Computation of Structures of Homologous Proteins. α-Lactalbumin from Lysozyme†

P. K. Warme,‡ F. A. Momany,§ S. V. Rumball,¶ R. W. Tuttle, and H. A. Scheraga*

ABSTRACT: The amino acid sequence of bovine α-lactalbumin has been shown to be remarkably similar to that of hen egg-white lysozyme, leading to the suggestion that the three-dimensional structures of these proteins are also similar. Browne *et al.* constructed a wire skeletal model of α-lactalbumin based on the X-ray structure of lysozyme and demonstrated that the amino acid sequence of α-lactalbumin is compatible with the assumption of similar structures for these two proteins. A considerable number of experimental comparisons of α-lactalbumin and lysozyme have also substantiated this conclusion. We have adopted our procedures for refinement of X-ray coordinates of proteins to compute a low-energy conformation for α-lactalbumin based on the known structure of lysozyme. In order to select appropriate conformations for the side chains of nonhomologous amino acids, we have also made use of recent conformational energy calculations on oligopeptides. Using this approach we have computed a low-energy conformation for α-lactalbumin which is quite similar to the X-ray structure of lysozyme. This computed structure for α-lactalbumin is consistent with most of the available experimental evidence.

In recent years, it has become apparent from comparisons of the amino acid sequences of a large number of proteins that many proteins have homologous sequences and may therefore have similar three-dimensional structures (Brew *et al.*, 1967; Brew and Campbell, 1967; Hartley, 1970; Dickerson, 1971). The computation of similar helix- and bend-probability profiles for homologous proteins (Lewis and Scheraga, 1971; Lewis *et al.*, 1971) supports this view. It is therefore of interest to adapt the procedures recently developed for the refinement of X-ray structures by conformational energy minimization (Warme *et al.*, 1972; Warme and Scheraga, 1973, 1974) to the computation of the structure of a protein based on the X-ray structure of a homologous protein. In considering the homologies in amino acid sequence among related proteins, the maximum extent of homology becomes apparent only by proper alignment of the sequences of the various proteins; *i.e.*, insertions and/or deletions of amino acids must be introduced in comparing the various sequences. Such insertions and deletions must be taken into account in computing a structure for the *complete* homologous protein, and recent conformational energy calculations on oligopeptides (Lewis *et al.*, 1973a,b) provide a basis for incorporating these effects.

This initial paper is devoted to the computation of the structure of bovine α-lactalbumin from that of hen egg-white lysozyme (Blake *et al.*, 1967). These two proteins have striking sequence homologies (Brew *et al.*, 1967; Brew and Campbell, 1967). After allowing for several deletions and insertions, 47 of the 123 amino acids in α-lactalbumin, including the eight half-cystine residues, occur in identical positions in the two proteins, and three other corresponding pairs are Asn-Asp or Gln-Glu differences. Many other corresponding residues, though different, are similar in their polar or nonpolar nature. Many similarities between these two proteins which strongly indicate structural homology have been observed by a variety of experimental approaches, as outlined in the Discussion. Browne *et al.* (1969) constructed a wire skeletal model of α-lactalbumin which closely approximates the known structure of lysozyme (Blake *et al.*, 1967). We have used the approximate backbone dihedral angles of Browne *et al.* (1969) for α-lactalbumin, corrected the sequence in those places where the identities of the amino acids were uncertain at the time of their work, taken account of insertions and deletions, and applied our geometrical fitting procedure (stage I refinement) (Warme *et al.*, 1972) to fit the α-lactalbumin sequence to the X-ray coordinates of the backbone atoms of lysozyme (Phillips, 1970), and thereby have obtained an initial backbone conformation. The side chains were then added and their conformations were obtained by methods to be described in section A2. Selected segments (especially in the regions of insertions and deletions) were then modified by utilizing information about the probabilities of occurrence of helical and bend conformations (Lewis and Scheraga, 1971; Lewis *et al.*, 1971) and, as a final step, our stage II energetic refinement procedure (Warme and Scheraga, 1973) was applied. A conformation of low energy was thus achieved.[1] We report here the techniques

† From the Department of Chemistry, Cornell University, Ithaca, New York 14850. *Received June 29, 1973.* This work was supported by research grants from the National Science Foundation (GB-28469X2), from the National Institute of General Medical Sciences of the National Institutes of Health, U. S. Public Health Service (GM-14312), from the Eli Lilly Grants Committee, and from Walter and George Todd.

‡ Postdoctoral Fellow of the National Institute of General Medical Sciences, National Institutes of Health, 1969–1971. Present address: Department of Biochemistry, The Pennsylvania State University, University Park, Pa. 16802.

§ Special Fellow of the National Institute of General Medical Sciences, National Institutes of Health, 1968–1969.

¶ American Association of University Women International Fellow, 1970–1971. Present address: Department of Chemistry, Biochemistry and Biophysics, Massey University, Palmerston North, New Zealand.

[1] As pointed out in the accompanying paper (Warme and Scheraga, 1974), it is possible to carry the computations to an additional stage of refinement (*i.e.*, through stage III). However, we will not be able to carry out the stage III refinement for α-lactalbumin in the immediate future because of the present shortage of research grant funds. In any event, from the experience with lysozyme (Warme and Scheraga, 1974), it is probable that the atomic coordinates reported here after stage II will change very little when the stage III refinement is applied, although the energy will probably decrease substantially. In other words, the final structure reported here corresponds to the correct local minimum of the potential energy well, although small conformational fluctuations within this minimum are to be expected upon further energy minimization.

used in the computation, a description of the low-energy structure, and a correlation of the computed structure with experimental evidence relating to the structure. The nomenclature used is that of a recent IUPAC–IUB Commission (1970). Throughout this manuscript, the numbering of the amino acids in the sequence of lysozyme is given in bold-face type and that of α-lactalbumin in italics.

Procedure

A. Acquisition of a Starting Conformation (Stage I). In our computations, the bond lengths and bond angles are held fixed at the values observed in crystal structures of amino acids and peptides (Momany *et al.*, 1973),[2] *i.e.*, each type of amino acid has its own geometry. In addition, the amide groups are restricted to a planar trans conformation. The fixed Cartesian coordinate system is established for the first amino acid as described by Warme and Scheraga (1974). The conformation of the protein can then be altered by rotation about the single bonds of the backbone and side chains. In the techniques used here, it is essential to first acquire a conformation for the entire polypeptide chain utilizing our standard bond lengths, bond angles, and planar trans amide groups before the conformational energy of the protein can be computed. This was accomplished as follows.

(1) *Computation of a Starting Conformation for the Backbone.* Starting with our amino acid geometry and the backbone dihedral angles of Browne *et al.* (1969) for α-lactalbumin, a full chain was generated by adjusting these dihedral angles to obtain a least-squares best fit of the computed coordinates of the backbone atoms to the X-ray coordinates of lysozyme (Phillips, 1970). This was accomplished with a modification of our stage I refinement procedure (Warme *et al.*, 1972), using the minimization procedures described therein; *i.e.*, the dihedral angles of α-lactalbumin were adjusted to minimize the function

$$F = \sum_i \kappa_i \delta_i^2 + \sum_j \lambda_j \alpha_j^2 \qquad (1)$$

In this equation, δ_i is the deviation of a computed atomic coordinate of α-lactalbumin from the corresponding X-ray coordinate of lysozyme, α_j is the deviation of a computed dihedral angle of α-lactalbumin from the initial value for the same protein as taken from Browne *et al.* (1969), and κ_i and λ_j are weighting factors designed to attain a proper balance between the two constraints (*i.e.*, to the X-ray coordinates of lysozyme and to the initial dihedral angles of α-lactalbumin). The range of i is over all fitted backbone atoms (N, C^α, C^β, C′, and O) with κ_i taken as 10 kcal/Å^2, and the range of j is over all fitted backbone dihedral angles (ϕ,ψ) with λ_j taken as 80 kcal/rad^2 for all fitted residues up to *86* in α-lactalbumin and as zero for residues *87–111* because the dihedral angles of Browne *et al.* (1969) did not provide a sufficiently close fit to the X-ray coordinates in the latter region. The adjustment of the backbone conformation of residues *112–123* (undefined by Browne *et al.* (1969)) was carried out by a different procedure, which will be described in section B2. Special treatment was required in the regions of deletions of certain amino acids from lysozyme (residues **14**, **15**, **47**, **100**, **117**, **125**, and **126**) and in the region of an insertion (between residues **127** and **128**) (Figure 1). In order to provide sufficient freedom for the backbone chain to adjust and remain continuous in these regions of deletions and insertions, (*i.e.*, to enable the backbone coordinates to deviate from those of lysozyme), κ_i was taken as zero for all atoms of one residue preceding each deletion and also for the two residues following each deletion. The primary reason for including the second term of eq 1 was to avoid excessively large movements (during stage I) of the atoms which were unconstrained in some regions by the first term (because $\kappa_i = 0$); *i.e.*, the constraint that the dihedral angles in these regions should remain similar to the angles of Browne *et al.* (1969) was imposed. It should be recognized that the first and second terms of eq 1 are not necessarily independent in all cases, since the angles reported by Browne *et al.* (1969) were also obtained by adjusting their model to fit the X-ray coordinates of lysozyme. Thus, in regions where α-lactalbumin and lysozyme are apparently similar in conformation, both terms of eq 1 will tend to maintain the close approximation to the X-ray coordinates. It should also be noted that the second term of eq 1 was omitted after completion of the stage I procedure, in order to allow all dihedral angles to change in response to the energy terms introduced in stage II (see below).

Equation 1 was applied to 20-residue segments of the backbone, chosen so that there was an overlap of two residues between segments to avoid distortions arising from end effects, as described previously (Warme *et al.*, 1972). Later, smaller segments were treated independently to take into account information available on the preferences for chain reversals and other types of bends (Lewis *et al.*, 1971) among particular amino acid sequences; these regions are discussed separately in section C.

```
                    5                   10
A:   Lys-Val-Phe-Gly-Arg-CYS-GLU-Leu-Ala-Ala-Ala-Met-LYS Arg-His-
B:   Glu-Gln-Leu-Thr-Lys-CYS-GLU-Val-Phe-Arg-Glu-Leu-LYS-XXX-XXX-
                    5                   10

                   20                  25                  30
A:  -Gly-LEU-Asp-Asn-TYR-Arg-GLY-Tyr-SER-LEU-Gly-Asn-TRP-VAL-CYS-
B:  -Asp-LEU-Lys-Gly-TYR-Gly-GLY-Val-SER-LEU-Pro-Glu-TRP-VAL-CYS-
       15                  20                  25

                   35                  40                  45
A:  -Ala-Ala-Lys-Phe-Glu-SER-Asn-Phe-Asn-THR-Gln-ALA-Thr-Asn-Arg-
B:  -Thr-Thr-Phe-His-Thr-SER-Gly-Tyr-Asp-THR-Glu-ALA-Ile-Val-Glu-
       30                  35                  40

                   50                  55                  60
A:  -ASN-Thr-Asp-Gly-SER-THR-ASP-TYR-GLY-Ile-Leu-GLN-ILE-ASN-Ser-
B:  -ASN-XXX-Asn-Gln-SER-THR-ASP-TYR-GLY-Leu-Phe-GLN-ILE-ASN-Asn-
          45                  50                  55

                   65                  70                  75
A:  -Arg-Trp-TRP-CYS-Asn-Asp-Gly-Arg-Thr-PRO-Gly-SER-Arg-ASN-Leu-
B:  -Lys-Ile-TRP-CYS-Lys-Asn-Asp-Gln-Asp-PRO-His-SER-Ser-ASN-Ile-
          60                  65                  70

                   80                  85                  90
A:  -CYS-ASN-ILE-Pro-CYS-Ser-Ala-Leu-LEU-Ser-Ser-ASP-Ile-THR-Ala-
B:  -CYS-ASN-ILE-Ser-CYS-Asp-Lys-Phe-LEU-Asn-Asn-ASP-Leu-THR-Asn-
          75                  80                  85

                   95                 100                 105
A:  -Ser-Val-Asn-CYS-Ala-LYS-LYS-ILE-Val-Ser-ASP-Gly-Asp-GLY-Met-
B:  -Asn-Ile-Met-CYS-Val-LYS-LYS-ILE-Leu-XXX-ASP-Lys-Val-GLY-Ile-
          90                  95                 100

                  110                 115                 120
A:  -ASN-Ala-TRP-Val-ALA-Trp-Arg-Asn-Arg-CYS-Lys-Gly-Thr-Asp-Val-
B:  -ASN-Tyr-TRP-Leu-ALA-His-Lys-Ala-Leu-CYS-Ser-XXX-Glu-Lys-Leu-
         105                 110                 115

                  125
A:  -Gln-Ala-TRP-Ile-Arg-Gly-CYS-XXX-Arg-LEU
B:  -Asp-Gln-TRP-Leu-XXX-XXX-CYS-Glu-Lys-LEU
                  120
```

FIGURE 1: Comparison of the amino acid sequences of lysozyme (A) (Rees and Offord, 1972) and α-lactalbumin (B) (Brew *et al.*, 1970). Identical amino acids are printed in capital letters. XXX denotes a deletion.

