1. Swanson, R.V., Schuster, S. C. & Simon, M. I. *Biochemistry*, **32**, 7623–7629 (1993).
2. Morrison, T.B. & Parkinson, J. S. *Proc. Nat. Acad. Sci. USA* **91**, 5485–5489 (1994).
3. Zhou, H., McEvoy, M. M., Swanson, R. V., Simon, M. I. & Dahlquist, F. W. *Biochemistry* **35**, 433–443 (1996).
4. Zhou, H., Lowry, D. F., Swanson, R.V., Simon, M. I. & Dahlquist, F. W. *Biochemistry* **34**, 13858–13870 (1995).
5. McEvoy, M. M. *et al. Biochemistry* **34**, 13871–13880 (1995).
6. Surette, M. G. *et al. J. Biol. Chem.* **271**, 939–945 (1996).
7. Levit, M., Liu, Y., Surette, M. & Stock, J. *J. Biol. Chem.* **271**, 32057–32063 (1996).
8. McEvoy, M. M., de la Cruz, A. F. A. & Dahlquist, F. W. *Nature Struct. Biol.* **4**,–9 (1997).
9. McEvoy, M. M., Muhandiram, D. R., Kay, L. E. & Dahlquist, F. W. *Biochemistry* **35**, 5633–5640 (1996).
10. Chothia C. Janin, J. *Nature* **256**, 705–708 (1975).
11. Li, J., Swanson, R. V., Simon, M. I. & Weis, R. M. *Biochemistry* **34**, 14626–14636 (1995).
12. Swanson, R. V. *et al. Nature Struct. Biol.* **2**, 906–910 (1995)
13. Volz, K. *Biochemistry* **32**, 11741–11753 (1993).
14. Stock, J. B., Ninfa, A. J. & Stock, A. M. *Microbiol. Rev.* 53, 450–490 (1989).
15. Parkinson, J.S. & Kofoid, E. C. *Annu. Rev. Genet.* **26**, 71–112 (1992).
16. Bischoff, D. S., Bourret, R. B., Kirsch, M. L. & Ordal, G. W. *Biochemistry* **32**, 9256–9261 (1993).
17. Kuo, S. C. & Koshland, D. E. Jr. *J. Bact.* **169**, 1307–1314 (1987).
18. Volz, K & Matsumura, P. *J Biol Chem* **266**, 15511–15519 (1991).
19. Stock, A. M. *et al. Biochemistry* **32**, 13375–13380 (19
20. Moy, F. J. *et al. Biochemistry* **33**, 10731–10742 (1994).
21. Ganguli, S., Wang, H., Matsumura, P. & Volz, K. *J. Biol. Chem.* **270**, 17386–17393 (1995).
22. Zhu, X., Rebello, J., Matsumura, P. & Voltz, K. *J. Biol. Chem.* **272**, 5000–5006 (1997).
23. Shukla, D. & Matsumura, P. *J. Biol. Chem.* **270**, 24414–24419 (1995).
24. Garzon, A. & Parkinson, J. S. *J. Bact.* **178**, 6752–6758 (1996).
25. Amsler, C. D. & Matsumura, P. Chemotactic signal transduction in *Escherischia coli* and *Salmonella typhimurium* in *Two-component signal transduction* (eds Hoch, J. A. & Silhavy, T. J.) 89–99 (American Society for Microbiology, Washington, DC–20005; 1995).
26. Leslie, A.G.W. in *Computational Aspects of Protein Crystal Data Analysis, Proceedings of the Daresbury Study Weekend* (eds Helliwell, J.R., Machin, P.A. & Papiz, M.Z.) 39–50 (SERC, Daresbury Laboratory, Warrington,UK; 1987).
27. Collaborative Computational Project 4. *Acta Crystallogr. D* **50**, 760–763 (1994).
28. La Fortelle, E. de & Bricogne, G. Maximum-Likelihood Heavy-Atom Parameter Refinement in the MIR and MAD Methods in *Methods in Enzymology, Macromolecular Crystallography*, Vol. 276 (eds Sweet, R.M. & Carter, C. W. Jr.) 472–494 (Academic Press New York; 1997).
29. Cowtan, K. 'DM' an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, Vol. 31 (eds Bailey, S. & Wilson, K.) 34–38 (Daresbury Laboratory, Hamburg,–Germany; 1994).
30. Navaza, J. *Acta Crystallogr. A* **50**, 157–163 (1994).
31. Brünger, A.T. *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR* (Yale University Press, New Haven, CT; 1992).
32. Brünger, A.T. *Nature* **355**, 472–475 (1992).

