Science Highlights
from the National Synchrotron Light Source

BEAMLINE
X6A

PUBLICATION
Tommi Kajander, Aitziber L. Cortajarena, Ewan R. G. Main, Simon G. J. Mochrie, and Lynne Regan, "A New Folding Paradigm for Repeat Proteins," *J. Am. Chem. Soc.*, **127**(29), 10188-10190 (2005).

FUNDING
Helsingin Sanomat Centennial Foundation (Finland) postdoctoral fellowship; Spanish Ministry of Culture, Education and Sports postdoctoral fellowship

FOR MORE INFORMATION
Simon Mochrie, Departments of Physics and Applied Physics, Yale University
simon.mochrie@yale.edu

A New Folding Paradigm for Repeat Proteins

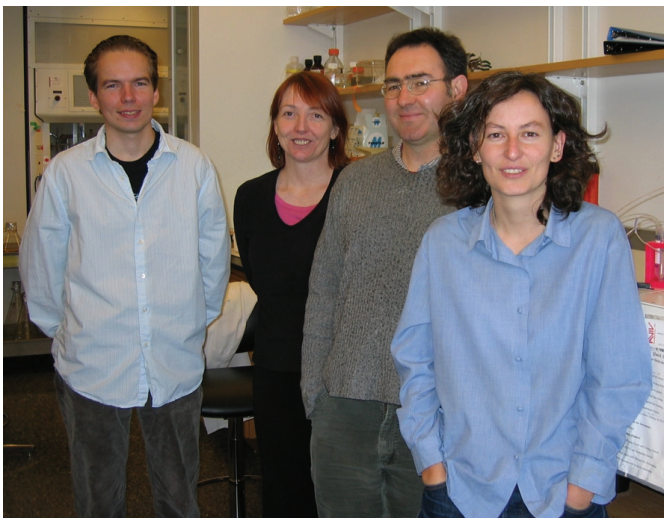
T. Kajander[1], A.L. Cortajarena[1], E.R.G. Main[2], S.G. J. Mochrie[3], and L. Regan[2,4]

[1]Department of Molecular Biophysics and Biochemistry, Yale University; [2]Department of Chemistry, University of Sussex, UK; [3]Departments of Physics and Applied Physics, Yale University; [4]Department of Chemistry, Yale University

Determining how a protein's amino acid sequence specifies its structure and properties stands as a grand challenge for biology in the post-genomic era. Repeat proteins, composed of tandem arrays of a basic structural motif, account for more than 5% of metazoan proteins in the Swiss-Prot database. It is therefore surprising that the folding of repeat proteins has been little studied, especially because their modular structures promise a more tractable folding problem than globular proteins. Our approach has been to design a consensus tetratricopeptide repeat (CTPRa) and from this construct a series of proteins that contain from two to 20 tandem copies of this repeat. We have shown that the folding of this family of TPR proteins can be described and predicted by the classical one-dimensional Ising model of statistical mechanics, which thus represents a new folding paradigm for repeat proteins.

We solved the x-ray crystal structure of the repeat protein CTPRa8, which contains eight identical TPR repeats, at 2.05Å resolution, using single wavelength anomalous diffraction data from cadmium. The structure of CTPRa8 (**Figure 1**) reveals that each repeat is comprised of two helices, which are arrayed to form a super helix. A key feature of CTPRa8, and of repeat proteins in general, is that, in contrast to globular proteins, there are no sequentially distant amino acid contacts. This pattern of local contacts suggests that a theoretical description based on the classical one-dimensional Ising model might be an appropriate way to understand the folding/unfolding behavior of repeat proteins, with the repeating units cast as Ising spins.

How can we test this hypothesis? Because the variation in behavior for different numbers of coupled subunits is a textbook signature of collective effects, the possibility of creating CTPRs in any multiplicity desired renders them an ideal system in which to test this model of repeat protein stability. Accordingly, we synthesized a series of CTPRs, for which the number of helices varied from five to 21. The stability of these proteins as a function of guanidine hydrochloride (GuHCl) concentration (GuHCl is an agent used to induce protein unfolding) is shown in **Figure 2**. The rapid change in each profile with increasing GuHCl concentration corresponds to the transition from the folded to the unfolded state. The best fits of the Ising model to these data are shown as the solid lines in **Figure 2**. Clearly, the model provides an excellent description of the folding/unfolding for all of the proteins studied.



