Computer-Aided Model-Building Strategies for Protein Design†

C. O. Pabo* and E. G. Suchanek

Department of Biophysics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received March 12, 1986; Revised Manuscript Received June 23, 1986

ABSTRACT: Model-building strategies for protein modification and design are developed. These strategies emphasize simple geometric aspects of protein structure, use local coordinate systems defined at particular residues, and systematically consider a large number of alternative sequences and conformations. We have written a computer program, PROTEUS, to implement these search methods. PROTEUS has been used to find positions where disulfide bonds could be added to the N-terminal domain of the λ repressor and to predict how a loop on the surface of repressor could be shortened.

Genetic engineering and DNA synthesis allow one to produce any desired polypeptide sequence, but it often is difficult to predict which sequences will be use useful. Any model-building strategy encounters several fundamental problems: (*1*) The combinatorial complexity of the design problem can be overwhelming. If any 1 of 20 amino acids can be at any position in a protein, the number of possible sequences is astronomical. (2) Our knowledge of the forces that stabilize protein structure is limited. (3) It often is difficult to state the structural problems in ways that are easy to visualize or that aid in planning an effective search strategy.

We are developing model-building strategies for protein modification and design. In an attempt to deal with these three problems, we have chosen to (1) *use a computer-based model-building system.* When changes at a number of residues are considered, there is a combinatorial explosion in the number of sequences and conformations to be tested. It becomes impossible to systematically analyze all the structures with physical models or a computer graphics system. A computer-based model-building strategy lets one test more possibilities. It also forces one to have explicit criteria for evaluating tentative models and allows one to keep track of precisely which structures have been considered. (2) *We used simple geometric aspects of protein structure.* Geometric data about bond angles and interatomic distances are the most fundamental information obtained from structural analyses of proteins. Even when the forces controlling protein structure are not fully understood, these simple geometric data may be a useful guide for model building. (3) *We used local coordinate systems so that the geometric relationships are easier to visualize.* Local coordinate systems may be defined at particular residues or side chains. They can make it easier to visualize the model-building process and to compare the environment of residues or side chains in different proteins.

We are developing a computer program, PROTEUS, that implements these model-building strategies. The program uses standard geometries for the peptide backbone and for the amino acid side chains (Momany et al., 1975). When modifying a protein, PROTEUS generates conformations that are similar to known proteins and tests how well the proposed modification fits with the remaining (unmodified) portion of the protein. The most plausible models can be examined in detail on a graphics system or evaluated with energy calculations. To test this general strategy, we have used PROTEUS to locate positions where disulfide bonds might be added to the N-terminal domain of the λ repressor and to determine how a surface loop of repressor might be altered. The following paper (Sauer et al., 1986) describes the construction and characterization of a covalent repressor dimer that contains one of these predicted disulfide bonds. Adding this disulfide enhanced the thermal stability and DNA binding activity of the repressor's N-terminal domain.

Experimental Procedures

Local Coordinate Systems. Our model-building strategies make extensive use of local coordinate systems, which are often defined by the orientation of particular residues or side chains. One standard coordinate system (Figure 1) has (1) a residue's α-carbon at the local origin, (2) a local *x* axis pointing toward the carbonyl carbon, (3) a local *y* axis in the plane defined by the α-carbon, the carbonyl carbon, and the nitrogen, (4) a *z* axis which completes a right-handed orthogonal system.

Our model-building strategies use these local coordinate systems in several ways. After a local coordinate system is established at one residue (residue i), the coordinates of any other residue (residue j) can be expressed in terms of this local system. This concisely describes the spatial relationship of the two residues. Comparing local coordinates provides a convenient way of comparing the spatial relationship of residues i and j with the spatial relationship of any other two residues (i′ and j′).

