

# Effect of mutations involving charged residues on the stability of staphylococcal nuclease: a continuum electrostatics study

Ulf Börjesson and Philippe H. Hünenberger<sup>1</sup>

Laboratorium für Physikalische Chemie, ETH Hönggerberg, HCI,  
CH-8093 Zürich, Switzerland

<sup>1</sup>To whom correspondence should be addressed.

E-mail: phil@igc.phys.chem.ethz.ch

**A continuum electrostatics model is used to calculate the relative stabilities of 117 mutants of staphylococcal nuclease (SNase) involving the mutation of a charged residue to an uncharged residue. The calculations are based on the crystallographic structure of the wild-type protein and attempt to take implicitly into account the effect of the mutations in the denatured state by assuming a linear relationship between the free energy changes caused by the mutation in the native and denatured states. A good correlation (linear correlation coefficient of ~0.8) is found with published experimental relative stabilities of these mutants. The results suggest that in the case of SNase (i) charged residues contribute to the stability of the native state mainly through electrostatic interactions, and (ii) native-like electrostatic interactions may persist in the denatured state. The continuum electrostatics method is only moderately sensitive to model parameters and leads to quasi-predictive results for the relative mutant stabilities (error of 2–3 kJ mol<sup>-1</sup> or of the order of  $k_B T$ ), except for mutants in which a charged residue is mutated to glycine.**

**Keywords:** continuum electrostatics/denatured state/  
electrostatic interactions/staphylococcal nuclease

## Introduction

The role of electrostatic interactions in determining the stability of proteins has been the subject of intensive research. Early experimental and theoretical studies suggested that charged residues contribute only little (or even unfavorably) to the overall stability of proteins (Akke and Forsén, 1990; Serrano *et al.*, 1990; Dao-pin *et al.*, 1991a,b; Sali *et al.*, 1991), whereas the opposite conclusion was reached in other studies (Anderson *et al.*, 1990; Tissot *et al.*, 1996). In specific cases, it has been possible to design mutant proteins that are more stable than the wild-type protein by introducing or removing charged residues on the protein surface, thus improving electrostatic interactions (Grimsley *et al.*, 1999; Loladze *et al.*, 1999; Spector *et al.*, 2000; Sanchez-Ruiz and Makhatadze, 2001). In contrast, a recent study concluded that surface charge–charge interactions are not essential for folding and stability of ubiquitin (Loladze and Makhatadze, 2002).

The most reliable sources of experimental information on this matter are probably mutational studies encompassing a large set of protein variants. For example, the denaturation equilibria of staphylococcal nuclease (SNase) and of a very large number of its mutants have been extensively studied

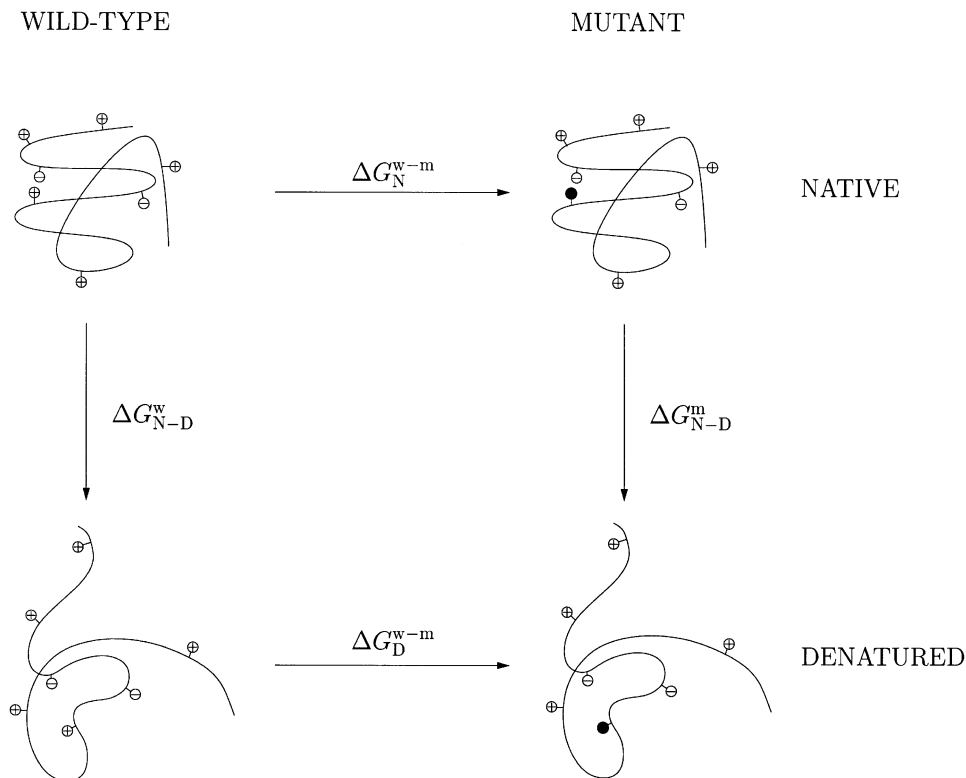
(Shortle *et al.*, 1990; Green *et al.*, 1992; Shortle, 1995; Meeker *et al.*, 1996; Schwehm *et al.*, 2003). In the most recent study, the disruption of specific electrostatic interactions was identified as being the dominant factor determining the stability change of charge-reversal and charge-deletion mutants (Schwehm *et al.*, 2003). In this work, and also in other studies [e.g., Pace *et al.* (Pace *et al.*, 2000)], the neglect of electrostatic interactions in the denatured state was suggested to be responsible for the occurrence of discrepancies between theoretical predictions and experimental results. In the present computational study, we show that it is possible implicitly to include these interactions in structure-based calculations, leading to quasi-predictive results for the relative stabilities of charge-deletion mutants of SNase.

The unfolding of SNase (a monomeric, single-domain protein consisting of 149 amino acid residues) can be described as a reversible equilibrium between a native state (N) and a denatured state (D), which may be viewed as two non-overlapping distributions of microstates (Lumry *et al.*, 1966). According to this simple two-state model, the free-energy difference  $\Delta G_{N-D} = G_D - G_N$  between the denatured and native states determines the relative populations of each state through the relationship

$$\Delta G_{N-D} = -RT \ln K_{N-D} \quad (1)$$

where  $K_{N-D}$  is the equilibrium constant for the denaturation process, defined as the ratio between the populations of the denatured and the native states.  $\Delta G_{N-D}$  can be measured experimentally, for example, through guanidinium hydrochloride denaturation of the protein and extrapolation of the measured denaturation free energy to zero denaturant concentration. The change in protein stability induced by a mutation (compared with that of the wild-type protein) is measured by the quantity  $\Delta\Delta G_{N-D} = \Delta G_{N-D}^{\text{mut}} - \Delta G_{N-D}^{\text{wt}}$ , where  $\Delta G_{N-D}^{\text{mut}}$  and  $\Delta G_{N-D}^{\text{wt}}$  are the denaturation free energies of the mutant and wild-type proteins, respectively. A negative value of  $\Delta\Delta G_{N-D}$  indicates that a specific mutation decreases the stability of the protein.

In a series of experimental studies (Shortle *et al.*, 1990; Green *et al.*, 1992; Meeker *et al.*, 1996), every residue of SNase has been mutated in a systematic way to either alanine or glycine. In this way, correlations between changes in protein stability ( $\Delta\Delta G_{N-D}$ ) and a number of structural features characterizing the mutated residue could be examined. It was found that replacing a charged residue by a smaller neutral one may destabilize the native state by up to 17 kJ mol<sup>-1</sup> (Meeker *et al.*, 1996), whereas replacing a large hydrophobic residue can destabilize the native state by up to 30 kJ mol<sup>-1</sup> (Shortle *et al.*, 1990). The descriptors of the local environment around the mutated residue that correlated best with the stability change induced by the mutation were those related to the extent of burial of the residue in the native state of the wild-type protein.



**Fig. 1.** Thermodynamic cycle used to estimate the stability change upon mutation of a charged residue; see Equation 2. Circles containing a plus or a minus sign symbolize charged residues. Filled circles indicate a charged residue mutated to an uncharged residue.

For instance, for mutations of large hydrophobic residues to glycine, the number of  $\alpha$ -carbon atoms located within 1 nm of the  $\alpha$ -carbon atom of the mutated residue correlated well with the stability change upon mutation (linear correlation coefficient  $r = 0.76$ ). Even in the case of mutations involving charged residues, modest correlations between  $\Delta\Delta G_{N-D}$  and side-chain burial were observed. However, no significant correlation between the stability change upon mutating a charged residue to alanine or glycine and the local electrostatic environment of the mutated residue (as probed by continuum electrostatics calculations) was found (Meeker *et al.*, 1996). These observations led to the conclusion that ionizable residues do not contribute significantly to the stability of SNase through electrostatic interactions, but predominantly through non-polar interactions.

