Optimization of Bivalent Glutathione S-Transferase Inhibitors by Combinatorial Linker Design

Sumit S. Mahajan,[‡] Liming Hou,[‡] Catalin Doneanu,[‡] Rajan Paranji,[¶] Dean Maeda,[§] John Zebala,[§] and William M. Atkins*,[‡]

Contribution from the Departments of Medicinal Chemistry, Box 357610, and Chemistry, Box 351700, University of Washington, Seattle, Washington 98195, and Syntrix Biosystems, Inc., 215 Clay Street NW, Suite B5, Auburn, Washington 98001

Received March 23, 2006; E-mail: winky@u.washington.edu

Abstract: Dimeric glutathione S-transferases (GSTs) are pharmacological targets for several diseases, including cancer. Isoform specificity has been difficult to achieve due to their overlapping substrate selectivity. Here we demonstrate the utility of bivalent GST inhibitors and their optimization via combinatorial linker design. A combinatorial library with dipeptide linkers emanating symmetrically from a central scaffold (bis-3,5-aminomethyl benzoic acid, AMAB) to connect two ethacrynic acid moieties was prepared and decoded via iterative deconvolution, against the isoforms GSTA1-1 and GSTP1-1. The library yielded high affinity GSTA1-1 selective inhibitors (70−120-fold selectivity) and with stoichiometry of one inhibitor: one GSTA1-1 dimer. Saturation Transfer Difference (STD) NMR with one of these inhibitors, with linker structure (Asp-Gly-AMAB-Gly-Asp) and $K_D = 42$ nM for GSTA1-1, demonstrates that the Asp-Gly linker interacts tightly with GSTA1-1, but not P1-1. H/D exchange mass spectrometry was used to map the protein binding site and indicates that peptides within the intersubunit cleft and in the substrate binding site are protected by inhibitor from solvent exchange. A model is proposed for the binding orientation of the inhibitor, which is consistent with electrostatic complementarity between the protein cleft and inhibitor linker as the source of isoform selectivity and high affinity. The results demonstrate the utility of combinatorial, or "irrational", linker design for optimizing bivalent inhibitors.

Introduction

"Multivalency" is a useful strategy in drug design, wherein multiple pharmacophores, or binding elements, on a single scaffold bind simultaneously to distinct protein sites.[1−6] This is a particularly obvious strategy for oligomeric protein targets, but the principle also holds for monomers with discrete "subsites" that recognize individual binding elements. An example of a a single protein subunit targeted by multivalent inhibitor design is acetylcholinesterase, wherein a potent bivalent inhibitor targeted toward the enzymes's two "anionic" binding sites on a single subunit was identified.[1,2] In principle, multivalency is a general strategy to enhance ligand affinity and selectivity.[3−6] In some cases, binding affinity is increased by several orders of magnitude for multivalent ligands compared to their monovalent counterparts. Although multivalent interactions could be a general strategy in theory, in practice, multivalency may be limited by the requirement for drugs with molecular weights below ∼1000 amu. Thus, bivalent or trivalent drugs may be a practical upper limit to the use of multivalency for design of clinically useful drugs.[1,5,7,8] In all multivalent ligands, characteristics of the linker play a critical role, as they have to be of appropriate length and flexibility to achieve a maximal benefit.[3,4] However, linker design is still a poorly understood process. For example, it is still not clear whether a flexible linker with length greater than the distance between active sites is preferable over a conformationally restricted linker with the exact distance between the sites. The former type is capable of adopting several different conformations to span the binding sites, thereby increasing the chance of potent multivalent inhibition at the entropic cost of restricting the linker. On the other hand, inflexible linkers may be advantageous, as there is less entropy loss upon binding. However, this advantage is offset by uncertainty about optimal length and steric features of the linker during early stages of iterative drug design. Irrespective of the uncertainty about length and flexibility of the linker, the existing paradigm for design of multivalent inhibitors usually ignores interactions between linker and the target molecule;

‡ Department of Medicinal Chemistry, University of Washington.
¶ Department of Chemistry, University of Washington.
§ Syntrix Biosystems, Inc.

(1) Greenblatt, H. M.; Guillou, C.; Guenard, D.; Argaman, A.; Botti, S.; Badet, B.; Thal, C.; Silman, I.; Sussman, J. L. *J. Am. Chem. Soc.* **2004**, *126*, 15405−15411.
(2) Guillou, C.; Mary, A.; Renko, D. Z.; Gras, E.; Thal, C. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 637−639.
(3) Fan, E.; Zhang, Z.; Minke, E. W.; Hou, Z.; Verlinde, C.; Hol, G. J. W. *J. Am. Chem. Soc.* **2000**, *122*, 2663−2664.
(4) Gargano, J. M.; Ngo, T.; Kim, J. Y.; Acheson, D. W.; Lees, W. J. *J. Am. Chem. Soc.* **2001**, *123*, 12909−12910.
(5) Loidl, G.; Groll, M.; Musiol, H. J.; Huber, R.; Moroder, L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5418−5422.
(6) Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2755−2794.
(7) Du, D. M.; Carlier, P. R. *Curr. Pharm. Des.* **2004**, *10*, 3141−3156.
(8) Dudkin, V. Y.; Orlova, M.; Geng, X.; Mandal, M.; Olson, W. C.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 9560−9562.
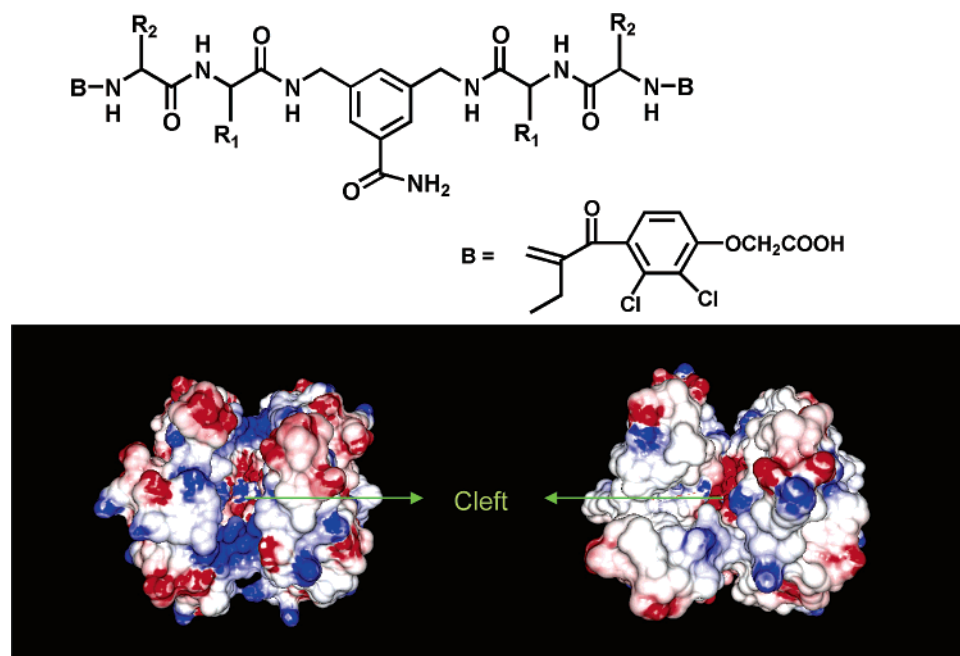
Figure 1. (Top) Bivalent linker scaffold with binding element, B, on each end. Here B = ethacrynic acid. (Bottom) Electrostatic potential map of GSTA1-1 (left) and GSTP1-1 (right). Red shows negative surface and blue positive. The intersubunit cleft of A1-1 is positively charged, whereas the P1-1 cleft is neutral.

linker/target interactions are often not considered in the design efforts. In some cases, the overall structure of the protein and relevant binding sites does not result in significant linker/protein interactions, such as with linkers that lie in solvated channels between pharmacophore binding sites.[3] However, in other cases, the linker would be expected to interact with the protein surface.