[2] Data are also given in Lewis *et al.* (1973b).

(2) *Computation of a Starting Conformation for the Side Chains.* The dihedral angles of the side chains of the residues which were identical with those of lysozyme were adjusted first, to fit the X-ray coordinates of lysozyme, using the fitting procedure described earlier (Warme *et al.*, 1972). The remaining side chains were treated in one of the following two ways to obtain *initial* conformations.

In the first type of treatment, those residues of α-lactalbumin that had some side-chain atoms which could be superimposed on corresponding atoms of a similar, though different, type of side chain in lysozyme were positioned. For example, Phe-3 of lysozyme is replaced by Leu-*3* in α-lactalbumin. Thus, χ_1 of Leu was determined initially by assuming that the C^γ atoms of both residues coincide, and χ_2 by assuming approximate coincidence of the C^δ atoms; the fit to χ_2 is only approximate because $\tau(C^\beta C^\gamma C^\delta)$ is 109.5° in Leu, although it is 120° in Phe.

The second type of treatment, which was applied to the remaining side chains, made use of the tables of minimum-energy conformations of "dipeptides" of all the naturally occurring amino acids (Lewis *et al.*, 1973b), in order to select suitable initial side-chain conformations (compatible with the already determined backbone conformations). In cases where more than one side-chain conformation is possible, the lowest energy "dipeptide" conformation was chosen in stage I (see section B1 for alternate choices of side-chain dihedral angles).

B. Application of the Stage II Energy Minimization. The stage II procedure, in which the nonbonded, hydrogen bond, rotational energies, and fitting energies are minimized, is that described by Warme and Scheraga (1973); electrostatic and solvent energies are omitted for reasons discussed in the reference cited.[1] The energy functions, parameters, and methods are those described by Warme and Scheraga (1973, 1974). In subsection 1, the side chains are adjusted one at a time while holding the backbone conformation fixed, whereas in subsections 2 and 3, the backbone and side-chain conformations are adjusted simultaneously in 20-residue segments.

(1) *Initial Adjustment of Side Chains to Optimize Hydrogen Bonds and to Relieve Atomic Overlaps.* After generating the α-lactalbumin molecule by the stage I procedure (described in section A), some polar side chains still had poor hydrogen bonds and some side chains were involved in atomic overlaps resulting in very high nonbonded energies. In order to optimize the side chain hydrogen-bond energy and also to relieve these overlaps, an initial optimization was carried out, residue by residue, in a specific series of steps designed to preserve a close correspondence to the conformation of lysozyme. The first group of residues whose conformations were initially altered [for reasons cited elsewhere (Warme and Scheraga, 1973)] consisted of those which can participate in hydrogen bonding, *viz.*, serine, threonine, tyrosine, asparagine, and glutamine. The side-chain dihedral angles of these residues were adjusted to optimize the hydrogen-bond energy of each of these residues. The remaining large steric overlaps were then relieved by adjusting the responsible side chains one at a time in the following order from (a) to (f): (a) side chains which were not homologous to those in lysozyme, and had energies greater than 1000 kcal; (b) side chains which differed from those of lysozyme (or, if identical with those of lysozyme, exhibited deviations of greater than 1 Å in the coordinates of homologous atoms), and had energies between 100 and 1000 kcal; (c) side chains which differed from those of lysozyme, and had energies less than 100 kcal; (d) side chains which differed from those of lysozyme but had some superimposable homologous atoms (*e.g.*, C^γ of Leu and Phe), and the initial fit to the lysozyme side chain exhibited deviations in coordinates greater than 1 Å; (e) side chains of the type in category (d), but for which the fit to the lysozyme side chain was close (*i.e.*, <1 Å); (f) side chains which were identical with those of lysozyme, but for which the fit to the lysozyme side chain was poor (*i.e.*, >1 Å).

It should be noted that the progression from category (a) to (f) is from high energy to low energy and also from distant homology to close homology. By this procedure, it was possible to maintain a close correspondence to those parts of the lysozyme structure which were most similar in both proteins. In treating each side chain, starting at those side chains in category (a) above, all minimum-energy side-chain conformations of single residues (for the already determined backbone conformation) having energies within 3 kcal/mol of the lowest energy dipeptide conformation (Lewis *et al.*, 1973b) were considered in turn. If none of the dipeptide conformations reduced the energy by at least 100 kcal/mol, the conformation found to be of lowest energy was then altered in 20° increments from the previous lowest energy conformation until a reduction of at least 100 kcal could be achieved. When a low-energy conformation was found, it was fixed before proceeding to the next residue in the given category. This approach was utilized for all amino acid side chains in all six categories. Since side chains were adjusted one at a time by this procedure, it is possible that some low-energy conformations could be missed, in particular, those which would require simultaneous changes in two or more side-chain conformations. For this reason, the adjustments of all residues within a given category were carried out a second time before proceeding to the next category. As a result, only one side chain in ten underwent further change in conformation in the second cycle, which suggests that in most cases the side chains being altered were sufficiently widely separated to move independently, without affecting other side chains undergoing conformational readjustments.

(2) *Initial Conformation of the C-Terminal Residues 112–123.* There are three deletions, one insertion, and a general lack of homology to lysozyme in the region of residues **116–129** of lysozyme, and, consequently, Browne *et al.* (1969) did not assign a conformation to this portion of the chain in α-lactalbumin. For this reason, the initial conformation of this C-terminal sequence was selected by the following procedure which bypasses stage I. The conformation of the backbone of residues *112–118* of α-lactalbumin was initially constructed to be α helical [since the sequence in this section has a high helix probability (Lewis and Scheraga, 1971), and in fact residues **120–124** in lysozyme are nearly helical] and the side-chain dihedral angles of this section were placed in the conformation of lowest energy (Lewis *et al.*, 1973b) for a right-handed α helix. The nonbonded energy of this seven-residue segment was then minimized in the protein (by the procedure of section B3), subject to the constraints (1) that the backbone atoms of residue *115–118* must fit (using eq 1) the corresponding backbone atoms of residues **120–123** of lysozyme and (2) that the side chain of Trp-*118* of α-lactalbumin must coincide with that of Trp-**123** of lysozyme [constraint (2) was later removed in the minimization described in section B3]. It should be noted that, for maximum homology, Gln-*117* and Trp-*118* of α-lactalbumin may be paired with Gln-**121** and Trp-**123** of lysozyme, with a deletion in α-lactalbumin corresponding to Ala-**122** of lysozyme. However, we feel that the limited homology in this region precludes any definite sequence correlations between α-lactalbumin and lysozyme, and we have thus

correlated residues *115–118* directly (without any deletions or insertions) with residues **120–123** of lysozyme. During the course of the conformational adjustment of residues *112–123*, residues *112* and *113* departed from the α-helical conformation and became more extended, but residues *114–118* remained approximately helical. Residues *119–123* were omitted during this preliminary energy minimization.

When the attempt was made to extend the helix to include the remaining residues, *119–123*, it was found that the residues beyond Cys-*120* overlapped severely with other parts of the protein whenever the disulfide bond length was reasonable. Furthermore, in spite of numerous attempts to find a conformation for residues *121–123* which would allow only residues *119* and *120* to remain helical while avoiding overlaps of the remaining residues, our search was unsuccessful. For this reason, a large number of favorable but nonhelical conformations (Lewis *et al.*, 1973b) for residues *119* and *120* were examined in order to select those which would allow closure of the disulfide bond. Several allowed conformations for residues *119* and *120* were found in this way. Starting from each of these conformations, the chain was then elongated to include the residues *121–123*. Again, many combinations of favorable conformations for residues *121–123* were obtained from the list of favorable dipeptide conformations (Lewis *et al.*, 1973b). Although it would have required an excessive amount of computer time to exhaustively search all possible conformations for residues *119–123*, we were able to find one conformation which adhered to all of the constraints outlined above. The nonbonded energy of this C-terminal segment (residues *119–123*) was then subjected to a preliminary (10-cycle) energy minimization according to the procedure of section B3. Several other alternative conformations for residues *121–123* were later found by the methods outlined in section D.

(3) *Minimization of Energy of Twenty-Residue Segments of α-Lactalbumin.* At this point, a complete conformation of α-lactalbumin was available from the procedures described previously, but a considerable number of high-energy steric overlaps were still present. The energy of this conformation was minimized by our stage II procedure (Warme and Scheraga, 1973), in which the sum of the nonbonded (E_{NB}), hydrogen bond (E_{HB}), rotational (E_{ROT}), disulfide loop-closing (E_{SS}), and fitting energies (E_{FP}) is minimized by allowing simultaneous variation of backbone and side-chain dihedral angles in 20-residue segments. Even though the dihedral angles of only 20 residues are varied at a time, the interactions between all interacting pairs of atoms in the complete protein are computed as described by Warme and Scheraga (1973). The overlapping segments chosen for this stage of refinement were *1–20*, *19–38*, *37–56*, *55–74*, *73–92*, *91–110*, and *109–123*. In using eq 1 (fitting energy) to maintain a close fit to the X-ray coordinates of lysozyme, κ_i was taken as zero for the backbone atoms near deletions (residues **14**, **15**, **47**, **100**, **117**, **125**, and **126**) and also for side-chain atoms whose positions had deviated by more than 1 Å from the corresponding X-ray coordinates of lysozyme during the preliminary optimization of the side-chain dihedral angles (described in section B1). For all other atoms, κ_i was taken as 50 kcal/Å^2, and λ_j was taken as zero (see discussion in section A1) for *all* dihedral angles at this stage of refinement. Each 20-residue segment was carried through 10 cycles of minimization with these constraints and then, after completing these calculations on all segments, another 10 cycles of minimization were performed with κ_i taken as 10 kcal/Å^2.

C. Adjustment of Structure to Eliminate High-Energy Local Minima. After the energy minimization described in section B3, the resulting structure was examined residue by residue in order to locate high-energy regions in which steric overlaps had not been eliminated. In certain regions, particularly on both sides of amino acid deletions, it was found that one or more atoms were trapped in local high-energy conformational minima from which they could not escape (in a minimization algorithm) because of steric hindrance from neighboring atoms. In such cases, further insight was gained by studying a molecular model of α-lactalbumin constructed from Lab-Quip molecular models (LABQUIP, 18 Rosehill Park Estate, Caversham, Reading RG48XE, England) and alternative conformations for certain backbone and side-chain dihedral angles were tested as follows.

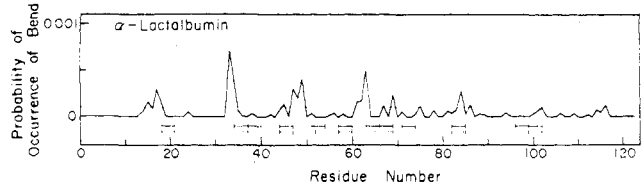

FIGURE 2: Probability profile for bends (beginning at position *i*) in α-lactalbumin, *i.e.*, for the occurrence of an $i \rightarrow i + 3$ bend, computed from the amino acid sequence, using the procedure of Lewis *et al.* (1971). (|——|) and (|· · · ·|) refer to the observed and calculated (Lewis *et al.*, 1971) bends in lysozyme, for comparison.