However, the conclusions reached in the above study (Meeker *et al.*, 1996) with respect to this point were drawn from continuum electrostatics calculations that may not be entirely relevant. The reason is that this study relied on theoretical electrostatic charging free energies estimated as half the product of the residue charge and the electrostatic potential at the residue site, computed based on the native wild-type protein structure. This approach neglects two effects: (i) the free-energy contribution arising from the mutation of the residue in the denatured state of the protein, and (ii) the change in the electrostatic potential within the whole system (solvent reorganization) associated with the removal of a charged side chain. As a consequence, a more accurate estimate of the electrostatic free energy change upon mutation of a charged residue should involve not only a calculation of the electro-

static potential for the solvated wild-type protein, but also additional calculations of the electrostatic potential for the solvated mutant proteins. In addition, the contribution associated with the charge change in the denatured state should also be taken into account when estimating the overall free energy change.

The thermodynamic cycle presented in Figure 1 accounts for the denaturation equilibria of the wild-type and a mutant protein, both of which are assumed to be adequately described as two-state processes. Because the free energy is a state function, the stability change  $\Delta\Delta G_{N-D}$  induced by a mutation, can be calculated in either of two ways:

$$\begin{aligned}\Delta\Delta G_{N-D} &= \Delta G_{N-D}^m - \Delta G_{N-D}^w \\ &= \Delta G_D^{w-m} - \Delta G_N^{w-m}\end{aligned}\quad (2)$$

where  $\Delta G_D^{w-m} = G_D^m - G_D^w$  and  $\Delta G_N^{w-m} = G_N^m - G_N^w$  represent the free-energy differences between mutant and wild-type proteins in the denatured and native states, respectively.

In the absence of structural information about the mutant protein, the quantity  $\Delta G_N^{w-m}$  can be estimated by comparing free energies (including internal and solvation contributions) calculated using the experimental wild-type protein structure and a mutant protein structure modeled from the wild-type structure. This approximation is valid under the assumption that no major structural or protonation-state changes occur in the native state upon mutating the specific charged residue. On the other hand,  $\Delta G_D^{w-m}$  cannot be calculated explicitly because the structure (or ensemble of structures) defining the denatured state is unknown. However, if one assumes that the free-energy

difference  $\Delta G_D^{w-m}$  between the wild-type and mutant proteins in the denatured state is linearly related to the corresponding free-energy difference  $\Delta G_N^{w-m}$  in the native state, Equation 2 can be rewritten as

$$\begin{aligned}\Delta\Delta G_{N-D} &= \Delta G_D^{w-m} - \Delta G_N^{w-m} \\ &= \alpha\Delta G_N^{w-m} + \beta - \Delta G_N^{w-m} \\ &= (\alpha - 1)\Delta G_N^{w-m} + \beta\end{aligned}\quad (3)$$

The parameters  $\alpha$  and  $\beta$  are empirical constants that can be determined (for a given system) by calibration using a large set of mutants. Once  $\alpha$  and  $\beta$  have been determined, only  $\Delta G_N^{w-m}$  needs to be calculated in order to estimate  $\Delta\Delta G_{N-D}$ .

The basic idea behind this approach is the assumption that there may be some similarity between the native and denatured states of the protein in terms of electrostatic interactions, i.e. that these interactions are qualitatively similar in the two states but differ in their magnitude. In the specific case of SNase, this hypothesis is made plausible by several experimental observations about the denatured state of this protein. Measurements of the hydrodynamic radius of the denatured state have revealed a compact structure (Eftink and Ramsay, 1997; Baskakov and Bolen, 1998). There is evidence from paramagnetic relaxation enhancement experiments carried out using spin labels that the overall topology of  $\Delta 131\Delta$  (a 131-residue fragment of SNase that is used as a model for the denatured state of wild-type SNase) is similar to that of the native protein (Gillespie and Shortle, 1997a,b). This native-like topology of  $\Delta 131\Delta$  persists even under strongly denaturing conditions, as evidenced by recent residual dipolar coupling measurements (Shortle and Ackerman, 2001). Furthermore, several residues in the denatured state have  $pK_a$  values that are close to those in the native state, which is evidence for native-like electrostatic interactions in the denatured state of SNase (Whitten and García-Moreno, 2000). Ultimately, the validity of the assumption of electrostatic similarity between the native and denatured states can be tested by comparing theoretical estimates of mutant stabilities obtained through Equation 3 with experimental data. Note, however, that the assumption will not be true for all proteins and, even in cases where the assumption holds, the empirical parameters  $\alpha$  and  $\beta$  are not likely to be transferable from system to system.

In the present continuum electrostatics study, the calculation of the quantity  $\Delta G_N^{w-m}$  is based on a single conformation (the crystallographic structure) and set of protonation states (determined for a pH of 7 according to the  $pK_a$ s of isolated residues; histidines are discussed separately), assumed to be representative of the ensemble of configurations and protonation states characterizing the solvated native (wild-type or mutant) protein. In this single configuration, the free energy  $G$  of the solute-solvent system is decomposed into an electrostatic (solute-solute and solute-solvent) contribution  $G_{el}$  and a non-polar (solute-solvent) contribution  $G_{np}$ . Defining a reference state, for which  $G = 0$ , as the state in which all solute atomic partial charges are zero and the solvent is of low, alkane-like dielectric permittivity  $\epsilon_i$ , the free energy of the solute-solvent system may be written as

$$G = G_{el} + G_{np} = G_{Cb} + G_{rf} + G_{np}\quad (4)$$

The Coulomb contribution  $G_{Cb}$  represents the electrostatic work required to create the solute atomic partial charges in a homogeneous medium of permittivity  $\epsilon_i$ , i.e.

$$G_{Cb} = \frac{1}{4\pi\epsilon_o\epsilon_i} \sum_i \sum_{j>i} \frac{q_i q_j}{r_{ij}}\quad (5)$$

Note that charges  $i$  and  $j$  belonging to the same mutated residue must be excluded from the summation to avoid artifactual intra-group contributions to the free energies of mutation. The reaction-field contribution  $G_{rf}$  represents the electrostatic work required to transfer the charged solute from a solvent of permittivity  $\epsilon_i$  into water (dielectric permittivity  $\epsilon_w$ ). This quantity is calculated by solving the Poisson equation in the medium of heterogeneous permittivity using a finite-difference algorithm. The non-polar contribution  $G_{np}$  represents the non-electrostatic work required to transfer the neutral solute from an alkane-like solvent into water. This quantity should account both for the hydrophobic effect and for differences between the two solvents with respect to their solute-solvent van der Waals interactions. The contribution  $G_{np}$  is assumed to be proportional to the solvent-accessible surface area  $A$  of the solute through an empirical coefficient  $\gamma$  (effective microscopic interfacial tension):

$$G_{np} = \gamma A\quad (6)$$

The quantities  $\Delta G_N^{w-m} = G_N^w - G_N^m$  to be used in Equation 3 are obtained by evaluating, through Equation 4, the quantity  $G_N^w$  for the wild-type protein and the quantities  $G_N^m$  for all mutants considered.

In the present study, we evaluated the correlation between measured relative stabilities of charge mutants of SNase and corresponding values calculated using continuum electrostatics, in order to investigate (i) the specific role of electrostatic interactions in determining the stability of SNase, (ii) the validity of the assumed linear relationship between the free energy changes caused by charge mutations in the native and denatured states, (iii) the sensitivity of the calculations to model parameters, and (iv) the predictive ability of the model.

## Materials and methods

All calculations were performed using the crystal structure of unliganded staphylococcal nuclease (SNase) refined at 0.17 nm resolution (Hynes and Fox, 1991), entry code 1STN of the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977). This structure does not include the coordinates of the disordered residues 1–5 and 142–149. These 13 residues were omitted from the calculations and the N- and C-termini of the remaining 136-residue protein fragment were considered to be neutral. All titratable side chains except the four histidine residues were modeled in their ionized forms as appropriate for pH 7, at which the stability measurements were performed (Meeker *et al.*, 1996). Several sets of calculations were performed in order to evaluate the dependence of the results on the protonation states of the histidine side chains, whose  $pK_a$  values are close to the experimental pH. These calculations led to an optimal set of histidine protonation states (third entry in Table I) that was used for all calculations unless specified otherwise. The structures of the mutant proteins were modeled by removing or replacing atoms based on the coordinates of the wild-type crystal structure of SNase.

