Dissecting Structural and Electrostatic Interactions of Charged Groups in α-Sarcin. An NMR Study of Some Mutants Involving the Catalytic Residues†

Mª Flor García-Mayoral,‡ José Manuel Pérez-Cañadillas,‡,§ Jorge Santoro,‡ Beatriz Ibarra-Molero,|| José Manuel Sanchez-Ruiz,|| Javier Lacadena,⊥ Álvaro Martínez del Pozo,⊥ José G. Gavilanes,⊥ Manuel Rico,‡ and Marta Bruix*,‡

Departamento de Espectroscopía y Estructura Molecular, Instituto de Química Física "Rocasolano", CSIC, Serrano 119, 28006 Madrid, Spain, Departamento de Química-Física, Universidad de Granada, 18071 Granada, Spain, and Departamento de Bioquímica y Biología Molecular I, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain

Received June 6, 2003; Revised Manuscript Received September 11, 2003

ABSTRACT: The cytotoxic ribonuclease α-sarcin is the best characterized member of the ribotoxin family. Ribotoxins share a common structural core, catalytic residues, and active site topology with members of the broader family of nontoxic microbial extracellular RNases. They are, however, much more specific in their biological action. To shed light on the highly specific α-sarcin activity, we have evaluated the structural and electrostatic interactions of its charged groups, by combining the structural and pK_a characterization by NMR of several variants with theoretical calculations based on the Tanford−Kirkwood and Poisson−Boltzmann models. The NMR data reveal that the global conformation of wild-type α-sarcin is preserved in the H50Q, E96Q, H137Q, and H50/137Q variants, and that His137 is involved in an H-bond that is crucial in maintaining the active site structure and in reinforcing the stability of the enzyme. The loss of this H-bond in the H137Q and H50/137Q variants modifies the local structure of the active site. The pK_a values of active site groups H50, E96, and H137 in the four variants have been determined by two-dimensional NMR. The catalytic dyad of E96 and H137 is not sensitive to charge replacements, since their pK_a values vary less than ±0.3 pH unit with respect to those of the wild type. On the contrary, the pK_a of His50 undergoes drastic changes when compared to its value in the intact protein. These amount to an increase of 0.5 pH unit or a decrease of 1.1 pH units depending on whether a positive or negative charge is substituted at the active site. The main determinants of the pK_a values of most of the charged groups in α-sarcin have been established by considering the NMR results in conjunction with those derived from theoretical pK_a calculations. With regard to the active site residues, the H50 pK_a is chiefly influenced by electrostatic interactions with E96 and H137, whereas the effect of the low dielectric constant and the interaction with R121 appear to be the main determinants of the altered pK_a value of E96 and H137. Charge−charge interactions and an increased level of burial perturb the pK_a values of the active site residues of α-sarcin, which can account for its reduced ribonucleolytic activity and its high specificity.

The structure, stability, and function of many proteins are governed by the protonation equilibria of ionizable groups (*1*−*5*). Most of these groups exhibit altered pK_a values whose shifts are the result of a variety of possible interactions, such as charge−charge interactions, charge−dipole interactions, hydrogen bonding, desolvation effects, and others. It is well-known that the formation of salt bridges can greatly alter the pK_a values of the participating groups and make substantial contributions to protein stability (*6*). Hydrogen bonding of a charged group with uncharged partners in the folded protein can also perturb its pK_a and reinforce the stability of the native state. Thus, for instance, in ovomucoid third domain, an altered carboxylate pK_a results from hydrogen bonding and not from electrostatic interactions with nearby lysine residues (*7*) as was originally predicted (*8*). Three hydrogen bonds between uncharged groups and buried charged Asp76 in RNase T1 (*9*) and Asp33 in RNase Sa (*10*) severely decrease their pK_a values. Hydrogen bonding between two buried histidines (*11*) and two glutamates near the surface of CD2 protein (*12*) also produces remarkable pK_a changes. The low dielectric constant in the interior of a protein is also responsible for altered pK_a values as elegantly demonstrated by Garcia-Moreno and co-workers by producing Ile to Glu and Ile to Lys mutations at the most buried position of staphylococcal nuclease. The pK_a values of these two residues are highly perturbed by this hydrophobic environment (*13*). Also, a buried cysteine residue in thioredoxin has a pK_a of >11 (*6*) and a buried non-hydrogen-bonded aspartic acid in RNase Sa a pK_a of 7.4 (*10*).

† This work was supported by Grants BIO2002-720 and BMC2000-0551 from the Ministerio de Ciencia y Tecnología and Educación y Cultura (Spain), respectively. M.F.G.-M. is the recipient of a fellowship from the Comunidad Autónoma de Madrid (CAM, Spain). J.M.P.-C. acknowledges the European Union and EMBO for support through their long-term postdoctoral fellowship programs.

* To whom correspondence should be addressed. E-mail: mbruix@iqfr.csic.es. Phone: 34 91 561 94 00. Fax: 34 91 564 24 31.

‡ CSIC.

§ Present address: Center for Protein Engineering, Medical Research Council, Hills Road, Cambridge CB2 2QH, England.

|| Universidad de Granada.

⊥ Universidad Complutense.

10.1021/bi0349773 CCC: $25.00
Published on Web 10/21/2003

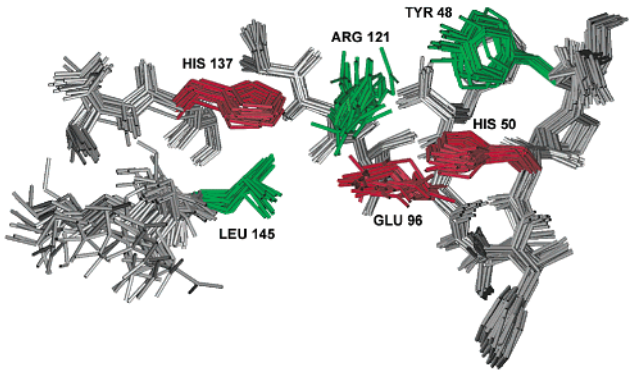

FIGURE 1: Superposition of the active site residues of α-sarcin in 20 NMR structures. Catalytic side chains are depicted in red; side chains for which an additional role has been suggested during the catalytic reaction are in green, and the rest of the side chains and the backbone atoms are in gray. This figure was generated with MolMol (*61*, *62*).

Since the pK_a values of ionizable residues reflect the complex relationship between structure and energetics within a protein, the experimental determination of their values and the structural features that influence them are key in understanding the physicochemical basis of the protein stability and function. Considerable efforts to determine and predict protein pK_a values are being made by experimentalists and theoreticians. NMR spectroscopy is currently the most powerful technique for obtaining individual pK_a values for most of the ionizable groups in biomolecules, and it has been successfully applied to identify catalytic (*14*, *15*), ligand binding (*12*), and structurally relevant (*16*) residues in a number of proteins. Additionally, as NMR signals are sensitive to the changes in the environment of the monitored nuclei, the observed titration curves reflect both proton release or binding and conformational changes that may occur when the pH is changed (*17*). Many charged groups involved in catalysis exhibit altered pK_a values whose shifts result from a variety of possible interactions (*2*, *18*, *19*). In a recent work, the relationship between protein structure and carboxylate group ionization was investigated for 115 aspartic acids and 96 glutamic acids' pK_a values in 24 proteins of known structure (*20*). Interestingly, carboxylate groups located in active sites or ligand binding sites exhibit, in nearly all cases, elevated pK_a values, suggesting that this decreased acidity plays a general functional role in proteins.

α-Sarcin is an extracellular cytotoxin produced by the mold *Aspergillus giganteus* (*21*) whose structure, dynamics, and electrostatics have been extensively characterized (*22*−*26*). This small ribosome-inactivating protein inhibits protein biosynthesis by cleaving a single phosphodiester bond in a strictly conserved RNA sequence of the largest ribosome subunit (*27*, *28*) located in the "sarcin/ricin" loop (*29*, *30*). α-Sarcin is a highly charged protein, with a high isoelectric point. The high content of positively charged residues is likely required for recognizing and binding of its highly negatively charged target. In a previous work, the pK_a values of all aspartic acid, glutamic acid, and histidine residues of α-sarcin were determined by NMR, revealing that many residues, including several at the active site, exhibit highly perturbed pK_a values (*24*). The active site of α-sarcin is composed of residues His50, Glu96, His137, Arg121, and Tyr48 (Figure 1), although only the first three are directly involved in proton transfer steps in the catalytic mechanism. They are located in the central β-sheet, and their side chains point toward the concave face of the protein structure. The most representative characteristics of the α-sarcin active site can be summarized as follows: (i) high density of charged residues, (ii) unusual pK_a values of His50, Glu96, and His137, (iii) unusual Nδ tautomeric forms adopted by His50 and His137, a common feature of microbial RNases, (iv) the presence of a structurally important hydrogen bond between the catalytic His137 and a backbone oxygen in loop 5, and (v) low surface accessibility of all titratable atoms (see Table 1).

Despite recent progress in the field, the origin of pK_a perturbations, both in general and in the case of α-sarcin, remains an open question and more effort is needed for a complete understanding of it. Here we have used two independent approaches to identify the interactions perturbing the pK_a values in α-sarcin. First, we have used computational methods based on the Tanford−Kirkwood model and numerical integration of the Poisson−Boltzmann formalism (*31*−*33*). Although it cannot be said that these methods yield, in general, accurate pK_a values for buried groups, they are useful for identifying interactions that potentially perturb pK_a values as well as for predicting the pK_a values of groups that cannot be measured experimentally. Second, we have systematically measured the pK_a values of titratable residues in a series of variants (E96Q, H50Q, H137Q, and H50/137Q) designed to remove potential interactions that might alter the pK_a values of active site residues. This approach leads to precise pK_a values, with the caveat that their accuracy may be affected by subtle changes in the structure associated with the mutations.

