Cys-X Scanning for Expansion of Active-site Residues and Modulation of Catalytic Functions in a Glutathione Transferase*

Malena A. Norrgård‡, Ulf Hellman§ and Bengt Mannervik¶†

From the ‡Department of Biochemistry and Organic Chemistry, the §Ludwig Institute for Cancer Research, Biomedical Center, Uppsala University, SE-75123 Uppsala, and the ¶Department of Neurochemistry, Stockholm University, SE-10691 Stockholm, Sweden

Running head: Cys-X scanning for enzyme redesign

* This work was supported by the Swedish Research Council

To whom correspondence should be addressed: Bengt Mannervik, Department of Biochemistry and Organic Chemistry, Uppsala University, Biomedical Center, Box 576, SE-75123, Uppsala, Sweden.
Fax: +46 18 558431; E-mail: Bengt.Mannervik@biorg.uu.se

We propose Cys-X scanning as a semi-synthetic approach to engineer the functional properties of recombinant proteins. As in the case of Ala scanning, key residues in the primary structure are identified and one of them replaced by Cys via site-directed mutagenesis. The thiol of the residue introduced is subsequently modified by alternative chemical reagents to yield diverse Cys-X mutants of the protein. This chemical approach is orthogonal to Ala or Cys-scanning and allows the expansion of the repertoire of amino acid side chains far beyond those present in natural proteins. In its present application, we have introduced Cys-X residues in human glutathione transferase (GST) M2-2, replacing Met212 in the substrate-binding site. In order to achieve selectivity of the modifications, the Cys residues in the wild-type enzyme were replaced by Ala. A suite of simple substitutions resulted in a set of homologous Met derivatives ranging from normethionine to S-heptyl-cysteine. The chemical modifications were validated by HPLC and mass spectrometry. The derivatized mutant enzymes were assayed with alternative GST substrates representing diverse chemical reactions: aromatic substitution, epoxide opening, transnitrosylation, and addition to an ortho-quinone. The Cys substitutions had different effects on the alternative substrates and differentially enhanced or suppressed catalytic activities depending on both the Cys-X substitution and the substrate assayed. As a consequence, the enzyme specificity profile could be changed among the alternative substrates. The procedure lends itself to large-scale production of Cys-X modified protein variants.

The 20-odd proteogenic amino acids provide combinations in protein structures that bring forth a perplexing variety of functions. Nevertheless, it is a reasonable assumption that expansion of the repertoire of amino acid residues in a protein could result in changes of functional properties that cannot be accomplished with the natural set of building blocks. In biological systems the limited number of amino acids that can be assembled by ribosomal protein synthesis is complemented by a wide variety of post-translational modifications catalyzed by specific enzymes (1). Such modifications are nonetheless subject to the restrictions set by specific recognition motifs in the target proteins as well as by the selectivities of the modifying enzymes. On the other hand, recombinant proteins have the distinct advantage that they can be subjected to chemical modifications \textit{ex vivo} and may, under appropriate circumstances, undergo numerous chemical modifications beyond the scope of biological modifying agents. In general, Cys residues in proteins can be selectively modified with an extremely broad range of electrophilic chemical reagents, and the cysteine residues can be introduced in strategic positions by means of site-directed mutagenesis.

Ala-scanning is a well established method, by which target residues in a peptide sequence are substituted by Ala (2). A similar Cys-scanning has been employed for analytical studies of the topology of membrane proteins (3, 4). In both cases the aim is to identify functionally important positions in the primary structure of a polypeptide chain. We have chosen a structurally orthogonal approach in which a Cys residue in a targeted position is substituted by a variety of alternative chemical groups. Our assumption was that favorable S-substitutions would be a means to chemically...
tune both catalytic efficiency and substrate selectivity of an enzyme, such as a promiscuous glutathione transferase (GST). We designate this orthogonal method “Cys-X scanning” and demonstrate its usefulness with a human GST (hGST M2-2). GSTs catalyze the conjugation of glutathione with a broad range of alternative electrophiles and are structurally and functionally well characterized (5, 6). hGST M2-2 is well suited as an example of the potential of Cys-X scanning since it has widely different catalytic efficiencies with alternative electrophilic substrates, which are undergoing divergent types of chemical reactions (7, 8). We demonstrate that different substitutions of Cys in the substrate-binding site in some cases actually enhance the catalytic activity with a particular substrate. In addition, the Cys-X modifications differentially change the relative rates by which alternative substrates are conjugated.

