Journal of Biomolecular NMR, **13:** 311–324, 1999.
KLUWER/ESCOM

A high quality nuclear magnetic resonance solution structure of peptide deformylase from *Escherichia coli*: Application of an automated assignment strategy using GARANT

John F. O'Connell*, KellyAnn D. Pryor, Stephan K. Grant & Barbara Leiting
Department of Biochemistry, Merck Research Laboratories, mail code RY80Y-103, P.O. Box 2000, Rahway, NJ 07065, U.S.A.

Received 3 November 1998; Accepted 17 December 1998

Key words: automated assignment, distance geometry, peptide deformylase, protein structure, structure refinement

Abstract

The NMR structure of the peptide deformylase (PDF) (1–150) from *Escherichia coli*, which is an essential enzyme that removes the formyl group from nascent polypeptides and represents a potential target for drug discovery, was determined using $^{15}N/^{13}C$ doubly labeled protein. Nearly completely automated assignment routines were employed to assign three-dimensional triple resonance, ^{15}N-resolved and ^{13}C-resolved NOESY spectra using the program GARANT. This assignment strategy, demonstrated on a 17 kDa protein, is a significant advance in the automation of NMR data assignment and structure determination that will accelerate future work. A total of 2302 conformational constraints were collected as input for the distance geometry program DYANA. After restrained energy minimization with the program X-PLOR the 20 best conformers characterize a high quality structure with an average of 0.43 Å for the root-mean-square deviation calculated from the backbone atoms N, C_α and C′, and 0.81 Å for all heavy atoms of the individual conformers relative to the mean coordinates for residues 1 to 150. The globular fold of PDF contains two α-helices comprising residues 25–40, 125–138, six β-strands 57–60, 70–77, 85–88, 98–101, 105–111, 117–123 and one 3_{10} helix comprising residues 49–51. The C-terminal helix contains the HEXXH motif positioning a zinc ligand in a similar fashion to other metalloproteases, with the third ligand being cysteine and the fourth presumably a water. The three-dimensional structure of PDF affords insight into the substrate recognition and specificity for N-formylated over N-acetylated substrates and is compared to other PDF structures.

Abbreviations: PDF, peptide deformylase; 2D, 3D, two-, three-dimensional; NOESY, NOE spectroscopy; COSY, correlation spectroscopy; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser effect; TOCSY, total correlation spectroscopy; PFG, pulse field gradient; ct, constant time; ppm, parts per million; $^3J_{HN\!H\alpha}$, vicinal spin-spin coupling constant between the amide proton and the α-proton; $^3J_{NH\beta}$, vicinal spin-spin coupling constant between the backbone amide nitrogen and one of the β-protons; H/D, hydrogen/deuterium; rmsd, root mean square distance.

Introduction

Bacterial protein synthesis is initiated with formylmethionine-tRNA which results in the synthesis of N-terminally formylated polypeptides (Meinnel et al., 1993). However, mature bacterial proteins do not retain the N-formyl group (Marcker and Sanger, 1964) and most do not retain the N-terminal methionine (Waller, 1963). Deformylation is part of the methionine cycle and is performed by the metallopeptidase peptide deformylase (PDF) and the maturation of the polypeptide completed by methionine aminopeptidase (MAP) (Meinnel et al., 1993). Deformylation of peptides after translation is required for bacterial survival

*To whom correspondence should be addressed. E-mail: oconnell@merck.com

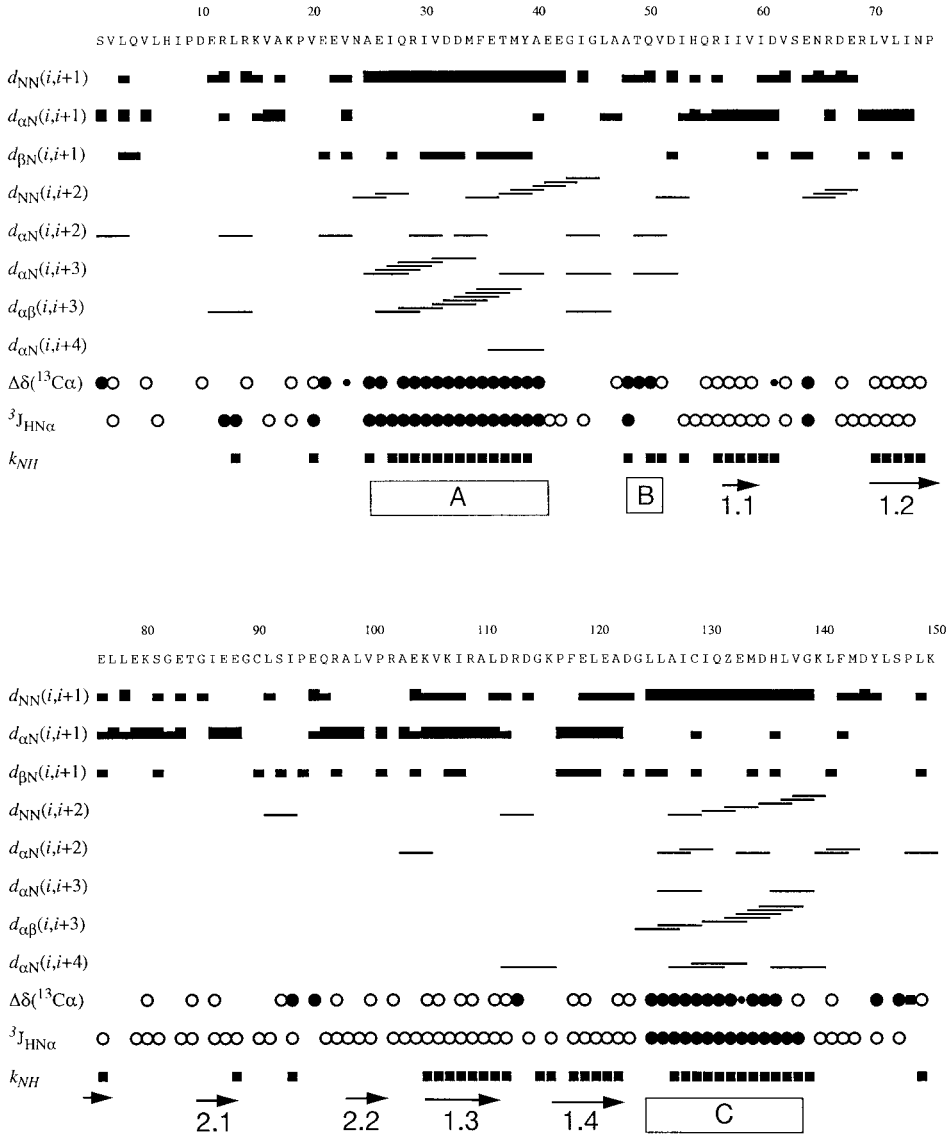

Figure 1. Amino acid sequence of PDF and survey of the sequential connectivities and additional data collected for secondary structure identification. The sequential NOE connectivities $d_{\alpha N}$, d_{NN} and $d_{\beta N}$ are indicated with thick, medium or thin black bars for strong, medium and weak NOEs, respectively. Medium-range connectivities $d_{NN}(i, i+2)$, $d_{\alpha N}(i, i+2)$, $d_{\alpha N}(i, i+3)$, $d_{\alpha\beta}(i, i+3)$ and $d_{\alpha N}(i, i+4)$ are shown by lines starting and ending at the positions of the residues related by the NOE. In the row $\Delta\delta(^{13}C_\alpha) = \delta(^{13}C_\alpha)_{obs} - \delta(^{13}C_\alpha)_{rc}$, where $\delta(^{13}C_\alpha)_{obs}$ and $\delta(^{13}C_\alpha)_{rc}$ denote the observed and random coil chemical shifts, large filled, small filled and large open circles indicate residues with $\Delta\delta(^{13}C_\alpha) > 2.0$ ppm, $\Delta\delta(^{13}C_\alpha) > 1.5$ ppm and $\Delta\delta(^{13}C_\alpha) < -1.5$ ppm, respectively. In the row $^{3}J_{HNH\alpha}$, filled and open circles denote residues with $^{3}J_{HNH\alpha} < 5.5$ Hz and $^{3}J_{HNH\alpha} > 8.0$ Hz, respectively. In the row k_{NH} filled squares identify residues with sufficiently slow amide exchange rates to enable observation of the ^{15}N-^{1}H cross-peak in a [^{15}N-^{1}H]-COSY spectrum recorded 24 h after dissolving the protein in D_2O at 45 °C. The sequence locations of regular secondary structure elements are indicated at the bottom, with A, B and C indicating the α-helices and 3_{10} turn and 1.1, 1.2, 1.3, 1.4, 2.1, 2.2 and 2.3 the β-strands and their connections.

and is apparently unique to bacterial cells (Mazel et al., 1994). Thus, inhibition of PDF presents an attractive target for a new class of antibiotics.

