ChemPubSoc Europe

DOI: 10.1002/cbic.201402192

# Repositioning the Substrate Activity Screening (SAS) Approach as a Fragment-Based Method for Identification of Weak Binders

Rafaela Gladysz,[a] Matthias Cleenewerck,[a] Jurgen Joossens,[a] Anne-Marie Lambeir,[b] Koen Augustyns,[a] and Pieter Van der Veken*[a]

Fragment-based drug discovery (FBDD) has evolved into an established approach for "hit" identification. Typically, most applications of FBDD depend on specialised cost- and time-intensive biophysical techniques. The substrate activity screening (SAS) approach has been proposed as a relatively cheap and straightforward alternative for identification of fragments for enzyme inhibitors. We have investigated SAS for the discovery of inhibitors of oncology target urokinase (uPA). Although our results support the key hypotheses of SAS, we also encountered a number of unreported limitations. In response, we propose an efficient modified methodology: "MSAS" (modified substrate activity screening). MSAS circumvents the limitations of SAS and broadens its scope by providing additional fragments and more coherent SAR data. As well as presenting and validating MSAS, this study expands existing SAR knowledge for the S1 pocket of uPA and reports new reversible and irreversible uPA inhibitor scaffolds.

## Introduction

Over the last decade, fragment-based drug discovery (FBDD) has become an established methodology, delivering several high-quality leads and clinical candidates and at least one FDA-approved drug (vemurafenib/PLX4032).[1] FBDD emerged as an alternative to high-throughput screening (HTS) for "lead" discovery. FBDD approaches have in common that they construct "lead" molecules from smaller fragments, typically containing fewer than 12 heavy atoms and possessing relatively low individual affinities. Such fragments in general allow orthogonal optimisation to meet predefined criteria for target affinity ("ligand-efficiency") and biopharmaceutical behaviour, as proposed in, for example, the "rule of three" for fragments.[3] Another advantage of the approach is that the compound libraries used for fragment identification can be multiple orders of magnitude smaller in size than those required for HTS. This relates to the fact that drug-like diversity space can be much more efficiently probed with small fragments than with the typically larger molecules found in drug-like HTS libraries.[2–5] There are several recent reviews that extensively document these concepts with relevant examples from case studies.[6]

Nonetheless, as summarised in a recent opinion article by Murray et al., FBDD still faces a number of fundamental challenges.[5] One of these relates to the need for better methodology to detect and to study weakly binding fragments. X-ray crystallography and protein NMR are among the most established techniques for this purpose.[7,8] However, their main drawback is the requirement for substantial amounts of highly purified target protein and specialised infrastructure that runs at high financial cost and with medium throughput capacity. Other biophysical techniques that bypass these limitations have been applied. These include, among others, surface plasmon resonance (SPR), $^{19}$F NMR, thermal denaturation, mass spectrometry, thermal electrophoresis and isothermal titration calorimetry.[5,9,10]

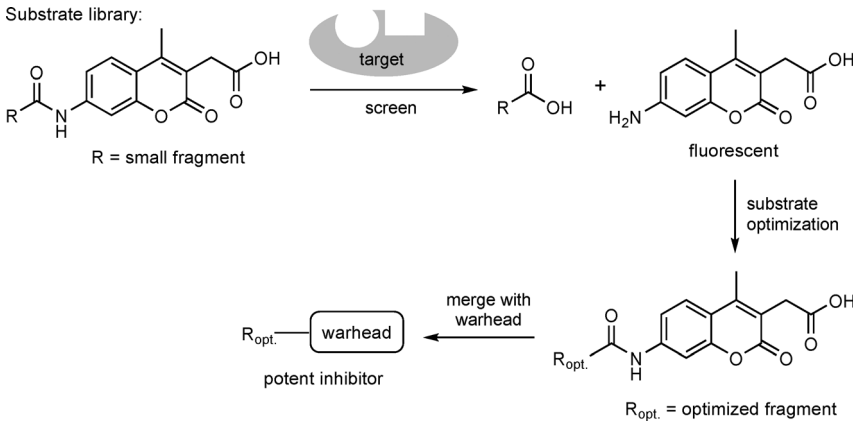
Specifically for enzyme targets, substrate activity screening (SAS) was proposed by Ellman and co-workers as an attractive fragment-based approach to inhibitor discovery.[11–16] Promising results with different classes of enzyme targets have been reported, mainly by the groups of Ellman and, more recently, Seebach.[17] Notable examples include identification of non-peptidic inhibitors for serine and cysteine proteases, receptor tyrosine kinases and the protein tyrosine phosphatase PtpB of *Mycobacterium tuberculosis*.[11–13,17] The SAS approach, demonstrated for a protease target, consists of three steps. First of all, a library of small, fragment-sized molecules, each linked to a scissile fluorogenic amide bond, is screened for substrates of a target protease (step 1, Scheme 1). Next, the identified substrates are optimised in a separate cycle (step 2) and finally transformed into inhibitors by replacing the scissile amide bond with a warhead functionality (step 3).[11] Diversity in the substrate library comes from drug-like, fragment-sized groups that function as potential affinity-conferring recognition units

[a] R. Gladysz, M. Cleenewerck, Dr. J. Joossens, Prof. Dr. K. Augustyns, Prof. Dr. P. Van der Veken
Medicinal Chemistry (UAMC)
Department of Pharmaceutical Sciences, University of Antwerp
Universiteitsplein 1, 2610 Wilrijk (Belgium)
E-mail: pieter.vanderveken@uantwerpen.be

[b] Prof. Dr. A.-M. Lambeir
Laboratory of Medical Biochemistry
Department of Pharmaceutical Sciences, University of Antwerp
Universiteitsplein 1, 2610 Wilrijk (Belgium)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201402192.




**Scheme 1.** The SAS approach, demonstrated for a protease target.

for the target enzyme. These are linked to a functionality that can be processed by the target enzyme, thereby releasing a quantifiable reporter molecule. The main rationale of the SAS methodology is that the cleavage efficiencies (expressed as the $k_{cat}/K_M$ ratio) for the individual library members are positively correlated with a fragment's affinity for the enzyme's transition-state-stabilising conformation and hence with its potential for inhibitor design. It is worth mentioning that this principle had already been recognised decades ago and has been applied extensively for discovery of substrate-derived enzyme inhibitors. SAS, however, does not rely on library molecules that are direct analogues of a target's natural substrates. In this way, it is not biased to deliver inhibitors with an overall biomolecule-derived architecture but has the unique potential to provide fragments with favourable, more drug-like structures immediately.

In practice, it is necessary to optimise fragments after the first stage. This is done by creating additional, directed chemical diversity around the best substrates identified. Optimised substrates are ultimately transformed into inhibitors by direct replacement of the enzyme-processed functionality in a substrate molecule with a mechanism-based warhead or pharmacophore.[11,18] Although only superficially examined for this purpose, SAS fragments could also be subjected to a standard FBDD-optimisation strategy for obtaining small-molecule inhibitors that do not draw upon a warhead functionality to gain target affinity.

