A Single-Point Mutation in the Extreme Heat- and Pressure-Resistant Sso7d Protein from *Sulfolobus solfataricus* Leads to a Major Rearrangement of the Hydrophobic Core[†,‡]

Roberto Consonni,*[,§] Laura Santomo,[§] Paola Fusi,[||] Paolo Tortora,[||] and Lucia Zetta[§]

Istituto di Chimica delle Macromolecole, Lab. NMR, CNR, Via Ampère 56, 20131 Milano, Italy, and Dipartimento di BioTecnologie e Bioscienze, Università di Milano−Bicocca, Pza delle Scienze 2, 20126 Milano, Italy

Received May 17, 1999; Revised Manuscript Received July 12, 1999

ABSTRACT: Sso7d is a basic 7-kDa DNA-binding protein from *Sulfolobus solfataricus*, also endowed with ribonuclease activity. The protein consists of a double-stranded antiparallel β-sheet, onto which an orthogonal triple-stranded antiparallel β-sheet is packed, and of a small helical stretch at the C-terminus. Furthermore, the two β-sheets enclose an aromatic cluster displaying a fishbone geometry. We previously cloned the Sso7d-encoding gene, expressed it in *Escherichia coli*, and produced several single-point mutants, either of residues located in the hydrophobic core or of Trp23, which is exposed to the solvent and plays a major role in DNA binding. The mutation F31A was dramatically destabilizing, with a loss in thermo- and piezostabilities by at least 27 K and 10 kbar, respectively. Here, we report the solution structure of the F31A mutant, which was determined by NMR spectroscopy using 744 distance constraints obtained from analysis of multidimensional spectra in conjunction with simulated annealing protocols. The most remarkable finding is the change in orientation of the Trp23 side chain, which in the wild type is completely exposed to the solvent, whereas in the mutant is largely buried in the aromatic cluster. This prevents the formation of a cavity in the hydrophobic core of the mutant, which would arise in the absence of structural rearrangements. We found additional changes produced by the mutation, notably a strong distortion in the β-sheets with loss in several hydrogen bonds, increased flexibility of some stretches of the backbone, and some local strains. On one hand, these features may justify the dramatic destabilization provoked by the mutation; on the other hand, they highlight the crucial role of the hydrophobic core in protein stability. To the best of our knowledge, no similar rearrangement has been so far described as a result of a single-point mutation.

Archaebacteria and eubacteria contain abundant small, basic proteins which are believed to play histonelike roles, such as DNA packing, maintenance, and control (*1*−*4*). Among archaebacteria, several histonelike proteins from *Sulfolobus* strains have been identified and grouped into three molecular mass classes of 7, 8, and 10 kDa (*5*−*9*). In particular, several members from the 7 kDa class were identified and characterized; in *S. acidocaldarius* they were referred to as Sac7a, Sac7b, Sac7c, Sac7d, and Sac7e, in order of increasing basicity (*7*, *10*). Furthermore, a homologous member from *S. solfataricus* was referred to as Sso7d (*8*). All these proteins were shown to be closely related to one another, with a sequence similarity of at least 65%. Particularly, Sso7d consists of 62 residues, has deduced pI and M_r of 10.2 and 7020, respectively, and is devoid of histidines, asparagines and cysteines.

Searching for ribonucleases from *S. solfataricus*, we subsequently identified three proteins which we named P1, P2, and P3 (*11*). Surprisingly, the primary structure of P2 was then shown to be identical to that of Sso7d, with the exception of an additional lysine at the N-terminus of Sso7d. Thus, we considered the two proteins as the same molecule and adopted the denomination Sso7d (*12*) to conform the more generally accepted nomenclature. P3 was shown to consist of 75 residues, and interestingly, its sequence was identical to that of P2 over a length of 57 residues starting from the N-terminus, whereas the C-terminus stretches completely diverged (*13*).

The structure of Sso7d from its natural source was solved by NMR[1] by Baumann and co-workers (*14*). They showed that the protein is folded into a compact globular unit consisting of a double-stranded antiparallel β-sheet, onto which an orthogonal triple-stranded antiparallel β-sheet is

[†] Work supported by a grant from the Italian Ministry for University and Scientific and Technological Research (PRIN, biocatalysis and bioconversions) and by the scientific Foundation "Antonio De Marco".

[‡] PDB file name: 1B4O.

* To whom correspondence should be addressed. Phone: +39/02/70643555. Fax: +39/02/70643557. E-mail: roberto@labnmr.icmnmr.mi.cnr.it.

[§] CNR.

[||] Università di Milano−Bicocca.

[1] Abbreviations: NMR, nuclear magnetic resonance; COSY, two-dimensional correlated spectroscopy; DQF, double quantum filtered; TOCSY, two-dimensional total correlation spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhanced spectroscopy; SA, simulated annealing; RMSD, root-mean-square deviation; photo-CIDNP, photochemically induced dynamic nuclear polarization; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight.

10.1021/bi9911280 CCC: $18.00
Published on Web 09/03/1999

packed, and of a small helical stretch at the C-terminus. Furthermore, the protein displayed a largely hydrophilic surface with many basic residues (i.e., 2 arginines and 13 lysines) and a very compact hydrophobic core made up by side chains at the interface of the two β-sheets, in particular the aromatics Phe5, Phe31, and Tyr33. In contrast, Trp23 was completely exposed to the solvent and was shown to play a major role in DNA binding (*15*, *16*). Monomethylation of several lysine residues at the N- and C-termini was also found (*14*).

In our effort to further characterize Sso7d, we subsequently cloned a synthetic gene encoding this protein and expressed it in *Escherichia coli* (*17*). The recombinant protein was indistinguishable from its natural counterpart on the basis of heat stability and RNA digestion pattern. Its structure, as solved by NMR spectroscopy (1JIC PDB code), showed an overall folding substantially identical to that of the natural form, with marginal differences in the more mobile regions, such as loops and helix; however, no lysine monomethylation was found (Consonni, R., unpublished results).

To gain more insight into the structural features responsible for the thermodynamic stability of Sso7d, we recently produced three single-point mutants, namely F31A, F31Y, and W23A, and assessed the changes in thermo- and piezostabilities resulting from these amino acid substitutions (*12*, *18*, *19*). In particular, Phe31 mutants were selected on the basis of structural data which indicated that this residue is located at the core of the aromatic cluster and has tight contacts with the side chains of several residues in the cluster, i.e., Phe5, Tyr33, Val14, Ile19, and Ile29 (*18*).

Furthermore, the cluster displayed a fish-bone geometry also found in several proteins and believed to play a major stabilizing role in these molecules (*20*). In keeping with our expectations, the single-point mutation F31A led to an unprecedented loss in thermo- and piezostabilities, i.e., by at least 10 kbar and 27 K, respectively (*18*).