Local coordinate systems also allow us to use "turtle geometry" (Abelson & diSessa, 1980) for model building. One can visualize local coordinate systems in terms of the position and orientation of a "turtle", with the local origin giving the "position", the local *x* axis giving the "heading", the local *y* axis pointing "left", and the local *z* axis pointing "up". (The turtle's position is defined simply by giving its *x*, *y*, and *z* coordinates with respect to the global coordinate system; the vectors *heading*, *left*, and *up* are described by giving their direction cosines with respect to the global system.) Motions of the turtle can be used to build segments of the protein. The *heading* vector is usually aligned with a bond, and moving the turtle in the direction of the *heading* vector translates the turtle along a bond from one atom to the next. Dihedral angles are generated by rotating the turtle around the *heading* vector. Bond bend angles are generated by rotating around the *up* vector. The program PROTEUS is written in Common LISP (Steele, 1984) and FORTRAN and runs on a VAX 11/750 computer. Programming in LISP usually consists of defining a large number of functions. PROTEUS contains approximately

† This research was supported by a University Exploratory Research Grant from Procter and Gamble. Computing facilities in the Department of Biophysics were supported by grants from the National Institutes of Health, the National Science Foundation, and the Richard King Mellon Foundation.

0006-2960/86/0425-5987$01.50/0

500 LISP functions that can be used to build and analyze protein structures. A few functions which involve extensive numerical computations, such as the disulfide search and loop-building procedure described below, have been written in FORTRAN to improve program efficiency; VAX-LISP allows PROTEUS to call these functions from the LISP environment. For convenience in maintaining and modifying PROTEUS, it has been divided into a number of sections. Some of the most frequently used sections are the following.

(1) TURTLE contains functions that interchange coordinate systems and functions that manipulate the turtle. For example, MOVE translates the turtle in the direction of the *heading* vector, moving the turtle along a bond from one atom to the next. ROLL generates dihedral angles, and YAW generates bond bend angles.

(2) GEOMETRY contains functions which perform basic geometric calculations (dot products, cross-products, etc.).

(3) AMINO stores fundamental data about properties of the 20 amino acids. These data include the covalent structure, the hydrophobicity (Wolfenden, 1983), the Chou–Fasman secondary structure parameters (Chou & Fasman, 1978), the volume of each side chain (Chothia, 1975), and other properties.

(4) PRODAT creates and manages the data structures needed to store information (such as dihedral angles and Cartesian coordinates) about particular protein structures. This section contains functions that can build a structure from a list of dihedral angles, check particular ϕ and ψ values against a table of allowed ϕ–ψ combinations (Ramachandran & Sasisekharan, 1968), generate random conformations with acceptable ϕ–ψ angles, etc. Functions in this section make frequent use of the functions MOVE, ROLL, and YAW.

Other sections contain the functions used for particular model-building problems or strategies. For example, DISULFIDE contains routines to build and analyze disulfide bonds. LOOP contains functions that generate and evaluate conformations for surface loops. When PROTEUS is run, all these sections are loaded into the computer, and functions from any section are accessible.

Results

To test our model-building strategies, we used PROTEUS to find positions where disulfide bonds might be added to the N-terminal domain of λ repressor. Finding locations for disulfide bonds seemed like a reasonable initial test for computer-aided model-building strategies. Since two residues may be changed, the problem is difficult enough that a systematic analysis on an interactive graphics system becomes tedious and unreliable. Our search used the known three-dimensional structure of the N-terminal dimer (Pabo & Lewis, 1982), which contains no cysteines. The goal was to find any pair of residues in the model that might, after introducing the required cysteines, be connected by a disulfide bond.

Our strategy assumes that *if* the backbone atoms of any two residues in the repressor dimer have the same spatial relationship as the backbone atoms of two cysteines in a disulfide bond, *then* these residues in repressor might provide a plausible position for introducing a disulfide. Disulfide bonds from known protein structures were characterized by describing the relationship of the backbone atoms of the two half-cystine residues. One residue was used to establish a local coordinate system, and coordinates for the backbone atoms of the second residue were given in this system (Figure 1). Ninety disulfide bonds from structures in the Proten Data Bank (Bernstein et al., 1977) were used to prepare a data base of known disulfide conformations (Figure 2), and the N-terminal fragment of

