Biochem. J. (1991) 275, 227–231 (Printed in Great Britain)

Molecular cloning, sequencing and characterization of cDNA to rat liver rhodanese, a thiosulphate sulphurtransferase

Katherine L. WEILAND and Thomas P. DOOLEY*
Upjohn Laboratories, Cancer and Infectious Diseases 7252-25-12, Kalamazoo, MI 49001, U.S.A.

Rhodanese (EC 2.8.1.1), a mitochondrial thiosulphate sulphurtransferase, is involved in the formation of iron–sulphur complexes and cyanide detoxification. By screening a rat liver cDNA library with oligonucleotide probes complementary to portions of the published bovine rhodanese peptide sequence, rat rhodanese cDNA clones were obtained and sequenced. Comparison of the rat rhodanese cDNA open reading frame with the bovine peptide sequence demonstrated in the rat open reading frame the presence of 27 amino acid substitutions, only five of which are highly non-conservative. Thus the rat enzyme is approx. 91% identical with bovine rhodanese, or about 98% similar when conservative substitutions are considered. In addition, the rat translation product contains a Gly-Lys-Ala *C*-terminal tripeptide that was not observed in the bovine peptide sequence. All cysteine and proline residues are invariant between the two mammalian proteins. Computer-generated structural modelling of rat rhodanese indicated that few amino acid substitutions were present within close proximity to the active site or within the hinge region (connecting loop) between the A and B domains. Furthermore, evidence is presented showing that rhodanese is highly conserved at the DNA level among rodents, primates and a variety of other vertebrates.

INTRODUCTION

Rhodanese (EC 2.8.1.1), a thiosulphate sulphurtransferase present in both eukaryotes and prokaryotes, has been extensively studied at the biochemical level. In eukaryotic cells, this mitochondrial enzyme, found in numerous tissues, is the product of a nuclear-encoded gene. Although the biochemical roles of rhodanese *in vivo* are not fully understood, this enzyme has been implicated in the formation of iron–sulphur complexes (Cerletti, 1986) and cyanide detoxification. Rhodanese accomplishes this by covalently transferring the sulphane sulphur atom of thiosulphate to cyanide, forming thiocyanate, thus preventing cyanide-mediated inhibition of electron transport at the cytochrome *c* oxidase step (see reviews by Westley, 1973, 1981).

The primary amino acid sequence of bovine liver rhodanese has been determined. This enzyme is expressed as a single polypeptide chain of 293 amino acid residues (Russell *et al.*, 1978). X-ray crystallography has demonstrated that bovine rhodanese consists of two globular domains, A and B (residues 1–142 and 159–293 respectively), separated by a short connecting hinge region (Ploegman *et al.*, 1978). The A and B domains exhibit a high degree of tertiary structural similarity, although at the primary sequence level virtually no similarity is noticeable. The active site of bovine rhodanese is contained within the B domain, where Cys-247 acts as an intermediate acceptor of the sulphur atom from thiosulphate. Even though a considerable amount is known about rhodanese enzymology, little is known to date about the contribution of the A domain in catalysis and mitochondrial localization, and before the present work the cloning of a rhodanese gene from any species had not yet been published.

We have obtained cDNA clones to the rat liver rhodanese gene by screening a cDNA library with oligonucleotide probes complementary to portions of the known bovine rhodanese peptide sequence, which was assumed to be highly homologous between the two mammalian species. In the present paper we describe sequence and structural comparisons of rhodanese from rat and cow, based on the cDNA sequence derived from rat.

EXPERIMENTAL

General

Methods for λ phage propagation and library screening, plasmid isolation and buffers used were as previously described (Maniatis *et al.*, 1982). A rat liver cDNA library was purchased from Clontech Laboratories (Palo Alto, CA, U.S.A.). This λgt11 '5′ stretch' cDNA library was constructed from polyadenylated mRNA from adult male Sprague–Dawley rat livers.