This finding raised the question of what structural changes resulting from the single-point mutation of the protein are associated with its dramatic destabilization. Preliminary structural characterization of the F31A mutant did not allow us to evidence considerable changes in its overall structure (*18*). Here, we report the three-dimensional solution structure of this mutant, obtained from analysis of multidimensional NMR spectra. Unexpectedly, we observed that the mutation led to a reorientation of the Trp23 side chain: in the mutant it points inside the hydrophobic core, thus filling the cavity created by the replacement of Phe31 by alanine, which also results in distortion of the backbone. This is a novel mechanism of adaptation to cavity-creating mutations.

MATERIALS AND METHODS

Materials. Mutants F31A, F31Y, and W23A were produced by cassette mutagenesis as previously reported (*12*, *18*). The same procedure was adopted to produce the mutant Y33A using suitable oligonucleotides. Wild-type and mutant proteins were purified as described previously (*18*). Sample concentration for NMR measurements was 2 mM, at pH 4.5 and at 300 K. All NMR samples were prepared in 90% H_2O and 10% D_2O.

NMR Spectroscopy. Two-dimensional ^{1}H NMR spectra, DQF-COSY (*21*, *22*), GE-NOESY (*23*, *24*), and GE-TOCSY (*23*, *24*) were collected on a 500 MHz Bruker Avance-DMX spectrometer equipped with a *z* gradient coil with a proton frequency of 500.13 MHz. All spectra were recorded with a spectral width of 8096 Hz, and the NOESY spectra, with a 150 ms mixing time. Standard Bruker library pulse programs were employed for acquisition: bidimensional NMR spectra were acquired in the phase sensitive mode, with quadrature detection in both dimensions by using the time proportional phase increment method (TPPI) (*25*) and pulsed-field gradients for water elimination (WATERGATE) (*24*, *25*). Prior to complex Fourier transformation, the data were weighted with a $\pi/3$ squared shifted sine bell in t_2 and a Kaiser window function in t_1. Baseline corrections were applied by using baseline convolution in the F_2 dimension with 4 and 40 points as noise and peak size, respectively. The final matrixes were 1024×1024 real points. All spectra were processed on Personal Iris 4TG30 by using Felix_ND a module of InsightII package from Molecular Simulation.

Spectral Analysis and Structure Calculations. Resonance assignments were performed by using the resonance spin pattern of Trp23 as entry point in the DQF and TOCSY spectra, and the other residues were identified by using the standard assignment procedures (*26*). The sequential assignment was achieved via NOESY connectivities following standard two-dimensional methods (*27*). Stereospecific assignment for the β-methylene protons and χ_1 torsion angles were obtained considering NOE(Hα, Hβ), NOE(HN, Hβ), and ^3J(Hα, Hβ) (*28*, *29*). The volumes of all assigned NOE cross-peaks from the NOESY spectra were measured using the FELIX integration routines and converted into distances as described previously (*30*). The geminal NOE's observed between the protons of Gly36 (1.78 Å) and the Hβ protons of Leu54 (1.75 Å) were used to calibrate the conversion of volumes into distances. A total of 744 derived distance constraints comprising 410 intraresidue, 156 sequential, 53 medium-range, and 125 long-range restraints were employed, afterwards divided into three classes: strong ($<$2.7 Å); medium ($<$3.5 Å); weak ($<$5.0 Å). Pseudoatoms with appropriate distance corrections were created for nonspecifically assigned methylene protons, methyl groups, and aromatic ring protons (*27*). Lower bounds were taken to be the sum of the van der Waals radii for the interacting protons in all cases. A set of 61 chiral restraints was used. Three-dimensional structures were determined using a modified dynamic simulated annealing method (*31*) implemented within the InsightII program by using the CVFF force field. The maximum force of the distance constraint potential was set to 15 kcal mol^{-1} Å^{-2} while either lower bound (K_L) and upper bound (K_U) force constants were set to 1 kcal mol^{-1} Å^{-2}. The simulations were carried out in five stages: (1) 100 steps Powell energy minimization to remove bad nonbonded contacts; (2) 15 ps of dynamics at 1000 K with normal van der Waals radii and repulsive force constant fixed at 0.001 kcal mol^{-1} Å^{-4}; (3) 10 ps of dynamics at 1000 K with repulsive force constant incremented at 0.01 kcal mol^{-1} Å^{-4}; (4) cooling to 300 K during 13 ps (13 steps of 1 ps with 8.3 K cooling/step) with repulsive force constant of 0.25 kcal mol^{-1} Å^{-4} and van der Waals radii scaled by 0.85; (5) energy minimization with 1200 steps of Steepest Descent algorithm followed by 1200 steps of Conjugate Gradient.

A minimized linear conformation of the protein was employed as starting structure. The simulations were repeated

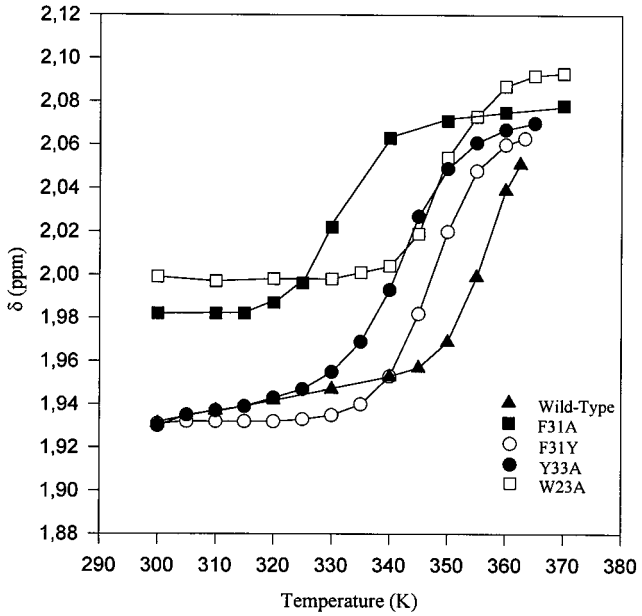

FIGURE 1: NMR thermal denaturation profiles of the Sso7d wild type and its F31Y, F31A, W23A, and Y33A mutants determined by monitoring changes in ^{1}H chemical shifts of the methyl resonances of Met57 (*18*).

several times, and finally 50 structures were obtained: 22 best structures were selected on the basis of best RMSD's, smallest constraint violations, and lowest total energy. An average-minimized structure was calculated from these 22 structures by using the same potential as in the last stage of the simulated annealing protocol. The 22 structures were afterward geometrically checked (in terms of $\chi_1-\chi_2$ plots, main chain bond lengths, main chain bond angles, and RMS distances from planarity), and a Ramachandran plot (*32*) was used to analyze the ϕ and ψ angles clustering. A validation structure check was performed by ProCheck V3.0 at Biotech Check Server-EMBL Heidelberg. The cavity volume calculations were obtained by Voidoo Program from Uppsala Software Factory (*33*).