The continuum electrostatics calculations were performed using a modified version of the GROMOS96 program (van Gunsteren *et al.*, 1996; Scott *et al.*, 1999) including the routines

**Table I.** Linear correlation coefficients relating the experimental relative mutant stability  $\Delta\Delta G_{\text{N-D}}^{\text{exp}}$  to the theoretical mutation free energy in the native state  $\Delta G_{\text{N-m}}^{\text{w}}$  (see Equation 3) for 48 mutants of SNase involving the mutation of a charged residue to alanine (Meeker *et al.*, 1996)

His8	His46	His121	His124	$r''$	$r'$	$r$
N <sub>ε</sub>	N <sub>ε</sub>	N <sub>ε</sub>	N <sub>ε</sub>	-0.51	-0.57	-0.65
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.66	-0.64	-0.64
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.72	-0.72	-0.76
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>ε</sub>	N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.56	-0.67	-0.70
N <sub>δ</sub> ,N <sub>ε</sub>	3/4 N <sub>δ</sub> ,3/4 N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.68	-0.67	-0.75
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	3/4 N <sub>δ</sub> ,3/4 N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.58	-0.63	-0.63
N <sub>δ</sub> ,N <sub>ε</sub>	3/4 N <sub>δ</sub> ,3/4 N <sub>ε</sub>	3/4 N <sub>δ</sub> ,3/4 N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.56	-0.63	-0.72
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.65	-0.64	-0.76
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>ε</sub>	3/4 N <sub>δ</sub> ,3/4 N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.60	-0.68	-0.72
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>ε</sub>	N <sub>δ</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.40	-0.58	-0.61
N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> (flipped)	N <sub>δ</sub> ,N <sub>ε</sub>	N <sub>δ</sub> ,N <sub>ε</sub>	-0.72	-0.72	-0.76

Values are reported for different combinations of protonation states of the histidine residues, i.e. protonation at the  $\delta$  nitrogen atom (N<sub>δ</sub>), at the  $\epsilon$  nitrogen atom (N<sub>ε</sub>) or at both (N<sub>δ</sub>,N<sub>ε</sub>; overall residue charge +1e). An alternative protonation state combining the molecular topology of a doubly protonated histidine residue with an intermediate charge (3/4 N<sub>δ</sub>,3/4 N<sub>ε</sub>; overall charge +1/2e) is also considered. 'Flipped' indicates a configuration where the imidazole ring of the histidine was rotated by 180° (exchange of C<sub>δ</sub> ↔ N<sub>δ</sub> and C<sub>ε</sub> ↔ N<sub>ε</sub>) compared with the crystal structure. The linear correlation coefficient is calculated for all mutants ( $r''$ ), all mutants excluding D83A, D95A and H121A ( $r'$ ) and all mutants excluding D83A, D95A, H121A and partially charged or uncharged histidines ( $r$ ).

of the UHBD program (Davis *et al.*, 1991; Madura *et al.*, 1994, 1995) for solving the linearized Poisson–Boltzmann equation through a finite-difference algorithm and for computing the surface area-dependent non-polar term. The interaction-function parameters of the GROMOS96 force field (Daura *et al.*, 1998) were used to define atomic charges and radii. The atomic radii of the solute atoms were calculated from the Lennard–Jones C<sub>6</sub>- and C<sub>12</sub>-parameters defining the interaction between the specific atom and an SPC water oxygen atom (Berendsen *et al.*, 1981) as  $R = (2C_{12}/C_6)^{1/6} - 0.14$  nm (the approximate radius of a water molecule subtracted from the atom–water distance at the minimum of the Lennard–Jones curve). Hydrogen atoms were treated differently and assigned a common radius of 0.01 nm. The dielectric boundary was defined as the contact and re-entrant surface obtained by rolling a probe of radius 0.14 nm over the protein, and the solvent-accessible surface area was defined by the location of the center of the same probe. The protein was centered on a cubic grid of 6 nm edges with a uniform grid spacing of 0.05 nm and rotated to maximize the solute-to-wall distance (>0.5 nm). Unless specified otherwise, the value of the relative dielectric permittivity  $\epsilon_i$  of the protein interior was set to 2, the ionic strength  $I$  was set to 0 M and the effective microscopic interfacial tension  $\gamma$  was set to 10.46 kJ mol<sup>-1</sup> nm<sup>-2</sup> (Hünenberger *et al.*, 1999). A value of 78 was used for the dielectric permittivity  $\epsilon_w$  of water.

A first series of calculations was dedicated to the evaluation of the accuracy and the optimization of the continuum-electrostatics calculations. These included: (i) a comparison of two models to compute  $\Delta G_{\text{N-m}}^{\text{w}}$ , either by considering only the electrostatic potential computed for the wild-type protein [as in Meeker *et al.* (Meeker *et al.*, 1996)] or by performing additional continuum electrostatics calculations for all modeled mutant proteins; (ii) an assessment of the dependence of the results on the protonation states of the histidine residues;

(iii) an optimization of selected empirical parameters of the continuum electrostatics calculations so as to achieve a more quantitative (predictive) model. Because Equation 3 suggests the existence of a linear correlation between  $\Delta G_{\text{N-m}}^{\text{w}}$  and  $\Delta\Delta G_{\text{N-D}}$ , the linear correlation coefficient  $r$  between the computed  $\Delta G_{\text{N-m}}^{\text{w}}$  and the experimental  $\Delta\Delta G_{\text{N-D}}^{\text{exp}}$  for a given set of charge mutants was used as a measure of the accuracy and predictive ability of the model. The set of mutant proteins considered comprises the mutants D19A, D21A, D40A, D77A, D83A, D95A, E10A, E43A, E52A, E57A, E67A, E73A, E75A, E101A, E122A, E129A, E135A, K6A, K9A, K16A, K24A, K28A, K45A, K48A, K49A, K53A, K63A, K64A, K70A, K71A, K78A, K84A, K97A, K110A, K116A, K127A, K133A, K134A, K136A, H8A, H46A, H121A, H124A, R35A, R81A, R87A, R105A and R126A. For reasons detailed below and unless specified otherwise, the mutants H46A, H121A, D83A and D95A were excluded from the correlation analysis determining  $r$  and the least-squares fit lines displayed in the figures.

Finally, the ability of the optimized model (after determination of the  $\alpha$  and  $\beta$  parameters in Equation 3) to predict relative mutant stabilities was tested for the corresponding glycine mutants (Meeker *et al.*, 1996) and for other mutants of SNase [D19C, E52C, E57C, K28C, K64C, K71C, K78C, K84C, K97C, K116C and R105C (Byrne and Stites, 1995; Gillespie and Shortle, 1997a); D19N, D21N, D77N, E73Q, E75Q, E135Q, K63Q and K70Q (Schwehm *et al.*, 2003); E43S and E43S + R87G (Weber *et al.*, 1991)]. All together, a total of 117 charge mutants were considered in this study.