Oligonucleotide probes

Oligonucleotides designed to isolate the rat rhodanese cDNA were based on the bovine rhodanese peptide sequence (Russell *et al.*, 1978). Two oligonucleotides were synthesized corresponding to regions of the amino acid sequence that had the least redundancy at the third position of the codon. The nucleotides inserted at the third position of redundant amino acids were chosen on the basis of rat codon usage. RhoI, 5′-GTGTGGT-GGATGTTC$^{A}/_{C}$GIGTGTTCGG-3′, was synthesized and corresponds to amino acid residues 111–119 of the bovine rhodanese peptide sequence ('I' refers to inosine), and RhoII, 5′-GTCAGGAAGTCCATGAAIGGCATGTTCAC-3′, corresponds to amino acid residues 208–217 of the cow enzyme. The oligonucleotides were end-labelled with [γ-^{32}P]ATP with the use of T4 polynucleotide kinase. The specific radioactivity of each probe was 5×10^8 c.p.m./μg.

DNA sequencing

The cDNAs from positive plaques were subcloned into pGEM-3Z (Promega Corp., Madison, WI, U.S.A.) and transformed into *Escherichia coli* JM109 (Yanisch-Perron *et al.*, 1985). Transformants were grown in LB broth containing 100 μg of ampi-

* To whom correspondence should be sent, at present address: Department of Genetics, Southwest Foundation for Biomedical Research, West Loop 410 at Military Drive, San Antonio, TX 78228, U.S.A.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X56228.

cillin/ml, and the plasmids were isolated by the alkaline lysis method (Maniatis *et al.*, 1982). Double-stranded plasmid DNA was denatured in 0.2 M-NaOH/0.02 mM-EDTA at room temperature for 5 min. The solution was adjusted to 0.3 M-sodium acetate and the DNA was precipitated with ethanol. The nucleotide sequence was determined by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977) with the use of Sequenase DNA sequencing kit (United States Biochemical Corp., Cleveland, OH, U.S.A.) and [^{32}P]dATP. The reaction products were electrophoresed on 6–8% acrylamide gels containing 8 M-urea.

Three different rhodanese cDNA clones (pRhoB2, 0.67 kb; pRhoC, 0.99 kb; pRhoD, 1.0 kb) were sequenced to obtain a composite rat rhodanese DNA sequence. The ends of each subclone were sequenced by using SP6 or T7 primers complementary to priming sites of the vector. Restriction analysis of subclones revealed a unique *Kpn*I site that was used to delete the vector *Kpn*I–*Kpn*I portion of the pRhoC gene, followed by religation for internal sequencing (RhoK subclone, 0.44 kb). Additional oligonucleotides were synthesized according to the derived DNA sequence and used as primers in reactions to obtain additional sequence information. The complete sequence was obtained for both strands of DNA and translated into polypeptide sequence for comparison with the cow rhodanese peptide.

Southern-blot analysis

A Zooblot nylon membrane containing 5 μg of *Eco*RI-cut genomic DNA from various species was obtained from Clontech Laboratories. The sources of the genomic DNA were: lane 1, human; lane 2, rhesus monkey; lane 3, Sprague–Dawley rat; lane 4, Balb/c mouse; lane 5, dog; lane 6, cow; lane 7, rabbit; lane 8, chicken. The blot was pre-hybridized at 65 °C for 1.5 h in 10 ml of 6 × SSPE/5 × Denhardt's solution/0.5% SDS containing 0.5 mg of boiled and sheared salmon sperm DNA. [1 × SSPE is 0.15 M-NaCl/10 mM-sodium phosphate buffer, pH 7.4, containing 1 mM-EDTA; 1 × Denhardt's solution is 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% BSA.] The probe was labelled by random oligonucleotide primers, [α-^{32}P]dATP and DNA polymerase I Klenow fragment with approx. 100 ng of a 1 kb *Eco*RI fragment containing the rat rhodanese cDNA (pRhoC, extending from Leu-5 to the 3′ untranslated portion). This boiled probe (approx. 10 μCi/100 ng) was then added for hybridization for 26 h at 65 °C. Following the hybridization, the filter was rinsed twice in 2 × SSPE (65 °C, 15 min) and once in 2 × SSPE/0.1% SDS (65 °C, 15 min). Kodak XAR-5 film was exposed to this blot for 5 days at −70 °C with an intensifier screen.

