The nature of the ion binding interactions in EF-hand peptide analogs: free energy simulation of Asp to Asn mutations

Blaise Prod'hom[1] and Martin Karplus

Department of Chemistry, Harvard University, Cambridge, MA 02138, USA

[1]Present address: Institut de Physiologie, Faculté de Medecine, Université de Lausanne, CH-1005 Lausanne, Switzerland

The binding of the La^{3+} ion to a tridecapeptide, which is a model for the EF-hand in calcium-binding proteins, is studied in solution by free energy simulations. The calculations analyze the effect on the La^{3+} ion binding of the mutation of Asp to Asn for side chains that interact directly with the ion. The results are compared with the measurements of Marsden,B.J., Hodges,R.S. and Sykes,B.D. (1989) Biochemistry, 28, 8839, on the same system. They found that the Asp to Asn mutation has only a small effect on the binding; the observed differences in the free energies on changing one Asp to an Asn are between −0.3 and 1.8 kcal/mol. This result is analyzed by alchemical simulations for the tridecapeptide in the bound (loop) structure and free (extended) form. The free energy changes due to the mutation of an Asp to an Asn are large and positive for both the bound and free forms. However, since the values of the free energy changes are calculated to be similar in the two forms, the difference in the binding free energy of Asp and Asn peptides is found to be small, in agreement with experiment. By use of thermodynamic integration, the various contributions to the free energy changes are estimated. In the complexed form, the Asp to Asn mutation is favored by the reduction in the repulsive interaction with other charged residues of the peptide; it is disfavored by the reduction of the stabilization of the ion and the surrounding water has a small effect. When the peptide adopts an extended conformation in the absence of the ion, the mutation Asp to Asn is strongly disfavored by the interactions with the water and is favored by the interactions within the peptide. The results demonstrate the essential role of contributions to the binding of EF-hands from interactions other than those between the ion and the charged amino acid side chains. The results obtained from the simulations suggest, in accord with crystal structures of La^{3+} bound to various ligands, that the calcium-binding loop complexed with La^{3+} in solution has a significantly different structure from that observed in proteins.

Key words: calcium binding motif/ion binding/EF hands/thermodynamic integration

Introduction

Free energy simulations in macromolecular systems have been shown to be useful for understanding the mechanisms of drug selectivity, protein stability and pathological mutations (Brooks et al., 1988; Karplus and Petsko, 1990; Straatma and McCammon, 1992). Such simulations provide a bridge between biochemical and thermodynamic measurements at the macroscopic level and descriptions of phenomena at the molecular level.

A problem of biological interest that can be analyzed by free energy simulations is ion binding to proteins. Since calcium plays an essential role in the control of cellular mechanisms from metabolism and cell signalling to secretion and contraction (Evered and Whelon, 1986; Smith and Augustin, 1988; Rasmussen, 1989; Papahadjopoulos et al., 1990; Wier, 1990), it is of importance to have a quantitative understanding of the binding mechanism. Among several families of calcium-binding proteins, the troponin C superfamily, which includes calmodulin and parvalbumin, is by far the most thoroughly studied (Forsén, 1989; Strynadka and James, 1989). The proteins of this family have binding sites with characteristic sequences and structures (Kretsinger and Nockolds, 1973; Herzberg and James, 1985; Statyshur et al., 1988; Godzik and Sander, 1989). Generally, the binding sites consist of a loop of 12 consecutive amino acids flanked by two α-helices. These supersecondary structures, often called EF-hands (Kretsinger and Nockolds, 1973; Godzik and Sander, 1989), are grouped in pairs in many calcium-binding proteins. As crystal structures have improved [e.g. compare the recent structures of carp parvalbumin (Swain and Amma, 1989; Swain et al., 1989; Kumat et al., 1990) with the pioneering study of Kretsinger and Nockolds (1973)], the ion coordination has been shown to be more similar throughout the superfamily (Strynadka and James, 1989).

It has been observed that the affinity of these proteins for calcium depends in part on the number of acidic (negatively charged) residues of the loop participating in the coordination of the ion (Reid and Hodges, 1985; Marsden et al., 1990). However, the binding is not simply related to the number of carboxylate groups involved. So far no mutant studies have been made to quantitate this observation. Marsden et al. (1988) have conducted NMR studies of the interaction between rare earth ions and several tridecapeptides with sequences based on that of the calcium-binding site III of rabbit skeletal muscle troponin C. They showed that such peptides in solution do not form well-defined complexes with Ca^{2+}, although at high calcium concentration there is evidence from circular dichroism that the peptide structure is affected. When trivalent lanthanide ions are used, the peptides adopt a conformation similar to the loop of an EF-hand (Marsden et al., 1989). Using nuclear Overhauser enhancement (NOE) measurements they were able to estimate 33 hydrogen−hydrogen distances of the peptide in solution in the presence of the lanthanide ion. With gadolinium-induced ^1H relaxation measurements, they determined 31 gadolinium−hydrogen distances. The root mean square difference (r.m.s.d.) between these distances and those measured in the crystal structure of turkey troponin C was 1.1 Å. Minimization and molecular dynamics simulations in vacuo with the NOE- and Gd-induced relaxation constraints yielded mean structures that had r.m.s.d.s from the crystal structure of 1.1 Å for the main chain and 1.8 Å for all heavy atoms. Small differences in the conformations of the peptides have been observed with different lanthanides (Gariepy et al., 1983). There was no evidence for a stable loop structure in the absence of ions, suggesting that the charge repulsions of the side chains leads to an extended random coil.

Since the solution structures of these peptides appear to be similar to that observed in calcium-binding proteins, Marsden *et al.* (1988) synthesized several tridecapeptide homologs of the loop segment of site III of rabbit skeletal troponin C with some aspartate residues of the binding site substituted by asparagine. By studying a series of related peptides, they were able to estimate the contribution of each of these ligands to the lanthanum ion affinity. Although no detailed structural studies of the modified peptides were made, it appears likely that the structures of the liganded loops are close to that of the peptide sequence whose structure was studied by NMR. Based on this assumption, Marsden *et al.* (1988) estimated the relative contribution of an Asn versus an Asp at the various positions and obtained values in the range −0.3 to 1.8 kcal/mol.

