Monovalent cation binding to cubic insulin crystals

Olga Gursky, Youli Li, John Badger, and Donald L. D. Caspar
Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254-9110 USA

ABSTRACT Two localized monovalent cation binding sites have been identified in cubic insulin from 2.8 Å-resolution difference electron density maps comparing crystals in which the Na^+ ions have been replaced by Tl^+. One cation is buried in a closed cavity between insulin dimers and is stabilized by interaction with protein carbonyl dipoles in two juxtaposed alternate positions related by the crystal dyad. The second cation binding site, which also involves ligation with carbonyl dipoles, is competitively occupied by one position of two alternate His B10 side chain conformations. The cation occupancy in both sites depends on the net charge on the protein which was varied by equilibrating crystals in the pH range 7–10. Detailed structures of the cation binding sites were inferred from the refined 2-Å resolution map of the sodium-insulin crystal at pH 9. At pH 9, the localized monovalent cations account for less than one of the three to four positive counterion charges necessary to neutralize the negative charge on each protein molecule. The majority of the monovalent counterions are too mobile to show up in the electron density maps calculated using data only at resolution higher than 10 Å. Monovalent cations of ionic radius < 1.5 Å are required for crystal stability. Replacing Na^+ with Cs^+, Mg^{++}, Ca^{++} or La^{+++} disrupts the lattice order, but crystals at pH 9 with 0.1 M Li^+, K^+, NH_4^+, Rb^+ or Tl^+ diffract to at least 2.8 Å resolution.

INTRODUCTION

Insulin, free of zinc, crystallizes from alkaline solutions of sodium salts at near physiological ionic strength as a symmetric dimer in the cubic space group $I2_13$ with unit cell edge $a = 78.9$ Å (Dodson et al., 1978). The structure of pig insulin in cubic crystals grown from 0.2 M Na_2HPO_4 at pH 9 has been solved at 1.7-Å resolution (Badger et al., 1991). Furthermore, the average electron density distribution of all the solvent in these crystals has been determined applying a method for refining maps of disordered parts of well-ordered crystal structures (Badger and Caspar, 1991). Fluctuations in the measured density distribution throughout the solvent space indicate nonrandom arrangements of the water molecules extending out to distances at least 15 Å from the protein surface. This long-range ordering of the solvent is presumably a consequence of electrostatic interactions among the water dipoles and the charge distribution in the protein and solvent.

The number and distribution of the counter ions in the crystal solvent depend on the number and arrangement of the charged groups in the protein. There are 16 ionizable groups in the 51 amino acid insulin molecule: four glutamic acids, two carboxyl termini, four tyrosines, one arginine, one lysine, two histidines, and two amino termini. At pH 8–9, all six carboxyl groups should be negatively charged, the four tyrosines and two histidines neutral, the arginine and lysine positively charged, and the two amino termini partially or largely unprotonated, depending on their pK in the crystal and the exact pH. Thus, under the conditions in which the crystals are grown, the net charge on an insulin molecule should be between -3 and -4, which should be neutralized by 3–4 sodium counterions. In equilibrium with 0.2 M salt, the solvent in the crystal should also contain about one molecule of the salt per insulin molecule.

To investigate the effect of cation size, charge and polarizability on crystal stability, the crystals were equilibrated with various concentrations of Li^+, Na^+, K^+, Rb^+, Cs^+, Tl^+, NH_4^+, Mg^{++}, Ca^{++}, or La^{+++}. X-ray diffraction data from crystals treated with the monovalent cation solutions that did not destroy or disorder their structure were collected to 2.8 Å resolution using the Brandeis electronic area detector (Kalata, 1985). The experiments were carried out with bovine insulin, which differs from pig insulin by two conservative amino acid replacements: A8 Thr → Ala and A10 Ile → Val. Measurement of the small differences between the electron density maps of the two insulins provided a calibration for assessing differences among crystals with the various cations. Monovalent cation binding sites could be located from difference electron density maps computed from diffraction data comparing crystals containing heavy and light cations.

Although Na^+ is abundant in physiological environments, Na-protein interactions have not been directly documented by x-ray protein crystallography because of the near equivalence in scattering from a Na^+ ion and a water molecule. By using the measured positions of the

0006-3495/92/03/604/08 $2.00

binding of the heavier monovalent cations Tl^+ and Rb^+ as a guide, Na^+ binding sites could be identified in the high resolution electron density maps obtained from Na-insulin.

MATERIALS AND METHODS

Crystallization

Cubic crystals of bovine insulin were grown at room temperature using a modified version of the protocol reported by Dodson et al. (1978) for porcine insulin. Dialysis buttons containing a filtered solution made up from 15 mg/ml bovine insulin, 0.018 M Na_2HPO_4, 0.01 M Na_3 EDTA at pH 10.5 were placed in a solution containing 0.15 M Na_2HPO_4, 0.01 M Na_3 EDTA, 0.5% w/v xylene at pH 9.2. After 24 h, a small quantity of 1 M Na_2HPO_4 was added to the liquid outside the dialysis buttons to raise the total Na cation concentration to 0.5 M. Crystals with cubic or rhombic dodecahedral morphology up to 1 mm in size were obtained over the next few days.