RESULTS

Thermostability of Sso7d and Its Mutants. In our previous investigations aimed at elucidating the determinants of thermodynamic stability in Sso7d, we produced two mutants of Phe31, a residue located at the hydrophobic core, i.e., F31A and F31Y (*18*), as well as one of the exposed residue Trp23, namely W23A (*12*). In the present work, we mutagenized the residue Tyr33, also found at the hydrophobic core, thus producing the mutant Y33A. Thermal denaturation profiles of these mutants, along with that of the recombinant wild type, were detected by NMR as previously reported (*18*) by monitoring changes in ^{1}H chemical shifts of the methyl resonances of Met57 (Figure 1). Interestingly, at room temperature, the chemical shift of the methyl resonance of Met57 strongly differs from that of Met28 ($\Delta\delta = -0.05$ ppm) in the wild-type protein and in all mutants except for F31A and W23A, in which the methyls of the two methionine are almost isochrone. Indeed, Met57 is quite close to the aromatic ring of Trp23 and experiences a deshielding effect, which is obviously absent in W23A mutant. The absence of the diamagnetic shift observed in F31A mutant suggests a possible reorientation of the Trp23 side chain with respect to that of Met57.

A drop of midpoint transition temperatures ($T_{1/2}$) was revealed in all of these mutants; however, F31A showed an impairment in thermostability by as much as 27 K (with a $T_{1/2}$ of 361 K for the wild type as compared to 334 K for the mutant), although our preliminary investigations based on the CHα chemical shift indexes did not show substantial changes in the overall fold of the protein as a consequence of the mutation (*18*). All the other mutations affected protein thermostability to a lesser extent: among them, Y33A was the most destabilizing, with a $T_{1/2}$ of 343 K. Interestingly, the mutation F31A also led to a dramatic loss in piezostability, i.e., of at least 10 kbar (*18*). To our knowledge, the dramatic impairment in pressure and temperature stabilities found in this latter mutant are unprecedented. These findings prompted us to undertake its structural characterization, which is reported in this paper.

Sequential Assignment and Secondary Structure. The secondary structure elements of F31A were identified by the detection of specific patterns of short-, medium- and long-range NOE's. The α-proton chemical shift indexes (*34*) were employed to initially explore the existence of a specific secondary structure (*18*); in Figure 2 the short and medium range NOE connectivities of backbone protons are summarized. From these data we identified the presence of five β-strands and one helix: in particular a double-stranded antiparallel β-sheet, made up of residues 2−7 and 10−15 (strands I and II), onto which a strongly distorted triple-stranded antiparallel β-sheet is packed in an orthogonal manner, involving the residues 21−25, 28−33, and 41−46 (strands III−V, respectively). Residues 16−19 make a tight turn as documented by the presence of $d_{\alpha N}$ (i, $i+3$), d_{NN} (i, $i+2$), $d_{\alpha N}$ (i, $i+2$) NOE's, that connects the two β-strands. Furthermore, a "classic" β-bulge (*35*, *36*) involving residues Phe5 and Lys6 was identified on the basis of the NOE scheme reported in Figure 3 where the carbonyl oxygens of Lys4, Phe5, and Lys6 point in about the same direction, as is typical in an α-helical conformation. Also, the side chains of the residues involved in the β-bulge (Phe5, Lys6, and, on the opposite strand, Lys12) were found to be at the same side of the β-sheet, which results in a locally accentuated twist of the irregular β-strands.

At the C-terminus, residues 47−58 form a distorted helix interrupted by Pro51. In this stretch, the residues 47−50 give rise to a complete turn of a 3_{10} helix, whereas the C-terminal region residues 52−58 make up an amphipathic α-helix that is well defined by medium $d_{\alpha N}$ (i, $i+3$) and $d_{\alpha\beta}$ (i, $i+3$), and small $d_{\alpha N}$ (i, $i+4$) NOE's. It was also observed that a hydrogen bond between the carboxyl oxygen of Glu53 and the N-terminal group of Ala1 connects the C-terminal helix to the double stranded antiparallel β-sheet, thus keeping them close to each other.

The residues not involved either in helices or in β-strands are clustered in loop regions. Residues of the segment Tyr7−Glu10 are involved in distorted type I β turn, probably better classified as a type IV on the basis of the ϕ, ψ plots (data not shown). Residues comprised between Val25 and Met28 seem to be involved in a type II$^{\prime}$ turn. Residues comprised between Asp34 and Thr40 form a flexible loop without a defined conformation, as confirmed by the high RMSD values.

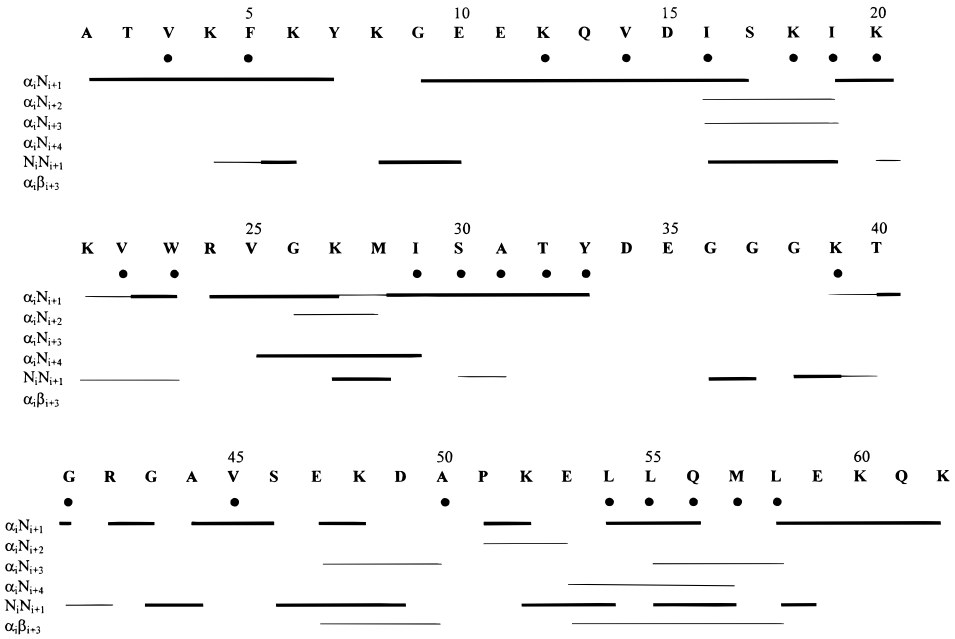

FIGURE 2: Summary of short- and medium-range NOE's for F31A in 90% H_2O and ^2H_2O at pH 4.5 and 300 K observed in a 2D NOESY experiment spectrum recorded with a mixing time of 150 ms. These structural parameters were used for the sequence-specific assignments and the identification of secondary structure elements in F31A. The NOE's are classified into strong and weak according to the thickness of the filled bar. The amide protons not displaying any isotope exchange after 4 h in deuterated solvent are shown as little black balls.

