

CHAPTER 17

A Link Means a Lot: Disulfide Tethering in Structure-Based Drug Design

JEANNE A. HARDY

The University of Massachusetts Amherst, Department of Chemistry, 710 N. Pleasant St., Amherst, MA 01003, USA

17.1 Introduction: What is Disulfide Tethering?

Tethering is a disulfide-based drug-discovery technology that grew out of the era when combinatorial chemistry was a burgeoning new development. What separates Tethering from other combinatorial approaches is its site-directed nature. The disulfide link that is formed allows a wide variety of advantages and analysis not available with other drug methods, thus in Tethering a link means a lot. Site-directedness makes it ideal for interrogating traditionally difficult drug targets, such as protein–protein interaction surfaces, and new binding sites. It has also led to candidates for linking in a combinatorial manner, new fragments for addition to existing scaffolds, and entirely new classes of pharmacophores for several different protein targets. At the very essence of tethering is the concept that structural information can and should inform drug design and discovery.

Tethering was conceptualized and developed by a team assembled and directed by James Wells at Sunesis Pharmaceuticals and first publicly reported by Erlanson and coworkers in 2000.¹ The progress of Tethering and its many applications has been extensively reviewed,^{2–4} thus overlap in content between this chapter and those reviews are bound to exist. Nevertheless, this review focuses on the practical considerations during Tethering and on the strong role that structural information plays during Tethering interrogation.

The inaugural report of Tethering focused on the cysteine-containing active site of thymidylate synthase (TS), a protein that is essential to synthesis of deoxythymidine monophosphate (dTMP) from deoxyuridine monophosphate

(dUMP) that has been identified as a cancer target.¹ To probe TS and develop Tethering, a library of 1200 disulfide-containing molecules was synthesized to allow proof of concept. Each molecule in the library consisted of two elements: a monophore, which is the unique chemical entity, and a linker region containing a thiol for disulfide exchange (Figure 17.1). Tethering screening was performed in pools of ten compounds per pool, where each compound has a unique molecular weight. Pools of compounds were incubated with TS under high reductant concentrations where disulfide exchange was rapid. Compounds that interacted even weakly with the surface, a pocket, or a binding site on the protein near a cysteine residue, and thus had an increased residence time, could promote the formation of disulfide bond between the cysteine thiol and the compound thiol. Binding of the monophores to TS entropically stabilized the interactions and favored the disulfide-bonded protein.

One of the most fascinating aspects of screening for drug leads by Tethering is that it is a tunable activity. At moderate reductant concentrations, monophores with too weak an inherent activity are reduced off, so that only interactions driven by interactions of the monophore with the protein remain intact. Covalently modified protein can then be analyzed by mass spectrometry or a functional assay so that bound compounds can be identified. In the presence of

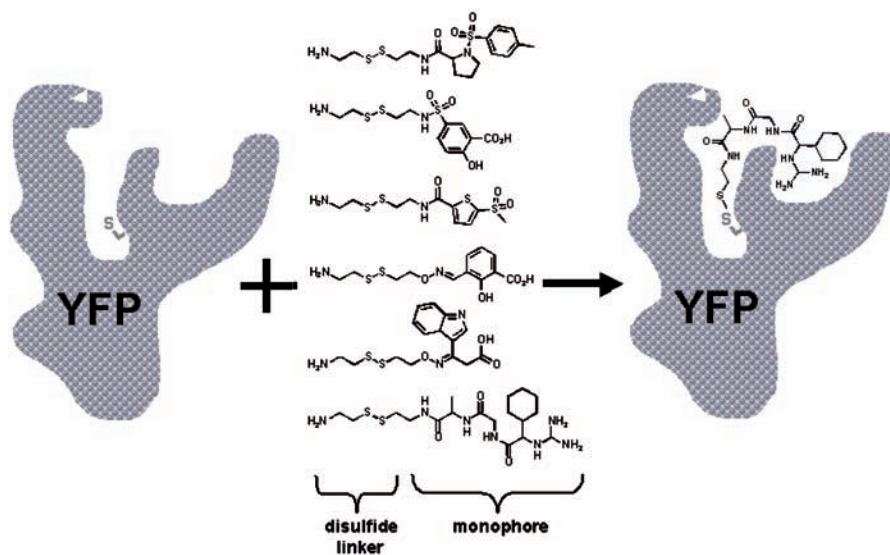


Figure 17.1 Tethering schematic. The general scheme of Tethering is shown for a YFP (your favorite protein), which contains an exposed and free cysteine thiol. Disulfide-containing compounds are incubated with various concentrations of reductant and YFP. Monophores that specifically bind to surfaces or cavities on YFP facilitate the formation of a disulfide bond between the native or introduced cysteine residue. Non-specific interactors are reduced off by the moderate amount of reductant present. Captured fragments can be easily identified by mass spectrometry of the covalent protein–small molecule complex.

very high levels of reductant, any observed binding or inhibition is relieved. This aspect of the technique is very useful for showing that the mechanism of inhibition occurs through binding at the targeted site, via the disulfide, and not through some non-specific mechanism of inhibition such as micelle formation or protein denaturation. In the case of TS, the results were analyzed by mass spectrometry. The most strongly interacting compound was *N*-tosyl-D-proline (TP). The specificity and affinity of TP for TS was so significant that even when 100 compounds were included in the pool with TP, TS was still bound selectively by TP, showing that exquisite specificity is a result of the monophore only.

Although TS has five cysteines, the active site was thought to be the most reactive and the only site at which compounds bound. Numerous subsequent studies have confirmed the finding that compound binding is specific for a single cysteine residue or a unique binding orientation. In addition, it has been generally observed that fragments that bind to one protein are specific for that protein and do not bind to other proteins. In the case of TS, the active site cysteine was mutated to serine (C146S) and TP binding was abolished. However, in that context, if a neighboring leucine residue was also mutated to cysteine (C146S/L143C) TP could still bind. Crystal structures of the wild-type TS and C146S/L143C show the TP to be in the same orientation (Figure 17.2). Introduction of a cysteine residue at a different adjacent position did not result in productive binding of TP. This suggests that there is some flexibility afforded by the linker portion of the disulfide-containing compounds. When the thiol was removed from TP to generate thiol-free-TP, that compound still bound to TS ($K_I = 1.1$ mM) and the conformation of the TP moiety in the crystallized complex was the same as in the covalently bound compound. Thus, the interaction is exquisitely driven by binding of the monophore region and is largely independent of the linker.

From the first success with TS, Tethering has been developed into a robust, rapid, and highly successful protocol for deriving new pharmacophores as pharmaceutical leads. During Tethering, fragments are equilibrated under conditions that promote disulfide exchange. Typically, these experiments are performed at concentrations of 0.1–2 mM β -mercaptoethanol (β -ME). There are two standard ways that the strength of the interaction between the small-molecule fragment and the protein are measured. The first method is to determine the β -ME₅₀, which measured the amount of β -ME needed to cause 50% labeling of the cysteine-containing protein. Another way to measure the strength of a Tethering interaction is by assessment of a dose–response-50 (DR₅₀) for a compound titration at a given level of reductant (usually 1–4 mM β -ME). Small molecule–protein pairs with very favorable interaction energies are less likely to dissociate. Typical β -ME₅₀ values for strongly selected compounds range from 0.7 to 5 mM.

In developing Tethering into a successful commercial venture, one consideration was the composition of pools that would be ideal for decoding the binding molecule while still yielding tractable drug leads. During Tethering the average molecular mass in disulfide-containing molecules is 250 Da. This size

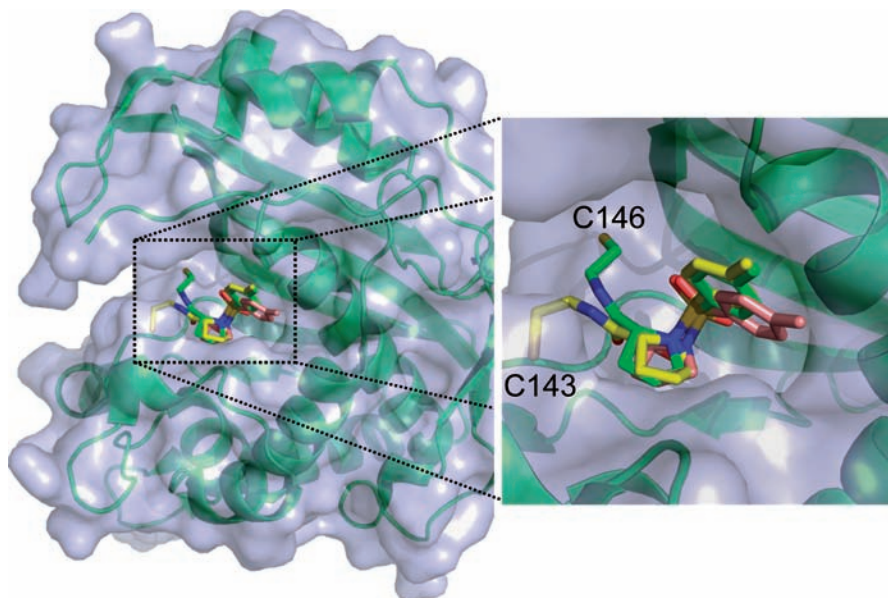


Figure 17.2 Tethering to thymidylate synthase. Thymidylate synthase (green ribbons and gray surface) is shown bound to *N*-tosyl-D-proline at active-site cysteine C146 (green sticks, PDB ID 1F4C), or at introduced cysteine C143 (yellow sticks, PDB ID = 1F4D) or non-covalently associated with the active site (orange sticks, PDB ID = 1F4E). The overlap in the conformations of the *N*-tosyl-D-proline moieties shows that binding is determined by the monophore and is independent of the cysteine linker.

allows for fragment recombination and many avenues for future optimization. In development of the pools, attention was paid to the chemical nature of the compounds so that each compound exhibited drug-like properties as well as structural and chemical diversity. Thus, overall the frequency of hits is low. For screening 12 cysteine mutants in interleukin-2 (IL-2), the frequency of a hit was 0–1% overall, where the highest hit rate for an individual cysteine was 1.3% and four of the positions produced 0–0.08% hit rates.⁵ In C5a the hit rate for the four cysteine residues studied ranged from 0 to 0.6%.⁶ In screening the caspase-3 active-site, no hits out of a 10 000 compound library were found.⁷ Screening by Tethering is a readily tunable activity, which is a hallmark of its adaptability, and thus varying the concentration of reductant dramatically changes the number of hits observed. The success of Tethering is monitored by stringent hit rate, biological sense of hit sites (either from other assay data or based on crystal structures), expectation that hits display sharp structure–activity relationships (SARs) – typically an enantiomer is expected to have at least ten-fold or more decreased activity as the originally identified compound – and the requirement that a hit must be reversible by addition of high concentrations of reductant.