Crystals were also obtained by continued dialysis at various pH conditions from 8.9–9.5 and Na_2HPO_4 concentrations from 0.15–0.3 M, without abruptly increasing the Na_2HPO_4 concentration in a later step. However, in these cases (or when the Na_2HPO_4 concentration was increased after only a few hours) the usual result was the production of small crystals less than 0.5 mm wide.

Equilibration against salt solutions

Dialysis buttons containing insulin crystals were transferred to solutions of Li^+, Na^+, K^+, Rb^+, Cs^+, NH_4^+, Tl^+, Mg^{++}, Ca^{++}, or La^{+++} acetates, chlorides or carbonates in Tris/HCl or Na carbonate/bicarbonate buffer at ~pH 9. A number of different trials, with cation concentrations between 1 mM and 1 M and buffer concentrations between 3 mM and 0.1 M, were carried out. Additional soaking experiments were made with Tl solutions at pH 9.5 and pH 7. The conditions surveyed are summarized in Table 1. The insulin concentration in the liquid surrounding grown crystals, and the liquid surrounding crystals equilibrated against various salt solutions was measured by UV absorption at $\lambda = 276$ nm.

X-ray data collection and processing

X-ray data were routinely collected to 2.8-Å resolution from crystals equilibrated with various monovalent cations using the Brandeis TV Area Detector system (Kalata, 1985; Kalata et al., 1990) with an Elliott rotating anode generator operated at 35 kV, 40 mA. For some of the experiments, crystal slippage during data collection was prevented by coating the crystal in the capillary with a Formvar film (Rayment et al., 1977). Nearly complete data sets were recorded from single crystals (typically over a rotation range of 60°) over time-spans of 20 h per data set with 0.1° frames and 120 s exposure time per frame. The oscillation frames were processed using a version of the MADNES program (Pflugrath and Messerschmidt, 1987) modified for the Brandeis detector system. The integrated peak intensities were estimated by three-dimensional profile fitting and subsequently corrected for radiation damage and crystal absorption (Kabsch, 1988).

We have also collected data extending to 2.0-Å resolution from Na-insulin crystals at pH 7, 9, and 10.2, and the refinement of the resulting atomic models is close to completion. Details of the pH-dependent structural changes will be presented elsewhere (manuscript in preparation), but the high resolution results relevant to the structure of the monovalent cation binding sites are described here.

RESULTS

Crystal stability

The appearance of crystals dialyzed against 0.1 M of the monovalent cations Li^+, Na^+, K^+, Rb^+, NH_4^+ or Tl^+ at pH 9 in 10–50 mM Tris/HCl buffer generally remained unchanged over periods of several days whereas crystals

TABLE 1 Summary of biophysical data and conditions surveyed for stability of cubic insulin crystals in various cation solutions

Ion	Number of electrons	Ionic radius* (Å)	Cation concentration	pH	Soaking time	Crystal appearance	Diffraction resolution ($Å^{-1}$)
Li^+	2	0.68	0.1–0.5 M	7.0–9.5	1–7 d	unchanged	>1/2.5
Na^+	10	0.97					
K^+	18	1.33		<7.0	1–3 d	turn brown	<1/4
NH_4^+	10	1.47					
Rb^+	36	1.47	0.1–0.3 M	9.0	1 d	unchanged	>1/2.8
				7.0	1 d	turn brown	~1/3.6
Tl^+	80	1.47		9.0	3–7 d	unchanged	<1/10
			>0.3 M	9.0	1 d	turn brown	<1/4
Cs^+	54	1.67	0.05–0.3 M	9.0	2 h	crack	<1/10
Mg^{++}	11	0.82	0.01–0.3 M	9.0	1 d	turn brown	<1/10
Ca^{++}	19	1.18					
La^{+++}	56	1.02	0.001–0.1 M	9.0	2 h	crack, turn brown	<1/10

Three categories are distinguished by the horizontal lines: (*a*) monovalent cations that maintain stability under conditions similar to the Na^+ used to grow the crystals, (*b*) heavier monovalent cations that can replace Na^+ under more restricted conditions and (*c*) cations that destroy or disorder the cubic insulin structure. *Values for the ionic radii are taken from *CRC Handbook of Chemistry and Physics,* 71st edition.

dialyzed with Cs^+ cracked after a few hours. When soaked in <0.05 M concentrations of these cations, the crystals dissolved within a few days, and in the absence of cations the crystals dissolved within a few hours. When 10 mM Na carbonate/bicarbonate was used as buffer, crystals also dissolved at 0.05 M cation concentrations. The crystals soaked in solutions of the di- and trivalent cations Mg^{++}, Ca^{++} and La^{+++} gradually became opaque and La-soaked crystals sometimes cracked. After a period of several weeks, crystals dialyzed against 0.3 M Tl^+ at pH 9 also became opaque, and at pH 7 crystals dialyzed against 0.1 M Tl^+ discolored in about a day. The observations on the effects of the ionic conditions surveyed on the cubic insulin crystals are summarized in Table 1.

