Journal & Proceedings of the Royal Society of New South Wales, vol. 152, part 3, 2020, pp. 343–345. ISSN 0035-9173/20/030343-03

One sequence, one structure: demise of a dogma, or fake news

Ronald J. Clarke

School of Chemistry, University of Sydney, Sydney, Australia

E-mail: ronald.clarke@sydney.edu.au

Abstract

For the last 50 years Anfinsen's dogma of one sequence, one structure has been a central principle of biochemistry and molecular biology. However, the inability of X-ray crystallography to resolve structures of many proteins and protein segments, leads naturally to the conclusion that proteins can also exist in intrinsically disordered states. This necessarily requires a revision to Anfinsen's dogma. Thus, the amino acid sequence of a protein must encode for both order and disorder.

Introduction

Since the discovery of the double-helical structure of DNA by Watson and Crick (1953) and the subsequent resolution of the genetic code (Leder and Nirenberg, 1964), it has been known that each amino acid residue in a protein is encoded by a sequence of three nucleic acid bases (a codon). Thus, the sequence of DNA bases in a gene determines the amino acid residue sequence, i.e., the primary structure, of the protein for which the gene encodes. In 1956, Christian Anfinsen, a biochemist with the National Institutes of Health U.S.A., carried out a very influential experiment linking the amino acid sequence to the protein's final three-dimensional folded structure, i.e. its tertiary structure (the term secondary structure is reserved for local polypeptide chain conformations). What Anfiinsen did was to first denature and reduce the protein ribonuclease by treating with the denaturing agent urea and the reducing agent β-mercaptoethanol. Following this, he dialysed the protein to remove both the urea and β-mercaptoethanol. After the dialysis he found that the protein regained its full activity and, therefore, must have spontaneously re-folded into its native active state. Based on this result, Anfinsen proposed that the entire information for the correct folding of a protein is contained in its amino acid sequence. This can be simply summarized in the phrase "one sequence, one structure". For this work Anfinsen received the 1972 Nobel Prize in Chemistry. "One sequence, one structure" has been a central principle of biochemistry and molecular biology ever since, and it is commonly referred to as "Anfinsen's dogma". However, evidence is now accumulating that Anfinsen's dogma may be in need of a revision.

Protein X-ray crystallography

The first three-dimensional protein structure to be determined was that of myoglobin (Kendrew et al., 1958), for which John Kendrew of the University of Cambridge received the 1962 Nobel Prize in Chemistry. This structural determination was based on the fundamental foundation of X-ray crystallography that exposure of an ordered array of atoms produces an X-ray diffraction pattern. Every different structure creates its own characteristic pattern. Therefore, if

the mathematical relationship between the structure and the pattern created is known, by analysis of the pattern one can solve the structure. The mathematical relationship is known as Bragg's Law, and was discovered in the early 20th century by the father and son team William and Lawrence Bragg. At one stage in his career William Bragg had been professor of physics at the University of Adelaide and his son Lawrence grew up in Adelaide. Lawrence was later the head of the Cavendish Laboratory in Cambridge where Kendrew and also Watson and Crick worked.

Since the determination of the myoglobin structure, three-dimensional structures of numerous proteins have been solved at the level of atomic resolution. Membrane proteins proved to be particularly hard nuts to crack because of the surrounding lipid. However, by replacing native lipids with detergent, even membrane protein structures are now being solved at an ever increasing rate. The first to be solved was that of the photosynthetic reaction centre of the bacterium *Rhodopseudomonas viridis*, which was published by Hartmut Michel, Johann Deisenhofer and Robert Huber from the Max-Planck-Institute of Biochemistry in Munich (Nobel Prize for Chemistry, 1988). The first crystal structure of the Na^+,K^+-ATPase, the ion pump responsible for maintaining cell volume in all animal cells, was published in 2007 (Morth et al.).

In spite of the success of protein X-ray crystallography and all the protein structures which have been successfully solved, it has become increasingly clear that some proteins or some segments of proteins still elude structural determination. For example, although structures of the Na^+,K^+-ATPase have now been published in a number of different conformational states, no-one has yet succeeded in resolving the structure of the cytoplasmic N-terminal tail of the protein's α-subunit. If this segment of the protein doesn't produce a diffraction pattern, the most likely reason is that it is too flexible on the time-scale of the X-ray crystallographic method, i.e. it can be classified as an intrinsically disordered segment of the protein.