Structural model of rat rhodanese

The predicted structure of rat rhodanese was generated using the Mosaic molecular modelling system, developed in Computational Chemistry at Upjohn Laboratories (J. W. Howe, J. R. Blinn, J. P. Moon, G. J. White, T. R. Hagadone & M. W. Shulz, unpublished work) and is based upon the Macromodel program (developed by C. Still and co-workers, Columbia University, New York, NY, U.S.A.). The X-ray crystal structure of bovine rhodanese was extracted from the Brookhaven National Laboratory protein data bank. The amino acid residue changes between rat and cow were then substituted manually. An assumed Met-Val dipeptide was incorporated at the *N*-terminus, this being based on the presumed bovine *N*-terminal sequence. The structures of domains A (plus the hinge region) and B were then optimized separately by soft-shell minimization around the modified residues. Then a composite of the optimized A and B domains was generated in stereo, which displayed only the side chains of the substituted rat residues on an α-carbon backbone of rhodanese.

RESULTS AND DISCUSSION

In order to obtain cDNA clones to the rat liver rhodanese mRNA, we probed a rat liver cDNA library with radiolabelled oligonucleotide probes complementary to portions of the known bovine peptide sequence (Russell *et al.*, 1978). The *Eco*RI DNA inserts from the positive clones were subcloned and subjected to DNA sequencing by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977). The strategy to obtain the rat rhodanese cDNA sequence is shown in Fig. 1. The cDNA and derived polypeptide sequences for rat rhodanese are presented in Fig. 2. The translation product of the rat open reading frame extends for at least 885 nucleotide residues followed by a short 3′ untranslated sequence. Although attempts were made to obtain a full-length cDNA encoding the *N*-terminus of rat rhodanese, none of the positive clones that were obtained from this library extended 5′ beyond His-2 of the bovine sequence.

A comparison of the rat amino acid sequence derived from its open reading frame and the published bovine peptide sequence is shown in Fig. 3. A high degree of primary sequence similarity is evident between the two mammals. The cDNA sequence data provide further proof that mammalian rhodanese is expressed as a single polypeptide chain encoding both A and B domains. Assuming that the bovine rhodanese sequence obtained by peptide sequencing accurately represents its cDNA open reading frame, rhodanese from the two species differs at only 27 out of 289 amino acid residues, not including the *C*-terminal tripeptide. In other words, they share 91% identical amino acid sequence. However, if one takes into account the degree of conservative amino acid substitutions made at these sites, only five of the 27 differences are strongly non-conservative (Dayhoff *et al.*, 1978). These non-conservative substitutions occur at positions Leu-48,

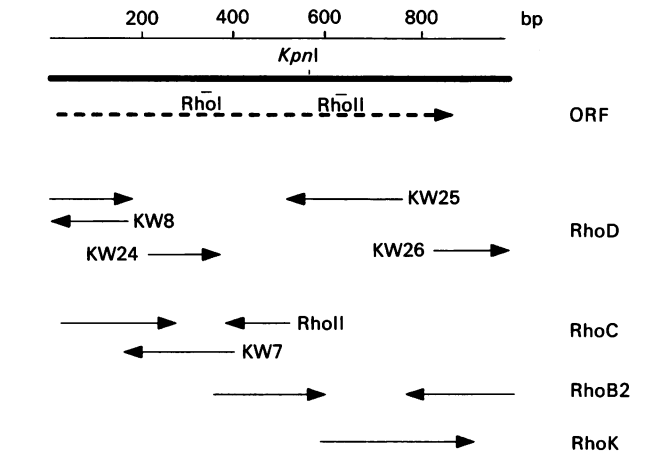

Fig. 1. Sequence strategy of the rat rhodanese gene

The broken line represents the mRNA open reading frame (ORF) of 885 nucleotide residues. The sites corresponding to oligonucleotide probes (RhoI and RhoII) used to obtain the cDNA clones are presented. Three different cDNA subclones in pGem-3Z (pRhoD, 1.0 kb; pRhoC, 0.99 kb; pRhoB2, 0.67 kb) and a deleted clone (pRhoK, 0.44 kb) were sequenced with the use of SP6 and T7 primers in dideoxy chain-termination reactions. The arrows indicate the regions and directions in which the nucleotide chain was synthesized. Those reactions that were primed by oligonucleotides are indicated with the name of the oligonucleotide next to the arrow (RhoII, KW7, KW8, KW24, KW25, KW26).