MATERIALS AND METHODS

NMR Samples. Samples were prepared by dissolving the wild-type (WT) protein and four α-sarcin variants (H50Q, E96Q, H137Q, and H50/137Q) in a 90% H_2O/10% D_2O mixture or D_2O, producing ca. 1 mM solutions containing 0.2 M NaCl as previously described (*24*). Gln was chosen as the substitute residue in all mutants to facilitate the analysis and comparison, and because it is similar in size to the replaced side chains. This is important because the active site residues of the wild-type α-sarcin are tightly packed and buried (*26*). In addition to the conservative substitution of Glu with Gln, replacement of His with Gln has also been shown to be conservative because of the similar size together with the ability of Gln to mimic some of the hydrogen bonds of the imidazole ring (*34*). All spectra were acquired on a Bruker AV-600 spectrometer at 35 °C. ^{1}H−^{1}H TOCSY (*35*) and NOESY (*36*) spectra with mixing times of 60 and 50 ms, respectively, were recorded by using standard pulse sequences. The spectra of the mutant variants were assigned at pH 6.0, following the standard NOE-based methodology as well as the reported assignment of the wild-type protein (*37*).

The effect of pH titration on the chemical shifts was determined by analyzing a series of TOCSY and NOESY spectra recorded at pH values ranging from 3.0 to 8.5. Chemical shifts were corrected for the effect of the TSP titration (*38*). Spectra were processed with the Bruker package and analyzed using version 3.3 of the program ANSIG (*39*).

Table 1: Surface Accessibilities and Distances between Side Chain Atoms of Residues of the α-Sarcin Active Site[a]

	His50		Glu96		Arg121	His137	
	Nδ	Nε	$Oε_1$	$Oε_2$	Nε	Nδ	Nε
ASA (%)	51	5	5	3	9	0	10
distance							
His50							
Nδ	−	−	6 ± 1	5 ± 1	6.3 ± 0.4	10.7 ± 0.4	9.2 ± 0.4
Nε	−	−	3.5 ± 0.6	4.8 ± 0.7	4.4 ± 0.6	9.6 ± 0.4	7.8 ± 0.4
Glu96							
$Oε_1$	6 ± 1	3.5 ± 0.6	−	−	4.6 ± 0.9	8 ± 1	6 ± 1
$Oε_2$	5 ± 1	4.8 ± 0.7	−	−	3.9 ± 0.7	7.7 ± 0.6	6.3 ± 0.6
Arg121							
Nε	6.3 ± 0.4	4.4 ± 0.6	4.6 ± 0.9	3.9 ± 0.7	−	7.0 ± 0.6	4.9 ± 0.6

[a] Mean values with their standard deviations over 20 NMR structures (PDB entry 1DE3).

Calculation of the Experimental pK_a Values. pK_a values were determined by measuring the changes in the ^{1}H resonances as a function of pH. Experimental pH titration curves were fitted to different models derived from the Henderson−Hasselbach equation, assuming a rapid equilibrium between protonated and unprotonated forms (*40*). The pK_a values were calculated as the weighted average of the analysis of various titration curves of different protons as previously described (*24*). The typical reproducibility of the pK_a values is comparable to the standard error of fits of that equation to experimental data. All reported errors reflect the precision of the data fitting and do not include the uncertainty in the pH determination of the sample, which is estimated to be ±0.1 unit.

Theoretical pK_a Calculations. pK_a calculations for wild-type α-sarcin were performed on a family of 20 NMR structures (Protein Data Bank entry 1DE3) on the basis of two different theoretical approaches. The first method follows the Tanford−Kirkwood (TK) model (*41*, *42*), corrected for solvent accessibility and with the mean field approximation of Tanford and Roxby (*43*), and was implemented using the computer program written by Ibarra-Molero et al. (*32*). The dielectric constants were 80 and 4 for the solvent and the protein, respectively. We further used a continuum electrostatic approach that relies on the finite difference solution of the Poisson−Boltzmann equation (FDPB) as implemented in the University of Houston Brownian Dynamics software package (UHBD) (*31*, *44*) with the optimized protocol of Wade and co-workers (*33*), which allows for different histidine tautomer forms and contains a protocol for addressing multiple titration problem by cluster analysis. In agreement with the experimental results, His36, His50, and His137 were considered NHδ tautomers, whereas all others were fixed as NHε (*45*) in the UHBD calculations. Two different values, 15 and 78.5, for the dielectric constant in the protein interior were used for either the mostly buried or the mostly solvent-exposed ionizable site (*33*). The dielectric constant for the solvent was set to 78.5. All the calculations were done by matching the experimental conditions (35 °C and an ionic strength of 200 mM). The intrinsic pK_a values used in the calculation arise from measured model compound pK_a values of Nozaki and Tanford (*46*): 3.8 for the C-terminus, 4.0 for Asp, 4.4 for Glu, 6.3 for His, 7.5 for the N-terminus, 9.6 for Tyr, 10.4 for Lys, and 12.0 for Arg. The reported pK_a value is the mean over the individual pK_a values calculated in the 20 NMR structures.

Theoretical pK_a values were also obtained for the titratable groups in the E96Q, H50Q, H137Q, and H50/137Q α-sarcin variants using both TK and UHBD approximations with similar protocols as described above. For the purposes of the electrostatic calculation, the charge deletion mutations were taken into account by simply setting the charge of the mutated residue to zero in the input file of the TK and UHBD programs. TK calculations for these mutants were carried out on the basis of the four lowest-energy NMR structures of wild-type α-sarcin of the 20 proposed. Also, the starting pK_a values used as input of the TK calculation were optimized on the basis of the comparison of experimental and predicted pK_a values for wild-type α-sarcin.

RESULTS

Assignment of ^{1}H NMR Spectra of E96Q, H50Q, H137Q, and H50/137Q α-Sarcin Mutants. To determine if the tertiary structure of the wild-type enzyme (*26*) is preserved in these mutants, beyond small local conformational changes near the mutated site, a structural characterization was performed by NMR. A comparison of the backbone NH and αH chemical shift deviations of E96Q, H50Q, and H137Q variants with respect to the chemical shifts of the wild-type protein is shown in Figure 2. Overall, chemical shift differences ($\Delta\delta = \delta_{mut} - \delta_{wt}$) are small. Profiles corresponding to H50Q and E96Q mutants show that their chemical shifts are within 0.05 ppm of their values in wild-type α-sarcin. More significant variations are found in residues near the mutated one. Thus, for these variants, the results indicate that the global fold is preserved with minor rearrangements at the mutated sites. A rather different situation is observed for the H137Q mutant. The magnitude of the chemical shift changes and the large number of protons affected by this mutation suggest there are larger changes in the environment of the mutated residue than in the other two cases. Important differences in the Hα chemical shifts are located not only in the vicinity of Gln137 but also in all residues belonging to loop 5 (139−143), strand β_5 (120−124), and residues in the N-terminal β-hairpin (2−26). The largest deviations are found in the protons of residues next to Gly143: 142 Hα, 0.34 ppm; 142 HN, 0.24 ppm; 143 Hα, −0.22 ppm; 143 HN, −0.17 ppm; and 144 HN, −0.34 ppm. For H50/137Q, the observed changes approximately correspond to the sum of effects of the individual mutations (data not shown). The large number of protons affected by the two substitutions throughout the protein sequence means that the possibility that significant structural rearrangements have occurred in this mutant cannot be ruled out.

Qualitative information about the side chain conformation of the mutated residues can also be obtained from NMR data.