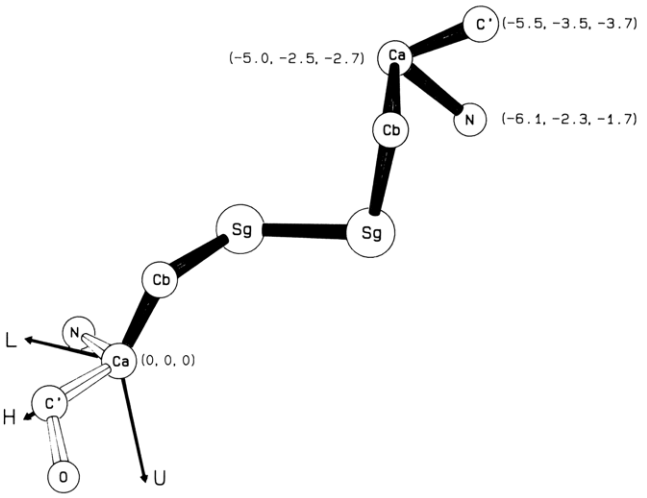

FIGURE 1: Local coordinate system used to describe the spatial relationship of residues in a disulfide bond. A local coordinate system is defined by the residue on the left. The origin is at the α-carbon, the local x axis or "heading" vector (H) points toward the carbonyl carbon, and the local y axis or "left" vector (L) is in the plane containing the α-carbon, the carbonyl carbon, and the nitrogen. The local z axis or "up" vector (U) completes a right-handed coordinate system. Coordinates of the backbone atoms in the distal residue are given in terms of this local coordinate system.

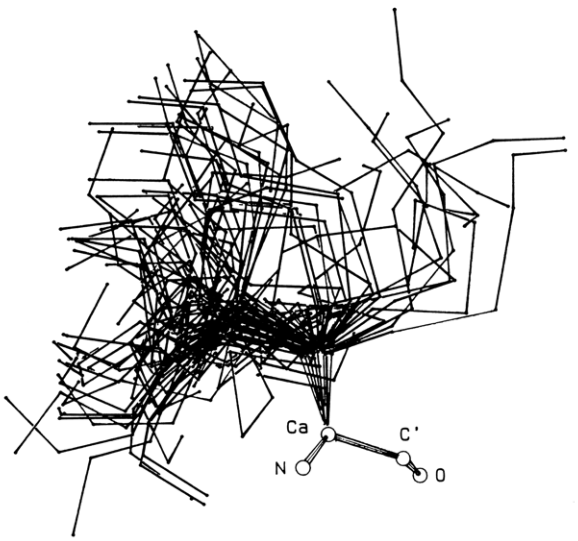

FIGURE 2: Disulfide bond conformations from the Protein Data Bank (Bernstein et al., 1977). These have been plotted with respect to a common origin to suggest how each of the conformations was tested at every position of repressor.

repressor was systematically searched for positions where any of these could be accommodated. The search proceeded as follows: (1) A local coordinate system was defined at one residue (the ith residue) of λ repressor, and coordinates for all other backbone atoms in the dimer were converted to this local system. (2) Local coordinates for the backbone atoms of every residue ($j \neq i$), expressed with respect to residue i, were compared with local coordinates of distal residues in the disulfide data base (expressed with respect to the proximal residue of the disulfide). Whenever the root mean square (rms) difference in coordinates was small enough (typically a cutoff of 1.5 Å was used), it suggested that a disulfide bond with this conformation might connect residue i and residue j. When the fit appeared reasonable, the orientation of residue j was also tested by building the disulfide from the other direction. To do this, the distal residue of the disulfide from the data base was superimposed on residue j of the repressor (by superimposing the N, C_α, and C atoms). The program

Table I: Possible Disulfide Bonds in λ Repressor[a]

proximal	distal	error	source
88	88′	0.3	left-handed spiral
91	73′	1.3	insulin 6–11 (C chain)
14	77	1.2	insulin 7–7 (C–D chains)
74	81	0.7	phospholipase A2 (61–91)

[a] The first two columns show which residues would be connected by a new disulfide bond. The error column indicates the average of the rms error, in angstroms, when fitting the three distal backbone atoms (N, C_α, and C) after building from the proximal side and of the error after building in the opposite direction. The fourth column gives the source of the disulfide conformation with the best fit at this position.

then checked the rms distance between the backbone atoms of the proximal residue and the backbone atoms of the residue i. (3) The process was repeated with $i = 4, 5, 6, ..., 92$. (The N-terminal fragment has 92 residues, but residues 1, 2, and 3 are not well-ordered in the electron density map.) This initial screen for plausible disulfide locations was written in FORTRAN, and the search of the repressor dimer required 15 min of CPU time on a VAX 11/750 computer.

