Bivalent Enzyme Inhibitors Discovered Using Dynamic Covalent Chemistry

This is the peer-reviewed version of the following article:

Clipson, A. J., Bhat, V. T., McNae, I., Caniard, A. M., Campopiano, D. J., & Greaney, M. F. (2012). Bivalent Enzyme Inhibitors Discovered Using Dynamic Covalent Chemistry. *Chemistry - A European Journal*, 18(34), 10562-10570.

which has been published in final form at http://dx.doi.org/10.1002/chem.201201507

Manuscript received: 30/04/2012; Article published: 10/07/2012

Bivalent Enzyme Inhibitors Discovered Using Dynamic Covalent Chemistry**

Alexandra J. Clipson,[2] Venugopal T. Bhat,[1] Iain McNae,[2] Anne M. Caniard,[2]
Dominic J. Campopiano[2] and Prof. Michael F. Greaney[1,*]

[1]School of Chemistry, University of Manchester, Oxford Rd., Manchester, M13 9PL, UK.

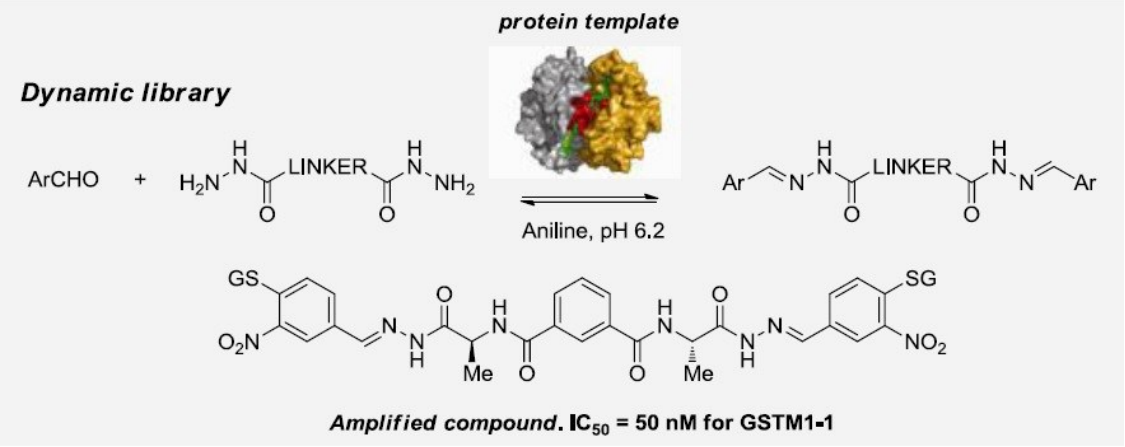
[2]EaStCHEM, School of Chemistry, Joseph Black Building, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JJ, UK.

[*]Corresponding author; e-mail: Michael.Greaney@manchester.ac.uk

[**]We thank CRUK, EaStCHEM and the EPSRC for funding (studentship for A.C., scholarships to V.T.B. and A.M.C., leadership fellowship to M.F.G., respectively) and the EPSRC mass spectrometry service at the University of Swansea. Dr. Elizabeth Blackburn is thanked for assistance with ITC measurements and Dr. Scott Baxter for assistance with protein purification.

Supporting information:
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201201507

Graphical abstract:



Keywords:
drug design; dynamic covalent chemistry; enzymes; hydrazones; inhibitors

Abstract
A bivalent dynamic covalent chemistry (DCC) system has been designed to selectively target members of the homodimeric glutathione-*S*-transferase (GST) enzyme family. The dynamic covalent libraries (DCLs) use aniline-catalysed acylhydrazone exchange between bivalent hydrazides and glutathione-conjugated aldehydes and the bis-hydrazides act as linkers to bridge between each glutathione binding site. The resultant DCLs were found to be compatible and highly responsive to templating with different GST isozymes, with the best results coming from the M and *Schistosoma japonicum* (Sj) class of GSTs, targets in cancer and tropical disease, respectively. The approach yielded compounds with selective, nanomolar affinity (K_i=61 nm for mGSTM1-1) and demonstrates that DCC can be used to simultaneously interrogate binding sites on different subunits of a dimeric protein.

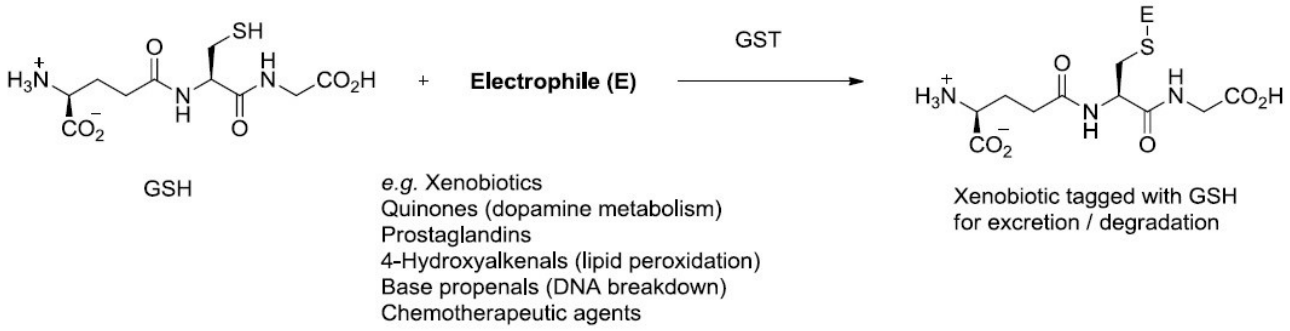
Introduction

The use of dynamic covalent chemistry (DCC) to discover new binding motifs for protein targets is a growing area of research.[1, 2] The DCC experiment involves an assembly of small molecules, termed a dynamic combinatorial library (DCL), in dynamic exchange through a reversible chemical reaction. In the presence of a target, such as a protein, the composition of the DCL will adjust to minimise the global free energy of the system. With suitable design parameters,[3] this adaptive behaviour can favour the strongest binding protein component(s) at the expense of the weaker ones. The approach is thus distinct from conventional structure-based drug design (SBDD) in that it requires no structural information and the synthesis event is accompanied by a qualitative binding assay.

Pioneering work in biological DCC has featured protein targets that were necessarily well understood in terms of medicinal chemistry structure–activity relationships (SAR), with DCLs producing hit structures largely comprising known binding motifs.[4] This was required to correlate the observed behaviour of DCLs in terms of amplifications with binding affinities. The future development of protein-directed DCC will depend on exploiting its complementarity to conventional methods,[5] and applying it to targets that would challenge standard drug discovery approaches. We describe one such approach here, in which we use DCC to discover selective bivalent inhibitors of glutathione-*S*-transferase (GST) isozymes, including the most potent inhibitor to date for the M-class of this enzyme.

The GSTs are a superfamily of enzymes responsible for conjugating the cellular thiol glutathione (GSH) to an array of endogenous and xenobiotic electrophiles (Scheme 1).[6] The majority of GSTs exist as homodimers containing a conserved GSH binding site and a hydrophobic substrate binding region at the interface of the two monomers (Figure 1). As GSH conjugation is central to phase II metabolism, modulation of GST isozyme activity has been identified in a range of therapeutic areas, for example, oncology (multidrug resistance and

cell proliferation),[7] asthma[8] and inflammation.[9] Parasite GSTs are also drug targets in tropical diseases, such as schistosomiasis and malaria, where they play a pivotal role in the life cycle of the parasite.[10]



Scheme 1. GSTs and GSH conjugation. Electrophiles include, for example, xenobiotics, quinones (dopamine metabolism), prostaglandins, 4-hydroxyalkenals (lipid peroxidation), base propenals (DNA breakdown), chemotherapeutic agents.

