); the initial and final energy constants for this term were set to of 1 and 2 kcal mol^{-1}, respectively; and distance-symmetry restraints, which maintain the relative orientation of the monomers symmetric, with an energy constant of 5 kcal mol^{-1} throughout. The unambiguous intra-monomer NOEs were used with initial and final energy constants of 50 and 75 kcal mol^{-1}, respectively. An additional inter-helical restraint was used to keep the geometric centres of seven consecutive Cα-atoms on one helix within 12.5 Å from the centres of the equivalent Cα-atoms on the other helices (Nilges and Brünger, 1991). This prevented the monomers from diverging, but allowed their separation to be a self-adjusting parameter. The initial and final energy constants for this restraint were 0.2 and 2.0 kcal mol^{-1}, respectively. Prochiral pairs were treated by repeatedly swapping assignments in the course of the calculation and selecting the lower energy conformation (Folmer et al., 1997). Since no

Table 1. Summary of the conformational constraints

Iteration	p	Number of restraints		
		Ambiguous	Unambiguous intramonomer	Unambiguous intermonomer
1	0.999	328	57	0
2	0.99	266	119	0
3	0.98	232	152	1
4	0.96	169	199	17
5	0.93	122	236	27
6	0.90	105	249	31
7	0.80	66	275	44
8[a]	–	66	275	44

[a]Refinement in water.

stereo-specific assignments were available, this swapping strategy was found to be important to achieve the proper side-chain packing of the valines and leucines in the trimer interface. An initial ensemble of 200 trimers was calculated.

The drawback of ADRs is that they decrease the convergence towards the global minimum. Consequently, the calculated ensemble may have very few or no low energy structures. To increase convergence, we iteratively assigned the symmetry-ADRs with a version of ARIA (Nilges et al., 1997; Nilges and O'Donoghue, 1998) adapted to the calculation of symmetric oligomers, interfaced to X-PLOR. ARIA iteratively removes those assignment possibilities from an ambiguous NOE that are judged to be very unlikely, based on the structure ensemble calculated in the previous iteration. The cutoff value, p, controlling the assignment was set to a value close to one ($p = 0.999$) for the first round of assignments, based on the initial ensemble. Values close to one exclude only very unlikely possibilities, while values close to zero exclude all but the dominant contribution. Useful values range between $p = 1.0$ and $p = 0.8$. 100 structures were calculated in each iteration. A total of eight assignment and refinement iterations were performed, using the values for the cut-off value p given in Table 1.

In addition to the partial assignment in ARIA, we improved convergence further with the interface filter method (O'Donoghue et al., 1996). Residues with remaining ambiguous NOEs after the first assignment iteration were considered to belong to the intermonomer interface. In subsequent iterations, starting structures were selected that satisfied the criterion that the d^{-6}-summed distance between the Cα-atom of each interface residue and the Cα-atoms of all interface residues on other monomers should be less than 9.0 Å.

The lowest energy structures in the eighth ensemble were then further refined in a 9 Å shell of solvent water (Linge and Nilges, 1999). The quality of the final ensemble was assessed with PROCHECK (Laskowski et al., 1993).

Initial model calculations

We also calculated model structures of the SP-D trimer, independent of the NMR data, using the method of Nilges and Brünger (1993), which starts from a template representing an idealized coiled coil and refines the model by molecular dynamics. The template consisted of three identical α-helices arranged in a parallel and symmetric orientation with infinite pitch (i.e., no super coiling). Each of the helices had a ratio of 3.5 residues per turn, which means that they were slightly under wound in comparison to a standard helix. Side-chains were built automatically on the backbone. During the following short molecular dynamics calculation the characteristic coiling developed, but the position of the helical subunits with respect to one another was not well defined.

Coiled coil parameters

Coiled coil parameters were determined using the program ThreadCoil (S.I. O'Donoghue, unpublished). Given a three-dimensional protein structure, the program uses a grid search method to find the coiled coil with ideal geometry (Crick, 1953) with the closest match.

Accesssion numbers

Coordinates for the 21 final structures, and the NMR restraints used to calculate them have been deposited with the Broohaven Protein Data Bank (accession code 1M7L).

Results

Spectral analysis

A qualitative analysis of the spectra allowed a clear distinction between unstructured regions of the peptide, characterized by intense, narrow cross peaks

