Catalytic activation of pre-substrates via dynamic fragment assembly on protein templates

Edyta Burda[1] & Jörg Rademann[1,2]

Sensitive detection of small molecule fragments binding to defined sites of biomacromolecules is still a considerable challenge. Here we demonstrate that protein-binding fragments are able to induce enzymatic reactions on the protein surface via dynamic fragment ligation. Fragments binding to the S1 pocket of serine proteases containing a nitrogen, oxygen or sulphur nucleophile are found to activate electrophilic pre-substrates through a reversible, covalent ligation reaction. The dynamic ligation reaction positions the pre-substrate molecule at the active site of the protein thereby inducing its enzymatic cleavage. Catalytic activation of pre-substrates is confirmed by fluorescence spectroscopy and by high-performance liquid chromatography. The approach is investigated with 3 pre-substrates and 14 protein-binding fragments and the specific activation and the templating effect exerted by the enzyme is quantified for each protease–fragment–pre-substrate combination. The described approach enables the site-specific identification of protein-binding fragments, the functional characterization of enzymatic sites and the quantitative analysis of protein template-assisted ligation reactions.

[1] Institute of Pharmacy, Medicinal Chemistry, University of Leipzig, Brüderstraße 34, 04103 Leipzig, Germany. [2] Institute of Pharmacy, Medicinal Chemistry, Freie Universität Berlin, Königin-Luise-Straße 2 + 4, 14195 Berlin, Germany. Correspondence and requests for materials should be addressed to J.R. (email: joerg.rademann@fu-berlin.de).

Fragment-based development of protein ligands has been recognized over the last years as a powerful strategy to identify molecules with an optimized fit to protein-binding pockets, providing ligands with reduced molecular weight and with higher ligand efficiency, thereby possibly improving the bioavailability, cell permeability and metabolic stability of drug molecules[1–8]. Scrutinizing the chemical space of small molecule fragments—instead of screening large libraries of drug-like molecules—requires considerably smaller chemical libraries for covering the chemical space and thus enables to identify and optimize hit ligands with less effort for testing and chemical synthesis.

There are, however, specific challenges implicated in the fragment-based drug discovery process. Many primary fragment hits possess low, often millimolar affinities to the protein target resulting from the small contact area and the low number of binding interactions of fragments compared with larger molecules. As a consequence, high concentrations of fragments have to be used in the bioassays, which can interfere with the reliable read-out of classical biochemical test systems. To circumvent this problem, biophysical methods have been employed for the detection of protein-binding fragments, using, for example, nuclear magnetic resonance (NMR) spectroscopy[9,10] or X-ray diffraction of protein crystals[11–14]. Next, primary fragment hits with low affinity have to be extended to fragment combinations to attain sufficient activity—a process requiring the systematic modification of hits by applying tools and methods of medicinal chemistry. Realistically, this optimization process can be more time-consuming and uncertain in its outcome than the initial discovery of fragment hits. Fragment combinations have been synthesized and tested not only for direct inhibitor screening, but also for the screening of substrate libraries. Active substrates can then serve as a starting point for the development of enzyme inhibitors following a seminal concept in medicinal chemistry[15,16].

Dynamic ligation screening (DLS) has been introduced as an approach for the sensitive, site-selective detection of protein-binding fragments[17]. In this method, a chemically reactive protein ligand serves as a directing probe for the screening of fragment libraries at a spatially defined protein site[18]. Reversible formation of a covalent bond between the directing probe and a nucleophilic fragment enables the sensitized detection of low-affinity fragments with millimolar activity while testing at lower concentrations of 10–100 μM. This sensitizing effect results from the templated assembly of protein ligands on the protein surface. As a consequence, standard biochemical assays can be applied for the detection of low-affinity fragments, for example, the enzymatic conversion of a fluorogenic substrate or fluorescence polarization for a binding assay[19]. DLS provides fewer hits than biophysical assays, as only fragments with a fitting spatial orientation in the ligation product are detected as actives. Active fragment combinations were identified and optimized iteratively, with respect to affinity or, if second-site-specific fragments are scrutinized with sets of proteins, with respect to specificity[20]. Other than in biophysical assays, nanomolar protein concentrations are sufficient and large collections of nucleophilic fragments can be tested in small volumes of liquids on microtitre plates. Conceptually, DLS constitutes a specific application of dynamic covalent chemistry, which has been defined as that branch of organic chemistry studying compounds formed reversibly in thermodynamic equilibrium[6,21]. Therefore in principal all dynamic ligation reactions should be applicable to fragment ligation assays, yielding, for example, imines[22–24], hydrazones[25,26], hemiacetals[27,28], hemithioacetals[29], acetals[30,31], thioesters[32] or Michael addition products[33].

In this paper we wish to explore a novel approach for the identification of protein-binding fragment assemblies formed by dynamic covalent chemistry introduced here as the 'pre-substrate concept'. Central idea of this approach is to employ completely inactive reporter probes denominated as 'pre-substrates' for the detection of protein-binding fragments and/or to use such probes for the detection or discovery of hitherto unknown reactive sites on protein surfaces. Pre-substrates are supposed to contain a reactive site, for example, an electrophile like aldehyde or Michael acceptor moiety for covalent ligation with various nucleophiles and an N-acylated reporter dye such as the fluorophore 7-amino-4-methylcoumarin (AMC) for detection. As N-deacylation of the acylated 7-amino group in AMC shifts the fluorescence emission of the coumarin dye to longer wavelengths, the cleavage of this amide bond by a hydrolytic enzyme is detectable as a time-dependent increase in fluorescence. In this contribution we demonstrate that electrophilic pre-substrates are valuable tools for the detection of biologically active fragments in a fluorescence-based enzyme assay and that the assay system enables the quantification of the templating effect exerted by the target protein for each fragment–pre-substrate combination.

Results

The concept of pre-substrate activation. To establish the assay system, serine proteases were selected as model targets (Fig. 1). Nucleophilic fragments, which bind to the S1 substrate-binding pocket of a serine protease were expected to assemble with a pre-substrate in a protein-templated reaction under formation of a covalent bond between the fragment nucleophile with the electrophile of the pre-substrate. In case of an optimal fit to the protein, this fragment assembly should position the amide bond of the pre-substrate at the active site of the protein and thus induce cleavage of the pre-substrate by the enzyme, which would result in the release of the coumarin fluorophor and generate a time-dependent increase in fluorescence. The experimental set-up should enable the parallel testing of several fragments, for example, in microtitre plates. Determination of the rate of the enzymatic reaction and its concentration-dependent saturation will then allow us to determine the stability of the enzyme–substrate complex (apparent K_M) consisting of a protease, the S1-binding nucleophile fragment and the pre-substrate. The recorded K_M values reflect the thermodynamic stabilization of the fragment–pre-substrate–enzyme complex (through $K_M \leq K_D$ and $G \leq RT \ln K_M$) and thereby quantify the templating effect exerted by the enzyme, resulting in the shift of the ligation equilibrium. Thereby, the pre-substrate concept will enable the quantitative analysis of dynamic ligation reactions of nucleophiles and electrophiles on protein templates.

