Substrate Protection in Controlled Enzymatic Transformation of Peptides and Proteins

Yan Zhao*[a]

[a] Prof. Yan Zhao
Department of Chemistry
Iowa State University
Ames, IA 50011-3111, U.S.A.
Phone: (+1) 515-294-5845
E-mail: zhaoy@iastate.edu

Abstract: Proteins are involved in practically every single biological process. The many enzymes involved in their synthesis, cleavage, and posttranslational modification (PTM) carry out highly specific tasks with no usage of protecting groups. Yet, the chemists' strategy of protection/deprotection potentially can be highly useful when a specific biochemical reaction catalyzed by a broad-specificity enzyme needs to be inhibited, during infection of cells by enveloped viruses, in the invasion and spread of cancer cells, and upon mechanistic investigation of signal-transduction pathways, for example. Doing so requires highly specific binding of peptide substrates in aqueous solution with biologically competitive affinities. Recent development of peptide-imprinted cross-linked micelles allows such protection and affords previously impossible manners of manipulating peptides and proteins in enzymatic transformations.

Introduction

Enzymatic efficiency and selectivity represent the ultimate goals of chemists who seek to develop catalysts for their interested reactions. Indeed, under largely ambient conditions in neutral aqueous solutions, enzymes hydrolyze particular amide bonds, selectively oxidize hydrocarbons, convert nitrogen into ammonia, and perform all kinds of transformations vital to the biological world.

The high selectivity of enzymatic catalysis allows cells to carry out desired biochemical transformations from exceedingly complex mixtures without usage of any protecting groups. Glycosyltransferases and glycosidases, for example, effortlessly make complex glycans and cleave them at specific locations.[1] In contrast, to synthesize even relatively simple glycans, chemists generally have to employ extensive protective/deprotective chemistry to deal with the many hydroxyl groups on the sugar building blocks that have little or no difference in intrinsic reactivity.[2] Only in this year of 2021, synthetic catalysts appeared in the literature that could hydrolyze oligo- and polysaccharides with a reasonable level of selectivity.[3]

Protective groups have been an indispensable tool in modern organic chemistry, not only in the synthesis of biomolecules such as carbohydrates, peptides, and nucleic acids full of degenerate functional groups, but also in total synthesis of almost any complex, multifunctional molecules.[4] Whenever chemists want to perform a chemical reaction that has compatibility issues with existing functional groups in the molecule, a straightforward and often the most reliable method is to protect the incompatible groups prior to the reaction and deprotect them at a suitable stage later on.

It seems, with the abundance of highly selective and even substrate-specific enzymes, protection/deprotection is neither necessary nor useful in biology. However, this is not the case, at least when it comes for researchers to intervene and interrogate certain biological processes.

A good example is in the proteolysis of peptides and proteins. Cancer cells rely on over-expressed proteases during their invasion and spread because of the need to remodel tissues.[5] Since the same proteases may be used by normal cells to perform their cellular functions, traditional protease inhibitors tend to have high toxicity. Many enveloped viruses depend on a critical proteolytic activation step in their cellular infection including coronavirus,[6] HIV-1,[7] and influenza virus A.[7] Selective inhibition of a specific proteolytic reaction instead of all proteolysis is again vital to the antiviral treatment. Antibodies can be used to block proteolytic cleavage sites on proteins[8] but they are expensive and fragile molecules made of polypeptides, which are susceptible to broad-specificity proteases themselves.

One other example is in the post posttranslational modification (PTM) of proteins. Kinases catalyze phosphorylation of proteins, a reaction critical to numerous processes in cell signaling, regulation, and development.[9] However, a vast number of potential phosphorylation sites exist in a cell, ~700,000 by one estimation.[9b] Even if most of these sites are buried and kinases have their own preferred substrates, a cell still has a large number of substrates for a given kinase.[9] Traditional enzymatic inhibition is again facing a problem in this case, because unintended consequences will emerge when a multisubstrate kinase is shut down.[10]

A long-recognized solution to the above problems lies in the selective inhibition of the peptide or protein substrates.[11] If a particular substrate of a protease or kinase can be selectively protected from the enzyme, one would be able to shut down a specific biological reaction with high precision. Such protection can help researchers understand the biological ramifications of the masked biological reaction, useful in mechanistic biology and also potentially in disease treatment.[8] Over the years, a few research groups have reported protection of peptides from chemical or enzymatic reactions, mainly using small-molecule synthetic receptors. The reactions involved include proteolysis,[11-12] acetylation of lysine side chains,[13] tyrosine phosphorylation,[14] and demethylation of methylated lysine side chains.[15]

Molecular Recognition of Peptides

The scarcity of peptide protection in the literature points to a great need in peptide recognition, especially of complex biological

peptides. To protect a peptide sequence from its enzymatic catalyst, one needs a receptor to bind the peptide with high affinity and selectivity in aqueous solution. Supramolecular chemistry in the last several decades largely have stayed in organic solvents where directional noncovalent forces such as hydrogen bonds are effective.[16] Although examples of strong synthetic receptors in aqueous solution exist,[17] they are exceptions rather than rules and a general strategy for effective molecular recognition of complex biological molecules in water is missing.[18] For peptides, a particular challenge is in the distinction of the 20 possible side chains of a peptide, some of which differ minutely. Leucine (L) and isoleucine (I), for example, differ in the position of a single methyl group by one carbon. Glutamic acid (E) has one extra methylene than aspartic acid (D), and tyrosine (Y) has one extra hydroxyl in comparison to phenylalanine (F).

Chemists have developed many scaffolds to build peptide receptors,[19] often focusing on specific residues with good supramolecular handles such as acidic and basic amino acids.[20] Tryptophan (W) and phenylalanine are also popular targets because their aromatic side chains can fit into appropriate macrocycles[21] such as cyclodextrins[22], cucurbiturils,[23] or self-assembled nanocages.[24] Other interesting platforms include molecular tweezers and clips,[25] pseudopeptidic cages[26], and gold nanoparticles that can be functionalized on the surface.[27]

Principles of complementarity and preorganization are the central dogma of supramolecular chemistry.[16, 28] It is impractical, however, to apply them in peptide recognition with a molecularly synthesized receptor. This is because, to bind a guest with multipoint noncovalent interactions, the host generally is larger than the guest and needs to possess a complementary, often concave-shaped binding interface. For a biological peptide with 10 to 20 amino acid residues that have subtly different side chains, a complementary host, if constructed step-by-step, would be too difficult to design and synthesize.

A potential solution to the above problem comes from molecular imprinting, a simple and powerful method to create guest-complementary binding sites in a cross-linked polymer network.[29] The method involves formation of a covalent or noncovalent complex between a template (often the guest molecule itself) and polymerizable functional monomers (FMs) in the presence of a large amount of a cross-linker. Polymerization fixes the FMs around the template molecules in the polymer network. Cleavage of the covalent bonds between the FMs and the templates or, in the case of noncovalent imprinting, washing off the noncovalently trapped templates leaves behind “imprints” or guest-shaped voids in the polymer. The FMs turn into binding groups in the imprinted sites during polymerization and can increase the selectivity and binding affinity for the template molecules during rebinding.

