FINAL REPORT

GRANT #: N00014-94-1-0192 R & T CODE: 3412-179

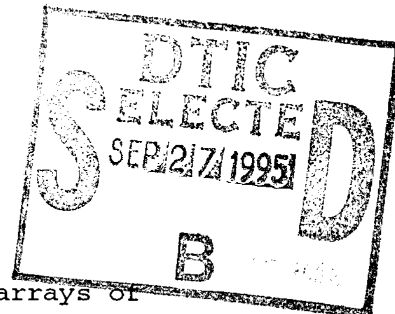
PRINCIPAL INVESTIGATOR: Donald Hilvert

INSTITUTION: The Scripps Research Institute

GRANT TITLE: Generation of Novel Protein Receptors

REPORTING PERIOD: 1 December 1993 - 30 November 1994

AWARD PERIOD: 1 December 1993 - 30 November 1994



OBJECTIVE: To develop a general strategy for creating arrays of relatively small protein receptors via mutagenesis and selection.

APPROACH: Families of protein receptors will be created from a small, stable template. The experimental approach involves synthesizing the gene for the template protein and cloning it into a bacteriophage vector to allow phage display of the protein. In this way, replication and recognition can be coupled. Repeated rounds of mutagenesis and affinity chromatography will be employed to select the desired binding properties out of the diverse population created from the original protein scaffold.

ACCOMPLISHMENTS: As a starting structure for building arrays of protein receptors, we have selected a *de novo* designed protein (NBP), abstracted from the NAD^+ binding domain of lactate dehydrogenase.[1] The gene for NBP was synthesized chemically, expressed in *E. coli,* and the resulting protein recovered as inclusion bodies in high yield (30-40 mg of 85-90% pure protein from a 1L culture). The protein was purified in 6 M urea by a combination of ion exchange and size exclusion column chromatography techniques. A method of refolding the protein and forming the single disulfide bond was developed and the chemical integrity of the protein was confirmed by mass spectrometry. The conformation of the refolded protein was probed using a variety of physical techniques, including fluorescence spectroscopy, circular dichroism spectroscopy, size exclusion chromatography, and NMR. The characterization suggested that the protein does not adopt a defined, compact, and monomeric structure in solution.

In collaboration with Dr. Jeffrey Skolnick and coworkers, the designed protein was evaluated using inverse folding techniques[2] and several key areas selected for modification. Eleven variants of NBP were designed and the genes for these variants constructed by site directed mutagenesis. Expression of the eleven modified proteins revealed that four of these variants, unlike NBP, were expressed as soluble protein in high yield. In all four variants, the single disulfide bond that had been introduced in the original design had been removed. Moreover, the two most highly expressed variants of these four shared in common the reintroduction of an α/β unit that had been removed in the original design. The latter two proteins were studied in detail. Both proteins were purified from large scale cultures by affinity column chromatography after incorporating a $(His)_6$ tag at the C terminus of the proteins. The purified proteins were characterized using the same techniques as before. Increased secondary structure was observed by

19950925 126

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DTIC QUALITY INSPECTED 5

circular dichroism spectroscopy. In addition, fluorescence spectroscopy indicated that the tryptophan probe was more buried than in the original protein, as anticipated. Denaturation studies suggested that an ensemble of conformations is sampled by these proteins and size exclusion column chromatography revealed that the proteins are significantly swollen relative to natural proteins of the same molecular weight. Both proteins appeared to be more stable than the original protein, as judged by resistance to proteolytic digestion.

As a comparison to these proteins, a variant bearing the reintroduced α/β unit yet still retaining the disulfide bond was also studied. As above, the variant protein was purified from large scale culture by affinity column chromatography after incorporating a $(His)_6$ tag at the C terminus. The single disulfide bond was oxidized as in the original protein and the conformation studied using the same techniques. Although a more compact structure was observed by size exclusion chromatography (presumably due to the covalently constrained topology of the polypeptide chain), the protein exhibited less secondary structure than the variants lacking the disulfide bond as judged by circular dichroism spectroscopy and less burial of the tryptophan probe as judged by fluorescence spectroscopy.