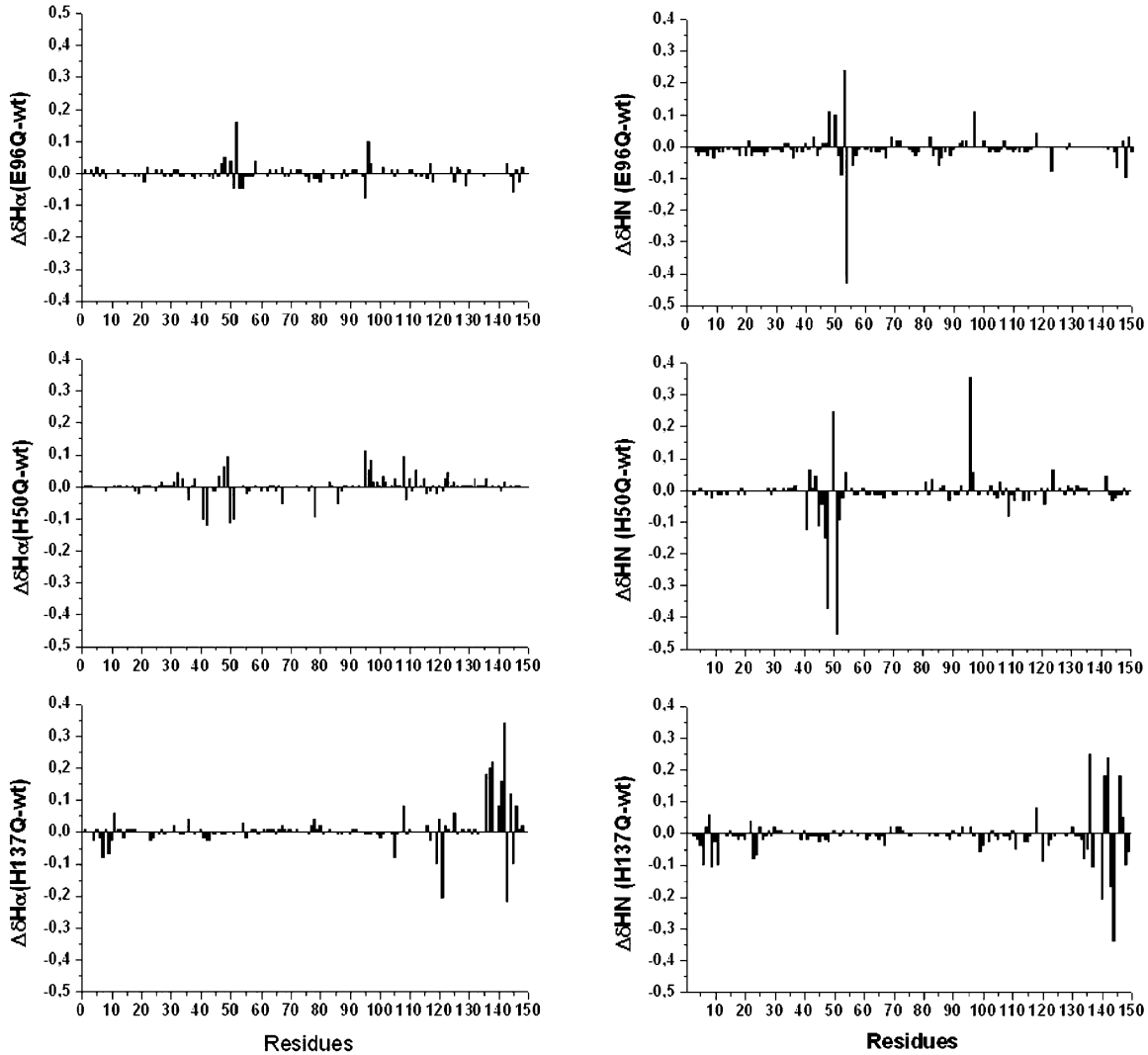

FIGURE 2: Chemical shift differences of Hα (left) and HN (right) protons of E96Q, H50Q, and H137Q mutants with respect to those in wild-type α-sarcin.

Table 2: Experimental pK_a Values of the Histidines and Glutamic Acid 96 of the Wild Type and E96Q, H50Q, H137Q, and H50/137Q α-Sarcin Mutants

protein	H35	H36	H50	H82	H92	E96	H104	H137	H150
α-sarcin[a]	6.3 ± 0.2	6.8 ± 0.2	7.7 ± 0.2	7.3 ± 0.1	6.9 ± 0.1	5.2 ± 0.1	6.5 ± 0.2	5.8 ± 0.1	7.6 ± 0.1
E96Q	6.3 ± 0.2	7.3 ± 0.2	6.4 ± 0.2	nd[b]	7.2 ± 0.1	−	nd[b]	5.9 ± 0.1	7.9 ± 0.1
H50Q	6.2 ± 0.2	nd[b]	−	7.4 ± 0.1	7.3 ± 0.1	5.0 ± 0.1	nd[b]	6.1 ± 0.1	7.3 ± 0.1
H137Q	6.3 ± 0.2	nd[b]	>8	7.4 ± 0.1	7.1 ± 0.1	5.7 ± 0.1	6.6 ± 0.2	−	7.7 ± 0.1
H50/137Q	6.3 ± 0.2	7.3 ± 0.2	−	7.0 ± 0.1	7.4 ± 0.1	5.4 ± 0.1	7.3 ± 0.2	−	7.6 ± 0.1

[a] Values from ref *24*. [b] Not determined because of signal overlap.

The chemical shifts of the NH_2 protons of Gln at positions 50, 96, and 137 in the different mutants are 7.48 and 6.69 ppm for Gln96 in E96Q, 6.68 and 6.54 ppm for Gln50 in H50Q, 7.32 and 6.50 ppm for Gln137 in H137Q, and 6.50 and 6.40 ppm for Gln50 and 7.29 and 6.53 ppm for Gln137 in H50/137Q. Gln96 and Gln137 side chains have chemical shifts in the range of the random coil values (7.52 and 6.85 ppm) (*47*). On the other hand, the Gln50 ϵ-protons have chemical shifts shifted upfield with respect to the random coil values, probably due to the ring current effects of the neighboring Tyr48. In addition, none of these hydrogen atoms show protection against solvent exchange, indicating that they do not participate in hydrogen bonds (*48*).

Titration Curves and pK_a Values of H50Q, E96Q, H137Q, and H50/137Q α-Sarcin Mutants. The pK_a values of all histidine residues and glutamic acid 96, obtained after fitting of all curves, are summarized in Table 2. In general, these mutations do not produce large changes in the pK_a values of histidines not located at the active site. Thus, His35, His82, and His150 maintain the same pK_a value in all the variants, and the pK_a value of His104 is shifted only in the double mutant H50/137Q (0.8 pH unit). Finally, the pK_a values of His36 and His92 are slightly increased in all variants (0.2−0.5 pH unit); this effect is more pronounced for His92. H50/137Q is the variant with the largest pK_a shifts: 0.5 pH unit for H36, 0.3 pH unit for H82, 0.5 pH unit for H92, 0.8 pH unit for H104.

However, remarkable differences in both the shape of the curves and the pK_a values with respect to those of the wild-type protein were observed for the active site groups (Figure

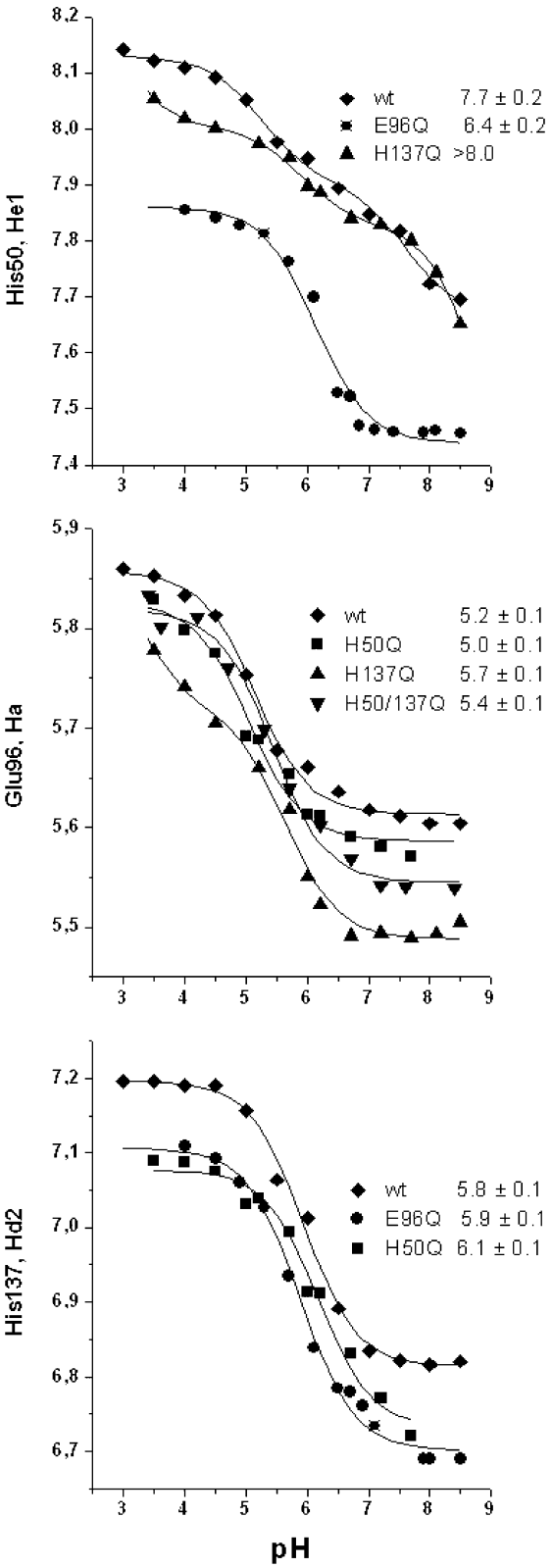

FIGURE 3: Chemical shift titration curves of different protons in H50Q, E96Q, H137Q, and H50/137Q α-sarcin mutants. Data points correspond to the experimental data, and the solid lines are the least-squares best fits to the data.

3). The titration of His50 was followed by observing the chemical shift of its Hδ_2 and Hϵ_1 protons. In the native protein, two transitions are clearly apparent in the His50 Hϵ_1 curves, which can be attributed to the pK_a of its own imidazole group and to the carboxylate group of its neighbor Glu96. As expected, in the E96Q mutant, the His50 data fit a single titration group model very well. The single pK_a observed (6.4) that can be attributed unambiguously to His50 decreases by more than 1 unit compared to the one in the wild-type protein (7.7). On the other hand, titration of His50 in the H137Q mutant shows, in addition to the titration corresponding to Glu96 (pK_a = 5.7), two incomplete inflections at low and high pH. The one at the basic end (pK_a > 8) can be attributed to its own pK_a.

As in wild-type α-sarcin, the titration of Glu96 in H50Q, H137Q, and H50/137Q mutants could be followed by the chemical shift of its HN, Hα, and Hγ protons. In the H50Q and H50/137Q variants, the experimental data could be fitted to a single titration model. On the contrary, in the H137Q mutant, an additional incomplete inflection in the titration was detected at acidic pH, as can be seen in the Hα titration (Figure 3). The pK_a value of Glu96 is 0.2 pH unit lower in the H50Q mutant and 0.5 and 0.2 pH unit higher in the mutants containing Gln at position 137 than in the wild-type protein.

