

Cite this: *Chem. Sci.*, 2022, 13, 210

All publication charges for this article have been paid for by the Royal Society of Chemistry

## Discovery of phosphotyrosine-binding oligopeptides with supramolecular target selectivity†

Ana S. Pina,<sup>id</sup>\*<sup>abc</sup> Leonor Morgado,<sup>id</sup><sup>bc</sup> Krystyna L. Duncan,<sup>ad</sup> Sara Carvalho,<sup>id</sup><sup>bc</sup> Henrique F. Carvalho,<sup>id</sup><sup>bc</sup> Arménio J. M. Barbosa,<sup>id</sup><sup>bc</sup> Beatriz de P. Mariz,<sup>bc</sup> Inês P. Moreira,<sup>bc</sup> Daniela Kalafatovic,<sup>a</sup> Bruno M. Morais Faustino,<sup>id</sup><sup>e</sup> Vishal Narang,<sup>a</sup> Tong Wang,<sup>af</sup> Charalampos G. Pappas,<sup>id</sup><sup>ad</sup> Isabel Ferreira,<sup>id</sup><sup>e</sup> A. Cecília A. Roque<sup>id</sup>\*<sup>bc</sup> and Rein V. Ulijn<sup>id</sup>\*<sup>agh</sup>

We demonstrate phage-display screening on self-assembled ligands that enables the identification of oligopeptides that selectively bind dynamic supramolecular targets over their unassembled counterparts. The concept is demonstrated through panning of a phage-display oligopeptide library against supramolecular tyrosine-phosphate ligands using 9-fluorenylmethoxycarbonyl-phenylalanine-tyrosine-phosphate (Fmoc-FpY) micellar aggregates as targets. The 14 selected peptides showed no sequence consensus but were enriched in cationic and proline residues. The lead peptide, KVYFSIPWRVPM-NH<sub>2</sub> (P7) was found to bind to the Fmoc-FpY ligand exclusively in its self-assembled state with  $K_D = 74 \pm 3 \mu\text{M}$ . Circular dichroism, NMR and molecular dynamics simulations revealed that the peptide interacts with Fmoc-FpY through the KVYF terminus and this binding event disrupts the assembled structure. In absence of the target micellar aggregate, P7 was further found to dynamically alternate between multiple conformations, with a preferred hairpin-like conformation that was shown to contribute to supramolecular ligand binding. Three identified phages presented appreciable binding, and two showed to catalyze the hydrolysis of a model *para*-nitro phenol phosphate substrate, with P7 demonstrating conformation-dependent activity with a modest  $k_{\text{cat}}/K_M = 4 \pm 0.3 \times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$ .

Received 12th August 2021  
Accepted 5th December 2021

DOI: 10.1039/d1sc04420f

rsc.li/chemical-science

## Introduction

Oligopeptides are involved in many functions of biological relevance, including self-assembly, molecular recognition and catalysis, holding much promise as designed components for molecular biotechnology and biomimetic materials research.<sup>1–6</sup>

This notion has inspired efforts to design functional peptide modalities as functional (gene-encodable) tags for protein-based materials. Approaches for the discovery of short functional peptide modules include rational (computational) design,<sup>7–9</sup> combinatorial screening methods,<sup>10,11</sup> dynamic peptide libraries,<sup>12</sup> phage display<sup>13–16</sup> and hybrid computational-experimental methods<sup>17,18</sup> that may be supported by machine learning algorithms.<sup>19–21</sup> Phage display has successfully led to the identification of binding sequences towards protein surfaces,<sup>22–24</sup> to a range of organic and inorganic nanostructures<sup>13,15,16,25–29</sup> and small molecules bound to a surface, sometimes *via* flexible linkers.<sup>14–16,30–32</sup> The identification of short peptides that select for self-assembled (amyloid) structures<sup>33,34</sup> or precursors for catalytic self-assembly<sup>15</sup> has clearly shown that short peptides can selectively bind stable supramolecular structures over unassembled counterparts.

In this work, we have developed methodology that brings together phage display with the field of designed supramolecular materials. We focus on an example from a popular class of supramolecular materials where aromatic groups are used to aid the self-assembly of biomolecules, here exemplified by 9-fluorenylmethoxycarbonyl-phenylalanine-tyrosine-phosphate (Fmoc-FpY). Although there are affinity reagents (*e.g.*,

<sup>a</sup>Advanced Science Research Center (ASRC) at the Graduate Center, City University of New York (CUNY), NY 10031, USA. E-mail: ana.pina@fct.unl.pt

<sup>b</sup>Associate Laboratory i4HB – Institute for Health and Bioeconomy, NOVA School of Science and Technology, NOVA University Lisbon, 2829-516 Caparica, Portugal. E-mail: cecilia.roque@fct.unl.pt

<sup>c</sup>UCIBIO – Applied Molecular Biosciences Unit, Department of Chemistry, School of Science and Technology, NOVA University of Lisbon, 2829-516 Caparica, Portugal

<sup>d</sup>Department of Pure & Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow G1 1XL, UK

<sup>e</sup>CENIMAT/I3N, Department of Materials Science, School of Science and Technology, NOVA University of Lisbon, 2829-516 Caparica, Portugal