Most screens based on Tethering have used covalent modification as measured by mass spectrometry as the readout. Two examples demonstrate that Tethering can also be monitored by functional assays. A functional screen for the inhibition of the intrinsic apoptotic pathway was measured as a function of caspase-3 activity, a cell-lysate based assay. One compound (FICA, see Section 17.7), which was subsequently shown to be an allosteric inhibitor of caspase-3 and -7, was discovered based on that screen.⁷ Tethering has also been applied to the integral membrane protein, the C5a receptor. It was possible to adapt traditional cell-based assays for IP₃ accumulation and C5a ligand-binding assays with membrane fractions containing C5a receptor to conditions where the level of reductant could be varied so that Tethering experiments could be performed.^{6,8}

In developing a site-directed drug-discovery platform, other types of reversible bond-forming moieties could have been selected, but much of the success of Tethering probably lies in the selection of a disulfide as the covalent link, as it can form and break under mild conditions. An additional advantage is that, since Tethering is reversible, it is easy to rule out non-specific modes of inhibition, such as aggregation or denaturation. In the vast majority of cases where the strength of the interaction between the monophore and the protein is sufficient to bind in the absence of a disulfide, the crystal structure of the disulfide-bonded and disulfide-free versions of the compounds are virtually identical.

17.2 Success of Native Cysteine Tethering

Many of the targets probed using Tethering have been enzymes with active-site cysteine residues [caspase-1, caspase-3, caspase-7, protein tyrosine phosphatase 1B (PTP1B), and TS (Table 17.1)]. In these cases the site of disulfide interaction is often presumed to be the catalytically important cysteine. Using native cysteines of any ilk for Tethering, an important consideration is whether the selected cysteine will react more readily with the library of thiol-containing compounds than any other cysteine residue on the protein. This concern was directly tested in TS where four non-conserved buried cysteine residues are present. These residues appeared far less reactive to disulfide exchange than the active site cysteine. The active-site cysteine readily reacted with cystamine (S-CH₂-CH₂-NH₂), however the active-site mutant wherein cysteine was replaced with serine (C146S) did not react with cystamine, suggesting that the active-site cysteine was the only exposed and disulfide-exchange-reactive residue.¹

The active-site cysteine is not necessarily the point of interaction in Tethering. If binding of the disulfide compounds was purely driven by electrophilic propensities, then the most activated cysteines should always preferentially be the disulfide-containing compounds. Highly electrophilic moieties, such as methylthiolsulfonates, fluoro- and chloroketones, and nitrosylating compounds, such as *S*-nitroso-*N*-acetyl-D,L-penicillamine (SNAP), label exposed

Table 17.1

<i>Protein</i>	<i>Tethering</i>	<i>Extended Tethering</i>	<i>Breakaway Tethering</i>	<i>Native Cysteine</i>	<i>Engineered Cysteine</i>	<i>MS Screen</i>	<i>Functional Screen</i>	<i>Conformational Change by Binding</i>
Thymidylate Synthase	+			+		+		+
IL-2	+				+	+		+
caspase-1		+		+		+		
caspase-3	+	+		+		+	+	
caspase-3 allosteric site	+			+		+	+	+
caspase-7 allosteric site	+			+		+		+
PTP-1B			+		+	+		
PTP-1B allosteric site	+			+		+		+
C5a	+			+	+	+	+	ND

cysteines as a function of the pK_a and surface exposure of the cysteine residue. During Tethering experiments, binding of the disulfide-containing compounds is driven by interactions of the monophore with the protein rather than by the electrophilicity of the thiol, so the active site is not the only site that can be labeled. An example of this is in PTP1B where an introduced cysteine could be preferentially alkylated even in the presence of the active-site cysteine residue.⁹

Native cysteine residues have also been used productively for extended Tethering in caspase-3^{10–12} and caspase-1^{13,14} even in the presence of other exposed cysteine residues (see Section 17.5). One native cysteine Tethering success occurred serendipitously at a non-active-site cysteine in a cavity distal from the active site. Tethering to this distal cysteine resulted in the discovery of a new allosteric site that is present in caspase-3 and caspase-7 (see Hardy *et al.*⁷ and Section 17.7). Similarly, in probing PTP1B with a series of alkylating agents, one thiol-selective fluorogenic compound, 4-(aminosulfonyl)-7-fluoro-2,1,3-benzoxadiazole (ABDF), was discovered that selectively modified a single native cysteine (C121) outside the active site and allosterically inactivated the protein.¹⁵

17.3 Role of Structure in Engineered-cysteine Tethering

A major advantage that Tethering holds over other methods of drug discovery is the ability to specifically target one particular region of the protein that may be the active site, an allosteric site, a novel cavity, or a protein binding surface. If the region one hopes to target is not near a native cysteine, a cysteine residue must be engineered into the protein to allow Tethering to proceed. Thus, a big question in using engineered cysteines is where to engineer the cysteines. Very often visual inspection is used, and all of the exposed residues within the appropriate distance (4–8 Å) of the active site are exhaustively or selectively mutated to cysteine.¹⁶ Typically, mutation of one surface residue to cysteine does not have dramatic deleterious effects on activity, so it is possible to introduce cysteine residues at various locations around the site of interest.

An early case study in the introduction of new cysteine residues for Tethering was carried out in IL-2 (T-cell cytokine interleukin-2).^{5,16–20} Arkin and co-workers selected a series of ten residues by visual inspection to mutate to cysteine for Tethering at an adaptable binding site that had been predicted to contact the IL-2 receptor (Figure 17.3).^{5,17} None of the introduced cysteine residues disrupted production or folding of IL-2, however, the same residues that had been shown by previous mutagenesis studies to disrupt IL-2 binding to the IL-2 receptor also disrupted binding when mutated to cysteine. Not all of the introduced residues had the same hit rate in a Tethering assay, underscoring the general observation that some residues are intrinsically more fruitful for Tethering. This is likely because of the makeup of the library, the positioning of the residues, and the conformations that are accessible from that residue conjugated with small molecules. In the case of IL-2, residues that display the highest conformational flexibility also have the highest hit rate.⁵

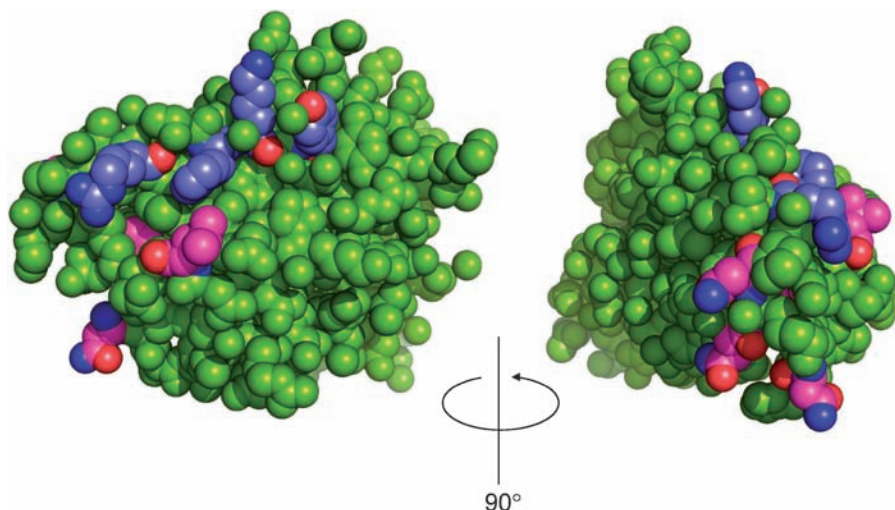


Figure 17.3 Residues mutated to cysteine for Tethering on IL-2. Residues that had lower reactivity (captured fewer than seven compounds per residue) were in the highly ordered region of the IL-2 binding site. These residues are blue and include, from left to right, in the panel on the left R38, F42, K43, and Y45. The residues that had higher reactivity (captured greater than 20 compounds per residue) were in the adaptable region of the IL-2 binding site. These residues are drawn in pink and include residues L72 and N77 in the panel on the left. In the panel on the right the pink residues, from top to bottom, are L72, N33, Y31, N30, and N77.

Computational methods can also be used to suggest which residues might be best for cysteine introduction and subsequent Tethering interrogation. In these calculations probes based on the linker regions present in the compounds in the disulfide library are attached to the candidate residues *in silico* and subjected to molecular dynamics simulations. The trajectories of the linkers toward or away from the region of interest can be used to determine which residues are the most likely to result in successful Tethering.

Depending on the position of the engineered cysteine and the flexibility of the linker, fragments discovered at one cysteine can also bind when an adjacent amino acid is substituted by cysteine. In the case of TS, the tosyl-proline fragment bound with similar affinity and in the same orientation to the active-site cysteine (C146) and to a cysteine introduced at position 143 (L143C), but not at other residues H147C, presumably because the geometry was non-optimal¹ (Figure 17.1). For C5a, approximately 10% of the hits were reactive with more than one of the engineered cysteine residues.⁶ In IL-2, no two introduced cysteine residues captured the same disulfide-containing fragments.⁵ This suggests, in general, that hits bind to a single cysteine with discriminating orientation.

Reactivity of the compounds from the disulfide library with other non-target cysteine residues is sometimes a consideration. Although Tethering is

predominantly driven by interactions of the monophore with the surface of the protein, the reactivity of the compound library is, to some extent, governed by the pK_a of the cysteines on the surface of the protein. Activated or hyper-exposed cysteines can interfere with Tethering at the desired location. The general reactivity of non-candidate cysteine residues can often be monitored by reaction with small thiol-containing compounds such as 2-ME, oxidized glutathione, or cystamine followed by peptide mapping with mass spectrometry. If additional non-candidate cysteine residues show strong reactivity, this often simplifies the analysis of the disulfide-library screening to “scrub” all other cysteine residues from the surface of the protein by mutation to alanine or serine. In screens on cysteine-free variants any covalent adducts can readily be attributed to the cysteine of interest rather than to other surface-exposed cysteine residues.

