Schmid, M. F., Weaver, L. H., Holmes, M. A., Gruetter, M. G., Ohlendorf, D. H., Reynolds, R. A., Remington, S. J., & Matthews, B. W. (1981) *Acta Crystallogr. A37*, 701–710.
Srere, P. (1966) *J. Biol. Chem. 241*, 2157–2165.
Srere, P. (1972) *Curr. Top. Cell. Regul. 5*, 245–283.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489–501.
Weitzman, P. D. J., & Danson, M. J. (1976) *Curr. Top. Cell. Regul. 10*, 161–204.
Wiedman, S. W., Drysdale, G. R., & Mildvan, A. S. (1973) *Biochemistry 12*, 1874–1883.
Wiegand, G., & Remington, S. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 97–117.
Wiegand, G., Remington, S., Deisenhofer, J., & Huber, R. (1984) *J. Mol. Biol. 174*, 205–211.

Crystal Structure of an Open Conformation of Citrate Synthase from Chicken Heart at 2.8-Å Resolution[†,‡]

Der-Ing Liao, Mihail Karpusas, and S. James Remington*
Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403
Received October 12, 1990; Revised Manuscript Received February 22, 1991

ABSTRACT: The X-ray structure of a new crystal form of chicken heart muscle citrate synthase, grown from solutions containing citrate and coenzyme A or L-malate and acetyl coenzyme A, has been determined by molecular replacement at 2.8-Å resolution. The space group is $P4_3$ with $a = 58.9$ Å and $c = 259.2$ Å and contains an entire dimer of molecular weight 100 000 in the asymmetric unit. Both "closed" conformation chicken heart and "open" conformation pig heart citrate synthase models (Brookhaven Protein Data Bank designations 3CTS and 1CTS) were used in the molecular replacement solution, with crystallographic refinement being initiated with the latter. The conventional crystallographic R factor of the final refined model is 19.6% for the data between 6- and 2.8-Å resolution. The model has an rms deviation from ideal values of 0.034 Å for bond lengths and of 3.6° for bond angles. The conformation of the enzyme is essentially identical with that of a previously determined "open" form of pig heart muscle citrate synthase which crystallizes in a different space group, with one monomer in the asymmetric unit, from either phosphate or citrate solution. The crystalline environment of each subunit of the chicken enzyme is different, yet the conformation is the same in each. The open conformation is therefore not an artifact of crystal packing or crystallization conditions and is not species dependent. Both "open" and "closed" crystal forms of the chicken heart enzyme grow from the same drop, showing that both conformations of the enzyme are present at equilibrium in solution containing reaction products or substrate analogues.

In higher organisms citrate synthase is a dimer of two identical subunits related by a 2-fold axis and has a total molecular weight of about 100 000. For extensive reviews of the physical–chemical properties of the enzyme from a wide variety of sources, see Srere (1972), Weitzman and Danson (1976), Beeckmans (1984), and Wiegand and Remington (1986). The structures of three different crystal forms of pig and chicken heart muscle citrate synthase have previously been determined [Remington et al., 1982; Wiegand et al., 1984; Karpusas et al., 1990, 1991 (preceding paper in this issue)] for unliganded enzyme and a variety of ternary and binary complexes. Each subunit can be described as consisting of two domains, a large and a small one, with the substrate binding site in the cleft between the two domains. The enzyme was the first shown to consist almost entirely of α helices.

Citrate synthase from pig heart has been shown to be capable of very large conformational changes. Two of the crystal structures (space groups $C2$ and $P4_32_12$) are very similar and are classified as "closed" (Remington et al., 1982; Wiegand et al., 1984; Wiegand & Remington, 1986). The two closed forms differ primarily in the arrangement of internal side chains while the third one (space group $P4_12_12$) is "open" (Remington et al., 1982).

In the "closed" forms ligands are buried inside the cleft and inaccessible to bulk solvent. The change in conformation between the open and closed form can be roughly approximated by a rigid body rotation of 19° of the small domain relative to the large one; however, there are substantial changes in the packing of side chains in the interior of the small subunit (Wiegand et al., 1984; Lesk & Chothia, 1984) upon the transition between conformations. Lesk and Chothia have argued that the conformational change is best described as a concerted series of shifts between adjacent helices (a "plastic" deformation) rather than a rigid body rotation. This description implies that a continuous range of stable conformations may be available to a given molecule, which we feel may be incorrect and additionally yield a misleading picture as to the mechanism of conformational change. In this instance, we present evidence which suggests that there are only two stable conformational states of citrate synthase relevant to the enzymatic mechanism.

