

13. Chou, P. Y. and Fasman, G. D. (1978) Empirical predictions of protein conformation. *Annu. Rev. Biochem.* **47**, 251–276.
14. Gamier, J., Osguthorpe, D. J., and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**, 97–120.
15. Rost, B. and Sander, C. Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* **232**, 584–599.
16. Rost, B. and Sander, C. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* **19**, 55–72.
17. Rost, B., Sander, C., and Schneider, R. (1994) PHD—an automatic mail server for protein secondary structure prediction. *Comput. Appl. Biosci.* **10**, 53–60.
18. Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying an hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132.
19. Hopp, T. P. and Woods, K. R. (1983) A computer program for predicting protein antigenic determinants. *Mol. Immunol.* **20**, 483–489.
20. Benson, D. A., Boguski, M., Lipman, D. J., and Ostell, J. (1994) GenBank. *Nucleic Acids Res.* **22**, 3441–3444.
21. Emmert, D. B., Stoehr, P. J., Stoesser, G., and Cameron, G. N. (1994) The European Bioinformatics Institute (EBI) databases. *Nucleic Acids Res.* **22**, 3445–3449.
22. Bairoch, A. and Boeckmann, B. (1994) The S,WISS-PROT protein sequence data bank: current status. *Nucleic Acids Res.* **22**, 3578–3580.
23. George, D. G., Barker, W. C., Mewes, H. W., Pfeiffer, F., and Tsugita, A. (1994) The PIRInternational Protein Sequence Database. *Nucleic Acids Res.* **22**, 3569–3573.
24. Pearson, W. R. and Lipman, D. J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
25. Lipman, D. J. and Pearson, W. R. (1985) Rapid and sensitive protein similarity searches. *Science* **227**, 1435–1441.
26. Pearson, W. R. (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* **183**, 63–98.
27. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
28. Smith, T. F. and Waterman, M. S. (1981) Identification of common molecular subsequences. *J. Mol. Biol.* **147**, 195–197.
29. Sturrock, S. S. and Collins, J. F. (1993) MPsrch version 1. 3. Biocomputing Research Unit, University of Edinburgh, UK.
30. Waterman, M. S. and Vingron M. (1994) Rapid and accurate estimates of statistical significance for sequence database searches. *Proc. Natl. Acad. Sci. USA* **91**, 4625–4628.
31. Karlin, S. and Altschul, S. F. (1993) Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc. Natl. Acad. Sci. USA* **90**, 5873–5877.
32. Karlin, S. and Altschul, S. F. (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. USA* **87**, 2264–2268.