An approach to optimizing the active site in a glutathione transferase by evolution *in vitro*

Lars O. HANSSON, Mikael WIDERSTEN and Bengt MANNERVIK¹

Department of Biochemistry, Uppsala University, Biomedical Center, Box 576, S-751 23 Uppsala, Sweden

A glutathione transferase (GST) mutant with four active-site substitutions (Phe¹⁰ → Pro/Ala¹² → Trp/Leu¹⁰⁷ → Phe/Leu¹⁰⁸ → Arg) (C36) was isolated from a library of active-site mutants of human GST A1-1 by the combination of phage display and mechanism-based affinity adsorption [Hansson, Widersten and Mannervik (1997) *Biochemistry* **36**, 11252–11260]. C36 was selected on the basis of its affinity for the transition-state analogue 1-(S-glutathionyl)-2,4,6-trinitrocyclohexadienyl. C36 affords a 10⁵-fold rate enhancement over the uncatalysed reaction between reduced glutathione and 1-chloro-2,4-dinitrobenzene (CDNB), as evidenced by the ratio between k_{cat}/K_m and the second-order rate constant k_2 . The present study shows that C36 can evolve to an even higher catalytic efficiency by an additional site-specific mutation. Random mutations of the fifth active-site residue 208

allowed the identification of 18 variants, of which the mutant C36 Met²⁰⁸ → Cys proved to be the most active form. The altered activity was substrate selective such that the catalytic efficiency with CDNB and with 1-chloro-6-trifluoromethyl-2,4-dinitrobenzene were increased 2–3-fold, whereas the activity with ethacrynic acid was decreased by a factor of 8. The results show that a single-point mutation in the active site of an enzyme may modulate the catalytic activity without being directly involved as a functional group in the enzymic mechanism. Such limited modifications are relevant both to the natural evolution and the *in vitro* redesign of proteins for novel functions.

Key words: activity screening, active-site mutation, glutathione transferase, mutant library, random mutagenesis.

INTRODUCTION

A glutathione transferase (GST) mutant with four active-site substitutions (Phe¹⁰ → Pro/Ala¹² → Trp/Leu¹⁰⁷ → Phe/Leu¹⁰⁸ → Arg; residues numbered according to [1]) (C36) was isolated in a previous phage-display selection [2]. The mutant gave a rate acceleration in a nucleophilic aromatic substitution (S_NAr) reaction by a factor of 10⁵ compared with the uncatalysed reaction, but was approximately 20 times less efficient than the wild-type GST A1-1 enzyme. With the aim of improving the S_NAr activity of C36, a fifth active-site residue was altered, and 18 mutant variants were screened for improved activity with reduced glutathione (GSH; γ -Glu-Cys-Gly) and 1-chloro-2,4-dinitrobenzene (CDNB). The question was raised whether it is possible to evolve the catalytic function of C36 by further complementary mutations, or if back mutations to the wild-type residues would be the only alternative to recover full catalytic efficiency.

The family of GSTs (EC 2.5.1.18) are detoxication enzymes that catalyse the conjugation of GSH with a large number of substrates with various functional groups, such as hydroxy-alkenals [3], α,β -unsaturated aldehydes [4], *o*-quinones [5] and chloronitrobenzenes [6]. The GSTs have been divided into classes on the basis of sequence similarity [7]. The Alpha class is one of them, and the structure of human GST A1-1 has been elucidated by X-ray crystallography [1]. Each monomer in the dimeric enzyme contains two distinct binding pockets, the G-site and the H-site. The G-site binds GSH with high specificity, while the H-site interacts with the second, hydrophobic substrate in a more unspecific manner. Some 13 residues contribute to the H-site and are situated in three regions of the primary structure, an N-terminal loop, the α 4-helix and the C-terminal α 9-helix (Figure

1). The C-terminal α 9-helix forms a 'lid' over the active site of the Alpha class GST.

The four mutated positions in C36 (residues 10, 12, 107 and 108) are located in the N-terminal loop and the α 4-helix of the H-site of GST A1-1. Met²⁰⁸, the fifth residue, which was targeted for random mutagenesis, is positioned adjacent to the C-terminal helix in the third, unaltered, region of the H-site in C36. This side chain also is closest in proximity to S-benzylglutathione in the three-dimensional structure of GST A1-1 [1], in comparison with other side-chains in the H-site. A number of substitutions of residue 208 have been analysed in a previous study of the wild-type enzyme [9], but the presence of four other substitutions in the binding pocket may alter the relative effect on kinetic properties in S_NAr reactions of Met²⁰⁸ mutants compared with the wild-type enzyme. The present study of multiple mutants of GST A1-1 shows that C36 can be catalytically improved by mutagenesis of an additional active-site residue.