The titration curves of His137 were determined from its δ_2 and ϵ_1 protons, and those of the E96Q and H50Q mutants are quite similar to that of the wild-type protein both free and complexed with 2′GMP (*24*). All of them fit a model with a single titrating group well. The pK_a value of His137 remains unchanged compared to that of the wild-type protein, indicating that the substitution of His50 or Glu96 does not substantially affect the electrostatic environment of His137.

The fact that the titration curves of H137Q mutant differ in shape with respect to those of the wild type and the other variants reinforces the idea that there is a structural alteration at the active site caused by the His/Gln substitution. In this new local conformation, the titrations of His50 and Glu96 are affected by a new acid group. According to the three-dimensional (3D) structure, the residue that could most likely affect their titration is Asp41. This is the only acidic residue located within 7 Å of the active site in the wild-type structure, and its pK_a value is unusually low (<3.0).

pK_a Values of WT α-Sarcin Calculated on the Basis of Theoretical Methods. To have a reference set of pK_a values for evaluating and comparing the corresponding ones in the mutants, we first undertook the calculation of the theoretical pK_a values of all titratable groups of WT α-sarcin. Table 3 shows the results of the pK_a calculations obtained from the UHBD and TK formalisms. All Arg residues show minor deviations from their nominal pK_a values, with the remarkable exception of Arg121 which has an unusual low solvent accessibility (6%) and interacts electrostatically with other active site residues: His50, Glu96, and His137.

The calculated pK_a values of lysine residues are in general close to the value for model compounds, in agreement with their high level of solvent exposure and the absence of structural restrictions. Deviations from the intrinsic pK_a are in all cases positive, indicating small stabilizations of the charged forms. Lys114 and Lys129 exhibit the highest pK_a values among all lysine residues (almost 1 pH unit above the intrinsic pK_a). Interestingly, the high pK_a of Lys114 may be related to the RNA recognition role of this side chain.

All tyrosine residues in α-sarcin are buried in the hydrophobic core of the protein. The NMR structure shows that some of these Tyr side chains are involved in hydrogen bonds within the protein core (i.e., Tyr106) and that their

Table 3: pK_a Values of Charged Side Chains of α-Sarcin[a] Calculated by Two Different Theoretical Methods

residue	UHBD	Tanford−Kirkwood	expt[b]	ASA (%)	**H-bond**/electrostatic interaction	H-bond restriction[c]
Arg22	12.1 ± 0.1 (12.0−12.2)	12.2 ± 0.1 (12.1−12.2)	nd[f]	49		
Arg66	12.3 ± 0.1 (12.0−12.5)	12.3 ± 0.2 (12.1−12.8)	nd[f]	43		
Arg78	12.6 ± 0.1 (12.3−12.7)	12.6 ± 0.1 (12.3−12.7)	nd[f]	28	D75, H92	
Arg121[d,e]	14.0 ± 0.1 (13.4−14.9)	13.1 ± 0.1 (12.8−13.3)	nd[f]	6	Y48, H50, E96, H137	G118, A120
Lys11	10.8 ± 0.1 (10.7−11.1)	10.9 ± 0.1 (10.7−11.2)	nd[f]	35	Y18	
Lys14	10.4 ± 0.1 (10.4−10.5)	10.6 ± 0.1 (10.6−10.9)	nd[f]	58		
Lys17	10.6 ± 0.1 (10.5−10.8)	10.7 ± 0.1 (10.6−10.8)	nd[f]	46		
Lys21	10.8 ± 0.1 (10.6−11.0)	10.9 ± 0.1 (10.7−11.1)	nd[f]	29	E19	
Lys29[d]	10.1 ± 0.4 (9.9−10.7)	10.9 ± 0.1 (10.8−11.1)	nd[f]	23		
Lys43	11.0 ± 0.1 (10.9−11.5)	10.9 ± 0.1 (10.8−11.4)	nd[f]	44	D85	
Lys61	10.6 ± 0.1 (10.4−10.9)	10.8 ± 0.1 (10.5−11.0)	nd[f]	49		
Lys64	10.4 ± 0.1 (10.3−10.5)	10.6 ± 0.1 (10.5−10.7)	nd[f]	66		
Lys70	10.7 ± 0.1 (10.5−10.8)	10.8 ± 0.1 (10.6−11.1)	nd[f]	49		
Lys73[d]	10.4 ± 0.1 (9.6−10.6)	10.8 ± 0.1 (10.7−11.0)	nd[f]	33		
Lys81	10.8 ± 0.2 (10.3−11.1)	10.9 ± 0.1 (10.6−11.1)	nd[f]	36	H92	P80
Lys84	10.7 ± 0.1 (10.7−10.9)	10.8 ± 0.1 (10.8−11.0)	nd[f]	60		
Lys89	10.5 ± 0.1 (10.2−10.5)	10.7 ± 0.1 (10.5−10.7)	nd[f]	63		D90
Lys107	10.7 ± 0.1 (10.6−10.8)	11.0 ± 0.1 (10.9−11.2)	nd[f]	37		
Lys111	10.6 ± 0.1 (10.5−10.7)	10.6 ± 0.1 (10.5−10.7)	nd[f]	47		
Lys112	10.4 ± 0.1 (10.3−10.5)	10.6 ± 0.1 (10.5−10.6)	nd[f]	58		
Lys114	11.2 ± 0.1 (11.1−11.3)	11.4 ± 0.2 (10.8−11.7)	nd[f]	24	Y48, Y106	
Lys129[d]	11.3 ± 0.5 (9.4−11.4)	11.1 ± 0.1 (11.0−11.2)	nd[f]	7	H35, Y124	A37, H35
Lys139	10.9 ± 0.1 (10.8−11.1)	11.1 ± 0.1 (11.0−11.3)	nd[f]	30	Y18, D9	
Lys146	10.6 ± 0.1 (10.5−10.7)	10.7 ± 0.1 (10.6−10.8)	nd[f]	34		Q138
Tyr18	9.9 ± 0.1 (9.6−10.1)	10.4 ± 0.1 (10.2−10.6)	nd[f]	21	E140, K139, K11	
Tyr25[d]	11.3 ± 0.2 (10.7−11.4)	10.3 ± 0.1 (10.3−10.4)	nd[f]	1		
Tyr48[d,e]	10.0 ± 0.4 (9.3−10.2)	10.0 ± 0.1 (9.7−10.2)	nd[f]	6	H50, K114, R121	
Tyr56	10.5 ± 0.2 (10.1−10.8)	10.5 ± 0.2 (10.0−10.7)	nd[f]	4	D77	
Tyr93	10.7 ± 0.1 (10.5−10.8)	11.0 ± 0.1 (10.7−11.3)	nd[f]	0	D91, Y126	G86
Tyr106[d]	12.2 ± 0.6 (10.8−12.7)	10.1 ± 0.1 (9.9−10.3)	nd[f]	1	**K114**	
Tyr124[d]	12.2 ± 0.3 (11.5−12.5)	10.3 ± 0.1 (10.2−10.4)	nd[f]	0	K129	
Tyr126	10.3 ± 0.2 (10.1−10.6)	10.6 ± 0.1 (10.3−10.9)	nd[f]	8	D91, Y93	
His35[d]	6.3 ± 0.2 (6.0−6.6)	6.4 ± 0.1 (6.3−6.4)	6.3 ± 0.1	34	E31, K129	
His36	7.1 ± 0.1 (7.0−7.3)	6.5 ± 0.1 (6.4−6.6)	6.8 ± 0.1	43	D105	
His50[d,e]	7.3 ± 0.1 (6.8−7.6)	6.4 ± 0.1 (6.1−6.6)	7.7 ± 0.1	18	Y48, E96, R121	W51
His82	7.7 ± 0.1 (7.2−7.8)	6.8 ± 0.1 (6.6−6.9)	7.3 ± 0.1	11	**D41**, D91	D41
His92	6.6 ± 0.1 (6.4−6.6)	6.1 ± 0.1 (5.8−6.2)	6.9 ± 0.1	18	R78, K81	
His104[d]	5.4 ± 0.1 (5.1−5.7)	6.7 ± 0.1 (6.5−6.7)	6.5 ± 0.1	27	D102, E115	D102, E115
His137[d,e]	5.6 ± 0.2 (5.3−5.9)	6.1 ± 0.1 (6.0−6.2)	5.8 ± 0.1	6	**G143**, R121	
His150[e]	6.8 ± 0.1 (6.4−6.9)	6.5 ± 0.1 (6.4−6.6)	7.6 ± 0.1	46	C-terminus	
Glu19	4.2 ± 0.1 (4.1−4.3)	4.1 ± 0.1 (3.9−4.2)	4.6 ± 0.1	22	K21	
Glu31[d]	4.8 ± 0.3 (4.2−5.2)	3.8 ± 0.1 (3.7−3.9)	4.6 ± 0.1	17	H35	Q27
Glu96[d,e]	5.2 ± 0.4 (3.8−5.3)	3.5 ± 0.1 (3.2−3.7)	5.2 ± 0.1	1	H50, R121	
Glu115	4.5 ± 0.1 (4.3−4.6)	4.1 ± 0.1 (3.9−4.2)	4.9 ± 0.1	36	H104	H104, K107
Glu140	4.6 ± 0.2 (4.4−4.6)	4.2 ± 0.1 (4.0−4.3)	4.2 ± 0.1	26	Y18	
Glu144	4.3 ± 0.2 (4.2−4.5)	4.2 ± 0.1 (4.0−4.3)	4.3 ± 0.1	44		
Asp9	4.1 ± 0.1 (3.9−4.2)	3.6 ± 0.1 (3.3−3.7)	3.9 ± 0.1	10	K139	
Asp41	4.7 ± 0.1 (4.6−4.7)	3.4 ± 0.1 (3.3−3.6)	<3.0	5	H82, W51	H82, W51
Asp57	4.1 ± 0.1 (4.0−4.3)	3.8 ± 0.1 (3.6−3.8)	4.2 ± 0.1	33		N59, K61
Asp59	3.9 ± 0.1 (3.7−4.2)	3.6 ± 0.1 (3.3−3.8)	4.0 ± 0.1	35		
Asp75	3.6 ± 0.1 (3.4−3.7)	3.4 ± 0.1 (3.2−3.6)	3.8 ± 0.1	22	R78	
Asp77	4.2 ± 0.1 (4.0−4.4)	3.6 ± 0.1 (3.4−3.8)	3.0 ± 0.1	19	Y56	F71, G72
Asp85	4.2 ± 0.1 (4.1−4.6)	3.6 ± 0.1 (3.1−3.7)	3.8 ± 0.1	26	K43	
Asp91[d]	4.6 ± 0.5 (4.2−5.1)	3.6 ± 0.1 (3.4−3.7)	<3.0	3	H82, Y93, Y126	S83, G88
Asp102	4.0 ± 0.1 (4.0−4.1)	3.5 ± 0.1 (3.4−3.6)	<3.0	46	H104	H104
Asp105	3.8 ± 0.1 (3.7−4.0)	3.5 ± 0.1 (3.4−3.6)	<3.0	30	H36	N33
Asp109	3.9 ± 0.1 (3.9−4.0)	3.7 ± 0.1 (3.7−3.8)	3.7 ± 0.1	37		G45