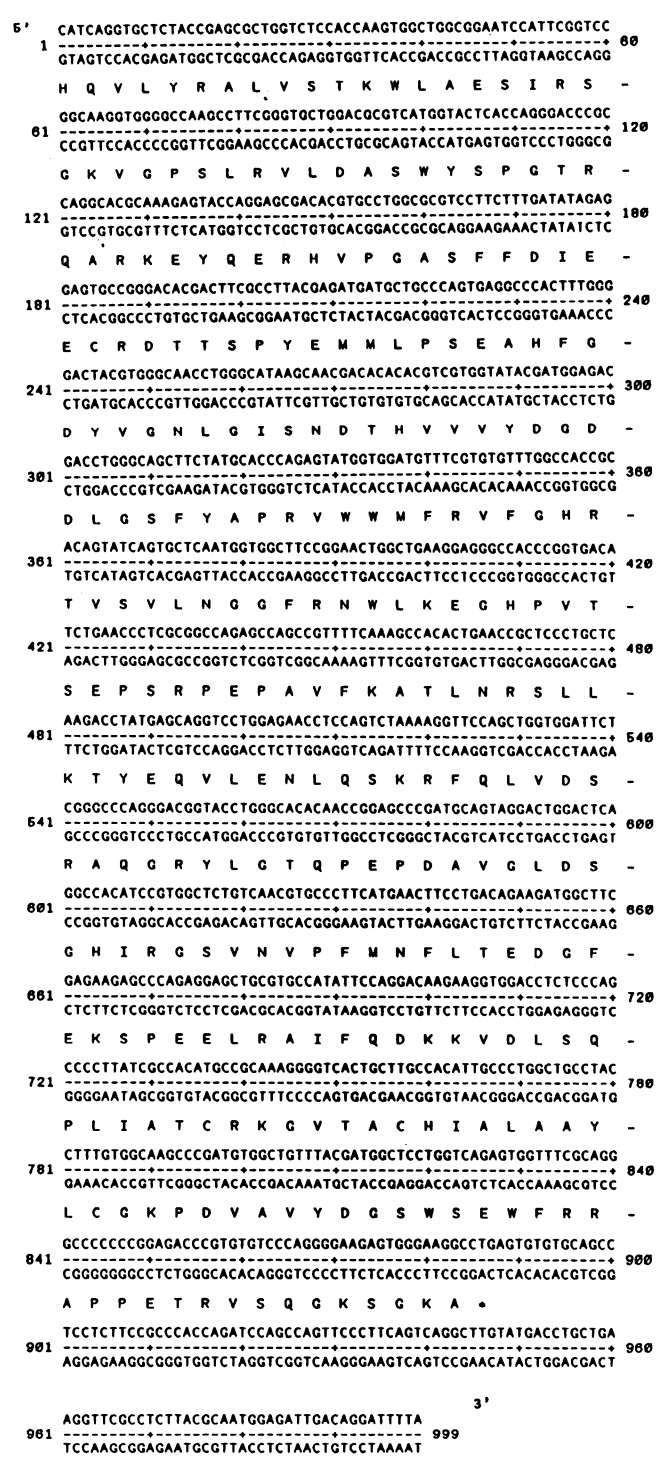

Fig. 2. Rat liver rhodanese cDNA sequence

The rat rhodanese composite sequence was obtained by sequencing cDNA inserts from several phage clones. The 295-amino acid-residue open reading frame is represented by the one-letter code for amino acid sequences. The 5′-end of the longest cDNA (pRhoD) was preceded by the following linker nucleotides, GAATTCCG. The precise 3′-end of the cDNA was identical in two different clones and is immediately followed by an *Eco*RI site. Position 1 of this cDNA sequence corresponds to the first position of the His-2 codon of bovine rhodanese. The unique *Kpn*I site is located at position 555 in the rat sequence.

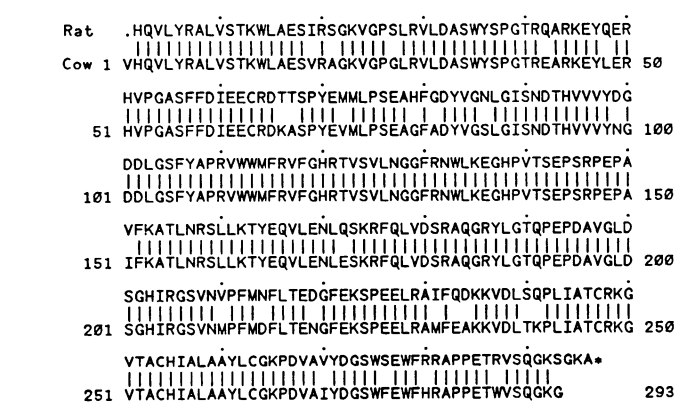

Fig. 3. Comparison of rat and cow rhodaneses

The polypeptide sequence of cow rhodanese (Russell *et al.*, 1978) was compared with the rat cDNA open reading frame. The 3′ tripeptide, derived from DNA sequence of rat rhodanese, was not present in the cow polypeptide sequence.