EXPERIMENTAL PROCEDURES

Cloning and purification — The human GST M2-2 wild type enzyme optimized by silent mutations for high-level expression (9) was used as template for site-directed mutagenesis. The gene was cloned into the plasmid vector pGΔEtac (10) by using the SalI and EcoRI restriction sites. The three Cys residues in hGST M2-2 (C87, C115, C174; designating initiator Met as number 1) were replaced with Ala through standard long-chain PCR in a sequential manner using Phusion Hot Start (Finnzymes, Espoo, Finland). The following primers were used (read in the 5’ to 3’ direction, mutations underlined): C87A-forward GCACAACCTGGCGGGGAATC and C87A-reverse TTGCGGGCAATGTACCGCAG; C115A-forward GCCCAAACTCggGGGGAATC and C115A-reverse TGGCGGGCAAATGTACCGCAG; C174A-forward GCCCAAACTCggGTATGAC and C174A-reverse TGCCGGGCAATGTACCGCAG; C115A-forward GCACAACCTGGCGGGGAATC and C115A-reverse TTGCGGGCAATGTACCGCAG; C174A-forward GCCCAAACTCggGGGGAATC and C174A-reverse TGGCGGGCAAATGTACCGCAG. All primers were 5’-phosphorylated and manufactured by Thermo Fisher Scientific (Ulm, Germany). PCR products were purified using QIAEX II Gel Extraction Kit (QIAGEN, Hilden, Germany) and ligated using T4 DNA ligase (Stratagene, La Jolla, CA). Mutations were confirmed through DNA sequencing at Uppsala Genome Center (Rudbeck Laboratory, Uppsala, Sweden). E. coli XL1 Blue cells (Stratagene, La Jolla, CA) were transformed with the mutant DNA sequence C87A/C115A/C174A/M212C (AAAM212C) by electroporation using a 0.1-cm electroporation cuvette at 1.25 V. The enzyme was expressed, purified and quantified as previously described (11, 12) with the following exceptions: overnight cultures were diluted 100-fold; 10 mM Tris-HCl pH 7.8 was used as equilibration buffer for affinity chromatography, and the GST protein was eluted using 10 mM GSH in the equilibration buffer. The purified protein was dialyzed against the equilibration buffer containing 1 mM 2-mercaptoethanol. Thermal inactivation studies were made in PBS for one hour to assess the stability of the AAAM212C mutant at 48 °C and the remaining activity was assayed with CDNB (7).

In vitro chemical modifications — The AAAM212C variant was chemically modified with different alkyl iodides as previously described (13). Reactions were performed in the dark overnight (15 h) with gentle shaking on a Vortex-Genie. As a control the AAAM212C variant was also subjected to the same procedure without the addition of any alkylating agent. The chemically modified variants were purified on a PD-10 column, eluted with 10 mM Tris-HCl buffer pH 7.8, and concentrated by centrifugation with PALL Centrifugal Devices Nanosep 10K Omega. The modified and purified enzyme variants were subjected to analyses on the day of purification and were shown to be stable during the experiments.

Proteolytic digestion and mass spectrometry — The presence of the desired alkylation on Cys-212 was confirmed by in-solution proteolytic digestion of GST M2-2. The primary structure around Cys212 is such that the conventional choice of protease, trypsin, would produce a fragment of only seven amino acids (212-CAVWGNK-C-terminus, MH+ = 777.36 Da), which is unfavorable in MALDI/ToF-MS. Also, cleavage near the modified Cys could present a challenge for the trypsin protease. Instead Glu-C protease (Sequencing grade, Roche
Diagnostics, Mannheim, Germany) was used for fragmentation of the alkylated GST protein. The resulting peptides were analyzed by matrix-assisted-laser-desorption/ionization time-of-flight-mass-spectrometry (MALDI-ToF-MS), on an Ultraflex III TOF/TOF, (Bruker Daltonics, Bremen, Germany) following the manufacturer’s instructions and using 2,5-dihydroxybenzoic acid as matrix. The spectra were calibrated externally using a mixture of seven peptides. The data obtained were analyzed using GPMAW ver. 8.1 (Lighthouse Data, Odense, Denmark).

HPLC– The chemically modified AAAM212C enzyme was separated from the unmodified variant on HPLC using a Jupiter 5µ C18 300A column (dimensions 250x4.60 mm) to determine the degree of modification. Typically 50 µl of the modified and purified enzyme mixture, with an enzyme concentration of approximately 0.4 mg/ml, was filtered using Acrodisc LC 13 mm syringe filter with 0.2 µM PVDF Membrane (PALL Life Sciences). The mixture was loaded onto the Hitachi Transgenomic L-7200 Autosampler and the chromatography was run with a Hitachi Transgenomic L-7100 Pump using a 1 ml/min flow rate for 125 min. The gradient was formed using buffer A (H2O with 0.1 % TFA) and buffer B (HPLC-grade acetonitrile with 0.1 % TFA). The temperature was kept at 30 °C by a Hitachi Transgenomic L-7300 Column Oven, and the protein products were detected at 280 nm using a Hitachi Transgenomic L-7400 UV Detector. The separation of chemically modified variants from the unmodified variant provides a quantitative and qualitative measure of the modification degree in a straightforward manner. When peaks corresponding to modified and unmodified enzyme were not resolved into discrete peaks, an approximate deconvolution was used rather than an actual measurement to obtain the degree of modification.

Activity measurements– Five electrophilic substrates were used to investigate the effect of the chemical modifications of GST M2-2. Specific activities with 1-chloro-2,4-dinitrobenzene, CDNB (14), 2-cyano-1,3-dimethyl-1-nitrosoguanidine, cyanoDMNG (14), (2S,3S)-(−)-3-(4-nitrophenyl)-glycidol, NPG (11), aminochrome (15) and azathioprine (16) were determined by spectrophotometric assays at 30 °C under standard conditions as described.