Some of the high-energy regions which were most difficult to adjust by energy minimization (and thus required resort to the molecular model) involved bends or chain-reversal conformations. Of the seven bends which existed at this stage of the calculations on α-lactalbumin, four were of relatively high energy. Two of these high-energy bend conformations occur near amino acid deletions, thus introducing additional uncertainties in deducing a reasonable starting conformation. As a guide for selecting possible alternative bend conformations, we have used the method of Lewis *et al.* (1971) for predicting the location of bends, together with recent results (Lewis *et al.*, 1973b) from conformational-energy calculations on various bends. The probability profile for the occurrence of bends in α-lactalbumin is shown in Figure 2, together with the observed and calculated bends for lysozyme (scaled by residue number to account for deletions). Three of the bends in the computed structure (after the energy minimization described in section B3) were in locations which agreed with the profile of Figure 2 and also exhibited conformations similar to the most stable bend conformations predicted by independent calculations on a tetrapeptide of alanine and other tetrapeptides known to occur in bend conformations (Lewis *et al.*, 1973b). However, the four computed bend conformations at residues *15–18*, *33–37*, *45–49*, and *62–70* either exhibited high-energy conformational minima, or else single bends existed where the predictions of Figure 2 showed the possibility of multiple bends. In order to select alternative bend conformations and ultimately reduce the conformational energy in these four bend regions, the following procedure was used.

(a) The several amino acid residues involved in the bend were detached from the Lab-Quip model, and this portion of the chain was then folded in various alternative conformations. Many of these conformations were rejected because the altered chain, when reinserted into the model, could not easily be made to bridge the gap introduced when the segment was detached from the model.

(b) The approximate bend conformations which were not rejected by criterion (a) were subjected to a stage I fitting

procedure as described in Al, but allowing only the backbone dihedral angles in the short segment of interest to be variables. The fit, as in section Al, was to the X-ray coordinates of lysozyme. Those conformations which did not have reasonable fits to the X-ray coordinates, or were grossly distorted from the initial bend conformation, were rejected at this stage.

(c) The bend conformations which satisfied criterion (b) were then subjected to stage II energy minimization in the whole protein (see section B3) by allowing all dihedral angles of the pertinent segments of the polypeptide chain (Cl–C4) (including those of the single residues preceding the first and following the last residue of the segment) to be varied. In most cases (see Cl–C4 for details), this preliminary energy minimization yielded a single bend conformation with a total energy (in the short segment) much lower than that of any other alternative conformation examined (including the conformation which resulted from the original stage II minimization).

In Sections Cl–C4, we describe how the procedures were applied to four specific bends.

(1) *Bend at Residues 15–18.* The amino acids at positions **14** and **15** of lysozyme are deleted in α-lactalbumin, thus causing a major disruption of the conformation in this region. After the initial energy minimization (of section B3), residues *15* (Leu) through *18* (Tyr) of α-lactalbumin were folded in a tight bend, in good agreement with the computed high probability of occurrence of a bend at this position (Figure 2). Also, the existence of a bend in this region is compatible with the termination of the helical segment in residues *5–15*. However, the minimized energy (235 kcal) of the segment from Leu-*12* to Gly-*19* was rather high, indicating that this portion of the chain was trapped in a high-energy local conformational minimum. From an inspection of the Lab-Quip model, it was discovered that a clockwise rotation (looking along the chain from the N to the C terminus) of the planar amide group between Lys-*16* and Gly-*17* by ~60° resulted in the formation of a distorted type II bend (Venkatachalam, 1968; Lewis *et al.*, 1973b) with residue *15* at position *i* (Lewis *et al.*, 1971); the resulting bend appeared to contain a favorable $i + 3 \rightarrow i$ hydrogen bond. By an adjustment of a few degrees in the dihedral angles of residues *11–15*, it was possible to incorporate this modified bend easily into the Lab-Quip model without disrupting the continuity of the chain. The new conformation for this segment was introduced into the computed structure for α-lactalbumin by again using the stage I fitting procedure and minimizing the fit to the coordinates. A short (five cycles) preliminary stage II energy minimization of segment *12–19* (incorporated into the molecule) resulted in an energy of 107 kcal (compared to the original energy of 235 kcal). The conformational alteration described above led to a double bend in the region between residues *15* and *20*, the first bend involving residues *15–18* and the second involving residues *17–20*. It was clear that a significant improvement of the conformation in this region had been obtained, and that the energy of this region would decrease even further when the entire 20-residue segment spanning this region was minimized (section E).

(2) *Bend at Residues 33–37.* The amino acid sequence of α-lactalbumin shows little homology to lysozyme in the region of Thr-*29* to Ser-*34*. This section is helical in lysozyme, but the helix probability profile (Lewis and Scheraga, 1971) indicates that the helix extends only to Phe-*31* in the case of α-lactalbumin. Also, the probability of occurrence of a bend at Thr-*33* to Gly-*35* in α-lactalbumin is very high (Figure 2). After the initial energy minimization of section B3 [in which residues *29–34* of α-lactalbumin had been placed *initially* in a helical conformation (Browne *et al.*, 1969)], residues *29–34* were still nearly helical, but several high-energy overlaps remained. Therefore, even though this was not a region of a deletion or an insertion, this region was adjusted with the aid of the Lab-Quip model—specifically to determine whether any of the types of bends described by Lewis *et al.* (1937b) could be accommodated in region *33–37* without disturbing the position of the side chain of Tyr-*36* (Phe-**38** in lysozyme) which lies in a pocket closely surrounded by other residues. In order to increase the freedom of the chain in this region during the subsequent energy minimization, residues *34* and *35* were not constrained to the corresponding X-ray coordinates of lysozyme. Only type I′ and III′ bends (Lewis *et al.*, 1973b) could accommodate to these constraints. After minimizing the energy of residues *33–37* (in contact with the whole protein), both bend conformations had a lower energy (42 and 41 kcal, respectively) than that (81 kcal) of the original (helical) segment. Since the backbone atoms of Thr-*33* corresponded more closely to the X-ray coordinates of lysozyme in a type III′ bend than in a type I′ bend, the type III′ bend was selected for the final energy minimization of section E.

(3) *Bend at Residues 44–47.* As can be seen in Figure 2, there is a considerable likelihood of the occurrence of two bends somewhere in the region between Asn-*44* and Asp-*49* of α-lactalbumin. Also, there is a deletion at residue **47** of lysozyme. The results of the previous stage II minimization had yielded a single bend with the amino acids preceding and following the bend involved in antiparallel pleated sheet formation. We investigated the possibility of incorporating other bends of various types in this region of α-lactalbumin, despite the absence of high-energy local minima in this region. However, all exploration failed to yield an alternative conformation with a lower energy or a better fit to the X-ray coordinates of lysozyme than the original bend between residues *44* and *47* along with the original antiparallel pleated sheet structure between residues *41–44* and *47–50*, despite the prediction of a double bend in Figure 2.

(4) *Bends in Region of Residues 62–70.* The results of the minimization of section B3 for residues *62–70* of α-lactalbumin were not satisfactory. The original type II bend at Pro-*67* gave high-energy overlaps in this region, which were not relieved.

In the sequence *62–70*, there are high bend probabilities for $i \rightarrow i + 3$ bends whose *i*th residues are Asn-*63*, Pro-*67*, and Ser-*69*. It was found that a type I bend could be produced with nearly the same fit to the X-ray data (except for the carbonyl oxygen) by rotating the planar amide group between Asp-*64* and Gln-*65* by 180°. In addition, the bend at Pro-*67* (found in the minimization procedure of section B3) had a high energy which could be reduced (without disrupting the fit to the X-ray data) by rotating the amide group between His-*68* and Ser-*69* clockwise (looking along the chain from the N to the C terminus) by 30°. When these two modifications (*i.e.*, the resulting bends at Asn-*63* and Asp-*66*, respectively) were incorporated, a short (five cycles) preliminary energy minimization reduced the energy of residues *62–70* (calculated for the segment while in contact with the rest of the molecule) from the original value of 506 kcal from the stage II results to 75 kcal. The resulting conformation of this segment was used in the final energy minimization described in section E.

D. Modification of C-Terminal Section (121–123). After the energy minimization of section B3, residues Glu-*121* through Leu-*123* had an extended conformation, designated here as T1, with this segment sticking out away from the main body of the protein. Further examination of this terminal section revealed

the possible existence of other conformations with energies lower than that (−12.2 kcal) of T1.

A second conformation, obtained by using the Lab-Quip model to search through possible dipeptide minima, placed residues *121* and *123* in an approximately helical conformation with the terminal carboxyl group close to Lys-*13*. This conformation (designated T2) had an energy of −18.5 kcal after ten cycles of energy minimization, and resembles that of lysozyme in that residues *121–123* are nearly helical, but differs from lysozyme in that the terminal carboxyl group is near Lys-*13* rather than near Asp-*116*, which would be close to the equivalent terminal position in lysozyme. In conformation T2 the side chain of Leu-*123* is partially exposed to the solution.

Two more energetically nearly equivalent conformations were obtained by positioning the three terminal residues in a small cleft defined by Lys-*5*, Cys-*6*, His-*32*, Phe-*33*, Trp-*118*, and Val-*119*. The first (designated T3) had Leu-*123* near Trp-*118* and Lys-*5*, with Glu-*121* near Trp-*118*, and exhibited an energy of −15.6 kcal after ten cycles of energy minimization. The second (designated T4) is obtained from T3 by varying ϕ of Leu-*123* to place the Leu side chain up into a hydrophobic pocket. In the further minimization of section E, T1 was discarded because of its relatively higher energy, and conformations T2, T3, and T4 were all considered.

E. Final Energy Minimization of Stage II. After carrying out the energy minimization procedure of section B3, and the subsequent adjustments of sections C and D, the stage II energy minimization procedure was then repeated on the whole protein, in 20-residue segments. The segments chosen for this minimization were the same as those listed in section B3. After completing ten cycles of energy minimization on each of these segments, a second iteration of ten cycles of minimization was performed on each segment. In these two final minimizations, κ_i was taken as 10 kcal/Å^2 and λ_i as zero in order to allow greater freedom for the conformation to change in response to the energetic requirements of the system.

Results

Table I summarizes the contributions of various energy terms to the total energy for the final structure of α-lactalbumin, computed by the procedure of section E. The magnitudes of the RMS values, also given in Table I, provide an indication of how close the fit of the calculated structure of α-lactalbumin is to the X-ray structure of lysozyme. The corresponding values for lysozyme, at the end of a single iteration of stage II refinement, are given in parentheses in Table I.

The various contributions to the total energy show that the computed conformation has (a) no serious interatomic overlaps, (b) a favorable set of hydrogen bonds, (c) a rotational energy of ~6 kcal per residue on the average (see Warme and Scheraga, 1974), and (d) fairly good stereochemistry at the disulfide bonds.

The nonbonded energy of each residue (see Warme and Scheraga, 1973, 1974 for definition) is plotted in Figure 3. Comparison with Figure 2 of Warme and Scheraga (1973) indicates that, after stage II, the nonbonded energies of α-lactalbumin are lower than those of lysozyme. However, only ten cycles of energy minimization had been performed on each segment of lysozyme, whereas at least 20 cycles of energy minimization was performed on each segment of α-lactalbumin at this point. The energy of lysozyme was also significantly reduced by a second application of stage II refinement

TABLE I: Results from Final Stage II Energy Minimization.

RMS deviations of computed coordinates of α-lactalbumin from the X-ray coordinates of lysozyme	
460 backbone atoms[a] (including C^β atoms)	0.79 Å (0.53 Å)[c]
163 side-chain atoms[b]	0.64 Å (0.64 Å)
623 backbone and side-chain atoms[a,b]	0.76 Å (0.57 Å)
Energy contributions[d] (kcal)	
Nonbonded (E_{NB})	59 (1300)
Hydrogen bond (E_{HB})	−73 (−60)
Rotational (E_{ROT})	730 (840)
Disulfide loop (E_{SS})	63 (40)
Total[e]	779 (2120)

[a] The RMS deviation of the backbone atoms was taken over only those atoms to which eq 1 was applied; *i.e.*, the residues on each side of a "deletion" or "insertion" were not included in the computed RMS value. However, those residues which were readjusted to construct low-energy bends *are* included in the computed RMS value. [b] This RMS value was computed only for atoms of homologous and near-homologous side chains. [c] The values in parentheses pertain to the stage II refinement of lysozyme (Warme and Scheraga, 1973). [d] These energy contributions (defined by Warme and Scheraga, 1973) are summed over *all* 123 residues of α-lactalbumin. [e] This value does not include the fitting potential, which was found to be 1900 kcal for α-lactalbumin.