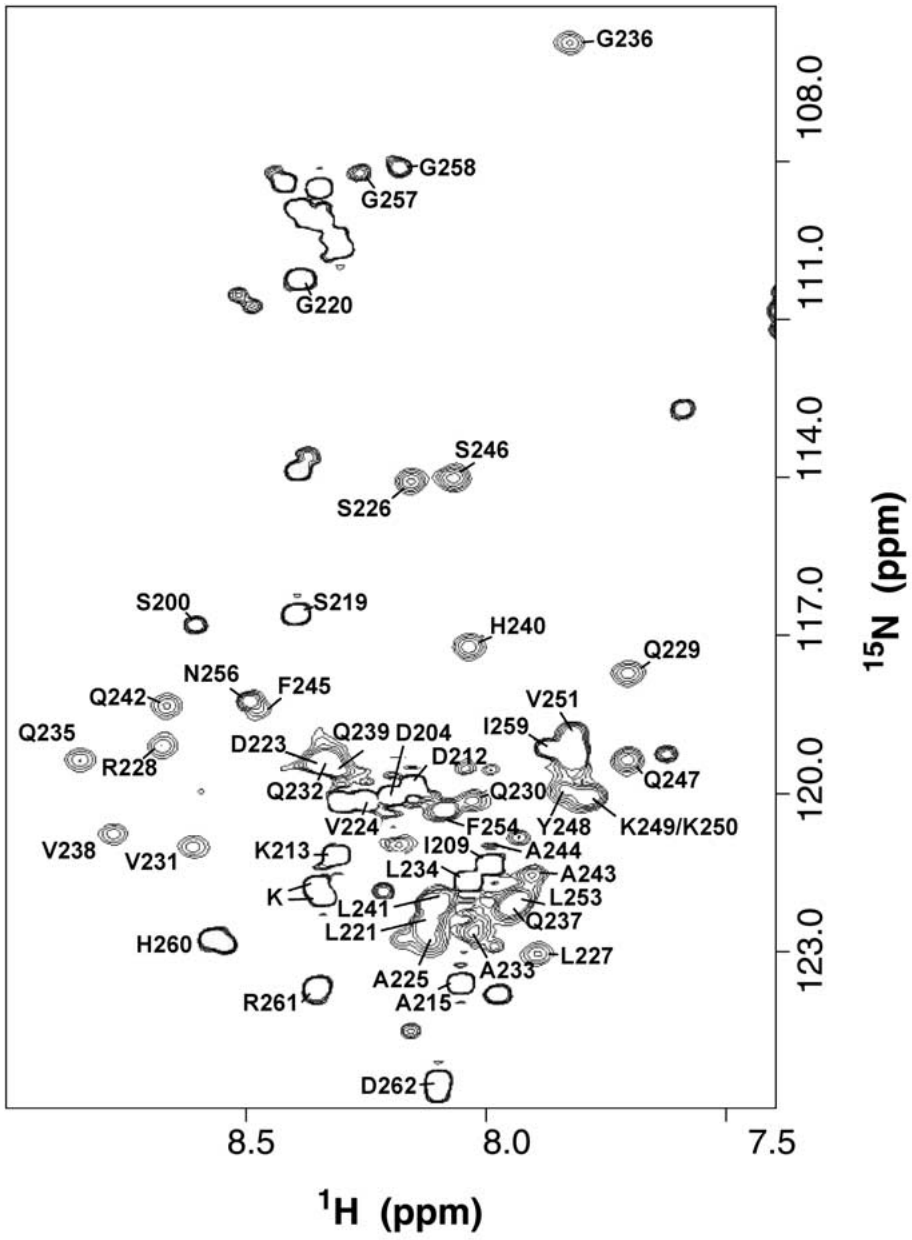

Figure 2. The ^{15}N-correlated HSQC spectrum the 64-residue polypeptide fragment at 600 MHz and 30 °C.

complicating the spectral analysis, and a folded segment, characterized by broader, more dispersed cross peaks. The number of amide proton resonances in the fingerprint region of the TOCSY and in the ^{15}N-HSQC spectrum, cf. Figure 2, corresponded to a single set of resonances for each residue in the monomeric peptide. This indicates point group 3 symmetry and therefore parallel orientation of the monomers. The broader cross-peaks were assigned to residues D223 to K250. For this part of the trimer, we observed the NOE connectivities characteristic of a regular α-helix (see Figure 3). The residues that could be identified within the N- and C-terminal segments have for the most part random coil chemical shifts and only intra-residue and sequential connectivities. Hence, the termini, including the collagen repeats, are flexible and disordered in solution. We included therefore only the central region from G220 to I258 in the subsequent structure calcula-

Figure 3. Display of intra-residue, sequential, and medium-range NOE connectivities in the coiled coil-region of SP-D. The plot distinguishes between weak and strong sequential $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ constraints (the cutoffs are 3.0, 3.6, and 3.6 Å, respectively). Figure prepared with DYANA (Güntert et al., 1996).

tion. Overlap of the H^{α} and the side chain resonances of V231 and V238, and of the methyl signals of L234 and V251, complicated the identification of distance restraints for these residues located in the trimer interface and lowered the precision of their side chain positions.

Structure modelling and distance analysis

Analogous to the study of O'Donoghue et al. (1993) for dimeric coiled coils, we surveyed backbone-backbone inter-proton distances for a number of known coiled coil trimers (the MBLs 1hup and 1rtm), and 33 model SP-D structures generated as described in the Methods section. In each of these trimeric coiled coils, we found no inter-helical backbone-backbone distances shorter than 6 Å. We concluded that any backbone-backbone interactions observed in the NOESY spectra, such as HN-HN, HN-H^{α} or HN-H^{β}, must be intra-monomer. In this way, we assigned 239 NOEs as unambiguously intra-monomer.