Many molecularly imprinted polymers (MIPs) have been created for peptides since the conception of the technique.[30] One of the earliest examples of noncovalent imprinting used amino acid derivatives as templates.[31] Traditional MIPs are insoluble polymeric materials. Nonetheless, when prepared under precipitation polymerization, MIP nanoparticles, 10–100 nm in size, are obtained that have great biological compatibility.[32] Materials imprinted against mellittin (the major component of bee venom) in this way could be used to remove the toxin from the bloodstream of living mice.[33] MIP nanoparticles can be prepared for hydrophilic peptides as well, if they are first functionalized with a fatty acid acyl chain and anchored at the interface of inverse microemulsion.[34] One other way to produce water-soluble imprinted materials is to perform imprinting on the surface of (diacetylene-containing) vesicles, which after polymerization could report the binding event by fluorescence.[35]

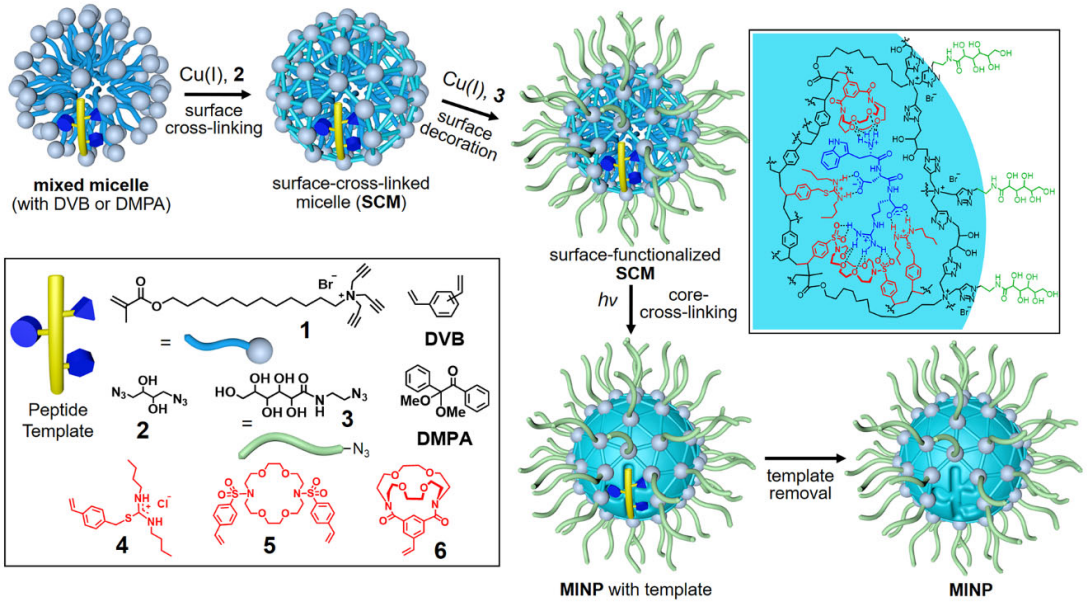
Micellar Imprinting of Peptides

To selectively protect a peptide under many biological settings, a 30–100 nm nanoparticle is probably still too large. Often times, it is a specific sequence of a long peptide to be protected in an enzymatic reaction, and the remaining peptide sequences could be part of a protein tertiary structure. Other times, a long peptide has several reaction sites and a specific site is to be protected while others remain accessible to their enzymatic catalysts. For all these situations, a high precision of protection is required that demands a peptide protector much smaller in size.

Our group in 2013 developed a method of molecular imprinting within doubly cross-linked surfactant micelles.[36] The so-called molecularly imprinted nanoparticles (MINPs) are ~5 nm in diameter with surface ligands and ~ 4 nm without. They are, hence, similar to a medium-sized protein in dimension and quite a bit smaller than typical antibodies (~10 nm). Their surface charge can be tuned by different types of cross-linkable surfactants.[37] Micellar imprinting was first used to create selective receptors for bile salt derivatives and then quickly expanded to a wide range of biologically interesting small molecules/drugs,[36-38] carbohydrates,[39] and peptides[40], *all in water*. Most recently, they are converted into artificial enzymes to catalyze a range of chemical reactions.[3, 41]

As shown in Scheme 1, micellar imprinting starts with spontaneous formation of micelles using a cross-linkable surfactant (**1**) in the presence of the interested peptide as the template molecule, divinylbenzene (DVB, a free-radical cross-linker), and a small amount of 2,2-dimethoxy-2-phenylaceto-phenone (DMPA, a photo initiator). The surface of the micelle is covered with a dense layer of alkyne groups from tripropargylammonium headgroup of the surfactant, and can be cross-linked by diazide **2** in the presence of Cu(I) catalysts via the highly efficient click reaction. Another round of click reaction with monoazide **3** installs a layer of hydrophilic ligands on the surface of the micelle. (The sugar-derived surface ligands are highly hydrophilic but insoluble in organic solvents such as acetone and methanol, and thus help the isolation and purification of the final MINPs.)

Micelles are highly dynamic, with diffusion-controlled intermicellar exchange of surfactants.[42] Covalently tethered on the surface, the surface-cross-linked micelle (SCM) becomes a nanoconfined space for molecular imprinting, as UV irradiation initiates free-radical polymerization/cross-linking around the template molecule in the micellar core.[43] The nanoconfinement is found to be extremely important to the large imprinting factors obtained from micellar imprinting (often in the hundreds[39d, 41c] and sometimes up to 10,000).[40e] In addition, MINPs can easily detect the addition,[43] removal,[43] and shift[40a] of a single methyl (or methylene) group in the guest binding.



Scheme 1. Preparation of peptide-binding MINP from molecular imprinting of a cross-linked micelle, with a schematic representation of the cross-linked structure containing WDR bound by polymerized FMs. (Reprinted with permission from Ref 51. Copyright 2021, the American Chemical Society.)

For peptide binding, we initially focused on those rich in hydrophobic amino acids because they have a strong driving force to enter the micelle.[40a] Our reasoning was that the hydrophobic side chains of amino acids have different degrees of hydrophobicity. For common hydrophobic amino acids, their side chains—shown schematically as blue shapes in Scheme 1—differ in size, shape, and hydrophobicity. Thus, a "hydrophobic code" exists for each peptide that describes the number, size, shape, and distribution of its hydrophobic side chains.

Micellar imprinting, indeed, was found to create a complementary array of hydrophobic indentations or "dimples" on the cross-linked micelles, essentially "encoding" the MINPs with supramolecular information to match the hydrophobic "code" of the peptide. These imprinted hydrophobic "dimples" turned out highly discriminating in their binding, to the point that the shift of a single methyl in leucine and isoleucine in isomeric di- and tripeptides could be distinguished, as well as phenylalanine and tyrosine.[40a] The binding was also highly selective. When 5 MINPs were created for 5 biological peptides, negligible cross-reactivity was observed when a particular peptide was titrated with the 4 nonmatching MINPs or, conversely, a particular MINP with the 4 nonmatching peptides.