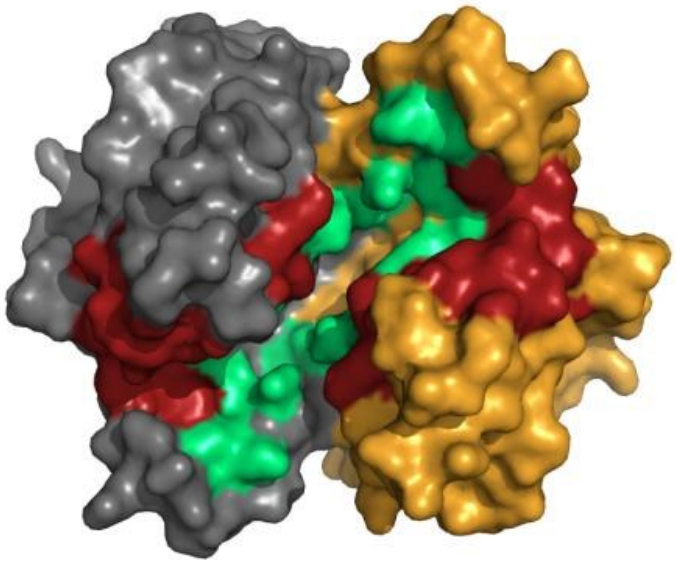
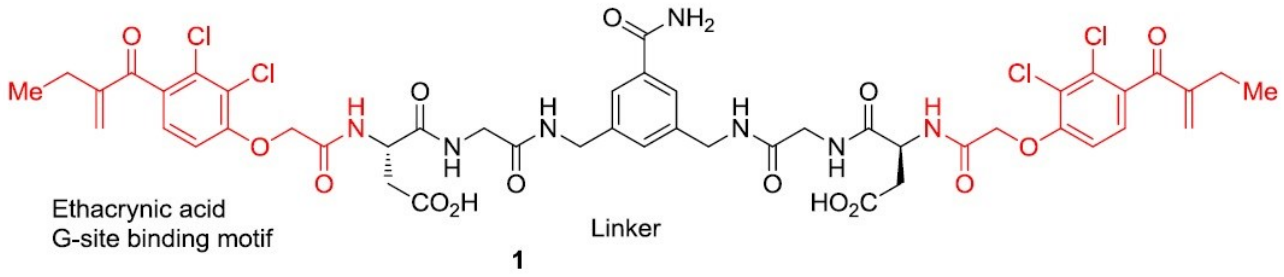


Figure 1. Structure of hGSTM1-1 that illustrates the H- and the G-sites (grey: monomer 1; yellow: monomer 2; green: G-site; red: H-site); PDB ID: 1XWK; the Figure was generated by using PyMOL.

The development of GST inhibitors is hampered by the plasticity of the H-site, which is large and has evolved to accept a range of small molecules in a variety of different conformations.[11] Thus, despite the GSTs being amenable to X-ray crystallography, conventional SBDD methods have yielded relatively little in the way of potent, isoform-specific inhibitors. A notable exception is Atkins' reports of bivalent inhibitors for the GSTα and π classes.[12]

As part of a multivalent strategy, non-selective GST inhibitors, such as ethacrynic acid, were appended to either end of a peptidic linker designed to span across the dimer interface. Using molecular modelling and split-and-pool combinatorial chemistry, the linker could be optimised to afford potent inhibitors with excellent selectivity between the GSTα and π isozymes (Figure 2). We have previously studied GST inhibition by DCC and discovered selective, monovalent G-site binders with low micromolar affinity.[13] A DCC approach based around a bivalent strategy has exciting potential both for the discovery of more potent and selective GST inhibitors, as well as the development of new concepts for protein-directed DCC in general. The DCL would produce bivalent molecules that span relatively large surface areas across two protein subunits—a new development that contrasts with existing DCL systems directed by proteins. The DCL design would feature linking molecules and G-site binders, affording the opportunity to discover novel combinations with immediate qualitative information on their binding affinity.



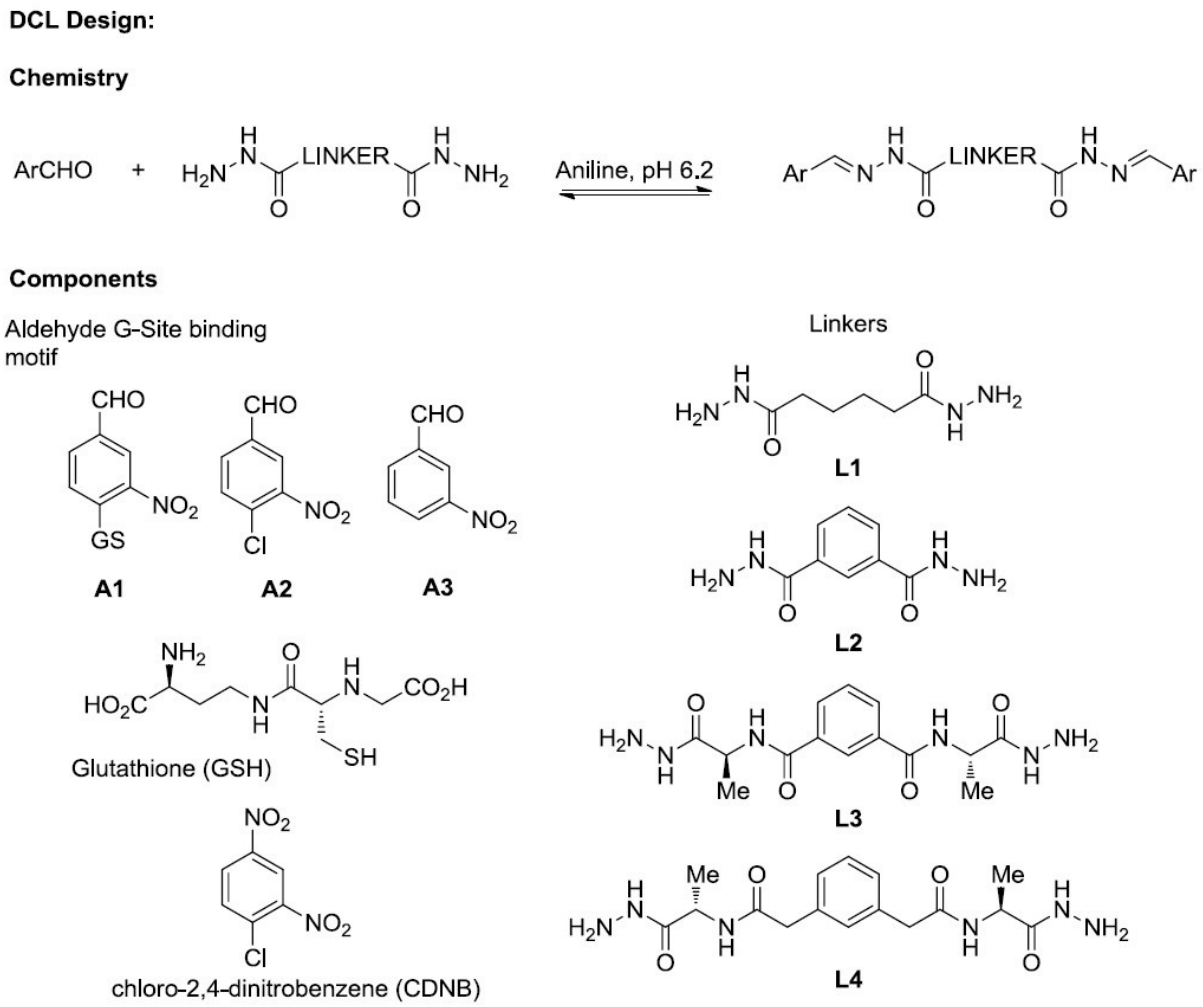
1

Figure 2. Atkins' bivalent GSTA1-1 inhibitor, 1; K_D=42 nm with 100-fold selectivity over GSTP1-1.