Structure calculation

With the 239 intra-monomer NOES, we calculated a set of monomer structures. The remaining 385 NOEs were used as symmetry-ADRs in the subsequent trimer calculations, starting from the monomer structures. A part of the NOEs were also ambiguous due to the chemical shift dispersion degeneracy. Thus the ADRs were set up to not only include the symmetry ambiguities but also those due to chemical shift. After assigning NOEs based on the initial ensemble, all *a*, *d*, *e*, and *g* heptad position residues (see Figure 2) were recognized as being part of the inter-monomer interface and used to filter starting structures in subsequent iterations. In the course of the ARIA iterations, the convergence successively improved, as illustrated in Figure 4, which displays the 21 lowest energy structures from the iterations 2, 5 and 8. The corresponding distributions of NOE restraints are reported in Table 1. By the eighth ARIA iteration, 44 NOEs had been assigned as entirely inter-monomer, while 66 NOEs remained ambiguous (cf. Table 1). This was sufficient to define a rather tight structural ensemble. 100 structures were generated within the final dataset, and of these, the 24 lowest energy structures were refined in explicit water. The 21 structures with lowest NOE energy were then selected as the final ensemble, which is depicted in Figure 5. The dihedral angle distribution of the final ensemble is shown in Figure 6.

The SP-D solution trimer

The final ensemble of 21 refined structures (Figure 5) is well defined in the helical region, as illustrated by Figures 5 and 6. The ensemble shows no significant NOE violations, and it has no ϕ/ψ angles in disallowed regions (cf. Table 2). Consistent with our model distance analysis (above), the ensemble shows no proton backbone-backbone distances shorter than 6 Å, supporting our assumption that NOEs of this class can be unambiguously assigned as intra-monomer. Figure 7 presents the 44 NOEs assigned as exclusively inter-monomer. Among these, 16 are between interface residues (heptad positions a and d), 24 are between

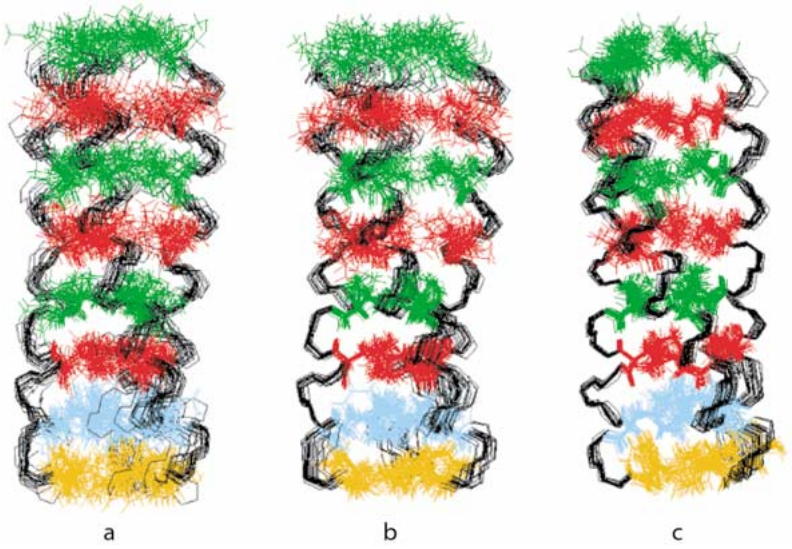

Figure 4. Sets of the 21 best structures from iterations 2, 5, and 8. The backbone of the coiled coil region V224-Y248 and all heavy atoms of the interface residues (positions *a* and *d*) are displayed. Side chains of valine residues are coloured green; leucine, red; phenylalanine, blue; tyrosine, yellow; the backbone, black. Figure prepared with MOLMOL (Koradi et al., 1996).

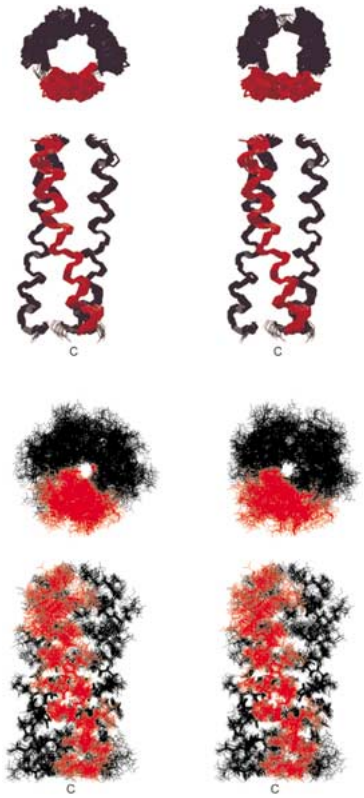

Figure 5. The solution structure of SP-D coiled coil represented by the 21 lowest energy conformers. The stereo figure was prepared with MOLMOL (Koradi et al., 1996). In the sideview the trimer is depicted with the N-terminus pointing up and the C-terminus pointing down. The cross section is depicted at the N-terminus, viewing down toward the C-terminus.

