Chemistry & Biology, Vol. 9, 667–672, June, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S1074-5521(02)00160-6

Previews

Building a Metal Binding Domain, One Half at a Time

The recently determined structure of a zinc binding peptide reveals that a particular sequence can adopt one stable fold as an isolated peptide but adopt an alternative structure as part of a larger protein domain.

Although the amino acid sequence of a protein contains all of the information necessary to determine its three-dimensional structure, this information is not distributed uniformly among the amino acids, making the prediction of protein fold from sequence a challenge. Nonetheless, rules that govern folding have begun to emerge for certain types of proteins, including small metalloproteins. Many tetrahedral metal binding sites consist of two halves, with each contributing two ligands to the metal center. Some of these half sites recur in a range of proteins [1]. For example, a structurally conserved fragment including a metal-chelating Cys-X_{2-4}-Cys sequence was found to occur in rubredoxin, aspartate transcarbamoylase, and other metal binding proteins. Similarly, numerous metalloproteins contain sequences of the form His-$X_{3,4}$-His,Cys that form helical metal-chelating structures. These two substructures combine around a metal ion to form the classical TFIIIA-like zinc finger domain with three key hydrophobic residues stabilizing the folded structure. As the structures of more zinc binding domains have been determined, these and a few more recurring substructures have been found to act as building blocks for the extant structures.

Mixing and matching these recurring substructures should result in the generation of novel folds. Mackay and coworkers demonstrated this concept recently in the journal Structure [2]. The Mackay group reports an NMR structure of a zinc binding domain that corresponds to the CH1 domain of CBP (CREB-binding protein). CBP is a transcriptional coregulator that contains three putative zinc binding domains, CH1 (or TAZ1), CH2 (a PHD motif), and CH3 (or TAZ2). The Mackay group identified two Cys-X_4-Cys-X_8-His-X_3-Cys (CCHC) potential metal binding sites in the CH1 region and structurally characterized the zinc bound form of a peptide corresponding to one of these sites. The secondary structure of this domain included two short but well-defined α helices corresponding to the Cys-X_4-Cys and His-X_3-Cys ligand sets, creating a novel overall fold that the researchers coined the CHANCE fold. Before Mackay and coworkers had completed their structure, Dyson, Wright, and coworkers reported the structure of the CH3 domain of CBP [3]. Because of the high level of sequence similarity between the CH1 and CH3 domains, one would expect the two structures to be analogous; surprisingly, this was not the case. Dyson, Wright, and coworkers had expressed the entire CH3 domain, which is significantly larger than the peptide prepared by the Mackay group. The ligand set for the CH3 domain included additional Cys and His ligands both N- and C- terminal to the CCHC domains identified by Mackay, such that zinc ions could bind to several different combinations of Cys and His ligands. A structure with three long α helices was formed with three HCCC zinc binding sites. Dyson, Wright, and coworkers also presented biochemical data suggesting that this fold is the physiologically relevant structure. Although these two structures have a dissimilar fold overall, the Mackay group showed that the individual zinc binding modules have a high degree of structural conservation. The smaller peptide of the Mackay group had combined two substructures that bind to two different zinc ions in the full CH3 domain around a single zinc ion to generate an entirely new fold (see Figure 1).

Mackay's group conducted a series of experiments to tease out the requisite features to get the CHANCE fold. They were able to substitute alanine for many of the amino acids of the CH1 domain and thus created a minimalist CHANCE domain akin to one previously described for a TFIIIA-like zinc finger domain [4]. Mackay's group was still able to retain folding, provided that they kept the metal binding and key hydrophobic residues in place. They were also able to prepare a segment of the CH3 domain that included two CCHC zinc coordination sites, and thus just one HCCC site. They found that this peptide bound two metals with high α-helical content and, presumably, formed the CHANCE fold.

The ease with which zinc binding domains can be mixed and matched to generate novel folds may have evolutionary implications. Since the first identification of a zinc binding domain in 1985, more than 15 classes of zinc binding domains have been identified [5]. Ranging from transcriptional regulation to protein folding and assembly [6], their functions have been found to be equally broad. The abundance of zinc-binding domains in eukaryotic genomes may be related the relatively simple requirement of the presence of two sets of metal-chelating ligands with properly spaced hydrophobic residues for the generation of a folded structure. The probability of these two features occurring by chance is reasonably high and thus evolutionarily favored. In fact, as Mackay points out, work by Keefe and Szostak to generate new functional (ATP binding) proteins from random libraries of polypeptide sequence pulled out four novel proteins, one of which included a zinc binding domain [7]. Mackay's report of the facile generation of a new fold for the CH1 domain speaks to this point; by preparing a piece of the CH1 domain that contained the requisite, but non-physiological zinc modules, his group generated a novel zinc binding fold.