We decided to apply SAS to inhibitor discovery for urokinase plasminogen activator (uPA), a trypsin-like serine protease that is overexpressed in metastasising solid tumours.[19,20] The enzyme is a valuable oncology target, but clinical development of its inhibitors has been problematic. This is most probably related to the doubtful biopharmaceutical performance of compounds developed so far and their insufficient selectivity with respect to other, phylogenetically related trypsin-like proteases. Nonetheless, the field of urokinase inhibitor discovery still sees highly interesting developments, such as with recent approaches based on bicyclic peptide constructs.[21]

Earlier, our group described selective, irreversible inhibitors of uPA with significant anti-metastatic activity in a rodent model of breast cancer.[22] Discovering structurally novel uPA inhibitors therefore continues to raise our interest. During our exploration of SAS we encountered a number of limitations of the reported approach. Here we describe a simple and effective alternative for SAS, which we have named "MSAS" (modified substrate activity screening). We demonstrate that screening the library for inhibitors of a target enzyme rather than for its substrates avoids false negatives: that is, fragments with high potential for inhibitor discovery that are not identified in a SAS assay. We also show that MSAS avoids false positives that can surface during a regular SAS assay, and runs with better cost and time efficiency. Furthermore, an FBDD strategy is reported to transform identified fragments into inhibitors that do not rely on a warhead functionality for target affinity. Finally, we also demonstrate with the aid of experimental data that the classical SAS step in which substrates are translated into inhibitors by addition of a warhead, although intrinsically highly valuable, does not per se lead to compounds of practical biopharmaceutical quality. Finally, the results show that adoption of the MSAS approach can not only circumvent limitations of the parent methodology, but can also offer additional potential for FBDD on enzyme targets.

## Results and Discussion

### Library synthesis

The start of our investigations was the synthesis of a SAS library of 137 fluorogenic *N*-acyl-7-amino-3-methylcoumarin substrates (*N*-acyl AMCs). All the compounds in the library contained fragment-sized *N*-acyl residues (MW < 150). Selection of around 90 of these residues was done in a non-target-biased manner, aiming to cover as much of "drug-like" chemical space as possible: steric, electronic and electrostatic parameters were taken into account. In addition, several target-biased subsets were prepared, containing moieties of known uPA inhibitors and/or fragments that might reasonably be anticipated to bind to the active centres of trypsin-like enzymes. Inclusion of these fragments as positive controls was considered most helpful for investigation of the intrinsic performance of SAS during fragment identification and the internal coherence of results obtained. Although highly interesting, potential issues of this type have not been investigated earlier. It is also worth mentioning that on the basis of the dimensions of fragments and SAS's reliance on enzymatic activity, processed substrates can reasonably be expected to be accommodated in the S1 region of the enzyme (i.e., the S1 pocket and the parts of the active centre immediately surrounding it). This consideration was taken into account during the selection of the positive control

set: this set contains mainly basic groups and S1-binding substituents of known uPA inhibitors.

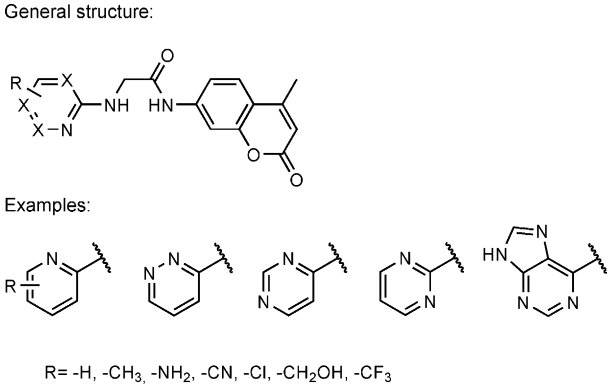
A substantial part of this library could be prepared by a one-pot protocol, by starting from the individual acyl residues and 7-amino-4-methylcoumarin and using Ghosez's reagent as the coupling mediator.[23,24] Of a large series of mild coupling reagents that we evaluated (DCC, EDC, TBTU, HATU, TFFH, PyBrop, Ghosez's reagent), only the last was found capable of cleanly and efficiently promoting the reaction with the very weakly nucleophilic 7-amino-4-methylcoumarin. For several compounds, additional steps (protection, homologation, functionalisation) were necessary in order to obtain the desired derivatives. The structures, the synthetic preparation and the characterisation data of all library members are listed in the Supporting Information.

### SAS experiment for the library of *N*-acyl AMCs

The prepared library of *N*-acyl AMCs was first screened for uPA substrates by a typical SAS protocol. All these experiments were conducted in duplicate in HEPES buffer at pH 8.2, thus allowing near-maximum enzymatic activity to be combined with minimal aspecific hydrolysis of *N*-acyl AMCs. An initial screening of the library was performed with 200 nM of recombinant human uPA and the highest substrate concentration allowed by compound solubility. Although these concentrations varied for the individual library members, they were generally in the 100–500 μM range. We reasoned that the use of high substrate concentrations in the exploratory phase of the project would allow identification of all library members that are processed by uPA, even those characterised by low $k_{cat}$ values.

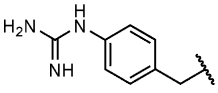
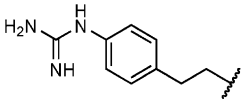
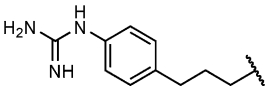
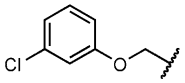
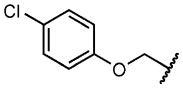
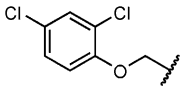
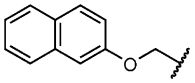
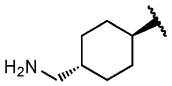
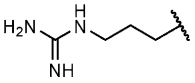
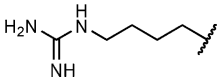
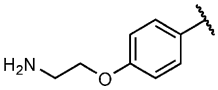
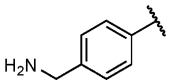
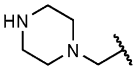
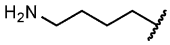
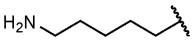
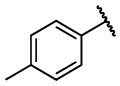
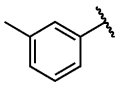
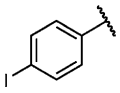
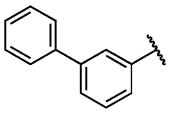
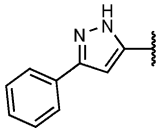
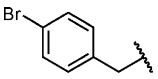
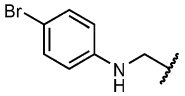
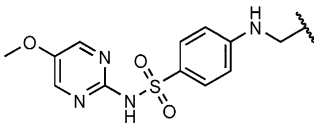
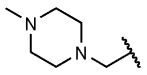
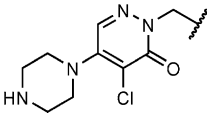
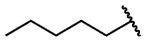
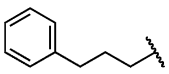
Under the initial conditions, eleven *N*-acyl AMCs from the library were found to behave as substrates of uPA, albeit with large differences in cleavage rate. To allow reliable ranking of cleavage efficiencies, assays for these compounds were repeated at subsaturating substrate concentrations ($[S] < K_M$). To avoid the need to determine $K_M$ values for all the obtained hits, we followed the approach proposed by Ellman et al.[11] Here, only the $K_M$ value of the optimal substrate in the series is determined. Subsequently, all initially obtained hits are investigated again at a concentration below the $K_M$ value of the best substrate, with this serving as a reference relative to which cleavage efficiencies are reported. We considered the guanidinophenyl-based compound **2** (Table 1), displaying a $K_M$ value of 120 μM, to be the best substrate. Subsequently, all initially obtained hits were rescreened at 100 μM concentration, and this reconfirmed guanidinophenyl derivative **2** as the most efficiently cleaved substrate in the series. The cleavage efficiencies of all eleven hits, relative to compound **2**, are summarised in Table 1 ("substrate screen" columns).