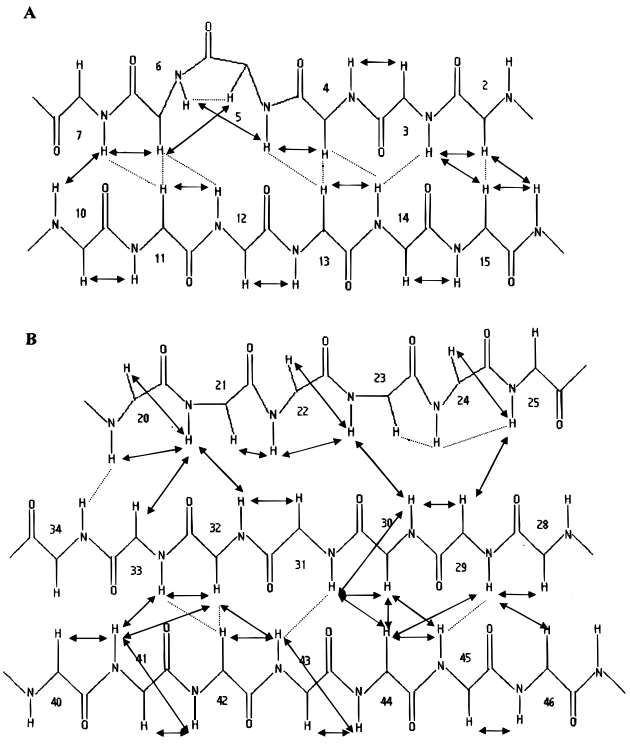

FIGURE 3: Schematic diagram illustrating the two β-sheet regions for F31A: (A) double-stranded β-sheet; (B) triple-stranded β-sheet. Observed NOE's are indicated by solid double-headed arrows. The ambiguous NOE's due to cross-peak overlaps are indicated by dotted lines.

The diagonal plot of NOE data (Figure 4) summarizes in a compact presentation the distance constraints of the secondary structure of F31A (*28*), whose classification is achieved on the basis of the different diagonal NOE distribution. Even though the elements of the secondary structure are easily recognized, they are not so well defined when compared with the regular and compact structure of the wild type (Figure 4); in particular, a comparison of the two structures highlights the double-stranded β-sheet and the C-terminal helix as the less ordered regions and shows the new NOE contacts resulting from the mutation.

To help in determining the degree of flexibility for F31A, amide exchange rates were studied. Amide hydrogens involved in intramolecular H-bonding or otherwise shielded from solvent could be identified by their relatively slow rate of H/D exchange. Very fast exchange (in 5 h all amides exchanged) was detected for all the amide protons of F31A by dissolving the protein in D_2O, which indicates that only few NH's are involved in H-bonds, while most amides seem to be completely solvent-exposed. This is in agreement with a floppy structure for F31A and with the loss of the H-bond network recognized in the compact and stable wild-type structure. Indeed, as reported in Figure 2, the NH's that have not yet exchanged in deuterated solvent after 4 h belong to the best defined and ordered regions of F31A. For instance, in the stretch Ile29−Tyr33 of the β-strand IV all the amide protons are in slow exchange with solvent and this poor accessibility fits well with the structured conformation of the strain. Also, the Leu54−Leu58 region of the α-helix has less labile protons, as confirmed by the two H-bonds found in the mutant structure (Met57:HN-Glu53:O; Leu58:HN-Glu59:O). Very few H-bonds were found in F31A, and only two of these belong to the β-sheet (Lys20:HN-Tyr32:O, Ala31:HN-Gly43:O) contributing to the strand packing. Probably other driving forces are involved in preserving the global fold of the protein, in particular hydrophobic interactions contributed by residues in the aromatic cluster.

Quality of the Calculated Structures. We also assessed in terms of structural statistics the quality of the structure we resolved in NMR; some results from the analysis of the minimized average structure using ProCheck-NMR are summarized in Table 1. No disallowed regions of the Ramachandran plot (not reported) are populated. The major-

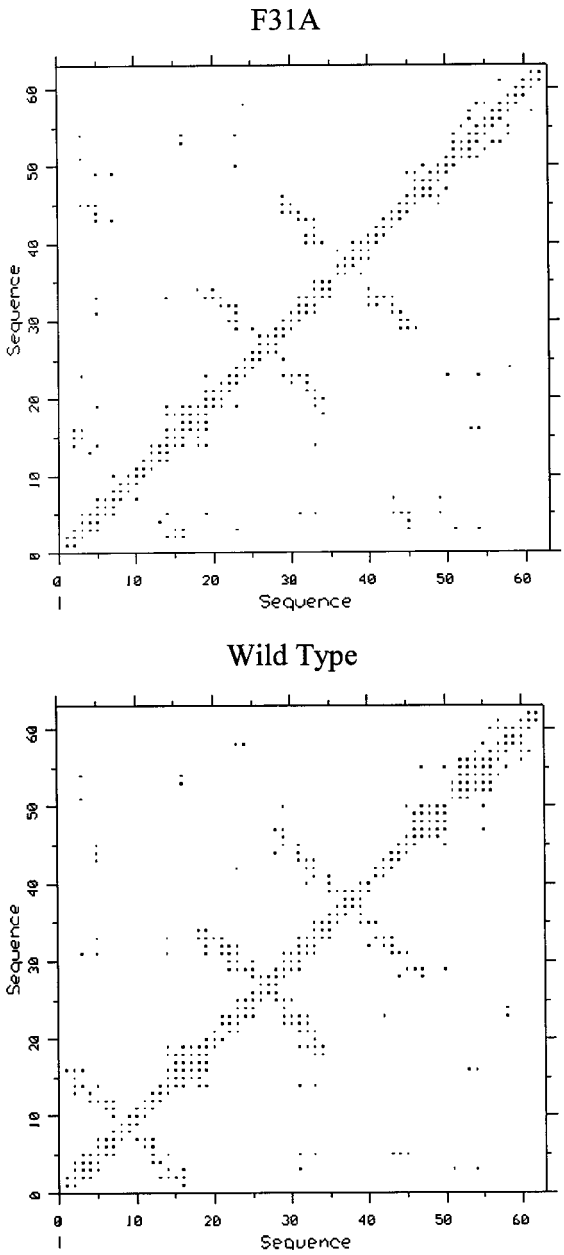

FIGURE 4: Diagonal plot of all intra- and inter-residue NOE's observed in Sso7d and in its F31A mutant. Each point can represent more than one NOE between different protons of the same pair of residues.

ity of the residues (73.1%) are within the most favored regions, and 21.2%, in the additional allowed regions. Only three residues (Lys21, Glu47, Glu59) have backbone conformations that fall within generously allowed regions of the Ramachandran plot. Except for Lys21, these residues are indeed located in the most flexible regions of the protein (see above).