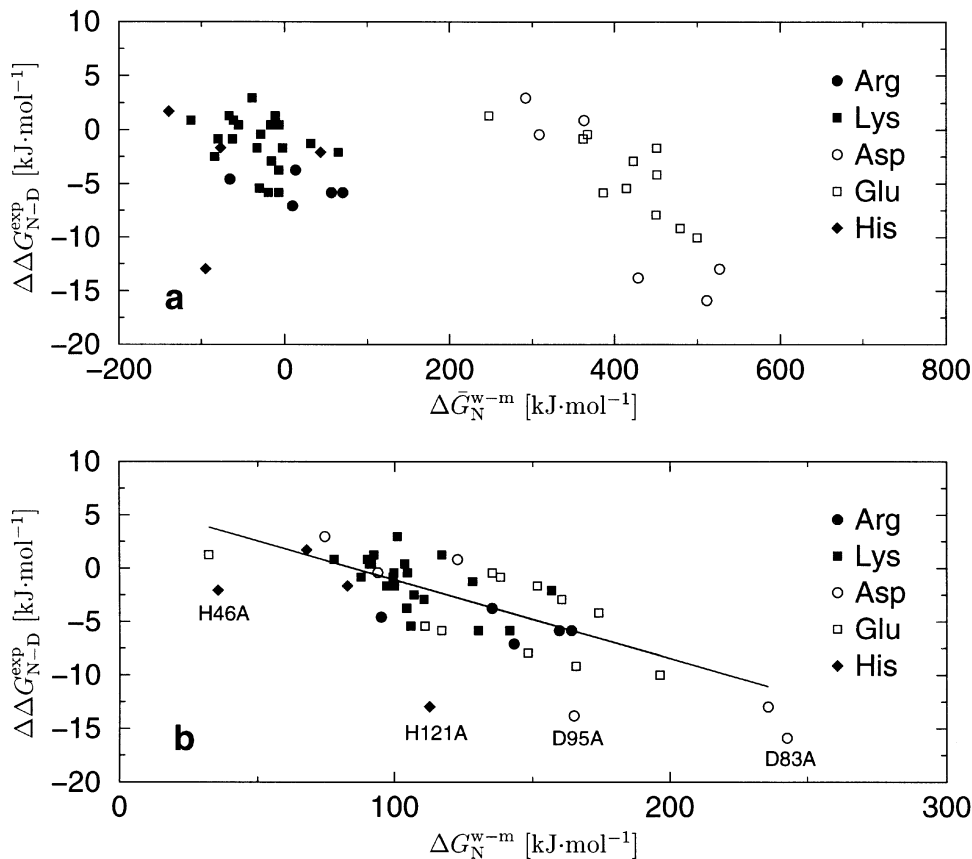
## Results

As an attempt to explain the previously reported absence of correlation between relative mutant stabilities and local electrostatic environment of the mutated charged residues (Meeker *et al.*, 1996), the free energy difference  $\Delta\tilde{G}_{\text{N-m}}^{\text{w}}$  upon mutating a charged residue to alanine was first estimated for each mutant as

$$\Delta\tilde{G}_{\text{N-m}}^{\text{w}} = \frac{1}{2} \sum_i \Delta q_i \phi^{\text{w}}(\mathbf{r}_i) \quad (7)$$

where  $\mathbf{r}_i$  is the atomic coordinate vector of atom  $i$  in the wild-type protein,  $\phi^{\text{w}}(\mathbf{r}_i)$  is the corresponding electrostatic potential at this site and  $\Delta q_i$  is the difference in atomic charge induced by the mutation (non-zero solely for atoms of the mutated residue). These free energy differences indeed do not show any correlation with the corresponding experimental stability changes  $\Delta\Delta G_{\text{N-D}}^{\text{exp}}$  (data not shown), in agreement with previous results (Meeker *et al.*, 1996). This absence of correlation is probably due to (i) the neglect of solvent reorganization upon mutating a charged residue (i.e. the electrostatic potentials corresponding to mutant and wild-type proteins are not identical), and (ii) artifacts arising from the use of a finite-difference estimate for the Coulombic free energy contribution  $G_{\text{Cb}}$ . The latter problem may be alleviated by performing two separate calculations of the electrostatic potential in which the solvent permittivity is set to either  $\epsilon_w = 78$  or  $\epsilon_w = \epsilon_i = 2$ . The free energy difference  $\Delta\tilde{G}_{\text{N-m}}^{\text{w}}$  is then estimated as





**Fig. 2.** Correlation between the calculated mutation free energy in the native state  $\Delta G_N^{w-m}$  and the experimental relative stability  $\Delta\Delta G_{N-D}^{exp}$  for 48 SNase mutants involving a charged residue mutated to alanine (Meeker *et al.*, 1996); see Equation 3. In (a), Equation 8 is used to calculate  $\Delta G_N^{w-m}$ . In (b), Equation 9 is used to calculate  $\Delta G_N^{w-m}$ .

$$\Delta\bar{G}_N^{w-m} = \frac{1}{2} \sum_i \Delta q_i [\phi_{78}^w(\mathbf{r}_i) - \phi_2^w(\mathbf{r}_i)] + \frac{1}{4\pi\epsilon_o\epsilon_i} \times \left\{ \sum_i \sum_{j>i} \frac{q_i^m q_j^m}{r_{ij}^m} - \sum_i \sum_{j>i} \frac{q_i^w q_j^w}{r_{ij}^w} \right\} \quad (8)$$

As shown in Figure 2a, neglecting solvent reorganization leads to a clear separation in the calculated values of  $\Delta\bar{G}_N^{w-m}$  depending on the sign of the mutated charged residue. The calculated values of  $\Delta\bar{G}_N^{w-m}$  for the arginine, lysine and histidine mutants are in the range  $-140$  to  $70$   $\text{kJ mol}^{-1}$ , whereas those for the aspartic and glutamic acid mutants are in the range  $250$ – $530$   $\text{kJ mol}^{-1}$ . This observation can be understood in view of the large overall positive charge ( $+13e$ ) of the SNase protein fragment. Owing to the neglect of solvent relaxation,  $\Delta\bar{G}_N^{w-m}$  is dominated by the direct Coulomb interaction between side chains. This contribution largely disfavors the mutation of a negative to a neutral residue and favors the mutation of a positive to a neutral residue. Overall, only a very weak correlation between  $\Delta\bar{G}_N^{w-m}$  and experimental stability changes  $\Delta\Delta G_{N-D}^{exp}$  is found (linear correlation coefficient  $r = -0.44$ ).

Finally, the most accurate treatment requires that one also takes the solvent reorganization into account. In this case,  $\Delta G_N^{w-m}$  is calculated as the difference between the electrostatic free energies of the (modeled) mutant and wild-type proteins:

$$\Delta G_N^{w-m} = G_N^m - G_N^w = \frac{1}{2} \sum_i q_i^m [\phi_{78}^m(\mathbf{r}_i) - \phi_2^m(\mathbf{r}_i)] - \frac{1}{2} \sum_i q_i^w [\phi_{78}^w(\mathbf{r}_i) - \phi_2^w(\mathbf{r}_i)] + \frac{1}{4\pi\epsilon_o\epsilon_i} \left\{ \sum_i \sum_{j>i} \frac{q_i^m q_j^m}{r_{ij}^m} - \sum_i \sum_{j>i} \frac{q_i^w q_j^w}{r_{ij}^w} \right\} + \Delta G_{N,np}^{w-m} \quad (9)$$

This evaluation is computationally more expensive and requires the modeling of each mutant protein, but greatly improves the correlation with the experimental stability changes, as shown in Figure 2b (linear correlation coefficient  $r = -0.76$  for the set of 44 mutants considered in the regression analysis). This correlation is significantly larger than those calculated for all other descriptors (e.g. the degree of burial of the mutated residues) considered previously (Meeker *et al.*, 1996). Notably, taking the solvent reorganization into account eliminates the clear separation that is observed for  $\Delta\bar{G}_N^{w-m}$  between mutants of positively charged residues and mutants of negatively charged residues (Figure 2a). The values of  $\Delta G_N^{w-m}$  are in the range  $30$ – $240$   $\text{kJ mol}^{-1}$  for all mutants. Since the Coulomb contributions in Equations 8 and 9 are identical and the non-polar contribution  $\Delta G_{N,np}^{w-m}$  in Equation 9 is small (see below), this effect is caused by a change in the reaction-field contribution to the free energy. Allowing the solvent to relax upon mutation of a positively charged residue is strongly

disfavorable, while the opposite is true for the mutation of a negatively charged residue. The corresponding differences  $\Delta G_{\text{N}}^{\text{w-m}} - \Delta \tilde{G}_{\text{N}}^{\text{w-m}}$  are between 90 and 210 kJ mol<sup>-1</sup> for mutants involving positively charged residues and between -310 and -215 kJ mol<sup>-1</sup> for mutants involving negatively charged residues. This observation can be qualitatively rationalized by applying the Born model for solvation in the case of a spherical solute with a positive overall charge  $Q$ . If the charge of the protein is decreased to  $Q - 1$  (mutation of a positively charged residue), the reaction-field energy is proportional to  $-(Q - 1)^2$  if the solvent is allowed to relax to the charge change, whereas it is proportional (same proportionality constant) to  $-Q(Q - 1)$  if the solvent is not allowed to relax. In this case, the solvent relaxation contribution to the overall free energy is proportional to  $Q - 1$ , i.e. disfavorable. A similar reasoning shows that the solvent relaxation contribution is proportional to  $-(Q + 1)$ , i.e. favorable, when the charge of the sphere is increased to  $Q + 1$  (mutation of a negatively charged residue).