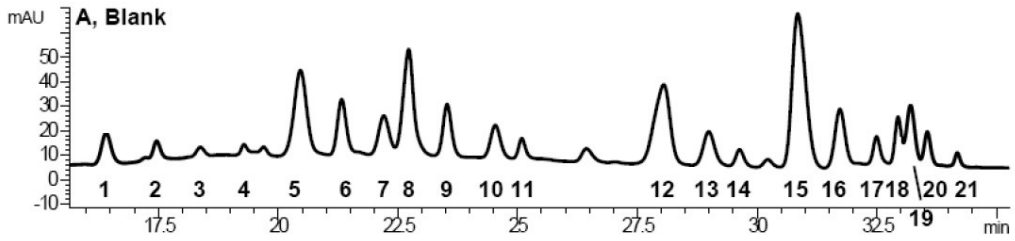
Results

We designed our DCL using aniline-catalysed acylhydrazone exchange[14] as the reversible reaction. We have recently shown that aniline catalysis permits this reaction to be used at pH values greater than 6, enabling its use with biomolecular targets, such as the GSTs.[13b, 15] The physicochemical and structural features of the acylhydrazone linkage have proven highly versatile in DCC, being exemplified across a range of systems and applications.[16] Crucially, the quaternary structure of the dimeric enzyme remains intact in the presence of aniline, enabling ligand selection to take place from the DCL. We chose the three nitro-substituted benzaldehydes **A1**–**A3** as our ligands. The compounds are analogues of the chloro-2,4-dintrobenzene (CDNB) structure, an archetypal weak-binding substrate of all GSTs.[11a, 17] Our linkers **L1**–**L4** were designed to have differing lengths and lipophilicity, thus maximising the opportunity to discriminate across the H-sites of different GST homodimers.

We constructed a DCL using the seven components **A1**–**A3** and **L1**–**L4** (Scheme 2), potentially forming a mixture of 24 homo- and hetero-bis-acylhydrazones (plus additional mono-acylhydrazones). Under reaction conditions with aniline (5 mm) in ammonium acetate buffer (0.1 m) containing 15 % by volume DMSO, pH 6.4 at 25 °C, we observed equilibration in 6 h. Pleasingly, we were able to resolve the majority of bis-acylhydrazone components using two HPLC columns in series (Figure 3). Identification of each peak was done by deconvolution in tandem with LC-MS analysis.



Scheme 2. DCL design for bivalent inhibitor synthesis; GS=*S*-glutathionyl. Our DCL design uses the aniline-catalysed reversible hydrazone formation between three aldehydes (**A1**–**A3**) and four bivalent hydrazides (**L1**–**L4**).



1	A1-L1-A1	**2**	A1-L4-A1	**3**	A1-L3-A1
4	A1-L2-A1	**5**	A1-L1-A3	**6**	A1-L4-A3
7	A1-L3-A3	**8**	A1-L2-A3 A1-L1-A2	**9**	A1-L4-A2
10	A1-L3-A2	**11**	A1-L2-A2	**12**	A3-L1-A3 A3-L4-A3
13	A3-L3-A3	**14**	A3-L2-A3	**15**	A2-L1-A3 A2-L4-A3
16	A2-L3-A3	**17**	A2-L2-A3	**18**	A2-L4-A2
19	A2-L1-A2	**20**	A2-L3-A2	**21**	A2-L2-A2

Figure 3. DCL consisting of linkers **L1–L4** and aldehydes **A1**–**A3**.

The DCL was then templated with four separate GST isozymes: mGSTM1-1, hGSTP1-1, SjGST and mGSTA4-4 (Figure 4). Best responses were obtained for the M and Sj isozymes, with the most notable amplification being product **3** (**A1-L3-A1**), amplified over 600 % by the M-class protein (Figure 5). Smaller amplifications were observed corresponding to hydrazones containing the hydrazide linker **L2**, the shortest of the linkers, and a single compound was amplified (**2**, >100 %) containing the longest linker, **L4**. For the SjGST isoform several peaks were amplified to a similar extent over the blank equilibrium. With the exception of **A1-L3-A1**, all of the other amplifications were due to acylhydrazones containing the hydrazide linker **L2**.

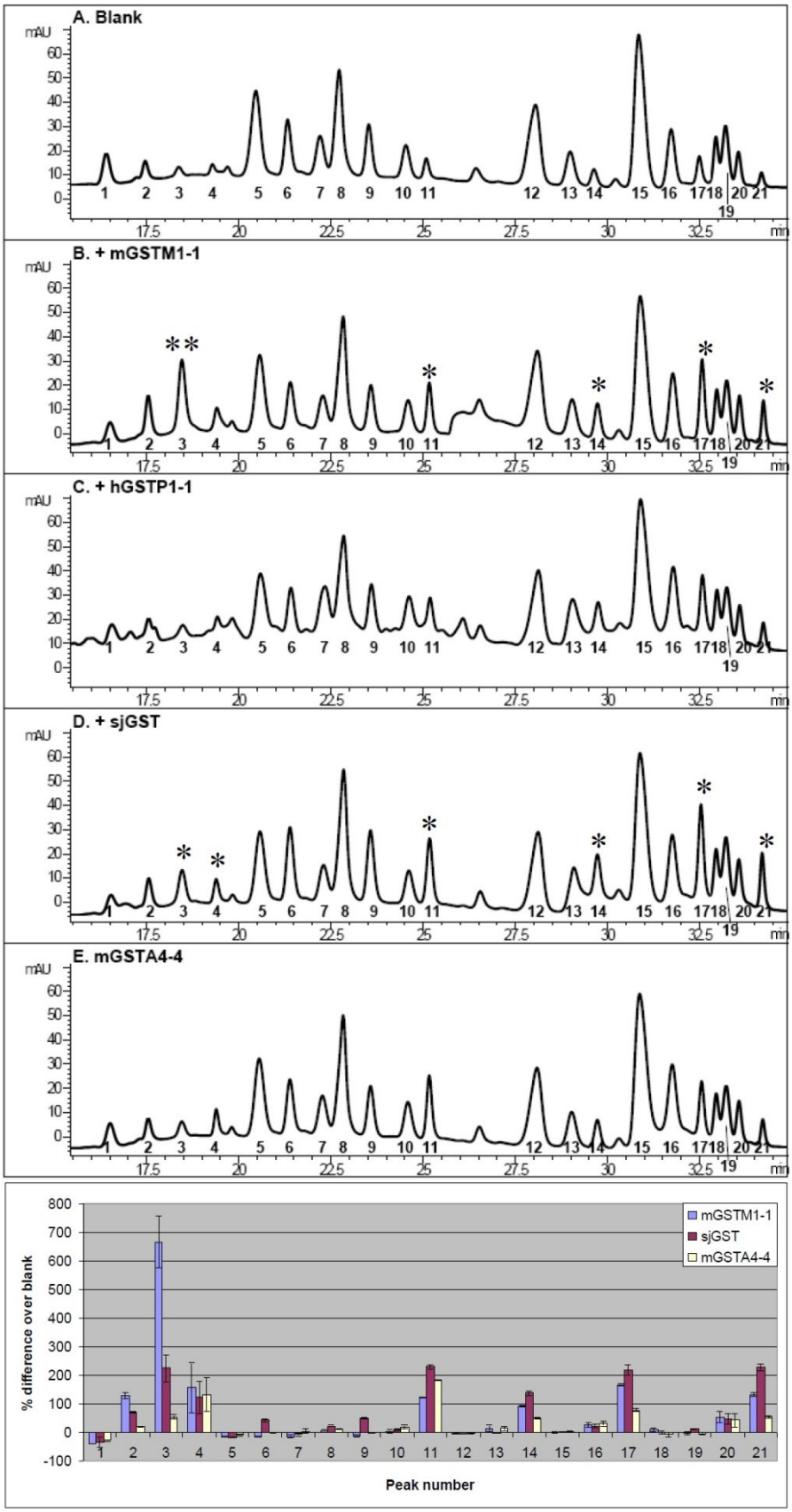


Figure 4. DCL targeted with four different GST isoforms: A, blank equilibrium. B, mGSTM1-1. C, hGSTP1-1. D, sjGST. E, mGSTA4-4. The changes in DCL composition are represented in the bar chart as %

differences of peak area over blank equilibrium. Error bars represent standard deviation over two experiments. The principal amplified peaks for mGSTM1-1 and SjGST are starred on the HPLC trace.

