HyCoSuL as a tool to dissect protease substrate specificities – engineering of highly selective substrates and probes

Marcin Poreba[1,2], Guy S. Salvesen[1], Marcin Drag[2]

[1]Sanford Burnham Prebys Medical Discovery Institute, 10901 North Torrey Pines Road, La Jolla, CA, USA, [2]Division of Bioorganic Chemistry, Faculty of Chemistry, Wroclaw University of Science and Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland

E-mail addresses: marcin.poreba@pwr.edu.pl, gsalvesen@sbpdiscovery.org, marcin.drag@pwr.edu.pl

Corresponding authors: marcin.poreba@pwr.edu.pl, marcin.drag@pwr.edu.pl

Phone: +48 71 320 4526, Fax: +48 71 320 2427

ABSTRACT

Protease substrate specificity profiling is a very active area of research. To date many biological- and chemical-based approaches have been developed in order to design new, highly active and selective protease substrates and probes. However, most of these techniques use only natural amino acids, which limits the ability to distinguish between proteases with overlapping substrates specificities. To overcome this limitation we developed novel chemical approach called HyCoSuL (Hybrid Combinatorial Substrate Library), which incorporates a large panel of unnatural amino acids coupled to a positional scanning protocol. Here, we describe the synthesis and screening of HyCoSuL toward human caspases and legumain. We also discuss possible modifications and adaptations of this approach, making it a useful tool for developing highly active and selective reagents for a wide variety of proteolytic enzymes. This entire protocol can be divided into three major parts: (1) HyCoSuL synthesis, (2) protease screening and (3) synthesis of an activity-based probe. The first part describes the combinatorial solid phase synthesis of fluorescent-labeled tetrapeptide substrates (HyCoSuL). In the second step this library is used to elucidate caspase and legumain preferences at P4-P2 positions followed by the synthesis and analysis of the most selective and/or most active sequences. The third part describes a detailed protocol for the synthesis of a biotin-labeled activity based probe with the selected peptide sequence. Beginning with the new library design, the entire protocol can be completed in 4-6 weeks (HyCoSuL synthesis: 3-5 weeks; HyCoSuL screening per enzyme: 2-4 days; and activity-based probe synthesis: 1-2 weeks).

Keywords: protease, substrate specificity, activity-based probe, HyCoSuL, fluorescent substrates, solid phase peptide synthesis

1. INTRODUCTION

1.1. Protease substrate specificity

Proteases are enzymes that catalyze peptide bond hydrolysis in peptides and proteins, thus are important regulators of many biological processes including cell death and cell cycle, blood coagulation, signal transduction, response to viral and parasitic infections and many others[1,2]. In the early days of proteases research these enzymes were considered to have nonspecific degradation functions, however nowadays it is well-established that most proteases are selective for those substrates complementary to their active site architecture[3]. The concept of substrate specificity, disseminated by Schechter and Berger in 1967[4], has proteases cleaving a target substrate between the P1 and P1` positions creating two new cleavage products: N-terminal Pn-…-P3-P2-P1-COOH and C terminal $H_2N$-P1`-P2`-P3`-…-Pn`. Understanding the detailed substrate specificity of proteases is very helpful for exploration of new putative protein substrates, and also is routinely used for designing small molecule substrates to monitor a particular protease (of group of proteases) in biological systems (for example in cells undergoing apoptosis)[5,6]. To date many biological- and chemical-based approaches have been proposed to determine substrate specificity of proteases (for reviews please see references[7-9]). Classical approaches such as phage displays or PS-SCL (Positional Scanning Synthetic Combinatorial Libraries) are the most popular strategies for profiling proteolytic enzymes[7]. Phage display (a biological approach) produces label-free peptide mixtures that are subjected to protease digestions and an optimal substrate sequence is selected after several cycles[10]. On the other hand, PS-SCL (a chemical approach) usually employs a tetrapeptide mixture labeled with a fluorophore that produces a strong fluorescence signal upon substrate cleavage[11,12]. Hundreds of proteases have been screened with these tools resulting in the discovery of potent and selective substrates for many proteases. However, biological diversity approaches are innately composed of only natural amino acids, and most chemical diversity approaches utilize natural amino acids during peptide synthesis, making it very difficult to obtain individual selectivity for closely related enzymes with overlapping substrate specificities (caspases, cathepsins, DUBs, etc.)[13-15]. These difficulties create a great opportunity to develop new strategies for profiling proteases in order to find substrates with better selectivity and higher activity.

1.2. Protease inhibitors and activity based probes

The objective of profiling protease specificity preferences is not limited for finding of new substrates. The optimal peptide sequence can be further converted into an inhibitor, by attaching a mechanism-based warhead, or into an activity-based probe (ABP) equipped with both a warhead and an easily detected tag (biotin, isotope, fluorophore, etc.). Both, inhibitors and ABPs are essential for monitoring protease activities in a wide scope of biological systems, from *in vitro* test tubes, through cell cultures to whole organisms[16-18].

1.3. Development of the protocol

We hypothesized that the chemical space in protease active sites can be explored and substantially expanded by the use of unnatural amino acids[19,20], thus we developed the HyCoSuL (Hybrid Combinatorial Substrate Library) approach (**Figure 1**). This chemical tool is built on a classical PS-SCL scaffold where P4-P2 positions in the peptide substrate are varied with natural amino acids[12,21]. In HyCoSuL the P4-P2 positions of a tetrapeptide utilize a wide set of unnatural amino acids (>100 derivatives)[18,22] conjugated with the ACC fluorescent tag, composed of three sub-libraries:

- P4 sub-library: Ac-Aaa-Mix-Mix-P1-ACC,
- P3 sub-library: Ac-Mix-Aaa-Mix-P1-ACC,
- P2 sub-library: Ac-Mix-Mix-Aaa-P1-ACC,

where Mix is an equimolar mixture of 19 natural amino acids (cysteine is omitted due to oxidation problems, Nle is used instead of Met; defined by Ostreh[23]), Aaa is one of the 19 natural or unnatural (over 100) amino acids fixed in the position of interest. In many protease families (such as serine protease and caspase clans) the primary specificity is defined by the S1 pocket that occupies the P1 residue, allowing us to fix P1 as a predefined amino acid. Once, the substrate is cleaved by a protease, the fluorophore is released and the increase of fluorescence signal is monitored over time. The use of HyCoSuL was reported for the first time by Kasperkiewicz et al. to design a highly active substrate and probe for human neutrophil elastase (P1 position in the library was occupied by Ala)[18]. Shortly after, Poreba et al. reported the application of HyCoSuL to discriminate between human apoptotic caspases (Asp at P1 position)[22]. These two primary studies using serine and cysteine proteases on HyCoSuL demonstrated that this approach is very useful in order to obtain highly active and specific substrates and substrate-derived activity based probes for important human proteases.

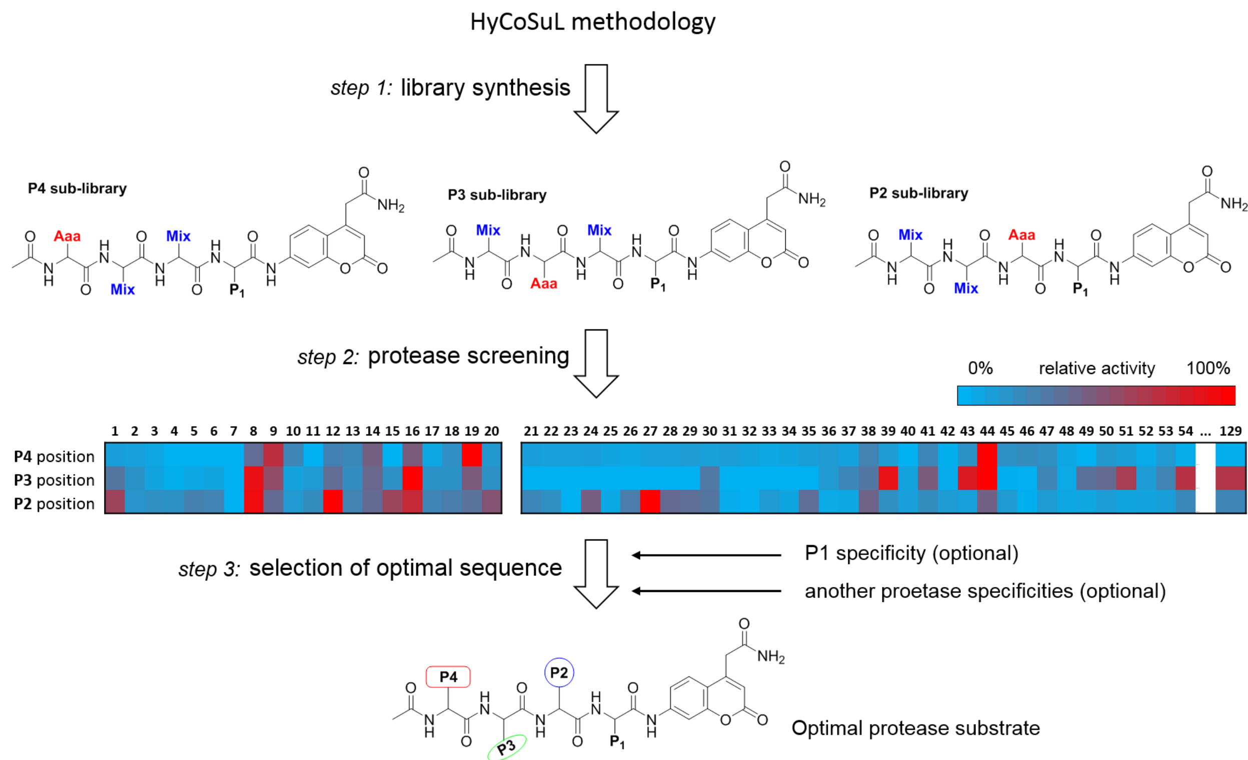


Figure 1 Outline of the Hybrid Combinatorial Substrate Library (HyCoSuL) method

Here we present the HyCoSuL screening protocol to obtain highly active and selective substrates of human caspases. The entire technique consists of several major steps: (1) solid phase synthesis of ACC-labeled hybrid combinatorial P4-P2 library with predefined Asp at P1 position, (2) screening and analysis of the caspase and legumain preferences at P4-P2 positions, (3) selection of optimal sequences (selective and active) for further enzyme investigation. The key feature unique to this protocol is the use of unnatural amino acids in the library structure to create a detailed map of the interactions between substrates and substrate binding pockets of caspases, legumain and other proteases. In this protocol we use the term "unnatural amino acids" for all amino acids except the 20 proteinogenic ones. This would include post-translationally modified and chemically synthesized amino acids, highlighting the context-specific differences between PS-SCL and HyCoSuL approaches.

1.4. Advantages and limitations of the protocol

Since HyCoSuL is a new concept in protease research, it is still being validated and improved in order to create a chemical model that covers structural requirements of multiple proteases. However, we have already found HyCoSuL to be versatile tool to profile proteases of different classes, catalytic mechanisms, origins, or activities[18,22,24-27]. The library structure allows for detailed exploration of the protease active site, thus the substrate specificity map is much more informative than the classical analysis based on PS-SCL, phage display or natural substrate cleavages. Furthermore, the possible number of individual peptides that can be synthesized and

validated is much higher than traditional approaches where only natural amino acids are used. For example the traditional PS-SCL can create 19x19x19 (P4xP3xP2 all natural - around 7000) individual structures that can be selected for further analysis, while HyCoSuL can produce 120x120x120 (natural + unnatural - around 1.7million) structures or more, depending on the number of unnatural amino acids used. Moreover, the selection of unnatural amino acids used for the synthesis is protease- and library-dependent. In our studies we selected a very broad range of diverse chemical structures to gain a detailed knowledge of protease interaction sites, however this selection can be more channeled. For example more *D*-amino acids can be selected to study some bacteria proteases, or only bulky and hydrophobic amino acids can be applied if the protease active site has strongly hydrophobic character. Unnatural amino acids can be also used to study posttranslational modifications in proteins that influence protease substrate specificity and activity. At present, there is a large number of commercially available Solid Phase Peptide Synthesis-suited amino acids that have various PTMs in their structures. The other advantage of our protocol is that the synthesis of peptide libraries of different types on solid phase (or in solution) is very well described in the literature. All the chemicals (including resins, coupling reagents, or natural and unnatural amino acids) are commercially available and affordable.

HyCoSuL (and PS-SCL) utilize positional substrate libraries, where some positions are fixed with one amino acid, while others are randomized. Each position (P4, P3, and P2) is screened separately, and the protease substrate specificity map is the sum of three sub-library analyses. Such methodology does not disclose protease potential subsite cooperativity, thus after determination of protease substrate specificity, several individual substrates are required to be synthesized in order to validate the screening results. In this protocol we address this issue in detail and provide the best solutions. Importantly, the method by its nature utilizes unnatural amino acids and therefore cannot be used to predict the location of cleavage sites in natural occurring proteins. On the other hand, the use of unnatural amino acids promotes the discovery of sites on proteases that natural amino acids are incapable of exploring, as demonstrated in the structure of a HyCoSuL-derived ABP in complex with neutrophil elastase[28].

1.5. Comparison with existing methods for protease screenings

The last 20-25 years yielded multiple methods for profiling protease substrate preferences[7]. The most prominent two of them are phage display developed by Smith[29] and adapted into protease research by Matthews and Wells[10] and PS-SCL developed by Rano, Thornberry and coworkers[11,21]. These methods allow for the precise determination of proteases preferences at the prime and non-prime region (phage display) or only at the non-prime region (PS-SCL). The great advantage of phage display is that up to 10^{10} individual peptides can be displayed and subjected to protease analysis, which would be very difficult to obtain through chemical synthesis. However, these peptides are label-free, and the additional steps to determine the optimal substrate sequences are needed, making this method labor-intensive. The protease analysis through PS-SCL is much faster and reliable, as peptides are equipped with a fluorescent

reporter tag. However, this method is suitable only for determination of non-prime region of protease catalytic cleft. Other methods for protease substrate specificity investigations include internally quenched fluorescent substrate libraries (to profile prime and non-prime regions)[30], microarrays that transform fluorescent substrate libraries into micro-scale formats[31], and multiple proteomic approaches using mass spectrometry for fishing out protease substrates from biological samples[32-34]. All these methods have their pros and cons, i.e. biological proteomics-based methods are more informative of natural targets, while phage display and related techniques, in common with chemical methods, are more comprehensive. Nevertheless, all these methods are based on natural amino acids. The HyCoSuL concept is based on the use of unnatural amino acids that allows for a more extensive exploration of the protease active site. Since this approach (similarly to PS-SCL) is of chemical origin, it provides a fast and reliable protease screening and can be easily adapted to study most proteolytic enzymes.

1.6. Applications and modifications of the HyCoSuL

Since this methodology consist of two main parts: (1) chemical synthesis of HyCoSuL and (2) its use for dissecting protease substrate specificity, its application is broad and depends on the context of the research conducted. Here we present several examples of HyCoSuL applications that we have successfully used in protease studies.