The strictly conserved HEXXH motif present in all known bacterial deformylases (Meinnel et al., 1995) is characteristic of zinc metalloproteases (Vallee and Auld, 1990). However, since PDF from *E. coli* and other bacteria have little resemblance to other known proteins, it represents a new family of metalloproteases with interesting catalytic properties (Meinnel et al., 1995). It has been shown that the two histidines of the HEXXH motif and Cys^{90} are ligands for the metal ion (Meinnel et al., 1995). An important feature of PDF is that it is selective for N-formylated substrates over acetylated substrates (Meinnel and Blanquet, 1995). In addition, PDF shows a strong preference for methionine and the non-naturally occurring norleucine side-chains in substrates and aldehyde inhibitors (Grant, S., personal communication). A detailed tertiary structure of the enzyme may provide insight into the substrate selectivity of PDF, its functionality and P1 binding of the enzyme in the region of the zinc and an S1 subsite. The NMR studies in this work were performed on the catalytic core composed of 150 residues. The C-terminal 18 residues have previously been shown to be dispensable for PDF activity and were disordered in solution (Meinnel et al., 1996). Indeed, the C-terminal residue chemical shifts were observed at random coil frequencies and a truncated form (1–150) was prepared and studied. While this work was in progress, the low resolution structure of a truncated PDF (1–147) was solved by NMR (Meinnel et al., 1996b), a new structure was deposited in the data base 2DEF and a full length form (1–170) was solved by X-ray crystallography (Chan et al., 1997; Becker et al., 1998). A comparison of these structures is discussed.

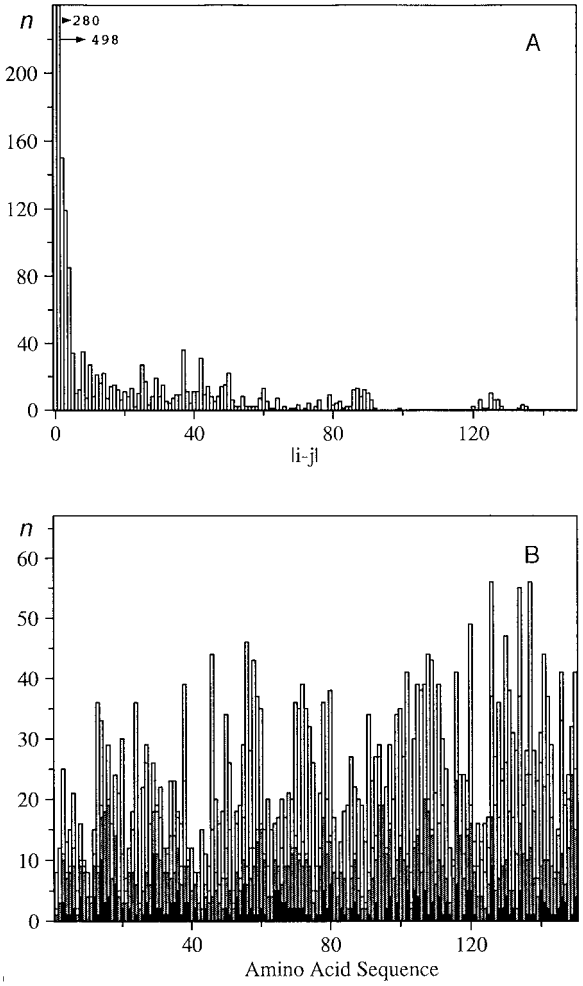

Figure 2. (A) Plot of the number of NOE distance constraints, n, versus their range along the amino acid sequence. The entries for the intra-residue constraints and constraints between protons in sequentially neighboring residues would go off-scale, the numbers at the top indicate their actual heights. Along the horizontal axis $|i - j|$ is the sequence separation of the two residues containing the protons considered. (B) Plot of the number of NOE distance constraints per residue, *n*, versus the amino acid sequence of PDF. The constraints are specified as follows: black, intra-residue; dark grey, constraints between protons in sequentially neighboring residues; light grey, constraints between protons located in residues separated by 2 to 5 positions along the sequence; white, all longer-range constraints.

Methods

Protein expression and purification

The PDF gene encoding the N-terminal 150 residues was cloned from genomic DNA of *E. coli* strain K37. The gene was amplified by PCR using primers 5′-CGCATATGTCAGTTTTGCAAGTGTTACA-3′ and 5′-TAGTCGACTTATTTCAGCGGTGACAGATA-3′, and cloned into the *Nde*I/*Sal*I sites of pET30a(+) (Novagen). The sequence was confirmed by DNA sequencing and matched the published PDF sequence (EMBL access code X7780). The protein was produced in minimal media (Pryor and Leiting, 1997) supplemented with 348 μM $ZnSO_4$ and 0.5 mM IPTG for 20 h at 18 °C. Cells were resuspended in 10 mM Tris/HCl pH 7.4 and lysed by french press (20K cell from SLM Aminco, 1200 PSI, single run at 4 °C). PDF was purified from the soluble fraction in two steps, first by ion exchange chromatography (Frac-

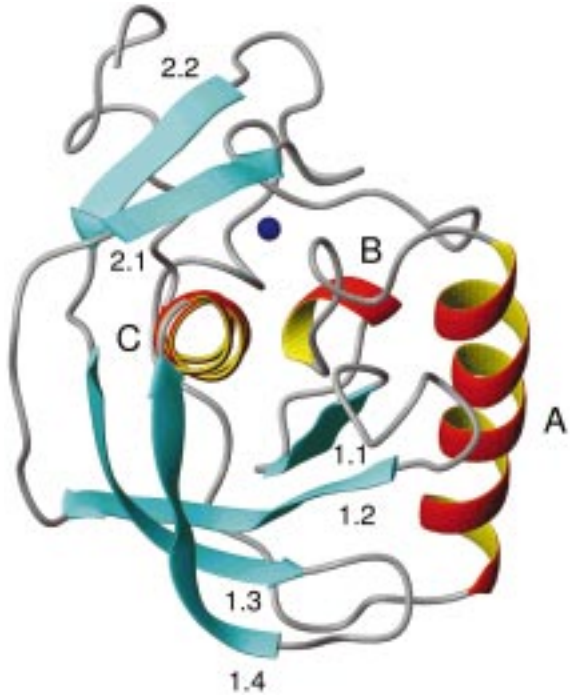

Figure 3. Representation showing the two α-helices and the β-sheets of PDF and their location using the MOLMOL (Koradi et al., 1996) program.

togel EMD DEAE 650 (S), EM Scientific) and second by size exclusion chromatography (superdex 75, Pharmacia).

NMR sample preparation

The purified protein was concentrated on an anion exchange column using 20 mM potassium phosphate buffer at pH 6.7 and potassium chloride to elute the protein. After dialysis against 20 mM potassium phosphate buffer at pH 6.8 and further concentration in a Biomax 5k device (Millipore), the NMR measurements were made with 1.5 mM solutions of PDF in 99.9% D_2O or a mixture of H_2O/D_2O (90:10 v/v) at pH 6.7. For the D_2O samples, the protein was lyophilized repeatedly from D_2O.

NMR spectroscopy

All NMR spectra were recorded at 30.0 °C on a Varian 600 MHz Unity or 600 MHz Inova system with triple resonance PFG (^{1}H, ^{15}N, ^{13}C) probes using either uniformly ^{13}C/^{15}N doubly labeled PDF, uniformly ^{15}N-labeled PDF or unlabeled PDF. Quadrature detection in the indirect dimensions was achieved using States-TPPI (Marion et al., 1989). Spectra were processed with NMRPipe (Delaglio et al., 1995) using forward-backward linear prediction and zero filling in each indirectly detected dimension using an IBM SP2 computer. The processed data were analyzed with XEASY (Bartels et al., 1995) for visualization of NMR data, peak-picking and peak-integration using an IBM SP2 computer. In addition, the program NMRVIEW (Johnson and Blevins, 1995) was also used for data visualization.