Removal of surface cysteine facilitated the screening of caspase-3, a heterotetrameric enzyme composed of two small and two large subunits. The most reactive cysteine residue on the large subunit was the active-site cysteine (C163), however a second cysteine on the small subunits was fully exposed in the existing caspase-3 crystal structure and was also reactive. It was thus necessary to mutate that residue to serine (C184S) to make interpretation of the mass spectral data unambiguous (unpublished data).

Whereas sometimes overly reactive cysteine residues can confound or complicate a Tethering experiment, there is also at least one example of an introduced cysteine being successfully probed in the presence of an active-site cysteine. In PTP1B a R47C mutation allowed alkylation by an extender molecule preferentially, even though the catalytic cysteine residue (C215) was still present in the protein construct. This is likely because C215 sits at the bottom of a deep hole,⁹ and can thus be protected from easy disulfide exchange. This recapitulates the results of a naïve screen of caspase-3, where no hits were found against the active-site cysteine residue. The caspase-3 active site is believed to be very mobile and mostly unstructured in the absence of substrate. Optimal caspase substrates contain a negatively charged aspartate moiety. Since the Tethering library was designed to have drug-like properties and charged entities are usually considered to be non-drug-like, it is not surprising that no hits against the caspase-3 active site were isolated. Together the examples of PTP1B and caspase-3 underscore the conclusion that not all active-site cysteines are structurally optimal for a Tethering reaction to occur. On the other hand, many introduced cysteines are perfectly competent for capturing hits via Tethering.

The vast majority of Tethering experiments have been on proteins of known structure so that sites for Tethering could be chosen based on inspection of the binding pocket to be probed. Particularly when introducing cysteine residues, it is very useful, if not obligatory, to have structural information available. One question of direct interest is whether Tethering is also applicable to proteins where a high-resolution structure is not available. G-Protein coupled receptors (GPCRs) are of tremendous importance in drug discovery as up to half of marketed drugs target GPCRs,²¹ however structures exist for just a handful of

GPCRs. For example, there is no structure available for the C5a receptor. A homology model of the C5a receptor based on the rhodopsin structure has been published²² and conclusive mutagenic studies have also pointed to the location of a ligand-binding pocket.^{22–24}

With only indirect structural data on the binding surface of the C5a receptor in hand, Buck and Wells opened a new avenue for Tethering by asking whether this was sufficient structural detail to mount a successful Tethering campaign.⁸ They first sought to determine that the location implicated by mutagenesis was, indeed, the ligand-binding site. Based on the homology model, they selected four sites at which cysteine residues were introduced.⁸ Previous data had also indicated that six amino acid peptides from the C-terminal sequence of the C5a ligand were themselves sufficient to agonize or antagonize the receptor, dependent on their precise sequence.²² Buck and Wells used the three amino acid sequence of the agonist or antagonist peptides and added an N-terminal cysteine residue. The homologous three amino acid peptides themselves were too short to either agonize or antagonize C5a like the six amino acid peptides had. The addition of the cysteine allowed for Tethering of these peptides to the cysteine-containing receptors. The four amino acid cysteine-containing peptides maintained the agonist or antagonist properties of the parent peptides, and also confirmed the location of the ligand-binding site to be near residues P113, G262, and L117. Thus, they concluded that these residues were appropriate residues for a screen for small molecules by Tethering.

In the case of C5a, four mutant cysteine positions were each screened against a library of 10 000 compounds. The three sites that were most appropriate for binding of the cysteine-containing peptides – P113C with 65 inhibitors, G262C with 36 inhibitors, and K117C with 24 inhibitors identified – were also successful for identifying small-molecule agonists and antagonists of C5a activity. An interesting feature of the C5a receptor is that structurally similar peptides can function as either agonists or antagonists. Of the small-molecule inhibitors identified by Tethering nearly equal amounts functioned as agonists and antagonists of C5a ligand binding. This is particularly interesting given that the small-molecule inhibitors are only 1/4 to 1/2 of the size of the minimally active peptide inhibitors.

17.4 Cooperative Tethering

Cooperative Tethering utilizes engineered cysteine residues on proximal regions to interrogate binding fragments that interact cooperatively with one another or to interrogate cooperative interactions between known binding elements and newly discovered disulfide-containing monophores. Many binding sites are very adaptive, which, stated differently, means that the energy barriers between various conformations are very small so a relatively small change in the system (*i.e.* binding of a Tethering monophore) can favor another conformational state. In these adaptive sites, it is therefore not surprising that there is a great

deal of inherent cooperativity that can be accessed during binding at adjacent sites.

Cooperative Tethering has been most successfully applied in the case of IL-2.¹⁶ IL-2 has historically been classified as a difficult target for drug discovery because drug molecules would have to bind to the interface of IL-2 that interacts with the IL-2 receptor. Compared to enzyme active sites that are often designed for binding and recognizing small molecules and that often contain rich functionalities in the catalytic residues, protein surfaces that promote contact with other proteins are comparably featureless and flat (Figure 17.4). Methods for probing such surfaces are few. Tethering is one of the few direct ways that a protein–protein interface can be specifically probed. Use of cooperative Tethering led to the development of a series of compounds that bound to IL-2 and strongly antagonized IL-2 receptor binding.

The interrogated IL-2 binding site had previously been identified as being involved in the IL-2–IL-2 receptor interface. This binding site was in an adaptive region, the conformation of which changed upon binding to non-peptidic small molecules such as **1** (Figure 17.4).⁵ Hyde and coworkers observed that the adaptive region extended beyond the residues that were occupied by early compounds, and selected several adjacent residues for cysteine substitutions (N30C, Y31C, and N33C) to probe a larger portion of the adaptive region with Tethering.¹⁶ These residues were selected from the crystal structure and none had any effect on the structure or function of IL-2.

Screening IL-2 by Tethering at these residues yielded 132 hits in this extended adaptive region. To probe the likelihood that these fragments could be coupled with **1** to generate a stronger binding inhibitor, binding of the identified disulfide-containing small molecules was tested in the presence of **1**. Fully 33% of the disulfide compounds showed an increase in conjugation efficiency (positive cooperativity), 44% showed a decrease in conjugation efficiency (negative cooperativity), and 23% showed no change in conjugation efficiency (no cooperativity). This cooperativity was confirmed using two-site Tethering with a disulfide-containing version of **1** that could be covalently coupled to the K43C mutation in IL-2. This allowed direct assessment of the fractional binding of **1** and hits from Tethering. Surface plasmon resonance (SPR) also confirmed the interaction between **1** and disulfide-free versions of the compound identified by Tethering and demonstrated that the cooperativity was reciprocal between **1** and hits derived from Tethering. Moreover, these non-covalent results emphasized that the binding of **1** did not simply change the accessibility of the disulfide required for Tethering, but induced important changes in the binding surface that contributed to binding of the monophore. The cooperativity between **1** and hits from Tethering was significant and contributed $-2 \text{ kcal mol}^{-1} \Delta\Delta G$ to synergistic binding.

Ultimately, several classes of compounds were developed by combining known fragments, such as **1**, and fragments discovered by Tethering.^{5,16–18,20} The most potent of these displayed a 60 nM IC₅₀.¹⁷ The crystal structure of this compound with IL-2 revealed that the compound induced massive structural rearrangement to promote binding to the adaptive region,¹⁹ emphasizing why

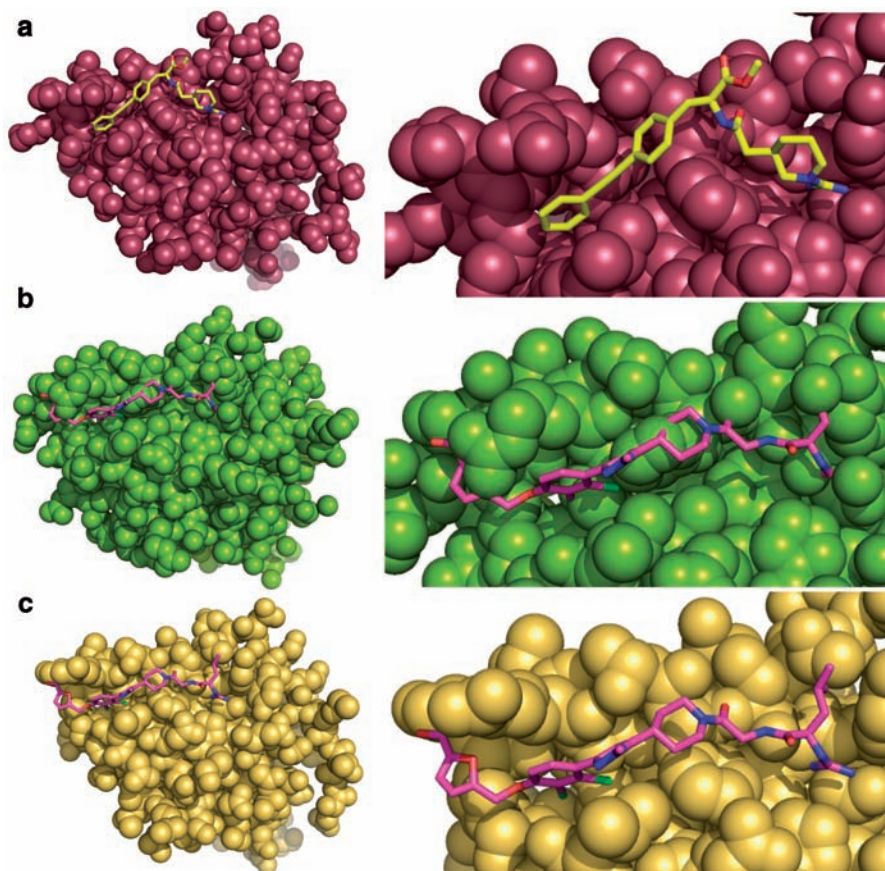


Figure 17.4 The adaptability of the IL-2 surface. (a) The structure of IL-2 (red) bound to **1** (yellow sticks) shows conformational changes in the adaptive site, which is occupied by the hydrophobic biaryl moiety (PDB ID 1M48). (b) The steric clashes that occur when a 60 nM inhibitor (pink sticks) is superimposed on the unliganded IL-2 structure (PDB ID 1M47, green). (c) Many conformational changes are wrought by the binding of the 60 nM inhibitor derived from tethering (PDB ID 1PY2, pink sticks on yellow protein surface), which carves out its own binding interface.

such strong cooperativity was observed. This work emphatically demonstrates that cooperative Tethering is a useful way to develop potent inhibitors, especially on dynamic regions of a protein surface.