Here, we consider the possibility that linkers between binding elements within the classic paradigm of multivalent inhibitor design are an attractive target for exploitation of increased binding interactions. Modifications of the linker can readily affect the affinity, selectivity, or pharmacokinetic properties of these multivalent ligands. As a proof of principle of this idea, we have used split-pool combinatorial methods to synthesize a library of compounds with linkers composed of variable amino acids as potential inhibitors for glutathione S-transferase (GST). Previous studies have demonstrated that bivalent inhibitors have increased affinity for GSTs, compared to their monovalent analogues.[9]

GSTs have recently invoked considerable interest as a drug target in humans as well as in infectious agents.[10] Human GST isoforms (mainly GSTP1-1) are often found to be overexpressed in multidrug resistance tumors and might be responsible for development of drug resistance either by direct detoxification or by inhibition of the MAP kinase pathway.[11] In fact, inhibition of GSTP1-1 has been shown to modulate myeloproliferation and to increase the efficacy of antitumor photodynamic therapy in a model system.[12,13] GSTA1-1 also has emerged as a potential target because it has been implicated in oxidative stress and

inflammation.[14] GSTs are also critical for survival of parasites responsible for diseases, such as malaria and schistosomiasis, and parasite-specific GST inhibitors could have therapeutic utility.[15,16] These aspects of GST function have led to increased interest in novel GST inhibitors with potential for therapeutic intervention. Although numerous X-ray structures exist for various GSTs from many organisms, it has been difficult to obtain isoform-specific GST inhibitors. This may be due to the apparent degeneracy of GST binding sites, wherein several subsites comprise a large promiscuous active site, which allows for a single drug binding in multiple locations, or several structurally unrelated drugs binding in overlapping sites. New design strategies are, therefore, needed to identify highly selective GST inhibitors.

The quaternary structure of GSTs are ideally suited for bivalent inhibition. The dimeric structure exhibits a solvent-accessible cleft between the two active sites on each monomer. A comparison of the intersubunit cleft of GSTA1-1 with GSTP1-1 indicates a modest difference in distances between active sites and a striking difference in the electrostatic surface (Figure 1). The GSTA1-1 intersubunit cleft is markedly less hydrophobic than GSTP1-1 and includes significant positive charge density. These differences suggest the potential utility of multivalency, with emphasis on the linker moiety, to identify isozyme-selective GST inhibitors. The rationale is 2-fold: (1) increasing enthalpic interactions from the linker component with the cleft for improved affinity; (2) better selectivity owing to significant differences in clefts of the GST isoforms.

(9) Lyon, R. P.; Hill, J. J.; Atkins, W. M. *Biochemistry* **2003**, *42*, 10418−10428.
(10) Mahajan, S.; Atkins, W. M. *Cell Mol. Life. Sci.* **2005**, *62*, 1221−1233.
(11) Tew, K. D. *Cancer Res.* **1994**, *54*, 4313−4320.
(12) Gate, L.; Majumdar, R. S.; Lunk, A.; Tew, K. D. *J. Biol. Chem.* **2004**, *279*, 8608−8616.
(13) Dabrowski, M. J.; Maeda, D.; Zebala, J.; Lu, W. D.; Mahajan, S. S.; Kavanagh, T. J.; Atkins, W. M. *Arch. Biochem. Biophys.* **2006**, *449*, 94−103.

(14) Zhao, T.; Singhal, S. S.; Piper, J. T.; Cheng, J.; Pandya, U.; Clark-Wronski, J.; Awasthi, S.; Awasthi, Y. C. *Arch. Biochem. Biophys.* **1999**, *367*, 216−224.
(15) Srivastava, P.; Puri, S. K.; Kamboj, K. K.; Pandey, V. C. *Trop. Med. Int. Health* **1999**, *4*, 251−254.
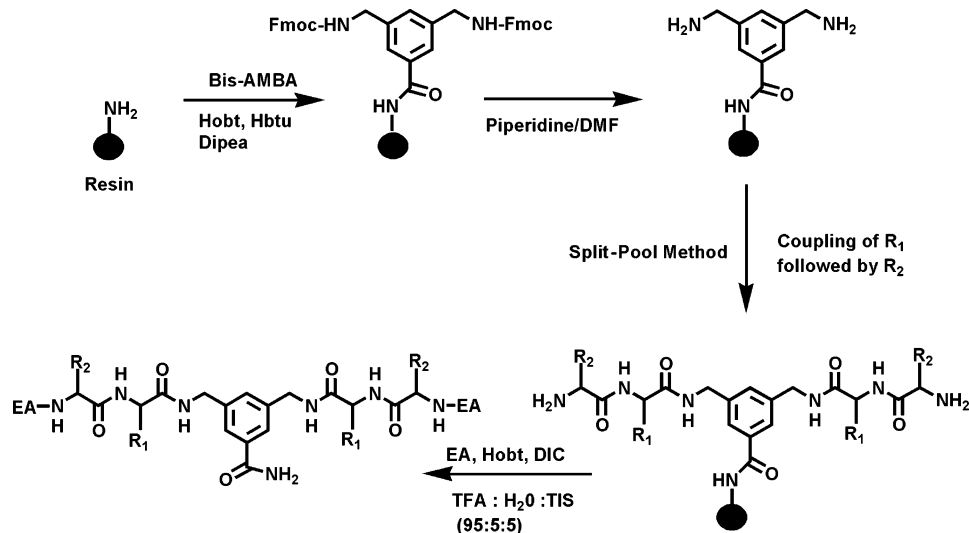(16) McTigue, M. A.; Williams, D. R.; Tainer, J. A. *J. Mol. Biol.* **1995**, *246*, 21−27.
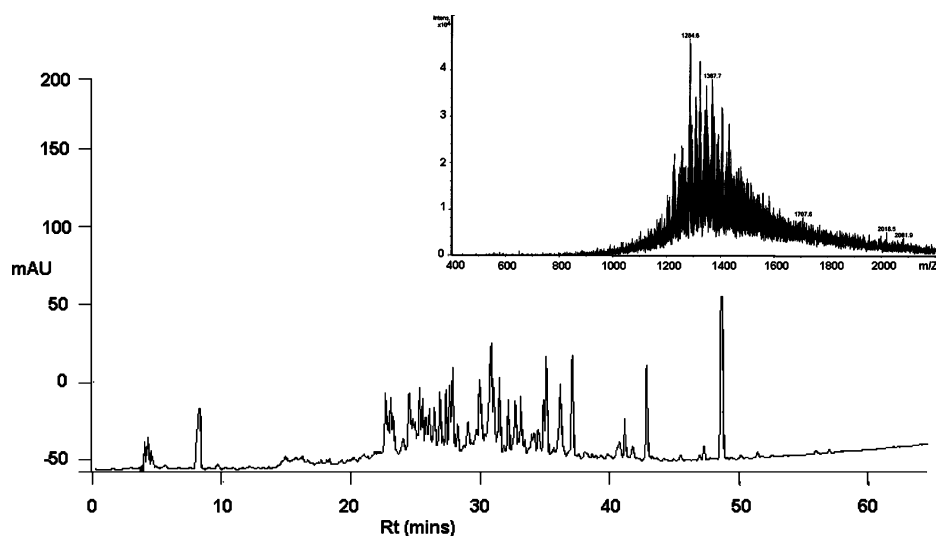
Figure 2. Scheme for synthesis of compounds.

Figure 3. LC trace of a library mixture with (R_2 = His, R_1 = X). Inset: MALDI-MS of the library. The mass envelope is in the range for this library of compound.

Results

Design Strategy: Crystal structures of GSTA1-1 and P1-1 with bound ethacrynic acid were used to determine the range of lengths of the linker. On the basis of available crystal structures, a linker of ~18−22 Å was estimated to be sufficiently long to span the cleft region of both A1-1 and P1-1. Modeling indicates that linkers composed of the 1,3-bis(amino)-methyl benzoic acid nucleus and two amino acid residues span a distance of 21 Å. Ethacrynic acid was selected as the binding element to be attached at the ends of this linker as it is a known inhibitor and has been shown in several clinical studies to increase remission from cancer resistance.[17] More importantly, it's a nonselective GST inhibitor ($K_D \sim 3.3$ μM for P1-1 and 6 μM for A1-1); hence the ethacrynic acid-based bivalent inhibitors designed and described here have minimal inherent bias toward a particular GST isoform.[18,19]

Library and Screening: To explore the "functional group space" of the linker, we synthesized a library of compounds using the split-pool combinatorial method. Twenty different amino acids, including a few unnatural amino acids, were used with two cycles of "split and pool" to give 400 compounds (Figure 2). The unnatural amino acids were cyclohexyl-L-alanine (cyx-A), pyridyl-L-alanine (Pyr-A), and biphenyl-L-alanine (Bip). Because iterative deconvolution was used to identify the most potent inhibitor, compounds were not pooled after the second round of coupling. This resulted in 20 different pools, each with a mixture of 20 different compounds, but with a known unique amino acid at the R_2 position. The mixtures were characterized using LC/MS as described in Materials and Methods. A typical LC trace and MALDI spectrum of the pool with R_2 = His are shown in Figure 3, which demonstrates a complex mixture of compounds with the expected mass range and chromatographic properties. Each mixture was then screened for inhibitory potential using the standard colorimetric CDNB assay, which monitors enzyme-dependent production of 1-glutathionyl-2,4-dinitrobenzene. Mixtures containing Asp and Glu at position R_2 resulted in 40 and 36% inhibition for GSTA1-1, respectively.

