Free Rad. Res. Comms., Vols. 12-13, pp. 363-370
Reprints available directly from the publisher
Photocopying permitted by license only

THE RATE OF Cu,Zn SUPEROXIDE DISMUTASE EVOLUTION

JAN KWIATOWSKI*, RICHARD R. HUDSON and FRANCISCO J. AYALA

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, U.S.A.

The rate of amino acid replacement in Cu,Zn SOD greatly departs from the expectations of the molecular clock. We examine 27 Cu,Zn SOD sequences available and conclude that: (1) the SOD enzymes from different mammal families differ from each other by roughly the same number of replacements, which is consistent with a simultaneous mammalian radiation; (2) over the most recent 60 million years (MY) the rate of SOD evolution is fairly high (15 aa/100 aa/100 MYR) and may be considered constant; (3) the rate of accumulation of amino acid replacements since the divergence of fungi, plants and animals to the present is inconstant along different branches of the evolutionary tree; moreover it steadily decreases with time, to the same extent in all lineages; (4) some comparisons exhibit divergences that are in any case greater than expected from a Poisson process on the assumption of a molecular clock; (5) plant chloroplast enzymes display fewer differences from each other than cytoplasmic ones; (6) bacteriocuprein (from *Photobacterium leiognathi)*, fluke and human extracellular SOD are all three extremely remotely related to one another and to the SOD of other eukaryotes.

The process of consistent decline of the rate of evolution of Cu, Zn SOD can be described by a number of mathematical functions. We explore simple models that assume constant rates and might be applicable to other proteins or genes that apparently evolve at disparate rates.

KEY WORDS: Molecular evolution, amino acid replacement rate, evolutionary clock, rates of evolution.

INTRODUCTION

The superoxide dismutases (SOD) are a group of three different proteins containing either Mn, Fe, or Cu and Zn in the active site.[1] The former two SODs can be traced to a common ancestor but the Cu, Zn enzyme displays no sequence homology with them, representing a separate evolutionary lineage and, therefore, an example of convergence at the molecular level. Mn and Fe SOD are encountered in both prokaryotic and eukaryotic organisms, while Cu, Zn SOD is found almost exclusively in eukaryotic cytoplasm, although a chloroplastic form is present in higher plants.[2]

The comparison of Cu, Zn SOD sequences from a wide range of organisms reveals an extreme inconstancy in the rate of amino acid (or nucleotide) substitutions.[3,4] Using all Cu, Zn SOD sequences available in the literature, we herein examine the rate of amino acid replacement of this enzyme.

MATERIALS AND METHODS

We analyze 27 amino acid sequences of Cu, Zn SOD from human, rat, mouse, horse,

*Corresponding author.

