# The presence of a copper/zinc superoxide dismutase in the bacterium *Photobacterium leiognathi*: A likely case of gene transfer from eukaryotes to prokaryotes

(amino acid mutation/secondary structure/model building)

JOE V. BANNISTER AND MICHAEL W. PARKER

Inorganic Chemistry Laboratory and Laboratory of Molecular Biophysics, University of Oxford, Oxford OX1 3QR, England

ABSTRACT     The free-living bacterium *Photobacterium leiognathi* is also known to be a symbiont of ponyfish. The presence of a copper/zinc superoxide dismutase in *P. leiognathi* has been considered to be a case of gene transfer from eukaryotes to prokaryotes because this form of superoxide dismutase is normally present only in higher eukaryotic species. However, the amino acid sequence of the enzyme from the bacterium exhibited low identities (25–30%) with the same enzyme from eukaryotes. When amino acid mutations are taken into consideration, the weighted sequence similarity increases significantly; furthermore, the bacterial enzyme has the same active site residues and similar predicted secondary structure as the eukaryotic enzymes. The possibility of convergence is ruled out and the case of divergence is considered unlikely because of the observed phylogenetic distribution of the enzyme. This indicates that the presence of the copper/zinc superoxide dismutase in *P. leiognathi* can indeed be considered a case of gene transfer from eukaryotic species to prokaryotic species.

The first line of defense against the toxicity of active oxygen species is removal of the superoxide anion radical [$O_2^-$] (1, 2). The enzymes called superoxide dismutases (EC 1.15.1.1), discovered by McCord and Fridovich (3), scavenge the superoxide radical [$2O_2^- + 2H^+ \rightarrow H_2O_2 + O_2$]. These enzymes are a group of metalloproteins containing either copper/zinc, manganese, or iron as the prosthetic group. The copper/zinc protein is now considered to be characteristically found in all higher eukaryotic cells. A report indicating its absence in adipocytes (4) has been shown to be in error (5). The manganese protein is present in prokaryotes and in the mitochondria of eukaryotic cells, while the iron protein is usually present in prokaryotes. A number of exceptions to these observations have recently come to light. The iron enzyme has, in addition to the copper/zinc and manganese enzymes, been isolated from the mustard plant *Brassica campestris* (6), and a survey of 43 plant families for the presence of the iron enzyme showed that it is present in three isolated families (7). Two bacterial species, *Photobacterium leiognathi* (8) and *Caulobacter crescentus* (9), have been shown to contain the copper/zinc superoxide dismutase although this form of the enzyme has been found to be absent from eukaryotic species more primitive than green algae such as protozoans and *Euglena* species (10).

The presence of a copper/zinc superoxide dismutase in *C. crescentus* is as yet an unexplained anomaly. However, the presence of this enzyme in *P. leiognathi* has been considered to be a case of gene transfer from eukaryotic species to prokaryotic species (11, 12). This is because this bacterial species, although found free-living, is also a symbiont of ponyfish (13). Using the amino acid sequences of the copper/zinc

superoxide dismutases isolated from *P. leiognathi* (14) and from the swordfish *Xiphias gladius* liver (15), we compare the two sequences. The swordfish liver enzyme has been shown to be closely related to the unsequenced enzyme of the bacterial host, the ponyfish (12). Considering amino acid mutations, predicted secondary structure, and model building, a case for convergent evolution is shown to be unlikely. On the basis of phylogenetic distribution the case for the bacterial enzyme evolving prior to the prokaryote–eukaryote split is considered also unlikely. The most plausible case is natural gene transfer from the host fish to the symbiotic bacterium.

## MATERIALS AND METHODS

Complete amino acid sequences of copper/zinc superoxide dismutases from *P. leiognathi* (14) swordfish liver, (15) bovine erythrocyte (16), horse liver (17), and yeast (18) are known. The sequence alignment of the eukaryotic enzymes is taken from the literature (19), and the alignment of the bacterial to the other enzymes was carried out by maximizing identity between the sequences together with the use of gap penalties (penalty of −2 identities). For ease of comparison, the corresponding residue numbers in the bovine erythrocyte enzyme sequence have been included here.