The protein concentration was 2.5 mM in the initial crystallizing solution and was measured to be 0.61 mM in the mother liquor surrounding crystals that had stopped growing in 0.25 M Na_2HPO_4. The protein concentration surrounding crystals dialyzed for 24 h against solutions of Li^+, Na^+, K^+, Rb^+, Cs^+ or Tl^+ at 0.2 M ionic strength with 10 mM or 40 mM Tris/HCl buffer at pH 9 was 0.64 mM (±5%) and was 1.8 mM (±5%) at 0.1 M ionic strength. Further measurements, made after longer soaking times, showed no significant difference. Protein solubility measurements were also made on the mother liquor surrounding somewhat smaller crystals obtained from a crystallization protocol in which the Na^+ concentration in the dialyzing solution was maintained at 0.3 M rather than raised to 0.5 M. In this case the protein concentration in the solution surrounding the grown crystals was 1.1 mM.

X-ray diffraction data

The space group ($I2_13$) and cell dimension (78.9 Å) of the bovine insulin crystallized from 0.25 M Na_2HPO_4 were the same as for porcine insulin crystals grown under similar conditions (Dodson et al., 1978). The cubic bovine insulin crystals dialyzed against 0.1 M solutions of Li^+, Na^+, K^+, Rb^+, NH_4^+ and Tl^+ cations in 10–50 mM Tris/HCl buffer for 24 h at ~pH 9 diffracted to at least 2.8-Å resolution and showed no change in lattice constant. Diffraction data were also collected from a crystal dialyzed in 0.3 M Tl^+ at pH 9.5, a crystal dialyzed against 0.02 M Tl^+ at pH 9 and from a crystal dialyzed in 0.1 M Tl^+ at pH 7. The Tl-insulin crystal at pH 7 was somewhat discolored and measurements were unreliable beyond 3.6-Å resolution. When dialyzed for one week against 0.1 M Tl^+ or Rb^+ solutions at pH 9, the crystals did not diffract beyond ~10 Å resolution, although the morphology was unchanged. Crystals dialyzed against Ca^{++}, Mg^{++} and La^{+++} were tested before deterioration could be detected visually, but these crystals diffracted very poorly. Attempts were made to obtain data from crystals soaked in Cs^+ solutions before the crystals cracked but they failed to give significant diffraction. The resolution of the diffraction from crystals following the treatments tested is indicated in Table 1; and the quantity and quality of the data collected from crystals in which Na^+ was replaced by Li^+, K^+, Rb^+, and Tl^+ under various conditions is summarized in Table 2.

Bovine and porcine insulins compared

Fourier difference maps between data obtained from bovine and porcine crystals (grown from 0.25 M Na_2HPO_4 solutions), and phased from the previously solved porcine insulin structure (Badger et al., 1991; Badger and Caspar, 1991), clearly showed the expected differences due to the two amino acid substitutions A8 Thr → Ala and A10 Ile → Val. This comparison demonstrated that porcine and bovine insulin crystals are isomorphous and that localized structural differences of <10 electrons can be detected from the data collected to 2.8-Å resolution. The electron density distribution near B10 His in bovine insulin also indicated an alternative conformation for this side chain with ~50% occupancy, which had not been apparent in the studies of porcine cubic insulin crystals. This interpretation has subsequently been confirmed by refinement against 2.0-Å resolution data ($R = 0.159$) collected from a bovine cubic insulin crystal at pH 9. A re-examination of electron density maps (Badger et al., 1991; Badger and Caspar, 1991) computed form the porcine insulin data (Dodson et al., 1978) showed similar but weaker density in the position of this alternative side chain conformation with an estimated occupancy of ~30%.

TABLE 2 Summary of data collected and processing statistics

Data set	R_{merge} (%)	Resolution ($Å^{-1}$)	Total number data	Unique number data (% possible)
0.1 M Li^+ pH 9	4.4	1/2.8	5,996	1,794 (84)
0.1 M K^+ pH 9	4.5	"	7,411	1,854 (87)
0.1 M Rb^+ pH 9	3.6	"	5,104	1,689 (80)
0.1 M Tl^+ pH 9	5.6	"	6,191	1,666 (78)
0.02 M Tl^+ pH 9	6.7	"	5,575	1,596 (75)
0.3 M Tl^+ pH 9.5	4.6	"	5,966	1,812 (85)
0.1 M Tl^+ pH 7	5.3	1/3.6	4,142	992 (98)

The merging R-factor for symmetry-related intensity measurements, resolution and number of data recorded are shown. Except for the Tl-insulin data measured at pH 7, the maximum resolution corresponds to the edge of the detector and is not the diffraction limit of the crystals. Most of the missing data are in the resolution interval below 10 Å or greater than 3 Å. Crystals were dialyzed for 1 d in Li^+, K^+, Rb^+ or Tl^+.

Locations of bound cations

Difference maps calculated from 2.8-Å resolution data for Tl-insulin (0.1 M Tl^+, pH 9) and comparable Na-insulin crystals showed two cation binding sites (with peak heights 23x and 9x the rms map fluctuation) per insulin molecule. The more fully occupied site lies close to a crystallographic two-fold symmetry axis between insulin molecules from different insulin dimers. (The molecular dimer is a pair of insulin molecules that are associated by antiparallel β-sheet structure in a similar way to the noncrystallographic insulin dimer found in 2Zn insulin crystals [Baker et al., 1988].) At the resolution of this map, the pair of symmetry-related density peaks appear to be connected across the two-fold axis (Fig. 1). Positional parameters for the Tl^+ cation obtained from phased heavy atom least-squares refinement (Dickerson et al., 1968) place these two symmetry-related sites 2 Å apart. Electrostatic repulsion will prevent more than one Tl^+ from occupying this region at any given time, placing an upper limit of 50% on each site occupancy.