17.5 Extended Tethering

The idea of Tethering with extenders (also called extended Tethering in some publications) is that a disulfide-containing small-molecule anchor can be used

in the active site or binding site to probe adjacent regions for chemical moieties that bind a little further afield. This anchor is often an irreversible alkylator that reacts specifically with a single cysteine (native or engineered) on the protein of interest.

Extended Tethering has most widely been used against the caspases where the active-site cysteine is prime for alkylation by an “extender” molecule. Caspases have high specificity for cleaving after aspartate residues so extenders in which a thiol-reactive alkylating agent is placed adjacent to an aspartic acid moiety readily and specifically modify the active site. Although the 13 human caspases may all prove to be valuable drug targets, extended Tethering has, to date, only been performed with caspase-3 (an apoptotic executioner caspase)¹² and on caspase-1 (an inflammatory caspase).¹⁴ Because of a family-wide specificity for binding to aspartic acid elements, the same extender could be used for both caspase-1 and caspase-3. These extenders were based on the structural and biochemical knowledge that all caspases need an acidic moiety to mimic the natural aspartate substrate that binds in the S1 pocket. Whether from a substrate or from some other source, an acidic moiety is necessary to nucleate formation of a correctly positioned active site and catalytic residues. Small, charged molecules, such as malonate, are capable of nucleating this conformation.²⁵ The extender used for both caspase-1 and caspase-3 mimics the natural substrates. Because the methods and results are similar, we focus here on the work on caspase-1, which is more recent. Nevertheless, the dramatic success of extended Tethering on caspase-1 relies heavily on the successful precedent set by extended Tethering in caspase-3, in which a salicylic acid discovered from an extended Tethering screen combined with medicinal chemistry optimization resulted in a series of inhibitors with K_{iS} of 20 nM.^{10–12}

Tethering with extenders was essential in the development of potent inhibitors of caspase-1.^{13,14} After modification of the extender, the crystal structure of caspase-1 demonstrated that the extender specifically alkylated the active-site cysteine, and not any of the other four cysteines in the same subunit. The structure also showed the extender pointing, as predicted, toward the outer edge of the binding cavity (this region is termed the S3 and S4 binding pockets because the third and fourth amino acids of the peptide typically bind here, Figure 17.5). Screening by Tethering identified ten fragments that bound with high affinity. When these fragments were converted to non-covalent analogs, however, the affinity was poor, generally greater than 100 μ M, demonstrating that their affinity was dependent on the interactions between the extender and the S1 pocket in the active site. The extender itself exhibited a binding affinity of only 110 μ M.

One of the strengths of extended Tethering is that it suggests a defined way that the extender plus the hit can be converted to a useful inhibitor. The alkylating cyclooxy-methylketone on the extender can be replaced by an aldehyde, which is reversible although covalent. The disulfide region can be converted atom by atom for replacement by methylene units. This procedure resulted in an initial lead compound with 150 nM affinity for caspase-1 and 57-fold selectivity over caspase-5.

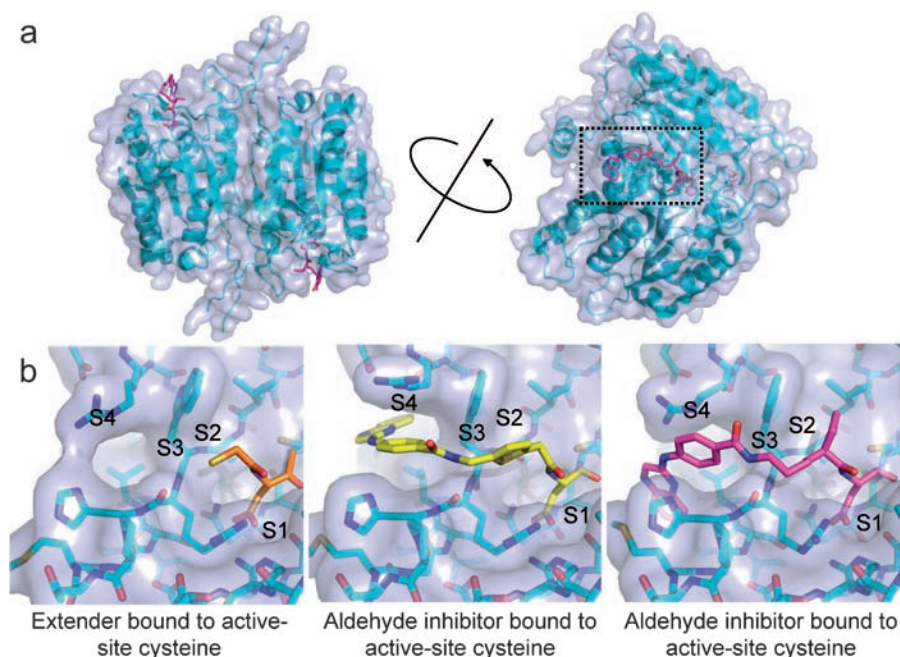


Figure 17.5 Extended Tethering in caspase-1. (a) A typical view of caspase-1 (ribbons) with an aldehyde inhibitor bound in the two active sites. Rotation around the indicated axis orients caspase-1 to look into one active site. The zoom region shown in (b) is marked by the dashed square. (b) The extender bound to caspase-1 occupies the S1 pocket (PDB ID 1RWK, left panel). S1–S4 binding pockets are marked. A compound derived from tethering generated by atom-for-atom replacement and addition of an ethyl moiety to bind in the S2 pocket results in a 150 nM inhibitor that orients the terminal tricyclic moiety in a downward orientation in the S4 pocket (PDB ID 1RWN, central panel). On the other hand, replacement of the disulfide with a bulkier thiophene group that does not enter the S2 pocket results in a 340 nM inhibitor that orients the terminal tricyclic moiety in a upward orientation in the S4 pocket (PDB ID 1RWM, right panel).

Guided by work in caspase-3 extended Tethering, the caspase-1 investigators knew that additional affinity could be added by modification of the region that bound to the S2 pocket. Based on the specificity of caspase-1, hydrophobic groups were introduced in this region. These additions resulted in sets of related compounds that contained the same constituents for binding to the S3 and S4 pockets, but which could bind to two different cavities outside of the S4 pocket that had not been previously exploited for drug discovery. Viewing the SARs of the compounds, it became clear that modification of the region binding to S2 was critical in determining the conformation of the distal end of the inhibitors. This suggests a cooperative coupling between the S2 and S4 regions of the

caspase-1 binding site that had not previously been recognized and that does not seem to be present in caspase-3. The success of extended Tethering in this situation is that the tricyclic fragment discovered by Tethering that led to the highest affinity inhibitors had modest-enough affinity on its own that the authors concluded it could not have been derived from any traditional functional screen. In this case Tethering was seminal in the identification of this fragment.

In the caspases, a known binding moiety was used as the extender, but another variation on Tethering with extenders would utilize a high-affinity hit, derived from naïve Tethering, as the extender to probe distal regions of the binding pocket. In principal this could be repeated more than once, generating longer and longer extender molecules. The presumptive reason this has not been attempted is that third-generation extended tethers would be too long and have too high a molecular weight to generate a fruitful drug lead. Another future development that will be exciting to witness is the application of Tethering with extenders to an *in situ* situation, a course that is ripe with possibilities.

In early studies on caspase-3, naïve Tethering against the active-site cysteine did not result in any hits. Why, then, was Tethering with extenders so successful? A lack of hits from naïve Tethering is likely both because the length of molecules present in the compound collection was insufficiently long to bypass the P1 pocket, and because it had been culled to contain fragments with drug-like properties (charged molecules are generally considered to be non-drug like), so the library contained very few fragments composed of aspartate mimics. For both caspase-3 and caspase-1 it is clear that a charged fragment that mimics the aspartate that is present in all caspase substrates is necessary to enhance binding affinity.

17.6 Breakaway Tethering

In proteins with cysteines at the active site, native-cysteine Tethering is of great utility. Tethering with extenders has been useful for caspases, which contain active-site cysteine, largely because the available chemistries for modifying the active-site cysteine are robust and specificity elements for substrate binding are near the catalytic residue. One concern with engineering cysteine residues in active sites or binding sites is that the residue that would ideally be mutated for fragment discovery may be important for binding or catalysis. To circumvent these complications, Erlanson and co-workers developed a strategy they termed breakaway Tethering,⁹ which is useful for probing sites that are narrow, deep, fragile, or significantly and negatively impacted by the introduction of a cysteine residue for Tethering. In short breakaway Tethering is ideal for any protein where modification within the active site is not desirable.

