References

Gilson,M.K. and Honig,B. (1987) *Nature*, **330**, 84–86.
Linse,S., Johansson,C., Brodin,P., Grundström,T., Drakenberg,T. and Forsén,S. (1991) *Biochemistry*, **30**, 154–162.
Mehler,E.L.and Solmajer,T. (1991) *Protein Engng*, **4**, 903–910.
Svensson,B., Jönsson,B. and Woodward,C. (1990) *Biophys. Chem.*, **38**, 179–183.
Szebenyi,D.M.E. and Moffat,K. (1986) *J. Biol. Chem.*, **261**, 8761–8777.

Johan Åqvist
Department of Molecular Biology
Uppsala University
Uppsala Biomedical Centre, Box 590
S-75124 Uppsala
Sweden

Reply to 'Comment on models for electrostatic effects in proteins' by J.Åqvist

In his comment to our paper (Mehler and Solmajer, 1991), Åqvist raises some points concerning the calculation of pK shifts in calbindin D_{9k}. In repeating one of our calculations, he finds that the electrostatic interaction energy substantially overestimates the experimental value, and the results given in our figure 2(c) suggest that an additional factor of 0.5 has been applied to the screened Coulomb potential (SCP) calculations. It is, in fact correct that all the electrostatic energies we calculated for the calbindin case were divided by a factor of two in order to determine the change in electrostatic free energy between the mutant and wild-type proteins. This factor does not arise from any electrostatic considerations, but must be included because of the stoichiometry of the reaction.

Since multiple site equilibria in proteins are discussed in several texts, and the definition of the free energy was precisely given by Linse *et al.* (1991), we did not discuss it further in our paper. To clarify this point, we outline our procedure very briefly below, primarily relying on the discussion in Tanford (1961).

For a system of identical and independent binding sites it can be shown that the free energy can be expressed in the form

$$\Delta G = \Delta G^\circ + RT\ln\left(\frac{\theta}{(1-\theta)C}\right) \tag{1}$$

where θ is the degree of association and C is the concentration of the binding species. In obtaining equation (1), it is essential to note that ΔG and ΔG° represent free energy changes per mole of combining sites. The form of equation (1) can be extended to identical, but interacting, binding sites by allowing ΔG° and the equilibrium constant to vary with the average number of molecules associated with each macromolecule. With the above definition, one can now express ΔG° in terms of an intrinsic and electrostatic contribution, that is,

$$\Delta G_i^\circ = \Delta G_i^\circ(\text{int}) + N_0 w$$

where N_0 is Avogadro's number, $N_0 w$ is the electrostatic contribution to the free energy and ΔG_i° (int) is the intrinsic contribution to the standard free energy obtained from a consideration of the hypothetically totally discharged system. The subscript i refers to the ith form of the protein, i.e. wild-type or mutant in the present application. To compute the change in free energy between two forms of the system one has

$$\Delta\Delta G_{ij}^\circ = [\Delta G_j^\circ(\text{int}) - \Delta G_i^\circ(\text{int})] + N_0\Delta w \tag{2}$$

Here the term in brackets is just the difference in the intrinsic free energies between systems i and j, which is assumed to be negligible, and the term $N_0\Delta w$ is the difference in electrostatic free energy between i and j. It is emphasized, however, that the quantities in equation (2) still represent free energy changes per mole of interaction site.

For calbindin D_{9k}, the change in electrostatic free energy due to mutating one or more charged residues was calculated using the SCP with different charge models and the finite difference method for solving the Poisson–Boltzmann (FDPB) equation (Gilson and Honig, 1987) with one charge model. In order to use these values to estimate shifts in pK_1K_2, which can be compared to the experimental results, it should be noted that Linse *et al.* (1991) defined $\Delta G_{tot} = -RT\ln(K_1K_2)$ for two moles of Ca^{2+} ions. Therefore, to bring $N_0\Delta w$, defined in equation (2) into coincidence with $\Delta\Delta G_{tot}$, it is necessary to take $N_0\Delta w = 2 \times 2.303RT\Delta p(K_1K_2)$. In view of this requirement, Åqvist's interpretation to consider $\epsilon(r)$ as scaled up by a factor of two is clearly incorrect.