Evidence for conformational change upon addition of ligands has been obtained from a variety of spectroscopic studies of the protein in solution (Srere, 1966; Bayer et al., 1981; Weidman et al., 1981; Kollmann-Koch & Eggerer, 1989). These and crystallographic studies suggest that the binding

[†] This research was supported in part by a grant from the National Science Foundation (DMB-8817438) to S.J.R. and a grant to the Institute of Molecular Biology from the Lucille P. Markey Charitable Trust.
[‡] Crystallographic coordinates have been submitted to the Brookhaven Protein Data Bank with designations as given in the text.
* To whom correspondence should be addressed at the Institute of Molecular Biology.

0006-2960/91/0430-6031$02.50/0

of oxaloacetate or oxaloacetate analogues induces a change in conformation that creates the binding site for acetylCoA[1] (Remington et al., 1982). Recent spectroscopic studies (Kollman-Koch & Eggerer, 1989) suggest that acetylCoA or an acetylCoA analogue can also induce a conformational change which can be completed by the addition of oxaloacetate or an oxaloacetate analogue. However, the nature of this change and its relation to the oxaloacetate-induced change remains obscure.

Until recently, crystals of the chicken heart enzyme could only be obtained in space group $C2$ (in the closed conformation), creating the possibility that the open form of the pig heart enzyme is an artifact of crystal packing or crystallization conditions or is species dependent and may thus be irrelevant to the enzymatic mechanism. The appearance of a new crystal form in a series of crystallization experiments with the chicken heart enzyme created the opportunity to determine the structural state of the enzyme from a different species in a different crystalline environment.

Materials and Methods

Chicken heart citrate synthase, citrate, and coenzyme A were purchased from Sigma. The crystals were grown by hanging drop vapor diffusion experiments at room temperature or at 4 °C. The drop had an initial volume of 10 μL and initial concentrations of 10 mg/mL protein, 0.5 M citrate, and 10 mM CoA, pH 6.0, while the well contained 1 mL of 1.0 M sodium citrate, pH 6.0. These conditions are almost identical with the conditions used for growth of monoclinic $C2$ crystals (Remington et al., 1982). Tetragonal crystals grew after about 2 weeks and were bipyramids of maximum dimensions 0.4 × 0.4 × 0.8 mm^3. In a few cases both monoclinic and tetragonal crystals grew out of the same drop. Both crystal forms have also been obtained in the same drop by using 1.0 M L-malate as the precipitant in the presence of 10 mM acetyl coenzyme A, pH 6.0. The structure of the monoclinic crystal form is discussed in the preceding paper (Karpusas et al., 1991).

Data were collected by oscillation photography, using a graphite monochromator (Schmidt et al., 1981) on an Elliot GX-21 X-ray generator. The collimator diameter was 0.2 mm. Crystals were rotated around their c axis, and photographs were taken at intervals of 3°, on Kodak No-Screen X-ray films mounted in flat cassettes. The crystal to film distance was 100 mm. The film data were processed by using a program based on that of Rossmann (Rossmann, 1979; Schmidt et al., 1981).

Initial inspection of precession photographs indicated that the crystals belong to the tetragonal system with cell dimensions a = 58.85 Å and c = 259.22 Å and were initially believed to be space group $P4_122$ or enantiomorph. Molecular replacement techniques were used in order to determine the space group and the position of the molecule in the unit cell. A polar angle fast rotation function programmed by Wolfgang Kabsch, based on the theory of Crowther and Tanaka (Crowther, 1972; Tanaka, 1977), was used to determine the orientation of the molecule. Both fast translation function (Crowther & Blow, 1967) and correlation R-factor translation function (Cygler & Anderson, 1988; S. J. Remington, unpublished program) calculations were used to determine the position of the molecule. For the initial rotation and translation function calculations a dimeric model based on the $C2$ chicken heart crystal structure (Remington et al., 1982; Brookhaven Protein Data Bank designation 3CTS) that included only the large domains of each subunit (residues 1–274 and 381–437) was used. Data in the range between 10- and 4-Å resolution were used for rotation function calculations and between 6- and 4-Å resolution for translation function calculations.

The translation function revealed that the correct hand of the screw axis was 4_3; however, we encountered many difficulties during the solution of the translation problem. Later it was discovered that the space group was in fact $P4_1$ or enantiomorph with an entire dimer in the asymmetric unit. The 2-fold axis of the dimer is only 0.4° from parallel to the diagonal in the a–b plane, which introduces nearly perfect pseudosymmetry into the diffraction pattern. The symptoms which indicated that the space group had been incorrectly identified were that calculations of the correlation R-factor translation function gave two large peaks on the Z = 0 section (related by the approximate 2-fold axis through the a–b diagonal) and that the fast translation function calculation failed to yield a significant peak for the Z coordinate. An hll precession photograph indicated that the diffraction pattern had only approximate mm symmetry on this level. The true Laue symmetry was 4/m.