PTP1B, a negative regulator of insulin-receptor phosphorylation and signaling, a pharmaceutical target for type-2 diabetes, is such a protein. The active site of PTP1B is deep, and because it binds to doubly charged phosphotyrosine

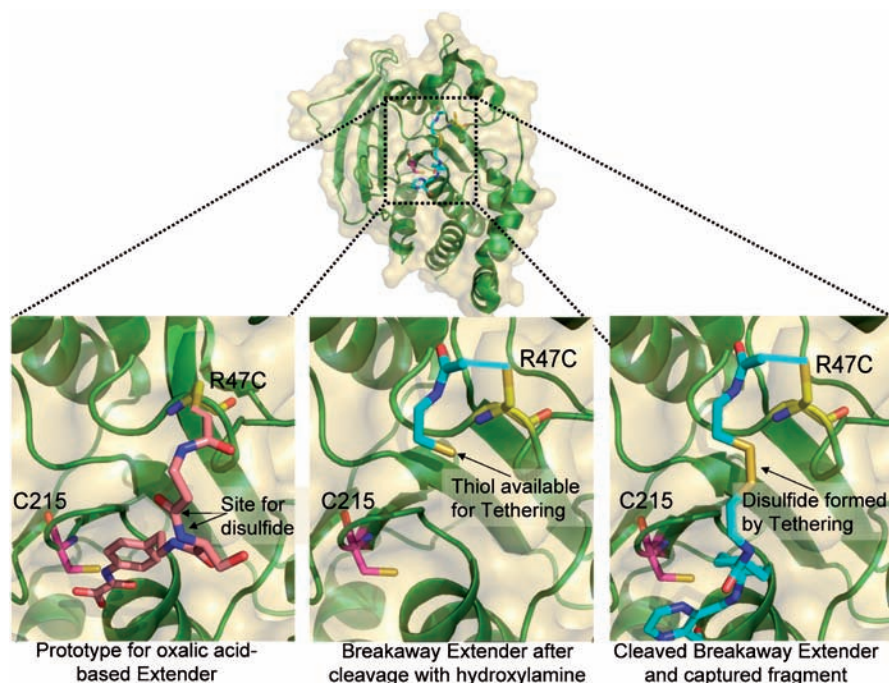


Figure 17.6 Breakaway Tethering in PTP1B. PTP1B (green ribbons, PDB ID 1NWE) is shown modified with a prototype oxalic acid-containing extender molecule attached to a cysteine residue that has been specifically engineered for Tethering (R47C, lower left panel). The prototype extender was the inspiration for the breakaway extender, which contained a disulfide bond at the positions equivalent to the carbons marked with arrows at the site for disulfide. Unfortunately, no structure of the oxalic acid breakaway extender is available. Treatment of the breakaway-extender-modified PTP1B with hydroxylamine exposes the free thiol which can then be interrogated with a library of disulfide-containing compounds (lower center panel). Compounds that bind to the active-site pocket, which is in a deep and fragile cavity, are able to form a disulfide bond with the cleaved extender, as was the case for the fragment captured in the PTP1B active site in the lower right panel (PDB ID 1NWL).

residues, could be severely impacted by introduced cysteines in the pocket, so it was a good candidate for the development of breakaway Tethering.⁹ To avoid disruption of the binding pocket, a cysteine residue was introduced outside of the binding pocket (R47C), which was roughly 10 Å from the active site and which was predicted to point toward the active site when mutated (Figure 17.6). This introduced cysteine could be alkylated by a breakaway extender. The breakaway extender was designed as a derivative of a known active-site binder oxalic acid, a phosphotyrosine mimetic. The oxalic acid was linked to the extender via a thioester so that it could be cleaved under mild conditions by exposure to hydroxylamine, leaving the extender conjugated to the distal

cysteine.⁹ It was impressive that this approach worked despite the fact that PTP1B contains an active-site cysteine. Alkylation occurred preferentially at the introduced cysteine, probably because the active-site cysteine was protected from alkylation by the oxalic acid. The charges on the oxalic acid mimic phosphotyrosine so that binding occurs preferentially in the direction that orients the extender toward the external cysteine at position 47. This again underscores the importance of non-covalent binding of the monophore driving the formation of the covalent linkage.

Once the oxalic acid was released, the extender-modified PTP1B was used to screen the library of disulfide-containing compounds. Novel phosphotyrosine mimics, of different classes than had been previously discovered through medicinal chemistry efforts or from traditional high-throughput screening (HTS), were ultimately developed based on the breakaway Tethering effort. In the crystal structures of the hits observed from breakaway Tethering, the monophores all sat in the deep active-site pocket as expected. The non-covalent monophores competitively inhibited PTP1B with a K_i of 4.1 mM, which is a notable improvement over extant phosphotyrosine mimetics with K_i s of greater than 10 mM and comparable to phosphotyrosine, the native substrate, which binds PTP1B with a K_m of 4.9 nM. This work established breakaway Tethering as yet one more useful adaptation of a clearly powerful technology, enabling discovery efforts in one more category of target – those with fragile active sites.

17.7 Discovery of Novel Allosteric Sites with Tethering

Within the past several years, several prominent drugs that have entered the marketplace have been structurally shown to act at allosteric sites [*e.g.* Gleevec (Glivec) which binds to c-Abl²⁶ and the human immunodeficiency virus (HIV) non-nucleoside reverse transcriptase inhibitors (NNRTIs)²⁷]. These developments have dramatically increased interest in exploiting allosteric sites in other proteins. Tethering is an ideal technology for both discovery and exploitation of novel allosteric sites. Tethering has been applied to the apoptotic executioner caspases, caspase-3 and -7. Caspases derive their name from their properties as cysteine aspartate proteases because they contain an active-site cysteine and cleave substrates after aspartate residues. An initial screen of caspase-3 against the disulfide-containing compound library using mass spectrometry as a read out was expected to elucidate compounds that bound the caspase-3 active-site cysteine in the large subunit. (Active caspases are heterotetramers composed of two large and two small subunits. The catalytic histidine–cysteine dyad is in the large subunit.) Surprisingly, no compounds were found that bound to the large subunit. One compound, DICA, was identified that bound strongly to a native non-active-site cysteine (C264) in the small subunit. In a separate functional screen for inhibitors of the intrinsic apoptotic pathway (see Section 17.1) another small molecule (FICA) was identified that bound to C264. This small-subunit cysteine exists in the bottom of a deep cavity at the dimer

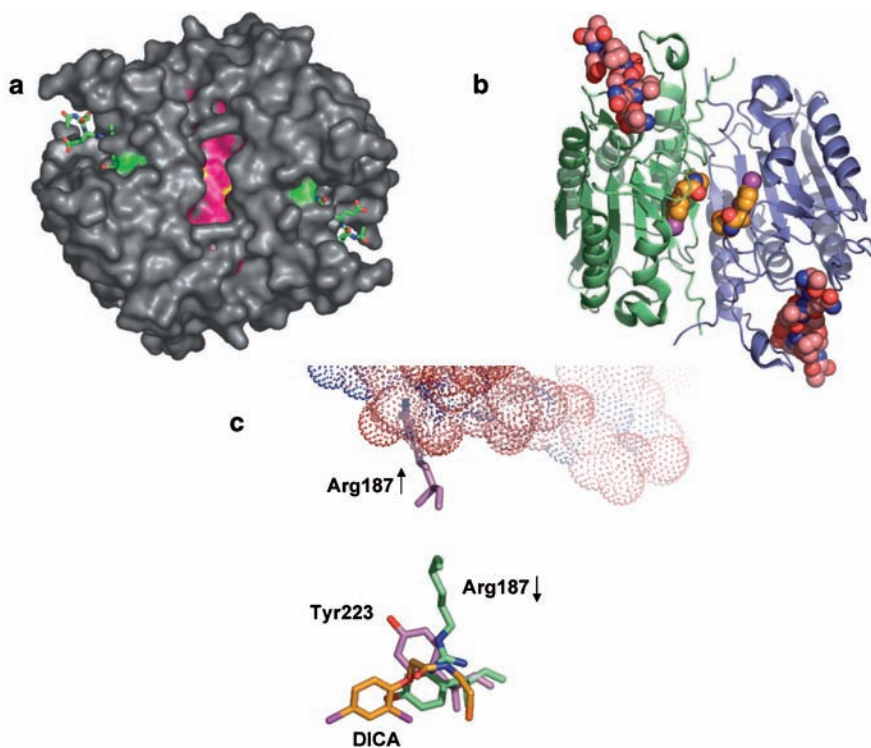


Figure 17.7 Allosteric Tethering in caspase-3 and -7. (a) The structure of caspase-7 (PDB ID 1F1J) bound at the active-site cysteine (green surface patch) to peptide-based inhibitor DEVD-FMK (green sticks) shows a large cavity (pink surface patch) at the dimer interface. Two native cysteine residues (yellow surface patches at C290 in caspase-7) are present in the cavity and were productive for tethering in the mass spectrometry-based and functional assays. (b) The new caspase-7 allosteric binding site (yellow spheres for compounds in the allosteric site) is at the dimer interface of the two caspase monomers (blue and green) and is spatially distant from the substrate-binding groove (peptide inhibitor in orange spheres). (c) Allosteric inhibitor DICA (yellow sticks) sterically clashes with the active conformation of caspase-7 (green sticks) and forces caspase-7 into an inactive conformation (purple sticks). This is accomplished when DICA contacts Tyr223 (green sticks), forcing Tyr223 (purple sticks) into the up position. This conformation of Tyr223 is in steric contact with Arg187 (green sticks), which subsequently causes Arg187 (purple sticks) to the up position. The up position of Arg187 (purple sticks) observed in the presence of the allosteric inhibitor is incompatible with binding of peptide (orange dots) in the active site.

interface (Figure 17.7a). Both of these compounds acted as covalent stoichiometric inhibitors of caspase-3 and caspase-7 activity.

The crystal structure of caspase-7 with the allosteric inhibitors revealed the mechanism of inhibition (Figures 17.7b and 17c). The stoichiometric nature of

inhibition was apparent when the crystal structure of these compounds was compared with the structure of caspase-7 in the presence of a substrate mimic (the covalent inhibitor z-DEVD-FMK). When DEVD binds to cleaved caspase-7, the presence of peptide orders the loops that together compose the substrate binding cleft. Part of that ordering includes the movement of the L2 loop (the N-terminal end of a loop that is cleaved to convert caspase from the inactive zymogen to the cleaved and active caspase) toward the core of the protein, burying an arginine residue (R187, immediately adjacent to the active-site residue C186) in the core of caspase-7. When R187 is in this downward conformation, it sterically constrains the position of tyrosine Y223 into a downward conformation. This active conformation is incompatible with binding of FICA or DICA at the dimer interface cavity. When FICA or DICA bind to caspase-7 they constrain Y229 to adopt only the up conformation. This Y229 conformation is incompatible with binding of substrate because there is not room for R187 to bury in the protein core.

