Richardson, J. S., & Richardson, D. C. (1988a) *Proteins: Struct., Funct., Genet. 4*, 229–239.
Richardson, J. S., & Richardson, D. C. (1988b) in *Prediction of protein structure and the principles of protein conformation* (Fasman, Ed.) G. D., Plenum, New York.
Roder, H., Elöve, G. A., & Englander, W. (1988) *Nature 335*, 700–704.
Rose, M. D. (1987) *Methods Enzymol. 152*, 481–507.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Serrano, L., Bycroft, M., & Fersht, A. R. (1991) *J. Mol. Biol. 218*, 465–475.
Sherman, F., Taber, H., & Campbell, W. (1965) *J. Mol. Biol. 13*, 21–39.
Sherman, F., Stewart, J. W., Parker, J. H., Inhaber, E., Shipman, N. A., Putterman, G. J., Gardisky, R. L., & Margoliash, E. (1968) *J. Biol. Chem. 243*, 5446–5456.
Sherman, F., Stewart, J. W., Jackson, M., Gildord, R. A., & Parker, J. H. (1974) *Genetics 77*, 255–284.
Sherman, F., Jackson, M., Liebman, S. M., Schweingruber, A. M., & Stewart, J. W. A. (1975) *Genetics 81*, 51–73.
Shortle, D. (1989) *J. Biol. Chem. 264*, 5315–5318.
Shortle, D., & Lin, B. (1985) *Genetics 110*, 539–555.
Sikorski, R. S., & Hieter, P. (1989) *Genetics 122*, 19–27.
Takano, T., & Dickerson, R. E. (1981a) *J. Mol. Biol. 153*, 79–94.
Takano, T., & Dickerson, R. E. (1981b) *J. Mol. Biol. 153*, 94–115.
Vanfleteren, J. R., Evers, E. A. I. M., Van de Werken, G., & Van Beeumen, J. J. (1990) *Biochem. J. 271*, 613–620.
Verdiere, J., & Petrochilo, E. (1979) *Genetics 175*, 209–216.
Vieira, J., & Messing, J. (1987) *Methods Enzymol. 153*, 3–11.
Wells, J. A. (1990) *Biochemistry 29*, 8509–8517.
Zamyatnin, A. A. (1972) *Prog. Biophys. Mol. Biol. 24*, 107–123.
Zoller, M. J., & Smith, M. (1983) *Methods Enzymol. 100*, 468–500.

Quantitating and Engineering the Ion Specificity of an EF-Hand-like Ca^{2+} Binding Site†

Joseph J. Falke,* Eric E. Snyder, Kay C. Thatcher, and C. Stefan Voertler
Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215
Received March 28, 1991; Revised Manuscript Received June 13, 1991

ABSTRACT: The *Escherichia coli* D-galactose and D-glucose receptor, an aqueous periplasmic receptor that triggers sugar sensing and transport, possesses a single Ca^{2+} binding site similar in structure and specificity to the EF-hand class of sites found in eukaryotic Ca^{2+} signaling proteins including calmodulin and its homologues. A universal feature of these sites is the use of a pentagonal bipyramidal array of seven oxygens to coordinate bound Ca^{2+}. Here we investigate the mechanisms used by this coordinating array to control ion specificity. To vary the cavity size and charge of the array, we have replaced axial glutamine 142 in the prokaryotic site with asparagine, glutamate, and aspartate. The ion selectivities of the resulting engineered sites have been quantitated by measuring dissociation constants for a series of spherical metal ions, differing in increments of radius and charge, from groups Ia, IIa, and IIIa and the lanthanides. Dramatic specificity changes are observed: sites containing an engineered smaller side chain (Asn or Asp) bind the largest cations up to 50-fold more tightly than the native site; and sites containing an engineered negative side chain (Glu or Asp) exhibit preferences for trivalent over divalent cations up to 1900-fold higher than the native site. The results indicate that the cavity size and negative charge of the coordination array play key roles in selective Ca^{2+} binding and that the array can be engineered to preferentially bind other cations.

Protein Ca^{2+} binding sites selectively bind Ca^{2+} even in the presence of up to 10^5-fold higher concentrations of Na^+, K^+, and Mg^{2+}. This specificity is particularly important for regulatory sites in Ca^{2+} signaling proteins, which must remain unoccupied until a Ca^{2+} signal appears (Persechini et al., 1989), and for Ca^{2+} channel sites, which provide selective Ca^{2+} fluxes (Tsien et al., 1987; Alsobrook, 1988; Catteral et al., 1990). We are using protein engineering to probe the molecular basis of Ca^{2+} selectivity. Ca^{2+} sites are well suited for the protein engineering approach for two reasons. First, they exhibit substantial heterogeneity in primary structure and are thus likely to tolerate engineered sequence changes, and, second, the spherical ion is the simplest type of substrate. Such a substrate requires optimization of relatively few geometrical variables, and its electrostatic field is likely to impose optimal ordering of the coordinating side chains. In a favorable system such as this, it may be possible to apply rational design methods to generate predictable changes in substrate selectivity.

The D-galactose and D-glucose receptor of *Escherichia coli* possesses a single Ca^{2+} site that shares key structural elements with the important EF-hand class of Ca^{2+} sites, the latter commonly found in eukaryotic Ca^{2+} signaling proteins [see previous structural comparisons by Vyas et al. (1987) and Snyder et al. (1990)]. The structure of the receptor, determined to 1.9-Å resolution by Quiocho and co-workers (Vyas et al., 1987), reveals a coordinating array of seven oxygens arranged in a pentagonal bipyramid geometry around the bound Ca^{2+}, as typically observed in EF-hand sites (Strynadka & James, 1989). In the prokaryotic site the coordinating array, illustrated in Figure 1A, consists of four coordinating oxygens provided by three side-chain carboxylates (mono-

† Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, and to the National Institutes of Health for support of this work.

