(2)

AD-A210 096

DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. SECURITY CLASSIFICATION AUTHORITY: U	1b RESTRICTIVE MARKINGS: NA
2a. SECURITY CLASSIFICATION AUTHORITY: NA	3 DISTRIBUTION / AVAILABILITY OF REPORT: Distribution Unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE: NA	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): 5L406	5 MONITORING ORGANIZATION REPORT NUMBER(S): NA
6a. NAME OF PERFORMING ORGANIZATION: Johns Hopkins University; 6b OFFICE SYMBOL (If applicable): NA	7a NAME OF MONITORING ORGANIZATION: Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code): Department of Molecular Biology & Genetics, 725 North Wolfe Street, Baltimore, MD 21205	7b ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217-5000
8a. NAME OF FUNDING / SPONSORING ORGANIZATION: Office of Naval Research; 8b OFFICE SYMBOL (If applicable): ONR	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-86-K-0196
8c. ADDRESS (City, State, and ZIP Code): 800 N. Quincy St., Arlington, VA 22217-5000	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 61153N; PROJECT NO RR04106; TASK NO NR441g003; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Computer-Aided Design of Thermostable Proteins

12. PERSONAL AUTHOR(S)
Carl O. Pabo

13a. TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
final	FROM 3/15/86 TO 3/14/89	June 7, 1989	

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	protein design / thermostability directed mutagenesis / protein engineering

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	U
22a. NAME OF RESPONSIBLE INDIVIDUAL: M. Marron	22b TELEPHONE (Include Area Code): 202-696-4760; 22c OFFICE SYMBOL: ONR

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

We are developing methods for computer-aided protein design and are testing these strategies by constructing thermostable variants of the lambda repressor. Repressor's DNA-binding domain normally denatures at 54°C, and we have constructed a quadruple mutant that is stable to 71°C and binds DNA as well as the wild type protein.

Our fundamental goal is to develop methods for *de novo* protein design, and we are proceeding by treating the problem of protein design as an "inverted" version of the protein folding problem (Pabo, 1983). In protein folding, one is given an amino acid sequence and must predict how this folds in three dimensions. Protein design can be approached in quite a different way - one can begin by choosing a folded arrangement of the polypeptide backbone and then try to pick an amino acid sequence that will stabilize this structure. "Inversion" eliminates the problem of predicting long-range interactions, since residues which will interact in the final tertiary or quaternary structure already are close in space when they are added to the prefolded backbone. One should be able to pick residues which will have favorable interactions with their neighbors.

We are developing a program, called PDB_PROTEUS, for computer-aided protein design. Our program uses simple geometric aspects of protein structure and frequently uses local coordinate systems so that the geometric relationships are easier to visualize (Pabo and Suchanek, 1986). There are many advantages to using a program rather than relying on simple visual inspection when designing changes: a program can easily check millions of possible sequences and conformations. Using a program also makes it easy to try several variations of a particular search strategy or to apply the same strategy to many different proteins.

PDB_PROTEUS

Much of our effort has focussed on developing and refining the PDB_PROTEUS system. Although the program is written in FORTRAN, we have tried to develop a programming strategy that will be very flexible. The core of the system is a library of subroutines. Each performs a discrete operation - like adding a residue or changing the

coordinate system - and our library now contains several hundred subroutines. These subroutines are used in two different ways: 1) The main programs use the subroutines (almost like a higher level programming language). 2) A menu driven system (called DESIGN_TOOLS) allows convenient, interactive access to individual subroutines. Since most subroutines proceed by reading, modifying and rewriting files written in the Protein Data Bank format, DESIGN_TOOLS can be used as a "high level editor" for modifying coordinate files.

Disulfide Bonds

In our attempts to stabilize the lambda repressor, we have written programs to search the repressor structure for the best residues to change and then have experimentally tested each of these predictions. When searching for places to introduce disulfide bonds, the search program used all the disulfide bond conformations found in the Protein Data Bank and also used a library of conformations that were closely related to the left handed spiral configuration (Richardson, 1981). Modeling suggested that an intermolecular disulfide bond could be introduced by changing Tyr 88 to Cys (Pabo and Suchanek, 1986). Experimental studies showed that this disulfide bond forms spontaneously, stabilizes repressor against thermal denaturation, and increases the affinity for DNA (Sauer *et. al.*, 1986).