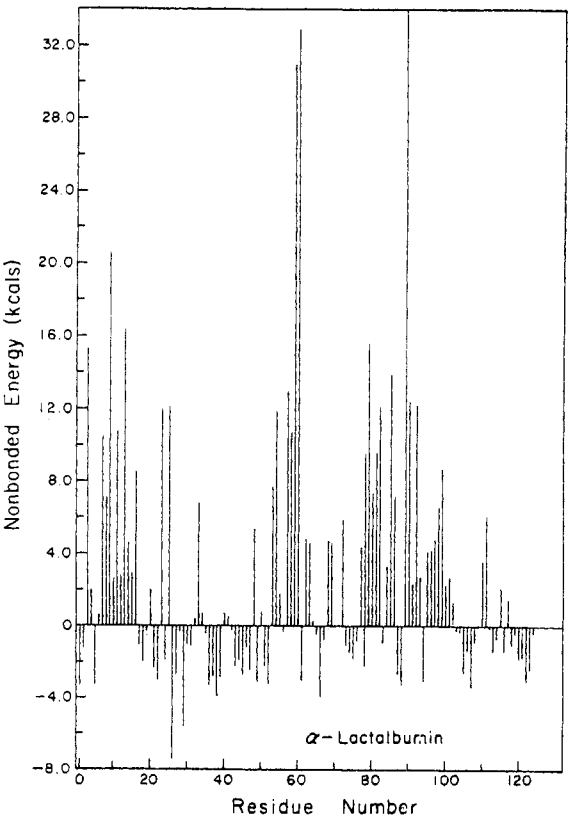

FIGURE 3: Total nonbonded energies ($E_{NB} + E_{HB} + E_{ROT}$) for all atoms in each residue of α-lactalbumin after the final stage II energy minimization.

TABLE II: Computed Dihedral Angles of Bovine α-Lactalbumin.[a]

Residue		ϕ	ψ	χ_1	χ_2	χ_3	χ_4
1	GLU	-180.00	127.25	-162.05	-95.73	12.72	
2	GLN	-126.24	67.84	-179.43	168.80	81.77	
3	LEU	-47.13	168.67	75.63	-177.30		
4	THR	-95.68	-179.79	37.48	163.51		
5	LYS	-64.70	-59.04	-63.85	179.42	-177.81	-179.48
6	CYS	-62.09	-48.28	-89.64			
7	GLU	-44.66	-60.02	-168.90	171.48	-177.53	
8	VAL	-55.90	-47.25	66.19			
9	PHE	-31.53	-85.57	179.64	-99.44		
10	ARG	-47.53	-46.84	-76.15	172.04	-77.92	176.84
11	GLU	-37.64	-69.87	-155.31	69.42	-100.22	
12	LEU	-79.12	-14.44	-166.08	84.32		
13	LYS	-23.30	-53.70	-158.07	-79.32	163.34	67.75
14	ASP	-40.53	-63.12	-79.02	77.85		
15	LEU	-42.77	-59.92	-173.60	66.68		
16	LYS	-17.36	113.06	164.76	-130.88	178.81	-174.97
17	GLY	118.67	-19.30				
18	TYR	-67.52	123.97	-157.87	75.56	0.01	
19	GLY	75.84	-1.98				
20	GLY	82.98	19.86				
21	VAL	-99.13	132.44	-178.97			
22	SER	-55.66	159.28	-100.00	-60.00		
23	LEU	-69.39	-44.92	117.90	87.39		
24	PRO	-57.58	-41.08				
25	GLU	-49.89	-37.31	-88.68	-175.92	-85.24	
26	TRP	-74.74	-34.91	-65.66	94.10		
27	VAL	-67.14	-33.76	158.98			
28	CYS	-70.67	-39.98	-178.80			
29	THR	-62.63	-51.67	-59.87	165.40		
30	THR	-67.46	-25.07	41.06	165.89		
31	PHE	-64.12	-63.64	-179.51	-104.91		
32	HIS	-80.32	27.20	-70.62	-61.58		
33	THR	-150.63	96.55	-61.31	-38.31		
34	SER	86.37	-11.76	-46.77	-132.95		
35	GLY	54.80	58.20				
36	TYR	50.49	44.44	-64.22	169.69	-173.57	
37	ASP	-133.54	107.85	-166.98	-8.85		
38	THR	-67.33	-17.27	53.11	-179.93		
39	GLU	-84.42	-33.58	-94.68	-168.69	-60.55	
40	ALA	-32.93	142.62				
41	ILE	-148.79	148.98	51.17	156.81		
42	VAL	-151.34	146.00	153.29			
43	GLU	-138.83	142.17	127.40	135.78	-172.48	
44	ASN	-94.14	-68.52	-174.95	25.60		
45	ASN	-144.04	-61.66	-60.97	102.08		
46	GLN	-108.75	-65.01	-72.64	168.22	-85.22	
47	SER	-69.65	140.76	122.62	60.90		
48	THR	-117.76	159.35	-68.34	60.84		
49	ASP	-111.70	131.53	-55.24	168.60		
50	TYR	-133.86	157.62	-46.42	-100.51	0.02	
51	GLY	85.96	168.14				
52	LEU	-62.73	-49.61	-122.02	140.39		
53	PHE	-75.58	20.02	-71.23	103.84		
54	GLN	23.26	81.93	-67.70	-108.46	132.95	
55	ILE	-90.31	133.90	-56.31	165.10		
56	ASN	-66.70	158.22	-139.77	50.44		
57	ASN	-117.19	26.26	100.32	-73.57		
58	LYS	-96.57	-8.38	165.15	165.83	-165.45	156.29
59	ILE	-169.40	-25.10	15.52	179.65		
60	TRP	-76.49	-71.69	-77.37	163.20		
61	CYS	-111.87	133.51	50.25			
62	LYS	-79.49	158.37	167.26	-144.30	102.88	-76.85
63	ASN	-152.22	-165.37	-126.82	68.58		
64	ASP	-123.49	8.47	-171.50	-158.68		
65	GLN	-113.21	-4.74	-112.58	176.66	-86.38	
66	ASP	-117.72	97.43	-150.06	-177.32		
67	PRO	-57.58	-37.81				
68	HIS	-81.13	29.80	-171.23	-84.86		
69	SER	-55.79	145.42	82.92	-63.61		
70	SER	-155.98	56.20	-122.47	60.75		
71	ASN	-37.60	109.14	-140.29	-28.73		
72	ILE	-103.75	-26.56	-76.58	-63.09		
73	CYS	-98.83	-7.39	-89.62			
74	ASN	53.43	69.32	-124.35	111.21		
75	ILE	-161.09	156.42	-161.73	63.04		
76	SER	-92.09	114.51	-173.02	-60.17		
77	CYS	-31.64	-55.94	-53.95			
78	ASP	-38.84	-55.22	-80.00	-90.00		
79	LYS	-38.29	-45.88	-163.42	78.28	-163.05	172.37
80	PHE	-55.55	-47.48	-76.41	-77.31		
81	LEU	-46.19	-20.70	-64.99	-176.89		
82	ASN	-50.28	155.19	-163.89	39.60		
83	ASN	-74.74	-5.97	63.68	-66.24		
84	ASP	-102.17	113.49	-134.36	86.02		
85	LEU	-72.48	6.41	66.07	143.52		
86	THR	-49.89	-62.00	-81.99	179.37		
87	ASN	-59.85	-56.98	-66.32	97.64		
88	ASN	-37.09	-78.61	-76.80	-83.28		
89	ILE	-46.93	-52.74	98.20	179.33		
90	MET	-42.36	-30.08	-171.19	56.30	63.70	
91	CYS	-97.54	-31.16	-178.38			
92	VAL	-62.48	-15.61	88.76			
93	LYS	-75.21	-35.09	-72.04	-97.73	179.23	85.17
94	LYS	-85.45	-33.01	-60.64	-136.76	129.64	-69.85
95	ILE	-36.98	-52.73	-95.49	81.05		
96	LEU	-52.81	80.86	-92.03	69.06		
97	ASP	-78.57	-11.07	53.03	-65.19		
98	LYS	131.29	11.85	-162.74	-155.27	104.18	-69.23
99	VAL	-112.72	7.23	161.64			
100	GLY	57.21	-147.99				
101	ILE	-79.66	3.95	-49.56	-151.84		
102	ASN	-64.68	-22.87	-90.15	-57.55		
103	TYR	-68.14	-6.66	-172.32	-97.32	-0.17	
104	TRP	-106.41	89.33	-80.95	-74.32		
105	LEU	-35.33	-53.10	-177.38	89.44		
106	ALA	-45.77	-81.47				
107	HIS	-28.00	-73.74	175.71	86.57		
108	LYS	-36.21	-59.69	-176.59	-179.22	-53.01	179.82
109	ALA	-52.61	-84.02				
110	LEU	-94.24	87.52	-3.36	134.17		
111	CYS	164.89	139.28	-86.15			
112	SER	179.62	-81.70	-153.49	74.47		
113	GLU	-116.54	178.65	-74.44	170.70	-86.57	
114	LYS	54.03	26.42	-77.13	178.20	-93.21	177.62
115	LEU	-80.26	-17.68	-70.48	156.80		
116	ASP	-61.04	-25.06	53.10	-65.26		
117	GLN	-60.07	-11.86	-90.72	100.66	-89.73	
118	TRP	-90.73	11.63	-68.51	111.73		
119	LEU	-67.75	-77.56	-166.44	118.37		
120	CYS	-161.75	46.47	-89.14			
121	GLU	-115.26	-56.11	174.27	-110.44	-73.36	
122	LYS	-52.15	-48.26	-76.08	-79.23	151.57	-87.90
123	LEU	70.90	-45.88	-136.86	-69.31		
				Conformation	T4		
121	GLU	-124.03	-55.78	177.22	-125.74	-113.22	
122	LYS	-49.64	-49.02	-75.49	-77.54	161.77	-85.10
123	LEU	-54.93	-52.87	-171.40	82.11		
				Conformation	T2		
104	TRP	-106.00	90.02	-80.63	-74.43		
105	LEU	-35.40	-52.54	-177.35	89.40		
106	ALA	-44.85	-81.08				
107	HIS	-27.03	-73.95	175.60	86.63		
108	LYS	-36.50	-59.87	-176.27	-179.65	-53.19	179.81
109	ALA	-53.63	-83.96				
110	LEU	-95.17	86.57	-3.81	134.38		
111	CYS	164.61	139.01	-86.04			
112	SER	179.42	-82.06	-153.55	74.46		
113	GLU	-117.14	178.76	-74.44	170.73	-86.57	
114	LYS	54.27	26.85	-76.94	178.19	-93.16	177.63
115	LEU	-80.17	-17.33	-70.25	157.17		
116	ASP	-61.00	-25.13	53.10	-65.25		
117	GLN	-61.22	-12.58	-91.19	100.84	-89.53	
118	TRP	-90.64	10.74	-67.85	111.07		
119	LEU	-66.40	-70.88	-172.52	117.90		
120	CYS	-161.81	153.19	-90.19			
121	GLU	-92.04	-57.04	-75.23	178.69	-124.94	
122	LYS	-59.66	-45.53	-161.14	-89.55	141.33	-67.80
123	LEU	-95.08	-51.45	-72.92	106.77		

[a] The data for the complete 123-residue protein pertain to conformation T3 of the terminal section. Those for conformation T4 and T2 are given at the end of the table.