(17) Petrini, M.; Conte, A.; Caracciolo, F.; Sabbatini, A.; Grassi, B.; Ronca, G. *Br. J. Haematol.* **1993**, *85*, 409−410.
(18) Ploemen, J. H.; van Ommen, B.; van Bladeren, P. J. *Biochem. Pharmacol.* **1990**, *40*, 1631−1635.
(19) Ahokas, J. T.; Davies, C.; Ravenscroft, P. J.; Emmerson, B. T. *Biochem. Pharmacol.* **1984**, *33*, 1929−1932.

Table 1. Apparent K_D's of Bivalent Inhibitors[a]

compounds with $R_2 =$ Asp and $R_1 =$	$K_D \pm$ SE A1-1 (nM)	$K_D \pm$ SE P1-1 (nM)
Asp	7 ± 1.4	1221 ± 174
Lys	130 ± 14.5	1676 ± 343
Arg	110 ± 9.5	1759 ± 144
Glu	22 ± 1.5	951 ± 41
Bip	3.4 ± 1.6	344 ± 16
Phe	9 ± 2.5	860 ± 105
Cyx-A	7.2 ± 2	530 ± 147
Tyr	10 ± 2	380 ± 13
Trp	45 ± 6	4500 ± 1491
Ala	13.7 ± 1	204 ± 21
Val	27 ± 1.1	285 ± 58
Leu	56 ± 3	589 ± 49
Pyr-A	140 ± 49	1273 ± 94
Asn	40 ± 5	675 ± 45
Gln	31 ± 3.1	700 ± 114
Gly, **1**	42 ± 7.3	429 ± 30
His	6 ± 1.6	445 ± 52
Pro	6 ± 1.8	504 ± 30
Ser	20 ± 3.4	1708 ± 258
Thr	25 ± 4	186 ± 15
compounds with $R_2 =$ Arg and $R_1 =$		
Gly, **2**	740 ± 50	1800 ± 163
Arg	1118 ± 55	3103 ± 123

[a] Bip = biphenylalanine; Cyx-A = cyclohexylalanine; Pyr-A = pyridyl-alanine.

Table 2. IC$_{50}$'s of Monovalent Analogues

compounds with $R_2 =$ Asp and $R_1 =$	IC$_{50} \pm$ SE A1-1 (μM)	IC$_{50} \pm$ SE P1-1 (μM)
Phe	2.9 ± 0.5	5.8 ± 1.5
Gly	9.6 ± 3	21 ± 5

Interestingly, none of the other mixtures showed significant inhibition for A1-1. In contrast, all mixtures showed 15−20% inhibition for GSTP1-1.

On the basis of these results, compounds with Asp at position R_2 and differing amino acids at R_1 were then synthesized individually using solid-phase synthesis. Table 1 reports the apparent K_D values determined for GSTA1-1 and GSTP1-1, based on IC$_{50}$ values and fitting to the eq 1 (Material and Methods). Modifications at R_1 did not alter the affinity significantly for A1-1. However, the presence of positively charged groups, such as Arg or Lys, resulted in 10-fold decrease in inhibition potency. A negatively charged group or a phenyl ring at R_1 results in much higher selectivity for GSTA1-1. Position R_2 is more sensitive to change than R_1 as there was a significant drop in affinity when Asp was replaced by Arg (Table 1). Hydrogen bond donors (Thr, Tyr) or small hydrophobic residues increased the potency toward P1-1. However, both positive and negative charge at this position resulted in considerable loss in activity for P1-1. Importantly, the monovalent analogues of the two most potent bivalent inhibitors ($R_2 =$ Asp and $R_1 =$ Phe or Gly) were synthesized and exhibited ∼100−150-fold decrease in their potency as compared to that of their bivalent counterparts (Table 2). In summary, the striking result of the initial screen was that negatively charged residues at position R_2 conferred significantly higher affinity and selectivity for GSTA1-1. Furthermore, bivalent inhibitors were consistently higher affinity than the monovalent paradigm analogues.
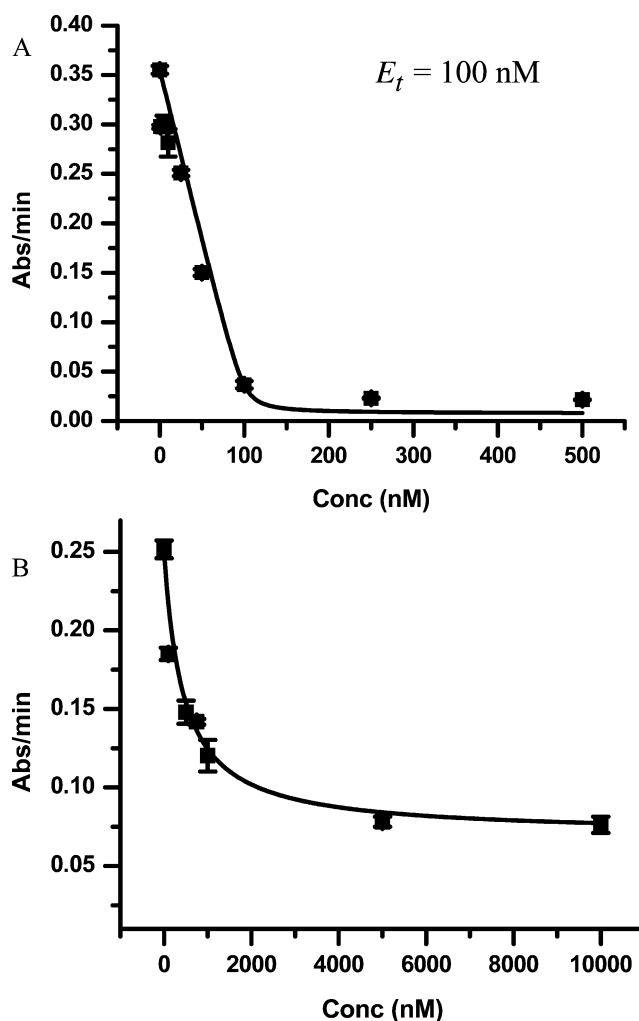
Figure 4. Binding curves obtained for inhibitors for A1-1 and P1-1 after fitting to the quadratic equation. Plot A represents the curve for the compound at total enzyme concentration, $E_t = 100$ nM. Complete saturation is obtained at 100 nM of inhibitor indicating 1:1 stoichiometry; χ^2 for fitting of A = 0.96, B = 0.92

Kinetics: The IC$_{50}$ values for all the compounds were determined with an enzyme concentration of 20 nM. This resulted in inhibitor depletion conditions for compounds with K_D's in the low nanomolar range (especially for A1-1). Under these circumstances, Michaelis−Menten kinetics fail to comply as they are based on the assumption that inhibitor remains in stoichiometric excess with respect to the protein. Such compounds are classified as "tight binding inhibitors", and their affinities (K_D's) must be determined with more complex quadratic equations (eq 1, Materials and Methods). However, for such tight binding compounds, one can readily determine the stoichiometry of binding as shown in Plot A, Figure 4. Plot A shows the data obtained at higher enzyme concentration ($E_t = 100$ nM) and fitted to the quadratic eq 1. According to the plot, at 100 nM of inhibitor, the enzyme is saturated, which suggests 1:1 stoichiometry for A1-1.

For P1-1 inhibition, the data resulted in biphasic kinetics for most of the compounds as shown in Plot B (Figure 4). Significant loss in activity was observed at low concentrations (20−500 nM); however, at higher concentrations (1000 nM to 20 μM), the rates did not decrease proportionally. It is striking and interesting that the first tight binding phase results in 50% inhibition, which may suggest single site occupancy. In fact, it

is possible that GSTP1-1 exists as a monomer−dimer equilibrium in solution and these inhibitors selectively inhibit either the monomer or dimer. This would lead to the "partial antagonist" behavior observed. The existence of monomer in solution has been suggested but not unambiguously demonstrated. Regardless, overall, much higher K_D's for P1-1 were obtained as compared to A1-1.

As noted above, EA has previously been shown to be a less potent inhibitor of both isoforms, with K_D values ∼1−5 μM. However, it was also necessary to consider the affinity of the linker itself, with no EA on the ends. Inhibition assays with Asp-Gly-AMAB-Gly-Asp, which is the linker contained within the bivalent inhibitor **1** used in subsequent studies, yielded negligible inhibition at 30 μM. This places a lower limit in its K_D nearly 3 orders of magnitude above the K_D of the bivalent analogue with the same linker, compound 1. Thus, both linker and EA moieties are required for high affinity.