An additional layer of inhibition is also suggested by the FICA and DICA crystal structures. The binding of FICA and DICA appears to be driven nearly entirely by hydrophobic (entropically driven) interactions, partitioning the small molecules into the solvent-protected cavity. The hydrophobic nature of the compounds is certainly influential in their further protection from solvent by the conformation of one of the active-site loops (L2', the C-terminal side of the loop cleaved upon conversion of the zymogen to the active caspase). In addition to interactions between FICA and DICA with the dimer cavity, there are strong hydrophobic interactions of the residues on the L2' loop with the compounds. Either of these mechanisms appears to be sufficient to inactivate caspases: caspases that are incapable of binding substrate because of indirect competition between the L2 loop residue 187 burial and the allosteric site are inactive; caspases with the L2' loop locked over the allosteric site instead of in the position that can allow substrate binding are also inactive.

An allosteric site in PTP1B, a negative regulator of the insulin receptor, was also discovered using Tethering, but in a somewhat less traditional format. In the context of searching for alkylating agents that could specifically label the active-site cysteine residue in PTP1B for use in Tethering with extenders, one cysteine reactive compound, ABDF, was discovered.¹⁵ Modification of PTP1B by ABDF was rapid, reversible, and quantitative. Many of the other alkylating agents tested were equally reactive with three or more of the cysteine residues in PTP1B, presumably two solvent-exposed cysteines C32 and C92 and the active-site cysteine (C215). ABDF was unique and noteworthy in that it selectively modified just one residue (C121, Figure 17.8), initially assumed to be the active-site cysteine, since the active-site cysteine has a pK_a of ~ 5 .

Binding of ABDF quantitatively to PTP1B resulted in a 7.4-fold decrease in V_{max} , with no significant change in K_m . ABDF binding also did not result in full catalytic inhibition, even at quantitative labeling conditions, leading Hansen and co-workers to suspect that ABDF was not binding to the active site.¹⁵ Electrospray ionization (ESI) mass spectrometry and peptide mapping confirmed that ABDF was selectively modifying the non-active-site residue C121.

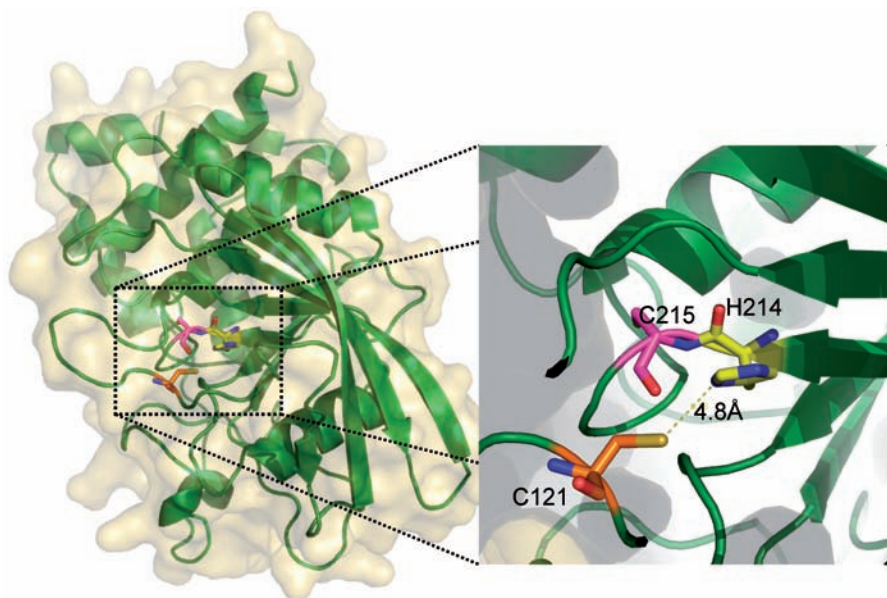


Figure 17.8 PTP1B ABDF-binding site. The serendipitous allosteric site discovered using ABDF at C121 is depicted as orange sticks. In this unliganded structure, C121 is pointed into the core of the protein. It is likely to change conformation upon binding to ABDF. The active-site cysteine (C215 pink sticks) and adjacent histidine (214 yellow sticks) are 4.8 Å from the cysteine that is modified.

C121 is conserved in most related phosphatases, including LAR and TCPTP. ABDF also inhibits LAR and TCPTP in a time- and dose-dependant manner, as would be expected for a rapidly binding, covalent compound. ABDF does not inactivate CD45, suggesting that there is some selectivity of binding and inhibition of ABDF. Although PTP1B readily crystallizes in the absence of ABDF or in the presence of other allosteric inhibitors which lock the WPD loop in an inactive conformation,²⁸ crystals of PTP1B in complex with ABDF could not be obtained, so the detailed mechanism of inhibition is not clear. Nevertheless, looking at the structure of PTP1B it is clear that the cysteine side-chain of C121 points toward the core, making hydrophobic interactions with Y124, which is in a hydrogen bonding network with H214, the residue adjacent to the active site. The inward direction of C121 in the crystal structures would not have been predicted to be ripe for tethering. Thus, conformational flexibility of PTP1B must allow the C121 thiol to become solvent accessible. It seems likely that binding of ABDF causes conformational changes in this region which, in Rube Goldberg fashion, is similar to the situation with FICA and DICA bound to caspase-7, where the conformation of the active site is disturbed. In contrast with the caspase-7 allosteric mechanism, any changes conferred to the active site do not have any effect on substrate binding.

Discovery of allosteric sites involving cysteine residues is intriguing because it immediately suggests a mechanism by which native proteins could be targeted for drug discovery.

From the point of view of the discovery of new allosteric sites, it is interesting to note that examination of the structure of caspase-3 or -7 could have suggested that this site would be useful for small-molecule binding given its size, concavity, and through-protein distance to the active site (13 Å). It is likely that the new allosteric site in caspase-3 and -7 could have been selected based on geometric properties that were apparent in structure of caspase-7 with an active-site inhibitor or with no ligands bound.^{29,30} The ideal future development of Tethering at allosteric sites is the ability to identify previously unexplored sites based on geometrical and proximity considerations and then exploit them as serendipitous allosteric sites.

Traditional HTS is poorly suited to probe allosteric sites such as the caspase and PTP1B allosteric sites because there is no mechanism to either target compounds toward any particular site, nor is there a ready means to determine the site of interaction. Tethering is by far the drug-discovery tool best suited to probe this site. Tethering has the unique ability first to validate the ability of this site to propagate an allosteric signal to the active site and second to identify fragments that bind to and modulate that site.

17.8 Tethering as a Validation Tool

Small-molecule hits derived from HTS that do not act as quantitative competitive inhibitors are often discarded because they have a higher-than-acceptable rate of non-drug-like or artifactual inhibition (e.g. McGovern *et al.*^{31,32}). This means that many legitimate and potentially useful allosteric inhibitors have likely been discarded as well. Because Tethering can probe protein interaction surfaces in a site-specific manner, it is the perfect tool for studying the location and mechanism of inhibition or activation of a small molecule that does not work as a competitive active-site inhibitor. In principle any small molecule could be converted to a disulfide-containing molecule and used with a series of cysteine mutants to determine the site of interaction. In the case of IL-2, a crystal structure of the complex between **1** and IL-2 had been determined so it was possible to engineer a disulfide-containing version of **1** with a three-methylene spacer that was competent to bind to a K43C mutant of IL-2 and antagonize binding in the same way that compound **1** behaved. It is likely that this same conversion would be successful in many other protein–small-molecule pairs.¹⁶ In practice, if the site of interaction is completely unknown it would be a colossal task to find the site of interaction *ab initio*. Nevertheless, when some evidence suggests a site of interaction, Tethering is a useful technique for defining the site of interaction. Perhaps the best example of Tethering being used as a validation tool is with C5a receptors.

As previously noted, GPCRs are the most frequently targeted class of proteins for pharmaceutical control. Unfortunately, this class of proteins has

been recalcitrant to structural studies due to difficulties of crystallization of integral membrane proteins. Tethering facilitates the use of disulfide capture as a means of determining the location of binding sites in the absence of other concrete structural information on C5a⁸ (see Section 17.3). Tethering is particularly well-suited for determination of the site of binding when determination of a crystal structure of the complex is impossible to attain. The work on C5a validated the proposed site of peptide interaction on the C5a receptor and underscores the strength of Tethering to determine the site of action in difficult classes like GPCRs. For proteins interacting with peptides, this approach is readily accessible.

Tethering was also useful in the validation of two sub-sites within the interface region of IL-2 that contacts the IL-2 receptor. Based on the number and type of hits that were observed across 12 engineered cysteine sites, the adaptive region was dissected into a rigid sub-site and an adaptive sub-site. This kind of information could be useful in directing lead optimization work in predicting what types of fragments are more likely to be successful.

17.9 Tethering vs. Traditional Medicinal Chemistry

Two hallmarks of Tethering are that it can probe regions of protein space that are unapproachable for all practical purposes by traditional HTS combined with medicinal chemistry. Tethering can also rapidly identify binding fragments that would not be suggested by other methods because of the increased dynamic ranges of interaction energies accessible through Tethering. For example, a team applying Tethering to PTP1B were able to develop a new aryl-oxalamic acid pharmacophore in a way that was not obvious from traditional medicinal chemistry, and was much more rapid than would have been possible without Tethering.⁹

A previously addressed strength of Tethering is the predictive ability in how to recombine discovered fragments from screening. In studies on caspase-1, direct atom-for-atom replacement of atoms in an extended Tethering scheme gave the highest inhibition constant for caspase-1,¹³ demonstrating the exquisite orientation specificity of the compounds selected by Tethering. However, when a series of linkers was introduced to replace the linker, the compound with the highest affinity was a benzenoid linker that had previously been exploited by Cytovia and Vertex in their development of caspase-1 inhibitors. Thus, even after using Tethering, a traditional medicinal chemistry approach is often required to improve the affinity of the monophore element. Is Tethering categorically better than traditional medicinal chemistry? Tethering certainly enjoys some distinct advantages over traditional medicinal chemistry, specifically in predicting recombination orientations.