Two methods were used to calculate the degree of sequence homology between the higher eukaryotic and *P. leiognathi* enzymes. The first method calculates the percentage of perfect identity between two aligned sequences. The second takes into account mutations in the gene coding for the proteins. This was carried out according to McLachlan (20). The method is based on frequences of observed substitutions in families of homologous proteins, and the percentage similarity between the superoxide dismutases was determined from McLachlan's frequency matrix. For example, identities in the sequences gained a score of 8 or 9 out of 9 depending on the residue. A score of 3 was average and a score of 0 was given for gaps. Common gaps in the aligned sequences were excluded from the calculation. This method is referred to here as a measure of the weighted degree of similarity.

Secondary structures were calculated using algorithms that predict conformation from amino acid sequence. Predictions were carried out according to Garnier *et al.* (21), Lim (22), and Chou and Fasman (23). An overall combined prediction was derived from the individual methods based on the requirement that at least two out of the three methods agreed for each of the structural elements predicted.

Model building was carried out using an interactive computer graphics system (Evans and Sutherland picture system II) equipped with the FRODO program (24). The program was used to delete, insert, and replace residues and also allowed for manual rotations about bonds so as to eliminate bad contacts. A model was built for the bacterial enzyme using the coordinates for the bovine erythrocyte copper/zinc

```
                    1                    10                      20                        30                      40
Swordfish       V L K A V C[V]L R G A G E[T]T G T[V]- Y F[E]G E G[N]A N A[V]G K G I I[L]K G -[L]
Bacterial       - - Q D L T[V]K M T D L Q[T]G K K[V]G T I[E]L S Q[N]K Y G[V]V F T P E[A]A D -[L]
                o o ● ■ ● ■ Δ ▲ Δ Δ Δ ● Δ ■ ▲ Δ ● ● ▲ ● ● ▲ o Δ ● ▲ o ▲ Δ ■ ▲ ▲ Δ ● ▲ Δ o Δ ▲ ▲ ● ● - ▲

                       50                          60                      70                          80
Swordfish       [T][P][G][E][H][G][F][H]V[H]G F[G]D N T N G[C]I[S]- - - - - - - - - - - - -[A][G][P][H]F N[P]
Bacterial       [T][P][G][M][H][G][F][H]I[H]Q N[G]- - S - -[C][A][S]S E K D G K V V L G G A[A][G][G][H]Y D[P]
                ▲ ▲ ▲ ▲ Δ ▲ ▲ ▲ ▲ ● ■ ▲ Δ o ▲ o o ● ■ o o ▲ ▲ ▲ o o o o o o o o o o o o o ▲ ▲ ● ▲ ■ ● ▲

                           90                          100                      110                      120
Swordfish       A - S K[K][H]A G[P]K D E[D][R][H]V[G][D][L]G N V T A D[A][N][G]V[A]K I D I T D - K I S
Bacterial       E H T N K[K][H]G F[P]W T D[D][N][H]K[G][D][L]P A L F V S[A][N][G][L][A]T N P V L A P - - R
                ■ o ■ ■ ▲ ▲ ● o ▲ ▲ ● ■ ▲ ● ■ ▲ ▲ ▲ ● ■ ▲ o ● ▲ ▲ ▲ ▲ ■ ▲ o ▲ ● ■ ● o o o ●

                               130                      140                      150                      160
Swordfish       [L][T]G P Y S I I[G]R T M V[T][H]E K A[D]D L G R G G N E E S L K T[G]N A[G][S][R][L][A]
Bacterial       [L][T]- L K E L K[G]H A I M[T][H]A G G[D]N - - H - - S D M P K A L[G]G G[G][A][R]V[A]
                ▲ ▲ o Δ Δ ■ ■ ▲ ■ ■ ▲ ● ● ● ■ ▲ ● ● ● ▲ ▲ ● ● ● ▲ ■ o o ■ o o ■ ■ ▲ ● Δ ● ▲ ● ● ▲ ■ ▲ ▲

                   168
Swordfish       [C][G][V][I]G T E -
Bacterial       [C][G][V][I]- - Q -
                ▲ ▲ ▲ ▲ o o ■ -
```
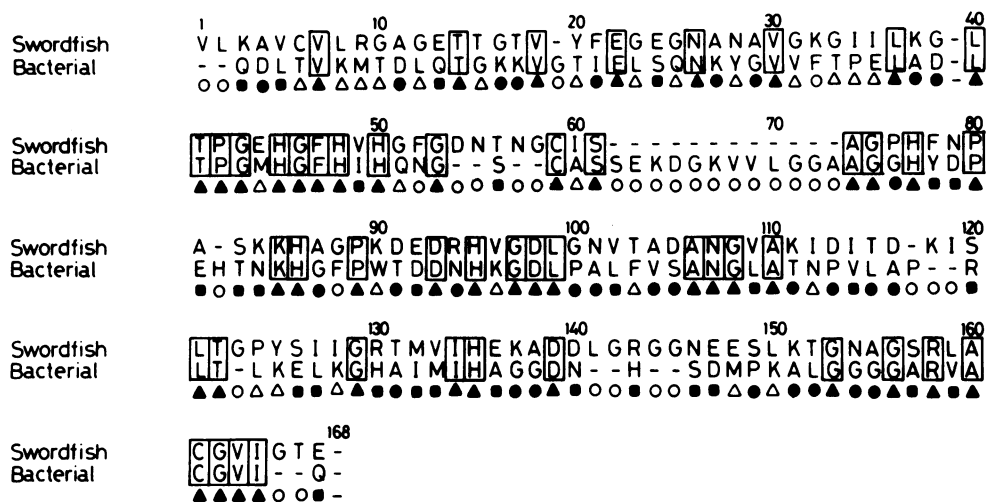
FIG. 1. Amino acid sequences of swordfish liver and bacterial (*P. leiognathi*) copper/zinc superoxide dismutases. Identical residues are boxed. Symbols refer to the frequency of substitutions determined according to McLachlan (20): ▲, identity; ■, frequently found substitution; ●, neutral substitution; Δ, infrequent substitution; o, rarely found or no substitution. Common gaps are due to alignment to other eukaryotic enzymes (not shown).