The DCL was generally unresponsive to the mGSTA4-4 isozyme as a template, with some small amplifications being noted for components containing **L2**. The hGSTP1-1 isoform, by contrast, appeared to interfere with the DCL equilibrium for certain components. Bishydrazones containing two **A1** units were degraded by the enzyme, with monohydrazones being favoured by the enzyme (see the Supporting Information). As a result, this isozyme was not used further in templating studies.

With the amplified components identified, it was necessary to establish that the observed amplifications were due to a genuine templative effect from the GST active site. First, the addition of bovine serum albumin (BSA) to the DCL had little effect on the equilibrium composition (see the Supporting Information). Second, we examined the effect of adding GST inhibitors to the DCL mixture to see if amplification would be suppressed. This experiment is complicated in the case of the GSTs first by the lack of any potent inhibitors that are commercially available, and second by the propensity of such inhibitors that are known to bind in varying locations across the large H-site. A variety of non-selective, weakbinding inhibitors are known, however (Figure 5), and their effect on the DCL was examined.

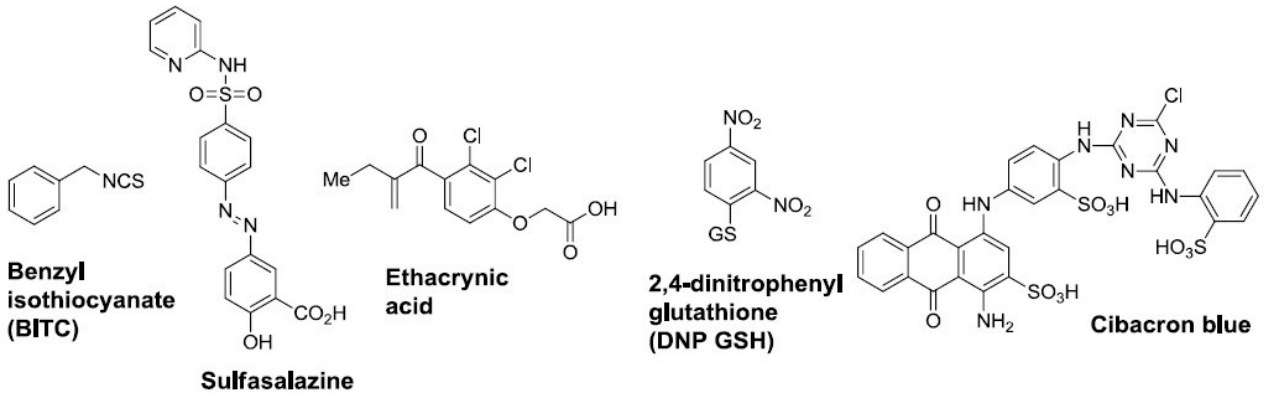


Figure 5. Known inhibitors of GSTs used to block the target's binding pocket in DCL studies.

Benzyl isothiocyanate (BITC) binds in the H-site of GSTs and is known to covalently modify the active site.[18] Sulfasalazine binds in the H-site and is a competitive inhibitor. The crystal structure of sulfasalzine bound to hGSTP1-1 also shows that the pyridinyl ring binds in a shallow pocket on the surface of the protein.[19] Ethacrynic acid is a non-specific inhibitor of GSTs widely used in GST studies, and binds in the H-site. There is evidence that ethacrynic acid can bind in more than one orientation within this hydrophobic binding pocket.[20] Similarly, the GSH-conjugate of CDNB, DNP–GSH, is believed to bind in a different

position to CDNB within the hydrophobic pocket.[19] A final inhibitor, cibacron blue, which binds to the ligand in site of GSTs,[19] was found to disrupt the background DCL equilibrium and could not be used as a control.

Initial control experiments with smaller DCLs established that most of the inhibitors (at 200 μm) had little effect on the enzyme-templated DCL distribution. DNP–GSH, however, was effective in suppressing amplification in a dose-dependent fashion. Adding increasing amounts of DNP–GSH to the main DCL reduced the amplification of **A1-L3-A1** (peak 3) and **A1-L4-A1** (peak 2) in a dose-dependent fashion (see the Supporting Information). The amplification of peak numbers 11, 14, 17 and 21, corresponding to hydrazones not containing a terminal GSH residue, remained unaffected by the addition of DNP–GSH. This could indicate that the external inhibitor was blocking a GSH-binding interaction for the **A1**-containing ligands. To see if those compounds not containing GSH were binding in a different location within the enzyme, other inhibitors, such as ethacrynic acid and CDNB, were also added to the DCL. Unfortunately, the inhibitors eluted in the middle of the library of bis-acylhydrazones, making it very difficult to compare the GST-templated library with the blank.

To assess the biological activity of the amplified components we separately synthesised a selection of symmetrical bis-acylhydrazones for assay, to identify trends between amplification and binding affinity. The compounds **A1-L1-A1**, **A1-L2-A1**, **A1-L3-A1** and **A1-L4-A1**containing the GS-tagged aldehyde **A1** were strongly responsive to mGSTM1-1 in the DCL, and include the strongest amplified compound (**A1-L3-A1**). Compounds containing the **L2** linker were also amplified to a lesser extent, whereas the remaining hydrazones containing **L1**and **L4** were not selected in the DCL and could serve as useful negative controls. The compounds were tested in the chloro-2,4-dinitrobenzene (CDNB) assay, a standard protocol that measures the inhibition of GSH conjugation to CDNB by S_NAr reaction.[17] Table 1shows the IC_{50} data for a subset of symmetrical bis-acylhydrazones containing aldehyde component **A1**.

Table 1. Table of IC_{50} data of DCL products with different GST isoforms measured by using the CDNB assay.

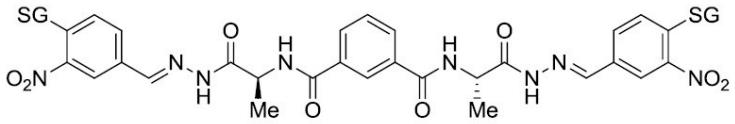
Hydrazone	**IC_{50} [μm]**			
	mGSTM1-1	**hGSTP1-1**	**SjGST]**	**mGSTA4-4**
A1-L1-A1	1.207	126.5	3.471	>100
A1-L2-A1	0.337	11.81	0.252	>100
A1-L3-A1	0.050	13.45	0.989	>100
A1-L4-A1	0.413	0.356	1.800	>100
aldehyde **A1**	341.7	≥500	265.6	>500
hydrazide **L3**	>500	>500	–	–

The IC_{50} data correlated well with the amplifications observed in the DCL in the case of mGSTM1-1. The product with the greatest amplification was **A1-L3-A1** (IC_{50}=50 nm) an almost tenfold greater inhibition of GST activity relative to other hydrazone products. Gratifyingly, excellent isoform specificity was observed, with **A1-L3-A1** showing approximately 20-fold greater inhibitory activity for mGSTM1-1 over SjGST and 270-fold greater inhibitory activity over hGSTP1-1. Both **A1-L2-A1** and **A1-L4-A1** showed smaller amplifications in the DCL, and inhibited mGSTM1-1 at the sub-micromolar level. Compound **A1-L1-A1** was reduced in concentration in the DCL and, accordingly, had the weakest activity at 1.2 μm.

