Investigations of Sso7d Catalytic Residues by NMR Titration Shifts and Electrostatic Calculations†

Roberto Consonni,*,‡ Ivana Arosio,‡ Barbara Belloni,‡ Federico Fogolari,§ Paola Fusi,|| Erlet Shehi,|| and Lucia Zetta‡

Istituto per lo Studio delle Macromolecole, Lab. NMR, CNR, v. Ampère 56, I-20131 Milan, Italy, Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Pz. ale delle Scienze 2, I-20126 Milan, Italy, and Dipartimento Scientifico e Tecnologico, Università di Verona, Cà Vignal, 1 Strada le Grazie 15, I-37134 Verona, Italy

Received July 25, 2002; Revised Manuscript Received November 5, 2002

ABSTRACT: Sso7d is a small basic protein consisting of 62 amino acids isolated from the thermoacidophilic archeobacterium *Sulfolobus solfataricus*. The protein is endowed with DNA binding properties, RNase activity, and the capability of rescuing aggregated proteins in the presence of ATP. In this study, the electrostatic properties of Sso7d are investigated by using the Poisson−Boltzmann calculation of the surface potential distribution and following by NMR spectroscopy the proton chemical shift pH titration of acidic residues. Although the details of the catalytic mechanism still have to be defined, the results from NMR experiments confirm the possible involvement of Glu35 as the proton acceptor in the catalytic reaction, as seen by its abnormally high pK_a value. Poisson−Boltzmann calculations and NMR titration shifts suggest the presence of a possible hydrogen bond between Glu35 and Tyr33, with a consequent rather rigid arrangement at these positions. Comparison with RNase T1 suggests that Tyr7 may be a good candidate for acting as a proton donor in the active site of Sso7d as shown by its low phenolic pK_a of ∼9.3. Titration experiments performed with the UpA, a RNA dinucleotide model, showed that the protein residues affected by the interaction are mainly located in a different region with respect to the surface affected by DNA recognition, in good agreement with the surface potential distribution found with electrostatic calculations.

In recent years, small, basic, DNA binding proteins from the archaeum *Sulfolobus solfataricus* (Sso7d) or *Sulfolobus acidocaldarius* (Sac7d) have been characterized, either by NMR[1] spectroscopy (*1*, *2*) or by X-ray crystallography (*3*). The structures of these strongly homologous proteins share a common fold (backbone rmsd = 3.80 Å), and they both possess the capability of binding to the minor groove of DNA and causing a kink in DNA of several degrees. This is in agreement with the generally accepted hypothesis concerning their role in genomic DNA protection against thermal denaturation and nuclease digestion. Only one protein from *S. acidocaldarius*, termed Sac7e, displays both DNA binding and RNase activity; however, the residues involved in the catalytic reaction remain unknown. For Sac7d, only a Glu-Lys transition at the C-terminal part of the molecule has been suggested to be responsible for recognition and cleavage of RNA (*4*). Particularly intriguing is the lack in both Sso7d and Sac7e of the two histidine residues usually involved in the catalytic mechanisms of most ribonucleases; this makes them a good model for investigations concerning the RNase catalytic mechanism that has been adopted.

In the past several years, the recombinant form of Sso7d and some single-point mutants have been extensively characterized by our group, to address (i) the thermal and piezo stability and (ii) the RNase activity of this thermophilic protein.

Recently, several papers have appeared regarding the factors determining the thermal stability of thermophilic proteins. Several factors have been found to play a relevant role in the protein stability such as increased hydrophobic and aromatic contacts (*5*), optimization of charge−charge interactions, and side chain packing (*6*, *7*); however, the mechanisms by which these proteins attain thermostability remain unclear. The NMR solution structure of Sso7d from both natural (PDB entry 1SSO) and recombinant (PDB entry 1JIC) sources consists of a compact globular unit that includes a double-stranded antiparallel β-sheet onto which an orthogonal triple-stranded antiparallel β-sheet is packed, and a small helical stretch at the C-terminus. The protein displays a very compact hydrophobic core consisting of side chains at the interface of the two β-sheets, in particular, the aromatic residues Phe5, Phe31, and Tyr33.

Inspection of the recombinant Sso7d NMR structure reveals the presence of a large number of hydrogen bonds and several salt bridges and, furthermore, the presence of the so-called "fish bone geometry" (*8*) assumed by the aromatic cluster, thus conferring a more compact and globular structure. All these interactions are related to the enhanced stability of the protein according to the demon-

† This work was supported by grants from the Italian National Research Council (Target Project on Biotechnology, to R.C.) and the scientific foundation "Antonio De Marco".

* To whom correspondence should be addressed: Istituto per lo Studio delle Macromolecole, Lab. NMR, CNR, v. Ampère 56, 20131 Milano, Italy. Phone: +0039-2-70643555. Fax: +0039-2-70643557. E-mail: roberto@icmnmr.mi.cnr.it.

‡ CNR.

§ Università di Verona.

|| Università di Milano-Bicocca.

[1] Abbreviations: NMR, nuclear magnetic resonance; UpA, uridylyl-(3′,5′) adenosine ammonium salt; TOCSY, total correlation spectroscopy.

10.1021/bi0265168 CCC: $25.00
Published on Web 01/25/2003

strated optimization of electrostatic interactions of proteins from hyperthermophilic microorganisms (*9*). Thirteen lysines are present in the amino acid sequence and are all well exposed to the solvent, giving a high degree of solubility and resulting in a pI as high as 9.

As far as RNase activity of Sso7d is concerned, the pH−activity profile of Sso7d was determined at 60 °C using yeast RNA as a substrate and different buffers in the pH range of 5−10 (*10*). The maximal activity was detected in the pH range of 6.7−7.7 and was not affected by the buffer. Investigations of the cleavage specificity had previously shown that this endonuclease is endowed with a narrow substrate specificity, the cleavage products of which are 3′-phosphooligonucleotides. The protein is stable up to 80 °C and exhibits only 16% inactivation at 85 °C, while complete inactivation is observed at 95 °C.

In this work, we investigate the electrostatic properties of Sso7d both theoretically and experimentally, by using Poisson−Boltzmann calculations and following pH titration via NMR, respectively. Theoretical predictions are compared with experimental findings, and implications are discussed. Ionization constants in proteins can be measured using NMR spectroscopy (*11*), and their determination represents an important tool in characterizing electrostatic interactions. The knowledge of these interactions at the microscopic level has been demonstrated to be crucial to understanding the biological behavior of a protein and especially to selecting mutants with enhanced activity. These experimental measurements are used in combination with a number of computational algorithms which enable us to calculate the pK_a values of titratable groups in proteins. Theoretical approaches based on the Poisson−Boltzmann equation (*12*, *13*) are well-established, and a semiquantitative agreement with experimental results is found.