To understand the origin of the small difference in ion binding resulting from Asp to Asn substitutions, we have performed alchemical free energy simulations (Gao *et al.*, 1989; Tidor and Karplus, 1991; Straatma and McCammon, 1992) on a peptide, whose sequence corresponds to peptide I in the terminology of Marsden *et al.* (1988),

```
 1    2    3    4    5    6    7    8    9   10   11   12   13  (I)
Asp−Arg−Asp−Ala−Asp−Gly−Tyr−Ile−Asp−Ala−Glu−Glu−Leu
```

Free energy simulations were performed using the exponential formula and thermodynamic integration (Tidor and Karplus, 1991) in which residues 3 or 5 [peptide N and III in the Marsden *et al.* (1988) nomenclature], were transformed from Asp to Asn; both of these residues participate in the ion coordination. The strong Ca^{2+}-binding site (III) in rabbit troponin C, in fact, has an Asn instead of an Asp at position 3. In turkey and chicken troponin C, for which crystal structures are available (Strynadka and James, 1989), Tyr7 is replaced by Phe and Ala10 by Ile. The residues participating in the coordination shell of the calcium in these structures correspond to positions 1, 3, 5, 7 (all monodentate) and 12 (bidentate); the ligands of 1, 3, 5 and 12 to Ca^{2+} are side chain oxygens, while for Tyr7 the backbone carbonyl is ligated to the Ca^{2+} (Herzberg and James, 1985; Satyshur *et al.*, 1988). In addition, one water is coordinated with the calcium. The coordination number is 7 and the ligand atoms have an approximately pentagonal bipyramidal arrangement. The most recent refinements of the carp parvalbumin coordinates (Swain and Amma, 1989; Swain *et al.*, 1989; Kumat *et al.*, 1990) show a similar seven-coordinate arrangement of the CD loop around the Ca^{2+} ion, which differs somewhat from the earlier results (Kretsinger and Nockolds, 1973).
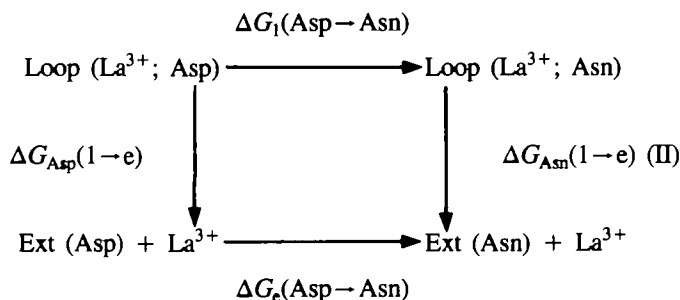
Although the overall structure of the loop appears to be similar in the crystal structure of the protein and in the peptide in solution, there is no evidence that the detailed geometry and the ion coordination are the same. The NMR results give little information on the side chain geometry and, in particular, the distances between the ion and possible ligands have not been measured. It is likely that when the Ca^{2+} ion is replaced by La^{3+} the coordination will change; the ionic radius of Ca^{2+} is 0.99 Å versus 1.15 Å for La^{3+} (Pauling, 1948) and the charge increases from +2 to +3. Significant differences were observed between the NMR spectra of Lu^{3+}- and La^{3+}-bound peptides (Gariepy *et al.*, 1983). In addition, there is the possibility that alterations in the loop structure are induced by the absence of the protein. Thus, the peptide loop studies of bound La^{3+} must be considered as approximate models of the Ca^{2+}-binding site in proteins, although they are of interest in themselves.

The simulations show that the free energy of mutation of a charged amino acid (Asp) into a polar amino acid (Asn) leads to nearly the same free energy change in two environments as different as a relatively protected cation-binding site and an extended chain where the amino acid is fully exposed to aqueous solution. To analyze this result, the individual contributions to the change in free energy of binding were estimated by the thermodynamic integration method (Gao *et al.*, 1989; Tidor and Karplus, 1991).

Methodology

Alchemical free energy simulations (Gao *et al.*, 1989; Tidor and Karplus, 1991; Straatma and McCammon, 1992) of the mutation of Asn into Asp at positions 3 and 5 of peptide I were performed. This direction was chosen because it was simpler to orient the oxygen of Asn to coordinate with the La^{3+}, although in an exact treatment the direction of the mutation should make no difference. We use the terms 'alchemical' and 'computer alchemy' for this type of simulation (Gao *et al.*, 1989; Tidor and Karplus, 1991; Straatma and McCammon, 1992) because we transmute one amino acid into another along a reversible path that makes it possible to calculate the free energy difference. Since no such path exists for the real system, the calculated free energy is not directly accessible by experiment. However, when alchemical simulations are done for the peptide in two different conformations (a loop structure corresponding to the bound system and an extended structure corresponding to the system without an ion), the difference between the two alchemical free energy changes can be measured. This is made clear by the thermodynamic cycle:

$$\Delta G_1(Asp \rightarrow Asn)$$

Loop $(La^{3+};$ Asp) $\longrightarrow$ Loop $(La^{3+};$ Asn)

$\Delta G_{Asp}(1 \rightarrow e)$ $\qquad\qquad\qquad$ $\Delta G_{Asn}(1 \rightarrow e)$ (II)

Ext (Asp) + La^{3+} $\longrightarrow$ Ext (Asn) + La^{3+}

$$\Delta G_e(Asp \rightarrow Asn)$$

where $\Delta G_1(Asp \rightarrow Asn)$ and $\Delta G_e(Asp \rightarrow Asn)$ correspond to the alchemical free energy change associated with the transformation Asp$\rightarrow$Asn in the loop conformation (with the ion) and the extended conformation (without the ion) respectively and $\Delta G_{Asn}(1 \rightarrow e)$ and $\Delta G_{Asp}(1 \rightarrow e)$ correspond to the chemical free energy change associated with the transition from the loop conformation (with the ion) to the extended conformation (without the ion) for the peptides containing Asn and Asp respectively.

We use the notation L1 for the La^{3+} binding loop with the peptide I sequence (see Introduction) and L3, L5 for the loop with Asp3$\rightarrow$Asn and Asp5$\rightarrow$Asn respectively. The corresponding notation E1, E3 and E5 is used for the extended chain without La^{3+}. Thus, the top horizontal line in scheme II corresponds to the mutation L1 to L3 or L5 and the bottom horizontal line to the mutation E1 to E3 or E5.