The values of  $\Delta G_{\text{N}}^{\text{w-m}}$  reported in Figure 2b arise from the partial cancellation of two large contributions. The Coulomb contributions to the difference  $\Delta G_{\text{N}}^{\text{w-m}}$  between mutant and wild-type protein are between -450 and -120 kJ mol<sup>-1</sup> for mutants involving positively charged residues and between 350 and 780 kJ mol<sup>-1</sup> for mutants involving negatively charged residues. On the other hand, the reaction-field (solvation) contributions to  $\Delta G_{\text{N}}^{\text{w-m}}$  are between 280 and 540 kJ mol<sup>-1</sup> for mutants involving positively charged residues and between -580 and -320 kJ mol<sup>-1</sup> for mutants involving negatively charged residues. The non-polar contributions to  $\Delta G_{\text{N}}^{\text{w-m}}$  are small in comparison with the two previous terms, ranging from -2.1 to 3.6 kJ mol<sup>-1</sup> for non-lysine mutants (-21.6 to -0.1 kJ mol<sup>-1</sup> for mutants involving the larger lysine residue). In the following discussion, all calculations of  $\Delta G_{\text{N}}^{\text{w-m}}$  were performed using Equation 9.

The mutants H121A, D83A, D95A and H46A are specifically indicated in Figure 2b and are not included in the regression analysis for the following reasons: (i) H121A and D83A are the two mutants that were found to be less than 96% native even in the absence of denaturant (Meeker *et al.*, 1996), (ii) D95A is the mutant with the largest  $m$ -value (Meeker *et al.*, 1996), which suggests that its energetics in the denatured state are severely altered compared with the wild-type protein (Shortle, 1995; Wrabl and Shortle 1999), and (iii) H46 is not charged in the present calculations (see below). If these mutants are included in the regression analysis, the correlation coefficient in Figure 2b changes slightly from  $r = -0.76$  to  $-0.72$ .

The experimental  $pK_{\text{a}}$ s of the histidine side chains in SNase are of the order of 5.7–6.8 (Alexandrescu *et al.*, 1988), which suggests a fractional extent of ionization at pH 7. The consequences are that (i) the best way to represent the charge states of the histidine residues in the present calculations must be investigated, and (ii) it may not be meaningful to consider the results for the histidine mutants on the same footing as the rest of the charge mutants of SNase. The results obtained using several charge-state combinations for the four histidine residues are summarized in Table I. For each combination, the linear correlation coefficients between the experimental  $\Delta \Delta G_{\text{N}}^{\text{w-m}}$  and the computed  $\Delta G_{\text{N}}^{\text{w-m}}$  are reported, calculated for all mutants ( $r''$ ), all mutants except D83A, D95A and H121A ( $r'$ ) and for all mutants except D83A, D95A, H121A and partially charged or uncharged histidines ( $r$ ). The coefficient  $r$  is systematically larger in magnitude than the correl-

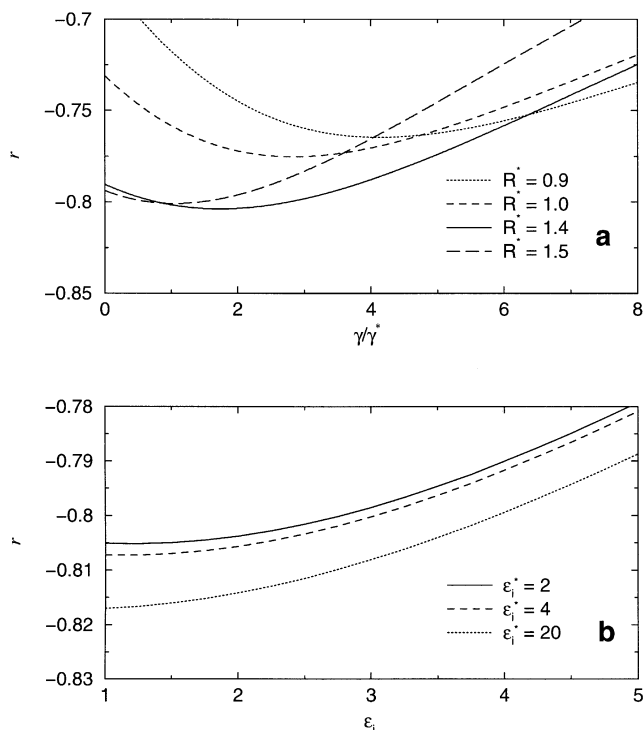
**Table II.** Linear correlation coefficients relating the experimental relative mutant stability  $\Delta \Delta G_{\text{N}}^{\text{w-m}}$  to the theoretical mutation free energy in the native state  $\Delta G_{\text{N}}^{\text{w-m}}$  (see Equation 3) for 44 mutants of SNase involving the mutation of a charged residue to alanine (Meeker *et al.*, 1996)

$\epsilon_i$	$R^*$	$I$ (M)	$r$	$\epsilon_i$	$R^*$	$I$ (M)	$r$
2.0	0.9	0	-0.72	4.0	0.9	0	-0.74
	1.0	0	-0.76		1.0	0	-0.76
	1.0	0.150	-0.73		1.0	0.150	-0.77
	1.1	0	-0.78		1.1	0	-0.78
	1.2	0	-0.79		1.2	0	-0.78
	1.3	0	-0.80		1.3	0	-0.79
	1.4	0	-0.80		1.4	0	-0.79
20.0	1.5	0	-0.80	20.0	1.5	0	-0.79
	0.9	0	-0.57		0.9	0.150	-0.75
	1.0	0	-0.58		1.0	0.150	-0.74
	1.1	0	-0.59		1.1	0.150	-0.75
	1.2	0	-0.59		1.2	0.150	-0.71
	1.3	0	-0.59		1.3	0.150	-0.73
	1.4	0	-0.60		1.4	0.150	-0.71
	1.5	0	-0.61				

Values are reported for different combinations of the solute internal relative permittivity  $\epsilon_i$ , scaling factor  $R^*$  applied to the atomic radii and ionic strength  $I$ .

ation coefficients from the analysis including all mutants ( $r''$ ). The differences in correlation are mainly caused by the mutants D83A, D95A and H121A and by a strong deviation from linearity for the uncharged (and partially charged) histidine residues. This observation suggests that the linear relationship postulated in Equation 3 is reasonable for charge mutants, but is not likely to hold for mutations that do not alter the charge of a residue. The results excluding uncharged (and partially charged) histidine residues and also the two aspartate mutants ( $r$ ) are relatively insensitive to the protonation states selected for the four histidine residues. Because the protonation-state combination involving charged histidines 8, 121 and 124 and uncharged (N<sub>E</sub>-protonated) histidine 46 (third entry in Table I) consistently shows the highest magnitude of the correlation coefficient irrespective of which residues are included or excluded in the least-squares fit analysis, this combination was chosen for the rest of the calculations.

Because continuum electrostatic models essentially rely on the application of a macroscopic theory at a microscopic level, a number of parameters involved in these models can only be given effective values, ultimately derived by calibration against experimental data. These parameters include the atomic charges and radii, the exact definition of the solute-solvent dielectric boundary, the dielectric permittivity of the solute and the interfacial tension coefficient. Although standard empirical values appear to work well for the present application (Figure 2b), it is of interest to (i) investigate the sensitivity of the results to these parameters so as to assess the reliability of the results and (ii) try to refine them further for the present problem so as to improve the predictive power of the model. Table II displays the linear correlation coefficient between the experimental  $\Delta \Delta G_{\text{N}}^{\text{w-m}}$  and the computed  $\Delta G_{\text{N}}^{\text{w-m}}$  calculated for different values of the internal permittivity  $\epsilon_i$ , a scaling factor  $R^*$  applied to the atomic radii and the ionic strength  $I$  of the solution. The model parameters have only a limited influence on the correlation coefficient, the exception being the internal permittivity. The best results are obtained for  $\epsilon_i = 2$  and 4 (at zero ionic strength) or  $\epsilon_i = 20$  (at an ionic strength  $I = 0.150$  M). Thus, it appears that the combinations of a low  $\epsilon_i$  with a low  $I$  or



**Fig. 3.** Linear correlation coefficient relating the experimental relative mutant stability  $\Delta\Delta G_{\text{N-D}}^{\text{exp}}$  to the theoretical mutation free energy  $\Delta G_{\text{N-D}}^{\text{w-m}}$  (Equation 3), extrapolated for changes in the model parameters (Equation 10). In (a), the interfacial tension coefficient  $\gamma$  is varied from its reference value  $\gamma^*$  for four values of the scaling factor  $R^*$  applied to the atomic radii. Here, the internal permittivity is  $\epsilon_i = 2$ . In (b), the internal permittivity  $\epsilon_i$  is varied for three values of the reference internal permittivity  $\epsilon_i^*$ . Here,  $R^* = 1.4$  and  $\gamma/\gamma^*$  is optimized at each  $\epsilon_i$  point for the best correlation.

a high  $\epsilon_i$  with a high  $I$  are both adequate to achieve the proper balance between direct Coulomb interactions and the reaction-field contribution. A high internal permittivity ( $\epsilon_i$  of  $\geq 20$ ) has often been used in continuum-electrostatics calculations after the observation that the accuracy of  $pK_a$  predictions in proteins is enhanced by increasing  $\epsilon_i$  (Antosiewicz *et al.*, 1994). In the present model, the fact that the interactions in the denatured state are implicitly taken into account may be the reason why a lower value of  $\epsilon_i$  also works well for predicting relative mutant stabilities.

