Superoxide dismutase in *Drosophila melanogaster*

Mutation site difference between two electromorphs

Young Moo Lee and Francisco J. Ayala

Department of Genetics, University of California, Davis, CA 95616, USA

Received 8 October 1984

Two electrophoretically distinguishable variants of superoxide dismutase (SOD) are common in natural populations of *Drosophila melanogaster*. We have earlier comparatively characterized these two electromorphs, SOD^F and SOD^S. By peptide mapping in high-performance liquid chromatography (HPLC), we now show that the difference between the two electromorphs is due to the replacement of Asn-96 (SOD^F) by Lys-96 (SOD^S). It is far from clear how this replacement causes the biochemical differences (in thermostability, specific activity, and others) observed between these two forms of the enzyme.

Superoxide dismutase Drosophila *Structural polymorphism* *Genetic variation* *Evolution*

1. INTRODUCTION

Genetic variation is a necessary condition for evolution. Electrophoretic studies have demonstrated that genetic variation is pervasive. It is now estimated that sexually reproducing organisms are heterozygous with respect to electrophoretically detectable variants at 5–25% of all gene loci. A central issue at present is what is the adaptive significance of this variation.

It has been possible to demonstrate in some cases that allelic variants affect the adaptive fitness of organisms in different environments. It is important in such cases to ascertain the underlying structural and biochemical bases of such adaptive differences. The structural basis of biochemical polymorphisms is known in a number of cases where the variants are responsible for human infirmities. An example is the sickle-cell polymorphism, where the replacement of valine for glutamic acid at position six of the β polypeptide modifies the affinity for oxygen of the hemoglobin molecules.

Among genetic polymorphisms not associated with disease, one exceptional case for which the structural basis is known is ADH in the fruit fly *Drosophila melanogaster*. Two electrophoretically distinguishable forms of this enzyme exist in substantial frequencies in natural populations. The evidence indicates that this polymorphism has an adaptive basis [1,2]. The two common electrophoretically distinguishable forms of this enzyme differ by a single amino acid substitution [3].

The enzyme superoxide dismutase (SOD; EC 1.15.1.1) is polymorphic in various *Drosophila* species [4,5]. In *D. melanogaster* the two more common electrophoretic variants (SOD^S and SOD^F) differ in various properties, such as p*I*, specific ativity in a cytochrome-*c* assay, thermal stability, and other biochemical properties, as well as in amino acid composition [6,7]. In addition, we have obtained evidence in our laboratory that the two electrophoretic variants affect in vivo the adaptation of the organisms to different environmental conditions. Here, we report the structural basis of the two variants: the replacement of Asn-96 in SOD^F by Lys-96 in SOD^S.

2. MATERIALS AND METHODS

Strains of *Drosophila melanogaster* can be made homozygous for any part of the genome by ap-

Published by Elsevier Science Publishers B.V.
00145793/85/$3.30

propriate crosses with laboratory stocks. In our case, we obtained flies identically homozygous for the complete third chromosome by crosses with 'balancer' stocks [8]. The gene coding for copper-zinc SOD is located at 32.5 on the third chromosome of *D. melanogaster*; flies homozygous for the third chromosome, hence, carry two copies of the same allele at this locus. Using flies collected in various localities (California and Florida in the United States, Tunisia, and Japan), we obtained a variety of strains homozygous for either the SOD^S or the SOD^F forms of the enzyme. Flies homozygous for a given chromosome were multiplied by the hundreds of thousands in order to obtain the biomass necessary for the studies herein reported. The flies were collected 2–3 days after eclosion from the puparium.

The procedure for SOD purification was essentially as described in [6]. The protein was reduced and carboxymethylated with doubly recrystallized iodoacetic acid as in [9]. Carboxymethylated SOD (CM-SOD) was digested with a trypsin to protein ratio (w/w) of 1:100 in 0.2 M ammonium bicarbonate buffer (pH 8.2) for 6 h at 37°C. Following the digestion, the buffer was removed by repeated lyophilization.