Protein-ligating groups in bovine cubic insulin, which are within 3.2 Å of the Tl^+ site, are the main and side chain keto oxygen atoms from A5 Gln and the main chain keto oxygen atom from two copies of A10 Val. The side chain oxygen from A9 Ser is at slightly greater distance, whereas the side chain oxygens from A15 Gln and A12 Ser appear too distant to ligate the Tl^+. A water molecule close to the symmetry-related position across the two-fold axis would complete the coordination of the Tl^+ and could form two hydrogen bonds with two ligating protein keto oxygen atoms. Distances between the calculated Tl^+ location and neighboring protein atoms are listed in the Fig. 1 caption.

The high resolution electron density maps for porcine (Badger et al., 1991) and bovine Na-insulin crystals (pH 9, grown from 0.5 M Na) show sharp elongated features close to each Tl^+ binding site (Fig. 2). We interpret this region as containing one Na^+ cation and one water molecule. The electron density is elongated because the bound Na^+ may approach within ~2.4 Å of the coordinating protein oxygen atoms, whereas the water oxygen contact distances should be ~2.8 Å. Thus, the alternate Na^+ and water binding sites, which are not exactly coincident, superimpose on averaging with the crystal symmetry to give the elongated density peaks.

After refinement without restraining the van der Waals contact distances, the Na^+ cation was located 2.3 Å from the A5 Gln and A10 Ile main chain keto oxygens and the water molecule oxygen in the symmetry-related site. The water molecule was located at normal (2.6–3.3 Å) hydrogen bond distances from the A5 Gln and A10 Ile main chain keto oxygen and the side chain oxygen of A5 Gln as detailed in the caption to Fig. 2. In the high resolution maps of both porcine and bovine cubic insulin crystals, the peak electron density in this feature is comparable to the density for well-ordered protein. When maximum occupancy (i.e., 0.5 Na^+ and 0.5 water in the two lobes of each density streak) is assumed, the individual atomic temperature factors for Na^+ and water refine to values similar to the most ordered parts of the surrounding protein ($B \sim 10$ Å^2).

The second cation binding site that we have been able to identify from the Tl-insulin crystal data is overlapped by one conformation of the two alternate His B10 side chain positions (Fig. 3). This site is only available for cation binding when the His B10 side chain is rotated into the nonoccluding conformation. The cation in this site is contained within a small surface niche between the main chain atoms of A17 Glu and the side chain of A14 Tyr. The main chain keto oxygen of A14 Tyr and the side chain keto oxygen of A18 Asn ligate the cation. Distances to neighboring protein atoms are given in the Fig. 3 caption. In the 2.0-Å resolution difference map computed between the bovine Na-insulin data at pH 9 and the refined model with the two His B10 conformations, a small density peak is found at the site partially occupied by Tl^+. We identify this small peak as a Na^+ cation partially occupying the binding sites vacated by B10 His.

Difference maps computed using 2.8 Å resolution data collected from crystals dialyzed against solutions of other cations (0.1 M Li^+, K^+, Rb^+, all at pH 9) and the Na-insulin data all showed similar but smaller features to those seen in the Tl-insulin difference maps. In the case of the difference map between K-insulin and Na-insulin the peaks are only at the noise level of the map, but the differences are seen more clearly in the difference map calculated between K-insulin and Li-insulin. Thus, Tl^+ and all the Group 1 cations which do not damage the crystal bind at the same two sites on the protein.

Cation occupancy

To assess the effect of changing pH on the occupancy of bound cations, we have compared diffraction data measured from bovine insulin crystals dialyzed against Tl^+ at pH 7, pH 9, and pH 9.5. We have also obtained high resolution diffraction data and refined models for Na-insulin crystals at pH 7, pH 9, and pH 10.2 (Gursky et al., manuscript in preparation) to provide detailed information on the protein and solvent conformation.

Difference maps computed between the data collected from a TL-insulin crystal at pH 9 in 0.1 M Tl^+ and that from crystals at pH 9.5, 0.3 M Tl^+ and pH 7, 0.1 M Tl^+ showed that the amount of Tl^+ at the site near the