Resonance assignments and the input for the structure calculation were obtained from gradient-enhanced (Muhandiram and Kay, 1994) versions of the following experiments. Using ^{13}C/^{15}N-labeled PDF: 3D CBCANH (Grzesiek and Bax, 1992a). Time domain data size 48 × 32 × 1024 complex points, $t_{1,max}$(^{13}C) 7.8 ms, $t_{2,max}$ (^{15}N) 14.5 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 3D CBCA(CO)NH (Grzesiek and Bax, 1992b). Time domain data size 48 × 32 × 1024 complex points, $t_{1,max}$ (^{13}C) 7.8 ms, $t_{2,max}$(^{15}N) 14.5 ms, $t_{3,max}$(^{1}H) 128.0 ms. 3D HNCA (Grzesiek and Bax, 1992a; Yamakazi et al., 1994). Time domain data size 40 × 24 × 1024 complex points, $t_{1,max}$ (^{13}C) 9.1 ms, $t_{2,max}$ (^{15}N) 10.9 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 3D HN(CO)CA (Grzesiek and Bax, 1992a). Time domain data size 40 × 24 × 1024 complex points, $t_{1,max}$ (^{13}C) 9.1 ms, $t_{2,max}$ (^{15}N) 10.9 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 3D HBHA(CO)NH (Grzesiek and Bax, 1993). Time domain data size 64 × 32 × 1024 complex points, $t_{1,max}$ (^{1}H) 12.8 ms, $t_{2,max}$ (^{15}N) 16.0 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 3D HNCO (Grzesiek and Bax, 1992a). Time domain data size 32 × 32 × 1024 complex points, $t_{1,max}$ (^{13}C) 14.5 ms, $t_{2,max}$ (^{15}N) 14.5 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 3D C(CO)NH (Grzesiek et al, 1993). Time domain data size 48 × 24 × 1024 complex points, $t_{1,max}$ (^{13}C) 4.6 ms, $t_{2,max}$ (^{15}N) 10.9 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 3D HC(CO)NH (Grzesiek et al, 1993). Time domain data size 48 × 24 × 1024 complex points, $t_{1,max}$ (^{13}C) 6.1 ms, $t_{2,max}$ (^{15}N) 10.9 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 3D HCCH_TOCSY (Bax et al., 1990; Sattler et al., 1995) Time domain data size 50 × 64 × 1024 complex points, $t_{1,max}$ (^{1}H) 6.25 ms, $t_{2,max}$ (^{13}C) 5.33 ms, $t_{3,max}$ (^{1}H) 128.0 ms. 2D [^{13}C, ^{1}H]-COSY (Bodenhausen and Ruben, 1980; Kay et al., 1992). Time domain data size 196 × 1024 complex points, $t_{1,max}$ (^{13}C) 10.0 ms, $t_{2,max}$ (^{1}H) 128.0 ms. 2D ct-[^{13}C, ^{1}H]-COSY (Vuister and Bax, 1992). Time domain data size 196 × 1024 complex points, $t_{1,max}$ (^{13}C) 10.0 ms, $t_{2,max}$ (^{1}H) 128.0 ms. 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY (Ikura et al., 1990). Time domain data size 96 × 24 × 1024 complex points, $t_{1,max}$ (^{13}C)

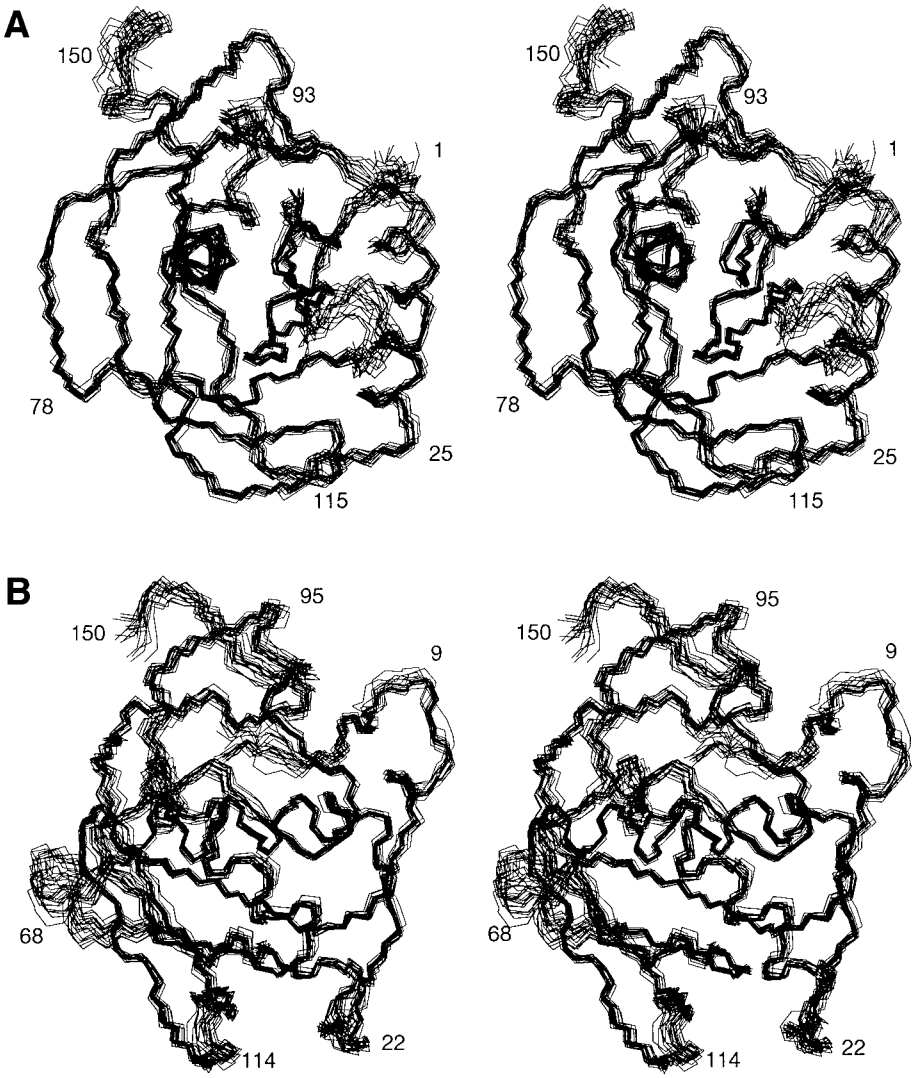

Figure 4. Stereo view of the polypeptide backbone of the final 20 energy refined DYANA conformers of PDF used to represent the NMR structure in solution. The conformers were superimposed for pairwise minimum rmsd of the backbone atoms N, C_α and C' of residues 1–150. Some positions have been identified with the residue number of the sequence location. (A) Front view. (B) Side view after rotation by 90° about the vertical axis relative to the front view.

10.0 ms, $t_{2,max}$ (^{15}N) 20.0 ms, $t_{3,max}$ (^{1}H) 128 ms. 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY (Fesik and Zuiderweg, 1988). Time domain data size 96 × 24 × 1024 complex points, $t_{1,max}$ (^{13}C) 10.0 ms, $t_{2,max}$ (^{15}N) 20.0 ms, $t_{3,max}$ (^{1}H) 128 ms. 3D ^{15}N-resolved [^{1}H,^{1}H]-TOCSY (Fesik and Zuiderweg, 1988). Time domain data size 96 × 24 × 1024 complex points, $t_{1,max}$ (^{13}C) 10.0 ms, $t_{2,max}$ (^{15}N) 20.0 ms, $t_{3,max}$ (^{1}H) 128 ms. 2D [^{15}N, ^{1}H]-COSY (Bodenhausen and Ruben, 1980). Time domain data size 196 × 1024 complex points, $t_{1,max}$ (^{13}C) 10.0 ms, $t_{2,max}$ (^{1}H) 128.0 ms.

The carrier position was set to 4.74 ppm for ^{1}H, 114.3 ppm for ^{15}N, 55.5 ppm, 43 ppm or 175 ppm for ^{13} C_α, $^{13}C_{aliph}$ or $^{13}C'$ carbon atoms. The ^{1}H chemical shifts are relative to internal 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS). The ^{15}N and ^{13}C chemical shifts are relative to DSS using the conversion factors that have been reported by Wishart et al. (1995).