Lys-66, Gly-79, Ala-234 and Phe-276 relative to the bovine sequence. Thus at the primary amino acid sequence level rhodanese sequences from rat and cow are approx. 98% similar. When the rat cDNA is compared with the assumed bovine DNA sequence (derived from its polypeptide sequence), all but three of the 27 differences could be due to single base-pair mutations (or possibly peptide-sequencing errors); the exceptions are at bovine amino acid residues Gly-79, Phe-276 and Trp-287. Thus 24 changes are probably due to single base-pair mutations, of which 17 occurred at position 1 of the codon, assuming the published peptide sequence of bovine rhodanese is identical with its cDNA open reading frame.

Various aspects of this inter-species comparison at the protein level are worth noting. Of the 27 amino acid residue differences between rat and cow rhodanese, no proline or cysteine residues are affected, including the presence of a cysteine in the rat gene at the same site as bovine Cys-247, indicating that this active-site sulphane sulphur acceptor is conserved. Also, bovine residues Arg-186 and Lys-249 have been suggested previously to be involved in substrate binding, and these residues are conserved in the rat enzyme as well (Ploegman *et al.*, 1978). X-ray crystallography of bovine rhodanese demonstrated the presence of two parallel β-sheet structural motifs contained within the A and B domains (Ploegman *et al.*, 1978). Only five of the 54 amino acid residues differ within the parallel β-sheets. These differences occur at bovine residues 27, 99, 210, 241 and 270, and all are conservative substitutions. This observation suggests that the β-sheet motifs within rat rhodanese are not likely to be significantly altered, with respect to the bovine enzyme's structure. Most surprisingly, the hinge region between the A and B domains (bovine amino acid residues 143–158), which is located predominantly on the external surface of the A domain, was unaffected by amino acid substitutions with the exception of a single Ile→Val conservative substitution (amino acid residue 151). We have also discovered a *C*-terminal Gly-Lys-Ala tripeptide sequence (adjacent to the UGA translational termination codon in the mRNA) in rat rhodanese that had not been found in the bovine peptide sequencing efforts.

In order to predict the effects of the 27 amino acid substitutions and terminal additions on the three-dimensional structure of rat rhodanese, a computer-generated structural analysis was performed on rat rhodanese. By introducing the rat substitutions into a model of the known X-ray-crystallographic structure of bovine rhodanese, followed by soft-shell minimization around