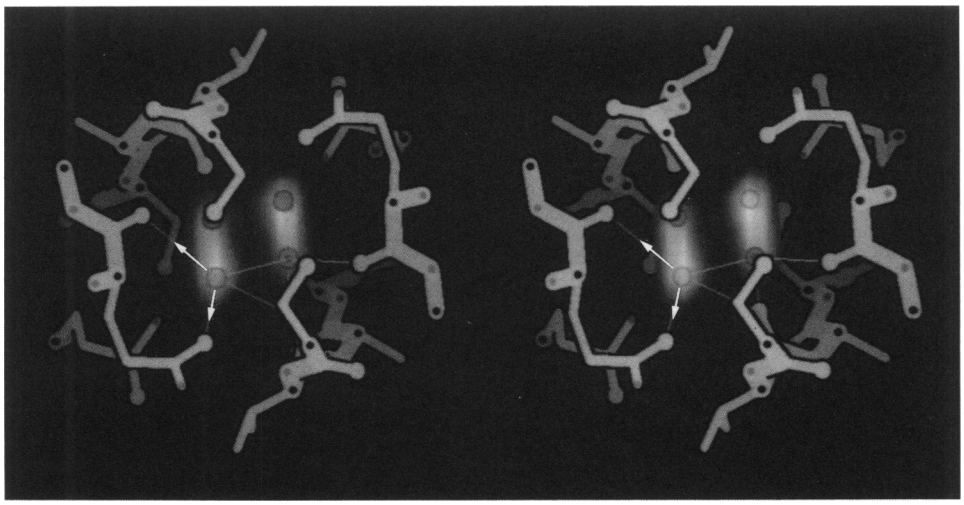

FIGURE 1 Stereo pair showing difference Tl^+ density and the model of the protein environment near the crystal dyad. For Tl^+ at this site ($X = 0.2456$, $Y = 0.5418$, $Z = 0.4879$), the protein ligating atoms (connected by thin lines) are the main and side chain keto oxygens of A5 Gln (2.7 Å and 3.2 Å distant) and two copies of the main chain keto oxygen of A10 Val (2.8 Å and 3.1 Å). The side chain oxygen in A9 Ser above the Tl^+ lies at a slightly longer distance (3.4 Å). Other side chains forming the bottom of the cavity are A15 Gln and A12 Ser. At any given time, only one of the two symmetry-related sites may be occupied by Tl^+, and a water molecule probably occupies a position close to the symmetry-related site. Oxygen atoms are shown as enlarged spheres; Cα positions are marked by black dots, and nitrogen atoms are marked with grey dots. The difference map was computed from diffraction data extending to 2.8-Å resolution from Tl-insulin and Na-insulin and was displayed by projecting electron density greater than 6 × the rms map fluctuation.

crystal dyad axis was somewhat increased at pH 9.5 and was decreased at pH 7. A difference map calculated between the data from the crystals with Tl^+ at 0.02 M and 0.1 M (both at pH 9) showed slightly less Tl^+ bound at the lower concentration. The fractional occupancy factors for the Tl^+ sites obtained from phased heavy atom refinements (Dickerson et al., 1968) using the differences between the various Tl-insulin data sets and the Na-insulin data (0.5 M, pH 9) were consistent with these maps (Table 3).

At pH 9.5 in 0.3 M Tl^+, the Tl^+ occupancy at the site near the crystal dyad is 0.46, which is close to the maximum allowed occupancy of 0.5, but at pH 7 in 0.1 M Tl^+ the refined occupancy is only 0.23. A comparison of high resolution Na-insulin data obtained at pH 9 and pH 10.2 showed no significant structural difference at the site of the bound Na^+/water pair. However, Na-insulin diffraction data obtained at pH 7 showed that the side chain of A5 Gln is predominantly rotated into a different conformation and the space that is vacated is filled by an additional water molecule. The density streaks corresponding to the Na^+/water pair at pH 9 become spherical and are at standard water hydrogen bond lengths from the surrounding protein keto oxygen, the additional water and each other.

The Tl-insulin difference maps also showed that at the Tl^+ binding site associated with B10 His, the amount of bound Tl^+ was much greater in the crystal at pH 9.5 in 0.3 M Tl^+ than at pH 9 in either 0.02 M Tl^+ or 0.1 M Tl^+. At pH 7 the Tl^+ was almost unobservable at this site. These maps and the refined high resolution Na-insulin structures showed that the weights of the two histidine conformations vary with pH. At pH 7 the histidine conformation that occludes the cation site is predominant with scarcely any significant density in the alternative conformation. Compared to the Na-insulin structure at pH 9, the structure at pH 10.2 showed a small decrease in the weight of the histidine conformation that occludes the cation binding site. Electron density maps obtained from Tl-insulin data at pH 9 and pH 9.5 did not show any significant difference in the weights of the two histidine conformations. Although the weights of the two histidine conformations are almost equal at pH 9–9.5, the refined Tl^+ occupancy in the crystal at pH 9.5 in 0.3 M Tl^+ approaches the maximum allowed occupancy of 0.5 but is significantly less at pH 9 (Table 3).