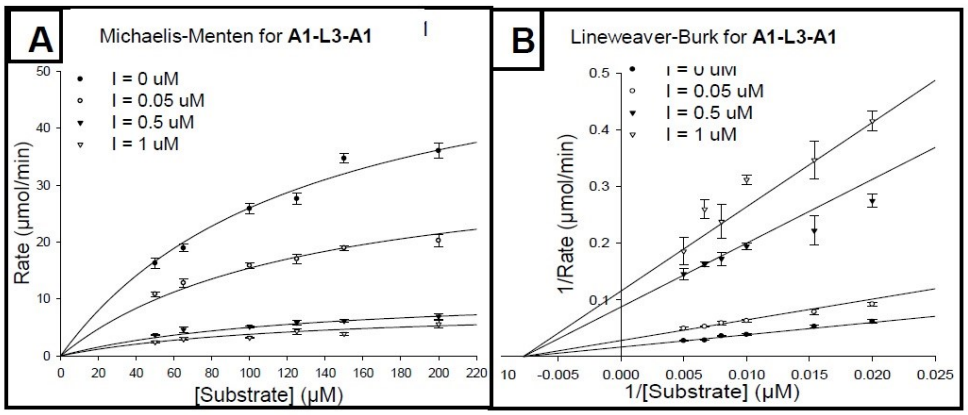
The data for SjGST was less well correlated between IC_{50} and amplification. Linker **L1** was again selected against in the DCL, consistent with DCL distribution, and proved the weakest inhibitor (IC_{50}=3.5 μm). Compound **A1-L2-A1**, however, was the most potent (IC_{50}=0.25 μm) despite being amplified to a lesser extent than **A1-L3-A1** (IC_{50}=0.99 μm). Amplification is not always correlated to binding affinity, a phenomenon identified by Severin[3a] and Sanders et al.[3b, c] that arises from DCLs equilibrating to the global energy minimum. The dynamic exchange of DCL components means this global equilibration can be at the expense of any one local energy minima for a particular protein-binder complex, even the strongest one. For the SjGST-templated library, the majority of amplified peaks (peaks 4, 11, 14, 17 and 21) contained the **L2** linker as a common component. It is possible that competition for this shared building block reduced the**A1-L2-A1** amplification factor relative to **A1-L3-A1** (peak 3), despite the stronger affinity of **A1-L2-A1** for SjGST.

The IC_{50} data for hGSTP1-1 could not be related to the DCL composition due to the lack of clear equilibration. The best inhibitor was **A1-L4-A1** with an IC_{50} value of 0.36 μm, similar to that for mGSTM1-1 (0.41 μm), and considerably more potent than the hydrazones containing shorter linkers. These data, taken with the amplification of monohydrazones in the hGSTP1-1 DCL, are consistent with the linkers **L1**–**L3**not being long enough to effectively span the two G-sites in hGSTP1-1.[12]

The principle of multivalent binding was validated in general for all GST isoforms—there was a large increase in inhibitory activity when the two units of aldehyde **A1** were connected by a linker, regardless of the length or shape of the linker, compared to the single aldehyde unit. This was strikingly illustrated for **A1-L3-A1**, the most potent inhibitor of mGSTM1-1, whereby the two components **A1** and **L3** inhibited the enzyme at 342 and 500 μm, respectively, and the assembled bis-acylhydrazone product was active at 50 nm (10 000-times the potency of the isolated hydrazide linker). More specifically, varying the linker length and structure proved an effective strategy for discriminating between each GST isozyme;[12b] the weak, non-specific inhibitor **A1** became a potent, selective inhibitor when linked with **L2** (SjGST), **L3**(GSTM1-1) and **L4** (GSTP1-1), respectively.



A1-L3-A1



C

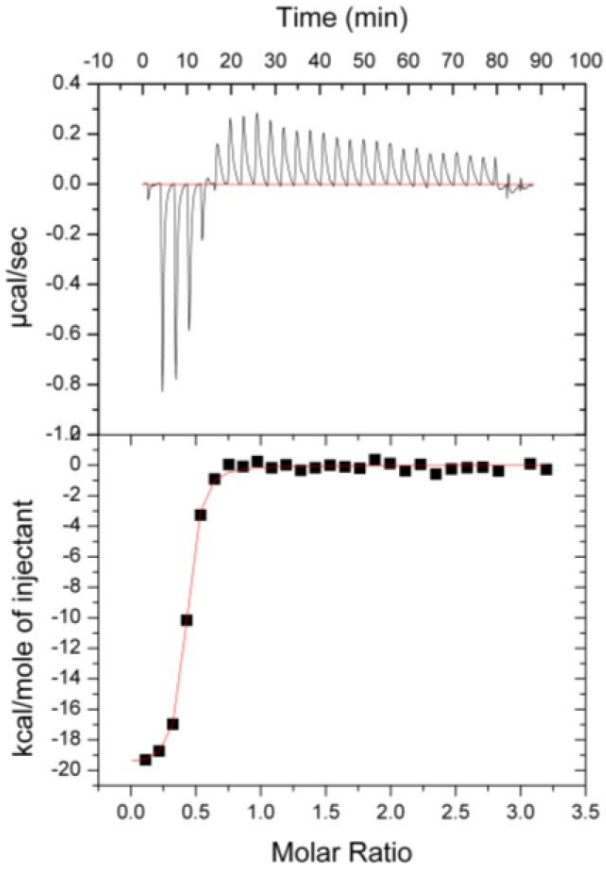


Figure 6. Assay of most potent compound, **A1-L3-A1**, with mGSTM1-1. A) Michaelis–Menten, and B) Lineweaver–Burk plots. The data best fit to a non-competitive (partial) equation. Standard errors are from three repeat experiments. C) ITC data: the top panel shows heat effects associated with the injection of ligand

into the cell containing GST. The lower panel shows the binding isotherm, generated from the integrated heats in the top panel, and the best fitted curves (see the Supporting Information for ITC data).

The inhibition data for mGSTA4-4 shows that each of the hydrazone products had little effect on the enzyme activity, in line with the minimal templating effect of the enzyme on the DCL equilibrium. This made the mGSTA4-4 isoform an excellent negative control, demonstrating that changes in the distribution of hydrazone products was due to specific binding site interactions within the mGSTM1-1 isoform and not due to any non-specific effects from the introduction of the hydrophobic protein to the DCL.

The inhibition mechanism of the bivalent inhibitors for mGSTM1-1 was further characterised by using the CDNB competition assay and the three best binding bis-acylhydrazones **A1-L2-A1**, **A1-L3-A1** and **A1-L4-A1**. The data show that the hydrazones were non-competitive partial inhibitors of CDNB (Figure 6 a and b). The K_i values were in close agreement with IC_{50} measurements for **A1-L3-A1** and **A1-L4-A1**, meeting the Cheng–Prusoff criteria of $IC_{50}=K_i$ for non-competitive inhibitors (Table 2).[21]

Table 2. Binding data for DCL components with mGSTM1-1.

Hydrazone	**IC_{50} [nm]**	**K_i[a] [nm]**	**K_D [nm]**
1. [a] Standard error shown for K_i data taken from three replicates.			
A1-L1-A1	1207	–	–
A1-L2-A1	337	646±80.7	525
A1-L3-A1	50	61±4.1	107
A1-L4-A1	413	634±23.8	–

Isothermal calorimetry (ITC) was then used to determine the binding stoichiometry of the most potent compound **A1-L3-A1**. The results gave a binding stoichiometry of 0.4 with the GST monomer, supporting our hypothesis that the linker is binding across the dimeric GST structure (Figure 6 c). Finally, we examined the effect of the GS moiety on the aldehyde with respect to binding affinity. The symmetrical compounds **A2-LX-A2** and **A3-LX-A3** were synthesised and their IC_{50} values determined by using the CDNB assay. As shown in Table 3, the removal of glutathione from the products resulted in a 100-fold decrease in inhibition for the longer linkers **L3** and **L4**, demonstrating the importance of glutathione for binding affinity. The loss of glutathione for the shortest linker **L2** was better tolerated, with a sixfold reduction in inhibition, suggesting that these shorter compounds were binding in a different mode to the longer linkers. Assaying **A2-L2-A2**, the

most potent bivalent inhibitor without glutathione, against CDNB again showed a non-competitive, partial inhibitor profile withK_i=1.98(±0.47) μm.