superoxide dismutase available from the Brookhaven Protein Data Bank (25).

## RESULTS AND DISCUSSION

Copper/zinc superoxide dismutase isolated from swordfish liver (26) and from *P. leiognathi* (12) has been found to have the same catalytic activity, metal content, and subunit size as the bovine erythrocyte enzyme. The amino acid sequences for the bacterial and swordfish enzymes are shown in Fig. 1. The alignment indicates that a number of residues previously determined to be important for metal binding, enzymatic activity, and maintenance of tertiary structure in the bovine enzyme (14, 15, 18, 27) are conserved in the bacterial protein as well as in all the other sequenced eukaryotic enzymes (19). These are His-48 (bovine 44), -50 (bovine 46), -77 (bovine 61), and -135 (bovine 118) as ligands for the copper; His-77 (bovine 61), -86 (bovine 69), and -95 (bovine 78), and Asp-98 (bovine 81) as ligands for the zinc; and Arg-158 (bovine 141) and Asp-139 (bovine 122), which are considered to be essential for activity and the intrachain disulfide bridge formed by Cys-59 and -161 (bovine 55 and 144). Out of 49 glycine residues in the bacterial and swordfish enzymes, 20 are conserved while, out of 13 proline residues, 6 are conserved.

The degree of identity and weighted similarity for the sequences are given in Fig. 2. The amino acid sequences of bacterial and swordfish liver enzymes indicate little identity relative to identities between the eukaryotic enzymes. However, when mutations are taken into account long stretches of closely related sequences are obtained (Fig. 1) and the weighted similarity between the two enzymes rises to 44%. This corresponds to 78 identical or closely related (above average substitution) residues out of 136 residues excluding deletions and insertions.

The predicted secondary structures for the bovine erythrocyte, swordfish liver, and bacterial enzymes are presented in Fig. 3. The predicted secondary structure for the bovine erythrocyte enzyme agrees well with the structure obtained from x-ray data (27) and the prediction for the swordfish liver enzymes is similar to the bovine erythrocyte enzyme structure. The bacterial enzyme, despite having a low identity with the two enzymes, also has, with two exceptions, a similar secondary structure. The first exception is the 12-residue insertion, residues 62–73 (after bovine 57) in the bacterial sequence, which is, in part, predicted to form a β-strand. The methods of Garnier *et al.* (21) and Chou and Fasman (23) predict an α-helix between residues 125 and 129 (bovine 108–112), whereas Lim (22) predicts nothing and a turn is present in this region for the bovine enzyme.