Vicinal $^{3}J_{HNH\alpha}$ coupling constants were determined by inverse Fourier transformation of in-phase multiplets (Szyperski et al., 1992) from a 2D [^{15}N, ^{1}H]-COSY spectrum recorded with a data size of 256

$\times$ 1024 complex points, $t_{1,max}$ (^{15}N) 10.0 ms, $t_{2,max}$ (^{1}H) 128.0 ms. Vicinal $^{3}J_{NH\beta}$ scalar coupling constants were estimated from a 3D ct-HNHB spectrum (Archer et al., 1991) with a time domain data size of 96 $\times$ 24 $\times$ 1024 complex points, $t_{1,max}$ (^{1}H) 10.0 ms, $t_{2,max}$ (^{15}N) 20.0 ms, $t_{3,max}$ (^{1}H) 128.0 ms.

For the estimation of the amide proton exchange rates, a 1.6 mM solution of fully protonated, uniformly ^{13}C/^{15}N-labeled, PDF was lyophilized. The protein was redissolved in D_2O and a 2D [^{15}N, ^{1}H]-COSY was recorded with a time domain data size of 128 $\times$ 1024 complex points, $t_{1,max}$ (^{15}N) 10.0 ms, $t_{2,max}$ (^{1}H) 128.0 ms. The sample was then placed in a 45 °C bath for 24 hr and the 2D [^{15}N, ^{1}H]-COSY was recorded. Those amide protons that remained were considered to have exchange rates $\ll 1$ s^{-1}.

Assignment strategy

The GARANT version 2.0 calculations were run on an IBM SP2 with eight 591 processors, modified to include the C(CO)NH and HC(CO)NH 3D experiments. Input consisted of the 12 XEASY peak-picked 3D spectra as defined in the NMR methods and the 2D [^{15}N, ^{1}H]-COSY. The population size for the evolutionary genetic algorithm of GARANT was typically 100 to 150 using the standard optimization macro. Structure based NOESY probabilities were set to 1.0, 0.5 and 0.25 for distances less than 3.0 Å in 20/20 structures, 4.0 Å in 15/20 structures and 5.0 Å in 10/20 structures, respectively. Chemical shift tolerances were set to 0.3 ppm for heteronuclear dimensions, 0.05 ppm and 0.03 ppm for indirect and direct detected proton dimensions, respectively. Later GARANT calculations included defined secondary shifts for the C′, C_α and C_β carbons for well defined secondary structure residues. Typical calculation times were between 100 and 200 h.

Structure calculation and refinement

The 20 structures with the lowest DYANA (v. 1.4; Güntert et al., 1997) target functions resulting from calculations with the final input data set starting with 100 random structures using the torsion-space molecular dynamics annealing were subjected to restrained energy minimization with the programs X-PLOR (Brünger, 1988) version 3.843 using the CHARM22 force field (MacKerell et al., 1998) and with FANTOM (Schauman et al., 1990; Braun, 1992; von Freyberg et al., 1993) version 3.2. The DYANA and X-PLOR calculations were run on an eight-processor

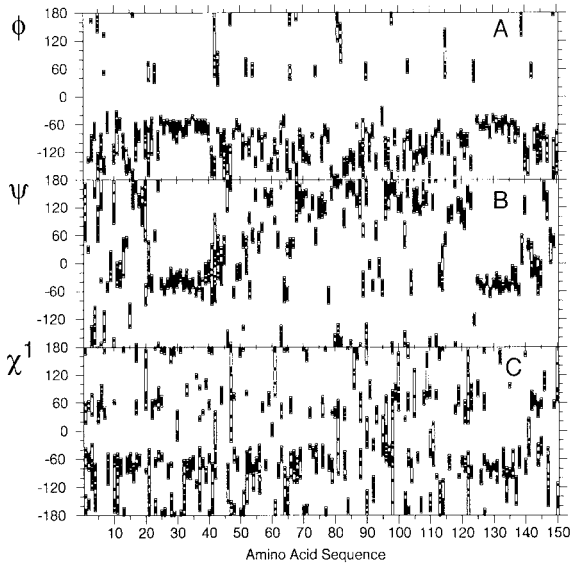

Figure 5. In plots versus the sequence, the range of dihedral angles ϕ (A), Ψ (B) and χ^1 (C) in the 20 X-PLOR(FANTOM) energy refined DYANA solutions of PDF are shown. The value for each of the NMR conformers is represented by a point.

Cray T916-8256 computer and the FANTOM calculations on an IBM SP2. In both minimization programs the pseudo-energy was proportional to the sixth power of the distance constraint violations, and was adjusted such that violations of 0.15 Å for the distance constraints and 2.5° for the dihedral angle constraints corresponded to 1/2kBT at room temperature (Billeter et al., 1990). During the minimization, a cutoff of 10.0 Å was used for pairs of interacting atoms in the energy evaluation. The resulting 20 X-PLOR(FANTOM) energy-minimized DYANA conformers are used to represent the solution conformation of PDF. All ϕ, Ψ backbone torsion angles within the ordered parts of the protein lie within the allowed regions of the Ramachandran plot. The Ramachandran values of the complete protein for all comformers is 64% in the most favored regions, 29% in the additionally allowed, 5% in the generously allowed and 2% in the disallowed regions. The structure coordinates and constraint data will be deposited in the PDB and BMRB data bases.

Structure analysis

For visual comparison of the structures, stereo views were produced with either the molecular graphics program MOLMOL (Koradi et al., 1996) or with the structure analysis program XAM (Xia, 1992;

O'Connell and Xia, unpublished) on an IBM RS6000 workstation. For pairs of conformers, global superpositions and rmsd values for various subsets of atoms were computed (McLachlan, 1979). The mean solution conformation was obtained by first superimposing the 20 energy minimized DYANA conformers so as to minimize the rmsd for the backbone atoms N, C_α, and C' of residues 1–150, and then averaging the Cartesian coordinates of the corresponding atoms in the 20 globally superimposed conformers. Displacements, D (Billeter et al., 1989), were used to quantify the local precision of the solution structure (Figure 6). In the group of conformers used to represent the solution structure, D was obtained as the mean standard deviation for the atom positions of N, C_α and C' after global superposition of the individual conformers with the average coordinates for the backbone residues.

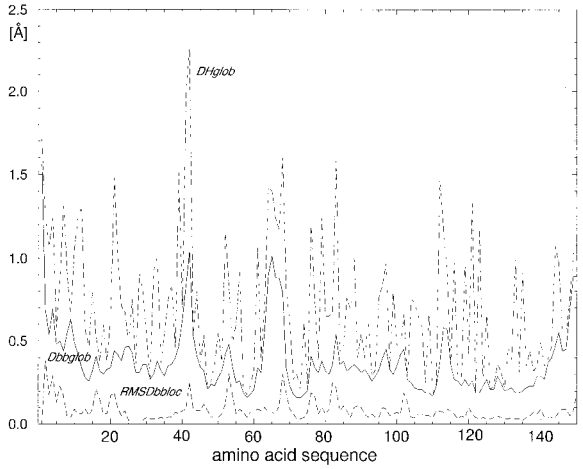

Figure 6. In plots versus the amino acid sequence, the mean global backbone displacements per residue, D^{bb}_{glob}, of the 20 energy-minimized DYANA coordinates are given relative to the mean NMR structure calculated after superposition of the backbone heavy atoms N, C_α and C' for minimal rmsd (solid line). The dashed line represents the mean global backbone displacement per residue, D^{h}_{glob}. The long dashed line represents the mean local rms deviation $rmsd^{bb}_{glob}$, calculated for the backbone superposition of all tripeptide segments along the sequence onto the mean NMR structure, with the rmsd values for the tripeptide segments plotted at the positions of the central residues.

