Organic & Biomolecular Chemistry

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 5714

Tuning activity-based probe selectivity for serine proteases by on-resin 'click' construction of peptide diphenyl phosphonates†

Sevnur Serim, Susanne V. Mayer and Steven H. L. Verhelst*

Received 1st May 2013,
Accepted 9th July 2013

DOI: 10.1039/c3ob40907d

www.rsc.org/obc

Activity-based probes (ABPs) are powerful tools for functional proteomics studies. Their selectivity can be influenced by modification of a recognition element that interacts with pockets near the active site. For serine proteases there are a limited number of simple and efficient synthetic procedures for the development of selective probes. Here we describe a new synthetic route combining solid and solution phase chemistries to generate a small library of diphenyl phosphonate probes. Building blocks carrying a P1 recognition element and an electrophilic phosphonate warhead were prepared in solution and 'clicked' on-resin onto a tripeptide. We show the ability to modulate the activity and selectivity of diphenyl phosphonate ABPs and demonstrate activity-dependent labeling of endogenous proteases within a tissue proteome. The herein described synthetic approach therefore serves as a valuable method for rapid diversification of serine protease ABPs.

Introduction

Activity-based protein profiling (ABPP) has played an increasing role in the functional analysis of enzymes within complex biological systems.[1,2] With the ever-growing applicability of this technique there is constant interest in designing activity-based probes (ABPs), which are the workhorses of ABPP. The ability to modulate probe selectivity is crucial for their applicability: while general probes are sufficient for some analysis methods, other techniques require probes with high selectivity.[3] Most ABPs derive their selectivity from two probe features. (1) An electrophilic warhead, which covalently binds to a nucleophile in the active site. Depending on the nature of the electrophile, this warhead can react with a variety of nucleophilic amino acid residues or can be specific for a certain type of side chain, such as the catalytic serine or cysteine in serine and cysteine proteases, respectively. (2) A recognition element that modulates the affinity of the probe and directs the warhead to a subset of enzymes within the targeted class.

ABPP has been especially important for the study of proteases[4] since their activities are tightly regulated by several post-translational mechanisms. For cysteine cathepsin proteases, general[5] and selective ABPs[6,7] have been designed based on the epoxysuccinate and vinyl sulfone electrophile. Both general[8] and selective[9] acyloxymethyl ketone-based probes have been reported for caspases. The aforementioned probes can all be constructed on solid support, which accelerates the synthesis and optimization of the selectivity. This convenient strategy has also been employed for enzymes other than cysteine proteases.[10]

For serine proteases, general fluorophosphonate[11,12] and sulfonyl fluoride probes[13] and more selective isocoumarins[14,15] and peptidyl diphenyl phosphonates[16–18] (DPPs) are available (Fig. 1). However, these have been made using solution phase chemistries.

We implemented a solid phase strategy for the synthesis and selectivity modulation of serine protease ABPs. We chose diphenyl esters of α-aminophosphonates because they are low molecular weight, irreversible inhibitors exhibiting explicit selectivity[19] for serine proteases and therefore represent useful warheads for ABPs. DPPs have been validated as covalent irreversible inhibitors of serine proteases and they can be fine-tuned both *via* modification of the group in the P1 position, which interacts with the enzyme S1 pocket and *via* introduction of additional recognition sites that interact with more distal non-primed site pockets. The herein reported approach allows rapid diversification of the recognition moiety and convenient synthesis of DPP ABPs. In addition, their applicability for selective labeling of endogenous proteases within complex proteomes is demonstrated.

Lehrstuhl für Chemie der Biopolymere, Technische Universität München, Weihenstephaner Berg 3, 85354 Freising, Germany. E-mail: verhelst@wzw.tum.de

†Electronic supplementary information (ESI) available: Graphical representation of probe **10** docked into bovine beta-trypsin, and copies of spectra and chromatograms. See DOI: 10.1039/c3ob40907d

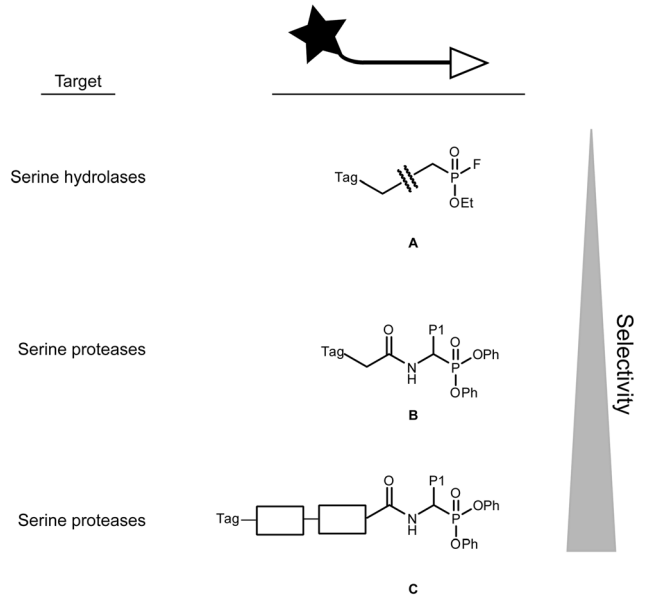
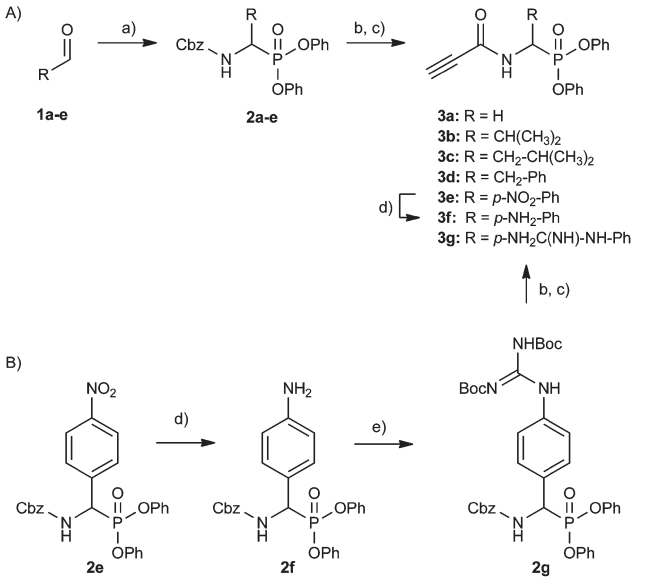


Fig. 1 Influencing the selectivity of phosphonate ABPs. The warhead (triangle) and spacer or recognition element (black line) determine the selectivity of the probe. Fluorophosphonates (A) are general probes that label most serine hydrolases. Peptide diphenyl phosphonates (B and C) label serine proteases. The selectivity is steered by the side chain in the P1 position and optional additional elements at distal sites.