MATERIALS AND METHODS

Library construction

A primer designed to randomize residue 208, Met208X2: 5'-CCG TCT AGA TTA AAA TCG GAA GAT TTT ACG TGC TTC TTC TAA AGA TTT CTC ATC MNN GGG AGG CTT CCT (where M = A/C and N = G/A/T/C), and M13 Forward primer were used in PCR mutagenesis in order to construct the Met²⁰⁸ → X library with the cDNA encoding C36 (pGΔETacA1:Phe¹⁰ → Pro/Ala¹² → Trp/Leu¹⁰⁷ → Phe/Leu¹⁰⁸ → Arg; [2]) as template. The reaction mixture (100 μ l) contained 1 μ M of each primer, 10 mM Tris/HCl (pH 8.8), 50 mM KCl,

Abbreviations used: GST, glutathione transferase; C36, GST mutant with four active-site substitutions (Phe¹⁰ → Pro/Ala¹² → Trp/Leu¹⁰⁷ → Phe/Leu¹⁰⁸ → Arg); CDNB, 1-chloro-2,4-dinitrobenzene; oCF₃CDNB, 1-chloro-6-trifluoromethyl-2,4-dinitrobenzene; GSH, reduced glutathione; K_p , formation constant; S_NAr , nucleophilic aromatic substitution; TNB, 1,3,5-trinitrobenzene.

¹ To whom correspondence should be addressed (e-mail Bengt.Mannervik@biokem.uu.se).

Figure 1 Representation of S-benzylglutathione bound to the hydrophobic binding pocket (H-site) of the active site of wild-type GST A1-1 [1]

The four residues altered in mutant C36 (Phe¹⁰, Ala¹², Leu¹⁰⁷ and Leu¹⁰⁸), which are situated in the N-terminal loop and α 4-helix, are shown as well as the Met²⁰⁸ residue preceding the C-terminal α 9-helix. The image was made with the program Molscript [8].

1.5 mM MgCl₂, 0.2 mM of each dNTP, 2 units of *Taq* DNA polymerase (MBI Fermentas, Lithuania) and approx. 1 ng of template. Forty PCR cycles were run (1 min at 95 °C; 1 min at 55 °C; 2 min at 72 °C). The product was digested with *Eco*RI and *Xho*I, separated on agarose gel, and the cDNA fragment was ligated into a pGAETac5His vector [10] that had been digested with *Eco*RI and *Spe*I and purified on agarose gel. Electrocompetent *Escherichia coli* XL-1 Blue cells (Stratagene, La Jolla, CA, U.S.A.) were transformed with the ligation mixture by electroporation (Gene-Pulser; Bio-Rad, Richmond, CA, U.S.A.).

Screening for clones with high CDNB activity

Randomly chosen library clones were inoculated into 2 ml of 2TY medium [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl] supplemented with 50 μ g/ml ampicillin. Protein expression was induced with 0.3 mM β -D-thiogalactopyranoside when log phase was reached. Cultures were grown overnight and harvested by centrifugation at 1300 g. Pellets were resuspended in 300 μ l of PBS and lysed by ultrasonication. Cell debris was removed by centrifugation at 14000 g for 15 min. The enzymic activity of 70 cell lysates with GSH and CDNB was measured at A_{340} in a plate reader (Molecular Devices, Menlo Park, CA, U.S.A.). Each well contained 50 μ l of cell lysate, 1 mM GSH, 2 mM CDNB and 0.1 M sodium phosphate, pH 6.0, in a total volume of 300 μ l.

A total of 19 clones corresponding to all possible substitutions in residue 208 (except for Asn and Tyr) were expressed from 20 ml of 2TY medium cultures treated as described above, except that they were resuspended in 950 μ l of PBS supplemented with 0.02% NaN₃. The activity with CDNB was measured in duplicate for all 19 clones simultaneously in a microtitre plate reader [10 μ l of a 100-fold dilution of lysate (the Arg mutant lysate was used undiluted) in 0.1 M sodium phosphate, pH 6.5, 1 mM GSH, 1 mM CDNB and 0.1 M sodium phosphate, pH 6.5, in a total volume of 300 μ l at 25 °C]. The non-enzymic reaction rate was determined simultaneously ($n = 18$; the non-enzymic rate constituted 15% of the measured activity with diluted C36) and was subtracted from the enzymic rates. The total protein concentration of the lysates was determined [11]. The relative expression level of the mutant enzymes was determined by Western blot.

Equal volumes of crude lysate (approx. 50 μ g according to protein determination) from each mutant were subjected to SDS/PAGE and Western-blot analysis using affinity-purified rabbit anti-(human GST A1-1) [12], horseradish-peroxidase-linked anti-rabbit antibody (Amersham Life Science Ltd., Little Chalfont, Bucks., U.K.), and detection by ECL[®] (Amersham Life Science Ltd.).

Enzyme expression, purification and activity measurements

The entire coding regions of the Cys, His, Leu and Trp mutants of C36Met²⁰⁸ were subjected to DNA analysis [13] in order to rule out PCR artifacts. Cell lysates containing the variant enzymes were obtained essentially as described above, with the exception that a volume of 1.5 litres of bacterial culture was used and 20 mM sodium phosphate (buffer A), pH 7.0, was used instead of PBS. The cell lysate was desalted on Sephadex G-25 (Amersham Pharmacia Biotech, Uppsala, Sweden), equilibrated with buffer A, and the mutant enzyme was purified on carboxymethyl-Sepharose (Amersham Pharmacia Biotech). The column was washed with buffer A and eluted with a salt gradient (0–1 M NaCl) in buffer A. Pooled mutant fractions were concentrated by ultrafiltration and desalted on a PD-10 column (Amersham Pharmacia Biotech). The purity of the mutants was checked by SDS/PAGE with silver staining.

