Cell, Vol. 53, 675, June 3, 1988, Copyright © 1988 by Cell Press

Letter to the Editor

Fingering Too Many Proteins

"Zinc finger" domains, first identified in transcription factor TFIIIA (Miller et al., EMBO J. *4*, 1609–1614, 1985; Brown et al., FEBS Lett. *186*, 271–274, 1985), are found in many eukaryotic regulatory proteins. More than twenty proteins have already been reported to contain sequences that are clearly related to the repeated domains of TFIIIA (see reviews by Klug and Rhodes, Trends in Biochem. Sci. *12*, 464–469, 1987; and Evans and Hollenberg, Cell *52*, 1–3, 1988), and it is likely that these sequences form structures involved in nucleic acid recognition. These proteins have several features in common: each contains multiple repeats of the 30-amino-acid zinc finger domain, and each repeat conserves two cysteine and two histidine residues which coordinate the metal ion, *as well as* two aromatic residues and one leucine residue. Berg (PNAS *85*, 99–102, 1988) has proposed that this very specific pattern of conserved amino acids allows a zinc finger domain to fold into a structure containing an antiparallel β-ribbon and an α-helix.

Unfortunately, the term "zinc finger" has not been restricted to domains that contain two cysteines, two histidines, two aromatic residues, and one leucine with the appropriate spacing. In fact, it has been applied to almost any regulatory protein that happens to have several cysteines within a short stretch of sequence, and it is presumed that these regions will form zinc-dependent DNA binding domains. Proteins considered to have zinc fingers by this loose definition include the nuclear receptors, the adenovirus E1A proteins, and GAL4 and related proteins. While the cysteine-rich regions of these proteins are likely to form metal-binding sites, there is no reason to believe that their three-dimensional structures will resemble the structure of the zinc finger domain, and there is no a priori reason to expect that these regions will be involved in nucleic acid recognition.

Recent studies of two proteins, previously described as zinc finger proteins, suggest that their metal-binding sites *are not* used to form nucleic acid binding domains. Coleman and coworkers found that zinc is essential for the folding of bacteriophage T4 gene 32 protein (Giedroc et al., PNAS *83*, 8452–8456, 1986), but they also found that the apoprotein and the metalloprotein bind to a single DNA site with similar affinities (Giedroc et al., Biochemistry *26*, 5251–5259, 1987). The zinc-binding domain is necessary only for cooperative binding to longer DNAs. Here, this domain may maintain protein–protein contacts between DNA-bound proteins rather than being directly involved in nucleic acid recognition. The tat protein from human immunodeficiency virus, which has seven cysteines within a stretch of sixteen residues, also has been misclassified as a zinc finger protein. The cysteine-rich region does bind metals, but these metal ions are used to stabilize the protein dimer, presumably by bridging cysteine ligands from each monomer (Frankel et al., Science *240*, 70–73, 1988).

The nuclear receptors are an important group of proteins for which the zinc finger analogy may also be misleading. Sequence alignments show a strictly conserved pattern of cysteines which is quite different from that seen in the zinc finger domains. While it is likely that the three-dimensional structures of the nuclear receptor domains are similar to each other, these structures will probably be unrelated to the structure of the zinc finger domain in TFIIIA. The results of many experiments strongly suggest that nuclear receptors bind DNA and that the DNA-binding domain is localized within the cysteine-rich region (region C) of the receptors (e.g., Green and Chambon, Nature *325*, 75–77, 1987; Rusconi and Yamamoto, EMBO J. *6*, 1309–1315, 1987; Hollenberg et al., Cell *49*, 39–46, 1987; and references therein), but no published experiments show that the cysteines bind metals or that specific binding to DNA is dependent on metals. Assuming that metals do bind to region C, it is possible that some metals organize a DNA-binding domain while others mediate protein–protein interactions. Introns divide region C of the progesterone receptor into two parts (Huckaby et al., PNAS *84*, 8380–8384, 1987; Jeltsch et al., PNAS *83*, 5424–5428, 1986), one containing four cysteines (C_4) and the other containing five (C_5). This observation led Evans and Hollenberg (op. cit.) to suggest that the two regions might be structurally distinct, and that the C_5 motif might bind more than one metal ion. Such a metal cluster might stabilize protein dimers by using cysteine–metal interactions to bridge receptor subunits (as seen with tat), or might form independent dimerization domains. Metal-binding domains might also mediate the formation of complexes between receptors and other proteins such as transcription factors.

Questions can also be raised about the structure and function of the "zinc finger" domains in the adenovirus E1A protein. The cysteine-rich region in E1A clearly is important for transactivation since a synthetic 49-amino-acid peptide containing this region activates an E1A-inducible promoter (Lillie et al., Cell *50*, 1091–1100, 1987). However, E1A does not seem to bind a specific site on DNA (Berk, Ann. Rev. Gen. *20*, 45–79, 1986), and, as suggested by others (Berg, Cold Spring Harbor Symp. Quant. Biol. *52*, 579–585, 1987), it is possible that the metal-binding domains mediate protein–protein interactions, perhaps with transcription factors.

In summary, while it appears that metal-binding sites are an important component of many regulatory proteins, several distinct three-dimensional structures seem to be involved. We should be careful to consider, and experimentally test, a broader range of possibilities before discussing these sites in functional terms.

Alan D. Frankel and Carl O. Pabo
Department of Molecular Biology and Genetics
and Howard Hughes Medical Institute
The Johns Hopkins University School of Medicine
Baltimore, Maryland 21205