After the correct space group had been identified, the film data were reprocessed. As 3 different crystals were used for the data collection, the correct relative orientation of two crystals had to be determined with respect to the third. Data merging statistics were inconclusive, so this was accomplished by calculation of R-factor translation functions for data from each crystal in turn, for each choice of $+l$. This allowed a self-consistent selection of crystal orientation.

To the extent that the conformational change from open to closed can be regarded as rigid body rotation of the small domain relative to the large, a series of models were created with conformations intermediate between open and closed. This follows as a result of the fact that any orthogonal transformation can be decomposed into a translation along a particular vector combined with a rotation about that vector (W. S. Bennett, unpublished program). For each of these intermediate models the translation function was calculated and the peak heights were compared. The highest value was for that of the model corresponding to the previously solved open form of the pig heart enzyme. The conventional crystallographic R factor was also minimal when the open form was used as a model.

The open ($P4_12_12$, Brookhaven Protein Data Bank designation 1CTS) form of the pig heart model was oriented accordingly. To distinguish between the two molecules in the asymmetric unit, 500 was added to the residue numbers of one of the molecules. These were used as an initial model for rigid body positional refinement of four separate domains (two large and two small) using the TNT program package (Tronrud et al., 1987), followed by conventional model building and crystallographic refinement. Electron density maps were averaged about the local 2-fold axis for inspection in the initial stages of model building, but in the final stage unaveraged electron density maps were inspected as well.

In addition to the usual refinement restraints on bond lengths and angles, a restraint on noncrystallographic 2-fold symmetry was included. To the overall error function minimized by the TNT package (E_0) was added a term which restrained each atom in one subunit to the coordinates obtained by applying the local 2-fold operator to its counterpart in the other subunit:

$$E = E_0 + W\sum[X_i - (RX'_i + T)]^2$$

Here, E is the overall error function, W is a relative weight

[1] Abbreviations: CS, citrate synthase; CoA, coenzyme A; CMCoA, carboxymethyl coenzyme A; rms, root mean square.

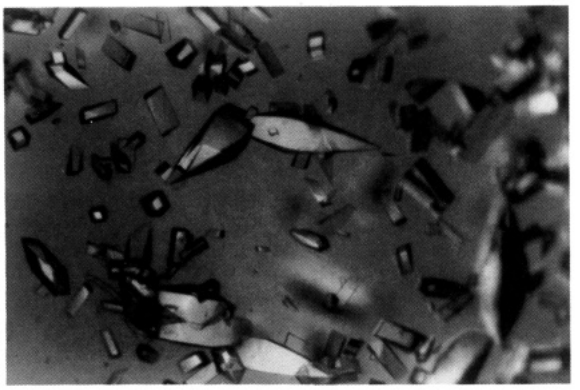

FIGURE 1: Photograph of a crystallization experiment that produced both monoclinic $C2$ (closed conformation, crystals are approximately rectangular) and tetragonal $P4_3$ (open conformation, bipyramidal) crystals. The $P4_3$ crystals are approximately 0.8 mm from tip to tip (the longest dimension) while the $C2$ crystals have a longest dimension of about 0.4 mm.

Table I: Data Collection Statistics

	mother liquor	
	CS/citrate/CoA	CS/L-malate/ AcCoA
intensity measurements	53 801	28 151
independent reflections	15 130	15 040
R_{merge} (%)	13.9	15.8
resolution (Å)	2.7	2.7
completeness of data[a] (%)		
6.0–2.7-Å range	62.4	62.5
2.8–2.7-Å range	22.0	26.5

[a] Completeness is the ratio of the measured to the possible number of reflections for this resolution shell.

X_i is the coordinate of the ith atom in one subunit and X'_i its counterpart, and R and T are the rotation matrix and translation vector relating one subunit to the next [determined after rigid body refinement by the algorithm of Kabsch (1978)]. The weighting factor W was determined empirically such that the rms deviation from local symmetry did not increase beyond 0.3 Å during the course of the refinement.

There is as yet no available sequence for the chicken heart enzyme; however, the side chains that were thought to be different, on the basis of work described in Remington et al. (1982), were removed from the pig heart enzyme model. The R factor dropped sharply from 32% to 18% in one cycle of least-squares crystallographic refinement for data from 6- to 3.5-Å resolution. After 4 cycles of refinement, data to 3-Å resolution were used for 8 additional cycles of refinement. The model and $2F_o - F_c$ electron density map were inspected, and the model was altered as necessary. For subsequent refinement, data up to 2.8-Å resolution were included and the weighting on the stereochemical constraints was increased. Individual atomic temperature factors were not refined, nor were water molecules included in the final model. In the last few cycles of coordinate refinement the weight on noncrystallographic symmetry was set to zero. Finally, two cycles of limited thermal factor refinement was performed in which each residue was assigned a single temperature factor by the TNT programs.

