

Genome Annotation and Protein Structure

Steven E. Brenner

University of California, Berkeley, CA

Abstract

Structural genomics aims to provide a good experimental structure or computational model of every tractable protein in a complete genome. Underlying this goal is the immense value of protein structure, especially in permitting recognition of distant evolutionary relationships for proteins whose sequence analysis has failed to find any significant homologue. A considerable fraction of the genes in all sequenced genomes have no known function, and structure determination provides a direct means of revealing homology that may be used to infer their putative molecular function. The solved structures will be similarly useful for elucidating the biochemical or biophysical role of proteins that have been previously ascribed only phenotypic functions. More generally, knowledge of an increasingly complete repertoire of protein structures will aid structure prediction methods, improve understanding of protein structure, and ultimately lend insight into molecular interactions and pathways.

References

(available from compbio.berkeley.edu):

1. Brenner SE. *A tour of structural genomics*. Nature Reviews Genetics 2:801-9 (2001).
2. Zupicich J, Brenner SE, Skarnes WC. *Computational prediction of membrane-tethered transcription factors*. Genome Biology 2: research0050.1-0050.6 (2001)
3. Brenner SE, Levitt M. *Expectations from structural genomics*. Protein Sci. 9:197-200 (2000).
4. Brenner SE. *Errors in genome annotation*. Trends in Genetics 15:132-133 (1999).
5. Brenner SE, Chothia C, Hubbard TJP. *Population statistics of protein structures*. Current Opinion in Structural Biology 7:369-376 (1997)
6. Murzin AG, Brenner SE, Hubbard T, Chothia C. *SCOP: a structural classification of proteins database for the investigation of sequences and structures*. Journal of Molecular Biology 247:536-540 (1995).