We considered the possibility that the inhibitor could react either enzymatically or nonenzymatically with GSH and lead to time-dependent inhibitor concentrations. However, the nonenzymatic rate constant for formation of the Michael adduct from inhibitor 1 and GSH is very low, $1.2 \times 10^{-5} \mu$M^{-1} min^{-1}, and we observe no enzyme-dependent acceleration of this adduct formation (not shown). Under the assay conditions used to determine IC$_{50}$, and at the high range of inhibitor concentration where the reaction would be fastest (100 nM), the inhibitor concentration would only decrease by 0.4 nM during the assay. Furthermore, we explicitly looked by mass spectrometry for adduct formation with GSTP1-1, which is known to be adducted slowly by EA at Cys47. We observe small amounts of adduct (<10%) after 2−3 h of incubation with the inhibitor, but virtually no adduct formation under conditions identical to the assay conditions, during the time course of 2−3 min (not shown). Thus, although interesting side reactions occur with these inhibitors, they are insufficiently rapid to significantly alter the relative IC$_{50}$ values for GSTA1-1 versus GSTP1-1.

Binding Epitope of the Ligand by STD NMR: Binding epitopes of the inhibitors were analyzed using saturation transfer difference (STD).[20,21] STD NMR has been used as a powerful method for studying binding between ligand and receptor proteins, in which saturation of the protein is "transferred" differentially to the protons on the bound ligand, thereby identifying the most intimate ligand interactions with the protein. The application of the difference spectroscopy cancels out the signals of nonbinding building blocks of the ligand or nonligand molecules. We chose the ligand with R$_2$ = Asp and R$_1$ = Gly (**1**) for the STD NMR studies because this compound is readily soluble in the aqueous buffer used for the GST assay, without any requirement for additional solvents. In addition, the absence of a side chain on Gly allows for a direct focus on the binding affinity contributed by the Asp at R$_2$ position. The STD spectra of this compound with both GSTA1-1 and GSTP1-1 proteins show differential saturation transfer to most of the protons of the ligand molecule, which strongly supports the close contact of the entire ligand with the proteins. The 1D NMR spectrum of this compound and its STD spectrum in the presence of GSTA1-1 are shown in Figure 5B. The largest STD effect was

observed for the CH$_3$ group of the ethacrynic acid moiety, indicating that these protons have the most intimate contact with the surface of GST proteins. In contrast, STD spectra of the control compound containing Arg at position R$_2$ (R$_1$ is still Gly, **2**) resulted in no significant transfer of saturation with GSTA1-1, indicating that the positively charged Arg side chain does not have a significant residence time on the protein surface (data not shown).

In STD NMR, the protons of the ligand that are in closest contact with protons inside the protein-binding pocket are saturated to the highest degree. Thus, a quantitative analysis of STD effects allows for determination of the binding epitope on the ligand with atomic resolution.[22,23] The relative degree of saturation transfer for each individual proton was determined by the ratio of the integral peak intensity in the STD spectrum (I_{STD}) and the reference 1D NMR spectrum (I_0) (Figure 5A). For both GSTA1-1 and GSTP1-1, the CH$_3$ on the EA moiety of **1** displays the largest STD signal. Thus, the I_{STD}/I_0 of the methyl protons was set to 100%, and the I_{STD}/I_0 of other protons are normalized accordingly. Protons of the EA binding element display slightly higher relative saturation with GSTA1-1 than with GSTP1-1 but in a comparable range. However, the I_{STD}/I_0 ratios of protons on the central benzoic acid nucleus, and particularly the peptide linker arms, are considerably higher in the presence of GSTA1-1 than in the presence of GSTP1-1. In particular, for the αHs of both amino acids on the linker arms, the I_{STD}/I_0 values with GSTA1-1 are more than double compared to those with GSTP1-1. The β-carbon protons on Asp also exhibit greater saturation transfer on GSTA1-1 than on GSTP1-1. These results imply that the interactions between the negatively charged Asp side chains of the linker arms and the positively charged protein surface of GSTA1-1 contribute to a more intimate interaction between the entire linker and the protein. It is notable that the EA pharmacophore on the bivalent inhibitor exhibits very similar saturation transfer when bound to either protein. Thus, the linker is mainly responsible for the increased affinity for GSTA1-1 versus GSTP1-1.

Binding Epitope of the Protein by H/D Exchange Mass Spectrometry: Hydrogen/Deuterium exchange mass spectrometry (HDX MS) was used to compare the relative deuterium incorporation into the amide "backbone" of free and **1**-bound or EA-bound GSTA1-1.[24,25] The degree of HDX, especially during short incubations in D$_2$O, can be used to infer information about the dynamic environment of protein regions upon ligand binding. Pepsin digestion and peptide analysis were performed on the protein with and without deuterium labeling. For the deuterium labeling, the proteins were incubated for time periods ranging from 0.5 to 15 min in deuterated buffer followed by quenching and pepsin digestion. The sequences of the peptic fragments were obtained from tandem mass spectrometry (MS/MS) analysis of the digested nondeuterated protein, and 22 peptides were identified, which cover 89% of the entire backbone. The deuterated fragments were assigned to their nondeuterated precursors according to their charge and average mass (Figure 6A).

(20) Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108−6117.
(21) Claasen, B.; Axmann, M.; Meinecke, R.; Meyer, B. *J. Am. Chem. Soc.* **2005**, *127*, 916−919.

(22) Sandstrom, C.; Berteau, O.; Gemma, E.; Oscarson, S.; Kenne, L.; Gronenborn, A. M. *Biochemistry* **2004**, *43*, 13926−13931.
(23) Meinecke, R.; Meyer, B. *J. Med. Chem.* **2001**, *44*, 3059−3065.
(24) Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 1−25.
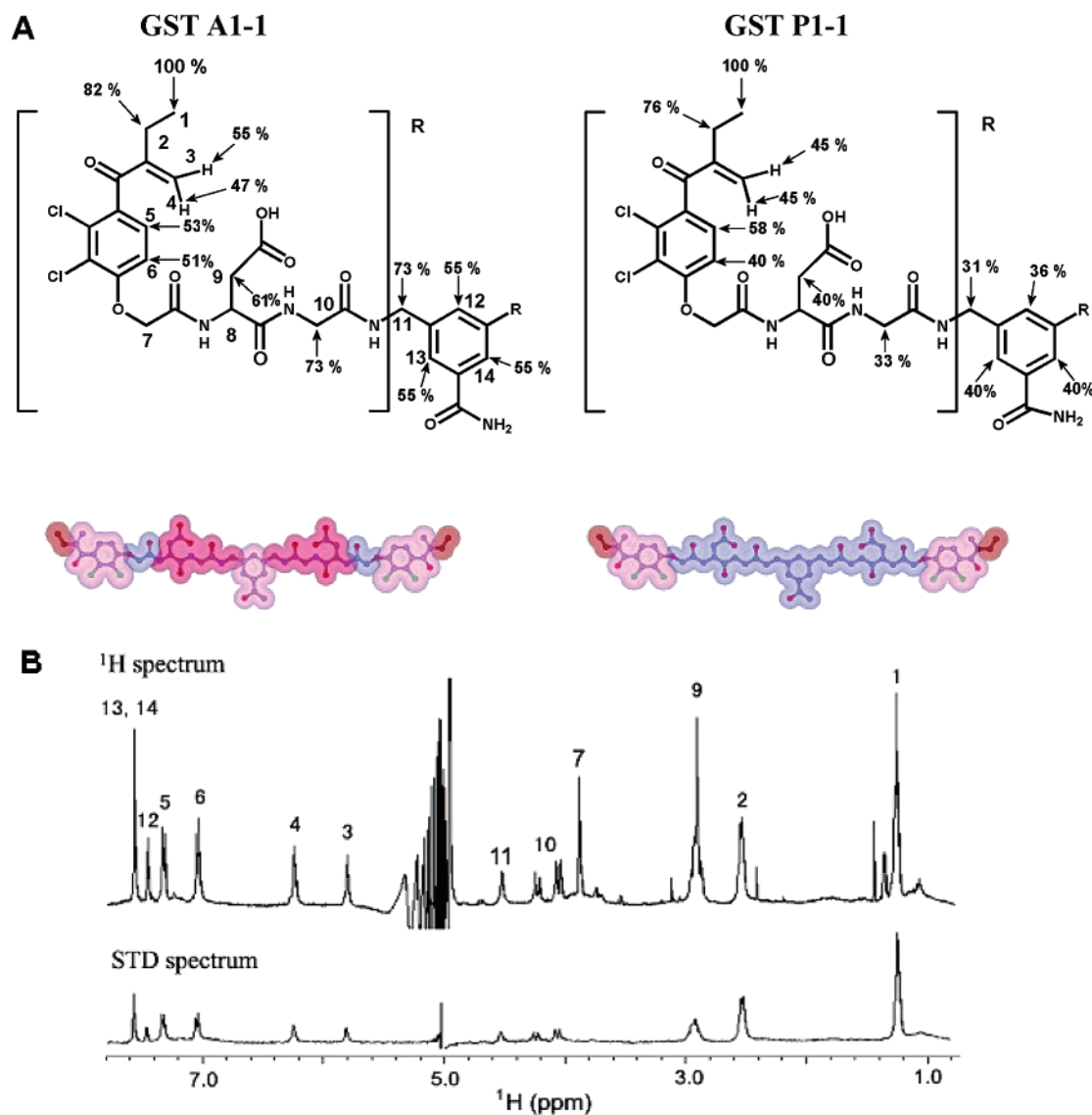(25) Eyles, S. J.; Kaltashov, I. A. *Methods* **2004**, *34*, 88−99.

Figure 5. (A) Top: The relative percentage of saturation (I_{STD}/I_0 %) of the individual protons of the ligand **1** in the presence of GSTA1-1 and GSTP1-1, respectively. The values are normalized to that of the CH_3 protons on the EA moiety, which has the highest I_{STD}/I_0 %. The I_{STD}/I_0 % for methylene protons at position 7 was not listed because their STD signals were too weak to be measured accurately. The molecular model of ligand **1** in solution shows that these two protons are sterically hindered from direct contact with the protein by the surrounding atoms (Supporting Information). The H8 peak is buried under the water signal, therefore it cannot be analyzed. Bottom: Graphic representation of the percentages shown above. Parts of the ligand are color coded according to the I_{STD}/I_0 %: red 75−100%, pink 60−75%, light pink 40−60%, blue <40%. (B) Top: 1D ^{1}H NMR spectrum of ligand (2 mM). Bottom: The STD spectrum in the presence of GSTA1-1 (0.005 mM). The unambiguous assignments were obtained from the 2D ROESY experiment (Supporting Information).

All of the labeled peptides were examined for dissimilarities in the deuterium incorporation among apo, **1**-bound, and EA-bound GSTA1-1. Most peptides exhibited minor differences or no differences in all three situations (Supporting Information). Six of the 22 peptides, including 2−16, 31−51, 106−123, 108−123, 110−123, and 216−222, displayed apparently less deuterium uptake when **1** or EA is bound. However, three more peptic fragments, 66−74, 49−79, and 95−102, exhibited significantly decreased deuterium incorporation for **1** but not EA (Figure 6A). The reduction of deuterium uptake reflects the effect of solvent protection by the ligand. The measured deuterium incorporation of the peptides at individual time points is plotted for three fragments, 66−74, 216−222, and 124−136, representing each of the situations mentioned above (Figure 6B): 66−74 is specifically protected by **1**; 216−22 is protected by both EA and 1; 124−136 is protected by neither. It has to

be noted that this technique characterizes each peptide as a whole, and the specific locations of deuterium within the peptide cannot be determined from peptide mass alone. Compared to **1**, free EA has a more limited effect on solvent protection shown by a smaller magnitude of $\Delta\Delta D$ (where ΔD is the difference in the number of deuterons exchanged and $\Delta\Delta D = \Delta D_{apo} - \Delta D_{ligand}$). Also, unlike ligand **1**, most of the differences caused by EA binding are negligible after 15 min of exchange (Supporting Information).

A structural model of the GSTA1-1/ligand **1** complex was constructed based on the GSTA1-1/EA crystal structure, using DS ViewerPro software. Free energy minimization of the isolated ligand **1** with a generic force field, DREIDING, resulted in a U-shaped molecule, which is docked into the GSTA1-1 binding site by tethering the four chlorine atoms of the two EA moieties of **1** with those from the two free EA molecules at the
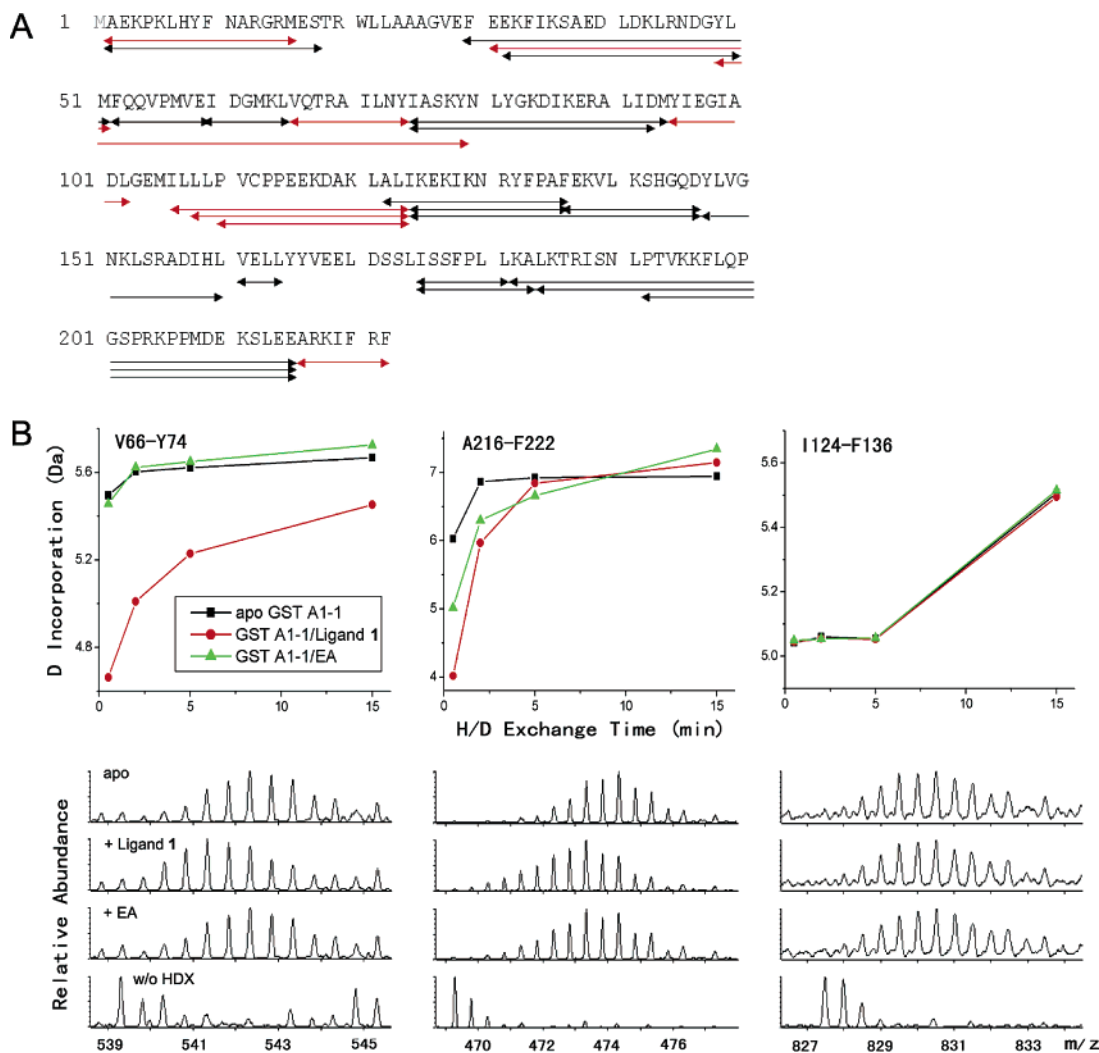
A

```
  1   MAEKPKLHYF NARGRMESTR WLLAAAGVEF EEKFIKSAED LDKLRNDGYL

 51   MFQQVPMVEI DGMKLVQTRA ILNYIASKYN LYGKDIKERA LIDMYIEGIA

101   DLGEMILLLP VCPPEEKDAK LALIKEKIKN RYFPAFEKVL KSHGQDYLVG

151   NKLSRADIHL VELLYYVEEL DSSLISSFPL LKALKTRISN LPTVKKFLQP

201   GSPRKPPMDE KSLEEARKIF RF
```

B

Figure 6. (A) Identified GSTA1-1 fragments from pepsin digestion. The sequences that exhibit reduced deuterium uptake upon binding of ligand **1** are highlighted in red. The Met1 is absent in the protein sample we used for studies. (B) Representative peptic fragments which had significant (left and middle) or no difference (right) of deuterium incorporation upon ligand binding. Peptide 216−222 showing reduced deuterium uptake of either ligand **1** or EA was bound, but the difference was observed for peptide 66−74 only when ligand **1** was bound. Top: Deuterium incorporation of apo (black squares), **1**-bound (red circles), and EA-bound (green triangles) GSTA1-1 after 0.5, 2, 5, and 30 min of exchange. Bottom: The mass spectra of the +2 charge states of the peptides after 0.5 min of HDX and without HDX. The mass and charge state of the nondeuterated fragments were used for the assignment of the deuterium-labeled peptides.

relevant positions. The peptide fragments that displayed significant protection from deuterium incorporation are highlighted in the model. The common regions observed for both **1** and EA (cyan) nicely outline a binding pocket around the EA molecule, where the EA elements of the bivalent ligand are also located. From the structural model, the linker of **1** is suggested to make additional contact with the bottom of the intersubunit cleft. On the basis of the HDX MS data, residues 65−73 and 94−101 highlighted in magenta and located at the bottom of the cleft are protected from solvent (Figure 7). These two sequences are located at the bottom region of the cleft, at the interface of the two GST subunits. The data suggested that the bivalent ligand protects the interface from solvent.

Discussion

Glutathione S-transferases (GSTs) belong to a family of Phase II detoxification enzymes, which catalyze the conjugation of the nucleophilic tripeptide glutathione (GSH) with endogenous and exogenous electrophiles.[26] Apart from the catalytic role involving GSH conjugation, GSTs have been implicated also in regulation of apoptosis, cell proliferation, and oxidative stress.[10]

Considerable effort has been directed toward identifying isoform-specific GST inhibitors. Examples include peptide or peptidomimetic analogues of GSH, GSH S-conjugates, and a wide range of nonpeptides, such as ethacrynic acid.[27,28] Our lab recently reported a novel class of GST inhibitors based on multivalency.[9] The design of these compounds included established GST inhibitors or substrates attached at both ends of linkers with varying lengths. The compounds were designed so that the linker would span the solvent-accessible intersubunit cleft and the binding element, or "pharmacophore", of the inhibitor would interact with the active sites on each GST

(26) Armstrong, R. N. *Chem. Res. Toxicol.* **1997**, *10*, 2−18.
(27) Burg, D.; Mulder, G. J. *Drug Metab. Rev.* **2002**, *34*, 821−863.
(28) Lyttle, M. H.; Hocker, M. D.; Hui, H. C.; Caldwell, C. G.; Aaron, D. T.; Engqvist-Goldstein, A.; Flatgaard, J. E.; Bauer, K. E. *J. Med. Chem.* **1994**, *37*, 189−194.
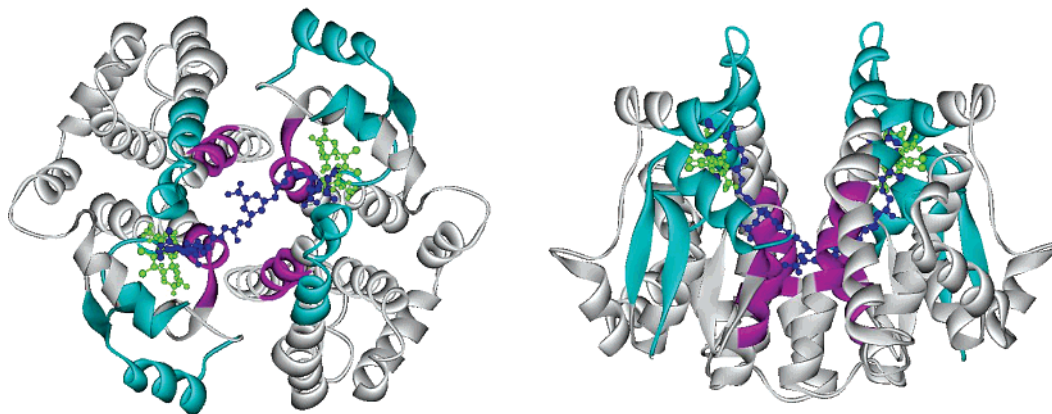
Figure 7. Ribbon diagram of regions of GSTA1-1 that displayed significant changes of deuterium incorporation upon binding of **1** or EA. The crystal structure of EA-bound GSTA1-1 was used, and the EA molecules are shown in green. **1** (blue) is docked into the structure using DS ViewerPro software. Common peptides exhibiting decreased deuterium exchange in the presence of either EA or **1** are highlighted in cyan (2−16, 31−51, 106−123, and 216−222). Peptides exhibiting decreased exchange only in the presence of **1** are in magenta (65−73 and 95−102). Peptic fragment 49−79, which was also observed to have reduced deuterium uptake only upon binding of ligand **1**, covers a long sequence, but the $\Delta\Delta D_{49-79}$ is contributed from $\Delta\Delta D_{65-73}$ and part of $\Delta\Delta D_{31-51}$ (Figure 6A and Supporting Information). There were no peptides that exhibited decreased exchange in the presence of EA but not in the presence of **1**. Apparently, inhibitor **1** protects the cleft region from solvent exchange.

monomer simultaneously. These inhibitors bind with higher affinity than their monovalent analogues, and they also exhibited some selectivity depending on the binding elements. Theoretically, the increase in affinity of multivalent ligands can be attributed to entropic cooperativity, as the entropy cost associated with binding of the second half of the inhibitor is decreased after the first half binds.[6] Interestingly, calorimetric studies with the bivalent GST inhibitors showed that the entropic gain observed is only partially responsible for the increased affinity of compounds. In fact, enthalpic interactions contributed significantly to the differential affinity of bivalent versus monovalent analogues. In our ongoing efforts to develop selective inhibitors, we have focused on the linker of these bivalent compounds. We reasoned that, apart from increasing binding affinity by virtue of higher enthalpic contributions from the functional groups on the linkers, we might also incur selectivity for GST isoforms. The approach of linker optimization via combinatorial methods could be generally useful in the design of multivalent inhibitors, but it has not been exploited previously.

Screening of the library identified two pools with potent inhibition of A1-1, R_2 = Asp or Glu. In contrast, each of these pools afforded similar inhibition of GSTP1-1. This result is consistent with the relative positive charge in the cleft region for A1-1 versus P1-1. As noted in Figure 1, an electrostatic potential surface map shows a striking difference between the charge distribution of A1-1 and P1-1. Presumably, the Asp/Glu at R_2 interacts with a residue in this part of the cleft resulting in enhanced selectivity for A1-1. This is supported by results with compounds that have R_2 = Arg and R_1= Arg or Gly. An 80-fold decrease in binding affinity is observed when R_2 = Arg. It is also interesting that changes in R_1 did not result in significant differences in affinities for A1-1, as all of the compounds with R_2 = Asp show low nanomolar K_D's. Apparently, the binding cleft is more promiscuous toward the C_2 axis of symmetry. Interestingly, the changes at R_1 did influence the binding toward P1-1 to some extent. Charged side chain residues on either of two positions results in dramatic loss in affinity for P1-1 and better selectivity for A1-1. Small hydrophobic residues, especially with hydrogen bond donating side chains,

tend to increase affinity toward P1-1, which might be due to the uncharged cleft surface. To determine whether bivalency is an advantage with this new class of inhibitors, the monovalent analogues were synthesized and tested. Both the compounds showed much lower affinity than the bivalent compounds.

A critical aspect of multivalency as a design principle is the need to experimentally verify the stoichiometry of binding. For A1-1, it was determined to be 1:1, suggesting that these inhibitors likely are binding in a bivalent manner, with one molecule per GST dimer. In the case of P1-1, the data interpretation was not straightforward as the binding curves were biphasic. Such behavior has previously been reported for P1-1 and estrogen-based inhibitors.[29] At this time, we do not know the molecular basis for this behavior.

The STD NMR provided useful information regarding the binding epitope of these compounds. That the STD effect was observed for all parts of the inhibitor **1** suggests that the entire molecule comes in close contact with the protein, which in turn is indicative of a bivalent binding mode. The lack of STD effect on the linker region of inhibitor 1 when complexed with GSTP1-1 emphasizes the importance of the linker in high affinity for GSTA1-1. The fact that no STD transfer was observed when Asp is replaced by Arg at R_2 emphasizes that there is likely a charge−charge interaction between the side chain and cleft surfaces. HDX MS data gave additional support for the contact between the entire bivalent ligand **1** with the intersubunit cleft region.