Application / case study 1 (human neutrophil elastase)

Human neutrophil elastase is a serine protease that is released by neutrophils during inflammation [35]. For many years the activity of this enzyme was measured by classic short substrates containing the Ala-Ala-Pro-Val (P4-P1) tetrapeptide sequence. Screening of elastase with HyCoSuL revealed that the chemical space of amino acids that occupy its active site can be substantially expanded to reveal exquisitely sensitive substrates, one of the best contained all unnatural amino acids Nle(*O*-Bzl)-Met(O_2)-Oic-Abu[18]. This unnatural substrate was almost 10,000-fold more sensitive than the classical natural epitope (Ac-AAPV-ACC k_{cat}/K_M=4.92x10^3 M^{-1}s^{-1}, and Ac-Nle(*O*-Bzl)-Met(O_2)-Oic-Abu-ACC k_{cat}/K_M=4.70x10^{7}M^{-1}s^{-1}). In the next step this substrate was converted into a biotin-labeled phosphonate activity-based probe that selectively detected active elastase in neutrophils during neutrophil extracellular trap formation[18]. Structural analysis of PK101 bound to elastase revealed that some of the side-chains explored pockets on the enzyme not available to natural amino acids, accounting for the enhanced activity and specificity[28]. This example demonstrates that an optimized peptide sequence from HyCoSuL profiling can be used as a scaffold to develop ultrasensitive probes for proteases detection.

Application / case study 2 (human apoptotic caspases)

Apoptotic caspases were first to be characterized with a PS-SCL approach by Thornberry and coworkers[11]. This analysis revealed that these enzymes have overlapping substrate specificity, and cannot be distinguished from each other by substrates consisting of natural amino acids[13]. The prototypic caspase substrates and inhibitors (DEVD for caspase-3, LEHD for caspase-

9, or IETD for caspase-8) are commonly used as “caspase specific” tools which has led to severe data misinterpretation in several studies[36]. Several broad studies have clearly demonstrated that these tetrapeptide-based sequences display a limited degree of selectivity, thus are not appropriate in dissecting individual caspase in complex mixtures [13,37,38]. Recently we profiled six human apoptotic caspases demonstrating that the use of over 100 unnatural amino acids can overcome the overlapping preferences of this group of enzymes[22]. Several substrates with unnatural amino acids were demonstrated to display much higher selectivity that commercially available structures. We validated their utility in a paradigm of cell-free apoptosis by demonstrating that one of the caspase-9 substrates (Ac-Oic-Tle-His-Asp-ACC) is hydrolyzed only by the initiator caspase-9, and not by executioner caspases-3, -6, and -7. This was the first example describing that caspase-9 activation can be monitored by a small molecule fluorescent substrate.

Application / case study 3 (ZIKA virus NS2B-NS3 protease)

Recently we used HyCoSuL to profile the P4-P2 positions of ZIKA virus NS2B-NS3 protease[25]. In order to do this we used a hybrid library with Arg at P1 position. We found that the best recognized amino acids by this protease are non-proteinogenic ornithine (Orn) at P2, lysine at P3 and unnatural *D*-Arg at P4. The screening results mirrored the k_{cat}/K_M for several substrates that were synthesized based on screening data (Ac-*D*-Arg-Lys-Orn-Arg-ACC substrate displayed the highest cleavage efficiency). In this study we explored the HyCoSuL concept by synthesizing a P1 library with over 100 unnatural amino acids. The structure of this library (Ac-Ala-Arg-Leu-P1-ACC) was not optimal for ZIKA virus NS2B-NS3 protease, as we wanted to make it useful also for P1 screening of many other proteases. Nevertheless, the P4-P1 profiling of NS2B-NS3 protease with unnatural amino acids allowed us to develop the first potent and irreversible activity-based probe for this enzyme.

These three examples illustrate the scope of the HyCoSuL as a broad and diverse discovery and development platform. On the one hand it can be used to design much more active substrates and probes for proteases of interest, and in another case it can be used to distinguish between closely related proteases. **However, HyCoSuL is not only a simple Ac-P4-P3-P2-P1-fluorophore library. We propose HyCoSuL as a general concept for the use of wide range of unnatural amino acids in an organized manner to dissect protease active site preferences.** The HyCoSuL architecture can be general and unified (as we presented in our case studies), but it can also be adapted to protease structural requirements, for example dipeptide (for diaminopeptidases), pentapeptide (for caspase-2), extended into prime region of the protease active sites (for example matrix metalloproteases). The choice of the unnatural amino acids in the library structure can also be channeled depending on the protease preferences. Moreover, the use of unnatural amino acids in proteases screening does not have to be limited to combinatorial peptide mixtures. HyCoSuL was published in 2014, however before that our group presented HyCoSuL-related strategies where individual substrate libraries with unnatural amino acids can be successfully applied in protease substrate screens (H_2N-P1-ACC for

aminopeptidases[39,40] or H_2N-P2-P1-ACC for dipeptidyl dipeptidases[41]). Recently we have also reported on the use of tripeptide libraries with unnatural amino acids tailored for ClpP proteases[42,43].

We have already demonstrated that HyCoSuL is a concept that opens doors for biochemical and biological experiments, providing for chemists who synthesizes substrate libraries of various size and length, biochemists who need new tools for the more accurate and rapid protease active site analysis, and finally biologists that use chemical tools (substrates, activity-based probes) for the protease detection *in cellulo*, or *in vivo*.

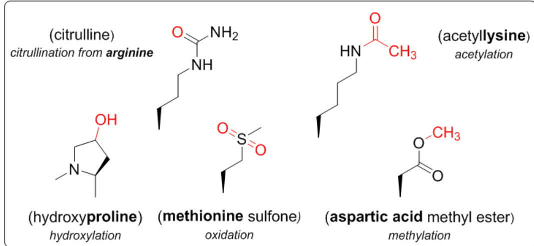
1.7. Experimental design

ACC fluorophore. The classical combinatorial fluorescence substrate libraries were equipped with AMC fluorescence tag[21]. In 2000 Ellman and Craik synthesized the first PS-SCL library with the ACC fluorophore[12]. The bifunctional characteristic of this molecule allows for the solid-phase synthesis of the entire library with virtually all amino acids at P1 position. In this protocol we use an ACC tag synthesized according to the protocol described by Maly[44]. However, in another *Nature Protocols* article Patterson et al. described the synthesis and use of another bifunctional fluorophore (AMCA, *N*-acyl 7-amino-4-methylcoumarin acetic acid), which can be used in HyCoSuL approach as well[45].

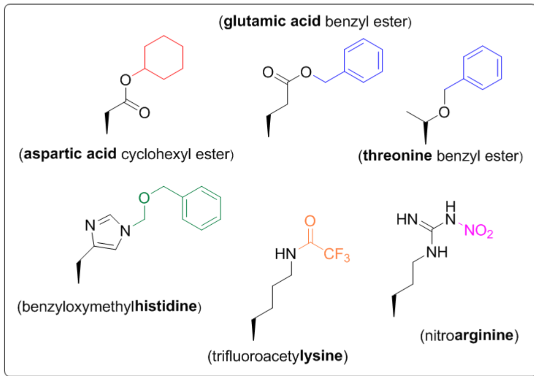
BOX 1. NATURAL and UNNATURAL AMINO ACIDS

1. Natural (proteinogenic) amino acids

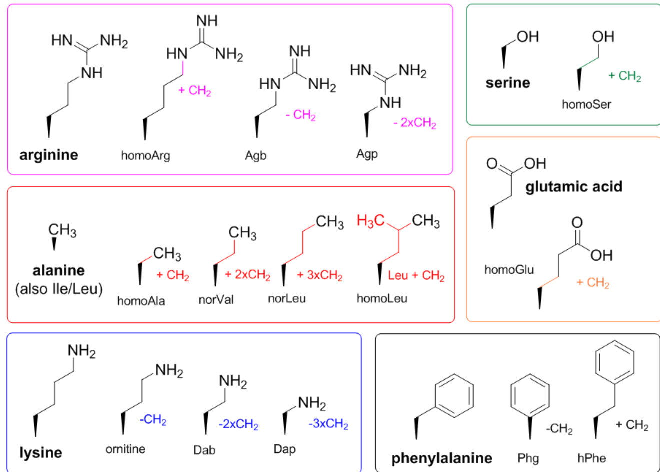
2. Postranslationally modified amino acids



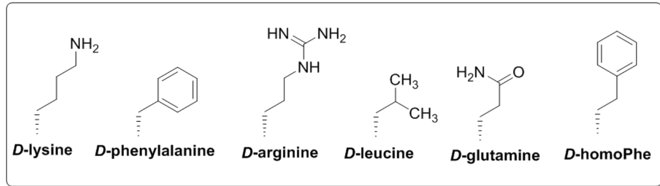
3. Amino acids with TFA/piperidine resistant protecting groups



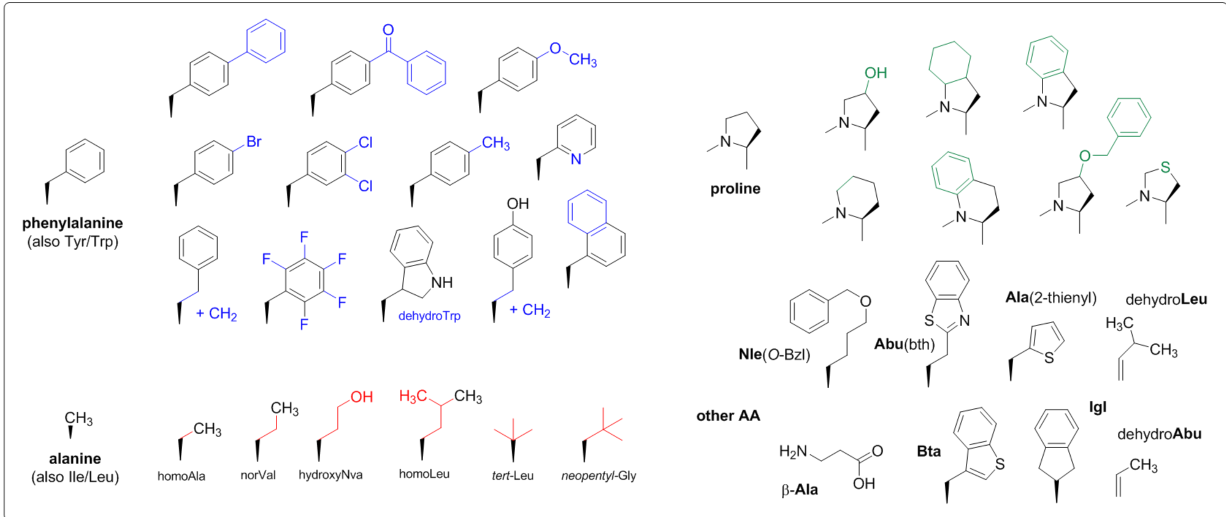
4. –CH_2- group homologs of natural amino acids



5. *D*-amino acids



6. Other amino acids



Box 1 Multiple examples of unnatural amino acids that can be applied for HyCoSuL synthesis and used to scan the substrate specificity of proteases.

Unnatural amino acids. The main concept of HyCoSuL is to use unnatural amino acids to better explore protease active site preferences. In this protocol we describe the use of over 100 unnatural amino acids in P1-Asp library for screening caspases. However, the pallet of these amino acids can be customized to the protease of interest. Based on the selection criteria we can divide all amino acids in different groups[46], however for our purposes we usually distinguish six groups of amino acids (see Box 1). One group are amino acids with different protecting groups (Asp(benzyl), Arg(methyl), etc.). These amino acids can be considered as unnatural analogues, if they are stable under conditions of the library synthesis. Since our protocol is based on the widely used Fmoc/tBu strategy, these groups must be resistant to reagents for Fmoc- (piperidine) and tBu- (trifluoroacetic acid, TFA) de-protection. For instance, aspartic acid protected with –methyl, -cyclohexyl, –benzyl, or –β-menthyl groups will stay intact during the whole synthesis, while –*tert*-butyl or -2-phenylisopropyl groups are TFA-labile. On the other hand, basic amino acids, such as arginine, can be protected with –methyl, -benzyl, or –nitro groups that are not hydrolyzed, but several other protecting groups as –Pbf, -bis-Boc are easily removed upon TFA treatment. A comprehensive list of chemical groups for amino acid protection, and the conditions for their removal, was reviewed by Isidro-Llobet et al.[47].

Library synthesis. In this protocol we use orthogonal Fmoc/tBu chemistry for the synthesis of tetrapeptide ACC-labeled combinatorial library. This library contains aspartic acid at P1 (suitable for caspase, legumain and proteasome caspase-like subunit screening), but any other amino acid (natural or unnatural) can be used. This synthesis is calculated for 3 sublibraries: P4, P3, P2, each of them containing 19 natural and 110 unnatural amino acids – the size of the library depends on the number of unnatural amino acids used. This synthesis is performed in 48-well cartridges for solid-phase and thus substrates are synthesized in parallel. If the number of amino acids (natural + unnatural) exceeds 48, the synthesis for other amino acids (49-96, 97-144, and so on) can be repeated under the same experimental setup as for the first (1-48) set, or by using three 48-well cartridges at the same time. The general procedure for the synthesis of peptides on the solid support is known for many years and very well described in protocols for individual peptides synthesis. In this protocol we applied this general procedure and optimized it for the synthesis of combinatorial peptides with unnatural amino acids. The key element of this procedure is to obtain 100% coupling of the ACC fluorophore, individual amino acids or isokinetic mixture to the solid support. Here we highlight some critical points. *(1)* The best reagent for coupling Fmoc-ACC-OH into Rink Amide resin is the HOBt/DICI pair. Although HATU/collidine is more potent, the final product is not pure, as there are some other side reactions between Fmoc-ACC-OH and HATU. *(2)* The optimal reagents for coupling of the first amino acid to the H_2N-ACC-resin is the HATU/2,4,6-trimethylcollidine pair. However, even using HATU some amino acids are not coupled completely to the resin[44]. In these cases, the unreacted N-terminal amine of the H_2N-ACC-resin must be acetylated following the protocol described by Maly et al.[44]. *(3)* In our primary studies (published in 2014) we used HBTU/DIPEA for coupling of the individual amino acids and isokinetic mixtures to the solid support, however as the price of HATU is getting lower, this reagent can be used in place of HBTU (which makes the coupling time shorter). *(4)* In the

protocols describing Fmoc-removal with piperidine in DMF solution we found several ways to do this (different piperidine concentration, numbers of de-protection cycles, timing). As the results of HyCoSuL are combinatorial peptide mixture that cannot be purified on HPLC, we must be sure that all the Fmoc-groups are removed after every cycle. Thus, for Fmoc de-protection we routinely use 20% piperidine in DMF in three cycles (5min, 5min, and 30min).

Isokinetic mixture. An isokinetic mixture contains 19 amino acids at a ratio that corresponds to their individual reaction kinetic of coupling to the free amine (for example the slower rate of the coupling with amine the higher ratio of such amino acid in the mixture). Using such a mixture ensures the equal distribution of all amino acids in the product after coupling. In our protocol we use the isokinetic mixture determined by Ostresh[23]. Since HyCoSuL can contain various number of unnatural amino acids, the amount of reagents (including isokinetic mixture) must be scaled to the numbers of positions in the library. To create our library we used the following ratios of Fmoc-protected amino acids (ratio in %): Fmoc-*L*-Ala-OH 3.4; Fmoc-*L*-Arg(Pbf)-OH 6.5; Fmoc-*L*-Asn(Trt)-OH 5.3; Fmoc-*L*-Asp(*t*Bu)-OH 3.5; Fmoc-*L*-Gln(Trt)-OH 5.3; Fmoc-*L*-Glu(*t*Bu)-OH 3.6; Fmoc-Gly-OH 2.9; Fmoc-*L*-His(Trt)-OH 3.5; Fmoc-*L*-Ile-OH 17.4; Fmoc-*L*-Leu-OH 4.9; Fmoc-*L*-Lys(Boc)-OH 6.2; Fmoc-*L*-Nle-OH 3.8; Fmoc-*L*-Phe-OH 2.5; Fmoc-*L*-Pro-OH 4.3; Fmoc-*L*-Ser(*t*Bu)-OH 2.8; Fmoc-*L*-Thr(*t*Bu)-OH 4.8; Fmoc-*L*-Trp(Boc)-OH 3.8; Fmoc-*L*-Tyr(*t*Bu)-OH 4.1; Fmoc-*L*-Val-OH 11.3.