Peptide mapping was performed by HPLC, using a Varian model 5000 equipped with a Rheodyne model 7125 sample injector. The reverse phase column used was Ultrasphere ODS (4.6 × 250 mm), 5 μm, from Beckman. For the analytical peptide mapping, we injected a sample of 0.6–1 nmol in 10 μl of 0.1% trifluoroacetic acid (TFA); the column was washed with 0.1% TFA for 5 min with a flow rate of 0.7 ml/min. Elution of the peptides was achieved in a linear gradient from 0.1% TFA to 60% of 0.075% of TFA in acetonitrile with a rate of solvent change of 1.1%/min. The eluting peptide peaks were recorded with a UV monitor (Perkin-Elmer model LC-75) at 218 nm. For sequencing the peptides we injected into the column samples of 5–10 nmol of the tryptic digests dissolved in 0.1% TFA using the Rheodyne model 7125 injector with a 100 μl sample loop.

We carried out the acid hydrolysis of peptides in 5.7 N HCl within sealed evacuated ignition tubes at 110°C for 24 h [10]. The hydrolyzates were analyzed using a single-column, three-buffer elution program in a Durrum model D-500 amino acid analyses. Sequence analysis was performed by automated Edman degradation in a Beckman model 890 sequencer. The PTH-amino acids were identified by at least two independent methods, including HPLC.

3. RESULTS

The copper-zinc superoxide dismutase in *D. melanogaster* is a dimer with two identical subunits, each with an M_r of approx. 16000 [7]. In natural populations, two forms of the enzyme exist, SOD^F and SOD^S, readily distinguished by standard starch-gel electrophoresis [11], because of their different migration rate. The frequency of SOD^F varies from one locality to another, mostly within the range 0.85–0.98, whereas SOD^S ranges from 0.02 to 0.15. Variants with slower electrophoretic migration than SOD^S or faster than SOD^F are occasionally found, but their frequencies are around 0.001 or lower.

Fig.1 shows the tryptic peptide maps of carboxymethylated SOD^F and SOD^S obtained under the conditions described. Seven SOD^F's obtained from different isogenic strains all yield peptide maps that are identical (in the number of peaks and overall configuration). The SOD^S map shown in fig.1 was obtained with a strain from Japan, which has the same amino acid composition as the SOD^S from two other strains originating from California and Tunisia.

The most significant differences between the maps in fig.1 are (i) the presence in the SOD^F map but absence in the SOD^S map of a peak at retention time (RT) of 29.33 min on the reverse phase column; and (ii) the presence in SOD^S but absence in SOD^F of a peak at RT 22.75. These two peaks are marked by arrows in the figure. In addition, the peak at RT 37.37 has a larger integrated value in the SOD^S map than in SOD^F (where the peak shows at RT 37.29). This would indicate that the SOD^S peak contains either a larger peptide or more peptides than the SOD^F peak.

For chemical characterization of the peaks in the SOD^F map, we injected 5–10 nmol of the tryptic digests to the column and used 0.5–1 nmol of the recovered peptides for amino acid composition analysis. The results are shown in table 1. The location of the peptides within the primary structure of the polypeptide was possible because we