(Warme and Scheraga, 1974). Only three residues of α-lactalbumin (Ile-*59*, Trp-*60*, and Ile-*89*) have nonbonded energies greater than 25 kcal. The high energies of Ile-*59* and Trp-*60* arise from a backbone interaction between the amide hydrogens of these two residues. The high energy of Ile-*89* results from an internal interaction between the γ_2 methyl group and the backbone amide hydrogen of this residue. Each of these high-energy interactions arises from a conformation caught in a local energy minimum, and could be eliminated by the stage III refinement procedure[1] (Warme and Scheraga, 1974). All remaining moderately large nonbonded energies should also be reduced in stage III, probably causing only slight alterations in the atomic coordinates. The computed heavy-atom dihedral angles (for both back-

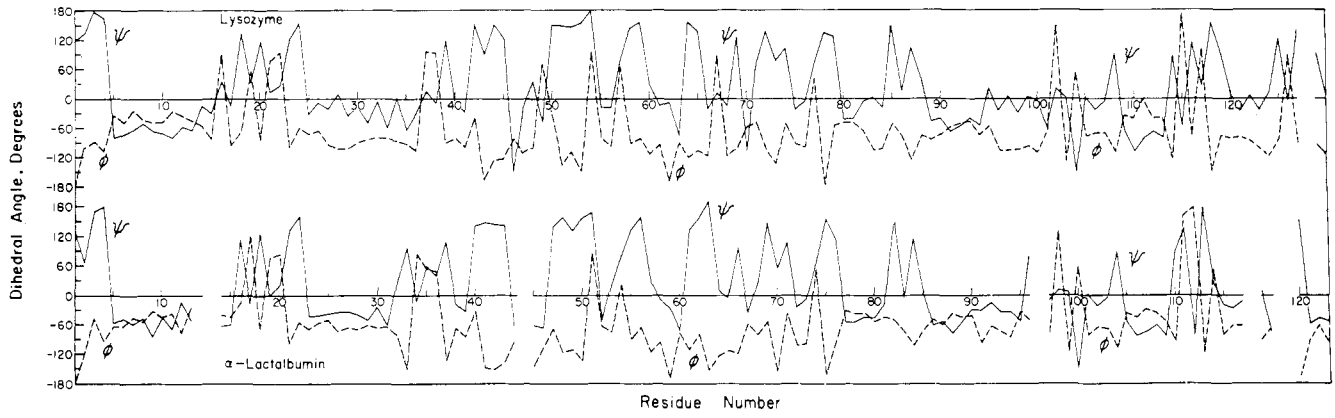

FIGURE 4: A comparison of the dihedral angles of lysozyme (after stage II) and α-lactalbumin at the same stage of refinement.

bone and side-chain atoms) are given in Table II, and the computed atomic coordinates of all heavy atoms have been submitted for documentation.[3]

Discussion

It must be emphasized that the structure of α-lactalbumin computed here was obtained on the *assumption* that it is structurally homologous to lysozyme. The determination of the crystal structure of α-lactalbumin is now in progress (Aschaffenberg *et al.*, 1972a,b). The purpose of carrying out the calculations described in this paper was to assess the predictive capabilities of our computational procedures and refiuement techniques by applying them to a protein whose structure is unknown, up to the present time, but will become known in the near future. In the absence of any definitive X-ray evidence, the best test of the validity of our computed structure for α-lactalbumin is to examine it in the light of the large body of experimental evidence, summarized below, regarding the structure of α-lactalbumin and its relationship to the structure of lysozyme.

A. General Structural Homology between Lysozyme and α-Lactalbumin. Figure 4 shows the distinct homology of the backbone dihedral angles in the refined structures of α-lactalbumin and lysozyme. This representation of the values of ϕ and ψ indicates quite clearly the regions of the polypeptide chain having certain commonly occurring conformations. For example, the α-helical regions in α-lactalbumin at residues *5–15*, *23–31*, *77–81*, *86–95*, *105–109*, and *115–117* are characterized by negative values of both ϕ and ψ. In each of these cases, helices also occur in the corresponding regions of lysozyme. Sections of extended chain, characterized by large negative values of ϕ and positive values of ψ, occur at residues *1–4*, *21–22*, *40–43*, *47–50*, *61–63*, *69–71*, and *75–76* in α-lactalbumin, and also occur at similar positions in lysozyme. Bend regions are generally characterized by abrupt changes in dihedral angles over several adjacent residues. Residues *16–20*, *33–37*, *44–45*, *51–52*, *98–100*, and *110–114* stand out as bends or chain reversal loops in this representation. Thus, the homologies in dihedral angles between lysozyme and α-lactalbumin are closely retained in most cases.

Although this homology is quite apparent for the backbone chains, a visual comparison of Lab-Quip models of lysozyme and α-lactalbumin immediately reveals significant differences. Lysozyme appears to be a prolate ellipsoid, while α-lactalbumin appears to be flattened into an oblate ellipsoid. A stereoview (presented in color to enable different kinds of side chains to be distinguished from each other and from the backbone) is shown in Figure 5 for α-lactalbumin, with its tail in the T3 conformation. The view in Figure 5 is toward the cleft (front surface) and, in this orientation, the front-to-back thickness is ~23 Å, the height is ~37 Å, and the width is ~32 Å; the corresponding dimensions of a Lab-Quip lysozyme model are 28 × 38 × 33 Å. The reduced thickness of α-lactalbumin arises in part from deletions at residues **14** and **15** of lysozyme on the back and at residue **47** in the front. Side-chain replacements occurring on the surface of the protein also account for the reduced thickness of α-lactalbumin. For example, in several cases, the very large arginine side chains of lysozyme are either deleted entirely or else are replaced by considerably smaller side chains in α-lactalbumin (*e.g.*, Arg-**14** of lysozyme is deleted, Arg-**21** is replaced by Gly-*19*, Arg-**45** by Glu-*43*, and Arg-**73** by Ser-*70*). A number of other amino acid replacements also contribute in less obvious manner to the difference in overall dimensions, particularly the front-to-back thickness.

Experimental evidence on the overall size and shape of α-lactalbumin has been obtained by several methods. From measurements of polarization of fluorescence, Rawitch (1972) obtained rotational relaxation times which correlated well with either larger effective volume or increased asymmetry of α-lactalbumin as compared to lysozyme. Since the difference in hydration between these proteins was too small to account entirely for the observed discrepancies, Rawitch (1972) concluded that significant differences in shape would explain these data. On the other hand, data based on diffusion and sedimentation experiments (Barel *et al.*, 1972) seem to indicate that a close similarity exists in the shapes when these proteins are studied by hydrodynamic methods. Low angle X-ray scattering experiments (Krigbaum and Kugler, 1970) have indicated that α-lactalbumin is an oblate ellipsoid with radius of gyration significantly larger than that of lysozyme, while the scattering results for lysozyme were entirely consistent with its known crystal structure. Achter and Swan (1971) have reinterpreted the observed differences between the X-ray scattering results for these two proteins in terms of

[3] The Cartesian coordinates of the computed structure (stage II) of α-lactalbumin have been deposited as document No. NAPS 02263 with the ASIS National Auxiliary Publication Service, c/o Microfiche Publications, 305 East 46th Street, New York, N. Y. 10017. A copy may be secured by citing the document number and by remitting $1.50 for microfiche or $5.00 for photocopies. Advance payment is required. Make checks or money orders payable to Microfiche Publications.

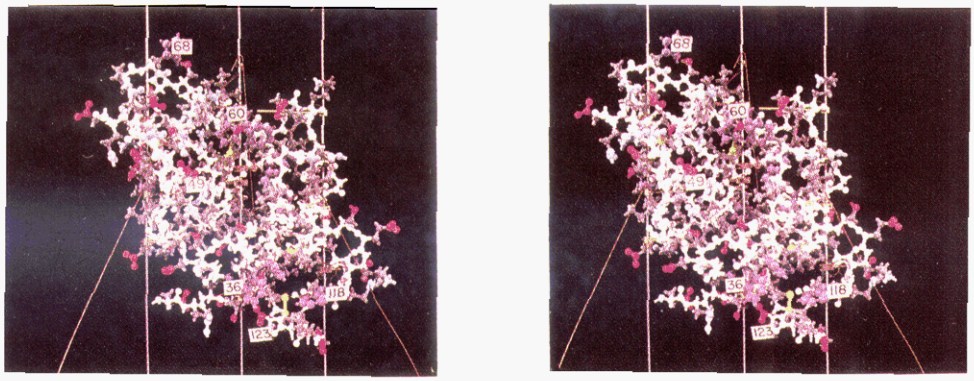

FIGURE 5: A stereoview of α-lactalbumin built from Lab-Quip molecular models. The backbone is white, aliphatic side chains are grey, carboxylic acid groups are red, amino, guanidino, and amide groups are blue, sulfurs are yellow, and ring structures (histidine, phenylalanine, tyrosine, and tryptophan) are purple. This model was constructed with the T3 tail conformation.

a small amount (<5%) of dimerization of α-lactalbumin under the given experimental conditions. In summary, our computed representation of α-lactalbumin as an oblate ellipsoid with overall dimensions somewhat different from lysozyme seems to be in general accord with experimental observations.

B. The Amino-Terminal Section. In our computed model of α-lactalbumin, the α-amino group of Glu-*1* and one of the side-chain amide hydrogens of Gln-*2* interact quite strongly with the side-chain carboxyl oxygens of Asp-*37*. The α-amino group also interacts favorably with the side-chain carboxyl group of Glu-*39*. These interactions contribute to the stabilization energy of the first five residues, keeping the N-terminus tightly bound along the bottom of the molecule (as seen at the bottom of Figure 5).

C. The Cleft Region. The cleft region is illustrated in Figure 6, which is a view looking down into the cleft from the top of the model (top being defined by Figure 5), and is also shown in Figure 5, in which the viewer is looking directly at the cleft. In lysozyme, this cleft is the active site region and contains Asp-**52** and Glu-**35**, both of which are thought to be necessary for enzymatic activity (Blake *et al.*, 1967). In α-lactalbumin, Thr-*33* replaces Glu-**35** and Asp-*49* corresponds to Asp-**52**. In Figures 5 and 7, it can be seen that the His-*32* side chain points down away from the cleft region, and does not occupy the position corresponding to Glu-**35** of lysozyme; this differs from the conclusion of Browne *et al.* (1969) who suggested that His-*32* may substitute for Glu-**35** in the active site. In our model of α-lactalbumin, Thr-*33* appears to correspond more closely in position to Glu-**35** of lysozyme.

Tyr-*103* lies in the upper portion of the cleft region (Figures 5 and 6) and tends to block Trp-*60* and Trp-*104* from contact with substrates. This result is in agreement with the conclusion of Browne *et al.* (1969) that the sugar binding sites A and B of lysozyme are blocked in α-lactalbumin.

D. The Carboxyl-Terminal Section. While only a limited search has been made among the many possible terminal conformations, this segment [which is left undefined in the model constructed by Browne *et al.* (1969)] was found to have at least three distinctly favorable conformations for the residues *121–123*, all of nearly equivalent energy. One of them, T3, is shown schematically in Figure 7 and a perpendicular projection is shown in Figure 8. This section lies just below Trp-*118* (see Figures 5 and 7) and shields this residue from the solvent. The salt bridge between the ϵ-NH_3^+ of Lys-*5* and the terminal COO^- of Leu-*123* helps to stabilize this conformation; at the same time, the Leu-*123* side chain lies below the Cys-*6*–Cys-*120* disulfide bond and may tend to shield it from the solvent.

In a second low-energy tail conformation, T4, the Leu-*123* side chain is inserted into the region in front of the Cys-*6*–Cys-*120* disulfide bond (from the viewpoint of Figure 7), and forms a hydrophobic pocket with Trp-*118* and other surrounding residues. This conformation results from T3 by varying ϕ of Leu-*123* (see Table II).

A third low-energy conformation, T2, is shown in Figure 9. In this arrangement, the tail is directed toward the back of the protein, from the viewpoint of Figure 5, and folds up in a manner similar to that of the tail in lysozyme. This conformation can be obtained from conformaton T4 by varying ψ of Cys-*120* which is the same as ψ of Cys-*120* in T3 (see Table II). It should be noted (as a point to be discussed below) that, in conformation T2, Trp-*118* is completely exposed to the solvent, and also the front lower right-hand side of the molecule (as viewed in Figure 5) is nearly devoid of charged groups, because of the removal of the COO^- of Leu-*123*, and the NH_3^+ of Lys-*122*.

E. Nonpolar Character of Front Right-Hand Surface. From the viewpoint of Figure 5, much of the front right-hand surface of our computed structure for α-lactalbumin is nonpolar (except for the presence of Lys-*108*), especially when the tail adopts conformation T2, in which Tyr-*36*, Phe-*31*, Trp-*118*, and His-*32* are exposed to solvent (*cf.* Figures 7–9). Thus, there are almost no polar groups on the front right-hand side of the molecule from the lower portion up through the cleft region, which may be the reason that α-lactalbumin tends to aggregate at low and high pH (Kronman and Andreotti, 1964; Kronman *et al.*, 1964). Disruption of the salt bridge of conformation T3 (see section B) by protonation of the COO^- of Leu-*123* or deprotonation of the NH_3^+ group of Lys-*5* should enable the tail to adopt conformation T2, thereby further enhancing the nonpolar character of the front right-hand side of the molecule. In conformation T2, the side chain of Lys-*5* might flip down into the opening left by the tail, thus further exposing Tyr-*36* to the solvent (see Figure 7).