Scheme 1 (A) Synthesis of building blocks with hydrophobic side chains. (a) $CbzNH_2$, $P(OPh)_3$, AcOH, 80 °C. (b) 33% HBr–AcOH. (c) Propiolic acid, DIC (pre-activation in THF, 0 °C), TEA, DMF. (B) Synthesis of building blocks with basic side chains. (d) Fe powder, AcOH, 70 °C. (e) 1,3-Di-Boc-2-methylisothiourea, $HgCl_2$, TEA, DCM.

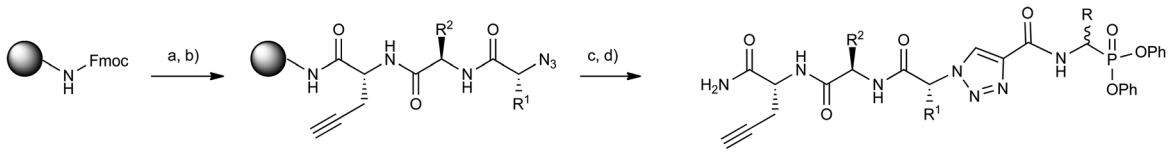
Results and discussion

Synthetic approach

Diphenyl esters of phosphonates are labile under basic conditions, which limits the reaction conditions that can be used during the synthesis. Our synthetic strategy towards phosphonate ABPs with extended peptide recognition elements therefore comprises the connection of a DPP warhead to a solid support during the last step using mild and chemoselective click chemistry. In this route, the peptidic portions of the probes are generated by solid phase peptide synthesis and the amino group of the last amino acid is converted to an azide by on-resin diazo transfer.[20] The azide allows for the introduction of the DPP building blocks *via* on-resin click reaction[21,22] forming a 1,4-substituted 1,2,3-triazole. Triazoles have been used before in selective probe design for cathepsin S.[23] Moreover, they are proteolytically and metabolically stable and provide good analogues of peptide bonds due to their strong dipole and H-bond accepting and donating properties.[24,25] We anticipated that the interactions at the P3 and P4 positions would provide the desired modulation of selectivity.

We synthesized alkynylated versions of diphenyl α-aminoalkylphosphonates in three easy steps. Cbz-protected derivatives with hydrophobic side chains or a hydrogen in the P1 position, such as glycine (**2a**), valine (**2b**), leucine (**2c**), phenylalanine (**2d**) and *p*-nitrophenylglycine (**2e**), were prepared from commercially available aldehydes (**1**) as described by Oleksyszyn *et al.*[26,27] The condensation products were obtained as pure solids after crystallization with yields ranging from 26 to 86%. The introduction of an alkyne handle to form the building blocks **3a–e** was achieved by coupling of propiolic acid after Cbz deprotection of the phosphonates (Scheme 1A). As a basic P1 residue we chose a diphenyl phosphonate with a *p*-guanidinophenyl side chain, which was made by using two additional reaction steps.[28] The *p*-nitrophenyl moiety of **2e** was reduced to an aniline (**2f**) using iron powder in acetic acid in quantitative yield. The aniline was then transformed into a Boc-protected guanidine in a $HgCl_2$-catalyzed reaction with 1,3-di-Boc-2-methylisothiourea giving compound **2g** in 57% yield (Scheme 1B). The introduction of the propiolic acid to both **2f** and **2g** was then performed as for the other building blocks.

The extended recognition elements of the probes were generated by solid phase peptide synthesis on a Rink amide resin (Scheme 2). All probes contain an L-propargylglycine to enable visualization of target proteases after labeling using click chemistry. This gives the flexibility to choose between different reporter tags such as biotin or fluorophores. Further elongation took place by coupling of two amino acids which form the P4 and P3 positions. We chose either two alanine residues or two residues according to the substrate specificities of different serine proteases.[29,30] The N-terminal amine was then transformed into an azide (**4**) *via* on resin diazo transfer. Subsequently, the alkyne DPP building blocks were reacted with the azide *via* on resin Cu(I)-catalyzed 1,3-dipolar cycloaddition. In principle, cross-coupling of an azide on one peptide chain and an alkyne on a neighboring peptide chain is possible. However, after cleavage from the resin, we did not observe any intramolecular cyclo-addition products, which may be due to the excess of the alkyne-containing phosphonate building



5 (AA-L):	R = CH_2-$CH(CH_3)_2$	$R^1 = R^2$ = Me	
6 (AA-F):	R = CH_2-Ph	$R^1 = R^2$ = Me	
7 (AA-Gua):	R = *p*-$NH_2C(NH)$-NH-Ph	$R^1 = R^2$ = Me	
8 (AS-Gua):	R = *p*-$NH_2C(NH)$-NH-Ph	R^1 = CH_2-OH	R^2 = Me
9 (AE-V):	R = $CH(CH_3)_2$	R^1 = CH_2-CH_2-COOH	R^2 = Me
10 (nT-Gua):	R = *p*-$NH_2C(NH)$-NH-Ph	R^1 = $CH(CH_3)$-OH	R^2 = CH_2-CH_2-CH_2-CH_3
11 (MF-F):	R = CH_2-Ph	R^1 = CH_2-Ph	R^2 = CH_2-CH_2-S-CH_3
12 (LF-Gua):	R = *p*-$NH_2C(NH)$-NH-Ph	R^1 = CH_2-Ph	R^2 = CH_2-$CH(CH_3)_2$

Scheme 2 Solid phase peptide synthesis. (a) Piperidine–DMF (1/4, v/v); then, Fmoc-aa-OH, HOBt, DIC, DMF. (b) Piperidine–DMF (1/4, v/v); then, TfN_3, $CuSO_4$, DCM–MeOH (9/1, v/v). (c) Building blocks **3**, $CuSO_4$, sodium ascorbate, TBTA, DMF–H_2O (10/1, v/v). (d) 95% TFA, 2.5% TIS, 2.5% H_2O. The abbreviations indicated behind each compound number correspond to the residues in the P4, P3 and P1 positions.

block in solution or to unfavorable distances between the resin-bound reaction partners. Final cleavage from the resin resulted in eight different DPP ABPs (**5–11**) with yields varying from 3 to 40% after purification (Scheme 2).