[a] Mean values and their standard deviations over 20 NMR structures (PDB entry 1DE3). Numbers in parentheses correspond to the minimum and maximum calculated pK_a values. [b] Values from ref *24*. [c] α-Sarcin numbering. [d] Low dielectric constant sites detected by the UHBD program. [e] UHBD pK_a calculated with His50, His137, and His150 Hδ1 tautomers as experimentally determined (*45*). Other histidines were considered Nϵ2 tautomers. [f] Not determined.

phenolic protons are partially buried, which might provide a structural explanation for their high pK_a values.

The TK model predicts pK_a values for the histidine residues close to that of the random coil value (6.3). The UHBD method yields high pK_a values for His36 (7.1), His50 (7.3), and His82 (7.7), whereas low pK_a values are predicted for His104 (5.4) and His137 (5.6). Low pK_a values are calculated for all glutamic acids with the Tanford−Kirkwood methodology as well as for the 11 aspartic acids, thus suggesting a general stabilization of the charged forms as a consequence of charge−charge interactions with positive side chains.

For most groups, the differences between the pK_a values calculated with the Tanford−Kirkwood and UHBD methods are less than 0.5 pH unit (Figure 4), and the pK_a values are in the range found in other protein systems. For Arg121,

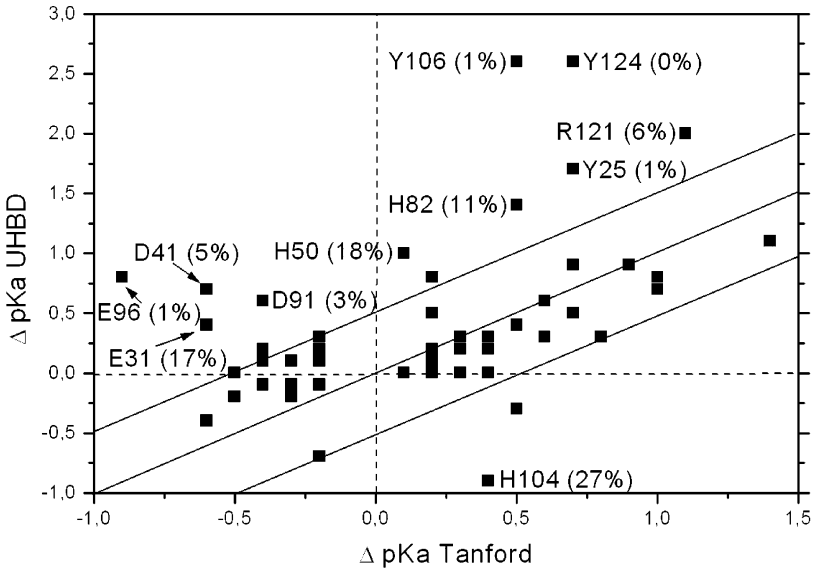

FIGURE 4: Comparison of pK_a differences in Tanford−Kirkwood (TK) and UHBD calculations with respect to the model compounds of Nozaki and Tanford [ΔpK_a = pK_a(method) − pK_a(model)]. Residues with different pK_a values predicted by both methods to be larger than 0.5 pH unit are labeled together with their percent of accessible surface.

Tyr25, Tyr106, Tyr124, His50, and His82, significant differences are found for the pK_a values calculated by the two methods, although they are always higher than the intrinsic values, thus indicating some stabilization of the protonated form. For His104, Glu31, Glu96, and most Asp groups, both methods also predict different pK_a values. In general, the UHBD package tends to yield higher pK_a values. The main differences (Figure 4) are found in deeply buried residues (accessible surfaces of less than 10%).

pK_a Values of H50Q, E96Q, H137Q, and H50/137Q α-Sarcin Variants Calculated on the Basis of Theoretical Methods. The pK_a values were also calculated for all titratable groups in the α-sarcin mutants. According to the NMR data, meaningful information could be obtained only for the H50Q and E96Q mutants, as the H137Q and H50/137Q α-sarcin variants do not maintain the native conformation of the active site. As expected, differences in the predicted pK_a values for most residues in the mutants are small with respect to those of the WT (Figure 5) and the pK_a values of residues at and near the active site (i.e., Tyr48 and Arg121) are the ones most strongly affected in the different variants. The TK- and UHBD-calculated pK_a values of these groups in the E96Q and H50Q mutants are summarized in Table 4. With the exception of Arg121, pK_a values calculated by both methods are in complete agreement. For Arg121, the UHBD method systematically yields a pK_a 1 pH unit higher than the one calculated by the TK method. Regardless of the mutated residue, the predictions point to a pK_a near the random coil value for Tyr48, a high pK_a for His50, Glu96, and Arg121, and a low pK_a with respect to the model for His137.

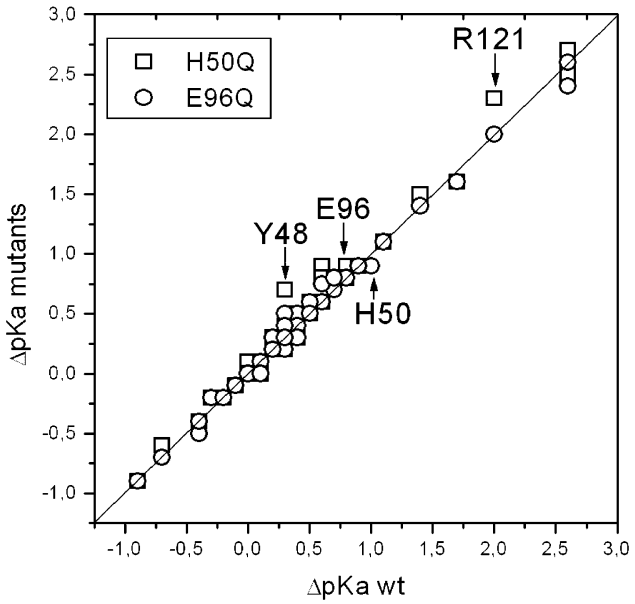

FIGURE 5: Comparison of calculated pK_a values (UHBD calculations) between the WT and the H50Q and E96Q α-sarcin mutants. pK_a values refer to model compounds of Nozaki and Tanford [ΔpK_a = pK_a(calculated) − pK_a(model)]. Residues whose values deviate by more than 0.3 pH unit and residues E96 and H50 in H50Q and E96Q mutants, respectively, are labeled.

DISCUSSION

Identifying the interactions responsible for changes in the intrinsic pK_a values of titrating groups is one of the more refined ways to describe the electrostatic properties of a protein, which can play decisive roles in the biological function. However, this is not an easy task. First, one needs to know the 3D structure of the protein very precisely to assign specific interactions. Moreover, few proteins are stable over a wide enough pH range for the experimental determination of all pK_a values. Consequently, the best approach is to use a combination of theoretical calculations and mutagenesis together with the experimental determination of the ionization equilibria (*49*). The cytotoxic ribonuclease α-sarcin is an excellent system for these analyses, because it is very well characterized structurally, its experimental pK_a values are known, and some mutants have been obtained and studied. On the other hand, consideration of the pK_a values in α-sarcin is a formidable task because the large number of charged groups leads to large charge−charge interactions. Moreover, many charges are not exposed, so desolvation effects and charge−dipole interactions can be

Table 4: Calculated pK_a Values of Active Site Residues of H50Q and E96Q α-Sarcin Mutants by Two Different Theoretical Methods[a]

	Tyr48		His50		Glu96		His137		Arg121	
mutant	TK	UHBD	TK	UHBD	TK	UHBD	TK	UHBD	TK	UHBD
E96Q	9.9 ± 0.1	10.0 ± 0.3	7.4 ± 0.1 (−1.0)	7.2 ± 0.2 (−0.8)	−	−	5.7 ± 0.1 (0.2)	5.6 ± 0.1 (0.3)	12.6 ± 0.1	14.0 ± 0.4
H50Q	10.0 ± 0.1	10.3 ± 0.3	−	−	5.4 ± 0.1 (−0.4)	5.4 ± 0.3 (−0.4)	5.9 ± 0.1 (0.2)	5.7 ± 0.1 (0.4)	13.0 ± 0.1	14.3 ± 0.5

[a] TK, Tanford−Kirkood method; UHBD, University of Houston Brownian Dynamics package. The pK_a differences with respect to the experimental values are shown in parentheses.

expected to shift individual pK_a values substantially as occurs in other proteins (*7*, *9*, *10*, *15*, *16*).