Table 3. IC_{50} data for homo-acylhydrazone linkers with mGSTM1-1.[a]

Hydrazide linker	**A1-LX-A1 IC_{50} mGSTM1 [μm]**	**A2-LX-A2 IC_{50} mGSTM1 [μm]**	**A3-LX-A3 IC_{50} mGSTM1 [μm]**
1. [a] Data were acquired by using the CDNB assay.			
aldehyde no linker	341.7	–	>500
L1	1.21	–	–
L2	0.337	2.02	3.38
L3	0.050	9.32	106.4
L4	0.413	49.89	>100

Discussion

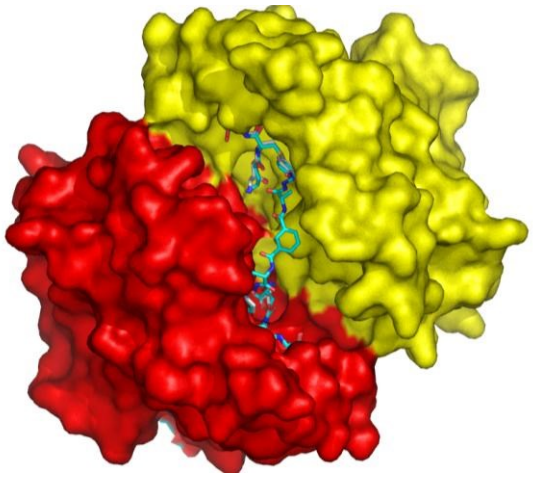
DCC has been used to identify novel bivalent inhibitors of GSTs, with the most potent compound (**A1-L3-A1**) inhibiting mGSTM1-1 with nanomolar activity and excellent isoform specificity over three other isoforms studied. Inhibition of the M-class of GSTs is relevant to oncology, where resistance to anticancer therapies has been correlated with altered expression levels of this enzyme (along with GSTP1).[7, 22] Resistance mechanisms are twofold: GSTM1 can catalyse GSH conjugation to anticancer agents, and it can also down-regulate the p38-MAPK pro-apoptotic pathway.[23]

In this second mechanism, GSTM1 has been shown to bind to apoptosis signal-regulating kinase 1 (ASK1), which is a MAP kinase that activates both the JNK and p38 signalling pathways leading to apoptosis. GSTM1 acts as an inhibitor of ASK1, which dissociates in response to cell stress resulting in the activation of ASK1. Literature reports of M-class GST inhibitors are sparse, with the natural product curcumin being the most potent compound described to date (IC_{50}=0.3 μm for GSTM1-1).[24]

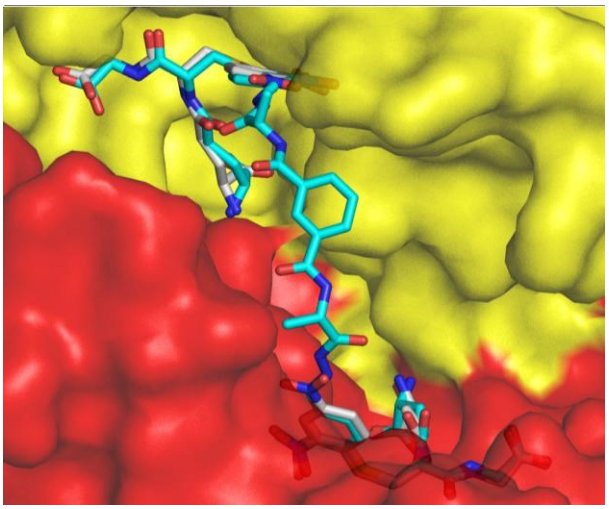
To better understand the role of the selected linker **L3** in spanning the two G-sites in the GSTM1-1 homodimer, we constructed an overlay model based on the crystal structure of human glutathione-*S*-transferase M1A-1A (PDB ID: 1XWK) containing the ligand DNP–GSH from Listowsky and co-workers.[25] The model shows clearly that the **L3** linker can comfortably span the dimer interface without introducing any unfavourable contacts (Figure 7 a). The distance between the two GS moieties in the crystal structure (15.4 Å) and the equivalent in the model (15.5 Å) allows the GS moiety to nearly perfectly overlay in both binding sites within the dimer maintaining all contacts with the protein (Figure 7 b). The slightly longer **L4** and the

shorter **L2** and **L1** linkers have correspondingly weaker affinity in their hydrazone adducts, requiring different linker conformation or protein movement to allow the GS moiety to maintain its binding mode in both G-sites.

(a)



(b)



(c)

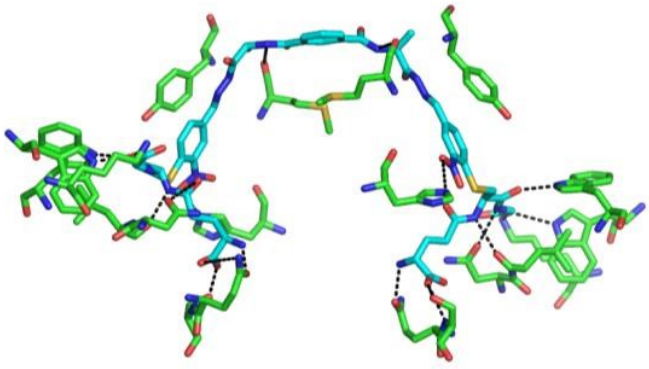


Figure 7. **a)** and **b) A1-L3-A1** overlaid onto GSTM1-1 crystal structure containing GS-DNB bound at each G-site. The protein surface across the dimer interface is shown with the two halves of the dimer coloured

yellow and red, respectively. The GS-DNB moiety of the crystal structure is shown in stick form with white carbons and the modelled **A1-L3-A1** ligand is shown with cyan carbons. The ligand is seen to span the dimer interface whilst introducing no unfavourable contacts. **C)** Residue environment around the **A1-L3-A1** ligand. All polar contacts to the protein come from the GS moiety and are identical to that found in the crystal structure. Polar contacts are indicated by black dashes.

The presence of glutathione at the terminal of each linker was of particular importance for binding affinity for the longer linkers **A1-L3-A1**and **A1-L4-A1**. The glutathione was less important for linkers made from **L2**, with all hydrazones featuring **L2** giving some amplification in the GST-templated DCLs. The activity of **A1-L2-A1** against the M, P and SjGST is somewhat surprising. The **L2** linker should not be long enough to span the distance between the two G-sites without substantial distortion of the dimer structure, suggesting a different, possibly monovalent, binding mode. It should be noted, however, that the IC_{50} value of **A1-L2-A1** was very favourable in comparison to other monovalent inhibitors based on the **A1** scaffold. Previous work from our laboratory has described acylhydrazone GST inhibitors formed from DCLs using **A1** and monovalent acylhydrazides designed to explore the hydrophobic binding pocket.[13b] The best of the compounds from that work had IC_{50} values of 50 μm for hGSTP1-1 and 20 μm for SjGST (**2** and **3**, Figure 8). This is a fivefold reduction in inhibition for hGSTP1-1 and a tenfold reduction for SjGST when compared to bivalent inhibitor **A1-L2-A1**. Compounds **2** and **3** showed little inhibitory effect against mGSTM1-1 with IC_{50} values being greater than 100 μm.