The chemical and alchemical free energy values can be compared by use of the double free energy difference, $\Delta\Delta G$,

$$\Delta\Delta G = \Delta G_{Asn}(1 \rightarrow e) - \Delta G_{Asp}(1 \rightarrow e) = \Delta G_e(Asp \rightarrow Asn) - \Delta G_1(Asp \rightarrow Asn) \tag{1}$$

Since the atomic coordinates of rabbit skeletal muscle troponin C were unavailable when we undertook this study and the published NMR data are not sufficient to generate accurate coordinates, the crystal coordinates of the CD site (residues 51–63) of carp muscle parvalbumin (Kretsinger and Nockolds, 1973) were used to generate the starting structure of the bound system; the coordinates were obtained from the Brookhaven Protein Data Bank (the structure used has the file name 3PTP; Bernstein et al., 1977). The r.m.s.d. between the main chains of the corresponding loops of troponion C and carp muscle β parvalbumin is stated to be 0.428 Å (Herzberg and James, 1985; Satyshur et al., 1988), suggesting that the choice of main chain coordinates is appropriate. To correspond to the chemical structure of the synthetic peptides, an acetyl group was introduced at the N-terminus, an amide group was introduced at the C-terminus and the side chains of the residues were modified to match the sequence of peptide I that was used for the simulation. The changes required are Glu2 → Arg, Lys4 → Ala, Ser5 → Asp, Phe7 → Tyr, Glu9 →Asn, Glu10 →Ala and Asp11 → Glu. In the original crystal structure (Kretsinger and Nockolds, 1973), which we used to build the loop model, the Ca^{2+} ion is coordinated approximately octrahedrally by six oxygen ligands (OD1 of Asp1, OD1 of Asp3, OG of Ser5, O of Phe7, OE1 of Glu9 and OE1 of Glu12 following the numbering system in scheme I). As already mentioned, this ligation scheme differs from the more recent structural data (Swain and Amma, 1989; Kumat et al., 1990), in which the ion is seven-coordinated (OD1 of Asp1, OD1 of Asp3, OG of Ser5, O of Phe7, OE1 of Glu9 and OE1 and OE2 of Glu12), which is essentially the same as in troponin C(III). Although the interactions in the model-built tridecapeptide differ significantly from those in the loop of troponin C, the structural rearrangements that occur in the simulation and the similarity of the resulting structures to those found in crystal complexes of La^{3+} (Hoard et al., 1965; Lind et al., 1965) suggest that the procedure used is meaningful. In going from parvalbumin to troponin C all of the side chains of the altered residues except Asp5 and Glu9 point towards the outside of the loop and do not interact significantly with the ion or the rest of the peptide. The modified residues were built so that common atoms coincide and the remaining atoms were added in the most probable conformation and then minimized with the rest of the peptide fixed. Thus, for Asp5 in place of Ser, the Cβ position was overlapped. Polar hydrogen atoms were constructed with the HBUILD facility (Brünger and Karplus, 1988). The loop was embedded in a sphere of water with a radius of 14 Å centered on the ion. The system includes 351 molecules of water introduced by three overlays with a box of TIP3P water and was confined by use of the stochastic boundary method (Brooks and Karplus, 1989); an 11 Å molecular dynamics region and 3 Å Langevin dynamics boundary region were used. While the water molecules close to the surface of the sphere were acted on by a boundary potential (Brooks and Karplus, 1989), the peptide loop was far enough from the border of the sphere (i.e. all of it inside the 11 Å radius) to be simulated solely by molecular dynamics. For long-range truncation, the shift option was used for electrostatics and switch for van der Waals with a cut-off of 9.0 Å for both electrostatic and van der Waals interactions. The shift option has been shown to yield satisfactory results in comparison with simulations without truncation (Loncharich and Brooks, 1989; Stote and Karplus, 1991). The SHAKE algorithm (van Gunsteren and Berendsen, 1977) was used to constrain the bonds between hydrogen atoms and heavy atoms. The system was minimized and equilibrated at 327 K. The water was first

equilibrated for 1 ps while the peptide atoms were kept fixed. Then 5 ps of simulations were done while imposing the NOE- and Gd-induced relaxation constraints of Marsden et al. (1988) on the hydrogen–hydrogen and ion–hydrogen distances. Finally, 7 ps of equilibration were done without constraints. The free energy simulations were performed with no constraints.

As already mentioned, Marsden et al. (1989) have shown that peptide I does not form a loop in the absence of a lanthanide ion. We made the assumption that the side chains of the acidic residues would be fully exposed to solvent and that the value of the free energy of mutation in an extended peptide would be representative of the mean free energy change calculated from an ensemble of structures. An extended peptide was built by setting the ϕ and ψ dihedral angles of the main chain to 180° and minimizing the structure in vacuo; the other parameters had the default values in the CHARMM topology file (Brooks et al., 1983). The peptide was placed in a 13 Å radius sphere of water; the region between 11 and 13 Å was used as a buffer region. Because of its extended structure, only five residues were within the sphere; the rest of the peptide was deleted. The conditions and protocols of the simulations of the peptide in the loop conformation, were also applied to the extended conformation simulations; a total of 310 water molecules was included.

The simulations were carried out with the BLOCK facility (Tidor and Karplus, 1991) in the program CHARMM (Brooks et al., 1983). The force constants, atomic partial charges and Lennard–Jones parameters of the peptide atoms were taken from the parameters of the CHARMM 19 polar hydrogen set. A modified TIP3P model (TIP3P/C) (A.MacKerrell and M.Karplus, personal communication), which differs slightly from the original Jorgensen TIP3P model (Jorgensen, 1981), was used to represent the water molecules. A simple point charge model was used for La^{3+} since it has no f electrons outside the core. More generally it has been suggested on structural and thermodynamic grounds that the f electrons play a relatively small role in the rare earth chelates (MacKay et al., 1962; Hoffmann et al., 1977; Cossy et al., 1989). The $R_{min}/2$ and ϵ Lennard–Jones parameters of the lanthanum ion were taken to be 1.71 Å and −0.12 kcal/mol respectively. These parameters yield a lanthanum–TIP3P/C interaction of −75.5 kcal/mol at 2.35 Å. For comparison, a recent ab initio effective core potential calculation of the lanthanum–water interaction gives a minimum of −83.5 kcal/mol at 2.42 Å (Meier et al., 1990). Although electronic polarization effects are expected due to the presence of La^{3+}, the calculations are simplified by choosing parameters that yield the approximately correct interaction energies without explicit polarization. This is analogous to the introduction of implicit polarization by the use of large charges in TIP3P water, for example (Jorgensen, 1981).