In some cases, the predicted pK_a shifts of α-sarcin groups correlate well with either electrostatic interactions or solvent inaccessibility (*24*). This is the case of Tyr25, for example, whose high pK_a value can be attributed to its low surface accessibility. However, this is not the general case. For other residues, changes in pK_a can be assigned to two effects that perturb the pK_a in the same direction. For example, the reduced accessible surface and the repulsion with negatively charged groups can account for the stabilization of the neutral OH form in tyrosines 56, 93, and 126. In some other cases, opposing effects partially cancel to produce the observed pK_a value, i.e., Glu96 and Arg121.

Many electrostatic interactions are concentrated at the α-sarcin active site. The residues belonging to this site, including His50, Glu96, Arg121, and His137, have been shown to have altered pK_a values, which in some cases (His50 and Glu96) have not received an appropriate theoretical explanation. This fact prompted us to measure the effect of removing nearby charges on those pK_a values by means of mutation. In general, the response to charge replacements is complex and difficult to anticipate. In some cases, substitutions that alter the active site charge balance may lead to structural rearrangements of the active site that restore the charge balance present in the wild-type protein (*50*). On the other hand, recent studies have shown that some residues are insensitive to charge replacements in their environment, and the results are not easily explained (*7*). Despite these caveats, charge substitution mutations are still a useful approach for isolating and experimentally quantifying individual interactions (*51*).

Structure of the Active Site in the Mutants on the Basis of NMR Data. Four α-sarcin variants, H50Q, E96Q, H137Q, and H50/137Q, were previously constructed and characterized (*52*). According to the published spectroscopic data (*52*), the structural changes produced by the mutations are small. NMR data, chemical shifts and NOEs, confirm the global structural similarity shared by mutants and wild-type proteins. Additionally, the pK_a values of other residues, six histidine groups (H35, H36, H82, H92, H104, and H150) far away (more than 7 Å) from the active site, are additional evidence that the global 3D structure is conserved. Consequently, we will discuss the effect of the substitutions in terms of local changes at the active site and its environment. There is no NMR evidence for conformational arrangements in the H50Q and E96Q mutants. In addition, both mutants and WT α-sarcin have similar conformational stabilities as indicated by the closeness of their T_m values at pH 7 (*52*). A different situation is found for mutants involving His137, where the NMR data suggest a conformational change as it was found in the crystal structure of the related H92Q mutant of RNase T1 (*53*). In the native protein, the His137 Nδ_1-H proton forms a hydrogen bond with the oxygen of Gly143, which helps fix the orientation of loop 5 with respect to the central β-sheet. In the mutants, no evidence for an equivalent hydrogen bond is found, since the observed chemical shifts of the side chain protons of Gln137 are close to the random coil values, and also its primary amide protons are unprotected against exchange (*48*). The lower stability of these mutants ($\Delta\Delta G \sim 2$ kcal/mol at pH 7.0) (*52*) with respect to the wild-type protein can most probably be ascribed to the loss of this specific hydrogen bond.

Interpretation of the Experimental pK_a Values of the Active Site Residues in the Mutants. Substitution of Glu96 with Gln provokes a decrease in the pK_a value of His50 of >1 pH unit. On the basis of only electrostatic considerations, this is the expected result since, in the mutant, the charge−charge interaction between the carboxylate and the imidazolium groups, which favors the protonated state of the histidine ring, is missing. In the mutant, the pK_a of His50 is close to the value reported for model peptides as expected for a side chain with moderate surface accessibility. From the pK_a of 6.4 for His50 in the E96Q mutant and that of 7.7 in the wild type, a $\Delta\Delta G°$ of 1.83 kcal/mol is obtained. This energy difference arises from the most favorable contribution of the charge−charge (His50−Glu96 and Arg121−Glu96) electrostatic interactions present in the wild type, relative to the charge−dipole (His50−Gln96 and Arg121−Gln96) interactions in the mutant. The pK_a of His137 does not change in the mutants within experimental error, indicating that the E96Q mutation does not affect its microenvironment as a consequence of the large distance between His137 and Glu96. In summary, the information obtained from the E96Q mutant demonstrates that the presence of Glu96 is the main cause of the relatively high pK_a of His50, whereas it does not affect the pK_a of His137.

A completely different picture can be drawn from the H50Q mutant data. Surprisingly, no significant variation in the pK_a value of Glu96 is observed when His50 is substituted with Gln, contrary to the increase expected on the basis of the E96Q mutant data. A slight decrease in the pK_a of Glu96 is observed instead, showing a hardly significant stabilization of the charged form. This result leads us to consider that some sort of compensating effect must be counteracting the loss of the positive charge of His50 to keep the pK_a of Glu96 unchanged. The NMR data clearly demonstrate that the $CONH_2$ protons of Gln50 are not involved in hydrogen bonds, so their participation in the stabilization of the charged form of Glu96 in the mutant can be discarded. The more reasonable explanation would be then to accept a conformational change that cannot be observed on the basis of the

NMR data. This could be as follows: replacing His with a smaller side chain, Gln, may increase the access of the solvent to the active site, which may interact favorably with the carboxylate group compensating partially for the interaction with the missing imidazole group. Additionally, since the unfavorable His50−Arg121 interaction has been suppressed in this mutant, it is likely that Arg121 modifies its position to produce a more favorable interaction with Glu96. As in the previous mutant, the pK_a variation observed for His137 in H50Q is small. The slight increase in its pK_a can be explained by a combination of effects: the loss of a distant positive charge (50 and 137 side chains are at 10 Å, in the wild type) and the higher solvent accessibility of the active site. The increase in the solvent accessibility in the H50Q mutant active site is more likely to occur than in E96Q because of the larger difference in the size of the substitution of His with Gln than in the substitution of Glu with Gln.

We have seen that His137 is not largely affected by the replacement of His50 or Glu96 with Gln. Consequently, the elimination of the positive charge at this position was expected to slightly contribute to the stabilization of the charged or neutral forms of His50 and Glu96. The pK_a value of His50 however increases upon mutation, reflecting the stabilization of its ionic form. The conformational change detected in this variant suggests a new electrostatic interaction, absent in wild-type α-sarcin, with an acidic residue, probably Asp41, which provides an additional contribution to the increase in the pK_a of His50. It is more difficult to interpret the pK_a increase for Glu96 of 0.5 unit, which is probably caused by a combination of effects. In this respect, the elimination of the positive charge of His137 and the possible electrostatic repulsion with Asp41 should be determinant factors.

Finally, the effects of the single mutations on the pK_a of Glu96 appear to be additive in the context of the double mutant. On one hand, H50Q decreases the pK_a of Glu96, and on the other, H137Q increases it, resulting in an intermediate value in the double mutant. This indicates that both mutations act independently, in agreement with published stability and spectroscopic data (*52*).

From the previous analysis, we may conclude, in contrast to the results obtained for His50, that the pK_a values of Glu96 and His137 are not sensitive to positive or negative charge replacements in their environment. In the case of His137, it appears reasonable to accept that the relative large distance separating it from the two other residues weakens the expected effects. The null effect of replacing His with Gln at position 50 on the determined pK_a of Glu96 may be the result of two counteracting effects. However, the electrostatic interactions among these three active site residues, and Tyr48 and Arg121, whose pK_a values cannot be obtained experimentally by NMR, still remain unknown. This information is essential for a complete interpretation of pK_a perturbations and may be analyzed on the basis of the calculated values.

Comparison between Experimental and Calculated pK_a Values. Calculated pK_a values for His, Glu, and Asp are compared with those determined experimentally for wild-type α-sarcin (*24*) and its mutants in Figure 6. Many experimentally measured pK_a shifts are very satisfactorily reproduced by the calculations. However, important deviations are clearly observed in many groups both in the wild type and in the mutants. First, on the basis of the known 3D

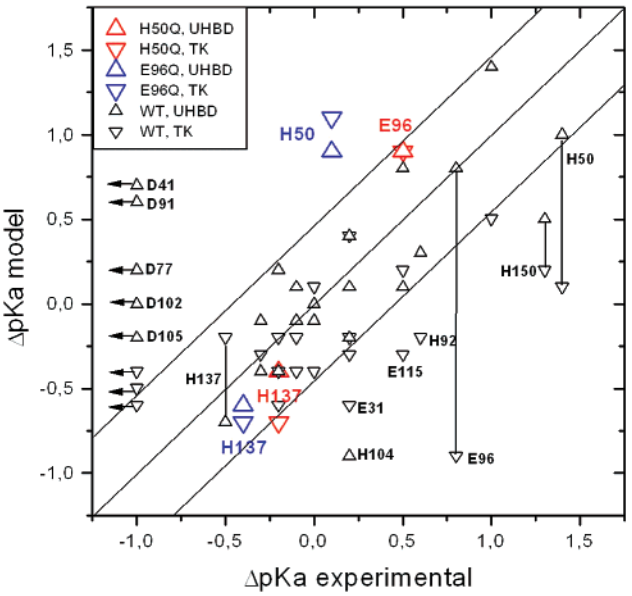

FIGURE 6: Comparison of calculated and experimental pK_a values for α-sarcin residues. Experimental and calculated data refer to model compounds of Nozaki and Tanford [ΔpK_a = pK_a(method)/exp pK_a(model)]. Residues that deviate by more than 0.5 pH unit in the WT protein are labeled. pK_a values corresponding to active site residues of H50Q are in red and those of E96Q in blue. ΔpK_a values of aspartic acid residues with pK_a values of less than 3.0 are out of the plotted range and are represented by arrows.

