

REFERENCES

References

23.1

- Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. & Weng, J. (1987). *Protein Data Bank*. In *Crystallographic databases – information content, software systems, scientific applications*, edited by F. H. Allen, G. Bergerhoff, & R. Sievers, pp. 107–132.
- Artymiuk, P. J., Mitchell, E. M., Rice, D. W. & Willett, P. (1989). *Searching techniques for databases of protein structures*. *J. Inf. Sci.* **15**, 287–298.
- Artymiuk, P. J., Rice, D. W., Poirrette, A. R. & Willett, P. (1995). *Beta-glucosyltransferase and phosphorylase reveal their common theme*. *Nature Struct. Biol.* **2**, 117–120.
- Barton, G. J. (1997). *3Dee: database of protein domain definitions*. <http://barton.ebi.ac.uk/servers/3Dee.html>.
- Bennett, M. J. & Eisenberg, D. (1994). *Refined structure of monomeric diphtheria toxin at 2.3 Å resolution*. *Protein Sci.* **3**, 1464–1475.
- Bork, P. (1992). *Mobile modules and motifs*. *Curr. Opin. Struct. Biol.* **2**, 413–421.
- Brenner, S. E., Chothia, C., Hubbard, T. J. & Murzin, A. G. (1996). *Understanding protein structure. Using SCOP for fold interpretation*. *Methods Enzymol.* **266**, 635–643.
- Brown, N. P., Orengo, C. A. & Taylor, W. R. (1996). *A protein structure comparison methodology*. *Comput. Chem.* **20**, 359–380.
- Chothia, C. (1993). *One thousand families for the molecular biologist*. *Nature (London)*, **357**, 543–544.
- Crippen, G. (1978). *The tree structural organization of proteins*. *J. Mol. Biol.* **126**, 315–332.
- Flores, T. P., Orengo, C. A. & Thornton, J. M. (1993). *Conformational characteristics in structurally similar protein pairs*. *Protein Sci.* **7**, 31–37.
- Gibrat, J. F., Madej, T., Spouge, J. L. & Bryant, S. H. (1997). *The VAST protein structure comparison method*. *Biophys. J.* **72**, MP298.
- Go, M. (1981). *Correlation of DNA exonic regions with protein structural units in hemoglobin*. *Nature (London)*, **291**, 90–92.
- Hogue, C. W., Ohkawa, H. & Bryant, S. H. (1996). *A dynamic look at structures: WWW-Entrez and the molecular modelling database*. *Trends Biochem. Sci.* **21**, 226–229.
- Holm, L. & Sander, C. (1993). *Protein structure comparison by alignment of distance matrices*. *J. Mol. Biol.* **233**, 123–138.
- Holm, L. & Sander, C. (1994a). *Searching protein structure databases has come of age*. *Proteins*, **19**, 165–173.
- Holm, L. & Sander, C. (1994b). *Parser for protein folding units*. *Proteins*, **19**, 256–268.
- Holm, L. & Sander, C. (1995). *Evolutionary link between glycogen phosphorylase and a DNA modifying enzyme*. *EMBO J.* **14**, 1287–1293.
- Holm, L. & Sander, C. (1996). *Mapping the protein universe*. *Science*, **273**, 595–602.
- Holm, L. & Sander, C. (1997). *Enzyme HIT*. *Trends Biochem. Sci.* **22**, 116–117.
- Holm, L. & Sander, C. (1998). *Dictionary of recurrent domains in protein structures*. *Proteins*, **33**, 88–96.
- Holm, L. & Sander, C. (1999). *Protein folds and families: sequence and structure alignments*. *Nucleic Acids Res.* **27**, 244–247.
- Islam, S. A., Luo, J. & Sternberg, M. J. (1995). *Identification and analysis of domains in proteins*. *Protein Eng.* **8**, 513–525.
- Jones, S., Stewart, M., Michie, A. D., Swindells, M. B., Orengo, C. A. & Thornton, J. M. (1998). *Domain assignment for protein structures using a consensus approach: characterisation and analysis*. *Protein Sci.* **7**, 233–242.
- Kikuchi, T., Nemethy, G. & Scheraga, H. A. (1988). *Prediction of the location of structural domains in globular proteins*. *J. Protein Chem.* **88**, 427–471.
- Kraulis, P. J. (1991). *MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures*. *J. Appl. Cryst.* **24**, 946–950.
- Lesk, A. M. & Rose, G. D. (1981). *Folding units in globular proteins*. *Proc. Natl Acad. Sci. USA*, **78**, 4304–4308.
- Levitt, M. & Chothia, C. (1976). *Structural patterns in globular proteins*. *Nature (London)*, **261**, 552–558.