The nonpolar character of the front right-hand face of our computed structure for α-lactalbumin strongly suggests that this is the site of its interaction with galactosyl transferase ("A" protein). Several experimental studies support this conclusion. For example, Castellino and Hill (1970) showed that carboxymethylation of Met-*90*, which is on the back of the model as viewed in Figure 5, had little or no effect on the ability of α-lactalbumin to specify lactose synthesis or to in-

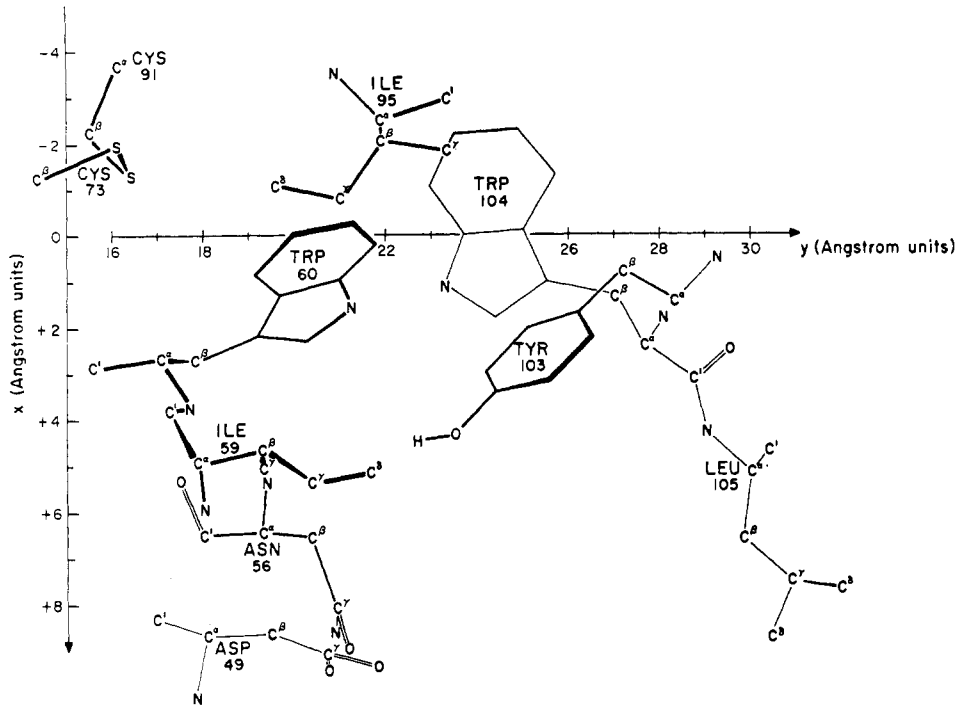

FIGURE 6: Environment of cleft region, including residues Trp-*60*, Tyr-*103*, and Trp-*104* of α-lactalbumin (*cf.* Figure 5).

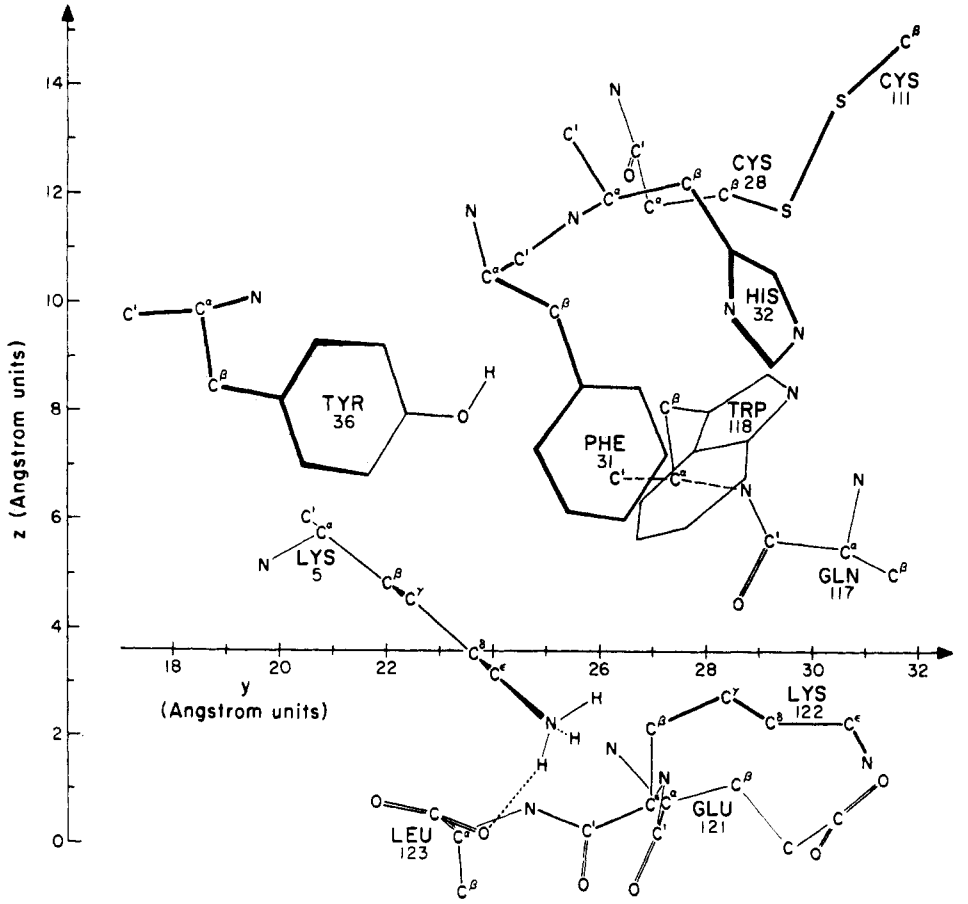

FIGURE 7: View of C-terminal section with tail in T3 conformation, from same direction as in Figure 5.

hibit *N*-acetyllactosamine synthesis when combined with galactosyl transferase. They also demonstrated that this activity was partially lost upon carboxymethylation of histidine residues, of which one (His-*32*) is located on the borders of this front nonpolar surface. The observation by Denton and Ebner (1971) that tyrosine modification prevents formation

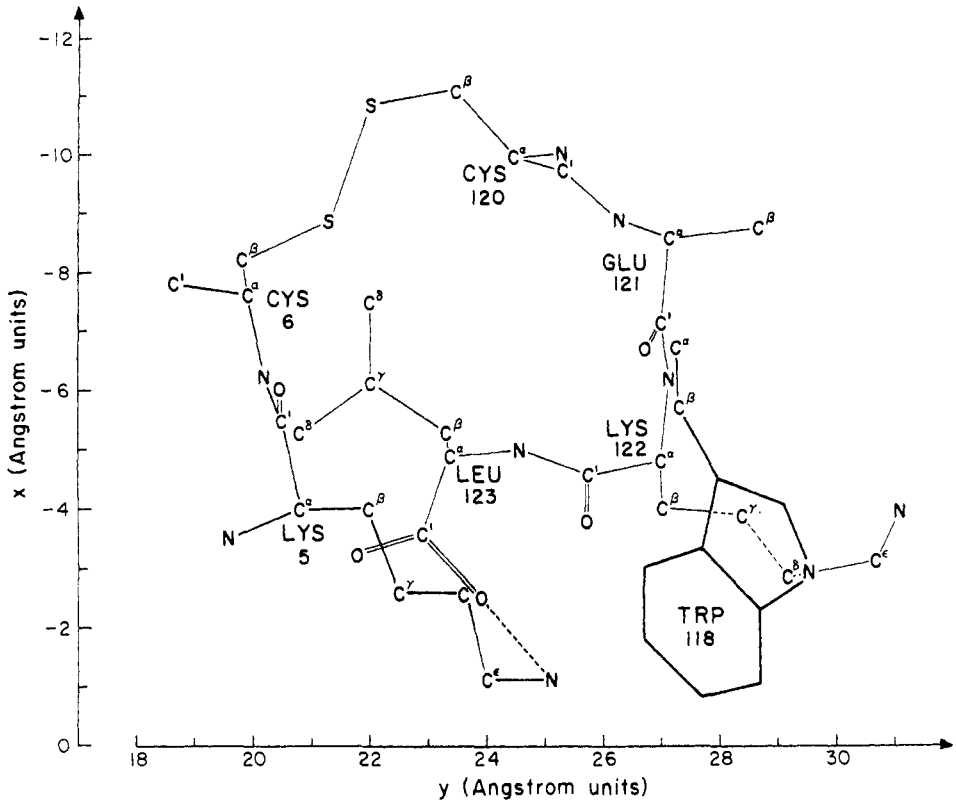

FIGURE 8: View of C-terminal section with tail in T3 conformation, looking down from top of molecule as viewed in Figure 5.

of the lactose synthetase complex suggests the involvement of Tyr-*36* and Tyr-*103* which are located on this front nonpolar surface. Lin (1970) showed that the presence of galactosyl transferase protected only about 4 out of 20 carboxyl groups from reacting with a water-soluble carbodiimide reagent, which suggests that these two proteins interact at a nonpolar surface nearly devoid of carboxyl groups, such as the front right lobe of α-lactalbumin (viewed as in Figure 5). In summary, the experimental evidence seems to suggest that the front surface of our computed structure for α-lactalbumin may be the site of interaction with galactosyl transferase.

F. Disulfide Bonds. Of the four disulfide bonds in our computed structure for α-lactalbumin, three were found to have a CS–SC dihedral angle of about +90 ± 20° (*i.e.*, a right-handed conformation). The fourth, that between Cys-*6* and Cys-*120*, is involved in the tail section and is nearly extended (~180°). While the latter is not a favorable conformation, our freedom to alter the conformation of the chain is somewhat limited in this region by the inability of our procedure to investigate both half-cystines simultaneously (*i.e.*, Cys-*6* is on the first segment and Cys-*120* is on the last segment in the minimization cycle). In lysozyme, two disulfide bonds are left handed (−90°) while the other two are right handed (+90°). Cowburn *et al.* (1972) conclude from optical activity measurements that there may be one left- and three right-handed disulfide bonds in α-lactalbumin, in agreement with our model.

Atassi *et al.* (1970) and Habeeb and Atassi (1971) reported that about 2.5 disulfide bonds of native α-lactalbumin are reduced by 2-mercaptoethanol at neutral pH, while none of the disulfide bonds of lysozyme are reduced under the same conditions. On the other hand, Iyer and Klee (1973) found that all four disulfide bonds of native α-lactalbumin can be reduced by dithiothreitol at pH 8 and 25° while, at 0°, only one disulfide bond is reduced. Since the disulfide bonds of lysozyme were reduced much more slowly under the same conditions, they interpret these results in terms of conformational differences, or perhaps a difference in degree of flexibility, between α-lactalbumin and lysozyme.

Our model indicates that the disulfide bond between Cys-*6* and Cys-*120* is exposed to the solvent (similarly to Cys-**6** and Cys-**127** in lysozyme) when the tail is in the T3 or T4 conformation, although it would be somewhat shielded from the solvent if the tail takes on the T2 conformation. Thus, this disulfide bond could be reduced easily when the tail is in the T3 or T4 conformation, but not in the T2 conformation. As noted below, interconversions between these alternative tail conformations may be induced by pH effects and/or solvent perturbation effects. Of the three remaining disulfide bonds, Cys-*28*–Cys-*111* is more exposed to the solvent than the corresponding disulfide bond of lysozyme, because of reorientation and substitution of the surrounding residues (*e.g.*, Arg-**114** in lysozyme is replaced by Leu-*110* in α-lactalbumin). In summary, two disulfide bonds of our computed structure for α-lactalbumin appear to be relatively exposed to the solvent, while the remaining two are no more exposed than the corresponding ones of lysozyme.

G. Environment of Tryptophan Residues. The exposure of the tryptophan residues of α-lactalbumin to various reagents and solvent perturbants has been studied extensively (Herskovits, 1965; Guire, 1970; Barman, 1970; Kronman *et al.*, 1971, 1972a,b; Tamburro *et al.*, 1972; Robbins and Holmes, 1972). However, it is not yet clear from these experiments which tryptophan residues are exposed. In our computed model, Trp-*26* is buried in a hydrophobic pocket (similar to Trp-**28** of lysozyme). Trp-*60* and Trp-*104* are located together in the cleft region in close proximity. This agrees with the observation of Sommers *et al.* (1973) that two tryptophan

residues of α-lactalbumin are close enough to cause quenching of fluorescence. Trp-*104* is partially shielded from the solvent, while Trp-*60* is almost completely buried by the nearby residues Tyr-*103*, Asn-*56*, and Asp-*97* and by the Cys-*73*–Cys-*91* disulfide bond. The degree of exposure to solvent of the fourth tryptophan residue (*118*) is dependent on the particular conformation (T2, T3, or T4) of the tail residues *118*–*123*. In conformation T2, Trp-*118* is on the surface of the protein, exposed to the solvent on at least one side of its ring, whereas in conformations T3 and T4 the other tail residues form a hydrophobic pocket surrounding Trp-*118*. Thus, the computed model predicts that two tryptophan residues (*28* and *60*) are relatively inaccessible to solvent, one (Trp-*104*) is on the surface, while a fourth (Trp-*118*) may be either exposed or buried, depending on the tail conformation.