As a final modification to the variant exhibiting the most structure (bearing the α/β unit yet lacking the disulfide bond), the tryptophan probe that had been introduced into the core of the protein as a fluorescence probe was removed by site-directed mutagenesis, thereby restoring the core sequence to that found in lactate dehydrogenase. Large scale production and purification of the $(His)_6$ tagged protein and conformational analysis suggested a slight increase in secondary structure as judged by circular dichroism spectroscopy, but still no compact tertiary structure as judged by size exclusion chromatography.

In order to obtain a protein exhibiting a stable tertiary fold that recognizes and binds NAD^+, the gene for the most stable variant described above was randomly mutagenized by error-prone PCR.[5] The resulting ensemble of $>10^7$ variants was then displayed on the surface of filamentous bacteriophage and phage bearing variants that recognize NAD^+ sought by carrying out multiple rounds of affinity chromatography coupled with amplification of the cognate phage.[4] After five such rounds, no increase in the number of cognate phage was observed, suggesting the need for either further refinement of the procedure or more extensive mutagenesis of the protein.

CONCLUSION: Based on the conformational characterization of the original designed protein, we conclude that the abstraction of this single domain from a larger, multimeric protein resulted in a species lacking the necessary enthalpic stability to overcome the entropic cost involved in forming a stable tertiary structure. Analysis by inverse folding provided a means of analyzing the design and identifying areas that required further refinement. Although variants with improved structural properties were obtained, these variants did not exhibit the necessary stability to function as small-molecule receptors. Further improvement of these variants may be achieved in the future by exploiting the power of random mutagenesis coupled with *in vitro* selection using phage display technology.

perletter
Distribution/
Availability Codes
Avail and/or
Dist Special
A-1

SIGNIFICANCE: Combining the latest advances in *de novo* protein design and powerful *in vitro* selection methods has the potential to become a general approach for creating novel receptor arrays and catalysts. These would be invaluable for a variety of chemical and biological applications, for example as sensors, diagnostic agents, and enzyme mimics. In addition to such practical benefits, characterization of the receptors will aid future efforts at protein design by illuminating the complex relationships between protein structure, function, folding and stability.

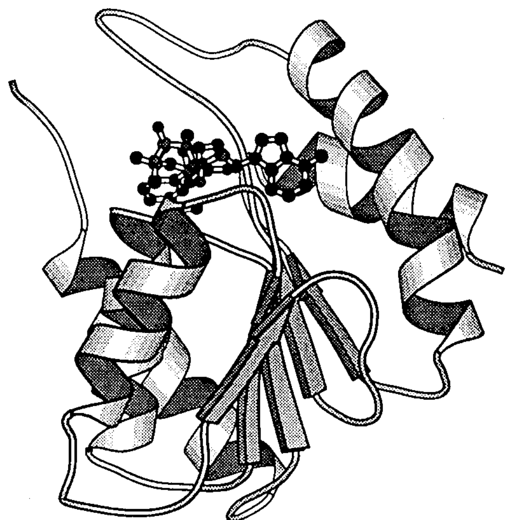
REFERENCES:

1. Sander, C. *et al.* (1992) *Proteins: Struct. Funct. Genet.* **12**, 105.
2. Godzik, A. *et al.* (1992) *J. Mol. Biol.* **227**, 227.
3. Jecht, M. *et al.* (1994) *Protein Sci.* **3**, 411.
4. Marks, J.D. *et al.* (1992) *J. Biol. Chem.* **267**, 16007.
5. Cadwell, R.C. and Joyce, G.F. (1992) *PCR Methods and Applications* **2**, 28.

GENERATION OF NOVEL PROTEIN RECEPTORS

D. Hilvert, The Scripps Research Institute; 1994

Model of best variant



Accomplishments

- NBP synthesized, purified, and characterized
- Design evaluated by inverse folding techniques and 11 variants constructed and screened
- Best two variants studied in detail and two other variants analyzed for comparison
- *In vitro* evolution of the best variant initiated

Objectives

- Synthesize, purify, and characterize a designed nicotinamide-binding protein (NBP)
- Improve NBP by rational approaches and *in vitro* evolution
- Generate arrays of novel NBP-based protein receptors, catalysts

Significance

- Uncover principles of protein structure and stability
- Combine *de novo* design with *in vitro* evolution
- Develop general methods of generating arrays of receptors
- Potential biosensors, catalysts