Results

Figure 1 is a photograph showing a crystallization experiment in which two different crystal forms of chicken heart citrate synthase were obtained in the presence of citrate and coenzyme A as described under Materials and Methods. The film data set collected from the tetragonal crystal form shown

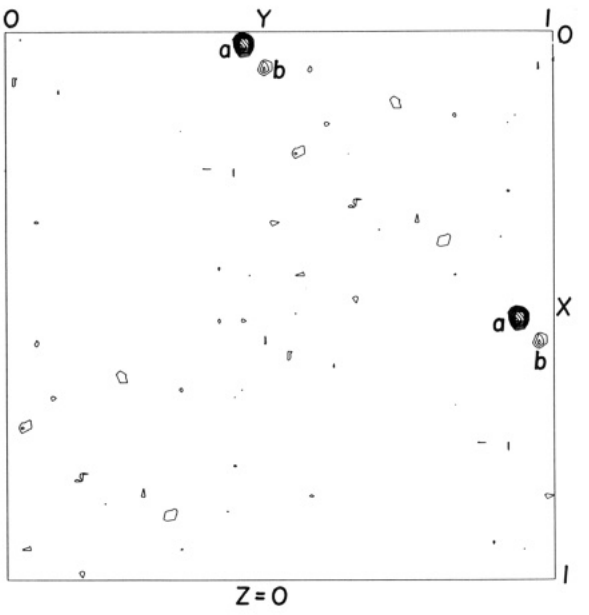

FIGURE 2: Correlation R-factor translation function for a full unit cell in X and Y on the $Z = 0$ section. The contour levels indicate increasing correlation of calculated and observed structure amplitudes and are in steps of 1σ, from 2σ. Peaks labeled "a" in the figure and in Table II are two equivalent solutions related by choice of origin. The weaker ghost peaks labeled "b" are a consequence of the strong noncrystallographic pseudosymmetry and are related to the main peak by a 180° rotation about the a–b diagonal of the crystal.

Table II: R-Factor Translation Function Peaks for Final Merged Data

X (Å)	Y (Å)	peak height (σ)[a]	label
37.73	67.30	10.45	a
40.75	70.07	4.65	b
51.97	2.27	3.50	

[a] σ is the unit of peak height defined as the root mean square value of all points of the function calculated.

included 15 130 independent reflections that extended to 2.7-Å resolution. Data statistics are in Table I.

The rotation function revealed a single large peak with peak to background ratio of 5. This proved to be the correct orientation as indicated by the translation function results summarized in Table II and Figure 2. After crystallographic refinement of this solution as described, the conventional crystallographic R factor the final model is 19.6%, and the rms deviation from ideal values of bond lengths is 0.034 Å and of angles is 3.6°. This geometry is admittedly somewhat distorted, but a number of factors should be taken into account when assessing the quality of the model. The film data were not of the highest quality due to the long c cell dimension and constraints on camera geometry. The amino acid sequence is unknown, and our guessed sequence is likely to be incorrect in many places. The resolution of the data is also not as high as one would like. Taking these points into account, the model represents a compromise between many competing factors. Coordinates of the refined structure have been deposited to the Brookhaven Protein Data Bank (Bernstein et al., 1977; Brookhaven Protein Data Bank designation 5CSC). Electron density was very weak for three regions in each monomer, presumably due to flexibility, and these parts of the molecule have not been modeled. These are residues 83, 292–294, 434–437, 583, 792–794, and 934–937.

The rms deviation for all atoms after superimposing the two monomers in the asymmetric unit is 0.40 Å. The molecular 2-fold axis resulting from this superposition is only 0.4° from