In an attempt to optimize the value of the effective solute permittivity  $\epsilon_i$  and microscopic interfacial-tension coefficient  $\gamma$ , the empirical expression

$$G(\epsilon_i, \gamma) = \frac{\epsilon_i^*}{\epsilon_i} G_{\text{Cb}}(\epsilon_i^*) + \frac{\left(\frac{1}{\epsilon_i} - \frac{1}{\epsilon_w}\right)}{\left(\frac{1}{\epsilon_i^*} - \frac{1}{\epsilon_w}\right)} G_{\text{rf}}(\epsilon_i^*) + \frac{\gamma}{\gamma^*} G_{\text{np}}(\gamma^*) \quad (10)$$

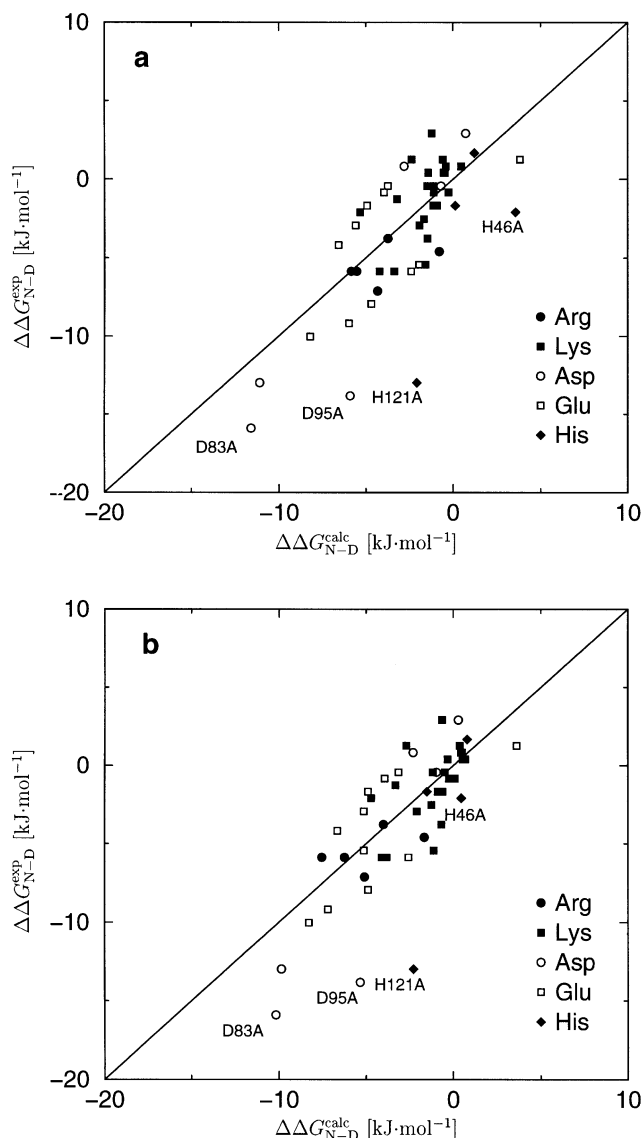
was used to extrapolate the electrostatic free energy  $G$  computed through Equation 4 using a given set of parameters  $\gamma^*$  and  $\epsilon_i^*$  to different parameter values  $\gamma$  and  $\epsilon_i$ . While the first and third terms in Equation (10) are exact, the second term is approximated by an expression derived from the Born model of ionic solvation (i.e. the solute is approximated by a sphere for estimating the reaction-field free-energy contribution). The results of this extrapolation are shown in Figure 3a for the

interfacial tension coefficient  $\gamma$  (for four different values of the scaling factor  $R^*$  applied to the atomic radii) and in Figure 3b for the solute internal permittivity  $\epsilon_i$  (based on three different reference values of the internal permittivity  $\epsilon_i^*$ ). Figure 3a shows that the correlation coefficient reaches a minimum for  $R^* = 1.4$  and  $\gamma/\gamma^* \approx 1.7$ , i.e. the correlation is improved by using larger atomic radii and an increased interfacial-tension coefficient. Note that the optimal value of  $\gamma$  is not independent of the choice of  $R^*$ , since a similar change in the non-polar free-energy contribution can be achieved by increasing either the surface area (through  $R^*$ ) or the interfacial tension coefficient. In fact, there exists a quasi-linear relationship between the optimal value of  $\gamma/\gamma^*$  and  $(1/R^*)^2$ . Figure 3b shows that an internal permittivity of 1–2 leads to the best correlation with experiment. This analysis suggests an ‘optimized’ parameter set defined by  $R^* = 1.4$ ,  $\gamma = 17.8 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  and  $\epsilon_i = 2$ , yielding a correlation coefficient of  $-0.80$ . It should be stressed, however, that the alteration of the parameters results in only a very modest gain of accuracy. This, again, indicates a moderate sensitivity of the theoretical results on the model parameters.

Adopting this optimized parameter set, the optimal values of  $\alpha$  and  $\beta$  to be used in Equation 3 are 0.95 and 2.3  $\text{kJ mol}^{-1}$ , respectively (excluding D83A, D95A, H46A and H121A in this determination). For the non-optimized parameter set, the corresponding values are 0.93 and 6.2  $\text{kJ mol}^{-1}$ . These values were used to compute estimated relative stabilities of mutants  $\Delta\Delta G_{\text{N-D}}^{\text{calc}}$  from the theoretical estimates of  $\Delta G_{\text{N-D}}^{\text{w-m}}$  through Equation 3. In Figure 4, the experimental stability changes  $\Delta\Delta G_{\text{N-D}}^{\text{exp}}$  are compared with the calculated stability changes  $\Delta\Delta G_{\text{N-D}}^{\text{calc}}$  for the 48 mutants involving the mutation of a charged residue to alanine (Meeker *et al.*, 1996), for calculations using either the non-optimized parameter set ( $R^* = 1$ ,  $\gamma = 10.46 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  and  $\epsilon_i = 2$ ; Figure 4a) or the optimal parameter set ( $R^* = 1.4$ ,  $\gamma = 17.8 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  and  $\epsilon_i = 2$ ; Figure 4b). The calculated energies are in good agreement with experimental values, yielding root-mean-square errors of 3.1 and 2.9  $\text{kJ mol}^{-1}$ , respectively, when considering all 48 mutants and 2.3 and 2.1  $\text{kJ mol}^{-1}$ , respectively, when omitting the mutants D83A, D95A, H46A and H121A.

Using Equation 3 together with the optimized parameter set and the values of  $\alpha$  and  $\beta$  determined for the alanine mutants,  $\Delta\Delta G_{\text{N-D}}^{\text{calc}}$  was evaluated for the corresponding 48 glycine mutants reported in (Meeker *et al.*, 1996) and the results are displayed in Figure 5. In the case of the glycine mutants, the root-mean-square error is as high as 4.9  $\text{kJ mol}^{-1}$ . The increased error is probably related to the significant increase in conformational flexibility induced by mutating a large residue to glycine. Increased flexibility is expected to destabilize the native state and therefore render  $\Delta\Delta G_{\text{N-D}}^{\text{exp}}$  more negative. Because such an effect is not accounted for by the present model, it is not surprising that the magnitudes of the calculated stability changes tend to be underestimated compared with their experimental counterparts.