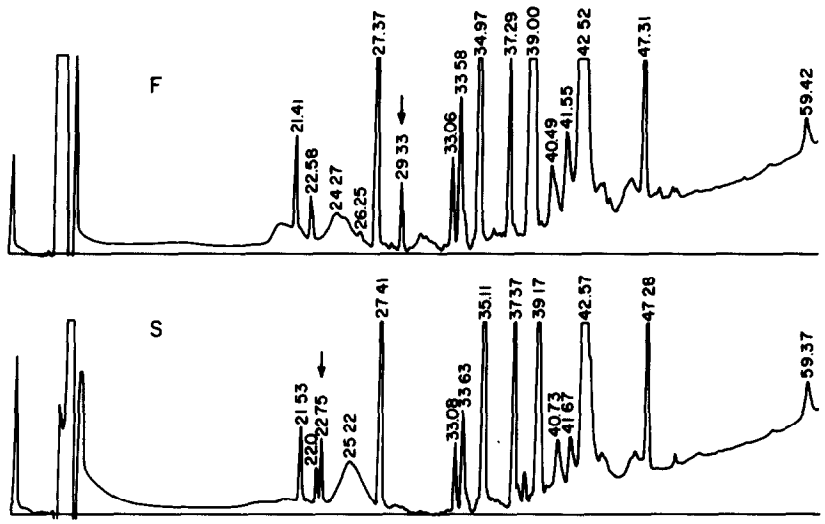

Fig.1. Peptide map of the tryptic digests of the SOD^F (above) and SOD^S (below) variants of the fruit fly *D. melanogaster*. The arrows indicate the most significant differences: the presence of a peak at 29.33 in SOD^F and at 22.75 in SOD^S. Numbers represent retention time in min. The graph is a UV scan at 218 nm.

Table 1

Amino acid composition of the tryptic peptides of carboxymethylated SOD^F from *D. melanogaster*

Retention time (min):	21.41	22.58	27.37	29.33	33.06	33.58	34.97	37.29	39.00[b]	42.52	47.31	37.02[c]
Tryptic peptide[a]	T-11	T-1	T-6	T-8	T-2	T-4	T-10	T-7	T-3, T-7′	T-5	T-9	T-12[c]
Residues[a]	134–141	1–3	68–77	95–101	4–13	29–38	114–133	78–94	–	39–67	102–113	142–151
CM-Cys					0.92(1)	0.70(1)		0.0 (1)		1.21(1)		0.62(1)
Asx	1.0 (1)		1.97(2)	1.90(2)	1.87(2)		2.96(3)	2.55(3)	2.52	3.0 (4)	1.03(1)	
Thr	0.94(1)			0.87(1)			0.80(1)	1.42(2)	3.05	0.89(1)	0.90(1)	
Ser	0.88(1)			1.12(1)		0.98(1)	0.93(1)		2.05	1.32(2)	0.89(1)	
Glx			1.97(2)			1.31(1)	1.98(2)	1.19(1)	3.28	1.10(1)		
Pro			1.0 (1)					0.91(1)	1.51	2.48(2)		
Gly	1.91(2)		1.0 (1)		1.12(1)	2.31(2)	3.0 (3)	2.48(3)	4.54	5.22(6)	2.01(2)	2.77(3)
Ala	1.90(2)		1.09(1)		1.71(2)	0.94(1)	2.13(2)	1.16(1)	1.0		1.0 (1)	1.01(1)
Val		1.77(2)	1.10(1)	0.97(1)	1.79(2)	1.90(2)	2.64(3)	0.78(0)	2.63	1.04(1)		1.29(1)
Met										0.76(1)		
Ile				0.76(1)	0.64(1)			0.94(1)	1.98		2.44(3)	2.41(3)
Leu						1.25(1)	1.88(2)	1.56(2)	1.14	0.90(1)	1.01(1)	
Tyr										0.99(1)		
Phe									1.83	3.22(3)	0.99(1)	
His			1.0 (1)				1.80(2)	0.72(1)	0.55	4.02(4)		
Lys		1.04(1)		1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	2.09	1.0 (1)		1.0 (1)
Arg	0.85(1)		0.87(1)								0.67(1)	
% yield	60	70	78	67	80	59	56	17	–	–	63	–

[a] Peptide and residue number and amino acid composition in parentheses are from the complete sequence determined by us

[b] T-7′, residues 78–101 due to tryptic resistant Lys 94

[c] RT of the peptide recovered by elution gradient of 0.1%–0.075% TFA in a 2-propanol system

know the complete amino acid sequence of SOD in *D. melanogaster* [12].