Given uPA's substrate preferences and the intentional inclusion of a substantial number of basic compounds in the library, it is not surprising that most other identified hits contain a basic functionality, most likely accommodated in the acidic S1 pocket of uPA. Additionally, the distance between this basic functionality and the acyl-AMC group in each substrate roughly equals the corresponding distance between the guanidine group and the scissile amide bond in a typical P1(Arg)-containing peptide substrate of the enzyme. The lipophilic aryl derivatives **6** and **17** are the only non-basic hits in the series, with **6** being processed by uPA with similar efficiency to **2**. Notably, even a small deviation from a hit compound's structure was observed to cause a total loss of substrate properties. This is illustrated by the cleavage efficiency pattern within the series **1**–**3**, **4**–**6**, **9** and **10**, and indicates that the robustness of SAS as a method to identify useful fragments for inhibitor discovery is not optimal. It is indeed highly conceivable that "unbiased" SAS libraries will overlook potentially interesting fragments if either 1) the linker distance between the fragment and the acyl-AMC functionality or 2) the fragment substitution pattern does not precisely fit the requirements for stabilisation of the transition state of substrate conversion. In our opinion, these findings demonstrate that unmodified application of the SAS protocol can lead to loss of relevant information and hence "false negatives". In addition, predictive application of the obtained processing data by construction of structure–cleavage efficiency relationships (analogous to structure–activity relationships in traditional inhibitor discovery) seems compromised by the use of a readout system that is so sensitive to minute structural changes. Furthermore, we observed that the SAS protocol requires substantial amounts of target protein (≈2.5 μg per well), together with long screening times. Both experimental parameters were used according to Ellman's reports and were found to be crucial for detection of slowly degraded library members. It is also worth highlighting the critical importance of enzyme purity. In a separate screen of our library with commercial uPA obtained from human urine, a series of additional hits characterised by very high turnover efficiencies was obtained. These compounds, however, were not cleaved to any extent by the recombinant enzyme (Scheme 2). During further investigations of this overt discrepancy between the two uPA preparations, we were able to show that the processing of those compounds could not be inhibited by addition of a nanomolar uPA inhibitor that we had reported earlier (UAMC-00122).[22] Application of chromatographic and gel electrophoretic techniques were not helpful for identifying the catalyst responsible for cleavage in the human uPA preparation. Nonetheless, these results also indi-




**Scheme 2.** Compounds that were processed by a commercial human uPA preparation obtained from urine, but not by recombinant human uPA.

**Table 1.** Hits obtained after screening of a 137-compound library of *N*-acyl AMCs).

| Cpd | Structure R= | Substrate screen cleavage efficiency[a] [%] | Inhibitor screen [I] [μM] | inhibition [%][b] | Cpd | Structure R= | Substrate screen cleavage efficiency[a] [%] | Inhibitor screen [I] [μM] | inhibition [%][b] |
|---|---|---|---|---|---|---|---|---|---|
| **1** |  | –[c] | 100 | 63.4 | **2** |  | 100 | 100 | 50 |
| **3** |  | – | 250 | 58 | **4** |  | – | 50 | 25.7 |
| **5** |  | – | 100 | 33.6 | **6** |  | 98.7 | 400 | 50 |
| **7** |  | – | 50 | 25 | **8** |  | 72.3 | 500 | 50 |
| **9** |  | – | 500 | 64.3 | **10** |  | 65 | 500 | 45 |
| **11** |  | 34.5 | 500 | 30.4 | **12** |  | 30 | 500 | 39 |
| **13** |  | 20.5 | 500 | 28 | **14** |  | 14 | 500 | 20 |
| **15** |  | 9 | 500 | 17 | **16** |  | – | 100 | 27.5 |
| **17** |  | 8.2 | 100 | 27 | **18** |  | – | 50 | 23 |
| **19** |  | – | 100 | 31 | **20** |  | – | 100 | 32.6 |
| **21** |  | – | 50 | 21 | **22** |  | – | 250 | 27.3 |
| **23** |  | – | 250 | 20.3 | **24** |  | 3.2 | 500 | 23 |
| **25** |  | – | 500 | 21.5 | **26** |  | – | 100 | 27 |
| **27** |  | – | 100 | 24 | **28** |  | – | 100 | 19 |

[a] Cleavage efficiency is defined as the cleavage rate of a compound relative to the "best" substrate in the library (compound **2**). [b] Inhibition is defined as the % decrease in the processing rate of reference uPA substrate *pyro*-Glu-Gly-Arg-pNA. [c] "–" indicates that no uPA-mediated cleavage of the compound was observed.

cate that the published SAS protocol is susceptible to possible occurrence of false positives, with the presence of other catalytically active species (e.g., other enzymes occurring as impurities) being responsible.

Inspired by these hitherto unaddressed but relevant findings, we devised a fundamentally different experimental setup for library evaluation. In this approach the inhibitory properties of the library members are investigated, rather than their substrate properties. A protocol strongly related to the archetypical assay normally used for enzyme inhibitor evaluation was elaborated. Here, the library members' potential to inhibit degradation of a known, peptide-derived chromogenic substrate of the target is evaluated. We hypothesised that this setup should be able to uncover all fragments with affinity for uPA, and not only those characterised by an ideal linker distance or an optimal substitution pattern. Additionally, selection of the efficiently processed chromogenic *pyro*-Glu-Gly-Arg-pNA ($K_M$ = 80 μM) for this assay allowed the enzyme concentration to be lowered tenfold relative to the SAS protocol.[25] The application of a single, kinetically well-characterised substrate in our opinion also avoids the possibility of "false positive" results and removes the need to verify whether the library members are processed by the actual target enzyme or by another catalytically active species occurring in the enzyme preparation.

In this experiment, with use of 20 nM uPA, the 137 compounds of the library were screened again at 50–500 μM, with concentrations depending upon compound solubility as before. The readout consisted of the evaluation of uPA-mediated *para*-nitroaniline release from the chromogenic substrate *pyro*-Glu-Gly-Arg-pNA, at 100 μM concentration. The results, expressed as percentage inhibition of *pyro*-Glu-Gly-Arg-pNA cleavage at a given compound concentration, are also summarised in Table 1 ("inhibitor screen" columns). In general, it deserves mentioning that the affinities displayed by the "hits" are well within the range that is generally reported for fragments (high micromolar). Furthermore, our alternative approach identifies all eleven "hits" initially revealed by the traditional SAS protocol. The relative affinities observed for these eleven compounds in the inhibition experiment (extrapolated from their inhibitory potencies) roughly reflect the cleavage efficiencies of the compounds.