Sso7d is stable and properly folded over a wide range of pHs, from approximately pH 1.5 to 12, allowing pK_a values for ionizable groups to be obtained. Series of two-dimensional total correlation experiments (TOCSY) carried out at different pH values ranging from 1.5 to 11.5 have been analyzed, giving rise to individual pK_a values of acidic residues. The titration curves obtained from the proton chemical shift variations as a function of pH were analyzed by nonlinear least-squares fitting to simple ionization models. On the basis of the individual pK_a values of all carboxyl groups in Sso7d, the relationship between the observed pK_a values and the tertiary structure was discussed in an effort to provide a comprehensive evaluation of the protein charge status that is relevant to the rationalization of the structure−RNase activity profile. The main goal of this investigation has been to provide key insight into the catalytic site residues.

MATERIALS AND METHODS

NMR Measurements. Sso7d samples were prepared at a concentration of 2 mM in an H_2O/D_2O (90:10) solution. The pH was adjusted to the required values by adding small amounts of HCl or KOH. All ^{1}H spectra were acquired at 11.7 T on a Bruker Advance DMX spectrometer equipped with an SGI INDY computer and a *z* gradient coil with a proton frequency of 500.13 MHz. Spectra were recorded at 300 K, with a spectral width of 8096 Hz and 4K data points, and referenced to sodium trimethylsilyl [2,2,3,3-$^2H_4$]propionate (TSP). Time proportional phase incrementation (*14*) was used to achieve quadrature detection in both dimensions. Solvent suppression was achieved by including the WATERGATE module (*15*) in the original two-dimensional TOCSY pulse sequence. All the spectra were analyzed on an SGI Octane workstation equipped with the MSI Felix package. ^{1}H chemical shifts were referenced to the TSP frequency, and pH-dependent shifts of reference protons have been corrected (*16*, *17*). The curve fitting of chemical shifts was performed with a logistic function consisting of four standard parameters, according to the equation derived from the Henderson−Hasselbach equation (*18*). For titration experiments with UpA dinucleotide, the pH was adjusted to 7 in the presence of 50 μL of 100 μM sodium phosphate buffer, and progressive additions of 4 μL of 26.2 mM dinucleotide solution were made to the protein solution (0.5 μmol of protein in 400 μL) up to a protein:dinucleotide molar ratio of 1:2.

Theoretical Computation of pK_as. The pK_as of titratable sites have been computed using a Poisson−Boltzmann-based methodology (*12*, *13*). In this approach, pK_a shifts are assumed to be due to the difference in the free energy of protonation of a titratable site in the protein and in a model compound for which the pK_a is known. This difference may arise from (i) desolvation terms (for instance, for buried titratable groups), (ii) the interaction of the titratable group with other partial charges (for instance, interaction with backbone amide partial charges), and (iii) charge−charge interactions between ionized groups.

Conversely, any deviation from model compound pK_as implies one or more of the structural electrostatic effects listed above. For this reason, theoretical computation of pK_as and comparison with experimental data may highlight model inaccuracies.

We have followed the procedure of Antosiewicz et al. (*13*), with minor modifications described in ref *19*. It is worth mentioning that Antosiewicz et al. were able to obtain better accuracy with respect to previous methods by increasing the inner dielectric constant to 20. This value takes implicitly into account the flexibility of protein structures, resulting in increased level of solvation of polar groups.

We have computed pK_as both for the two NMR structures of Sso7d deposited in the Protein Data Bank (1JIC and 1SSO) and for the corresponding DNA−protein complex (1BBX). Furthermore, pK_as have been also computed for the strongly homologous protein Sac7d from *S. acidocaldarius* in both the free and DNA-bound state (PDB entries 1SAP and 1AZP, respectively). To simulate the presence of a hydrogen bond between the COOH proton of Glu35 and the OH oxygen of Tyr33, 200 steps of constrained steepest descent and conjugate gradient minimization imposing this hydrogen bond have been performed using Charmm (version 27b2) keeping all atoms fixed at their deposited positions, except for residues 33 and 35. The electrostatic free energy of protonating Glu35 in the protein with all other titratable charged groups and in the isolated residue was computed employing a dielectric constant of 4.0, typical of organic molecules. This choice would imply a rather stable hydrogen bond between Glu35 and Tyr33. This protocol was applied to all structures for Sso7d considered here (PDB entries 1JIC, 1SSO, and 1BBX). The same procedure has been applied for all Sac7d models considered here (residues Tyr33 and