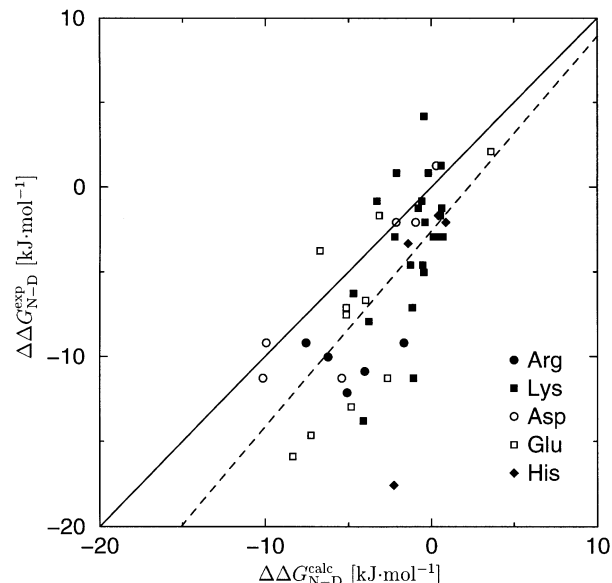
Figure 6 displays a comparison between computed and experimental stability changes for the following sets of SNase charge mutants: D19C, E52C, E57C, K28C, K64C, K71C, K78C, K84C, K97C, K116C, R105C [set 1 (Gillespie and Shortle, 1997a; Byrne and Stites, 1995)], D19N, D21N, D77N, E73Q, E75Q, E135Q, K63Q, K70Q [set 2 (Schwehm *et al.*, 2003)] and E43S and the double mutant E43S + R87G [set 3 (Weber *et al.*, 1991)]. The root-mean-square error is 3.0  $\text{kJ mol}^{-1}$  for the 21 mutants.



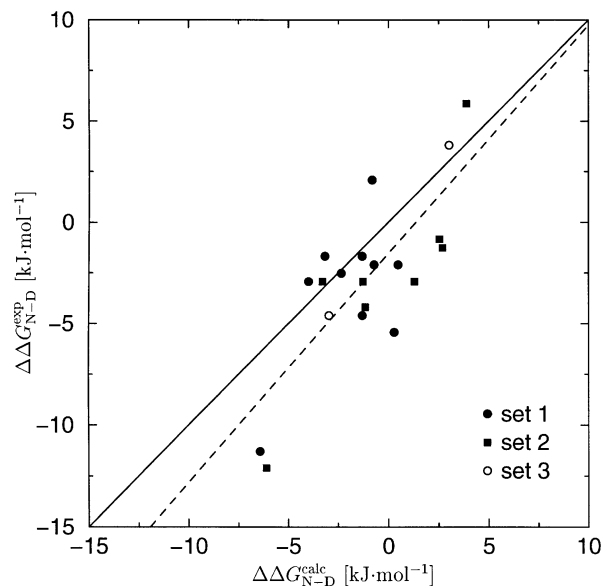
**Fig. 4.** Experimental relative stabilities  $\Delta\Delta G_{N-D}^{\text{exp}}$  versus calculated values  $\Delta\Delta G_{N-D}^{\text{calc}}$  for 48 SNase mutants involving a charged residue mutated to alanine (Meeker *et al.*, 1996). In (a), the non-optimized parameter set ( $\gamma = 10.46 \text{ kJ mol}^{-1}$ ,  $R^* = 1$ ,  $\epsilon_i = 2$ ) is used and  $\Delta\Delta G_{N-D}^{\text{calc}}$  is calculated from Equation 3 with  $\alpha = 0.93$  and  $\beta = 6.2 \text{ kJ mol}^{-1}$ . In (b) the optimized parameter set ( $\gamma = 17.8 \text{ kJ mol}^{-1}$ ,  $R^* = 1.4$ ,  $\epsilon_i = 2$ ) is used and  $\Delta\Delta G_{N-D}^{\text{calc}}$  is calculated from Equation 3 with  $\alpha = 0.95$  and  $\beta = 2.3 \text{ kJ mol}^{-1}$ . The line of perfect correspondence (which is identical with that of the best linear least-squares fit) is also displayed.

## Discussion

The aim of the present study was to investigate the role played by electrostatic interactions in determining the stability changes of staphylococcal nuclease (SNase) upon mutation of charged to uncharged residues. This was done using continuum-electrostatics calculations that attempt to account for the full thermodynamic cycle (Figure 1), something that is not routinely done today. The present calculations of the relative stabilities  $\Delta\Delta G_{N-D}^{\text{calc}}$  of charge mutants involve calculating the electrostatic potential both for the wild-type protein and for each modeled mutant protein and rely on the assumption that the electrostatic interactions in the denatured and native states are linearly related in the case of SNase.



**Fig. 5.** Experimental relative stabilities  $\Delta\Delta G_{N-D}^{\text{exp}}$  versus calculated values  $\Delta\Delta G_{N-D}^{\text{calc}}$  for 48 SNase mutants involving a charged residue mutated to glycine (Meeker *et al.*, 1996). The optimized parameter set ( $\gamma = 17.8 \text{ kJ mol}^{-1}$ ,  $R^* = 1.4$ ,  $\epsilon_i = 2$ ) was used to calculate  $\Delta\Delta G_{N-D}^{\text{calc}}$  through Equation 3 with  $\alpha = 0.95$  and  $\beta = 2.3 \text{ kJ mol}^{-1}$ . The lines of perfect correspondence (solid) and of the best linear least-squares fit (dashed) are also displayed.



**Fig. 6.** Experimental relative stabilities  $\Delta\Delta G_{N-D}^{\text{exp}}$  versus calculated values  $\Delta\Delta G_{N-D}^{\text{calc}}$  for 21 charge mutants of SNase. Set 1 corresponds to mutants in which a charged residue is mutated to a cysteine (Byrne and Stites, 1995; Gillespie and Shortle, 1997a) and set 2 to mutants in which a charged residue is mutated to a large polar residue (Schwehm *et al.*, 2003). Set 3 correspond to the active-site mutants E43S and E43S + R87G (Weber *et al.*, 1991). The optimized parameter set ( $\gamma = 17.8 \text{ kJ mol}^{-1}$ ,  $R^* = 1.4$ ,  $\epsilon_i = 2$ ) was used to calculate  $\Delta\Delta G_{N-D}^{\text{calc}}$  through Equation 3 with  $\alpha = 0.95$  and  $\beta = 2.3 \text{ kJ mol}^{-1}$ . The lines of perfect correspondence (solid) and of the best linear least-squares fit (dashed) are also displayed.

Neglecting the effect of solvent relaxation upon mutation, i.e. considering only the electrostatic potential of the wild-type protein, lacks consistency and leads to poor results. The good correlation between experimental and calculated stability



changes indicates that electrostatic interactions play a dominant role in determining the effect of this type of mutation on protein stability. This finding is consistent with the conclusions from a recent study of charge-deletion and charge-reversal mutants of SNase (Schwehm *et al.*, 2003) and one may conclude that charged residues contribute to the overall stability of the native state of SNase mainly through electrostatic interactions.

Because continuum-electrostatics calculations rely on a number of empirical parameters, it is essential to assess the dependence of the results on specific parameter values. In the present case, the accuracy of the predictions, as measured by the linear correlation coefficient  $r$  between the calculated and experimental stability changes, was found to be rather insensitive to model parameters such as atomic radii, solute permittivity and interfacial tension coefficient within the ranges considered. Increasing the atomic radii and the solute-solvent interfacial tension coefficient slightly and using an internal permittivity of 2 led to optimal accuracy of the calculations, although the improvement over standard parameters was only moderate. The results for the mutations of charged residues are also only weakly affected by the choice of the charge state of the four histidine residues in SNase. The root-mean-square error of the calculated relative stabilities for the 69 non-glycine mutants is in the range 2–3 kJ mol<sup>-1</sup>, i.e. of the order of  $k_B T$  at room temperature. The method can therefore be considered to reach quasi-predictive accuracy, the predictive power being limited, however, by the small range of the experimental stability values.

The calculations were performed under the assumption that the electrostatic interactions in the denatured state of SNase are attenuated, but qualitatively similar to those in the native state. This assumption was included in the calculations by assuming an approximately linear relationship between the free energy changes upon mutation of the residue in the native and denatured states (and thus between  $\Delta G_N^{\text{mut}}$  and  $\Delta\Delta G_{\text{ND}}^{\text{mut}}$ ). The fact that, under this assumption, a high correlation was found between calculated and experimental data for various combinations of model parameters is a hint towards its validity. Interestingly, the parameter  $\alpha$  relating  $\Delta G_D^{\text{mut}}$  to  $\Delta G_N^{\text{mut}}$  was determined using linear regression to be 0.95 for substituting charged residues with alanine. Such a high number suggests that the free energy change upon mutation is, on average, only 5% smaller in magnitude in the denatured state than the native state, which indicates that the electrostatic environment of charged residues in the denatured state is very similar to their environment in the native state. The value of  $\alpha$  is, however, sensitive to parameters such as  $\epsilon_i$  and  $I$  used in the calculations. Using an internal relative permittivity of 20 at an ionic strength of 0.150 M leads to a value of  $\alpha$  of only ~0.25. Note, however, that the validity of Equation 3 is independent of the magnitude of  $\alpha$ . Regardless of the actual value of  $\alpha$ , the suggestion that there might be a similarity between electrostatic interactions in the native and denatured states is in line with the conclusions drawn by Whitten and García-Moreno (Whitten and García-Moreno, 2000) from their observation of native-like  $pK_a$  values in the denatured state of SNase. What this energetic similarity means exactly in terms of the ensemble of structures characterizing the denatured state remains uncertain. However, these findings are in line with the set of structures determined for the fragment  $\Delta 131\Delta$ , considered as a model of the denatured state of SNase (Gillespie and Shortle, 1997b).

