Enliven Archive
Intensifying Thoughts

Short Communication www.enlivenarchive.org Enliven: Bioinformatics

ISSN: 2376-9416

Got Crystals? Try Racemic Protein Crystallography

Jack Merrin

Staff Scientist, Institute of Science and Technology, Austria

*Corresponding author: Jack Merrin, Staff Scientist, Institute of Science and Technology, AM Campus 1, 3400 Klosterneuburg, Austria, E-mail: Jack.Merrin@ist.ac.at

Received Date: 11 Mar 2014

Accepted Date: 15 Mar 2014

Published Date: 20 Mar 2014

Citation:Merrin J (2014) Got Crystals? Try Racemic Protein Crystallography. Enliven: Bioinform 1(1): 001.

Abstract

Verification of bioinformatics structure predictions depends heavily on the protein data bank. The best protein structures are generally acquired through X-ray crystallography, but the main difficulty of this technique is crystallization. One approach to generating crystals is to start with a solution containing one half the original protein and one half the mirror image protein, a racemic mixture. The mirror image protein can be directly synthesized from D-amino acids, up to 50 amino acids long. These molecules can be combined together with a technique called chemical ligation to create larger proteins. Racemic crystals generally form centrosymmetricachiral crystals which is a constraint that aids in the computation of the structure. Achiral racemic crystals are also energetically favored to crystallize relative to chiral crystals. Racemic protein crystallography is an essential approach for proteins that do not form a normal crystal.

Short Communication

The protein databank [1,2] is a database of protein structures and is vital in the support of the bioinformatics enterprise. The PDB is particularly useful in the problem of predicting protein structure, the so called protein folding problem [3-5]. Without the PDB, bioinformatics could not proceed because theoretical predications need to be verified by experiment. Computers are now powerful enough to simulate small proteins for short times until they fold [6]. Success by this approach is rare, but I still emphasize these must be checked against an actual experimental determination of the structure.

There are a variety of methods for determining protein structures which have their advantages and disadvantages. Generally the best method to produce the highest resolution structures is X-ray crystallography which is not limited by the size of the protein or beam damage. The other useful methods are protein NMR [7] and cryoelectron microscopy [8]. NMR is limited to small proteins and the resolution is slightly less than with X-ray crystallography. Cryoelectron microscopy potentially works on single molecules but the beam damage is severe. Cryoelectron microscopy can be used to find the out shape of a protein by coating the protein in electron dense elementsor can be used to get a picture of how molecular complexes form. The main difficulty with X-ray crystallography is how to form the crystals. There is no general knowledge a priori for what conditions to use, whether a crystal will actually form, or what the space group of the crystal will be. Typically a large synthesis takes place and then a robot aliquots different conditions in multiwall plates. Some proteins are difficult to crystallize and resist potentially 1000's of different conditions. Sometimes only microcrystals can be formed, but these can be studied with robotic manipulation and micro x-ray beams [9,10]. How are we to study protein structures if we can't get protein crystals? Recent developments called racemic protein crystallography facilitate the easy formation of crystals and the analysis of their structure.

The first separation of left and right handed molecules, so called chiral molecules, was performed by Louis Pasteur with tartaric acid from wine. Chiral molecules rotate light differently depending on whether they are left of right handed. Natural amino acids are left handed except glycine which is neither left nor right handed. Proteins are polymers of amino acid monomers and are therefore also left handed and chiral. There is no natural machinery to produce a right handed version of a protein. They must be directly synthesized using D-amino acids by organic chemistry. The L-Protein and D-protein are called a pair of enantiomers. An equal mixture of the enantiomers is racemic and does not rotate light. One would not expect a D-protein to have any function, at worst it might be toxic. For example,the protein plectasin is anantimicrobrial agent, while D-plectasinis inactive[11]. D-amino acids can have some function though. For example in bacteria, certain D-amino acidsact as a signal for decomposing biofilms [12].

There are 65 space groups that chiral molecules can form crystals in. When you make a racemic crystal it opens up another 165 possible achiral space groups for a total of all known 230 space groups. Proteins fall into these space groups in certain proportions. An explanation for the proportions was suggested by Wukovitz and Yeates [13]. In brief, there are a number of degrees of freedom in which proteins in a crystal interact. When the number is higher, then that geometry is favored. For chiral space groups, this number can range from 4 to 7. For achiral space groups the number can range up to 8 for the P1 (bar) space group which is optimally favored. For achiral crystals, P21/c and C2/c have 7. Most racemic crystals fall into one of those 3 groups. Since the number is high it also signifies that it is more energetically favorable to form a crystal.

