A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition GDCh Chemie

www.angewandte.org

Accepted Article

Title: Molecularly Imprinted Polymer Nanogels for Protein Recognition: Direct Proof of Specific Binding Sites by Solution STD and WaterLOGSY NMR Spectroscopies

Authors: Alejandra Mier, Irene Maffucci, Franck Merlier, Elise Prost, Valentina Montagna, Guillermo U. Ruiz-Esparza, Joseph V. Bonventre, Pradeep K. Dhal, Bernadette Tse Sum Bui, Peyman Sakhaii, and Karsten Haupt

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202106507

Link to VoR: https://doi.org/10.1002/anie.202106507

WILEY-VCH

10.1002/anie.202106507

Molecularly Imprinted Polymer Nanogels for Protein Recognition: Direct Proof of Specific Binding Sites by Solution STD and WaterLOGSY NMR Spectroscopies

Alejandra Mier,[a] Irene Maffucci,[a] Franck Merlier,[a] Elise Prost,[a] Valentina Montagna,[a] Guillermo U. Ruiz-Esparza,[b] [c] Joseph V. Bonventre,[b] [c] Pradeep K. Dhal,[d] Bernadette Tse Sum Bui,*[a] Peyman Sakhaii,*[e] and Karsten Haupt,*[a]

[a] A. Mier, Dr. I. Maffucci, F. Merlier, E. Prost, Dr. V. Montagna, Dr. B. Tse Sum Bui, Prof. Dr. K. Haupt
CNRS Enzyme and Cell Engineering Laboratory
Université de Technologie de Compiègne
Rue du Docteur Schweitzer, CS 60319, 60203 Compiègne Cedex, France
E-mail: alicia.mier@utc.fr, irene.maffuci@utc.fr, franck.merlier@utc.fr, elise.prost@utc.fr, valentina.montagna@utc.fr, jeanne.tse-sum-bui@utc.fr, karsten.haupt@utc.fr

[b] Dr. G. U. Ruiz-Esparza, Prof. Dr. J. V. Bonventre
Divisions of Engineering in Medicine and Renal Medicine, Department of Medicine
Harvard Medical School, Brigham and Women's Hospital
Boston, MA, 02115, USA
E-mail: gruizeh@bwh.harvard.edu, jbonventre@bwh.harvard.edu

[c] Dr. G. U. Ruiz-Esparza, Prof. Dr. J. V. Bonventre
Division of Health Science and Technology
Harvard University - Massachusetts Institute of Technology
Cambridge, MA, 02139, USA
E-mail: gruizeh@bwh.harvard.edu, jbonventre@bwh.harvard.edu

[d] Dr. P. K. Dhal
US Early Development
Sanofi Global R&D
153 Second Avenue, Waltham, MA02451, USA
E-mail: Pradeep.Dhal@sanofi.com

[e] Dr. P. Sakhaii
R&D Global CMC Development - Synthetics - Early Development Frankfurt
Sanofi-Aventis Deutschland GmbH
Industriepark Hoechst, Frankfurt am Main, Germany
E-mail: Peyman.Sakhaii@sanofi.com

Supporting information for this article is given via a link at the end of the document

Abstract: Molecularly imprinted polymers (MIPs) are tailor-made synthetic antibodies possessing specific binding cavities designed for a target molecule. Currently, MIPs for protein targets are synthesized by imprinting a short surface-exposed fragment of the protein, called epitope or antigenic determinant. However, finding the epitope par excellence that will yield a peptide 'synthetic antibody' cross-reacting exclusively with the protein from which it is derived, is not easy. We propose a computer-based rational approach to unambiguously identify the 'best' epitope candidate. Then, using Saturation Transfer Difference (STD) and WaterLOGSY NMR spectroscopies, we prove the existence of specific binding sites created by the imprinting of this peptide epitope in the MIP nanogel. The optimized MIP nanogel could bind the epitope and cognate protein with a high affinity and selectivity. The study was performed on Hepatitis A Virus Cell Receptor-1 protein, also known as KIM-1 and TIM-1, for its ubiquitous implication in numerous pathologies.

INTRODUCTION

Molecularly imprinted polymers (MIPs), often dubbed 'synthetic antibodies' contain recognition sites which bind target molecules (the antigen) with a high affinity and selectivity.[1–3] The molecular imprinting process consists of co-polymerizing functional and cross-linking monomers in the presence of a template molecule. The template can be the target molecule or a derivative thereof. Functional monomers assemble around the template, followed by co-polymerization with a cross-linker. Subsequent removal of the template molecule from the three-dimensional polymer network reveals imprinted cavities complementary to the template in terms of size, shape, and positioning of chemical groups. Thus, a molecular memory is introduced in the imprinted polymer, enabling the recognition and binding of the target molecule at the imprinted sites.

Lately, MIPs for protein recognition have raised growing interest because of their potential applications in nanomedicine and medical diagnostics.[4–11] The major drawbacks in protein imprinting are the risk of alteration of the protein's native conformation, and the presence of its numerous functionalities for which the generation of selective imprinted sites is difficult. To address this shortcoming, epitopes, which are fragments of the protein generally exposed at its surface (in analogy to the antigenic determinant of the protein which binds to an antibody), have been proposed as templates,[12] generating MIPs capable of recognizing the entire protein.[13,14] The epitope can be linear,[12–15] *i.e.* a continuous sequence of amino

acids in the protein, or conformational, *i.e.* discontinuous amino acids that come together in a 3D conformation.[16,17] Nowadays, free online access to protein databanks and bioinformatics tools can help to predict linear and conformational epitopes[18–20] of a given protein sequence or structure.

So far, epitopes employed for MIP preparation have been selected using various approaches: (i) simply taking the C- or N-terminus of the protein,[14,21,22] (ii) extracting the antigenic region bound to the antibody from the published 3D structure of the protein-antibody complex,[23] (iii) from experimental data,[24,25] such as digesting the immobilized protein linked to the MIP followed by analyzing the fragments retained on the MIP,[25] and (iv) using a rational approach with the help of databases associated with bioinformatics tools.[26,27] The last method was demonstrated in detail for the cardiac biomarker protein, NT-proBNP, which consisted of downloading the protein sequence from UniProt, followed by *in silico* digestion with site-directed proteolytic enzymes and chemicals. The most promising resultant peptides of lengths between 7 to 12 residues, were aligned to protein sequences in the whole protein database using BLAST software, to select the peptides unique to the target protein. The chosen linear epitopes effectively produced MIPs, which recognized the digested peptide template in a complex sample, but no evidence of the MIP's ability to bind selectively the native intact protein was shown. A succinct bioinformatic approach was also reported, predicting conformational epitopes for HER2 protein,[16] a biomarker of cancer, using ElliPro[28] and PyMOL softwares on the non-digested protein. Nevertheless, no evidence that the resulting MIP could bind selectively the epitope or the whole protein was shown. Recently, a rational approach based on molecular dynamic calculations to identify stable epitopes (surface-exposed α-helices) as templates for the synthesis of electropolymerized MIP films for the recognition of neuron specific enolase, a cancer biomarker, was proposed. Binding of the MIP film to both the template peptide and the protein was demonstrated, although with very different affinity constants.[29]

Thus there is no general consensus on which method to apply, although the careful selection of the epitope par excellence remains one of the crucial elements to obtain a MIP that will solely recognize the peptide epitope and its cognate protein. The creation of imprinted sites by the peptide epitope in the MIP is fundamental to reach this goal. We use Saturation Transfer Difference (STD) and Water-Ligand Observed *via* Gradient Spectroscopy (WaterLOGSY) NMR spectroscopies, to prove for the first time, the evidence of specific binding sites created by the imprinting of the peptide epitope in a MIP nanogel. First, we describe a computer-based rational approach, outlining the essential criteria to unambiguously predict continuous epitope candidates and the 3D conformation of their structures.[18,19,30,31] The protein Hepatitis A Virus Cell Receptor-1 (HAVCR-1), also known as T-cell Immunoglobulin Mucin Receptor-1 (TIM-1) or Kidney Injury Molecule-1 (KIM-1) was chosen as target. This protein serves as a cellular entry receptor for hepatitis A virus,[32] and the filoviruses Ebolavirus[33] and Marburgvirus,[34] which cause hemorrhagic fever and death in humans. Moreover, it is overexpressed in numerous cancers,[35] and its ectodomain has been shown to be a urinary biomarker of kidney diseases.[36]

Then, a solid-phase synthesis method using the selected epitope was applied, yielding water compatible MIP nanogels (MIP-NGs) of particle size ~48 nm, which could bind the epitope and parent protein with high affinity and selectivity. We believe that our method for the identification of epitope candidates together with the study of their interaction with the resultant MIP by solution STD and WaterLOGSY NMR can be applied to a vast array of proteins present in the protein databases to identify the best epitope that will generate the best MIP-NGs. As synthetic antibody mimics, MIP-NGs[37] may serve as potential antiviral therapies by blocking the binding site of the protein, thereby preventing filoviruses from binding to host cell and cause infection. Similarly, these MIP-NGs could potentially function as anticancer therapeutics.

RESULTS AND DISCUSSION

Rational selection of epitope templates. The protein HAVCR-1/TIM-1/KIM-1 was chosen for its ubiquitous role in numerous pathologies. For simplicity, the term HAVCR-1 will be employed for this protein throughout the text. HAVCR-1 is a type 1 transmembrane glycoprotein composed of an extracellular signal peptide (SP) domain, an Ig variable domain (residues 21-121), a highly glycosylated mucin domain (residues 122-295), a transmembrane domain (TM) and a cytoplasmic tail (Figure 1A). The primary sequence of the protein (UniProtKB Q96D42) (MW: 39,250 Da) is shown in Figure 1B, and the Ig variable domain, for which there is a crystal structure resolved to 1.30 Å (PDB 5DZO), corresponding to residues 22-127, is colored in red.[33] The candidate epitopes were selected by analyzing the 3D structure, using an approach originally developed to predict antigen epitopes for raising peptide antibodies.[38,39] The following criteria were adopted: (i) they must be protruding regions of the protein,[40] exposed to solvent and accessible, (ii) they should preferably correspond to flexible loops rather than to α-helixes and β-sheets,[38,41] (iii) they must mimic the structural conformation of the fragment in the native protein, and (iv) they must be unique to the protein of interest, meaning that the sequence is not present in other proteins. Moreover, epitopes that are composed of 8 to 20 amino acid residues are preferred,[19] since shorter epitopes could lead to low specificities and longer peptides could adopt secondary structures, which could be different from the original conformation of the epitope in the native protein. Accordingly, we identified three peptide candidates, which meet these criteria. The structures of these peptides are shown in Figure 2A and Table 1.

Using these three epitopes, a second selection was carried out to assess whether they contain post-translational (glycosylation) and other amino acid modifications, which would compromise the subsequent interaction of the protein with the MIP. Glycosylation sites and Cys-Cys interactions were searched using UniprotKB Table of amino acid modifications and from glycosylation prediction servers NetOGlyc 4.0 (http://www.cbs.dtu.dk/services/NetOGlyc/) and NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/). HAVCR-1 has one N-glycosylation site at residue 65 and disulfide bonds are formed between C36 and C105, C46 and C57, as well as C52 and C104 (Table 1). The presence of disulfide bonds between the given epitope and another region of the protein would constrain the epitope within the protein in

10.1002/anie.202106507

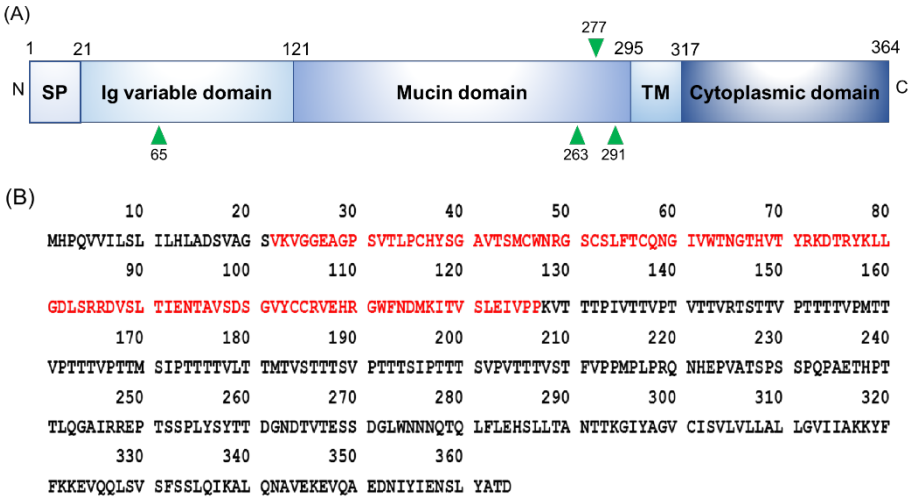


Figure 1. (A) Schematic representation of HAVCR-1 showing the signal peptide (SP), Ig variable, mucin, transmembrane (TM), and cytoplasmic domains. The triangles indicate the locations of N-linked glycosylation sites. (B) Primary sequence of the protein (UniProtKB Q96D42) with residues 22-127, whose crystal structure (PDB accession code 5DZO) is known and is highlighted in red.

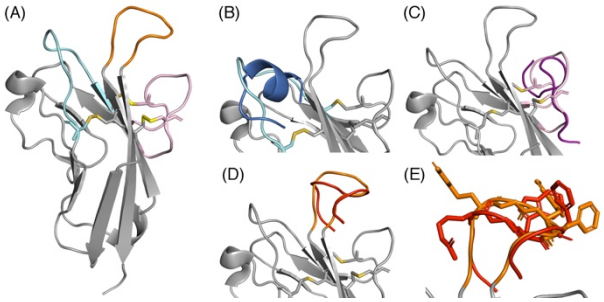


Figure 2. Crystal structure (PDB 5DZO), visualized in PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) of (A) Ig variable domain with candidate epitope peptides highlighted in light blue (P1), light violet (P2), and orange (P3); cysteine residues are shown in yellow. (B-D) Superposition of the predicted 3D structure (from PEP-FOLD3) aligned and studied in PyMOL, of P1 (dark blue), P2 (dark violet) and template peptide P3 (flanked by 2 Cys, red), onto the native protein epitopes in light blue, light violet and orange, respectively. (E) Comparison of the side chains of the structure of template peptide (red) and the fragment in the native protein (orange).

Table 1. Candidate epitope peptides corresponding to the flexible loops of the HAVCR-1 protein.

	Sequence	Amino acids UniProt	N° of amino acids	Cys bond	Glycosylation sites
P1	CHYSGAVTSMC	36-46	11	C36-C105 C46-C57	NONE
P2	RGSCSLFTCQNG	49-60	12	C52-C104	NONE
P3	EHRGWFND	108-115	8	NONE	NONE

a conformation which might not be reproduced by the free epitope. Accordingly, epitopes P1 and P2 were eliminated. This choice was further supported by analyzing the secondary structure patterns of the three peptides. The most probable 3D structures of the free epitopes were predicted with the PEP-FOLD3 server (https://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD3), to establish that, once excised from the protein, they could maintain a conformation compatible with the native one.[42] The analysis was carried out by superposing the predicted 3D structure of the fragments onto the native protein using the software PyMol (Figure 2B-D). Together with the most probable peptide 3D structure prediction, PEP-FOLD3 provides the probability of each residue to adopt a coil (*i.e.* unstructured), helix or β-sheet conformation, which are shown in Figure 3. Since the probabilities are computed using a structural alphabet of 4-residues fragments, the last 3 residues of the modelled peptide are not shown graphically. From Figure 3, we observe that P1 has preferentially an unstructured conformation at the N terminus (sequence CHYS), adopting a helical conformation afterwards, as also supported by the 3D structure of the best model (Figure 2B). P2 shows a high percentage of β-sheets, although the best model shows a small helix involving residues SLF (Figure 2C). Conversely, P3 shows a constant and high percentage of coil structure, which might match correctly the loop conformation (Figure 2D). Loops are more likely to be unstructured and therefore more flexible. Thus the resulting peptides can easily adopt the same conformation as within the protein.

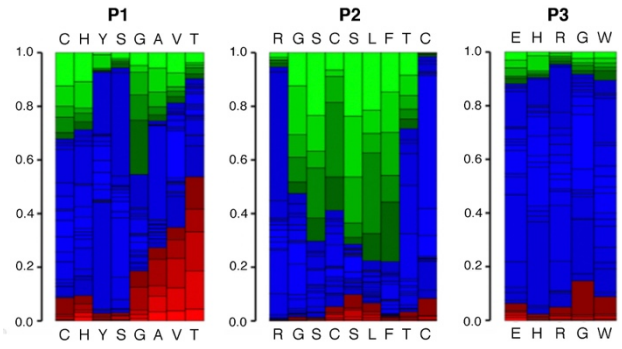


Figure 3. PEP-FOLD3 prediction of amino acids residues of P1, P2 and P3. Green: β-sheet, Blue: coil and Red: helix.

10.1002/anie.202106507

Figure 3 definitely rules out P1 and P2, thus leaving P3 as the best mimic of the structure of the corresponding flexible loop in the native protein. Additionally, the side chains of the central residues of the peptide are well solvent-exposed, similar to those within the loop of the native protein (Figure 2E). Finally, the specificities of the three epitope candidates towards the HAVCR-1 protein were investigated using the "peptide search" tool from UniProtKB (https://www.uniprot.org/peptidesearch/). This step enables to evaluate the potential cross-reactivity of MIPs with other proteins.[19] The three peptide sequences were found to be unique to the HAVCR-1 protein.

Based on the above rationale, P3 (EHRGWFND) was considered the best candidate and was selected for the preparation of a MIP. Since P3 adopts a coil conformation in the native protein (Figure 2D), a cyclic template peptide was synthesized in order to mimic closely the conformation of the epitope in its native form. Thus, the template consists of P3 with cysteine residues added at both ends for cyclization. Furthermore, a modified lysine residue bearing an azide moiety was incorporated at the C-terminal Cys to immobilize the template peptide to a solid support *via* click chemistry. The final cyclic peptide template presents the following sequence: CEHRGWFNDC-K(N_3) (C-C cyclic) (Figure 4 and Figure S1).

Solid-Phase Synthesis of MIP-NGs. The synthesis of MIP-NGs was carried out using a solid-phase approach on glass beads (GBs) as support, following previously described procedures.[10,14,43,44] Briefly, GBs were activated with NaOH, followed by surface modification using (3-aminopropyl)triethoxysilane (APTES) to introduce amino groups. The amino-functionalized GBs were reacted with (1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-ylmethyl N-succinimidyl carbonate (BCN-NHS) to introduce a strained cyclononyne. The template peptide was conjugated to the resulting GBs *via* a copper-free click chemistry reaction, the so-called strain-promoted alkyne azide cycloaddition (SPAAC), forming a stable triazole[45,46] (Figure 4). The copper-catalyzed azide-alkyne cycloaddition (CuAAC) click chemistry reaction that we used previously for the immobilization of templates to GBs[14,47] was not suitable here, as the reducing environment of the CuAAC reaction led to the cleavage of ~ 35% of the disulfide bonds of the cyclic template peptide. The amount of immobilized cyclic peptide was determined by the quantification of the sulfhydryl groups of the cysteine residues with Ellman's reagent after the disulfide bonds of the peptide were reduced with tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (Figure S2). From a calibration curve of cysteine (Figure S3), the amount of immobilized template was found to be 86 ± 3 nmol per gram of GBs. A monomer mixture comprising *N*-isopropylacrylamide (NIPAM),

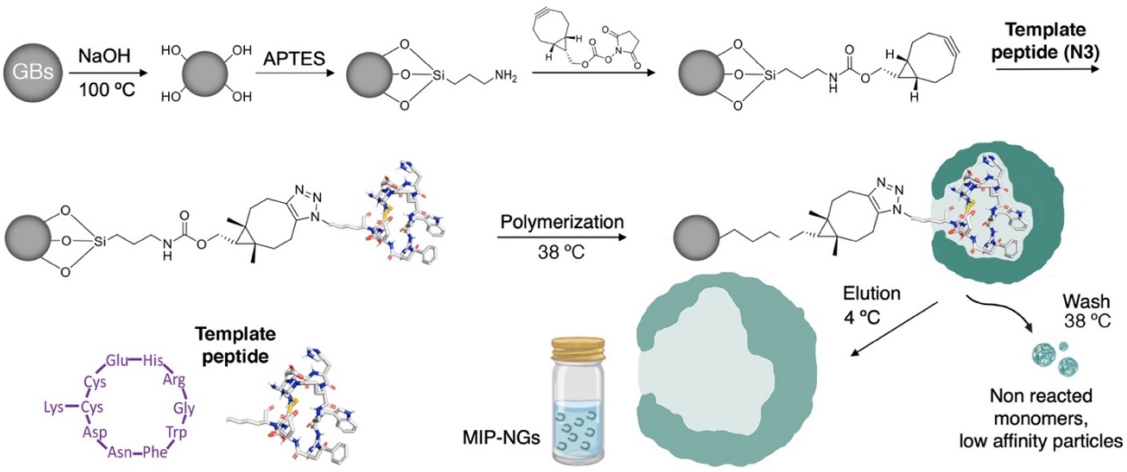


Figure 4. Solid-phase synthesis of MIP-NGs. The template peptide CEHRGWFNDC-K(N_3) is immobilized onto the alkyne-derivatized GBs *via* strain-promoted alkyne azide cycloaddition, followed by polymerization to form thermoresponsive MIP-NGs (green) around the peptide at 38 °C. After washing away the non-reacted monomers and low-affinity polymers, the MIP-NGs are released from the solid-phase by a temperature change from 38 °C to 4 °C. The 3D model of the template peptide was generated using PEP-Fold 3.

N-phenylacrylamide (PAA), *N*-*tert*-butylacrylamide (TBAM) and 4-acrylamidophenyl(amino)-methaniminium acetate (AB) was used for the synthesis of MIP-NGs.[14] The benzamidine group of AB is known to form salt bridges with carboxylates,[47–49] and can thus associate with aspartate and glutamate of the template peptide, in polar solvents. PAA provides π-π interactions, TBAM hydrophobic interactions and NIPAM hydrogen bonding.[50] NIPAM was added at a high molar ratio (80%) to impart thermoresponsive characteristics to the MIP-NGs, which facilitates the subsequent elution of the polymer from the solid support. *N*,*N*'-methylenebis(acrylamide) (BIS) (5% molar ratio) was used to provide low crosslinking. The MIP was synthesized by free radical polymerization at 38 ºC in 25 mM sodium phosphate buffer, pH 7.0 (Buffer A), overnight. At this temperature, the polymers adopt a collapsed state, encapsulating the immobilized peptide template. After polymerization, the reactor was washed with Buffer A at 38 ºC to remove any unreacted reagents and low affinity polymers. The high affinity MIP-NGs were eluted with water at 4 ºC, which is below the lower critical solution temperature of the polymers (LCST of pNIPAM ~32 °C).[50] A control polymer (CONT-NGs), which is a MIP targeting a different peptide sequence, SLAPAEG, was also synthesized with similar monomer combination, under similar conditions.

10.1002/anie.202106507

Physicochemical characterization of polymers. The hydrodynamic size (Figure S4) and zeta potential of MIP-NGs and CONT-NGs were determined by DLS measurements at 38 ºC. The results along with the concentration of the polymers and yield of synthesis are summarized in Table S1.

Binding specificity of MIP-NGs. The interaction of MIP-NGs with the peptide template was evaluated by equilibrium binding assays at 38 °C. To monitor binding, the peptide epitope was fluorescently labeled with fluorescein dibenzocyclooctyne (FAM DBCO), *via* SPAAC (details of synthesis and fluorescence characterizations in Figures S5-S7). Equilibrium binding isotherms showed that MIP-NGs bind specifically to the fluorescent template, as no binding was observed with CONT-NGs (Figure 5). These results suggest the creation of imprinted sites at the surface of the nanogels. Additionally, we showed the importance of the functional monomers for interaction with the template. Indeed, the suppression of the hydrophobic monomers, TBAM or PAA led to polymers with lower binding, although PAA seems to be more favorable than TBAM, probably because it can form π–π interactions with Trp and Phe. Most notably, omission of AB resulted in no binding, reflecting how AB is essential for interaction (Figure 5).

Binding selectivity of MIP-NGs. The interactions of MIP-NGs with recombinant HAVCR-1 protein, the template peptide CEHRGWFNDC-K(N_3) and other peptides, were determined by competitive binding assays. We employed a recombinant HAVCR-1 protein expressed in human embryonic kidney (HEK293) cells to reflect more likely the relevant glycosylations present in the wild-type native protein. Competing peptides were the linear epitope EHRGWFND, as well as three non-related peptides, one linear, DWVIPPI[14] and two cyclic, CFGRVMQIGSRC-K(N_3) and CMDYKGSYLC-K(N_3) (Table 2). The non-related peptides were chosen for their similar MWs or pIs and for the presence of some of the important amino acids involved in MIP binding. Increasing concentrations of free competitors competed with a fixed amount of fluorescent epitope peptide (6 nM) bound to 25 μg/mL of MIP-NGs, to determine their affinity. A non-linear regression fit of the curves (Figure S8) was used to determine the IC_{50} (the concentration of competing ligand required to displace 50% of fluorescent template peptide from the MIP-NGs). The IC_{50} values for the template peptide and the HAVCR-1 protein were found to be 13 and 14 nM respectively, indicating high affinity of the MIP-NGs towards the epitope and the protein. This confirms our hypothesis from the molecular modeling step concerning the choice of the epitope. The MIP-NGs also exhibited a high affinity (IC_{50} 14 nM) towards the linear form of the template peptide, EHRGWFND, suggesting that the linear non-constrained peptide may adopt a conformation in solution close to that of the cyclic template, possibly induced by the MIP-NGs. Previous studies have shown that MIPs prepared with a cyclic peptide as template can facilitate the protein folding of an unstructured linear peptide to its cyclic form.[51,52] The cross-reactivities of the MIPs towards the non-template peptides DWVIPPI, CFGRVMQIGSRC-K(N_3) and CMDYKGSYLC-K(N_3) were negligible, reflecting the high selectivity of the MIP-NGs.

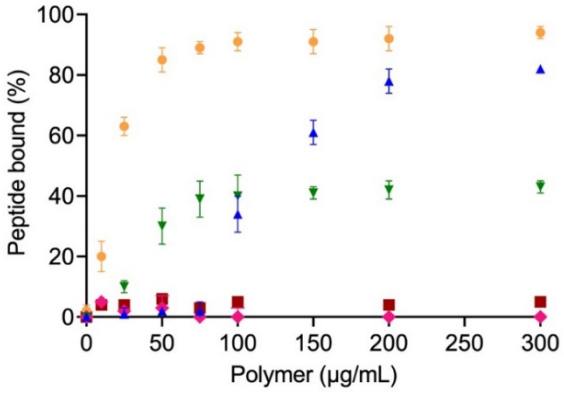


Figure 5. Equilibrium binding isotherms of fluorescently-labeled peptide (6 nM) and MIP-NGs (orange), CONT-NGs (red), MIP-NGs without PAA (green), MIP-NGs without TBAM (blue) and MIP-NGs without AB (pink), in 25 mM sodium phosphate buffer pH 7, 38 ºC.

Table 2. IC_{50} and cross-reactivity values of different peptides and HAVCR-1 protein towards MIP-NGs (n=4).

	MW (g/mol)	pI	IC_{50} (nM)	Cross-reactivity (%)
Template cyclic peptide CEHRGWFNDC-K(N_3)	1417.6	5.2	13	100
Linear peptide EHRGWFND	1059.5	5.2	14	93
Recombinant HAVCR-1 protein	~90,000[a]	6.6	14	93
Linear peptide DWVIPPI	838.5	3.1	> 1000	< 1
Cyclic peptide CFGRVMQIGSRC-K(N_3)	1507.7	9.7	> 1000	< 1
Cyclic peptide CMDYKGSYLC-K(N_3)	1333.6	5.9	> 1000	< 1

[a] Glycosylated protein expressed in HEK293 cells, as given by supplier.

NMR spectroscopy. In order to confirm the specific binding of MIP-NGs to its template peptide, we employed STD and WaterLOGSY NMR spectroscopies as label-free complementary analytical tools. Both NMR methods have been successfully used to study ligand binding to high molecular weight species like proteins, viruses and DNA.[53,54] This prompted us to explore the potential utility of these NMR techniques to investigate binding selectivity of MIP-NGs towards the template. While STD is more adapted for non-polar interactions, WaterLOGSY is used to assess polar interactions.

STD NMR. STD NMR spectroscopy consists of selectively irradiating the macromolecule, in our case MIP-NGs, with a weak ^{1}H saturation radiofrequency pulse train causing the saturation to propagate by proton spin diffusion across the macromolecular network. In case of MIP-peptide interaction, the saturation is transferred from MIP to peptide, with the peptide protons in closest contact with the MIP receiving the most saturation (ON resonance experiment). The transmitter frequency of the saturation pulse is carefully chosen to irradiate

10.1002/anie.202106507

only the MIP's protons without affecting those of the peptide. Towards that end, a careful search for a non-overlapping region is required. Generally, two experiments are recorded, ON and OFF resonance. In our experiments, the ON resonance irradiation of the MIP backbone was set to 1.1 ppm. The OFF resonance transmitter frequency was set to the dark spectral region, where usually no proton signal exists (-40 ppm) (see SI for details). The difference between ON and OFF resonance STD spectra reveals only signals from the peptide that binds to the MIP.

Semi solid-state STD high resolution magic angle spinning NMR has been used in the past to study the differences in the interaction of a low molecular weight ligand bupivacaine (MW 288 Da), with a monolithic MIP and NIP.[55] In this study, due to the excellent solubility of the MIP-NGs in aqueous medium, solution NMR could be used, which affords significantly improved spectral resolution with higher sensitivity. As can be seen from Figure 6B, a positive STD response is evident, suggesting an interaction of the cyclic template peptide, CEHRGWFNDC-K(N_3) with MIP-NGs, whose profile matches with the ^{1}H NMR spectrum of the cyclic peptide (Figure S10A). To ensure that no peptide irradiation was taking place, a control STD experiment with peptide alone was carried out (Figure 6A). As expected, no significant STD signal intensities was detected. A similar binding pattern was observed with the corresponding linear peptide EHRGWFND-K(N_3) (Figure 6E and Figure S10B), which confirms our hypothesis that upon binding, MIP-NGs can induce the linear peptide into a cyclic conformation. The control experiment using CONT-NGs showed no significant STD signal intensities, suggesting that the binding of the cyclic template peptide to the MIP-NGs is highly selective (Figure 6C). Similarly, no difference spectrum was observed when the cyclic non-template peptide CFGRVMQIGSRC-K(N_3) was studied with MIP-NGs (Figure 6G).

An additional experiment to confirm the specific binding of MIP-NGs to the cyclic template was performed by exploiting the thermoresponsiveness of the polymer. As presented in Figure S11, the STD NMR spectrum recorded at 290 K (below LCST of the polymer) showed significantly reduced STD response (see for aromatic region 7-8 ppm), compared to the STD spectrum recorded at 312 K. The rationale behind this observation is attributed to the fact that at a temperature below the LCST, the spatial structure of the MIP is swollen, and as a result, peptide binding to the MIP is lost.

WaterLOGSY NMR. To further support the STD NMR results, WaterLOGSY NMR spectroscopy was employed. In this technique, water signal is used to propagate the dipolar cross relaxation to ligands and receptors.[56,57] Water molecules act as a mediator for the transfer of magnetization between ligand and receptor *via* the nuclear Overhauser effect (NOE). Since intermolecular NOE transfer is favored for functional groups with high water affinity, this NMR technique is particularly useful for enhanced detection of binding from hydrophilic compartments of a ligand including exchangeable proton sites. WaterLOGSY NMR spectra are generally analyzed by comparing the strengths and signs of signal intensities of free ligand with those of the receptor-bound ligand (details in SI). Direct comparison of both spectra provides structural insights about the bound state of the ligand.

At first, a reference WaterLOGSY of the cyclic template peptide was recorded and the sign and strength of signals were determined. Subsequently, a mixture of template peptide and MIP-NGs was subjected to WaterLOGSY analysis under the same experimental conditions. The WaterLOGSY NMR spectrum of the free template peptide is shown in Figure 7A. In this case, proton signal intensities in downfield regions are negative. Upon addition of MIP-NGs to the template, the downfield ^{1}H signal intensities associated with the peptide are flipped positive in the same direction as the residual MIP signal, marked by asterisks, indicating a clear binding event (Figure 7B). A similar pattern was observed with the linear peptide (Figures 7D and E). In order to unequivocally attribute the observed phenomenon to template binding to the MIP-NGs, a control experiment was performed with a mixture containing the cyclic template and CONT-NGs (Figure 7C). Spectral analyses revealed no changes to the signs of ^{1}H signals in the downfield regions compared to those observed in the presence of MIP-NGs (Figure 7B). Furthermore, no changes to the signs of signals of the protons in the downfield regions of the non-template cyclic peptide CFGRVMQIGSRC-K(N_3) were observed in the presence of MIP-NGs (Figure 7F and Figure 7G), suggesting peptide binding by MIP-NGs is highly selective towards its template. These systematic experiments further enhanced the level of confidence of these NMR methods to assess template selective binding features of MIP-NGs.

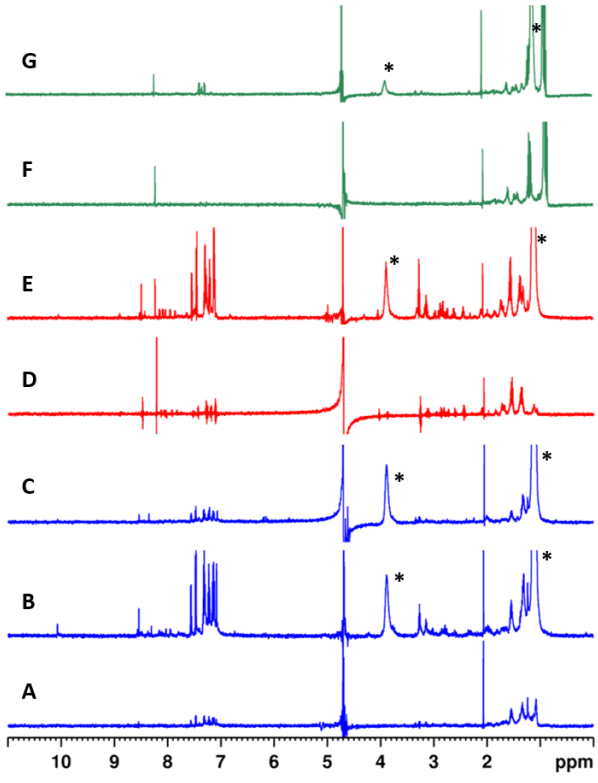


Figure 6. STD NMR spectra at 312 K of (A) cyclic template alone, (B) cyclic template + MIP-NGs, (C) cyclic template + CONT-NGs, (D) linear peptide alone, (E) linear peptide + MIP-NGs, (F) non-template cyclic peptide alone and (G) non-template cyclic peptide + MIP-NGs. Spectra were obtained by applying a gaussian saturation pulse train for 5s and subsequent excitation sculpting water suppression.[58] A 100 ms spin lock filter was applied in the STD pulse sequence to suppress residual MIP signals in the final spectra. ON resonance irradiation of the MIP was set to 1.1 ppm and OFF resonance irradiation was set to -40 ppm. Proton resonances marked with asterisks (*) originate from MIP signal contributions outbreaking the spin lock filter.

10.1002/anie.202106507

Taken together, the STD and WaterLOGSY NMR results clearly show the specific interaction of the entire target peptide epitope with the MIP-NGs. Looking at the spectral peak assignments of ^{1}H NMR spectrum of the peptide template (Figure S10A), the contributions of the amino acids to the spectral change in the presence of the MIP can be clearly observed with one or the other method (or both), depending on the hydrophobic or hydrophilic nature of the residue. The ligand chemical groups in direct contact with the MIP show the stronger STD signals. Noteworthy are the strong signals observed for the peak cluster between 7 and 8 ppm in Figure 6B, attributed mainly to Trp (7.56/7.47) and Phe (7.31), which indicates their close proximity to the binding groups in the imprinted sites of the MIP. Similarly, the inversion of the sign of the ^{1}H peaks at 10.06, 9.05, 8.58, 8.10, and 7.40/6.74 ppm in the WaterLOGSY spectra (Figure 7B) suggests the involvement of Trp, Glu, His, Asp and Asn, respectively, in the binding process. In the two methods, the contribution of the Trp side chain seem predominant and can be attributed to the apical position of this amino acid in the template peptide (see Figure S1 for a 3D model), causing it to be deeply buried within the binding site. These spectral changes were not observed during binding with a non-template cyclic peptide or with the compositionally similar CONT-NGs lacking specific binding sites for the template. Finally when the MIP-NGs were treated with the template peptide below the LCST, no binding event was evident from these NMR studies (Figure S11).

CONCLUSION

MIP-NGs for the selective recognition of HAVCR-1 protein were synthesized by using a solid-phase epitope imprinting method. The epitope was selected *in-silico* by using a computer-based rational approach, with the main guidelines being that it has a flexible solvent-exposed loop structure (vs helix or β-sheet structure), it has a 3D structure with similar structural conformation as that of the native protein, and it is unique to the protein. The selected template peptide was synthesized and was immobilized on a solid support (GBs) by strain-promoted alkyne azide cycloaddition (SPAAC). Such linking chemistry has been introduced for the first time to immobilize templates for solid-phase molecular imprinting. This approach enabled us to achieve an oriented immobilization of the template, upon which thermoresponsive MIP-NGs were synthesized. After polymerization, the MIP-NGs were released from the support by a simple temperature change, thus generating template-free polymers in a single step. The binding sites of the resulting MIPs have similar orientation, and could thus be considered analogous to monoclonal antibodies, explaining their high affinity and selectivity towards the template and the whole HAVCR-1 protein.

In order to confirm and gain further insights on the specific binding of MIP-NGs with their template peptide, STD and WaterLOGSY solution NMR spectroscopies were employed. The combined strengths of STD and WaterLOGSY experiments to study MIP-peptide interaction enabled us to obtain rapid access to detailed binding information, with only a small quantity of MIP (1-2 mg/experiment) and without the need for any labeling chemistry. The results clearly show that all amino acids of the peptide epitope were involved in MIP binding. WaterLOGSY was found to be particularly valuable for the detection of hydrophilic features of the peptide, since water interacts more specifically with polar groups like carboxylic acids or amines. On the other hand, STD NMR is more suited for the investigation of hydrophobic parts of the peptide, since in this method, the saturation pulses are invariant with respect to affinity towards water molecules. The results show that both methods can discriminate between binding and non-binding peptides at atomic level resolution, opening new opportunities for these label-free analytical tools to investigate detailed molecular recognition information of MIP-NGs.

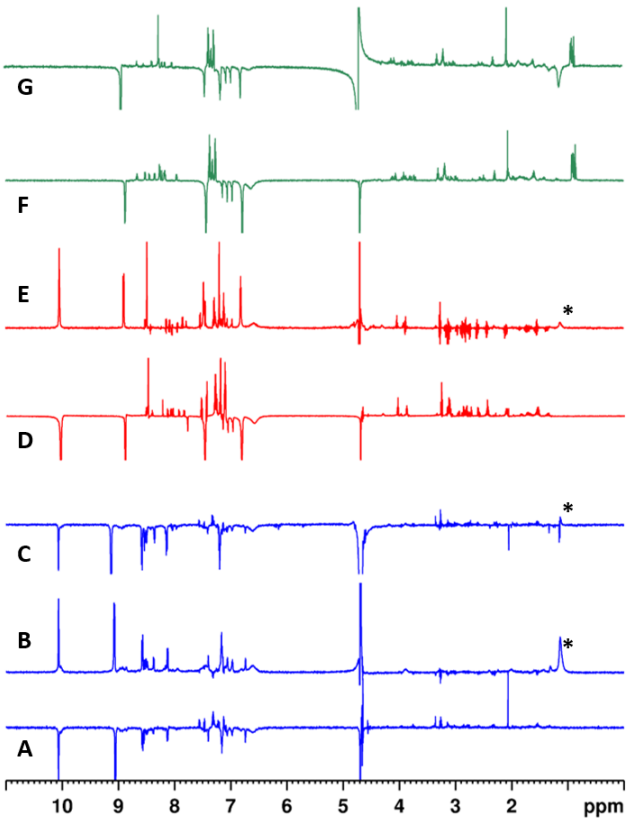


Figure 7. WaterLOGSY NMR spectra of (A) cyclic template alone, (B) cyclic template + MIP-NGs, (C) cyclic template + CONT-NGs, (D) linear peptide alone (E) linear peptide + MIP-NGs, (F) non-template cyclic peptide alone and (G) non-template cyclic peptide + MIP-NGs. Spectra were obtained by application of WaterLOGSY pulse cascade. A mixing time of 4s was used to propagate the intermolecular dipolar cross relaxation of proton spins. Upon ligand binding, for (B) and (E), the sign of ^{1}H signals in downfield spectral regions are flipped positive. Asterisks (*) represent residual MIP-NG signals outbreaking the spin-lock filter.

ACKNOWLEDGMENTS

A.M. thanks the Mexican National Council for Science and Technology (CONACYT) and the Instituto de Innovación y Transferencia de Tecnología de Nuevo Leon for PhD scholarship. The authors acknowledge financial support from the Region of Picardy and the European Union (co-funding of equipment under CPER 2007–2020 and project POLYSENSE). K.H. acknowledges financial support from Institut Universitaire de France. J.V.B. is supported by US NIH grants R37D039773 and R01DK072381.

Keywords: molecularly imprinted polymer • rational design • epitope • NMR • HAVCR-1

REFERENCES

[1] G. Vlatakis, L. I. Andersson, R. Müller, K. Mosbach, *Nature* **1993**, *361*, 645–647.

[2] K. Haupt, A. V Linares, M. Bompart, B. Tse Sum Bui, *Top. Curr. Chem.* **2012**, *325*, 1–28.

[3] B. Sellergren, *Molecularly Imprinted Polymers : Man-Made Mimics of Antibodies and Their Applications in Analytical Chemistry*, Amsterdam ; New York : Elsevier, **2001**.

[4] W. Chen, Y. Ma, G. Pan, in *Theranostic Bionanomaterials Micro Nano Technol.* (Eds.: W. Cui, X. Zhao), **2019**, pp. 113–142.

[5] T. Takeuchi, Y. Kitayama, R. Sasao, T. Yamada, K. Toh, Y. Matsumoto, K. Kataoka, *Angew. Chem.* **2017***, 129, 7194–*7198; *Angew. Chem. Int. Ed.* **2017**, *56*, 7088–7092.

[6] K. Haupt, P. X. Medina Rangel, B. Tse Sum Bui, *Chem. Rev.* **2020**, *120*, 9554–9582.

[7] H. Zhang, *Adv. Mater.* **2020**, *32*, 1806328.

[8] J. Xu, H. Miao, J. Wang, G. Pan, *Small* **2020**, *16*, 1906644.

[9] M. Gast, H. Sobek, B. Mizaikoff, *Trends Anal. Chem.* **2019**, *114*, 218–232.

[10] J. Xu, F. Merlier, B. Avalle, V. Vieillard, P. Debré, K. Haupt, B. Tse Sum Bui, *ACS Appl Mater Interfaces* **2019**, *11*, 9824–9831.

[11] I. A. Nicholls, J. G. Wiklander, *Aust. J. Chem.* **2020**, *73*, 300–306.

[12] A. Rachkov, N. Minoura, *Biochim. Biophys. Acta* **2001**, *1544*, 255–266.

[13] H. Nishino, C.-S. Huang, K. J. Shea, *Angew. Chem. Int. Ed.* **2006**, *45*, 2392–2396.

[14] P. X. Medina Rangel, E. Moroni, F. Merlier, L. A. Gheber, R. Vago, B. Tse Sum Bui, K. Haupt, *Angew. Chem.* **2020**, *132,* 2838–2844; *Angew. Chem. Int. Ed.* **2020**, *59*, 2816–2822.

[15] D.-F. Tai, Y.-F. Ho, C.-H. Wu, T.-C. Lin, K.-H. Lu, K.-S. Lin, *Analyst* **2011**, *136*, 2230–2233.

[16] H. Hashemi-Moghaddam, S. Zavareh, S. Karimpour, H. Madanchi, *React. Funct. Polym.* **2017**, *121*, 82–90.

[17] Y. Zhang, C. Deng, S. Liu, J. Wu, Z. Chen, C. Li, W. Lu, *Angew. Chem.* **2015***, 127,* 5246–5249; *Angew. Chem. Int. Ed.* **2015**, *54*, 5157–5160.

[18] J. V. Ponomarenko, M. H. V. Van Regenmortel, in *Struct. Bioinforma.* (Eds.: J. Gu, P.E. Bourne), John Wiley & Sons, **2009**, pp. 847–877.

[19] N. H. Trier, G. Houen, *Adv. Clin. Chem.* **2017**, *81*, 43–96.

[20] W. Fleri, K. Vaughan, N. Salimi, R. Vita, B. Peters, A. Sette, *J. Immunol. Res.* **2017**, *2017*, 1–13.

[21] R. Xing, Y. Ma, Y. Wang, Y. Wen, Z. Liu, *Chem. Sci.* **2019**, *10*, 1831–1835.

[22] D. Dechtrirat, K. J. Jetzschmann, W. F. M. Stöcklein, F. W. Scheller, N. Gajovic-Eichelmann, *Adv. Funct. Mater.* **2012**, *22*, 5231–5237.

[23] F. Canfarotta, L. Lezina, A. Guerreiro, J. Czulak, A. Petukhov, A. Daks, K. Smolinska-Kempisty, A. Poma, S. Piletsky, N. A. Barlev, *Nano Lett.* **2018**, *18*, 4641–4646.

[24] H. Bagán, T. Zhou, N. L. Eriksson, L. Bülow, L. Ye, *RSC Adv.* **2017**, *7*, 41705–41712.

[25] S. Piletsky, E. Piletska, F. Canfarotta, D. Jones, *Methods and Kits for Determing Binding Sites*, **2018**, WO 2018/178629A1.

[26] L. Pasquardini, A. M. Bossi, *Anal. Bioanal.Chem.* **2021**, doi.org/10.1007/s00216-021-03409-1.

[27] A. M. Bossi, P. S. Sharma, L. Montana, G. Zoccatelli, O. Laub, R. Levi, *Anal. Chem.* **2012**, *84*, 4036–4041.

[28] J. Ponomarenko, H.-H. Bui, W. Li, N. Fusseder, P. E. Bourne, A. Sette, B. Peters, *BMC Bioinformatics* **2008**, *9*, 514.

[29] Z. Altintas, A. Takiden, T. Utesch, M. A. Mroginski, B. Schmid, F. W. Scheller, R. D. Süssmuth, *Adv. Funct. Mater.* **2019**, *29*, 1807332.

[30] D. C. Hancock, N. J. O'Reilly, in *Methods Mol. Biol.*, **2005**, *295* 13–25.

[31] T. M. Shinnick, J. G. Sutcliffe, N. Green, R. A. Lerner, *Annu. Rev. Microbiol.* **1983**, *37*, 425–446.

[32] S. Jemielity, J. J. Wang, Y. K. Chan, A. A. Ahmed, W. Li, S. Monahan, X. Bu, M. Farzan, G. J. Freeman, D. T. Umetsu, R. H. Dekruyff, H. Choe, *PLoS Pathog.* **2013**, *9*, e1003232.

[33] S. Yuan, L. Cao, H. Ling, M. Dang, Y. Sun, X. Zhang, Y. Chen, L. Zhang, D. Su, X. Wang, Z. Rao, *Protein Cell* **2015**, *6*, 814–824.

[34] A. S. Kondratowicz, N. J. Lennemann, P. L. Sinn, R. A. Davey, C. L. Hunt, S. Moller-Tank, D. K. Meyerholz, P. Rennert, R. F. Mullins, M. Brindley, L. M. Sandersfeld, K. Quinn, M. Weller, P. B. J. McCray, J. Chiorini, W. Maury, *Proc. Natl. Acad. Sci.* **2011**, *108*, 8426–8431.

[35] E. Telford, W. Jiang, T. Martin, *Histol. Histopathol.* **2017**, *32*, 121–128.

[36] W. K. Han, V. Bailly, R. Abichandani, R. Thadhani, J. V. Bonventre, *Kidney Int.* **2002**, *62*, 237–244.

[37] P. Liu, C. M. Pearce, R.-M. Anastasiadi, M. Resmini, A. M. Castilla, *Polymers* **2019**, *11*, 353.

[38] N. D. Rubinstein, I. Mayrose, D. Halperin, D. Yekutieli, J. M. Gershoni, T. Pupko, *Mol. Immunol.* **2008**, *45*, 3477–3489.

[39] M. H. V. Van Regenmortel, *J. Mol. Recognit.* **2006**, *19*, 183–187.

[40] J. M. Thornton, M. S. Edwards, W. R. Taylor, D. J. Barlow, *EMBO J.* **1986**, *5*, 409–413.

[41] N. H. Trier, P. R. Hansen, G. Houen, *Methods* **2012**, *56*, 136–144.

[42] P. Thévenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux, P. Tufféry, *Nucleic Acids Res.* **2012**, *40*, W288-W293.

[43] S. Ambrosini, S. Beyazit, K. Haupt, B. Tse Sum Bui, *Chem. Commun.* **2013**, *49*, 6746–6748.

[44] J. Xu, S. Ambrosini, E. Tamahkar, C. Rossi, K. Haupt, B. Tse Sum Bui, *Biomacromolecules* **2016***, 17*, 345–353.

[45] Y. Zhang, C. Fang, R. E. Wang, Y. Wang, H. Guo, C. Guo, L. Zhao, S. Li, X. Li, P. G. Schultz, Y. J. Cao, F. Wang, *Proc. Natl. Acad. Sci.* **2019**, *116*, 15889–15894.

[46] R. van Geel, M. A. Wijdeven, R. Heesbeen, J. M. M. Verkade, A. A. Wasiel, S. S. van Berkel, F. L. van Delft, *Bioconjug. Chem.* **2015**, *26*, 2233–2242.

[47] P. X. Medina Rangel, S. Laclef, J. Xu, M. Panagiotopoulou, J. Kovensky, B. Tse Sum Bui, K. Haupt, *Sci. Rep.* **2019**, *9*, 3923.

[48] M. Panagiotopoulou, Y. Salinas, S. Beyazit, S. Kunath, A. G. Mayes, L. Duma, E. Prost, M. Resmini, B. Tse Sum Bui, K. Haupt, *Angew. Chem.* **2016**, *128,* 8384–8388; *Angew. Chem. Int. Ed.* **2016**, *55*, 8244–8248.

[49] S. Nestora, F. Merlier, S. Beyazit, E. Prost, L. Duma, B. Baril, A.

10.1002/anie.202106507

Greaves, K. Haupt, B. Tse Sum Bui, *Angew. Chem.* **2016**, *128,* 6360–6364; *Angew. Chem. Int. Ed.* **2016**, *55*, 6252–6256.

[50] R. Pelton, *J. Colloid Interface Sci.* **2010**, *348*, 673–674.

[51] L. Cenci, G. Guella, E. Andreetto, E. Ambrosi, A. Anesi, A. M. Bossi, *Nanoscale* **2016**, *8*, 15665–15670.

[52] S. Liu, Q. Bi, Y. Long, Z. Li, S. Bhattacharyya, C. Li, *Nanoscale* **2017**, *9*, 5394–5397.

[53] M. Mayer, B. Meyer, *Angew. Chem.* **1999**, *111,* 1902–1906; *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 1784–1788.

[54] A. J. Benie, R. Moser, E. Bäuml, D. Blaas, T. Peters, *J. Am. Chem. Soc.* **2003**, *125*, 14–15.

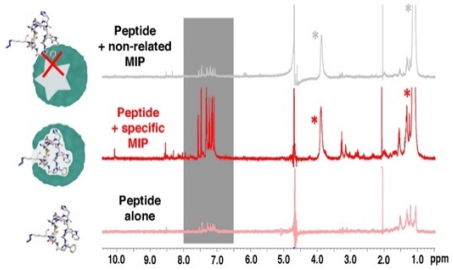
[55] J. Courtois, G. Fischer, S. Schauff, K. Albert, K. Irgum, *Anal. Chem.* **2006**, *78*, 580–584.

[56] C. Dalvit, P. Pevarello, M. Tatò, M. Veronesi, A. Vulpetti, M. Sundström, *J. Biomol. NMR* **2000**, *18*, 65–68.

[57] C. Raingeval , O. Cala , B. Brion, M. Le Borgne , R. E. Hubbard, I. Krimm, *J. Enzyme Inhib. Med. Chem.* **2019**, *34*, 1218–1225.

[58] T. L. Hwang, A. J. Shaka, *J. Magn. Reson. Ser. A* **1995**, *112*, 275–279.

Entry for the Table of Contents



MIPs for protein recognition are often synthesized by the epitope approach. The epitope is a short surface-exposed peptide of the protein. Identification of the 'best' epitope will generate a selective MIP. Solution STD and WaterLOGSY NMR spectroscopies were used to prove the evidence of specific binding sites created by the imprinting of the peptide epitope in the corresponding MIP nanogel. No binding was observed with a non-related peptide.