All spectrophotometric measurements were carried out at 30 °C in 0.1 M phosphate on either a Shimadzu UV-2501PC or a Varian 2290 spectrophotometer. The specific activities of the enzyme with CDNB [14], 1-chloro-6-trifluoromethyl-2,4-dinitrobenzene (oCF₃CDNB) [15], 4-nitrobenzyl chloride [14] or ethacrynic acid [14] as substrates were measured at pH 6.5; the activity with 1,2-dichloro-4-nitrobenzene [6] as substrate was determined at pH 8.0. In the determination of the kinetic parameters, k_{cat} , K_m and k_{cat}/K_m , the GSH concentration in the reaction mixture was kept at a saturating value of 5.0 mM, except with ethacrynic acid, when 1.0 mM GSH was used. The rates of the uncatalysed reactions were in agreement with those presented elsewhere [2,9,15]. For mutants C36, C36Met²⁰⁸ $\rightarrow$ Cys, C36Met²⁰⁸ $\rightarrow$ His and C36Met²⁰⁸ $\rightarrow$ Leu the molar absorption coefficient (at 280 nm) 30000 M⁻¹·cm⁻¹ was

used (determined for GST A1-1:Met²⁰⁸ → Trp; [9]), and ϵ 35300 M⁻¹·cm⁻¹ was used for C36Met²⁰⁸ → Trp (estimated by addition of the difference between GST A1-1:Met²⁰⁸ → Trp and the GST A1-1 wild-type enzyme).

The enzymic stabilization of the σ -complex between 1,3,5-trinitrobenzene (TNB) (kindly provided by Nobel Chemicals AB, Karlskoga, Sweden) and GSH was analysed by difference spectroscopy as described previously [16]. The GSH concentration was maintained at 5 mM in 0.1 M sodium phosphate, pH 6.5, and the TNB concentration was varied in the range 0.064–1.1 mM. The mutant sample (6–16 μ M) was included in the sample cuvette. For the C36Met²⁰⁸ → Cys mutant, the experiment was repeated under the same conditions, except that GSH was excluded from both cuvettes.

The possible alkylating effect of ethacrynic acid on the C36Met²⁰⁸ → Cys mutant, as well as on the parental mutant C36, was studied by analysing the time course of the specific activity with ethacrynic acid (0.2 mM ethacrynic acid, 0.25 mM GSH, 0.1 M sodium phosphate, pH 6.5, at 30 °C) by comparing the mutant preincubated with ethacrynic acid and the mutant preincubated without ethacrynic acid, which instead was added at the start of the activity measurement. The activity was measured, in triplicate, after incubations of 0 min, 70 min, 130 min, 280 min and 19 h.

The parameters k_{cat}/K_m and K_t (formation constant) were determined by fitting a hyperbola [$y = A \cdot [S]/(1 + B \cdot [S])$] to the experimental data. Enzyme-kinetics data were analysed with GraphPad Prism (GraphPad Software Inc., San Diego, CA, U.S.A.).

Molecular modelling

The structures of human GST A1-1 [1,17] were analysed with the InsightII package (Molecular Simulations, San Diego, CA, U.S.A.) on Silicon Graphics O2 or Indigo2 Extreme Workstations and Weblab Viewer (MSI) on Windows NT PC.

RESULTS

Screening of mutant library of enzyme randomized at residue 208

A library was constructed in which the active-site quadruple mutant C36 of human GST A1-1, previously derived from a phage-display selection [2], was randomized in codon 208. The PCR primer for introducing random mutations in codon 208 was designed to restrict the size of the library to 32 (codon randomized as NNG/T) in order to reduce the number of distinct codons, and hence limit the number of clones to be examined.

The enzymic activity with CDNB (see Figure 2A) was screened in crude cell lysates of 70 mutants randomly chosen from the library. The assay was carried out at pH 6.0 in order to decrease the non-enzymic reaction rate. At the first stage of analysis mutants exhibiting high activity in the lysate were selected for DNA sequence analysis, but only a limited number of Met²⁰⁸ mutants were identified (Ala, Asp, Cys, Gln, His, Ile, Leu, Met, Phe, Pro, Ser, Trp and Val). Clones with low CDNB activity, which could be due to poor expression of the mutant protein, were consequently analysed and six additional mutants were found (Arg, Glu, Gly, Lys, Thr and Stop). The Asn and Tyr substitutions were not identified, even though 47 of the 70 clones tested for activity were subjected to DNA sequence analysis.

Comparison of residue-208 variants of mutant C36

All identified residue-208 variants of mutant C36 (Asn and Tyr replacements not included) were expressed simultaneously at a larger scale and analysed more rigorously. Since several alterna-

Figure 2 Substrates and reactions catalysed by GST A1-1 variants

(A) S_NAr reaction of GSH and CDNB. (B) Formation of the equilibrium product of GSH and TNB. (C) Structural formulae of oCF₃CDNB, 1,2-dichloro-4-nitrobenzene, 4-nitrobenzyl chloride and ethacrynic acid.