Salt Bridges

The program searches for any position where a new salt bridge could be introduced by changing a single residue. The best position appeared to be at the C-terminal end of the helix 5, where changing Ser 92 to Lys should allow a salt bridge with Glu 89. This has no effect on the thermal stability of repressor, but introducing Lys 93 (effectively adding a residue to the C-terminal end of the helix) does stabilize the protein by about 0.5°C. It is possible that salt bridges on the surface do not contribute much to thermal stability, but we need to test additional positions and also should search for places where two amino acid substitutions would give a good salt bridge.

Aromatic

Studies of aromatic-aromatic interactions in proteins suggest that these can stabilize a protein if the aromatic rings are about 5.5 A apart and are approximately perpendicular to each other (Burley and Petsko, 1985). We have searched for places where aromatic residues could be added to make favorable contacts with an existing aromatic residue. Unfortunately, the only position that appears plausible (residue 33) changes a key residue involved in nonspecific contacts with the DNA. Studies in Robert Sauer's laboratory at M.I.T. have shown that this mutation increases the stability to thermal denaturation, (Hecht *et. al.* 1984) but it is not useful to us because it disrupts DNA binding.

Glycine to Alanine Changes

Hecht and Sauer (1986) have shown that repressor can be stabilized by changing both glycine 46 and glycine 48 to alanine. We have set up a program that automatically searches for places that Gly to Ala changes might be made. The program does not find any other plausible positions in repressor, it confirms that the backbone angles and side chain accessibility at positions 46 and 48 are favorable for introducing alanine, and the program should be useful with other proteins.

Proline

Proline residues may stabilize proteins by reducing the conformational entropy of the unfolded protein. Obviously, they can introduce unfavorable strain if they are put at the wrong positions, but we have written a program to search repressor for positions where the backbone conformation and side chain accessibility should allow a proline residue to be introduced. Two positions appeared plausible and have been tested. We found that changing Tyr 60 to Pro has a mild destabilizing effect, but changing Gln 9 to Pro stabilizes the repressor by 0.6°C.

Combining Stabilizing Mutations

Since a set of changes may be needed to dramatically stabilize a protein, it was important to determine whether the effects of multiple mutations were additive. Our initial results are quite encouraging. To test the effects of multiple mutations, we

combined our disulfide mutant with the two glycine to alanine changes in helix 3. We found that the wild type protein denatured at 54°, the Cys 88 mutant denatured at 62°, the Ala46Ala48 double mutant denatured at 62°, and the Ala46Ala48Cys88 mutant was stable to 70° (Stearman *et. al.*, 1988). More recently, we have shown that the Ala46Ala48Cys88Lys93 quadruple mutant is stable to 71°.

Metal-binding Sites

Although this requires more drastic changes in repressor, we are trying to introduce metal binding sites (using tetrahedral coordination with cysteines and/or histidines) to stabilize repressor. As a first step, a set of structural "rules" were established by examining metal-binding proteins in the Brookhaven Data Bank. We then wrote a program (based on the PROTEUS subroutines) that "builds" cysteines and histidines off each position in the backbone and finds sets of residues that can form a reasonable site. The program identified two positions in the lambda repressor as possible candidates for a tetrahedral metal binding site. Both sites required two additional substitutions in order to sterically accomodate the binding site. To date, one of these proteins (which required that we change 6/92 amino acids in the N-terminal domain!) has been constructed. This protein precipitates when expressed at high levels *in vivo*, but it can be resolubilized and studied. Preliminary work suggests that the protein does bind zinc. Unfortunately, Zn binding appears to make the protein less soluble, and preliminary experiments do not show any DNA binding.

Although we have just begun to explore this approach, our first experiments suggest that the *de novo* introduction of metal sites will be difficult. The introduction of a tetrahedral metal-binding site almost inevitably requires the removal of buried, hydrophobic residues. The removal of such residues, which often play a critical role in the folding and stability of proteins, certainly complicates the design problem. One alternative approach that might be useful is to try designing tetrahedral coordination sites near the surface which have only three amino acid ligands and use water as the fourth ligand. (Such sites are frequently found in enzymes which use metals as an element in the active site.)

