April 20, 1987

The Path of the Polypeptide Backbone of Ribulose-1,5-bisphosphate Carboxylase-Oxygenase from Nicotiana tabacum[1]

David Eisenberg*, Michael S. Chapman, Se Won Suh,
Duilio Cascio, and W.W. Smith

Department of Chemistry and Biochemistry and
Molecular Biology Institute
University of California, Los Angeles
Los Angeles, California 90024, U.S.A.

[1] Supported by NIH grant GM 31299 and USDA grant 83-CRCR-1324 in early stages

[2] Abbreviations:
RuBisCO, ribulose-1,5-bisphosphate carboxylase-oxygenase (EC 4.1.1.39)
L, large subunit;
S, small subunit;
CABP, 2-C-carboxy-D-arabinitol-1,5-bisphosphate
MIR, multiple isomorphous replacement

ABSTRACT

A three-dimensional atomic model for the polypeptide backbone of RuBisCO[2] from Nicotiana tobacum has been determined by x-ray diffraction methods. The 8 large and 8 small subunits are arranged with 422 symmetry in the shape of a keg with an open but narrow central channel along the 4-fold axis of symmetry (1). The large subunit is made up of two major folding domains. The larger of these, the C-domain, is an $(alpha\text{-}beta)_8$ barrel, of the type first found in triosephosphate isomerase and subsequently in more than half a dozen enzymes. The smaller folding domain of the large subunit, the N-domain, contains a twisted beta sheet flanked by alpha helices, and may be considered of the "open-face sandwich" type (2). The remaining electron density of the molecule has been assigned to the small subunits. Four small subunits are clustered about the 4-fold axis of symmetry at high z-value and four others at low z-value. Each set of four small subunits fills the conical crevice formed within a layer of four large subunits, and makes extensive contacts with four barrel domains. This organization suggests that the small subunits may function as a partial scaffold. The mouth of each barrel is "guarded" by an N-domain, in a position that could facilitate a regulatory role. The eight N-terminal domains interdigitate about the plane of the the four 2-fold axes, implying that the upper and lower halves of the molecules are held firmly in register.

As described in other contributions to this symposium as well as elsewhere (e.g. 3), RuBisCO catalyzes the initial steps in both the competing pathways of photosynthesis and photorespiration. This presents a tantalizing challenge for genetic engineering of RuBisCO. Partly to provide a structural

basis for these efforts, and partly to illuminate general features of RuBisCO function and assembly, we have worked for some years to learn the three-dimensional structure of plant RuBisCO. This work has involved four different crystal forms (4-8). Early work on forms I and II showed that RuBisCO's large (L) and small (S) subunits are assembled in a L_8S_8 subunit stoichiometry; it demonstrated 422 molecular symmetry (4), and provided the overall molecular shape (5). X-ray diffraction studies on form III (6,7,8) has progressed to the point that we can report details of the path of the polypeptide backbone in plant RuBisCOs.

METHODS

Crystals, heavy atoms, and x-ray diffraction data. The structure reported here was determined from the form III crystals, grown at high salt and low pH in the absence of CO_2, and Mg^{2+}. Thus the crystals contain RuBisCO in its unactivated state. The space group is I422, with one LS subunit pair in the crystallogrpahic asymmetric unit. The unit cell dimensions are a = 148.7 A; c = 137.5 A. All data were collected by area detector. Data collection, as well as methods of data reduction, heavy atom refinement, and preliminary phase determination have been described elsewhere (1).

X-ray phase determination and improvement. From diffusion experiments with about 200 heavy atom compounds, three acceptable heavy atom adducts were discovered (1). These gave a multiple isomorphous replacement figure of merit of 0.39 for data to 3.1 A resolution. The electron density revealed some helices and extended polypeptide chain, but required improvement for interpretation.

Some improvement came from solvent flattening, using the algorithm of Cascio and Williams (9), a reciprocal space version of the method of Wang

(10).