Labeling experiments

With both the P1 probes and the extended probes in hand we set out to test their labeling capability using commercially available, purified serine proteases of different specificities: bovine chymotrypsin, known for its P1 specificity for large hydrophobic residues; human cathepsin G with tryptic–chymotryptic "dual specificity";[31] human neutrophil elastase, known for its P1 specificity for small hydrophobic residues; bovine trypsin and human urokinase-type plasminogen activator (uPA) known for their basic P1 specificity. For comparison we included a nonselective fluorophosphonate ABP, FP–rhodamine[32] (FP–R). Proteases tagged by DPP ABPs **3a–g** and **5–12** were visualized by clicking a tetramethylrhodamine (TAMRA) derivative carrying an azide function, using conditions similar to those that have been optimized by Speers and Cravatt for efficient labeling in tandem ABPP.[33]

Both the P1 probes **3a–g** and the extended probes **5–12** resulted in labeling of proteases (Fig. 2). The reactions between the probes and proteases occur in an activity-dependent manner, since pretreatment with an active site-directed inhibitor competes away the labeling. The influence of the extended peptide recognition element on the probes' affinity towards their target proteases can be clearly observed from the overall increase of the labeling intensities in comparison to the P1 probes.

The general serine hydrolase ABP FP–R labeled all proteases regardless of their P1 preferences while a change in the extended recognition elements resulted in different selectivities of the different probes. Probe **7**, with a Phe in the P1 position, mainly labels chymotrypsin. Unexpectedly, a change in the P3 and P4 position in probe **8**, led to labeling of trypsin and uPA as well. Another example of the modulation of the selectivity was observed for probes with a basic P1 side chain. Trypsin, uPA and human cathepsin G were all equally labeled with the Gua P1 probe (**3g**) and the extended probe nT-Gua (**11**). This result is in line with previously reported phosphonate inhibitors of these enzymes.[34,35] In contrast, AS-Gua (**10**) displayed a high intensity band for trypsin and only a weaker band for uPA.

To gain insight into the binding mode of the extended probes that carry a triazole in their backbone, we performed

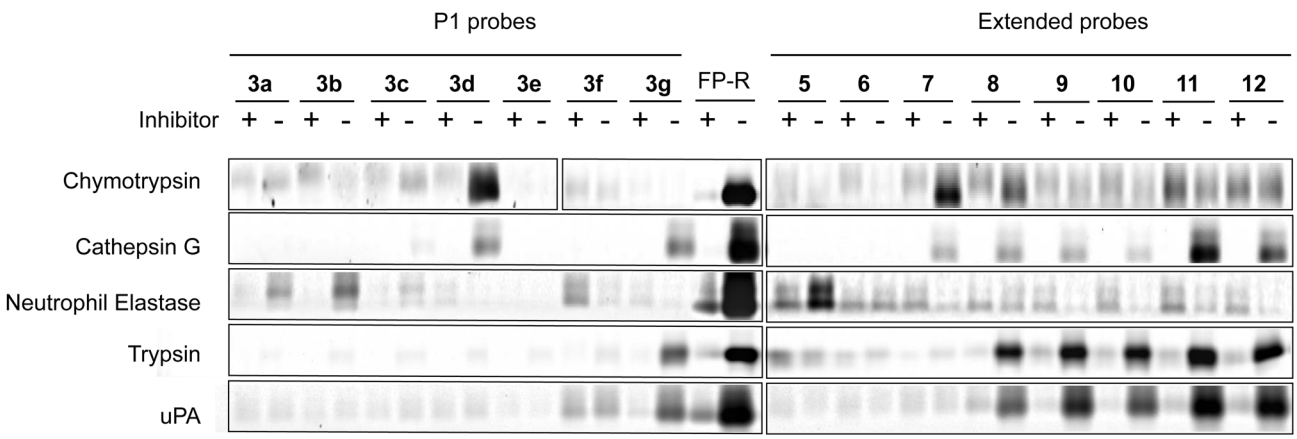


Fig. 2 Labeling of purified proteases with FP–rhodamine, P1 and extended DPP probes. Probe concentrations: 5 μM for DPPs, 1 μM for FP–R.

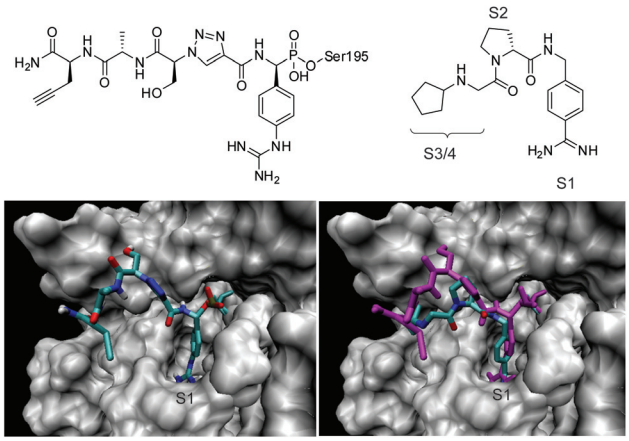


Fig. 3 Docking of probe **10** bound to Ser195 in bovine beta-trypsin. The enzyme surface is represented in white, except for the S195 residue, which is shown as a stick model as are the small molecules. Left: probe **10** (colored by element: red = oxygen, blue = nitrogen, cyan = carbon, brown = phosphorus) docks with its *para*-guanidino-phenyl group into the S1 pocket. Right: probe **10** is colored in magenta and overlaid with a non-covalent inhibitor.[34] The triazole of probe **10** takes a similar position as the proline in the inhibitor, while the other residues interact with various sites on the non-primed site.