Table 1 Codon 208 of C36 variants selected for comparison of activities with CDNB as substrate

Residue	Codon
Ala	GCT
Arg	CGG
Asp	GAT
Cys	TGT
Glu	GAG
Gln	CAG
Gly	GGG
His	CAT
Ile	ATT
Leu	TTG
Lys	AAG
Met	ATG
Phe	TTT
Pro	CCG
Ser	TCT
Thr	ACG
Trp	TGG
Val	GTG
Stop	TAG

tive codons had been identified for a number of amino acid substitutions, it was possible, for the replacements with Ala, Pro and Ser, to choose codons that are known to be preferred in highly expressed proteins [18]. Codons for residue-208 mutants selected for analysis are presented in Table 1. The relative activities with CDNB were determined with diluted cell lysates (Figure 3A), and the relative amounts of mutant protein were estimated from Western blots (Figure 3B). Both the original mutant C36 and the identical protein derived from random mutagenesis (C36Met²⁰⁸ → Met) were present, and therefore comparison allows an estimate of the variance in the level of

Figure 3 Semi-quantitative analysis of specific activity with CDNB in cell lysates containing C36 variants differing in the substitution of residue 208

(A) Enzymic activity with GSH and CDNB in lysates containing Met²⁰⁸ mutants of C36. Values were normalized to the volume of cell lysate and given as means $\pm$ S.E.M. ($n = 2$). (B) Western blot of equal volumes of lysate containing Met²⁰⁸ mutants of C36. The residues are shown by single-letter code. The gels in the two panels are not fully comparable because of a slight difference in exposure time.

enzyme expression. Of the C36 variants (wild-type Met²⁰⁸ included) analysed, 18 out of 19 gave rise to an enzyme with a complete polypeptide chain, as ascertained from the Western blot (Figure 3B). Only the truncated sequence, with a stop codon at position 208, did not express a detectable amount of protein and was probably degraded within the cell. The enzyme activities

in bacterial lysates of all mutants remained constant over at least 1.5 h, showing that significant differences in stability were not evident in the variants studied.

The CDNB activity of the Cys, Met and Phe mutants was notably higher than that of mutants with other substitutions. The CDNB activity of the Met²⁰⁸ $\rightarrow$ Cys mutant was clearly highest, and this mutant was expressed at a moderate level. All mutants were expressed, judging from the Western blot (Figure 3B), and migrated an equal distance, indicating identical length of the polypeptide chain. The Trp and Arg mutants appeared to be least well expressed. The Arg mutant seemed to be a poor catalyst, since it did not exhibit detectable activity with CDNB, whereas the Trp mutant displayed moderate activity and therefore could be assumed to have a reasonably high specific activity with CDNB. The Met²⁰⁸ $\rightarrow$ Glu mutant was expressed at the highest level of all, but showed low activity with CDNB. In fact, all mutants with charged-residue substitutions displayed low catalytic activity, indicating that a charge in this part of the active site is unfavourable. Since the Cys and Trp variants exhibited the highest specific activity with CDNB, these mutants were selected for further characterization; the His and Leu variants were also analysed.

Catalytic properties of mutants

The purified C36Met²⁰⁸ $\rightarrow$ Cys mutant showed a specific activity with CDNB which was almost twice that of the parental four-residue mutant C36 of GST A1-1 (Table 2), and which fits well with the result from the preliminary lysate studies (Figures 3A and 3B). This mutant displayed a two-fold enhancement in catalytic efficiency (k_{cat}/K_m) with CDNB when compared with C36, and an even greater effect was seen with oCF₃CDNB (Figures 2A and 2C; Table 3). The effects on these kinetic parameters in the Trp mutant were less marked with the same substrates in S_NAr reactions. The K_m with CDNB was larger for the Trp variant than for the Cys mutant. Both mutants showed only a minor increase in specific activity compared with C36 (Table 2) when 1,2-dichloro-4-nitrobenzene (Figure 2C) was used as substrate. The His and Leu variants were catalytically less efficient with CDNB when compared with the parental C36 (Table 2).

The reaction mechanism of the S_NAr is believed to include a σ -complex as an intermediate, and the formation of the σ -complex of GSH and TNB (Figure 2B) was studied as a measure of the ability to afford stabilization in the active site [16]. The Cys

Table 2 Specific activities of active-site mutants of GST A1-1 with different substrates

Abbreviation: nd, not determined. S.E.M. was less than 12% for each of the specific activities given ($n \geq 5$).

Enzyme	Substrate...	Specific activity ($\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$)			
		CDNB	1,2-Dichloro-4-nitrobenzene	4-Nitrobenzyl chloride	Ethacrynic acid
C36M208C		19	0.007	0.050	0.22
C36M208H		7.4	nd	nd	2.2
C36M208L		9.1	nd	nd	1.6
C36M208W		15	0.01	0.097	1.5
C36		10*	0.006*	0.12	1.8*
GST A1-1		80†	0.2†	0.3‡	0.24*

* Derived from [2].

† Derived from [25].

‡ Derived from [9].

Table 3 Kinetic parameters of active-site mutants of GST A1-1Values $\pm$ S.E.M. were obtained by non-linear regression analysis. EA, ethacrynic acid.