RESULTS

Cloning a variant of hGST M2-2 optimized for chemical modification– The three Cys residues in wild type hGST M2-2 were replaced with Ala, and the enzyme variant C87A/C115A/C174A was then further modified by a Met212Cys-mutation to give mutant AAAM212C. A close-up of the hGST M2-2 active site displays Met212 located near the C-terminus and situated between the catalytically significant Tyr116 and Ser210 (Fig. 1a). The nature of residue 210 was previously shown to influence the activity and substrate selectivity profile of the enzyme (7). We hypothesized that by chemically modifying the Cys sulfur and thereby changing the topology of the active site we would modulate the catalytic efficiency and the discrimination among different substrates. Methylation of Cys gives rise to normethionine and other alkylations generate related Met analogs. Fig. 1b illustrates the modifications made to the AAAM212C-variant and shows how the unnatural residues in position 212 are extended by methylene groups in a homologous series as well as by other substitutions.

These Met analogs based on Cys212 render the sulfur closer to the protein backbone by one -CH2- group than the sulfur atom in the wild type Met212. From a protein redesign perspective this can be expected to facilitate insertion of slightly larger chemical groups in the protein. Furthermore, residue 212 is located in the C-terminal region near the surface of the protein and the chemically attached groups could therefore be expected not to cause excessive perturbation of the tertiary structure.

Synthesis of Met analogs through alkylation of Cys212– Different n-alkyl iodides were used for chemical modification of Cys212 in mutant AAAM212C. The reagents have different reactivities with the Cys sulfur, mostly correlated to the length of the carbon chain, but also dependent on any additional functional groups. Nonetheless, the different reactions were run in a uniform manner in order to make use of the same control and simplify the comparison of results. The time frame for the chemical modifications was set so that even the more slowly reacting thiol modifying agents would afford ≥50% derivatization of the enzyme.
MALDI/ToF MS verifies the chemical modifications—Glu-C was utilized for proteolysis of the modified GST mutants, and peptide fragments of suitable size for MALDI/ToF MS analysis were produced (KISAYMKSSRFPLRPVFXTKC(X)AVWGNK, MH+ = 3014.62+X). The analyses of the various alkylated peptide fragments revealed in all cases a small remnant of the unmodified Glu-C peptide as well as the modified counterpart with an additional mass of 14, 28, 42, 56, 70, 84, 98, 58, 44 and 57 Da corresponding to the substituents methyl, ethyl, propyl, butyl, pentyl, hexyl, heptyl, carboxymethyl, hydroxyethyl and carbamidomethyl, respectively (see Fig. 2a for a typical example). The mass spectra did not indicate alkylation of any peptide fragments other than the Cys-containing sequence. No indication for oxidation of the Cys sulfur was observed. The desired alkylations were thus confirmed, but the heights of the relevant mass peaks could not be used to accurately determine the extent of chemical modification. 

HPLC provides a quantitative measure of the chemical modifications—The chemically modified enzyme variants were chromatographed in undigested form on a C18 HPLC column to determine the degree of chemical modification. Most of the modified enzyme variants were separated from the unmodified AAAM212C on the HPLC column when the covalently bound alkyl was larger than an ethyl group (see Fig. 2b for a typical example). Proteins derivatized with the more hydrophilic groups, such as hydroxyethyl, could not be distinguished from the unmodified form in the HPLC chromatograms. Since these more hydrophilic reagents are more reactive than the hydrophobic counterpart (ethyl iodide in this case), we assume that these reactions have run to completion. All chemical modifications resulted in ≥50 % modification (see Table 1), as calculated from the area under the curve in the corresponding HPLC chromatograms.

Differential effects on the enzyme activity with alternative substrates—Despite the three Cys-Ala replacements and the introduction of Cys in position 212 the enzyme was remarkably active and stable (see entry for AAAM212C (H-) substitution in Table 1). The specific activities of the AAAM212C mutant and the alkylated enzyme variants were determined with five different substrates, the general GST substrate CDNB, the two GST M2-2 distinctive substrates cyanoDMNG and aminochrome, the epoxide substrate NPG, and the pro-drug azathioprine (Fig. 3). Wild-type hGST M2-2 has modest activity (0.05 µmol·min⁻¹·mg⁻¹) with azathioprine (16), but the AAAM212C variant displayed no detectable activity. The chemical modification of Cys212 did not restore activity with azathioprine with any of the reagents tested, and the results with this substrate are therefore not recorded in Table 1. However, it is noteworthy that azathioprine activity of wild-type hGST M2-2 could be enhanced 30-fold by two point mutations of other residues in the active site (our unpublished data).

The specific activities of mutant C87A/C115A/C174A were comparatively high with three of the alternative substrates: CDNB, NPG and aminochrome, but somewhat lower (50 – 70%) than those of wild-type hGST M2-2. An exception was cyanoDMNG, which gave only 2.5% of the wild-type activity. This difference indicates a role for Cys in the transnitrosylation reaction of cyanoDMNG that is not required in the aromatic substitution (CDNB), epoxide opening (NPG), and addition to theortho-quinone (aminochrome). Cys115 could possibly play a role in the cyanoDMNG reaction, since it is adjacent to Tyr116, which contributes to other catalytic reactions of Mu class GSTs (17).

AAAM212C, which served as the parental enzyme for the alkylated variants, was more active than mutant C87A/C115A/C174A with CDNB, cyanoDMNG, and aminochrome, but showed only half of the NPG activity. The aminochrome activity was raised 2-fold to the level of wild-type hGST M2-2 by the Met212Cys substitution; the cyanoDMNG activity was elevated 10-fold but was still only 25% of the wild-type activity.