Specific FMs (**4**–**6**) can be included in the MINP preparation to further improve the binding. They generally contain a polymerizable vinyl group and a molecular recognition motif targeting specific functional groups on the peptide (see the schematic representation of WDR bound by polymerized FMs in Scheme 1). FM **4**, for example, binds carboxylic acids through the hydrogen bond-reinforced thiouronium–carboxylate salt bridge.[40c] FM **5**, with abundant hydrogen-bond acceptors in the structure, prefers the guanidinium side chain of arginine.[40b] FMs **6** is selective for the amino group on the side chain of lysine and also on the N-terminus.[40d] With these FMs, we can target the hydrophobic, acidic, and basic groups of a peptide simultaneously, greatly enhancing both the binding selectivity and affinity of the MINP.[40d] The functionalized MINPs have been shown to distinguish closely related hydrophilic residues such as aspartic acid/glutamic acid and lysine/arginine.

One might be surprised by how well these hydrogen-bonded FMs work in MINP formation and binding. The mechanism is the same as how proteins and nucleic acids use these noncovalent forces in water—i.e., in a relatively nonpolar microenvironment where water is largely excluded. Although hydrogen-bonds are weakened by strong solvent competition in an aqueous solution, they are much stronger in the hydrophobic core of a micelle.[44]

Most recently, we discovered that, instead of specially designed FMs, commercially available amide-containing cross-linkers such as *N*,*N*'-methylene-bisacrylamide (MBAm) can be used instead of DVB during micellar imprinting (Figure 1).[40e] The radical initiator (DMPA), being hydrophobic, strongly prefers to reside within the nonpolar core of the micelle. Once the initiating radical reacts with the methacrylate of the cross-linkable surfactant (**1**) inside the SCM, the propagating radical is covalently attached to the micellar core and can polymerize only the MBAm molecules diffused to the surfactant/water interface of the micelle. As a result, a belt of hydrogen-bonding groups is formed at the interface, around the peptide template residing in the same area by its amphiphilicity.

When we compared the binding properties of MINPs prepared with DVB (our normal core-cross-linker) plus FMs and those prepared with MBAm (essentially as a hydrogen-bonding functional cross-linker), the MBAm-cross-linked MINPs were pleasantly found to outperform the DBV-cross-linked, functionalized MINPs consistently (Table 1). The binding constants for a number of complex biological peptides (**7**–**12** in Figure 1) ranged from 60 to 90 $\times 10^5$ M^{-1}, corresponding to 110–170 nM of binding affinity. Excellent binding selectivity was also observed (Figure 2).[40e]

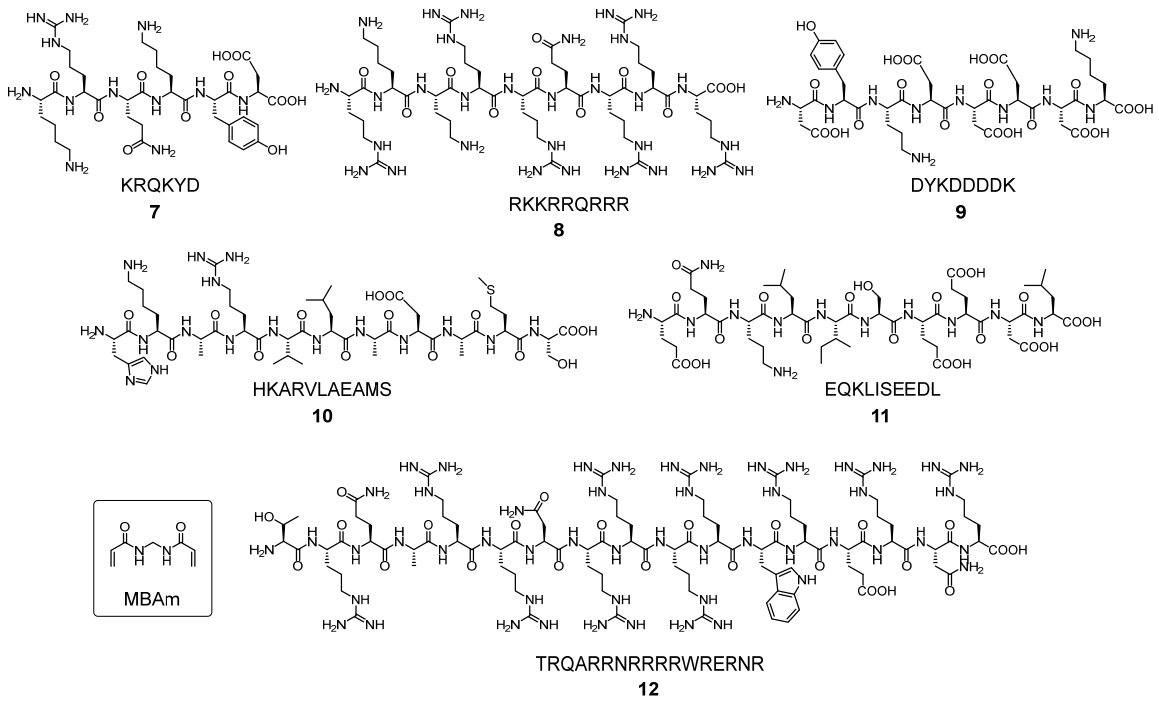


Figure 1. Structures of peptide templates used in the synthesis of MBAm-functionalized MINPs.

Table 1. Binding data for biological peptides **7**–**12** by MINPs prepared with DVB and FMs, and by MINPs prepared with MBAm without FMs.[a]

entry	template	cross-linker	K_a (×10^5 M^{-1})	-ΔG (kcal/mol)	N[b]
1	**7**	DVB	34.4 ± 1.73	8.91	0.9 ± 0.1
2		MBAm	62.2 ± 2.32	9.26	0.9 ± 0.1
3	**8**	DVB	45.3 ± 2.85	9.07	1.1 ± 0.1
4		MBAm	67.50 ± 2.66	9.31	1.1 ± 0.1
5	**9**	DVB	59.2 ± 0.31	9.23	1.1 ± 0.1
6		MBAm	73.10 ± 2.47	9.36	1.2 ± 0.1
7	**10**	DVB	82.3 ± 2.29	9.43	0.9 ± 0.1
8		MBAm	89.10 ± 2.47	9.47	1.1 ± 0.1
9	**11**	DVB	66.4 ± 2.65	9.30	0.8 ± 0.1
10		MBAm	72.50 ± 1.27	9.35	0.9 ± 0.1
11	**12**	DVB	53.40 ± 1.84	9.17	1.1 ± 0.1
12		MBAm	66.20 ± 3.36	9.30	1.0 ± 0.1

[a] The titrations were performed in HEPES buffer (10 mM, pH 7.4) in duplicates at 298 K and the errors between the runs were <10%. For MINPs prepared with FMs, the following stoichiometry was used in the formulation: 1.5:1 for **4**/carboxylate, 1:1 for **6**/amine, and 1:1 for **5**/arginine. [b] N is the number of binding sites per nanoparticle determined by isothermal titration calorimetry (ITC). (Reprinted with permission from Ref 40e. Copyright 2020, the American Chemical Society.)