0006-2960/91/0430-8690$02.50/0

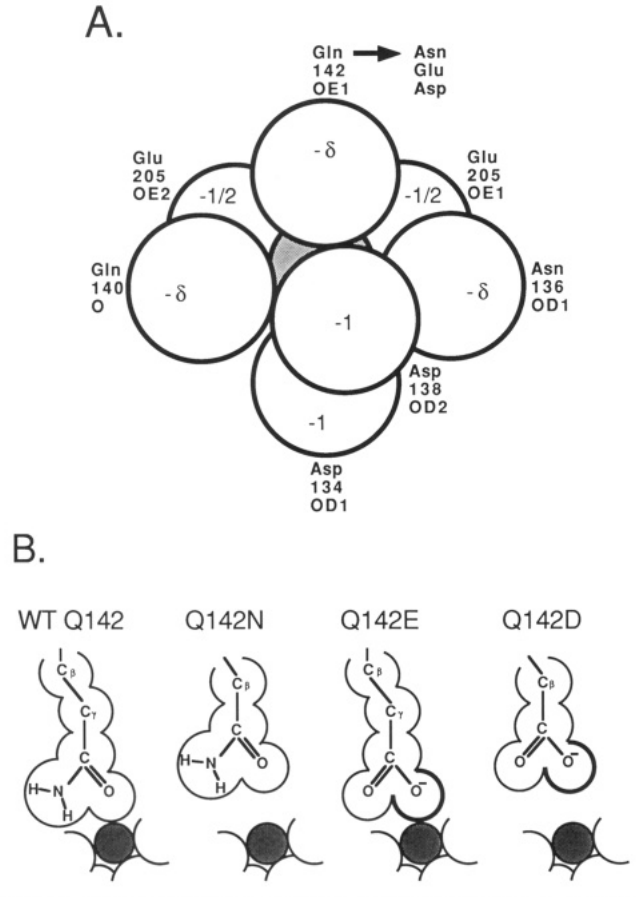

FIGURE 1: Coordination structure of the prokaryotic calcium site. (A) Schematic space-filling view of the seven coordinating oxygens (van der Waal's radius = 1.4 Å) arranged in a pentagonal bipyramid about the central Ca^{2+} ion (effective ionic radius = 1.06 Å). The approximate charge of each oxygen is indicated; the partial negative charge associated with each carbonyl dipole is indicated by δ. The coordinating oxygens are two from monodentate side-chain carboxylates (Asp 134 OD1, Asp 138 OD2), two from a bidentate side-chain carboxylate (Glu 205 OE1, OE2), two from side-chain amide carbonyls (Asn 136 OD1, Gln 142 OE1), and one from a backbone carbonyl (Gln 140 O). From the crystal structure by Vyas et al. (1987). (B) Schematic representation of the native and substituted side chains at position 142 illustrating the effects of substitution on side-chain size and charge.

dentate Asp 134 and 138; bidentate Glu 205), as well as two oxygens from side-chain carbonyls (Asn 136 and Gln 142) and one oxygen from a backbone carbonyl (Gln 140). The seven oxygens are tightly packed around the Ca^{2+} ion, such that each oxygen is in or near van der Waal's contact with the central ion, while the nearest-neighbor oxygens in the pentagon are within 0.1–0.2 Å of contact with each other and the axial oxygens lie within 0.3–0.8 Å of the pentagon oxygens.

In addition to site structure, electrostatic interactions play a key role in determining the ion specificity of EF-hand sites. Electrostatic repulsion between the coordinating oxygens is likely to be particularly important. In the prokaryotic site, the carboxylate oxygens provide a net negative charge up to −3 and the carbonyl oxygens provide additional partial negative charges associated with their carbonyl dipole moments. As in eukaryotic EF-hand sites, the net negative charge is altered by hydrogen bonding between the coordinating oxygens and nearby donor groups (Strynadka & James, 1989), enabling delocalization of an undetermined fraction of the negative charge away from the site. The remaining electrostatic repulsion within the array may be at least partially responsible for disrupting the structure of the empty site: indirect evidence suggests that the coordinating array opens and is exposed to solvent when bound cation is removed (Luck & Falke, 1991b), as in the empty sites I and II of troponin C and as suggested for other empty EF-hand sites (Herzberg & James, 1988; Satyshur et al., 1988; Drabikowski et al., 1982). In this picture the binding of a cation to the open, empty site serves to counteract the electrostatic repulsion between coordinating oxygens, enabling closure of the site and the formation of the coordinating array around the bound cation.

The prokaryotic site is particularly well suited for quantitative studies of ion specificity: it is an independent site and exhibits simple binding behavior (cooperativity and other multiple-site effects are absent; Vyas et al., 1989; Snyder et al., 1990). Quantitation is further facilitated by the availability of a range of similar metal ion substrates differing only by small radius increments or discrete charge increments (Snyder et al., 1990). In particular the cations of groups Ia, IIa, and IIIa and the lanthanides all possess spherically symmetrical filled s or p outer electronic subshells that are insensitive to coordination geometry; therefore, these ions behave like simple spheres of different size and charge. Previous studies of the native site indicated that both size and charge are important determinants of ion binding affinity, yielding a range of dissociation constants spanning 8 orders of magnitude (Snyder et al., 1990).

Two simple molecular ideas have been proposed to explain control of specificity by the native coordinating array: size specificity could be controlled by the size of the substrate cavity surrounded by the coordinating oxygens (Macmanus et al., 1989; Snyder et al., 1990), while charge specificity could be controlled by the negative charge density on the oxygens (Snyder et al., 1990). To test the importance of cavity size and negative charge density, we have varied the size and charge of the side chain at position 142, which is glutamine in the native array and provides a carbonyl oxygen for axial coordination (Figure 1B). This position corresponds to position nine in the coordinating loop of EF-hand sites and was chosen as least likely to cause loss of structural integrity when substituted because it exhibits the greatest sequence variability of the coordinating side-chain positions (Strynadka & James, 1989; Marsden et al., 1990). The present study replaces glutamine 142 with asparagine (Q142N), glutamate (Q142E), and aspartate (Q142D). The ion selectivities of these sites are quantitated by determining the free energy of binding as a function of substrate radius and charge. The results indicate that substitution of a smaller coordinating side chain at position 142 facilitates binding of larger cations, while substitution of an acidic side chain favors binding of more highly charged cations.

Materials and Methods

Construction of Engineered Ca^{2+} Sites. Engineered sites were generated as follows. Phagemid plasmid pSF5 was constructed by insertion of a *Pst*I–*Sac*I fragment containing the previously constructed receptor gene (Snyder et al., 1990) into the corresponding sites of the phagemid vector pTZ18U (Mead et al., 1986), yielding pSF5. Mutagenic oligonucleotides 19–21 bases in length were synthesized on an Applied Biosystems DNA synthesizer and purified by polyacrylamide gel electrophoresis. Uracil-labeled single-stranded template pSF5 DNA was obtained (Kunkel et al., 1987; Vieira & Messing, 1987) and used to generate engineered receptor genes via a standard protocol (Bio-Rad Laboratories). The mutagenesis reaction products were transformed into *E. coli* and minilysate DNA was isolated from transformants; then the resulting plasmids were screened for the desired mutation by dideoxy DNA sequencing (Tabor & Richardson, 1987;

Table I: Effect of Ca^{2+} Site Substitutions on the Sugar Binding Site and on Receptor ^{19}F NMR Resonances

		^{19}F NMR resonances of 5-fluorotryptophan-labeled receptor[b] (ppm)				
receptor	D-galactose K_D[a] (μM)	W127	W133	W183	W195	W284
WT	0.5 ± 0.1	46.0, 46.2	51.2, 51.8	42.8	48.3	52.5
Q142N	0.5 ± 0.1	45.7, 45.8	51.4, 52.0	42.7	48.3	52.5
Q142E	0.3 ■ 0.1	46.1, 46.2	51.0, 51.5	42.8	48.3	52.4
Q142D	0.6 ± 0.1	46.3, 46.5	51.3, 52.0	42.7	48.3	52.5

[a] Measured using an intrinsic tryptophan fluorescence assay described by Boos et al. (1972). Data were best fit by nonlinear least-squares to determine the K_D ± 1 standard error. [b] Standard deviations are ±0.1 ppm, measured for typical resonances by determining triplicate values with independent samples. The W127 and W133 5-fluorotryptophan resonances are each doublets due to the presence of two stable receptor structures on the NMR time scale. The W183, W195, and W284 positions sample regions outside the Ca^{2+} site. For details see Luck and Falke (1991a,b).