Library screening and data analysis. Library screening of proteases of interest can be performed on any multi-well (96 or 384) plates appropriate for fluorescence assays. We pipette our libraries manually (1μL of library in DMSO) thus we use 96-well plates (see **Figure 2**). However, 384-well plate format can also be utilized if automatic pipetting equipment is available. In such a case the whole sub-library with over 100 unnatural amino acids can be screened on one plate. Once the proper plates are chosen, the library is pipetted into wells, enzyme is added with multichannel pipette and the increase of fluorescence is monitored over time. The crucial factor in the library screening is to identify and utilize only the linear part of the plot for determination of reaction rates. This is important especially for enzymes with narrow specificity at certain positions. A good example here is caspase 3. This enzyme is highly stringent for Asp at P4, since the activity for other amino acids is below 10%. However it is important to determine the activity for all of these amino acids, for example in order to distinguish between caspase 3 and other caspases (see Box 2). Here we show that the specificity of an enzyme does not depend on its concentration if only the linear part of the plot for each amino acid is selected for the analysis. Nevertheless, we recommend using an enzyme concentration at which even 1% of the signal from the best amino acids can be accurately detected.

Plate A
natural + unnatural amino acids

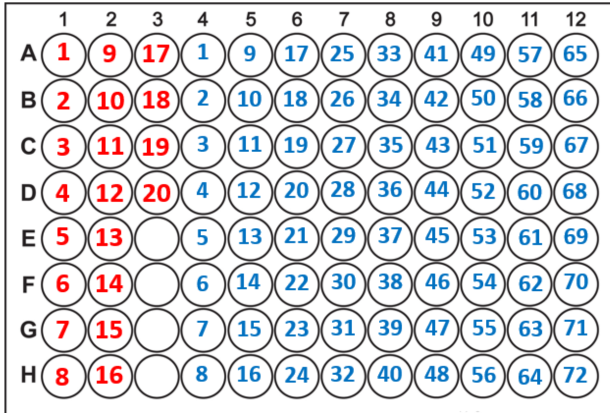


Plate B
4 natural (control) + unnatural amino acids

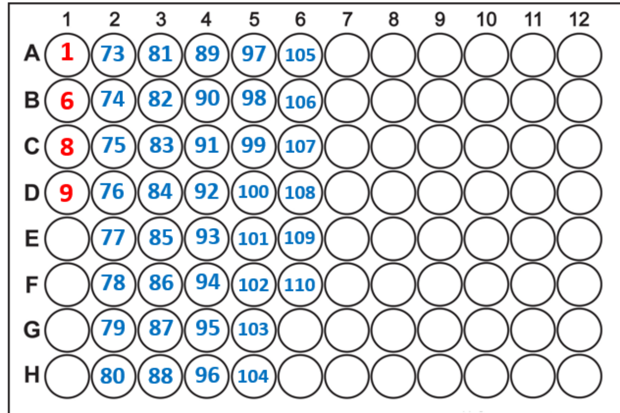
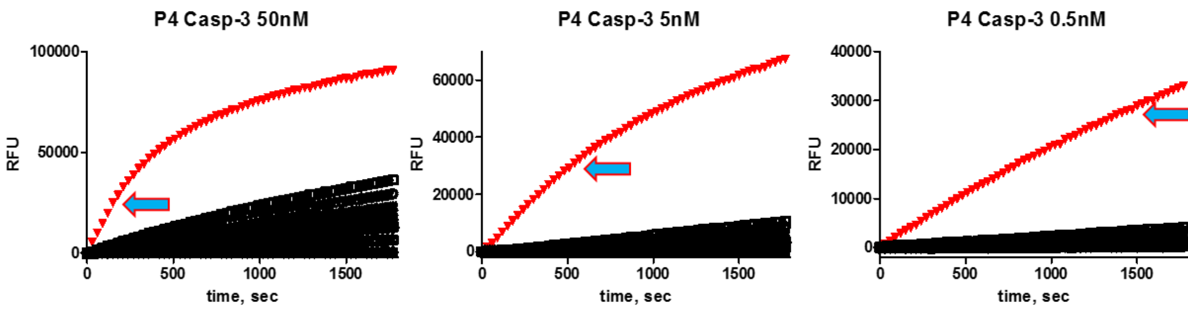
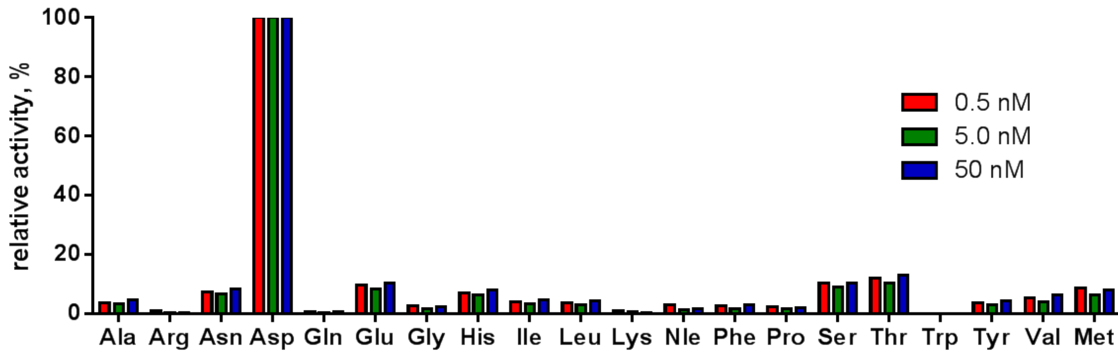


Figure 2 Diagram of 96-well plate for placement of HyCoSuL substrates. Red numbers indicate natural amino acids (cysteine was replaced with norleucine) and blue numbers indicate unnatural amino acids. On Plate B several good natural substrates are screened again in order to combine the results from two plates into one diagram.





Box 2 The P4 analysis of caspase-3 shows that the specificity does not depend on the enzyme concentration. In these screenings the enzyme concentration was 50nM, 5nM, and 0.5nM. Each screen was run for 30 min (library concentration 100μM). In the top panel the red line is the hydrolysis of Asp (RFU – relative fluorescence units), the black lines are the other amino acids. The best readout for weakly recognized amino acids is when 50nM of enzyme was used. However, the linear part for the P4 Asp cleavage is only 200 sec, so only this part can be taken

for the RFU/s calculations. When caspase 3 was used at 0.5nM the line for P4 Asp cleavage is straight for almost 30min, however other amino acids produce only poor fluorescence. Nevertheless, all three experiments produce the same caspase 3 P4 specificity matrix.

Validation of screening results and selection of optimal sequence. As we have mentioned previously, one of the drawback of using positional scanning libraries to profile protease preferences is the lack of information regarding potential subsite cooperativity. It can turn out that the best amino acids form P4, P3, and P2 screening when combined into an individual substrate does not generate the optimal (most active or most selective) molecule. Thus, in order to overcome this limitation we recommend to synthesize several individual substrates with the selected amino acids and determine the kinetic parameters toward the enzyme of interest. Although strong subsites cooperativity is not a common issue in proteases we strongly encourage validation of screening results, especially when new proteases are tested. More information about protease subsites cooperativity can be found elsewhere[48].

In this protocol we describe the use of HyCoSuL in order to design short substrates that can discriminate between human apoptotic caspases and legumain (AEP, asparaginyl endopeptidase). After screening the whole P4-P2 library toward six apoptotic caspases and legumain, and creating a specificity matrix (heat-maps) for all of these enzymes (see Figure 3) we were able to design selective sequences for almost all tested enzymes.

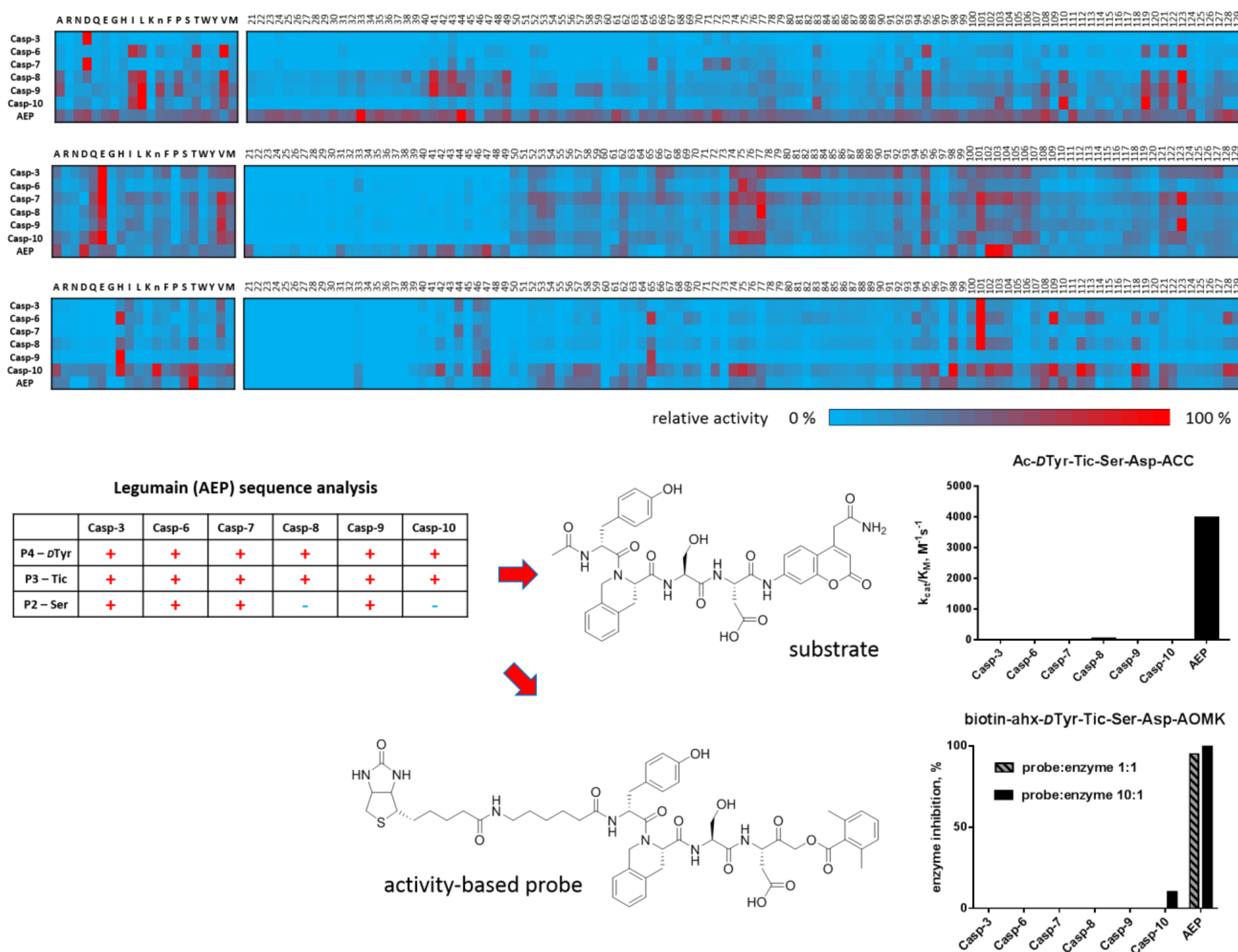


Figure 3 The selection process of protease specific substrates and probes. The top panel presents the P4, P3 and P2 specificity profiles for six human apoptotic caspases and legumain (AEP) determined through HyCoSuL. Red squares represent the amino acids that are hydrolyzed very well by the enzymes, and blue squares indicate amino acids that are not recognized by the enzymes. The bottom panel describes selection of a specific legumain substrate. The unnatural amino acids *D*-Tyr at P4 and Tic at P3 are selective for legumain over all other enzymes (red plus). The most legumain selective amino acid at P2 position is Ser, which is not recognized by caspases-3, -6, -7, and -9 (red plus) and was only accepted by caspases-8 and -10 (blue minus). After selection, the Ac-*D*-Tyr-*L*-Tic-*L*-Ser-*L*-Asp-ACC substrate and biotin-6-ahx-*D*-Tyr-*L*-Tic-*L*-Ser-*L*-Asp-AOMK were synthesized and demonstrated to be the first legumain specific substrate and probe with Asp at P1 position reported to date [24].

One size does not fit all. This protocol focuses on the screening of caspase preferences with a P1-Asp hybrid combinatorial library. This library can be also used for other proteases that accept Asp at P1 position (legumain or the proteasome caspase-like subunit). However, for other enzymes different libraries are needed. We have already synthesized and validated two other libraries with Arg and Ala amino acids at P1 positions (Figure 4). These three libraries satisfy the primary requirements of several proteolytic enzymes, however in some cases other P1 libraries are also needed. On the other hand, many proteases display broad P1 preferences, thus unnatural amino acids can be also incorporated at the P1 position. Since it is very challenging to

synthesize combinatorial libraries with different P1 positions (especially with unnatural amino acids) we have proposed another strategy with the synthesis of individual substrate libraries with predefined P4-P2 positions and various P1. Such libraries were successfully applied for P1 human neutrophil elastase and ZIKA virus NS2B-NS3 protease screenings. In Figure 5 we present a simplified algorithm for the synthesis of (1) P1-defined combinatorial library and (2) individual libraries with various P1 position.

From substrate to inhibitor and activity-based probe. Highly selective and active substrates for proteolytic enzymes can be converted into inhibitor and/or activity-based probe that can be used to block or track selected protease in biological systems (for example cells or whole organisms). There are multiple protocols and procedures for the synthesis of inhibitors/probes with different warheads and tags for all classes of proteolytic enzymes[18,24,26,49-53].

In our protocol we describe the synthesis of the first P1-Asp legumain specific biotin-labeled activity based probe for legumain. The synthesis of the biotin-labeled activity based probe is divided into three parts: (block A) synthesis of biotin-6-ahx-*D*-Tyr(*t*Bu)-*L*-Tic-*L*-Ser(*t*Bu)-OH peptide on solid phase (using 2-chlorotrityl chloride resin), (block B) synthesis of H_2N-Asp(Bzl)-AOMK electrophilic warhead (in solution) and (block C) joining these molecules into the final product. This example presents a classical approach for the synthesis of protease inhibitors/activity-based probes.

1.8. Conclusions

In conclusion, the HyCoSuL protocol is a convenient chemical tool for determining the substrate specificity of proteases in the non-prime region of the active site. The use of a wide range of unnatural amino acids in the combinatorial library structure (P4-P2 positions) allows for both very precise and extensive explorations of the chemical space surrounding protease active sites. Thus, it has become possible to distinguish between closely related enzymes (apoptotic caspases and legumain), or to develop new, highly active substrates (human neutrophil elastase). We also demonstrate that the use of unnatural amino acids is not only limited to P4-P2 region, but it can be also applied into P1 position (individual libraries) to dissect primary specificity of multiple proteases (for example ZIKA virus NS2B-NS3 protease). In summary, we have developed a new highly-adaptable chemical concept where a large set of unnatural amino acids are used in an ordered manner to investigate proteolytic enzymes.