Results and discussion

Assignment of resonances and NOESY spectra

The spectral assignments and cross-validation of the sequential assignments were facilitated by using the program GARANT (Bartels et al., 1996, 1997). GARANT is a very powerful and flexible approach to the automated assignment of NMR data. It produces lists of expected peaks based on the primary structure, the magnetization transfer pathways in the spectra used and, if available, short distances present in a set of preliminary structures. The algorithm assigns each expected peak in the model a probability of locating the corresponding peak in the measured data. It then optimizes 'expected' versus experimentally 'observed' peaks using a combination of an evolutionary genetic algorithm and a local optimization routine (Bartels et al., 1997). GARANT can utilize information other than the chemical shifts and coherance pathways to generate a list of 'expected' peaks such as distances derived from tertiary structures or secondary structure information. These features allow for a very robust approach in producing both sequential and NOE assignments simultaneously. GARANT does not depend on the identification of matching chemical shifts in a defined set of spectra. Although matching assignments throughout several spectra score higher in the assignment routine it allows for NOESY assignments without complete chemical shift assignments. This offers a solution for the problem of noise inherent in experimental peak lists and promotes successful assignments of regions with low signal to noise ratios. GARANT can search through a practically unlimited number of spectra simultaneously which is manually very time consuming. One can either conservatively initiate sequential assignments with a partial set of spectra first and add additional data later or use all input data at once. The assignment progress can also be controlled by a systematic modification of the GARANT search parameters while observing the convergence of the assignments. Using GARANT in cycles and feeding it preliminary structures from previous assignments rapidly drives an NMR structure calculation towards both complete sequential and NOE assignments and high resolution structures simultaneously. The PDF structure determination presented in this manuscript is the first application of GARANT using multiple triple resonance spectra and NOESY spectra to obtain chemical shift and NOE assignments simultaneously. It is also the first application of GARANT/DYANA cycles to derive complete assignments and a high resolution structure simultaneously in an automated fashion with interactive validation analysis.

The overall strategy used for PDF is summarized in Table 1 which lists the spectra used as different stages

Table 1. Outline of the GARANT/DYANA assignment procedure

Step I.	**Initial resonance assignments**
	GARANT using 3D spectra set A,
	interactive control
Step II.	**Collection of constraints for initial structure calculations**
	3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY + spectra set A
	interactive control
Step III.	**Refinement of resonance assignments**
	upper limit constraints,
	structures
	GARANT+DYANA using spectra set A and B
	coupling constants,
	interactive analysis of constraint violations-cycle
Spectra set A	
HNCO, HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH, ^{15}N-resolved[^{1}H,^{1}H]-TOCSY, ^{15}N-resolved [^{1}H,^{1}H]-NOESY, HBHA(CO)NH, C(CO)NH, HC(CO)NH, 2D [^{15}N-^{1}H]-COSY, 2D [^{13}C-^{1}H]-COSY	
Spectra set B	
^{13}C-resolved [^{1}H,^{1}H]-NOESY, HCCH-TOCSY.	

This outline is a simplification, e.g. coupling constants were measured later and therefore used in the later stages of the refinement (step III). In addition, initially a subset of spectra set A was used during the development stage.

to finally obtain a complete set of ^{1}H, ^{15}N and ^{13}C chemical shifts. Among the labile side-chain protons, all amide groups of the 3 Asn and the 6 Gln residues were assigned as well as the ϵ-protons resonances of 6 out of 10 Arg residues. Some representative automated assignments were cross-validated with manual assignments using the standard strategy (Wüthrich, 1986; Powers et al., 1992). Success with GARANT was greatly dependent on the quality of the peak-picking. The program XEASY (Bartels et al., 1995) provided excellent results with 3D data due to the very robust peak-picking algorithm and greatly reduced the time to prepare the input peak lists as compared to other programs. Multiple GARANT assignments were run with different random seeds and compared until they were consistent between calculations. This convergence of assignments improved as more spectral data were added.

Initial structures from Step II (Table 1) were characterized by rmsd values in the 2.5 Å to 2.9 Å range for the backbone residues 1 to 150. At this stage GARANT NOESY probabilities which function as search criteria (Bartels et al., 1996, 1997) were set to 1.0 and 0.1 for distances less than 3.0 Å in 20 structures and 4.0 Å in 5/20 structures, respectively. However, further development and testing of the GARANT procedures showed more complete NOESY assignments when the parameters described in the Methods section were used which afforded 50% more NOESY assignments and improved the overall fit of derived structures to the experimental data. After including the 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectra and several rounds of GARANT and DYANA calculations the final assignments were made using preliminary, but well defined, DYANA structures with low target functions using all spectral data (Step III, Table 1).

Assignment and structure calculation time line

Excluding the time for the collection of the NMR spectra, the initial sequential assignments and preliminary assignment of the ^{15}N-resolved [^{1}H,^{1}H]-NOESY to produce the initial structures took approximately 6 weeks. This involved both GARANT calculations and partial manual sequential and NOEs assignments for comparison and initial structures. Including the ^{13}C-resolved [^{1}H,^{1}H]-NOESY data refined the structures to approximately 1.0 Å rmsd and took approximately 4 weeks. Inclusion of the scalar coupling constants and complete incorporation of the triple resonance NMR spectra in the GARANT calculations took approximately 4 weeks. At this stage inspection of the structures indicated a twist in the catalytic helix that was due to miss assignment of the C-terminal residue

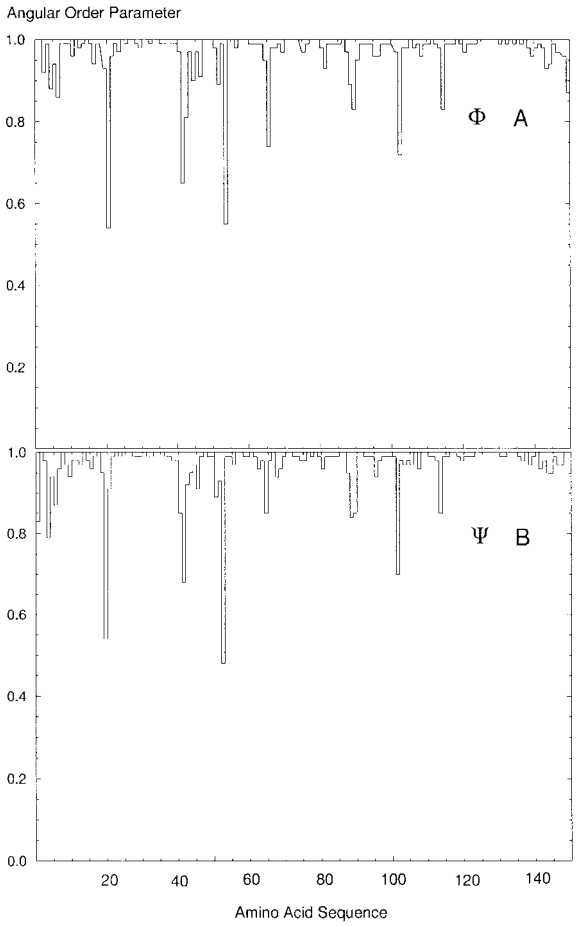

Figure 7. In plots versus the sequence, the backbone dihedral angular order parameters (Hyberts et al., 1992) for ϕ (A) and Ψ (B) are shown. Standard deviations of a torsion angle of 10, 20, 40 and 80° correspond to angular order parameters of 0.985, 0.942, 0.768, and 0.397, respectively.

142 long range NOEs to β-sheet residue 82. This was not noted during the violation analysis since few other long range NOEs were involved in the C-terminal residues that could not be satisfied. Correction of these distance constraints using the more through GARANT NOESY probabilities described in the Methods section produced the final structures. This last step was performed over an 8 week period involving a large number of GARANT/DYANA calculation cycles to test the methods and to obtain nearly complete assignments of the 3D NOESY data. The total time for this PDF structure determination was less than one year including the installation and testing of the programs GARANT, DYANA, XEASY to explore the robustness of the described strategy with respect to the quantity and quality of input data and GARANT parameters. During this time this strategy using GARANT was also applied to, and validated with structure determinations of three polypeptides of about 10 kDa each which are to be published elsewhere. Using these methods and procedure outlined in Table 1, one could probably solve an equivalent structure on the order of three months. Other structure determinations and further development of this strategy are in progress.