Most interestingly, though, our modified screening procedure also discloses an additional 17 molecules that inhibit the release of *para*-nitroaniline. Inspection of the compounds with inhibitory properties immediately provides a more coherent image of structural classes that possess potential for uPA inhibitor discovery within the library. As an example, all guanidinophenyl (**1**–**3**) and guanidinoalkyl (**9**, **10**) homologues present in the library were identified as inhibitors, whereas SAS had only selected one of either class. Analogously, all chlorophenyl (**4**–**6**) and closely related lipophilic phenyl derivatives (**7**, **16**–**19**, **21** and **22**) that were present in the library turned up as potentially valuable constituents for new uPA inhibitors. Again, from the results of the SAS approach, one would conclude that **6** and **17** are two singletons of interest within the library, whereas they instead belong to a group of closely related structures that could all be valuable for uPA inhibitor design. Furthermore, our alternative method avoids false positives due to other catalytically active species present in the enzyme preparation. False positive results caused by compound-induced aggregation or denaturation of the enzyme were also not observed during the inhibitor screen.

These findings confirm that, to identify useful fragments in a given SAS library, it could be more efficient to evaluate the inhibitory properties of the library members rather than their substrate characteristics. More emphasis can in this way be placed on creating a structurally diverse library because the need for a number of homologues or close analogues around each structural feature is reduced. A point-by-point comparison of the two screening modes is given in Table 2.

**Table 2.** Comparison of SAS and inhibitor screening protocols.

| Parameter | SAS | Inhibitor screen |
|---|---|---|
| enzyme amount per well | 2.5 μg (400 IU) | 0.25 μg (40 IU) |
| screening times | 6 h[a] | 10 min |
| false positives[b] | yes[b] | not observed |
| false negatives | yes[c] | not observed |

[a] Long screening times can cause errors due to, for example, enzyme denaturation, autoproteolysis or buffer evaporation.[18] [b] Due to other catalytically active species present in the enzyme preparation. [c] Resulting in incomplete SAR data.

When using a library of AMC amides for fragment identification, one might nonetheless speculate on the possibility that the AMC moiety might interfere during the process of target binding. Although peptidyl-AMC amides have been used successfully for decades in inhibitor discovery, this is not completely inconceivable in, for example, hypothetical cases in which the AMC ring is not accommodated in the S1′ region of the enzyme. Theoretically, such interference could consist either of 1) a net impeditive effect on fragment binding (e.g., by steric hindrance) or, alternatively, 2) a net supportive effect (e.g., if the AMC ring were to contribute to affinity). We expect that interference of the first type would surface in the form of incoherent SAR data, involving "outliers" that unexpectedly do not show target affinity. The identification of all positive controls in our library and the near complete coverage of compounds that belong to inhibiting structural subgroups of the library do indicate that, in general, the AMC ring is not significantly hampering fragment binding. Similarly, the obtained results also do not suggest that the AMC ring contributes significantly to affinity, because most of the library members evaluated were devoid of measurable uPA affinity. Taking all these considerations into account, we expect the AMC portion of the library members not to interfere significantly with the inhibitory properties of the fragments. Equally illustrative of this is the fact that the isolated guanidinobenzene fragment **38** (vide infra) has an affinity broadly comparable to those of AMC-linked fragments **1**–**3**. So far we have also not observed indications of interference in a number of other ongoing projects dealing with inhibitor design for caspases and autophagins during which the same library was screened for inhibiting fragments.

It also deserves mentioning that the highest affinities observed within the structural subgroups with inhibitory properties do not necessarily belong to the compounds that are also substrates. This is seen, for example, on comparing the slightly higher inhibitory potency of guanidinophenyl derivative **1** with that of homologue **2**, of which only the latter is processed as a substrate. We therefore investigated whether apparently lower affinities of library members displaying substrate properties might be accounted for by their gradual consumption during the inhibition experiment. Quantification of cleavage during the course of an inhibition experiment was determined by monitoring the release of AMC. The process was, however, found to be too slow to interfere significantly with determination of inhibitory potency. This is read out after 10 min, typically during the linear phase of peptide substrate consumption (Figure 1). In hypothetical cases in which library members with exceptionally high $k_{cat}/K_M$ values relative to the peptide substrate used might be present, however, such an effect could not be excluded. In addition, therefore, two control experiments were also carried out for each inhibitory library member to ascertain whether or not the identified inhibitor series could include non-competitive, allosteric or irreversible compounds. Again, the SAS protocol would not allow identification of such fragments, although they could certainly be of interest to specific inhibitor discovery programs. For all compounds (**1**–**28**), inhibition was found to decrease with increasing concentration of the chromogenic substrate, thus indicating competition for the enzyme's active site. The percentage inhibition increased with inhibitor concentration and did not change significantly with longer inhibitor preincubation times, as would be the case for slowly and irreversibly binding compounds.

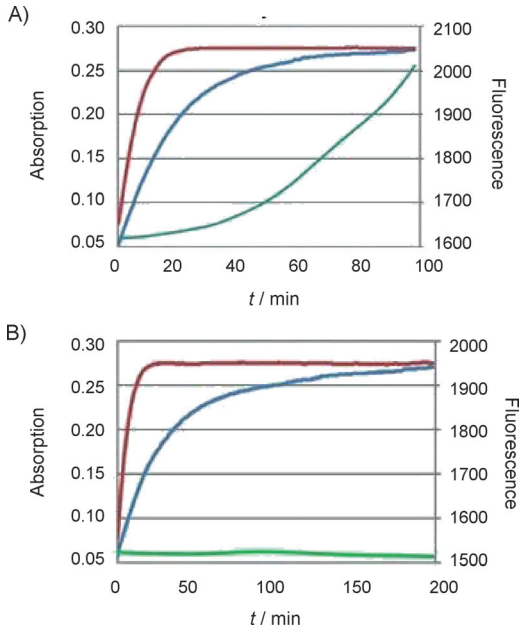
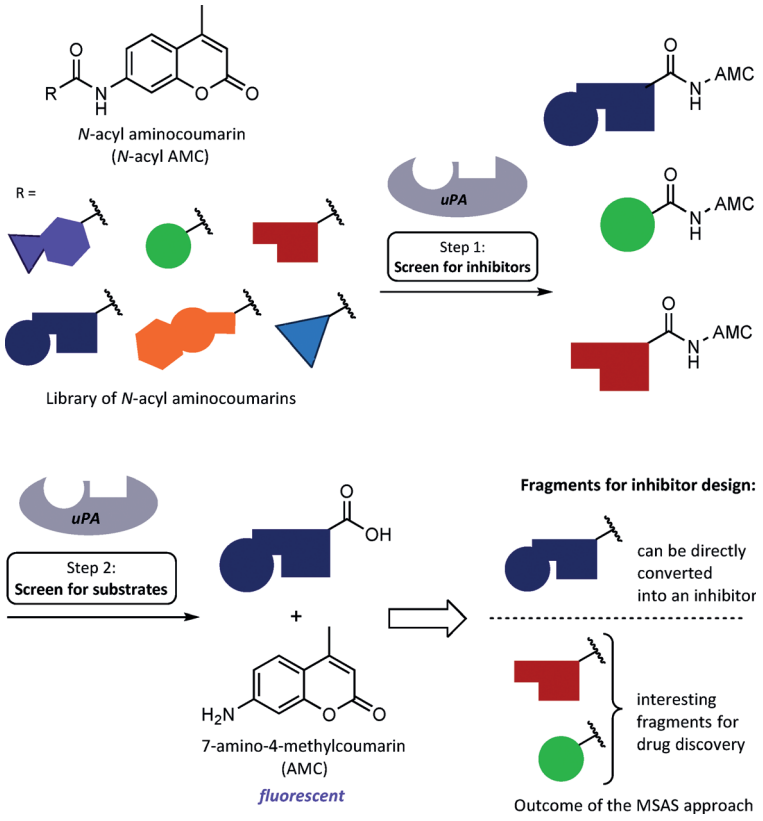



**Figure 1.** Inhibitory profiles of two selected structurally related MSAS hits. A) Inhibitory profile of **2**, a library member with substrate properties. At $t = 10$ min, consumption of **2** is minimal, and read-out of inhibitory potency is not significantly influenced. B) Inhibitory profile of **3**, a library member that is not a substrate. ——: uPA + *pyro*-Glu-Gly-Arg-pNA (control): release of pNA; ——: uPA + *pyro*-Glu-Gly-Arg-pNA + **2**/**3**: release of pNA; ——: uPA + *pyro*-Glu-Gly-Arg-pNA + **2**/**3**: release of AMC.