Authors (from left) Tommi Kajander, Lynne Regan, Simon Mochrie, and Aitziber Cortajarena (Ewan Main is absent from the photo).

We believe that several aspects of this work are particularly exciting. First, the Ising model is predictive: Because thermodynamic data from just two different CTPRs in a series are sufficient to predict the behavior of all additional CTPRs in that series, we may view four of the model profiles in Figure 2 as predictions. These are the first examples of successfully predicting the stabilities of proteins and the shapes of their unfolding curves.

Second, the Ising model requires a new microscopic picture: In the usual two-state transition, a protein is essentially always completely folded or completely unfolded with only brief transient behavior. In contrast, the Ising description implies that

near the transition midpoint, partially-folded configurations occur with significant probability.

Finally, the structure of CTPRa8 reveals both of the atomic details of an individual repeat, and the arrangement of repeats in the superhelix, and thus promises the possibility of being able to assign the Ising model parameters to specific structural elements. Re-designs are in progress to test such ideas.

In a more general sense, beyond repeat proteins, does the behavior of the CTPRa series tell us anything about the folding of *globular* proteins? The folding of several small, single-domain globular proteins has been well-studied. As for individual repeat proteins, their thermodynamic behavior can usually be well-described by the traditional two-state model. However, in many cases, hydrogen-exchange measurements reveal a richer free-energy landscape than simply two free-energy minima, hinting that it may be sensible to envision each element of secondary structure itself as a degree of freedom that can be either folded or unfolded, *i.e.* as an Ising spin. Such observations suggest important similarities between the thermodynamics of repeat proteins and globular proteins.

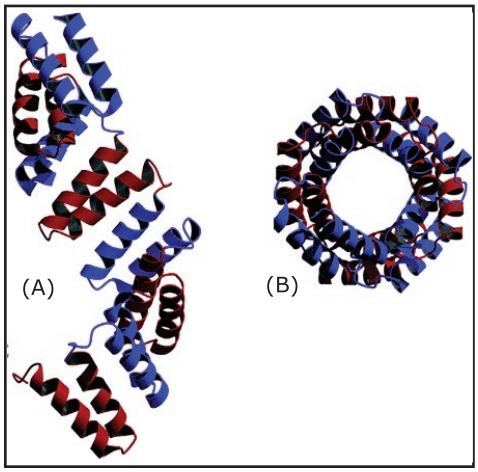


Figure 1. *Crystal structure of CTPRa8. (A) View perpendicular to the superhelical axis. Each TPR repeat is colored either red or blue. (B) View along the superhelical axis.*

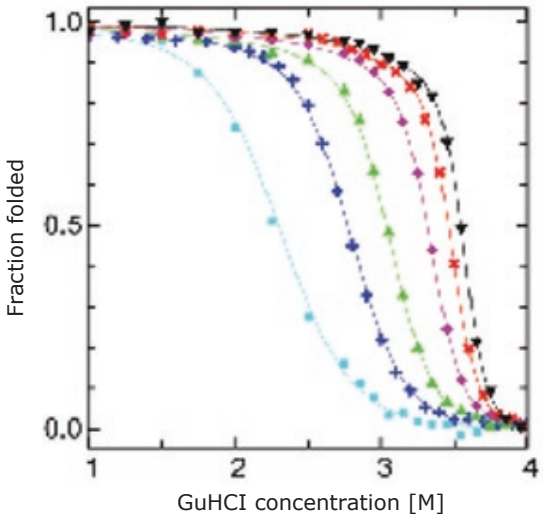


Figure 2. *Thermodynamics of CTPRa unfolding. (A) Measured ellipticity and (B) fraction folded vs. [GuHCl] for CTPRa2 (squares), CTPRa3 (crosses), CTPRa4 (triangles), CTPRa6 (diamonds), CTPRa8 (crosses), CTPRa10 (inverted triangles) in 50 mM sodium phosphate, pH 6.5, 150 mM sodium chloride, at -25 °C. Root-mean-square errors of 0.8 millidegrees are smaller than the plotting symbols, except (when transformed) for the fraction folded of CTPRa2 in (B). Solid lines correspond to the best-fits to the predictions of the one-dimensional N-spin Ising model. Analogously to the Ising-Zimm-Bragg treatment of the polypeptide helix-coil transition, our model ascribes a reduced stability to the endmost helices, relative to those not at the end, by incorporating interactions with ficticious down spins at position*