Collection of conformational constraints

The input for the distance geometry calculations with the program DYANA consisted of upper distance limits derived from NOESY cross-peak intensities with the program CALIBA (Güntert et al., 1991a,b), and dihedral angle constraints obtained using the program HABAS (Güntert et al., 1989, 1991a,b) from an initial interpretation of the spin-spin coupling constants $^3J_{HNH\alpha}$ and $^3J_{NH\beta}$, together with backbone dihedral angle constraints derived from conformation-dependent $^{13}C_\alpha$ chemical shifts (Luginbühl et al., 1995). HABAS also provided a number of stereospecific assignments of β-methylene protons. Upper and lower distance constraints were included for the zinc-nitrogen and zinc-sulfur bonds with distances derived from ab initio calculations (GAUSSIAN 94; Frisch et al., 1995). No additional constraints were used to enforce hydrogen bonds implicated by NOEs or by amide proton exchange data were used at any time during the DYANA structure calculations nor restrained energy minimizations. The program GLOMSA was also used (Güntert et al., 1991a,b) to obtain stereospecific assignments of both additional β-methylene protons and other pairs of diastereotopic substituents.

Structure determination

Following the procedures described, an input of NOE distance constraints for PDF was collected from the 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY in H_2O and 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY in D_2O recorded with a mixing time, τ_m, of 100 ms. For the calibration of ^{1}H-^{1}H upper distance limits, r, versus the NOESY cross peak intensities with the program CALIBA (Güntert et al., 1991a,b) as implemented in DYANA, a $1/r^6$ dependence was used for backbone proton-backbone proton NOEs, and a $1/r^4$ dependence for NOEs involving side chain protons and methyl groups using the standard protein calibration macro. These calibration curves were refined based on plots of cross peak volume versus average proton-proton distances in sets of preliminary structures. The proton-proton upper

distance bounds before application of pseudoatom corrections were in all cases confined to the range 2.4–5.0 Å in order to obtain reasonable upper bounds for both strong and weak NOEs. A total of 2570 NOESY cross-peaks were assigned and used for the generation of the input of upper-limit distance constraints for the structure calculation. Of these, 1119 resulted from the 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY and 1451 from 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY. In addition, a total of 246 vicinal scalar coupling constants were determined, including 117 $^3J_{HNH\alpha}$ and 129 $^3J_{NH\beta}$ coupling constants. The catalytic zinc was constrained by five upper distance constraints with a zinc containing histidine residue for the DYANA calculations and bond force constants in the energy minimization.

A total of 2205 NOE distance constraints were collected. Of these, 245 constraints were found to be irrelevant on the basis that they are either independent of conformation or that there exists no conformation that would violate the constraint. From the HABAS interpretation of the backbone-backbone intra-residual and sequential NOE upper distance limits together with the spin-spin coupling constants and $^{13}C_\alpha$ chemical shift indices, a total of 130 ϕ and 130 Ψ backbone dihedral angle constraints were generated, along with 82 side-chain χ^1 dihedral angle constraints. These stereospecific assignments were then used for the preparation of the final input which included 1960 NOE upper distance constraints and 342 dihedral angle constraints. The range and sequence distributions of the NOE distance constraints are shown in Figure 2. The NOE constraints obtained may be classified into four classes with $|j-i|=0$, $|j-i|=1$, $2 \leq |j-i| \leq 5$, $|j-i|>5$, where j and i are the residue positions in the sequence that contain the two protons for which the NOE is observed. The numbers of distance constraints in these four classes are, respectively, 280, 498, 387, 795. The large number of long range constraints is expected for a protein with a high content of β-sheets (Wagner and Wüthrich, 1982).

Structure refinement

The 20 conformers with the lowest target function values resulting from the final DYANA calculation were subjected to restrained energy-minimization with the programs X-PLOR and FANTOM. An analysis of the structures before and after restrained minimization is afforded by Table 2. The best 20 DYANA structures before energy-minimization have low target function values, and satisfy the NOE distance constraints and dihedral angle constraints nearly perfectly. Energy-minimization with the program FANTOM located low energy conformations very near the DYANA molecular geometries, with only a moderate increase in the sum of constraint violations (not listed in Table 2). Since FANTOM works in dihedral angle space and X-PLOR in Cartesian space and the two programs use different force fields, subsequent X-PLOR minimization of the FANTOM-refined structures afforded yet lower energies (Table 2), which are also much lower than those obtained by direct X-PLOR minimization (not listed).

The 20 best DYANA conformers after energy minimization with FANTOM and X-PLOR are used to represent the solution conformation of PDF. The X-PLOR CHARM H22 energies range from 1149 to 1237 kcal/mol with van der Waals energies in the range -201 to -164, which is a reduction of about 3000 kcal relative to the unrefined DYANA conformers, respectively (Table 2). In the set of structures, all residual violation of upper distance limits were less than 0.33 Å, with not more than three violations per structure exceeding 0.20 Å. All dihedral angle violations were less than 4.0° (Table 2).

A visual impression of the solution structure and the quality of the structure determination is afforded in Figure 4, which shows the superposition of the polypeptide backbone in the final 20 representation conformers. The global fold is well defined, as indicated by the average pair-wise rmsd values between the individual conformers and their mean of 0.43 Å for all backbone atoms and 0.83 Å for all heavy atoms for residues 1 to 150.

Structure of PDF and enzyme substrate selectivity

A survey of the sequential NOEs and additional data used to derive the secondary structure of PDF is shown in Figure 1. There is a close correlation of the secondary structure with the sequential NOEs, $^{13}C_\alpha$ chemical shift indices and scalar $^3J_{HNH\alpha}$ coupling constants as well as slow H/D exchanging backbone amides. A large number of backbone amides have very slow exchange rates that do not exchange at 45 °C for 48 h, most of those do not exchange after several months at 30 °C. This is indicative of the tertiary stability of PDF. This was also observed with circular dichroism spectroscopy in the thermal denaturation of PDF which demonstrated a melting point of >78 °C (Leiting, B., personal communication). In addition, the stability of the tertiary structure is evident

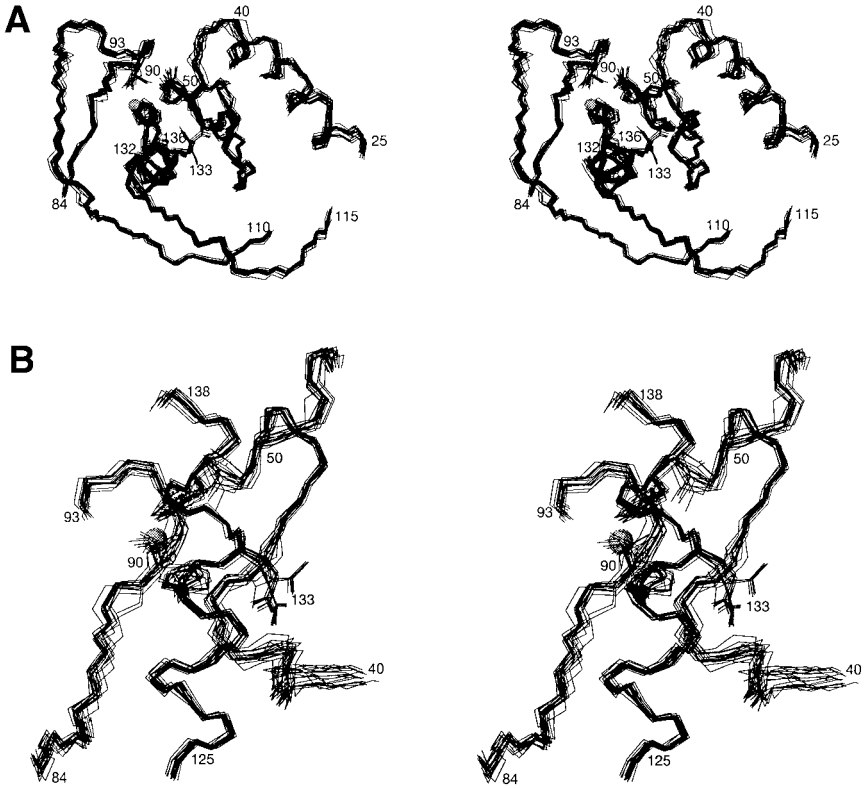

Figure 8. Stereo views of selected structural elements of PDF. The amino acid side chains are shown for selected residues with selected residue numbers. The stereo views were generated by superposition for minimal rmsd of the backbone N, C_α and C' atoms of the amino acid residues 1–150.

in the large number of long range NOE constraints (Figure 2). A large number of hydrogen bonds were located in the calculated structures. A MOLMOL ribbon diagram of the PDF structure (1–150) is shown in Figure 3. The secondary structure assignments determined using MOLMOL and the NMR data are as follows: α-helix, (A) 25–40 and (C) 125–138; 3–10 helix (B) 49–51; β-strands (1.1) 57–60, (1.2) 70–77, (2.1) 84–88, (2.2) 98–102, (1.3) 105–111 and (1.4) 117–123. Figure 5 shows that the range of observed dihedrals in the structure is highly restricted for most of the residues.