MINP contains hydrogen-bonding groups including triazole, hydroxyl, and ester. Although these "background" interactions cannot be defined as precisely as the supramolecular "codes" defined by the specifically shaped hydrophobic dimples and the specially designed FMs, they are expected to be optimized to some extent during the imprinting process, for both the peptide backbone and side chains. These secondary interactions also can play important roles, evident from the binding of peptides containing glycine that lacks a side chain.[40a, 45]

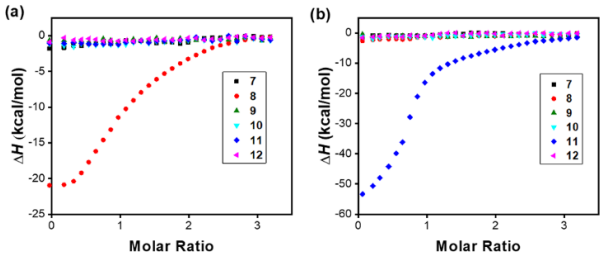


Figure 2. (a) ITC titration of peptides **7**–**12** to (a) MINP(**8**) and (b) MINP(**11**), showing only the desired peptide bound by the MINP. [MINP] = 5.0 μM. [peptide] = 75 μM in 10 mM HEPES buffer. The MINPs were prepared with 1:1 [**1**]/[MBAm]. (Reprinted with permission from Ref 40e. Copyright 2020, the American Chemical Society.)

Sequence-Selective Protection of Peptides by MINPs in Enzymatic Reactions

The nanodimension of MINPs and their strong and selective bindings for complex biological peptides bode well for their usage as protective agents for peptides. The Michaelis constants for common proteases are in the submillimolar to millimolar range[46] and those for kinases range from micromolar to millimolar.[47] The 100–200 nanomolar binding affinities (sometimes as low as 20

nM) of MINPs for hydrophobic and hydrophilic biological peptides[40] suggest that selective binding should be totally achievable.

Our first model peptide for protected proteolysis was Angiotensin III (A-III, RVYIHPF),[48] cleavable by two common endopeptidases —trypsin after arginine (R) and by chymotrypsin after tyrosine (Y). LCMS analysis showed that MINP(A), i.e., MINP prepared with A-III as the template, suppressed the proteolysis of the peptide to ≤10% during a period of 2 h at 1 equiv in the trypsin proteolysis and 2 equiv in the chymotrypsin proteolysis. Nonimprinted nanoparticles (NINPs) prepared without templates only slowed down the reaction slightly. A strong correlation between binding and protection was observed when MINPs targeting the first 4, 5, and 6 amino acids of the *N*-, and *C*-terminal sequences were used for the protection. The protection factor, defined as the ratio between the yield in buffer at 2 h and the yield in the presence of the MINP, showed a linear relationship to the binding free energy.

Interestingly, the proteolytic yield of A-III in the presence of MINP(A) fitted well to a 1:1 binding isotherm against the MINP concentration. The apparent "binding constant" obtained for trypsin inhibition (K_a = (2.35 ± 0.31) × 10^7 M^{-1}) was quite close to the actual binding constant determined by ITC (K_a = (1.89 ± 0.13) × 10^7 M^{-1}), suggesting the protection happened almost strictly with a 1:1 stoichiometry. Although the protection-based "binding constant" for chymotrypsin was a few times lower than the ITC-determined value, a strong binding–protection correlation was still observed. MINP protection was also found to work well for hydrophilic peptides (LRRASLG, PAGYLRRASVAQLT, and TGHGLRRSSKFCLK), if suitable FMs were used in the MINP preparation.

β-Amyloid peptides are released through proteolysis and implicated in Alzheimer's disease.[49] We decided to use $Aβ_{1-28}$ to demonstrate selective protection of a fragment of a long peptide because it contains two cleavable sites by trypsin—arginine at AA5 and lysine at AA16 (marked in green in Figure 3). Two MINPs, MINP($β_{1-14}$) and MINP($β_{15-28}$), were prepared, targeting the first and second halves of the parent peptide. ITC showed that the two MINPs bound the parent peptide strongly in pH 7.4 phosphate buffer, with K_a = 1.97 × 10^7 and 3.06 × 10^7 M^{-1}, respectively.

In the phosphate buffer (Figure 3a) or in the presence of NINPs (Figure 3b), trypsin hydrolyzed $Aβ_{1-28}$ to afford the expected peptide products (**13**–**15**). In addition, two peptides (**16** and **17**), with only the arginine or lysine cleaved, showed transiently in the first 2 h of reaction time. NINPs slowed down the proteolysis somewhat but the product distribution curves were similar in shape as those in the buffer.

A totally different product distribution was obtained when $Aβ_{1-28}$ was treated with trypsin in the presence of 1 equiv MINP($β_{1-14}$) or MINP($β_{15-28}$). The formerly transiently observed **16** (Figure 3c) and **17** (Figure 3d) were produced continually depending on which MINP protector was employed.

MINP protection did slow down the proteolysis of the exposed site, especially if the site was close to the protected sequence. For example, lysine 16 in $Aβ_{1-28}$ was only two residues away from the protected sequence of $Aβ_{1–14}$; its (selective) proteolysis in the presence of MINP($β_{1–14}$) took ~24 h to complete (Figure 3c), instead of 4 h in buffer (Figure 3a) and 6 h with NINP (Figure 3b). Arginine 5, on the other hand, was 9 residues away from $Aβ_{15–28}$ bound by MINP($β_{15–28}$) and its (selective) hydrolysis in $Aβ_{1-28}$ took approximately 12 h (Figure 3d).

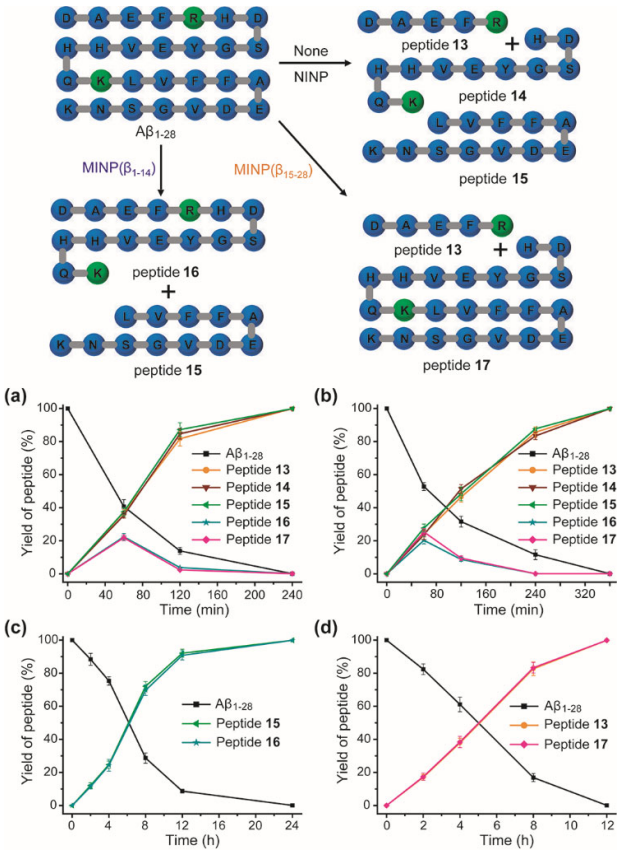


Figure 3. Product distribution curves in the trypsin digestion of $Aβ_{1-28}$ in buffer (a) and in the presence of 1 equiv NINP (b), MINP($β_{1–14}$) (c), and MINP($β_{15–28}$) (d). (Reprinted with permission from Ref 48. Copyright 2021, John Wiley & Sons, Inc.)