Cro, CAP and λ repressor led the way

When a new protein is discovered, its gene cloned, the first questions are: is it related to anything? Does it have any motifs we can understand? A helix-turn-helix? A zinc-finger? An ankyrin repeat?

But this 'motif mindset' is a modern invention. In the case of DNA binding proteins, this idea came out of three major structural determinations, of Cro[1] (from bacteriophage λ), CAP[2] (the catabolite activator protein from *E. coli*) and the DNA binding domain of λ repressor[3]. Prior to these studies, models had been proposed for both α-helices and pairs of β-strands interacting with the major groove of DNA, but no one had really anticipated that basic structural scaffolds such as those mentioned above would be used over and over again to recognize DNA or other proteins.

Cro and CAP both form dimers, and their backbone structures only (because of a low effective resolution of Cro and an incomplete identification of the CAP sequence) were published in 1981, showing a mix of α-helices and β-sheets in both proteins. The Cro dimer displayed two particularly interesting α-helices, perfectly separated and tilted to match the 34 Å distance between DNA major grooves (Fig. 1), and therefore the Cro–DNA complex could be modeled easily[1]. The CAP dimer had a pair of helices separated by a similar distance, but because of their different tilt (Fig. 1), they did not easily fit into right handed B-DNA major grooves, and thus the first model of the CAP–DNA complex postulated that CAP bound left handed B-DNA[2]. However, subsequent analysis, allowing flexibility of the protein and bending of the binding site, suggested that CAP could indeed bind right handed B-DNA[3,10]. Following closely on the heels of these studies was the report of the structure of λ repressor's all α-helical DNA binding domain at 3.2 Å resolution in 1982[3] (Fig. 1). Although the λ repressor DNA binding domain did not form a dimer in solution, it was thought to bind DNA as a dimer. A reasonable dimer structure was located in the protein crystal by looking for features that would be complementary to the known DNA binding site.

Direct comparisons of the their sequences and structures[4–6] suggested that each single α-helix that was predicted to contact DNA was actually part of a related set of helices, what we now call the helix-turn-helix motif, in each of these proteins. But, according to Brian Matthews, people were at first somewhat skeptical: how could you make a case based on correspondence between such short segments of sequence[4] or structure[5,6]? Nevertheless, the helix-turn-helix motif gained more acceptance soon after, as it was also found by sequence analysis alone in many other DNA binding proteins, including the Lac and Gal repressors[7–9].

Cro, CAP and the DNA binding domain of λ repressor were different in overall structure, yet each contained this particular DNA recognition motif. These discoveries dramatically increased our understanding of DNA binding proteins — and added a new member to our collection of protein motifs. TLS

1. Anderson, W.F., Ohlendorf, D.H., Takeda, Y. & Matthews, B.W. *Nature* **290**, 754–758 (1981).
2. McKay, D.B. & Steitz, T.A. *Nature* **290**, 744–749 (1981).
3. Pabo, C.O & Lewis, M. *Nature* **298**, 443–447 (1982).
4. Anderson, W.F., Takeda, Y., Ohlendorf, D.H. & Matthews, B.W. *J Mol. Biol.* **159**, 745–751 (1982).
5. Steitz, T.A., Ohlendorf, D.H., McKay, D.B., Anderson, W.F. & Matthews, B.W. *Proc. Natl. Acad. Sci. USA* **79**, 3097–3100 (1982).
6. Ohlendorf, D.H., Anderson, W.F., Lewis, M., Pabo, C.O & Matthews, B.W. *J. Mol. Biol.* **169**, 757–769 (1983).
7. Matthews, B.W., Ohlendorf, D.H., Anderson, W.F. & Takeda, Y. *Proc. Natl. Acad. Sci. USA* **79**, 1428–1432 (1982).
8. Sauer, R.T., Yocum, R.R., Doolittle, R.F., Lewis, M. & Pabo, C.O. *Nature* **298**, 447–451 (1982).
9. Weber, I.T., McKay, D.B. Steitz, T.A. *Nucleic Acids Res.* **10**, 5085–5102 (1982).
10. Takeda, Y., Ohlendorf, D.H., Anderson, W.F. & Matthews, B.W. *Science* **221**, 1020–1026 (1983).

Fig. 1 Schematic diagrams of B form DNA, the Cro dimer and the DNA binding domains of the λ repressor dimer and the CAP dimer. The α3 helices of Cro and λ repressor and the αF helix of CAP were predicted to contact DNA based on the first structures of these proteins[1–3]. The helix-turn-helix motifs consist of α2 and α3 or αE and αF. Figure used with permission from ref. 10.