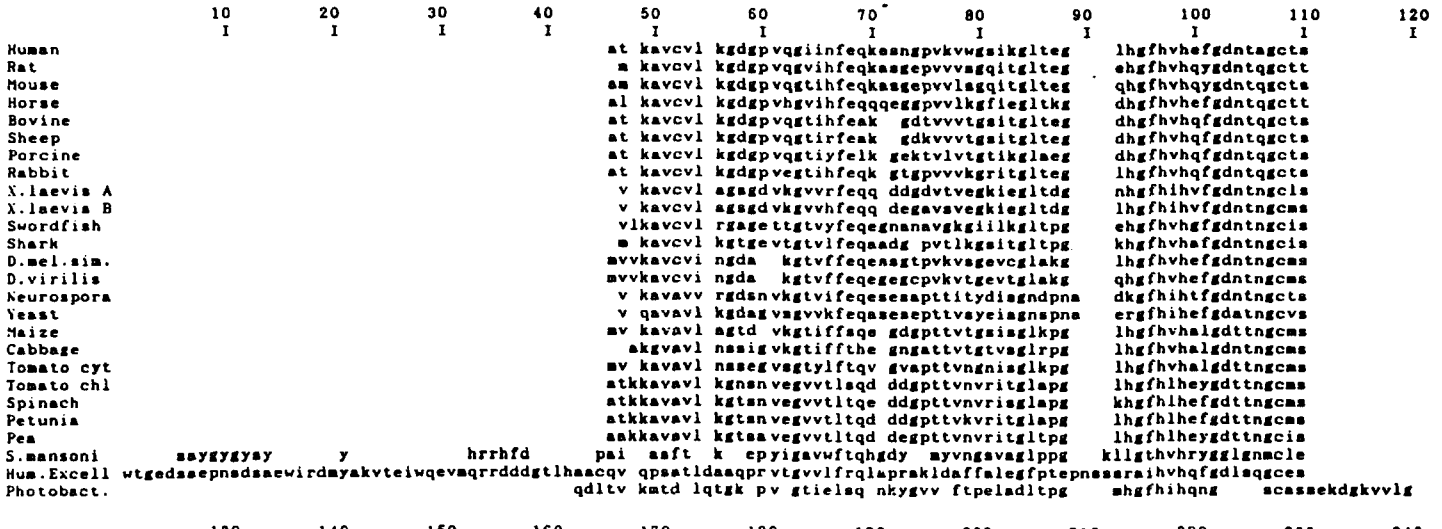

FIGURE 1 Aligned amino acid sequences of 26 Cu, Zn superoxide dismutases, *Drosophila melanogaster* and *D. simulans* have identical sequences.

TABLE I

Amino acid replacement among Cu, Zn SOD enzymes.
Differences shown were calculated for the whole molecules of approximately 150 aa; gaps introduced to maximize the alignments were not counted.

1 Human	2 Rat	3 Mouse	4 Horse	5 Cow	6 Sheep	7 Pig	8 Rabbi	9 Frog A	10 Frog B	11 Fish	12 Shark	13 D.mel	14 D.vir
1	26	25	30	26	28	26	21	51	48	47	45	58	60
2	Rat	5	29	24	26	28	25	51	51	50	41	60	59
3		Mouse	29	22	25	26	23	51	51	48	38	60	58
4			Horse	27	29	36	25	47	47	48	44	63	61
5				Cow	5	22	22	49	49	40	39	60	56
6					Sheep	25	24	47	49	42	41	61	57
7						Pig	26	54	51	46	49	62	60
8							Rabbit	49	49	46	39	60	59
9								Frog A	18	49	45	64	63
10									Frog B	46	48	58	59
11										Fish	46	61	62
12											Shark	60	57
13												D. melanogaster	20 D.virilis

TABLE I Cont...

	15 Mold	16 Yeast	17 Maize	18 Cabbg	19 Tomcy	20 Tomch	21 Spnch	22 Petun	23 Peach	24 Fluke	25 HumEC	26 Pbact
1	68	68	57	66	68	65	64	64	65	82	86	87
2	69	66	58	68	67	67	68	83	83	86		
3	67	68	57	67	66	68	68	69	66	84	84	85
4	68	65	63	72	72	68	68	67	68	86	85	86
5	65	68	60	67	69	69	67	67	68	81	83	82
6	65	67	60	66	69	67	65	65	66	80	82	82
7	68	71	57	61	67	70	67	69	71	84	83	85
8	66	66	57	65	66	60	59	57	60	80	84	83
9	66	63	55	63	63	63	61	61	61	77	82	82
10	64	62	52	58	60	60	59	58	58	78	81	83
11	63	65	61	69	66	71	68	70	67	80	81	80
12	66	65	59	66	62	62	58	56	58	80	83	84
13	65	69	49	57	55	71	67	69	71	85	89	91
14	65	69	50	56	56	65	63	65	69	83	88	89
15	N. crassa	47	63	66	64	68	65	63	69	86	87	85
16		Yeast	64	70	66	64	61	62	63	82	85	86
17			Maize cyt.	31	27	52	48	52	53	80	80	83
18				Cabbage cyt.	30	57	51	54	55	78	85	82
19					Tomato cyt.	56	50	55	54	77	82	87
20						Tomato chl.	16	11	16	79	86	86
21							Spinach chl.	15	18	79	84	89
22								Petunia chl.	18	79	85	88
23									Pea chl.	78	88	86
24										S. mansoni	102	91
25											Human excell.	94

TABLE 2
Amino acid differences among Cu, Zn superoxide dismutases and approximate time since their divergence

Organims compared	*aa* diff ± SE[1]	Div. time (Myr)
D. simulans - *D. melanogaster*	0	3
Mouse-Rat	5	8
Sheep-Cow	5	10
X. laevis isozymes A & B	18	30
D. virilis-D.melanogaster	20	40
Pig-Ruminants	23.5 1.5	45
Mammals	26.2 0.6	65
Flowing plants (chloropl.)	16.6 0.6	140
Flowing plants (cytopl.)	29.3 1.2	140
Frog-Mammals	49.4 0.5	380
Fish-Tetrapods	46.2 1.0	420
Shark-Bony fishes	43.2 1.1	480
Yeast-Mold	47	(500)?
SOD transfer to chloroplasts	52.6 0.8	(500)?
Vertebrata-Arthropoda	60.0 0.4	650
Plants-Animals	63.7 0.5	1200
Fungi-Plants	64.9 0.7	1200
Fungi-Animals	66.3 0.4	1200
Fluke-Cellular	80.9 0.6	(900)?
Human Extracell. -Cellular	84.2 0.5	?
Extracellular-Cellular	82.5 0.5	?
Baterial-Cellular	85.1 0.7	?
Extracellular-Bacterial	92.5	?