The accuracy of the refined occupancy factors is limited by possible scaling errors between the data, lack of strict isomorphism between crystals (particularly near B10 His, where the conformational equilibrium between the occluding and nonoccluding conformations might be shifted by pH, ionic strength and cation type) and errors and incompleteness in the intensity measurements. Because the data we have measured from the Tl-insulin crystals extend only to 2.8 Å, the magnitudes of structure factor components calculated for the Tl^+ cations are

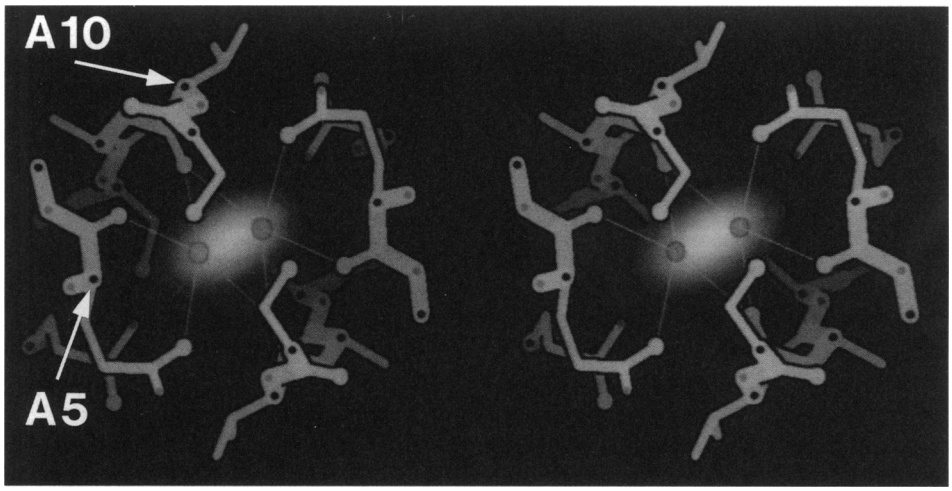

FIGURE 2 Stereo pair showing electron density for the two elongated Na^+/water peaks and the model of the protein near the crystal dyad. Refinement of the Na-bovine cubic insulin structure at pH 9 (currently $R = 0.159$) was by standard restrained least-squares methods (manuscript in preparation). There are two possible symmetry-related positions for the Na^+-water pair. The ligating groups for the Na^+ are the main chain keto oxygens of A5 Gln, A10 Val and the water molecule (all 2.3 Å distant). For the water molecule, the ligating groups are the main chain keto oxygens of A5 Gln (3.1 Å) and A10 Val (3.3 Å), the side chain keto oxygen of A5 Gln (2.6 Å) and the Na^+ cation (2.3 Å). Stereochemical restraints were not imposed in determining the coordinates of the bound Na^+-water pair. The most likely orientation of the water is with hydrogens pointed toward the main and side chain keto oxygens of A5 Gln (as indicated by the arrows) so that the negative end of the water dipole is directed towards the Na^+. The difference map was computed from the diffraction data extending to 2.0 Å resolution after subtracting the structure factors calculated from model coordinates with the Na^+ and water omitted. Details of the model display are as in Fig. 1.

fairly insensitive to the associated atomic temperature factors, and these are not accurately determined. However, refinement tests in which the temperature factors were fixed at likely values or allowed to vary over a plausible range demonstrated that the resulting occupancy factors were not affected by more than ~10%.

DISCUSSION

The initial objective of this study was to determine the hydrated counterion distribution around the insulin molecules by exchanging the Na^+ with Tl^+. For this purpose accurate measurements of the intense low-resolution diffraction data is required (work in progress). From the analysis of the high resolution data reported here (which excludes most data below 10 Å resolution), we have identified two monovalent cation binding sites localized on the protein. One cation is buried in the cavity between insulin dimers, inaccessible to bulk solvent, and the other is contained within a surface niche between insulin molecules.

Crystal stability

Solutions of Cs^+ ions and of the polyvalent cations Ca^{++}, Mg^{++} and La^{+++} were found to rapidly damage the insulin crystals. Crystals equilibrated with Tl^+ and Rb^+ over extended time periods also ceased to diffract. Disordering of the crystal lattice could be due to binding of the large monovalent or multiply charged cations to the protein molecules in the crystal. It may be that disordering of the crystal lattice by the heavy monovalent cations occurs as the larger cations slowly replace the Na^+ ion in the cavity between insulin dimers.

Structure of cation binding sites

The cation binding site in the cavity on the crystal two-fold axis is created by protein keto oxygen groups and one water molecule, without the close involvement of any charged group. Furthermore, this cavity is completely enclosed by the protein groups and there is no direct contact with external solvent. A similar stabilization of buried anions and amino acid zwitter-ions by polarized peptide units has been observed in periplasmic binding proteins (Quiocho et al., 1987). In the cubic insulin example, however, there is no possible contribution from a helix macrodipole, because the helices are not pointing towards the cation binding site.

Metal ion binding sites are frequently associated with groups of polar atoms (oxygen and nitrogen) embedded in hydrophobic environments (Yamashita et al., 1990). Simple electrostatic calculations, using the Coulomb