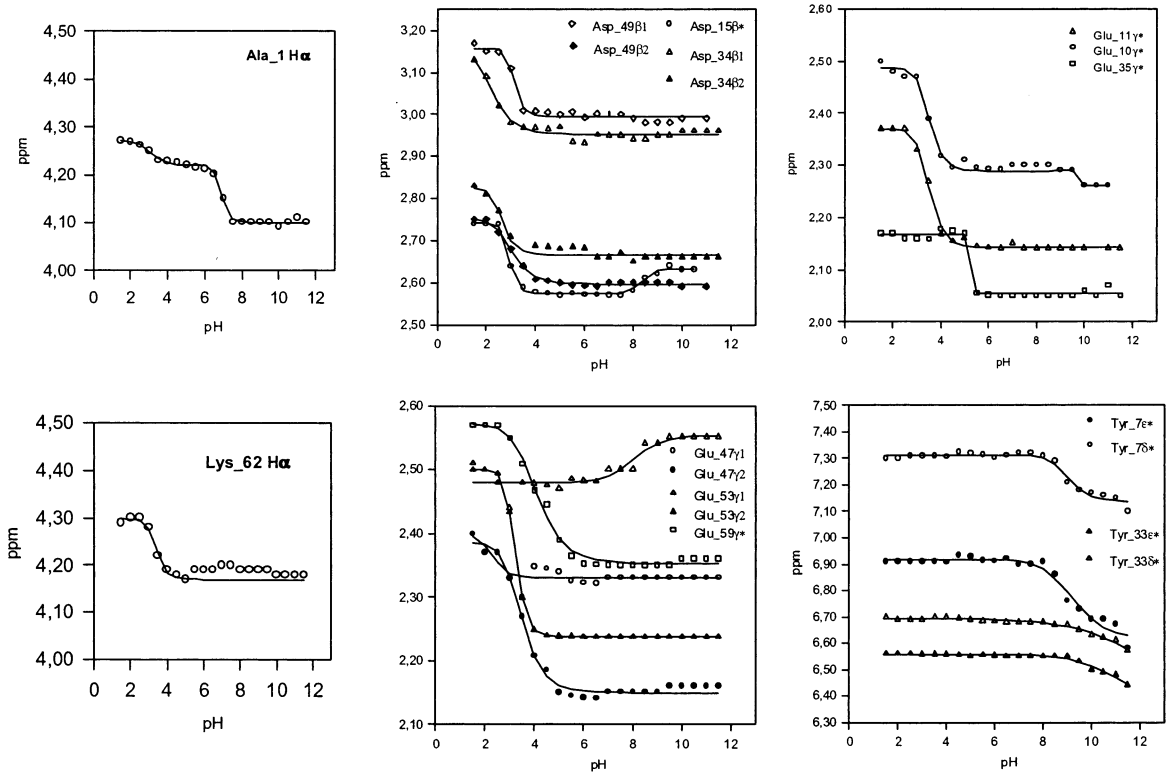

FIGURE 1: pH dependence chemical shifts for titratable residues of recombinant Sso7d. Experimental titration curves as determined by ^{1}H NMR spectroscopy in the pH range of 1.5−11.5.

Glu35 align with Tyr34 and Asp36, respectively) (PDB entries 1SAP and 1AZQ).

Electrostatic potential at solvent accessible surface has been computed for the protein with all basic groups protonated and all acidic groups deprotonated, except where explicitly stated. All details about electrostatic calculations are essentially as previously described (*19*).

RESULTS AND DISCUSSION

pH Titration. The advantage of monitoring pH titration experiments via proton NMR is that the detailed electrostatic properties of the protein can be attained by using all protons as probes distributed all over the protein globule. This approach represents a great help in the description of substrate recognition and catalytic processes.

On the basis of our previous NMR investigations on Sso7d, ~500 assigned protons have been monitored at different pHs; for most proton resonances, the chemical shift differences between the consecutive spectra were small, allowing assignments by comparison with the spectrum under the reference condition. Resonances displaying small chemical shift changes (<0.07 ppm) were considered pH-independent and discarded. According to these criteria, 72 of all the analyzed protons were considered pH-dependent, and for these, sigmoid titration curves were determined. Most of the titration curves presented simple sigmoid shapes between the ionized and neutral states of the titratable groups. On the basis of the intrinsic pK_a values, 11 residues were expected (*20*) to titrate in the pH range that was analyzed (pH 1.5−11.5). These were Glu10, Glu11, Glu35, Glu47, Glu53, Glu59, Asp15, Asp34, Asp49, and C- and N-terminal groups (Figure 1).

The pK_a of a given ionizable residue is determined from the chemical shift variation of protons close to the titrating group. In addition to this effect, other protons not adjacent in the covalent structure to any titratable group may feel other titrating groups, which are spatially close. It is therefore useful to distinguish the pK_a for each proton belonging to its titratable side chain of a given residue and the pK_a'' which reflects the pK_a of nearby residues, affecting the chemical shift of other protons of the same residue. In this way, we can monitor the ionization of titrating nearby groups on protons of different residues but in close contact. These values are summarized and compared to the theoretical ones in Table 1. In particular, pK_as are calculated for residues with a $\Delta\delta$ of $\geq$0.07, while pK_a'' values are indicated only for residues affected by other titrating groups, showing a $\Delta\delta$ of $\leq$0.07.

On the basis of the pK_a values, it was possible to sort all the acidic ionizable residues into a small number of groups localized in well-defined regions of the protein surface.

Low-pH Titrating Groups (Glu11, Asp15, Asp49, and Asp34). Residues Asp15, Asp49, and Asp34 exhibit particularly low pK_a values, 2.9, 3.0, and 2.6 respectively, while Glu11 exhibited a sigmoidal titration curve at pH 3.5. These pK_a values together with structural analysis indicate the presence of both a salt bridge and a H-bond involving Asp15 in the middle and Lys18 and Ser17 on opposite sides, as confirmed by pK_a'' values of 2.9 and $\cong$3 for Ser17 and Lys18, respectively. The titration of the N-terminal residue, in close contact with Asp15 and Ser17, is reflected on both residues giving pK_a'' values of 8.5 and 7.8, respectively. A salt bridge is present between Asp49 and Lys6 and between Asp34 and Lys20, again confirmed by pK_a'' values observed for the two lysines. No salt bridge was observed for Glu11,

Table 1: Ionization Constants[a]

residue	pK_a	pK_a''	pK_a^{theo1}	pK_a^{theo2}	pK_a^{theo3}	pK_a^{theo4}	pK_a^{theo5}
N-terminus	7.0	3.0	8.0	7.6	6.9	6.5	7.7
E10	3.5	9.7	3.5	4.1	3.9	3.9	3.4
E11	3.5	ND	4.0	3.9	4.2	3.9	3.7
D15	2.9	8.5	2.7	2.9	2.6	2.5	3.1
D34	2.6	4.6	2.7	3.0	3.8	2.3	2.4
E35	5.4	ND	4.2	3.8	0.8	3.2	2.9
E47	3.5	ND	4.4	3.9	2.8	3.5	3.0
D49	3.0	ND	2.2	2.6	2.9	3.2	2.5
E53	3.2	8.0	3.3	3.7	4.5	3.8	3.6
E59	4.1	ND	3.9	4.2	*c*	4.2	*c*
C-terminus	3.4	ND	3.1	3.0	4.0	3.6	3.3
Y7[b]	>9.3	ND	9.6	9.8	9.6	9.0	9.1
Y33[b]	>10.0	4.8	11.0	10.6	13.6	11.9	11.2
K6	>11.5	3.7	11.6	11.1	10.4	10.8	10.5
K12	>10.0	3.5	11.4	12.1	10.7	11.3	12.5
K18	>10.0	<3	11.3	11.0	10.8	11.1	11.5

[a] pK_a intrinsic values of ionizable groups based on NMR experiments. pK_a'' stands for values due to spatial proximity of other ionizable side chains. pK_a^{theo} stands for theoretical values determined following the methodology of Antosiewicz et al. (*13*) for different Protein Data Bank structures; in particular, pK_a^{theo1}, pK_a^{theo2}, pK_a^{theo3}, pK_a^{theo4}, and pK_a^{theo5} are calculated for recombinant Sso7d (1JIC), natural Sso7d (1SSO), recombinant Sac7d (1SAP), the Sso7d−DNA complex (1BBX), and the Sac7d−DNA complex (1AZP), respectively. Both pK_a and pK_a'' have been obtained by fitting proton chemical shift titration curves. The standard deviation error for the theoretical values is ±0.7. ND stands for not determined. [b] Obtained for ϵ-aromatic protons. [c] Noncharged residue present in the Sac7d sequence.

at least with distances of <4 Å between the side chain carboxyl oxygen atoms and the nitrogen atoms of neighboring arginines or lysines. If distances between the nitrogen atoms of neighboring lysines and arginines of up to 6 Å were considered, an extensive network of salt bridges and H-bonds could involve nine residues, with Glu11 located between Lys4 and Lys6, confirmed by a pK_a'' value of 3.7 for Lys6 (see Figure 2A and Table 1).