- Li, M., Dyda, F., Benhar, I., Pastan, I. & Davies, D. R. (1996). *Crystal structure of the catalytic domain of Pseudomonas exotoxin A complexed with a nicotinamide adenine dinucleotide analog: implications for the activation process and for ADP ribosylation*. *Proc. Natl Acad. Sci. USA*, **93**, 6902–6906.
- Lionetti, C., Guanziroli, M. G., Frigerio, F., Ascenzi, P. & Bolognesi, M. (1991). *X-ray crystal structure of the ferric sperm whale myoglobin: imidazole complex at 2.0 Å resolution*. *J. Mol. Biol.* **217**, 409–412.
- Mizuguchi, K., Deane, C. A., Blundell, T. L. & Overington, J. P. (1998). *HOMSTRAD: a database of protein structure alignments for homologous families*. *Protein Sci.* **7**, 2469–2471.
- Moult, J. & Unger, R. (1991). *An analysis of protein folding pathways*. *Biochemistry*, **30**, 3816–3824.
- Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). *SCOP: a structural classification of the protein database for the investigation of sequences and structures*. *J. Mol. Biol.* **247**, 536–540.
- Nemethy, G. & Scheraga, H. A. (1979). *A possible folding pathway of bovine pancreatic Rnase*. *Proc. Natl Acad. Sci. USA*, **76**, 6050–6054.
- Orengo, C. A., Jones, D. T., Taylor, W. & Thornton, J. M. (1994). *Protein superfamilies and domain superfolds*. *Nature (London)*, **372**, 631–634.
- Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B. & Thornton, J. M. (1997). *CATH – a hierarchic classification of protein domain structures*. *Structure*, **5**, 1093–1108.
- Orengo, C. A., Pearl, F. M. G., Bray, J. E., Todd, A. E., Martin, A. C., LoConte, L. & Thornton, J. M. (1999). *The CATH database provides insights into protein structure/function relationships*. *Nucleic Acids Res.* **27**, 275–279.
- Phillips, D. E. (1970). *British biochemistry, past and present*, p. 11. London Biochemistry Society Symposium. Academic Press.
- Rashin, A. A. (1976). *Location of domains in globular proteins*. *Nature (London)*, **291**, 85–87.
- Richardson, J. S. (1981). *The anatomy and taxonomy of protein structure*. *Adv. Protein Chem.* **34**, 167–339.
- Rose, G. D. (1979). *Hierarchic organization of domains in globular proteins*. *J. Mol. Biol.* **134**, 447–470.
- Rossmann, M. G. & Argos, P. (1975). *A comparison of the heme binding pocket in globins and cytochrome b5*. *J. Biol. Chem.* **250**, 7525–7532.
- Rossmann, M. & Liljas, A. (1974). *Recognition of structural domains in globular proteins*. *J. Mol. Biol.* **85**, 177–181.
- Russell, R. B. & Barton, G. J. (1993). *Multiple protein sequence alignment from tertiary structure comparisons. Assignments of global and residue level confidences*. *Proteins*, **14**, 309–323.
- Sali, A. & Blundell, T. B. (1990). *The definition of general topological equivalences in proteins: a procedure involving comparison of properties and relationships through simulated annealing and dynamic programming*. *J. Mol. Biol.* **212**, 403–428.
- Sander, C. (1981). *Physical criteria for folding units of globular proteins*. In *Structural aspects of recognition and assembly in biological macromolecules*, Vol. I. *Proteins and protein complexes, fibrous proteins*, edited by M. Balaban, pp. 183–195. Jerusalem: Alpha Press.
- Sander, C. & Schneider, R. (1991). *Database of homology-derived protein structures and structural meaning of sequence alignments*. *Proteins*, **9**, 56–68.
- Schulz, G. E. & Schirmer, H. (1979). *Principles of protein structure*, ch. 5. New York: Springer Verlag.
- Siddiqui, A. S. & Barton, G. J. (1995). *Continuous and discontinuous domains: an algorithm for the automatic generation of reliable protein domain definitions*. *Protein Sci.* **4**, 872–884.
- Sowdhamini, R., Rufino, S. D. & Blundell, T. L. (1996). *A database of globular protein structural domains: clustering of representa-*

23. STRUCTURAL ANALYSIS AND CLASSIFICATION

23.1 (cont.)