Analysis of disulfide bonds has suggested that a left-handed spiral conformation ($\chi_1 = -60°$, $\chi_2 = -90°$, $\chi_3 = -90°$, $\chi_{2'} = -90°$, $\chi_{1'} = -60°$) is especially stable (Richardson, 1981; Thornton, 1981). A right-handed hook conformation ($\chi_1 = -60°$, $\chi_2 = 120°$, $\chi_3 = 90°$, $\chi_{2'} = -50°$, $\chi_{1'} = -60°$) also appears to be favorable. To ensure that these conformations were thoroughly tested, we also prepared a data base containing 243 conformations closely related to the left-handed spiral (each dihedral angle was varied by ±15°) and 243 conformations related to the right-handed hook. The search was repeated with each of these data bases.

Plausible locations for disulfide bonds are listed in Table I. The search indicated that a cysteine at residue 88 could form an intermolecular disulfide bond with residue 88 of the other monomer (Figure 3B). This appeared to be the best location for a disulfide: it used the most favorable disulfide conformation (a left-handed spiral) and also had the lowest rms error. However, the search strategy used to compile Table I only ensured that the backbone atoms would have an appropriate orientation for disulfide bond formation. To check for van der Waals collisions and other problems, the disulfide bond was added to the model. The tyrosine side chains were removed, and the left-handed spiral conformation which had given the best fit ($\chi_1 = -45°$, $\chi_2 = -105°$, $\chi_3 = -105°$, $\chi_{2'} = -105°$, $\chi_{1'} = -45°$) was built first from one side and then from the other. A weighted average of coordinates was used to estimate the positions of C_β, S_γ, $S_{\gamma'}$, and $C_{\beta'}$. The resulting "disulfide bond" was examined on the PS 300 graphics system, and these coordinates were also used as a starting point for energy refinement (Brooks et al., 1983). Refinement only led to small shifts in the coordinates, indicating that a disulfide bond could be accommodated at this position without significant distortion. This prediction was also tested experimentally (Sauer et al., 1986). Oligonucleotide-directed mutagenesis (Dalbadie-McFarland et al., 1982) was used to change Tyr-88 to Cys. The protein was produced in *Escherichia coli* as a fragment containing residues 1–102 and was purified. The disulfide bond forms spontaneously in vitro. At low salt concentrations, the disulfide-bonded N-terminal fragment binds λ operator DNA more tightly than does the wild-type N-terminal fragment. The disulfide-bonded dimer is 6–10° more stable than the wild-type protein and also is more stable to urea denaturation.

We have also tested similar model-building strategies for planning deletions in loop regions. Designing loops is formally