N-Glyoxalyl-AMC 1 was designed and synthesized as the first potential pre-substrate, (Fig. 1, Supplementary Fig. 1). In aqueous solution, 1 is hydrated as indicated by the corresponding acetal proton at 5.99 pm in the ^{1}H-NMR in acetone/D$_2$O and by the absence of the aldehyde proton, which can be observed at 10.08 ppm in dimethyl sulfoxide (DMSO) solution. Hydration of 1 is favoured thermodynamically in aqueous solution due to the electrophilicity of 1 and due to the high concentration of the water nucleophile (55 M). As a result, at physiological pH of 7.4 the equilibrium of ligation reactions of 1 with most O-, N- or S-nucleophiles providing hemiacetals, imines or thiohemiacetals as ligation products is on the side of the hydrate—with few exceptions like aromatic amines. We hypothesized, however, that despite the favoured hydration of 1, a templating effect exerted by a fitting protein surface might still enable the formation of fragment ligation products of 1 and a nucleophile. If these ligation products can be formed, they might position the amide bond of 1
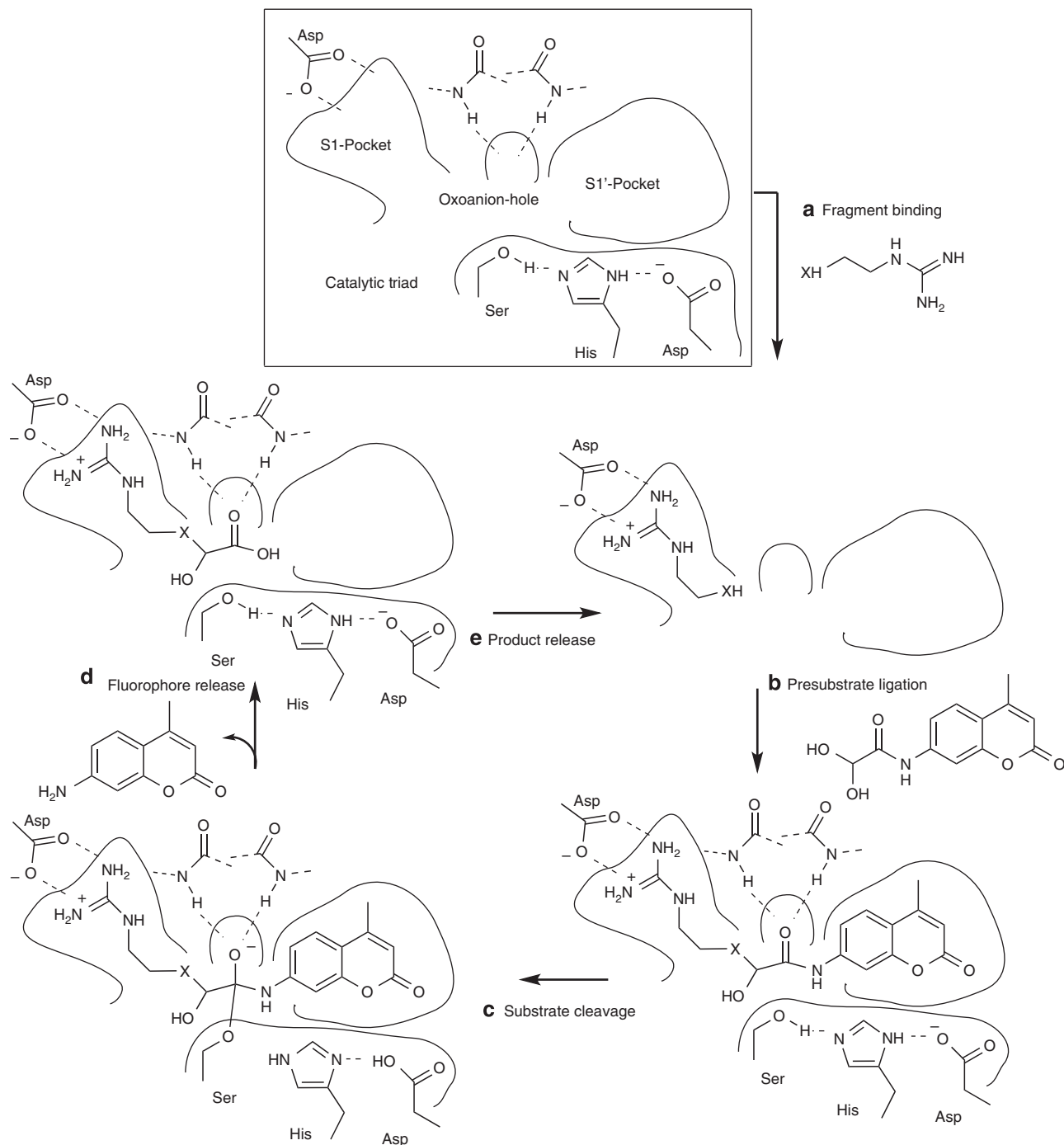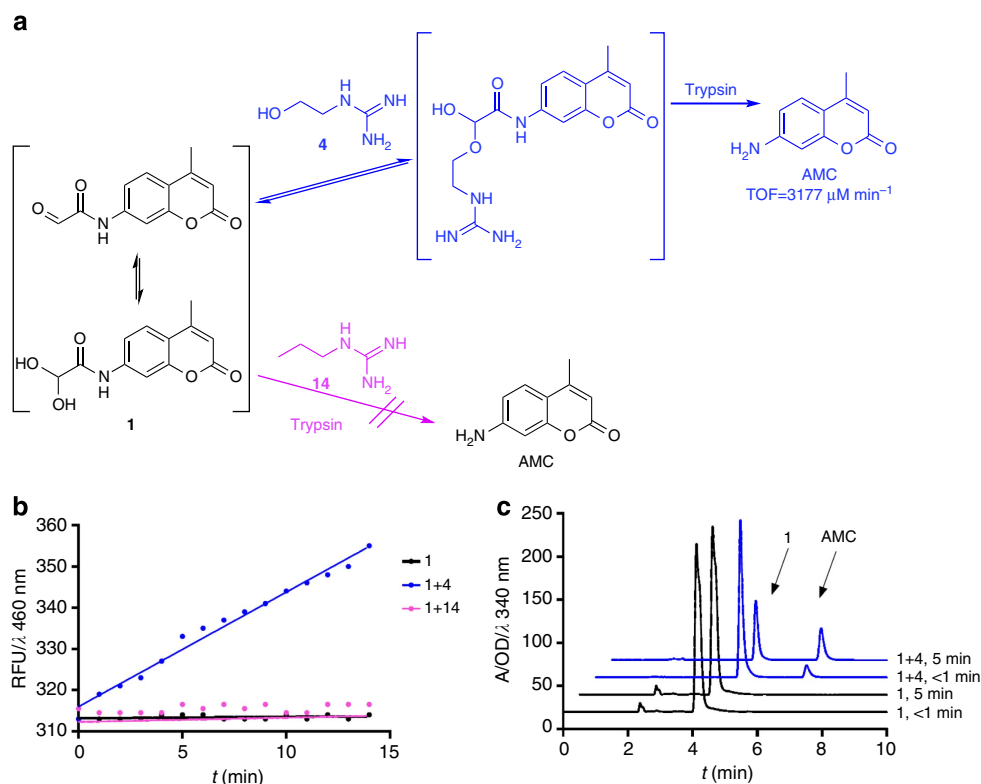
Figure 1 | The concept of pre-substrate activation by protein-binding fragments. Fragments bind to the S1 pocket of trypsin-like proteases with reported affinity (**a**). Only in case where the nucleophile is positioned correctly, ligation of a pre-substrate occurs (**b**), resulting in cleavage (**c**) of the formed substrate and release of the cleavage products (**d,e**).

in the active site of the enzyme and thereby induce the cleavage of the pre-substrate. Thus, observing amide bond cleavage of **1** will prove both the covalent ligation of both fragments and the templating effect of the protein surface.