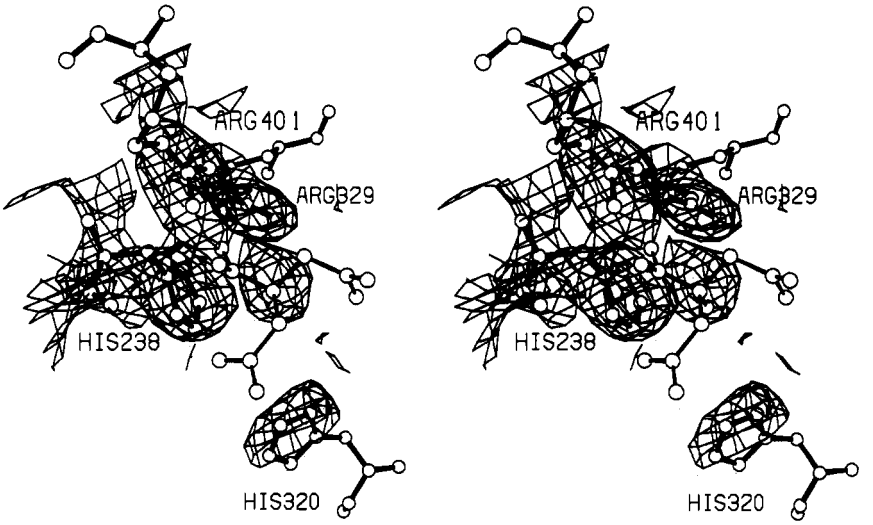

FIGURE 3: Final $2F_o - F_c$ electron density map superimposed on a model of citrate (thinner bonds) bound in the active site. Selected active site side chains which are citrate ligands in the "closed" conformation are also shown. The density cage is contoured at 1 standard deviation of the map. The position of the citrate molecule is based on the *C*2 crystal structure which contains bound citrate, after superimposing the side chains of arginine 401. Although the crystals were grown from 1.0 M citrate solution, the density feature at the expected binding site is only about the size expected for a water molecule, indicating that citrate is either not bound at all, has low occupancy (consistent with very low affinity), or is poorly ordered when bound to the open conformation.

parallel to the diagonal of the *a*–*b* plane.

Table I includes also information and statistics about X-ray data collected from tetragonal crystals grown with L-malate as the precipitant and containing acetylCoA instead of citrate and coenzyme A. An $F_{o,malate} - F_{o,citrate}$ map was calculated by using the refined structure as a model for the calculation of phases. No significant peak was observed in the map, indicating the two crystal structures to be indistinguishable, presumably since in neither case are ligands bound in a well-ordered manner (discussed later).

The overall molecular conformation of the enzyme is essentially identical with that of the $P4_12_12$ open form previously determined for the pig heart enzyme. The rms coordinate difference of main-chain atoms between the two structures, after superposition of the pig heart monomer on each independent monomer in the chicken heart structure in turn using the algorithm of Kabsch (1978), is 0.45 Å for each subunit. This is in fact close to the combined estimated error of the two structure determinations. The matrices which transform each of the first and second subunits of the chicken heart model onto the subunit of the pig heart $P4_12_12$ atomic model are

$$\begin{vmatrix} X'_1 \\ Y'_1 \\ Z'_1 \end{vmatrix} = \begin{vmatrix} -0.9130 & 0.0929 & -0.3972 \\ 0.0815 & -0.9125 & 0.4008 \\ -0.3997 & -0.3984 & -0.8256 \end{vmatrix} \begin{vmatrix} X_1 \\ Y_1 \\ Z_1 \end{vmatrix} + \begin{vmatrix} 82.2 \\ 66.1 \\ 83.5 \end{vmatrix}$$

$$\begin{vmatrix} X'_2 \\ Y'_2 \\ Z'_2 \end{vmatrix} = \begin{vmatrix} -0.0811 & 0.9121 & 0.4018 \\ 0.9144 & -0.923 & 0.3942 \\ 0.3967 & 0.3994 & -0.8265 \end{vmatrix} \begin{vmatrix} X_2 \\ Y_2 \\ Z_2 \end{vmatrix} + \begin{vmatrix} 11.2 \\ -4.8 \\ 14.8 \end{vmatrix}$$

where X_1 is the coordinate vector of an atom in subunit 1 in Å, etc. Inspection of superimposed models reveals no evidence for differences in the relative arrangements of the large (residues 1–274 and 381–437) and small (residues 275–380) domains. Also, the temperature factors show no evidence for a difference in the vibrational behavior of atoms in the domains.

Inspection of crystal contacts in the two different "open" crystal forms is quite interesting (see Table III) as there is not a single interaction in common between the two forms. The pig heart $P4_12_12$ crystals contain extensive contacts between the large domain of one molecule and the small domain of a neighbor, forming several patches of 3–8 residues each.

Table III: Crystal Contacts in the Two Different Crystals of the "Open" Conformation[a]

crystal	reference contact	neighbor contact	symmetry operator
$P4_3$	206	365*	$X+1, Y, Z$
	219–221	790*	$-Y-1, X, Z-{}^1/_4$
	195–196	795–797*	$-Y-1, X, Z-{}^1/_4$
	153	527	$X+1, Y, Z$
	82–84	561–562	$-Y+1, X, Z-{}^1/_4$
$P4_12_12$	432	293*	$-Y+{}^1/_2, X-{}^1/_2, Z+{}^1/_4$
	116–119	296–299*	$-Y-{}^1/_2, X-{}^1/_2, Z+{}^1/_4$
	160	325*	$-Y-{}^1/_2, X-{}^1/_2, Z+{}^1/_4$
	17–28	290–296*	$Y+{}^1/_2, -X-{}^1/_2, Z+{}^1/_4$
	163–165	337–342*	$Y+{}^1/_2, -X-{}^1/_2, Z+{}^1/_4$
	14–23	342–352*	$Y+{}^1/_2, -X-{}^1/_2, Z+{}^1/_4$
	414–416	342–343*	$Y+{}^1/_2, -X-{}^1/_2, Z+{}^1/_4$
	163	223	$Y+{}^1/_2, -X-{}^1/_2, Z+{}^1/_4$