Enzyme	Substrate	k_{cat} (s^{-1})	K_{m} (mM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \cdot \text{mM}^{-1}$)	$k_{\text{cat}}/K_{\text{m}}$ ratio (CDNB/EA)
C36M208C	CDNB	24 ± 2	1.6 ± 0.3	14 ± 0.6	2.1
	oCF_3CDNB	27 ± 5	1.0 ± 0.3	26 ± 2	
	EA	0.51 ± 0.04	0.079 ± 0.02	6.8 ± 0.4	
C36M208H	CDNB	— [‡]	—	3.6 ± 0.7	0.095
	EA	0.79 ± 0.05	0.021 ± 0.005	38 ± 8	
C36M208L	CDNB	16 ± 5	3.8 ± 1.5	4.2 ± 0.5	0.14
	EA	0.85 ± 0.03	0.027 ± 0.004	31 ± 4	
C36M208W	CDNB	26 ± 6	3.4 ± 1.2	7.0 ± 0.4	0.18
	oCF_3CDNB	—	—	20 ± 0.5	
	EA	0.76 ± 0.03	0.020 ± 0.003	38 ± 3	
C36*	CDNB	—	—	6.7 ± 0.2	0.12
	oCF_3CDNB	9.8 ± 1.9	1.1 ± 0.3	9.1 ± 3.1	
	EA	0.83 ± 0.04	0.016 ± 0.004	53 ± 12	
GST A1-1	CDNB [†]	88 ± 3	0.56 ± 0.04	160 ± 8	41
	$\text{oCF}_3\text{CDNB}^{\dagger}$	69 ± 3	0.084 ± 0.008	820 ± 60	
	EA*	0.19 ± 0.02	0.048 ± 0.016	3.9 ± 1.4	

* Derived from [2].

[†] Derived from [15].[‡] Separate values of k_{cat} and K_{m} could not be determined because of the low degree of saturation with the substrate.**Table 4** K_{i} and absorption coefficient (ϵ_{450}) of the enzyme-bound GSH and TNB σ -complexValues $\pm$ S.D. were obtained by non-linear regression analysis.

Enzyme	K_{i} (mM^{-1})	ϵ_{450} ($\text{mM}^{-1} \cdot \text{cm}^{-1}$)
C36M208C	5.4 ± 0.5	8.8 ± 0.3
C36M208W	1.5 ± 0.1	12 ± 0.6
C36*	0.41 ± 0.11	7.1 ± 0.8
GST A1-1 [†]	8.9 ± 1.4	12 ± 0.4

* Derived from [2].

[†] Derived from [15].

mutant provided a 10-fold higher formation constant (K_{i}) than did the parental C36 mutant, and the Trp variant displayed an intermediate value. The K_{i} value of the wild-type enzyme GST A1-1 was only twice as high as that of the five-residue substituted mutant C36Met²⁰⁸ $\rightarrow$ Cys (Table 4). From experiments conducted in the absence of GSH it was shown that the thiol group of the Cys²⁰⁸ residue in the active site did not, by itself, form a σ -complex with TNB.

The four-residue mutant C36, when compared with the wild-type GST A1-1, was found to exhibit a 30-fold improvement in catalytic efficiency with ethacrynic acid as substrate [2] (Figure 2C). The His, Leu and Trp variants of C36Met²⁰⁸ showed a similarly high efficiency, whereas the Cys mutant displayed a notably lower activity. The decrease was brought about by both a decrease in k_{cat} and an increase in K_{m} . Since the Cys²⁰⁸ sulphydryl group could possibly be alkylated and the enzyme thereby inactivated, the activity with ethacrynic acid as substrate

was analysed for various times of incubation with the alkylating substrate. The loss of activity in the Met²⁰⁸ $\rightarrow$ Cys mutant was not significantly different from that of C36 (results not shown). Thus the lower activity with ethacrynic acid cannot be explained by alkylation of residue 208.

Both the Cys and Trp variants of mutant C36 exhibited decreased activity in the substitution reaction with 4-nitrobenzyl chloride under standard conditions (Figure 2C; Table 2), in contrast with the improved activity seen with $\text{S}_{\text{N}}\text{Ar}$ substrates. However, the reaction with 4-nitrobenzyl chloride occurs by a different reaction mechanism, an $\text{S}_{\text{N}}2$ mechanism, which does not involve a σ -complex as an intermediate.

DISCUSSION

Random mutagenesis of an active-site residue

Mutations of an H-site residue in C36 were shown to result in improvements in catalytic activity. The Met²⁰⁸ $\rightarrow$ Cys mutant displayed two- and three-fold increased catalytic efficiency with the $\text{S}_{\text{N}}\text{Ar}$ substrates CDNB and oCF_3CDNB respectively (Table 3). The σ -complex stabilization was also enhanced ten-fold compared with the parental enzyme C36 (Table 4). The H-site of C36 apparently offers a potential for the evolution of the enzyme function by additional point mutations.

Even though more than 80% of the analysed mutants were considered to be less active with CDNB than C36, this finding does not necessarily imply that these mutants are inferior with other substrates. The Cys mutant, which was superior with CDNB, gave lower activity with ethacrynic acid, whereas the His and Leu mutants, that had decreased activity with CDNB, retained the high activity with ethacrynic acid and was four-fold more efficient than the Cys variant. Table 3 shows that there is a > 20 -fold difference in preference for CDNB over ethacrynic acid among the analysed point mutants, as measured by the

k_{cat}/K_m ratios. Wild-type GST A1-1 was less active with ethacrynic acid than any of the mutants shown in Table 3, and had a corresponding k_{cat}/K_m ratio which was > 200-fold higher than those of other mutants, except for the Cys mutant. Also, the Cys and Trp mutants, that showed improved activity with $S_N\text{Ar}$ substrates, were less active with the S_N2 -reaction substrate 4-nitrobenzyl chloride. C36 thus allows a high variability in side chains in the active-site position 208, while retaining the ability to fold properly and also retaining at least a fraction of the enzymic activity with CDNB.