For the free energy simulations, a hybrid potential energy function of the form $V(\lambda) = (1 - \lambda) V_A + \lambda V_B$ was introduced where λ is a parameter that goes from zero [$V(0) = V_A$, the energy function for the initial system] to one [$V(1) = V_B$, the final system]. To determine the alchemical free energy changes, $\Delta G(A \rightarrow B)$, both the exponential formula and thermodynamic integration (Mezei and Beveridge, 1986; Gao et al., 1989; Tidor and Karplus, 1991) were used; the exponential formula is

$$\Delta G = G_B - G_A = -k_B T \sum_i \ln < \exp(- \frac{(\Delta V)(\Delta \lambda_i)}{k_B T}) >_{\lambda_i} \quad (2)$$

and thermodynamic integration formula is

$$\Delta G = G_B - G_A = -\int_0^1 <\Delta V>_\lambda \, d\lambda \cong \sum_i <\Delta V>_{\lambda_i}\Delta\lambda_i \quad (3)$$

In equations (2) and (3) the expression $<\;>_\lambda$ represents a simulation with the energy function $V(\lambda)$; $\Delta V = V_B - V_A$. The mutations were done in five steps, corresponding to $\lambda = 0.1$, 0.3, 0.5, 0.7 and 0.9. For each step, the system was equilibrated for 10 ps and then sampled for 10 ps.

To obtain a correctly oriented Asn side chain with its oxygen coordinating with the ion, the simulations were started with a model-built structure for Asn and mutating it to Asp. In the starting structure for the free energy simulation, the duplicated side chain had the common atoms of Asn and Asp in the same positions. Thus, the only difference between V_A and V_B is that the former has an Asn side chain and the latter an Asp side chain. Although the Asn and Asp side chains interact with the rest of the system during the simulation, they do not interact with each other.

Results

We first describe the structural results and then consider the binding thermodynamics and its analysis.

Structural aspects

Figure 1a and b shows the loop and extended structures with the Asn and Asp copies of the mutated side chain. The main chain r.m.s.d.s between the initial model-built structure (L1) and the L3 and L5 structures were 1.68 Å and 2.40 Å respectively after equilibration in a sphere of water. Asp9 has moved out of the ion coordination sphere and is exposed to solvent. In all three peptide loops (L1, L3 and L5), the La^{3+} ion coordinates 10 oxygens. In L3, both Asp1 and Asp5 interact in a bidentate mode with the La^{3+} ion; Glu12 is coordinated in a monodentate mode, although the second oxygen is also relatively close (see Table I). Asn3 has its oxygen coordinated to the ion. In addition, the ion is coordinated with the main chain carbonyl oxygen of Tyr7 and three water molecules. The mean distance between La^{3+}

and the three waters is 2.64 Å (2.55, 2.49 and 2.88 Å), while the Tyr7 carbonyl oxygen is 2.35 Å from the ion. In L5, there is a corresponding coordination involving Asp1, Asp3, Asn5 and Glu12, though the roles of OE1 and OE2 are interchanged. Also, the backbone carbonyl oxygen of Tyr7 is not in direct contact with the ion. It is replaced by a water oxygen, bringing the number of water molecules in the coordinates sphere to four. The mean ion−water distance is 2.60 Å (2.67, 2.55, 2.55 and 2.64 Å).

Since the simulations done were Asn3→Asp for L3 and Asn5→Asp for L5, both L3 and L5 have the peptide sequence L1 as their final state. Thus, the structures of these two L1 loops can be compared. They are similar but not identical (the main chain r.m.s.d. is 1.89 Å) and the La^{3+} ion is not coordinated in exactly the same way. In both cases, the ion is coordinated by 10 oxygen atoms, but the oxygens involved are not all the same. L1 (L5) coming from the mutation Asn5→Asp has three carboxylate groups liganding the ion in a bidentate fashion [Asp1, Asp3 and Glu12; Glu12 changes from unidentate to bidentate while going from L5→L1 (L5)], the mutated Asp5 stays in a unidentate mode and three water molecules supply the three additional oxygen ligands [mean ion−water distance, 2.53 Å (2.56, 2.52 and 2.51 Å)]. L1 (L3) resulting from the mutation Asn3→Asp has two carboxylate groups in a bidentate mode (Asp1 and Glu12), while Asp3, the mutated residue, and Asp5 interact with the ion in a unidentate mode; the remainder of the coordination sphere is made up by the Tyr7 main chain oxygen (ion−oxygen distance 2.63 Å) and three water molecules with mean ion−oxygen distance of 2.57 Å (2.55, 2.49 and 2.88 Å). Thus, the L1 loops essentially retain the structure that they had in the L3 and L5 form, except that in L5 the additional oxygen of Glu12 replaces one of the water molecules. It is likely that the two different L1 structures reflect two possible forms of nearly equal free energy. Because the essential coordination is very similar, the effect on the free energy simulation results is expected to be small.
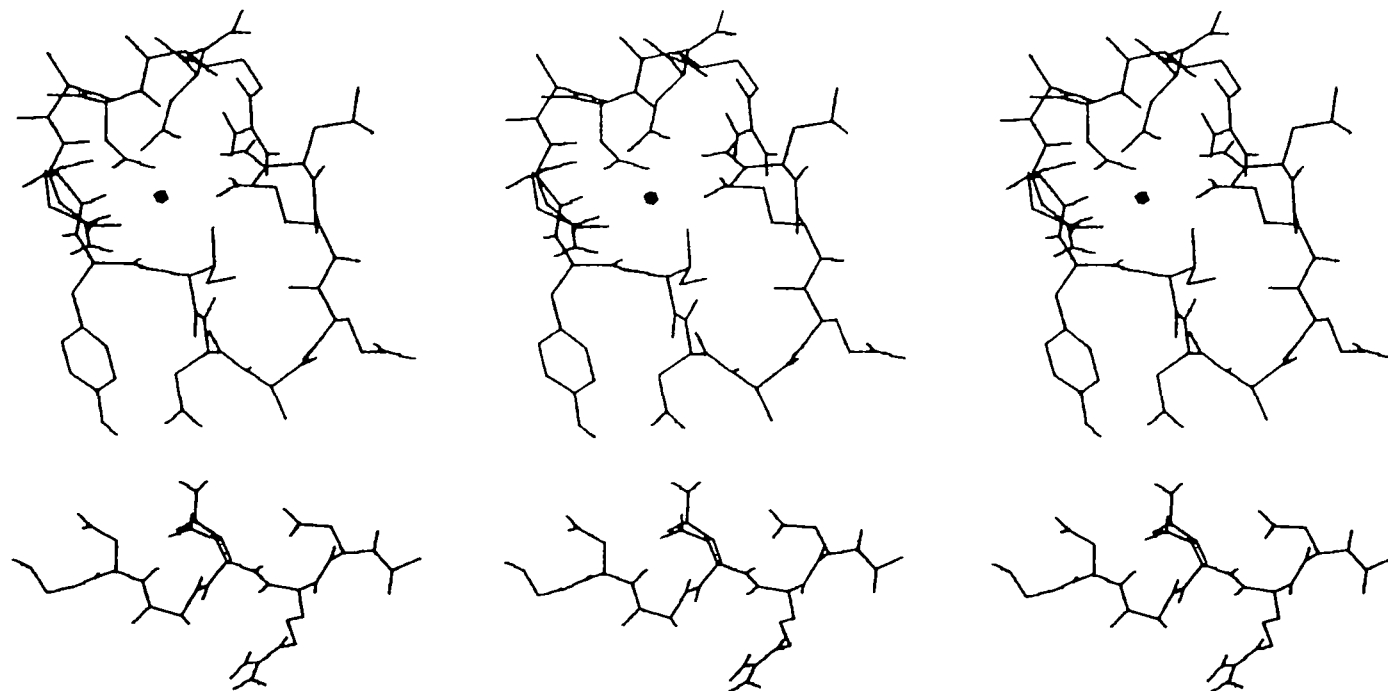