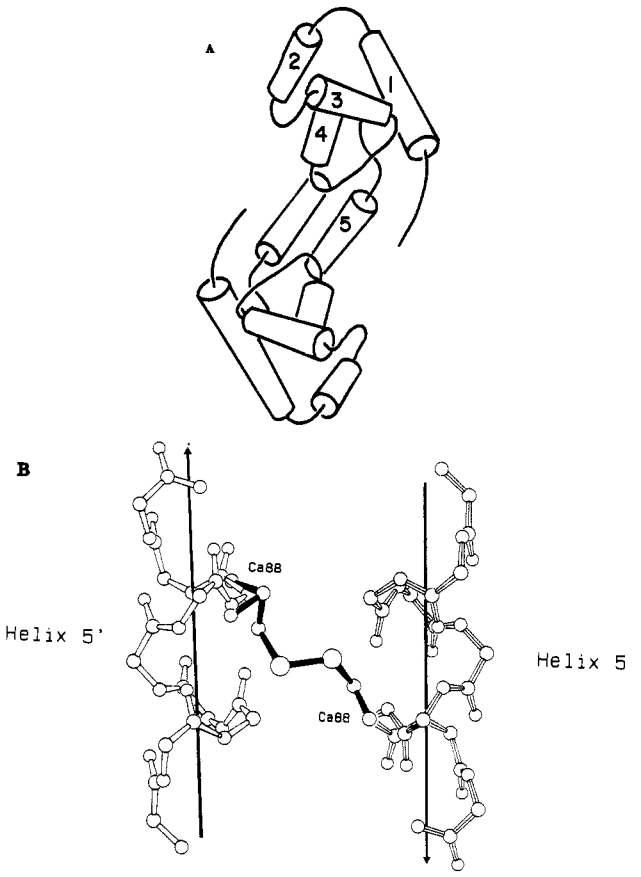

FIGURE 3: (A) Sketch showing the N-terminal dimer of λ repressor, with cylinders used to represent α-helices. Helices in one monomer are numbered, and the 2-fold axis would be next to helix 5. [Used with permission from Pabo & Lewis (1982).] (B) Sketch showing how a disulfide with a left-handed spiral configuration could connect residue 88 on one monomer to residue 88 on the other. In this model, the disulfide is attached to helix 5 on the right, and the distal atoms superimpose almost perfectly on residue 88 of helix 5′. The helices are oriented as in (A), and the arrows point toward the C-terminal end of each helix.

analogous to designing disulfide bonds: in each case, one wishes to know whether there is an allowed conformation that can connect two residues, and local coordinate systems can be used to characterize the relative orientation of residues. Loop design introduces several new complications: (1) A systematic analysis must consider many different loop lengths. Because loops will have several residues, the number of allowed backbone conformations may be extremely large. (2) The allowed backbone conformations will depend on the sequence of the loop. The standard Ramachandran plot can be used for most residues, but a larger range of dihedral angles will be acceptable for a glycine residue (Ramachandran & Sasisekharan, 1968), and cis peptide bonds should be considered when prolines are introduced.

As a test for our loop-building strategy, we tried to shorten the loop between helix 1 and helix 2 of repressor (Figures 3A and 4). This seemed reasonable because helices 1 and 2 of repressor have a limited structural homology to helices 1 and 2 of cro, which are connected by a much shorter loop (Ohlendorf et al., 1983). We decided to remove residues 22 through 32 from the repressor model and to build the shortest loop that might connect helix 1 to helix 2. The last residue *before* the gap (residue 21) was used to define a local coordinate system, and the first residue of a proposed loop was also placed at the origin of this local system (thereby superimposing it on residue 21). The program then constructed loops by randomly picking ϕ–ψ values from a list of allowed pairs

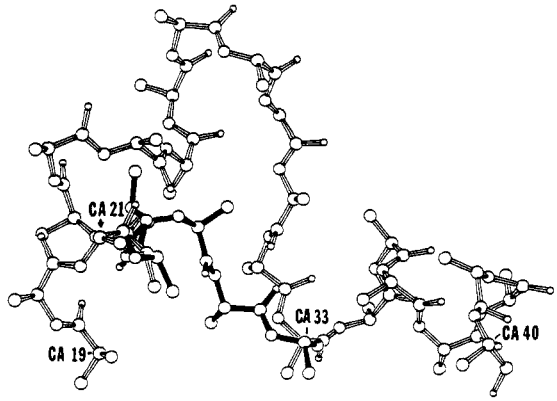

FIGURE 4: Model for loop that might connect helices 1 and 2 of repressor. Helices have roughly the same orientation as in Figure 3A. The original loop containing residues 22–32 has open bonds; the proposed, shortened loop has solid bonds. α-Carbons of selected residues are labeled.

[assigned at 10° intervals from the plots of Ramachandran & Sasisekharan (1968)] and using PROTEUS to build backbone segments with idealized geometry. (The program uses a separate ϕ–ψ list for glycine, cis peptide bonds are considered before proline, and the user can decide whether to try only the fully allowed regions of the ϕ–ψ tables or to try both fully and partially allowed regions.) For each loop configuration, the program determined whether the backbone atoms of the last residue in the loop were close to the backbone atoms of residue 33. (A rms cutoff of 1.5 Å was typically used.) Whenever this happened, the program checked whether the loop could really be attached at the other side of the gap (residue 33). To do this, the N, C_α, and C atoms from the last residue in the loop were superimposed on the corresponding atoms from residue 33. Using the same loop configuration, the program then checked that the backbone atoms of the first residue in the proposed loop were close to the backbone atoms of residue 21.