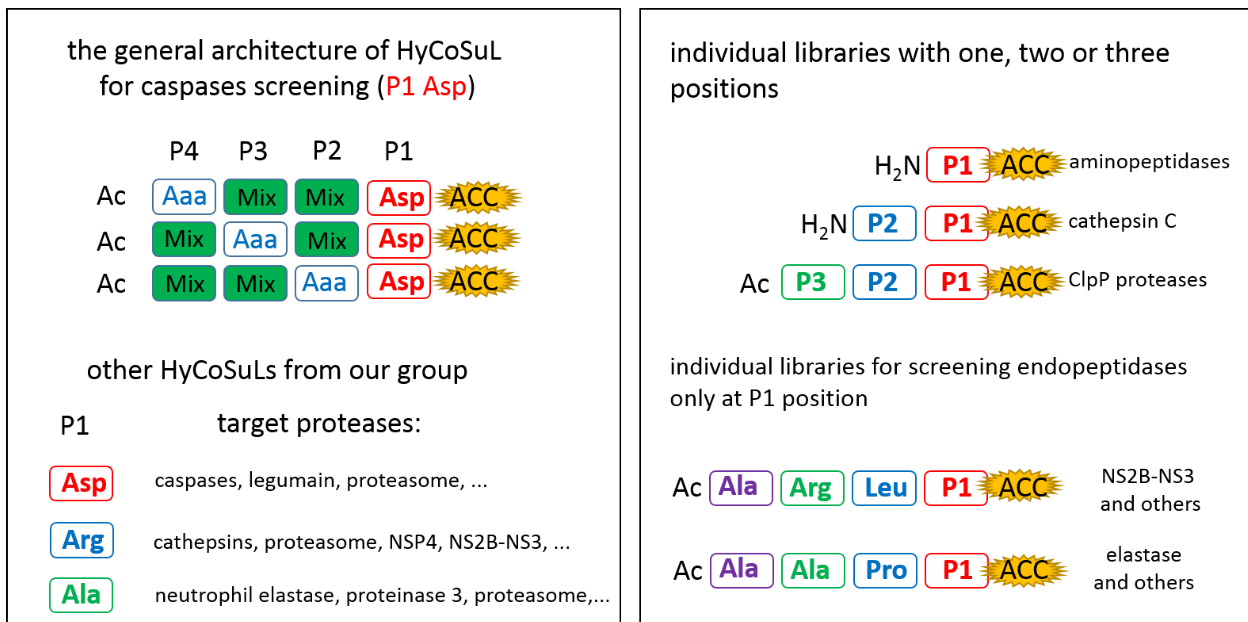


Figure 4 Unnatural amino acids can be used for various types of peptide substrate libraries. Left panel – the application of unnatural amino acids in combinatorial libraries (HyCoSuL). Right panel – the use of unnatural amino acids in the individual libraries for exo- and endo-proteases.

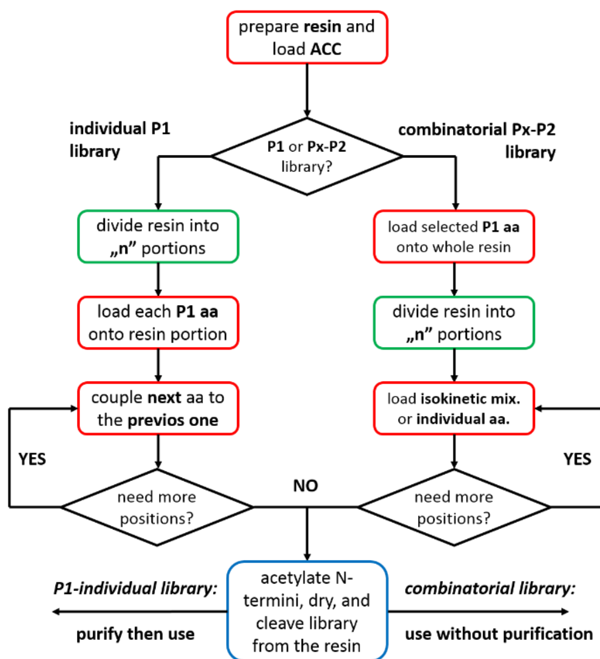


Figure 5 The simplified algorithm for the synthesis of combinatorial either individual substrate libraries with unnatural amino acids.

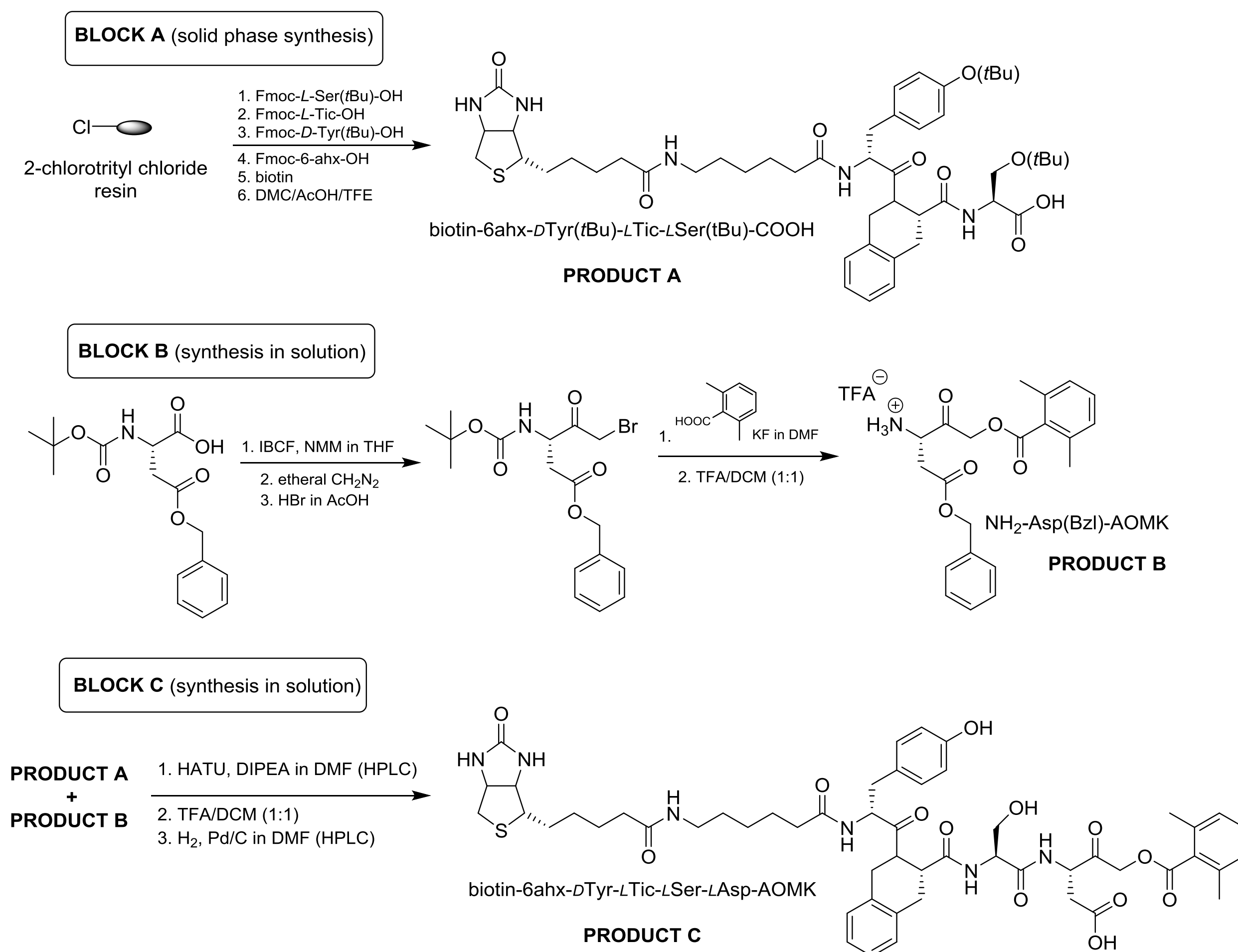


Figure 6 General scheme of the building blocks synthesis of legumain selective activity-based probe

2. MATERIALS

REAGENTS

! CAUTION Most of the reagents used in this protocol require proper gloves, googles, and lab coats.

IMPORTANT Below we present the list of reagents and suppliers that we used to complete this protocol, however all the reagents can be purchased from other suppliers as long as the display at least the same level of purity as these indicated below.

- Fmoc-Rink-Amide AM resin (Iris Biotech GmbH, cat. no. BR-1340)
- 2-chlorotrityl chloride resin (Iris Biotech GmbH, cat. no. BR-1065)
- Fmoc-protected amino acids (various suppliers)
- Boc-*L*-Asp(Bzl)-OH (for the synthesis of inhibitor and activity-based probe)

- 7-Fmoc-aminocoumarin-4-acetic acid (Fmoc-ACC-OH or Fmoc-ACA-OH, Iris Biotech GmbH, cat. no. RL-1170). This fluorophore can be also synthesized according to the method described by Maly et al.[44].
- *N*-hydroksybenzotriazole (HOBt monohydrate; FisherScientific, cat. no. NC9894735) **! CAUTION** It is an irritant.
- Diisopropylcarbodiimide (DICI, Iris Biotech GmbH, cat. no. RL-1015) **! CAUTION** It is extremely flammable and toxic.
- *N,N*-diisopropylethylamine (DIPEA, Iris Biotech GmbH, cat. no. SOL-003) **! CAUTION** It is corrosive and highly flammable.
- 2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, Iris Biotech GmbH, cat. no. RL-1030) **! CAUTION** It is an irritant/harmful.
- 2-(1H-7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronoium hexafluorphosphate (HATU, Iris Biotech GmbH, cat. no. RL-1190) **! CAUTION** It is an irritant/harmful.
- 2,4,6-trimethylpyridine (2,4,6-trimethylcollidine, Sigma-Aldrich, cat. no. 142387) **! CAUTION** It is corrosive and highly flammable.
- *N,N*-dimethylformamide (DMF, Sigma-Adrich, cat. no. 437573) **! CAUTION** It is toxic.
- Dichloromethane / methylene chloride (DCM, POCh, cat. no. 628410114) **! CAUTION** It is toxic.
- Methanol (MeOH, Avantor Materials, cat. no. 9070-03) **! CAUTION** It is highly flammable and toxic.
- Acetonitrile (ACN, Sigma-Aldrich, cat. no. 360457) **! CAUTION** It is highly flammable.
- Diethyl ether (Et_2O, Sigma-Aldrich, cat. no. 676845) **! CAUTION** It is extremely flammable.
- Piperidine (Iris Biotech GmbH, cat. no. SOL-010) **! CAUTION** It is toxic and highly flammable.
- Ninhydrin (Sigma-Aldrich, cat. no. 151173) **! CAUTION** It is harmful.
- *p*-Chloranil (Sigma-Aldrich, cat. no. 232017) **! CAUTION** It is harmful.
- Acetaldehyde (Sigma-Aldrich, cat. no. 402788) **! CAUTION** It is highly corrosive.
- Trifluoroacetic acid (TFA, Iris Biotech GmbH, cat. no. SOL-011) **! CAUTION** It is highly corrosive. Extra precautions are required.
- Trifluoroethanol (TFE, Iris Biotech GmbH, cat. no. SOL-015) **! CAUTION** It is harmful.
- Triisopropylsilane (TIPS, Iris Biotech GmbH, cat. no. RL-1102) **! CAUTION** It is an irritant.
- Acetic acid (AcOH, Sigma-Aldrich, cat. no. 695092) **! CAUTION** It is highly corrosive.
- Phosphorus pentoxide (P_2O_5, Sigma-Aldrich, cat. no. 298220) **! CAUTION** It is highly corrosive.
- DMSO, for molecular biology (Sigma-Aldrich, cat. no. D8418) **! CAUTION** It can transport chemical compounds into the body
- UtraPure™ Sucrose (Invitrogen, cat. no. 15503-022)
- DTT (Dithiothreitol, BioPioneer, cat. no. C0041) **! CAUTION** It is an irritant/harmful.
- 1,4-piperazinediethanesulfonic acid (PIPES, Sigma-Aldrich, cat. no. P6757)
- Sodium chloride for molecular biology (NaCl, Sigma-Aldrich, cat. no. S3014)

- EDTA disodium dehydrate (Genesee Scientific, cat. no. 20-147)
- Sodium Citrate, dehydrate (MACRON fine chemicals, cat. no. 0754-06)
- Tetrahydrofuran (THF, anhydrous, Sigma-Aldrich, cat. no. 401757) **! CAUTION** It is highly flammable.
- 4-methylmorpholine (NMM, Sigma-Aldrich, cat. no. 407704) **! CAUTION** It is flammable and corrosive
- Isobutyl chloroformate (IBCF, Sigma-Aldrich, cat. no. 177989) **! CAUTION** It is toxic and corrosive
- Ethereal diazomethane solution generated according to the Aldrich Technical Bulletin (AL-180) protocol **! CAUTION** It is highly toxic and explosive
- Hydrogen bromide in acetic acid (HBr 30% wt. in CH_3COOH; Sigma-Aldrich cat. no. 18735) **! CAUTION** It is very corrosive
- Sodium bicarbonate ($NaHCO_3$, Sigma-Aldrich cat. no. S5761)
- Magnesium sulfate ($MgSO_4$, anhydrous, Sigma-Aldrich cat. no. M7506)
- Potassium fluoride (KF, anhydrous, Sigma-Aldrich cat. no. 449148) **! CAUTION** It is toxic
- 2,6-dimethylbenzoic acid (2,6-DMBA, Sigma-Aldrich, cat. no. 156906)
- Biotin (Sigma-Aldrich, cat. no. B4501)
- Ethyl acetate (Sigma-Aldrich, cat. no. 437549) **! CAUTION** It is flammable and irritant
- Palladium on carbon (Pd/C, 10 wt. % loading, Sigma-Aldrich cat. no. 205699) **! CAUTION** It is flammable solid
- Citric acid (Sigma-Aldrich, cat. no. 251275)
- Caspase and legumain buffers (see REAGENT SETUP)
- Recombinant human caspases (-3, -6, -7, -8, -9, -10). These enzymes can be purchased from commercial suppliers (Sigma-Aldrich) or expressed and purified according to published protocols[54,55].
- Recombinant human legumain. This enzyme can be purchased from commercial suppliers (R&D Systems) or expressed and purified according to published protocols[24].
- Inhibitors for the active site titration of the enzymes: Cbz-VAD-fmk for caspases and Cbz-ATN-AOMK for legumain.