Solvent perturbation studies indicate that approximately two of the tryptophan residues of α-lactalbumin are relatively exposed to solvent (Herskovits, 1965; Kronman *et al.*, 1972a). The results of Guire (1970) indicate that none of the tryptophan residues of α-lactalbumin are capable of face-to-face stacking with *N*-methylnicotinamide chloride; however, Robbins and Holmes (1972) have reported that one tryptophan residue is capable of such interactions. The latter authors also report that, at pH 2, two tryptophans are capable of forming charge-transfer complexes with *N*-methylnicotinamide chloride, while three tryptophans are exposed at pH 11. One possible explanation for these pH effects would involve exposure of Trp-*118* when tail conformations T3 or T4 open up into conformation T2. As mentioned earlier, tail conformations T3 and T4 are stabilized by a salt bridge from the ε-amino group of Lys-*5* to the free carboxylate group of Leu-*123* (Figure 7). This salt bridge would be weakened at high or low pH, thus favoring a transition to tail conformation T2, in which Trp-*118* is much more exposed to the solvent.

Studies on the photooxidation of the tryptophan residues of α-lactalbumin, in the presence of proflavine as a sensitizer, suggest that two tryptophan groups are oxidized relatively quickly, a third group oxidizes more slowly, while the fourth tryptophan is unreactive (Tamburro *et al.*, 1972). Treatment of α-lactalbumin with 2-hydroxy-5-nitrobenzyl bromide in 10% acetone and 0.17 M acetic acid (pH 2.7) caused considerable reaction of tryptophan residues *26*, *104*, and *118*, whereas Trp-*60* was almost completely unreactive (Barman, 1970). Although the "buried" tryptophan residue **28** of lysozyme also reacts with 2-hydroxy-5-nitrobenzyl bromide to a limited extent (6.8%) under these conditions, the much greater reactivity (43.5%) of the corresponding Trp-*26* in α-lactalbumin leaves some doubt about whether this tryptophan is actually buried, as suggested by the computed structure. It is quite possible that the rather vigorous reaction conditions of this experiment may allow penetration of the reagent to a normally inaccessible site or that prior reaction of the other tryptophan residues with this bulky reagent induces conformational changes which permit reaction of Trp-*26*. Treatment of α-lactalbumin with tetranitromethane has been shown to cause nitration of about 1.4 tryptophan residues (Habeeb and Atassi, 1971).

In summary, most of the experimental evidence suggests that two of the four tryptophan residues of α-lactalbumin are relatively exposed to solvent. In terms of our computed model for α-lactalbumin, these exposed residues are probably Trp-*104* and Trp-*118*, which is quite exposed in tail conformation T2. Since tail conformations T2, T3, and T4 exhibit nearly equal conformational energies, we may anticipate that they

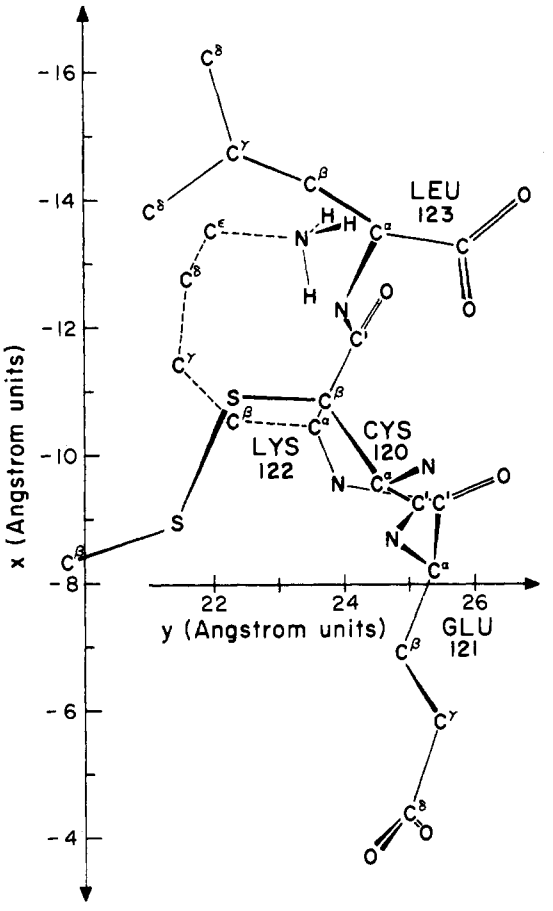

FIGURE 9: View of C-terminal section with tail in T2 conformation.

are readily interconvertible and that various perturbing influences could easily favor one form over another, thus altering the environment of Trp-*118*.

H. Environment of Tyrosine Residues. The reactivity of the four tyrosine residues of α-lactalbumin toward various reagents has also been examined experimentally by a wide variety of techniques, as described below. In our computed structure of α-lactalbumin, Tyr-*18* (similar to Tyr-**20** in lysozyme) and Tyr-*103* (Figures 5 and 6) are both easily accessible to solvent. Try-*36* is buried, with its hydroxyl group pointing inward and hydrogen bonded to the backbone carbonyl of Val-*27*. However, the hydrogen in the ring of Tyr-*36* which is ortho to its OH is somewhat exposed to the solvent, and the entire side chain of Tyr-*36* could be exposed to solvent by reorientation of the tail section (from a T3 to a T2 conformation) if Lys-*5* were to change its conformation simultaneously by folding down. Tyr-*50* (similar to Tyr-**53** in lysozyme) is also buried, with its hydroxyl group pointing toward the surface.

The p*K*'s of the four tyrosines of α-lactalbumin are all in the normal range (between 10 and 11), thus suggesting that they are not deeply buried (Gorbunoff, 1967; Robbins *et al.*, 1967; Kronman *et al.*, 1972a). In lysozyme, two tyrosines ionize with an apparent p*K* of 10.4–10.5, while a third (presumably Tyr-**53**) ionizes very slowly, with a midpoint in its titration curve at 12.8 (Inada, 1961). Gorbunoff (1967) reports that, at pH 9.3, three tyrosyl groups of α-lactalbumin react with cyanuric fluoride, while at pH 10.0 all four groups react. This increased reactivity at higher pH may be interpreted as a manifestation of alkaline denaturation (Robbins

et al., 1967; Kronman *et al.*, 1972a,b). Kronman *et al.* (1971, 1972a) find that two of the four tyrosines of α-lactalbumin are easily acetylated with *N*-acetylimidazole, while the third and fourth tyrosine residues are acetylated only at higher concentrations of reagent. Habeeb and Atassi (1971) find 2.5 tyrosines modified upon nitration with tetranitromethane; however, Denton and Ebner (1971) report that all four tyrosines are destroyed by the nitration reaction, even though only two are converted to nitrotyrosine. The latter authors also find that the ability of α-lactalbumin to bind to galactosyl transferase is lost at a rate which parallels the rate of destruction of tyrosine residues by nitration, iodination, or tyrosinase treatment.

There are several complications in the interpretation of these experiments dealing with tyrosine reactivity. First of all, the identities of the reactive and unreactive tyrosine residues have not been established in any of the above cases. Furthermore, iodination or nitration of a particular tyrosine residue does not conclusively demonstrate that the tyrosine is exposed to solvent in the native state of the protein, since these reagents are known to react with "buried" residues in some proteins, such as cytochrome *c* (McGowan and Stellwagen, 1970; Skov *et al.*, 1969). The observation that α-lactalbumin undergoes major conformational changes in alkaline solutions (Robbins *et al.*, 1967; Kronman *et al.*, 1972a,b) further complicates the interpretation of the results of titration of the phenolic hydroxyl groups and of the reaction with cyanuric fluoride at alkaline pH values. With these reservations in mind, the experimental results on tyrosine exposure are in reasonable agreement with our computed model of α-lactalbumin. Based on this model, we would conclude that tyrosine residues *18* and *103* are those most easily acetylated by acetylimidazole, while Tyr-*36* and Tyr-*50* are the ones which react more slowly. Since the "buried" location of Tyr-*36* is stabilized by a hydrogen bond to the backbone carbonyl of Val-*27*, it may be the one which reacts least rapidly with cyanuric fluoride. On the other hand, it should be noted that Tyr-*50* of α-lactalbumin is in a position homologous to Tyr-**53** of lysozyme, which is unreactive to iodination (Hayashi *et al.*, 1968), and thus is probably also the one which is nonreactive with cyanuric fluoride (Kurihara *et al.*, 1963). Reasoning by analogy to lysozyme, the latter results would suggest that Tyr-*50* of α-lactalbumin may be the one which reacts most slowly with cyanuric fluoride.

I. Environment of Histidine Residues. The three histidine residues of our model appear to be exposed to solvent and reagents to varying degrees. His-*68* is completely exposed to solvent (see Figure 5), His-*32* is partially buried [near Phe-*31*, Trp-*118*, the Cys-*28*–Cys-*111* disulfide bond, and covered somewhat by Val-*110* (see Figure 5)], and His-*107* is partially buried (more so than His-*32*) by Glu-*25* Ser-*112*, and the backbone of residues *101–112*. The only histidine residue in lysozyme is His-**15**, and since the corresponding residue in α-lactalbumin is deleted, no analogies can be drawn.

The reactivity of α-lactalbumin toward iodoacetic acid indicates that carboxymethylation proceeds in the order His-*68* > His-*32* > His-*107* (Castellino and Hill, 1970), which is in good agreement with our model.

J. Environment of Methionine-90. The single methionine residue, Met-*90*, lies in an exposed position on the back surface of our model and is homologous in position to Asn-**93** of lysozyme, which is also exposed to the solvent. Photooxidation and carboxymethylation of Met-*90* (Tamburro *et al.*, 1972; Castellino and Hill, 1970) indicate that it is readily accessible to reagents, in agreement with our model. In apparent conflict with this conclusion is the observation that the circular dichroism spectra of α-lactalbumin and its monosulfoxide differ substantially in the aromatic region (Tamburro *et al.*, 1972), thus suggesting enhanced exposure of tyrosine and tryptophan residues upon oxidation of Met-*90*. However, carboxymethylation of this same Met-*90* had no effect on the ability of α-lactalbumin to form an active lactose synthetase complex (Castellino and Hill, 1970).

K. Reactivity of Carboxyl Groups. The side-chain carboxyl groups in our model of α-lactalbumin are all exposed to solvent to some extent. The results of Lin (1970), in which lysozyme and α-lactalbumin were modified by reaction with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide, indicate that most probably all 20 carboxyls of α-lactalbumin react at pH 4.75, while only 8 of 11 carboxyl groups of lysozyme react under similar conditions. At pH 7, 10 carboxyl groups of α-lactalbumin react, while only 2–3 of lysozyme react. These results, showing that more carboxyls are exposed in α-lactalbumin than in lysozyme at both pH values, are compatible with the models of both proteins.

L. Other Experimental Results. Immunochemical studies of the cross-reactivity of antibodies to lysozyme and α-lactalbumin suggest that these proteins are quite dissimilar with respect to surface antigens (Arnon and Maron, 1970; Atassi *et al.*, 1970; Habeeb and Atassi, 1971). As we have noted above, there are a large number of differences between these proteins, particularly for the exterior residues, and thus, it is not very surprising that they differ immunologically.

Comparative studies of the helix contents of lysozyme and α-lactalbumin suggest that these proteins are quite similar in helical content (Robbins and Holmes, 1970; Barel *et al.*, 1972). Other indications of structural homology between these proteins may be found in the infrared studies of the amide I and II bands and the proton nuclear magnetic resonance spectra of Cowburn *et al.* (1970).

M. Comparison of Our Low-Energy Conformation with the Model of Browne et al. (1969). A direct comparison of the low-energy conformation obtained here with that of Browne *et al.* (1969) is possible only for the backbone dihedral angles up through residue *111*, since they did not report a backbone conformation past this point nor did they report side-chain dihedral angles for any part of their structure. We will therefore concentrate on these data, even though some of the most interesting physical properties of the molecule arise primarily from the unique side-chain conformations and tail positions which we have already described.