On the basis of the low pK_a values for Asp15, Asp49, and Glu11, the network seems to be present over the complete range of pHs that was analyzed since these acidic residues remain in their anionic state down to very low pHs. Most likely, the extensive network is only transiently populated, resulting in an averaged NMR structure whose NOE and coupling constant pattern alone do not reflect all of the salt bridge and H-bond interactions that can be monitored by the much more sensitive titration shift measurements. It is possible that the complete network could be formed given minor changes in the NMR structures. The network, if present, could confer high stability to the protein, being exposed to solvent and connecting the double β-strand to the helix. This hypothesis is in good agreement with the suggestion by Graziano et al. (*21*), according to which two carboxyl groups possessing very low pK_a values are present in the native structure of Sso7d, probably involved in salt bridges on the surface of the protein.

Asp34 exhibits a very low pK_a value (2.6), in good agreement with its involvement in an isolated salt bridge with Lys20, combined with a H-bond with Thr40, which is in turn close (around 4 Å) to Thr32 (Figure 2B).

This salt bridge seems to confer high stability to the protein, being at the surface, exposed to solvent, and holding together the triple-β-strand itself.

C-Terminal Titrating Groups (Glu47, Glu53, and Glu59). This group is confined in the C-terminal region where the helix, which extends from residue 47 up to residue 57, is divided into two small parts by Pro51. These two small helical segments are anchored to the triple- and double-β-sheet elements probably by salt bridges (Glu47−Arg24, Asp49−Lys6, Glu53−Ala1, and Glu59−Arg24), giving rise to a compact globular structure (Figure 2C).

Indeed, $H_{\gamma 1}$ of Glu53 exhibits a pK_a'' of 8.0, due to the titration of the salt-bridged Ala1, while titration of Glu59 is reflected over the H_δ protons of Arg24, with a pK_a'' of 4.1, not reported in Table 1, thus confirming the presence of a salt bridge. The two H_γ protons of Glu47 experience different behaviors upon pH titration; the chemical shift of one of them is essentially independent of pH over the entire pH range, while that of the other one shows a sigmoidal titration with a pK_a value of 3.5.

Glu10 and Glu35. The side chains of Glu10 and Lys12 are oriented in a parallel manner interacting with a possible salt bridge, thus explaining a pK_a'' of 3.5 for Lys12 and a pK_a'' of 9.7 for Glu10.

A particular consideration is necessary for Glu35, which exhibits a pK_a of 5.4, a value approximately 1−2 pH units above typical pK_as of glutamate in folded proteins. The carboxyl group of this residue is most likely involved in a H-bond with the hydroxyl group of Tyr33, which resulted in some trajectories of dynamics for the recombinant Sso7d; furthermore, Tyr33 H_N and $H_{\beta 1}$ protons exhibited pK_a'' values of 4.4 and 4.8, respectively, strongly supporting the presence of a transient side chain H-bond interaction (Figure 2D).

The resulting stabilization of the protonated form of Glu35 should be responsible for the increase in pK_a of ∼1−2 units. The experimental value of 5.4 for the pK_a of this residue makes it a candidate for playing a prominent role in the ribonucleolytic activity of the protein. For this reason, we further investigated the effect of the protonation state of residue Glu35 on the surface potential of the protein. Most interestingly, mutants E35Q and E35L were almost completely inactive and completely inactive, respectively (*22*).

Electrostatic Calculations. The results of theoretical pK_a shift computation are reported in Table 1. There is a general agreement between computed and measured pK_a values, taking into account the approximations involved in the computation. The largest discrepancy between experimental and theoretical values is observed for Glu35. This prompted us to examine, through more accurate computational protocols, different molecular models which could be compatible with such a large shift in pK_a. Therefore, we have additionally computed the free energy of protonation of Glu35 assuming that a hydrogen bond between Glu35 and the nearby Tyr33 is present and that the molecular environment around Glu35 is not easily permeated by water molecules. For this reason, we used an inner dielectric of 4.0 and we kept all other titratable groups in their charged state. The calculation was performed (as described in Materials and Methods) for all the PDB structures considered here for Sso7d (PDB entries 1JIC, 1SSO, and 1BBX) and for Sac7d (PDB entries 1AZQ and 1SAP) imposing a hydrogen bond between the carboxylic proton of Glu35 (Asp36 for Sac7d) and the phenolic OH of Tyr33 (Tyr34 for Sac7d). Theoretical pK_a shifts ranging from 0.0 for 1BBX to 1.1 for 1JIC have been obtained for Sso7d Glu35, which are fully compatible with the observed value of 5.4. For Sac7d, pK_a shifts ranging from −1.9 for 1SAP to 0.2 for 1AZP were found. The large range of shifts

