Vol. 182, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1992 Pages 355-360

DROSOPHILA GLUTATHIONE *S*-TRANSFERASES HAVE SEQUENCE HOMOLOGY TO THE STRINGENT STARVATION PROTEIN OF *ESCHERICHIA COLI*

Yann-Pyng S. Toung and Chen-Pei D. Tu§

Department of Molecular and Cell Biology
The Pennsylvania State University, University Park, PA 16802

Received November 30, 1991

SUMMARY: The Drosophila glutathione *S*-transferase D genes encode a family of isozymes. We have determined the amino acid sequence of a new member of this family by nucleotide sequence analysis of a genomic DNA clone. The open reading frame of this intronless gene should encode an isozyme subunit of 211 amino acids. This sequence has significant homology to the *E. coli* stringent starvation protein, SSP, which is also a protein of two identical 211 amino acid subunits. The two proteins have very similar overall amino acid composition as well. It is possible that SSP may be a glutathione *S*-transferase(s) in *E. coli* or is evolutionarily related to glutathione S-transferases. Because SSP is known to be tightly associated with the RNA polymerase holoenzyme during purification, it is conceivable that Drosophila glutathione *S*-transferase(s) may potentially interact with the transcription machinery in a fashion similar to SSP's interaction with *E. coli* RNA polymerase holoenzyme.

The glutathione *S*-transferases (GSTs, EC 2.5.1.18) are dimeric proteins which are multifunctional in xenobiotic metabolism, drug biotransformation and protection against peroxidative damage (for recent reviews, see ref. 1,2). *Drosophila melanogaster* genome encodes at least two classes of GSTs which are immunologically unrelated to each other (3). We have previously isolated and characterized a Drosophila GST cDNA and its genomic DNA for isozyme DmGST 1-1 (4,5). The gene is intronless, and maps between the two hsp70 gene families at 87B on the right arm of chromosome 3 (5). We have completed the nucleotide sequence of a genomic clone containing several GST genes and pseudogenes. In this communication, we report the sequence of a second Drosophila GST gene, DmGST27 and its unusual homology to the *E. coli* stringent starvation protein, SSP. The implications of this sequence homology are discussed.

§To whom correspondence should be addressed.

0006-291X/92 $1.50

EXPERIMENTAL PROCEDURES

The genomic clone λGTDm101 has been reported earlier (4,5). The complete analysis of this part of the GST D gene family will be reported elsewhere (manuscript in preparation). DmGST27 is located at the end of λGTDm101 opposite to that of the DmGST1 gene (5). The DNA sequence analysis was performed on plasmid subclones with the dideoxynucleotide chain termination method and [α-^{35}S]dATP (specific activity >600 Ci/mmole, Amersham Corp., Arlington Heights, IL] (6-8). Exonuclease deficient T7 DNA polymerase was kindly provided by Dr. Kenneth A. Johnson of this Department. Nucleotide sequence data analysis was carried out with the DNA Inspector IIe program (Textco, Inc., W. Lebanon, NH) and the IntelliGenetics suite program via the Biotechnology Institute at The Pennsylvania State University.

RESULTS

An open reading frame of 212 amino acids can be deduced from the DNA sequence of a plasmid subclone which has sequence homology to the cDNA probe pGTDm1 (4). Therefore, a dimeric GST of 211 amino acid per subunit should result from this GST gene, designated DmGST27. After extensive search of the GenBank, N-Geneseq., and EMBL data bank for homologous genes, we found that DmGST27 has significant homology to the *E. coli* stringent starvation protein, SSP (9) as shown in Figure 1. Interestingly, *E. coli* SSP is also a dimeric protein of two identical 211 amino acid subunits (9,10). It is the major protein synthesized in *E. coli* under stringent starvation conditions (10,11). The amino acid compositions of DmGST27 and SSP are also very similar (Table 1).