The active site comprises the C-terminal α-helix, the β-strands 2.1 and 2.2 as well as the residues 40–48 which extend up along the helix and near Cys^{90} (Figure 8). The C-terminal helix (C) contains the conserved HEXXH motif, which comprises the zinc binding site along with residue Cys^{90}. In chemical shift perturbation experiments using 2D [^{15}N,^{1}H]-COSY spectra with several PDF inhibitors identified in our laboratory the consistent perturbation of these active site residues was observed (data not shown, will be published separately). These observations are also consistent with the shift perturbations observed in PDF in the presence of excess substrate (Meinnel et al., 1996b).

Residues 43–48 as they appear in the present NMR structure are likely to comprise a P1 binding pocket that rises slightly above the catalytic helix. The two glycine residues in the loop and its hydrophobic nature could be used to bind straight, hydrophobic side-chains that the enzyme prefers and confer the observed substrate selectivity of PDF. It has been noted that residues Leu^{46} and Leu^{91} form a cleft (Figure 9) that perhaps forms a physical barrier against larger substrates such as the methyl group in an acetyl group (Chan et al., 1997). This region also contains conserved residues from different bacteria. This may account for the selectivity of PDF for formyl substrates and not acetylated N-termini nor peptide bonds.

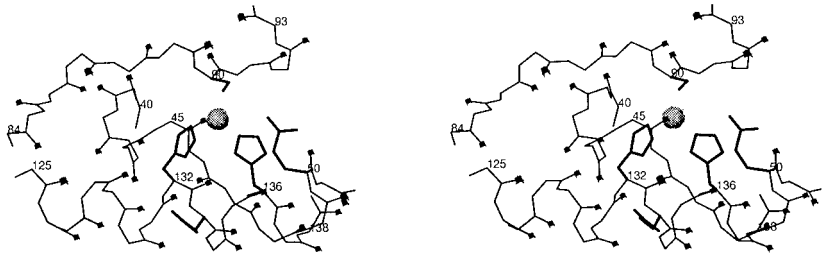

Figure 9. Stereo view of the active site residues of PDF. The amino acids involved in the zinc coordination and selected side-chains of residues near the active site are shown. The side-chains are shown with thicker lines and the backbone oxygen atoms are shown as a small ball using one of the representation conformers.

Table 2. Analysis of the 20 best DYANA conformers of PDF after restrained energy minimization with the program XPLOR

Quantity	Average value ± standard deviation (range)			
	Before energy minimization		After energy minimization	
DYANA target function (Å^2)[a]	8.12 ± 0.91	(5.79 .. 9.89)		
XPLOR energy (kcal/mol)			1183 ± 6	(1149 .. 1237)
XPLOR VDW energy (kcal/mol)			−188 ± 3	(−201 .. −164)
Residual NOE distance constraint violations				
Sum (Å)	33.1 ± 2.2	(29.2 .. 39.3)	92.9 ± 2.62	(88.9 .. 97.2)
Maximum (Å)	0.55 ± 0.16	(0.34 .. 0.82)	0.31 ± 0.01	(0.30 .. 0.34)
rmsd (Å)	0.05 ± 0.01	(0.04 .. 0.05)	0.01 ± 0.001	(0.01 .. 0.01)
Residual dihedral angle constraint violations				
Sum (°)	97.6 ± 9.8	(61.7 .. 98.4)	32.2 ± 4.20	(24.6 .. 40.6)
Maximum (°)	8.8 ± 0.2	(4.5 .. 18.7)	2.6 ± 0.57	(1.9 .. 4.6)
rmsd (°)	0.91 ± 0.15	(0.7 .. 1.30)	0.34 ± 0.04	(0.26 .. 0.42)

[a]The final structure calculation was started with 100 randomized conformers. The 20 DYANA conformers with the lowest residual target function values were refined by energy minimization and are used to represent the NMR structure.

Structure comparison

While this work was in progress Meinnel et al. (1996b) reported the solution structure of PDF (2–147) and Chan et al. (1997) as well as Becker et al. (1998) presented the crystal structure (1–168), all of which have different N- and C-termini as compared to this work (1–150). The coordinates for the Meinnel NMR structure are available from the Protein Data Bank (PDB) (code 1DEF). Subsequently, the same group deposited a new set of coordinates (code 2DEF) and later an accompanied publication (Dardel et al., 1998). The coordinates from the Chan et al. (code 1DFF) were also recently released which comprise unusually high B-factors. Therefore, the comparison presented here focuses on the more well defined regions. The overall topology of all PDF structures is very similar and the C-termini defined by residues 100–140 are almost identical. It was noted by Chan et al. (1997) that the original solution coordinates 1DEF differ significantly in the active site, particular in the positioning of Cys^{90} and the 90's loop. It appears that this is corrected in their newly refined solution structure 2DEF. Changes between the structures occur primarily in surface exposed loop regions and most notably at the N- and C-termini due to actual sequence differences. The greatest rmsd difference between all structures occurs in the N-terminal domain, especially residues 1–6 while the same helical turn is found in residues 12–14 and an α-helix formed by residues 25–40. 2DEF starts with a Val residue while the present structure was obtained on PDF with its native N-terminal Ser residue. The Val^1 side chain in 2DEF interacts with Ala^{38} while Ser^1 in the present work is positioned close to Asp^{41}. Preliminary unrestrained, solvated molecular dynamics calculations (publication in progress) indicate a stabilizing hydrogen bond between Ser^1 and Asp^{41} that may explain the observed NOE differences. Another difference at the N-terminus is the assignment of a *cis*-Pro in 2DEF and Becker et al. while 1DFF and the present work contain a *trans*-Pro. In DYANA test calculations using the same upper distance con-

straints either with a *cis*- or *trans*-Pro both isomers were tolerated, though the *cis* resulted in higher target functions. A *cis*-Pro^{9} violates a number of sequential and medium range upper distance constraints in the presented restraint set and even when the consistently violated restraints are removed the target functions for the *cis* derived structures are still twice as high. The resulting structural differences are very local and it should be pointed out that the overall number of NOEs in this turn is small compared to the rest of the structure (Figure 2) resulting in higher local rmsd values and lower angular order parameters (Figures 6 and 7).

Other structural differences between these PDF structures occur in loop regions such as the 40's loop which contains a β-strand for residues 44–48 in 2DEF and residues 45–47 in 1DFF which lie against β-strand-1.1. In this PDF structure residues Ile^{44} and Gly^{45} are positioned closer to the catalytic zinc (Figure 9) and consequently none of these residues are in a stable β-conformation. The careful analysis of slow H/D exchange, $^{13}C_{\alpha}$ chemical shift indices and $^{3}J_{HNH\alpha}$ coupling constants did not exhibit characteristic values consistent with a β-strand. The NOE data also places these residues too far from the 1.1 β-strand of residues 57–60 to form hydrogen bonds although residues 46–47 do occur in an extended conformation. The backbone carbonyl of Gly^{45} points into the active site (Figure 9) and this conformation has been implicated in hydrogen-bonding the substrate amide nitrogen as noted in thermolysin (Chan et al., 1997). In addition, residues 45–49 comprise one of the conserved sequence elements amongst PDF sequences from different bacteria. They have been shown to participate in substrate binding (Meinnel et al., 1996b) and interact with inhibitors of PDF (unpublished) and are therefore more likely to be located close to the active site metal ion as described in the present structure. Further differences in the loop regions are the 60's loop composed of residues 63–69. Residues 62–68 are also described to be flexible in the crystal structure (Becker et al., 1998). This is a solvent exposed loop that is defined by few long range NOEs, with residues 64–68 containing no long range NOEs. Small changes in the sequential NOEs could thus have a large effect on the conformation and subsequent differences in the superposition of these residues. Though these residues adopt a similar conformation in all sets of structures, they appear to be displaced relative to one another. The 2DEF coordinates do not have the corresponding restraint data deposited, but it was noted that 86 distance constraints for 43 hydrogen bonds, at unspecified positions, were used in the structure calculation. Since no hydrogen bond constraints were used to calculate the representation structures shown here, this could account for some of the differences.