EQUIPMENT

- Vacuum line with trap
- 250mL glass reaction vessel for solid phase synthesis (Chemglass Life Sciences, cat. no. CG-1866)
- 25mL glass reaction vessel for solid phase synthesis (Chemglass Life Sciences, cat. no. CG-1860)
- Analytical lab balance
- 50-mL conical centrifuge tubes
- 15-mL conical centrifuge tubes
- 1L glass baker (for ice/acetone bath)

- 1mL syringe and needles
- 0.2μm syringe filter
- Magnetic stirrer
- 250mL separatory funnel
- Rotary evaporator with water bath
- 10mL and 100mL round-bottom flasks
- The Aldrich Mini Diazald® Apparatus with Clear-Seal® joints (Sigma-Aldrich cat. no. Z108898) – this equipment is not necessary if you can purchase ethereal solution of diazomethane
- Parafilm® (Bemis Company, HS234526A)
- Teflon spatula
- Glass tubes and glass sticks
- Heating block for 95°C
- -80°C freezer (Eppendorf)
- Lyophilizer with fast-freeze flasks
- Milli-Q water deionizing purification system
- MultiChem 48-wells synthesis apparatus (FlexChem, SciGene)
- Centrifuge with adaptors for 15mL conical tubes (Eppendorf)
- Rocker
- Vortex mixer
- Reverse-phase HPLC (Waters system with UV lamp – 220nm and 254nm)
- Semi-preparative HPLC C8 Discovery® column (particle size 10μm, Sigma-Aldrich)
- Analytical HPLC C8 Discovery® column (particle size 5μm, Sigma-Aldrich)
- Microplate spectrofluorimeter with adjustable temperature (CLARIOstar, BMG Labtech)
- 96-well assay plates, white, flat bottom (Corning, cat. no. 3912)
- Automatic pipettes (1-10μL, 10-100μL, 100-1000μL)
- Multichannel pipette (20-200μL)
- 25mL reagent reservoir (pipette.com, P8025)
- 250mL separatory funnel with stockpot
- 250mL Erlenmeyer flask
- 100mL of 200mL Buchner funnel with fritted disc and vacuum adapter
- Pipette tips (0.1-10μL, 10-200μL, 100-1000μL)

REAGENT SETUP

Caspase buffer 10% w/v sucrose, 20mM PIPES, 10mM NaCl, 1mM EDTA, 10mM DTT (pH 7.2-7.4) in deionized water[22]. The buffer for initiator caspases (8, 9, and 10) is supplemented with 0.75M sodium citrate (to allow caspases for dimerization)[56,57]. Buffers were prepared at room

temperature. Add DTT to the buffer just before the assay. The buffer without DTT can be stored at room temperature up to several weeks, or at +4°C for several months.

Legumain buffer 40 mM citric acid, 1 mM EDTA, 120 mM Na_2HPO_4, and 10 mM DTT, pH 5.8[24]. Add DTT to the buffer just before the assay. The buffer without DTT can be stored at room temperature up to several weeks, or at +4°C for several months.

Extraction reagents For the extraction steps mentioned in the procedure use the following reagents: **(1)** brine – saturated solution of NaCl in deionized water; **(2)** saturated solution of $NaHCO_3$ in water; **(3)** 5% solution of $NaHCO_3$ in deonized water (for 5% solution weigh 5g of $NaHCO_3$ in a glass beaker and fill it with water up to 100mL, mix until all powder is dissolved); **(4)** 5% citric acid (prepare in the same way as 5% $NaHCO_3$).

EQUIPMENT SETUP

Spectrofluorimeter setup Set the spectrofluorimeter to an excitation wavelength of 355nm and an emission wavelength of 455nm (cutoff 405nm). Adjust the sensitivity of the plate reader to the ACC fluorescence[6].

Semi-preparative HPLC Use a RP-HPLC-gradient system equipped with an UV detector (220nm and 254nm) and C8 column (semi-preparative Discovery® column, particle size 10μm). Run a linear gradient of 0-100% B in 40 min (flow 10ml per min). Solvent A 0.1% TFA in water, solvent B 0.1% TFA in acetonitrile.

Analytical HPLC Use a RP-HPLC-gradient system equipped with an UV detector (220nm and 254nm) and C8 column (analytical Discovery® column, particle size 5μm). Run a linear gradient of 0-100% B in 30 min (flow 1mL per min). Solvent A 0.1% TFA in water, solvent B 0.1% TFA in acetonitrile.

3. PROCEDURE

Loading of Fmoc-ACC-OH onto Rink AM resin TIMING: 2 days

Step 1. Weigh 12g (5.76mmol, 1eq) of Rink AM resin (0.48mmol/g) into 250mL glass reaction vessel (cartridge) and add 100mL of DCM to the dried resin for swelling (30 min).

Step 2. Remove DCM by vacuum filtration and wash the resin three times with DMF (60mL for each washing).

Step 3. Remove the Fmoc-protecting group from the resin by using 20% piperidine in DMF (10mL of piperidine and 40mL of DMF = 1 wash). Pour this mixture onto resin and agitate for 5 min. Remove piperidine solution by vacuum filtration, and add the next portion (50mL) of this solution. Agitate for another 5 min. Wash the resin, and add the third portion of piperidine

solution (50mL). Gently agitate for 30 min. Use this protocol (with proper volume of piperidine solution) for each Fmoc de-protection during the library synthesis (5min, 5min, and 30min).

Step 4. Wash the resin six times with DMF (60mL per wash) to remove all piperidine and side reaction products

CRITICAL STEP. It is important to remove all the piperidine from the resin, as even small amount of this amine can de-protect the Fmoc-group from amino acids used in the next step or catalyze side reactions during the coupling of amino acids.

Step 5. Perform a "ninhydrin test" to confirm the Fmoc de-protection. Dissolve 5g of ninhydrin in 100mL of ethanol. Add 1ml of ninhydrin solution to a glass tube and transfer (using glass stick) several beads into the glass tube. Heat the mixture for 3 min at 95°C in a heating block. A positive test (free amine groups) is indicated by dark blue color of resin beads. If the resin is colorless to pale-yellow the result is negative (Fmoc protecting group is still present).

Step 6. In a 50mL falcon tube weigh 6.4g of Fmoc-ACC-OH (14.4 mmol, 2.5eq), 2.16g of HOBt (14.4 mmol, 2.5eq) and dissolve in a minimal amount of DMF. Then, add 1.9 mL of DICI (14.4 mmol, 2.5eq) and pre-activate this mixture for 5 min by gentle stirring.

Step 7. Pour this mixture onto resin and shake gently until all liquid and resin are mixed well. Add DMF if needed. Protect the cartridge from light by covering it with aluminum foil.

CRITICAL STEP. It is very important to dissolve all reagents in a minimal amount of DMF, to increase the reagent concentrations in mixture and allow for the high-yield coupling. However, the mixture must be diluted enough to allow easy mixing.

PAUSE POINT Turn the shaker on, and gently shake the reaction vessel for 24 hours.

Step 8. Remove the mixture from the resin by vacuum filtration and wash the resin three times with DMF (60mL per wash).

Step 9. To ensure the high yield of Fmoc-ACC-OH coupling to the resin, repeat the coupling using ½ of the reagents from the first coupling. In a 50mL falcon tube place 3.18 g of Fmoc-ACC-OH (7.2 mmol, 1.25 eq), 1.1 g of HOBt (7.2 mmol, 1.25eq) and dissolve in a minimal amount of DMF. Then, add 1mL of DICI (7.2 mmol, 1.25eq) and pre-activate this mixture for 5 min by gentle stirring.

Step 10. Pour this mixture onto resin and shake gently until all liquid and resin are mixed well. Add DMF if needed. Protect the cartridge from light by covering it with aluminum foil.

PAUSE POINT. Gently shake the reaction vessel for 24 hours on the shaker.

Step 11. Remove the mixture from the resin by vacuum filtration and wash the resin three times with DMF (60 mL per wash).

Step 12. Perform a ninhydrin test for unreacted amines.

? TROUBLESHOOTING

Step 13. Remove the Fmoc-protecting group from ACC with 20% piperidine in DMF (see Step 3).

Step 14. Perform a ninhydrin test for free NH_2-ACC-resin. This is an aromatic amine so the free amine groups are indicated by orange to red color of resin beads.

Step 15. Wash the resin six times with DMF (60mL per wash) to remove all piperidine and side reaction products.

Loading of the first amino acid (Fmoc-*L*-Asp(*t*Bu)-OH) onto NH_2-ACC-resin TIMING: 3 days

Step 16. In a 50mL falcon tube place 5.93g of Fmoc-*L*-Asp(*t*Bu)-OH (14.4 mmol, 2.5eq), 5.47g of HATU (14.4 mmol, 2.5eq) and dissolve them in a minimal amount of DMF. Then, add 1.9mL of 2,4,6-trimethylcollidine (14.4 mmol, 2.5eq) and pre-activate this mixture for 1 min by gentle stirring.

Step 17. Pour this mixture onto resin and shake gently until all liquid and resin are mixed well. Add DMF if needed. Protect the cartridge from light by covering it with aluminum foil.

PAUSE POINT Gently shake the reaction vessel for 24 hours on the shaker.

Step 18. Remove the mixture from the resin by vacuum filtration and wash the resin three times with DMF (60mL per wash).

Step 19. To ensure the high yield of Fmoc-*L*-Asp(*t*Bu)-OH coupling to the ACC, repeat the coupling using ½ of the reagents from first coupling. In a 50mL falcon tube place 2.96 g of Fmoc-*L*-Asp(*t*Bu)-OH (7.2 mmol, 1.25 eq), 2.74 g of HATU (7.2 mmol, 1.25eq) and dissolve in a minimal amount of DMF. Then, add 0.95mL of 2,4,6-trimethylcollidine (7.2 mmol, 1.25eq) and pre-activate this mixture for 1 min. by gentle stirring.

Step 20. Pour this mixture onto resin and shake gently until all liquid and resin are mixed well. Add DMF if needed. Protect the cartridge from light by covering it with aluminum foil.

PAUSE POINT Gently shake the reaction vessel for 24 hours on the shaker.

Step 21. Remove the mixture from the resin by vacuum filtration and wash the resin three times with DMF (60mL per wash).

Step 22. Perform a ninhydrin test for unreacted amines.

? TROUBLESHOOTING

Step 23. Remove the Fmoc-protecting group from Asp with 20% piperidine in DMF (see Step 3).

Step 24. Wash the resin six times with DMF (60mL per wash) to remove all piperidine and side reaction products.

Step 25. Wash the resin three times with DCM (60mL per wash) and three times with MeOH (60mL per wash).

Step 26. Dry the resin in the desiccator over P_2O_5 overnight. If you desiccator is too small for the cartridge, transfer all the resin with plastic/Teflon spatula into a beaker. Cover a beaker with Parafilm® and make several holes with a needle. Exchange of P_2O_5 is needed.

PAUSE POINT Dried H_2N-*L*-Asp(*t*-Bu)-ACC-resin can be stored in a vacuum desiccator at room temperature for about a month, or at -20°C for several months.

Synthesis of a P2 HyCoSuL sub-library (Ac-Mix-Mix-Aaa-Asp-ACC) TIMING: 4-5 days for each sub-library

Step 27. Each sub-library (P2, P3, or P4) can be synthesized separately. Here we present the P2 sub-library synthesis in details. This library consists of 129 members. Weight 80 mg (0.04 mmol, 1eq) of resin into each well of a MultiChem 48-wells synthesis apparatus (here we used three such apparatuses to synthesize the whole P2 sub-library, see Experimental Design). For this library you can also use any other multi-well device suitable for solid phase synthesis.

Step 28. Swell the resin in each well in 1mL of DCM for 30 min.

Step 29. Wash the resin three times with DMF (1-3 mL per wash) using wash bottle.

CRITICAL STEP When you work with multi-well cartridge, every time you wash the resin you can pour DMF directly from the wash bottle, instead of pipetting 1-2mL into each well using Pasteur pipette. This makes all the washing steps much faster and straightforward.

Step 30. In a 1.5mL Eppendorf tube place 2.5eq of Fmoc-amino acid (0.1mmol). For this library we prepared 119 such vials with 119 different (19 natural and 110 unnatural) amino acids.

Step 31. In a 50mL falcon tube place 0.72g of HOBt (4.8mmol, 120eq) and dissolve in 48mL of DMF (per one multi-well cartridge).

Step 32. Transfer 1mL of HOBt/DMF solution into the 1.5 mL Eppendorf tube with Fmoc-amino acid and dissolve it. Here we prepared 119 such samples.

Step 33. Add 15μL of DICI (0.1 mmol, 2.5eq) into each tube, mix and pour the solution onto the resin (1 sample into 1 well). The best way to do this is: add DICI to the first tube, mix, and pour onto resin (1. well), add DICI to the second tube, mix, and pour onto resin (2. well) and so on. It is recommended by the authors to perform this step by two persons. When all wells are supplied with reagents in DMF, cover the MultiChem apparatus with top lid.

PAUSE POINT Gently shake the multi-well cartridge for 3 hours on the shaker.

Step 34. Wash all the wells three times with DMF (1-3 mL per wash) using wash bottle and filtrate using vacuum. Then perform a ninhydrin test (see Step 5). Select resin from around 10 wells and using the glass stick transfer them to the 10 glass tubes filled with 1mL of ninhydrin solution.

? TROUBLESHOOTING

Step 35. Remove the Fmoc-protecting group from P2 amino acid by using the piperidine solution (see Step 3). For multi well system you can use wash bottle with 20% piperidine in DMF or plastic transfer pipette. Add approximately 1-2 mL of de-protecting solution to each well.

Step 36. Wash all the wells six times with DMF (1-3 mL per wash) using wash bottle and filtrate using vacuum.

CRITICAL STEP For some amino acids (proline, and its derivatives such as hydroxyproline, azetidine, etc.) ninhydrin test is not a suitable method for the detection of free N-terminus. Thus, another method (acetaldehyde/chloranil test) must be applied here[58].

Step 37. Perform a ninhydrin test on resin samples from several randomly selected wells. In order to detect free proline (and its derivatives) perform acetaldehyde/chloranil test. In two separate tubes prepare 2% solution of *p*-chloranil in DMF and 2% solution of acetaldehyde in DMF. Then place a few beads of resin in glass tube and add 1-3 drops of each solution. Mix and incubate the resin at room temperature for 5 min. Beads that turn blue indicate the free N-terminal proline. Perform this test for every sample that contains proline (or one of its derivative) on the N-terminus.

? TROUBLESHOOTING

Step 38. Prepare an isokinetic mixture of natural amino acids for P3 position. In a 50mL falcon tube place isokinetic mixture (5eq, 9.22 mmol), 1.4g of HOBt (9.22 mmol, 5eq) and dissolve it with DMF to 48mL (per one multi-well cartridge). Add 1.2mL of DICI (9.22mmol, 5eq) to the mixture and activate it for 3 min.

CRITICAL STEP It is very important to activate the mixture for 3 min to ensure that all amino acids from the isokinetic mixture are activated equally.

Step 39. Aliquot the pre-activated mixture into the multi-cartridge wells (1mL per well) using Pasteur or plastic pipette. Close the multi-cartridge with top lid.

PAUSE POINT Gently shake the multi-well cartridge for 3 hours on the shaker.

Step 40. Wash all the wells three times with DMF (1-3 mL per wash) using wash bottle and filtrate using vacuum.

Step 41. Perform a ninhydrin test.

? TROUBLESHOOTING

Step 42. Remove the Fmoc-protecting group from P3 amino acid by using the procedure from Step 3. For multi-well system you can use wash bottle with 20% piperidine in DMF or plastic transfer pipette. Add approximately 1-2 mL of de-protecting solution to each well.

Step 43. Wash all the wells six times with DMF (1-3 mL per wash) using wash bottle and filtrate using vacuum.

Step 44. Perform a ninhydrin test on resin samples from several randomly selected wells.

? TROUBLESHOOTING

Step 45. Prepare an isokinetic mixture of natural amino acids for P4 position. In a 50mL falcon tube place isokinetic mixture (9.22 mmol, 5eq), 1.4g of HOBt (9.22 mmol, 5eq) and dissolve it with DMF to 48mL (per one multi-cartridge). Add 1.2mL of DICI (9.22mmol, 5eq) to the mixture and activate it for 3 min.

CRITICAL STEP It is very important to activate the mixture for 3 min to ensure that all amino acids from the isokinetic mixture are activated equally.

Step 46. Aliquot the pre-activated mixture into the multi-cartridge wells (1mL per well) using Pasteur or plastic pipette. Close the multi-cartridge with top lid.

PAUSE POINT Gently shake the multi-well cartridge for 3 hours on the shaker.

Step 47. Wash all the wells three times with DMF (1-3 mL per wash) using wash bottle and filtrate using vacuum.