- tive family members into similar folds. *Structure Fold. Des.* **1**, 209–220.
- Swindells, M. B. (1995). A procedure for detecting structural domains in proteins. *Protein Sci.* **4**, 103–112.
- Taylor, W. R. & Orengo, C. A. (1989). Protein structure alignment. *J. Mol. Biol.* **208**, 1–22.
- Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y. & Steitz, T. A. (1996). Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. *EMBO J.* **15**, 5739–5751.
- Vriend, G. & Sander, C. (1991). Detection of common three-dimensional substructures in proteins. *Proteins*, **11**, 552–558.
- Wernisch, L., Hunting, M. & Wodak, J. (1999). Identification of structural domains in proteins by a graph heuristic. *Proteins*, **35**, 338–352.
- Wetlaufer, D. B. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl Acad. Sci. USA*, **70**, 697–701.
- Wodak, J. & Janin, J. (1981). Location of structural domains in proteins. *Biochemistry*, **20**, 6544–6552.
- Zehfus, M. H. (1994). Binary discontinuous compact protein domains. *Protein Eng.* **7**, 335–340.
- Zehfus, M. H. (1997). Identification of compact, hydrophobically stabilized domains and modules containing multiple peptide chains. *Protein Sci.* **6**, 1210–1219.
- Zehfus, M. H. & Rose, G. D. (1986). Compact units in proteins. *Biochemistry*, **25**, 5759–5765.