The first region of interest occurs at the deletions of residues **14** and **15** in lysozyme. The energy-refined dihedral angles on each side of this deletion are in the same conformational space as those of Browne *et al.* (1969), although deviations of between 30 to 50° in ϕ and ψ occur frequently [*cf.* Table II with p 66–69 of Browne *et al.* (1969)]. In particular, these authors report that ψ of Asp-*14* is −25° (−63° in our refined conformation), ϕ of Leu-*15* is −85° (−42° here), and that ψ of Leu-*15* is −10° (−60° here). In the latter case, the position of the Leu-*15* side chain is quite critical in determining the backbone conformation, because the hydrophobic core in which it is located is tightly packed.

A second region of interest, near the folded region of residues *32–34*, is important because it may involve the active site. There are differences here between our structure and the model of Browne *et al.* (50° in ψ of His-*32;* 70° and 150° in ϕ and ψ, respectively, of Thr-*33;* 130° in ϕ of Ser-*34*); the His-*32* side chain is not found in the cleft region but in the region surrounding the disulfide Cys-*28*–Cys-*111*, where it is

not readily available for interaction with residues in the cleft region.

The backbone conformations in the region of the deletion of residue **47** are similar in both models. No major differences in backbone dihedral angles occur along the chain up to the region of the deletion at residue **100**. From this point on, many rather large differences are found, making a comparison to the model of Browne *et al.* difficult. We have already described in detail the different tail conformations found in this work, which were impossible to assign from model building studies.

Conclusion

The computed conformation of α-lactalbumin seems to be a reasonable one when examined in light of the available experimental data. However, we must emphasize that we have placed rather stringent constraints on the structure by assuming a homology to lysozyme. The computed structure, which is consistent with these constraints, would not necessarily be the one of lowest energy if these constraints were removed. It will be of interest to see how closely this computed structure will correspond to the X-ray structure when it becomes available.

We expect that the X-ray structure will differ in small details from our computed one, but that the overall conformation will agree with our model. It should be noted, however, that α-lactalbumin by itself has no known enzymatic activity, and thus its active conformation in the lactose synthetase complex may differ somewhat from the structure of α-lactalbumin in the absence of galactosyl transferase. Indeed, some experiments on α-lactalbumin in aqueous solution seem to suggest that α-lactalbumin is a rather flexible molecule.

The energy minimization procedures used here, together with related information on the probability of occurrence of bends and helical sections (used *prior* to the final energy minimization), seem well suited as a general approach to the computation of the structures of homologous proteins. On the basis of our experience with lysozyme (Warme and Scheraga, 1974), it seems reasonable to expect that only minor modifications of the structure will be found, when we are able[1] to carry out the stage III energy minimization.

Acknowledgments

We wish to thank Dr. P. K. Ponnuswamy and Dr. D. Rasse for help in the preparation and testing of the parameters of the energy functions used here, and Professors M. Loudon and M. J. Kronman for helpful discussion of the experimental data on α-lactalbumin and lysozyme.

References

Achter, E. K., and Swan, I. D. A. (1971), *Biochemistry 10*, 2976.

Arnon, R., and Maron, E. (1970), *J. Mol. Biol. 51*, 703.

Aschaffenburg, R., Fenna, R. E., Handford, B. O., and Phillips, D. C. (1972a), *J. Mol. Biol. 67*, 525.

Aschaffenburg, R., Fenna, R. E., and Phillips, D. C. (1972b), *J. Mol. Biol. 67*, 529.

Atassi, M. Z., Habeeb, A. F. S. A., and Rystedt, L. (1970), *Biochim. Biophys. Acta 200*, 184.

Barel, A. O., Prieels, J. P., Moes, E., Looze, Y., and Leonis, J. (1972), *Biochim. Biophys. Acta 257*, 288.

Barman, T. E. (1970), *J. Mol. Biol. 52*, 391.

Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1967), *Proc. Roy. Soc., Ser. B 167*, 365.

Brew, K., and Campbell, P. N. (1967), *Biochem. J. 102*, 258.

Brew, K., Castellino, F. J., Vanaman, T. C., and Hill, R. L. (1970), *J. Biol. Chem. 245*, 4570.

Brew, K., Vanaman, T. C., and Hill, R. L. (1967), *J. Biol. Chem. 242*, 3747.

Browne, W. J., North, A. C. T., Phillips, D. C., Brew, K., Vanaman, T. C., and Hill, R. L. (1969), *J. Mol. Biol. 42*, 65.

Castellino, F. J., and Hill, R. L. (1970), *J. Biol. Chem. 245*, 417.

Cowburn, D. A., Bradbury, E. M., Crane-Robinson, C., and Gratzer, W. B. (1970), *Eur. J. Biochem. 14*, 83.

Cowburn, D. A., Brew, K., and Gratzer, W. B. (1972), *Biochemistry 11*, 1228.

Denton, W. L., and Ebner, K. E. (1971), *J. Biol. Chem. 246*, 4053.

Dickerson, R. E. (1971), *J. Mol. Evol. 1*, 26.

Gorbunoff, M. J. (1967), *Biochemistry 6*, 1606.

Guire, P. (1970), *Biochim. Biophys. Acta 221*, 383.

Habeeb, A. F. S. A., and Atassi, M. Z. (1971), *Biochim. Biophys. Acta 236*, 131.

Hartley, B. S. (1970), *Phil. Trans. Roy. Soc. London, Ser. B 257*, 77.

Hayashi, K., Shiroda, T., Inoto, T., and Funatsu, M. (1968), *J. Biochem. (Tokyo) 64*, 365.

Herskovits, T. T. (1965), *J. Biol. Chem. 240*, 628.

Inada, Y. (1961), *J. Biochem. (Tokyo) 49*, 217.

IUPAC-IUB Commission on Biochemical Nomenclature (1970), *Biochemistry 9*, 3471.

Iyer, K. S., and Klee, W. A. (1973), *J. Biol. Chem. 248*, 707.

Krigbaum, W. R., and Kugler, F. R. (1970), *Biochemistry 9*, 1216.

Kronman, M. J., and Andreotti, R. E. (1964), *Biochemistry 3*, 1145.

Kronman, M. J., Andreotti, R. E., and Vitols, R. (1964), *Biochemistry 3*, 1152.

Kronman, M. J., Hoffman, W. B., Jeroszko, J., and Sage, G. W. (1972a), *Biochim. Biophys. Acta 285*, 124.

Kronman, M. J., Holmes, L. G., and Robbins, F. M. (1971), *J. Biol. Chem. 246*, 1909.

Kronman, M. J., Jeroszko, J., and Sage, G. W. (1972b), *Biochim. Biophys. Acta 285*, 145.

Kurihara, K., Horinishi, H., and Shibata, K. (1963), *Biochim. Biophys. Acta 74*, 678.

Lewis, P. N., Momany, F. A., and Scheraga, H. A. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2293.

Lewis, P. N., Momany, F. A., and Scheraga, H. A. (1973a), *Biochim. Biophys. Acta 303*, 211.

Lewis, P. N., Momany, F. A., and Scheraga, H. A. (1973b), *Isr. J. Chem. 11*, 121.

Lewis, P. N., and Scheraga, H. A. (1971), *Arch. Biochem. Biophys. 144*, 584.

Lin, T. Y. (1970), *Biochemistry 9*, 984.

McGowan, E. B., and Stellwagen, E. (1970), *Biochemistry 9*, 3047.

Momany, F. A., McGuire, R. F., and Scheraga, H. A. (1973), *J. Phys. Chem.*, to be submitted.

Phillips, D. C. (1970), private communication of X-ray coordinates of lysozyme.

Rawitch, A. B. (1972), *Arch. Biochem. Biophys. 151*, 22.

Rees, A. R., and Offord, R. T. (1972), *Biochem. J. 130*, 965.

Robbins, F. M., Andreotti, R. E., Holmes, L. G., and Kronman, M. J. (1967), *Biochim. Biophys. Acta 133*, 46.
Robbins, F. M., and Holmes, L. G. (1970), *Biochim. Biophys. Acta 221*, 234.
Robbins, F. M., and Holmes, L. G. (1972), *J. Biol. Chem. 247*, 3062.
Skov, K., Hofmann, T., and Williams, G. R. (1969), *Can. J. Biochem. 47*, 750.
Sommers, P. B., Kronman, M. J., and Brew, K. (1973), *Biochem. Biophys. Res. Commun. 52*, 98.
Tamburro, A. M., Jori, G., Vidali, G., Scatturin, A., and Saccomani, G. (1972), *Biochim. Biophys. Acta 263*, 704.
Venkatachalam, C. M. (1968), *Biopolymers 6*, 1425.
Warme, P. K., Go, N., and Scheraga, H. A. (1972), *J. Comput. Phys. 9*, 303.
Warme, P. K., and Scheraga, H. A. (1973), *J. Comput. Phys. 12*, 49.
Warme, P. K., and Scheraga, H. A. (1974), *Biochemistry 13*, 757.

Composition, Fluorescence, and Circular Dichroism of Rat Lysozyme†

Roderick S. Mulvey, Richard J. Gualtieri, and Sherman Beychok*

ABSTRACT: The fluorescence and circular dichroism of rat lysozyme have been investigated. The fluorescence emission spectrum has λ_{max} 328 nm and Q = 0.034, at pH 7.5. These values are almost identical with those of human lysozyme. The effects of the inhibitor tetrasaccharide $(AcGlcN)_4$ and pH on the rat lysozyme fluorescence were also determined. The results parallel corresponding studies with human lysozyme and suggest that the fluorescence of the rat enzyme is also dominated by Trp-109 (108). The near- and far-ultraviolet circular dichroism (CD) spectra of rat lysozyme are described and compared to those of human lysozyme. A striking similarity is observed, also, between the CD spectra of these two proteins. In particular, the near-ultraviolet CD spectrum of rat lysozyme develops a negative band near 313 nm at alkaline pH. The amino acid composition of rat lysozyme was determined, the tyrosine and tryptophan contents being established by a spectrophotometric method. These data are compared to the compositions of human and mouse lysozyme. Discussion of the spectral properties of rat lysozyme is based upon previous studies with human lysozyme and on the known sequences of the human and mouse enzymes.

Since the determination of the three-dimensional structure of hen egg-white lysozyme (Blake *et al.*, 1965) considerable interest has been shown in the homologous lysozymes and α-lactalbumins. In particular, the chemical and physical properties of human lysozyme and bovine α-lactalbumin have been examined in a number of laboratories. Constructive analysis of these properties is greatly facilitated by comparison of the homologous sequences and by reference to the hen egg-white lysozyme structure (see, for example, Sommers *et al.*, 1973; Mulvey *et al.*, 1973, 1974). The validity of the approach rests on the assumption of a high degree of similarity in tertiary structures among homologous proteins. With respect to the lysozymes, the assumption is, for the most part, supported by X-ray crystallographic studies of human lysozyme (Blake and Swan, 1971; Banyard *et al.*, 1973).

Recently the nearly complete sequence of mouse lysozyme has been described (Riblet, 1974). We considered it worthwhile, therefore, to examine some of the spectroscopic properties of the closely related rat lysozyme. Here, we report the amino acid composition of rat lysozyme, including a spectrophotometric determination of tryptophan and tyrosine content, and compare it with the composition of the mouse enzyme. The circular dichroism and fluorescence emission spectra have been determined, as well as the effects of pH and inhibitor binding on the latter.

Experimental Section

Materials. Three times crystallized hen egg-white lysozyme was purchased from Pentex and used without further purification. Human and rat lysozymes were a generous gift of Dr. E. F. Osserman. The human lysozyme was prepared as described previously (Mulvey *et al.*, 1973). The rat enzyme had been isolated from the urine of rats bearing a transplantable Shay chloroleukemia by elution from bentonite in 5% aqueous pyridine at pH 5. This material was received in a lyophilized state, and was further purified on a CM-32 cellulose column and finally deionized on Bio-Rad AG 501-X8 mixed bed resin. The enzyme had a tendency to precipitate on the deionizing resin, and after lyophilization its solution showed considerable light scattering at long wavelengths (350–400 nm). Since no other proteins were detectable by either chromatography or electrophoresis on polyacrylamide gels the light scattering was probably caused by partially denatured aggregates of lysozyme. Stock solutions of rat lysozyme were routinely passed through Millipore filters before use.

† From the Departments of Biological Sciences and Chemistry, Columbia University, New York, New York 10027. *Received September 24, 1973.* This work was supported in part by Grants GB 29481 from the National Science Foundation and 5 R01 CA 13014 from the National Institutes of Health. Part of the work presented here belongs to the thesis submitted by R. S. M. to Columbia University in partial fulfillment of the requirements for the Ph.D. degree.