A Synthetic Lattice for Structure Determination of Uncharacterized Proteins

Julie E. Norville NORVILLE@CSAIL.MIT.EDU
MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge MA, 02139 USA

1. Introduction

Proteins are constructed from amino acid chains which are folded into complex three-dimensional structures. Structural knowledge provides a molecular snapshot of the protein. Molecular pictures can assist scientists who wish to understand proteins, engineers who wish to modify either the proteins' structures or functions, and drug designers who wish to cause large effects within an organism due to small perturbations involving proteins at the molecular scale. The structure of many proteins remains unexplored. In order to determine a protein's structure it is not sufficient to examine only one molecule because it will be destroyed during the imaging process. Instead, one must examine an ensemble of molecules. One type of easily interpreted ensemble of molecules for structural determination is a crystalline lattice.

Unfortunately, protein crystallization is an art rather than a science. The target protein is screened against a diverse array of solutions until crystals are observed. Usually the first crystals that appear are not suitable for structure determination. To grow larger or more ordered crystals, it is usually necessary to change the pH, ionic strength, or temperature from the starting conditions. The process is painful, unreliable, and time consuming, and membrane proteins, which do not readily dissolve in water, are even more difficult to crystallize.

This paper describes a general method for preparing two-dimensional crystals from either monomeric proteins or symmetric protein complexes. Though formation of protein crystals is challenging, several classes of proteins form two-dimensional crystalline arrays in nature. I propose using such a natural crystalline array as a template; proteins can be bound to this lattice, using the natural structure of the array to form two-dimensional crystals. While common structural determination techniques require three-dimensional crystals, there are standard methods for which two-dimensional crystals are sufficient (Murata et al., 2000), (Saibil, 2000). In each of the following sections, the key parameter is that the three-dimensional structure of the target protein is undetermined.

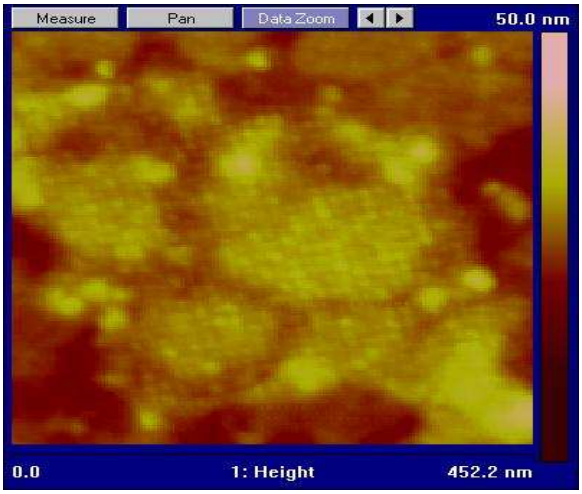

Figure 1. The protein we are using forms a crystalline lattice.

2. Making a Crystal out of a Crystalline Protein

A crystalline lattice is an ordered ensemble of molecules sufficient for structural determination. In this case the natural crystalline lattice will act as a support scaffold for the target molecule. Together they will create a new "synthetic" crystalline lattice. There exist natural proteins that form crystalline lattices. In some cases practical concerns prohibit the use of certain crystalline lattices. For example, a number of membrane proteins form two-dimensional crystalline lattices. However, they are hard to produce. In addition, other proteins require being at the right state before they can be formed into a lattice. Instead it was desirable to choose a protein that could be purified in large quantities and easily forms a lattice. S-layer proteins, or surface layer proteins, are a class of proteins that satisfy many of the criteria for an ideal crystalline lattice scaffold. In nature these proteins exist as a crystalline sheet on the outermost layer of many types of bacteria (Sleytr & Sara, 1997). The s-layer protein that best met these first conditions, *SbpA* did not yet have a determined structure. Thus, now that we have mastered the purification of the protein, we are currently in the process of determining its structure

from two-dimensional crystals.

3. Choosing a Target Protein

The choice of the scaffold lattice constrains the choice of target protein for structure determination. It must fit within the lattice of the crystalline scaffold formed by the protein *SbpA*. *SbpA* is a tetrameric protein with a unit cell of size 13nm. Thus the target protein must either be less than 6.5nm X 6.5nm in size or it must be a symmetric tetramer smaller than 13 nm in size. Proteins that do not fit within these constraints would require a different crystalline scaffold. (The class of S-layer proteins provides many possible candidates with different lattice spacing and symmetry groups.) The dimensions of the target protein can be determined by examining single particles with electron microscopy. Symmetry can be determined using a combination of gel electrophoresis and analytical centrifugation.

4. Binding the Target Protein to the Crystalline Lattice

There are many ways in which the target protein can be bound to the crystalline lattice. A suitable binder must attach the target to the lattice so that it assumes one position. Currently it is unclear whether to use a structurally flexible binder that would it allow the target protein to find its own spatial equilibrium or to use a structurally rigid binder that will constrain the target protein in place. I am currently using a strategy in which a small peptide tag which binds to the crystalline lattice is genetically engineered onto the target protein. Ideally the binding tag will bring the target protein down to specific positions on the scaffold lattice. Currently we are in the process of finding a small peptide that binds to the crystalline lattice using the technique called phage display. Phage display allows one to select binding sequences from a library of 10^9 possible sequences. The phage display determined tag could be bound to the target protein in either a structurally loose or constrained fashion and has the possibility of being genetically engineered onto many types of target proteins.

5. Conclusions and Remaining Challenges

This paper provides the starting point for building synthetic crystals for structure determination. Currently, we will need to complete the structural determination of the protein *SbpA* which forms the crystalline lattice. If it does not diffract to high resolution, then another crystalline protein may be more suitable. We will also need to explore experimentally the binders which will connect the target protein and the lattice.

6. Acknowledgements

I thank my advisors Angela Belcher and Tom Knight. For assistance with the AFM imaging of *SbpA*, shown in the figure I thank Dan Solis. The author received graduate fellowships from the National Science Foundation, Merck, and Vinton Hayes. This work is supported by a Packard Fellowship, DARPA/ONR N00014-01-1-1060, and SRC task ID 1267.001.

References

Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J., Engel, A., & Fujiyoshi, Y. (2000). Structural determinants of water permeation through aquaporin-1. *Nature*, *407*, 599–605.

Saibil, H. (2000). Macromolecular structure determination by cryo-electron microscopy. *Acta Cryst. D*, *56*, 1215–1222.

Sleytr, U., & Sara, M. (1997). Bacterial and archaeal S-layer proteins: structure-function relationships and their biotechnological applications. *Trends in Biotechnology*, *15*, 20–26.