Following alkylation of Cys212 in AAAM212C the four reactive substrates displayed varying degrees of activity depending on the modifying group (Table 1 and Fig. 4). When a methyl group was attached to Cys212, a >20-fold decrease in the CDNB activity was noted, the NPG activity was undetectable, whereas the cyanoDMNG and aminochrome activities were considerably less affected. Ethyl and propyl groups gave significantly higher activities with all substrates except aminochrome, which gave essentially the same activity as after the methyl substitution. In
particular, the oxirane ring opening activity with NPG was raised from immeasurable with the methyl substitution to the activity of wild-type hGST M2-2.

The effects of the homologous alkyl groups were not monotonic. A butyl group attached to Cys212 gave lower activities with CDNB, aminochrome, and cyanoDMNG, but higher activity with NPG than expected from the closest alkyl homologs. By contrast, the butylated enzyme showed 2-fold higher activity with NPG than the wild-type value, and showed the highest activity of all the variants with this substrate (Table 1). Of the unsubstituted alkyl groups, hexyl gave the highest activity with both CDNB and cyanoDMNG; the activity with aminochrome was relatively insensitive to the length of the carbon chain. The values for the longest alkyl chains are biased towards lower values due to incomplete modification of the enzyme, but the qualitative conclusions remain unchanged.

In the cases of long-chain substitutions giving approximately 50% modification of the protein, the result in Table 1 show that dimeric enzyme displays subunit cooperativity, since the observed specific activities are not additive. For example, the CDNB activity of the heptyl derivative is < 50% of the value for the unmodified AAAM212 mutant; by contrast, the NPG activity is increased 2-fold. Thus, the enhanced activities with some substrates exclude the interpretation that 50% of the enzyme is inactivated by long-chained alkyl groups. Considering the extensive reaction time for modification, the data suggest that only one of the subunits of the dimeric protein has been chemically modified rather than half of the dimer molecules. Half-of-the-sites behavior has been observed for many oligomeric proteins and previously been noted with GSTs (18).

Cys212 was also alkylated with the more hydrophilic hydroxyethyl, carboxymethyl, and carbamidomethyl groups. The effects of carboxymethyl and carbamidomethyl substitutions were similar on the four alternative substrates and in no case were the observed activities higher than for the most effective alkyl substituent; with NPG no activity could be detected. The hydroxyethyl substitution gave higher activities, and the value for aminochrome exceeded the value of wild-type hGST M2-2 (Table 1).

The divergent effects of the different substituents at Cys212 in AAAM212C on the activities with the alternative substrates are illustrated in radar plots (Fig. 4a), which can be considered as activity fingerprints resulting from the diverse substitutions. The shapes of the trajectories linking the different substitutions in the graphs show obvious similarities between CDNB and cyanoDMNG, substrates that both undergo substitution reactions. The trajectory of NPG has the most divergent pattern.

Altered substrate selectivity profiles via diverse Cys-X substitutions—Glutathione transferases are generally promiscuous enzymes, as indicated by the promiscuity index (19), and they accept a wide range of alternative substrates representing many different types of reaction. For biotechnological applications, such as fine chemical synthesis, promiscuity will not be a desirable trait for an engineered enzyme. Substrate selectivity may be more important than absolute catalytic efficiency, as is also the case for enzymes involved in natural biosynthetic reactions in a cell. Fig. 4b illustrates the changes in relative activities with the four reactive substrates caused by the diverse substitutions of Cys212 in the AAAM212C mutant. In comparison with the other substrates the activity with NPG is negligible, and no activity at all could be observed with azathioprine. However, the discrimination between the other three substrates is modulated by the chemical substitutions of Cys212. A methyl group gives an enzyme with a clear preference for aminochrome among the substrates. In general, the relative rate with CDNB is significantly more variable than the other substrates. Additional substitutions of Cys212 may bias the substrate selectivity further in different useful directions.

**DISCUSSION**

**Limitations set by Nature**—Sequence space is sparsely populated by functional protein structures. However, the emerging fitness landscape is generally regarded as having select regions that are relatively dense with functional structures (20, 21). Redesign of such functional structures by exploring their neighborhood in sequence space has proved to be a fruitful approach to protein engineering. Even though
the recombination of the naturally available residues gives rise to a virtually limitless number of functional proteins, the chemical repertoire they represent is still limited. Twenty canonical amino acids have sustained life since the divergence of bacteria and man and they are sufficient for life as we know it (22). However, the proteogenic set of amino acids was established at a time in evolution when the selection pressure was quite different from present day conditions, and a contemporary de novo selection of amino acids would not necessarily result in the same set of amino acids (23). However, molecular genetics and biochemistry can accomplish diversification of the chemical space of proteins by additions to the set of the proteogenic amino acids. Proteins containing novel unnatural amino acids can find valuable use in biotechnical applications, e.g., as novel catalysts in fine-chemical synthesis or in protein therapeutics and nanotechnology (24, 25).

Introduction of unnatural amino acids via ribosomal protein synthesis—Schultz and coworkers have designed methods to overcome the limitations set by the genetic code and added 70-odd novel amino acids with diverse side chains in order to enrich protein functions (25). Numerous related procedures have been developed, but each of them is hampered by particular limitations (26). Obstacles include amino acids that are incompatible with ribosomal translation or have structures toxic to cells. Searching for an orthogonal aminoacyl-tRNA synthetase/iso-tRNA pair (aaRS/isotRNA pair) that does not cross-react with existing aaRS/isotRNA pairs can also present a challenge (27).