Table 2. Quality of the calculated structures

	21 best structures	<NMR>[b]
Violations of distance bounds (> 0.3 Å)[a]	0	0
Ramachandran analysis[c]		
Residues in the most favored regions	90.5%	
Additional allowed regions	7.2%	
Generously allowed regions	2.2%	
Disallowed regions	0.1%	
RMSD[d] (Å)		
Residues	Backbone N, C^{α}, and C'	All heavy atoms
V224-Y248	0.61 ± 0.21	0.99 ± 0.19
I heptad (V224-Q230)	0.46 ± 0.28	0.91 ± 0.33
II heptad (V231-Q237)	0.31 ± 0.09	0.84 ± 0.12
III heptad (V238-A244)	0.27 ± 0.12	0.64 ± 0.14
IV heptad (F245-Y248)	0.37 ± 0.09	0.96 ± 0.12

[a]Violated in 10 or more structures.
[b]The mean structure was obtained by averaging the coordinates of the 21 best structures after superposition of the N, C^{α}, C' backbone atoms in the structured regions comprising residues V224-Y248 and the side chain heavy atoms of the interface residues in heptad positions *a* and *d*.
[c]Determined by the program PROCHECK (Laskowski et al., 1993); including residues G220-I259.
[d]Root mean square deviation of the 21 best structures vs. the mean structure.

an interface and a neighbouring residue (heptad positions e and g), and 4 are outside the coiled coil region. Although inter-helical contacts occur along the entire coiled coil domain (see Figure 7), all 44 of the inter-monomer NOEs are found in the C-terminal half of the trimer. Hence, the packing of the aromatics appears to promote and stabilize inter-helical contacts. The 66 remaining ambiguous NOEs may be 'comonomer' (O'Donoghue et al., 1996), i.e., they contain contributions from intra- and inter-monomer NOEs.

The first heptad, and in particular the position of V224, is more disordered than the rest of the coiled coil. Interestingly, this valine is completely missing from the crystal coordinates. The individual layers of valines (*a* positions) and leucines (*d* positions) pack in the so called acute, or 'knobs into holes' manner, particularly in the best-defined third heptad (see Figure 8). In the fourth heptad, the *a* and *d* positions are occupied by the aromatic residues Phe and Tyr, respectively. These also pack in the acute manner, with the rings oriented parallel to the coil axis (Figure 8). The phenylalanines are completely buried, while the hydroxyl groups of the tyrosines are exposed to the solvent. The solvent accessible surface area for each residue in the coiled coil region (averaged over the ensemble) is less than 10% for the interface residues in position *a* and *d* (plus V251) and more than 20% for all others. Thus, the hydrophobic interface is very well buried. It is likely that towards the end of the coiled coil V251 is allowed to twist inward to a buried position due to the mobile unstructured C-terminus in the present peptide fragment. This does not necessarily occur in the complete protein, where the sequence continues into the CRD domain.

The ensemble superimposes closely onto the crystal structure, with a backbone RMSD of 0.69 Å over the last three heptads (the first heptad being only partially present in the crystal structure). Both structures have similar coiled coil parameters (cf. Table 3). The two structures have the same coiled coil radius and orientation angle of the *a* position. The solution structure of the SP-D coiled coil is, however, more extended than the SP-D crystal structure, indicated by the slightly higher number of residues per super-coil and higher super-coil pitch. The fit of the SP-D solution structure to an ideal coiled coil geometry over four heptad repeats is 0.80 Å, that of the crystal structure over three heptad repeats 0.59 Å. Thus, the crystal structure appears to be only slightly more regular than the solution structure.