The RMSD's for superposition of several subsets of residues onto the average structure are quoted in Table 1; if only the residues assigned to elements of β-sheet secondary structure are considered, the RMSD for the backbone is 0.72 $\pm$ 0.17. If the entire molecule is considered, two regions stand out as poorly defined by the ensemble of structures: the C-terminal residues (Ser46−Lys62 residues) and the loops (in particular the Tyr7−Glu10, Val25−Met28, and Asp34−Thr40 regions). The lower precision obtained for these regions can in most cases be attributed to rapid internal motions due to local fluctuations. As a possible consequence, these segments show only few NOE distance constraints (Figure 4).

Table 1. NMR Constraints and Structural Statistics for F31A Mutant

A. Constraints Used in the Final Structure Calculation	
long-range ($\|i-j\| > 4$)	125
medium-range ($1 < \|i-j\| \leq 4$)	53
sequential ($\|i-j\| = 1$)	156
intraresidue	410
total	744
torsion angle (χ^1)	5
B. Atomic RMS Difference Statistics for the 22 Converged Structures vs Average Structure[a]	
backbone RMS for the fragment superimposed 1−60	1.41 ± 0.25
backbone RMS for the fragment superimposed 46−60	0.94 ± 0.09
backbone RMS for the fragment superimposed 2−7, 10−15	0.84 ± 0.06
backbone RMS for the fragment superimposed 21−25, 28−34, 41−45	0.64 ± 0.12
C. Ramachandran Analysis of Minimized Average Structure (PROCHECK Analysis)	
residues in most favored regions	73.1%
residues in additional allowed regions	21.2%
residues in generously allowed regions	5.8%
residues in disallowed regions	0.0%
D. Structural Statistics of Minimized Average Structure (PROCHECK Analysis)	
no. of "significant" bond length violations (>0.05 Å)	0.0
no. of "significant" bond angles violation (>10.0°)	1.0
no. of residues in χ^1 and χ^2 unfavorable conformations	0.0
no. of aromatic residue with a deviation from planarity larger than 0.03 Å	1.0

[a] Mean RMS ± standard deviation (Å).

FIGURE 5: Best-fit superpositions of backbone atoms from residue 2 to 57 of the 15 wild-type Sso7d (A) and 22 F31A mutant (B) structures.

Comparison between the Structures of the Recombinant Wild-Type Sso7d and the F31A Mutant. The calculated average structures of the wild type and F31A possess a similar backbone folding (Figure 5); however, the superposition of the backbone atoms of the refined structures yields an RMSD of 2.65 ± 5.31 Å. The mutant has, in fact, a more flexible and less parallel β-structure and a disordered and truncated C-terminal helix when compared to the wild type (Glu47−Met58 stretch in F31A instead of Glu47−Gln61 in the wild type).

On the basis of the distributions of the experimental distance constraints reported in Figure 4, wild-type and mutant structures may be compared; thus, most of NOE distance restraints defining the elements of the secondary structure are observed both in the wild type and in the mutant,

FIGURE 6: Backbone superposition of wild type Sso7d (rectangular gray ribbon) and the F31A mutant (oval black ribbon). The figure shows the different orientation of the Trp23 residue in the two structures.

although some regions show a different NOE distribution. In particular, the double-stranded β-sheet is less defined in the mutant, and new interchain NOE's involving Trp23, not observed in the wild type, are recognized.

The most remarkable feature found in the structure of the mutant is the change in orientation of the Trp23 side chain (Figure 6). In the wild type, the aromatic ring of Trp23 is completely exposed to the solvent (Consonni, R., unpublished experiments), in keeping with our previous photo-CIDNP experiments (*37*); furthermore, it is located in a region corresponding to the DNA-binding surface of the protein where it can directly interact with the nucleic acid (*15*, *16*). In contrast, the Trp23 side chain in the mutant is largely buried in the aromatic cluster with Phe5 and Tyr33, which results in backbone distortion. This reorientation of the Trp23 side chain prevents formation of a cavity in the hydrophobic core of the mutant, which instead would arise in the absence of structural rearrangements. In addition, the side chains of residues Tyr7 and Phe5 also undergo reorientation: in the mutant, Tyr7 points toward the triple-stranded β-sheet, making the double-stranded β-sheet very distorted and curled even if the residue conserves its high solvent accessibility. The side chain of Phe5 points inside the new hydrophobic region where, together with Trp23 and Tyr33, it cooperates in restoring the aromatic core.

The calculated structure of F31A can account for the differences of chemical shifts observed between mutant and wild-type proteins. Deviations from wild-type values are illustrated in Figure 7 as a plot of the chemical shift difference ($\delta_{F31A} - \delta_{wild\ type}$) versus the residue number. Distinction is made between backbone amide and α protons.

Discrepancies in the NH chemical shifts higher than $\pm$0.2 ppm were found for 10 residues (Lys4, Lys6, Gln13, Val14, Lys20, Ala31, Tyr33, Gly43, Ala44, Val45) belonging to stretches of β-sheets. Located in the same regions, nine residues (Val3, Lys4, Phe5, Lys6, Glu11, Gln13, Ser30, Thr32, Ala44) present deviations higher than $\pm$0.2 ppm in the Hα chemical shifts. Many of these chemical shift deviations are located in the double stranded β-sheet. In particular, with respect to the value observed for the wild type (5.35 ppm), the α proton of Lys4, a residue belonging to the strand I, is upfield shifted at 4.69 ppm, probably due