Post-translational modifications of proteins—In part, the need for other functions of the amino acids has been solved by Nature via post-translational modifications (1). However, this process is restricted to certain amino acids and the chemistries that are introduced are functionally limited. A different approach to the generation of unnatural amino acids involves the chemical modification of Cys ex vivo, as developed in this work. The Cys residue can be naturally occurring in the protein or be introduced by means of site-directed mutagenesis. The chemical modification is usually expeditious and can be performed with native protein or under denaturing conditions, depending on the accessibility of the Cys residue. It should be added that with some proteins the native sulfhydryl groups are buried or otherwise unreactive, such that only the inserted Cys will be modified (34). Even though the method applies without difficulty only to Cys residues, this intrinsic limitation is offset by the rich chemistry of thiol modifications.

Ala- and Cys-scanning—The Wells laboratory introduced a strategy, Ala-scanning mutagenesis, by which amino acids in a certain region were successively replaced with Ala, in order to explore side-chain contributions to the interaction of growth hormone with its receptor (2). The method has subsequently been used successfully to map, e.g., functional epitopes in proteins and the energetics of molecular interactions (28, 29). A similar Cys-scanning mutagenesis was subsequently employed to study the formation of disulfide bridges in lysozyme (30). Cys-scanning has also been used to investigate the pore topology in ion channels and conformational changes of proteins associated with channel gating in membranes (3, 4) as well as in studies of other membrane proteins (31). In these studies chemical modifications have been limited, and primarily aimed at specifying the location of a certain amino acid, inside or outside a membrane. By an orthogonal approach, we and others have focused on diverse modifications of one or a few cysteines to alter the functional properties of GSTs and other proteins. The results include, e.g., altered catalytic efficiencies, enhanced and reversed enantioselectivities (32, 33, 34, 13) and boosted T-cell recognition via novel thiol-modified peptide variants (35).

In the present study we have focused on modulating the active-site functions in a promiscuous GST enzyme by mutating the naturally occurring methionine into a cysteine and subsequently modifying it with a suite of alkylating agents. A small library of unnatural amino acids was scanned with a set of substrates to find an unnatural amino acid capable of turning a mutant of hGST M2-2 into variants with diverging substrate selectivity profiles. We managed to decrease activities dramatically with some substrates and enhance activities with others, and introduce radar plots to display how different substrates are acted upon by the various enzyme derivatives (Fig. 4a). The spikes originating from the center of
the plot are the axes indicating the activities of the respective enzyme variants acting on a given substrate. The radar diagram is thus the complement to the plot of relative activities with the alternative substrates (Fig. 4b). For more incisive studies of structure-activity relationships the kinetic parameters of the modified enzymes should be studied, but the present investigation aims primarily at showing the potential of Cys-X scanning.

**The importance of substrate selectivity**—The unique and defining character of enzymes was long considered to be the specificity with which they catalyze chemical reactions, in contrast to conventional catalysts in organic chemistry where such specificity is a challenge to accomplish. Indeed, enzymes show specificity, but increasingly enzymes are unraveled as being catalytically promiscuous (36). Many forms of promiscuity with different classifications exist and have been reviewed elsewhere (37-39). Functional promiscuity may underlie various hidden metabolic functions where reactions are carried out at undetectable levels until they become important when substrate or enzyme concentrations change for some reason. Such backup functions can explain why organisms can be more robust than expected following a gene deletion in a major metabolic pathway (37). On the other hand, substrate promiscuity is an essential trait of members of a detoxification enzyme superfamily such as the GSTs. Promiscuous enzymes are excellent starting points for protein redesign experiments aimed at a biocatalyst for new reactions (with new chemistries) or for selective transformation of a particular substrate. Altering the substrate selectivity can prove useful in various applications in diverse areas of biotechnology. In fact, the selectivity profile of an enzyme is generally more important than the catalytic efficiency, since low activity can be compensated by increased amounts of enzyme, but undesired side-reactions cannot.

**Applications of Cys-X scanning**—We focus here on the aspect of substrate promiscuity, i.e. the ability, with a mutant of hGST M2-2, to catalyze the conjugation of the tripeptide GSH with various substrates. For certain biotechnological applications it can be crucial to find a highly specific enzyme with the ability to discriminate between similar substrates. GSTs are generally considered to have broad substrate acceptance, and there are successful examples of GST redesign, some with only one amino acid change (40-44). We have shown in this study that incorporation of unnatural amino acids in a mutant of hGST M2-2, by means of chemical modifications, can shift the preference from one substrate to the other (Fig. 4). This is a straightforward method where a chemically modified enzyme variant can subtly or dramatically alter the substrate selectivity profile of the enzyme. A scan with various functional groups of X attached to a cysteine allows for the mapping of a desired function. It should be added that according to conventional wisdom most mutations of an enzyme are deleterious to function. None the less, our results show that all modified enzymes have high activity with at least two of the substrates tested.