Fig. 1. (a) Cross-eyed stereo plot of the loop structure of peptide 1 with the ion (black dot) and peptide Asn5 superimposed on Asp5. (b) Cross-eyed stereo plot of the extended structure with Asn3 superimposed on Asp3; only the five residues included in the simulation sphere are shown.

The peptide in the bound conformations forms a bowl inside which the La^{3+} ion is located and the water molecules form a cap. The structure is thus significantly different from the Ca^{2+}-binding loops found in the proteins. However, it is reminiscent of La^{3+} complexes with EDTA for which crystal structures are available. In the crystal structure of the complexes $La^{3+}EDTA(H_2O)_3(26)$ and $La^{3+}EDTAH(H_2O)_4$ (Lind *et al.*, 1965), La^{3+} is coordinated by one oxygen of each of the four carboxyl groups and the two nitrogens, which form a bowl and three or four H_2O molecules which form a cap. The observed $La^{3+}-x$ distances are 2.51 (2.54) Å for x equal to O^-, 2.76 (2.87) Å for x equal to N and 2.58 (2.59) Å for x equal to O of H_2O; the two values given in each case correspond to those observed in the two different compounds. Thus, the overall structure and distances for the oxygen ligands are very similar to those found here. In lanthanide crystals formed with guanidinium acetate or nicotinate where the La^{3+} is surrounded by oxygen atoms the coordination number can vary from 7 to 10 (Sinha, 1976). Some considerations of the structure and energetics of the lanthanides as a function of ion size are given in MacKay *et al.* (1962) and Hoffmann *et al.* (1977). It has been observed that the same general type of cage structure is found for the EDTA chelates of the different lanthanides. However, the details of the structure (e.g. number of coordinated H_2O molecules) are sensitive to the ion radius. There is a decrease in coordination from 10 to eight as the ion size decreases from La^{3+} (ionic radius 1.15 Å) to Lu^{3+} (ionic radius 0.85 Å). These radii are to be compared with the value of 0.99 Å for Ca^{2+}. X-ray diffraction measurements of lanthanum chloride solutions give an ion to water oxygen distance of 2.48–2.58 Å and a coordination number of eight to nine (Smith and Wertz, 1975; Habenschuss and Spedding, 1980).

The mean numbers of water molecules, whose oxygen is at a distance smaller than 3.3 Å from OD, OE, ND of Asp, Asn or Glu were examined in the bound and extended structures. The mean number of water molecules around an Asp in the extended peptide is 6.4, slightly less than the value (7.1) found for a non-ligating Asp in the loop with bound La^{3+}. The ligating Asp residues in the loop structure have a mean of 3.2 water molecules around them, of which approximately half (mean 1.4) are also in the coordination sphere of the ion. In the extended conformation the mean number of water molecules around an Asp is 6.0; in the loop structures, all of the Asn side chains are coordinated with the La^{3+} ion. There are no experimental data or other simulation results for comparison.

Thermodynamics

To relate the alchemical simulation results to the measurements by Marsden *et al.* (1988) of the difference in the La^{3+}-binding free energy between L1 and L3 or L5, we make use of the thermodynamic cycle shown in Scheme II. The top line represents the mutations of the peptide in the loop conformation in the presence of the ion, while the bottom line represents the mutation in the extended conformation of the peptide in the absence of the ion. Since the solvation free energy of the free ion is constant, its contribution cancels in $\Delta\Delta G$, which appears on the right-hand side of equation (1). The vertical lines reflect the difference in free energy of binding for peptide I and the mutants determined from the experimental binding affinity data; this difference gives the $\Delta\Delta G$ on the left-hand side of equation (1). The measured values of the affinity constant and the corresponding free energies are shown in Table II. As can be seen L3 has the highest affinity, L1 is intermediate and L5 has the lowest affinity, though all of the differences are small.

The free energy differences calculated for the alchemical change (Asp to Asn) in the loop or extended conformation by thermodynamic integration (TI) and the exponential formula (EF) are presented in Table III. The differences between the free energy values calculated by the two methods vary between 0.7 and 2.9 kcal/mol. This sets a lower limit to the uncertainty of the results for this system with the present protocol. Although these values cannot be compared directly with the experiment, they are of considerable interest. The alchemical changes are sufficiently large that the inherent error (of the order of 10%) is tolerable. The mutation of the negatively charged Asp into the polar Asn is always very unfavourable, as indicated by the positive sign of the free energy differences. The values are rather similar in the loops and in the extended structures, ranging from 41.9 kcal/mol in E1 to E5 with the EF calculation to 48.1 kcal/mol in L1 (L5) to L5 with the TI calculation. The values are much smaller than the difference in the La^{3+} interaction energies for Asp and Asn. The calculated interaction between La^{3+} and a single acid group in the binding site varies between −246 kcal/mol for Asp3 in L1 (L3) to −327 kcal/mol for Asp1 in L5; the much weaker $La^{3+}-Asn$ interaction varies between −63 and −70 kcal/mol. These values are obtained by averaging the interaction energy over 10 ps of the simulation.