For MINP(A), MINP($β_{1-14}$), and MINP($β_{15-28}$), the nontemplating peptides showed very low cross-reactivities (0.06–0.13%) in the binding. This feature allowed us to carry out more advanced protections, using a 2:1 mixture of A-III and $Aβ_{1-28}$ for a proof of concept. Without any protector, the peptide mixture were digested by trypsin to afford peptides **13–15**, as well as **18** from A-III (Figure 4a). One equivalent of MINP(A) largely suppressed the proteolysis of A-III, while $Aβ_{1-28}$ hydrolyzed (Figure 4b). If MINP($β_{15–28}$) was used together with MINP(A), $Aβ_{1-28}$ underwent the anticipated selective cleavage after arginine 5 to afford **13** and **17** while A-III stayed largely intact (Figure 4c). Most interestingly, MINP($β_{1–14}$) and MINP($β_{15–28}$) could shied the long $Aβ_{1-28}$ together: after 4 h of reaction time, nearly 90 % of A-III hydrolyzed in the mixture while $Aβ_{1-28}$ persisted (Figure 4d). ITC confirmed that the long peptide indeed could bind both MINPs simultaneously, although the binding constants were several times lower than those measured with only one MINP, suggesting some steric/electrostatic repulsion existed when two MINPs came together to bind one long peptide.

Because the inhibition of the enzymatic reaction is driven strictly by selective binding, we expect that, anytime a peptide is bound more strongly by an MINP receptor than its enzyme catalyst, the enzymatic reaction can be inhibited. The prediction was confirmed recently in selective phosphorylation of peptide mixtures by cyclic AMP-dependent protein kinase (PKA), an enzyme with over 100 physiological substrates.[50] A particular

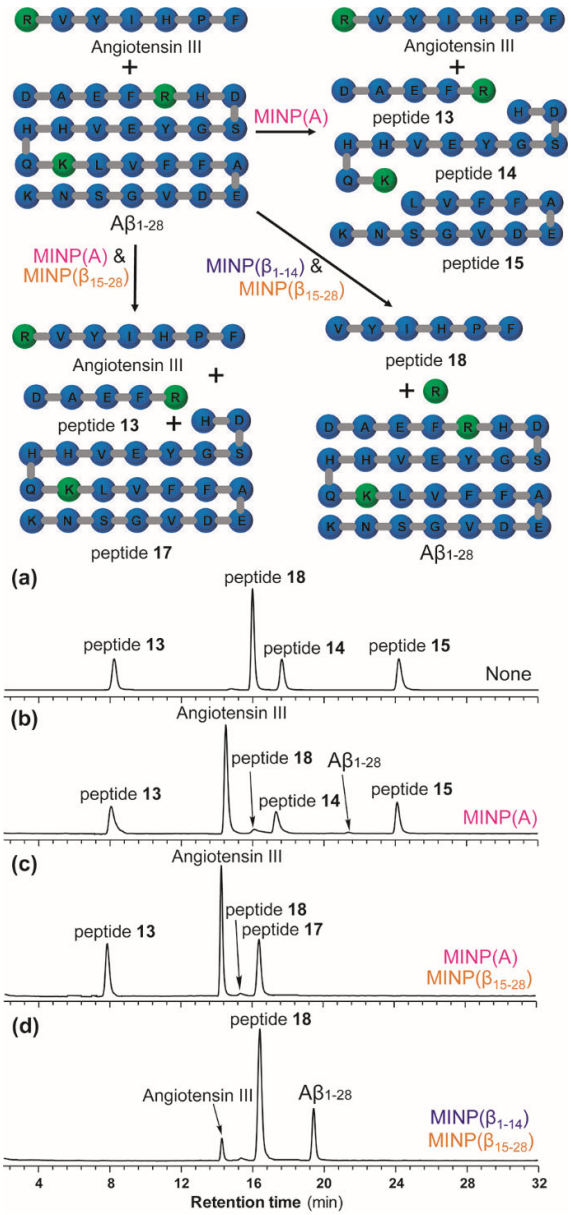


Figure 4. HPLC chromatograms of trypsin digestion of a 2:1 mixture of Angiotensin III and Aβ$_{1-28}$ by trypsin (a) without any protection, and in the presence of (b) MINP(A), (c) MINP(A) & MINP(β$_{15-28}$), and (d) MINP(β$_{1-14}$) & MINP(β$_{15-28}$). Reaction time was 4 h except in (c) which required 12 h for the selective hydrolysis of Aβ$_{1-28}$. (Reprinted with permission from Ref 48. Copyright 2021, John Wiley & Sons, Inc.)

challenge in controlled phosphorylation is that different substrates of a kinase generally have very similar or even identical "consensus motifs" surrounding the phosphorylation sites.[50b] PKA, for example, phosphorylates peptides with an RRXS motif (X = a variable amino acid). Yet, MINP was able to control the PKA-catalyzed phosphorylation of Kemptide (LRRASLG), β$_2$-adrenergic receptor peptide (TGHGLRRSSKFCLK), pyruvate kinase peptide (PAGYLRRASVAQLT), and cardiac myosin binding protein-C peptide (FRRTSLAGGGRRISDSHE) completely.[51] Note that Kemptide and pyruvate kinase peptide share identical consensus motifs, even the leucine (L) in front of the recognition motif. For cardiac myosin-binding protein-C peptide, selective protection of a fragment of the long peptide and cooperative protection of the entire sequence by two MINPs were both achieved.

Biological phosphorylation frequently occurs within a protein complex. One such example is the phosphotransfer step in the activation of the proline-rich tyrosine kinase 2 (Pyk2), a regulator of leukocyte motility, bone remodeling, and neuronal development.[52] As shown in Figure 5a, the Pyk2 activation occurs when tyrosine Y402 in the linker between the regulatory FERM and the kinase domain is autophosphorylated.[53] The intramolecular nature of the reaction makes it even more challenging to protect the substrate because MINP binding will have to compete with intramolecular protein–protein interactions.