Current approaches to computational design of enzymes call for experimental complements. In many cases, the computed active site cavity shows geometrical deficiencies that prevent an optimal fit of the transition state of the targeted reaction (45). It is possible that enhanced activities can be obtained by filling empty cavities by suitable amino acid substitutions in the active site, thus accomplishing proper transition-state stabilization. Cys-X scanning with a set of alternative X-substituents could have advantages over site-directed mutagenesis with conventional amino acids. This idea needs to be tested in practice. In the present study we have limited the substituents to simple alkyl groups. However, branched sidechains and aromatic residues have also been considered, and unless steric hindrance or other restrictions limit the access to the targeted Cys residues in the protein, there is an almost limitless number of substitutions that can be used in the Cys-X scanning. Some reagents used in the experiments may be less reactive resulting in incomplete modification. However, also a preparation with a partially derivatized protein may reveal the desirable properties in the screening experiments. Subsequent optimization of reaction conditions can lead to complete derivatization of the desired protein variants.

Mutations and the chemical substitutions may also influence other molecules than substrates that bind to a protein. In many GSTs
water molecules are located close to the active-site Tyr (46), and changing the topography of the active site by introduction of covalently bound groups may change the network of H-bonded water. Irrespective of whether they are directly involved in the catalytic process (47, 48), displacement of a water molecule or accommodation of an additional one can have functional consequences. Such indirect effects are generally difficult to predict, even with the most advanced computational methods (45) and have to be investigated directly by experiments. Cys-X scanning in combination with functional and structural studies can be a valuable tool in the investigation of the importance of hydration of protein binding sites.

Conclusion— Even though the natural complements of amino acids, by definition, presents solutions to extant biological systems at the molecular level, it is reasonable to assume that the fitness of many proteins can be enhanced by addition of novel chemical building blocks. Nature provides examples in which heme (49), FAD (50), and other prosthetic groups have been post-translationally and covalently linked to amino acid sidechains. The substitutions enable redox reactions in active sites of enzymes (51). The present study shows that even simpler substitutions, such as alkyl residues without reactive groups, can alter the functionality of an active site. Optimization of proteins will be valuable in protein engineering in the search for optimal properties and for targeted function.

The case of GST M2-2 provides proof of the concept that Cys-X scanning can effect the modulation of enzymatic activities. Residue 212 was selected for Cys-X modifications on the basis of its location in the structure of GST M2-2. In the general case, Ala- or Cys-scanning can be employed for identifying residues in the primary structure that are suitable for modifications. After an initial screening with different modifying agents, optimization may be required to secure complete derivatization of the protein with a promising reagent. Many inexpensive chemical reagents can provide facile modifications of a thiol, and Cys-X modifications can consequently be carried out in a scale that is compatible with biotechnical applications. It would appear that Cys-X scanning is a useful complement to other methods for expanding the repertoire of amino acid residues in proteins.
REFERENCES