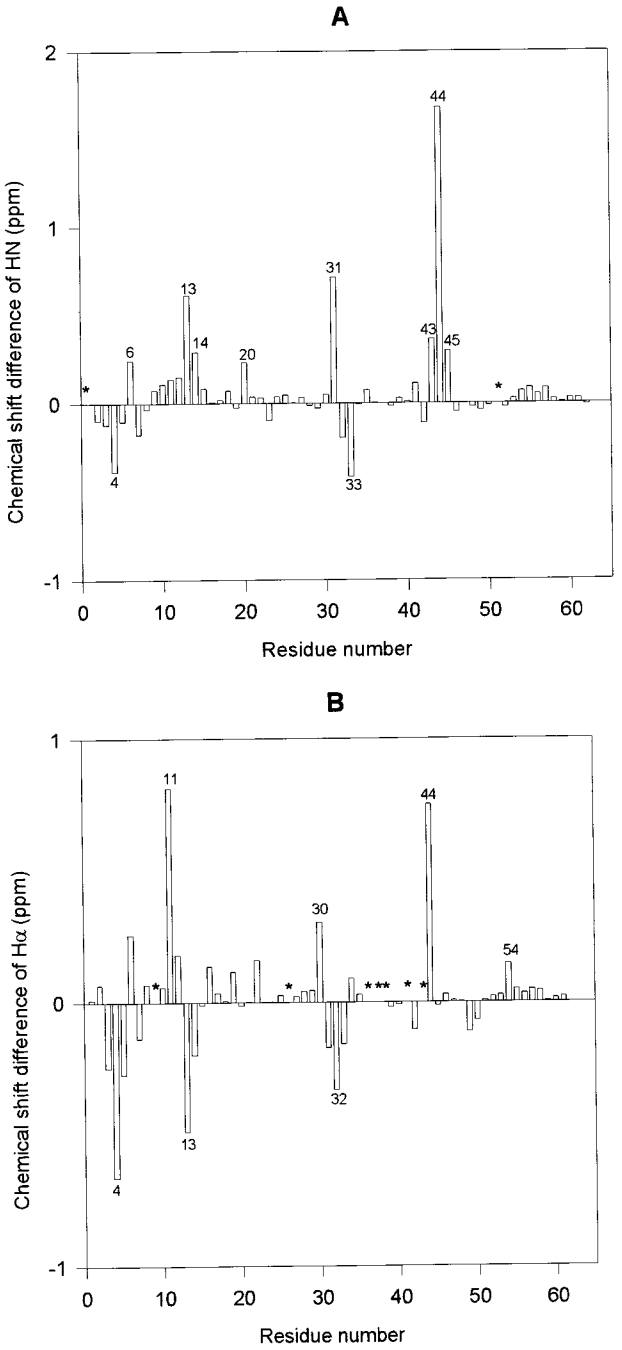

FIGURE 7: Plot of the chemical shift difference for F31A in 90% H_2O/10% ^2H_2O at pH 4.5 and 300 K. The chemical shifts of wild-type Sso7d were subtracted from those of the F31A mutant where available and plotted versus residue numbers. Chemical shift differences are for (A) backbone amide protons and (B) Hα protons. An asterisk indicates no chemical shift available.

to the reorientation of the Phe5 aromatic chain in the mutant; this proton is in fact in the region above the plane of the aromatic ring at a distance below 5 Å, which results in strong shielding. In contrast, the α proton of Lys4 in the wild type lies in the plane of the Phe5 ring and because of this experiences a low-field shift effect. The same effect was also observed for the NH chemical shift difference of Lys4.

The strand II also presents some significant chemical shift variations localized on Glu11 and Gln13. In particular, the α proton chemical shift of Glu11 is subjected to a strong deshielding effect: on inspection of the structure, the backbone carbonyl groups of neighboring Glu10 and Lys12

are both oriented in the same direction toward the Glu11 α proton at a distance of 2.72 and 2.53 Å, respectively. In this conformation it is reasonable that the α proton experiences a strong paramagnetic anisotropy responsible for the low-field shift. In addition, while in the wild type the Glu11 α proton was just inside the shielding cone of Phe5 aromatic ring at a distance of 6 Å, in the mutant it points far away from Phe5 at a distance of 12 Å.

In strand IV, the stretch Ser30−Tyr33 experiences a large low-field shift (Ser30:Hα and Ala31:HN) and upfield shift (Thr32:Hα and Tyr33:HN); all these protons point in the same direction toward the strand V and are influenced mainly by the aromatic substitution F31A, which causes the loss of the ring current effect for the HN of Ala31 and the Hα of Ser30. Furthermore, the aromatic ring reorientation of Phe5 gives rise to upfield shifts for Hα of Thr32 and HN of Tyr33. In addition, new NOE contacts (such as Phe5:Hz-Ile19:Hδ1; Leu54:Hδ1*-Trp23:Hδ1; Leu54:Hδ1*-Trp23:Hϵ1), caused by structural reorganization of Phe5 and Trp23, can contribute to the chemical shift variations observed.

In the sheet V, the spin system of Ala44 changes completely its chemical shift values; a general low-field shift of all signals, due to the ring current effect determined by the Tyr7 aromatic side chain, is observed.

In the terminal helix the available chemical shift perturbations are small, except for Leu54 whose α proton experiences in the mutant a downfield shift most likely determined by the Trp23 aromatic ring. All the chemical shifts of the Leu54 side chain are subjected to a deshielded effect, in good agreement with the new experimental NOE contacts between Leu54 and Trp23 (Leu54:Hδ1*-Trp23:Hδ1; Leu54:Hδ1*-Trp23:Hϵ1), that make the frequencies of Leu54:Hδ1* and Leu54:Hδ2* distinguishable (1.0 and 0.8 ppm) in the mutant with respect to the degenerate value (0.6 ppm) obtained in the wild type. This finding may be a consequence of the methyl reorientation just above the indolic plane at a distance of 4 Å.

DISCUSSION

In the past years, we put much effort in elucidating the structural features responsible for the thermodynamic stability of the small, basic Sso7d protein from *S. solfataricus*. To this end, we first expressed in *E. coli* the recombinant protein (*17*) and subsequently produced several mutants, which were characterized along with the wild type (*12*, *18*, *19*). Among the mutations assayed, F31A was the most destabilizing, with a decrease in thermo- and piezostabilities by 27 K and over 10 kbar, respectively. An effect of such a magnitude is, to our knowledge, unprecedented. This work was therefore undertaken to determine the solution structure of F31A and find out what structural changes may be related to this dramatic outcome.

The most remarkable difference we detected between the structures of the wild type and F31A is the reorientation of the tryptophan residue, which is completely exposed to the solvent and participates in DNA binding in the wild type (*14*−*16*). We found, in contrast, that in the mutant this residue points inside the aromatic cluster, thus filling the cavity which otherwise would arise from the replacement of Phe31 by alanine. This structural change can be quantified in terms of decrease in solvent accessible surface for Trp23, i.e., 111 and 27 Å^2 in wild type and mutant, respectively. This rearrangement allowed the reconstitution of the hydrophobic core so that, on the whole, no significant difference in cavity volumes between the wild type and mutant could be detected. Thus, additional factors should be invoked to account for the strong destabilization produced by the mutation.