[a] Due to the lack of sequence information for the chicken heart enzyme, a loose definition of crystal contacts was used. Intermolecular contacts were deemed to exist if an α-carbon of the reference molecule is within 8 Å of an α-carbon of a symmetry-related neighbor. The apparent overlap between some contact regions is due to one molecule simultaneously contacting two others. Contacts are further classified according to whether the two large domains form contacts or whether the large and small domains are in contact (residue numbers with asterisks). There are no instances where the small domains of two neighboring molecules contact in either crystal form. Residue numbers separated by a dash indicate that all residues in that sequence satisfy the contact criterion. (*) Indicates residues in the small domain.

In contrast, the $P4_3$ chicken heart crystal has far fewer sets of interactions. These appear to be smaller in area and are scattered over the surface of the molecule. Thus, the large and small domains in both "open" crystal forms maintain identical relative positions independently of the different constraints imposed on the molecule by different crystal contacts, indicating that the conformation is an energy minimum.

There is no observable density for coenzyme A or acetylCoA and very little density, insufficient to model the citrate molecule or malate, at the citrate binding site (Figure 3) in maps calculated from the two data sets. This is in remarkable agreement with the results of Remington et al. (1982) for the pig heart enzyme and suggests that the affinity of citrate and coenzyme A is low for the open forms. Also, both the pig heart ($P4_12_12$) and chicken heart ($P4_3$) open form crystals shatter

when oxaloacetate is added to the mother liquor, suggesting that oxaloacetate induces a conformational change incompatible with crystal contacts.

Discussion

An "open" conformation of citrate synthase from pig heart and chicken heart crystallizes in two different crystal forms, with three different crystalline environments for the individual monomers. Enzyme from both sources can also be crystallized in two different crystal forms under a wide variety of conditions in "closed" conformations. Indeed, the $P4_32_12$ crystal form of pig heart citrate synthase containing oxaloacetate and *S*-acetonyl coenzyme A described by Wiegand et al. (1984) also contains an entire dimer in the asymmetric unit, so that there are a total of six different crystalline environments in which these conformations have been observed.

This strongly suggests that the different conformations of the enzyme are stable states of the enzyme in solution and not artifacts of crystal contacts or crystallization conditions, nor are they species dependent. Further evidence for the above argument comes from the observation that $P4_3$ (open) and $C2$ (closed) crystals simultaneously grow out of the same drop (Figure 1). It seems therefore likely that the open and the closed conformations are the two most populated stable conformations of this enzyme and are in equilibrium on the time scale of crystal growth.

Kollmann-Koch and Eggerer (1989) stated that the formation of an open form crystal from citrate solution cannot be reconciled with their spectroscopic evidence that both oxaloacetate and citrate induce one and the same conformational change. Our results demonstrate conclusively that $P4_3$ crystals grow out of a (possibly minor) population of molecules in the open conformational state in a solution containing approximately 1.0 M citrate. The same is true of the pig heart enzyme, where either open ($P4_12_12$) and closed ($C2$) form crystals are obtained from 1.0 M citrate, depending on small differences in pH (Remington et al., 1982). The apparent contradiction between the spectroscopic work and the crystallographic results can easily be reconciled if one considers that formation of the crystal contacts of the open forms of citrate synthase from citrate solution may be more favorable thermodynamically than the citrate-induced conformational change. However, upon addition of oxaloacetate, both open crystals forms crack, implying the reverse for this substrate.

A series of crystallization experiments were performed at pH 6.0 with the chicken heart enzyme and indicated that the crystal form obtained (and therefore the enzyme conformation) depends on the choice of ligands. The enzyme gives invariably "closed" $C2$ crystals in the presence of carboxymethylCoA (CMCoA), a strong inhibitor that resembles a proposed transition state of the reaction (Karpusas et al., 1990). If other CoA analogues are used (such as CoA, or acetylCoA and L-malate) than both "open" and "closed" crystals can be grown. Since CMCoA is known to bind much more strongly (K_i = 0.07 μM in the presence of oxaloacetate; Bayer et al., 1981) than do other CoA analogues, this behavior must be a consequence of a solution equilibrium between the "open" and "closed" states which is driven toward the "closed" state in the presence of CMCoA or other tight-binding ligands.