α-sarcin structure, we will discuss the observed differences in the WT pK_a values of groups outside the active site. The Tanford−Kirkwood model accurately reproduces (within a 0.5 pK_a unit difference) the pK_a values of His35, His36, His82, and His104, while it fails for His92 and His150. The UHBD method accounts satisfactorily for the pK_a values of His35, His36, His82, and His92, although some of these pK_a values are highly perturbed. There is, however, no good agreement between the calculated UHBD and the experimental pK_a values for His104 and His150. His150 is an interesting case since neither computational method accurately reproduces its pK_a, even though it is partially exposed. As recently reported (*45*), the calculated pK_a of His150 strongly depends on its tautomeric form. In the UHBD calculations, the experimentally determined Nδ_1 tautomers were used, and the pK_a values produced are closer to the experimental ones. For His92, the TK model indicates a stabilization of the neutral form to avoid repulsive electrostatic interactions with Arg78 and Lys81, whereas in the UHBD calculation, the effect attributed to the desolvation term gives a value more in line with the experimental pK_a. The UHBD model predicts, for the Nδ_1 tautomer of His104, a pK_a closer to the experimental value than for the Nϵ_2 form (*45*). However, the tautomeric state of His104 could not be determined on the basis of NMR data (*45*), so the pK_a of the more common Nϵ_2 tautomer was employed in the current calculations.

In the case of the Glu side chains, the UHBD results match the experimental values more closely. TK-calculated pK_a values show significant differences with the experimental values of glutamic acids 31 and 115. In these cases, it is clear that the hydrophobic environment is best reproduced by the Poisson−Boltzmann approximation, and that the relatively high calculated pK_a values can be mainly attributed to this effect. For the aspartyl groups, none of the pK_a values

experimentally determined to be less than 3.0, namely, Asp41, -77, -91, -102, and -105, are reproduced by the two theoretical calculations employed here, although the differences are smaller when using the TK approach. All these aspartic acids show important electrostatic interactions with positively charged groups that should be the main determinants of their low pK_a values. The UHBD method appears to underestimate the electrostatic component, whereas the TK method produces more realistic values.

In relation to the active site groups, the analysis of their calculated pK_a values should be restricted to the WT, and the E96Q and H50Q α-sarcin variants, which are the only molecules conserving all the conformational features of the active site. In all cases, the pK_a of His137 is accurately predicted (within 0.5 pH unit) despite its highly perturbed pK_a. The absence of variations in this pK_a upon mutation points to desolvation being the main factor in determining its value. Glu96 has a high pK_a which is not accurately reproduced by the TK method in the wild-type protein. In the H50Q mutant, the values computed from the two different approaches for the pK_a of Glu96 are within ±0.5 pH unit of the experimental values. This means that the structural model used for the calculations is very similar to the real solution structure, since it reproduces quite satisfactorily the environment of the active site. Accordingly, the conformational shift suggested by the NMR data for H50Q must be small, or the effect on the pK_a should be efficiently compensated, i.e., by a new charge distribution.

Finally, it is well-known that the pK_a of His50 depends on its tautomeric form and also on the tautomeric state of His137. In the UHBD calculations, the experimentally determined tautomers were used, but they are only able to reproduce the pK_a value of His50 in the WT protein. The TK approach also fails in predicting the correct value of the pK_a of the latter residue. There is not a clear and unique explanation that accounts for all of these data. It appears that the modeled structure may not reproduce the critical balance between the new geometry and charge distribution caused by mutation in complex environments such as the α-sarcin active site.

Wild-Type α-Sarcin Active Site. The pK_a of 7.7 for His50 in WT α-sarcin is 1.4 units higher than that of the histidine model compound, which corresponds to a $\Delta\Delta G°$ of 1.98 kcal/mol. This value agrees with that found for the His50−Glu96 interaction obtained from the E96Q mutant data. Hence, it is clear that this electrostatic interaction contributes in a substantial way to the perturbed pK_a of His50 in the wild-type protein. The analysis of the mutants also shows that the pK_a values of Glu96 and His137 are insensitive to charge replacement, yet their pK_a values differ significantly from the corresponding intrinsic values, which suggests that they may be affected by some other effects apart from the interaction with His50. Both residues are deeply buried within the protein interior with global surface accessibilities of 1 and 6% for residues 96 and 137, respectively. The exposure of the titrating atoms is also very low: 3 and 5% for the oxygen atoms of Glu96 and 0 and 10% for the nitrogen ring atoms of His137 (Nδ and Nϵ, respectively). However, since the tautomer of His137 present in solution is the Nδ form, the nitrogen involved in the proton transfer of the catalytic reaction is the more exposed Nϵ. Charged groups with very hydrophobic environments may undergo dramatic pK_a shifts (4−6 pH units) (*10*, *54*−*56*), much greater than those found in α-sarcin. Dielectric constant (ϵ) values on the order of 10−12 (*14*, *57*), measured experimentally, have been reported for buried positions. In α-sarcin, the first question to answer is whether the desolvation of the active site is the only factor responsible for the perturbed pK_a values of Glu96 and His137. Assuming it to be true, and ignoring other electrostatic interactions, we calculate an ϵ_{prot} mean value of 43 ± 1 for the active site using the Born approximation:

$$\Delta\Delta G° = 164.9(1/\epsilon_{prot} - 1/\epsilon_{water})/r$$

where $\Delta\Delta G° = 1.13$ kcal/mol (from E96 pK_a) and 0.71 kcal/mol (from H137 pK_a), $\epsilon_{water} = 74.9$ at 35 °C, and r is the corrected cavity radius [1.50 and 2.17 Å for oxygen and secondary amine, respectively (*58*)]. The high value of the dielectric constant (43) calculated in that way is very close to that found near the solvent-exposed surfaces. This result is in satisfactory agreement with the position of Nϵ in His137 near the surface but is not consistent with the very low surface accessibility of the Glu96 α-sarcin catalytic side chain. Then, the assumption that desolvation is the sole determinant of the pK_a of Glu96 is not correct. As a consequence, the perturbed pK_a values should result from the combination of at least two effects: the low dielectric constant that tends to increase the pK_a of Glu96 and an opposite compensating effect. The simplest explanation for this second effect would be an electrostatic interaction with a nearby cationic residue. In the protein structure, Arg121 is strategically located more or less equidistant from all groups in the active site and interacts electrostatically with all of them. The pK_a calculations have shown that Arg121 has a high pK_a, 14.0, despite its low surface accessibility (6% global and 9% Nϵ), suggesting its participation in a salt bridge. The electrostatic interaction of Arg121 is expected to lower the pK_a value of Glu96 from a higher value to the measured value. But, as described above, Arg121 is also quite near His50 and His137. The effect of Arg121 on the pK_a of His137 should go in the same direction as that of desolvation, so it is probable that the combination of these two effects is responsible for the low pK_a of this catalytic side chain. Thus, everything suggests that several factors affect the pK_a values of Glu96 and His137, so they may act as the general base and the general acid, respectively, in the catalytic mechanism at physiological pH. This is interesting because it would imply that, although not directly involved in the catalytic mechanism, Arg121 is crucial for determining the electrostatic balance and properties of the catalytic residues in α-sarcin. The enzymatic characterization of R121K and R121Q mutants using assays with dinucleotides suggested that Arg121 is involved in catalysis rather than substrate binding (*59*), which is consistent with our current results. Mutation of Arg to Lys maintains the positive charge and some of the hydrogen bonding properties. Thus, it is not surprising that R121K can recognize and cleave dinucleotide substrates. However, the very different electrostatic properties of Lys (intrinsic pK_a of 10) and Arg (intrinsic pK_a of 12) and side chain shapes and volumes should affect the pK_a values of Glu96 and His137 in a different way, reducing this mutant's catalytic efficiency (*59*). It would be very interesting to determine the pK_a values of the active site

groups in the absence of the positive charge at position 121. Unfortunately, the R121Q variant could only be produced in very small amounts, and its conformational stability is too low to permit the experimental determination of pK_a values by NMR. An arginine side chain is a highly conserved group at the active site of microbial RNases, although its specific role has not yet been established. RNase T1, one of the best characterized ribonucleases, also has an arginine, Arg77, at the active site. However, all attempts to isolate RNase T1 Arg77 mutants have been unsuccessful (*60*), suggesting that this residue is important for the enzyme's folding and/or stability. In α-sarcin, the R121K and R121Q variants are less stable than the wild-type protein by 2.0 and 3.7 kcal/mol, respectively, which together with the high value of the calculated pK_a for Arg121, speak to the importance of the electrostatic interactions formed by this residue to the global stability of the protein.

In summary, we have examined the electrostatics of α-sarcin using a combination of theoretical calculations, mutagenesis, and NMR measurements of individual pK_a values. For many titratable groups, this holistic approach succeeds in identifying the factors influencing their ionization equilibria. For others, particularly some active site residues, this approach reveals a complex interplay between structure, electrostatics, and activity. We find that it is not possible to explain the electrostatic interactions at the active site of α-sarcin in terms of single additive effects. On the contrary, His50, Glu96, Arg121, and His137 form a structurally and electrostatically interdependent unit that is greater than the sum of its parts. To further tweeze apart the roles played by individual groups, experiments are now underway to examine the pH dependence of α-sarcin's conformational stability and ribonucleolytic activity.