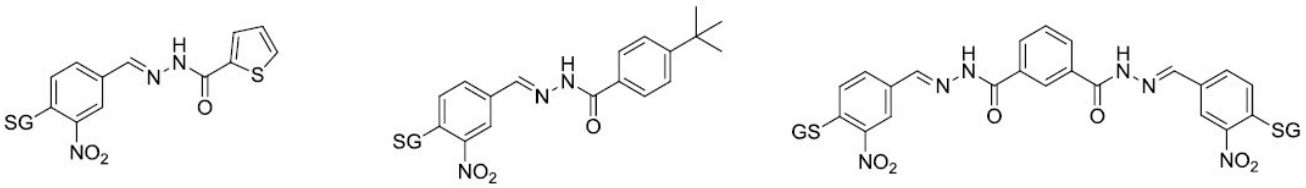


Figure 8. **A1-L2-A1** binding data relative to monovalent hydrazones previously prepared.

The monovalent acylhydrazones were competitive inhibitors of CDNB, and molecular docking studies indicated that they bound into the H-site. ITC was used to determine whether, due to the short linker length, two molecules of **A1-L2-A1** were in fact binding to the GST dimer—one in each active site. Unfortunately, the results were inconclusive (n=0.7) with the measurements being complicated by gellation effects from the compound at higher concentration (see the Supporting Information).

Conclusion

The versatility of DCC in discovering novel binding partners for proteins has been demonstrated, with the identification of a novel, isoform specific, bivalent inhibitor of mGSTM1-1. The success of the study rests on the ability of aniline to catalyse acylhydrazone exchange in the presence of proteins, without compromising tertiary and quaternary protein structure. The lead compound, **A1-L3-A1**, was amplified from a 24-membered library, and is the most potent (and selective) inhibitor described to date for the M-class of GSTs. DCC has a natural overlap with multivalent, fragment-type approaches to ligand discovery given that it is predicated on the assembly of small, weakly-binding fragments, in situ. Our study shows how this element of DCL design may be used very successfully in the discovery of potent enzyme inhibitors. We have also demonstrated that the addition of an external enzyme inhibitor to an equilibrating DCL can provide additional insights to simply switching amplification off. The large GST binding region interrogated by our DCLs enabled us to select a range of inhibitors that bind in different areas of the enzyme, and examine their effect (or lack of one) on the amplification of best-binding components. Finally, by designing DCLs to probe larger protein surface areas across two 28 kDa protein monomers, we have shown that the approach is suited to targets that are less tractable to conventional structure-based drug discovery methods. Future work will look to extrapolate these techniques to the study of protein–protein interactions.

Experimental Section

HPLC conditions for analysis of DCL libraries: Column, Luna 5μ C18(2), 50 mm×2.0 mm, and Luna 5μ C18(2), 30 mm×4.6 mm, in sequence; flow rate, 1 mL min^{-1}; wavelength, 254 nm; temperature, 30 °C; gradient H_2O/MeCN (0.01 % TFA) from 95 to 80 % over 15 min, then to 60 % over a further 15 min, and finally to 5 % over 5 min.

DCL blank equilibrium; aniline catalysis of reversible hydrazone formation: The four bishydrazide linkers **L1**–**L4** (4×2.4 μL, 10 mm, DMSO), aldehyde **A1** (2.4 μL, 10 mm, water), aldehydes **A2**-**A3** (2×2.4 μL, 10 mm, DMSO) and aniline (15 μL, 0.1 m, DMSO) were added to a mixture of DMSO (15.6 μL, 15 % by vol.) and ammonium acetate buffer (252.6 μL, 100 mm, pH 6.4). The DCL was allowed to incubate at room temperature with occasional shaking, and was monitored periodically by HPLC to establish the blank composition until the relative populations of the hydrazones became constant.

Enzyme-templated DCL assay: Protein, GST (to a final concentration of 20 μm, for mGSTM1-1; 152.8 μL, 1.1 mg mL^{-1}, in potassium phosphate buffer 0.1 m, pH 6.8) or BSA (to a final concentration of 20 μm, 199.5 μL, 2.0 mg mL^{-1}, in 0.9 % aqueous NaCl solution containing sodium azide), the four hydrazide linkers **L1**–**L4** (4×2.4 μL, 10 mm, DMSO), aldehyde **A1** (2.4 μL, 10 mm, water), aldehydes **A2**-**A3** (2×2.4 μL, 10 mm, DMSO) and aniline (15 μL, 0.1 m, DMSO) were added to a mixture of DMSO (15.6 μL, 15 % by vol.) and

ammonium acetate buffer (to a total volume of 255 μL, for mGSTM1-1; 99.8 μL, 100 mm, pH 6.4). The DCL was allowed to incubate at room temperature, with occasional shaking, for 48 h. HPLC analysis was performed and the traces were compared with the blank equilibrium.

Notes and references

[1] (a)V. T. Bhat, M. F. Greaney, in *Dynamic Combinatorial Chemistry: In Drug Discovery, Bioorganic Chemistry, and Materials Science*(Ed.: B. L. Miller), Wiley, New York, 2009, pp. 43–82; (b)O. Ramström, L. Amorim, R. Caraballo, O. Norberg, in *Dynamic Combinatorial Chemistry*, (Eds.: J. N. H. Reek, S. Otto), Wiley,New York, 2009, pp. 109–150; (c) M. Hochgürtel, J.-M. Lehn, in *Fragment-Based Approaches in Drug Discovery* (Eds.: W. Jahnke, D. A. Erlanson), Wiley, New York,2006, pp. 341–364.

[2] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* 2006, **106**, 3652–3711.

[3] (a) K. Severin, *Chem. Eur. J.* 2004, **10**, 2565–2580; (b) P. T. Corbett, S. Otto, J. K. M. Sanders, *Chem. Eur. J.* 2004, **10**, 3139–3143; (c) P. T. Corbett, J. K. M. Sanders, S. Otto, *Chem. Eur. J.* 2008, **14**, 2153–2166.

[4] Selected examples: (a) I. Huc, J.-M. Lehn, *Proc. Natl. Acad. Sci. USA* 1997, **94**, 2106–2110; (b) R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Angew. Chem. Int. Ed.* 2002, **41**, 3405–3407; (c) S. Zameo, B. Vauzeilles, J.-M. Beau, *Angew. Chem. Int. Ed.* 2005, **44**, 965–969; (d) S.-A. Poulsen, L. F. Bornaghi, *Bioorg. Med. Chem.* 2006, **14**, 3275–3284; (e) S. Zameo, B. Vauzeilles, J.-M. Beau, *Eur. J. Org. Chem.* 2006, 5441–5444.

[5] DCMS (dynamic combinatorial mass spectrometry) has been used to discover novel binding fragments via disulfide formation and protein mass spectrometry. Seminal work: D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon, J. A. Wells, *Proc. Natl. Acad. Sci. USA* 2000, **97**, 9367–9372; recent examples: (a) E. C. Y. Woon, M. Demetriades, E. A. L. Bagg, W. Aik, S. M. Krylova, J. H. Y. Ma, M. Chan, L. J. Walport, D. W. Wegman, K. N. Dack,M. A. McDonough, S. N. Krylov, C. J. Schofield, *J. Med. Chem.* 2012, **55**, 2173–2184; (b) B. M. R. Liénard, R. Huting, P. Lassaux, M. Galleni, J.-M. Frere, C. J. Schofield, *J. Med. Chem.* 2008, **51**, 684–688; (c) B. M. R. Liénard, N. Selevsek, N. J. Oldham, C. J. Schofield, *ChemMedChem* 2007, **2**, 175–179; (d) S. A. Poulsen, *J. Am. Soc. Mass Spectrom.* 2006, **17**, 1074–1080.

[6] J. D. Hayes, J. U. Flanagan, I. R. Jowsey, *Annu. Rev. Pharmacol. Toxicol.* 2005, **45**, 51–88.