Table I. Distances between La^{3+} and the oxygens of the acid residues in the loop conformations (Å)

	L3	L1(L3)[a]	L5	L(L5)[1]
Asp1 O1	2.58	2.59	2.48	2.53
Asp1 O2	2.38	2.45	2.41	2.52
Asp3 O1	2.29[b]	3.51	2.50	2.45
Asp3 O2	−[b]	2.34	2.46	2.39
Asp5 O1	2.40	2.32	2.36[b]	2.42
Asp5 O2	2.77	3.95	−[b]	4.32
Glu12 O1	2.39	2.59	2.97	2.58
Glu12 O2	2.94	2.46	2.33	2.52

[a]L1(L3) corresponds to the structure L1 obtained from a simulation that started with L3; L1(L5) is defined in the same way.
[b]The residue corresponds to an Asn in L3 or L5 and so has only a carbonyl oxygen that can coordinate the ion.

Table II. Experimental affinities[a]

Marsden *et al.* (1988)[a] nomenclature	Type of loop	$K(mol^{-1})$	$\Delta G(1 \rightarrow e)$ (kcal/mol)[a]
I	L1	1.4×10^5	7.11
N	L3	2.5×10^5	7.46
III	L5	1.8×10^4	5.88

[a]$\Delta G(1 \rightarrow e)$ refers to the vertical arrows in Scheme II.

Table III. Calculated alchemical free energy changes (kcal/mol)[a]

	Loop (with La^{3+})		Extended	
	TI	EF	TI	EF
$\Delta G(Asp3 \rightarrow Asn)$	47.6	46.4	44.3	47.2
$\Delta G(Asp5 \rightarrow Asn)$	48.1	45.3	42.6	41.9

[a]Although the simulations were done in the direction Asn to Asp, the free energy changes are listed as Asp to Asn.

Table IV shows the calculated and experimental $\Delta\Delta G$ values. The difference between the calculated $\Delta\Delta G$ values obtained with the two methods (TI and EF) is larger than the experimental binding free energy differences. Thus, predicting the relative affinity of the different peptides for the lanthanide ion is not possible. The only important result is that the differences in affinities are small and of the same magnitude as the experimental results. Nevertheless, we note that comparison of the $\Delta\Delta G$ values of L3/L1 and L5/L1 shows that the calculated order agrees with that found experimentally.

Decomposition of the free energy change resulting from the Asp to Asn mutation into contributions from the components of the system has been performed with the thermodynamic integration method, as described previously (Gao et al., 1989; Tidor and Karplus, 1991). The results are shown in Table V, which gives the contributions of the water, the peptide and the ion; the self-term is dominated by the change in the number of degrees of freedom and makes only a small contribution. Although there are quantitative differences between the results for the two mutants, the overall features are very similar and, as already mentioned, the differences may not be significant. Thus, we concentrate on aspects of the results that are common to the two sets of calculations.

Of particular interest is the possibility of interpreting the fact that the free energy of mutation of a charged residue (Asp) into a polar one (Asn) is similar in the aqueous solvent environment and in the cation-binding site (see Table V). In the bound form the largest free energy term is the loss of the interaction between La^{3+} and the Asp side chain, which is replaced by the much smaller interaction between La^{3+} and the neutral, although polar, Asn side chain. A significant fraction of this destabilization (approximately 100 out of 150 kcal/mol) is regained by the reduction in repulsive interactions between the closely spaced negatively charged side chains in the loop conformation in going from Asp to Asn (see also below). The solvent contribution is small in the loop configuration with the bound ion. There is some indication, in accord with the structural results (see above) that there is a compensation between the solvent and the peptide contributions for the two mutants in the bound form, i.e. the

solvent plus peptide values are very similar (-95.1 and -93.7 kcal/mol for L1$\rightarrow$L3 and L1$\rightarrow$L5 respectively), while the individual terms differ by more than 14 kcal/mol.

For the extended configuration in the absence of the La^{3+} ion, the qualitative features of the interactions are very different. There is of course no La^{3+} interaction term, the reduction in the repulsive interactions is much smaller (in the order of 30 rather than 100 kcal/mol) and the solvent contribution is dominant. This is not unexpected because the residues are further from each other and completely immersed in solvent. The calculated values (87.8 kcal/mol for L3 and 68.3 kcal/mol for L5) are of the order of the free energy of solvation of acetate (84 kcal/mol) (Yu and Karplus, personal communication) and similar to the value obtained in a free energy simulation form of acetate to acetamide (-71.6 kcal/mol) in aqueous solution (Bash et al., 1987).

Further decomposition of the peptide free energy change into contributions from the individual amino acids is shown in Table VI. The side chain atoms and the $C\alpha$ atom of each residue are included. The overall changes in backbone interactions (excluding $C\alpha$) are listed separately. Although not considered in the previous discussion, the latter make a significant contribution. The Asp to Asn mutation is favored due to the decrease in repulsion between the side chain and the main chain carbonyl groups. For Asp3$\rightarrow$Asn, the effect is to destabilize the bound form relative to the extended by 4.3 kcal/mol, while there is a reverse effect of 8.8 kcal/mol for Asp5$\rightarrow$Asn. The side chains which strongly favor the mutation Asp$\rightarrow$Asn are the amino acids with a carboxylic group that is close to the mutated residue. For analysis, we define the distance between two acidic side chains as the mean distance between the oxygens. In the folded state, the charged residues tend to be closer together than in the extended form. Thus, they have a large influence on the free energy of mutation; the values range from -16.2 kcal/mol for Glu12 at 5.4 Å from the mutated residue in L5 to -39.1 kcal/mol for Asp5 at 4.02 Å from the mutated residue of L3. Carboxylate residues which are far from the mutated residue (and from the ion) make only a small contribution, i.e. Asp9 and Glu11, whose oxygens are 10.0 and 12.7 Å way from Asn3 and 8.4 and 13.8 Å away from Asn5, have almost no effect. Correspondingly, the interaction of the Asp residues in the extended form of the peptide is smaller,

Table IV. Calculated and experimental $\Delta\Delta G$ values (kcal/mol)[a]

System	Calculation			Experiment
	TI	EF	Mean	
Asp3$\rightarrow$Asn	-3.3	$+0.8$	-1.3	$+0.35$
Asp5$\rightarrow$Asn	-5.5	-3.4	-4.5	-1.23

[a]$\Delta\Delta G$ is defined in equation (1); see also footnote to Table III.