The subunit structure of bovine erythrocyte copper/zinc superoxide determined from crystal structure has been found to be comprised of eight antiparallel β-strands that form a flattened cylinder plus three external loops (27). After establishing close similarity between the primary and secondary structures of the bovine erythrocyte, swordfish liver, and *P. leiognathi* enzymes, an attempt was made to see whether the bacterial enzyme could be built into the known bovine erythrocyte tertiary structure without serious disruption to the fold.

A schematic diagram of the model obtained for the bacterial enzyme is presented in Fig. 4. The model was built taking into consideration the positioning of essential and conserved residues (including closely related residues) and secondary structure predictions. All the prediction methods indicated a strand for the 12-residue insertion (residues 62–73) in the bacterial enzyme. However, despite exhaustive attempts a strand could not be built in the model so as to interact appropriately with nearby strands. The insertion was successfully built into the structure as a loop. The predicted α-helix also proved to be an unreliable prediction. The overall fold of the model could not have been retained if the α-helix had not been built as a loop of random coil. The loop positions the seventh predicted β-strand into favorable hydrogen-bonding contacts with strands six and eight and completes the fold of the β-barrel. This strand also positions two conserved resi-

| | Bacterial | Swordfish | Bovine | Horse | Human | Yeast |
|---|---|---|---|---|---|---|
| Bacterial | | 44 | 44 | 44 | 43 | 43 |
| Swordfish | 30 | | 73 | 72 | 72 | 63 |
| Bovine | 28 | 71 | | 78 | 80 | 65 |
| Horse | 26 | 66 | 78 | | 79 | 66 |
| Human | 25 | 66 | 82 | 78 | | 66 |
| Yeast | 26 | 53 | 53 | 56 | 55 | |

FIG. 2. Sequence similarities between copper/zinc superoxide dismutases. The left-hand side of the diagonal shows percentage identity figures and the right-hand side shows weighted similarities in percentages.

Evolution: Bannister and Parker

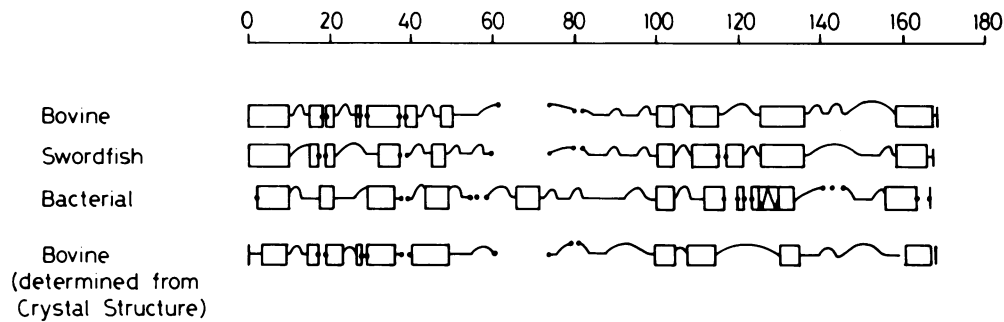*Proc. Natl. Acad. Sci. USA 82 (1985)* 151



FIG. 3. Comparison of predicted secondary structures of bovine erythrocyte, swordfish liver, and bacterial (*P. leiognathi*) copper/zinc superoxide dismutases. □, β-Strand; ▦, α-helix; ⌇, β-turn.

dues (His-135 and Asp-139), considered essential for activity in the bovine enzyme, near the active site. These cases demonstrate one of the problems encountered in secondary structure prediction methods. These methods do not take into account possible nonlocal interactions. The *P. leiognathi* sequence could be fitted into the fold of the bovine erythrocyte enzyme without serious disruption to the fold. No obvious steric problems were encountered as a result of the introduction of larger residues for smaller residues and all residue side chains pointing into the interior of the β-barrel were uncharged.

The presence of a copper/zinc superoxide dismutase in a symbiotic bacterial species initially appeared to be a case of direct gene transfer. However, although gene transfer among prokaryotes is a well-documented phenomenon, as is transfer from prokaryotic to eukaryotic species (28), gene transfer from eukaryotic to prokaryotic species has never been clearly proven. The amino acid sequence of *P. leiognathi* copper/zinc superoxide dismutase turned out to show low identity with bovine erythrocyte, human erythrocyte, horse liver, and yeast copper/zinc superoxide dismutases. Therefore, in 1983, the hypothesis of gene transfer was abandoned and an independent evolutionary line was suggested for the presence of the copper/zinc enzyme in a bacterial species (14). This hypothesis continued to be plausible when the sequence for the first fish copper/zinc superoxide dismutase, from

swordfish liver, was determined (15) and also showed very low identity with the bacterial enzyme.