Remington et al. (1982) suggested that the open–closed conformational change is required for the catalytic cycle as no enzymatic activity could be detected in either crystal form available at the time. They suggested that the open form is a "product release/substrate entry" form of the enzyme. This work demonstrates that the open form of the enzyme is a stable conformation of the chicken heart enzyme and is not an artifact of crystallization conditions. Therefore, it seems unlikely that this conformation or that conformational change has no function. Karpusas et al. (1990) proposed that the enzyme conformation corresponding to that in the $C2$ (closed) crystal form can catalyze all steps in the enzymatic reaction. Since substrates and products are isolated from bulk solvent in the closed forms, it seems very plausible that the conformational change is an induced fit process that is necessary for substrate entry and binding and product release. The apparent low affinity (see Figure 3) of both open form crystals for products or substrate analogues supports this argument.

The significance of the second "closed" conformation of the pig heart enzyme is unclear, since it was obtained in the presence of a potent inhibitor of the enzyme, *S*-acetonyl coenzyme A (Wiegand et al., 1984) and may thus be locked in an unnatural state. Those workers commented that the inhibitor appeared to bind to the enzyme in a disordered manner, since the electron density in the CoA binding site could not be interpreted. Whether this conformation is relevant to the enzymatic mechanism is unknown; however, Kollmann-Koch and Eggerer (1989) have provided spectroscopic evidence for three different conformational states in solution depending on the particular ligand(s) present. It appears that further solution work, perhaps under conditions similar to those required to grow the crystals, will be necessary to resolve this issue. It would be exceedingly interesting, possibly by NMR techniques, to determine the time scale at which the different conformations interconvert. Another experimental approach might be to create mutants in which the enzyme could be locked into various conformational states, for example, by introducing disulfide bridges between domains or subunits (Matsumura & Matthews, 1989). We are pursuing these and other ideas in the effort to shed light on the role of conformational changes in this system.

Acknowledgments

We thank Larry Weaver, Dale Tronrud, and Cai Zang for help in data collection and computing and Jeff Bell, Jim Endrizzi, and Enoch Baldwin for helpful comments on the manuscript.

Registry No. CS, 9027-96-7.

References

Bayer, E., Bauer, B., & Eggerer, H. (1981) *Eur. J. Biochem. 120*, 155–160.

Beekmans, S. (1984) *Int. J. Biochem. 16*, 341–351.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Schimanouchi, T., & Tasunmi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Crowther, R. A. (1972) in *The Molecular Replacement Method* (Rossmann, M. G., Ed.) pp 173–178, Gordon & Breach, New York, NY.

Crowther, R. A., & Blow, D. M. (1967) *Acta Crystallogr. 23*, 544–546.

Cygler, M., & Anderson, W. F. (1988) *Acta. Crystallogr. A44*, 300–308.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Kabsch, W. (1978) *Acta Crystallogr. Sect. A 34*, 827–828.

Karpusas, M., Branchaud, B., & Remington, S. J. (1990) *Biochemistry 29*, 2213–2219.

Karpusas, M., Holland, D., & Remington, S. J. (1991) *Biochemistry* (preceding paper in this issue).

Kollmann-Koch, A., & Eggerer, H. (1989) *Eur. J. Biochem. 185*, 441–447.

Lesk, M., & Chothia, C. (1984) *J. Mol. Biol. 174*, 175–191.

Matsumura, M., & Matthews, B. W. (1989) *Science 243*, 792–794.
Remington, S., Wiegand, G., & Huber, R. (1982) *J. Mol. Biol. 158*, 111–152.
Rossmann, M. (1979) *J. Appl. Crystallogr. 12*, 225–238.
Schmid, M. F., Weaver, L. H., Holmes, M. A., Gruetter, M. G., Ohlendorf, D. H., Reynolds, R. A., Remington, S. J., & Matthews, B. W. (1981) *Acta Crystallogr. A37*, 701–710.
Srere, P. A. (1966) *J. Biol. Chem. 241*, 2157–2159.
Srere, P. A. (1972) *Curr. Top. Cell. Regul. 5*, 245–283.
Tanaka, N. (1977) *Acta Crystallogr. A33*, 191–193.
Tronrud, D., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489–501.
Weitzman, P. D. J., & Danson, M. J. (1976) *Curr. Top Cell. Regul. 10*, 161–204.
Wiegand, G., & Remington, S. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 97–117.
Wiegand, G., Remington, S., Deisenhofer, J., & Huber, R. (1984) *J. Mol. Biol. 174*, 205–219.