As target proteins, two serine proteases were selected, factor Xa (FXa) and trypsin, two enzymes with well-described preferences for substrate recognition in the S1 pocket, which is adjacent to the protein's active site and thus might be accessible by fragments ligated to the aldehyde position of **1**. The first target, FXa, is a key-player in the blood coagulation cascade and an established drug target for inhibition of haemostasis. It prefers to cleave

substrates with arginine (Arg) next to the cleavage site[34]. The second target, trypsin, not only hydrolyzes peptides with arginine at P1, but also those endowed with lysine (Lys) at the cleavage site and plays its major role in digestion in the duodenum[35]. While trypsin has a broad substrate tolerance, FXa has rather specific substrate requirements.

Proof-of-concept. It was attempted to prove the validity of the pre-substrate concept with respect to the potential pre-substrate **1** (Fig. 2). The small peptide substrate Boc-Leu-Gly-Arg-AMC or

3

Figure 2 | Proof-of-concept. Pre-substrate **1** was completely inactive towards cleavage with trypsin (**a**, black). Addition of fragment **4**, a ligand that binds to the S1 pocket of trypsin, resulted in the activation of the pre-substrate, cleavage of the amide bond and the release of the reporter dye 7-amino-methyl coumarin (AMC; **a**, blue). No fluorescence was emitted in the presence of negative control fragment **14** (**a**, magenta). Formation of AMC was followed by fluorescence spectroscopy (**b**) and HPLC (**c**) and can be rationalized by formation of a hemiacetal ligation product and its recognition by the protein. The experiments were replicated twice by two different operators with duplicate repeated measures within each replication. Each data point represents the average of double determination.

the shortened amino acid substrate Z-Arg-AMC were incubated with one of the proteases (3.5 nM) in a 50 mM 3-(N-morpholino)propanesulfonic acid buffer (MOPS buffer) at pH 8 and room temperature. For both enzymes a distinct increase in fluorescence was recorded indicating the efficient turnover of both substrates. Next, the potential pre-substrate **1** was incubated with one of the proteases under the same conditions. Even at concentration up to 12.6 mM absolutely no conversion of **1** was recorded and thus **1** was proven not to be a substrate of either trypsin or FXa.

Next, **1** (3 mM) was incubated in the very same assay with trypsin (3.5 nM) and a strong excess of N-(2-hydroxyethyl) guanidine **4** (7.7 mM) as a potentially activating nucleophilic fragment. **4** was selected for this purpose as it was expected to bind to the S1 site of the protein and could be able to form an isosteric ligation product with **1**, which might position the AMC-amide bond at the active site of the protease. Indeed, the mixture of **1** and **4** with each of the proteases produced a time-dependent, linear increase in fluorescence at 460 nm (Fig. 2b). The observed effect depended critically on the presence of the protease: pre-substrate **1** and fragment **4** in buffer without addition of the protease did not affect cleavage of **1** and an increase in fluorescence. For another negative control, the non-nucleophilic control fragment propyl guanidine **14** was incubated with **1** and the protease at high concentrations. No cleavage of the pre-substrate was observed under these conditions, either. This experiment excludes the theoretical possibility that the guanidine fragment can induce cleavage of the pre-substrate by binding to S1 without ligating to the pre-substrate. For a secondary and independent confirmation of pre-substrate activation, the

reaction mixture consisting of pre-substrate **1** (0.3 mM), trypsin and bis-nucleophilic fragment **4** (2-aminoethanol) was analyzed by high-performance liquid chromatography (HPLC) at different time points. The fluorescence increase at 430 nm was found to correlate with cleavage of the amide bond in **1** and the release of non-acylated AMC (Fig. 2c, Supplementary Fig. 8). Additionally, possible side reactions of pre-substrate **1** with side chains of trypsin were excluded using MALDI-ToF/ToF mass spectrometry (Supplementary Figs 9–11). In summary, these experiments confirmed that pre-substrate **1**, which was completely inactive towards the protease alone, was activated by covalent dynamic ligation with nucleophile **4**.

Potential pre-substrates and fragments. The scope of the pre-substrate concept was to be investigated with respect to the selection of nucleophiles, which undergo the activating ligation reaction, and with respect to the pre-substrates to be used. In addition to the first pre-substrate, N-glyoxalyl-AMC **1** (2-oxo-N-(4-methyl-2-oxo-2H-chromen-7-yl)-acetamide), N-malonalyl-AMC **2** (3-oxo-(N-(4-methyl-2-oxo-2H-chromen-7-yl)-propanoylamide) and N-acryloyl-AMC **3** (N-(4-methyl-2-oxo-2H-chromen-7-yl)-acrylamide) were designed as potential alternatives (Fig. 3, Part a). In pre-substrate **2** a methylene bridge was inserted between the acyl-AMC and the aldehyde functionality, resulting in fragment ligation products elongated by one carbon. In addition, the CH acidity at C2 of the malonyl aldehyde should facilitate the enolization and thereby might favour the formation of ligation products compared with **1**. Acrylamide **3** was chosen to test the possibility of Michael additions with nucleophiles
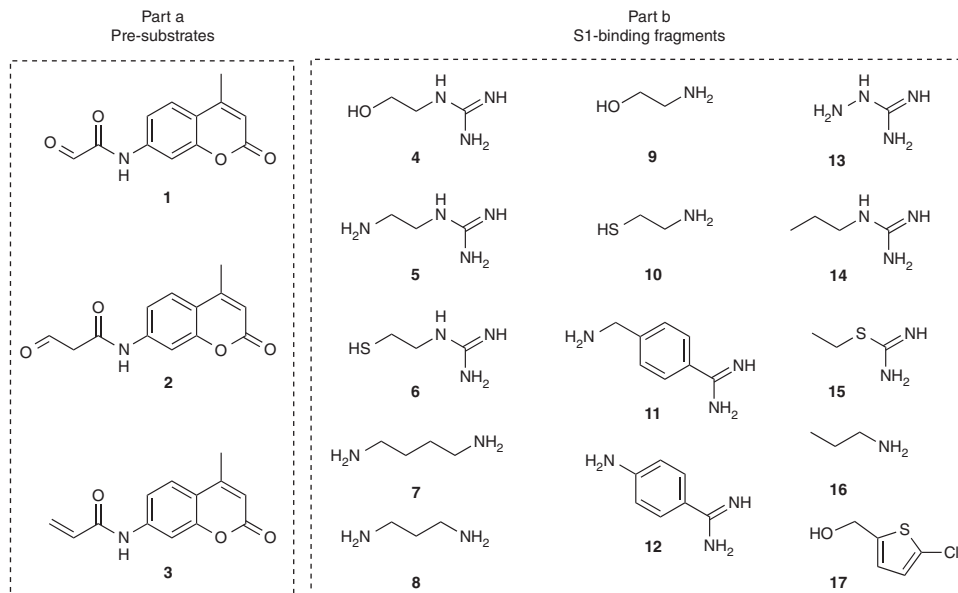
Figure 3 | Structure of potential pre-substrates 1–3 and S1-binding fragments 4–17. The details of the synthesis of compounds **1–17** as other methodological information are reported in the Supplementary Methods.