The present investigation reexamines the possibility of a direct evolutionary relationship and of gene transfer for the presence of copper/zinc superoxide dismutase in *P. leiognathi*. The possibility of convergent evolution between lowly related proteins is always difficult to exclude. However, no example in which convergent evolution has led to close similarities in structure and sequence is known. The oxygen-binding proteins—hemoglobin, hemerythrin, and hemocyanin—are examples of convergent evolution, but their primary and tertiary structures are totally unrelated.

The following evidence establishes convincingly that the *P. leiognathi* copper/zinc superoxide dismutase is directly related to the eukaryotic copper/zinc enzymes: (*i*) similar molecular weight, subunit size, metal content, and catalytic activity; (*ii*) the single disulfide bridge, the seven metal ligands, and the two other residues shown to be important in the mechanism of the bovine erythrocyte enzyme are all found in the bacterial enzyme, in the correct sequence order; (*iii*) analysis of the aligned sequences shows that more than half of the bacterial sequence has closely related or identical residues in common with the swordfish liver enzyme after deletions and insertions are excluded; and (*iv*) very similar secondary structure is predicted for bacterial, swordfish liver, and bovine erythrocyte enzymes. The case for a diver-
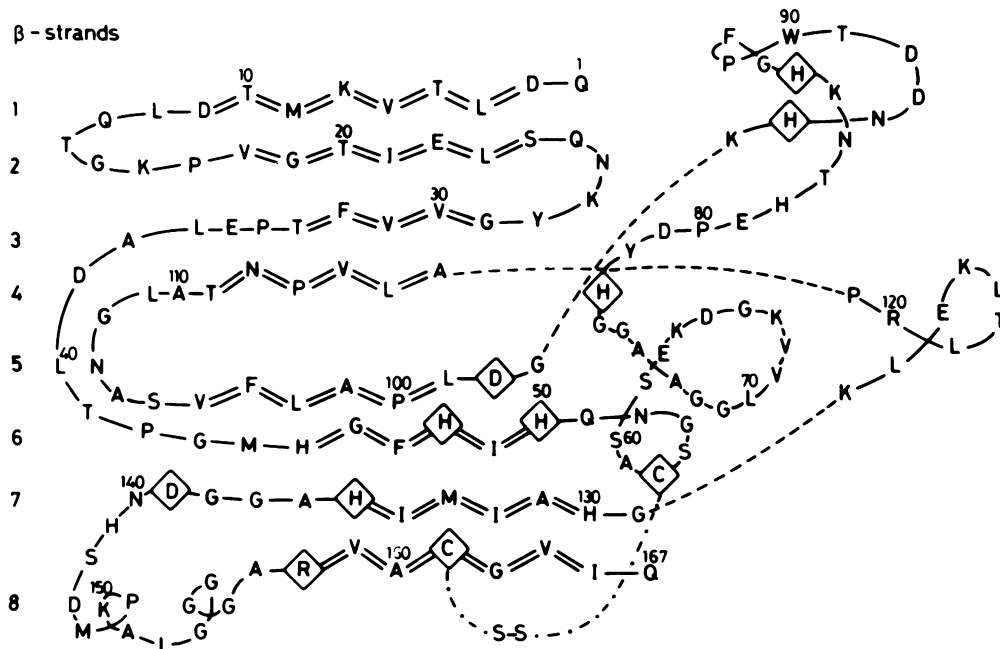


FIG. 4. Schematic topological diagram of bacterial (*P. leiognathi*) copper/zinc superoxide dismutase based on model building, by analogy with the bovine enzyme (27). β-Strands are drawn as zigzag lines with residues in the "valley" of the zigzag pointing into the interior of the β-barrel. Essential residues are boxed. Numbering is as in Fig. 1.