Role of H-site residue 208

Residue 208 contributes to the H-site of GST A1-1 and is situated at the link between the $\alpha 8$ -helix and the C-terminal $\alpha 9$ -helix (Figure 1). The C-terminal helix has been described as being mobile and not visible in the three-dimensional structure of the GST A1-1 apoenzyme [17], and it has been proposed to contribute to a 'soft' active site in high-pressure studies of the enzyme [19]. The degree of mobility of the $\alpha 9$ -helix may be modulated by residue 208, since at this residue the polypeptide backbone serves as a hinge to the helix. The results of recent studies suggest that the observed activation of the partially inactivated GST A1-1 Met²⁰⁸ → Lys mutant by *p*-propylbenzoic acid, and related negatively charged molecules, is caused by restoration of the hydrophobic milieu of the H-site and stabilization of the helix [20]. In relation to the binding of electrophilic substrates to the active site, Met²⁰⁸ is found in proximity to the position expected for the *para*-substituent of $S_N\text{Ar}$ substrates (Figures 1 and 2).

The replacement of Met²⁰⁸ in wild-type GST A1-1 with various amino acid residues has been analysed previously [9], but substitution with a Cys residue was not identified. The replacement of Met²⁰⁸ by Leu and His in C36 caused a decrease in $k_{\text{cat}}/K_m^{\text{CDNB}}$ (63 % and 54 % respectively), which is very similar to those reported for the corresponding substitutions in the wild-type enzyme, suggesting an additive effect for the Met²⁰⁸ substitutions. The Met²⁰⁸ → Trp mutation in wild-type GST A1-1 was the only substitution studied that showed an increase in $k_{\text{cat}}/K_m^{\text{CDNB}}$, and this mutation has also been found to have improved activity with oCF₃CDNB [15]. In both cases the improvement was due mainly to a decrease in K_m (three-fold). The corresponding substitution in C36 was not advantageous with CDNB, but did increase the catalytic efficiency with oCF₃CDNB two-fold. The Met²⁰⁸ → Trp replacement thus appears not to cause simple additive effects.

The Met²⁰⁸ → Cys mutant was the most efficient enzyme variant with both CDNB and oCF₃CDNB substrates, while the structurally similar Met²⁰⁸ → Ser mutant appeared to be catalytically poor. From the activity and Western-blot data in Figure 3 it can be estimated that the specific activity of the Ser mutant was less than one-fifth of that of the Cys mutant with CDNB as substrate.

Mechanistic implications

The present study focuses on the catalysis of $S_N\text{Ar}$, and the reaction of GSH and CDNB in particular. The $S_N\text{Ar}$ reaction mechanism has been postulated to involve a σ -complex intermediate (Meisenheimer complex; see Figure 2) close to the transition state(s) on the reaction co-ordinate. The equilibrium product of GSH and TNB is a σ -complex which has been used in several studies to probe the enzymic stabilization of $S_N\text{Ar}$ intermediates. The K_t of this complex describes to what extent the enzyme promotes σ -complex formation with GSH and TNB bound in the active site [16]. The development of the σ -complex

has been proposed to be the rate-limiting step in the uncatalysed reaction [21].

The active-site interactions that stabilize the σ -complex of GSH and TNB can be expected to contribute to the stabilization of intermediates formed in the reactions with CDNB and oCF₃CDNB. It has thus been suggested that the K_m for $S_N\text{Ar}$ substrates should be inversely correlated to the K_t of the GSH–TNB complex [15]. This relation seems to hold for CDNB with wild-type GST A1-1 and the Met²⁰⁸ → Cys and Met²⁰⁸ → Trp variants of C36. The Cys mutant has a three-fold higher K_t for the σ -complex when compared with the Trp mutant, but only half the K_t of the wild-type enzyme, which fits the inverse order of K_m^{CDNB} (Tables 3 and 4). However, in the reaction with oCF₃CDNB, the K_m value appeared to be unchanged upon the Met²⁰⁸ → Cys substitution in C36, which makes a ten-fold difference in K_t . The difference may be related to the higher reactivity of oCF₃CDNB, which may alter the rate-limiting step of the enzyme-catalysed reaction. The third substituent of oCF₃CDNB, a trifluoromethyl group, may introduce interactions in the framework of C36 that differ from those in the wild-type enzyme.

In the original identification of the four-residue mutant, C36, it was found that the lower catalytic efficiency, k_{cat}/K_m , for a number of $S_N\text{Ar}$ substrates was correlated to the decrease in K_t when compared with the GST A1-1 wild-type [2]. The k_{cat}/K_m parameter describes steps up to, and including, the first irreversible step in the mechanism [22]. The Met²⁰⁸ → Cys mutant used in the present study displayed a ten-fold enhancement in promotion of GSH–TNB complex formation, but the catalytic efficiencies with CDNB and oCF₃CDNB were increased two- and three-fold respectively. In the reaction catalysed by C36, the σ -complex K_t appeared to limit k_{cat}/K_m , but the results using the five-residue mutants indicate that there are additional mechanistic steps which markedly influence the catalytic rate.