Sanger et al., 1977) using Sequenase DNA polymerase (U.S. Biochemical Corp.). Receptor protein was isolated from the periplasmic shock fluid, and substrate Ca^{2+} and sugar were removed as previously described (Snyder et al., 1990). The overall structures of the resulting engineered proteins were compared to the native receptor by measurement of their D-galactose dissociation constants and by 1D ^{19}F NMR of fluorine-labeled proteins. D-Galactose binding was quantitated by use of the previously described increase in intrinsic tryptophan fluoresence that occurs when sugar binds to its substrate cleft (Boos et al., 1972). ^{19}F NMR experiments were carried out as previously described and were used to monitor biosynthetically incorporated fluorines at the five tryptophan positions of the 5-fluorotryptophan-labeled receptor. The ^{19}F NMR frequency of each resonance provides a sensitive probe of local structural perturbations resulting from engineered changes in the Ca^{2+} site (Luck & Falke, 1991a,b).

Quantitation of Binding Equilibria for Metal Ion Substrates. Binding free energies were quantitated as previously described by incorporating the fluorescent lanthanide Tb^{3+} into the Ca^{2+} site, then competitively displacing the bound Tb^{3+} by stepwise addition of a second metal ion (Me) of unknown affinity (Snyder et al., 1990). The resulting plot of bound Tb^{3+} fluorescence against the concentration of competing metal ion was best fit to the independent site model

$$F/F_{max} = 1 - \{[Me]/(K_{D_{appMe}} + [Me])\} \quad (1)$$

where F/F_{max} is the bound Tb^{3+} fluorescence relative to the fluorescence in the absence of competing ion, [Me] is the concentration of free competing ion, and $K_{D_{appMe}}$ is the apparent dissociation constant for the competing ion. The best fit apparent dissociation constant provided by eq 1 was then corrected for competition with Tb^{3+} to yield the true dissociation constant given by

$$K_{D_{Me}} = K_{D_{appMe}}/(1 + [Tb^{3+}]/K_{D_{Tb}}) \quad (2)$$

where K_{DMe} is the dissociation constant for metal ion in the absence of Tb^{3+} ion, $[Tb^{3+}]$ is the concentration of free Tb^{3+} ion, and $K_{D_{Tb}}$ is the dissociation constant for Tb^{3+} ion. Finally the dissociation constant was converted to the binding free energy according to

$$\Delta G^0_{Me} = RT \ln K_{D_{Me}} \quad (3)$$

for the 1 M standard state concentration. All measurements were at 25 °C with 2.5 μM protein, 100 mM KCl, and 10 mM PIPES, pH 6.0.

The free ion concentrations used in eqs 1 and 2 were determined as previously described (Snyder et al., 1990). The Tb^{3+} dissociation constant $K_{D_{Tb}}$ used in eq 2 was determined as follows. First, the apparent dissociation constant $K_{D_{appTb}}$ was interpolated from the measured apparent dissociation constants for Gd^{3+} and Dy^{3+}, which are the next smaller and larger ions in the lanthanide series, respectively, by use of the linear average $K_{D_{appTb}} = (K_{D_{appGd}} + K_{D_{appDy}})/2$. Then, $K_{D_{Tb}}$ was calculated by substituting $K_{D_{Me}} = K_{D_{Tb}}$ into eq 2 and solving the resulting quadratic equation. The dissociation constant $K_{D_{Tb}}$ obtained by this interpolation method was found to be in good agreement with the value measured directly by fluorescence quantitation of Tb^{3+} binding. The interpolation method was chosen for routine use because it utilizes the same sample and experimental conditions used to determine the $K_{D_{app}}$ values for all competing metal ions.

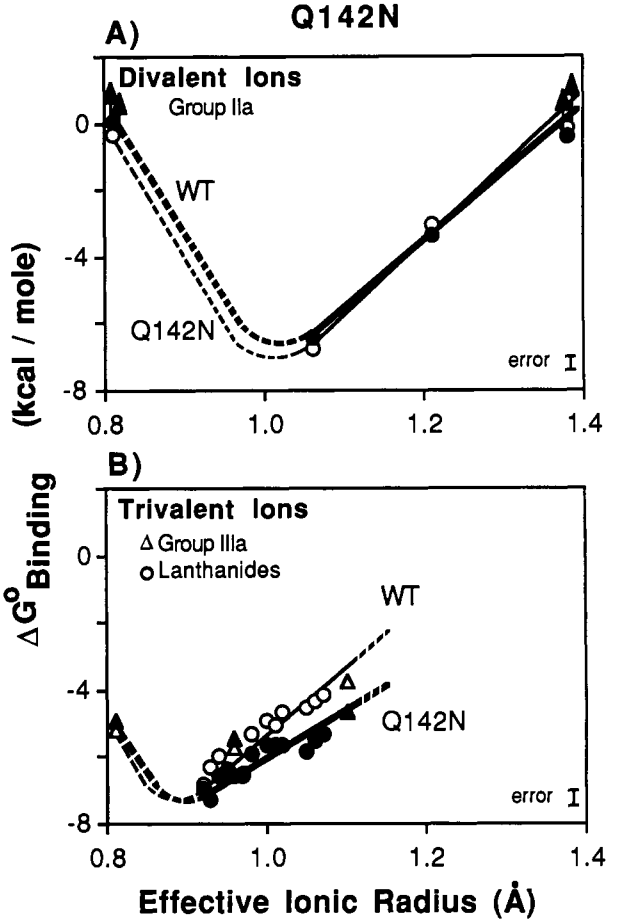

FIGURE 2: Dependence of binding free energy on cation charge and size: the Q142N engineered site. Plotted against effective ionic radius are binding free energies (symbols) or lower limit binding free energies (symbols with up arrows). Data are given for both the Q142N engineered site (filled symbols, bold curve) and the native site (open symbols, fine curve). An error bar (determined by RT ln [K_D ± 1 standard deviation]) is indicated for the largest error in each data set. Effective ionic radii are for 7-fold coordination, from Shannon (1976). (A) Divalent ions of group IIa: Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+} (in order of increasing radius). (B) Trivalent ions of group IIIa and the lanthanides: Sc^{3+}, Lu^{3+}, Yb^{3+}, Tm^{3+}, Er^{3+}, Ho^{3+}, Y^{3+}, Dy^{3+}, Tb^{3+}, Gd^{3+}, Eu^{3+}, Sm^{3+}, Nd^{3+}, Pr^{3+}, Ce^{3+}, La^{3+}. All measurements were at 25 °C and at 2.5 μM protein, 100 mM KCl, and 10 mM PIPES, pH 6.0.