[1] Taken as averages from Table 1

ox, sheep pig, rabbit, *Xenopus laevis* (two isozymes), swordfish, shark, *Drosophila melanogaster, D. simulans, D. virilis, Neurospora crassa,* yeast, maize, cabbage, tomato (two isozymes), spinach, petunia, green pea, *Schistosoma mansoni, Photobacterium leiognathi* and human extracellular SOD[4-14] and ref. therein.

The sequences were aligned using FASTA program[15] and corrected by visual inspection using the program written by R. Tyler (Dept. of Ecology and Evolutionary Biology, UCI). The same program was used to calculate amino acid differences among proteins. The functions describing the rate of evolution were fitted by minimizing the sum of the squared differences between the values observed and expected. For the function $\Delta aa = xN(1 - \exp(-0.02t) + (1 - x)N(1 - \exp(-0.0003t))$ the parameter x was determined. For the function $\Delta aa = At/(B + t)$ the constants A and B were determined after conversion to a straight line equation, $t/\Delta aa = t/A + B/A$.

RESULTS

The alignment of the 27 amino acid sequences of Cu, Zn SOD is shown in Figure 1. These sequences repersent organisms from different taxons for which the divergence time can be estimated at least to a first approximation.

The number of amino acid replacements between any two species is shown in Table

I; the mean value of amino acid differences and divergence times for taxonomic groups with similar divergence time is shown in Table II. All mammal families considered seem to differ from one another by roughly the same number of replacements. This agrees with the nearly simultaneous radiation of those families, which occurred about 65–75 million years ago.[16] However, some organisms show smaller (e.g. rabbit) or larger (e.g. horse) differences than expected. When we compare mammals versus frog, swordfish versus tetrapods, and shark versus bony fishes, the differences are inversely related to the divergence times between these organisms. This may reflect a chance fluctuation: a variance greater than expected is often observed for molecular data.[17] However, the data exhibit systematic deviations: the fly enzymes are unexpectedly similar to the plant cytoplasmic ones, expecially that from maize, which shows consistently smaller distances to animal sequences than the rest of the plant enzymes. Moreover, plant chloroplast enzymes display much smaller differences between one another than the enzymes from plant cytoplasm. This suggests a distinctly different rate of evolution of chloroplastic enzymes, as opposed to the cytoplasmic ones, although it may also reflect closer relationship among the plants from which the chloroplastic sequences are obtained. The families of higher plants are thought to have radiated about 145 million years ago, but abundant homoplasies makes their relationships difficult to resolve.[18] With respect to chloroplast SOD, tomato and petunia, which belong to the same family, give, as expected, the smallest difference. Human extracellular and fluke SOD are almost as remotely distant from the rest of the eukaryotic enzymes and from each other as the SOD from *P. leiognathi* is from all eukaryotic SODs. This suggests (1) that the first two enzymes are not homologous to animal cytoplasmic superoxide dismutases and (2) that the amino acid differences between these enzymes have reached the maximum possible. They are excluded from further analysis.

DISCUSSION

It was first proposed by Zuckerkandl and Pauling that the rate of amino acid (and nucleotide) replacements in evolving proteins may be constant over time.[19] When comparisons are made between living organisms, it becomes necessary to take into account reversed and superimposed replacements, as well as the possibility that there might be sites at which a protein's function would not allow any replacements. A varitey of methods have been developed to correct the observed differences between extant organisms and estimate the rate of replacement.[3] However, various corrections applied to SOD sequences fail to show a constant rate of amino acid replacement, even though diverse correction methods yield convergent results.[20]

The set of 28 SOD sequences now available shows the same patterns previously observed,[3,20] but they also suggest that organisms that have diverged for similar periods exhibit comparable numbers of amino acid replacements. Thus, lineages diverged within the last 60 MY (comparisons among mammals, or *Drosophila* species, or *Xenopus* enzymes) are consistent with an observable rate of 15 aa/100 aa/100 MY (Figure 2, Table II). Lineages that diverged about 400 MY ago give an observable rate of about 7 *aa*/100 *aa*/100 MY, which is considerably lower than the previous but is roughly the same for comparisons between fishes, frogs and mammals; and comparisons between organisms from different kingdoms, which diverged about 1200 MY ago, give apparent rates of 3 *aa*/100 *aa*/100 MY.