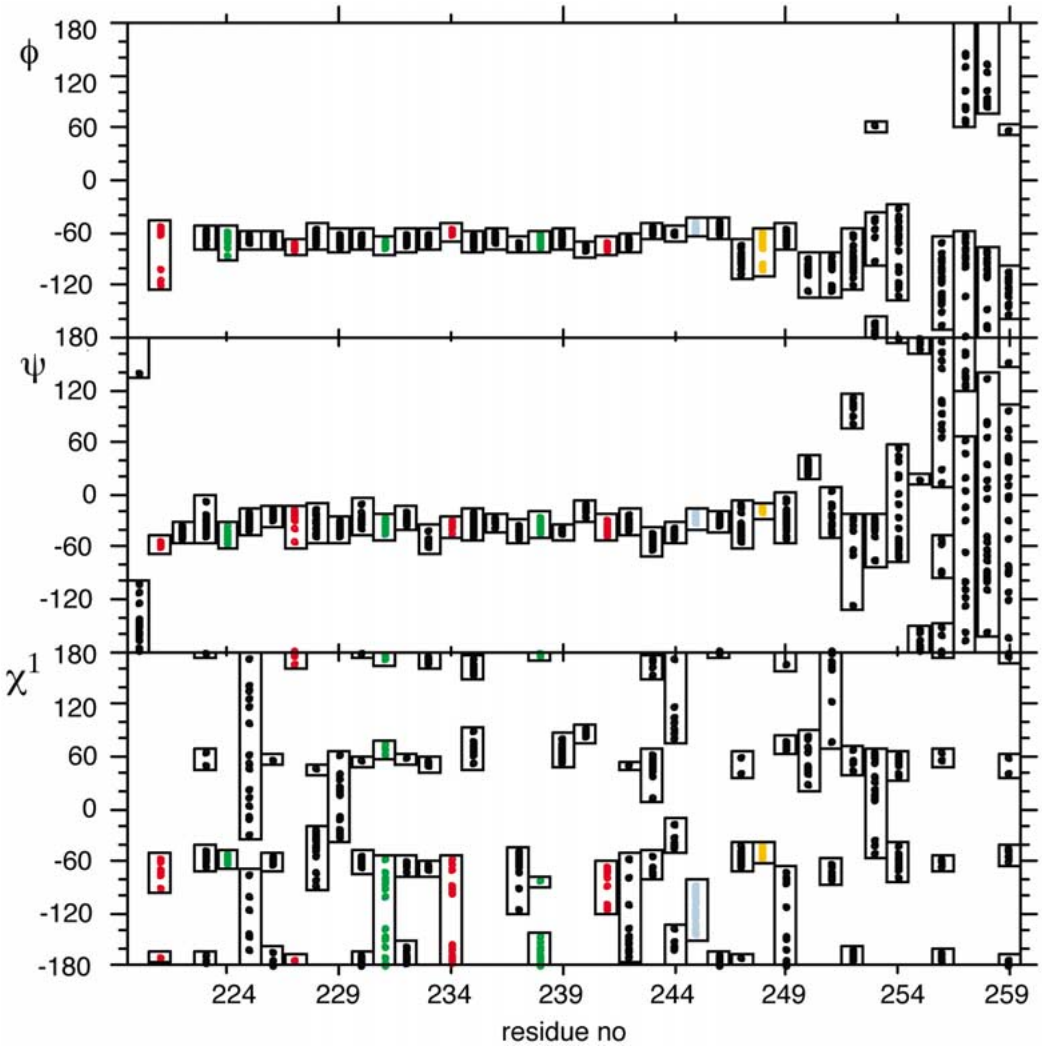

Figure 6. The distribution of the ϕ, ψ, and χ^1 dihedral angles in the final set of 21 structures. Each value is displayed as a point, and the complete range is indicated by a bar. The colour coding of the interface residues is the same as in Figure 3 (valine green; leucine red; phenylalanine blue; tyrosine yellow; backbone black).

The coiled coil parameters of the GCN4-V_aL_d trimer with a valine-leucine coiled coil interface (1coi, Ogihara et al., 1997) and the GCN4-pII trimer with a coiled coil interface consisting of isoleucines (1gcm, Harbury et al., 1994) are also presented in Table 3. The crystal structure of SP-D and the GCN4-V_aL_d are the most compact, indicated by the lowest number of residues per turn of the super-coil and the lowest pitch of the super-coil. The GCNA-pII trimer adopts loose packing in order to accommodate its all-isoleucines interface, characterized by a longer super-coil with a larger radius, and consequently a larger deviation from an ideal coiled coil trimer. The SP-D solution and crystal structures are intermediate cases between the other two trimer structures.

Discussion

Structure calculation

Calculation of a symmetric coiled coil trimer from NMR data is still challenging, particularly in the absence of X-filtered data. The simplest way to solve the problem would have been to use previously solved coiled coil trimers to assign the ambiguous NOEs. This approach has been used to solve some other