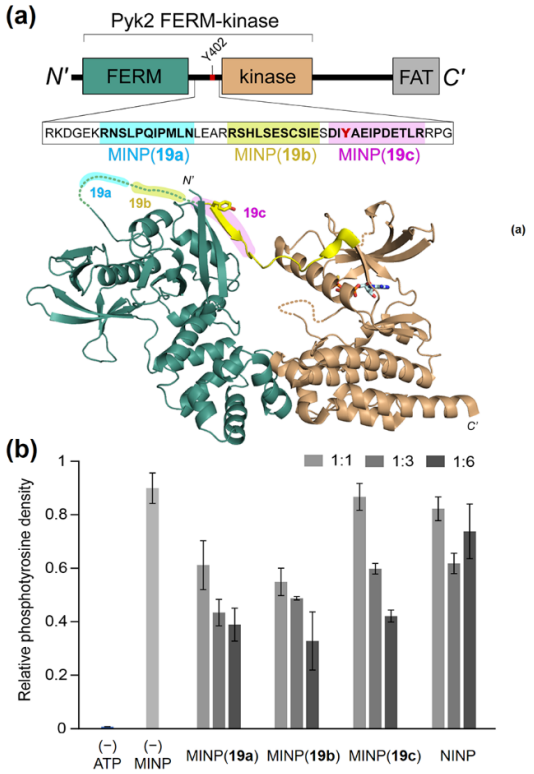


Figure 5. (a) Domain organization of Pyk2 and structural model depicting the Pyk2 FERM (PDB 4eku) and kinase (PDB 3fzp) aligned to the FAK FERM-kinase (PDB 2j0j). The FAK FERM--kinase linker is superimposed (yellow) to illustrate putative MINP binding sites. (b) Inhibition of Pyk2 autophosphorylation by MINPs at 1:1, 1:3, and 1:6 protein/MINP ratios, with the NINP as the control. [Pyk2] = 1.0 μM. (Reprinted with permission from Ref 51. Copyright 2021, the American Chemical Society.)

Since we can prepare MINPs conveniently to target different sections of a long peptide, we can essentially "scan" the linker by different MINPs and observe how the MINP binding affects the phosphorylation. Interestingly, when three MINPs targeting AA373–383, 388–398, and 400–411 of the linker, were added to the protein complex and ATP mixture, MINP(**19a**) and MINP(**19b**) turned out significantly more potent than MINP(**19c**) in the inhibition of the autophosphorylation (Figure 5b), even though it was MINP(**19c**) that directly impinged on Y402 in its binding. This unusual behavior might have resulted from a lower accessibility of **19c** sequence, since evidence exists that suggests the

autophosphorylation site could be part of an abbreviated β sheet.[54]

Conclusions and Outlook

MINPs have a remarkable ability to bind complex biological peptides in aqueous solution. With appropriate functional monomers[40d] and/or cross-linkers,[40e] they can frequently achieve tens of nanomolar binding affinities for peptides with 10–20 amino acid residues. Their ability to distinguish closely related residues including leucine/isoleucine,[40a] phenylalanine/tyrosine,[40a] glutamic acid/aspartic acid,[40c] and lysine/arginine[40b] makes them an extremely attractive class of materials for biological applications. Once the cross-linkable surfactant, cross-linker, and templates are available, their preparation takes less than 2 days and purification requires nothing other than precipitation and washing.

Because most enzymes bind their substrates with millimolar to micromolar affinities, MINPs are expected to compete effectively with many enzymes in the binding and, in turn, to shield their peptide substrates from enzymatic actions. Controlled proteolysis and phosphorylation are just examples chosen to illustrate the power of substrate protection, which should be quite general. Biology historically has been a great source of inspiration to chemists in their development of methods to recognize, transport, and transform molecules. Protection/deprotection, on the other hand, is largely an invention of chemists for the synthesis of complex organic molecules. Maybe, the strategy now is ready to find its way back into biology, as a way to return the favor.

Acknowledgements

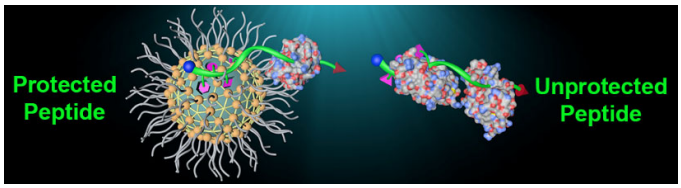
We thank NSF (DMR-2002659) for financial support of this research.