Entropy

Entropy, or in layman's terms disorder, is a fundamental concept of physical chemistry, being the topic of both the Second and Third Laws of Thermodynamics. Under constant temperature conditions, the change in free energy and consequently the distribution of a protein between two states (e.g. two protein states in an enzymatic cycle) are determined not only by the change in the strength of intermolecular bonding forces (reflected in the enthalpy change, ΔH), but also by the change in entropy (ΔS). Thus, a change in the degree of disorder of an intrinsically disordered segment of an enzyme could have a significant effect on an enzyme's kinetics.

Again using the Na^+,K^+-ATPase as an example, recent evidence (Garcia et al., 2017; Jiang et al., 2017; Nguyen et al., 2018) suggests that this protein's N-terminus switches between two states where it is either bound to the neighbouring membrane or freely moving in the cytoplasm, i.e. from a low to a high entropy state. Based on the amino acid content of the N-terminus, the free energy change associated with a transition from a completely ordered to a completely disordered state can be estimated using the PLOPS server of Baxa et al. (2014) to be -234 kJ mol^{-1}. This is a larger free energy decrease than that released by ATP hydrolysis, i.e., approximately -50 kJ mol^{-1} (Clarke et al., 2013). Although a partial rather than a complete disordering of the N-terminus is

more likely, this comparison demonstrates the potential of entropy changes to contribute significantly to the stability of protein conformational states.

Revision of Anfinsen's dogma

Just as amino acid sequences can cause protein segments to fold in particular ways to generate a three-dimensional structure, the existence of intrinsically disordered protein segments indicates that amino acid sequences can also code for disorder. Because disorder necessarily implies more than one possible local conformation of a protein segment, the concept of "one sequence, one structure" is no longer tenable. However, rather than discard Anfinsen's dogma entirely, a better solution would seem to be to expand it to include disorder as well as structure. Thus, one could say that the amino acid sequence of a protein contains all the information to generate both its order and disorder.

References

Baxa, M. C., Haddadian, E. J., Jumper J. M., Freed, K. F., and Sosnick, T. R. (2014) Loss of conformational entropy in protein folding calculated using realistic ensembles and its implications for NMR-based calculations; *Proc. Natl. Acad. Sci.*, **111**, 15396–15401.

Clarke, R. J., Catauro, M., Rasmussen, H. H., and Apell, H.-J. (2013) Quantitative calculation of the role of the Na^+,K^+-ATPase in thermogenesis; *Biochim. Biophys. Acta*, **1827**, 1205–1212.

Garcia, A., Pratap, P. R., Lüpfert, C., Cornelius, F., Jacquemin D., Lev, B., Allen, T. W., and Clarke, R. J. (2017) The voltage-sensitive dye RH421 detects a Na^+,K^+-ATPase conformational change at the membrane surface; *Biochim. Biophys. Acta*, **1859**, 813–823.

Jiang, Q., Garcia, A., Han, M., Cornelius, F., Apell, H.-J., Khandelia, H., and Clarke, R. J. (2017) Electrostatic stabilization plays a central role in autoinhibitory regulation of the Na^+,K^+-ATPase; *Biophys. J.*, **112**, 288–299.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. and Phillips, D. C. (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis; *Nature* **181**, 662–666.

Leder, P. and Nirenberg, M. W. (1964) RNA codewords and protein synthesis; *Proc. Natl. Acad. Sci.*, **52**, 1521–1529.

Morth, J. P., Pedersen, B. P., Toustrup-Jensen, M. S., Sørensen, T. L., Petersen, J., Andersen, J. P., Vilsen, B., and Nissen, P. (2007) Crystal structure of the sodium-potassium pump; *Nature*, **450**, 1043–1049.

Nguyen, K., Garcia, A., Sani, M.-A., Diaz, D., Dubey, V., Clayton, D., Dal Poggetto, G., Cornelius, F., Payne, R. J., Separovic, F., Khandelia H., and Clarke, R. J. (2018) Interaction of N-terminal peptide analogues of the Na^+,K^+-ATPase with membranes; *Biochim. Biophys. Acta*, **1860**, 1282–1291.

Watson, J. D. and Crick, F. H. C. (1953) Molecular structure of nucleic acids: A structure of deoxyribonucleic acid; *Nature,* **171**, 737–738.