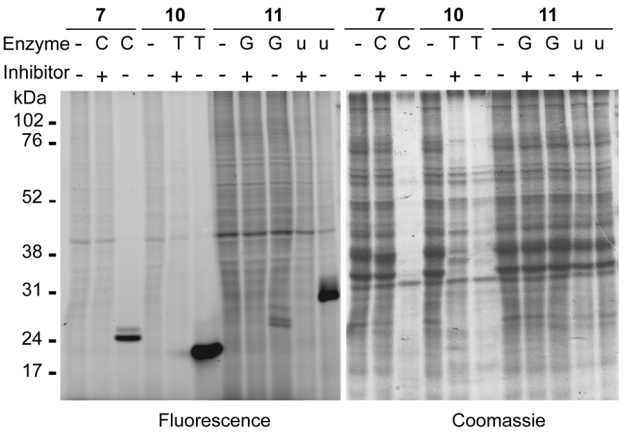


Fig. 4 In-gel fluorescence (left) of proteases labeled in the context of a proteome (a cell lysate of the human colon adenocarcinoma cell line HT-29). The Coomassie stain (right) shows protein loading. The apparent lower amount of protein in the uninhibited lanes for chymotrypsin and trypsin are due to the high digestive properties of these proteases. Note that for trypsin, there is still some digestion observed in the inhibited sample compared with the no trypsin control probably due to incomplete inhibition of the trypsin. Chy.: chymotrypsin; Try.: trypsin; Cat. G: cathepsin G; u: urokinase-type plasminogen activator.

molecular docking of a covalently bound probe **10** inside a bovine trypsin crystal structure. As expected, the *p*-guanidinophenyl ring in the P1 position of the probe occupies the S1 pocket (Fig. 3), where it forms a salt bridge with Asp189 (Fig. S1†). The triazole occupies the S2 site and has a similar position as a proline in the P2 position of a non-covalent trypsin inhibitor[36] (Fig. 3). The alanine and propargylglycine interact with the more distal S3/S4 sites. Overall, this structure suggests that the probes – despite the non-natural triazole and the reversed polarity of the P3–P5 backbone – interact with the non-primed site to give increased potency compared to the P1 only probes.

We next investigated whether the extended probes can be applied within the context of a complex proteome. To demonstrate that the presence of other proteins does not interfere with the labeling of the target proteases, we spiked selected proteases into a cell lysate and incubated with probes that displayed the most selective activity-based labeling of the purified proteases. In-gel fluorescence showed intense activity-dependent bands for each enzyme without much cross-reactivity with other proteins (Fig. 4). Only nT-Gua (**11**) when reacted with cathepsin G displayed some background bands which were not competed away by pretreatment with PMSF.

Finally, to illustrate the applicability of DPP ABPs in the labeling of endogenous serine proteases within a complex proteome, we used a rat pancreas lysate that was activated by treatment with enterokinase. Enterokinase activates trypsinogen into trypsin, the common activator of all pancreatic zymogens.[37] Both activated and unactivated proteomes were used in labeling experiments with decreasing concentrations of probes that displayed strong activity towards chymotrypsin and trypsin in the previous experiments. A selective labeling of chymotryptic and tryptic enzymes was detectable (as confirmed by pretreatment with chymotryptic inhibitor DAP22c and tryptic inhibitor TLCK) with probe concentrations down to 0.1 and 1 μM, respectively (Fig. 5). No labeling was observed in the case of unactivated or preinhibited lysates, as well as buffer containing only enterokinase.

Conclusion

Diphenyl esters of α-aminophosphonates are phosphonic analogues of naturally occurring amino acids. They selectively, covalently and irreversibly bind to the active site serine residue of serine proteases. These features qualify them as attractive warheads for ABPs. The selectivity of DPPs can be modulated *via* the adjustment of the group at the P1 position and other non-primed site residues. While synthetic strategies to generate DPPs in solution have been reported, an easy and rapid way to synthesize DPP ABPs remains a challenge. Here, we presented a new synthetic route that enables simple modification of the extended peptide sequence and an overall convenient construction of DPP ABPs. Key features of the synthesis include solid phase generation of extended recognition elements and on-resin click chemistry mediated introduction of the P1 phosphonate building blocks to form the final probes. Using this approach we synthesized eight different DPP ABPs, and labeled serine proteases of different specificities. We demonstrated that it is possible to tune the activity and selectivity of the DPP ABPs by varying the extended recognition elements. Molecular modeling suggests that the probes interact with the non-primed sites, which may explain their higher potency and capacity to influence the selectivity towards the protease targets. Furthermore, the probes display activity-dependent labeling of proteases within complex