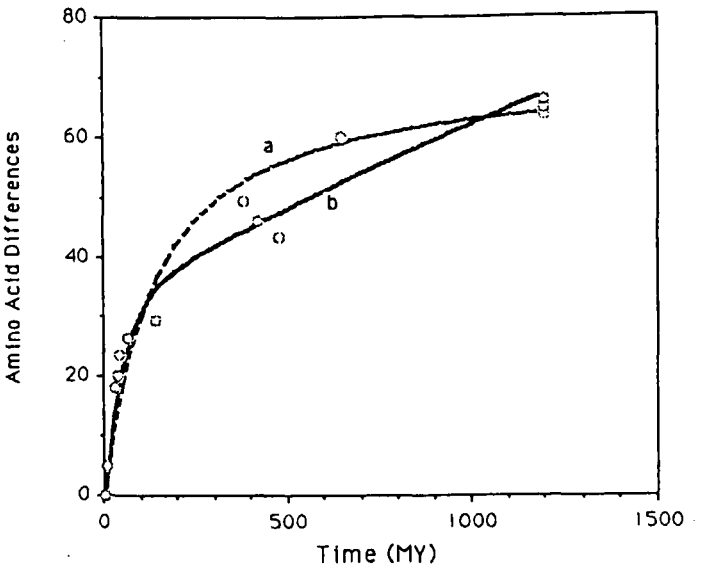

FIGURE 2 Amino acid replacements in cytoplasmic eukaryotic Cu, Zn superoxide dismutases. The two curves represent two different models of SOD evolution (see text).

We have explored a variety of models that assume constant rates of SOD evolution but would nevertheless fit the pattern of differentiation observed in SOD. Figure 2 shows the results obtained with two simple models. Curve *a* is the function of a rectangular parabola: number of amino acid differences, $\Delta aa = At/(B + t)$, where t is time, and A and B are constants estimated from the data. The model sees amino acid replacement as a saturation process that hence can be described by the Langmuir isotherm. Curve *b* is the equation $\Delta aa = xN(1 - \exp(-ut) + (1 - x)N(1 - \exp(-vt))$, where N is the length of a molecule (150 for SOD), x is the proportion of amino acids evolving at the neutral rate of evolution (u), and v is the rate of evolution of the other amino acids which is assumed to very highly constrained. We assume that u is the maximum rate observed in protein evolution and v is the smallest rate observed. Then, only one parameter, x, needs to be estimated from the data. This has been done for curve *b*, whereas we have assumed $u = 0.02$, $v = 0.0003$, and have estimated x as 0.21. Figure 2 shows that simple models that assume constant rates of evolution are able to account for the pattern of SOD evolution, at least with respect to the apparent large decrease of evolution rate with increasing divergence, which decrease does not seem readily accountable by available models.[3,20]

Cu, Zn superoxide dismutase is a relatively small dimeric molecule composed of two polypeptide chains of about 150 amino acids. Its tertiary structure has two characteristics. The overall topology has a beta-barrel configuration that provides the backbone of the molecule, which configuration is secured by about 25 glycine residues scattered throughout the sequence, a motif found in rapidly evolving proteins.[21] A second feature is the presence of a Y-shaped positively charged area, which incorporates the part of the molecule surface surrounding the active site.[22] This area, or

active site channel, is formed by two random coiled-loops of amino acid chain. The channel provides long range electrostatic guidance of the substrate anion molecule to the active center. The preservation of the function of this channel should not require strict preservation of specific amino acids in specific positions, but rather the overall charge topology of this region, which may be provided by different distribution of polar amino acid residues. These two structural features allow for the rapid evolution of SOD observed over the most recent 60 MY because many sites can be replaced without functional consequences in spite of the three-dimensional constraints imposed by the conservative configuration of the molecule.

Acknowledgement

We thank Mr. R. Tyler (UCI) for making available his LIGASE package of computer programs.