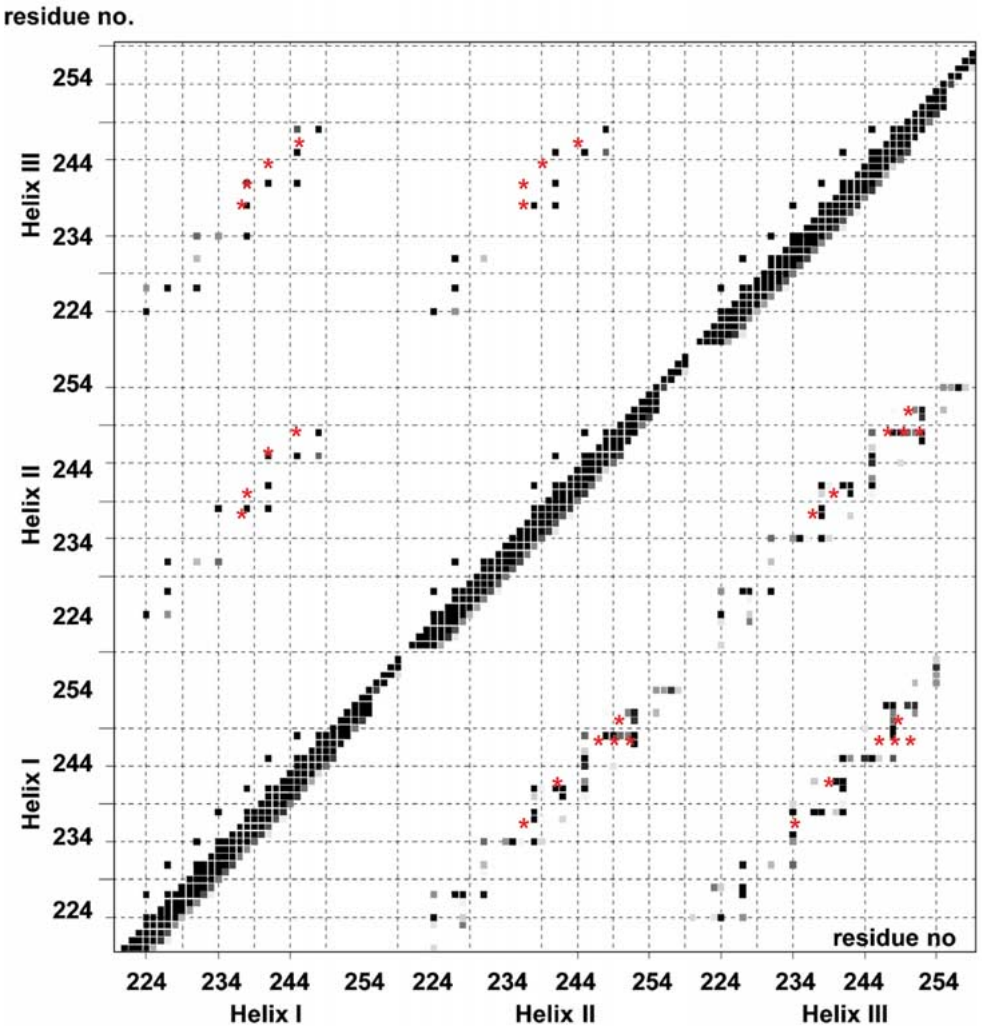

Figure 7. ^{1}H-^{1}H contact map with distances less than 4 Å shown in black and distances between 4 and 6 Å in fading grey-scale. The upper left half of the diagram displays side chain inter-helical contacts of the interface residues (heptad positions *a* and *d*) and the lower right half of the diagram shows all side chain contacts. The distance of each ^{1}H-^{1}H pair is calculated as an average over the distance in the 21 final structures. Figure prepared with MOLMOL (Koradi et al., 1996). The red stars indicate the NOE connectivities identified as predominantly inter-helical in the final iteration.

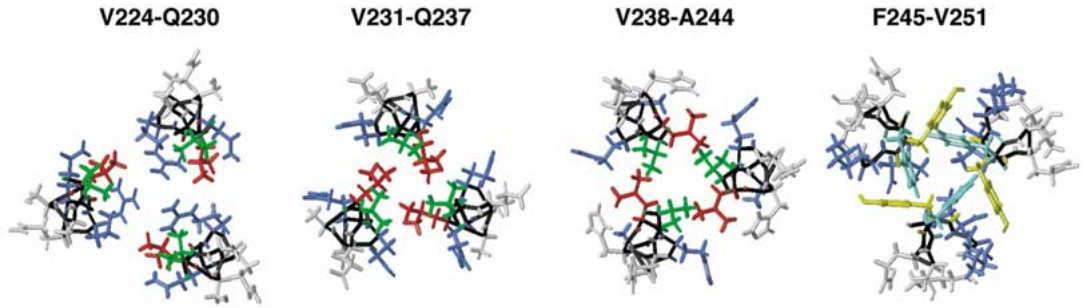

Figure 8. Packing of each heptad layer in the lowest energy structures. The following colour coding was used: backbone, black; valines, green; leucines, red; phenylalanine, cyan; tyrosine, yellow; residues in positions *e* and *g*, blue; residues in positions *b*, *c*, and *f*, grey. Figure prepared with MOLMOL (Koradi et al., 1996).

oligomers (see references in O'Donoghue and Nilges, 1999). The calculation from the ambiguous data with few assumptions proved to be rather sensitive to small errors in the input data or the calculation set-up. The entire iterative procedure had to be repeated many times to identify and remove these errors. We also found that we needed to make some general assumptions about the overall fold: The trimer consists of three parallel, interacting helices with a stagger less than one heptad. Such assumptions are used conservatively in the structure calculation, so that the detailed features of the final structure (the inter-helical separation, the coiling of the helices, the orientation of the residues to the coiled coil axis, and the precise degree or lack of stagger) are determined by the ambiguous NOE data and not by the additional assumptions. We solved the structure independently, since we assigned the ambiguous NOEs without any explicit reference to any previous structures, the final structure is independent. The solution ensemble we obtained is very close to the crystal structure, assessed by overall RMSD and in terms of the coiled coil parameters.

Comparison to the crystal structure

The only significant differences between the structures in crystal and solution are in the last heptad, which is asymmetric in the crystal structure. In one of the helices, the side-chain of the *d* position residue Tyr248 is buried in the centre of the helices with an orientation perpendicular to the symmetry axis; in the other two helices, the Tyr248 side-chains are in a different conformation, exposed, and hydrogen-bonded to water. Hakansson et al. (1999) suggested that the asymmetry seen in the crystal structure might be an artefact of crystallization. Alternatively, the asymmetry in the crystal structure may be induced by the presence of the CRD domain, absent in the peptide construct used here. The NMR spectra show no indication of multiple conformations (e.g., peak doubling).

However, the discrepancy between the crystal and solution studies could be reconciled if the two states observed in the crystal structure were exchanging rapidly on the NMR time scale. This occurs, for example, for the Asn residue in the central *a* position of the leucine zippers (MacKay et al., 1996; King, 1996). The two states in the SP-D crystal structure differ only in the conformation of a single Tyr side chain. However, the exchange between the two states would also involve considerable local rearrangement of the backbone. The symmetric orientation of the Tyr248 aromatic rings observed in the final ensemble would then be an incorrect consequence of one of our assumptions (symmetry of the trimer), which we incorporated into the calculation in the form of symmetry restraints; explicit assignment of NOEs in a symmetric fashion would obviously produce the same effect. The symmetric model is the simplest one consistent with the data. In particular, we observed no NOEs incompatible with the symmetric arrangement. In this case, the symmetry restraints could have been removed for the Tyr side chain. On the basis of the present structural data from X-ray crystallography and NMR, the question whether the coiled coil trimer is symmetric in solution or exchanging rapidly between different asymmetric conformations, cannot be resolved.