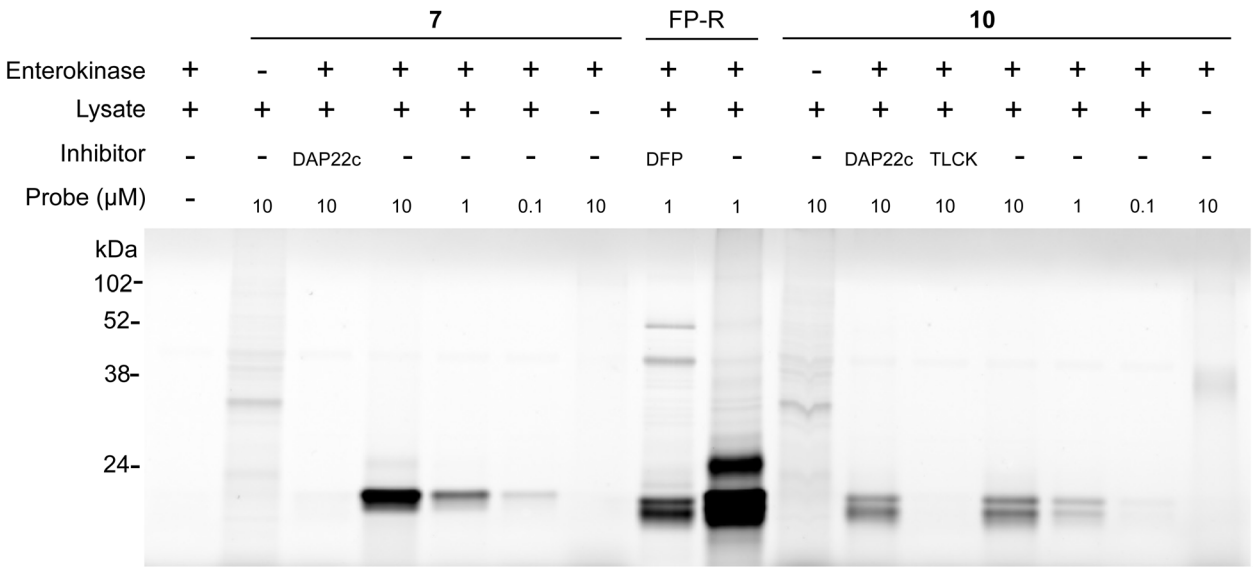


Fig. 5 Fluorescent labeling of endogenous proteases in enterokinase-activated rat pancreas lysate by DPP ABPs with extended recognition elements and FP–R. The relatively unstable inhibitor DFP did not completely inhibit all labeling by FP–R. The higher intensity of the two bands at approximately 42 and 50 kDa in the inhibited sample may be due to the higher availability of FR–R compared to the enterokinase treated sample, in which FP–R reacts with the activated proteases. Enterokinase is not labeled (rightmost lanes for probes **7** and **10**, respectively), probably due to the hydrophobic P1 of probe **7** and the preference of enterokinase for negatively charged residues in the P2 and P3 position, which are absent from the probes used.

proteomes. We envision that this strategy will allow for future probe optimization when selective serine protease ABPs are required.

Experimental

General

All starting materials, enzymes and inhibitors were obtained from Sigma-Aldrich with the exception of uPA (VWR) and FP–R (Thermo Scientific) and used without further purification. Thin layer chromatography (TLC) was performed on ALUGRAM sil G/UV254 TLC plates (Carl Roth). Solvents used for SPPS and purification *via* column chromatography were purchased from Applichem. Compounds were separated over silica gel with a grain distribution of 0.04–0.063 mm and a pore size of 60 Å (Carl Roth). Creosalus was the supplier for Fmoc-protected L-amino acids, Rink resin and coupling reagents for SPPS. LC-MS spectra were recorded by an Agilent 1100 Series LC system coupled to an Agilent 6210 ESI TOF mass spectrometer, with elution by solvents A (5% ACN–H_2O + 0.1% formic acid) and B (95% ACN–H_2O + 0.1% formic acid). A Zorbax SB C18 5 μm (0.5 × 150 mm) column was used to separate samples at r.t. with a flow rate of 20 μl min^{-1} and with a gradient of 2.57% ACN per min starting from A for 35 min. HPLC purifications were carried out on Waters Xbridge C18 5 μm (4.6 × 150 mm) and a Waters Xbridge BEH130 Prep C18 5 μm (19 × 150 mm) columns. ^1H and ^{13}C NMR spectra were measured on a Bruker 400 MHz DRX (400; 100 MHz). Chemical shifts (δ) are given in parts per million (ppm) relative to tetramethylsilane as an internal standard.

General methods for the synthesis of propynoylated diphenyl α-aminoalkylphosphonate building blocks

Birum–Oleksyszyn reaction. Benzyl carbamate (1 eq.) was combined with triphenylphosphite (1 eq.), the corresponding aldehyde (1.5 eq.) and acetic acid (0.1–0.2 mL per mmol aldehyde) and heated to 82 °C for *ca.* 2 h. The reaction was monitored by TLC until benzyl carbamate was consumed. Volatile components were removed and methanol was added to allow crystallization at −20 °C overnight. Crystals were collected by filtration, washed with cold methanol and dried. The resulting solids were sufficiently pure for the following reactions.

Cbz deprotection. Cbz-protected DPPs were treated with 33% HBr–AcOH solution for 1–2 h at room temperature. The solvent was removed and the oily residue was dissolved in a minimal amount of methanol; excess diethylether was added and overnight storage at −20 °C led to crystallization. The crystals were filtered, washed with cold diethylether and dried. The purity of the compounds was sufficient for further reactions.

Coupling of propiolic acid. Propiolic acid (2.6 eq.) was preactivated with DIC (1.3 eq.) in THF (0.3 mL per mmol propiolic acid) for 1 h at 0 °C. The DPP hydrobromide (1 eq.) was dissolved in DMF (0.75 mL $mmol^{-1}$), treated with DIEA (2 eq.) and then added to the preactivated propiolic acid solution and stirred overnight at room temperature. Base was neutralized with AcOH and the volatile components were removed. The residue was dissolved in ethylacetate and washed 2× with 1 M HCl, 2× with H_2O, 2× with a saturated $NaHCO_3$ solution, and brine. The organic phase was dried over $MgSO_4$ and concentrated at reduced pressure. If necessary, compounds were further purified *via* silica column chromatography.

Transformation of the NO_2 group to a Boc-protected guanidine. Acetic acid (4 mL $mmol^{-1}$) was added to a mixture of the nitro compound (1 eq.) and Fe powder (9 eq.). The reaction mixture was heated to 70 °C and stirred under a nitrogen atmosphere for 2 h. Afterwards, the acid was removed and the crude residue was dissolved in EtOAc. Fe_2O_3 was centrifuged down and the supernatant was concentrated to form the crude product. The aniline (1 eq.) was mixed with 1,3-bis-(*tert*-butoxycarbonyl)-2-methyl-2-thiopseudourea (1.1 eq.) and mercury(II) chloride (1.2 eq.). Subsequently, DCM and Et_3N (3 eq.) were added resulting in a yellowish suspension, which was stirred overnight. DCM was evaporated and the residue was redissolved in EtOAc. The remaining solid was centrifuged down; the supernatant was decanted and washed with 1 M $KHSO_4$, saturated $NaHCO_3$ solution and brine. The organic phase was dried over $MgSO_4$ and the solvent was removed. Further purification was carried out *via* silica column chromatography.

Diphenyl α-*N*-(benzyloxycarbonyl)amino-methylphosphonate (2a). A mixture of benzyl carbamate (1 eq.), acetic anhydride (1.25 eq.), and paraformaldehyde (1 eq.) in acetic acid (100 μL $mmol^{-1}$) was heated at 65 °C for 3 h. The resulting solution was treated with triphenyl phosphite (1 eq.) and heated at 115 °C for 2 h. The mixture was concentrated under high vacuum, and a small volume of diethyl ether was added to allow crystallization at −20 °C overnight. The precipitate was then collected by filtration. The resulting solid was sufficiently pure for the following reactions. The title compound was isolated as a white solid in 26% yield. ESI-MS: $[M + H]^+$ *m/z* 398.1194 (found), 398.1151 (calculated).