1. Walsh, C. T., Garneau-Tsodikova S., and Gatto, Jr., G. J. (2005) Angew. Chem. Int. Ed. 44, 7342–7372
2. Cunningham, B.C., and Wells, J.A. (1989) Science 244, 1081-1085
3. Kubo, Y., Yoshimichi, M., Heinemann, S.H. (1998) FEBS Letters 435, 69-73
4. Pinto, L.H., Dieckmann, G.R., Gandhi, C.S., Papworth, C.G., Braman, J., Shaughnessy, M.A., Lear, J.D., Lamb, R.A., and DeGrado, W.F. (1997) Proc. Natl. Acad. Sci. U.S.A. 94, 11301-11306
5. Josephy, P. D., and Mannervik, B. (2006) Molecular Toxicology, pp.333–364, Oxford University Press, New York
6. Armstrong, R.N. (1997) Chem. Res. Toxicol. 10, 2-18
7. Norrgård, M.A., Ivarsson, Y., Tars, K., and Mannervik, B. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 4876-4881
8. Emrén, L.O., Kurtovic, S., Runarsdottir, A., Larsson, A-K., and Mannervik, B. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 10866-10870
9. Johansson, A.-S., Bolton-Grob, R., Mannervik, B. (1999) Protein Expr. and Purif. 17, 105-112
10. Widersten, M., Mannervik, B., (1995) J. Mol. Biol. 250, 115–122
11. Ivarsson, I., Mackey, A.J., Edalat, M., Pearson, W.R., and Mannervik, B., (2003) J. Biol. Chem. 278, 8733-8738
12. Simons, P.C., Vander Jagt, D.L. (1981) Meth. in Enzymol. 77, 235-237
13. Hegazy, U.M., Hellman, U., and Mannervik, B. (2006) Chem and Biol. 13, 929-936
14. Jemth, P., and Mannervik, B. (1999) Curr. Prot. Toxicol. (Wiley, New York), pp. 6.4.1–6.4.10
15. Dagnino-Subiabre, A., Cassels, B.K., Baez, S., Johansson, A.-S., Mannervik, B., and Segura-Aguilar, J. (2000) Biochem. Biophys. Res. Commun. 274, 32–36
16. Eklund, B.I., Moberg, M., Bergquist, J., and Mannervik, B., (2006) Mol. Pharmacol., 70, 747-754
17. Johnson, W.W., Liu, S., Ji, X., Gilliland, G.L., Armstrong, R.N., (1993) J. Biol. Chem. 268, 11508-11511
18. Lien, S., Gustafsson, A., Andersson, A.-K., and Mannervik, B. (2001), J. Biol. Chem. 276, 35599-35605
19. Nath, A., and Atkins, W.M. (2008) Biochemistry 47, 157-166
20. Kauffman, S.A. (1993) The Origins of Order: Self-Organization and Selection in Evolution, Oxford University Press, New York
21. Romero, A., and Arnold, F.H. (2009) Nat. Rev. Mol. Cell Biol. 10, 866-876
22. Lu, Y., Freeland, S. (2006) Gen. Biol. 7, 102
23. Sella, G., Ardell, D.H. (2006) J. Mol. Evol. 63, 297-313
24. Desantis, G., and Jones, B. (1999) Curr. Opin. in Biotech. 10, 324-330
25. Liu, C.C., and Schultz, P.G. (2010) Annu. Rev. Biochem., 79, 413-444
26. Hegazy, U.M., and Mannervik, B. (2009) in Amino Acids, Peptides and Proteins in Organic Chemistry, Vol. 2, pp. 379-431 (A.B. Hughes, ed.), Wiley-VCH, New York
27. Wang, Q., Parrish, A.R., and Wang, L. (2009) Chem. and Biol. 16, 323-336
28. Weiss, G.A., Watanabe, C.K., Zhong, A., Goddard, A., Sidhu, S.S. (2000) Proc. Natl. Acad. Sci. U.S.A. 97, 8950-8954
29. Thorn, K.S., and Bogan, A.A. (2001) Bioinformatics 17, 284-285
30. Kanaya, E., Kanaya, S., and Kikuchi, M. (1990) Biochem. Biophys. Res. Commun. 173, 1194-1199
31. Sahin-Tóth, M., and Kaback, H.R. (1993) Protein Science 2, 1024-1033
32. Ivarsson, Y., Norrgård, M.A., Hellman, U., and Mannervik, B. (2007) Biochim. Biophys. Acta 1770, 1374-1381
33. DeSantis, G., and Jones, J.B. (1999) Bioorg. and Med. Chem. 7, 1381-1387
34. Wright, S.K., and Viola, R.E. (2001) J. Biol. Chem., 276, 31151-31155
35. Manning, T.C., Schodin, B.A., and Kranz, D.M. (1996) J. Immunol. Methods. 192 125-132
36. Khersonsky, O., and Tawfik, D.S. (2010) Annu. Rev. Biochem. 79, 471-505
37. Nobeli, I., Favia, A.D., and Thornton, J.M. (2009) Nature Biotechnology 27, 157-167
38. Hult, K., and Berglund, P. (2007) Trends Biotechnol. 25, 231-238
39. Jestin, J.-L., and Vichier-Guerre, S., (2005) *Res. Microbiol.* **156**, 961-966
40. Björnstedt, R., Tardioli, S., and Mannervik, B. (1995) *J. Biol. Chem.* **270**, 29705-29709
41. Nuccetelli, M., Mazzetti, A.P., Rossjohn, J., Parker, M.W., Board, P., Caccuri, A.M., Federici, G., Ricci, G., and LoBello, M. (1998) *Biochem. Biophys. Res. Commun.* **252**, 184-189
42. Nilsson, L.O., Gustafsson, A., and Mannervik, B. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9408-9412
43. Pettersson, P.L., Johansson, A.-S., and Mannervik, B. (2002) *J. Biol. Chem.* **277**, 30019-30022
44. Blikstad, C., Shokeer, A., Kurtovic, S., and Mannervik, B. (2008) *Biochim. Biophys. Acta* **1780**, 1458-1463
45. Kiss, G., Röthlisberger, D., Baker, D., and Houk, K.N. (2010) *Protein Sci.* **19**, 1760-1773
46. Grahn, E., Novotny, M., Jakobsson, E., Gustafsson, A., Grehn, L., Olin, B., Madsen, D., Wahlberg, M., Mannervik, B., and Kleywegt G.J. (2006), *Acta Crystallogr. D Biol. Crystallogr.* **62**, 197-207
47. Dourado, D.F.A.R., Fernandes, P.A., Mannervik, B., Ramos, M.J. (2008) *Chem. Eur. J.*, **14**, 9591-9598
48. Dourado, D.F.A.R., Fernandes, P.A., Ramos, M.J. (2010) *J. Phys. Chem. B* **114**, 12972-12980
49. Theorell, H., and Åkesson, Á. (1941) *J. Am. Chem. Soc.* **63**, 1804-1811
50. Singer, T. P., and McIntire, W. S. (1984) *Methods Enzymol.* **106**, 369-378
51. Davidson, V.L. (2011) *Mol. BioSyst.* **7**, 29-37

**FOOTNOTES**

The abbreviations used are: GST, glutathione transferase; GSH, glutathione; AAAM212C, A87C/A115C/A174C/M212C hGST M2-2; CDNB, 1-chloro-2,4-dinitrobenzene; cyanoDMNG, 2-cyano-1,3-dimethyl-1-nitrosoguanidine; NPG, (2S,3S)-(−)-3-(4-nitrophenyl)-glycidol

Keywords: Cys-X scanning, chemical modification, unnatural amino acids, substrate selectivity, normethionine

Acknowledgements – We thank Dr. Birgitta Tomkinson and Dr. Diana Lindberg for advice on the manuscript.