gent evolutionary relationship between the eukaryotic enzymes and the bacterial enzyme is not supported by the phylogenetic distribution of the enzymes. The copper/zinc superoxide dismutases have not been found in primitive eukaryotic species or in other bacteria [with one as yet unexplained exception (9)], including closely related marine bacterial species. The only simple explanation that remains is therefore gene transfer between the host ponyfish and the symbiotic bacterium *P. leiognathi*. This, however, leaves unexplained why the degree of identity between the bacterial enzyme sequence and the eukaryotic enzyme sequences is low. The presence of iron superoxide dismutase in the bacterial species implies that the introduction of the copper/zinc enzyme was not essential for the survival of the bacterium so that the introduced protein was able to accept mutation changes more readily in its new environment. The amino acid composition of the host ponyfish copper/zinc superoxide dismutase is much closer to that of the *P. leiognathi* enzyme than is the composition of the swordfish liver enzyme (12). The amino acid sequence of the ponyfish enzyme would be expected to show an even higher homology with the bacterial enzyme, as historical evidence of a gene transfer from eukaryotes to prokaryotes.

1. Bannister, J. V., Bannister, W. H., Hill, H. A. O. & Thornalley, P. J. (1982) *Life Chem. Rep.* **1**, 49–53.
2. Fridovich, I. (1972) *Acc. Chem. Res.* **5**, 321–326.
3. McCord, J. M. & Fridovich, I. (1969) *J. Biol. Chem.* **244**, 6049–6055.
4. Peeters-Joris, C., Vandervoorde, A. M. & Bandheim, P. (1975) *Biochem. J.* **150**, 31–39.
5. Bannister, W. H. & Bannister, J. V. (1982) *FEBS Lett.* **142**, 42–43.
6. Salin, M. L. & Bridges, S. M. (1980) in *Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase*, eds. Bannister, J. V. & Hill, H. A. O. (Elsevier/North-Holland, New York), pp. 176–284.
7. Bridges, S. M. & Salin, M. L. (1981) *Plant Physiol.* **68**, 275–278.
8. Puget, K. & Michelson, A. M. (1974) *Biochem. Biophys. Res. Commun.* **58**, 830–838.
9. Steinman, H. M. (1982) *J. Biol. Chem.* **157**, 10283–10292.
10. Asada, K., Kanematsu, S., Okaka, S. & Hayakawa, T. (1980) in *Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase*, eds. Bannister, J. V. & Hill, H. A. O. (Elsevier/North-Holland, New York), pp. 136–153.
11. Fridovich, I. (1975) *Annu. Rev. Biochem.* **44**, 147–159.
12. Martin, J. P. & Fridovich, I. (1981) *J. Biol. Chem.* **256**, 6080–6089.
13. Haneda, Y. (1950) *Pac. Sci.* **4**, 214–227.
14. Steffens, G. J., Bannister, J. V., Bannister, W. H., Flohe, L., Gunzler, W. A., Kim, S.-M. A. & Otting, F. (1983) *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 675–690.
15. Rocha, H., Bannister, W. H. & Bannister, J. V. (1984) *Eur. J. Biochem.*, in press.
16. Steinman, H. M., Naik, V. R., Abernethy, J. L. & Hill, R. L. (1974) *J. Biol. Chem.* **249**, 7325–7338.
17. Lerch, K. & Ammer, D. (1981) *J. Biol. Chem.* **256**, 11545–11551.
18. Steinman, H. M. (1980) *J. Biol. Chem.* **255**, 6758–6762.
19. Parker, M. W., Schinina, M. E., Bossa, F. & Bannister, J. V. (1984) *Inorg. Chim. Acta* **91**, 307–317.
20. McLachlan, A. D. (1971) *J. Mol. Biol.* **61**, 408–424.
21. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120.
22. Lim, V. I. (1974) *J. Mol. Biol.* **88**, 857–872; 873–894.
23. Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry* **13**, 211–222; 222–245.
24. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11**, 268–272.
25. Tainer, J. A., Getzoff, E. D., Richardson, J. S. & Richardson, D. C. (1980) in *2SOD:Cu,Zn Superoxide Dismutase Complete Atomic Co-ordinates*, eds. Richardson, D. C. & Richardson, J. S. (Brookhaven Protein Structure Data Bank, New York).
26. Bannister, J. V., Anastasi, H. & Bannister, W. H. (1977) *Comp. Biochem. Physiol. B* **56**, 235–238.
27. Tainer, J. A., Getzoff, E. D., Beem, K. M., Richardson, J. S. & Richardson, D. C. (1982) *J. Mol. Biol.* **160**, 181–217.
28. Drummond, M. (1979) *Nature (London)* **281**, 343–347.