In this process of phase refinement, we used both automatic and manual masks to outline the current estimate of the surface of the RuBisCO molecule, applied in alternative cycles and later combined in the same cycle. Details will be published elsewhere. The phases changed 39 degrees, and solvent-flattened map was more interpretable than the MIR map.

Model building and refinement. A preliminary polyalanine model (Model 1) of about 540 residues was built into the solvent-flattened electron density over the period December 1985-July 1986. Much of the electron density was represented by about 3300 atoms, but they were grouped into 35 fragments. Following the addition of side chain atoms to fill obvious electron density, Model 1 was refined by D.C. Rees' modification of Hendrickson's (10) restrained least squares method. The goal of this atomic refinement is to move the model into electron density by optimizing the fit of calculated to observed diffraction pattern, and to idealize protein geometry. Substitution of solvent-flattened phases for calculated phases during the initial cycles of refinement helped to ensure rapid convergence from the poor initial model. The R-factor dropped from 48% for the unrefined Model 1 to 31% for the refined model.

At this point Model 2 was built. To prevent bias from Model 1, which enters through phases, an omit map was computed. In an omit map, the unit cell is divided into boxes, each box having its phases determined from the model in all boxes except its own. The map was computed by a local implementation of an algorithm similar to that of Bhat and Cohen (12), which will be described elsewhere. Model 2 was built with the aid of both the omit map and a map with the coefficients proposed by Stuart and Artymiuk (13). The omit map is less biased by the current model, but in some places it showed

little electron density, and the Stuart and Artymiuk map was found more useful. In connecting the 35 fragments of Model 1 with protein-like topology, we took advantage of the similarity of our electron density in the C-domain of the L subunit to the beta barrels of pyruvate kinase and triosephosphate isomerase, and in the N-domain to the proposed alpha carbon diagram of Schneider et al. (14).

Model 2 contains 3900 atoms. Side chains known from the sequence (15-17) have been built into part of the C-domain. The model is being refined, the R factor currently being about 34% at 3.5 A resolution. Our plan is to continue refinement and to construct Model 3, with full side chains.

RESULTS

Tertiary structures of the polypeptide chains. A cluster of six alpha helices of our Model 1 were found to superimpose on the alpha carbon backbone of the alpha/beta barrel in pyruvate kinase (18). Further inspection of Map 1 revealed extended polypeptide chains connecting these helices in the manner of other alpha/beta barrels. A complete alpha/beta barrel was built, which has the C-termini of the beta chains pointing toward the solvent.

Adjacent to this barrel domain, and partially occluding the solvent entrance to the barrel is the N-domain. In this deomain, several helices pack against the barrel, and a five-stranded antiparallel beta sheet faces the solvent. As is discussed below, the N-domain which packs against the C-domain barrel is not part of the same L subunit. Rather this N-domain is is probably connected covalently to a C-domain in another L subunit.

The remaining electron density of the crystallographic protomer must be the S subunit. In our model it is composed of mainly beta strands and loops

with one prominent alpha helix.

Quaternary structure of the RuBisCO molecule. Four S subunits cluster together around the 4-fold axis, both at high z-value and at low z-value. These two sets of S subunit tetramers are separated by more than 30 A along the z-axis (Fig. 1), and thus make no contact with each other. The intervening space is occupied by the L subunits.

When viewed down the 4-fold axis, the RuBisCO molecule is seen (Fig. 2) to have four small subunits (in blue) grouped about the axis, enjoying contacts with each other. It is noteworthy that tetrameric aggregates of S subunits, located at the top and bottom of the RuBisCO molecule, were suggested more than 10 years ago by Bowien et al. (19) from electron micrographs of RuBisCO from Alcaligenes eutrophus. This aspect of their subunit-level model is consistent with our atomic-level model, although certain other aspects (such as a four distinct levels of subunits) are not. Contiguous small subunits were also correctly inferred by Roy et al. (20) on the basis of chemical cross-linking.

Four C-domains from four L subunits fill a crevice under the tetramer of S subunits at high z. Four other C-domains fill the corresponding crevice on top of the other tetramer of S subunits at low z. One C-domain binds in a staggered fashion between each pair of S subunits. These are shown in green in Fig. 2, and it can be seen that they make extensive contacts with the S subunits. Shown in red in the fugure are four N-domains. These domains protrude farther from the 4-fold axis than any other molecular features, and are represented by the ridges at the sides of our wooden keg model of the 5 A-resolution electron density map (1). In our present model these red N-domains are covalently connected to the four C-domains in the other half (lower z region) of the molecule, not to the green C-domains in the same (high

z). The connection between the two domains is not yet certain, and the exact connectivity must await further model building and refinement. The connection in our present model is consistent with the overall shape of the model for R. rubrum RuBisCO given by Schneider et al. (14).

Figures 3 and 4 show combinations of S and L subunits. Whereas Figs. 1 and 2 are views down the 4-fold axis. Figs. 3 and 4 are views down the 2-fold along y, with the 4-fold (z-axis) vertical. Fig. 3 is a partial RuBisCO molecule, with four blue S subunits at the top around the z-axis, and two C-domains in green, and two N-domains in yellow. The view of the darker green C-domain to the right is directly into the barrel axis, and the central cavity in this alpha carbon model appears open. In analogy to other alpha/beta barrel enzymes, and in accord with our tentative fitting of part of the amino acid sequence into the electron density map, this cavity is presumably the active site. The view of the lighter green C-domain to the left is at right angles to the barrel axis, which runs roughly horizontal. Both barrels open to the solvent at their faces, which are formed from the C-terminal ends of the eight beta strands.

In Fig. 3 the two yellow N-domains are connected to the adjacent C-domains in much the same way at connection proposed for the R. rubrum RuBisCO by Schnieder et al. (14). Notice by comparison with Fig. 4 that the covalent connection between these domains of the L subunits extends across the central (z=0) plane of the molecule. This is also true of the other six L subunits, four of them having their C-domains at the bottom (negative z).

Fig. 4 shows the front half of an entire RuBisCO model in the same orientation as Fig. 3. Added to the subunits of Fig. 3 are two turquoise C-domains at the bottom and two red N-domains just above the z=0 plane. The profile exhibits the keg shape of the low resolution model, with the N-domains

extending farthest to the sides (high and low x-values). These protrusions were modeled at low resolution as small spheres in the drawings of Baker et al. (5,6). Notice also that the C-domains extend almost as high and low along the z-axis as the S subunits.

DISCUSSION

Similarity of RuBisCO to other enzymes. The complexity of the arrangement of RuBisCO subunits tends to overshadow the fact that the individual domains bear similarities to protein structures already known. This is particularly true of the C-domain with its $(alpha/beta)_8$ barrel structure. This folding motif was first seen in triosephosphate isomerase (21), and now has been found also in pyruvate kinase (18) and half a dozen other enzymes. In all these enzymes, the active site is at the mouth of the barrel at the C-termini of the beta strands. The N-terminal domain has the feature of antiparallel beta structure found in many protein domains (2).

Comparison of the structure of RuBisCO from Nicotiana tobacum to that of other RuBisCOs. The structures of RuBisCos from several other species are under study in various laboratories (22-28). Presently the most advanced of these is that of R. rubrum from Brandin's laboratory in Upsalla. An alpha carbon backbone model has been proposed for this L_2 structure. The electron density map of tobacco RuBisCO is consistent with an L subunit that resembles this proposed polypeptide folding. In Fig. 3, the C- and N-domains have been connected to give an elongated L subunit with much the same overall shape as the structure proposed for the subunit of R. rubrum RuBisCO. At the present stage of refinement of the tobacco structure, other connections cannot be ruled out.

Earlier it was anticipated that the structures of the L subunit and of R. rubrum would be similar (26). This prediction was based on the correlation of amino acid hydrophobicities of the two polypeptide chains. Computer alignment of the two amino acid sequences suggest that about one-third of the residues in the two chains are identical. However, the hydrophobicities of even the non-identical residues show a very significant positive correlation, suggesting that the two proteins are folded the same way. This suggestion is strengthened by the findings of this paper.

Comparison must also be made of the present model and the R. rubrum model with the model proposed for RuBisCO from Alcaligenes eutrophyus (27). Holzenburg et al. have reported a massive conformational change for Alcaliglenes RuBisCO as it binds CABP, the transition-state analog. Their report is based on 5 A resolution x-ray data, with phases from single isomorphous replacement, for CABP-complexed RuBisCO and 17 A resolution electron microscopy data for the presumably activated form. Their proposal is that upon binding CABP, the high z and low z halves of the RuBisCO molecule each maintain their 4-fold symmetry, but that the upper half molecule is translated relative to the lower half by 36 A in a direction normal to the 4-fold axis, so that the two local 4-fold axes are separated by 36 A. From our model it appears that such a translation would be impeded by the tight interdigitation of the 8 N-domains through the plane z=0. Moreover, if the N and C-domains are connected as suggested by the R. rubrum model, the proposed translation would require either that the eight N-C connections are broken, or that the polypeptide chains unwind in some way to permit the movement.

These difficulties in reconciling the various models will undoubtedly disappear as more progress in atomic refinement is made. At UCLA we have grown an activated crystal form of tobacco RuBisCO (8) and we plan to compare

the structure of RuBisCO molecules in this form to the structure we observe in the form III unactivated crystals.

References

1. Chapman, M.S., W.W. Smith, S.W. Suh, D. Cascio, A. Howard, R. Hamlin, N.-h. Xuong, and D. Eisenberg, 1986 Structural studies of RuBisCO from tobacco. Phil. Trans. Roy. Soc. B313:367-378.

2. Richardson, J. 1981 The anatomy and taxonomy of protein structure. Adv. Protein Chem. 34:167-339.

3. Miziorko, H.M., and G.H. Lorimer 1983 Ribulose-1,5-bisphosphate carboxylase/oxygenase. Ann. Rev. Biochem. 52:507-535.

4. Baker, T.S., D. Eisenberg, F.A. Eiserling, and L. Weissman, 1975 The structure of form I crystals of D-Ribulose-1,5-Bisphosphate Carboxylase. J. Mol. Biol. 91:391-399.

5. Baker, T.S., D. Eisenberg, and F.A. Eiserling, 1977a Ribulose Bisphosphate Carboxylase: A two-Layered, square-shaped molecule of symmetry 422. Science 196:293-295.

6. Baker, T.S., S.W. Suh, and D. Eisenberg 1977 Structure of ribulose-1,5-bisphosphate carboxylase-oxygenase: Form III crystals. Proc. Natl. Acad. Sci. USA 74:1037-1041.

7. Eisenberg, D., T.S. Baker, S.W. Suh, and W.W. Smith 1978 Structural studies of Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase in photosynthetic carbon assimilation (H.W. Siegelman, and G. Hind, eds.) pp. 271-281, Plenum Press, New York.

8. Suh, S.W., D. Cascio, M.S. Chapman, and D. Eisenberg 1987 A crystal form of ribulose-1,5-bisphosphate carboxylase-oxygenase from Nicotiana tabacum in the Activated State. J. Mol. Biol. Submitted.

9. Cascio, D., and R. Williams 1987 In preparation.

10. Wang, B.C. 1985 Resolution of phase ambiguity in macromolecular crystallography. Methods Enzymol. 115:90-117.

11. Hendrickson, W.A. 1985 Stereochemically restrained refinement of macromolecular structure. Methods Enzymol. 115:252-270.

12. Bhat, T.N. and G.H. Cohen 1984 Omitmap: an electron density map suitable for the examination of errors in a macromolecular model. J. Appl. Cryst. 17:244-248.

13. Stuart, D., and P. Artimuik 1984. The use of phase combination in crystallographic refinement: the choice of amplitude coefficients in combined syntheses. Acta Cryst. A40:71-716.

14. Schneider, G., Y. Lindqvist, C.-I. Branden, and G. Lorimer 1986. Three-dimensional structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from Rhodospirillum rubrum at 2.9 A resolution EMBO J. 5:3409-3415.