**FIGURE LEGENDS**

**FIGURE 1.** Close-up of the hGST M2-2 T210S active site and the chemical substituents introduced at Cys in position **212**. *a*, The structure (PDBcode: 2C4J) displays a glutathione-styrene oxide conjugate GS-SO (styrene portion in blue). The catalytically significant residues Tyr116 and Ser210 (Thr210 in the wild type enzyme) shown in stick representation. The targeted residue Met212 in ball-and-stick representation mutated to Cys212 for the substitutions is situated close to the mu-loop (left) covering the active site. *b*, Met212 as well as Cys212 (H-) are shown above the different unnatural amino acid side chains created by chemical modification of Cys212. Abbreviations: Me- (methyl-), Et- (ethyl-), Pr- (propyl-), Bu- (butyl-), Pe- (pentyl-), Hex- (hexyl-), Hep- (heptyl-), Cox- (carboxymethyl-), HOEt- (hydroxyethyl-), and Cm- (carbamidomethyl-).

**FIGURE 2.** Analysis by mass spectrometry and HPLC of AAAM212C mutant of hGST M2-2 chemically modified by 1-iodohexane. *a*, Close-up of a peptide mass finger-print (MALDI/ToF-MS) following Glu-C digestion of the hexyl-modified Cys212 enzyme rendering a peptide of MH+= 3014.62 m/z for the unmodified protein and MH+= 3098.78 m/z for the derivatized protein (difference 84.2 m/z). *b*, HPLC chromatogram of the 1-hexane-treated AAAM212C mutant, showing modified and unmodified proteins.

**FIGURE 3.** Alternative reactions investigated with the different GST variants. CDNB and azathioprine represent aromatic substitutions, cyanoDMNG a transnitrosylation, aminochrome an
ortho-quinone addition, and NPG an oxirane ring opening. The deprotonated GSH can attack at two positions in NPG as indicated by numbers.

FIGURE 4. Effects of alkylation of the AAAM212C mutant on the substrate selectivity profile. a, Radar plots illustrate how different substrates are acted upon by the various enzyme derivatives. The spikes originating from the center of the plot are the axes showing the specific activities (µmol·min⁻¹·mg⁻¹) of the respective enzyme variants acting on a given substrate. As compared with the activity of the unmodified mutant AAAM212C at the top (H-), the activities on the different axes increase or decrease depending on the Cys212 modification, and each substrate displays a distinctive pattern: (1) CDNB, (2) aminochrome, (3) NPG, and (4) cyanoDMNG. Designations of enzyme variants: H- corresponds to unmodified Cys212 in mutant AAAM212C, and the following are S-substituents on Cys 212: Me- (methyl-), Et- (ethyl-), Pr- (propyl-), Bu- (butyl-), Pe- (pentyl-), Hex- (hexyl-), Hep- (heptyl-), Cm- (carbamidomethyl-). b, The relative activities of the different enzyme derivatives with respect to the alternative substrates investigated. The activities with NPG and azathioprine are several orders of magnitude lower than the other substrates and are invisible.
TABLE 1
Characteristics of hGST M2-2 variants showing the effects of different structures of residue 212 in the active site

Specific activities (µmol-min⁻¹·mg⁻¹) were determined with alternative substrates, extent of modification (in %) was estimated by HPLC analysis, and thermal stability was indicated by the half-life at 48 °C.

| Cys212-modification | Substrate | Modification | Half-life |
|---------------------|-----------|--------------|-----------|
|         | CDNB | NPG | cyanoDMNG | Aminochrome | (%) | (min) |
| Parental C87A/C115A/C174A | 538±18 | 0.19±0.06 | 9±0.1 | 207±26 | - | 10±2 |
| AAAM212C (H⁻) | 641±20 | 0.09±0.01 | 87±3 | 429±26 | - | 18±2 |
| Methyl- | 28±2 | ≤0.02ᵃ | 26±5 | 292±20 | 95ᵇ | - |
| Ethyl- | 131±6 | 0.22±0.02 | 36±2 | 286±6 | 95ᵇ | - |
| Propyl- | 160±27 | 0.25±0.03 | 56±2 | 306±20 | 95ᶜ | - |
| Butyl- | 68±9 | 0.47±0.07 | 23±1 | 103±11 | 90ᶜ | - |
| Pentyl- | 136±17 | ≤0.06ᵃ | 21±3 | 167±14 | 60 | - |
| Hexyl- | 388±7 | 0.23±0.06 | 67±5 | 286±21 | 60 | - |
| Heptyl- | 237±15 | 0.21±0.04 | 34±3 | 132±10 | 50 | - |
| Carboxymethyl- | 250±21 | ≤0.016ᵃ | 25±2 | 165±12 | 95ᵇ | - |
| Hydroxyethyl- | 663±52 | 0.02±0.01 | 116±5 | 541±21 | 95ᶜ | - |
| Carbamidomethyl- | 113±8 | ≤0.001ᵃ | 16±1 | 137±12 | 90ᶜ | - |
| hGST M2-2 | 860±20ᵈ | 0.27±0.16ᵈ | 354±8ᵈ | 450±35ᵈ | - | 220±120ᵈ |

ᵃ low activity that could not be quantified accurately
ᵇ considered to be reacted completely due to their reactivity but not verified
ᶜ estimate used to obtain degree of modification
ᵈ adopted from (7)
FIGURE 1.
FIGURE 2.
FIGURE 3.

![Chemical structures and reactions](image-url)
Cys-X scanning for expansion of active-site residues and modulation of catalytic functions in a glutathione transferase
Malena A. Norrgard, Ulf Hellman and Bengt Mannervik

J. Biol. Chem. published online March 23, 2011

Access the most updated version of this article at doi: 10.1074/jbc.M111.230078

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts