Sequence-Selective Protection of Peptides from Proteolysis

*Xiaowei Li, Kaiqian Chen, and Yan Zhao**
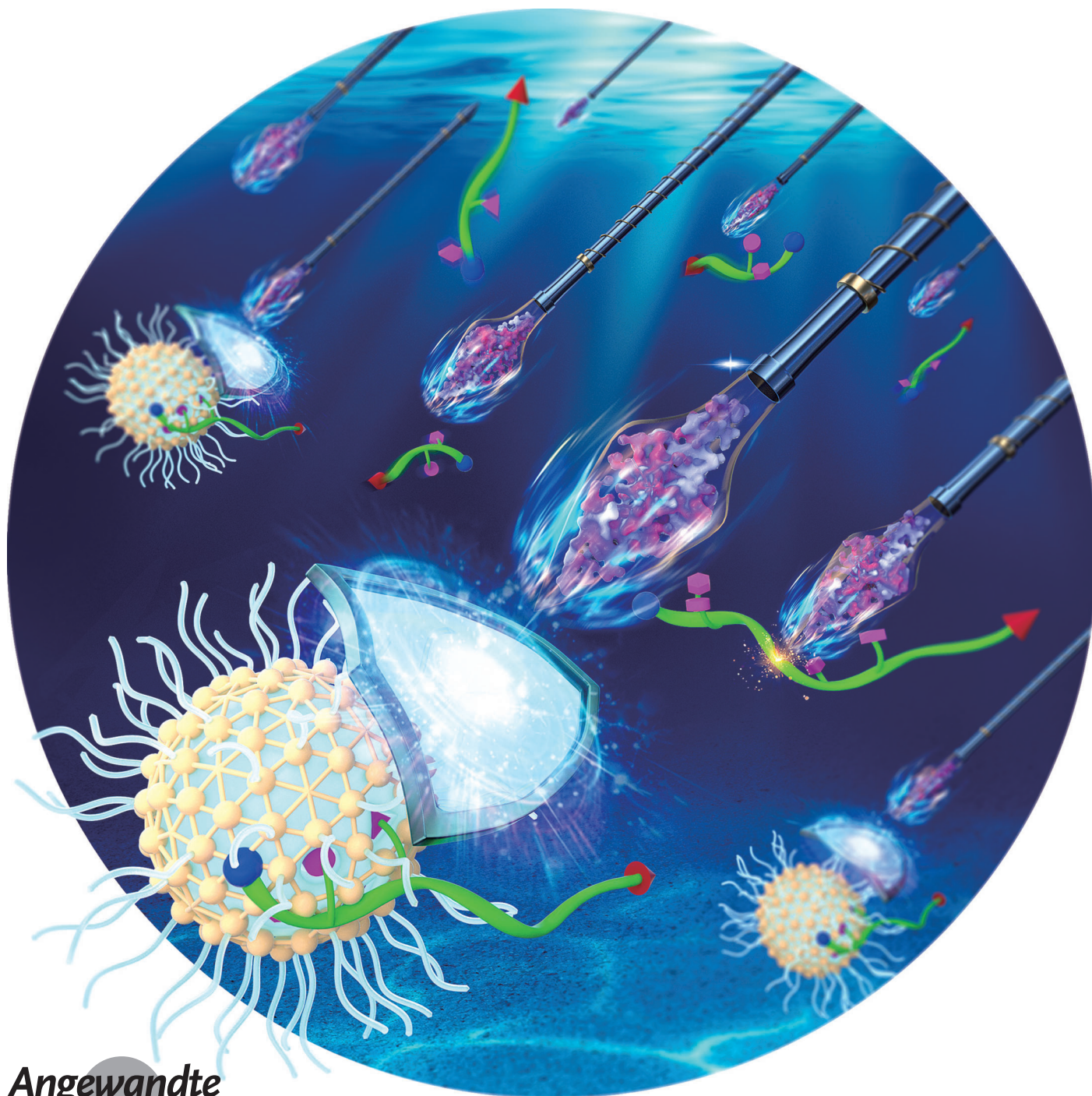
Abstract: Proteolysis of proteins and peptides is involved in the infection of cells by enveloped viruses and also in the invasion and spread of cancer cells. Shutting down broad-specificity proteases, however, is problematic because normal functions by these proteases will be affected. Herein, nanoparticle receptors were prepared from molecular imprinting for complex biological peptides. Their strong and selective binding enabled them to protect their targeted sequences from proteolysis in aqueous solution at stoichiometric amounts. Generality of the method was demonstrated by the protection of hydrophobic and hydrophilic peptides from different proteases, selective protection of a segment of a long peptide, and selective protection of a targeted peptide in a mixture. Most interestingly, two receptors targeting different parts of a long peptide could work in cooperation to protect the overall sequence, highlighting the versatility of the method.
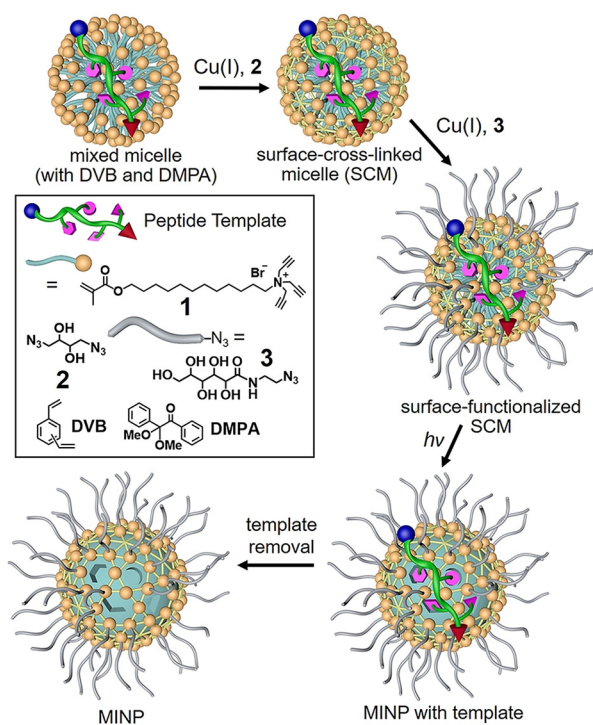
Total synthesis of complex organic molecules is inconceivable without the advancement of protective group chemistry.[1] Whenever a functional group needs to be preserved under an otherwise reactive condition, an appropriate protective group is often the most straightforward and reliable solution. The situation is very different in biology, with many enzymes available to carry out highly specific tasks in cleavage, coupling, and other manipulation of biomolecules. Although these enzymes enable researchers to perform selective transformations of biomolecules without protective groups, the specificity of enzymes has its own limitation. Widely used endopeptidases, for example, possess impressive selectivities: trypsin cleaves the peptide bond after a basic residue (K or R) and chymotrypsin after a hydrophobic aromatic residue (F, W, or Y). Nonetheless, if one wants to hydrolyze a particular K in a peptide with multiple Ks or hydrolyze a particular peptide among multiple peptides present that all contain Ks, selectivity of the protease is not enough.

The situations described above are not just of hypothetic value. Proteolytic activation is a critical step in cellular infection by enveloped viruses including HIV-1,[2] influenza virus A,[2] and coronavirus.[3] Over-expressed proteases are crucial in tissue remodeling associated with cancer invasion and spread.[4] Traditional protease inhibitors tend to have high toxicities in anticancer treatment because they affect normal cellular functions that also rely on these enzymes.

[*] Dr. X. Li, K. Chen, Prof. Y. Zhao
Department of Chemistry
Iowa State University
Ames, IA 50011-3111 (USA)
E-mail: zhaoy@iastate.edu
Homepage: https://www.chem.iastate.edu/people/yan-zhao

One solution to selective inhibition of proteolysis is to target the peptide substrates instead.[5] Antibodies indeed have been used to inhibit protease-activated receptors by shielding the targeted cleavage sites on the substrates.[6] Nonetheless, antibodies are expensive and fragile molecules made of polypeptides, which themselves are subject to proteolysis. Other materials for shielding peptides from proteolysis are library-encoded peptides/peptoids, but the binding affinities were in the micromolar range[5] and the identification of strong binders was not a trivial task.[7] Synthetic peptide-binding materials do exist[8] including those made through molecular imprinting.[9] However, less than a handful of reports appeared in the literature that demonstrated protection of peptides (from proteolysis[10] or other reactions[11]). With the reported protection so far dependent on the binding of specific amino acids or charged side chains, a general protection strategy for long, biological peptides is missing despite the critical needs in biology.

Herein, we describe synthetic receptors for precise, sequence-selective protection of peptides. With their nano-dimension, as well as strong and selective binding, these receptors can shield specific parts of a peptide while the rest of the chain and other unprotected peptides in the solution undergo proteolysis. Also, a long peptide can be protected by two receptors targeting different segments of the total sequence, making the protection highly versatile and predictable.

Our protective receptors were prepared through molecular imprinting[12] of cross-linked micelles (Scheme 1).[13] The facile one-pot synthesis consists of several key steps: (a) spontaneous micellization of cross-linkable surfactant **1**, (b) spontaneous inclusion of the peptide template, divinyl-

Scheme 1. Preparation of peptide-binding MINP from molecular imprinting of a cross-linked micelle.

benzene (DVB), and 2,2-dimethoxy-2-phenylacetophenone (DMPA, a photoinitiator) into the micelle, (c) covalent capture of the mixed micelle on the surface by diazide **2** using the click reaction, (d) surface decoration of the cross-linked micelle by monoazide **3** using the click reaction, and (e) an orthogonal, free radical core-cross-linking induced typically by UV irradiation. A highlight of micellar imprinting is the large imprinting factor obtained (frequently reaching hundreds[14] and sometimes 10 000[15]), as a result of templated polymerization in the confined nanospace.[16] The number of binding sites per molecularly imprinted nanoparticle (MINP) is conveniently controlled by the surfactant/template ratio.[13] Purification requires nothing other than precipitation and washing, due to the solubility properties of MINP imparted by surface ligand **3**.

The idea of using MINPs for peptide protection was based on the reasoning that the Michaelis constants for common proteases are usually in the submillimolar to millimolar range,[17] but MINPs can bind many biological peptides with tens of nanomolar affinities.[15,18] Selectivity of the protection comes from the high selectivity of binding displayed by MINPs for peptides. The imprinted pockets, for example,
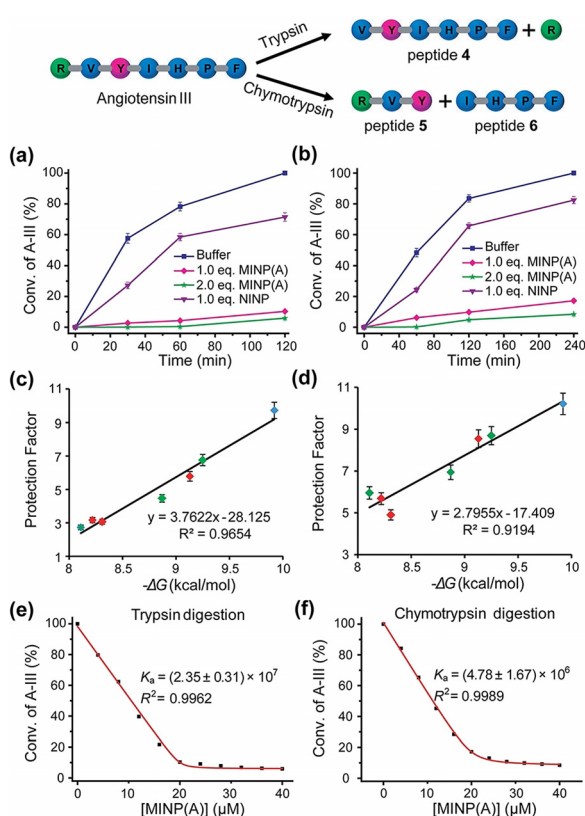
could distinguish leucine from isoleucine easily, as well as phenylalanine from tyrosine.[18a]

To demonstrate selective protection of peptides by MINP, we first tested the proteolysis of a biological peptide, Angiotensin III (A-III, RVYIHPF). The peptide was chosen as our first model substrate because it is cleavable both by trypsin after arginine (R) and by chymotrypsin after tyrosine (Y). The proteolysis was monitored by LC–MS and the results are summarized in Figure 1.

Figure 1 a,b shows that the peptide could be cleaved by trypsin and chymotrypsin in 120 and 240 minutes in buffer, respectively. Importantly, MINP(A), i.e., MINP prepared with A-III as the template, could suppress the reaction to $\leq 10\%$ at 1 equiv in the trypsin proteolysis and 2 equiv in the chymotrypsin proteolysis. Molecular imprinting was key to the protection, because nonimprinted nanoparticles (NINPs) prepared without templates only slowed down the proteolysis slightly. Isothermal titration calorimetry (ITC) showed that the binding constant (K_a) of MINP(A) for A-III was 1.89×10^7 M^{-1} in pH 7.4 phosphate buffer and that of NINP was only 1.99×10^4 M^{-1} (Table 1, entries 1 and 2). The imprint/non-imprint ratio (i.e., imprinting factor) was thus 950, under-scoring the strong template effect in the micellar imprinting.

To establish the correlation between binding and protection, we prepared a series of 6 MINPs for A-III, using the N- and C-terminal sequences containing the first 4, 5, and 6 amino acids from the chain ends, and measured their bindings for the parent peptide by ITC (Table 1, entries 3–8). We then studied the proteolysis of the parent peptide in the presence of the seven MINPs—6+ MINP(A)—and deter-mined the yields of the expected products after 2 h of digestion with trypsin (Table S1) and chymotrypsin

Figure 1. a,b) Proteolysis of A-III in 10 mM phosphate buffer (pH 7.4) under different conditions by a) trypsin and b) chymotrypsin. c,d) Correlation between the binding free energies of MINPs for A-III and the protection factors of the MINPs in proteolysis by c) trypsin and d) chymotrypsin. Blue data point was for the full sequence, green points were for the N-terminal sequences, and red for the C-terminal sequences. e,f) Nonlinear least squares curve fitting of the conversion of A-III by e) trypsin and f) chymotrypsin digestion in 10 mM phosphate buffer (pH 7.4).

Table 1: Binding data for MINPs obtained by ITC.[a]

Entry	Template for MINP	Guest	K_a [10^4 M^{-1}]	$-\Delta G$ [kcal mol^{-1}]	N[b]
1	A-III	A-III	1890 ± 130	9.92	0.95 ± 0.01
2	none	A-III	1.99 ± 0.68	5.86	—[c]
3	RVYIHP	A-III	605 ± 18	9.25	0.98 ± 0.01
4	VYIHPF	A-III	501 ± 65	9.13	1.07 ± 0.01
5	RVYIH	A-III	324 ± 18	8.87	0.98 ± 0.01
6	YIHPF	A-III	107 ± 4.58	8.22	0.99 ± 0.01
7	RVYI	A-III	89.4 ± 4.38	8.11	0.98 ± 0.01
8	IHPF	A-III	125 ± 6.46	8.31	1.16 ± 0.01
9	A III	Aβ_{1-28}	1.58 ± 0.33	5.72	—[c]
10	Aβ_{1-14}	Aβ_{1-28}	1970 ± 250	9.95	0.99 ± 0.01
11	Aβ_{1-14}	Aβ_{1-28}[d]	574 ± 31	9.22	0.94 ± 0.03
12	Aβ_{15-28}	Aβ_{1-28}	3060 ± 560	10.24	1.08 ± 0.01
13	Aβ_{15-28}	Aβ_{1-28}[e]	612 ± 60	9.26	0.90 ± 0.02
14	Aβ_{1-14}	A-III	2.44 ± 0.31	5.98	—[c]
15	Aβ_{15-28}	A-III	1.76 ± 0.45	5.79	—[c]

[a] The titrations were performed in duplicates in phosphate buffer (10 mM, pH 7.4) and the errors between the runs were < 10%. The titration curves are shown in Figures S6–S20 including the binding enthalpy and entropy. [b] N is the average number of binding sites per nanoparticle measured by ITC. [c] The number of binding sites per particle (N) is not listed because the much weaker binding made the determination of the value inaccurate. [d] The binding was performed by titrating MINP(β_{1-14}) with a 1:1 mixture of Aβ_{1-28} and MINP(β_{15-28}). [e] The binding was performed by titrating MINP(β_{15-28}) with a 1:1 mixture of Aβ_{1-28} and MINP(β_{1-14}).

(Table S2). To our delight, 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 (Figure 1 c,d, Table S3).

To further support the binding-driven protection, we monitored the proteolysis of A-III with various concentrations of MINP(A). The proteolytic yields at 2 h were found to fit well to a 1:1 binding isotherm and afforded an apparent "binding constant" of $K_a = (2.35 \pm 0.31) \times 10^7 \, \text{M}^{-1}$ for trypsin (Figure 1 e). This value was quite close to the actual binding constant determined by ITC between MINP(A) and A-III, i.e., $(1.89 \pm 0.13) \times 10^7 \, \text{M}^{-1}$. The sharp transition in Figure 1 e indicates that the protection happened nearly perfectly with a 1:1 stoichiometry.

For the proteolysis by chymotrypsin, the reaction-based "binding constant" was $(4.78 \pm 1.67) \times 10^6 \, \text{M}^{-1}$ (Figure 1 f), about 4 times lower than the ITC-determined binding constant. Thus, even though the degree of protection still correlated with the binding affinity, the protective effect was somewhat weaker than what a perfect 1:1 protection would predict. As demonstrated in Figure 1 b, a stronger protection could be obtained by using a higher amount of MINP.

Not only did the MINP protection work well for relatively hydrophobic peptides such as A-III, hydrophilic peptides such as LRRASLG, PAGYLRRASVAQLT, and TGHGLRRSSKFCLK can be protected from proteolysis as well, shown by similar ITC and LC–MS experiments (Table S9 and Figures S44–S46). Functional monomers (FMs) were needed for the MINP preparation for these hydrophilic peptides to achieve strong binding, targeting the amine,[18d] carboxylate,[18b] and guanidinium groups.[18c]

Having established the binding–protection correlation and the generality of the MINP protection, we set the next goal of selectively hydrolyzing β-amyloid peptide (1–28) or Aβ$_{1-28}$ at one of the two cleavable sites by trypsin, i.e., arginine at AA5 and lysine at AA16 marked in green in Figure 2. Since the two sites are only 11 amino acids apart, the selectivity was highly demanding for a 5 nm MINP.

We prepared MINP(β$_{1-14}$) and MINP(β$_{15-28}$), using the first and second halves of Aβ$_{1-28}$ as the template. Not only was Aβ$_{1-28}$ a disease-related peptide released through proteolysis,[19] its structure also helped us understand how MINP binding would affect *unbound* segment of a long peptide in a distance-dependent manner (see below). ITC showed that the two MINPs bound the parent peptide strongly, with $K_a = 1.97 \times 10^7$ and $3.06 \times 10^7 \, \text{M}^{-1}$, respectively (Table 1, entries 10 and 12).

In a phosphate buffer, the yields of the anticipated products (i.e., peptides **7**, **8**, and **9**) increased expectedly over a period of 4 h, when the proteolysis was complete (Figure 2 a). Also as expected, two intermediate products (i.e., peptides **10** and **11**), with only the arginine or lysine cleaved, showed transiently in the first 2 h of reaction time but disappeared when the reaction was allowed to proceed longer. The product distribution for the proteolysis in the presence of NINP was similar in shape (Figure 2 b), except that complete digestion took 6 instead of 4 h, consistent with the effect of NINP in the A-III proteolysis (Figure 1 a,b).
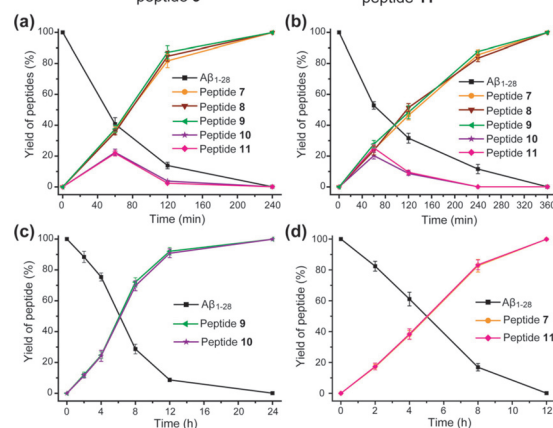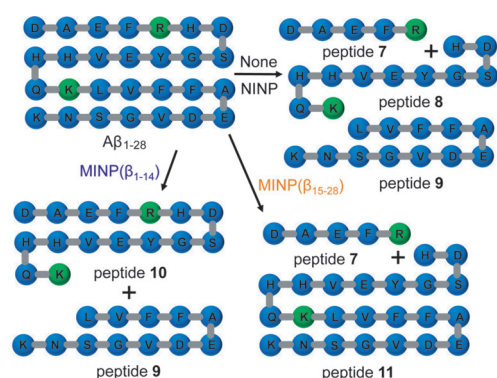

Figure 2. Product distribution curves in the trypsin digestion of Aβ$_{1-28}$ in a) buffer and in the presence of 1 equiv b) NINP, c) MINP(β$_{1-14}$), and d) MINP(β$_{15-28}$).

When Aβ$_{1-28}$ was hydrolyzed in the presence of the two MINPs, product distribution was totally different. With 1 equiv MINP(β$_{1-14}$), the arginine site was completely protected and the products observed were peptides **9** and **10** (Figure 2 c, see Figure S40 for the HPLC chromatogram), with the latter being the transient intermediate from uncontrolled hydrolysis of the parent peptide (Figure 2 a). In the presence of MINP(β$_{15-28}$), the only products were peptides **7** and **11** (Figure 2 d, Figure S41 for the HPLC chromatogram), again fully consistent with the intended protection.

It should be noted that MINP binding did slow down the cleavage of the exposed site but did not stop it, suggesting the dangling, unbound portion of the peptide was accessible to the protease. Unprotected lysine 16 in Aβ$_{1-28}$, for example, underwent trypsin cleavage in the presence of MINP(β$_{1-14}$) but the (selective) proteolysis took approximately 24 h to complete (Figure 2 c), instead of 4 h in buffer (Figure 2 a) and 6 h with NINP (Figure 2 b). Given that the cleavable lysine was only two residues away from the supposedly bound β$_{1-14}$, the precision of our protection was quite remarkable. Importantly, the proteolysis of the exposed cleavable site did seem to correlate with its distance to the bound sequence. The cleavable site at arginine 5, for example, was 9 amino acids away from β$_{15-28}$ bound by MINP(β$_{15-28}$), (selective) hydrolysis of Aβ$_{1-28}$ took approximately 12 h with this MINP (Figure 2 d).

The results so far demonstrated that MINP protection of peptides mainly depended on the binding affinity and the

distance of the cleavable site to the bound sequence. The predictability of the protection enabled us to manipulate the proteolysis of peptides in different ways, using a 2:1 mixture of A-III and $A\beta_{1-28}$ for a proof of concept. A 1:1 ratio was used for all MINP and its protected peptide. Table 1 shows that non-templating peptides displayed very low cross-reactivities (0.06–0.13%) in the binding (compare Table 1, entries 1 vs. 9, 10 vs. 14, and 12 vs. 15).

Figure 3a shows that unprotected hydrolysis of the mixture gave peptides **4**, **7**, **8**, and **9** as anticipated. In the presence of MINP(A), $7.4 \pm 1.8\%$ of the shorter A-III hydrolyzed, while the longer $A\beta_{1-28}$ underwent proteolysis nearly completely in a 99% yield (Figure 3b). This is the simplest protection one can imagine. Nonetheless, selective binding of MINPs also enabled other types of protections. When used together, MINP(A) and MINP(β_{15-28}) were

expected to protect the entire sequence of A-III and the second half of $A\beta_{1-28}$. Indeed, over a period of 12 h, only $8.8 \pm 2.1\%$ of A-III hydrolyzed while $A\beta_{1-28}$ underwent the anticipated selective cleavage after arginine 5 to afford **7** and **11** (Figure 3c).

The most interesting protection is shown in Figure 3d, when MINP(β_{1-14}) and MINP(β_{15-28}) were used simultaneously. Because of the steric hindrance anticipated for the two MINPs to bind the same peptide, we did not expect the protection would work as well as in the other scenarios. Yet, LC–MS analysis showed that, after 4 h of trypsin digestion, $87.1 \pm 4.4\%$ of A-III hydrolyzed and $A\beta_{1-28}$ stayed intact (Figure 3d).

To understand the double protection, we measured the binding between MINP(β_{1-14}) and $A\beta_{1-28}$ in the presence of 1 equivalent of MINP(β_{15-28}). Our ITC titration revealed that, when MINP(β_{1-14}) was titrated with a 1:1 mixture of $A\beta_{1-28}$ and MINP(β_{15-28}), a binding constant of 5.74×10^6 M^{-1} was obtained (Figure S16, Table 1, entry 11). The value was 3.4-times smaller than the number obtained in the absence of MINP(β_{15-28}) (Table 1, entry 10). Thus, the long peptide indeed could be bound by two MINPs simultaneously, apparently with some steric/electrostatic repulsion as judged from the weaker binding. A similar observation was made when MINP(β_{15-28}) was titrated with a 1:1 mixture of $A\beta_{1-28}$ and MINP(β_{1-14}) (compare Table 1, entry 13 with 12).

In summary, MINP protection of peptides is characterized by a broad range of substrates to be protected,[15,18] high stability of the cross-linked micelles,[13] and predictability of the protection (Figure 1c–f). Protection from common proteases can be expected as long as strong binding is obtained (e.g., $K_d < 1$ μM). Protection also depended on the location of the cleavable site, with stronger protection observed when the site was closer to the bound sequence. Cooperative protection of a peptide by two MINPs was also effective, as a result of double binding of the two segments by both MINPs.

Acknowledgements

Conflict of interest

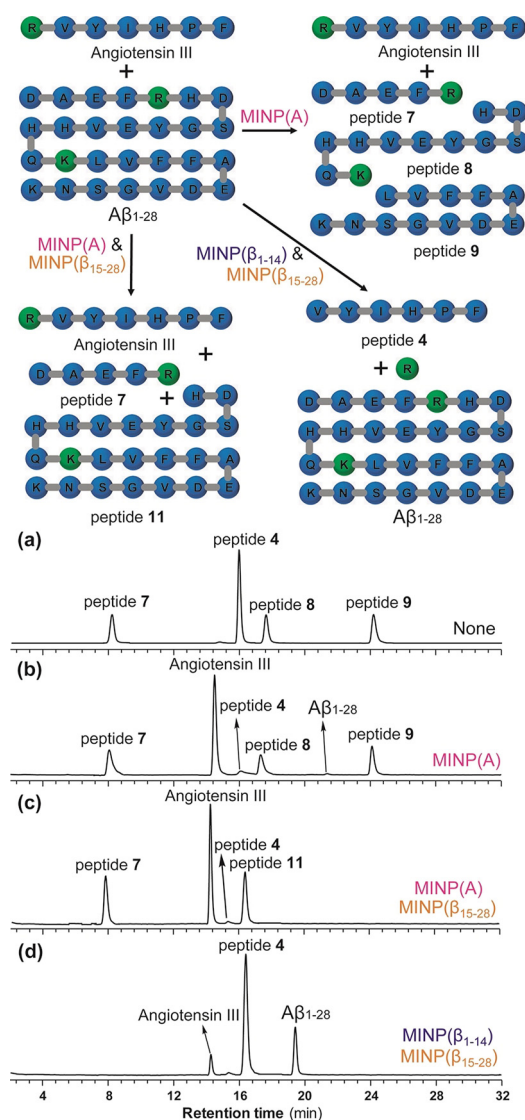The authors declare no conflict of interest.

Figure 3. HPLC chromatograms of trypsin digestion of a 2:1 mixture of Angiotensin III and $A\beta_{1-28}$ by a) trypsin 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 $A\beta_{1-28}$.

[1] P. G. M. Wuts, T. W. Greene, T. W. Greene, *Greene's protective groups in organic synthesis*, 5th ed., Wiley, Hoboken, New Jersey, **2014**.

[2] H. Kido, Y. Niwa, Y. Beppu, T. Towatari, *Adv. Enzyme Regul.* **1996**, *36*, 325–347.

[3] 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.

[4] Y. A. DeClerck, S. Imren, *Eur. J. Cancer* **1994**, *30*, 2170–2180.

[5] a) Z. Zhang, T. Ly, T. Kodadek, *Chem. Biol.* **2001**, *8*, 391–397; b) T. Kodadek, *Biopolym. Peptide Sci.* **2002**, *66*, 134–140.

[6] G. R. Sambrano, W. Huang, T. Faruqi, S. Mahrus, C. Craik, S. R. Coughlin, *J. Biol. Chem.* **2000**, *275*, 6819–6823.

[7] Z. Zhang, W. Zhu, T. Kodadek, *Nat. Biotechnol.* **2000**, *18*, 71–74.

[8] 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.

[9] a) J. U. Klein, M. J. Whitcombe, F. Mulholland, E. N. Vulfson, *Angew. Chem.* **1999**, *111*, 2100–2103; *Angew. Chem. Int. Ed.* **1999**, *38*, 2057–2060; b) H. Nishino, C. S. Huang, K. J. Shea, *Angew. Chem.* **2006**, *118*, 2452–2456; *Angew. Chem. Int. Ed.* **2006**, *45*, 2392–2396; c) Y. Hoshino, T. Kodama, Y. Okahata, K. J. Shea, *J. Am. Chem. Soc.* **2008**, *130*, 15242–15243; d) Y. Hoshino, H. Koide, T. Urakami, H. Kanazawa, T. Kodama, N. Oku, K. J. Shea, *J. Am. Chem. Soc.* **2010**, *132*, 6644–6645; e) Z. Zeng, Y. Hoshino, A. Rodriguez, H. Yoo, K. J. Shea, *ACS Nano* **2010**, *4*, 199–204; f) J. L. Urraca, C. S. A. Aureliano, E. Schillinger, H. Esselmann, J. Wiltfang, B. Sellergren, *J. Am. Chem. Soc.* **2011**, *133*, 9220–9223; g) S. Banerjee, B. König, *J. Am. Chem. Soc.* **2013**, *135*, 2967–2970; h) A. Cecchini, V. Raffa, F. Canfarotta, G. Signore, S. Piletsky, M. P. MacDonald, A. Cuschieri, *Nano Lett.* **2017**, *17*, 2307–2312.

[10] 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.

[11] F. Wang, A. Gomez-Escudero, R. R. Ramireddy, G. Murage, S. Thayumanavan, R. W. Vachet, *J. Am. Chem. Soc.* **2013**, *135*, 14179–14188.

[12] 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.

[13] J. K. Awino, Y. Zhao, *J. Am. Chem. Soc.* **2013**, *135*, 12552–12555.

[14] a) L. Duan, M. Zangiabadi, Y. Zhao, *Chem. Commun.* **2020**, *56*, 10199–10202; b) X. Li, Y. Zhao, *Chem. Sci.* **2021**, *12*, 374–383.

[15] M. Zangiabadi, Y. Zhao, *ACS Appl. Polym. Mater.* **2020**, *2*, 3171–3180.

[16] K. Chen, Y. Zhao, *Org. Biomol. Chem.* **2019**, *17*, 8611–8617.

[17] 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.

[18] a) J. K. Awino, R. W. Gunasekara, Y. Zhao, *J. Am. Chem. Soc.* **2017**, *139*, 2188–2191; b) S. Fa, Y. Zhao, *Chem. Eur. J.* **2018**, *24*, 150–158; c) S. Fa, Y. Zhao, *Chem. Mater.* **2017**, *29*, 9284–9291; d) S. Fa, Y. Zhao, *Chem. Mater.* **2019**, *31*, 4889–4896.

[19] S. S. Sisodia, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6075–6079.