Table 3. Coiled coil parameters

	R_0 (Å)	$^1/\omega_0$	P(Å)	ϕ	RMSD(Å)
SP-D, solution	6.5	106	157	13°	0.80
SP-D, crystal	6.5	95	139	11°	0.59
GCN4-V_aL_d 6.5	78	112	17°	0.98	
GCN4-pII	6.7	120	179	19°	0.59

R_0, coiled-coil radius; $1/\omega_0$, residues per turn; P, pitch; ϕ, position *a* orientation angle; gMSD, measures how closely the structures match the ideal coiled coil. Solution SP-D, average structure from the final ensemble; crystal SP-D, 1bo8 (Hakansson et al., 1999); GCN4-V_aL_d, 1coi (Ogihara et al., 1997); GCN4-pII, 1gcm (Harbury et al., 1994).

Trimer specificity

The SP-D coiled coil domain appears to be very specific for a symmetric trimer, as no other coiled coil oligomer states (dimers, tetramers etc.) have been observed. Thus, this domain may serve as a good scaffold for the design of trimers and for introducing trimerization regions. What causes this stabilization of the trimer state over the other possible states? Part of the explanation may lie in the pattern of *a* and *d* position residues. In the first three heptads, the pattern is exclusively V_aL_d (i.e., valines at *a* positions and leucines at *d* positions). Several mutation studies of short coiled coils have suggested that such a pattern of branched hydrophobic residues favours stabilization of trimers (Betz et al., 1995; Harbury et al., 1993; Ogihara et al., 1997; Lovejoy et al., 1993). The V_aL_d pattern is shared by some other members of the collectin family (e.g., conglutinin and collectin-43), but not all of them.

In contrast, the final heptad in SP-D contains F_aY_d (i.e., phenylalanine at *a* position and tyrosine at *d* position). It is unusual for aromatic residues to occur at these positions in coiled coils. Interestingly, Gonzalez et al. (1996b) showed that the GCN4(N16A) mutant forms stable trimers when a non-covalently bound aromatic ring is added. The mutation of the *a* position Asn16 to Ala creates a cavity between the three helices, which can be occupied by a single non-covalently bound benzene molecule. The benzene ring was orientated perpendicular to the symmetry axis, similar to the position of the buried Tyr248 side-chain in the SP-D crystal structure. These structures suggest a possible mechanism for the stabilization of stable coiled coil trimers by aromatic rings.

Conclusions

In this paper, we have successfully used the symmetry-ADR method to solve the structure of a trimeric coiled coil. Only few general assumptions were necessary for the structure determination (such as the symmetry of the structure and the fact that the monomers interact), but no NOEs needed to be assigned explicitly. Calculations of this type are now also possible with an adapted version the current release of ARIA (version 1.2), which is available on request. The structure of the SP-D coiled-coil domain shows that this domain forms an autonomously folded domain. The structure differs from the independently solved X-ray crystal structure in the symmetry of one single residue. This might be due to artefacts in either the crystal structure or the NMR structure determination. This questions, which has important biological consequences for the orientation of the attached domains, cannot be resolved based on the current experimental evidence.

References

Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **5**, 1–10.

Betz, S.F., Bryson, J.W. and DeGrado, W.F (1995) *Curr. Opin. Struct. Biol.*, **5**, 457–463.

Brünger, A.T. (1993) *X-Plor Version 3.1*, Yale University Press, New Haven, U.S.A.

Caffrey, M., Cai, M., Kaufman, J., Stahl, S.J., Wingfield, P.T., Gronenborn, A.M. and Clore, G.M. (1997) *J. Mol. Biol.*, **271**, 819–826.

Chan, D.C., Fass, D., Berger, J.M. and Kim, P.S. (1997) *Cell*, **89**, 263–273.

Clore, G.M., Omichinski, J.G., Sakaguchi, K., Zambrano, N., Appella, E. and Gronenborn, A. (1995) *Science*, **267**, 1515–1516.

Crick, F.H.C. (1953) *Acta Crystallogr.*, **6**, 589–697.

Dames, S.A., Kammerer, R.A., Wiltscheck, R., Engel, J. and Alexandrescu, A.T. (1998) *Nat. Struct. Biol.*, **5**, 687–691.

Fass, D., Harrison, S.C. and Kim, P.S. (1996) *Nat. Struct. Biol.*, **3**, 465–469.

Folmer, R.H.A., Hilbers, C.W., Konings, R.N.H. and Nilges, M. (1997) *J. Biomol. NMR*, **9**, 245–258.

Gonzalez, L., Plecs, J.J. and Alber, T. (1996a) *Nat. Struct. Biol.*, **3**, 510–515.

Gonzalez, L., Brown, R.A., Richardson, D. and Alber, T. (1996b) *Nature Struct. Biol.*, **3**, 1002–1009.