References

1. I. Fridovich (1986) Superoxide Dismutases. *Advances in Enzymology*, **58**, 61–97.
2. J. Kwiatowski (1987) Comparison of chloroplast and cytosolic Cu/Zn superoxide dismutase isozymes from tomato in relation to superoxide dismutase evolution. *Isozymes: Current Topics in Biological Medical Research*, **15**, 121–39.
3. Y.M. Lee, D.J. Friedman and F.J. Ayala (1985) Superoxide dismutase: An evolutionary puzzle. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 824–828.
4. G.J. Steffens, A.M. Michelson, K. Puget and L. Flohe (1986) The amino-acid sequence of rat Cu–Zn superoxide *Biological Chemistry Hoppe-Syeler*, **367**, 1017–1024.
5. J. Kwiatowski and F.J. Ayala (1989) *Drosophila virilis* Cu–Zn superoxide dismutase gene sequence. *Nucleic Acids Research*, **17**, 2133.
6. M.E. Shinina, D. Barra, S. Gentilomo, F. Bossa, C. Capo, G. Rotilio and L. Calabrese (1986) Primary structure of cationic Cu,Zn superoxide dismutase. The sheep enzyme. *FEBS Letters*, **207**, 7–10.
7. J. Kwiatowski, F. Gonzalez and F.J. Ayala (1989) *Drosophila simulans* Cu–Zn superoxide dismutase gene sequence. *Nucleic Acids Research*, **17**, 16735.
8. L. Calabrese, F. Polticelli, P. O'Neill, A. Galtieri, D. Barra, M. Shinina and F. Bossa (1989) Substitution of arginine for lysine 134 alters electrostatic parameters of the active site in shark Cu,Zn superoxide dismutase. *FEBS Letters*, **250**, 49–52.
9. J.R. Scioli and B.A. Zilinskas (1988) Cloning and characterization of a cDNA encoding the chloroplastic copper/zinc-superoxide dismutase from pea. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 7661–7665.
10. K. Reinecke, B. Wolf, A.M. Michelson, K. Puget, G.J. Steffens and L. Flohe (1988) The amino-acid sequence of rabbit Cu-Zn superoxide dismutase. *Biological Chemistry Hoppe-Seyler*, **369**, 715–725.
11. M.E. Schinina, D. Barra, F. Bossa, L. Calabrese, L. Montesano, M.T. Carri, P. Mariottini, F. Amaldi and G. Rotilio (1989) Primary structure from amino acid and cDNA sequences of two Cu,Zn superoxide dismutase variants from *Xenopus laevis*. *Archives of Biochemistry and Biophysics*, **272**, 507–515.
12. J. Tepperman, C. Katayama and P. Dunsmuir (1988) Cloning and nucleotide sequence of a petunia gene encoding a chloroplast-localized superoxide dismutase. *Plant Molecular Biology*, **11**, 871–872.
13. R. Perl-Treves, B. Nacmias, D. Aviv, E.P. Zeelon and E. Galun (1988) Isolation of two cDNA clones from tomato containing two different superoxide dismutase sequences. *Plant Molecular Biology*, **11**, 609–623.
14. M.C. Simurda, H. van Keulen, D.M. Rekosh and P.T. LoVerde (1988) *Schistosoma mansoni*: Identification and analysis of an mRNA and a gene encoding superoxide dismutase (Cu/Zn). *Experimental Parasitology*, **67**, 73–84.
15. W.R. Pearson and Lipman (1988) Improved tools for biological sequence analysis. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 2444–2448.
16. A.S. Romer (1966) *Vertebrate Paleontology* (3rd ed). University of Chicago Press, Chicago.
17. C.H. Langley and W.M. Fitch (1974) An estimation of the constancy of the rate of molecular evolution. *Journal of Molecular Evolution*, **3**, 161–177.

18. M. Syvanen, H. Hartman and P.F. Stevens (1989) Classical plant taxonomic ambiguities extended to the molecular level. *Journal of Molecular Evolution*, **28**, 536–544.
19. E. Zuckerkandl and L. Pauling (1965) Molecules as documents of evolutionary history. *Journal of Theoretical Biology*, **8**, 357–366.
20. F.J. Ayala (1986) On the virtues and pitfalls of the molecular evolutionary clock. *Journal of Heredity*, **77**, 226–235.
21. J.S. Richardson (1981) The anatomy and taxonomy of protein structure. *Advances in Protein Chemistry*, **34**, 168–339.
22. E.D. Getzoff, J.A. Tainer, P.K. Weiner, P.A. Kollman, J.S. Richardson and D.C. Richardson (1983) Electrostatic recognition between superoxide and copper, zinc superoxide dismutase. *Nature*, **306**, 287–290.

Accepted by Prof. G. Czapski.