A plausible structural model for the inhibitor complex, consistent with the experimental data, was constructed by docking. The free energy of the docked GSTA1-1/**1** complex was optimized using a DREIDING force field minimization of the DS ViewerPro software. Inclusion of the ligand has negligible effect on the overall structure of the protein. Figure 8 illustrates all of the atoms within 5 Å of inhibitor **1** and within 5 Å of monovalent EA. All these atoms are located within the regions that were demonstrated to have reduced solvent accessibility by HDX MS. Notably, R69, I96, and E97, which are within 5 Å of ligand **1**, but not EA, have decreased deuterium

(29) Abel, E. L.; Lyon, R. P.; Bammler, T. K.; Verlinde, C. L.; Lau, S. S.; Monks, T. J.; Eaton, D. L. *Chem. Biol. Interact.* **2004**, *151*, 21−32.
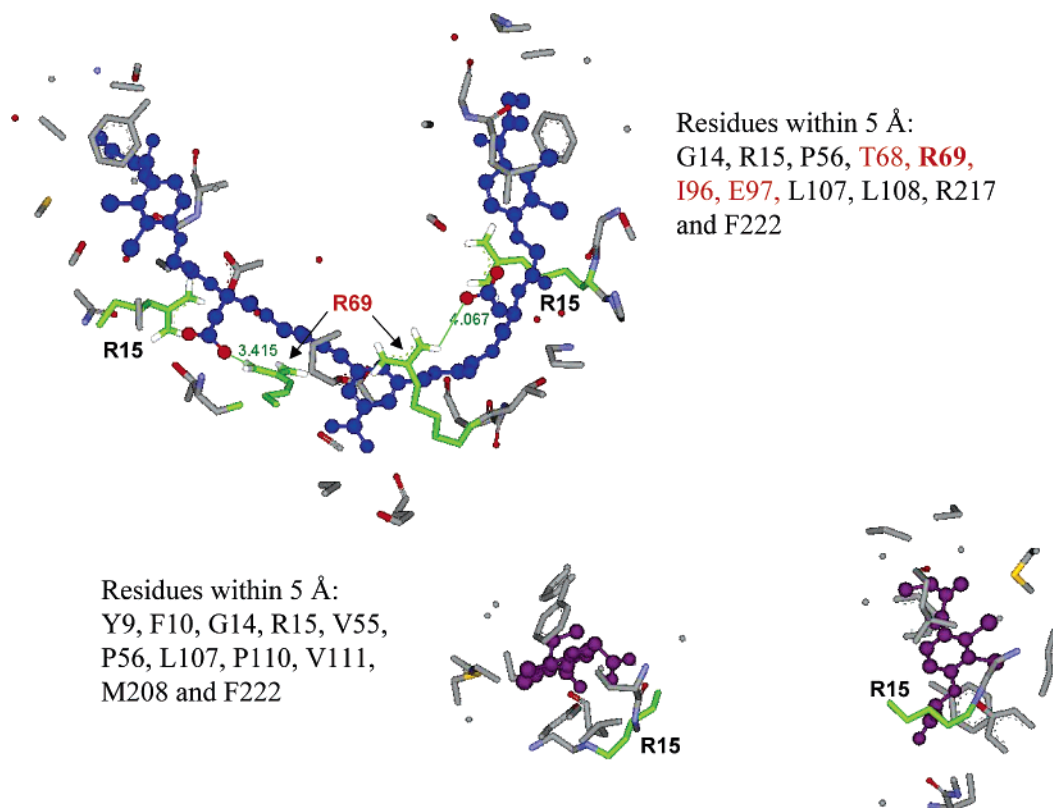
Figure 8. Structural models of atoms within 5 Å of ligand **1** (blue, top) and EA (purple, bottom). All the atoms belong to the residues located in the regions with reduced deuterium uptake in the HDX MS experiments. T68, R69, I96, and E97, which are close in space with **1** but not EA, are within the peptic fragments 66−74 and 95−102, where deuterium uptake differences were seen between apo and **1**-bound but not EA-bound GSTA1-1. The carboxyl oxygen atoms of the Asp side chains of ligand **1** are shown in red, and atoms from R15 and R69 are highlighted in green.

uptake only for the **1**-bound GSTA1-1. In the model, the distance between the guanidine N−H of the protein R69 and the CO_2^- of the ligand Asp side chain was measured to be 3.4 and 4.1 Å for the two subunits, respectively, indicating possible salt bridges between ligand **1** and GSTA1-1. This potential electrostatic interaction may contribute to the binding affinity and specificity observed.

In our model, both EA moieties of ligand **1** occupy the same site of free EA in the crystal structure but with different orientation. This is not surprising, as it is known that EA binds in multiple orientations in GSTA1-1.[30] In addition to R69, the positive guanidine group of R15 was also shown to be close in space with the Asp side chain of ligand **1** (Figure 8, top). Together these Arg residues form a positively charged binding environment at the GSTA1-1 cleft surface. On the basis of the H/D exchange, the STD analysis, and the modeling, we propose that charge−charge interactions of the linker assist in anchoring these inhibitors into the cleft region.

In conclusion, we have shown how linkers in multivalent ligands can be exploited to increase both affinity and selectivity toward the target molecules. Such a strategy could be useful in drug discovery research wherein the goal is often to either identify a lead compound or modify a lead compound for a more potent and selective molecule. Additionally, our compounds, despite being high molecular weight, possess much better solubility than EA. The compounds with R_1 = Asp, Arg, Gly, and Glu are readily soluble in water. Hence, it is

conceivable that the linker of bivalent drugs can be used to improve the physiochemical properties, as well as improve affinity and selectivity.

Materials and Methods

3,5-Dimethyl benzoic acid ester was purchased from Chem-Impex International. The resin Tentagel S AM was bought from Advanced Chemtech. F-moc-protected amino acids were purchased form Novabiochem. All other coupling reagents and solvents were obtained from Sigma Aldrich. GSTA1-1 and P1-1 were both expressed and purified using a previously reported procedure.[31] Analytical HPLC analyses were done on a Waters 600 HPLC (Phenomenex Jupiter C_{18} column, 4.6 × 250 mm) using a linear gradient of acetonitrile (5−95% over 30 min) at a flow rate of 1 mL/min. LC/MS was recorded on Waters 2690 separation module coupled with Micromass Platform LCZ (ESI-MS) using Agilent Zorbax SB-C_{18} column (5 μM, 2.1 × 150 mm) at a flow rate of 0.25 mL/min. MALDI-MS was carried out on Agilent Technologies's LC-MSD-Trap-XCT with built in MALDI source.

Synthesis of Compounds. Solution Phase: 3,5-Dimethyl amino benzoic acid ester was synthesized starting from 3,5-dimethyl benzoic acid ester using the procedure reported by Vladimir et al.[32] The corresponding acid was obtained by refluxing the ester using 2 N HCl. The dihydrochloride salt obtained was then protected using Fmoc succinimide.

Library Synthesis Using the Split-Pool Method and Iterative Deconvolution. Resin Tentagel SAM was used to synthesize the library. F-moc-protected 3,5-dimethyl benzoic acid was loaded on the resin using standard peptide coupling procedures (Figure 2). Twenty different

(30) Oakley, A. J.; Lo Bello, M.; Mazzetti, A. P.; Federici, G.; Parker, M. W. *FEBS Lett.* **1997**, *419*, 32−36.

(31) Ibarra, C.; Nieslanik, B. S.; Atkins, W. M. *Biochemistry* **2001**, *40*, 10614− 10624.

(32) Vladimir, V. M.; Laszlo, L.; Keana, J. *Org. Prep. Proc. Int.* **1995**, *27*, 117−120.

residues, including some unnatural amino acids, were selected to be incorporated in the linker. The resin was then split into 20 different tubes to be coupled with the amino acids. The completion of coupling was monitored by the Kaiser color test for unreacted amines. Coupling times varied depending on the amino acid, but on average it was 2 h. The resin from 20 different tubes was then pooled in a single column and mixed. The resin was again divided into 20 different tubes to be coupled individually with 20 amino acids. Since only two cycles were needed, coupling with ethacrynic acid (EA) was carried out for 2 h. The total number of compounds = (number of amino acids)^cycles of coupling = 20^2 = 400. Iterative deconvolution was used to identify the lead compound from the mixtures.[33] After coupling with ethacrynic acid, the resin was not pooled, but instead, it was retained in the individual tube with known amino acid at R_2. Each mixture was screened for inhibition using the standard CDNB assay for GSTs.[34] From these mixtures, the preferred amino acids at the R_2 position were determined. All 20 compounds in that mixture were then synthesized separately using the same procedure with fixed R_2 to identify the best amino acid at R_1 (average yield = 80%).