Step 48. Perform a ninhydrin test.

? TROUBLESHOOTING

Step 49. Remove the Fmoc-protecting group from P4 amino acid by using the procedure from Step 3. For multi well system you can use wash bottle with 20% piperidine in DMF or plastic transfer pipette. Add approximately 1-2 mL of de-protecting solution to each well.

Step 50. Wash all the wells six times with DMF (1-3 mL per wash) using wash bottle and filtrate using vacuum.

Step 51. Perform a ninhydrin test on resin samples from several randomly selected wells.

? TROUBLESHOOTING

Step 52. Acetylation of the N-terminal end of the peptide library. In a 50mL falcon tube place 3.5g of HBTU (9.22 mmol, 5eq) and dissolve it with DMF to 48mL (one falcon tube per cartridge). Then add 530µL of acetic acid (9.22 mmol, 5eq) and 1.6mL of DIPEA (9.22mmol, 5eq). Pre-activate this mixture for 1 min by gentle shaking.

Step 53. Aliquot the pre-activated mixture into the multi-cartridge wells (1mL per well) using Pasteur or plastic pipette.

PAUSE POINT Gently shake the multi-well cartridge for 45 minutes on the shaker.

Step 54. Wash all the wells three times with DMF (1-3 mL per wash) using wash bottle and filtrate using vacuum.

Step 55. Perform a ninhydrin test.

? TROUBLESHOOTING

Step 56. Wash all the wells three times with DCM (1-3 mL per wash) using wash bottle and filtrate using vacuum.

Step 57. Wash all the wells three times with methanol (1-3 mL per wash) using wash bottle and filtrate using vacuum.

Step 58. Dry the resin in the desiccator over P_2O_5 overnight. Replace P_2O_5 if needed.

CRITICAL STEP The resin must be dry in order to cleave the library from it.

Step 59. Cleave peptide substrates from the resin. For each multi-well cartridge (48 wells) prepare 100mL of cleavage solution (95mL of TFA, 2.5mL of H_2O, 2.5mL of TIPS) and cool it on ice for 1 hour.

Step 60. Aliquot 1mL of such mixture into each well using Pasteur pipette and shake the cartridge once per 10-15 min for 2 hours. Save remaining 50mL of cleavage mixture on ice or at +4°C.

Step 61. Filter and collect separately the mixture from each well into 15mL falcon tube (1 well – 1 tube).

CRITICAL STEP Use permanent marker to label all the 15-mL falcon tubes as these samples will be further frozen at -80°C and lyophilized. We even recommend to scratch the numbers on falcon tubes with sharp knife, scalpels or similar tools!!!

Step 62. Wash each well with the remaining cleavage mixture (1mL per well) using Pasteur pipette.

Step 63. Filter and collect the mixture from each well into the same 15ml falcon tube as previous. Now each tube contains around 2mL of substrate solution.

Step 64. Add 13 mL of ice cold diethyl ether to each 15mL falcon tube. Close the tube, shake it vigorously and allow the substrate for precipitation at -20°C for 30 min.

Step 65. Centrifuge each 15mL falcon tube for 5 min at 4°C (3000 x g) and remove the supernatant by decantation.

Step 66. Add 5mL of ice cold diethyl ether to each tube, shake it vigorously, and centrifuge again under the same conditions.

Step 67. Remove the supernatant by decantation and dry the crude product on air until all diethyl ether evaporates (couple of hours – overnight). Protect the substrates from light. We recommend to evaporate all the samples under the fume hood covered with aluminum foil.

Step 68. Prepare ACN/water mixture for lyophilization (150mL per 48 tubes/1 multi-well cartridge). Mix 100mL of ACN and 50mL of water, and aliquot 3mL of such solution into each tube. Dissolve each substrate by mixing on vortex and place them into -80°C overnight.

? TROUBLESHOOTING

Step 69. Lyophilize all substrates in 15mL falcon tubes.

? TROUBLESHOOTING

CRITICAL STEP. During the lyophilization protect the samples from light. We recommend to cover lyophilization (fast-freeze) flasks with aluminum foil.

Step 70. Weight each substrate and dissolve it in DMSO to the final concentration of 10 mM.

Step 71. Store at -80°C until use.

PAUSE POINT HyCoSuL can be stored in -80°C for at least 4 years.

Step 72. Perform the synthesis of HyCoSuL P3-sublibrary by repeating the cycle starting from Step 1 going to Step 71.

CRITICAL STEP. Change the coupling order for P3 sub-library: isokinetic mixture (P2), individual amino acids (P3), isokinetic mixture (P4), and N-acetylation.

Step 73. Perform the synthesis of HyCoSuL P4-sublibrary by repeating the cycle starting from Step 1 going to Step 71.

CRITICAL STEP. Change the coupling order for P4 sub-library: isokinetic mixture (P2), isokinetic mixture (P3), individual amino acids (P4), and N-acetylation.

Preparing assay for the initial screening TIMING: 4-6 hours

IMPORTANT. Before any kinetic assay each enzyme should be active site titrated in order to obtain reproducible data. In this protocol we titrated all caspases and legumain using Cbz-VAD-fmk and Cbz-ATN-AOMK inhibitors, respectively.

Step 74. Remove the P4-P2 HyCoSuL from -80°C freezer and allow it to warm to room temperature for at least 3-4h.

Step 75. Prepare the buffer for caspase-3. Add DTT into the buffer.

Step 76. Remove caspase-3 aliquot from -80°C freezer and thaw it on ice.

Step 77. Turn on the spectrofluorimeter and set the temperature to 37°C.

Step 78. Before you do the screening of the whole library, perform an initial screening (only natural amino acids) to determine the optimal enzyme concentration. Vortex each substrate and by using micropipette place 1μL of it onto 96-well plate (see the diagram). Such prepared plate can be stored at room temperature for 2-3 hours (avoid direct exposure to light).

Step 79. Incubate caspase-3 in 6mL of assay buffer for 15 min in 37°C water bath (or in the incubator). Use the information from previous studies from literature (substrate kinetics, inhibitor kinetics, etc) to select the appropriate enzyme concentration.

Step 80. Pour the pre-incubated assay buffer with caspase-3 onto 25mL reagent reservoir.

Step 81. Using eight-channel pipette transfer 99μL of the caspase-containing buffer into wells containing 1μL of substrate library.

Step 82. Shake the plate in the plate reader (spectrofluorimeter) for 10 seconds. Read the fluorescence (usually RFU – relative fluorescence units) over the time for 30 min (every 15 or 30 seconds). The ACC wavelength setup is: 355nm (excitation) and 460nm (emission).

? TROUBLESHOOTING

Step 83. Calculate the RFU/s value for each substrate.

CRITICAL STEP For each substrate cleavage use the appropriate time scale to select only linear portion of cleavage curve.

? TROUBLESHOOTING

Step 84. Transfer the RFU/s values corresponding to each substrate into a spreadsheet and make the substrate specificity profile for caspase-3 (column diagram or heat-map).

Preparing assay for entire HyCoSuL screening (P2 sub-library) TIMING: 2-3 hours per one screening

Step 85. Once you determine the optimal caspase concentration, perform the whole HyCoSuL screening. Start from P2 sub-library.

Step 86. If your sub-library contains up to 96 substrates you can do the whole sub-library screening on one 96-well plate. If you library is larger than 96 substrates, you have to do the assay on two (or more) plates, so make sure that the second (and the next one) plate contains several "good" substrates from first plate (control) thus you can combine all the results. Regardless of the size of the library (≤ 96 or > 96 substrates) each screening must be performed at least three times, and the result must be present as an average value (see Step 93).

Step 87. Vortex each substrate and by using micropipette place 1μL of it onto 96-well plate (see the diagram). Such prepared plate can be stored at room temperature for 2-3 hours (avoid direct exposure to light).

Step 88. Incubate caspase-3 in 10mL of assay buffer for 15 min in 37°C water bath (or in the incubator). Use the caspase concentration that you determined in the initial screening (natural library).

Step 89. Pour the assay buffer with caspase-3 onto reagent reservoir.

Step 90. Using eight-channel pipette transfer 99μL of the caspase-containing buffer into wells containing 1μL of substrate library.

Step 91. Shake the plate in the plate reader (spectrofluorimeter) for 10 seconds. Read the fluorescence over the time for 30 min (every 15 or 30 seconds). The ACC wavelength setup is: 355nm (excitation) and 460nm (emission).

? TROUBLESHOOTING

Step 92. Calculate the RFU/s value for each substrate from at least three independent screenings.

CRITICAL STEP For each substrate cleavage use the appropriate time scale to select only linear portion of cleavage curve.

? TROUBLESHOOTING

Step 93. Transfer the RFU/s values corresponding to each substrate into a spreadsheet and make the substrate specificity profile for caspase-3 (if the library contains a large number of amino acids the best graphical representation of enzyme specificity is a heat-map). For each substrate the RFU/s must be presented as an average value from at least three independent screenings.

Preparing assay for entire HyCoSuL screening (P3 sub-library) TIMING: 2-3 hours per one screening.

Step 94. Screen the HyCoSuL P3 sub-library by repeating the cycle starting from Step 86 going to Step 93.

? TROUBLESHOOTING

Preparing assay for entire HyCoSuL screening (P4 sub-library) TIMING: 2-3 hours per one screening

Step 95. Screen the HyCoSuL P4 sub-library by repeating the cycle starting from Step 86 going to Step 93.

? TROUBLESHOOTING

Preparing assay for entire HyCoSuL screening for other caspases TIMING: 6-8 hours per one screening per one enzyme

Step 96. Screen the whole HyCoSuL P4-P2 libraries toward other caspases and legumain by repeating the cycle starting from Step 75 going to Step 95.

CRITICAL STEP Incubate each enzyme in the optimal assay buffer, thus remember that some caspases (-8, -9, and -10) required sodium citrate in order to become active.

? TROUBLESHOOTING

Note that for many enzymes, the same concentration can be used for screening of each sub-library, however there may be some exceptions, thus the initial screening should be performed with all three "natural" sublibraries.

Step 97. Transfer all the data into a spreadsheet and make a heat-maps of caspase and legumain specificity profiles.

Select the optimal sequences for individual substrate synthesis and further kinetic assays **TIMING:** 1-2 hours (selection) + 3-4 days (substrate synthesis and purification) + 2-4 hours (determination of kinetic parameters of a substrate)

Step 98. Select the optimal sequence (most selective and/or most active) for the caspase of interest.

? TROUBLESHOOTING

Step 99. Synthesize the optimal substrate Ac-P4-P3-P2-Asp-ACC with selected natural/unnatural amino acids on solid phase using standard protocol [59,60].

Step 100. Determine the kinetic parameters (k_{cat}, K_M, k_{cat}/K_M) for selected substrate toward enzymes of interest. The detailed protocol for this step can be found elsewhere [59].

Synthesis of the most selective activity-based probe (biotin-6-ahx-*D*Tyr-Tic-Ser-Asp-AOMK)

BLOCK A (biotin-6-ahx-*D*-Tyr(*t*Bu)-*L*-Tic-*L*-Ser(*t*Bu)-COOH) **TIMING:** 2-3 days

Step 101. In a 25mL glass cartridge for solid phase weigh 200mg of dry 2-chlorotrityl chloride resin (1.3mmol/g, 1eq, 0.154mmol) and swell it in 5mL of dry DCM for 30 min.

Step 102. In a separate 15mL falcon tube weigh 180 mg of Fmoc-*L*-Ser(*t*Bu)-OH (0.47mmol, 3eq) and suspend it into minimal amount of dry DCM. Then add 780µL of DIPEA (0.71mmol, 4.5eq) and mix it on vortex mixer until the amino acid fully dissolves. Add DCM if needed.

CRITICAL STEP The amino acid will not dissolve in DCM until you add DIPEA. Thus, provide a minimal amount of DCM before you add this amine.

Step 103. Immediately after you dissolve Fmoc-*L*-Ser(*t*Bu)-OH, filter DCM from resin, and pour the amino acid mixture onto resin.

CRITICAL STEP It is very important to dissolve the first amino acid in a minimal amount of DCM, to increase the reagent concentrations in mixture and allow for the high-yield coupling. However, the mixture must be diluted enough to allow easy mixing.

PAUSE POINT Turn the shaker on, and gently shake the reaction vessel for 3 hours. Check every 15 min if the mixture is being mixed well. Add DCM if needed.

Step 104. Remove the mixture from the resin by vacuum filtration and wash the resin three times with DMF (2-4 mL per wash).

Step 105. Remove the Fmoc-protecting group from Fmoc-*L*-Ser(tBu)-resin with 20% piperidine in DMF (see Step 3).

Step 106. Wash the resin six times with DMF (2-4mL per wash) to remove all piperidine and side reaction products.

Step 107. Perform a ninhydrin test.

? TROUBLESHOOTING

Step 108. In a 15mL falcon tube weigh 154 mg of Fmoc-*L*-Tic-OH (0.39mmol, 2.5eq), 148 mg of HATU (0.39 mmol, 2.5eq) and dissolve it in a minimal amount of DMF. Next, add 51 μL of 2,4,6-trimethylcollidine (0.39 mmol, 2.5eq) and pre-activate this mixture for 1 min by gentle stirring.

Step 109. Pour this mixture onto resin and shake gently until all liquid and resin are mixed well. Add DMF if needed.

PAUSE POINT Turn the shaker on, and gently shake the reaction vessel for 3 hours.

Step 110. Remove the mixture from the resin by vacuum filtration and wash the resin three times with DMF (2-4 mL per wash).

Step 111. Perform a ninhydrin test

? TROUBLESHOOTING

Step 112. Remove the Fmoc-protecting group from Fmoc-*L*-Tic-*L*-Ser(*t*Bu)-resin with 20% piperidine in DMF (see Step 3).

Step 113. Wash the resin six times with DMF (2-4mL per wash) to remove all piperidine and side reaction products.

Step 114. Perform a ninhydrin test.

? TROUBLESHOOTING

Step 115. Using the same procedure (steps 108-114) attach the following amino acids to the resin: (1) Fmoc-*D*-Tyr(tBu)-OH (179mg, 0.39mmol, 2.5eq), and (2) Fmoc-6-ahx-OH (138mg, 0.39mmol, 2.5eq) to obtain H_2N-6-ahx-*D*-Tyr(*t*Bu)-*L*-Tic-*L*-Ser(*t*Bu)-resin.

Step 116. In a 15mL falcon tube weigh 95mg of biotin (0.39mmol, 2.5eq) and 148g of HATU (0.39mmol, 2.5eq). Dissolve it in a minimal amount of DMF:DMSO mixture (1:1 v/v). Then, add

51μL of 2,4,6-trimethylcollidine (0.39 mmol, 2.5eq) and pre-activate this mixture for 1 min by gentle stirring.

Step 117. Pour this mixture onto resin and shake gently until all liquid and resin are mixed well. Add DMF:DMSO (1:1 v/v) mixture if needed.

CRITICAL STEP Since the biotin is poorly dissolved in DMF, an addition of DMSO is needed.

PAUSE POINT Turn the shaker on, and gently shake the reaction vessel for 3 hours.

Step 118. Remove the mixture from the resin by vacuum filtration and wash the resin three times with DMF (2-4 mL per wash).

Step 119. Perform a ninhydrin test

? TROUBLESHOOTING

Step 120. Wash the resin three times with DCM (2-4 mL per wash) and three times with MeOH (2-4 mL per wash).

Step 121. Dry the resin in the desiccator over P_2O_5 overnight. Replace P_2O_5 if needed.

CRITICAL STEP The resin must be dry in order to cleave the final product from it.