potentially yielding bioactive substrate structures. Both compounds **2** and **3** were completely stable towards both proteases under assay conditions without nucleophilic fragments, establishing them as pre-substrates. A small collection of 14 fragments (**4–17**) was composed containing various N-, O- and S-nucleophiles (amines, alcohols and mercaptanes) and non-nucleophilic control fragments as potential mimetics of the arginine or Lys side chains (Fig. 3, Part b). The already mentioned N-(2-hydroxyethyl)guanidine **4** was complemented by its amino- and mercapto-derivatives (**5, 6**), the diamines **7** and **8** as potential mimetics of the Lys side chain were complemented by ethanolamine **9** and 2-mercaptoethanamine **10**. 4-(amino-methyl)benzamidine **11**, 4-aminobenzamidine **12** and amino-guanidine **13** represent further classical mimetics of the arginine side chain. N-propylguanidine **14**, S-ethylisothiourea **15** and propylamine **16** were selected as negative controls, while 5-chloro-2-thiophenemethanol **17** was included as the fragment binding to the S1 pocket of FXa in the drug rivaroxaban (Fig. 3, Part b).

Screening with different nucleophiles. All fragments were investigated separately at identical concentrations on 96-well microtitre plates with and without one of the proteases for activation of the three pre-substrates. No increase of fluorescence was observed for any of the fragments together without protease in the reaction buffer. *In situ* formation of substrates by dynamic ligation of a fragment and a pre-substrate was detected via release of the fluorescent reporter molecule, AMC by amide bond cleavage. AMC-release was followed by excitation at 380 nm and fluorescence detection at 460 nm. The fluorescence intensity at 460 nm was recorded in each well for a time period of 15 min and the slope was determined for a comparison of all fragment–pre-substrate combinations. For each of the three pre-substrates, several activating fragments were identified, whereas none of the pre-substrates was activated by one of the negative control fragments **14–16**. Glyoxalyl-AMC **1** was activated by 9 out of 14 fragments to a substrate of trypsin, and by 7 out of 14 fragments to a substrate of FXa. **9** (ethanolamine), **7** (butane-1,4-diamine), **11** (4-(aminomethyl)benzamidine) and **4** (2-hydroxyethylguanidine) were found to be most active for **1**
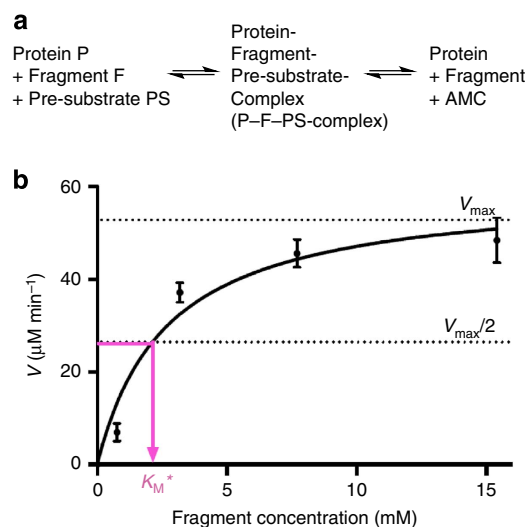
Figure 4 | Model for the activation of pre-substrates by nucleophilic protein-binding fragments. One protease molecule, a fragment and a pre-substrate form a three-component P–F–PS complex, which releases the coumarin dye AMC as a result of the enzymatic activity of the protease (**a**). The apparent K_M value, reflecting the stability of the P–F–PS complex, can be derived from the half maximal rate of the enzymatic reaction $V_{max}/2$ as exemplified here for the combination of protease FXa, fragment **7** and pre-substrate **1** (**b**).

and trypsin (Table 1, entry 6, 4, 7 and 1), whereas ethanolamine **9**, 2-hydroxyethylguanidine **4** and butane-1,4-diamine **7** were established as the most active molecules for **1** and FXa (Table 1, entry 6, 1 and 4). Malonalyl-AMC **2** was turned into a trypsin substrate by 11 out of 14 non-control fragments, and into a FXa substrate by 10 fragments. In the presence of trypsin, **2** was best activated by fragments **6** and **9** (Table 1, entry 3 and 6); in the presence of FXa by fragments **4**, **9** and **11** (Table 1, entry 1, 6 and 7). In general, as summarized in Table 1, both the activating effect of different nucleophiles varied between the pre-substrates and between the two proteases (Supplementary Figs 5 and 7). Glyoxalyl-AMC was most efficiently activated towards trypsin by

Table 1 | Quantitative determination of pre-substrate activation.

Entry	Fragment*	Trypsin				FXa			
		1[†]		2[‡]		1[†]		2[‡]	
		K_M^* (mM)	V_{max} (μM min^{-1})	K_M^* (mM)	V_{max} (μM min^{-1})	K_M^* (mM)	V_{max} (μM min^{-1})	K_M^* (mM)	V_{max} (μM min^{-1})
1	4	1.55 ± 0.6	27.49 ± 2.82	2.53 ± 1.56	33.09 ± 7.39	2.25 ± 0.65	12.91 ± 1.66	2.84 ± 0.56	86.59 ± 5.20
2	5	1.96 ± 1.5	40.43 ± 8.36	0.94 ± 0.14	52.37 ± 1.53	3.15 ± 0.71	15.00 ± 2.39	2.48 ± 0.83	73.90 ± 07.10
3	6	5.97 ± 0.4	26.75 ± 2.73	0.5 ± 0.26	91.24 ± 7.47	2.96 ± 0.9	43.82 ± 18.13	1.93 ± 0.71	120.01 ± 9.93
4	7	2.08 ± 0.5	17.54 ± 1.42	0.94 ± 0.32	136.33 ± 8.84	2.68 ± 0.91	58.48 ± 5.94	3.63 ± 0.98	242.28 ± 10.90
5	8	2.01 ± 0.4	26.63 ± 1.74	2.33 ± 1.21	79.32 ± 11.62	2.4 ± 1.2	19.06 ± 2.78	2.847 ± 1.1	174.15 ± 2.60
6	9	1.92 ± 0.8	22.61 ± 2.61	0.15 ± 0.06	66.11 ± 2.14	1.64 ± 0.11	35.18 ± 5.73	0.13 ± 0.07	73.06 ± 2.52
7	11	0.32 ± 0.2	38.55 ± 3.62	0.3 ± 0.15	39.31 ± 2.38	2.24 ± 1.1	13.18 ± 1.99	1.82 ± 0.73	29.13 ± 3.02

FXa, factor Xa.
Apparent K_M values and V_{max} rates for the various combinations of serine proteases, pre-substrates and nucleophilic fragments.
No activation of pre-substrates was detected with fragments **14**, **15** and **16**. Relative fluorescence units (RFUs; λ 460 nm) were determined as 47.465 RFU per μM_{AMC} (Supplementary Fig. 6).
The experiments were replicated twice by two different operators. Each data point represents for the average of double determination. Error bars indicate s.e. (Supplementary Figs 5 and 7).
*Only low activation was detected by fragments **10**, **12**, **13** and **17** (results not shown).
†Pre-substrate **1**.
‡Pre-substrate **2**.

the hydroxy-derivatives of the Lys and arginine mimetics, followed by the diamines (forming Lys isosters) and then the amino-substituted guanidines (forming Arg isosters). FXa was more selective in the acceptance of fragments (and ligation products), and had distinct preferences with respect to the strength of activation.