Diphenyl α-*N*-(benzyloxycarbonyl)amino-2-methylpropylphosphonate (2b).[26] The title compound was isolated as white crystals in 52% yield. ESI-MS: $[M + H]^+$ *m/z* 440.1577 (found), 440.1621 (calculated).

Diphenyl α-*N*-(benzyloxycarbonyl)amino-3-methylbutylphosphonate (2c).[26] The title compound was isolated as white crystals in 69% yield. ESI-MS: $[M + H]^+$ *m/z* 454.1737 (found), 454.1777 (calculated).

Diphenyl α-*N*-(benzyloxycarbonyl)amino-2-phenylethylphosphonate (2d).[26] The title compound was isolated as a white solid in 40% yield. ESI-MS: $[M + H]^+$ *m/z* 488.1566 (found), 488.1621 (calculated).

Diphenyl α-*N*-(benzyloxycarbonyl)amino-(4-nitro-phenyl)-methanephosphonate (2e).[27] The title compound was isolated as white crystals in 86% yield. ESI-MS: $[M + H]^+$ *m/z* 519.1321 (found), 519.1315 (calculated).

Diphenyl α-*N*-(benzyloxycarbonyl)amino-(4-amino-phenyl)-methanephosphonate (2f).[27] The title compound was isolated as an orange solid in 90% yield. ESI-MS: $[M + H]^+$ *m/z* 489.1622 (found), 489.1573 (calculated).

Diphenyl α-*N*-(benzyloxycarbonyl)amino-(4-diBocguanidinium-phenyl)methanephosphonate (2g).[27] The title compound was isolated as a white solid in 57% yield. ESI-MS: $[M + H]^+$ *m/z* 731.2890 (found), 731.2840 (calculated).

Diphenyl α-*N*-(propiolamido)-methylphosphonate (3a). ESI-MS: $[M + H]^+$ *m/z* 316.0735 (found), 316.0732 (calculated). ^{1}H-NMR (400 MHz, $CDCl_3$): δ = 7.40–7.32 (m, 4 H), 7.26–7.16 (m, 6 H), 6.66 (s, 1 H), 4.11 (dd, J = 12.0 Hz, J = 6.0 Hz, 2 H), 2.88 (s, 1 H).

Diphenyl α-*N*-(propiolamido)-2-methylpropylphosphonate (3b). The title compound was isolated as yellowish oil in 45% yield. ESI-MS: $[M + H]^+$ *m/z* 358.1237 (found), 358.1202 (calculated). ^{1}H-NMR (400 MHz, $CDCl_3$): δ = 7.40–7.28 (m, 4 H), 7.26–7.09 (m, 6 H), 6.46 (d, J = 10.4 Hz, 1 H), 4.83 (ddd, J = 19.2 Hz, J = 10.4 Hz, J = 4.3 Hz, 1 H), 2.90 (s, 1 H), 2.57–2.37 (m, 2 H), 1.19–1.08 (m, 6 H).

Diphenyl α-*N*-(propiolamido)-3-methylbutylphosphonate (3c). The title compound was isolated as yellowish oil in 29% yield. ESI-MS: $[M + H]^+$ *m/z* 372.1350 (found), 372.1358 (calculated). ^{1}H-NMR (400 MHz, $CDCl_3$): δ = 7.39–7.29 (m, 4 H), 7.26–7.08 (m, 6 H), 6.64 (d, J = 10.2 Hz, 1 H), 5.06–4.87 (m, 1 H), 2.84 (s, 1 H), 1.91–1.71 (m, 3 H), 1.06–0.90 (m, 6 H).

Diphenyl α-*N*-(propiolamido)-2-phenylethylphosphonate (3d). The title compound was isolated as a white solid in 52% yield. ESI-MS: $[M + H]^+$ *m/z* 406.1255 (found), 406.1202 (calculated). ^{1}H-NMR (400 MHz, $CDCl_3$): δ = 7.49–6.94 (m, 15 H), 5.38–5.01 (m, 1 H), 3.59–3.27 (m, 1 H), 3.15–2.97 (m, 1 H), 2.74 (s, 1 H).

Diphenyl α-*N*-(propiolamido)-(4-nitro-phenyl)methanephosphonate (3e). The title compound was isolated as an orange solid in 48% yield. ESI-MS: $[M + H]^+$ *m/z* 437.10 (found), 437.08 (calculated). ^{1}H-NMR (400 MHz, $CDCl_3$): δ = 8.19 (d, J = 8.0 Hz, 2 H), 7.71 (dd, J = 8.0 Hz, J = 2.0 Hz, 2 H), 7.41–7.33 (m, 2 H), 7.28–7.21 (m, 3 H), 7.20–7.12 (m, 3 H), 6.94–6.89 (m, 2 H), 6.03 (dd, J = 24.0 Hz, J = 8.0 Hz, 1 H), 2.83 (s, 1 H).

Diphenyl α-*N*-(propiolamido)-(4-amino-phenyl)methanephosphonate (3f). The title compound was isolated in 75% yield. ESI-MS: $[M + H]^+$ *m/z* 407.13 (found), 407.11 (calculated). ^{1}H-NMR (400 MHz, DMSO-d_6): δ = 10.14 (d, J = 9.7 Hz, 1 H), 7.41–7.30 (m, 6 H), 7.24–7.15 (m, 2 H), 7.07 (d, J = 7.5 Hz, 2 H), 6.94 (d, J = 7.5 Hz, 2 H), 6.74 (d, J = 7.5 Hz, 2 H), 5.77 (dd, J = 21.3 Hz, J = 9.7 Hz, 1 H), 4.34 (s, 1 H).

Diphenyl α-*N*-(propiolamido)-(4-guanidinium-phenyl)-methanephosphonate (3g). ESI-MS: $[M + H]^+$ *m/z* 449.1413 (found), 449.1372 (calculated). ^{1}H-NMR (400 MHz, DMSO-d_6): δ = 10.35 (d, J = 10.0 Hz, 1 H), 9.71 (s, 1 H), 7.71 (dd, J = 8.4 Hz, 2.0 Hz, 2 H), 7.46 (s, 3 H), 7.42–7.32 (m, 4 H), 7.29 (d, J = 8.4 Hz, 2 H), 7.26–7.19 (m, 2 H), 7.09 (d, J = 8.4 Hz, 2 H), 7.02 (d, J = 8.4 Hz, 2 H), 6.00 (dd, J = 22.6 Hz, J = 10.0 Hz, 1 H), 4.41 (s, 1 H).