Refinement

While these modeling and design projects were in progress, we also have continued with crystallographic refinement of repressor.

Our experience shows that a highly refined structure is very important for modeling and design. Our initial predictions had used the represssor structure obtained by fitting an isomorphous electron density map at 3.2 A resolution (Pabo and Lewis, 1982). We now have much better data from our repressor-operator cocrystals (Jordan and Pabo, 1988; Beamer, Jordan and Pabo, unpublished) and this structure has been refined to an R factor of 20.6% using data from 8.0 to 2.5 A resolution. Comparisons have shown that our model-building predictions are very sensitive to differences between these coordinate sets. The initial, less accurate, coordinates gave several predictions (not discussed above because they were not obtained with the better coordinates) that were thermally unstable.

Perpectives

Our experiences allow us to make several general conclusions, comments and suggestions about the prospects for rational protein design:

1) We have proven that it is possible to use computer-aided design to plan changes that will stabilize a protein. We were able to dramatically stabilize the lambda repressor without interfering with DNA-binding activity.

2) Stabilizing changes can be combined to make hyperstable proteins.

3) Modeling and design are significantly easier if a high-resolution structure is available, since relatively small changes in the coordinates can drastically affect the modeling.

4) Not all changes predicted by the modeling will actually stabilize the protein. As emphasized by recent calculations (Gao *et. al.* 1989), changing a single residue can have very complicated thermodynamic effects: It can {add and/or remove} {favorable and/or unfavorable} interactions with the {protein and/or solvent} in the {folded and/or unfolded state}. Given this complex balance of forces, we cannot expect simple modeling methods to give accurate predictions every time. It clearly will be necessary to test individual mutations and then combine the most favorable changes.

5) Although it is too early for a firm conclusion, our data suggest that it may be easiest to stabilize a protein by introducing mutations that reduce the entropy of the unfolded form. Designing appropriate sites for chemical crosslinks may be a promising strategy for future work.

Burley, S. K. and Petsko, G. A. (1985) Aromatic-Aromatic Interaction: A Mechanism of Protein Structure Stabilization. **Science 229**, 23-28.

Gao, J. Kuczera, K., Tidor, B., and Karplus, M. (1989) Hidden Thermodynamics of Mutant Proteins: A Molecular Dynamics Analysis. **Science 244**, 1069-1072.

Hecht, M. H., Sturtevant, J.M. and Sauer, R.T. (1984) Effects of Single Amino Acid Replacements on the Thermal Stability of the NH2-terminal Domain of Phage Lambda Repressor. **Proc. Natl. Acad. Sci. USA 81**, 5685-5689.

Hecht, M. H., Sturtevant, J. M. and Sauer, R.T. (1986) Stabilization of Lambda Repressor Against Thermal Denaturation by Site-Directed Gly to Ala Changes in Helix 3. **Proteins 1**, 43-46.

Jordan, S. R. and Pabo, C. O. (1988) Structure of the Lambda Complex at 2.5 A Resolution: Details of the Repressor-Operator Interactions. **Science 242**, 893-899.

Pabo, C. O. (1983) Molecular Technology: Designing Peptides and Proteins. **Nature 301**, 200.

Pabo, C.O. and Lewis, M. (1982) The Operator-binding Domain of Lambda Repressor: Structure and DNA Recognition. **Nature 298**, 443-447.

Pabo , C. O. and Suchanek, E. G. (1986) Computer-Aided Model-Building Strategies for Protein Design. **Biochemistry 25**, 5987-5991.

Richardson, J. (1981) Protein Anatomy and Taxonomy. **Advances Protein Chemistry 34**, 167-339.

Sauer, R. T., Hehir, K., Stearman, R. S., Weiss, M. A., Jeitler-Nilsson, A., Suchanek, E.G. and Pabo, C. O. (1986) An Engineered Intersubunit Disulfide Enhances the Stability and DNA-Binding of the N-terminal Domain of Lambda Repressor. **Biochemistry 25**, 5992-5998.

Stearman, R. S., Frankel, A. D., Freire, E., Liu, B. and Pabo, C. O. (1988) Combining Thermostable Mutations Increases the Stability of Lambda Repressor. **Biochemistry 27**, 7571-7574.