DISCUSSION

The DmGST27 gene is a member of the GST D multigene family. The gene is functional as judged by RNA hybridization experiment using the unique 3′ noncoding region of the gene as hybridization probe (manuscript in preparation). The amino acid sequence of DmGST27 subunit is 67% identical to that of DmGST1 as shown in Figure 1. Experiments for expression of this GST subunit in *E. coli* and characterization of its enzyme activities are currently underway.

The sequence homology between DmGST27 and *E. coli* SSP is rather unexpected. This homology may be of biological significance, especially because both proteins are dimeric and identical in size, and have very similar overall amino acid compositions. The major difference is in the contents of charged and polar amino acids; the SSP has more arginines (16 vs 5) whereas the DmGST27 has more aspartic acids and asparagines (27 vs 17). The two proteins have 53 identical residues (25.1%) with 20 of them in a region of 44 amino acids which are homologous to the maize GST III (12).

The identical subunit size, dimeric structure, and significant sequence homology between Drosophila GST27 and *E. coli* SSP provided a basis for functional speculations. First, the *E. coli* SSP may have GST activity or be evolutionarily related to GST(s). *E. coli* grown under normal conditions of nutrients has very low level of conjugation activity with 1-chloro-2,4-dinitrobenzene (CDNB), the "universal" substrate for GST activities (13). Shishido has reported the occurrence a GST activity in *E. coli*, but no structural information of *E. coli*

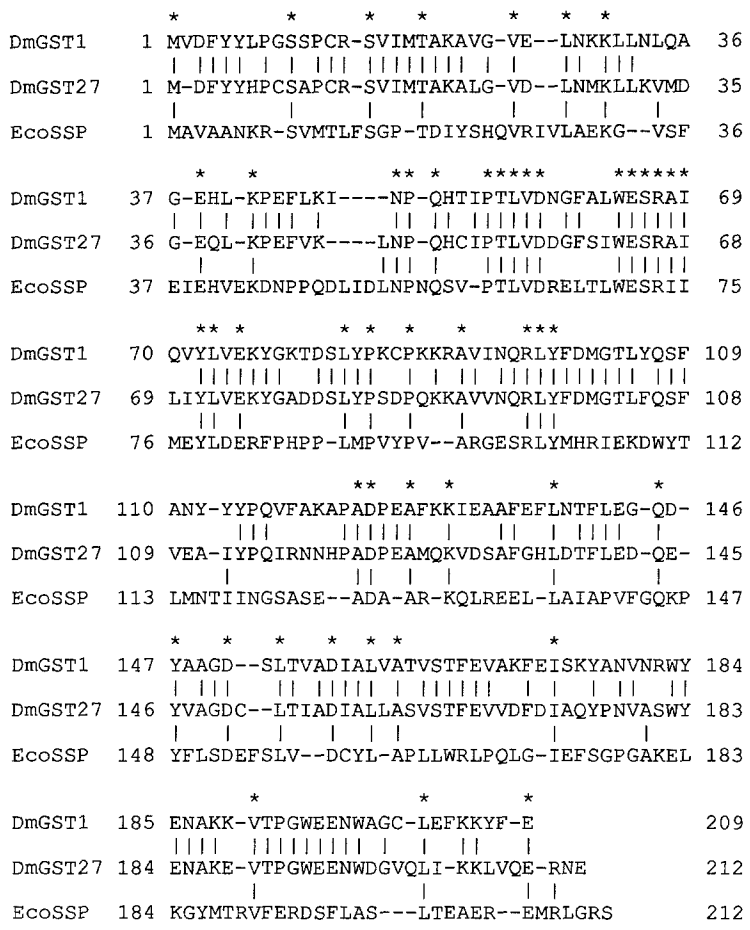

Figure 1. Amino acid sequence comparison among DMGST1, DMGST27, and *E. coli* SSP.

Vertical dashes between sequences indicate identical amino acids. Gaps to maximize homologies are represented by dashes. Conserved amino acids among all three sequences (a total of 48) are labeled by asterisk (*). Postitions of amino acids in three proteins are indicated on two sides.

GST is available in the literature (14). During the last few years, GSTs have been purified from several bacterial species such as *proteus mirabilis* and *Methylobacterium sp.* strain DM4 (15,16). The latter "GST" is a dichloromethane dehalogenase (16), which has sequence homology to Drosophila GST 1-1 but with a larger size (287 amino acids per subunit). It is likely that SSP is a GST, but one not very active against CDNB as substrate.

Since SSP is expressed abundantly under conditions of stringent starvation (11) it should be feasible to purify this protein to homogeneity and test it for GSH, or S-hexyl GSH binding and activities against a spectrum of GST substrates (13,17). It is interesting to note that SSP does not have near its N-terminal region (e.g. positions 4 and 5 of DmGST 1-1) any tyrosine residue(s) which is highly conserved among members of the GST gene superfamily (9,17). The tyrosine at position 7 in porcine GST π has been shown to be interacting with

Table 1. Amino Acid Compositions

Amino acid	DmGST27	SSP
Nonpolar		
Alanine	17	16
Isoleucine	11	11
Leucine	19	25
Methionine	6	8
Phenylalanine	10	9
Proline	12	15
Tryptophan	4	3
Valine	18	13
Total	97	100
Basic		
Arginine	5	16
Histidine	4	4
Lysine	12	8
Total	21	28
Polar		
Cysteine	4	1
Glycine	9	10
Serine	11	15
Threonine	7	8
Tyrosine	10	8
Total	41	42
Acidic/polar		
Asparagine	9	6
Aspartic acid	18	11
Glutamine	11	6
Glutamic acid	15	19
Total	53	42
Grand Total	212	212

the competitive inhibitor glutathione sulfonate by x-ray crystallography (18). Therefore, it is possible that SSP is evolutionarily related to GST, but has diverged from it to assume other functions under stressful conditions such as stringent starvation.

The second interesting speculation of this observed sequence homology originates from the observation by Ishihama and Saitoh (10). They reported that SSP was tightly associated with *E. coli* RNA polymerase holoenzyme but not with the core enzyme. The SSP copurified with the polymerase holoenzyme but was dissociated from it by phosphocellulose chromatography. More importantly, SSP influenced RNA synthesis *in vitro* by holoenzyme in a template dependent manner (10). Under stoichiometric conditions, for example, RNA synthesis *in vitro* from T7 DNA was inhibited up to 20% in the presence of SSP protein. However, SSP was not required for RNA synthesis *in vitro*. It is likely that such tight

interaction and its abundance under stringent starvation may influence the transcription specificity *in vivo*, conceivably by changing the selectivity of RNA transcription, among other possibilities. Using this tight association as the basis for further speculation, the homology between SSP and DmGST27 could suggest a possible interaction between GST27 (or other related isozymes) and Drosophila RNA polymerase(s) under stressful conditions. It is also interesting to note that the GST D genes are mapped at 87B on the polytene chromosome, a location between the two hsp70 gene clusters covered by a huge "puff" under heat shock conditions (19). This location is also in relative proximity to the gene encoding the second largest subunit of RNA polymerase II (88A/B), with which GST27 shares some nucleotide sequence homology near the 5' end (20). This possible interaction between GST and RNA polymerase(s) may have other interesting implications in several systems. For example, the biological significance of shistosomal GST(s) (21-23) has remained a mystery in parasitology. A report by Capron *et al.* (21) revealed that immunization of monkeys and several rodent species against shistosomal GST provided protection against shistosomiasis. It is possible that shistosomal GSTs may interfere with host functions by tight association with RNA polymerases and/or other important enzymes/proteins. Thus, prevention of shistosome GST interaction with host RNA polymerase(s) or other essential proteins by immunization may partially explain this protection. A second example involves an ethylene-responsive flower senescence-related gene from carnation reported by Meyer *et al.* (24). This gene encodes a protein of 220 amino acids which contain 44%, 53% and 52% homology with GSTs from maize, Drosophila, and man, respectively. It would be interesting to determine if this GST or GST-related protein is responsible for some aspects of the aging process or is synthesized as a stress response due to the major changes occurred in the aging process.

ACKNOWLEDGMENTS

We thank Randy Zauhar for directions in computer data search and Eileen McConnell for typing the manuscript. This work was supported by a Biomedical Research Support Grant to The Pennsylvania State University (2 S07 RR07082).

REFERENCES

1. *Glutathione S-Transferase and Drug Resistance* (eds. Hayes, J. D., Mantle, T. J., and Pickett, C. B.) Taylor and Francis, London, 1990.
2. Pickett, C. B. and Lu, A. Y. H. (1989) Ann. Rev. Biochem. 58, 743-764.
3. Tu, C.-P. D., Toung, Y. P. S., Hsieh, T.-s., Simkovich, N. M., and Tu, Y.-S. L. (1990) in *GSH S-Transferases and Drug Resistance* (eds. Hayes, J. D., Pickett, C. B., and Mantle, T. J.) Taylor and Francis Ltd., London and Philadelphia, pp. 379-386.
4. Toung, Y.-P. S., Hsieh, T.-s. and Tu, C.-P. D. (1990) Proc. Natl. Acad. Sci. USA 87, 31-35.
5. Toung, Y.-P. S., Hsieh, T.-s. and Tu, C.-P. D. (1991) Biochem. Biophys. Res. Commun. 178, 1205-1211.
6. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
7. Biggin, M. D., Gibson, T. J. and Hong, G. F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.

8. Chen, E. Y. and Seeberg, P. (1985) DNA 4, 165-170.
9. Serizawa, H. and Fukuda, R. (1987) Nucleic Acids Res. 15, 1153-1163.
10. Ishihama, A. and Saitoh, T. (1979) J. Mol. Biol. 129, 517-530.
11. Reeh, S., Pedersen, S., and Friesen, J. D. (1976) Molec. Gen. Genet. 149, 279-289.
12. Grove, G., Zarlengo, R. P., Timmerman, K. P., Li, N., Tam, M. F., and Tu, C.-P. D. (1988) Nucleic Acids Res. 16, 425-438.
13. Habig, W. H. and Jacoby, W. B. (1981) Methods Enzymol. 77, 231-235.
14. Shishido, T. (1981) Agri. Biol. Chem. 45, 2951-2953.
15. Di Ilio, C., Aceto, A., Piccolomini, R., Allocati, N., Faraone, A., Cellini, L., Ravagnan, G., and Federici, G. (1988) Biochem. J. 255, 971-975.
16. La Roche, S. and Leisinger, T. (1990) J. Bacteriol. 172, 164-171.
17. Lai, H.-C. J. and Tu, C.-P. D. (1986) J. Biol. Chem. 261, 13793-13799.
18. Reinemer, P., Dirr, H. W., Ladenstein, R., Shäffer, J., Gallay, U., and Huber, R. (1991) EMBO J. 10, 1997-2005.
19. Ashburner, M. and Bonner, J. J. (1980) Nucleic Acids Res. 8, 235-253.
20. Sitzler, S., Oldenburg, I., Petersen, G., and Bautz, E. K. (1991) Gene 100, 155-162.
21. Balloul, J.-M., Sondermeyer, P., Dreyer, D., Capron, M., Grzych, J. M., Pierce, R. J., Carvallo, D., Leconcq, J. P., and Capron, A. (1987) Nature (London) 326, 149-153.
22. Smith, D. B., Davern, K. M., Board, P. G., Tiu, W. V., Garcia, E. G., and Mitchell, G. F. (1986) Proc. Natl. Acad. Sci. USA 83, 8703-8707.
23. Taylor, J. B., Vidal, A., Torpier, G., Meyer, D. J., Roitsch, C., Balloul, J.-M., Southan, C., Sondermeyer, P., Pemble, S., Lecocq, L. P., Capron, A., and Ketterer, B. (1988) EMBO J. 7, 465-472.
24. Meyer, R. C., Goldsbrough, P. B., and Woodson, W. R. (1991) Plant Mol. Biol. 17, 277-281.