Further differences are the length of β-sheets 1.2, 2.1 and 2.2 which extend in the crystal structure from 70–81, 84–90 and 93–102, respectively. However, similar as in case of the 40's loop, residues 78–81 are implicated to be in an extended conformation (Figure 1), but do not form consistent hydrogen bonds reflected by H/D exchange. Between 2DEF and the present structure the end of β-sheets 1.3 and 1.4 formed by residues Arg^{113} and Asp^{114} are slightly displaced. The hydrogen-bonding patterns in residues 84–90 and 98–102 in which 2DEF does not have a β-sheet and the amide of residue 88 form a hydrogen-bond with residue 86 instead of residue 98. There are also some differences in the C-termini which may be accounted for by the different length of the proteins studied 1..150 versus 2..147. Overall, the rmsd between the two solution structures is for the well defined regions containing the secondary structure residues (10..60, 70..140; 120 residues, 80%) is 1.79 Å.

Conclusions

The well defined, high quality solution structure of PDF(1–150) was determined using 15.5 constraints per residue. The structure was solved using a number of 3D NMR data and the assignment of the spectral data was greatly assisted and accelerated through the use of the program GARANT. The use of GARANT/DYANA/violation analysis solved the structure of the 17 kDa PDF protein in less 20 weeks of spectral and structure analysis effort through the use of the methodology developed and described here. Application and further development of automated assignment strategies using GARANT is in progress.

Acknowledgements

We thank Christian Bartels for access to the program GARANT, many useful discussions and for careful reading of the manuscript.

References

Archer, S.J., Ikura, M., Torchia, D.A. and Bax, A. (1991) *J. Magn. Reson.*, **95**, 636–641.

Bax, A., Clore, M. and Gronenborn, A.M. (1990) *J. Magn. Reson.*, **88**, 425–431.

Bartels, C., Xia, T.H., Billeter, M. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1–10.

Bartels, C., Billeter, M., Güntert, P. and Wüthrich, K. (1996) *J. Biomol. NMR*, **7**, 207–213.

Bartels, C., Güntert, P., Billeter, M. and Wüthrich, K. (1997) *J. Comput. Chem.*, **18**, 139–149.

Becker, A., Schlichting, I., Kabsch, W., Schultz, S. and Volker Wagner, A.F. (1998) *J. Biol. Chem.*, **273**, 11413–11416.

Billeter, M., Schaumann, T., Braun, W. and Wüthrich, K. (1990) *Biopolymers*, **29**, 695–706.

Billeter, M., Braun, W. and Wüthrich, K. (1989) *J. Mol. Biol.*, **206**, 677–687.

Bodenhausen, G. and Ruben, D. (1980) *Chem. Phys. Lett.*, **69**, 185–188.

Braun, W. (1987) *Biopolymers*, **26**, 1691–1704.

Brünger, A.T. (1988) *J. Mol. Biol.*, **203**, 803–816.

Chan, M.K., Gong, W., Rajogopalan, P.T.R., Hao, B., Tsai, C.M. and Pei, D. (1997) *Biochemistry*, **36**, 13904–13909.

Dardel, F., Ragusa, S., Lazennec, C., Blanquet, S. and Meinnel, T. (1998) *J. Mol. Biol.*, **280**, 501–513 (PDB entry 2DEF).

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Fesik, S.W. and Zuiderweg, E.R.P. (1988) *J. Magn. Reson.*, **78**, 588–593.

Frisch, M.J., Trucks, G.W., Schlegel, H.B., Gill, P.M.W., Johnson, B.G., Robb, M.A., Cheeseman, J.R., Keith, T., Petersson, G.A., Montgomery, J.A., Raghavachari, K., Al-Laham, M.A., Zakrzewski, V.G., Ortiz, J.V., Foresman, J.B., Cioslowski, J., Stefanov, B.B., Nanayakkara, A., Challacombe, M., Peng, C.Y., Ayala, P.Y., Chen, W., Wong, M.W., Andres, J.L., Replogle, E.S., Gomperts, R., Martin, R.L., Fox, D.J., Binkley, J.S., Dfrees, D.J., Baker, J., Stewart, J.P., Head-Gordon, M., Gonzalez, C. and Pople, J.A. (1995) *Gaussian94*, Gaussian, Inc., Pittsburgh, PA.

Grzesiek, S. and Bax, A. (1992a) *J. Magn. Reson.*, **99**, 201–207.

Grzesiek, S. and Bax, A. (1992b) *J. Am. Chem. Soc.*, **114**, 6291–6298.

Grzesiek, S. and Bax, A. (1992c) *J. Magn. Reson.*, **96**, 432–438.

Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 185–204.

Grzesiek, S., Anglister, J. and Bax, A. (1993) *J. Magn. Reson.*, **B101**, 114–119.

Güntert, P., Braun, W., Billeter, M. and Wüthrich, K. (1989) *J. Am. Chem. Soc.*, **111**, 3997–4004.

Güntert, P., Braun, W. and Wüthrich, K. (1991a) *J. Mol. Biol.*, **217**, 517–530.

Güntert, P., Qian, Y.Q., Otting, G., Müller, M., Gehring, W.J. and Wüthrich, K. (1991b) *J. Mol. Biol.*, **217**, 531–540.

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

Hyberts, S.G., Goldberg, M.S., Havel, T.F. and Wagner, G. (1992) *Protein Sci.*, **1**, 736–751.

Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **86**, 204–209.

Johnson, B.A. and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.

Kay, L., Keifer, P. and Saarinen, T. (1992) *J. Am. Chem. Soc.*, **114**, 10663–10665.

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

Luginbühl, P., Szyperski, T. and Wüthrich, K. (1995) *J. Magn. Reson.*, **B109**, 229–233.

MacKerell, A.D., Bashford, D., Bellott, M., Dundrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., McCarthy, J.D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reirer, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Kuczera, W.J., Yin, D. and Karplus, M. (1998) *J. Phys. Chem.*, **102**, 3586–3616.

Marcker, K. and Sanger, F. (1964) *J. Mol. Biol.*, **8**, 835–840.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.

Mazel, D., Pochet, S. and Marliere, P. (1994) *EMBO J.*, **13**, 914–923.

McLachlan, A.D. (1979) *J. Mol. Biol.*, **128**, 49–79.

Meinnel, T., Mechulam, Y. and Blanquet, S. (1993) *Biochimie*, **75**, 1061–1075.

Meinnel, T. and Blanquet, S. (1995) *J. Bacteriol.*, **177**, 1883–1887.

Meinnel, T., Lazennec, C., Cardel, F., Schmitter, J.M. and Blanquet, S. (1996a) *FEBS Lett.*, **385**, 91–95.

Meinnel, T., Blanquet, S. and Dardel, F. (1996b) *J. Mol. Biol.*, **262**, 375–386.

Muhandiram, D.R. and Kay, L.E. (1994) *J. Magn. Reson.*, **B103**, 203–216.

Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1992) *Biochemistry*, **31**, 4334–4346.

Pryor, K.D. and Leiting, B. (1997) *Protein Express. Purif.*, **10**, 309–319.

Sattler, M., Schwendinger, M.G., Schleucher, J. and Griesinger, C. (1995) *J. Biomol. NMR*, **6**, 11–22.

Schaumann, T., Braun, W. and Wüthrich, K. (1990) *Biopolymers*, **29**, 679–694.

Szyperski, T., Güntert, P., Otting, G. and Wüthrich, K. (1992) *J. Magn. Reson.*, **99**, 552–560.

Vallee, B.L. and Auld, D.S. (1990) *Biochemistry*, **29**, 5647–5659.

Vuister, G.W. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.

von Freyberg, B., Richmond, T.J. and Braun, W. (1993) *J. Mol. Biol.*, **233**, 275–292.

Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 347–366.

Waller, J.P. (1963) *J. Mol. Biol.*, **7**, 483–496.

Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, H.J.D., Markley, J.L. and Sykes, B.C. (1995) *J. Biomol. NMR*, **6**, 135–140.

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

Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994) *J. Am. Chem. Soc.*, **116**, 11655–11666.

Xia, T.H. (1992) Ph.D. Thesis, ETH-Zürich.