In attempting to connect residue 21 to residue 33, we tested loops that contained 3, 4, 5, or 6 residues. Several million conformations were tested for each loop size. Glycines and prolines were inserted randomly, typically with frequencies of 0.1–0.3. A random-number generator was used to pick allowed ϕ–ψ pairs from the appropriate table. The search indicated that several closely related conformations of a four-residue loop (which did not contain glycine or proline) could make this connection and one of these conformations is shown in Figure 4. Additional model building and energy refinement of the resulting structure, which has a net deletion of seven residues, ensured that this loop did not give any bad van der Waals collisions. Interestingly, this loop would make the connection between helices 1 and 2 of repressor similar to the connection between helices 1 and 2 of cro (Ohlendorf et al., 1983). It is important to recognize that the loop-building protocol described here does not attempt to pick the best sequence for a loop—it merely determines the length of the loop region and determines which residues should be glycine or proline. Other model-building methods will be needed to pick the best sequence. Experimental tests may begin by synthesizing random codons through the loop region and attempting to select functional repressor mutants.

DISCUSSION

The model-building strategies described in this paper allow systematic analysis of changes that might be made within a protein of known structure. The program PROTEUS generates a large number of possible conformations and evaluates each with respect to clearly defined geometric criteria. Approaching the problem of protein modification and design in this manner has a number of advantages.

(1) The program can test many more conformations than a person could check at a graphics system, and this is especially useful when several amino acids are being changed. The program can also keep a precise record of all the conformations that have been considered.

(2) The program has explicit cutoffs that are used when screening possible structures, and one always knows exactly what criteria were used in the design process. If similar changes—such as the insertion of disulfide bonds—are tried in a number of different proteins, the use of such explicit criteria may provide a way of comparing structural predictions.

(3) Visualizing the model-building process in terms of a turtle, with the position and orientation of the turtle describing a local coordinate system, may have significant advantages as PROTEUS is developed. The turtle construct provides a powerful way of integrating the model-building process with information about the position of neighboring atoms and residues; model building is concisely described in terms of the turtle's motions, while the local coordinate system also provides a convenient framework for defining the position of neighboring atoms.

Does a good geometric fit ensure that a disulfide bond actually will form? A disulfide with a left-handed spiral configuration fits extremely well at position 88 of the repressor dimer (Figure 3; Table I). This prediction has been tested, and it has been shown that the disulfide forms spontaneously and stabilizes the protein against denaturation (Sauer et al., 1986). However, many modified proteins must be characterized before we can determine how effective this strategy is at picking locations for disulfide bonds. When the orientation of the residues is less favorable than it was for Cys-88, disulfide bond formation will depend on the "plasticity" of a protein, that is, how readily a protein can make small structural changes to accommodate the new disulfide bond. Energy minimization methods might be helpful at this stage. Clearly, many factors other than the quality of a rigid geometric fit must be considered when evaluating a proposed disulfide bond, and PROTEUS checks a number of other criteria. (1) van der Waals collisions with atoms in the disulfide are an obvious problem, and PROTEUS lists possible collisions. (2) For intramolecular bonds, free energies of loop closure are estimated (Poland & Scheraga, 1965). In general, a disulfide which closes a larger loop will have a greater stabilizing effect. (3) To assess overall packing differences, the van der Waals volume of a disulfide bond is compared with the sum of the volumes of the existing side chains (Chothia, 1975). The structural or functional roles of the existing residues must also be considered; these are currently checked by examining the proposed disulfide on the graphics system.

Given our limited understanding of the factors that allow disulfide bond formation in proteins, we have been forced to make a number of somewhat arbitrary decisions about modeling strategies, allowed conformations, and rms cutoffs. We have used rather stringent criteria for modeling disulfides and may have missed some locations in repressor that would be acceptable. As a control, the program has been tested with T4 lysozyme and *E. coli* dihydrofolate reductase, two proteins where disulfides have been added by directed mutagenesis (Villafranca et al., 1983; Perry et al., 1984). In T4 lysozyme, PROTEUS correctly predicts that residues 3 and 97 might form a disulfide bond. (The initial screening, which only checks the backbone orientation, also suggests that disulfides might

be added between residues 28 and 63 and between residues 1 and 158. However, evaluation on the graphics system suggests there may be other structural problems at these positions.) Residues 39 and 85 in dihydrofolate reductase, which can form a disulfide bond after treatment with dithionitrobenzoic acid, are not predicted by PROTEUS with a 2.0-Å rms cutoff. We also are testing PROTEUS by using it to analyze disulfide-containing proteins from the Protein Data Bank. We have not finished a systematic analysis, but PROTEUS finds most of the sites which actually do contain disulfide bonds. (This does not count the trivial case in which the conformation of a particular disulfide bond causes PROTEUS to flag that particular location.) Clearly, the accuracy of our predictions depends on the validity of our modeling strategy and on the quality of the disulfide data base. We expect our predictions to improve as more information becomes available from refined protein structures or from improved energy calculations.