It is worth noticing that the four variant residues providing the highest CDNB activity (Cys, Met, Trp and Phe) all contain polarizable groups. It can thus be speculated that the negative potential in the σ -complex developing during the $S_N\text{Ar}$ reaction might be balanced by an induced dipole of a polarized residue in the H-site.

The four substitutions in C36 contributed to a greater than ten-fold increase in catalytic efficiency with ethacrynic acid compared with the wild-type GST A1-1 [2]. With substitutions His, Leu and Trp in position 208 of C36 high activity was retained in this Michael addition. The stereochemical requirements for the addition reaction are different from those of the $S_N\text{Ar}$ mechanism, and thus it appears that residue 208 is not orientated so that it has a major influence on the ethacrynic acid reaction. Also the Michael addition of GSH to 4-nitrocinnamaldehyde has been found to be independent of the substitutions of Met²⁰⁸ to Ala or Trp in the wild-type GST A1-1 [15]. In contrast, the Met²⁰⁸ → Cys substitution in C36 displayed a seven-fold reduced catalytic efficiency with ethacrynic acid.

Evolution of enzymes *in vitro*

A mutated enzyme with reduced catalytic activity can be functionally restored by additional residue replacements. Removal of a side chain can be compensated for by insertion of another side chain at a different position in the protein. For example, a catalytically poor point mutant of triose phosphate isomerase was restored to high activity by a distinct compensatory single-point mutation [23]. Improvement was achieved by different amino-acid substitutions in the active-site region. Another example of functional recovery is the conversion of the active site

of dihydrofolate reductase from *E. coli* into that of the homologous enzyme from *Lactobacillus casei* in order to analyse the differences in their active-site cavities [24]. The conversion was performed in two steps of mutagenesis. First, a triple mutant of the enzyme was produced, which showed reduced catalytic efficiency, and by means of two additional mutations (one being an insertion) full activity was essentially restored.

In the case of GSTs, which have a very broad spectrum of substrates [25], directed evolution has the potential of giving rise to alternative variants with distinctly different specificities [26]. Thus, the Met²⁰⁸ → Cys and Met²⁰⁸ → Trp variants of C36 both have increased catalytic efficiency with S_NAr substrates, but diverge in activity with the Michael addition substrate ethacrynic acid, which gives rise to a > 10-fold difference in preference for CDNB over ethacrynic acid (Table 3). The GST scaffold thus allows analysis in more than one dimension of 'activity space'.

New approaches have been developed to search the protein-sequence space for potential catalysts. *In vitro* evolution is accomplished by mimicking processes in natural molecular evolution. Random recombination of protein fragments of homologous enzymes can be achieved by shuffling of DNA encoding these enzymes [27,28]. DNA shuffling has been applied in the redesign of a number of enzymes [27–31]. Point mutagenesis at random sites is a second approach to the *in vitro* evolution of proteins, frequently effected by error-prone PCR [32]. In the present study, the template C36 was, to some extent, a product of similar combinatorial protein chemistry. The four original substitutions are situated in two distinct regions of the polypeptide chain that are combined in the active site. The additional mutations of Met²⁰⁸ occur in a third structural element of the active site and thereby allow exploration of a new contribution to the combinatorial process, with a result similar to that of DNA shuffling.

The present study showed that, although the majority of substitutions in an active-site residue caused reduction of catalytic activity, a few of the alternative amino acids improved the activity. Thus one single mutation was sufficient to increase the catalytic efficiency by 100–200%. This limited substitution can be compared with the four residues which differ in the H-sites of GSTs A1-1 and A2-2 (11 different residues in the complete enzymes) that confer an approximately 100-fold difference in isomerase activity with Δ^5 -androstendione (L. O. Hansson, M. Widersten and B. Mannervik, unpublished results; [25]).

Under normal conditions the natural evolution of proteins may proceed by small steps. It has been proposed by several authors that *in vitro* evolution should similarly be limited to a small number of mutation sites at a time in order to obtain improved proteins [33,34]. Favourable mutations can either be accumulated in repeated rounds of selection or site-specifically combined in a new mutant protein. As noted by Wells [35], the energetic advantages of several distinct mutations are additive as long as the residues are not in contact with each other. The mutagenesis of residue 208 in C36 lends support to this notion. However, for drastic changes in functional properties to occur, major structural alterations may be required.

It should also be noted that attaining proper stability of novel protein variants is a stringent boundary condition in the evolution of altered functions. The available degrees of freedom in substitutions, insertions and deletions of amino acids are therefore restricted. However, in the present case, substitutions of Met²⁰⁸ with alternative amino acids were not found to affect the stability of C36 or the wild-type enzyme GST A1-1.