General methods for solid phase peptide synthesis of peptide diphenyl α-aminoalkylphosphonates

Amino acid coupling and Fmoc deprotection. Fmoc groups were deprotected by 20% piperidine–DMF (v/v) for 20 min. For amino acid couplings, Fmoc–aa–OH (3 eq., 0.25 M in DMF), DIC (3 eq.) and HOBt (3 eq.) were added to the resin and shaken at room temperature for 1–3 h. Completion of the reactions was checked by the Kaiser test.

Diazotransfer. Triflyl azide was freshly prepared from sodium azide (5 eq.) and triflic anhydride (1 eq.). NaN_3 was dissolved in H_2O (0.16 mL $mmol^{-1}$). At 0 °C, DCM (0.27 mL $mmol^{-1}$) was added to the clear solution under heavy stirring.

Tf_2O was added dropwise and the reaction mixture was stirred at 0 °C for 2 h. The water phase was extracted twice with DCM. The combined organic phases were washed once with a saturated $NaHCO_3$ solution. The maximal volume of the organic phase was 2.70 mL $mmol^{-1}$ TfN_3. For the diazotransfer, the Fmoc deprotected tripeptide on the resin (1 eq.) was treated with TfN_3 (0.37 M, 16.7 eq.) in DCM and $CuSO_4$ (12.6 mM, 0.1 eq.) in MeOH for 24 h while shaking. Before a Kaiser test was performed, the resin was washed with NMP (3 × 2 min), 0.5% DIEA–NMP (3 × 2 min), 0.05 M sodium diethyldithiocarbamate in NMP (3 × 10 min), NMP (5 × 5 min) and DCM (3 × 3 min) and Et_2O.

On-resin click reaction. The resin-tripeptide azide conjugate (1 eq.) was resuspended in DMF and the building block (3 eq.) was added as a solution in DMF (0.45 M), followed by the click reagents TBTA (0.2 eq.), $CuSO_4$ (0.1 eq.) and sodium ascorbate (3 eq.) in DMF–H_2O (10/1, v/v). The reaction was shaken at room temperature for about 24 h, followed by three washing steps each with DCM, DMF, 0.02 M sodium diethyldithiocarbamate in DMF, DMF, MeOH, DMF, DCM and Et_2O.

Resin cleavage. The resin was treated with 95% TFA, 2.5% H_2O and 2.5% TIS for 1 h. Supernatant was collected and compounds were precipitated by the addition of excess Et_2O. The precipitates were dried under a nitrogen stream and purified by HPLC.

PraAlaGlu(triazole)Val^P(OPh)$_2$ (5). The title compound was isolated as a white solid in 3% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 696.2496 (found), 696.2541 (calculated).

PraAlaAla(triazole)Leu^P(OPh)$_2$ (6). The title compound was isolated as a white solid in 40% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 652.2603 (found), 652.2642 (calculated).

PraAlaAla(triazole)Phe^P(OPh)$_2$ (7). The title compound was isolated as a white solid in 7% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 686.2456 (found), 686.2486 (calculated).

PraMetPhe(triazole)Phe^P(OPh)$_2$ (8). The title compound was isolated as a white solid in 4% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 822.2804 (found), 822.2833 (calculated).

PraAlaAla(triazole)Gua^P(OPh)$_2$ (9). The title compound was isolated as a white solid in 8% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 729.2622 (found), 729.2656 (calculated).

PraAlaSer(triazole)Gua^P(OPh)$_2$ (10). The title compound was isolated as a white solid in 9% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 745.2548 (found), 745.2605 (calculated).

PraNleThr(triazole)Gua^P(OPh)$_2$ (11). The title compound was isolated as a white solid in 8% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 801.3216 (found), 801.3231 (calculated).

PraLeuPhe(triazole)Gua^P(OPh)$_2$ (12). The title compound was isolated as a white solid in 7% yield after HPLC purification. ESI-MS: $[M + H]^+$ *m*/*z* 847.3404 (found), 847.3439 (calculated).

Docking experiments

The phosphonate **10** (with the same chirality at the P1 position as a natural substrate) was geometry optimized with an MMFF94 force field and defined as an extension of the side chain of the bovine trypsin S195 (PDB coordinates: 1MAX, from which the original covalent phosphonate inhibitor structure was deleted). Docking of the inhibitor as a flexible side chain of S195 was performed with AutoDock Vina.[38] A non-covalent inhibitor bound to bovine trypsin was taken as a comparison (PDB coordinates: 2ZFT). Pictures were generated using VMD 1.9.[39]

Labeling experiments

Labeling of purified enzymes. Each enzyme in PBS was treated for 30 min at room temperature with either a suitable inhibitor (0.1 mM DAP22c for chymotrypsin, 1 mM PMSF for cathepsin G, 1 mM DFP for uPA and neutrophil elastase, 1 mM TLCK for trypsin) or DMSO in a final volume of 50 μL. Pretreatment was followed by labeling with ABPs for 30 min at r.t. In the case of DPPs, TAMRA azide (25 μM) was subsequently clicked onto the alkynylated ABPs in the presence of $CuSO_4$ (1 mM), TBTA (50 μM) and sodium ascorbate (0.5 mM) for 30 min at room temperature in the dark. The reactions were stopped with 4× sample buffer. The samples were boiled at 95 °C for 3 min. Protein samples were resolved on a 12% AA SDS gel and gels were visualized by a Typhoon Trio+ scanner (GE Healthcare). Protein/lane: 100 ng.

Labeling of purified proteases in a proteome background. Each protease was diluted in a lysate of an HT-29 cell line (1 mg mL^{-1}) as 1% of total protein. Afterwards the protocol for the labeling of purified enzymes was followed. Protein loading: 10 μg per lane.

Labeling in enterokinase-activated rat pancreas lysate. Rat pancreas lysate in PBS (concentration of total protein: 3.5 mg mL^{-1}) was activated with enterokinase (1 U mg^{-1} protein) for 2 h on ice. Labeling with ABPs was carried out at a protein concentration of 3 mg ml^{-1} following the protocol for labeling of purified enzymes. Protein samples were resolved on a 15% AA SDS gel. Protein loading: 15 μg per lane.