On the basis of all these results, we propose a modified experimental strategy that combines optimal efficiency and maximum extraction of useful structural information during screening of SAS libraries. We have provisionally called this approach MSAS. Its key steps are represented in Scheme 3. In MSAS, the library is first screened for inhibitory fragments (Scheme 3, step 1). This experimental layer will provide SAR data for the interesting fragment types present within the library. As demonstrated, the hits obtained in the inhibition experiment will also include library members with substrate properties. Because step 1 runs with higher time- and cost-efficiency than a traditional SAS experiment, we propose to perform SAS during the second phase of MSAS and only for identifying the substrates within the set of hits identified during the inhibitor screening experiment (Scheme 3, step 2). Furthermore, it is important to stipulate that MSAS's experimental setup, like that of the parent methodology, is not limited to protease inhibitor discovery. The same strategy can directly be applied to any other type of enzyme target studied, provided that the members of the screened library each contain a suitable enzyme-processable functionality.

The affinity data obtained after step 1 and the substrate property data obtained after step 2 can theoretically be validated in several orthogonal approaches for transforming fragments into inhibitors. To investigate and validate the results of the MSAS setup further, representative examples of such approaches are preliminarily explored in the following part. Library member **2**, possessing both significant uPA affinity and substrate properties, was selected as the common starting point. The 4-guanidinophenethyl portion of this compound had already been reported earlier by us as a constituent of nanomolar and highly selective peptide-derived diaryl phosphonate inhibitors of uPA.[22] It was therefore included in the "biased" portion of the library. Its status as a positive control element also justified its selection for the experiments dealing with translating fragments into inhibitors.

For the validation of affinity data produced during step 1, we present a generally applicable strategy—not previously reported, to the best of our knowledge—for obtaining scaffold-based inhibitors. With specific regard to protease targets, such scaffold-based compounds have the potential to circumvent several of the inherent liabilities of classical, peptide-based protease inhibitors [mainly related to ADME (absorption, distribution, metabolism and excretion)]. Also, on a much more general level, many of the recently approved small-molecule drugs share an overall comparable architecture consisting of a central scaffold decorated with several substituents that confer additional affinity for the biomolecules they target. A generalised routine for obtaining such compounds is proposed. Firstly, a fragment identified during the first step of MSAS is chemically grafted onto a limited set of different, drug-like scaffolds (Scheme 4). If a compound with uPA affinity significantly higher than that of the original fragment is found within this small set of monosubstituted scaffolds, that compound is selected for further optimisation. During the optimi-

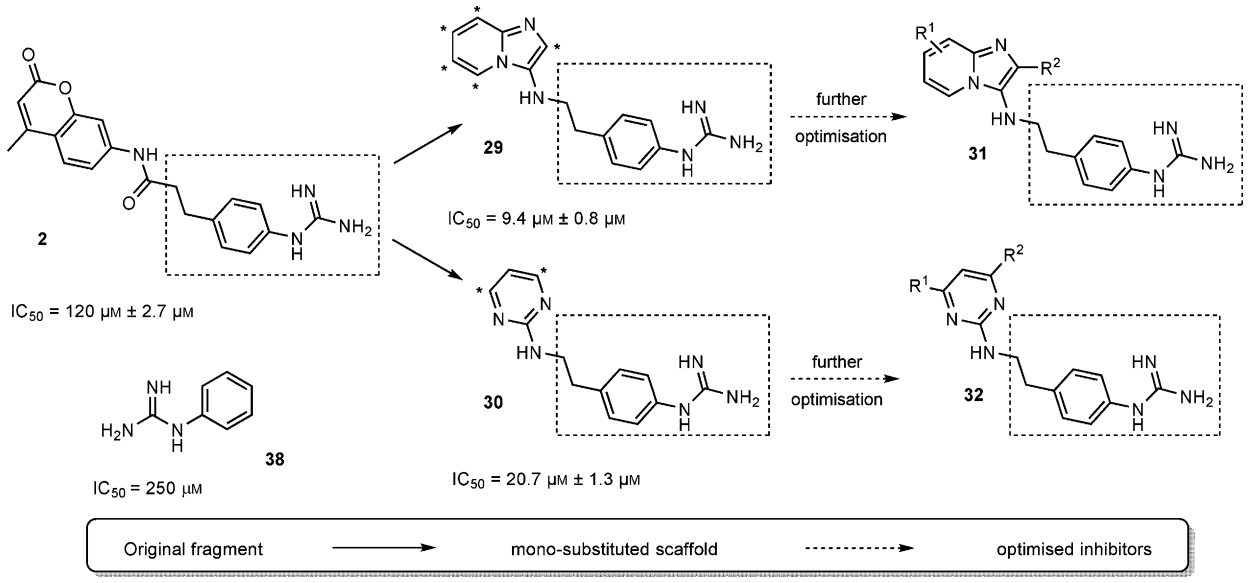



**Scheme 3.** Outline of the modified substrate activity screening approach, demonstrated for a protease target.

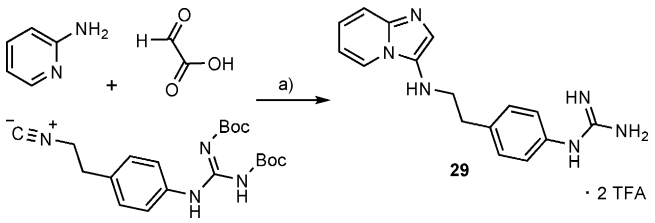
sation, one to several additional substituents are introduced on the selected monosubstituted scaffold to increase target affinity further. For maximum efficiency, it is advisable to select scaffold types onto which several additional substituents can readily be introduced, preferentially in a regioselective fashion and by combinatorial chemistry techniques.

To elaborate this concept, we chose the imidazopyridine and pyrimidine scaffolds, already present in, for example, the hypnotic drug zolpidem and the HIV-RT inhibitor rilpivirine.[26,27] A wealth of efficient chemical decoration strategies that allow efficient production of diversely substituted analogues exist in both cases.[28,29] Firstly, the guanidinophenethyl fragment was attached through an amine linker (Scheme 4). The two scaffolded inhibitor fragments **29** (Scheme 5) and **30** were then evaluated; they displayed roughly comparable affinities for uPA (9.4 ± 0.8 μM and 20.7 ± 1.3 μM, respectively). This corresponds to a relative increase in affinity of about one order of magnitude relative to library member **2** and indicates that both the imidazopyridine and pyrimidine scaffolds could be used for construction of scaffold-based uPA inhibitors. Hypothetically, additional steps necessary for obtaining molecules with further optimised affinity would consist of introducing one or several additional substituents.