Ten tryptic peptides (out of the 12 expected) were recovered to homogeneity in 0.1–0.075% TFA in acetonitrile gradient elution (table 1). By direct sequencing of the peak at RT 42.52 we verified that it contains two peptides, T-5 (80%) and T-12 (20%). The amino acid composition analysis of the peak at RT 39.00 is shown in the table. Retention orders of the tryptic peptides (consisting of 3–15 residues) derived from SOD^F on the reverse column calculated from the sum of the Rekker fragmental constants of the hydrophobic amino acid residues [13] indicate that the T-3 peptide should be eluted at around RT 39. We have confirmed that the peak at RT 39.00 contains the T-3 peptide, but it also includes the T-7′ peptide. The peaks at RT 24.27, 40.49, 41.55, 59.42 and those emerging before application of the gradient in the SOD^F profile are also seen in the blank sample, which only contained trypsin and digestion buffer. As expected, no amino acids were detected in any of these elutes after 5.7 N HCl hydrolysis.

The peak at 29.33 in SOD^F was identified as the T-8 peptide (table 1). The peak at 22.75 in SOD^S elutes considerably earlier, indicating that it contains a peptide smaller and/or less hydrophobic than T-8. Earlier analyses of the amino acid composition of the whole protein had shown that SOD^F has one less lysine than SOD^S [6]. We inferred that one amino acid in T-8 had been replaced by lysine in SOD^S, which would yield a small peptide and account for the earlier elution (RT 22.75). This inference was confirmed by amino acid analysis and sequencing of the peptides at RT 29.33 and 22.75.

The peptide at RT 22.75 was eluted from the tryptic digest (10 nmol) of carboxymethylated SOD^S with a 30% recovery. Its amino acid composition is compared to that of T-8 (SOD^F) in table 2. It is apparent that the RT 22.75 peptide lacks one Val and one Asx. We infer that the Asn known to be residue 96 in SOD^F is replaced by Lys in SOD^S (residue 95 in SOD^F is Val). This is confirmed by direct sequence analysis (see table 3).

One peculiarity is that no PTH-serine is identified at the fourth cycle (residue 100) of the degradation of the peptide from RT 22.75, whereas PTH-serine is identified at the sixth cycle (also residue 100) of T-8. We have no explanation for this result (table 3). The amino acid analysis yields one serine in RT 22.75: this residue may be modified during HCl hydrolysis in a form that yields the free amino acid. We have made no search to find the Val-Lys dipeptide which should be generated by the trypsin digestion of SOD^S.

Table 2

Amino acid composition of the peptides RT 29.33 and 22.75

Amino acid	Molar ratio of amino acids in peptide	
	RT 29.33 (F)	RT 22.75 (S)
Asx	1.90	0.98
Thr	0.87	0.80
Ser	1.12	0.77
Val	0.97	–
Ile	0.76	0.91
Lys	1.0	1.0

Table 3

Sequence of the peptides RT 29.33 and RT 22.75

Cycle number	Sequence number	HPLC of PTH	TLC of PTH	GLC of PTH	Assignment
A. Peptide RT 29.33 (F)					
1	95	Val		Val	Val
2	96	Asn		Asn	Asn
3	97	Ile		Ile	Ile
4	98	Thr	Thr		Thr
5	99	Asp	Asp		Asp
6	100	Ser	Ser		Ser
7	101	Lys	Lys		Lys
8	–	–	–		–
B. Peptide RT 22.75 (S)					
1	97	Ile		Ile	Ile
2	98	Thr		Thr	Thr
3	99	Asp	Asp		Asp
4	100				Ser
5	101	Lys	Lys		Lys

The quantities of peptides used for sequence analysis are: RT 29.33, 5 nmol; RT 22.75, 3 nmol

4. DISCUSSION

The complete amino acid composition analysis of the polypeptides shows that, in *D. melanogaster*, SOD^F has one more negatively charged amino acid residue than SOD^S; previous evidence suggested that Asx (SOD^F) is replaced by Lys (SOD^S) [6]. We have used HPLC peptide mapping for locating the replaced residue. This simple technique has led us to identify the mutation site in the fruit fly SOD: Asn 96 in SOD^F to Lys 96 in SOD^S.

The 3-dimensional structure of bovine SOD is known [14]. The polypeptide consists of 151 residues in both the cow and the fruitfly; 57% of the residues are identical in both; in addition, all metal-binding ligands are conserved [12]. It therefore seems reasonable to assume that the 3-dimensional structure of the fruit fly SOD is similar to that of the cow. If so, the fruit fly Asn-96 residue is located, like in the cow, in the middle of the beta strand 4f (residues 91–99), which forms the outside frame of the beta barrel structure of the molecule. This is one of the less conserved regions among the 5 SOD primary structures known in eukaryotes [12].

In the cow SOD, Asp 96 does not set any main-chain hydrogen bonds with the adjacent 3c (residues 26–34) or 5e (80–89) strands; does not have any side-chain Van der Walls contacts; and the Asp side chain is exposed to solvent [15]. If Asn 96 in the fruit fly is like that in the cow, its mutation to Lys (or vice versa) would not modify the SOD structure in any major way. However, the stability of the molecule might or not be affected, depending on what kind of amino acid is present. Some amino acids but not others might form a salt bridge with amino acids in adjacent beta strands: Lys 96 in SOD^S (but not Asn 96 in SOD^F) might form a salt bridge to Glu 86 in 5e (E.D. Getzoff and J. Tainer, personal communication). This might account for some of the biochemical differences between SOD^S and SOF^F [6], but it remains for the future to ascertain how the single replacement of Asn 96 to Lys 96 modifies the specific activity, thermostability and other biochemical properties of the enzyme.

ACKNOWLEDGEMENTS

We thank Lorraine G. Barr for providing the isogenic liens of *D. melanogaster*, David J. Friedman for assistance in the enzyme preparations, K. Kanagaki for operation of the Durrum 500 Analyzer, and Al Smith for valuable suggestions concerning the peptide sequence determination by the microsequencing approach. This research was supported by grant GM 22221 from the US National Institute of Health and contract PA 200-14 Mod 4 with the US Department of Energy.

REFERENCES

[1] McDonald, J.F., Chambers, G.K., Davis, J. and Ayala, F.J. (1977) Proc. Natl. Acad. Sci. USA 74, 4562–4566.
[2] Clarke, B.C. (1975) Genetics 79, 101–113.
[3] Fletcher, T.S., Ayala, F.J., Thatcher, D.R. and Chambers, G.K. (1978) Proc. Natl. Acad. Sci. USA 75, 5609–5612.
[4] Ayala, F.J., Powell, J.R. and Dobzhansky, T. (1971) Proc. Natl. Acad. Sci. USA 68, 2480–2483.
[5] Ayala, F.J. and Powell, J.R. (1972) Proc. Natl. Acad. Sci. USA 69, 1094–1096.
[6] Lee, Y.M., Misra, H.P. and Ayala, F.J. (1981) Proc. Natl. Acad. Sci. USA 78, 7052–7055.
[7] Lee, Y.M., Ayala, F.J. and Misra, H.P. (1981) J. Biol. Chem. 256, 8506–8509.
[8] McDonald, J.F. and Ayala, F.J. (1978) Genetics 89, 371–388.
[9] Gracy, R.W. (1977) Methods Enzymol. 47, 195–204.
[10] Moore, S. and Stain, W.H. (1963) Methods Enzymol. 6, 819–831.
[11] Ayala, F.J., Powell, J.R., Tracey, M.L., Mourao, C.A. and Perez-Salas, S. (1972) Genetics 70, 113–139.
[12] Lee, Y.M., Friedman, D.J. and Ayala, F.J. (1985) Proc. Natl. Acad. Sci. USA 82, in press.
[13] Rekker, R. (1977) The Hydrophobic Fragmental Constant, Elsevier, Amsterdam, p.201.
[14] Richardson, J.S., Thomas, K.A., Rubin, B.H. and Richardson, D.C. (1975) Proc. Natl. Acad. Sci. USA 72, 1349–1353.
[15] Getzoff, E.D. (1982) PhD dissertation, Duke University.