Assignments of Backbone ^1H, ^{13}C, and ^{15}N Resonances and Secondary Structure of Ribonuclease H from *Escherichia coli* by Heteronuclear Three-Dimensional NMR Spectroscopy

Toshio Yamazaki,*,‡ Mayumi Yoshida,§ Shigenori Kanaya,∥ Haruki Nakamura,∥ and Kuniaki Nagayama*,‡

Biometrology Lab, JEOL Limited, Musashino, Akishima, Tokyo 196, Japan, Tokyo Research Laboratories, Kyowa Hakko Kogyo Company, Limited, Machida, Tokyo 194, Japan, and Protein Engineering Research Institute, Furuedai, Suita, Osaka 565, Japan

Received January 7, 1991; Revised Manuscript Received March 20, 1991

ABSTRACT: The assignments of individual magnetic resonances of backbone nuclei of a larger protein, ribonuclease H from *Escherichia coli*, which consists of 155 amino acid residues and has a molecular mass of 17.6 kDa are presented. To remove the problem of degenerate chemical shifts, which is inevitable in proteins of this size, three-dimensional NMR was applied. The strategy for the sequential assignment was, first, resonance peaks of amides were classified into 15 amino acid types by ^1H–^{15}N HMQC experiments with samples in which specific amino acids were labeled with ^{15}N. Second, the amide ^1H–^{15}N peaks were connected along the amino acid sequence by tracing intraresidue and sequential NOE cross peaks. In order to obtain unambiguous NOE connectivities, four types of heteronuclear 3D NMR techniques, ^1H–^{15}N–^1H 3D NOESY-HMQC, ^1H–^{15}N–^1H 3D TOCSY-HMQC, ^{13}C–^1H–^1H 3D HMQC-NOESY, and ^{13}C–^1H–^1H 3D HMQC-TOCSY, were applied to proteins uniformly labeled either with ^{15}N or with ^{13}C. This method gave a systematic way to assign backbone nuclei (N, NH, $C^\alpha H$, and C^α) of larger proteins. Results of the sequential assignments and identification of secondary structure elements that were revealed by NOE cross peaks among backbone protons are reported.

Ribonuclease H (RNase H)[1] is an enzyme that cleaves the RNA moiety of an RNA–DNA hybrid endonucleolytically. It degrades an RNA strand into oligonucleotides with 5′-phosphate and 3′-hydroxyl termini with the aid of Mg^{2+} but is inactive with respect to other structures of nucleic acids, such as single- or double-stranded DNA and single- or double-stranded RNA. This function was first recognized in extracts from calf thymus (Stein & Hausen, 1969) and was later found in various organisms ranging from prokaryotes to higher eukaryotes (Crouch & Dirksen, 1982). The C-terminal domains of retroviral reverse transcriptases also have this function.

In the present study, RNase H from *Escherichia coli* has been investigated; this enzyme consists of 155 amino acid residues and has a molecular mass of 17.6 kDa (Kanaya & Crouch, 1983). A series of site-directed mutagenesis experiments (Kanaya et al., 1990) revealed the active site of this enzyme. Recently, a three-dimensional structure was proposed for the RNase H on the basis of X-ray crystallography data, at a 1.8-Å resolution (Katayanagi et al., 1990; Yang et al., 1990).

The nuclear magnetic resonance experiment has recently become one of the most efficient techniques for clarifying the structure–function relationship of proteins. For the atomic scale investigation of protein structures, each resonance peak in NMR spectra should first be assigned to a specific atom in the molecule. The assignment strategy for small proteins or peptides (<10 kDa) has been well-established (Wüthrich, 1986). The first step is the use of COSY to classify proton resonances into groups of amino acids according to the spin–spin coupling networks that reflect the amino acid chemical structures. The second step is the sequential connection of the classified spin systems along the primary structure. In most of the neighboring residues, pairs of protons, $C^\alpha H(i-1)$–$NH(i)$ [the proton at the α-carbon of $(i-1)$th residue and the amide

* To whom correspondence should be addressed.
‡ Biometrology Lab, JEOL Ltd.
§ Tokyo Research Laboratories, Kyowa Hakka Kogyo Co., Ltd.
∥ Protein Engineering Research Institute.

[1] Abbreviations: C^α, α-carbon; $C^\alpha H$, proton at the α-carbon; C^β, β-carbon; $C^\beta H$, proton at the β-carbon; $C^\gamma H_3$, methyl protons at the γ-carbon; COSY, two-dimensional correlation spectroscopy; HMQC, heteronuclear multiple-quantum correlation; N, backbone amide nitrogen; N^ϵ, ϵ-nitrogen; $N^{\epsilon 1}$, $\epsilon 1$-nitrogen; $N^\epsilon H$, proton at the ϵ-nitrogen; $N^{\epsilon 1}H$, proton at the $\epsilon 1$-nitrogen; NH, proton at the backbone amide nitrogen; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; RNase H, ribonuclease H; ROE, rotating frame NOE; TOCSY, total correlation spectroscopy; TSP, 3-(trimethylsilyl)tetradeuteriopropionate; TMS, tetramethylsilane; amino acids are denoted by standard one- or three-letter codes.

0006-2960/91/0430-6036$02.50/0