23.2

- Åqvist, J., Luecke, H., Quioco, F. A. & Warshel, A. (1991). Dipoles localized at helix termini of proteins stabilize charges. *Proc. Natl Acad. Sci. USA*, **88**, 2026–2030.
- Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967). On the conformation of the hen egg-white lysozyme molecule. *Proc. R. Soc. London Ser. B Biol. Sci.* **167**, 365–377.
- Bochkarev, A., Pfuetzner, R. A., Edwards, A. M. & Frappier, L. (1997). Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature (London)*, **385**, 176–181.
- Burd, C. G. & Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.
- Cheng, X. (1995). DNA modification by methyltransferases. *Curr. Opin. Struct. Biol.* **5**, 4–10.
- Cleland, W. W. & Kreevoy, M. M. (1994). Low-barrier hydrogen bonds and enzymic catalysis. *Science*, **264**, 1887–1890.
- Cotton, F. A., Hazen, E. E. Jr & Legg, M. J. (1979). *Staphylococcal nuclease: proposed mechanism of action based on structure of enzyme-thymidine 3',5'-bisphosphate-calcium ion complex at 1.5-Å resolution.* *Proc. Natl Acad. Sci. USA*, **76**, 2551–2555.
- Cusack, S., Yaremchuk, A. & Tukalo, M. (1996a). The crystal structure of the ternary complex of *T. thermophilus* seryl-tRNA synthetase with tRNA(Ser) and a seryl-adenylate analogue reveals a conformational switch in the active site. *EMBO J.* **15**, 2834–2842.
- Cusack, S., Yaremchuk, A. & Tukalo, M. (1996b). The crystal structures of *T. thermophilus* lysyl-tRNA synthetase complexed with *E. coli* tRNA(Lys) and a *T. thermophilus* tRNA(Lys) transcript: anticodon recognition and conformational changes upon binding of a lysyl-adenylate analogue. *EMBO J.* **15**, 6321–6334.
- Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A. L., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. (1998). *Science*, **280**, 68–77.
- Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R. & Steitz, T. A. (1988). Cocrystal structure of an editing complex of Klenow fragment with DNA. *Proc. Natl Acad. Sci. USA*, **85**, 8924–8928.
- Gerlt, J. A. & Gassman, P. G. (1993). *Understanding the rates of certain enzyme-catalyzed reactions: proton abstraction from carbon acids, acyl-transfer reaction, and displacement reactions of phosphodiester.* *Biochemistry*, **32**, 1943–1952.
- Glusker, J. P. (1991). Structural aspects of metal liganding to functional groups in proteins. *Adv. Protein Chem.* **42**, 1–76.
- Goldgur, Y., Mosyak, L., Reshetnikova, L., Ankilova, V., Lavrik, O., Khodyreva, S. & Safro, M. (1997). The crystal structure of phenylalanyl-tRNA synthetase from *Thermus Thermophilus* complexed with cognate tRNAPhe. *Structure*, **5**, 59–68.
- Harrington, R. E. & Winicov, I. (1994). New concepts in protein-DNA recognition: sequence-directed DNA bending and flexibility. *Prog. Nucleic Acid Res. Mol. Biol.* **47**, 195–270.
- He, J. J. & Quioco, F. A. (1993). Dominant role of local dipoles in stabilizing uncompensated charges on a sulphate sequestered in a periplasmic active transport. *Protein Sci.* **2**, 1643–1647.
- Hibbert, F. & Emsley, J. (1990). Hydrogen bonding and chemical reactivity. *Adv. Phys. Org. Chem.* **226**, 255–379.
- Hodel, A. E., Gershon, P. D. & Quioco, F. A. (1998). Structural basis for sequence non-specific recognition of 5'-capped mRNA by a cap-modifying enzyme. *Mol. Cell*, **1**, 443–447.
- Hodel, A. E., Gershon, P. D., Shi, X., Wang, S. M. & Quioco, F. A. (1997). Specific protein recognition of an mRNA cap through its alkylated base. (Letter.) *Nature Struct. Biol.* **4**, 350–354.
- Jacobson, B. L. & Quioco, F. A. (1988). Sulphate-binding protein dislikes protonated oxyacids: a molecular explanation. *J. Mol. Biol.* **204**, 783–787.
- Joachimiak, A., Schevitz, R. W., Kelley, R. L., Yanofsky, C. & Sigler, P. B. (1983). Functional inferences from crystals of *Escherichia coli* trp repressor. *J. Biol. Chem.* **258**, 12641–12643.
- Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell*, **63**, 579–590.
- Labahn, J., Schärer, O. D., Long, A., Ezaz-Nikpay, K., Verdine, G. L. & Ellenberger, T. E. (1996). Structural basis for the excision repair of alkylation-damaged DNA. *Cell*, **86**, 321–329.
- Ledvina, P. S., Tsai, A.-H., Wang, Z., Koehl, E. & Quioco, F. A. (1998). Dominant role of local dipolar interactions in phosphate binding to a receptor cleft with an electronegative charge surface potential: equilibrium, kinetic and crystallographic studies. *Protein Sci.* **7**, 2550–2559.
- Ledvina, P. S., Yao, N., Choudhary, A. & Quioco, F. A. (1996). Negative electrostatic surface potential of protein sites specific for anionic ligands. *Proc. Natl Acad. Sci. USA*, **93**, 6786–6791.
- Lindahl, T. (1982). DNA repair enzymes. *Annu. Rev. Biochem.* **51**, 61–87.
- Luecke, H. & Quioco, F. A. (1990). High specificity of a phosphate transport protein determined by hydrogen bonds. *Nature (London)*, **347**, 402–406.
- Marcotrigiano, J., Gingras, A. C., Sonenberg, N. & Burley, S. K. (1997). X-ray studies of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Nucleic Acids Symp. Ser.* pp. 8–11.
- Meador, W. E., George, S. E., Means, A. R. & Quioco, F. A. (1995). X-ray analysis reveals conformational adaptation of the linker in functional calmodulin mutants. (Letter.) *Nature Struct. Biol.* **2**, 943–945.
- Medveczky, N. & Rosenberg, H. (1971). Phosphate transport in *Escherichia coli*. *Biochim. Biophys. Acta*, **241**, 494–506.
- Nagai, K. (1996). RNA-protein complexes. *Curr. Opin. Struct. Biol.* **6**, 53–61.
- Nagai, K., Oubridge, C., Ito, N., Jessen, T. H., Avis, J. & Evans, P. (1995). Crystal structure of the U1A spliceosomal protein complexed with its cognate RNA hairpin. *Nucleic Acids Symp. Ser.* 1–2.
- Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association; insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.
- Orgel, L. E. (1966). *An introduction to transition-metal chemistry. Ligand-field theory*, 2nd ed. London: Methuen and New York: Wiley.