Gordon-Smith D.J., Carbajo R.J., Yang J.C., Videler H., Runswick M.J., Walker J.E. and Neuhaus D. (2001). *J. Mol. Biol.*, **308**, 325–339.

Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.

Hakansson, K., Lim, N.K., Hoppe, H.-J. and Reid, K.B.M. (1999) *Structure*, **7**, 1–10.

Harbury, P.B., Kim, P.S. and Alber, T. (1994) *Nature*, **371**, 80–83.

Harbury, P.B., Zhang, T., Kim, P.S. and Alber, T. (1993) *Science*, **262**, 1401–1407.

Hidalgo P., Ansari A.Z., Schmidt P., Hare B., Simkovich N., Farrell S., Shin E.J., Ptashne M. and Wagner G. (2001) *Genes. Dev.*, **15**, 1007–1020.

Holmskov, U., Malhotra, R., Sim, R.B. and Jensenius, J.C. (1994) *Immunol. Today*, **15**, 67–73.

Hoppe, H.-J. and Reid, K.B.M. (1994) *Prot. Sci.*, **3**, 1143–1158.

Hoppe, H.-J., Barlow, P.N. and Reid, K.B.M. (1994) *FEBS Lett.*, **344**, 191–195.

Jasanoff, A. (1998). *J. Biomol. NMR*, **12**, 299–306.

Johansson, J., Curstedt, T. and Robertson, B. (1994) *Eur. Respir. J.*, **7**, 372–391.

Junius, F.K., O'Donoghue S.I., Nilges, M., Weiss, A.S. and King, G.F. (1996) *J. Biol. Chem.*, **271**, 13663–13667.

King, G.F. (1996) *Biophys. J.*, **71**, 1–2.

Kishore, U., Wang, J.-Y., Hoppe, H.-J. and Reid, K.B.M. (1996) *Biochem. J.*, **318**, 505–511.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graphics*, **14**, 51–55.

Kraulis, P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.

Landschultz, W.H., Johnson, P.F. and McKnight, S.L. (1988) *Science*, **240**, 1759–1764.

Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Moss, D.S., Kaptein, R. and Thornton, J.M. (1993) *J. Biomol. NMR*, **8**, 477–486.

Linge, J.P. and Nilges, M. (1999) *J. Biomol. NMR*, **13**, 51–59.

Lovejoy, B., Choe, S., Cascio, D., McRorie, D.K., DeGrado, W.F. and Eisenberg, D. (1993) *Science*, **259**, 1288–1293.

MacKay, J.P., Shaw, G.L. and King, G.F. (1996) *Biochemistry*, **35**, 4867–4877.

Miyamura, K., Leigh, L.E.A., Hopkin, J., Lopez-Bernal, A. and Reid, K.B.M. (1994) *Biochem. J.*, **300**, 237–242.

Nicholls, A. and Honig, B. (1991) *J. Comput. Chem.*, **12**, 435–445.

Nilges, M. (1993) *Proteins*, **17**, 297–309.

Nilges, M. and Brünger, A.T. (1991) *Protein Eng.*, **4**, 649–659.

Nilges, M. and O'Donoghue, S.I. (1998) *Progr. NMR Spectrosc.*, **32**, 107–139.

Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988a) *FEBS Lett.*, **229**, 317–324.

Nilges, M., Gronenborn, A.M., Brünger, A.T. and Clore, G.M. (1988b) *Protein Eng.*, **2**, 27–38.

Nilges, M., Macias, M., O'Donoghue S.I. and Oschkinat, H. (1997) *J. Mol. Biol.*, **269**, 408–422.

Nooren, I.M., Kaptein R., Sauer R.T. and Boelens R. (1999) *Nature Struct. Biol.*, **6**, 755–759.
O'Donoghue, S.I. and Nilges, M. (1999) In *Biological Magnetic Resonance*, Vol. 17, Berliner, J.L. and Krishna, N.R. (Eds.), Plenum, New York pp. 131–161.
O'Donoghue, S.I., Chang, X., Abseher, R., Nilges, M. and Led, J.J. (2000) *J. Biomol. NMR*, **16**, 93–108.
O'Donoghue, S.I., Junius, F.K. and King, G.F. (1993) *Protein Eng.*, **6**, 557–564.
O'Donoghue, S.I., King, G.F. and Nilges, M (1996) *J. Biomol. NMR*, **8**, 193–206.
Ogihara, N.L., Weiss, M.S., Degrado, W.F. and Eisenberg, D. (1997) *Protein Sci.*, **6**, 80–88.
Pascual, J., Pfuhl, M., Walther, D., Saraste, M. and Nilges, M. (1997) *J. Mol. Biol.*, **273**, 740–751.
Sheriff, S. , Chang, C.Y. and Ezekowitz, R.A.B. (1994) *Nat. Struct. Biol.*, **1**, 789–794.
Thiel, S. and Reid, K.B.M. (1989) *FEBS Lett.*, **250**, 78–84.
Weaver, T.E. and Whitsett, J.A. (1991) *Biochem. J.* **273**, 249–264.
Weis, W.I. and Drickamer, K. (1994) *Structure*, **2**, 1227–1240.
Wilson, I.A., Skehel, J.J. and Wiley, D.C. (1981) *Nature* **289**, 366–373.