Next, a quantitative description of the observed activation of pre-substrates by nucleophilic fragments and of the templating effect exerted by the enzymes was to be established. For the purpose, a modified Michaelis–Menten model for enzymatic reactions was adapted to the special case of pre-substrate activation. In this model, the protease P was expected to bind a nucleophilic fragment F in its S1 pocket, which reacted in a dynamic and covalent ligation reaction to form the three-component protease–fragment–pre-substrate (P–F–PS) complex. Taking into account the S1 affinity of the employed nucleophiles and the non-detectable binding of the pre-substrates to the protein alone, one can assume that the binding of the nucleophilic fragment is the initial step of the complex formation. The P–F–PS complex is then converted by the enzyme yielding free AMC as the detectable product. Experimentally, the rate of the enzymatic reaction, V, corresponding to the cleavage of the ligation product can be determined. When V was measured at a fixed pre-substrate concentration (3 mM), the saturation of the reaction rate to V_{max} could be determined for increasing fragment concentrations. The fragment concentration at the half maximal rate ($V_{max}/2$) was considered as the apparent K_M value (K_M^*) of the ligation product and indicates the stability of the enzyme–ligation product complex, the P–F–PS complex (Fig. 4).

It should be noted that the thermodynamic stabilization of the P–F–PS complex expressed as K_M^* value is a direct measure for the templating effect exerted by the enzyme, which corresponds to the free energy of complex formation and describes the shift in the equilibrium from the free fragments to the ligated fragments. Determinations of both K_M and V_{max} values were conducted for the seven most active fragments, two pre-substrates and two enzymes (Table 1). The apparent K_M values of the three-component P–F–PS complexes varied between 130 μM (for the combination of PS **2**, F **9** and FXa as protease) and ca. 6 mM (for the combination of PS **1**, F **6** and trypsin). The V_{max} values were between 12.9 μM min^{-1} (PS **1**, F **4** and FXa) and 242.3 μM min^{-1} for the combination of PS **2**, F **7** and FXa. No stringent correlation between the K_M and V_{max} values was observed for both enzymes and pre-substrates, however, in most of the cases, the V_{max} was higher with PS **2**; in 11 out of 14 cases it

was more than twofold. The kinetic analysis of pre-substrate activation revealed that the reaction process was catalytic in a double sense. Not only the proteases but also the nucleophilic fragments themselves acted as catalysts in the conversion of the pre-substrate without being consumed or modified in the process. While the catalytic action of proteases as enzymes was expected, the organo-catalytic effect[36,37] of the protein-binding nucleophile was remarkable, as it constitutes to our best knowledge the first example of a catalytic activity of a small molecule in a protein-templated reaction.

Michael addition ligation. The third pre-substrate, N-acrylamido-AMC **3**, was selected as a model for irreversible pre-substrate activation. Activation of **3** was investigated in the presence of thiol nucleophiles, which were expected to yield Michael addition products being stable under assay conditions[38,39]. Thus, pre-substrate **3** was incubated with one of three thiol nucleophiles, 2-mercaptoethyl-guanidine **6**, cysteamine **10** and 4-amino benzenethiol in the assay buffer for 30 min at room temperature (20 °C). Subsequently, trypsin was added and time-dependent fluorescence was recorded at 480 nm (Fig. 5). The thiol fragments **6** and **10** were active and led to the release of AMC, while no activation was observed with the negative control fragment, 4-aminobenzenethiol. For an independent confirmation of the observed pre-substrate activation, the Michael addition product **18**, formed by reaction of cysteamine **10** and presubstrate **3** was synthesized at 35 °C (Supplementary Fig. 4), isolated and tested in the presence of trypsin. leading to a comparable release of AMC (Fig. 5, method B). Thus, it has been demonstrated that the activation of pre-substrates can be effected not only by reversible but also by irreversible ligation reactions.

Three-fragment assembly. Finally, it was investigated, if the pre-substrate concept could be extended further from two-fragment assemblies as demonstrated above to the protein template-assisted formation of bioactive assemblies consisting of three-components. This investigation was based on the following consideration. Pre-substrate **2** was found to be activated efficiently by butane-1,4-diamine **7**. If one assumes that this activation proceeds via formation of the amino-butyl-imine **19** as the substrate and reversibly formed two-fragment assembly, one can hypothesize that a third fragment might be able to add to the imine **19**, yielding an assembly of three fragments with potentially enhanced substrate activation (Fig. 6). To test this hypothesis,