The first synthesized racemic protein crystal was rubredoxin [14]. This crystal formed into the P1(bar) space group which was most likely. P1(bar) is centrosymmetric which means the phase angle to be determined by x-ray crystallography is either 0 or 180 degrees. For a normal protein crystal, the angle could be any number. This limitation facilitates the solution of the protein structure due the limited phase angle. Direct synthesis, like with rubredoxin's45 amino acids, can be performed up to about 50 amino acids. These pieces can be joined together by a method called chemical ligation [15]. Chemical ligation is expected to work for up to 8 pieces at a total length around 300 amino acids [16]. It is generally found that racemic crystals form crystals easier than either of the enantiomer alone. This has some theoretical support and experimental supportas mentioned in a recent survey [17]. If a protein crystallizes normally with zero percent chance, then it crystallizes closer to 50 percent chance with racemic protein crystallography for the limited number of cases tested.

Racemic protein crystallography seems quite useful. Why aren't more groups adopting this technique? There are only a few groups using racemic protein crystallography and the number of structures determined this way is in the tens compared to the 100,000 structures in the protein data bank. Perhaps most structural biology groups lack the nuts and bolts knowhow to do the tour de force of organic chemistry. Perhaps it is easier to move on to the next protein with the same crystallization methods than to learn a new synthesis method.

In science, we often don't hear about negative results. Not much is reported on the structural biology of proteins where crystals could not be formed. By reporting negative crystallization results, we could form a database which would be of high value. Primarily, people wouldn't waste their resources trying to crystallize proteins by the same methods that were used before. These difficult proteins would be prime candidates for racemic protein crystallography. We expect to see a high proportion of these no crystal proteins solved by racemic protein crystallography. Perhaps then some of the protein structures that have eluded us thus far will be more easily revealed.

Eventually the obstacles to mainstream racemic protein crystallography will be overcome. Perhaps, the techniques will become more widespread or a commercial racemic protein service will emerge. They might even crystalize it for you too. Then, it may only cost a few thousand dollars to get a crystal. Maybe also in the not to near future a standalone commercial machine will become available that can synthesize proteins akin to the new DNA printers. With such technology in place we would enter an era of high throughput structural proteomics. Protein structures from humans will be important for developing new medicines. The structural proteome of a given organism would be an interesting new level of understanding. Crystals on demand will be important if structural biology is to keep pace with gene sequencing. Also importantly, when we have a good data bank of protein structures with a wide variety of folds, our ability to make better bioinformatics predictions is enhanced.

References

1. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Brice MD, et al. (1977) The protein data bank. Eur J Biochem 80: 319-324.
2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The protein data bank. Nucleic Acids Res 28: 235-242.
3. Chan HS, Dill KA (1993) The protein folding problem. Physics today 46: 24-32.
4. Dill KA, Ozkan SB, Shell MS, Weikl TR (2008) The protein folding problem. Annu Rev Biophys 37: 289-316.
5. Dill KA, MacCallum JL (2012). The protein-folding problem, 50 years on. Science 338: 1042-1046.
6. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE (2011) How fast-folding proteins fold. Science 334: 517-520.
7. Roberts G, Lian LY (2011) Protein NMR spectroscopy: practical techniques and applications. John Wiley & Sons.
8. Hoenger A (2014) High-resolution cryo-electron microscopy on macromolecular complexes and cell organelles. Protoplasma 251: 417-427.
9. Abola E, Kuhn P, Earnest T, Stevens RC (2000) Automation of X-ray crystallography. Nat Struct Biol 7: 973-977.
10. Cusack S, Belrhali H, Bram A, Burghammer M, Perrakis A, et al. (1998) Small is beautiful: protein micro-crystallography. Nat Struct Biol 5: Suppl634-637.
11. Mandal K, Pentelute BL, Tereshko V, Thammavongsa V, Schneewind O, et al. (2009) Racemic crystallography of synthetic protein enantiomers used to determine the X-ray structure of plectasin by direct methods. Protein Sci 18: 1146-1154.
12. Kolodkin-Gal I, Romero D, Cao S, Clardy J, Kolter R, et al. (2010) D-amino acids trigger biofilm disassembly. Science 328: 627-629.
13. Wukovitz SW, Yeates TO (1995) Why protein crystals favour some space-groups over others. Nat Struct Biol 2: 1062-1067.
14. Zawadzke LE, Berg JM (1992) A racemic protein. J American Chemical Soc 114: 4002-4003.
15. Dawson PE, Muir TW, Clark-Lewis I, Kent SB (1994) Synthesis of proteins by native chemical ligation. Science: 266: 776-779.
16. Kent SB (2009) Total chemical synthesis of proteins. Chem Soc Rev 38: 338-351.
17. Yeates TO, Kent SB (2012) Racemic protein crystallography. Annu Rev Biophys 41: 41-61.