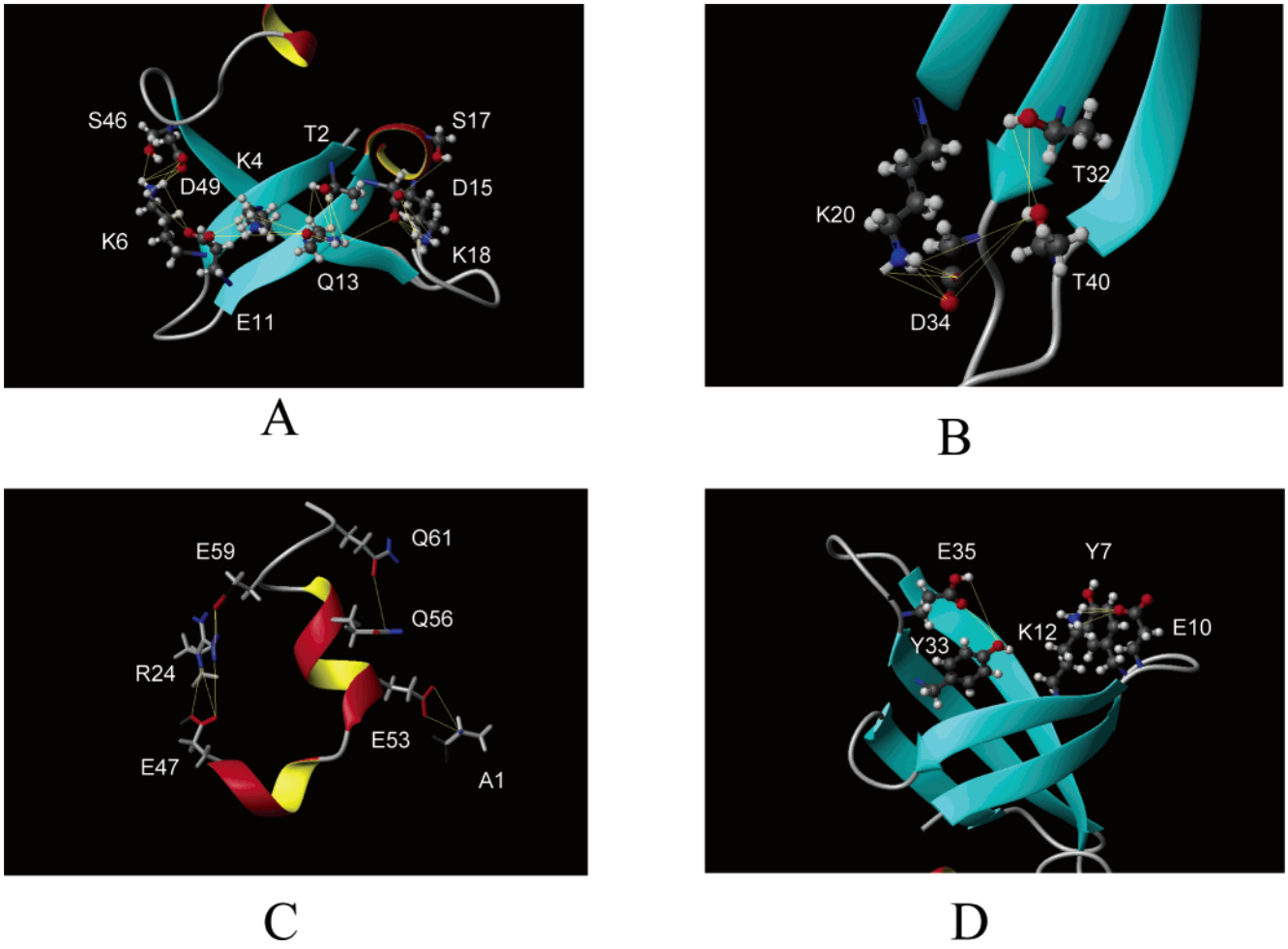

FIGURE 2: Recombinant Sso7d salt bridge networks. Spatial representation (yellow lines) for group 1 (A and B), group 2 (C), and group 3 (D). All the represented distances are taken to be within the range of 4−6 Å.

computed for different structures of the same molecule has been further examined by closer inspection of pK_a shift determinants. The calculated shifts correlate very well with the electrostatic microenvironment of Glu35 (Asp36 in Sac7d). The presence of the imposed hydrogen bond is responsible for the pK_a shift toward higher values, while the proximity of other positive charges, like those due to hydrogens belonging to polar groups or to positively charged groups, tends to shift the pK_a toward lower values. Consistent with the titration calculation, the structure with PDB entry 1SAP for Sac7d exhibits the largest downward shift (−1.9 pK_a units) for Asp36, and indeed, the carboxylic group is involved in two salt bridges. The same salt bridges lead to a computed pK_a value of 0.8 in the standard pK_a shift computation. Very large computed pK_a shifts are a well-known artifact of this computational methodology and can be ascribed to the assumption of a rigid molecular model. The largest upward pK_a shift is exhibited by 1JIC for Sso7 which displays a shift of 1.1 pK_a units. In this structure, not only are salt bridges absent but also the HH hydrogen of Tyr33 is pointing away from the carboxylic group, which is not the case for the other structural models. This discussion highlights the difficulties in making any detailed prediction for the behavior of groups the microenvironment of which is different in different PDB structures. It is worth noticing that it is likely that calculated shifts are artifactual in these cases because hydroxyl hydrogens are usually not observed by NMR and never observed by X-ray crystallography. Moreover, for NMR structures, in the absence of restraints, possible salt bridges will be present or absent according to the force field used for the final in vacuo restrained energy minimization usually performed in structure determination. Far from being predictive, these calculations provide, however, a rationale for the observed unusually large upward pK_a shift at Glu35. Indeed, the calculations strongly support the presence of a hydrogen bond with Tyr33, the absence of salt bridges involving the same group, and the absence of any other polar hydrogen close to the carboxylic hydrogen. It will be interesting to test whether Sac7d exhibits a similar shift at Asp36 (which was not found in our calculations for any structure) and whether this behavior correlates with the absence of RNAse activity. Note, however, that the salt bridges found at Asp36 for the 1SAP structure include basic residues involved in salt bridges with other residues both in other Sac7d structures and after alignment in other Sso7 structures.

Photo-CIDNP Experiments. In our previous study (*36*), recombinant Sso7d was characterized by means of ^{1}H NMR and photochemically induced dynamic nuclear polarization spectroscopy (photo-CIDNP) by using a photoexcitable flavin dye, the structure of which is reminiscent of that of a nucleic acid base. In those experiments, an unusually large amount of flavin was necessary to observe a yellow fluorescent solution, suggesting the existence of specific interactions