In all calculations, molar concentrations, rather than activities, were used. This approximation is justified because

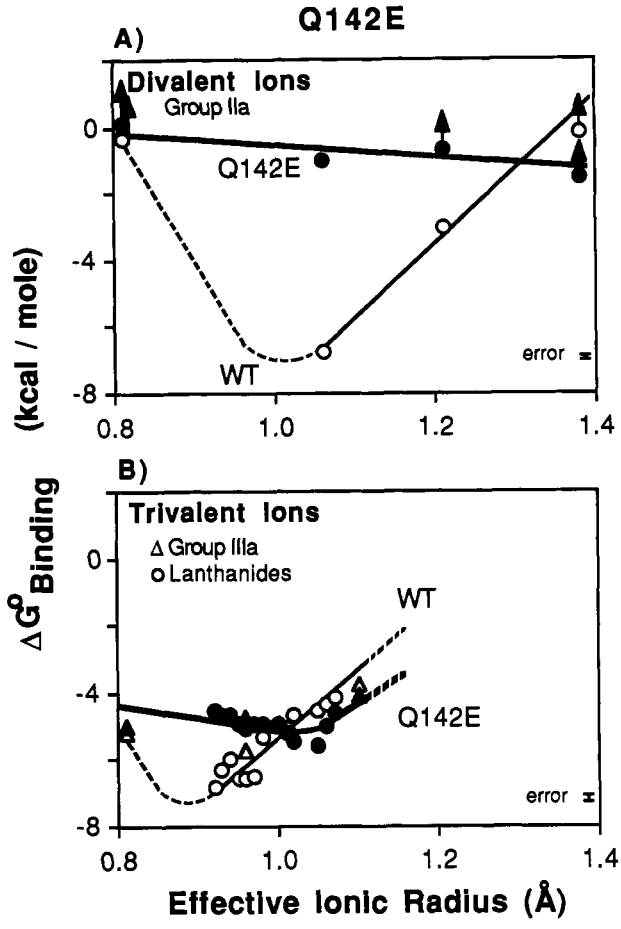

FIGURE 3: Dependence of binding free energy on cation charge and size: the Q142E engineered site. Legend as for parts A and B of Figure 2 except that data are for the Q142E engineered site (filled symbols, bold curve) and the native site (open symbols, fine curve).

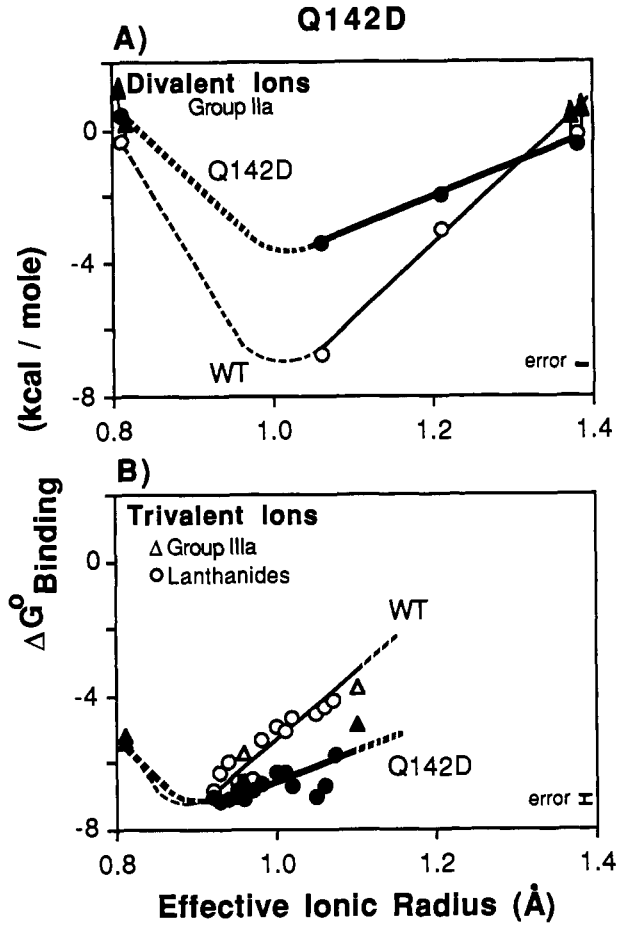

FIGURE 4: Dependence of binding free energy on cation charge and size: the Q142D engineered site. Legend as for parts A and B of Figure 2 except that data are for the Q142D engineered site (filled symbols, bold curve) and the native site (open symbols, fine curve).

the relevant experimental conditions yield activity coefficients the same order as unity, ranging between $0.5 \leq \gamma \leq 1.0$ for monovalent, divalent, and trivalent ions (Harned & Owen, 1950). Thus, activity corrections would alter the ΔG^0_{Me} values of Figures 2–4 by less than 0.2 kcal mol^{-1}.

Results

Effect of Side-Chain Replacement on Receptor Structure. Engineered sites containing asparagine, glutamate, or aspartate in place of glutamine 142 were constructed by oligonucleotide-directed mutagenesis of the receptor gene, and the resulting plasmids were used to overexpress receptor protein in *E. coli*. The apoprotein with empty Ca^{2+} and sugar sites was then isolated to a purity of ≥90% as judged by SDS–polyacrylamide gel electrophoresis. Engineered receptors exhibited ≤1.6-fold differences from the native receptor in their affinity for D-galactose, as measured by sugar binding titrations using an intrinsic tryptophan fluorescence assay (Table I). Similarly, ^{19}F NMR spectra of receptors labeled at five positions with 5-fluorotryptophan revealed no significant structural effects outside the Ca^{2+} site (Table I); only two tryptophans (Trp 127, 133) at the base of the Ca^{2+} binding loop yielded detectable but small frequency shifts (≤0.3 ppm, Table I). Thus, as designed, the engineered changes left the global structure of the receptor unaltered. Moreover, the following results reveal that each engineered Ca^{2+} site binds certain ions more tightly than the native site, indicating that key selectivity features in the site have been changed without disrupting the essential features of site architecture.

Quantitation of Binding Free Energies for Substrate Metal Ions. The ion binding equilibria of the engineered sites were quantitated by use of the previously described competition assay (Snyder et al., 1990) in which the fluorescent lanthanide Tb^{3+} is placed in the site, where it is selectively excited by resonance energy transfer from nearby tryptophan donors. The bulk of the energy transfer is from tryptophan 127, which lies 4.5 Å from the bound metal ion (Vyas et al., 1987), a distance similar to the characteristic distance for energy transfer from tryptophan to Tb^{3+} ($R_0 \sim 4$ Å; Horrocks & Collier, 1981). The resulting bound Tb^{3+} fluorescence emission was monitored while a given second metal ion was titrated into the system. If the second ion competitively displaced the Tb^{3+}, the bound Tb^{3+} emission decreased proportionally, yielding the competitive binding curve from which the apparent dissociation constant of the second ion was extracted by nonlinear least-squares best fit of the untransformed data. Finally, the apparent dissociation constant was corrected for Tb^{3+} competition to yield the true dissociation constant for the competing metal ion (K_D) and the corresponding binding free energy ($\Delta G^0 = RT \ln K_D$). Overall, this measurement scheme uses two half-reactions and the path independence of free energy to quantitate the desired equilibrium given by

$$P + Me\cdot(H_2O)_n \xrightleftharpoons{K_D,\ \Delta G^0} P\cdot Me + nH_2O \qquad (4)$$

which consists solely of contributions from (a) the ion-site interaction and (b) metal ion dehydration. As emphasized in this equation, the metal ion is fully dehydrated in the prokaryotic site [as indicated by the crystallographic studies of Vyas et al. (1987, 1989)].