Keywords: molecular imprinting • peptide • binding • proteolysis • protection

[1] A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, M. E. Etzler, *Essentials of glycobiology*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., **2009**.
[2] A. V. Demchenko, Wiley-VCH, Weinheim, , Germany, **2008**.
[3] a) X. Li, Y. Zhao, *Chem. Sci.* **2021**, *12*, 374-383; b) X. Li, M. Zangiabadi, Y. Zhao, *J. Am. Chem. Soc.* **2021**, *143*, 5172-5181.
[4] P. G. M. Wuts, T. W. Greene, T. W. Greene, *Greene's protective groups in organic synthesis*, 5th ed., Wiley, Hoboken, New Jersey, **2014**.
[5] Y. A. DeClerck, S. Imren, *Eur. J. Cancer* **1994**, *30*, 2170-2180.
[6] C. Burkard, M. H. Verheije, O. Wicht, S. I. van Kasteren, F. J. van Kuppeveld, B. L. Haagmans, L. Pelkmans, P. J. M. Rottier, B. J. Bosch, C. A. M. de Haan, *PLoS Pathog.* **2014**, *10*, e1004502.
[7] H. Kido, Y. Niwa, Y. Beppu, T. Towatari, *Adv. Enzyme. Regul.* **1996**, *36*, 325-347.
[8] a) J. A. Dent, S. D. Berkowitz, J. Ware, C. K. Kasper, Z. M. Ruggeri, *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 6306-6310; b) G. R. Sambrano, W. Huang, T. Faruqi, S. Mahrus, C. Craik, S. R. Coughlin, *J. Biol. Chem.* **2000**, *275*, 6819-6823; c) D. J. Shedlock, M. A. Bailey, P. M. Popernack, J. M. Cunningham, D. R. Burton, N. J. Sullivan, *Virology* **2010**, *401*, 228-235; d) K. Tiyanont, Thomas E. Wales, M. Aste-Amezaga, Jon C. Aster, John R. Engen, Stephen C. Blacklow, *Structure* **2011**, *19*, 546-554.
[9] a) L. A. Pinna, M. Ruzzene, *Biochim. Biophys. Acta* **1996**, *1314*, 191-225; b) J. A. Ubersax, J. E. Ferrell Jr, *Nat. Rev. Mol. Cell. Biol.* **2007**, *8*, 530-541.
[10] S. Müller, A. Chaikuad, N. S. Gray, S. Knapp, *Nat. Chem. Biol.* **2015**, *11*, 818-821.
[11] a) Z. Zhang, T. Ly, T. Kodadek, *Chem. Biol.* **2001**, *8*, 391-397; b) T. Kodadek, *Biopolym. Peptide Sci.* **2002**, *66*, 134-140.
[12] a) L. A. Logsdon, A. R. Urbach, *J. Am. Chem. Soc.* **2013**, *135*, 11414-11416; b) J. Mosquera, B. Szyszko, S. K. Y. Ho, J. R. Nitschke, *Nat. Commun.* **2017**, *8*, 14882.
[13] F. Wang, A. Gomez-Escudero, R. R. Ramireddy, G. Murage, S. Thayumanavan, R. W. Vachet, *J. Am. Chem. Soc.* **2013**, *135*, 14179-14188.
[14] a) E. Faggi, Y. Perez, S. V. Luis, I. Alfonso, *Chem. Commun.* **2016**, *52*, 8142-8145; b) C.-P. Li, Y.-X. Lu, C.-T. Zi, Y.-T. Zhao, H. Zhao, Y.-P. Zhang, *Int. J. Mol. Sci.* **2020**, *21*, 4979; c) L. Tapia, N. Solozabal, J. Sola, Y. Pérez, W. T. Miller, I. Alfonso, *Chem.-Eur. J.* **2021**, *27*, Accepted Author Manuscript. https://doi.org/10.1002/chem.202100990.
[15] H. Peacock, C. C. Thinnes, A. Kawamura, A. D. Hamilton, *Supramol. Chem.* **2016**, *28*, 575-581.
[16] a) J. L. Atwood, J. M. Lehn, *Comprehensive Supramolecular Chemistry*, Pergamon, New York, **1996**; b) J. W. Steed, P. A. Gale, *Supramolecular Chemistry: From Molecules to Nanomaterials*, Wiley, Weinheim, **2012**.
[17] a) M. V. Rekharsky, T. Mori, C. Yang, Y. H. Ko, N. Selvapalam, H. Kim, D. Sobransingh, A. E. Kaifer, S. Liu, L. Isaacs, W. Chen, S. Moghaddam, M. K. Gilson, K. Kim, Y. Inoue, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20737-20742; b) R. A. Tromans, T. S. Carter, L. Chabanne, M. P. Crump, H. Li, J. V. Matlock, M. G. Orchard, A. P. Davis, *Nat. Chem.* **2019**, *11*, 52-56.
[18] a) G. V. Oshovsky, D. N. Reinhoudt, W. Verboom, *Angew. Chem. Int. Ed.* **2007**, *46*, 2366-2393; b) E. A. Kataev, C. Müller, *Tetrahedron* **2014**, *70*, 137-167.
[19] a) M. W. Peczuh, A. D. Hamilton, *Chem. Rev.* **2000**, *100*, 2479-2494; b) D. Maity, C. Schmuck, in *Synthetic Receptors for Biomolecules: Design Principles and Applications*, The Royal Society of Chemistry, **2015**, pp. 326-368; c) S. van Dun, C. Ottmann, L.-G. Milroy, L. Brunsveld, *J. Am. Chem. Soc.* **2017**, *139*, 13960-13968.
[20] a) J. I. Hong, S. K. Namgoong, A. Bernardi, W. C. Still, *J. Am. Chem. Soc.* **1991**, *113*, 5111-5112; b) A. T. Wright, E. V. Anslyn, J. T. McDevitt, *J. Am. Chem. Soc.* **2005**, *127*, 17405-17411; c) C. Schmuck, P. Wich, *Angew. Chem. Int. Ed.* **2006**, *45*, 4277-4281; d) S. Niebling, H. Y. Kuchelmeister, C. Schmuck, S. Schlucker, *Chem. Sci.* **2012**, *3*, 3371-3377; e) T. Gersthagen, J. Hofmann, F. G. Klärner, C. Schmuck, T. Schrader, *Eur. J. Org. Chem.* **2013**, *2013*, 1080-1092.
[21] M. A. Hossain, H.-J. Schneider, *J. Am. Chem. Soc.* **1998**, *120*, 11208-11209.
[22] a) R. Breslow, Z. Yang, R. Ching, G. Trojandt, F. Odobel, *J. Am. Chem. Soc.* **1998**, *120*, 3536-3537; b) H. Yamamura, M. V. Rekharsky, Y. Ishihara, M. Kawai, Y. Inoue, *J. Am. Chem. Soc.* **2004**, *126*, 14224-14233.
[23] a) L. M. Heitmann, A. B. Taylor, P. J. Hart, A. R. Urbach, *J. Am. Chem. Soc.* **2006**, *128*, 12574-12581; b) M. E. Bush, N. D. Bouley, A. R. Urbach, *J. Am. Chem. Soc.* **2005**, *127*, 14511-14517; c) L. C. Smith, D. G. Leach, B. E. Blaylock, O. A. Ali, A. R. Urbach, *J. Am. Chem. Soc.* **2015**, *137*, 3663-3669.
[24] S. Tashiro, M. Tominaga, M. Kawano, B. Therrien, T. Ozeki, M. Fujita, *J. Am. Chem. Soc.* **2005**, *127*, 4546-4547.
[25] F.-G. Klärner, T. Schrader, *Acc. Chem. Res.* **2013**, *46*, 967-978.
[26] a) E. Faggi, A. Moure, M. Bolte, C. Vicent, S. V. Luis, I. Alfonso, *J. Org. Chem.* **2014**, *79*, 4590-4601; b) E. Faggi, C. Vicent, S. V. Luis, I. Alfonso, *Org. Biomol. Chem.* **2015**, *13*, 11721-11731.
[27] S. Yapar, M. Oikonomou, A. H. Velders, S. Kubik, *Chem. Commun.* **2015**, *51*, 14247-14250.
[28] D. J. Cram, *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 1039-1057.
[29] a) G. Wulff, *Chem. Rev.* **2002**, *102*, 1-28; b) K. Haupt, K. Mosbach, *Chem. Rev.* **2000**, *100*, 2495-2504; c) L. Ye, K. Mosbach, *Chem. Mater.* **2008**, *20*, 859-868.
[30] a) A. Rachkov, N. Minoura, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2001**, *1544*, 255-266; b) H. Nishino, C. S. Huang, K. J. Shea, *Angew. Chem. Int. Ed.* **2006**, *45*, 2392-2396; c) A. M. Bossi, P. S. Sharma, L. Montana, G. Zoccatelli, O. Laub, R. Levi, *Anal. Chem.* **2012**, *84*, 4036-4041; d) K. Yang, J. Liu, S. Li, Q. Li, Q. Wu, Y. Zhou, Q. Zhao, N. Deng, Z. Liang, L. Zhang, Y. Zhang, *Chem. Commun.* **2014**, *50*, 9521-9524; e) Y. Zhang, C. Deng, S. Liu, J. Wu, Z. Chen, C. Li, W. Lu, *Angew. Chem. Int. Ed.* **2015**, *54*, 5157-5160; f) S. Li, K. Yang, N. Deng, Y. Min, L. Liu, L. Zhang, Y. Zhang, *ACS Appl. Mater. Interfaces* **2016**, *8*, 5747-5751; g) S. Schwark, W. Sun, J. Stute, D. Lutkemeyer, M. Ulbricht, B. Sellergren, *RSC Adv.* **2016**, *6*, 53162-53169; h) A. Cecchini, V. Raffa, F. Canfarotta, G. Signore, S. Piletsky, M. P. MacDonald, A. Cuschieri, *Nano Lett.* **2017**, *17*, 2307-2312; i) G. Pan, S. Shinde, S. Y. Yeung, M. Jakštaitė, Q. Li, A. G. Wingren, B. Sellergren, *Angew. Chem. Int. Ed.* **2017**, *56*, 15959-15963; j) Y.-P. Qin, C. Jia, X.-W. He, W.-Y. Li, Y.-K. Zhang, *ACS Appl. Mater. Interfaces* **2018**, *10*, 9060-9068; k) R. Xing, Y. Ma, Y. Wang, Y. Wen, Z. Liu, *Chem. Sci.* **2019**, *10*, 1831-1835.
[31] B. Sellergren, M. Lepistoe, K. Mosbach, *J. Am. Chem. Soc.* **1988**, *110*, 5853-5860.
[32] Y. Hoshino, T. Kodama, Y. Okahata, K. J. Shea, *J. Am. Chem. Soc.* **2008**, *130*, 15242-15243.