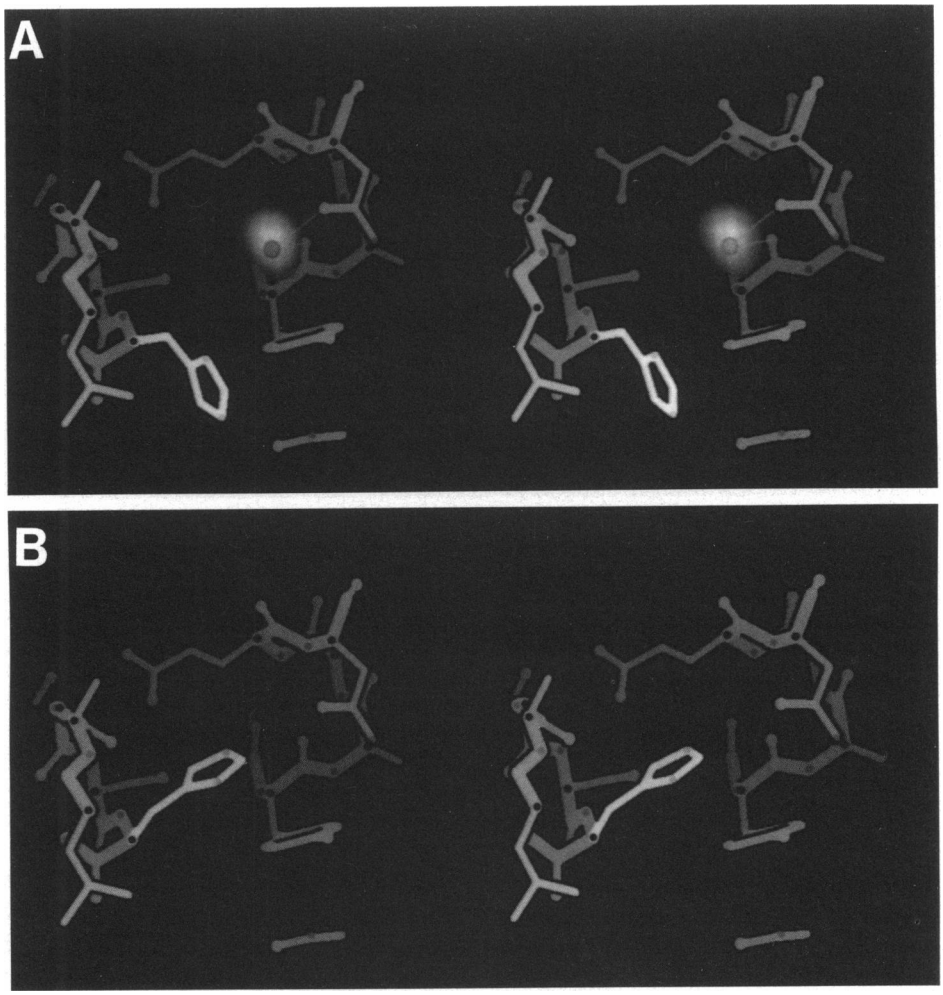

FIGURE 3 Stereo pairs showing (*A*) Tl^+ electron density and the protein model for the available site near B10 His. The cation is located at $X = 0.4245$, $Y = 0.5684$, $Z = 0.3341$ and is ligated by the main chain keto oxygen of A14 Tyr (2.8 Å) and the side chain oxygen of A18 Asn (2.6 Å). At pH 9 there is a second conformation for the side chain of B10 His (*B*) which occludes the cation binding site ~50% of the time. Details of the map calculation and model display are as for Fig. 1.

relation in which a set of partial charges (Weiner et al., 1984) were assigned to the protein atoms, showed that a relatively large electronegative potential is produced in the cation binding cavity of cubic insulin by the combined action of the peptide and side chain carbonyl dipoles. Two water molecules are unlikely to occupy this cavity because none of the surrounding protein atoms are hydrogen donors. However, at pH 7 the major conformation of the side chain of A5 Gln is rotated outside the Na^+ binding cavity, and its place is filled by a new water molecule. Stereochemical considerations imply that the Na^+/water pair is then replaced by two water molecules. Consistent with this result, the Tl-insulin data at pH 7 demonstrated reduced cation occupancy. Because the ionic radii for Na^+ is small, the Na^+ is able to closely approach three oxygens (two main chain keto oxygen atoms and the water molecule as shown in Fig. 2), whereas the Tl^+ is almost equidistant from four ligating oxygens (Fig. 1). Electrostatic energy calculations favor the observed Na^+ position. The interactions between the monovalent cation and the keto oxygens at this site are similar to those modeled for many ion channels (Eisenman and Dani, 1987); examples of heavy monovalent cation binding in protein crystals have been useful in the development of these models of ion channel structure.

The second bound cation occupies a space which is vacated when the side chain of B10 His assumes the nonoccluding conformation. Affinity of this site for a positive charge suggests that the pK of this histidine may be anomalously high. Electrostatic calculations do show that there is a relatively large electronegative field in this

TABLE 3 **Refined Tl^+ occupancy factors in bovine Tl-insulin crystals**

Dialysis conditions	Tl^+ Occupancy	
	Site 1	Site 2
pH 9.5 0.3 M Tl^+	0.46	0.45
pH 9.0 0.1 M Tl^+	0.34	0.20
pH 9.0 0.02 M Tl^+	0.30	0.15
pH 7.0 0.1 M Tl^+	0.23	—

The fractional Tl^+ occupancy at the site near the crystal dyad (site 1) and the site associated with B10 His (site 2) was refined (Dickerson et al., 1968) using the differences between each set of Tl-insulin data and the Na-insulin data. The Na-insulin structure factors were placed on an absolute scale by scaling against a set of structure factors calculated from an atomic model. The occupancy factors have been adjusted assuming that either a Na^+ or a water molecule (10 electrons) is present in the Na-insulin crystal at site 1 and one half of a Na^+ or water (five electrons) is present at site 2. The maximum possible Tl^+ occupancy at site 1 is 0.5 because a symmetry-equivalent site across the crystal dyad is only 2 Å away. The maximum expected occupancy at site 2 is ~0.5 at pH 9–9.5 because one conformation of the side chain of His B10, with occupancy of ~0.5, occludes the site. At pH 7 the occluding conformation of His B10 predominates and Tl^+ is excluded from this site.