15. Shinozaki, K., and M. Sugiura 1982 The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Gene 20:91-102.

16. Muller, K.-D., J. Salnikow, and J. Vater 1983 Amino acid sequence of the small subunit of D-ribulosebisphosphate carboxylase/oxygenase from Nicotiana tabacum. Biochim. Biophys. Acta 742:78-83.

17. Amiri, I., J. Salnikow, and J. Vater 1984 Amino-acid sequence of the large subunit of D-ribulosebisphosphate carboxylase/oxygenase from Nicotiana tabacum. Biochim. Biophys. Acta 784:116-123.

18. Muirhead, H., D.A. Clayden, D. Barford, C.G. Lorimer,

L.A. Fothergill-Gilmore, E. Schiltz, and W. Schmitt. 1986 The structure of cat muscle pyruvate kinase. EMBO J. 5:475-481.

19. Bowien, B., F. Mayer, G.A. Codd, and H.G. Schlegel 1976 Purification, some properties and quaternary structure of the D-ribulose-1,5-diphosphate carboxylase of Alcaligenes eutrophus Arch. Microbiol. 110:157-166.

20. Roy, H., A. Valer, D.H. Pope, L. Rueckert, and K.A. Costa 1978 Small subunit contacts in ribulose-1,5-bisphosphate carboxylase. Biochemistry, 17:665-668.

21. Petsko, G.A., D.C. Phillips, R.J.P. Williams, and I.A. Wilson 1978 On the problem of crystal chemistry of chloroplatinite ions: General principles of interactions and interactions with triose phosphate isomerase. J. Mol. Biol. 120:345-359.

22. Andersson, I., A.-C. Tjader, E. Cedergren-Zeppezauer, and C.-I. Branden 1983 Crystallization and preliminary x-ray studies of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase complexed with activator and a transition state analogue. J. Biol. Chem. 258:14088-14090.

23. Barcena, J.A., R.W. Pickersgill, M.J. Adams, D.C. Phillips, and F.R. Whatley 1983 Crystallization and preliminary x-ray data of ribulose-1,5-bisphosphate carboxylase from spinach. EMBO J. 2:2363-2367.

24. Bowien, B., F. Mayer, E. Spiess, A. Pahler, U. Englisch, and W. Saenger 1980 On the structure of crystalline ribulosebisphosphate carboxylase from Alcaligenes eutrophus. Int. J. Biochem. 106:405-410.

25. Nakagawa, H., M. Sugimoto, Y. Kai, S. Harada, K. Miki, N. Kasai, K. Saeki, T. Kakuno, and T. Horio 1986 Preliminary crystallographic study of a ribulose-1,5-bisphosphate carboxylase-oxygenase from Chromatium vinosum

J. Mol. Biol. 191:577-578

26. Janson, C.A., W.W. Smith, D. Eisenberg, and F.C. Hartman 1984 Preliminary structural studies of ribulose-1,5-bisphosphate carboxylase/oxygenase from Rhodospirillum rubrum. J. Biol. Chem. 259:11594-11596.