In conclusion, the present study can be seen in the perspective of an evolutionary design of enzyme function. The first step was initiated in the library of active-site mutants [2] containing

variations in two of the three regions forming the H-site that governs the substrate-specificity profile of the enzyme. This mutant library represents the population of molecular species upon which evolution could operate. The selection of C36 through a mechanism-based procedure [2] led to the development of a lineage which, in the present investigation, was developed further into a new generation involving a third structural region that contributes to the active site. In this evolutionary process, the selection was based on the search for variants with high catalytic activity in S_NAr reactions. The results with ethacrynic acid show that optimization for Michael reactions would have led to other choices in the evolutionary path. In general terms, the two generations of GST mutants represent a combinatorial process involving distinct segments of primary structure and point mutations. Rearrangements of gene fragments in combination with point mutations presumably constitute the structural basis for natural molecular evolution, and the results illustrate how they can be exploited for the redesign of protein functions *in vitro*. The difference between the natural process and protein engineering, however, is that the latter allows a rational choice of regions to be mutated, which presumably is more efficient than the stochastic processes operating in Nature.

This investigation was supported by the Swedish Research Council for Engineering Sciences.

REFERENCES

- Sinning, I., Kleywegt, G. J., Cowan, S. W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G. L., Armstrong, R. N., Ji, X., Board, P. G. et al. (1993) *J. Mol. Biol.* **232**, 192–212
- Hansson, L. O., Widersten, M. and Mannervik, B. (1997) *Biochemistry* **36**, 11252–11260
- Hubatsch, I., Ridderström, M. and Mannervik, B. (1998) *Biochem. J.* **330**, 175–179
- Berhane, K., Widersten, M., Engström, Å., Kozarich, J. W. and Mannervik, B. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1480–1484
- Baez, S., Segura-Aguilar, J., Widersten, M., Johansson, A.-S. and Mannervik, B. (1997) *Biochem. J.* **324**, 25–28
- Booth, J., Boyland, E. and Sims, P. (1961) *Biochem. J.* **79**, 516–524
- Mannervik, B., Ålin, P., Guthenberg, C., Jonsson, H., Tahir, M. K., Warholm, M. and Jörnvall, H. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7202–7206
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950
- Widersten, M., Björnstedt, R. and Mannervik, B. (1994) *Biochemistry* **33**, 11717–11723
- Widersten, M. (1998) *Protein Expression Purif.* **13**, 389–395
- Bradford, M. A. (1976) *Anal. Biochem.* **72**, 248–254
- Hao, X.-Y., Castro, V. M., Bergh, J., Sundström, B. and Mannervik, B. (1994) *Biochim. Biophys. Acta* **1225**, 223–230
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Habig, W. H., Pabst, M. J. and Jakoby, W. B. (1974) *J. Biol. Chem.* **249**, 7130–7139
- Widersten, M., Björnstedt, R. and Mannervik, B. (1996) *Biochemistry* **35**, 7731–7742
- Graminski, G. F., Zhang, P., Sesay, M. A., Ammon, H. L. and Armstrong, R. N. (1989) *Biochemistry* **28**, 6252–6258
- Cameron, A. D., Sinning, I., L'Hermite, G., Olin, B., Board, P. G., Mannervik, B. and Jones, T. A. (1995) *Structure* **3**, 717–727
- Bulmer, M. (1988) *J. Evol. Biol.* **1**, 15–26
- Atkins, W. M., Dietze, E. C. and Ibarra, C. (1997) *Protein Sci.* **6**, 873–881
- Gustafsson, A. and Mannervik, B. (1999) *J. Mol. Biol.* **288**, 787–800
- Ingold, C. K. (1953) *Structure and Mechanism in Organic Chemistry*, pp. 797–815, G. Bell and Sons Ltd., London
- Northrop, D. B. (1998) *J. Chem. Educ.* **75**, 1153–1157
- Hermes, J. D., Blacklow, S. C. and Knowles, J. R. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 696–700
- Li, L. and Benkovic, S. J. (1991) *Biochemistry* **30**, 1470–1478
- Mannervik, B. and Widersten, M. (1995) in *Advances in Drug Metabolism in Man* (Pacifci, G. M. and Fracchia, G. N., eds.), pp. 407–459, European Commission, Luxembourg

- 26 Mannervik, B., Björnstedt, R., Davey, E., Emahazion, T., Fernandez, E., Hansson, L. O., Kolm, R. H., Nilsson, L. O., Olin, B., Stenberg, G. et al. (1996) in *Glutathione S-Transferases. Structure, Function and Clinical Implications* (Vermeulen, N. P. E., Mulder, G. J., Nieuwenhuyse, H., Peters, W. H. M. and van Bladeren, P. J., eds.), pp. 3–11, Taylor and Francis Ltd., London
- 27 Cramer, A., Raillard, S.-A., Bermudez, E. and Stemmer, W. P. C. (1998) *Nature* (London) **391**, 288–291
- 28 Hansson, L. O., Bolton-Grob, R., Massoud, T. and Mannervik, B. (1999) *J. Mol. Biol.* **287**, 265–276
- 29 Stemmer, W. P. C. (1994) *Nature* (London) **370**, 389–391
- 30 Stemmer, W. P. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10747–10751
- 31 Yano, T., Oue, S. and Kagamiyama, H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5511–5515
- 32 Chen, K. and Arnold, F. H. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5618–5622
- 33 Kuchner, O. and Arnold, F. H. (1997) *Trends Biotechnol.* **15**, 523–530
- 34 Arkin, M. R. and Wells, J. A. (1998) *J. Mol. Biol.* **284**, 1083–1094
- 35 Wells, J. A. (1990) *Biochemistry* **29**, 8509–8517

Received 19 April 1999/25 June 1999; accepted 3 September 1999