region. Histidine B10 occupies the cation site almost exclusively at pH 7, but at pH 9, B10 His occupies both conformations with almost equal weight. The increased occupation of the binding site by the histidine when it is more protonated accounts for the absence of significant cation binding at pH 7. The larger Tl^+ occupancy found at this site at pH 9.5 could be explained by partial titration of A14 Tyr, which is very close to this site, or might also be due to the higher Tl^+ concentration used to dialyze the insulin crystal.

Conclusion

We have surveyed the effects of various salt solutions on the stability of cubic insulin crystals and have identified two sites for monovalent cations specifically bound to the protein molecules. The cations contained within these two sites account for less than one of the 3–4 cations we expect for electro-neutrality of the crystal. By refining the disordered solvent contribution, from analysis of both low and high resolution diffraction data (Badger and Caspar, 1991), it should be possible to map the average distribution of the mobile cations as well as determining more exactly the occupancies and coordinations of the monovalent cations in the specific binding sites.

This work was supported by United States Public Health Service grant CA47439 to D. L. D. Caspar from the National Cancer Institute. Funds to purchase and maintain the computer system were obtained from a Shared Instrumentation grant 1-S10-RR04671-01 awarded to D. J. DeRosier by the National Institutes of Health. Grant BBS8713278 from the National Science Foundation, awarded to K. Kalata, has supported the development of the Brandeis area detector system.

Received for publication 11 March 1991 and in final form 31 October 1991.

REFERENCES

Badger, J., and D. L. D. Caspar. 1991. Water structure in cubic insulin crystals. *Proc. Natl. Acad. Sci. USA.* 88:622–626.

Badger, J., M. R. Harris, C. D. Reynolds, A. C. Evans, E. J. Dodson, G. G. Dodson, and A. C. T. North. 1991. Structure of the pig insulin dimer in the cubic crystal. *Acta Cryst.* B47:127–136.

Baker, E. N., T. L. Blundell, J. F. Cutfield, S. M. Cutfield, E. J. Dodson, G. G. Dodson, D. M. C. Hodgkin, R. E. Hubbard, N. W. Isaacs, C. D. Reynolds, K. Sakabe, N. Sakabe, and N. M. Vijayan. 1988. The structure of 2Zn insulin crystals at 1.5 Å resolution. *Phil. Trans. R. Soc. Lond.* B319:365–456.

Dickerson, R. E., J. E. Weinzierl, and R. A. Palmer. 1968. A least-squares refinement method for isomorphous replacement. *Acta Cryst.* B24:997–1003.

Dodson, E. J., G. G. Dodson, A. Lewitova, and M. Sabesan. 1978. Zinc-free cubic pig insulin: crystallization and structure determination. *J. Mol. Biol.* 125:387–396.

Eisenman, G., and J. A. Dani. 1987. An introduction to molecular architecture and permeability of ion channels. *Annu. Rev. Biophys. Biophys. Chem.* 16:205–226.

Kabsch, W. 1988. Evaluation of single-crystal x-ray diffraction data from a position-sensitive detector. *J. Appl. Cryst.* 21:916–924.

Kalata, K. 1985. A general purpose, computer configurable television area detector for x-ray diffraction applications. *In* Diffraction Methods for Biological Macromolecules. Methods in Enzymology, Vol. 114. H. W. Wyckoff, C. H. W. Hirs and S. N. Timasheff, editors. Academic Press, New York. 486–510.

Kalata, K., W. C. Phillips, M. Stanton, and Y. Li. 1990. Development of a synchrotron CCD-based area detector for structural biology. *Proc. Soc. Photo-Opt. Instr. Eng.* 1345:270–280.

Pflugrath, J. W., and A. Messerschmidt. 1987. Fast system software. *In* Computational Aspects of Protein Crystal Data Analysis: Proceeding of the Daresbury Study Weekend, 23–24 January, 1987. J. R. Helliwell, P. A. Machin and M. Z. Papiz, editors. Science and Engineering Research Council, Daresbury Laboratory.

Quiocho, F. A., J. S. Sack, and N. K. Vyas. 1987. Stabilization of charges on isolated ionic groups sequestered in proteins by polarized peptide units. *Nature* (Lond.). 329:561–564.

Rayment, I., J. E. Johnson, and D. Suck. 1977. A method for preventing crystal slippage in macromolecular crystallography. *J. Appl. Cryst.* 10:356.

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, P. Salvatore, and P. Weiner. 1984. A new force field for molecular mechanical simulation of nucleic acids and Proteins. *J. Am. Chem. Soc.* 106:765–784.

Yamashita, M. M., L. Wesson, G. Eisenman, and D. Eisenberg. 1990. Where metal ions bind in proteins. *Proc. Natl. Acad. Sci. USA.* 87:5648–5652.