Table V. Contributions to the mutation free energies (kcal/mol)[a]

	ΔG(Asp3$\rightarrow$Asn)		ΔG(Asp5$\rightarrow$Asn)	
	Loop	Extended	Loop	Extended
La^{3+}	150.1	–	145.9	–
Solvent	9.6	87.8	-4.5	68.3
Peptide	-106.7	-38.5	-89.2	-20.8
Self	-5.4	-5.0	-4.0	-4.9
Total	47.6	44.3	48.1	42.6

[a]Values are calculated by the thermodynamic integration method (see text); see also footnote [a] of Table III.

Table VI. Peptide side chain contributions to the mutant free energies (kcal/mol)

	ΔG(Asp3$\rightarrow$Asn)		ΔG(Asp5$\rightarrow$Asn)	
	Loop	Extended	Loop	Extended
Asp1	-29.5	-16.0	-31.8	0.0
Arg2	3.4	1.6	0.0	1.6
Asp3	–	–	-30.9	-13.3
Ala4	2.1	2.9	2.0	1.4
Asp5	-38.4	-11.0	–	–
Gly6	0.1	1.0	4.1	3.0
Tyr7	0.2	–	4.2	–
Ile8	0.1	–	3.4	–
Asp9	-0.6	–	-2.2	–
Ala10	0.0	–	0.1	–
Glu11	0.0	–	0.0	–
Glu12	-31.4	–	-15.9	–
Le13	0.0	–	0.2	–
Main chain without $C\alpha$	-12.8	-17.1	-22.3	-13.5

[a]Values are calculated by the thermodynamic integration method (see text); see also footnote [a] of Table III.

−11.4 to −16.1 kcal/mol. There is an approximate inverse correlation between the distance separating a given carboxylate side chain from the mutated Asp and the contribution of the residue to the free energy of mutation. If an effective dielectric constant D_{eff} is used to describe the dependence of ΔG on the distance r (in Å) between two charges q_1 and q_2 (in electronic units) (Russell *et al.*, 1987; Bashford and Karplus, 1990),

$$\Delta G = \frac{332 \, q_1 q_2}{r \, D_{eff}}$$

the calculated free energies yield an effective dielectric constant between 2 and 4.

In both conformations, loop and extended, the contributions of the other side chains are very small (see Table VI). The largest contribution of the non-coordinating amino acids comes from Tyr7 in L5 with a value of 4.2 kcal/mol. The contribution of the van der Waals term to the alchemical free energy changes per side chain is also very small. The values range from 0.3 kcal/mol for $L1(L5) \rightarrow L5$ to 1.0 kcal/mol for $L1(L3) \rightarrow L3$. Thus, the van der Waals contribution is less than 3% of the total, which is dominated by the electrostatic term. However, for the $\Delta\Delta G$ values, the van der Waals contribution, which is −0.3 kcal/mol for Asp3 → Asn and 0.4 kcal/mol for Asp5 → Asn, is 10−25% of the total. As found in other studies (Gao *et al.*, 1989; Tidor and Karplus, 1991), the cancellation of the larger electrostatic term in $\Delta\Delta G$ is much greater than that for the van der Waals contribution.

The analysis of Marsden *et al.* (1988) and of Reid and Hodges (1985) on the effect of a mutation of Asp to Asn in different positions on the stability of the complex can be examined in light of the above results. The observation that the mutant Asp3 → Asn leads to stronger binding, in contrast to all other cases where there is a reduction in binding when an Asp is replaced by an Asn, is explained by Marsden *et al.* (1988) on the basis of the conclusion of Reid and Hodges (1985) that Asn3, which corresponds to the wild type, acts as a spacer to reduce the electrostatic repulsion between the charged groups, i.e. it is argued that the complex is stabilized by having Asn instead of Asp, even though the interaction with the ion is weaker. Although the simulations are not precise enough to draw conclusions concerning the difference in the overall binding free energy (the difference is only 0.3 kcal/mol), an analysis of the model of Reid and Hodges (1985) is possible. From Table V, we see that ΔG (Asp3 → Asn) in the bound loop has indeed a larger stabilizing effect due to the reduction of charged side chain repulsions (−106.7 kcal/mol) than the Asp5 → Asn mutation (−89.2 kcal/mol). The difference of −17.5 kcal/mol is much larger than the observed change in binding free energy. The calculations show essentially the same difference (−17.7 kcal/mol) in the extended form, so that the overall contribution of this term is small. Thus, the implicit assumption in the model of Reid and Hodges (1985) that only the bound configuration has to be considered is apparently not valid. From the decomposition in Table VI, it is clear that the major difference between the effect of the two mutations in the bound loop is that the repulsive interaction between Asp3 and Glu12, which are neighbors in the complex (at Y and −Z), is much larger than that of Asp5 and Glu12, which are more distant neighbors (at Z and −Z). In the extended form, neither Asp3 nor Asp5 interact with Glu12, which is far away. The differences arise in the extended form because Asp3 interacts repulsively with both Asp1 and Asp5 while Asp5 has a repulsive interaction only with Asp3,

again due to the distances involved. Thus, the phenomenon being considered involves a complex cancellation of several factors, suggesting that any simple rule may be valid for only a small number of cases.

Conclusion

Molecular dynamics free energy simulations of the mutation of aspartate residues into asparagine have been performed and used to analyze the binding of an La^{3+} ion to a tridecapeptide that is a possible model for a calcium-binding loop in proteins. The results are compared with NMR structural and binding studies of the same system. Although the NMR work was undertaken to model the EF-hands of calcium proteins such as troponin C, it is not clear that the solution structure of the calcium-binding site corresponds to that in a protein. Nevertheless, the results are of interest because they provide an understanding of the effect on ion binding of a charged to uncharged mutation. The differences in the alchemical free energy of the mutation Asp → Asn in two different conformations of the peptide are calculated; one corresponds to a loop structure for the peptide complexed with a lanthanum ion and the other to an extended structure that is a model for the peptide in the absence of the lanthanum ion. In qualitative agreement with the experiment, it was found that the overall change in the free energy of binding due to an Asp to Asn mutation is small.