Acknowledgements

We acknowledge financial support from the DFG (Emmy Noether program), the Fonds der Chemischen Industrie and the Center for Integrated Protein Science Munich. We thank Dr O. Frank for measuring NMR, Eliane Wolf for a critical reading of the manuscript and Prof. Dr D. Langosch for general support.

Notes and references

1 M. J. Evans and B. F. Cravatt, *Chem. Rev.*, 2006, **106**, 3279–3301.
2 A. M. Sadaghiani, S. H. L. Verhelst and M. Bogyo, *Curr. Opin. Chem. Biol.*, 2007, **11**, 20–28.
3 U. Haedke, E. V. Kuttler, O. Vosyka, Y. Yang and S. H. L. Verhelst, *Curr. Opin. Chem. Biol.*, 2013, **17**, 102–109.

4 S. Serim, U. Haedke and S. H. L. Verhelst, *ChemMedChem*, 2012, **7**, 1146–1159.
5 D. Greenbaum, K. F. Medzihradszky, A. Burlingame and M. Bogyo, *Chem. Biol.*, 2000, **7**, 569–581.
6 F. Yuan, S. H. L. Verhelst, G. Blum, L. M. Coussens and M. Bogyo, *J. Am. Chem. Soc.*, 2006, **128**, 5616–5617.
7 S. H. L. Verhelst and M. Bogyo, *ChemBioChem*, 2005, **6**, 824–827.
8 D. Kato, K. M. Boatright, A. B. Berger, T. Nazif, G. Blum, C. Ryan, K. A. H. Chehade, G. S. Salvesen and M. Bogyo, *Nat. Chem. Biol.*, 2005, **1**, 33–38.
9 A. B. Berger, M. D. Witte, J. B. Denault, A. M. Sadaghiani, K. M. Sexton, G. S. Salvesen and M. Bogyo, *Mol. Cell*, 2006, **23**, 509–521.
10 K. A. Kalesh, L. P. Tan, K. Lu, L. Gao, J. Wang and S. Q. Yao, *Chem. Commun.*, 2010, **46**, 589–591.
11 D. Kidd, Y. S. Liu and B. F. Cravatt, *Biochemistry*, 2001, **40**, 4005–4015.
12 Y. S. Liu, M. P. Patricelli and B. F. Cravatt, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 14694–14699.
13 D. A. Shannon, C. Gu, C. J. McLaughlin, M. Kaiser, R. A. van der Hoorn and E. Weerapana, *ChemBioChem*, 2012, **13**, 2327–2330.
14 U. Haedke, M. Gotz, P. Baer and S. H. L. Verhelst, *Bioorg. Med. Chem.*, 2012, **20**, 633–640.
15 J. C. Powers, C. M. Kam, L. Narasimhan, J. Oleksyszyn, M. A. Hernandez and T. Ueda, *J. Cell. Biochem.*, 1989, **39**, 33–46.
16 B. F. Gilmore, L. Carson, L. L. McShane, D. Quinn, W. A. Coulter and B. Walker, *Biochem. Biophys. Res. Commun.*, 2006, **347**, 373–379.
17 S. Mahrus and C. S. Craik, *Chem. Biol.*, 2005, **12**, 567–577.
18 Z. Pan, D. A. Jeffery, K. Chehade, J. Beltman, J. M. Clark, P. Grothaus, M. Bogyo and A. Baruch, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 2882–2885.
19 M. Sienczyk and J. Oleksyszyn, *Curr. Med. Chem.*, 2009, **16**, 1673–1687.
20 D. T. S. Rijkers, H. H. R. van Vugt, H. J. F. Jacobs and R. M. J. Liskamp, *Tetrahedron Lett.*, 2002, **43**, 3657–3660.
21 N. G. Angelo and P. S. Arora, *J. Org. Chem.*, 2007, **72**, 7963–7967.
22 V. Aucagne and D. A. Leigh, *Org. Lett.*, 2006, **8**, 4505–4507.
23 M. Verdoes, L. E. Edgington, F. A. Scheeren, M. Leyva, G. Blum, K. Weiskopf, M. H. Bachmann, J. A. Ellman and M. Bogyo, *Chem. Biol.*, 2012, **19**, 619–628.
24 Y. L. Angell and K. Burgess, *Chem. Soc. Rev.*, 2007, **36**, 1674–1689.
25 D. S. Pedersen and A. Abell, *Eur. J. Org. Chem.*, 2011, 2399–2411.
26 J. Oleksyszyn, L. Subotkowska and P. Mastalerz, *Synthesis*, 1979, 985–986.
27 J. Oleksyszyn and J. C. Powers, *Methods Enzymol.*, 1994, **244**, 423–441.
28 M. Sienczyk and J. Oleksyszyn, *Tetrahedron Lett.*, 2004, **45**, 7251–7254.
29 J. L. Harris, B. J. Backes, F. Leonetti, S. Mahrus, J. A. Ellman and C. S. Craik, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 7754–7759.
30 D. N. Gosalia, C. M. Salisbury, J. A. Ellman and S. L. Diamond, *Mol. Cell. Proteomics*, 2005, **4**, 626–636.
31 W. W. Raymond, N. N. Trivedi, A. Makarova, M. Ray, C. S. Craik and G. H. Caughey, *J. Immunol.*, 2010, **185**, 5360–5368.
32 M. P. Patricelli, D. K. Giang, L. M. Stamp and J. J. Burbaum, *Proteomics*, 2001, **1**, 1067–1071.
33 A. E. Speers and B. F. Cravatt, *Chem. Biol.*, 2004, **11**, 535–546.
34 M. Sienczyk, A. Lesner, M. Wysocka, A. Legowska, E. Pietrusewicz, K. Rolka and J. Oleksyszyn, *Bioorg. Med. Chem.*, 2008, **16**, 8863–8867.
35 M. Sienczyk and J. Oleksyszyn, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 2886–2890.
36 T. Brandt, N. Holzmann, L. Muley, M. Khayat, C. Wegscheid-Gerlach, B. Baum, A. Heine, D. Hangauer and G. Klebe, *J. Mol. Biol.*, 2011, **405**, 1170–1187.
37 X. L. Zheng, Y. Kitamoto and J. E. Sadler, *Front. Biosci. (Elite Ed)*, 2009, **1**, 242–249.
38 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.
39 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38, 27–38.