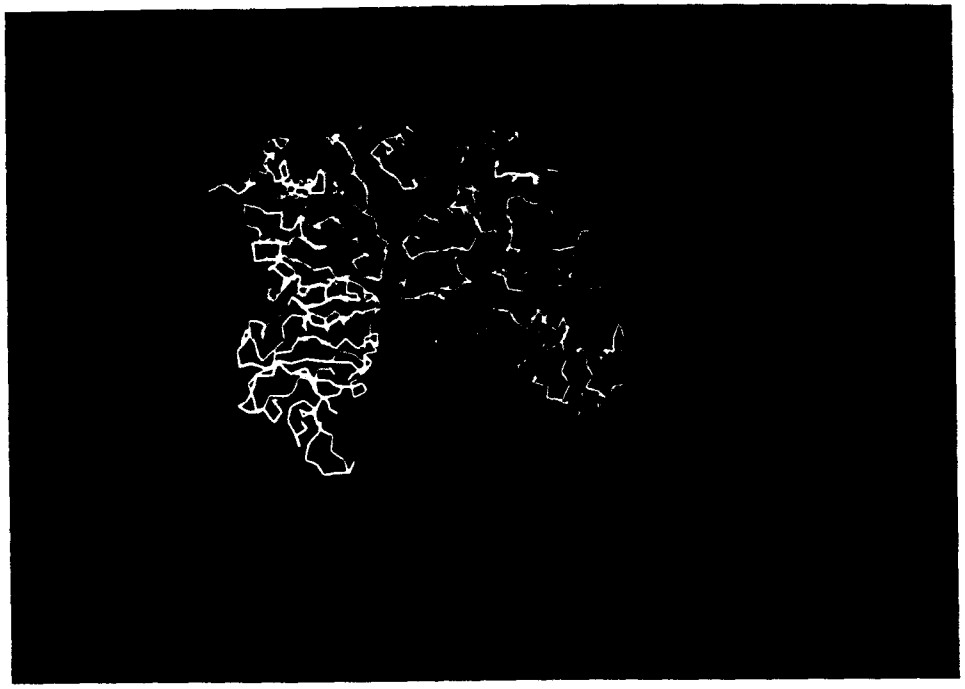
27. Schneider, G., C.-I. Branden, and G. Lorimer 1984 Preliminary x-ray diffraction study of ribulose-1,5-bisphosphate carboxylase from Rhodospirillum rubrum. J. Mol. Biol. 175:99-102.

Figure 1. The eight small subunits. The protein chains in this and the following figures are represented by line segments that link alpha carbon atoms. The z (4-fold) axis is vertical. Four small subunits are clustered about the 4-fold axis at high z-values and four at low z. The space between the two clusters is occupied by the large subunits.

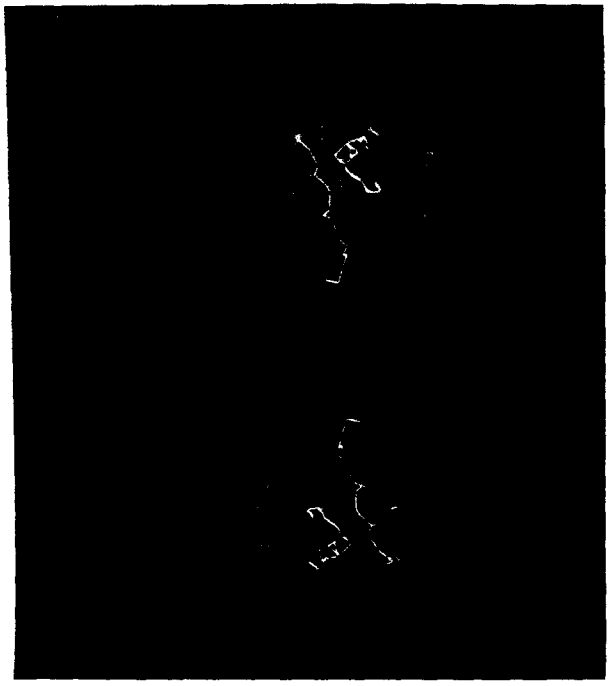
Figure 2. The RuBisCO molecule viewed down the z (4-fold) axis. Four small subunits are shown in blue, clustered about the central channel. Four C-domains are shown in green and four N-domains are in red.

Figure 3. A partial RuBisCO molecule viewed down the y-axis, with the z-axis vertical. Four S subunits are shown in blue at the top. Two large subunits are shown with green C-domains and yellow N-domains. The N-domain to the right is viewed down the barrel axis. The active site region is visible as the opening to the barrel.

Figure 4. The RuBisCO molecule viewed down the y-axis, with the z-axis vertical. The two clusters of S subunits are in blue, in the same orientation as they are in Fig. 1. Four C-domains are visible, in shades of green, next to the S subunits. Four N-domains are visible, two in red and two in yellow.



3



1

4

2