fragment **7** was incubated with one of the two proteases (trypsin or FXa), pre-substrate **2** and either **6** or **10** as third fragments. As described above, fragment **6**, **7** and **10** activated the
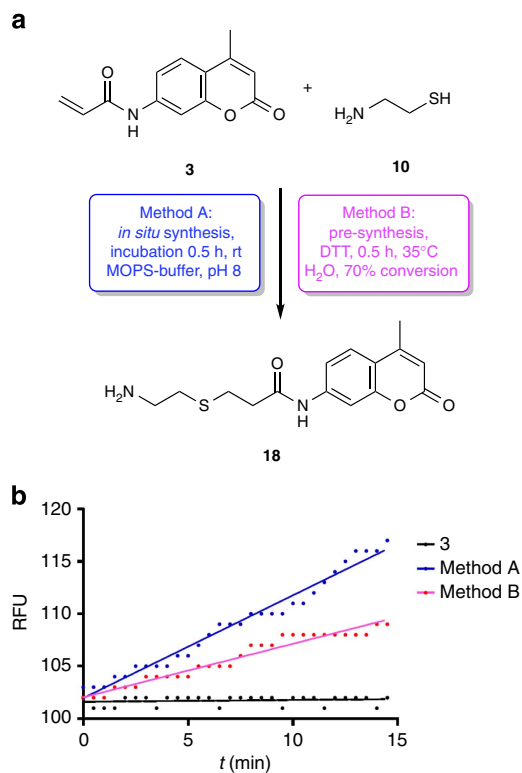
Figure 5 | Activation of pre-substrate 3. (**a**) Reaction scheme. (**b**) Recorded fluorescence of released AMC over time. Influence of ligation product (*in situ* formed from **3** and **10**, method A, blue) in comparison with isolated substrate **18** on the enzyme activity (method B, magenta). The experiments were repeated twice by two different operators with duplicates measured within each replication. Each data point represents the average of double determination. DTT, 1,4-dithio-D-threitol; rt, room temperature.
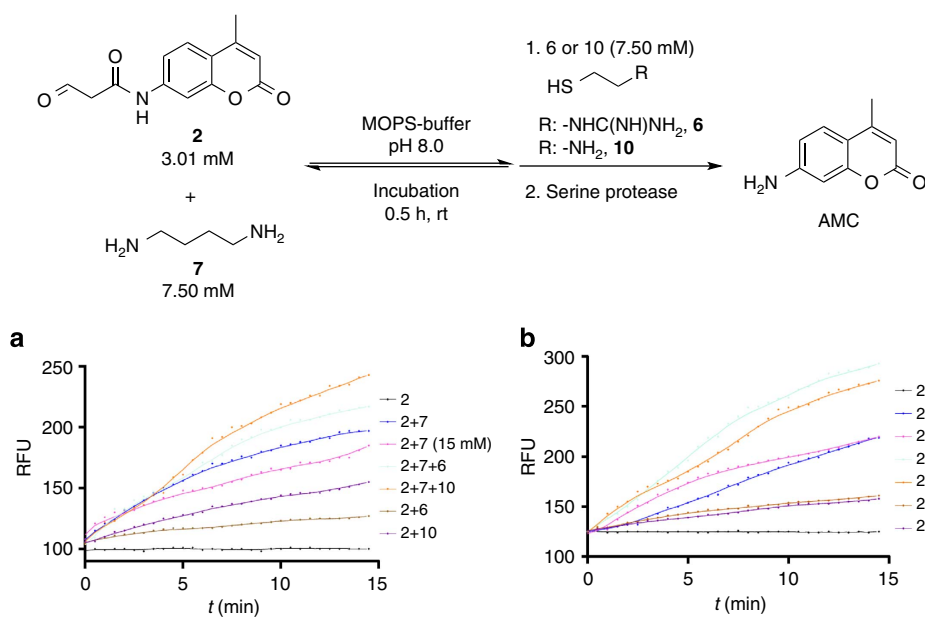
pre-substrate **2** when added as secondary fragment. Addition of the third fragment generated for both enzymes and for both thiol fragments stronger AMC fluorescence. For trypsin, the strongest activation of pre-substrate **2** was observed with fragments **7** and **10** by 56% or 50 relative fluorescence units (RFUs) after 15 min (Fig. 6a); for FXa, the combination of **2** with **7** and **6** resulted in the strongest activation by almost 100% or 80 RFU after the same time (Fig. 6b). In both cases, the activation by three fragments was compared with activation of a doubled concentration of fragment **7** (15 mM instead of 7.5), which yielded no or low increase in fluorescence, indicating that there was indeed an enhancing effect of the third fragment added.

Discussion

It has been demonstrated that protein-binding small molecule fragments can induce chemical reactions on the protein surface in a catalytic process. Nucleophilic, low molecular weight fragments binding to the S1 pocket of serine proteases were able to activate electrophilic pre-substrates in a protein-templated reaction. Activation of pre-substrates occurred via a dynamic covalent ligation reaction on the protein surface and both the free nucleophile and the presence of the enzyme were essential for this process. The reversible ligation reaction of the protein-bound amino-, hydroxy-, or mercapto-nucleophile with the electrophile positioned the pre-substrate into the active site of the enzyme and thereby induced the enzymatic hydrolysis of the protein–fragment–pre-substrate assembly. The activation process was studied quantitatively with three pre-substrates containing *N*-acylated AMC as a fluorogenic dye, which upon hydrolysis released the fluorophore 7-amino-methyl coumarin. The cleavage reaction was monitored by fluorescence spectrometry.

The mechanism of pre-substrate activation was confirmed by the results of several control experiments. Only S1-binding fragments containing a free nucleophile available for dynamic covalent ligation were able to activate pre-substrates. No cleavage of the pre-substrate was observed without presence of a protease. Second, HPLC analysis of the assay solution revealed that the increase in fluorescence indeed correlated with the hydrolytic cleavage of pre-substrates and the release of 7-amino-methyl

Figure 6 | Three-fragment assembly. Influence of nucleophilic fragments **7**, **6** and **10** and pre-substrate **2** on the enzyme activity; (**a**) trypsin, (**b**) factor Xa. The experiments were replicated twice by two different operators. Each data point represents the average of double determination.

coumarin. The rate of the pre-substrate cleavage (V) was measured as a function of the fragment concentration for all combinations of one of eight active fragments with one of two pre-substrates and proteases. In all cases, the substrate cleavage rate V increased with growing fragment concentrations and could be saturated when approaching the maximal conversion rate (V_{max}). The stability of the protease–fragment–pre-substrate (P–F–PS) complex, K_M^*, was determined for each of the tested P–F–PS combinations as the fragment concentration at the half maximal velocity ($V_{max}/2$). The stability of the P–F–PS complex is a direct measure for the free enthalpy of complex formation and thus the K_M^* values indicate quantitatively the shift in the ligation equilibrium exerted by the enzyme's templating effect. Both the V_{max} and the K_M^* values were specific for each of the tested P–F–PS combinations, suggesting both a specific recognition effect exerted by the protease and a specific catalytic activation effect of the bis-nucleophile. No correlation between the stability of the P–F–PS complex and the maximal cleavage rate was observed, however, the cleavage rates were consistently higher for pre-substrate **2** compared with pre-substrate **1**, suggesting that the latter compound ensures—at least for the activating fragments tested here—a better positioning of the cleaved amide bond in the active site of the enzyme.

The pre-substrate concept was further extended from reversible towards irreversible, covalent fragment ligations as demonstrated with the Michael addition of thiol nucleophiles to the acrylamide containing pre-substrate **3**. Finally, pre-substrate activation was conducted not only by dimeric ligations (with two fragments) but also using hetero-trimeric fragment assemblies, if one amine and one thiol nucleophile were employed.

The reported activation of pre-substrates constitutes a novel approach to identify small molecule fragments binding specifically to defined protein-binding pockets. The described assays—demonstrated here for the S1-pockets of two proteases, trypsin and FXa—enable the quantitative detection of protein-templated ligation reactions. The templating effects responsible for shifting the ligation equilibrium can be determined quantitatively. One would further expect that the activation of pre-substrates can be applied to other proteases as well and can be extended towards other secondary binding sites and towards other classes of hydrolytic enzymes, whenever a protein-bound primary fragment is capable to react with a secondary fragment and thereby direct cleavable bond at the active site of the protein. The assays can be extended to the screening of larger libraries of nucleophilic fragments and the sensitized detection of fragments binding to defined pockets of enzymes might serve as a starting point for the development of site-directed inhibitors of these enzymes as well. Libraries of nucleophilic fragments are smaller and less diverse than unbiased fragment libraries. On the other hand, presence of a nucleophilic functionality in all hits of a pre-substrate screen will facilitate the subsequent extension of hit fragments, possibly in an iterative ligation screening step.

Beyond its practical implications, pre-substrate activation by protein-binding nucleophiles constitutes an example for a coupling of dynamic covalent ligation reaction with an enzymatic reaction. The protein-binding nucleophilic fragments act as organo-catalysts for the hydrolysis of the pre-substrate by the enzyme. Such organo-catalytic reactions of protein-binding fragments in fact enlarge the substrate pool of enzymes by pre-substrates and thereby extend the realm of exploitable enzymatic reactions.