Is Tethering faster? Tethering, particularly extended Tethering or breakaway Tethering, often provides a more direct means of performing SAR studies because it does not require synthesizing each compound by hand. In the case of IL-2 the results of Tethering suggested a focused set of just 20 compounds from

which resulted an improvement in binding affinity from 3 μM to 60 nM.¹⁷ In contrast, when those same compounds were interrogated with the standard medicinal chemistry approach, the binding affinity was significantly reduced. Even after extensive medicinal chemistry interrogation in various regions of the compound, the only approach that yielded a dramatic boost in affinity was the addition of a furanoic acid fragment that was discovered by Tethering from an adjacent residue.²⁰

One arena where Tethering again offers a distinct advantage over other methods is that of non-active site binders. In these situations it is frequently difficult to conclude convincingly that the binding mode and mechanism are consistent with the inhibition or activation observed. When compounds are covalently linked via a Tethering interaction, much of the ambiguity is diminished. In addition, it is possible to derive structural information much earlier in the move from hit to lead with compounds that bind weakly. Because of disulfide stabilization, it is much more likely for the crystal structure of a low-affinity compound derived from Tethering to be determined than it is for one derived from a traditional high-throughput screen.

17.10 Tethering in Structural Determination

Determining the X-ray crystal or nuclear magnetic resonance (NMR) structure of small molecules in complex with the proteins to which they bind is often challenging. The concentrations of the complex required for crystallization or NMR spectroscopy are often prohibitively high and small-molecule inhibitors are often not soluble at concentrations required for crystallization or NMR. Even when compound solubility is not an issue, incomplete occupancy of the binding sites in all proteins in the sample can thwart structure determination.

The presence of small molecules can also affect crystal formation. In some cases, protein crystals will grow in the absence of a small-molecule inhibitor (or activator) but not in the presence of one, so the small molecule must be soaked into the crystalline lattice. Frequently this soaking disrupts and damages the crystal, so this is not a foolproof method for complex crystallization.

Crystallization of covalent complexes of compounds derived from Tethering offers several advantages. Because a covalent complex is formed stoichiometrically, one can be assured of full occupancy of all binding sites before the protein is crystallized or the solution structure determination proceeds. This confirmed occupancy dramatically increases the success of crystallization of compounds derived from Tethering. Additionally, because the investigator knows the site of modification the search for the location of the bound compound is simplified, which is especially helpful if the compound does not bind to the active site. Spurious binding of non-covalent small molecules to weak secondary sites sometimes occurs, but with disulfide-bound molecules this is not much of a risk, because small amounts of reducing agent can be included in the crystallization conditions to prevent any spurious binding.

17.11 The Challenge of Covalency

The advantages of Tethering as a tool for drug discovery in site determination, site-directedness, and the ability to combine monophores are clear. The most formidable challenge of the Tethering method is that of converting covalent fragments identified by Tethering to non-covalent molecules with drug-like properties. This necessary conversion requires empirical determination of a useful substituent to replace the disulfide moiety. This process usually requires the work of a team of medicinal chemists, so that hits derived from Tethering face some of the same challenges as hits derived from other HTS processes in terms of the effort required to produce a lead compound that can be used in clinical trials from a validated hit.

The success of converting a covalent disulfide-containing compound to a useful non-covalent compound depends exquisitely on the interaction energy of the monophore with the binding site on the protein. If the interaction energy of the monophore with the binding site is very high, then binding of the compound to the binding pocket is relatively free of the requirement for disulfide-bond formation to drive complex formation. For these compounds conversion to a non-covalent compound is very straightforward. Synthesis of an analog free of the disulfide bond is sufficient for conversion to a non-covalent analog. Although hits derived from the adaptive region of the IL-2 binding surface that were converted to non-covalent analogs by substitution of a methylene cap were competent to bind to IL-2 as assessed by SPR,¹⁶ in general examples of this simplistic type of conversion are rare. By definition and empirical observation through screening at various reductant concentrations, hits derived from Tethering have sufficient interaction energy between the monophore and the binding pocket that their resident time at the binding pocket promotes formation of a covalent disulfide bond between the compound and the protein. Nevertheless, for most Tethering hits, the interaction energy of the monophore with the binding pocket is not sufficient to promote meaningful binding on the time scale of biological assays in the absence of the covalent tether. In these situations when disulfide-free analogs are synthesized the biological activity (*e.g.* inhibition of caspase activity) is no longer observed. In these cases, the monophore from Tethering can be used as an anchor for extended Tethering, or as a starting point for traditional synthetic medicinal chemistry efforts to generate analogs with improved binding affinity.

A related challenge of covalency is the ability to rank interaction energies of covalent hits in a way that predicts the interaction energies of their non-covalent counterparts. A standard method is via measurement of a β -ME₅₀ value (the β -ME concentration at which 50% of the protein in the sample is conjugated by a disulfide-containing compound, at a fixed compound concentration) or via measurement of a DR₅₀ (the disulfide-containing compound concentration at which 50% of the protein sample is covalently conjugated by the compound, at a fixed reductant concentration). Both of these values are straightforward to measure using standard mass spectrometry, however neither of these measures has been shown to have a strict relationship to the energetics

of binding of the disulfide-free monophore. This is likely because the hydrophobic nature of the linker region contributes differently to the overall binding energetics for various compounds and their associated binding modes. This may be particularly true for introduced cysteine residues that can access more than one region of a binding site (*e.g.* caspase-7 FICA *vs.* DICA⁷ or IL-2.¹⁶)

Using extended Tethering, an irreversible warhead such as a chloro- or fluoro-methylketone attacks the active-site cysteine. This necessitates a modest medicinal chemistry effort to convert hits from extended Tethering to either non-covalent or reversibly covalent molecules. The caspase-1 team found that hits discovered from extended Tethering could be routinely converted from covalent to non-covalent molecules by simply performing atom-for-atom replacement of the atoms in the linker, and by replacement of the disulfide with an aldehyde moiety.¹⁴ Aldehydes have also been widely used as covalent but reversible electrophiles against cysteine proteases. This replacement approach, while straightforward in terms of maintaining potency against the caspases, is not completely foolproof. Converted non-covalent monophores from Tethering with extenders on caspase-1 revealed two distinct binding modes when the linker was converted with a rigid thiophene linker rather than with an ethylene unit. This region of the compound sits in the hydrophobic S2 pocket and appears to determine the orientation of the molecule in the S4 binding pocket. Thus the binding mode is not uniquely dependent on the monophore, but is also influenced by the linker portion of the molecule.¹⁴

17.12 Hydrophobic Binders

The successes of Tethering are impressive and numerous. One challenge of Tethering is that hydrophobic pharmacophores are often selected. A general rule seems to be that hydrophobicity drives binding affinity while hydrogen bonds and an overall shape complementarity drive specificity. Of the published molecules discovered from Tethering, a high proportion are relatively hydrophobic in nature, or derive their binding affinity largely through hydrophobic interactions. It is not surprising that a large number of small-molecule–protein interactions are driven largely by hydrophobic interactions.

Specific cases include C5a receptor for which the strongest compound was biaryl pyrrolidine. Though the exact mechanism of binding is not yet known, inspection of this molecule suggest that it would make predominantly hydrophobic interactions. The remaining reported hits for C5a were also hydrophobic in nature. When the adaptive region of IL-2 was probed by Tethering, the majority of the compounds selected were hydrophobic in nature.⁵ Similarly, the best compound identified for allosterically inhibiting caspase-3 and -7 was a hydrophobic dichlorophenyl moiety. In the structures of caspase-7 with both FICA and DICA no direct hydrogen bonds between the small molecule and any protein atoms were observed. (The crystal structures were not of sufficiently high resolution to model water molecules accurately, so it is possible that water-mediated hydrogen bonds were formed.) Nevertheless, shape

complementarity to the binding cavity and hydrophobic interactions appear to be the major driving force in this interaction.

Taking the published results of Tethering in sum it seems fair to conclude that hydrophobic constituents are also favored to bind to the more adaptable regions. This was certainly the case for caspase-1 extended Tethering, where the best fragment discovered was a hydrophobic tricyclic moiety. Most of the binding affinity seems to arise from hydrophobic interactions with the residues that make up the binding pocket.¹⁴ Is this apparent bias toward hydrophobic compounds a function of the compound collection or is it the inherent selectivity of the types of cavities that are being probed with Tethering?

Using cooperative Tethering on IL-2, fragments identified against Y31C and L72C were overlapping, meaning that fragments that were effective at Y31C were also effective when bound to L72C. Compound **1** itself was a hydrophobic biaryl moiety that caused a modest rearrangement of the surface relative to the unliganded structure (Figure 17.4). These fragments were predicted computationally to occupy a deep hydrophobic cavity on the surface of IL-2 that has been termed the “adaptive region”.¹⁷ The highest affinity compound produced from these studies (60 nM affinity) did not bind as computationally predicted, but caused a relatively dramatic rearrangement of this surface of IL-2 binding in a groove that had not previously been observed in any crystal structures (Figure 17.4).¹⁹ This is an example of a small molecule carving out a new binding site based on physical, in this case hydrophobic, interactions that are addressable with small changes in the energy of the system. Thus, as with other methods of drug discovery, structure determination is the only unambiguous means to determine the binding mode of discovered fragments, particularly since hydrophobic constituents which tend to partition to the core of the protein are prevalent.

17.13 Conclusions: The Future of Tethering

Tethering has already been successfully applied to a wide variety of protein surfaces (Table 17.1) – somewhat featureless, but adaptable protein surfaces (IL-2), deep and fragile active sites (PTP1B), novel allosteric sites (caspase-3, caspase-7, PTP1B), cysteine-containing active sites (TS, caspases, PTP1B), proteins of known structure (most) and of unknown structure (C5a). Indeed, this might be a case where “more of the same” would be a great advance and we can expect that both natural cysteines and cysteines introduced all over the surfaces of a great variety of proteins will lead profitably to both our biological and pharmacological understanding. Applying the wealth of Tethering technologies to more protein targets promises a rich and fruitful road ahead.

A biological maxim might be that it is more probable to find inhibitory compounds than to find activating compounds. This is probably because there are nearly infinite ways to disrupt protein function (*e.g.* blocking ligand binding either directly or indirectly, disrupting proper protein folding, disrupting dynamics and conformational changes necessary to catalysis or binding) and

any disruption can lead to loss of function. On the other hand, there are relatively fewer ways for a small molecule to activate a protein (*e.g.* bind and shift protein to a catalytically- or binding-competent form, stabilize the binding and transition states of the protein). Is it possible to identify activators using Tethering? Certainly it is possible due to the site-directed nature of Tethering and its unique ability to target and probe protein structure and conformation site specifically, Tethering is perhaps the means of drug discovery with the best chance of developing activators for biologically important process. In fact, both agonist and antagonist small molecules have been derived from Tethering against the C5a receptor.⁶ An exciting future frontier of Tethering will certainly be its application to the discovery of activating compounds.