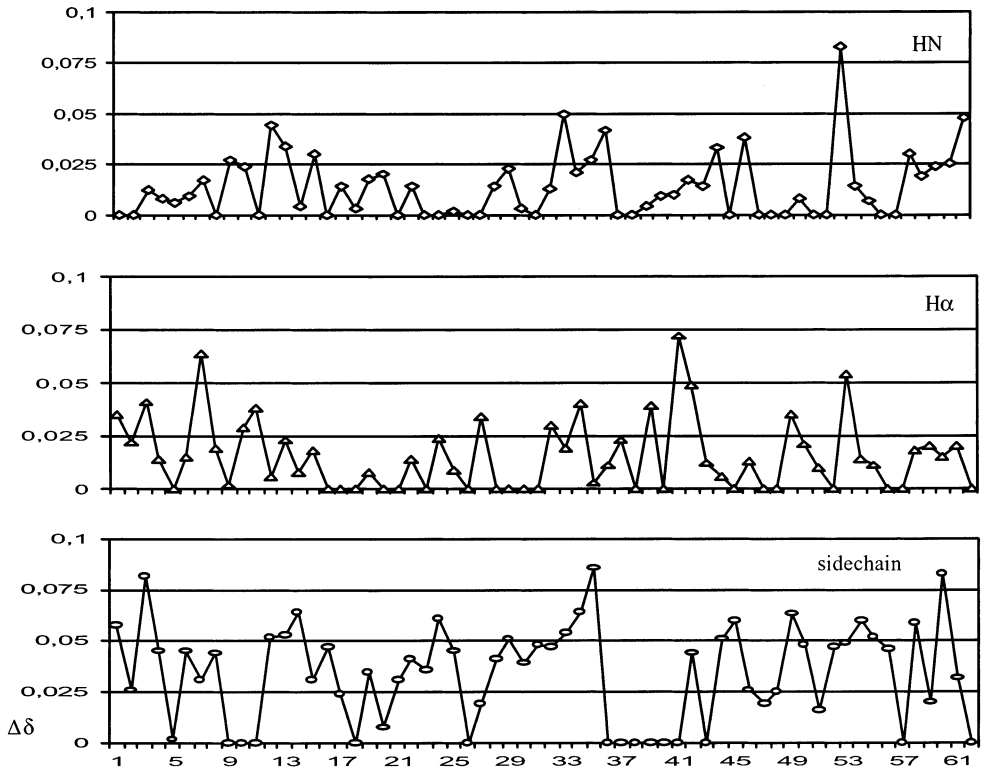

FIGURE 3: Chemical shift perturbations observed upon UpA titration. Amide, α, and side chain protons are represented; residues with chemical shift perturbation (Δδ) above 0.05 are considered (see the text).

between the flavin dye and the protein. In particular, all signals of Trp23 and Phe31 were strongly upfield shifted and broadened, indicating a specific interaction between the dye and each of the two aromatic moieties. In contrast, Phe5, Tyr7, and Tyr33 did not seem to be appreciably affected by the flavin, either in the chemical shift or in the line width. Inspection of the solution structure confirmed the complete exposure of Trp23, while Tyr7 and particularly Phe31 were part of the aromatic buried cluster, but nevertheless accessible to the dye. This result suggested, therefore, that a putative nucleotide involved in the cleavage of a phosphodiester bond could interact with the aromatic cluster, probably via ring stacking.

Dinucleotide Titration. Following the above preliminary suggestion, we performed Sso7d titration experiments with a dinucleotide RNA model, UpA. The choice of this dinucleotide was based on the observation that A+U sequences are preferentially cleaved, over other sequences, by Sso7d. The fact that Sso7d recognizes generic features of the substrate rather than any particular cognate sequence suggested that dinucleotides, the smallest possible RNA fragments to be cleaved, may be the most appropriate choice as model substrates for NMR investigations. Analysis of the NMR spectra recorded after each single addition showed the chemical shift perturbation of some selected residues. Side chain protons of Ala1, Val3, Lys12, Gln13, Val14, Arg24, Ile29, Tyr33, Glu34, Glu35, Ala44, Val45, Asp49, Leu54, Leu55, Leu58, and Lys60, α-protons of Tyr7, Gly41, and Glu53, and only one amide proton, Glu53, presented chemical shift variations larger than 25 Hz ($\Delta\delta > 0.05$ in Figure 3).

In particular, side chains of Val3, Glu35, and Lys60, the amide proton of Glu53, and α-protons of Gly41, being the most affected residues, experienced shifts larger than 35 Hz.

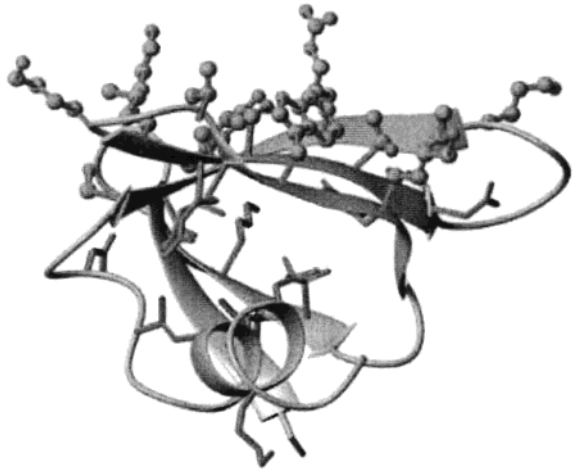

FIGURE 4: Spatial distribution of residues interacting with DNA and UpA. Ball and stick and neon representations for side chains are used, respectively.

Figure 4 clearly shows that all these residues are located on a surface opposite of that involved in DNA binding and furthermore that they are localized in specific secondary structural elements. On the basis of their spatial proximity, it has been possible to distinguish three different groups of residues interacting with the dinucleotide.

The first group, located on the double-stranded β-sheet, contains residues Ala1, Tyr7, Lys12, Gln13, and Val14. The second group is located on the head of the triple-stranded β-sheet and consists of residues Tyr33, Asp34, Glu35, and Gly41, where, in particular, the α-protons of Gly41 are in close contact (<4 Å) with the side chains of Glu35 and Tyr33. The last group is the most populated one and consists of residues Arg24, Ile29, Val45, Asp49, Glu53, Leu54, Leu55, Leu58, and Lys60. These residues are located mainly

on the C-terminal helix and on the tail of the triple-stranded β-sheet.

This titration indicates the possibility of the protein having selective interaction with a dinucleotide by using residues not involved in DNA binding; furthermore, it also shows that these residues are located on a different protein surface with respect to the well-characterized DNA binding surface. Analysis of the electrostatic potential at the surface of the protein shows that, as expected, the most positive and extended potential is found at the surface involved in DNA binding. However, upon protonation of Glu35, a region of positive potential is also present in the open end of the protein, encompassing side chains of residues Lys12, Val14, Tyr33, and Glu35, suggesting that this picture would require a consistent shift in the pK_a of this residue to make it active at physiological pH (vide infra).

As expected, single-point mutants at positions 12 (K12R and K12L) and 35 (E35Q and E35L) were found to be more stable than the wild type, but devoid of RNase activity, while a single-point mutant at position 33 (Y33F) still retains some residual activity (P. Tortora, unpublished observations). This finding agrees with the theory according to which active sites are not optimized for structural stability but for biological function, with the consequence that appropriate mutations can improve protein stability at the expense of its function.