The basic strategies developed in PROTEUS, which involve systematically generating and testing a large number of conformations, should be useful when considering other problems in protein modification and design. Minor changes in PROTEUS should allow it to find plausible locations for chemical cross-links between side chains, for salt bridges, or for favorable contacts between aromatic side chains (Burley & Petsko, 1985). For example, one could prepare a data base of known salt bridge conformations and then search a protein structure for any position where a pair of residues have the same spatial arrangement. PROTEUS should also be useful for de novo protein design (Drexler, 1981). We have proposed that de novo design could use an "inverted" approach to the protein folding problem, picking a suitable fold for the backbone and then picking the most appropriate side chain to add at each position (Pabo, 1983). The problem of de novo design may highlight differences between the rule-based approach taken by PROTEUS and approaches based on energy minimization strategies. It should be possible to use rule-based design when assembling a partial model, but energy-based design will not make much sense until the model is virtually complete. A systematic model-building approach to de novo design will also require simple methods for describing the spatial arrangement of secondary units. At this stage, it might be useful to define local coordinate systems with respect to particular α-helices or β-sheets. For example, a polar coordinate system might be used, with the z axis defined by the α-helical axis (Chou et al., 1984). The loop-building protocol may also be useful in de novo design. Plausible structures could be assembled by *first* packing units of secondary structure together and *then* finding loops to connect these units.

ACKNOWLEDGMENTS

We thank Bobby Barnett, John Sullivan, and Irwin Chaiken for helpful discussions and Brian Matthews for sending us refined coordinates of T4 lysozyme.

REFERENCES

Abelson, H., & diSessa, A. (1980) *Turtle Geometry*, MIT Press, Cambridge, MA.

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

Brooks, B., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Burley, S. K., & Petsko, G. A. (1985) *Science (Washington, D.C.) 229*, 23–28.

Chothia, C. (1975) *Nature (London) 254*, 304–308.

Chou, K.-C., Nemethy, G., & Scheraga, H. A. (1984) *J. Am. Chem. Soc. 106*, 3161–3170.

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.

Dalbadie-MacFarland, G., Cohen, L. W., Riggs, A. D., Morin, C., Itakura, K., & Richards, J. H. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6409–6413.

Drexler, K. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 5275–5278.

Momany, R. A., McGuire, R. F., Burgess, A. W., & Scheraga, H. A. (1975) *J. Chem. Phys. 79*, 2361–2381.

Ohlendorf, D. H., Anderson, W. F., Lewis, M., Pabo, C. O., & Matthews, B. W. (1983) *J. Mol. Biol. 169*, 757–769.

Pabo, C. O. (1983) *Nature (London) 301*, 200.

Pabo, C. O., & Lewis, M. (1982) *Nature (London) 298*, 443–447.

Perry, L. J., & Wetzel, R. (1984) *Science (Washington, D.C.) 226*, 555–557.

Poland, D. C., & Scheraga, H. A. (1965) *Biopolymers 3*, 379–399.

Ramachandran, G. N., & Sasisekharan, V. (1968) *Adv. Protein Chem. 23*, 283–437.

Richardson, J. (1981) *Adv. Protein Chem. 34*, 167–339.

Sauer, R. T., Hehir, K., Stearman, R. S., Weiss, M. A., Jeitler-Nilsson, A., Suchanek, E. G., & Pabo, C. O. (1986) *Biochemistry* (following paper in this issue).

Steele, G. L. (1984) *Common Lisp: The Language*, Digital Press.

Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261–287.

Villafranca, J. E., Howell, E. E., Voet, D. H., Strobel, M. S., Ogden, R. C., Abelson, J. N., & Kraut, J. (1983) *Science (Washington, D.C.) 222*, 782–788.

Wolfenden, R. (1983) *Science (Washington, D.C.) 222*, 1087–1093.