To date, Tethering using disulfides has only been applied in a large-scale way at Sunesis Pharmaceuticals. Compared with traditional high-throughput compound screening and medicinal chemistry efforts, which have been performed on a large number of protein targets at a huge number of pharmaceutical companies and an increasing number of academic facilities, the application of Tethering has been modest in terms of both numbers of compounds and numbers of protein targets screened [currently publications exist on eight protein targets (see Table 17.1)]. In most publications, the library size reported was 10 000–30 000 compounds, however the fraction of mostly hydrophobic compounds in the compound collection has not been discussed. A larger and more diverse library of disulfide-containing compounds might improve the already marked success of Tethering and could influence the nature of the hits derived from Tethering. Nevertheless, the original intent of developing Tethering was that by directing drug discovery toward the pocket of interest it would be possible to find or generate (through mechanisms such as extended Tethering) high-affinity leads by searching a limited space, and it has been clearly demonstrated on a number of protein targets that this can be done with 10 000–30 000 compounds. Overall, Tethering can be measured as a resounding success both in the development of a robust new strategy for the discovery of protein surfaces that are good drug targets and for small molecules to bind to those sites.

References

1. D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon and J. A. Wells, Site-directed ligand discovery, *Proc. Natl. Acad. Sci. USA.*, 2000, **97**(17), 9367–9372.
2. D. A. Erlanson and S. K. Hansen, Making drugs on proteins: site-directed ligand discovery for fragment-based lead assembly, *Curr. Opin. Chem. Biol.*, 2004, **8**(4), 399–406.
3. D. A. Erlanson, R. S. McDowell and T. O'Brien, Fragment-based drug discovery, *J. Med. Chem.*, 2004, **47**(14), 3463–382.
4. D. A. Erlanson, J. A. Wells and A. C. Braisted, Tethering: fragment-based drug discovery, *Annu. Rev. Biophys. Biomol. Struct.*, 2004, **33**, 199–223.

5. M. R. Arkin, M. Randal, W. L. DeLano, J. Hyde, T. N. Luong, J. D. Oslob, D. R. Raphael, L. Taylor, J. Wang, R. S. McDowell, J. A. Wells and A. C. Braisted, Binding of small molecules to an adaptive protein–protein interface, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**(4), 1603–1608.
6. E. Buck and J. A. Wells, Disulfide trapping to localize small-molecule agonists and antagonists for a G protein-coupled receptor, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**(8), 2719–2724.
7. J. A. Hardy, J. Lam, J. T. Nguyen, T. O'Brien and J. A. Wells, Discovery of an allosteric site in the caspases, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**(34), 12461–12466.
8. E. Buck, H. Bourne and J. A. Wells, Site-specific disulfide capture of agonist and antagonist peptides on the C5a receptor, *J. Biol. Chem.*, 2005, **280**(6), 4009–4012.
9. D. A. Erlanson, R. S. McDowell, M. M. He, M. Randal, R. L. Simmons, J. Kung, A. Waight and S. K. Hansen, Discovery of a new phosphotyrosine mimetic for PTP1B using breakaway tethering, *J. Am. Chem. Soc.*, 2003, **125**(19), 5602–5603.
10. D. A. Allen, P. Pham, I. C. Choong, B. Fahr, M. T. Burdett, W. Lew, W. L. DeLano, E. M. Gordon, J. W. Lam, T. O'Brien and D. Lee, Identification of potent and novel small-molecule inhibitors of caspase-3, *Bioorg. Med. Chem. Lett.*, 2003, **13**(21), 3651–3655.
11. I. C. Choong, W. Lew, D. Lee, P. Pham, M. T. Burdett, J. W. Lam, C. Wiesmann, T. N. Luong, B. Fahr and W. L. DeLano *et al.* Identification of potent and selective small-molecule inhibitors of caspase-3 through the use of extended tethering and structure-based drug design, *J. Med. Chem.* 2002, **45**(23), 5005–5022.
12. D. A. Erlanson, J. W. Lam, C. Wiesmann, T. N. Luong, R. L. Simmons, W. L. DeLano, I. C. Choong, M. T. Burdett, W. M. Flanagan, D. Lee, E. M. Gordon and T. O'Brien, *In situ* assembly of enzyme inhibitors using extended tethering, *Nat. Biotechnol.*, 2003, **21**(3), 308–314.
13. B. T. Fahr, T. O'Brien, P. Pham, N. D. Waal, S. N. Baskaran, B. C. Raimundo, J. W. Lam, M. M. Sopko, H. E. Purkey and M. J. Romanowski, Tethering identifies fragment that yields potent inhibitors of human caspase-1, *Bioorg. Med. Chem. Lett.*, 2006, **16**(3), 559–562.
14. T. O'Brien, T. Fahr, M. M. Sopko, J. W. Lam, N. D. Waal, B. C. Raimundo, H. E. Purkey, P. Pham and M. J. Romanowski, Structural analysis of caspase-1 inhibitors derived from Tethering, *Acta Crystallogr., Sect. F: Struct. Biol. Crystallogr. Commun.*, 2005, **61**(5), 451–458.
15. S. K. Hansen, M. T. Cancilla, T. P. Shiau, J. Kung, T. Chen and D. A. Erlanson, Allosteric inhibition of PTP1B activity by selective modification of a non-active site cysteine residue, *Biochemistry*, 2005, **44**(21), 7704–77012.
16. J. Hyde, A. Braisted, M. Randal and M. R. Arkin, Discovery and characterization of cooperative ligand binding in the adaptive region of interleukin-2, *Biochemistry*, 2003, **42**(21), 6475–64783.

17. A. C. Braisted, J. D. Oslob, W. L. Delano, J. Hyde, R. S. McDowell, N. Waal, C. Yu, M. R. Arkin and B. C. Raimundo, Discovery of a potent small molecule IL-2 inhibitor through fragment assembly, *J. Am. Chem. Soc.*, 2003, **125**(13), 3714–3715.
18. B. C. Raimundo, J. D. Oslob, A. C. Braisted, J. Hyde, R. S. McDowell, M. Randal, N. D. Waal, J. Wilkinson, C. H. Yu and M. R. Arkin, Integrating fragment assembly and biophysical methods in the chemical advancement of small-molecule antagonists of IL-2: an approach for inhibiting protein–protein interactions, *J. Med. Chem.*, 2004, **47**(12), 3111–3130.
19. C. D. Thanos, M. Randal and J. A. Wells, Potent small-molecule binding to a dynamic hot spot on IL-2, *J. Am. Chem. Soc.*, 2003, **125**(50), 15280–15281.
20. N. D. Waal, W. Yang, J. D. Oslob, M. R. Arkin, J. Hyde, W. Lu, R. S. McDowell, C. H. Yu and B. C. Raimundo, Identification of non-peptidic small-molecule inhibitors of interleukin-2, *Bioorg. Med. Chem. Lett.*, 2005, **15**(4), 983–987.
21. M. Gurrath, Peptide-binding G protein-coupled receptors: new opportunities for drug design, *Curr. Med. Chem.*, 2001, **8**(13), 1605–1648.
22. B. O. Gerber, E. C. Meng, V. Dotsch, T. J. Baranski and H. R. Bourne, An activation switch in the ligand binding pocket of the C5a receptor, *J. Biol. Chem.*, 2001, **276**(5), 3394–3400.
23. T. J. Baranski, P. Herzmark, O. Lichtarge, B. O. Gerber, J. Trueheart, E. C. Meng, T. Iiri, S. P. Sheikh and H. R. Bourne, C5a receptor activation. Genetic identification of critical residues in four transmembrane helices, *J. Biol. Chem.*, 1999, **274**(22), 15757–15765.
24. J. A. DeMartino, Z. D. Konteatis, S. J. Siciliano, G. Van Riper, D. J. Underwood, P. A. Fischer and M. S. Springer, Arginine 206 of the C5a receptor is critical for ligand recognition and receptor activation by C-terminal hexapeptide analogs, *J. Biol. Chem.*, 1995, **270**(27), 15966–1599.
25. M. J. Romanowski, J. M. Scheer, T. O'Brien and R. S. McDowell, Crystal structures of a ligand-free and malonate-bound human caspase-1: implications for the mechanism of substrate binding, *Structure*, 2004, **12**(8), 1361–1371.
26. C. Pargellis, L. Tong, L. Churchill, P. F. Cirillo, T. Gilmore, A. G. Graham, P. M. Grob, E. R. Hickey, N. Moss, S. Pav and J. Regan, Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site, *Nat. Struct. Biol.*, 2002, **9**(4), 268–272.
27. R. Esnouf, J. Ren, C. Ross, Y. Jones, D. Stammers and D. Stuart, Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors, *Nat. Struct. Biol.*, 1995, **2**(4), 303–308.
28. C. Wiesmann, K. J. Barr, J. Kung, J. Zhu, D. A. Erlanson, W. Shen, B. J. Fahr, M. Zhong, L. Taylor, M. Randal, R. S. McDowell and S. K. Hansen, Allosteric inhibition of protein tyrosine phosphatase 1B, *Nat. Struct. Mol. Biol.*, 2004, **11**(8), 730–737.

29. J. Chai, E. Shiozaki, S. M. Srinivasa, Qi Wu, P. Dataa, E. S. Alnemri and Y. Shi, Structural basis of caspase-7 inhibition by XIAP, *Cell*, 2001, **104**(5), 769–780.
30. Y. Wei, T. Fox T, S. P. Chambers, J. Sintchak, J. T. Coll, J. M. Golec, L. Swenson, K. P. Wilson and P. S. Charifson, The structures of caspases-1, -3, -7 and -8 reveal the basis for substrate and inhibitor selectivity, *Chem. Biol.*, 2000, **7**(6), 423–432.
31. S. L. McGovern, B. T. Helfand, B. Feng and B. K. Shoichet, A specific mechanism of nonspecific inhibition, *J. Med. Chem.*, 2003, **46**(20), 4265–4272.
32. S. L. McGovern, E. Caselli, N. Grigorieff and B. K. Shoichet, A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening, *J. Med. Chem.*, 2002, **45**(8), 1712–1722.