There is, of course, a valid objection which can be raised against the above procedure in that it assumes that the two sites are identical, although interaction is permitted. Linse *et al.* (1991) showed that there are various degrees of interaction between the Ca^{2+} binding sites in the different mutants, and structurally it is obvious that the two binding sites are not alike. At present this objection can only be answered by pointing out that the results do not seem to be overly sensitive to the differences in the binding sites. Moreover, it seems questionable whether such differences could be accounted for by the approximate model we are using, or whether the differences are greater than we showed to exist between the different charge models.

In the latter part of his comment, Åqvist discusses the use of an effective dielectric constant of 92.9 (in view of the above discussion the actual value should be 46.45) for screening the Coulombic interaction, and points out that this is the optimal value leading to r.m.s. errors smaller than any given in table II of Mehler and Solmajer (1991). This value is obtained by fitting to the experimental data of calbindin, and essentially reduces the dielectric permittivity to an arbitrary parameter. The justification for this alternative point of view, taken by several authors, is not clear. It suffers from at least two difficulties: (i) the effective dielectric constant would become available only after the measured values of ΔpK, for the specific protein in question, could be fitted to Coulomb's law and would thus not be useful as a theoretical predictive device; and (ii) the values which are obtained in each case become uncorrelated and are not appropriate for use in a generalized force field. Moreover, the dielectric permittivity is a well defined property of matter and its functional form at microscopic distances was derived long ago by Debye (1929) using purely theoretical arguments. The main objections to Debye's approach are the lack of explicit boundaries between the components, and the internal fields are calculated for locally homogeneously polarized matter using the Lorentz (1880) relationship. Using this approach, Onsager (1936) tried to calculate the dipole moment of water and obtained a value which was much too low. However, central to Onsager's approach was the assumption that the dipole moment of water is a constant, independent of environment. In the light of modern molecular

theory, which shows that overlap and short range forces are important in condensed states, this assumption no longer seems so valid, and therefore the objections to using Debye's approach for polar solvents are considerably weakened (see also Bucher and Porter, 1986).

The conceptual and computational simplicity of this approach makes it a very attractive alternative to computationally more demanding methods for calculating changes of equilibrium properties in proteins, and for its use in modeling bulk solvent effects in computer simulations (Ramstein and Lavery, 1988; Solmajer and Mehler, 1991). However, it is essential to explore the limitations of the technique. We have pointed out (Mehler and Solmajer, 1991) that in the case of closely lying charged groups, i.e. hydrogen-bonded and salt-bridged moieties, where quantum effects such as charge transfer and polarization become significant, the simple sigmoidal function becomes ill defined and a more rigorous microscopic treatment is more appropriate.

In his comment, Åqvist points out the promising results obtained with the Monte Carlo (MC) treatment described by Svensson *et al.* (1990). We certainly agree with this remark. Our error analysis (Mehler and Solmajer, 1991, table II) showed that for calculating pK shifts, the MC method yielded excellent results. More significantly, this approach has a wider generality than the SCP, which seems to be essentially limited to considering interactions between charged groups. For example, in the calbindin case the SCP combined with the Debye ionic screening reproduced the shifts in pK well, because of charge annihilation. However, we were unable to develop a satisfactory model for calculating the changes in binding in the wild-type protein because of changes in the ionic strength. These latter quantities were also well reproduced by Svensson *et al.*'s (1990) MC approach. Therefore, for the present it appears that there is no single method which is best suited for all applications, but that the appropriate approach will have to be selected for the application of interest based on available resources, required accuracy, etc.