There is also evidence of residual electrostatic interactions in the denatured states of other small proteins such as barnase (Oliveberg *et al.*, 1994), ribonuclease Sa (Pace *et al.*, 2000) and the N-terminal domain of L9 (Kuhlman *et al.*, 1999) and theoretical models attempting to include these interactions have proved to be valuable. In particular, predictions of the pH-dependence of protein stability have been greatly improved by modeling the electrostatic interactions in the denatured state (Elcock, 1999; Zhou, 2002a,b). For example, the denatured state may be modeled explicitly by a single configuration obtained by performing an energy minimization of the native protein structure using a molecular mechanics force field including artificial van der Waals interactions [e.g. with a minimum-energy distance for all atom-atom interactions set to 0.6 nm (Elcock, 1999)]. The denatured state is thus represented by an expanded structure that retains the overall topology of the native state. The Gaussian-chain model (Zhou, 2002a,b) follows a different principle by empirically relating the interaction energy between two residues in the denatured state to the number of peptide bonds separating the residues. Our model for including electrostatic interactions in the denatured state, however, introduces no assumption about the structure (or ensemble of structures) characterizing this state and depends only on the corresponding interactions in the native state.

## Acknowledgements

The authors thank Wesley Stites for useful suggestions and for early access to experimental data. Financial support was provided by the Swiss National Science Foundation (Project 2100-061939).

## References

- Akke, M. and Forsén, S. (1990) *Proteins*, **8**, 23–29.
- Alexandrescu, A.T., Mills, D.A., Ulrich, E.L., Chinabi, M. and Markley, J.L. (1988) *Biochemistry*, **27**, 2158–2165.
- Anderson, D.E., Becktel, W.J. and Dahlquist, F.W. (1990) *Biochemistry*, **29**, 2403–2408.
- Antosiewicz, J., McCammon, J.A. and Gilson, M.K. (1994) *J. Mol. Biol.*, **238**, 415–426.
- Baskakov, I.V. and Bolen, D.W. (1998) *Biochemistry*, **37**, 18010–18017.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. and Hermans, J. (1981) In Pullman, B. (ed.), *Intermolecular Forces*. Reidel, Dordrecht, pp. 331–342.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.*, **112**, 535–542.
- Byrne, M.P. and Stites, W.E. (1995) *Protein Sci.*, **4**, 2545–2558.
- Dao-pin, S., Sauer, U., Nicholson, H. and Matthews, B.W. (1991a) *Biochemistry*, **30**, 7142–7153.
- Dao-pin, S., Soderlind, E., Baase, W.A., Wozniak, J.A., Sauer, U. and Matthews, B.W. (1991b) *J. Mol. Biol.*, **221**, 873–887.
- Daura, X., Mark, A.E. and van Gunsteren, W.F. (1998) *J. Comput. Chem.*, **19**, 535–547.
- Davis, M.E., Madura, J.D., Luty, B.A. and McCammon, J.A. (1991) *Comput. Phys. Commun.*, **62**, 187–197.
- Eftink, M.R. and Ramsay, G.D. (1997) *Proteins*, **28**, 227–240.
- Elcock, A.H. (1999) *J. Mol. Biol.*, **294**, 1051–1062.
- Gillespie, J.R. and Shortle, D. (1997a) *J. Mol. Biol.*, **268**, 158–169.
- Gillespie, J.R. and Shortle, D. (1997b) *J. Mol. Biol.*, **268**, 170–184.
- Green, S.M., Meeker, A.K. and Shortle, D. (1992) *Biochemistry*, **31**, 5717–5728.
- Grimsley, G.R., Shaw, K.L., Fee, L.R., Alston, R.W., Huyghues-Despointes, B.M.P., Thurlkill, R.L., Scholtz, J.M. and Pace, C.N. (1999) *Protein Sci.*, **8**, 1843–1849.
- Hünenberger, P.H., Helms, V., Narayana, N., Taylor, S.S. and McCammon, J.A. (1999) *Biochemistry*, **38**, 2358–2366.
- Hynes, T.R. and Fox, R.O. (1991) *Proteins*, **10**, 92–105.
- Kuhlman, B., Luisi, D.L., Young, P. and Raleigh, D.P. (1999) *Biochemistry*, **38**, 4896–4903.
- Loladze, V.V. and Makhadze, G.I. (2002) *Protein Sci.*, **11**, 174–177.

- Loladze,V.V., Ibarra-Molero,B., Sanchez-Ruiz,J.M. and Makhatadze,G.I. (1999) *Biochemistry*, **38**, 16419–16423.
- Lumry,R., Biltonen,R. and Brandts,J.F. (1966) *Biopolymers*, **4**, 917–944.
- Madura,J.D., Davis,M.E., Gilson,M.K., Wade,R.C., Luty,B.A. and McCammon,J.A. (1994) In Lipkowitz,K.W. and Boyd,D.B. (eds), *Reviews in Computational Chemistry*, Vol. 4. VCH, New York, pp. 229–267.
- Madura,J.D. et al. (1995) *Comput. Phys. Commun.*, **91**, 57–95.
- Meeker,A.K., García-Moreno,E.B. and Shortle,D. (1996) *Biochemistry*, **35**, 6443–6449.
- Oliveberg,M., Vuilleumier,S. and Fersht,A.R. (1994) *Biochemistry*, **33**, 8826–8832.
- Pace,C.N., Alston,R.W. and Shaw,K.L. (2000) *Protein Sci.*, **9**, 1395–1398.
- Sali,D., Bycroft,M. and Fersht,A.R. (1991) *J. Mol. Biol.*, **220**, 779–788.
- Sanchez-Ruiz,J.M. and Makhatadze,G.I. (2001) *Trends Biotechnol.*, **19**, 132–135.
- Schwehm,J.M., Fitch,C.A., Dang,B.N., García-Moreno,E.B. and Stites,W.E. (2003) *Biochemistry*, **42**, 1118–1128.
- Scott,W.R.P., Hünenberger,P.H., Tironi,I.G., Mark,A.E., Billeter,S.R., Fennen,J., Torda,A.E., Huber,T., Krüger,P. and van Gunsteren,W.F. (1999) *J. Phys. Chem. A*, **103**, 3596–3607.
- Serrano,L., Horovitz,A., Avron,B., Bycroft,M. and Fersht,A.R. (1990) *Biochemistry*, **29**, 9343–9352.
- Shortle,D. (1995) *Adv. Protein Chem.*, **46**, 217–247.
- Shortle,D. and Ackerman,M.S. (2001) *Science*, **293**, 487–489.
- Shortle,D., Stites,W.E. and Meeker,A.K. (1990) *Biochemistry*, **29**, 8033–8041.
- Spector,A., Wang,M., Carp,S.A., Robblee,J., Hendsch,Z.S., Fairman,R., Tidor,B. and Raleigh,D.P. (2000) *Biochemistry*, **39**, 872–879.
- Tissot,A.C., Vuilleumier,S. and Fersht,A.R. (1996) *Biochemistry*, **35**, 6786–6794.
- van Gunsteren,W.F., Billeter,S.R., Eising,A.A., Hünenberger,P.H., Krüger,P., Mark,A.E., Scott,W.R.P. and Tironi,I.G. (1996) *Biomolecular Simulation: the GROMOS96 Manual and User Guide*. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, pp. 1–1042.
- Weber,D.J., Meeker,A.K. and Mildvan,A.S. (1991) *Biochemistry*, **30**, 6103–6114.
- Whitten,S.T. and García-Moreno,E.B. (2000) *Biochemistry*, **39**, 14292–14304.
- Wrabl,J. and Shortle,D. (1999) *Nat. Struct. Biol.*, **6**, 876–883.
- Zhou,H.-X. (2002a) *Proc. Natl Acad. Sci. USA*, **99**, 3569–3574.
- Zhou,H.-X. (2002b) *Biochemistry*, **41**, 6533–6538.

Received March 24, 2003; revised August 11, 2003; accepted September 12, 2003