[33] Y. Hoshino, H. Koide, T. Urakami, H. Kanazawa, T. Kodama, N. Oku, K. J. Shea, *J. Am. Chem. Soc.* **2010**, *132*, 6644-6645.
[34] Z. Zeng, Y. Hoshino, A. Rodriguez, H. Yoo, K. J. Shea, *ACS Nano* **2010**, *4*, 199-204.
[35] S. Banerjee, B. König, *J. Am. Chem. Soc.* **2013**, *135*, 2967-2970.
[36] J. K. Awino, Y. Zhao, *J. Am. Chem. Soc.* **2013**, *135*, 12552-12555.
[37] L. Duan, Y. Zhao, *J. Org. Chem.* **2019**, *84*, 13457-13464.
[38] a) J. K. Awino, Y. Zhao, *ACS Biomater. Sci. Eng.* **2015**, *1*, 425-430; b) M. D. Arifuzzaman, Y. Zhao, *J. Org. Chem.* **2016**, *81*, 7518-7526; c) L. Duan, Y. Zhao, *Bioconjugate Chem.* **2018**, *29*, 1438-1445; d) L. Duan, Y. Zhao, *Chem. Asian J.* **2020**, *15*, 1035-1038; e) L. Duan, Y. Zhao, *React. Funct. Polym.* **2021**, *158*, 104759.
[39] a) J. K. Awino, R. W. Gunasekara, Y. Zhao, *J. Am. Chem. Soc.* **2016**, *138*, 9759-9762; b) R. W. Gunasekara, Y. Zhao, *J. Am. Chem. Soc.* **2017**, *139*, 829-835; c) R. W. Gunasekara, Y. Zhao, *Chem. Commun.* **2019**, *55*, 4773-4776; d) L. Duan, M. Zangiabadi, Y. Zhao, *Chem. Commun.* **2020**, *56*, 10199-10202; e) M. Zangiabadi, Y. Zhao, *Nano Lett.* **2020**, *20*, 5106-5110.
[40] a) J. K. Awino, R. W. Gunasekara, Y. Zhao, *J. Am. Chem. Soc.* **2017**, *139*, 2188-2191; b) S. Fa, Y. Zhao, *Chem. Mater.* **2017**, *29*, 9284-9291; c) S. Fa, Y. Zhao, *Chem. -Eur. J.* **2018**, *24*, 150-158; d) S. Fa, Y. Zhao, *Chem. Mater.* **2019**, *31*, 4889-4896; e) M. Zangiabadi, Y. Zhao, *ACS Appl. Polym. Mater.* **2020**, *2*, 3171-3180.
[41] a) M. Arifuzzaman, Y. Zhao, *ACS Catal.* **2018**, *8*, 8154-8161; b) L. Hu, M. D. Arifuzzaman, Y. Zhao, *ACS Catal.* **2019**, *9*, 5019-5024; c) X. Li, Y. Zhao, *ACS Catal.* **2020**, *10*, 13800-13808; d) X. Li, Y. Zhao, *J. Am. Chem. Soc.* **2019**, *141*, 13749-13752; e) I. Bose, Y. Zhao, *ACS Catal.* **2020**, *10*, 13973-13977; f) I. Bose, S. Fa, Y. Zhao, *J. Org. Chem.* **2021**, *86*, 1701-1711.
[42] E. A. G. Aniansson, S. N. Wall, M. Almgren, H. Hoffmann, I. Kielmann, W. Ulbricht, R. Zana, J. Lang, C. Tondre, *J. Phys. Chem.* **1976**, *80*, 905-922.
[43] K. Chen, Y. Zhao, *Org. Biomol. Chem.* **2019**, *17*, 8611-8617.
[44] a) J. S. Nowick, J. S. Chen, *J. Am. Chem. Soc.* **1992**, *114*, 1107-1108; b) J. S. Nowick, J. S. Chen, G. Noronha, *J. Am. Chem. Soc.* **1993**, *115*, 7636-7644.
[45] Y. Zhao, *Chem.-Eur. J.* **2018**, *24*, 14001-14009.
[46] a) K. A. Walsh, in *Meth. Enzymol., Vol. 19*, Academic Press, **1970**, pp. 41-63; b) A. Himoe, P. C. Parks, G. P. Hess, *J. Biol. Chem* **1967**, *242*, 919-929.
[47] B. E. Kemp, R. B. Pearson, *Methods Enzymol.* **1991**, *200*, 121-134.
[48] X. Li, K. Chen, Y. Zhao, *Angew. Chem. Int. Ed.* **2021**, *60*, 11092-11097.
[49] S. S. Sisodia, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 6075-6079.
[50] a) D. A. Johnson, P. Akamine, E. Radzio-Andzelm, Madhusudan, S. S. Taylor, *Chem. Rev.* **2001**, *101*, 2243-2270; b) J. B. Shabb, *Chem. Rev.* **2001**, *101*, 2381-2412.
[51] X. Li, T. M. Palhano Zanela, E. S. Underbakke, Y. Zhao, *J. Am. Chem. Soc.* **2021**, *143*, 639-643.
[52] R. Naser, A. Aldehaiman, E. Diaz-Galicia, S. T. Arold, *Cancers* **2018**, *10*, 196.
[53] a) H. S. Loving, E. S. Underbakke, *Biochemistry* **2019**, *58*, 3767-3776; b) S. Y. Park, H. K. Avraham, S. Avraham, *J. Biol. Chem.* **2004**, *279*, 33315-33322.
[54] D. Lietha, X. M. Cai, D. F. J. Ceccarelli, Y. Q. Li, M. D. Schaller, M. J. Eck, *Cell* **2007**, *129*, 1177-1187.

Entry for the Table of Contents



Highly specific enzymes make protecting groups unnecessary in biology but some enzymes including protease and kinase have multiple biological substrates and there is strong interest to stop a specific enzyme-catalyzed reaction without inhibiting the enzyme itself. Recent developments in peptide-imprinted nanoparticles allow precise protection of peptide substrates and open up previously impossible ways to manipulate enzymatic reactions.