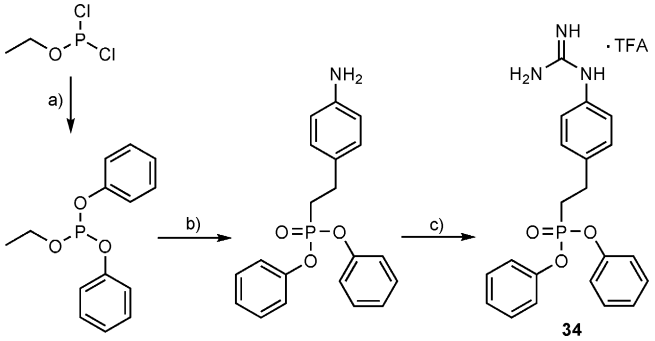
As pointed out earlier, attention was then devoted to the original data validation strategy of the SAS approach. To this end, we grafted the potentially irreversibly binding diphenyl phosphonate warhead onto the 4-guanidinophenethyl moiety (Scheme 6) and also, as a control, onto the homologous 4-guanidinophenylmethyl residue of **1** (Table 3). The latter fragment, although still possessing uPA affinity, was not processed as a substrate in the corresponding assay. As suggested by Ellman et al., this would translate either into an inhibitor with significantly reduced potency, or at least into a compound




**Scheme 4.** Introduction of the 4-guanidinophenethyl fragment onto an imidazopyridine and a pyrimidine scaffold to afford **29** and **30**, respectively. Scaffold positions that—from a synthetic point of view—allow easy substituent introduction for further optimisation are marked with asterisks; $R^1$, $R^2$ = additional substituents.

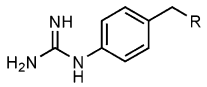


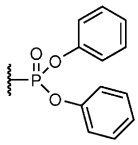
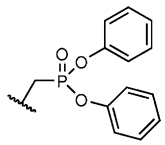
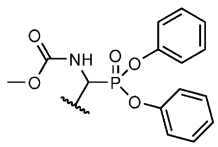
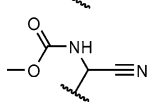

**Scheme 5.** Synthesis of scaffold-based inhibitor **29** (for synthesis and characterisation see the Supporting Information). Reagents: a) i: MeOH, $HClO_4$ (cat.), RT; ii: TFA/$CH_2Cl_2$ 1:1, RT.




**Scheme 6.** Synthesis of diaryl phosphonate inhibitor **34** (for synthesis and characterisation see Supporting Information). Reagents: a) phenol, $Et_3N$, toluene; b) i: 1-(2-bromoethyl)-4-nitrobenzene, pressure tube, 120 °C; ii): zinc dust, THF, 0 °C; c) i: *N*,*N*′-di-Boc-guanylpyrazole, $Et_3N$, THF, RT; ii: TFA/$CH_2Cl_2$ 1:1, RT.

**Table 3.** Comparison of $IC_{50}$ values for uPA inhibition of different benzylguanidine-containing compounds derived from **1** and **2**.



| Cpd | R= | $IC_{50}$ (uPA) [μM] | Inhibition type |
|---|---|---|---|
| **33** |  | 1.39 ± 0.06 | irreversible[a] |
| **34** |  | 0.0097 ± 0.0003 | irreversible[a] |
| **35**[b] |  | 0.0031 ± 0.0005 | irreversible[a] |
| **36** |  | 250 | reversible |
| **37** |  | 5.0 ± 0.01 | reversible |

[a] Validated experimentally by dilution assay.[25] [b] Previously reported as UAMC-00150.[22]

with compromised capability of mechanism-based enzyme blocking.[12] Compounds **33** and **34** indeed displayed an approximately 1000-fold difference in potency (Table 3). Furthermore, both compounds were subsequently demonstrated to be irreversible inhibitors. This typically implies that the inhibitory potencies of **33** and **34** should not be interpreted as a strict indication of intrinsic affinities but rather as a measure for their respective second-order rate constants of the irreversible step in enzyme inactivation. In any case, these results agree well with SAS's tenets, and on a broader level, with the fundamental assumptions of more canonical approaches in substrate-based drug discovery.

It is also remarkable that the potency displayed by compound **34** broadly compares with that of diaryl phosphonate **35** (UAMC-00150), reported earlier by us as one of the best representatives of known uPA inhibitors.[22] This reference inhibitor contains an additional methoxycarbamoyl group mimicking the P2–P1 amide bond of uPA's peptide substrates. On the basis of this higher degree of similarity with compound types that are naturally processed by uPA, one could reasonably assume that both the kinetic and the thermodynamic parameters of target binding could be favourable in the case of **35**, thus resulting in a higher net potency. However, the absence of the methoxycarbamoyl fragment in **34** does not seem to interfere significantly with the overall process of target recognition and irreversible covalent bond formation.

Additionally, compounds **36** and **37**, containing identical methoxycarbamoyl substituents, were evaluated as potential uPA inhibitors. The first, compound **36**, lacks a warhead functionality and served mainly to assess the contribution of the reactive functionality to inhibitory potencies, as discussed earlier. Compound **36** has the same $IC_{50}$ value as phenylguanidine fragment **38** ($IC_{50}$ = 250 μM); this indicates that the potencies observed for **33**–**35** are mainly driven by an efficiently occurring irreversible step after initial binding of the inhibitor to uPA.[30]

With this information at hand, compound **37** was evaluated as a second test case for the SAS protocol. This molecule contains a carbonitrile group, a potentially reversible, covalent warhead type very often used in inhibitors of serine proteases. Its low-micromolar affinity indicates that the nitrile group in this molecule might not be suitably oriented to allow covalent bond formation with uPA in a low-energy inhibitor conformation. This result, in our opinion, does not raise any critical doubts as to the validity of the SAS protocol. It nonetheless warns that for transformation of SAS "hits" into inhibitors, evaluation of several warheads might be mandatory in order to identify a type that performs well with the selected protease and inhibitor.

## Conclusion

In conclusion, we have investigated SAS for the discovery of inhibitors of oncology target urokinase (uPA). Although our results were supportive of the fundamental hypotheses formulated earlier for SAS, we also encountered a number of hitherto unreported limitations of the approach. In response, we pro-

pose a simple, efficient modified methodology: "MSAS" (modified substrate activity screening). This methodology not only circumvents limitations of the parent approach, but also broadens its scope by providing additional fragments and more coherent SAR data. As well as introducing MSAS as a generally applicable method for enzyme inhibitor discovery, this study has expanded existing SAR knowledge on S1-pocket-binding fragments of uPA. In addition, hitherto unreported uPA inhibitor scaffolds are presented and have been used to obtain new reversible and irreversible compounds.