Step 122. Cleave the peptide from the resin. Prepare 10mL of cleavage solution (8mL of DCM, 1mL of TFE, 1mL of AcOH), pour 5mL of it onto resin and shake the cartridge one per 10 min for 45 min.

Step 123. Filter and collect the mixture into 100mL round bottom flask, and wash the resin with the remaining 5mL of cleavage solution. Vacuum filter it and collect into the same flask.

Step 124. Remove the cleavage mixture on rotary evaporator under reduced pressure until the white/yellow oil forms.

Step 125. Dissolve the oil in 20mL of the water/acetonitrile mixture (1:1), freeze at -80°C and lyophilize to obtain biotin-6-ahx-*D*-Tyr(*t*Bu)-*L*-Tic-*L*-Ser(*t*Bu)-COOH as a white powder. Overall yield (based on 2-chlorotrityl chloride resin loading capacity > 70%), peptide purity > 90%.

BLOCK B (Boc-*L*-Asp(Bzl)-AOMK) **TIMING:** 2 days

Step 126. In a 100mL round bottom flask weigh 1.16 g of Boc-*L*-Asp(Bzl)-OH (5mmol, 1eq) and suspend it in 25mL of anhydrous THF (final Asp concentration is 0.2M). Close the flaks with a rubber stockpot and stir the mixture in an ice/acetone bath at -10°C for 10 min.

Step 127. Using a syringe with needle add 690μL of 4-methylmorpholine (6.25mmol, 1.25eq) and 750μL of isobutyl chloroformate (5.75 mmol, 1.15eq) to the flask, through rubber stockpot.

PAUSE POINT Carry out the reaction at -10°C for 45 min to obtain a mixture of mixed anhydrides.

Step 128. In a separate round bottom flask (with clear joints) placed in an ice/acetone bath prepare ethereal solution of diazomethane (15-25 mmol) according to the Sigma Technical Bulletin AL-180. Place a stir bar in the flask.

CRITICAL POINT Diazomethane is extremely explosive, thus a special precautions are needed. Ethereal solution of diazomethane is safe when stored at low temperature, thus the generation of this mixture can be outsourced. It must be stressed here that in some countries the generation and use of diazomethane is strictly regulated and prior authorization may be required. However, the final product of this synthesis (Boc-*L*-Asp(Bzl)-AOMK, Step 138) is safe in transport and use, thus its synthesis may be outsourced to labs that generate diazomethane routinely.

Step 129. Transfer dropwise the solution of mixed anhydrides into the diazomethane solution (within 10 min) and carry on the reaction for 10 more minutes.

Step 130. Remove the ice bath and carry out the reaction at room temperature until it is completed (Boc-L-Asp(Bzl)-CH_2N_2 is formed; usually 2-3hours). Monitor the reaction progress on analytical HPLC (254nm) (see Equipment setup).

CRITICAL POINT Boc-L-Asp(Bzl)-CH_2N_2 produces a very intense peak at 254nm (-CH_2N_2 group) while analyzing by HPLC.

Step 131. Prepare a 10mL solution of HBr (30% in AcOH) and water (1:2, v/v) and add it dropwise into the flask with Boc-L-Asp(Bzl)-CH_2N_2 over a period of 10min. Carry on the reaction for 10 more minutes and check the progress on analytical HPLC (254nm) (see Equipment setup).

? TROUBLESHOOTING

Step 132. Add 150mL of ethyl acetate into 500mL separatory funnel and transfer the reaction mixture to this funnel.

Step 133. Extract the mixture with 100mL of water (once), 100mL of saturated aqueous $NaHCO_3$ (twice) and 100mL of brine (twice).

CRITICAL POINT The extraction with saturated aqueous $NaHCO_3$ must be performed very gently, as the unreacted HBr and AcOH from Step 128 produce a very high pressure (CO_2 release) in contact with base.

Step 134. After the last extraction, collect the organic layer into 250mL Erlenmeyer flask, dry it over $MgSO_4$, filter the organic layer into 100mL round bottom flask using Buchner funnel and remove the ethyl acetate on rotary evaporator under reduced pressure to obtain a crude product (Boc-*L*-Asp(Bzl)-CH_2Br) as a white to pale yellow oil. Dry the oil in desiccator over P_2O_5 overnight to obtain white to pale yellow powder. Overall yield > 85%, product purity > 90%.

PAUSE POINT Since Boc-*L*-Asp(Bzl)-CH_2Br is a strong electrophile, it is not very stable at room temperature thus once is dry, it should be stored at -80°C or immediately transformed into AOMK analogue.

Step 135. In a 100mL round bottom flask with a stir bar weigh 100 mg of crude Boc-*L*-Asp(Bzl)-CH_2Br (0.25 mmol, 1eq) and dissolve it in a minimal volume of DMF. Add 44 mg of KF (0.75 mmol, 3eq) and 45 mg of 6-dimethylbenzoic acid (2,6-DMBA, 0.3 mmol, 1.2eq) and stir the reaction mixture for 30min in inert atmosphere of argon.

Step 136. Dilute the reaction mixture with 100mL of ethyl acetate, transfer it into 250mL separatory funnel and extract it with 80mL of 5% citric acid (twice), 80mL of 5% aqueous $NaHCO_3$ (twice) and 80mL of brine (twice). After each extraction remove the bottom aqueous phase.

Step 137. Collect the organic fraction in 250mL Erlenmeyer flask and dry it over $MgSO_4$.

Step 138. Filter the dried organic fraction into 100mL round bottom flask using Buchner funnel and remove the solvent on rotary evaporator under reduced pressure to obtain a final product (Boc-*L*-Asp(Bzl)-AOMK) as a yellow oil. Dry the oil in desiccator over P_2O_5 overnight to obtain white to pale yellow powder. Overall yield > 95%, product purity > 90%.

PAUSE POINT Aliquot the crude product into tubes and store them at -20°C either -80°C until use.

BLOCK C (biotin-6-ahx-*D*-Tyr-*L*-Tic-*L*-Ser-*L*-Asp-AOMK) TIMING: 3 days

Step 139. Weigh 100mg of crude Boc-*L*-Asp(Bzl)-AOMK (0.213 mmol, 1eq) into 10mL round-bottom flask and add 5mL of 50% solution of TFA in DCM. Stir the reaction for 30 min to de-protect Boc-group and obtain TFA x H_2N-*L*-Asp(Bzl)-AOMK.

Step 140. Remove the cleavage mixture on rotary evaporator under reduced pressure.

Step 141. Add 5mL of DCM into the tube, dissolve the crude product using sonicator and remove the DCM (and remaining TFA) under reduced pressure. Repeat this step is some TFA has still left. Dry the crude product in desiccator over P_2O_5 over 30 min.

CRITICAL POINT It is important to remove all the TFA from the product to ensure the high yield of coupling in the next step.

Step 142. Remove the 10mL round bottom-flask from desiccator and place it on magnetic stirrer.

Step 143. In a separate 15mL falcon tube weigh 225 mg of biotin-6-ahx-*D*-Tyr(*t*Bu)-*L*-Tic-*L*-Ser(*t*Bu)-OH (0.256 mmol, 1.2eq) and 97 mg of HATU (0.256 mmol, 1.2eq). Dissolve them in 1mL of DMF and add 67μL of 2,4,6-trimethylcollidine (0.512 mmol, 2.4eq). Pre-incubate the mixture for 1 min and transfer it into 10mL round bottom-flask containing TFA x H_2N-*L*-Asp(Bzl)-AOMK. Turn on the magnetic stirrer.

Step 144. Monitor the reaction progress on analytical HPLC (220nm) until is completed (the peptide peak shifts to the right in water/acetonitrile gradient) (see Equipment setup). The yield of this reaction (calculated based on HPLC chromatogram) is usually >80%.

? TROUBLESHOOTING

CRITICAL POINT It is important to keep the reaction pH at around 8 to ensure the high yield of coupling. If there is some TFA left (see Step 138) the pH drops down, thus additional amount of 2,4,6-trimethylcollidine is needed to capture the free acid.

Step 145. Dilute the reaction mixture to the 5mL with DMSO, filter it through 0.2μm syringe filter, and purify the product on semi-preparative HPLC (1mL of mixture per purification; 220nm, see Equipment setup) and lyophilize it to obtain white powder (biotin-6-ahx-*D*-Tyr(*t*Bu)-*L*-Tic-*L*-Ser(*t*Bu)-*L*-Asp(Bzl)-AOMK).

Step 146. Place the product into 10mL round-bottom flask (10-30mg – depending on the yield of synthesis) and add 5mL of 50% solution of TFA in DCM. Stir the reaction for 30 min to de-protect *t*Bu-groups from Tyr and Ser to obtain biotin-6-ahx-*D*-Tyr-*L*-Tic-*L*-Ser-*L*-Asp(Bzl)-AOMK.

Step 147. Remove the volatiles on rotary evaporator under reduced pressure.

Step 148. Add 5mL of DCM into the tube, dissolve the crude product using sonicator and remove the DCM (and remaining TFA) under reduced pressure. Repeat this step if some TFA is still left. Dry the crude product in desiccator over P_2O_5 over 30 min.

Step 149. Remove the round-bottom flask from desiccator and dilute the crude product in 1mL of DMF.

Step 150. Replace all the air in the flask with inert gas (argon or nitrogen), add around 10mg of palladium on carbon (Pd/C) to the flask and bubble the hydrogen through the reaction mixture. Monitor the benzyl group de-protection on analytical HPLC (220nm, see Equipment setup). The yield of the Bzl group de-protection should be at least > 80%.

? TROUBLESHOOTING

CRITICAL POINT Pd/C tends to settle down over the reaction time, thus ensure that the mixture is bubbling well.

Step 151. Once the reaction is complete, transfer the mixture into the syringe, filter the mixture through 0.2μm syringe filter, and purify the final product on semi-preparative HPLC in water/acetonitrile gradient (220nm, see Equipment setup).

Step 152. Lyophilize the final product to obtain white powder. Weigh it and dissolve in peptide grade DMSO to the final concentration of 10mM.

PAUSE POINT The final product (activity-based probe) can be stored (either as a powder or DMSO solution) at -80°C for at least couple of years

TIMING:

Steps 1-7: 3 h + 24 h coupling reaction

Steps 8-10: 1 h + 24 h coupling reaction

Steps 11-15: 1.5 h

Steps 16-17: 0.5 h + 24 h coupling reaction

Steps 18-20: 0.5 h + 24 h coupling reaction

Steps 21-26: 2 h + overnight drying time (12-16 h)

Steps 27-33: 5 h + 3 h coupling reaction

Steps 34-39: 2.5 h + 3 h coupling reaction

Steps 40-46: 2.5 h + 3 h coupling reaction

Steps 47-53: 3.5 h

Steps 54-58: 1 h + overnight drying time (12-16 h)

Steps 59-69: 10-12 h + lyophilization time (1-2 days)

Steps 70-71: 4-6 h

Step 72 (= steps 1-71) 9-11 days

Step 73 (= steps 1-71) 9-11 days

Steps 74-84: 4-6 h

Steps 85-93: 2-3 h

Step 94 (= steps 86-93): 2-3 h

Step 95 (= steps 86-93): 2-3 h

Steps 96-97: 6-8 h

Step 98: 1-2 h

Step 99: 3-4 days (this time includes drying the resin /overnight/ and lyophilization)

Step 100: 2-4 h

Steps 101-107: 5 h

Steps 108-114: 5 h

Step 115 (= steps 108-114, twice): 2 x 5 h

Steps 116-121: 4 h + overnight drying time (12-16 h)

Steps 122-125: 1.5-2 h + lyophilization time (12-24 h)

Steps 126-127: 1.5-2 h

Step 128: 4-6 h (if synthesized in house), 0.5 h (if purchased)

Steps 129-131: 1.5 h + overnight drying time 12-16 h)

Steps 132-135: 2.5 h + overnight drying time 12-16 h)

Steps 136-138: 1.5 h

Steps 139-141: 1-3 h

Step 142: 1-5 h (depends on the number of purification cycles) + lyophilization time (12-24 h)

Steps 143-145: 1.5-2 h

Steps 146-147: 2-8 h (depends on the kinetic of Asp-Bzl de-protection)

Steps 148-149: 1-1.5 h + lyophilization time (12-24 h)

? TROUBLESHOOTING

Table 1 Troubleshooting table

Step	Problem	Possible reason	Possible solutions
12	Some resin beads are dark blue	Coupling of Fmoc-ACC-OH is incomplete	A) Repeat Fmoc-ACC-OH coupling using the same amount of reagents as in step 9. B) Acetylate the remaining amine groups on resin in order to avoid side reactions in later steps
22	Some resin beads are dark blue	Coupling of Fmoc-Asp(*t*Bu)-OH is incomplete	Acetylate the remaining H_2N-ACC using the protocol described in [44]. For acetylation of H_2N-ACC a special procedure is required [44].
34, 111	Some resin beads are still blue	Coupling of some Fmoc-amino acids is incomplete	Repeat the ninhydrin test for all of the wells, select these with blue beads, and repeat the coupling with ½ of the reagents from first coupling. To wells where coupling was 100% add just DMF.
37, 44, 51, 107, 114	Some resin beads are still yellow/colorless	Removal of the Fmoc group is incomplete	Repeat the Fmoc de-protection with 20% piperidine in DMF for 5 and 15 min.
		The N-terminal amino acid is proline or one of its derivative	Perform the acetaldehyde/chloranil test in order to detect free proline (or its derivative)
41, 48	Some resin beads are still blue	Coupling of isokinetic mixture is incomplete	Repeat the whole sub-library synthesis starting from Step 1. Note: we have never observed incomplete coupling of the isokinetic mixture as we use 5x excess of reagents.
55	Some resin beads are still blue	Acetylation of the N-terminus of library is incomplete	Repeat the acetylation of N-terminus of the library using ½ of the reagents from first acetylation. Reaction time 15 min.
68	Some substrates are not fully dissolved	Substrates are insoluble in this amount of ACN/water mixture	Try to dissolve substrates by sonication. If this will not help, add 1mL of 100% ACN into the tube.