[7] Review: D. M. Townsend, K. D. Tew, *Oncogene* 2003, **22**, 7369–7375. Recent examples: (a) K. Tsuboi, D. A. Bachovchin, A. E. Speers, T. P. Spicer, V. Fernandez-Vega, P. Hodder, H. Rosen, B. F. Cravatt, *J. Am. Chem. Soc.*2011, **133**, 16605–16616; (b) J. Son, J. J. Lee, J. S. Lee, A. Schuller, Y. T. Chang, *ACS Chem. Biol.* 2010, **5**, 449–453; (c) L. Federici, C. Lo Sterzo, S. Pezzola, A. Di Matteo, F. Scaloni, G. Federici, A. M. Caccuri, *Cancer Res.* 2009, **69**, 8025–8034; (d) W. H. Ang, L. J. Parker, A. De Luca, L. Juillerat-Jeanneret, C. J. Morton, M. Lo Bello, M. W. Parker, P. J. Dyson, *Angew. Chem. Int. Ed.* 2009,

48, 3854–3857; (e) W.-S. Li, W. S. Lam, K.-C. Liu, C.-H. Wang, H.-C. Chang, Y.-C. Jen, Y. T. Hsu, S. S. Shivatare, S.-C. Jao, *Org. Lett.* 2010, **12**, 20–23.

[8] Selected examples: (a) I. Romieu, M. Ramirez-Aguilar, J. J. Sienra-Monge, H. Moreno-Macías, B. E. del Rio-Navarro, G. David, J. Marzec, M. Hernández-Avila, S. London, *Eur. Respir. J.* 2006, **28**, 953–959; (b) W. D. Carroll, W. Lenney, P. W. Jones, R. C. Strange, F. Child, M. K. Whyte, R. A. Primhak, A. A. Fryer, *Clin. Exp. Allergy* 2005, **35**,1155–1161; (c) F. Sampsonas, M.-A. Archontidou, E. Salla, K. Karkoulias, G. Tsoukalas, K. Spiropoulos, *Allergy Asthma Proc.* 2007, **28**, 282–286.

[9] B. Xue, Y. Wu, Z. Yin, H. Zhang, S. Sun, T. Yi, L. Luo, *FEBS Lett.* 2005, **579**, 4081–4087.

[10] S.-C. Jao, J. Chen, K. Yanga, W.-S. Lia, *Bioorg. Med. Chem.* 2006, **14**, 304–318.

[11] (a) B. Mannervik, U. H. Danielson, *Crit. Rev. Biochem. Mol. Biol.* 1988, **23**, 283–337; (b) D. Sheehan, G. Meade, V. M. Foley, C. A. Dowd, *Biochem. J.* 2001, **360**, 1–16.

[12] (a) R. P. Lyon, J. J. Hill, W. M. Atkins, *Biochemistry* 2003, **42**, 10418–10428; (b) D. Y. Maeda, S. S. Mahajan, W. M. Atkins, J. A. Zebala, *Bioorg. Med. Chem. Lett.* 2006, **16**, 3780–3783; (c) S. Mahajan, L. Hou, R. Paranji, D. Maeda, J. Zebala, W. M. Atkins, *J. Am. Chem. Soc.* 2006, **128**, 8615–8625.

[13] (a) B. Shi, R. Stevenson, D. J. Campopiano, M. F. Greaney, *J. Am. Chem. Soc.* 2006, **128**, 8459–8467; (b) V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat. Chem.* 2010, **2**, 490–497.

[14] (a) E. H. Cordes, W. P. Jencks, *J. Am. Chem. Soc.* 1962, **84**, 826–831; (b) A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, *J. Am. Chem. Soc.* 2006, **128**, 15602–15603.

[15] For alternative nucleophilic catalysis of hydrazone formation, see: (a) A. R. Blanden, K. Mukherjee, O. Dilek, M. Loew, S. L. Bane, *Bioconjugate Chem.* 2011, **22**, 1954–1961; (b) S. R. Beeren, M. Pittelkowz, J. K. M. Sanders, *Chem. Commun.* 2011, **47**, 7359–7361.

[16] (a) G. R. L. Cousins, S.-A. Poulsen, J. K. M. Sanders, *Chem. Commun.* 1999, 1575–1576; (b) T. Bunyapaiboonsri, O. Ramström, S. Lohmann, J.-M. Lehn, L. Peng, M. Goeldner, *ChemBioChem* 2001, **2**, 438–444; (c) R. L. E. Furlan, Y.-F. Ng, S. Otto, J. K. M. Sanders, *J. Am. Chem. Soc.* 2001, **123**, 8876–8877; (d) S. L. Roberts, R. L. E. Furlan, G. R. L. Cousins, J. K. M. Sanders, *Chem. Commun.* 2002, 938–939; (e) T. Bunyapaiboonsri, H. Ramström, O. Ramström, J. Haiech, J.-M. Lehn, *J. Med. Chem.* 2003, **46**, 5803–5811; (f) O. Ramström, S. Lohmann, T. Bunyapaiboonsri, J.-M. Lehn, *Chem. Eur. J.* 2004, **10**, 1711–1715; (g) J. Liu, K. R. West, C. R. Bondy, J. K. M. Sanders, *Org. Biomol. Chem.* 2007, **5**, 778–786; (h) M. G. Simpson, M. Pittelkow, S. P. Watson, J. K. M. Sanders, *Org. Biomol. Chem.* 2010, **8**, 1173–1180; (i) M. G. Simpson, M. Pittelkow, S. P. Watson, J. K. M. Sanders, *Org. Biomol. Chem.* 2010, **8**,

1181–1187; (j) S. R. Beeren, J. K. M. Sanders, *J. Am. Chem. Soc.* 2011, **133**, 3804–3807; (k) S. R. Beeren, J. K. M. Sanders, *Chem. Sci.* 2011, **2**, 1560–1567; (l) A. V. Gromova, J. M. Ciszewski, B. M. Miller, *Chem. Commun.* 2012, **48**, 2131–2133; (m) G. Deng, F. Li, H. Yu, F. Liu, C. Liu, W. Sun, H. Jiang, Y. Chen, *ACS Macro Lett.* 2012, **1**, 275–279.

[17] W. H. Habig, M. J. Pabst, W. B. Jakoby, *J. Biol. Chem.* 1974, **249**, 7130–7139.

[18] L. A. Ralat, R. F. Colman, *J. Biol. Chem.* 2004, **279**, 50204–50213.

[19] A. J. Oakley, M. Lo Bello, M. Nuccetelli, A. P. Mazzetti, M. W. Parker, *J. Mol. Biol.* 1999, **291**, 913–926.

[20] (a) A. J. Oakley, J. Rossjohn, M. Lo Bello, A. M. Caccuri, G. Federici, M. W. Parker, *Biochemistry* 1997, **36**, 576–585; (b) A. J. Oakley, M. Lo Bello, A. P. Mazzetti, G. Federici, M. W. Parker, *FEBS Lett.* 1997, **419**, 32–36.

[21] Y.-C. Cheng, W. H. Prusoff, *Biochem. Pharmacol.* 1973, **22**, 3099–3108.

[22] C. C. McIlwain, D. M. Townsend, K. D. Tew, *Oncogene* 2006, **25**, 1639–1648.

[23] Selected examples: (a) S.-G. Cho, Y. H. Lee, H.-S. Park, K. Ryoo, K. W. Kang, J. Park, S.-J. Eom, M. J. Kim, T. S. Chang, S.-Y. Choi, J. Shim, Y. Kim, M.-S. Dong, M.-J. Lee, S. G. Kim, H. Ichijo, E.-J. Choi, *J. Biol. Chem.* 2001, **276**, 12749–12755; (b) S. Dorion, H. Lambert, J. Landry, *J. Biol. Chem.* 2002, **277**, 30792–30797; (c) K. Ryoo, S.-H. Huh, Y.-H. Lee, K. W. Yoon, S.-G. Cho, E.-J. Choi, *J. Biol. Chem.* 2004, **279**, 43589–43594.

[24] R. Appiah-Opong, J. N. M. Commandeur, E. Istyastono, J. J. Bogaards, N. P. E. Vermeulen, *Xenobiotica* 2009, **39**, 302–311.

[25] Y. Patskovsky, L. Patskovska, S. C. Almo, I. Listowsky, *Biochemistry* 2006, **45**, 3852–3862.