The analysis of the pH dependence of the proton NMR chemical shifts indicated that the side chains of Glu35 and Tyr7 have pK_as of 5.4 and >9.3, respectively. If compared to those expected for random coil polypeptides, these values are unusually high and low, respectively, which is particularly surprising in light of their position at the protein surface. Unusual pK_a values are often observed in a large variety of proteins for many residues residing within an active site and normally implicated in catalytic functions. However, in most RNases, the catalytic mechanism consists of an acid−base catalysis accomplished by two histidines (*23*), and only in some exceptions, like in RNase T1 from the mould fungus *Aspergillus oryzae* (*24*), are histidines replaced with acidic residues, still retaining significant residual enzymatic activity. As in most known RNases, the Sso7d catalytic mechanism is a two-step reaction (*25*−*28*), consisting of a transphosphorylation followed by a hydrolysis, in which cleavage of the phosphodiester bond is achieved via a 2′,3′-cyclic nucleotide intermediate. The capability of Sso7d to hydrolyze 2′,3′-cyclic nucleotides has already been recognized (*29*), so these forms should represent reaction intermediates in archaea bacterial enzymes as well. In accordance with the established mechanisms of Heinemann and Saenger (*30*), a general acid is required for activating the phosphate bond and a basic residue for deprotonating the 2′-hydroxyl group of the ribose ring.

On the basis of the above titration results supported by the theoretical calculations, it seems that Glu35 can play a predominant role in the catalytic reaction, acting as a proton acceptor, due to its high pK_a value. At 60 °C, the maximal activity is in the pH range of 6.7−7.7; Glu35 exists in its ionized form and can act as a nucleophile, taking a proton from the RNA ribose ring to be cleaved. No other acidic residue exhibited a pK_a suitable for general acid behavior. However, Tyr7 seems to be a good candidate for such a role, due to its very low pK_a value. Structurally, Glu35 and Tyr7 are on opposite sides of the putative active center and have

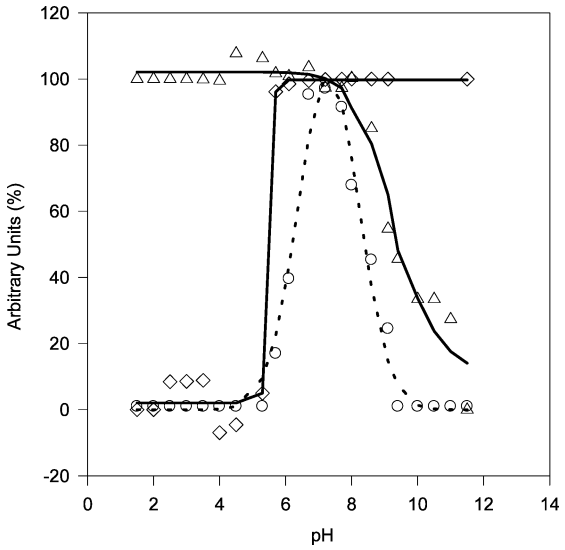

FIGURE 5: Comparison between the activity profile of Sso7d (○) and the pH titration curves for Glu35 (◇) and Tyr7 (△).

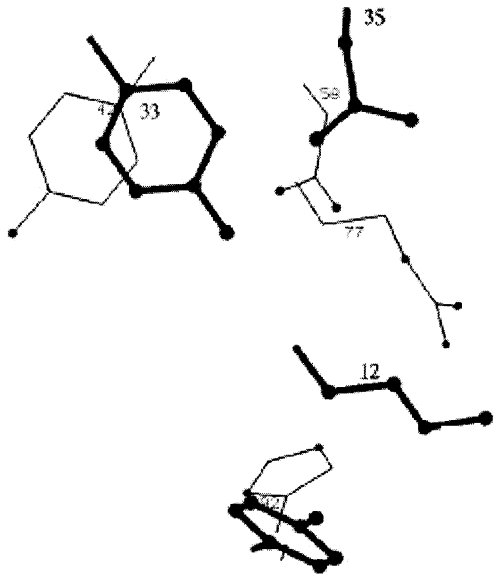

FIGURE 6: Superposition of the RNase T1 active site (lines) with the putative active site of Sso7d (balls and sticks). Heteronuclei of RNase T1 active site residues are represented by small balls.

correct geometry for the in-line cleavage of RNA. Moreover, the phenolic OH moiety of Tyr7 is exposed at the surface, available for a proton donation, as indicated by the exposure results obtained via photo-CIDNP experiments.

The pK_a values for the two hypothesized catalytic residues of Sso7d differ by 3.9 units, allowing the protein to be active over a wide pH range between 5.4 and 9.3, in remarkable agreement with the pH−activity profiles of the protein (Figure 5).

Indeed, a very interesting comparison can be made between RNase T1 and RNase A. By superposing C_β and C_γ of residues Glu58 and His92 of RNase T1, assumed to play a role in the catalytic reaction for RNase T1, with C_β and C_γ of residues Glu35 and Tyr7 of Sso7d, respectively, we observe very good agreement (Figure 6, rmsd = 1.6 Å).

In addition, Arg77 and Tyr42 of RNase T1 are aligned with Lys12 and Phe33 of Sso7d, respectively; the resulting geometry of the Sso7d putative active site is interestingly reminiscent of that of the RNase T1 active site. In contrast,

His40 of RNase T1 which has been shown to be not indispensable (*31*) and capable of acting as a base in place of Glu58, when replaced, resulted not overimposed with any putative residues of Sso7d. The comparison between Sso7d and some RNase T1 mutants is even more interesting since significant activity in the double mutant His92,His40/Asp has been demonstrated (*32*). This suggested that the catalytic reaction of RNase T1 can be carried out by residues capable of proton exchange, provided they are properly positioned as shown by the superposition in Figure 6.

In the RNase A cleavage reaction, a lysine residue (Lys41) facilitates catalysis by donating a hydrogen bond to the cyclic transition intermediate (*33*). A single-point mutant at this position decreases the activity by a factor of 10^4. The same holds true for Lys12 in Sso7d, as substantiated by the complete loss of catalytic activity resulting from the single-point mutant K12L (*22*). Most likely, the same "supporting" role is played by Tyr33, substitution with phenylalanine of which still retains residual activity.

In Sso7d for three residues, we have measured pK_a values lower than their intrinsic values, namely, Asp15, Asp34, and Asp49, the pK_a values of which range from less than 3 to less than 2. This finding suggests that the native structure favors the ionized side chains over the neutral protonated species and is consistent with the presence of networks of salt bridges and H-bonds. This was also suggested on the basis of thermodynamic measurements (*21*) and is in agreement with the analysis of a molecular simulation study on the homologous protein Sac7d from *S. acidocaldarius*, according to which three large clusters are present rather than individual salt bridges (*34*).

CONCLUSIONS

Our pH titration data highlighted the role of Glu35 and Tyr7 in the RNase activity of the recombinant Sso7d, as suggested by their unusual pK_a values, 5.4 and ~9.3, respectively. Comparison of the putative catalytic residues of Sso7d with those of RNase T1, containing only one histidyl residue, or those of its mutant in which the catalytic histidine has been substituted with an aspartic acid, still maintaining RNase activity, indicates a very good similarity. This agreement suggests that in the archaeum protein the two catalytic residues are located in a suitable geometry in which they are opposite each other and have a correct position for the in-line cleavage of RNA. Interestingly, the overall putative active site of Sso7d, reminiscent of the RNase T1 active site, showed two additional residues, Arg77 and Tyr42, aligned with Lys12 and Tyr33 of Sso7d, respectively.