First of all, the inclusion of the tryptophan side chain in the hydrophobic core necessarily involves an entropic cost associated with the immobilization of this aromatic, which obviously represents an unfavorable contribution to the thermodynamic stability of the mutant. Furthermore, changes in the overall structure of the aromatic cluster should be also taken into account. It has been indeed reported that in several proteins aromatic side chains participate in aromatic pairs which often form networks of three or more interacting side chains, where phenyl ring centroids are separated by a preferential distance ranging from 4.5 to 7 Å and dihedral angles approaching 90° are most common (*20*). Also, clusters displaying such geometry are believed to play a major stabilizing role, irrespective of whether they are buried in the hydrophobic core of the protein or exposed to the solvent (*20*). The aromatic cluster in wild-type Sso7d was found to fit well this pattern (*18*), whereas in the mutant it lost this typical fishbone geometry. Particularly, Phe5, Tyr33, and Trp23 were roughly coplanar, Phe5 being shifted toward the space vacated by the mutation and rotated by about 90° with respect to orientation it had in the wild type.

Our data also suggest that further factors may concur to the loss in stability produced by the mutation. We found indeed that the solvent accessible area was significantly larger in the mutant than in the wild type, i.e., 6124 Å^2 vs 5800 Å^2. This corresponds to an increase of 324 Å^2, mostly accounted for in terms of exposure of hydrophobic patches, which alone represent over 90% of this surface. These changes should necessarily result in a substantially reduced hydrophobic effect in the mutant with respect to the wild type, if this is quantified in terms of hydrophobic surface area exposed on denaturation (*37*, *38*). This implies that overall structural changes arisen from the mutation, besides the local ones regarding the aromatic cluster, are also involved in the destabilization. Particularly, the mutation also gave rise to significant modifications of the secondary structure. Although β-strands and the helical stretch could still be recognized in the mutant, they were not, on the whole, so regular and well defined as in the wild type, so that only two hydrogen bonds in the β-sheets and two in the helical stretch were preserved in the mutant. The C-terminal helical stretch was indeed largely unstructured in F31A as compared to the wild type. Furthermore, a major change resulting from the reorientation of the tryptophan residue was the distortion of the backbone at the level of the β-strand III, where this residue is located. In addition, a "classic" β-bulge (*35*, *36*) involving residues Phe5−Lys6 on strand I and Lys12 on strand II was observed. This altered the direction of the backbone modifying the shape of the loop (Tyr7−Glu10) and accentuating the twist of the strands. On the whole, these modifications of the secondary structure involve a much larger overall flexibility of the mutant with respect to the wild type. This is also apparent from a comparison of the RMSD, with a value for backbone atoms of 1.41 Å in the former and 0.93 Å in the latter. If only β-sheets are taken

into account, the values are 0.72 and 0.45 Å, respectively. This should in part justify the impaired stability of the mutant, as an inverse correlation between flexibility and both piezo- and thermostability has been clearly established (*39*, *40*).

Besides the increased flexibility, some local conformational strain should be also involved in destabilization of the mutant. This holds true at least for the residue Lys21 as shown by its relatively unfavorable position in the Ramachandran plot and by the low RMSD of its backbone atoms (*41*).

Before solving the structure of the mutant F31A, we did not expect dramatic structural modifications as a consequence of the mutation. It has been indeed reported that, on replacement of a bulky hydrophobic residue by alanine, a protein may or may not relax, depending on the context in which the mutation is introduced (*38*, *42*). In one extreme case the protein relaxes completely, thus filling the cavity resulting from the mutation. In such a case the thermodynamic destabilization has the lower theoretical value. At the other extreme, the protein structure is completely unchanged: a large cavity arises and the extent of thermodynamic destabilization reaches its highest theoretical limit. On the basis of these observations, it was plausible to expect the appearance in the mutant of a cavity filled to some extent by the collapse of the surrounding residues. As our data show that the real situation is completely different, we are forced to conclude that the rearrangements predicted on the basis of the aforementioned reports would be far more destabilizing than that actually associated with the inclusion of tryptophan in the aromatic cluster. This highlights the crucial role of a well-packed hydrophobic core for protein stability, even though in the mutant F31A a correct fishbone geometry of the aromatic cluster was not preserved. It should be also stressed, in this respect, that we could not detect significant overall differences between mutant and wild-type proteins in terms of cavity volumes; so the decreased stability does not result from a cavity-creating rearrangement.

To the best of our knowledge, no similar rearrangement has been so far described for other proteins. However, a finding relevant to our observations is that reported by Baldwin et al. (*43*). They produced several single-point mutants of T4 lysozyme in which a number of aliphatic residues were replaced by alanine and, in one case, by glycine. Such mutations gave rise to either cavities or crevices which could be filled by aromatic compounds present in the medium, such as benzene or *p*-xylene. This shows the propensity of cavities located in an apolar context of a protein to bind aromatic compounds. However, in the case of the inclusion of Trp23 in the hydrophobic core of F31A, this necessarily involves unfavorable contributions, as outlined above. We therefore conclude that the energetically favorable reconstitution of the hydrophobic core must outweigh these factors.

Besides the structural characterization of F31A, we also addressed the issue of how this mutation affects the biological properties of Sso7d. The ability of this molecule to bind DNA aspecifically has been well established, so that it is believed to act as a histonelike protein in *S. solfataricus* (*8*, *14*, *15*). We found in addition that it is endowed with ribonuclease activity. However intriguing this finding may be, it is unquestionable that this enzyme activity is an intrinsic property of Sso7d, as supported by the following evidence: (i) The protein isolated from its biological source was devoid of any detectable contamination, as shown in SDS−PAGE, sequencing experiments (*11*), and MALDI-TOF mass spectrometry (data not shown). (ii) The recombinant Sso7d, expressed in *E. coli*, also displayed catalytic activity indistinguishable from that of the natural protein (*17*). (iii) Some single-point mutants of the protein were devoid of any ribonuclease activity (Fusi, P., manuscript in preparation). Interestingly, ribonuclease activity was also found in the Sac7e protein from *S. acidocaldarius*, which is closely related to Sso7d (*44*).

The DNA-binding capacity of wild-type Sso7d and its mutant F31A was qualitatively assessed in double-stranded DNA−cellulose chromatography. Both proteins bound to the resin, but they were eluted at 0.5 and 0.2 M NaCl concentrations, respectively (data not shown); this highlights the impaired capacity of the mutant to bind DNA, in keeping with former reports which assign to the tryptophan residue a major role in this function (*15*, *16*). As expected, the mutant W23A also displayed an impaired DNA-binding capacity in that it was eluted at the same NaCl concentration as F31A (data not shown).