## Experimental Section

Reagents were obtained from Sigma–Aldrich or from Acros, Fluorochem or Apollo Scientific and were used without further purification. Synthesised compounds were characterised by $^{1}$H NMR, $^{13}$C NMR and mass spectrometry. $^{1}$H NMR and $^{13}$C NMR spectra were recorded with a 400 MHz Bruker Avance DRX 400 spectrometer, and analysed by use of MestReNova analytical chemistry software. ES mass spectra were obtained with an Esquire 3000plus ion-trap mass spectrometer from Bruker Daltonics. Purities were determined with two diverse HPLC systems based either on mass detection or on UV detection. A Waters acquity UPLC system coupled to a Waters TQD ESI mass spectrometer and a Waters TUV detector was used. Where necessary, flash purification was performed with a Biotage ISOLERA One flash system equipped with an internal variable dual-wavelength diode array detector (200–400 nm).

**Library of *N*-acyl aminocoumarins**: A library of 137 *N*-acyl aminocoumarins was prepared. A large part of the library was synthesised by means of single coupling reactions of the constituting moieties in the presence of a mild acyl chlorinating agent—namely tetramethyl-$\alpha$-chloroenamine ("Ghosez's reagent")[23,24]—as a coupling mediator. Further details on the library synthesis and the chemical characterisation of the obtained compounds can be found in the Supporting Information.

**Synthesis of scaffolded reversible inhibitors of uPA**: The scaffold-based inhibitors of uPA discussed in this report contain the imidazopyridine and pyrimidine scaffold types. The imidazopyridine-scaffold-containing inhibitor **29** was prepared by a general protocol for the Groebke–Blackburn–Bienaymé reaction for the synthesis of fused 3-aminoimidazoles.[31,32] The pyrimidine-scaffold-containing compound **30** was prepared by a standard nucleophilic aromatic substitution reaction protocol from previously prepared starting material. More detailed synthetic procedures and chemical characterisation of the structures can be found in the Supporting Information.

**Synthesis of diaryl phosphonate irreversible inhibitors of uPA**: Diaryl phosphonate inhibitors **33** and **34** were prepared by a general protocol for base-promoted alkylation of H-phosphonates (the Michaelis–Becker reaction) and by the modified version of the classical Arbuzov reaction protocol, respectively. More detailed synthetic procedures and the chemical characterisation can be found in the Supporting Information.

**General procedures for biochemical assays**: Enzymatic assays were performed with use of BioTek Microplate Reader (Synergy MX). Data collection and analysis were performed with Gen5 Microplate Software and Microsoft Excel. Human urokinase-type plasminogen activator (uPA) was obtained from Nodia. A fluorogenic substrate screen against uPA was performed with recombinant human uPA (R&D Systems). Inhibitor kinetic assays were carried out with urokinase chromogenic substrate BIOPHEN CS-61(44) (*pyro*-Glu-Gly-Arg-pNA) purchased from Nodia ($K_M = 80$ μM).[25] A HEPES buffer (Sigma–Aldrich, pH 8.2, 50 mM) was used. All enzymatic activity measurements were routinely performed in duplicate. *N*-Acyl AMC stock solutions and inhibitor stock solutions (10 mM) were prepared in DMSO and stored at −20 °C. Enzymatic assays contained not more than 5% (*v*/*v*) of DMSO.

**Inhibitor kinetic assays**: Enzymatic activity was measured over 5 min at 37 °C in the presence of urokinase chromogenic substrate BIOPHEN CS-61(44). Absorbance was monitored at $\lambda = 405$ nm. The assay mixture contained the *N*-acyl AMC compound (50–500 μM depending on the solubility), the non-recombinant urokinase solution in buffer (ca. 20 nM) and substrate BIOPHEN CS-61(44) ($K_M =$ 80 μM, 100 μM) in a final volume of 200 μL. The concentration of the chromogenic substrate used (100 μM) allowed for a sufficiently high initial substrate processing rate, while limiting competition between substrate and inhibitor.

**Fluorogenic substrate screen against uPA**: Screening of the library of *N*-acyl AMCs for substrates of uPA was performed over 6 h at 37 °C. The excitation wavelength was 383 nm, and the emission wavelength was 455 nm. Because of false positives appearing in assays based on a non-recombinant enzyme, the substrate screening was performed with a recombinant human uPA. Initial screening of the library was performed at the highest substrate concentration possible (for most of the library members 100–500 μM), uPA concentration was around 200 nM. Final screening was performed with subsaturating levels of the substrate.[11] Under these conditions cleavage is assumed to be a first-order process; hence $k_{cat}/K_M$ values are compliant with the relationship $S_t/S_0 = e^{-k_{obs}t}$, where $S_t =$ concentration of the substrate remaining at time $t$, $S_0 =$ initial substrate concentration, and $k_{obs} = k_{cat}[\text{enzyme}]/K_M$.[33] Final substrate screen and ranging hits based on the enzymatic cleavage efficiency were determined at 100 μM substrate and approximately 200 nM uPA concentration. Relative fluorescence units (RFUs) were measured for each substrate at regular intervals over a 6 h period of time with and without enzyme (blank). Blank was subtracted from the enzymatic activity measurements. The slope of the plotted line gave the relative $k_{cat}/K_M$ value for each substrate.[25] More information can be found in the Supporting Information.

**Determination of $IC_{50}$ values**: Enzymatic activity was measured at 37 °C with urokinase chromogenic substrate BIOPHEN CS-61(44). Absorbance was monitored at $\lambda = 405$ nm. Each reaction mixture had a volume of 200 μL and contained the chromogenic substrate (250 μM), the non-recombinant enzyme solution (ca. 20 nM) in buffer (145 μL) and the inhibitor (5 μL). An initial screening at three inhibitor concentrations (250 μM, 2.5 μM and 25 nM) was performed to estimate the range of the $IC_{50}$ value. For exact $IC_{50}$ determination, at least four inhibitor concentrations above and four concentrations below the estimated $IC_{50}$ value were used. $IC_{50}$ values were determined by fitting the obtained data with a four-parameter logistics equation with the aid of GraFit7. More information can be found in the Supporting Information.

**Determination of inhibition type**: To follow dissociation of the inhibitor·enzyme complex, aliquots of enzyme were incubated at 37 °C 1) without and 2) with the inhibitor, at a concentration 50 times higher than its $IC_{50}$. Enzyme was used at 2.5 times higher concentration than for the $IC_{50}$ determination. After 15 min, the aliquots were diluted 50-fold with the substrate (250 μM) solution in assay buffer. Dissociation of the enzyme·inhibitor complex was determined spectrophotometrically by monitoring hydrolysis of the chromogenic substrate over time.[25]

## Acknowledgements

*Rafaela Gladysz is grateful to the IWT–Vlaanderen for a scholarship. This work received support from the Fund for Scientific Research–Flanders (FWO–Flanders).*

**Keywords:** drug design · inhibitors · MSAS · SAS · urokinase

[1] G. Bollag, P. Hirth, J. Tsai, J. Zhang, P. N. Ibrahim, H. Cho, W. Spevak, C. Zhang, Y. Zhang, G. Habets, E. A. Burton, B. Wong, G. Tsang, B. L. West, B. Powell, R. Shellooe, A. Marimuthu, H. Nguyen, K. Y. J. Zhang, D. R. Artis, J. Schlessinger, F. Su, B. Higgins, R. Iyer, K. D'Andrea, A. Koehler, M. Stumm, P. S. Lin, R. J. Lee, J. Grippo, I. Puzanov, K. B. Kim, A. Ribas, G. A. McArthur, J. A. Sosman, P. B. Chapman, K. T. Flaherty, X. Xu, K. L. Nathanson, K. Nolop, *Nature* **2010**, *467*, 596 – 599.