The results of theoretical pK_as calculated for Sso7d and Sac7d, a highly homologous protein devoid of RNase activity, showed large pK_a shifts for Glu35 (1.1) and Asp36 (−1.9), respectively. It is tempting to correlate the large difference in pK_a shift to the absence of RNase activity in Sac7d.

With regard to RNase proteins from the archea, both Sso and Sac families (Sso7d and Sac7e, respectively) have shown poor similarity in the amino acid sequence with other known proteins, thus suggesting their independent evolution as well as DNA binding capacity and, whenever present, RNase activity. On the basis of the results of Edmonson and Gao (*3*, *35*), the structure of the protein−DNA complex showed the complete absence of any sequence specificity, placing the triple-stranded β-sheet across the DNA minor groove. The insertion of hydrogen bond-donating side chains, additionally stabilized by nonpolar and electrostatic interactions with the DNA backbone, explains the strong binding that can be achieved, causing a distortion of the DNA conformation and introducing significant unwinding of the helix. On the basis of our results, residues affected by UpA dinucleotide interaction are mainly located on a different surface of the protein with respect to that used for DNA recognition. Electrostatic calculations indicate the presence of a highly positive uniformly charged surface localized in the triple-stranded β-sheet-exposed surface, conferring the suitable potential for DNA recognition. However, as shown by our previous results (*29*), a saturating amount of DNA did not affect the digestion pattern of CI_{83} RNA by Sso7d. This result is in agreement with our data, confirming that the putative catalytic site is localized in a separate distinctive surface of the protein.

ACKNOWLEDGMENT

F. Greco and G. Zannoni are acknowledged for their helpful technical support.

REFERENCES

1. Robinson, H., Gao, Y. G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (1998) *Nature 392*, 202−205.
2. Agback, P., Baumann, H., Knapp, S., Ladenstein, R., and Hard, T. (1998) *Nat. Struct. Biol. 5*, 579−584.
3. Gao, Y. G., Su, S. Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (1998) *Nat. Struct. Biol. 5*, 782−786.
4. Kulms, D., Schafer, G., and Hahn, U. (1997) *Biol. Chem. 378*, 545−551.
5. Kannan, N., and Vishveshwara, S. (2000) *Protein Eng. 13*, 753−761.
6. Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M., and Makhatadze, G. I. (1999) *Biochemistry 38*, 16419−16423.
7. Gromiha, M. M. (2001) *Biophys. Chem. 91*, 71−77.
8. Burley, S. K., and Petsko, G. A. (1985) *Science 229*, 23−28.
9. Karshikoff, A., and Ladenstein, R. (2001) *Trends Biochem. Sci. 26*, 550−556.
10. Fusi, P., Tedeschi, G., Aliverti, A., Ronchi, S., Tortora, P., and Guerritore, A. (1993) *Eur. J. Biochem. 211*, 305−310.
11. Ebina, S., and Wüthrich, K. (1984) *J. Mol. Biol. 179*, 283−288.
12. Bashford, D., and Karplus, M. (1990) *Biochemistry 29*, 10219−10225.
13. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1994) *J. Mol. Biol. 238*, 415−436.
14. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−971.
15. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
16. De Marco, A. (1977) *J. Magn. Reson. 26*, 527−528.
17. Bundi, A., and Wüthrich, K. (1979) *Biopolymers 18*, 285−297.
18. Edsel, J. T., and Wyman, J. (1958) *Biophysical Chemistry*, Academic Press, New York.
19. Fogolari, F., Ragona, L., Licciardi, S., Romagnoli, S., Michelutti, R., Ugolini, R., and Molinari, H. (2000) *Proteins: Struct., Funct., Genet. 39*, 317−330.
20. States, D. J., and Karplus, M. (1987) *J. Mol. Biol. 197*, 122−130.
21. Graziano, G., Catanzano, F., and Nappa, M. (1999) *Int. J. Biol. Macromol. 26*, 45−53.
22. Shehi, H., Serina, S., Fumagalli, G., Vanoni, M., Consonni, R., Zetta, L., Dehò, G., Tortora, P., and Fusi, P. (2000) *FEBS Lett. 497*, 131−136.
23. Walsh, C. (1979) *Enzymatic Reaction Mechanisms*, W. H. Freeman and Co., San Francisco.
24. Landt, O., Thölke, J., Grunert, H. P., Saenger, W., and Hahn, U. (1997) *Biol. Chem. 378*, 553−558.

25. Findlay, D., Herries, D. G., Mathias, A. P., Rabin, B. R., and Ross, C. A. (1961) *Nature 190*, 781−784.
26. Findlay, D., Herries, D. G., Mathias, A. P., Rabin, B. R., and Ross, C. A. (1962) *Biochem. J. 85*, 152−153.
27. Cuchillo, C. M., Parés, X., Guasch, A., Barman, T., Travers, F., and Nogués, M. V. (1993) *FEBS Lett. 333*, 207−210.
28. Thompson, J. E., Venegas, F. D., and Raines, R. T. (1994) *Biochemistry 33*, 7408−7414.
29. Fusi, P., Grisa, M., Tedeschi, G., Negri, A., Guerritore, A., and Tortora, P. (1995) *FEBS Lett. 360*, 187−190.
30. Heinemann, U., and Saenger, W. (1982) *Nature 299*, 27−31.
31. Grunert, H. P., Zouni, A., Beuineke, M., Quaas, R., Georgalis, Y., Saenger, W. J., and Hahn, U. (1991) *Eur. J. Biochem. 197*, 203−207.
32. Landt, O., Thölke, J., Grunert, H. P., Saenger, W., and Hahn, U. (1997) *Biol. Chem. 378*, 553−558.
33. Messmore, J. M., Fuchs, D. N., and Raines, R. T. (1995) *J. Am. Chem. Soc. 117*, 8057−8060.
34. De Bakker, P. I. W., Hünenberger, P. H., and McCammon, J. A. (1999) *J. Mol. Biol. 285*, 1811−1830.
35. Edmonson, S. P., Qui, L., and Shriver, J. W. (1995) *Biochemistry 34*, 13289−13304.
36. Consonni, R., Limiroli, R., Molinari, H., Fusi, P., Grisa, M., Vanoni, M., and Tortora, P. (1995) *FEBS Lett. 372*, 135−139.

BI0265168