nature structural biology
January

Snapshots of the 'molten globule'

The first issue of this new monthly journal provides a variety of papers for structural biologists (and others) to get their teeth into, among them two dealing with the intriguing matter of protein folding.

GIVEN the number of possible conformations that a polypeptide chain can adopt it is remarkable that proteins fold to a unique tertiary structure at all, let alone achieve the feat in the order of seconds or minutes. So how do they do it? Reports in the January issue of *Nature Structural Biology*, by Feng *et al.*[1] and Redfield *et al.*[2], describe the structures of two partially unfolded proteins, apocytochrome b_{562} and interleukin-4, which provide hints about what is happening towards the end of the folding pathway.

Protein folding is clearly a directed process. The challenge is to divine the physico-chemical properties of proteins *in vitro*, and the role of the *in vivo* co-factors that aid and abet the transformation of a linear sequence of amino acids into a functioning, paid-up protein. One of the main difficulties in studying protein folding *in vitro* is the transient nature of the kinetic intermediates on the pathway. In consequence these species are for the most part only ever present in low stoichiometric amounts, making analysis of their structure extemely tricky. One solution is to look at intermediates between the folded and unfolded state that, under certain conditions, occur at equilibrium in sufficient quantity that they can be studied at leisure. One such species is the molten globule, which has properties similar to those of the kinetic intermediates.

Relatively stable, partially folded proteins can be generated by a number of means (variation of the solvent conditions, removal of a ligand or prosthetic group and so on). In general, a partially folded protein is considered to be a molten globule if it has the following properties: a well-defined secondary structure, some of which may be in a native-like conformation; none of the specific tertiary interactions that define the native state; and an expanded structure relative to the folded protein.

Also in *Nature Structural Biology* in January: substrate–ribozyme interactions; using isosteric base analogues as a means to map the interactions between the *trp* repressor and its cognate operator sequence; intact–cleaved human antithrombin III complex as a model for serpin–proteinase interactions; structure of the toxin from *Shigella dysenteriae* responsible for the severe form of dysentry in humans.

Cytochrome b_{562} is a haem-containing protein that consists of four helices with a binding pocket in which the large haem prosthetic group nestles (the holoprotein). Removal of the haem results in partial unfolding of the protein, which nonetheless maintains much of its structure under near-physiological conditions (the apoprotein).

The solution structure of the apoprotein, as determined by Feng *et al.*, is remarkably similar to the structure of the holoprotein. But although three of the four helices in the holoprotein are fully represented in the apoprotein, distortion of the C terminus of helix IV and the misalignment of helix I and II result in poor packing interfaces between these secondary structure elements.

The various derangements of the apoprotein result in the exposure of the haem-binding pocket to the surrounding solvent and the complete solvation of the amino-acid residues, Met 7 and His 107, that provide axial ligands to the haem. The residues that contact the haem group, and other residues that form the core of the protein, are not repacked in the apoprotein and remain exposed to solvent. The large cavern thus formed in the apoprotein is huge, and able to accommodate up to fifty water molecules.

But can the apoprotein be said to be a molten globule? Although Feng *et al.* wisely hesitate about claiming that it should have full molten-globule status, it is clear that the apoprotein has some of the appropriate features. But argument over that question is in some senses a side issue, for the structure is of interest from another point of view. As the authors point out, the NMR constraints were obtained under near-physiological conditions suggesting that the structure of the apoprotein represents a folding intermediate late on the pathway in haem protein assembly.

Redfield and colleagues' structure analysis is of interleukin-4 (IL-4), a four-helix-bundle protein, at low pH. Both the loss of the near-ultraviolet circular dichroism (CD) spectrum and the enhanced fluorescence of the hydrophobic dye 1-anilinonaphthalene-8-sulphonic acid (ANS) in the presence of IL-4 under acid conditions are characteristic of the canonical molten globule state.

Despite this, comparison of the protein in roughly neutral and acidic conditions again reveals that there is relatively little change in the NMR spectra for most of the amino-acid residues, suggesting that the structure at low pH is similar to that of the native form of the protein. Even so, more than one third of the residues are in regions of significant disorder. The four helices, which form the core of the protein, are for the most part highly ordered. It is the sequences which connect the helices that are substantially disordered.

Determination of order parameters, which provide an indication of the rigidity of the main chain of the protein, indicate that the region around the amino terminus of the C helix undergoes a local unfolding transition at low pH. So it would seem that the 'molten globule-like' ANS and near-ultraviolet CD characteristics are a consequence of increased disorder of localized portions of the structure, rather than the formation of a globally disordered state.

The structures of apocytochrome b_{562} and IL-4 are both highly ordered compared to some of the more disordered molten globules characterized so far. Nonetheless both proteins have some of the features of the molten globule state and, as Redfield *et al.* suggest for IL-4, both might reasonably be considered as 'highly ordered molten globules'.

Clearly, the label of molten globule can encompass a range of degrees of unfolding of a polypeptide chain. Those that lack most residual tertiary structure and are highly disorganized are probably related to kinetic intermediates formed early in the folding pathway. The 'highly ordered molten globules' characterized by Feng *et al.* and Redfield *et al.* are likely to be similar to the types of structures occurring late in the protein folding process.

The burning question is this — what happens *in vivo*? The various proteins that assist in folding (protein disulphide isomerase, chaperones and so on) do not seem to impart structural information to the polypeptide chains with which they interact. So one would hope that the lessons learnt *in vitro* will illuminate protein folding in the cell.

Guy Riddihough

Guy Riddihough is Editor of Nature Structural Biology.

1. Feng, Y., Sligar, S.G. & Wand, A.J. *Nature struct. Biol.* **1**, 30–35 (1994).
2. Redfield, C., Smith, R.A.G. & Dobson, C.M. *Nature struct. Biol.* **1**, 23–29 (1994).

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature*'s staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review Articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Progress articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research.

News and Views articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963 65). The first and last page numbers are included; reference to books includes publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value $5' to prevent the imposition of import duty and value-added tax.