[2] M. Congreve, G. Chessari, D. Tisi, A. J. Woodhead, *J. Med. Chem.* **2008**, *51*, 3661 – 3680.

[3] M. Congreve, R. Carr, C. Murray, H. Jhoti, *Drug Discovery Today* **2003**, *8*, 876 – 877.

[4] a) C. W. Murray, D. C. Rees, *Nat. Chem.* **2009**, *1*, 187 – 192; b) P. J. Hajduk, J. Greer, *Nat. Rev. Drug Discovery* **2007**, *6*, 211 – 219.

[5] C. W. Murray, M. L. Verdonk, D. C. Rees, *Trends Pharmacol. Sci.* **2012**, *33*, 224 – 232.

[6] D. C. Rees, M. Congreve, C. W. Murray, R. Carr, *Nat. Rev. Drug Discovery* **2004**, *3*, 660 – 672.

[7] C. W. Murray, T. L. Blundell, *Curr. Opin. Struct. Biol.* **2010**, *20*, 497 – 507.

[8] M. J. Hartshorn, C. W. Murray, A. Cleasby, M. Frederickson, I. J. Tickle, H. Jhoti, *J. Med. Chem.* **2005**, *48*, 403 – 413.

[9] a) W. Huber, F. Mueller, *Curr. Pharm. Des.* **2006**, *12*, 3999 – 4021; b) C. Lambruschini, M. Veronesi, E. Romeo, G. Garau, T. Bandiera, D. Piomelli, R. Scarpelli, C. Dalvit, *ChemBioChem* **2013**, *14*, 1611 – 16119; c) C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *50*, 259 – 261; d) G. Papeo, P. Giordano, M.-G. Brasca, F. Buzzo, D. Caronni, F. Ciprandi, N. Mongelli, M. Veronesi, A. Vulpetti, C. Dalvit, *J. Am. Chem. Soc.* **2007**, *129*, 5665 – 5672.

[10] T. Neumann, H. D. Junker, K. Schmidt, R. Sekul, *Curr. Top. Med. Chem.* **2007**, *7*, 1630 – 1642.

[11] W. J. L. Wood, A. W. Patterson, H. Tsuruoka, R. K. Jain, J. A. Ellman, *J. Am. Chem. Soc.* **2005**, *127*, 15521 – 15527.

[12] C. M. Salisbury, J. A. Ellman, *ChemBioChem* **2006**, *7*, 1034 – 1037.

[13] M. B. Soellner, K. A. Rawls, C. Grundner, T. Alber, J. A. Ellman, *J. Am. Chem. Soc.* **2007**, *129*, 9613 – 9615.

[14] M. Drag, M. Bogyo, J. A. Ellman, G. S. Salvesen, *J. Biol. Chem.* **2010**, *285*, 3310 – 3318.

[15] M. J. Leyva, F. Degiacomo, L. S. Kaltenbach, J. Holcomb, N. Zhang, J. Gafni, H. Park, D. C. Lo, G. S. Salvesen, L. M. Ellerby, J. A. Ellmann, *Chem. Biol.* **2010**, *17*, 1189 – 1200.

[16] T. D. Baguley, H.-C. Xu, M. Chatterjee, A. C. Nairn, P. J. Lombroso, J. A. Ellman, *J. Med. Chem.* **2013**, *56*, 7636 – 7650.

[17] J. Chapelat, F. Berst, A. L. Marzinzik, H. Moebitz, P. Drueckes, J. Trappe, D. Fabbro, D. Seebach, *Eur. J. Med. Chem.* **2012**, *57C*, 1 – 9.

[18] A. W. Patterson, W. J. L. Wood, J. A. Ellman, *Nat. Protoc.* **2007**, *2*, 424 – 433.

[19] a) T. Steinmetzer, *IDrugs* **2003**, *6*, 138 – 142; b) A. H. Mekkawy, D. L. Morris, M. H. Pourgholami, *Future Oncol.* **2009**, *5*, 1487 – 1500.

[20] a) K. Dass, A. Ahmad, A. S. Azmi, S. H. Sarkar, F. H. Sarkar, *Cancer Treat. Rev.* **2008**, *34*, 122 – 136; b) G. Abbenante, D. P. Fairlie, *Med. Chem.* **2005**, *1*, 71 – 104.

[21] a) R. Roodbeen, B. Paaske, L. Jiang, J. K. Jensen, A. Christensen, J. T. Nielsen, M. Huang, F. A. A. Mulder, N. Chr. Nielsen, P. A. Andreasen, K. J. Jensen, *ChemBioChem* **2013**, *14*, 2179 – 2188; b) S. Chen, D. Gfeller, S. A. Buth, O. Michielin, P. Leiman, P. G. Leiman, C. Heinis, *ChemBioChem* **2013**, *14*, 1316 – 1322.

[22] J. Joossens, O. M. Ali, I. El-Sayed, G. Surpateanu, P. Van der Veken, A.-M. Lambeir, B. Setyono-Han, J. A. Foekens, A. Schneider, *J. Med. Chem.* **2007**, *50*, 6638 – 6646.

[23] F. Munyemana, A.-M. Frisque-Hesbain, A. Devos, L. Ghosez, *Tetrahedron Lett.* **1989**, *30*, 3077 – 3080.

[24] C. A. G. N. Montalbetti, V. Falque, *Tetrahedron* **2005**, *61*, 10827 – 10852.

[25] J. Joossens, P. Van der Veken, G. Surpateanu, A.-M. Lambeir, I. El-Sayed, O. M. Ali, K. Augustyns, A. Haemers, *J. Med. Chem.* **2006**, *49*, 5785 – 5793.

[26] D. S. Johnson, J. J. Li, *The Art of Drug Synthesis*, Wiley, New York, **2007**.

[27] M. Sharma, L. D. Saravolatz, *J. Antimicrob. Chemother.* **2013**, *68*, 250 – 256.

[28] D. J. Brown, R. F. Evans, W. B. Cowden, M. D. Fenn, *Chemistry of Heterocyclic Compounds: The Pyrimidines*, Wiley, New York, **1994**.

[29] E. Ruijter, R. Scheffelaar, R. V. Orru, *Angew. Chem. Int. Ed.* **2011**, *50*, 6234 – 6246; *Angew. Chem.* **2011**, *123*, 6358 – 6371.

[30] J. C. Powers, J. L. Asgian, O. D. Ekici, K. E. James, *Chem. Rev.* **2002**, *102*, 4639 – 4750.

[31] H. A. Bienayme, *Angew. Chem. Int. Ed.* **1998**, *37*, 2234 – 2237; *Angew. Chem.* **1998**, *110*, 2349 – 2352.

[32] A. El Akkaoui, M.-A. Hiebel, A. Mouaddib, S. Berteina-Raboin, G. Guillaumet, *Tetrahedron* **2012**, *68*, 9131 – 9138.

[33] F. G. Gervais, *J. Biol. Chem.* **1998**, *273*, 17102 – 17108.

Received: April 24, 2014
Published online on August 25, 2014