69	Some of the substrates turn into oil after lyophilization	Lyophilization was not complete (water still present in samples)	Dissolve selected substrates in ACN/water mixture, freeze, and lyophilize again.
82, 91, 94, 95, 96	No substrate activity is detected (even for these substrates which are positive controls)	Enzyme is inactive	Repeat the assay with the active enzyme. Check the enzyme activity using one of the commercial/reference substrates.
	Slopes of all substrate activities (good and bad) decreases over the time	Enzyme inactivation due to prolonged activation	Decrease the time of enzyme pre-incubation; however this might decrease the rate of the substrate cleavages as not 100% of the enzyme is active
		Enzyme inactivation due to monomer dissociation (mainly for initiator caspases)	Make sure that you use the optimal concentration of sodium citrate.
83, 92, 94, 95, 96	Slope of the substrate cleavage is approx. 0 but the overall fluorescence is much higher than the background	Substrate was completely hydrolyzed by the enzyme within the first several seconds of the assay	Decrease the enzyme concentration and repeat the screening
	Only one (or several) substrate from the entire library are cleaved by the enzyme	Enzyme concentration is too low to cleave non-optimal substrates	Increase the enzyme concentration and repeat the screening. If the results will be similar, it means that enzyme has a very narrow specificity in this particular pocket/subsite.
98	There is no clear answer which sequence is the most active/selective	Several amino acids can be equally recognized by the enzyme at certain positions	Synthesize several substrates with most promising sequences and measure their kinetic parameters. Detailed kinetic analysis will provide the information about possible subsites cooperativity and will indicate the champion substrates.
131	Reaction is not 100% complete	10mL of HBr (30% in AcOH) and water (1:2, v/v) was not sufficient to transform diazomethylketone into bromomethylketone.	Prepare an additional 5mL of HBr (30% in AcOH) and water (1:2, v/v), and add it to the reaction flask dropwise over 5 min. Carry on the reaction for 5 more minutes and check the progress on analytical HPLC (254nm).
144	Reaction progress stops at some point (HPLC shows that both substrates are still present in the reaction mixture)	pH is below 8 as there is some TFA left over from Step 138	Adjust the reaction pH to the optimal value by adding base (2,4,6-trimethylcollidine)
150	Reaction progress stops at some point (HPLC shows that Asp(Bzl) is not fully de-protected)	Pd/C is quenched by one of the reaction product/substrate thus is loses its catalytic activity	Replace all the hydrogen in flask with inner gas, add additional portion of Pd/C and continue the bubbling of hydrogen through the reaction mixture
		Pd/C has been contaminated or has partially decomposed during storage	Use the Pd/C from freshly opened bottle

4. ANTICIPATED RESULTS

Library synthesis

The HyCoSuL is a combinatorial library that is used for the protease substrate specificities screening without any purification. This protocol provides the detailed information for the

synthesis of three combinatorial sub-libraries (P4: Ac-Aaa-Mix-Mix-Asp-ACC; P3: Ac-Mix-Aaa-Mix-Asp-ACC; P2: Ac-Mix-Mix-Aaa-Asp-ACC) that in each of P4-P2 positions contain 19 natural amino acids (without cysteine) and 110 unnatural amino acids (Figure 1).

Library screening

Using this library we determined the detailed substrate preferences of six human apoptotic caspases and legumain at P4-P2 pockets. The in depth analysis of these specificity profiles allowed us to design new, selective caspase and legumain substrates, as well as the potent and selective biotin-labeled activity based probe for legumain (Figure 3).

Activity-based probe

This protocol provides a detailed procedure for the synthesis and purification of biotin-labeled irreversible activity based probe containing aspartic acid at P1 position, thus it is useful for caspases and legumain investigation (Figure 6). This synthesis can be also easily adapted to any endoprotease that can be inhibited by AOMK warhead. The use of Pd/C catalyst in Bzl group de-protection make this protocol suitable for a wide range of amino acids (Glu(Bzl), Arg(Cbz)$_2$, Lys(Cbz), and more).

ACKNOWLEDGMENTS:

This work was supported by Ministry of Science and Higher Education in Poland (grant Iuventus Plus IP2012 040172 to MP), National Science Centre in Poland (grant 2014/14/M/ST5/00619 to MD) and National Institutes of Health in USA (grant R01GM099040 to GSS). This project has received founding from the European Union`s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 661187.

AUTHOR CONTRIBUTION:

M.P., G.S.S. and M.D. developed the protocol, designed research, interpreted data and wrote the protocol. M.P. carried out the experiments in the protocol.

COMPETING FINANCIAL INTERESTS:

The authors declare no competing financial interests.

5. REFERENCES

1 Drag, M. & Salvesen, G. S. Emerging principles in protease-based drug discovery. *Nat Rev Drug Discov* **9**, 690-701, doi:10.1038/nrd3053 (2010).

2 Turk, B. Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov* **5**, 785-799, doi:10.1038/nrd2092 (2006).

3 Rawlings, N. D., Barrett, A. J. & Finn, R. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* **44**, D343-350, doi:10.1093/nar/gkv1118 (2016).

4 Schechter, I. & Berger, A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157-162 (1967).

5 Lopez-Otin, C. & Overall, C. M. Protease degradomics: a new challenge for proteomics. *Nat Rev Mol Cell Biol* **3**, 509-519, doi:10.1038/nrm858 (2002).

6 Grootjans, S. *et al.* A real-time fluorometric method for the simultaneous detection of cell death type and rate. *Nat Protoc* **11**, 1444-1454, doi:10.1038/nprot.2016.085 (2016).

7 Poreba, M. & Drag, M. Current strategies for probing substrate specificity of proteases. *Curr Med Chem* **17**, 3968-3995 (2010).

8 Kasperkiewicz, P., Poreba, M., Groborz, K. & Drag, M. Emerging challenges in the design of selective substrates, inhibitors and activity-based probes for indistinguishable proteases. *FEBS J*, doi:10.1111/febs.14001 (2017).

9 Diamond, S. L. Methods for mapping protease specificity. *Curr Opin Chem Biol* **11**, 46-51, doi:10.1016/j.cbpa.2006.11.021 (2007).

10 Matthews, D. J. & Wells, J. A. Substrate phage: selection of protease substrates by monovalent phage display. *Science* **260**, 1113-1117 (1993).

11 Thornberry, N. A. *et al.* A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. *J Biol Chem* **272**, 17907-17911 (1997).

12 Harris, J. L. *et al.* Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc Natl Acad Sci U S A* **97**, 7754-7759, doi:10.1073/pnas.140132697 (2000).

13 McStay, G. P., Salvesen, G. S. & Green, D. R. Overlapping cleavage motif selectivity of caspases: implications for analysis of apoptotic pathways. *Cell Death Differ* **15**, 322-331, doi:10.1038/sj.cdd.4402260 (2008).

14 Choe, Y. *et al.* Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. *J Biol Chem* **281**, 12824-12832, doi:10.1074/jbc.M513331200 (2006).

15 Drag, M. *et al.* Positional-scanning fluorigenic substrate libraries reveal unexpected specificity determinants of DUBs (deubiquitinating enzymes). *Biochem J* **415**, 367-375, doi:10.1042/BJ20080779 (2008).

16 Withana, N. P. *et al.* Labeling of active proteases in fresh-frozen tissues by topical application of quenched activity-based probes. *Nat Protoc* **11**, 184-191, doi:10.1038/nprot.2016.004 (2016).

17 Oresic Bender, K. *et al.* Design of a highly selective quenched activity-based probe and its application in dual color imaging studies of cathepsin S activity localization. *J Am Chem Soc* **137**, 4771-4777, doi:10.1021/jacs.5b00315 (2015).

18 Kasperkiewicz, P. *et al.* Design of ultrasensitive probes for human neutrophil elastase through hybrid combinatorial substrate library profiling. *Proc Natl Acad Sci U S A* **111**, 2518-2523, doi:10.1073/pnas.1318548111 (2014).

19 Kasperkiewicz, P., Gajda, A. D. & Drag, M. Current and prospective applications of non-proteinogenic amino acids in profiling of proteases substrate specificity. *Biol Chem* **393**, 843-851, doi:10.1515/hsz-2012-0167 (2012).

20 Rut, W. *et al.* Recent advances and concepts in substrate specificity determination of proteases using tailored libraries of fluorogenic substrates with unnatural amino acids. *Biol Chem* **396**, 329-337, doi:10.1515/hsz-2014-0315 (2015).

21 Rano, T. A. *et al.* A combinatorial approach for determining protease specificities: application to interleukin-1beta converting enzyme (ICE). *Chem Biol* **4**, 149-155 (1997).

22 Poreba, M. *et al.* Unnatural amino acids increase sensitivity and provide for the design of highly selective caspase substrates. *Cell Death Differ* **21**, 1482-1492, doi:10.1038/cdd.2014.64 (2014).

23 Ostresh, J. M., Winkle, J. H., Hamashin, V. T. & Houghten, R. A. Peptide libraries: determination of relative reaction rates of protected amino acids in competitive couplings. *Biopolymers* **34**, 1681-1689, doi:10.1002/bip.360341212 (1994).

24 Poreba, M. *et al.* Counter Selection Substrate Library Strategy for Developing Specific Protease Substrates and Probes. *Cell Chem Biol* **23**, 1023-1035, doi:10.1016/j.chembiol.2016.05.020 (2016).

25 Rut, W. *et al.* Extended substrate specificity and first potent irreversible inhibitor/activity-based probe design for Zika virus NS2B-NS3 protease. *Antiviral Res* **139**, 88-94, doi:10.1016/j.antiviral.2016.12.018 (2017).

26 Kasperkiewicz, P. *et al.* Design of a Selective Substrate and Activity Based Probe for Human Neutrophil Serine Protease 4. *PLoS One* **10**, e0132818, doi:10.1371/journal.pone.0132818 (2015).

27 Lentz, C. S. *et al.* Design of Selective Substrates and Activity-Based Probes for Hydrolase Important for Pathogenesis 1 (HIP1) from Mycobacterium tuberculosis. *ACS Infect Dis* **2**, 807-815, doi:10.1021/acsinfecdis.6b00092 (2016).

28 Lechtenberg, B. C., Kasperkiewicz, P., Robinson, H., Drag, M. & Riedl, S. J. The elastase-PK101 structure: mechanism of an ultrasensitive activity-based probe revealed. *ACS Chem Biol* **10**, 945-951, doi:10.1021/cb500909n (2015).

29 Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317 (1985).

30 Legowska, M. *et al.* Ultrasensitive internally quenched substrates of human cathepsin L. *Anal Biochem* **466**, 30-37, doi:10.1016/j.ab.2014.08.010 (2014).

31 Gosalia, D. N., Denney, W. S., Salisbury, C. M., Ellman, J. A. & Diamond, S. L. Functional phenotyping of human plasma using a 361-fluorogenic substrate biosensing microarray. *Biotechnol Bioeng* **94**, 1099-1110, doi:10.1002/bit.20927 (2006).

32 Vizovisek, M., Vidmar, R., Fonovic, M. & Turk, B. Current trends and challenges in proteomic identification of protease substrates. *Biochimie* **122**, 77-87, doi:10.1016/j.biochi.2015.10.017 (2016).

33 Schilling, O. & Overall, C. M. Proteomic discovery of protease substrates. *Curr Opin Chem Biol* **11**, 36-45, doi:10.1016/j.cbpa.2006.11.037 (2007).

34 Staes, A. *et al.* Selecting protein N-terminal peptides by combined fractional diagonal chromatography. *Nat Protoc* **6**, 1130-1141, doi:10.1038/nprot.2011.355 (2011).

35 Korkmaz, B., Horwitz, M. S., Jenne, D. E. & Gauthier, F. Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases. *Pharmacol Rev* **62**, 726-759, doi:10.1124/pr.110.002733 (2010).

36 Janicke, R. U., Sohn, D., Totzke, G. & Schulze-Osthoff, K. Caspase-10 in mouse or not? *Science* **312**, 1874, doi:10.1126/science.312.5782.1874a (2006).

37 Berger, A. B., Sexton, K. B. & Bogyo, M. Commonly used caspase inhibitors designed based on substrate specificity profiles lack selectivity. *Cell Res* **16**, 961-963, doi:10.1038/sj.cr.7310112 (2006).

38 Pereira, N. A. & Song, Z. Some commonly used caspase substrates and inhibitors lack the specificity required to monitor individual caspase activity. *Biochem Biophys Res Commun* **377**, 873-877, doi:10.1016/j.bbrc.2008.10.101 (2008).

39 Drag, M., Bogyo, M., Ellman, J. A. & Salvesen, G. S. Aminopeptidase fingerprints, an integrated approach for identification of good substrates and optimal inhibitors. *J Biol Chem* **285**, 3310-3318, doi:10.1074/jbc.M109.060418 (2010).

40 Byzia, A., Szeffler, A., Kalinowski, L. & Drag, M. Activity profiling of aminopeptidases in cell lysates using a fluorogenic substrate library. *Biochimie* **122**, 31-37, doi:10.1016/j.biochi.2015.09.035 (2016).

41 Poreba, M. *et al.* Unnatural amino acids increase activity and specificity of synthetic substrates for human and malarial cathepsin C. *Amino Acids* **46**, 931-943, doi:10.1007/s00726-013-1654-2 (2014).

42 Zeiler, E. *et al.* Structural and functional insights into caseinolytic proteases reveal an unprecedented regulation principle of their catalytic triad. *Proc Natl Acad Sci U S A* **110**, 11302-11307, doi:10.1073/pnas.1219125110 (2013).

43 Gersch, M. *et al.* Barrel-shaped ClpP Proteases Display Attenuated Cleavage Specificities. *ACS Chem Biol* **11**, 389-399, doi:10.1021/acschembio.5b00757 (2016).

44 Maly, D. J. *et al.* Expedient solid-phase synthesis of fluorogenic protease substrates using the 7-amino-4-carbamoylmethylcoumarin (ACC) fluorophore. *J Org Chem* **67**, 910-915 (2002).

45 Patterson, A. W., Wood, W. J. & 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, doi:10.1038/nprot.2007.28 (2007).

46 Lu, Y. & Freeland, S. On the evolution of the standard amino-acid alphabet. *Genome Biol* **7**, 102, doi:10.1186/gb-2006-7-1-102 (2006).

47 Isidro-Llobet, A., Alvarez, M. & Albericio, F. Amino acid-protecting groups. *Chem Rev* **109**, 2455-2504, doi:10.1021/cr800323s (2009).

48 Ng, N. M., Pike, R. N. & Boyd, S. E. Subsite cooperativity in protease specificity. *Biol Chem* **390**, 401-407, doi:10.1515/BC.2009.065 (2009).

49 Kato, D. *et al.* Activity-based probes that target diverse cysteine protease families. *Nat Chem Biol* **1**, 33-38, doi:10.1038/nchembio707 (2005).

50 Winiarski, L., Oleksyszyn, J. & Sienczyk, M. Human neutrophil elastase phosphonic inhibitors with improved potency of action. *J Med Chem* **55**, 6541-6553, doi:10.1021/jm300599x (2012).

51 Edgington, L. E. *et al.* An optimized activity-based probe for the study of caspase-6 activation. *Chem Biol* **19**, 340-352, doi:10.1016/j.chembiol.2011.12.021 (2012).

52 Powers, J. C., Asgian, J. L., Ekici, O. D. & James, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev* **102**, 4639-4750 (2002).

53 Sanman, L. E. & Bogyo, M. Activity-based profiling of proteases. *Annu Rev Biochem* **83**, 249-273, doi:10.1146/annurev-biochem-060713-035352 (2014).

54 Stennicke, H. R. & Salvesen, G. S. Caspases: preparation and characterization. *Methods* **17**, 313-319, doi:10.1006/meth.1999.0745 (1999).

55 Wachmann, K. *et al.* Activation and specificity of human caspase-10. *Biochemistry* **49**, 8307-8315, doi:10.1021/bi100968m (2010).

56 Boatright, K. M., Deis, C., Denault, J. B., Sutherlin, D. P. & Salvesen, G. S. Activation of caspases-8 and -10 by FLIP(L). *Biochem J* **382**, 651-657, doi:10.1042/BJ20040809 (2004).

57 Boatright, K. M. *et al.* A unified model for apical caspase activation. *Mol Cell* **11**, 529-541 (2003).

58 Sainlos, M. & Imperiali, B. Tools for investigating peptide-protein interactions: peptide incorporation of environment-sensitive fluorophores through SPPS-based 'building block' approach. *Nat Protoc* **2**, 3210-3218, doi:10.1038/nprot.2007.443 (2007).

59 Poreba, M., Szalek, A., Kasperkiewicz, P. & Drag, M. Positional scanning substrate combinatorial library (PS-SCL) approach to define caspase substrate specificity. *Methods Mol Biol* **1133**, 41-59, doi:10.1007/978-1-4939-0357-3_2 (2014).

60 Schneider, E. L. & Craik, C. S. Positional scanning synthetic combinatorial libraries for substrate profiling. *Methods Mol Biol* **539**, 59-78, doi:10.1007/978-1-60327-003-8_4 (2009).