References

Bucher,M. and Porter,T.L. (1986) *J. Phys. Chem.*, **90**, 3406–3411.
Debye,P. (1929) *Polar Molecules.* Chemical Catalog Co., New York, p. 115.
Gilson,M.K. and Honig,B. (1987) *Nature*, **330**, 84–86.
Linse,S., Johansson,C., Brodin,P., Grundström,T., Drakenberg,T. and Forsen,S. (1991) *Biochemistry*, **30**, 154–162.
Lorentz,H.A. (1880) *Ann. Physik.*, **9**, 641.
Mehler,E.L. and Solmajer,T. (1991) *Protein Engng*, **4**, 903–910.
Onsager,L. (1936) *J. Am. Chem. Soc.*, **58**, 1486–1493.
Ramstein,J. and Lavery,R. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 7231–7235.
Solmajer,T. and Mehler,E.L. (1991) *Protein Engng*, **4**, 911–917.
Svensson,B., Jönsson,B. and Woodward,C. (1990) *Biophys. Chem.*, **38**, 179–183.
Tanford,C. (1961) *Physical Chemistry of Macromolecules.* John Wiley & Sons, New York, pp. 526–539.

E.L.Mehler and T.Solmajer
Department of Structural Biology
Biocenter of the University of Basel
CH-4056 Basel
Switzerland

Reply

The calculated change in electrostatic free energy, obtained with Coulomb's law, for doubly Ca^{2+}-loaded calbindin caused by the removal (mutation) of a single negative charge (in kcal/mol and Å) is

$$\Delta G_{el}^{calc} = 332 \left[\frac{Q}{r_l\ \epsilon(r_l)} + \frac{Q}{r_l\ \epsilon(r_2)} \right]$$

where Q is the charge of the calcium ions and r_i is the distance from the mutated group to the ith ion binding site. Here, we have written ϵ as a function of r to indicate a possible distance dependence. The calculated quantity above is expressed per mole of protein since the contributions from the two sites are added together. The experimentally observed quantity is

$$K_1K_2 = \frac{[PCa_2]}{[P][Ca]^2}$$

where $[PCa_2]$, [P] and [Ca] are the concentrations of doubly loaded protein, unloaded protein and Ca^{2+} respectively. As can be seen, the total binding constant is also defined per mole of protein. Hence, no correcting factor is required to make the two definitions coincide. By dividing the calculated electrostatic energy by n (i.e. the number of binding sites, which in this case is two), as was done in Mehler and Solmajer (1991), ΔG_{el}^{calc} becomes expressed per binding site instead. For the hypothetical case of identical non-interacting binding sites, the electrostatic contribution to the binding energy would then be independent of the number of sites per protein molecule. This definition contrasts, however, with the one used by Mehler and Solmajer (1991) (see Linse *et al.*, 1991) for the experimental free energy, namely, $\Delta G^{obs} = -RT\ln K_1K_2$. For identical non-interacting sites we would have $\Delta G^{obs} = -nRT\ln K$, K now being the same for all sites, and ΔG^{obs} would depend linearly on n.

The MC calculations by Svensson *et al.* (1990) (see also Svensson *et al.*, 1991) employ a high uniform dielectric constant of $\epsilon = 78.7$. The effective dielectric constant in these calculations becomes even higher because of the presence of counter ions (they were present even at 'zero ionic strength' in order to neutralize the net negative charge of the protein). Thus, the good agreement between the MC simulations and experiments also shows that a high dielectric constant is required to reproduce the observed shifts, and it is clear that the stoichiometric factor suggested by Mehler and Solmajer (1991) was not employed in the MC simulations of Svensson *et al.* (1990, 1991) either.

My example given in the comment above shows that an effective dielectric constant of $\epsilon = 92.9$, in combination with Debye–Hückel screening, is able to describe the calbindin experiments better than any of the methods discussed in Mehler and Solmajer (1991). This merely indicates that the physics (or electrostatics) of this particular problem is rather uncomplicated, and can be well described by a simple Coulomb law (and Debye–Hückel screening in the case of non-zero ionic strength). The problem with using the above approach for predictive purposes is, as pointed out by Mehler and Solmajer above, that one has to know the value of ϵ in order to be able to make any progress. Although the dielectric permittivity might be a well defined property of homogeneous matter, this does not seem to be the case for inhomogeneous systems such as protein solutions, judging from the numerous discussions in the literature on this subject. This is also probably the best argument for microscopic methods in which the concept does not need to be introduced.

References

Linse,S., Johansson,C., Brodin,P., Grundström,T., Drakenberg,T. and Forsén,S. (1991) *Biochemistry*, **30**, 154–162.