ACKNOWLEDGMENT

We thank Dr. R. C. Wade for providing the UHBD scripts.

REFERENCES

1. Tanford, C. (1962) *Adv. Protein Chem. 17*, 69−165.
2. Matthew, J. B., Gurd, F. R., Garcia-Moreno, B., Flanagan, M. A., March, K. L., and Shire, S. J. (1985) *CRC Crit. Rev. Biochem. 18*, 91−197.
3. Yang, A. S., and Honig, B. (1992) *Curr. Opin. Struct. Biol. 2*, 40−45.
4. Warshel, A., and Papazyan, A. (1998) *Curr. Opin. Struct. Biol. 8*, 211−217.
5. Honig, B., and Nicholls, A. (1995) *Science 268*, 1144−1149.
6. Anderson, D. E., Becktel, W. J., and Dahlquist, F. W. (1990) *Biochemistry 29*, 2403−2408.
7. Forsyth, W. R., and Robertson, A. D. (2000) *Biochemistry 39*, 8067−8072.
8. Forsyth, W. R., Gilson, M. K., Antosiewicz, J., Jaren, O. R., and Robertson, A. D. (1998) *Biochemistry 37*, 8643−8652.
9. Giletto, A., and Pace, C. N. (1999) *Biochemistry 38*, 13379−13384.
10. Laurents, D. V., Huyghues-Despointes, B. M., Bruix, M., Thurlkill, R. L., Schell, D., Newsom, S., Grimsley, G. R., Shaw, K. L., Trevino, S., Rico, M., Briggs, J. M., Antosiewicz, J. M., Scholtz, J. M., and Pace, C. N. (2003) *J. Mol. Biol. 325*, 1077−1092.
11. Henning, M., and Geierstanger, B. H. (1999) *J. Am. Chen. Soc. 121*, 5123−5126.
12. Chen, H. A., Pfuhl, M., McAlister, M. S., and Driscoll, P. C. (2000) *Biochemistry 39*, 6814−6824.
13. Fitch, C. A., Karp, D. A., Lee, K. K., Stites, W. E., Lattman, E. E., and Garcia-Moreno, E. B. (2002) *Biophys. J. 82*, 3289−3304.
14. Czerwinski, R. M., Harris, T. K., Massiah, M. A., Mildvan, A. S., and Whitman, C. P. (2001) *Biochemistry 40*, 1984−1995.
15. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) *Biochemistry 35*, 9958−9966.
16. Qin, J., Clore, G. M., and Gronenborn, A. M. (1996) *Biochemistry 35*, 7−13.
17. Kim, S., and Baum, J. (1998) *Protein Sci. 7*, 1930−1938.
18. Warshel, A., and Aqvist, J. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 267−298.
19. Juffer, A. H. (1998) *Biochem. Cell Biol. 76*, 198−209.
20. Forsyth, W. R., Antosiewicz, J. M., and Robertson, A. D. (2002) *Proteins 48*, 388−403.
21. Olson, B. H., and Goerner, G. L. (1965) *Appl. Microbiol. 13*, 314−321.
22. Martínez-Ruiz, A., García-Ortega, L., Kao, R., Lacadena, J., Oñaderra, M., Mancheño, J. M., Davies, J., Martínez del Pozo, A., and Gavilanes, J. G. (2001) *Methods Enzymol. 341*, 335−351.
23. Gasset, M., Mancheño, J. M., Lacadena, J., Turnay, J., Olmo, N., Lizarbe, M. A., Martínez del Pozo, A., Oñaderra, M., and Gavilanes, J. G. (1994) *Curr. Top. Pept. Protein Res. 1*, 99−104.
24. Pérez-Cañadillas, J. M., Campos-Olivas, R., Lacadena, J., Martínez del Pozo, A., Gavilanes, J. G., Santoro, J., Rico, M., and Bruix, M. (1998) *Biochemistry 37*, 15865−15876.
25. Pérez-Cañadillas, J. M., Guenneugues, M., Campos-Olivas, R., Santoro, J., Martínez del Pozo, A., Gavilanes, J. G., Rico, M., and Bruix, M. (2002) *J. Biomol. NMR 24*, 301−316.
26. Pérez-Cañadillas, J. M., Santoro, J., Campos-Olivas, R., Lacadena, J., Martínez del Pozo, A., Gavilanes, J. G., Rico, M., and Bruix, M. (2000) *J. Mol. Biol. 299*, 1061−1073.
27. Schindler, D. G., and Davies, J. E. (1977) *Nucleic Acids Res. 4*, 1097−1110.
28. Endo, Y., Huber, P. W., and Wool, I. G. (1983) *J. Biol. Chem. 258*, 2662−2667.
29. Correll, C. C., Munishkin, A., Chan, Y. L., Ren, Z., Wool, I. G., and Steitz, T. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 13436−13441.
30. Correll, C. C., Wool, I. G., and Munishkin, A. (1999) *J. Mol. Biol. 292*, 275−287.
31. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1996) *Biochemistry 35*, 7819−7833.
32. Ibarra-Molero, B., Loladze, V. V., Makhatadze, G. I., and Sánchez-Ruiz, J. M. (1999) *Biochemistry 38*, 8138−8149.
33. Demchuk, E., and Wade, R. C. (1996) *J. Phys. Chem. 100*, 17373−17387.
34. Lowe, D. M., Fersht, A. R., Wilkinson, A. J., Carter, P., and Winter, G. (1985) *Biochemistry 24*, 5106−5109.
35. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
36. Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
37. Campos-Olivas, R., Bruix, M., Santoro, J., Martínez del Pozo, A., Lacadena, J., Gavilanes, J. G., and Rico, M. (1996) *Protein Sci. 5*, 969−972.
38. Bundi, A., and Wüthrich, K. (1979) *Biopolymers 18*, 285−297.
39. Kraulis, P. J. (1989) *J. Magn. Reson. 24*, 627−633.
40. Forman-Kay, J. D., Clore, G. M., and Gronenborn, A. M. (1992) *Biochemistry 31*, 3442−3452.
41. Tanford, C., and Kirkwood, J. G. (1957) *J. Am. Chem. Soc. 79*, 5333−5339.
42. Matthew, J. B., and Gurd, F. R. (1986) *Methods Enzymol. 130*, 437−453.
43. Tanford, C., and Roxby, R. (1972) *Biochemistry 11*, 2192−2198.
44. Davis, M. E., Madura, J. D., Luty, B. A., and McCammon, J. A. (1991) *Comput. Phys. Commun. 62*, 187−197.
45. Pérez-Cañadillas, J. M., García-Mayoral, M. F., Martínez del Pozo, A., Gavilanes, J. G., Rico, M., and Bruix, M. (2003) *FEBS Lett. 534*, 197−201.
46. Nozaki, Y., and Tanford, C. (1967) *Methods Enzymol. 11*, 715−734.
47. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR 5*, 67−81.
48. Englander, S. W., Sosnick, T. R., Englander, J. J., and Mayne, L. (1996) *Curr. Opin. Struct. Biol. 6*, 18−23.
49. Sundd, M., Iverson, N., Ibarra-Molero, B., Sanchez-Ruiz, J. M., and Robertson, A. D. (2002) *Biochemistry 41*, 7586−7596.
50. Olson, M. A. (2001) *Biophys. Chem. 91*, 219−229.
51. Masip, M., García-Ortega, L., Olmo, N., García-Mayoral, M. F., Pérez-Cañadillas, J. M., Bruix, M., Oñaderra, M., Martínez del Pozo, A., and Gavilanes, J. G. (2003) *Protein Sci. 12*, 161−169.

52. Lacadena, J., Martínez del Pozo, A., Martínez-Ruiz, A., Pérez-Cañadillas, J. M., Bruix, M., Mancheño, J. M., Oñaderra, M., and Gavilanes, J. G. (1999) *Proteins 37*, 474−484.
53. De Vos, S., Doumen, J., Langhorst, U., and Steyaert, J. (1998) *J. Mol. Biol. 275*, 651−661.
54. Dwyer, J. J., Gittis, A. G., Karp, D. A., Lattman, E. E., Spencer, D. S., Stites, W. E., and Garcia-Moreno, E. B. (2000) *Biophys. J. 79*, 1610−1620.
55. Dao-pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., and Matthews, B. W. (1991) *Biochemistry 30*, 11521−11529.
56. Lambeir, A. M., Backmann, J., Ruiz-Sanz, J., Filimonov, V., Nielsen, J. E., Kursula, I., Norledge, B. V., and Wierenga, R. K. (2000) *Eur. J. Biochem. 267*, 2516−2524.
57. Garcia-Moreno, B., Dwyer, J. J., Gittis, A. G., Lattman, E. E., Spencer, D. S., and Stites, W. E. (1997) *Biophys. Chem. 64*, 211−224.
58. Rashin, A. A., and Honing, B. (1985) *J. Phys. Chem. 89*, 5588−5593.
59. Masip, M., Lacadena, J., Mancheño, J. M., Oñaderra, M., Martínez-Ruiz, A., Martínez del Pozo, A., and Gavilanes, J. G. (2001) *Eur. J. Biochem. 268*, 6190−6196.
60. Steyaert, J. (1997) *Eur. J. Biochem. 247*, 1−11.
61. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
62. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 29−32.

BI0349773