Methods
General methods. The details of the synthesis of compounds **1–18** (Supplementary Figs 1–4) as other methodological information are reported in the Supplementary Methods. The activity of the serine proteases was determined by measuring the release of AMC from the fluorogenic substrate synthesized *in situ* on 96-well microtitre plates using a Tecan Safire[2] wellplate reader. The excitation wavelength was set to 380 nm and emission was recorded at 460 nm. For the analysis of the data, GraphPad Prism 5 was used. Purity characterization of trypsin and FXa and also the investigation of side reactions between pre-substrate and proteins were analyzed using MALDI-MS. MALDI-ToF/ToF MS was performed on Ultraflex III, Bruker Daltonics 2007 equipped with a smartbeam laser (Nd:YAG). SDHB (mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid) or CA (alpha-cyano-4-hydroxycinnamic acid) from Bruker Daltonics were used as matrix materials. Buffers and reagents were purchased from Roth and Sigma-Aldrich. Water was purified by ultrafiltration (Millipore Simplicity 185). Human FXa was obtained as lyophilized powder by Enzyme Research Laboratories Ltd (Catalogue: HFXa 1011, activity: 228 IU mg^{-1}) and dissolved in BisTris-buffer (bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane, pH 6). Bovine trypsin was acquired by Sigma-Aldrich as lyophilized powder (T1426 Sigma, TPCK treated, essentially salt free, > 100,000 BAEE U mg^{-1} protein) and taken up in 1 mM HCl. MOPS-NaOH buffer pH 8 was prepared in ultra pure water. 3-(N-morpholino)-propanesulfonic acid (MOPS, 50 mM) and NaCl (154 mM) were dissolved in water; pH was adjusted with NaOH (stored at 4 °C).

DLS procedures. *First procedure.* Pre-substrate activation by nucleophilic protein-binding fragments. First screening procedure was performed using the pre-substrate **1** or **2** (3 μl of 233 mM stock solution in DMSO resulting in 3.07 mM assay concentration) and one of the various nucleophilic fragments **4–17** (3 μl of 0, 60, 240, 580 or 1,160 mM stock solution in DMSO resulting in 0, 0.77, 3.2, 7.7 or 15.4 mM assay concentration) in a total volume of 206 μl MOPS-NaOH buffer (pH 8, 50 mM) per well. Fluorescence was monitored starting immediately after the addition of 20 μl of the serine protease solution (3.5 nM trypsin or 3.5 nM FXa; 1/1,000 dilution in bovine serum albumin (BSA) 0.1% (wt/vol), NaCl 154 mM). The initial rates were determined and plotted against the fragment concentration (Supplementary Figs 5 and 7).

Second procedure. Three-fragment assembly screening. The three-fragment assembly screening was investigated using pre-substrate **2** (3 μl of 233 mM DMSO stock solution equates to 3.07 mM assay concentration) and fragment **7** (3 μl of 580 or 1,160 mM DMSO stock solution equates to 7.7 mM or 15.4 mM assay concentration) in a total volume of 206 μl MOPS-NaOH buffer (pH 8, 50 mM). After incubation (30 min, room temperature) the second nucleophilic fragment **6** or **10** (3 μl of 580 mM DMSO stock solution equates to 7.7 mM assay concentration; Fig. 2, Part b) and 20 μl of the serine protease solution (3.5 nM trypsin or 3.5 nM FXa; dilution in BSA 0.1% (wt/vol), NaCl 154 mM) was added. The fluorescence intensities were plotted over time (Fig. 6).

Third procedure. Michael addition screening. Third screening procedure was performed using pre-substrate **3** (3 μl of 233 mM DMSO stock solution equates to 3.07 mM assay concentration) and cysteamine **10** (3 μl of 580 mM DMSO stock solution equates to 7.7 mM assay concentration) in a total volume of 206 μl MOPS-NaOH buffer (pH 8, 50 mM). After incubation (30 min, 35 °C) 20 μl serine protease (3.5 nM trypsin or 3.5 nM FXa; dilution in BSA 0.1% (wt/vol), NaCl 154 mM) was added. Relative fluorescence was recorded and plotted against the time (Fig. 5). For comparison the addition product **18** was pre-synthesized. Twenty microlitres of serine protease (3.5 nM trypsin or 3.5 nM FXa; dilution in BSA 0.1% (wt/vol), NaCl 154 mM) was added to the isolated compound **18** (5 μl of 140 mM DMSO stock solution equates to 3.01 mM assay concentration) in a total volume of 205 μl MOPS-NaOH buffer (pH 8, 50 mM). The measurement was started without secondary incubation.

Determination of RFUs. For the determination of RFUs, the fluorescence of AMC was measured at different concentrations. RFUs (RFUs; λ 460 nm) were found to correlate linearly with the AMC concentration at concentration between 0 and 8 mM. The calculation factor was determined as being 47.465 RFU per μM$_{AMC}$ and the initial rates in μM min^{-1} were plotted against the fragment concentrations (Supplementary Fig. 6).

HPLC investigations. Twenty microlitres of trypsin (3.5 nM; dilution in BSA 0.1% (wt/vol), NaCl 154 mM) was added to pre-substrate **1** (20 μl of 4 mM stock solution equates to 0.307 mM assay concentration) and fragment **4** (20 μl of 12 mM stock solution equates to 0.92 mM assay concentration) in total volume of 240 μl distilled water and mixed over 0, 5 and 25 min. Next, 260 μl acetonitrile was added to the probes, mixed and analyzed by HPLC. The turnover frequency (TOF) refers to the trypsin and was calculated on the basis of the amount of the product formed as well as the amount of trypsin used per minute. The reaction was detected at 340 and 430 nm (Supplementary Fig. 8).

MALDI-TOF/TOF MS investigations. *Purity characterizations.* Trypsin (10.2 μM; dilution in water) and FXa (42.6 μM; dilution in water) were analyzed separately with SDHB (mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid) or cinnamic acid (CA) matrix material (Supplementary Figs 9 and 10).

Investigation of side reactions between trypsin and pre-substrate 1. Twenty microlitres of trypsin (10.2 μM, dilution in water) was incubated with 2 μl

pre-substrate **1** (100 μM, dilution in water). The reaction mixture was analyzed using CA matrix material.

*Investigation of side reactions between trypsin, pre-substrate **1** and fragment **4**.* Twenty microlitres of trypsin (10.2 μM, dilution in water) was incubated with 2 μl pre-substrate **1** (100 μM, dilution in water) and 2 μl fragment **4** (100 μM, dilution in water). The reaction mixture was analyzed using CA matrix material; (Supplementary Fig.11).