Decomposition of the calculated free energy values by the thermodynamic integration method shows that the small effect of the mutation arises from the cancellation of large free energy differences that exist in the bound loop and in the free extended peptide. In each case, the electrostatic contribution, which dominates, leads to a free energy change of approximately 40 kcal/mol. Although the overall magnitude of the free energy change is nearly the same, the contributions are very different in the two environments. With the bound ion, the largest effect of the Asp to Asn mutation is a decrease in the favorable interaction with the ion, which is partly cancelled by a reduction in the repulsion between charged side chains; the water contribution is small in comparison to the contribution of the ion and the peptide. In the unfolded peptide, the main contribution to the free energy change comes from the water surrounding the mutated aspartate and the peptide contribution is small, though not negligible. Clearly, the 'bound loop conformation' in the absence of the ion is very unstable, relative to an extended solvated form, although no direct simulation was done to determine this free energy difference. If a loop structure is preserved in the protein in the absence of the ion, strong destabilizing repulsions are expected, though it is possible that the pK_a values of the interacting carboxylic acids are sufficiently perturbed that some of them would be protonated at normal pH values.

The present results, in correspondence with earlier calculations on other mutants (Gao *et al.*, 1989; Tidor and Karplus, 1991), suggest that the cancellation of large effects is a general phenomenon when electrostatic interactions are involved. This complicates simple interpretations of the small differences in the observed free energy changes (the ion binding constants in this case). It is to be hoped that an understanding of the apparent simplicity observed in experimental studies will emerge from the insights provided by free energy simulations.

Acknowledgements

References

Bash,P.A., Singh,U.C., Langridge,R. and Kollman,P.A. (1987) *Science*, **236**, 564.

Bashford,D. and Karplus,M. (1990) *Biochemistry*, **29**, 10219.

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.

Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M. (1983) *J. Comput. Chem.*, **4**, 187.

Brooks,C.L. and Karplus,M. (1989) *J. Mol. Biol.*, **208**, 159.

Brooks,C.L.,III, Karplus,M. and Pettitt,B.M. (1988) *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*. Adv. Chem. Phys. LXXI, John Wiley & Sons, New York.

Brünger,A.T. and Karplus,M. (1988) *Proteins*, **4**, 148.

Cossy,C., Barnes,A.C., Enderby,J.E. and Merbach,A.E. (1989) *J. Chem. Phys.*, **90**, 3254.

Evered,D. and Whelon,J. (1986) *Calcium and the Cell*. Ciba Foundation Symposium 122, John Wiley & Sons, New York.

Forsén,S. (1989) In Jardetsky,O. (ed.), *Protein Structure and Engineering*. Plenum Press, New York, pp. 291–308.

Gao,J., Kuzcera,K., Tidor,B. and Martin,K. (1989) *Science*, **244**, 1069.

Gariepy,J., Sykes,B.D. and Hodges,R.S. (1983) *Biochemistry*, **22**, 1765.

Godzik,A. and Sander,C. (1989) *Protein Engng*, **2**, 589.

Habenschuss,A. and Spedding,F.H. (1980) *J. Chem. Phys.*, **73**, 442.

Herzberg,O. and James,M.N.G. (1985) *Nature*, **313**, 653–659.

Hoard,J.L., Lee,B. and Lind,M.D. (1965) *J. Am. Chem. Soc.*, **87**, 1612.

Hoffmann,R., Buer,B.F., Muetterties,E.C. and Rossi,A.R. (1977) *Inorg. Chem.*, **16**, 511.

Jorgensen,W.L. (1981) *J. Am. Chem. Soc.*, **103**, 335.

Karplus,M. and Petsko,G.A. (1990) *Nature*, **347**, 631.

Kretsinger,R.H. and Nockolds,C.E. (1973) *J. Biol. Chem.*, **248**, 3313.

Kumat,V.D., Lee,L. and Edwards,B.F.P. (1990) *Biochemistry*, **29**, 1404.

Lind,M.D., Lee,B. and Hoard,J.L. (1965) *J. Am. Chem. Soc.*, **87**, 1611.

Loncharich,R.J. and Brooks,B.R. (1989) *Proteins*, **6**, 32.

MacKay,J.L., Powell,J.E. and Spedding,F.H. (1962) *J. Am. Chem. Soc.*, **84**, 2047.

Marsden,B.J., Hodges,R.S. and Sykes,B.D. (1988) *Biochemistry*, **27**, 4198.

Marsden,B.J., Hodges,R.S. and Sykes,B.D. (1989) *Biochemistry*, **28**, 8839.

Marsden,B.J., Shaw,G.S. and Sykes,B.D. (1990) *Biochem. Cell. Biol.*, **68**, 587.

Meier,W., Bopp,P., Probst,M.M., Spohr,E. and Lin,J.L. (1990) *J. Phys. Chem.*, **94**, 4672.

Mezei,M. and Beveridge,D.L. (1986) *Ann. N.Y. Acad. Sci.*, **482**, 1.

Papahadjopoulos,D., Nir,S. and Nejat,D. (1990) *J. Bioenerget. Biomembrane*, **22**, 157.

Pauling,L. (1948) *The Nature of the Chemical Bond*. Cornell University Press, Ithaca, NY.

Rasmussen,H. (1989) *Sci. Am.*, **261**, 66.

Reid,R.E. and Hodges,R.S. (1985) *J. Theoret. Biol.*, **84**, 5298.

Russell,A.J., Thomas,P.G. and Fersht,A.R. (1987) *J. Mol. Biol.*, **193**, 803.

Satyshur,K.A., Rao,S.T., Pyzalska,D., Drendel,W., Greaser,M. and Sundaralingam,M. (1988) *J. Biol. Chem.*, **263**, 1628.

Sinha,S.P. (1976) *Struct. Bond.*, **25**, 69.

Smith,L.S. and Wertz,D.L. (1975) *J. Am. Chem. Soc.*, **97**, 2365.

Smith,S.J. and Augustin,G.J. (1988) *Trends Neurosci.*, **11**, 458.

Stote,R. and Karplus,M. (1991) *J. Chim. Phys.*, **88**, 2419.

Straatma,T.P. and McCammon,J.A. (1992) *Annu. Rev. Phys. Chem.*, **43**, 407.

Strynadka,N.C.J. and James,M.N.G. (1989) *Annu. Rev. Biochem.*, **58**, 951.

Swain,A.L. and Amma,E.L. (1989) *Inorg. Chim. Acta*, **163**, 5.

Swain,A., Amma,S. and Kretsinger,R.H. (1989) *J. Biol. Chem.*, **264**, 16620.

Tidor,B. and Karplus,M. (1991) *Biochemistry*, **30**, 3217.

van Gunsteren,W.F. and Berendsen,H.J.C. (1977) *Mol. Phys.*, **34**, 1311.

Wier,G. (1990) *Annu. Rev. Physiol.*, **52**, 467.