The information provided by a pool of incrementally varying

Table II: Metal Ion Dissociation Constants of the Sites WT, Q142N, Q142E, and Q142D[a]

	effective ionic radius (Å)	K_D (M)			
		WT	Q142N	Q142E	Q142D
			Group Ia		
Li(I)	0.84	≥5.9	≥1.6	≥90	≥26
Na(I)	1.12	≥2.7	≥13	≥26	≥4
K(I)	1.46	≥3.9	≥1.8	≥4.3	≥13
Rb(I)	1.56	≥3.3	≥16	≥90	≥2
			Group IIa		
Mg(II)	0.81	$(570 \pm 120) \times 10^{-3}$	≥7	≥1.1	≥1.6
Ca(II)	1.06	$(13 \pm 2) \times 10^{-6}$	$(15 \pm 2) \times 10^{-6}$	0.18 ± 0.2	$(2.8 \pm 0.1) \times 10^{-3}$
Sr(II)	1.21	$(6.1 \pm 1.2) \times 10^{-3}$	$(2.5 \pm 0.6) \times 10^{-3}$	≥0.33	$(28 \pm 2) \times 10^{-3}$
Ba(II)	1.38	$\geq 740 \times 10^{-3}$	≥0.33	≥0.76	≥0.35
			Group IIIa		
Sc(III)	0.81	$(170 \pm 30) \times 10^{-6}$	$(170 \pm 40) \times 10^{-6}$	$(250 \pm 25) \times 10^{-6}$	$(150 \pm 10) \times 10^{-6}$
Y(III)	0.96	$(74 \pm 12) \times 10^{-6}$	$(75 \pm 7) \times 10^{-6}$	$(390 \pm 80) \times 10^{-6}$	$(22 \pm 3) \times 10^{-6}$
La(III)	1.10	$(1.9 \pm 0.2) \times 10^{-3}$	$(400 \pm 20) \times 10^{-6}$	$(1.1 \pm 0.5) \times 10^{-3}$	$(380 \pm 10) \times 10^{-6}$
			Lanthanides		
Ce(III)	1.07	$(1.0 \pm 0.1) \times 10^{-3}$	$(140 \pm 11) \times 10^{-6}$	$(490 \pm 70) \times 10^{-6}$	$(122 \pm 10) \times 10^{-6}$
Pr(III)	1.06	$(730 \pm 80) \times 10^{-6}$	$(100 \pm 8) \times 10^{-6}$	$(230 \pm 40) \times 10^{-6}$	$(54 \pm 8) \times 10^{-6}$
Nd(III)	1.05	$(510 \pm 50) \times 10^{-6}$	$(61 \pm 11) \times 10^{-6}$	$(94 \pm 10) \times 10^{-6}$	$(10 \pm 1) \times 10^{-6}$
Sm(III)	1.02	$(400 \pm 30) \times 10^{-6}$	$(77 \pm 5) \times 10^{-6}$	$(110 \pm 14) \times 10^{-6}$	$(18 \pm 3) \times 10^{-6}$
Eu(III)	1.01	$(230 \pm 20) \times 10^{-6}$	$(81 \pm 5) \times 10^{-6}$	$(170 \pm 50) \times 10^{-6}$	$(36 \pm 11) \times 10^{-6}$
Gd(III)	1.00	$(270 \pm 15) \times 10^{-6}$	$(85 \pm 7) \times 10^{-6}$	$(270 \pm 40) \times 10^{-6}$	$(34 \pm 8) \times 10^{-6}$
Tb(III)	0.98	$(150 \pm 50) \times 10^{-6}$	$(51 \pm 15) \times 10^{-6}$	$(280 \pm 90) \times 10^{-6}$	$(49 \pm 10) \times 10^{-6}$
Dy(III)	0.97	$(20 \pm 1) \times 10^{-6}$	$(17 \pm 2) \times 10^{-6}$	$(280 \pm 60) \times 10^{-6}$	$(14 \pm 4) \times 10^{-6}$
Ho(III)	0.96	$(18 \pm 1) \times 10^{-6}$	$(17 \pm 2) \times 10^{-6}$	$(220 \pm 50) \times 10^{-6}$	$(9 \pm 2) \times 10^{-6}$
Er(III)	0.945	$(18 \pm 1) \times 10^{-6}$	$(25 \pm 4) \times 10^{-6}$	$(280 \pm 50) \times 10^{-6}$	$(15 \pm 5) \times 10^{-6}$
Tm(III)	0.94	$(45 \pm 10) \times 10^{-6}$	$(18 \pm 4) \times 10^{-6}$	$(410 \pm 80) \times 10^{-6}$	$(9 \pm 3) \times 10^{-6}$
Yb(III)	0.925	$(27 \pm 5) \times 10^{-6}$	$(6 \pm 1) \times 10^{-6}$	$(420 \pm 90) \times 10^{-6}$	$(8 \pm 2) \times 10^{-6}$
Lu(III)	0.92	$(12 \pm 2) \times 10^{-6}$	$(9 \pm 1) \times 10^{-6}$	$(500 \pm 10) \times 10^{-6}$	$(10 \pm 3) \times 10^{-6}$

[a] Measurements were at 25 °C with 2.5 μM protein, 100 mM KCl, 10 mM PIPES, pH 6.0. See text for details.

substrates is illustrated by an analysis of the native site (Snyder et al., 1990; confirmed in Table II). The native site effectively excludes the monovalent ions of group Ia, including Na^+ and K^+ (Table II). In contrast, the site binds the divalent ions of group IIa and the trivalent ions of group IIIa and the lanthanides with affinities that vary with both ionic radius and charge. When the binding free energy is plotted against effective ionic radius (coordination number = 7; Shannon, 1976), a deep well is observed in the binding free energy vs radius relationship. Divalent and trivalent ions yield distinct free energy wells and thus are displayed separately (Figure 2A,B). These free energy wells contain all the information necessary to quantitate substrate specificity: the depth of each well provides the optimal substrate affinity, the position of the bottom indicates the optimal substrate radius, and the steepness of the well quantitates the size selectivity of the site. Regions of the free energy wells that are poorly determined, due to the unavailability of spherical metal ions with appropriate radii, are indicated by dashed lines. When the binding of two ions to the same site is compared, the difference binding free energy $\Delta\Delta G^0$ is useful. This $\Delta\Delta G^0$ consists of two components, one stemming from changes in the ion-site interaction, the other from differences in the metal ion dehydration free energy. In contrast, when the binding of the same ion to two nonidentical sites is compared, the difference binding free energy $\Delta\Delta G^{0'}$ stems solely from changes in the ion-site interaction: the metal ion dehydration term is eliminated because it is a constant, independent of variations in binding site structure.