References

1. Rees, D. C., Congreve, M., Murray, C. W. & Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discov.* **3,** 660–672 (2004).
2. Hajduk, P. J. & Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* **6,** 211–219 (2007).
3. Erlanson, D. A., McDowell, R. S. & O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **47,** 3463–3482 (2004).
4. Fattori, D., Squarcia, A. & Bartoli, S. Fragment-based approach to drug lead discovery. *Drugs RD* **9,** 217–227 (2008).
5. Rademann, J. Organic protein chemistry: drug discovery through the chemical modification of proteins. *Angew. Chem. Int. Ed.* **43,** 4554–4556 (2004).
6. Corbett, P. T. *et al.* Dynamic combinatorial chemistry. *Chem. Rev.* **106,** 3652–3711 (2006).
7. Erlanson, D. A. Fragment-based lead discovery: a chemical update. *Curr. Opin. Biotechnol.* **17,** 643–652 (2006).
8. Schmidt, M. F. & Rademann, J. Dynamic template-assisted strategies in fragment-based drug discovery. *Trends Biotechnol.* **27,** 512–521 (2009).
9. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274,** 1531–1534 (1996).
10. Vongvilai, P., Angelin, M., Larsson, R. & Ramström, O. Dynamic combinatorial resolution: direct asymmetric lipase-mediated screening of dynamic nitroaldol library. *Angew. Chem. Int. Ed.* **46,** 948–950 (2007).
11. Nienaber, V. L. *et al.* Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* **18,** 1105–1108 (2000).
12. Blundell, T. L., Jhoti, H. & Abell, C. High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discov.* **1,** 45–54 (2002).
13. Sharff, A. & Jhoti, H. High-throughput crystallographyto enhance drug discovery. *Curr. Opin. Chem. Biol.* **7,** 340–345 (2003).
14. Sanders, W. J. *et al.* Discovery of potent inhibitors of dihydroneopterin aldolase using CrystaLEAD high-throughput X-ray crystallographic screening and structure-directed lead optimization. *J. Med. Chem.* **47,** 1709–1718 (2004).
15. Patterson, A. W., Wood, W. J. L. & Ellman, J. A. Substrate activity screening (SAS): a general procedure for the preparation and screening of a fragment-based non-peptidic protease substrate library for inhibitor discovery. *Nat. Protoc.* **2,** 424–433 (2007).
16. Mares-Guia, M. & Shaw, E. Substrate side chain guanidines as model of the binding of amidines and studies on the active center of trypsin. *J. Biol. Chem.* **240,** 1579–1585 (1965).
17. Schmidt, M. F. *et al.* Sensitized detection of inhibitory fragments and iterative development of non-peptidic protease inhibitors by dynamic ligation screening. *Angew. Chem. Int. Ed.* **47,** 3275–3278 (2008).
18. Schmidt, M. F., El-Dahshan, A., Keller, S. & Rademann, J. Selective identification of cooperatively binding fragments in a high-throughput ligation assay enables development of a picomolar caspase-3 inhibitor. *Angew. Chem. Int. Ed.* **48,** 6346–6349 (2009).
19. Schmidt, M. F., Groves, M. & Rademann, J. Dynamic substrate enhancement for the identification of specific, second-site-binding fragments targeting a set of protein tyrosine phosphatases. *Chembiochem* **12,** 2640–2646 (2011).
20. Al-Gharabli, S. I. *et al.* An efficient method for the synthesis of peptide aldehyde libraries employed in the discovery of reversible SARS coronavirus main protease (SARS-CoV M^{pro}) inhibitors. *Chembiochem* **7,** 1048–1055 (2006).
21. Reek, J. N. H. & Otto, S. (eds) *Dynamic Combinatorial Chemistry* (Wiley-VCH, 2010).
22. Osowska, K. & Miljanic, O. S. Oxidative kinetic self-sorting of a dynamic imine library. *J. Am. Chem. Soc.* **133,** 724–727 (2011).
23. Hermann, A. Dynamic mixtures and combinatorial libraries: imines as probes for molecular evolution at the interface between chemistry and biology. *Org. Biomol. Chem.* **7,** 3195–3204 (2009).
24. Meyer, C. D., Joiner, C. S. & Stoddart, J. F. Template-directed synthesis employing reversible imine bond formation. *Chem. Soc. Rev.* **36,** 1705–1723 (2007).
25. Rodriguez-Decampo, Z. & Otto, S. Orthogonal or simultaneous use of disulfide and hydrazone exchange in dynamic covalent chemistry in aqueous solution. *Chem. Commun.* 5301–5303 (2008).
26. Dirksen, A., Yegneswaran, S. & Dawson, P. E. Bisaryl hydrazones as exchangeable biocompatible linkers. *Angew. Chem. Int. Ed.* **49,** 2023–2027 (2010).
27. Soutullo, M. D., O'Brien, R. A., Gaines, K. E. & Davis, Jr J. H. Constitutional dynamic systems of ionic and molecular liquids. *Chem. Commun.* 2529–2531 (2009).
28. Drahonovsky, D. & Lehn, J.-M. Hemiacetals in dynamic covalent chemistry: formation, exchange, selection and modulation processes. *J. Org. Chem.* **74,** 8428–8432 (2009).
29. Caraballo, R., Dong, H., Ribeiro, J. P., Jimenez-Barbero, J. & Ramström, O. Direct STD NMR identification of β-galactosidase inhibitors from a virtual dynamic hemithioacetal system. *Angew. Chem. Int. Ed.* **49,** 589–593 (2010).
30. Cacciapaglia, R., Di Stefano, S. & Mandolini, L. Metathesis reaction of formaldehyde acetals: an easy entry into the dynamic covalent chemistry of cyclophane formation. *J. Am. Chem. Soc.* **127,** 13666–13671 (2005).
31. Berkovich-Berger, D. & Lemcoff, N. G. Facile acetal dynamic combinatorial library. *Chem. Commun.* 1686–1688 (2008).
32. Larsson, R., Pei, Z. & Ramström, O. Catalytic self-screening of cholinesterase substrates from a dynamic combinatorial thioester library. *Angew. Chem. Int. Ed.* **43,** 3716–3718 (2004).
33. Shi, B. & Greaney, M. F. Reversible Michael addition of thiols as a new tool for dynamic combinatorial chemistry. *Chem. Commun.* 886–888 (2005).
34. de Candia, M., Lopopolo, G. & Altomare, C. Novel factor Xa inhibitors: a patent review. *Expert Opin. Ther. Pat.* **19,** 1535–1580 (2009).
35. Schröder Leiros, H.-H. *et al.* Trypsin specificity as elucidated by LIE calculations, X-ray structures, and association constant measurements. *Protein Sci.* **13,** 1056–1070 (2004).
36. List, B., Lerner, R. A. & Barbas, III C. F. Proline-catalyzed direct asymmetric aldol reactions. *J. Am. Chem. Soc.* **122,** 2395–2396 (2000).
37. Wilhelms, N., Kuchat, S. & Lehn, J.-M. Organocatalysis of C=N/C=N and C=C/C=N exchange in dynamic covalent chemistry. *Helv. Chim. Acta* **95,** 2635–2651 (2012).
38. Khatik, G. L., Kumar, R. & Chakraborti, A. K. Catalyst-free conjugate addition of thiols to α,β-unsaturated carbonyl compounds in water. *Org. Lett.* **8,** 2433–2436 (2006).
39. Hoyle, C. E. & Bowman, C. N. Thiol-ene click chemistry. *Angew. Chem. Int. Ed.* **49,** 1540–1573 (2010).

Acknowledgements

Author contributions

J.R. and E.B. conceived and designed the experiments, E.B. performed the experiments. Both authors discussed the results and co-wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Burda, E. and Rademann, J. Catalytic activation of pre-substrates via dynamic fragment assembly on protein templates. *Nat. Commun.* 5:5170 doi: 10.1038/ncomms6170 (2014).