Characterization of Library Mixtures and Individual Compounds. Compounds were cleaved from the resin using TFA:H_2O: trimethylsilane 95:5:5. The solution was evaporated, and the residue was then precipitated with cold diethyl ether. The precipitate was then centrifuged and the pellet dried. Small amounts of each mixture were taken for characterization by LC/MS. HPLC showed approximately 20−25 peaks of more or less similar intensity (Figure 3). MALDI-MS yielded a mass envelope in the expected range of the all the compounds in the mixture (Figure 3 inset). In LC/MS, at least 15 of the possible 20 compounds in each mixture could be assigned based on retention times and MH$^+$ (data not shown). Some compounds in the library were not identified unambiguously due to overlapping retention times, MWs, and also due to the presence of several isotopic peaks of a single compound. All the compounds listed in Table 1 were characterized similarly using MS and HPLC.

Screening and Determination of K_D's. The residue obtained from each flask was dissolved in DMSO. It was diluted further (1:100) for screening in the CDNB assay. The total inhibitor concentration in each mixture was ∼3 μM (final assay concentration ∼ 150 nM), and the concentrations of GSH and CDNB were at their respective K_M values (A1-1 = 350 and 720 μM, P1-1 = 150 and 820 μM). The GST concentration was 20 nM. Rates of product formation were obtained by measuring absorption at 340 nM for 1 min on a Beckmann DU 7400 spectrophotometer. Inhibitor concentrations spanned 4 orders of magnitude, and experiments were carried out in triplicate. Initially, data were fit to a sigmoidal dose−response equation using Graphpad Prism to determine the 50% inhibitory concentration. However, for A1-1, the compounds exhibited tight binding and hence were fit to quadratic eq 1 to calculate K_D's. To estimate stoichiometry of binding with GSTA1-1, the CDNB assay was repeated for the compound with R_1 = Phe and R_2 = Asp (K_D = 10 nM) but with 10-fold higher enzyme concentration (E_t = 100 nM). The data obtained were fit to eq 1.

$$[ES] = \frac{([E]_t + [S]_t + K_D) - \sqrt{([E]_t + [S]_t + K_D)^2 - 4[E]_t[S]_t}}{2}$$ (1)

NMR Experiments. The NMR data were recorded at 295 K on a Varian Inova 500 MHz spectrometer equipped with a 5 mm inverse triple-resonance/pulse field gradient (PFG) probe. A basic saturation transfer difference (STD) sequence was used, with the on-resonance irradiation performed at a chemical shift of −1.5 ppm. The off-resonance irradiation was set at 40 ppm, where no protein signals are present. A train of 40 Gaussian-shaped pulses of 50 ms length separated by a 1 ms delay was employed, leading to a total saturation time of

2.04 s. The spectra were subtracted internally via phase cycling after every scan using different buffers for on- and off-resonance. Total scan number was 1024. The NMR samples of the ligands were prepared as 2 mM solutions in 10 mM sodium phosphate buffered D_2O containing 150 mM NaCl. The STD spectra were obtained with the presence of 0.005 mM of relevant proteins (ligand/protein molar ratio = 400:1) (pD = 7.8; the pD value was obtained by an addition of 0.4 units to the readings of the pH meter with a glass electrode in a deuterated solution).[35]

Hydrogen Exchange and Pepsin Digestion. Hydrogen exchange experiments were performed on the free enzyme or on the ligand-bound complex. Deuterium exchange was initiated by a 20-fold dilution of the enzyme with 10 mM Tris-DCl/D_2O buffer (pD 7) at 22 °C. The pD measurements are given as the values read from the pH meter with no adjustment for isotope effects. The exchange was quenched after 0.5, 2, 5, and 15 min by mixing aliquots of the incubation mixture with equal volume of ice-cold 50 mM potassium phosphate−sodium citrate buffer (pH 2.2) to drop the pH to 2.5. The quenched mixture was allowed to equilibrate on ice for 30 s, 1 μL of ice-cold pepsin solution (3660 units/mg solid, 24 mg/mL in H_2O) was added to 20 μL of the protein solution. The digestion proceeds for 1 min on ice, and the resulting peptides were injected immediately to the HPLC/MS system. The HPLC column and the entire injector assembly were packed in ice to minimize back exchange during analysis. The masses of resulting peptides were determined by LC/MS without correction of the back exchange because the experiments were performed and analyzed under identical conditions. Pepsin is a nonspecific protease, meaning its cleavages cannot be predicted from the sequence alone, but the cleavage sites for a given protein can be reproduced under identical digestion conditions. [36]

Electrospray Ionization Mass Spectrometry. Electrospray mass spectra (ESI-MS) of GSTA1-1 were recorded on a quadrupole/time-of-flight mass spectrometer−QTOF (Micromass, Manchester, UK). Instrument settings were the following: electrospray voltage 3.8 kV, extraction cone 1 V, cone voltage 65 V, and source temperature 100 °C. Data acquisition was carried out from m/z 800 to 2400 using a scan time of 2.4 s. Protein/peptide samples, with a concentration range of 1−5 μM, were injected using an 8125 injector (Rheodyne, Rohnert Park, CA) on a perfusion chromatography column (0.5 mm i.d. × 5 cm) packed in-house with 30 μm POROS R2 particles (PerSeptive Biosystems, Framingham, MA). A Shimadzu LC10AD solvent delivery module (Shimadzu Scientific Instruments, Columbia, MD) was used to deliver isocratically the mobile phase containing 20% solvent A (5% acetonitrile:2-propanol = 2:1, 0.1% TFA) and 80% solvent B (90% acetonitrile:2-propanol = 2:1 mixture, 0.1% TFA) at a flow rate of 4 μL/min. During the entire LC/MS analysis, the injector, column, solvent bottles, and all the connecting tubing were kept in ice-cold water to minimize H/D exchange. The zero-charge deconvoluted mass spectrum of GSTA1-1 was generated using the MaxEnt1 deconvolution program incorporated in MassLynx 4.0 data acquisition software (Waters, Milford, MA).

Nondeuterated pepsin-digested samples were sequenced by tandem mass spectrometry on a QTOF mass spectrometer using the following parameters: electrospray potential 3.5 kV, cone voltage 32 V, the extraction cone 1 V, and source temperature 100 °C. The MS survey scan was m/z 400−1600 with a scan time of 1 s and the collision energy set to 5 eV. For operation in the MS/MS mode, the scan time was increased to 2 s, and the isolation width was set to include the full isotopic distribution of each precursor (4 Da mass window). Doubly, triply, and quadruply protonated peptide ions selected by the data-dependent software were subjected to collision-induced dissociation (CID) using appropriate collision energies (16−40 eV). The GSTA1-1 pepsin digest was separated on a reversed phase Atlantis column (0.3

(33) Hougten, R. *J. Med. Chem.* **1999**, *42*, 3743−3778.
(34) Lyon, R. P.; Atkins, W. M. *Biochemistry* **2002**, *41*, 10920−10927.

(35) Glasoe, P. K.; Long, F. A. *J. Phys. Chem.* **1960**, *64*, 188−190.
(36) Resing, K. A.; Hoofnagle, A. N.; Ahn, N. G. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 685−702.

mm i.d. × 15 cm), packed with 3 μm C$_{18}$ particles (Waters, Milford, MA) operated at a flow rate of 3 μL/min. The gradient consisted of 5−50% solvent B for 45 min, followed by 50% B for 10 min (A = 5% acetonitrile, 0.1% TFA; B= 95% acetonitrile, 0.1% TFA). Upon the completion of the LC/MS/MS run, the MS/MS spectra were searched against the nonredundant NCBI protein database using MASCOT (Matrix Science, UK).

To minimize back exchange, the deuterated pepsin-digested samples were analyzed by LC/ESI-MS using the isocratic chromatographic conditions described above, which has a total elution time of ∼5 min. The deuterium incorporation of each peptic peptide was calculated by subtracting the average mass of the nondeuterated form from the centroid mass of the isotope-labeled form. Two data sets were averaged for each experiment.

Molecular Modeling. GRASP2 was used to determine the electrostatic potential surface map of A1-1 and P1-1. To generate the model in Figure 8, the 1D structure of ligand **1** was imported into DS VIEWER PRO 6.0. The software's Dreiding force field was used to generate the low energy conformation of the compound.[37] The crystal structure of GSTA1-1/EA complex (resolution = 2.70 Å, PDB entry 1GSF) was then used to dock the minimized structure. Tethers were created between chlorine atoms of EA bound to GST and the EA moiety of the bivalent ligand. Docking was then carried out using a molecular overlay module of the program, and the overall energy of the GSTA1-1/**1** was minimized using Dreiding force field for 10 000 iterations with the inclusion of intermolecular forces.

Supporting Information Available: ROESY spectrum for assignment of ligand 1, HDX data for peptides with less or no difference in exchange are included. This material is available free of charge via the Internet at http://pubs.acs.org.

JA061766N

(37) Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III. *J. Phys. Chem.* **1990**, *94*, 8897−8909.