For the engineered sites, dissociation constants and binding free energies are presented in Table II and Figures 2–4, respectively, together with the native site values for comparison. Like the native site, each of the engineered sites is observed to effectively reject the monovalent ions of group Ia (Table II). However, the engineered sites exhibit altered specificites for divalent and trivalent ions, revealed by changes in the free energy wells for each site (Figures 2–4). Since the best determined region of a given free energy well is provided by ions larger than the optimal radius, the following analysis focuses primarily on this large radius limit.

The Q142N Engineered Site. The Q142N substitution replaces glutamine 142 with asparagine (Figure 1B), yielding a decrease in side-chain volume (−31%) and length (−1.5 Å)(Creighton, 1983). This alteration has no detectable effect on the affinities of divalent ions for the site as illustrated by Figure 2A, although the data are limited due to the availability of only one cation, Ca^{2+}, with a radius near the optimal radius. In contrast, the engineered site exhibits higher affinities for trivalent ions in the large radius limit, as illustrated by Figure 2B. The difference is greatest for the five largest trivalent ions: Sm^{3+}, Nd^{3+}, Pr^{3+}, Ce^{3+}, and La^{3+} each bind ≥5-fold more tightly to the Q142N site than the native site, while the binding of small trivalent ions is unchanged (note that the indicated differences are large relative to experimental error; Table II and Figure 2). As a result, the steepness of the free energy well in the large ion limit exhibits a 30% decrease, where steepness is defined as the slope of a best-fit straight line through the free energies for trivalent ions larger than the optimal radius, demonstrating that the engineered site is less size-selective (Figure 2B). Thus, by decreasing the size of the coordinating side chain, a substrate cavity is generated that more easily adjusts to ions of increasing size, suggesting that expansion of the native cavity is constrained, in part, by an unidentified force requiring close packing of the surrounding residues. Subtracting the slopes of the native and engineered wells yields the magnitude of the constraining force, $F \sim 7.1$ kcal mol^{-1} $Å^{-1}$. This net force arises entirely from microscopic forces within the ion-occupied receptor that vary as the size of the bound ion increases, such as steric conflicts in the site, or decreases in favorable electrostatic interactions, or unfavorable changes in protein solvation. Metal ion dehydration

is not a factor because it is eliminated by the subtraction (see above discussion of $\Delta\Delta G^{0'}$).

It is instructive to compare the work required to expand the Ca^{2+} site with that required to distort a typical covalent bond or to expand a typical cavity in a protein. The constraining force deduced above yields 1.4 kcal mol^{-1} of work for a 0.2-Å increase in the Ca^{2+} site radius. More work is needed to distort a C–H bond by the same distance: a bond stretch of 0.2 Å requires 10 kcal mol^{-1}, while a bond compression of 0.2 Å uses 22 kcal mol^{-1} (calculated for a Morse potential; Levine, 1975). Less work is required to expand a typical protein cavity: in this case the work function is estimated to range up to 0.06 kcal mol^{-1} $Å^{-3}$ (Sandberg & Terwilliger, 1991), so that a 0.2-Å increase in cavity radius from 0.9 to 1.1 Å necessitates an input of 0.15 kcal mol^{-1}. Thus, although the Ca^{2+} site is 1 order of magnitude easier to distort than a covalent bond, it remains 1 order of magnitude more rigid than a typical protein cavity. This analysis indicates that the architecture of the site is specialized to generate an unusually large constraining force for a protein, yet this constraining force lies well within the range expected for simple steric interactions.

The Q142E Engineered Site. The Q142E substitution places glutamate at position 142, providing an additional negative charge to the coordination array as well as a minor decrease in side-chain volume (−7%) (Creighton, 1983). This substitution dramatically alters charge selectivity, such that Ca^{2+} and other divalent ions are effectively rejected while trivalent ions still bind, particularly those of radii larger than 1.05 Å that bind up to 5-fold more tightly than they bind to the native site (Table II and Figure 3A,B). The selectivity for trivalent ions over divalent ions, quantitated as the ratio of binding constants for the highest affinity trivalent and divalent substrates ($R = \{K_B(\text{best } M^{3+})/K_B(\text{best } M^{2+})\}$), is increased from $R = 1.1$ for the native site to $R = 1900$ for the engineered site (Table II).

These results provide evidence that charge specificity is controlled by the negative charge density of the coordinating array. In particular, the results implicate the significance of electrostatic repulsion between the coordinating oxygens, such that an increase in this repulsion requires a corresponding increase in the positive charge of substrate to stabilize the occupied site. Size selectivity is also altered by the glutamate substitution; in particular, for trivalent ions the free energy minimum is shifted ≥0.15 Å to a larger optimal radius (Figure 3B). This increase in optimal radius is larger than expected for the small reduction in side-chain volume (compare with Figure 2B), suggesting that the increased electrostatic repulsion within the coordinating array drives an expansion of the substrate cavity.

The Q142D Engineered Site. The Q142D substitution combines the essential features of the Q142N and Q142E substitutions by replacing glutamine 142 with an aspartate, thereby providing significant decreases in both side-chain volume (−39%) and length (−1.5 Å) (Creighton, 1983) as well as an additional negative charge. This substitution produces a substantial increase in selectivity for trivalent ions over divalent ions, yielding $R = 350$ (Table II). Contributing to this increased charge specificity is the tighter binding observed for trivalent ions, up to 50-fold tighter than by the native site, as well as weaker binding of divalent ions. Also observed is a dramatic loss of size selectivity as revealed by the decreased steepness of the free energy well, particularly for the trivalent ions where the steepness is decreased by 50% (Figure 4B).

Qualitative comparison of the Q142D site with the other two engineered sites suggests that side-chain size and charge make roughly additive contributions to ion selectivity, suggesting that the size and charge contributions are at least partially independent. In particular, the Q142D site (Figure 4A,B) qualitatively combines the attributes observed for the smaller side chain (Q142N, increased affinity for large trivalent ions; Figure 2B) and for the additional negative charge (Q142E, decreased affinity for divalent ions and increased affinity for trivalent ions; Figure 3A,B), yielding a site that prefers ions that are larger and trivalent. Further structural characterization of the engineered sites is needed for a more quantitative analysis of how the size and charge and other characteristics of the side chain at position 142 contribute to substrate selectivity.

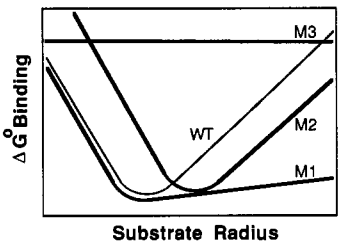

FIGURE 5: Schematic illustration of three fundamental types of substrate selectivity changes. The native site selectively binds substrate with the optimal radius corresponding to the bottom of the free energy well labeled WT. The engineered site M1 is more expandable than the native site and yields a lower barrier to increases in substrate radius, yielding a well that is less steep in the large substrate limit. The engineered site M2 has a larger equilibrium cavity size and thus shifts the well to a larger substrate radius. The engineered site M3 no longer binds the test substrates, thereby yielding an unfavorable free energy of binding over the entire range of radii.

Discussion

When a substrate binding site is engineered to alter its selectivity, a pool of incrementally varying substrates can quantitate the selectivity change and provide considerable insight into the molecular effects of altered primary structure. Figure 5 schematically illustrates the dependence of binding free energy on substrate radius for a native site and three fundamental classes of engineered sites. The native site exhibits an optimal radius for binding and a characteristic free energy well (Figure 5, WT) that quantitatively describes the selectivity of the starting site. Each engineered site yields a different free energy well; in the first example the steepness of the well is decreased in the large ion limit, indicating that the engineered site is more expandable or less rigid than the native site (Figure 5, WT-M1). In the second example, the position of the well shifts to a larger optimal radius, indicating that the equilibrium cavity size of the site has increased (Figure 5, WT-M2). And in the third case the well is completely eliminated, indicating that the site now excludes the entire range of substrates (Figure 5, WT-M3). Each of these selectivity changes has been observed in the three engineered Ca^{2+} sites at hand.

The Q142 and Q142D sites introduce a smaller coordinating side chain at the 142 position, yielding free energy wells that are less steep in the large ion limit, due to an increased affinity for large ions and a decrease in size selectivity (compare example M1 in Figure 5, Figure 2B, and Figure 4B). In both modified sites the constraining force that opposes expansion of the substrate cavity is reduced such that the smaller side chain gives the site more conformational freedom, rather than simply increasing the cavity size (compare Figure 5, examples M1 and M2). The constraining force estimated from the data falls well within the range expected for simple steric constraints such as bond stretching or compression (see Results), sug-

gesting that steric constraints contribute to the steepness of the large ion side of the well and that these constraints can be reduced by substitution with a less bulky side chain.

When an acidic coordinating side chain is introduced at the 142 position, as in the Q142E and Q142D sites, the increase in negative charge density dramatically alters the charge selectivity. In particular, the free energy well for divalent ions is either eliminated completely (Figure 5, example M3; compare with Q142E in Figure 3A) or becomes significantly shallower (Q142D in Figure 4A), indicating a decrease in affinity for divalent ions. Simultaneously observed is an increase in affinity for trivalent ions (Q142E in Figure 3B; Q142D in Figure 4B). These results fully support a simple molecular picture for the conformational events that occur upon ion binding: initially the empty site has an open conformation, in which the distance between coordinating oxygens is relative large (Herzberg & James, 1988; Satyshur et al., 1988; Luck & Falke, 1991a). The binding of a highly charged cation counteracts the electrostatic repulsion between these oxygens, enabling them to form the coordinating array around the bound ion, thereby closing the occupied site. As the negative charge of the coordinating array increases, a more highly charged substrate ion is needed to stabilize the closure event, explaining the observed correlation between cavity charge and substrate charge. The Q142E substitution also yields a shift of the free energy well to larger radius, suggesting that increased electrostatic repulsion between coordinating oxygens is more effective than reduced side-chain bulk in expanding the optimal cavity radius while retaining size selectivity (Figure 5, example M2; compare Q142E and Q142N in Figures 2B and 4B).

Further structural and analytical studies of the engineered sites will provide additional important information regarding the molecular mechanisms of ion selectivity. For example, two explanations can be offered for the decreased size selectivity when a smaller asparagine or aspartate side chain is substituted for the axial glutamine 142. One explanation is that the less bulky side chain creates a cavity, thereby increasing the conformational freedom of adjacent residues. Alternatively, a water molecule, which has additional degrees of freedom because it is not covalently attached to a side chain, could provide axial coordination by bridging the short side chain to the metal ion, as observed in EF-hand sites possessing a short side chain at this position (Strynadka & James, 1989). Another undetermined parameter is the effective negative charge of the coordinating oxygen array, which may be decreased if the carboxylates are not fully deprotonated or if charge is withdrawn by hydrogen bonds to the coordinating oxygens. Such factors can explain how different EF-hand sites possessing three and four coordinating carboxylates, respectively, are observed to bind Ca^{2+} with similar affinities (Marsden et al., 1990; Sekharadu & Sundaralinga, 1988).

The present results establish that the ion specificity of the prokaryotic EF-hand-like site is controlled by the interaction between the metal ion and the site rather than by the energetic contribution of metal ion dehydration. Two key findings support this view: (a) the binding of large ions to the site weakens as their ionic radius increases, even though the larger radius makes it easier to dehydrate the ion (Rashin & Honig, 1985); and (b) a change in the size or charge of the coordinating side chain at position 142 yields predictable changes in ion size and charge selectivity. The key selectivity control elements appear to be the cavity size and charge of the coordinating oxygen array. This picture explains the failure of Na^+, K^+, and Mg^{2+} to bind to the empty site as follows. Na^+ and K^+ are excluded because their single positive charge is insufficient to overcome the electrostatic repulsion between the coordinating array of oxygens that occurs as the oxygens pack tightly around the substrate ion. Mg^{2+} is excluded because the cavity size of the pentagonal bipyramid array (2.0 Å for close packing of oxygens) is substantially larger than the effective ionic diameter of Mg^{2+} (1.6 Å; Shannon, 1976), precluding close contact between the ion and the coordinating array. For Mg^{2+} ion the octahedral hydration shell of the free ion provides a smaller cavity and better metal–oxygen contact, thus the site cannot overcome the energetic cost of removing the ion from water, as evidenced by computational results (Sussman & Weinstein, 1989). The protein engineering approach, in conjunction with physical studies, will provide further insights into the molecular mechanisms underlying selective Ca^{2+} binding and will continue to generate new sites with altered ion specificities.

Acknowledgments

We thank Drs. Sunney Chan, William DeGrado, Jeffrey Gelles, and Jamie Williamson for helpful discussions; Dr. Linda Luck and Claire Careaga for technical advice; and Katherine Swaggert for the construction of plasmid pSF5.

Registry No. Ca^{2+}, 7440-70-2; Mg^{2+}, 7439-95-4; Sr^{2+}, 7440-24-6; Ba^{2+}, 7440-39-3; Sc^{3+}, 7440-20-2; Lu^{3+}, 7439-94-3; Yb^{3+}, 7440-64-4; Tm^{3+}, 7440-30-4; Er^{3+}, 7440-52-0; Ho^{3+}, 7440-60-0; Y^{3+}, 7440-65-5; Dy^{3+}, 7429-91-6; Tb^{3+}, 7440-27-9; Gd^{3+}, 7440-54-2; Eu^{3+}, 7440-53-1; Sm^{3+}, 7440-19-9; Nd^{3+}, 7440-00-8; Pr^{3+}, 7440-10-0; Ce^{3+}, 7440-45-1; La^{3+}, 7439-91-0; Gln, 56-85-9; Asn, 70-47-3; Glu, 56-86-0; Asp, 56-84-8; D-galactose, 59-23-4; D-glucose, 50-99-7.

References

Alsobrook, J. P., & Stevens, C. V. (1988) *Trends Neurosci. 11*, 1–2.

Babu, Y. S., Bugg, C. E., & Cook, W. J. (1988) *J. Mol. Biol. 204*, 191–204.

Boos, W., Gordon, A. S., Hall, R. E., & Price, D. H. (1972) *J. Biol. Chem. 247*, 917–924.

Catterall, W. A., Nunoki, I., Lai, Y., Dejongh, K., Thomsen, W., & Rossie S. (1990) *Adv. Second Messenger Phosphoprotein Res. 24*, 30–35.

Creighton, T. E. (1983) *Proteins*, W. H. Freeman & Co., New York.

Drabikowski, W., Brzeska, H., & Venyaminov, S. Y. (1982) *J. Biol. Chem. 257*, 11584–11593.

Harned, H. S., & Owen, B. B. (1950) *Physical Chemistry of Electrolytic Solutions*, Reinhold Publishing Corp., New York.

Herzberg, O., & James, M. N. G. (1988) *J. Mol. Biol. 203*, 761–779.

Horrocks, W. DeW., & Collier W. E. (1981) *J. Am. Chem. Soc. 103*, 2856–2862.

Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.

Luck, L. A., & Falke, J. J. (1991a) *Biochemistry 30*, 4248–4256.

Luck, L. A., & Falke, J. J. (1991b) *Biochemistry 30*, 4257–4261.

Macmanus, J. P., Hutnik, C. M. L., Sykes, B. D., Szabo, A. G., Williams, T. C., & Banville, D. (1989) *J. Biol. Chem. 264*, 3470–3477.

Marsden, B. J., Shaw, G. S., & Sykes, B. D. (1990) *Biochem. Cell Biol. 68*, 587–601.

Mead, D. A., Szczesna-Shorupa, E., & Kemper, B. (1986) *Protein Eng. 1*, 67–74.

Persechini, A., Moncreif, N. D., & Kretsinger, R. H. (1989)

Trends Neurosci. 12, 462-467.
Rashin, A. A, & Honig, B. (1985) *J. Phys. Chem. 89*, 5588-5593.
Sandberg, W. S., & Terwilliger, T. C. (1991) *Trends Biotechnol. 9*, 59-63.
Sanger, F., Miklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Satyshur, K. A., Sambhorao, T. R., Pyzalska, D., Drendall W., Greaser, M., & Sundaralingam, M. (1988) *J. Biol. Chem. 263*, 1628-1647.
Sekharadu, Y. C., & Sundaralingam, M. (1988) *Protein Eng. 2*, 139-146.
Shannon, R. D. (1976) *Acta Crystallogr. A32*, 751-767.
Snyder, E. E., Buosci, B. W., & Falke, J. J. (1990) *Biochemistry 29*, 3937-3943.
Strynadka, N. C., & James, M. N. G. (1989) *Annu. Rev. Biochem. 58*, 951-998.
Sussman, F., & Weinstein, H. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7880-7884.
Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767-4771.
Tsien, R. W., Hess, P., McCleskey, E. W., & Rosenberg, R. L. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 265-290.
Vieira, J., & Messing, J. (1987) *Methods Enzymol. 153*, 5463-5467.
Vyas, N. K., Vyas, M. N., & Quiocho F. A. (1987) *Nature 327*, 635-638.
Vyas, M. N., Jacobson, B. L., & Quiocho, F. A. (1989) *J. Biol. Chem. 264*, 20817-20821.

Determination of the Three-Dimensional Solution Structure of Barnase Using Nuclear Magnetic Resonance Spectroscopy

Mark Bycroft,*,‡ Svend Ludvigsen,§ Alan R. Fersht,‡ and Flemming M. Poulsen*,§

Cambridge Centre for Protein Engineering, University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW U.K., and Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK 2500 Valby, Copenhagen, Denmark

Received March 22, 1991; Revised Manuscript Received June 14, 1991

ABSTRACT: The solution conformation of the ribonuclease barnase has been determined by using ^{1}H nuclear magnetic resonance (NMR) spectroscopy. The 20 structures were calculated by using 853 interproton distance restraints obtained from analyses of two-dimensional nuclear Overhauser spectra, 72 ϕ and 53 χ_1 torsion angle restraints, and 17 hydrogen-bond distance restraints. The calculated structures contain two α-helices (residues 6-18 and 26-34) and a five-stranded antiparallel β-sheet (residues 50-55, 70-75, 85-91, 94-101, and 105-108). The core of the protein is formed by the packing of one of the α-helices (residues 6-18) onto the β-sheet. The average RMS deviation between the calculated structures and the mean structure is 1.11 Å for the backbone atoms and 1.75 Å for all atoms. The protein is least well-defined in the N-terminal region and in three large loops. When these regions are excluded, the average RMS deviation between the calculated structures and the mean structure for residues 5-34, 50-56, 71-76, 85-109 is 0.62 Å for the backbone atoms and 1.0 Å for all atoms. The NMR-derived structure has been compared with the crystal structure of barnase [Mauguen et al. (1982) *Nature* (*London*) *297*, 162-164].

Bacillus amyloliquefaciens produces a small extracellular ribonuclease (M_r 12382) known as barnase (Hartley, 1989). Barnase is not homologous with pancreatic ribonuclease but does have some sequence and structural homology to a number of other microbial purine-specific ribonucleases (Hill et al., 1983). The gene coding for barnase has been cloned and the protein has been expressed in *Escherichia coli* (Padden & Hartley, 1987). Barnase has proved a good model system for studying the factors affecting protein stability and the pathway of protein folding (Kellis et al., 1988, 1989; Sali et al., 1988; Serrano & Fersht, 1989; Matouschek et al., 1989, 1990; Bycroft et al., 1990a). To complement these studies, we have embarked on an investigation of barnase using nuclear magnetic resonance spectroscopy.

The development of 2D NMR methods (Ernst et al., 1987) has enabled essentially complete assignments to be made of the ^{1}H NMR spectra of small proteins (Wüthrich, 1986). Once assignments are available, it is possible to analyze two-dimensional NOE spectra to obtain a set of interproton distances that can, in conjunction with torsion angle data obtained from J coupling constants, be used to calculate a solution conformation of the protein (Wüthrich, 1989; Clore & Gronenborn, 1989).

We have recently reported assignments of the majority of the ^{1}H NMR spectrum of barnase (Bycroft et al., 1990b). In this paper, we describe the calculation of the solution conformation of the protein on the basis of NMR data.

Experimental Procedures

Recombinant barnase was purified from a culture of *E. coli* containing the plasmid pMT410 (Paddon & Hartley, 1987) as described previously (Mossakowska et al., 1989). Samples contained 5 mM protein, pH 4.5, in either 90% H_2O/10% D_2O or 99.96% D_2O. Spectra were recorded at 37 °C on a Bruker AM 500 spectrometer. Phase-sensitive spectra were obtained by using both the time-proportional phase incrementation method (Marion & Wüthrich, 1983) and the hypercomplex method (States et al., 1982). NOESY (Jeener et al., 1979) and DQF-COSY (Piantini et al., 1982; Rance et al., 1983) spectra were recorded with 512 t_1 increments each with 2048 complex data points. Mixing times of 50, 100, and 150 ms were used in the NOESY experiments. The experimental data were zero filled to give a 8K × 2K data matrix. The data were

‡ Cambridge Centre for Protein Engineering.
§ Carlsberg Laboratory.

0006-2960/91/0430-8697$02.50/0