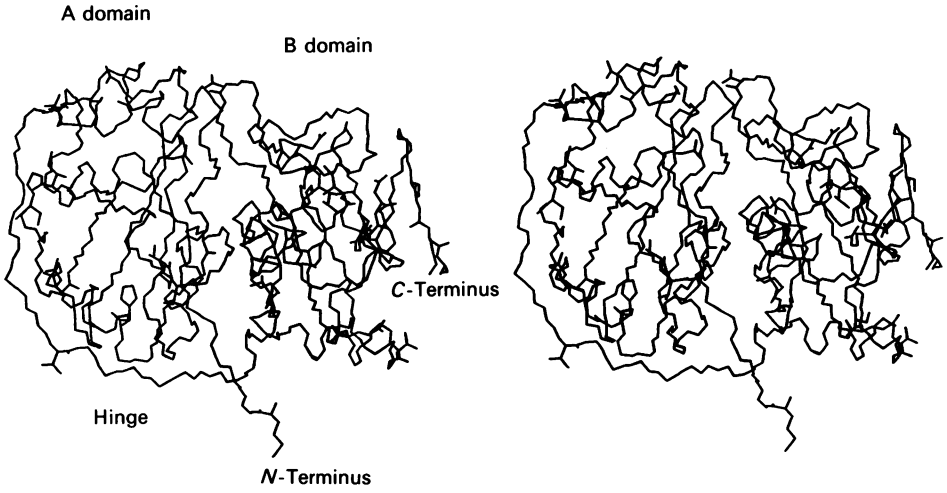

Fig. 4. Predicted molecular model of rat rhodanese

A stereo image of the computer-generated model of rat liver rhodanese is shown diagrammatically, displaying the α-carbon backbone and the side chains of the rat amino acid substitutions, relative to the bovine sequence (a detailed description of how this model was produced is found in the Experimental section). The A domain represents His-2–Ser-142, and includes a presumed Met-Val (−1 to +1) dipeptide at the *N*-terminus. The interdomain connecting loop (Hinge) consists of the region from Glu-143 to Arg-158. The rat conservative amino acid substitution side chain of Val-151 is visible on the periphery of this image. The B domain extends from Ser-159 to the *C*-terminal Gly-Lys-Ala tripeptide at positions 294–296.

these alterations, an energetically optimized model of the rat enzyme's structure was generated. The α-carbon backbone tracing of this molecule, with substituted side chains, is presented as a stereo image in Fig. 4. Only four of the amino acid differences between the two species are likely to be located with 1 nm (10 Å) of the active-site Cys-247 atoms. The active-site-proximal residues are Ser-276, Val-270, Val-210 and Met-72 in rat. None of these substitutions appears likely to exert profound effects on the active site of rhodanese. Ploegman *et al.* (1978) suggested that the following residues were likely to perform structural and catalytic roles in the bovine rhodanese active site: the hydrophobic residues Phe-212, Phe-106, Tyr-107, Trp-35 and Val-251 and the hydrophilic residues Asp-180, Ser-181, Arg-182, Arg-186, Glu-193, Arg-248, Lys-249 and Thr-252. Yet none of these presumably critical residues is altered in the rodent liver rhodanese open reading frame. Taken together, these data strongly suggest that the active sites of bovine and rat rhodanese are essentially identical.

Also noticeable in the rat rhodanese structural model are the solvent-exposed external domains of the *N*- and *C*-termini. The external location of these portions of rhodanese might allow them to be exposed to proteinases or other modifying enzymes, thus accounting for their absence from the derived bovine peptide sequence. Furthermore, by analogy with other nuclear-encoded gene products that are transported to the mitochondria, the *N*-terminus of rhodanese presumably plays a role in mitochondrial transport (reviewed in Hurt & van Loon, 1986). Also, the *N*-terminal region of rat rhodanese is identical with the bovine enzyme, extending for at least 17 amino acid residues. This rhodanese *N*-terminal region exhibits both basic and hydroxy-group-containing amino acids, and lacks acidic residues, all of which are important features for the mitochondrial transport of other nuclear-encoded gene products (Hurt & van Loon, 1986).

The apparent high degree of amino acid sequence similarity between cow and rat rhodanese prompted us to determine the extent of DNA sequence similarity between rat and other vertebrates. To accomplish this a Southern blot containing *Eco*RI-digested genomic DNA from various mammals and chicken was hybridized with a rat rhodanese radiolabelled probe (spanning from Leu-5 to the 3′ untranslated end). The

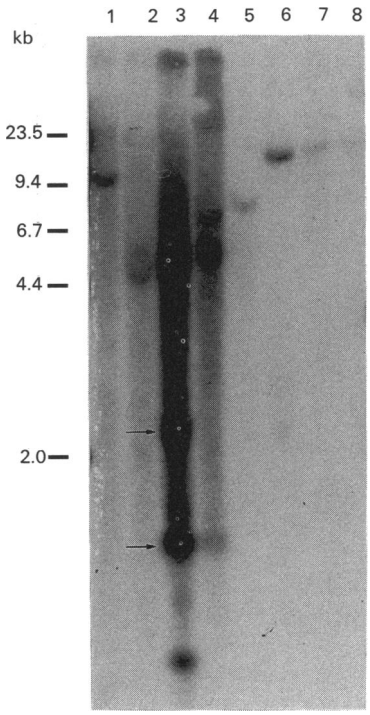

Fig. 5. Southern-blot analysis of rhodanese genes

A rat rhodanese cDNA probe was cross-hybridized to *Eco*RI-cut genomic DNA of various mammals and chicken. Lane 1, human; 1 and 2, monkey; lane 3, rat; lane 4, mouse; lane 5, dog; lane 6, cow; lane 7, rabbit; lane 8, chicken. The arrows indicate the positions of rodent repetitive sequence elements.