One must stop short, however, of concluding that the high probability of zinc-binding domains occurring by chance is the raison d'être for the abundance of zinc binding proteins. The evolution of zinc binding domains is probably more complicated. For example, the propensity for evolving zinc binding domains does not explain why eukaryotes have a much more varied set and higher abundance of zinc binding domains than do prokaryotes. Eukaryotes appear to have much more sophisticated mechanisms of acquiring and storing zinc in their cells than do prokaryotes. Research into these zinc reg-

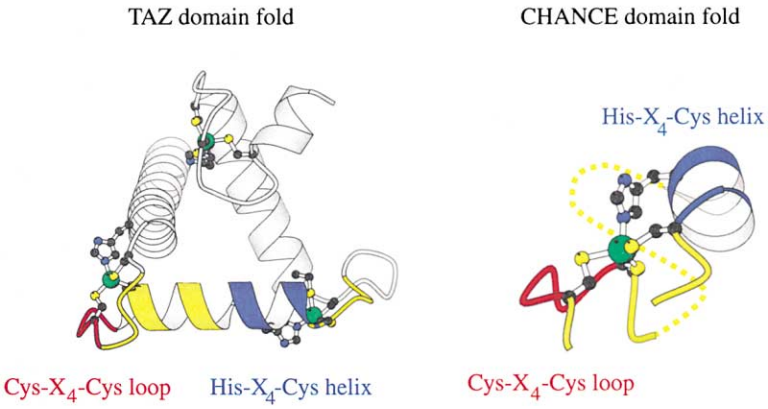

Figure 1. Alternative Zinc Binding Modes

In the TAZ domain structure, the Cys-C4-Cys loop and the His-X4-Cys helix bind to different metal ions. In the CHANCE domain fold, these two substructures come together to form a novel zinc-binding module.

ulatory systems is still in its infancy; however, it is possible that their development is part of the story of the evolution of zinc binding domains.

The concept of mixing domains to generate a new structure is not unique to zinc binding domains. Oligomeric proteins can be derived from monomeric proteins via a "3D domain swapping" mechanism [8]. Typically, a portion of the tertiary structure of one monomeric protein is replaced by the analogous structural element of a second monomeric protein, resulting in a dimer. For example, the removal of part of a surface loop in the monomeric staphylococcal nuclease resulted in a dimeric structure, stabilized by a swapping of the C-terminal α helices of two monomers [9]. As with the generation of novel zinc binding folds by mixing and matching metal binding substructures, 3D domain swapping may provide a route for the evolution of oligomeric proteins from monomers, and it has even been suggested to be a route for the formation of amyloid fibrils associated with neurodegenerative diseases [10]. The results of Mackay and coworkers reveal this process in reverse with the conversion of a single-chain trimeric structure of the TAZ domain into a well-defined folded monomer by truncation.

Sarah L.J. Michel and Jeremy M. Berg
Department of Biophysics and Biophysical Chemistry
Johns Hopkins University School of Medicine
Baltimore, Maryland 21205

Selected Reading

1. Berg, J.M. (1995). Zinc finger domains: from predictions to design. Acc. Chem. Res. *28*, 14–18.
2. Sharpe, B.K., Matthews, J.M., Kwan, A.H.Y., Newton, A., Gell, D.A., Crossley, M., and Mackay, J.P. (2002). A new zinc binding fold underlies the versatility of zinc binding modules in protein evolution. Structure *10*, 639–648.
3. De Guzman, R.N., Liu, H.Y., Martinez-Yamout, M., Dyson, H.J., and Wright, P.E. (2000). Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. J. Mol. Biol. *303*, 243–253.
4. Michael, S.F., Kilfoil, V.J., Schmidt, M.H., Amann, B.T., and Berg, J.M. (1992). Metal binding and folding properties of a minimalist Cys_2His_2 zinc finger peptide. Proc. Natl. Acad. Sci. USA *89*, 4796–4800.
5. Schwabe, J.W.R., and Klug, A. (1994). Zinc mining for protein domains. Nat. Struct. Biol. *11*, 345–349.
6. Laity, J.H., Lee, B.M., and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. Curr. Opin. Struct. Biol. *11*, 39–46.
7. Keefe, A.D., and Szostak, J.W. (2001). Functional proteins from a random-sequence library. Nature *410*, 715–718.
8. Bennett, M.J., Schlunegger, M.P., and Eisenberg, D. (1995). 3D domain swapping: a mechanism for oligomer assembly. Protein Sci. *4*, 2455–2468.
9. Green, S.M., Gittis, A.G., Meeker, A.K., and Lattman, E.E. (1995). One-step evolution of a dimmer from a monomeric protein. Nat. Struct. Biol. *2*, 746–751.
10. Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2001). A domain-swapped RNAse A dimer with implications for amyloid formation. Nat. Struct. Biol. *8*, 211–214.

Chemistry & Biology, Vol. 9, June, 2002, PII S1074-5521(02)00158-8

Dissecting Histone Deacetylase Function

Recent research describes the use of chromatin immunoprecipitation and intergenic chromosomal-DNA microarrays to analyze HDAC function genome-wide. The next step in realizing the full potential of these analyses will be to develop specific and temporal control over HDAC perturbation.

Histone acetylation was initially observed and characterized in the 1960s by Allfrey and colleagues. However, the unexpected catalyst for recent advances in our knowledge of histone deacetylase (HDAC) function was not initiated until 1990, when Itazaki and coworkers at Shionogi Laboratories set out upon a chemical genetic screen to look for small molecules with the ability to