Ribonuclease activity was assayed using yeast RNA as a substrate, as previously described (*11*). It was thus shown that the activity of the mutant F31A was approximately 40% of that determined in the recombinant wild type (data not shown). This may be understood on the basis of our previous photo-CIDNP experiments (*45*). They showed that the aromatic cluster is accessible to, and capable to interact with, a flavin dye used as a model ligand: this suggests that the cluster might participate in the catalytic mechanism by binding RNA bases. Thus, it is plausible that any modification of this structure also impairs the catalytic activity of the protein.

In conclusion, our results show a novel mechanism of adaptation to a cavity-creating mutation: the fact that, in the mutant F31A, the tryptophan is included in the hydrophobic core despite several unfavorable contributions associated with this rearrangement, highlights the importance of the preservation of the core itself for protein stability. This also implies that, in the absence of an exposed aromatic residue in a suitable location as in the case of Trp23, the mutation F31A would be even more destabilizing. Work is in progress in our laboratories to test this hypothesis in the double mutant W23AF31A.

ACKNOWLEDGMENT

The technical assistance of Fulvia Greco and Giulio Zannoni in the NMR measurements is acknowledged.

REFERENCES

1. Schmid, M. B. (1990) *Cell 63*, 451−453.
2. Dijk, J., and Reinhardt, R. (1986) in *Bacterial chromatin* (Gualerzi, C. O., and Pon, C. L., Eds.) pp 185−218, Springer-Verlag, Berlin.
3. Pettijohn, D. (1988) *J. Biol. Chem. 236*, 12793−12796.
4. Drlica, K., and Rouviere-Yaniv (1987) *Microbiol. Rev. 51*, 301−319.
5. Kimura, M., Kimura J., Davie, P., Reinhardt, R., and Dijk, J. (1984) *FEBS Lett. 176*, 176−178.
6. Grote, M., Dijk, J., and Reinhardt, R. (1986) *Biochim. Biophys. Acta 873*, 405−413.

7. Choli, T., Wittmann-Liebold, B., and Reinhardt, R. (1988) *J. Biol. Chem. 263*, 7087−7093.
8. Choli, T., Henning, P., Wittmann-Liebold, B., and Reinhardt, R. (1988) *Biochim. Biophys. Acta 950*, 193−203.
9. Reddy, T. R., and Suryanarayana, T. (1988) *Biochim. Biophys. Acta 949*, 87−96.
10. Edmonson, S. P., Qiu, L., and Shriver, J. W. (1995) *Biochemistry 34*, 13289−13304.
11. Fusi, P., Tedeschi, G., Alverti, A., Ronchi, S., Tortora, P., and Guerritore, A. (1993) *Eur. J. Biochem. 211*, 305−310.
12. Catanzano, F., Graziano, G., Fusi, P., Tortora, P., and Barone, G. (1998) *Biochemistry 37*, 10493−10498.
13. Fusi, P., Grisa, M., Tedeschi, G., Negri, A., Guerritore, A., and Tortora, P. (1995) *FEBS Lett. 360*, 187−190.
14. Baumann, H., Knapp, S., Lundbäck, T., Ladenstein, R., and Hard, T. (1994) *Nat. Struct. Biol.* 1, 808−819.
15. Agback, P., Baumann, H., Knapp, S., Ladenstein, R., and Härd, T. (1998) *Nat. Struct. Biol. 5*, 579−584.
16. Gao, Y.-G., Su, S.-Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmonson, S. P., Shriver, J. W., and Wang, A. H.-J. (1998) *Nat. Struct. Biol. 5*, 782−786.
17. Fusi, P., Grisa, M., Mombelli, E., Consonni, R., Tortora, P., and Vanoni, M. (1995) *Gene 154*, 97−102.
18. Fusi, P., Goossens, K., Consonni, R., Grisa, M., Puricelli, P., Vecchio, G., Vanoni, M., Zetta, L., Heremans, K., and Tortora, P. (1997) *Proteins: Struct., Funct., Genet. 29*, 381−390.
19. Mombelli, E., Afshar, M., Fusi, P., Mariani, M., Tortora, P., Connelly, J. P., and Lange, R. (1997) *Biochemistry 36*, 8733−8742.
20. Burley, S. K., and Petsko, G. A. (1985) *Science 229*, 23−28.
21. Piantini, U., Sørensen, O. W., and Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800−6801.
22. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1993) *Biochem. Biophys. Res. Commun. 117*, 479−485.
23. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−666.
24. Sklenar, V., Piotto, M., Lappik, R., and Saudek, V. (1993) *J. Magn. Reson. Ser. A* 102, 241−245.
25. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−971.
26. Englander, W., and Wand, J. (1987) *Biochemistry 26*, 5958−5962.
27. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
28. Bystrov, V. F. (1976) *Prog. NMR Spectrosc. 10*, 41−81.
29. Montelione, G. T., Winkler, M. E., Rauenbuehler, P., and Wagner, G. (1989) *J. Magn. Reson. 82*, 198−204.
30. Baleja, J. D., Marmorstein, R., Harrison, S. C., and Wagner, G. (1992) *Nature 356*, 450−453.
31. Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M. (1988) *Protein Eng. 2*, 27−38.
32. Laskowsky, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
33. Kleywegt, G. J., and Jones, T. A. (1994) *Acta Crystallogr. D50*, 178−185.
34. Wishart, D. S., Sykes B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
35. Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
36. Richardson, J., and Richardson, D. (1989) in *Prediction of Protein Structure, and the Principle of Protein Conformation* (Fasman, G., Ed.) Plenum, New York.
37. Matsumura, M., Becktel, W. J., and Matthews, B. W. (1988) *Nature 334*, 406−10.
38. Matthews, B. W. (1993) *Annu. Rev. Biochem. 62*, 139−160.
39. Cooper, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2740−2741.
40. Gekko, K., Tamura, Y., Ohmae, E., Hayashi, H., Kagamiyama, H., and Ueno, H. (1996) *Protein Sci. 5*, 542−545.
41. Karplus, P. A. (1996) *Protein Sci. 5*, 1406−1420.
42. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E., and Matthews, B. W. (1992) *Science 255*, 178−183.
43. Baldwin, E., Baase, W. A., Zhang, X. J., Feher, V., and Mathew, B. W. (1998) *J. Mol. Biol. 277*, 467−485.
44. Kulms, D., Schäfer, D., and Hahn, U. (1997) *Biol. Chem. 378*, 545−551.
45. Consonni, R., Limiroli, R., Molinari, H., Fusi, P., Grisa, M., Vanoni, M., and Tortora, P. (1995) *FEBS Lett. 372*, 135−139.

BI9911280