results of this Southern blot are shown in Fig. 5. The rat rhodanese probe hybridized strongly at moderate to stringent conditions to both rat and mouse genomic DNA sequences, as expected. In addition to the predominant cross-hybridizing *Eco*RI fragments (of about 5 kb), some smaller highly repetitive

elements were also detected in rodents (by comparison with the ethidium bromide-stained gel used in this blot; results not shown). The rhodanese genes of the remaining mammals and chicken were also detected with the rat rhodanese probe, but hybridization to the rhodanese gene(s) of these species was less intense, indicating that these genes are less complementary to the rodent-specific probe. Apparently, only one genomic *Eco*RI fragment is recognized by the probe in most species, suggesting that only one diploid copy of this gene is present in most mammals, and that the cDNA representing the open reading frame is contained within a relatively small piece of genomic DNA (4.5–20 kb) in most mammals.

The degree of cross-hybridization of a cloned gene probe from one species to the same gene(s) from other species has proven useful in studying phylogenetic relationships between species. The extent of cross-hybridization of a rodent rhodanese probe to other mammals (especially cow and man) indicates that, at the DNA sequence level, rhodanese genes exhibit a high degree of DNA (and thus amino acid) sequence similarity. The rat rhodanese probe was least complementary to the rhodanese genes of chicken, as expected for a non-mammal, and to rabbit, suggesting a significant amount of molecular divergence between rat and rabbit rhodanese genes. Molecular phylogenetic cladograms have been generated for a large number of mammals by comparison of amino acid residue or nucleotide changes within the derived sequences of certain genes, such as myoglobin and α-crystallin A-chain (McKenna, 1987). In general, interspecies comparisons of distantly related mammalian species have demonstrated approx. 8–20 % amino acid sequence dissimilarities within essential genes (Kimura, 1983; Numa, 1986). The approx. 9 % (or less) dissimilarity at the amino acid sequence level between rat and cow rhodanese is unremarkable and consistent with previous reports comparing the open reading frames of other genes between distantly related mammals, such as cattle and rodents. Even though some molecular divergence has occurred within the rhodanese genes, we can conclude that a high degree of selective pressure (in population dynamic terms) has been applied to the maintenance of critical amino acid residues involved in the structural and catalytic properties of rat rhodanese.

We are very grateful to R. L. Heinrikson and I. M. Reardon (Upjohn) for stimulating our interest in cloning rhodanese cDNAs, to P. Horowitz (University of Texas, San Antonio, TX, U.S.A.) for a mutual exchange of unpublished information about rhodanese cDNA sequences, to M. W. Schulz and J. R. Blinn (Upjohn) for technical assistance on computer-generated structural modelling, to R. Drong, D. Siemieniak and N. Hatzenbuhler (Upjohn) for technical advice, and to C. Falany (University of Rochester, Rochester, NY, U.S.A.) and G. Johnson (Upjohn) for helpful discussions regarding sulphurtransferase-related enzymes.

REFERENCES

Cerletti, P. (1986) Trends Biochem. Sci. **11**, 369–372
Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978) Atlas of Protein Sequence and Structure, vol. 5, suppl. 3, chapter 22, National Biochemical Research Foundation, Silver Springs
Hurt, E. C. & van Loon, A. P. G. M. (1986) Trends Biochem. Sci. **11**, 204–211
Kimura, M. (1983) in Evolution of Genes and Proteins (Nei, M. & Koehn, R., eds.), pp. 208–233, Sinauer Associates Publishers, Sunderland
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor
McKenna, M. C. (1987) in Molecules and Morphology in Evolution: Conflict or Compromise? (Patterson, C., ed.), Cambridge University Press, Cambridge
Numa, S. (1986) Chem. Scr. **26B**, 173–178
Ploegman, J. H., Drent, G., Kalk, K. H., Hol, W. G., Heinrikson, R. L., Keim, P., Weng, L. & Russell, J. (1978) Nature (London) **273**, 124–129
Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Russell, J., Weng, L., Keim, P. S. & Heinrikson, R. L. (1978) J. Biol. Chem. **253**, 8102–8108
Westley, J. (1973) Adv. Enzymol. Relat. Areas Mol. Biol. **39**, 327–368
Westley, J. (1981) Methods Enzymol. **77**, 285–291
Yanisch-Perron, C., Viera, J. & Messing, J. (1985) Gene **33**, 103–119

Received 9 July 1990/15 October 1990; accepted 23 October 1990