<sup>f</sup>Imaging Facility of CUNY ASRC, 85 St Nicholas Terrace, New York 10031, USA

<sup>g</sup>Hunter College of CUNY, Department of Chemistry and Biochemistry, 695 Park Avenue, New York 10065, USA

<sup>h</sup>PhD Programs in Chemistry and Biochemistry, The Graduate Center of CUNY, New York 10016, USA. E-mail: rulijn@gc.cuny.edu

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1sc04420f



antibodies/SH2 domains) that recognize phosphotyrosine moieties with high affinity and selectivity,<sup>35–38</sup> the methodology proposed herein targets self-assembled structures that present the ligand of interest, the phosphotyrosine moiety, on a dynamic supramolecular surface. Although there is no known biological equivalent of this phosphorylated self-assembled material, the high-density presentation of functional groups is common in several supramolecular materials approaches. For our system, we show a first example where we target tyrosine-phosphate moieties exclusively in the self-assembled state.

The ability to identify peptide sequences that can bind to specific moieties presented by designed dynamic supramolecular structures can be useful in several contexts. For example, these peptide sequences can be incorporated into proteins that then bind to designed supramolecular structures; or identified peptide modalities may inform multi-component self-assembly strategies of relevance to dynamic structures that may incorporate significant disorder. For these, reversibility and dynamics can be important, so modest  $K_D$  values can be beneficial over the tight binding structures typically sought.

We also hypothesized that selection for supramolecular ligands may give rise to peptides that bind to ligands only in the supramolecular context; such an approach would be useful to target chemical functional groups to both biological and synthetic supramolecular structures.

To demonstrate our concept, we chose Fmoc-FpY as a simple target that was previously shown to form micellar aggregates<sup>39</sup> through aggregation of the hydrophobic Fmoc-moieties and consequent presentation of phosphotyrosine on the surface. At 20 mM Fmoc-FpY, forms aggregated micelles with a size distribution between 50–200 nm in diameter (Fig. 1 and S3†).

A commercially available dodecapeptide M13 phage display library was used to screen against self-assembled Fmoc-FpY (Fig. 1a) and subsequently tested for binding this target and for the ability to hydrolyze the phosphate-ester bond (Fig. 1b and c). The selected phages were tested for binding and chemical reactivity, and the binding mode and efficiency of the top three lead sequences alongside appropriate controls were analyzed by spectroscopy techniques and molecular dynamics (MD) simulations. A number of strategic mutants of the lead P7 peptide sequence were also investigated to gain understanding of the mode of interaction.

## Results and discussion

Phage display peptide libraries were panned against Fmoc-FpY above the critical micelle concentration (CMC) of 20 mM.<sup>39</sup> During panning, the self-assembled Fmoc-FpY target was free in solution, thus maximizing interactions between phages and targets, which contrasts previous biopanning strategies where self-assembling targets were typically pre-immobilized onto surfaces (*e.g.* polystyrene plates, glass substrates).<sup>33,34</sup>

A total of 33 colonies were obtained from the 3<sup>rd</sup> round output (Table S1†), where an enrichment of binders was observed. The 33 phage colonies were subsequently sequenced and analyzed leading to 14 new peptide sequences (Table 1). While there was no sequence consensus in the hits obtained,

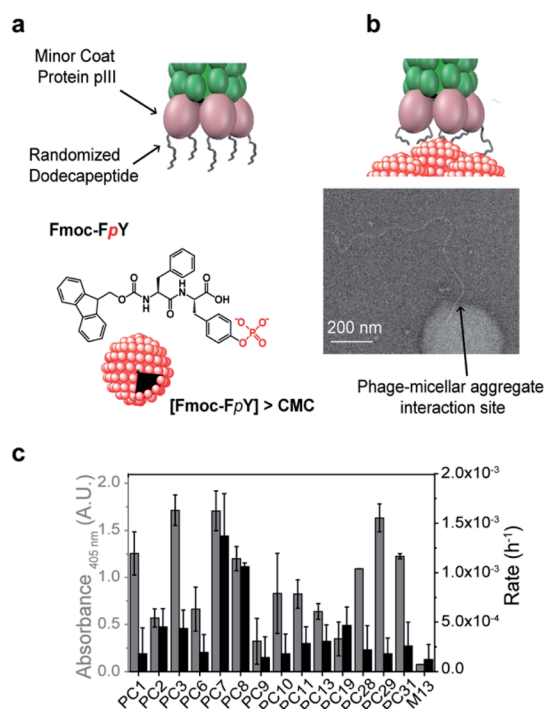


Fig. 1 Phage display panning with supramolecular target selectivity. (a) Schematic representation of the M13 phage and phosphotyrosine supramolecular target, Fmoc-FpY, (b) TEM imaging of the interaction of the tip of the phage of the lead phage clone PC7 towards Fmoc-FpY (c) Phage clone data regarding binding assay by the colorimetric ELISA and phosphatase activity using the phosphatase model substrate *p*-nitrophenol phosphate (pNPP).

sequences PC11 and PC13 were repeated four and two times, respectively. The chemical diversity obtained provides general features of binders.

The phage panning led to the selection of 14 dodecapeptide sequences (Table 1, Fig. S1a†). The diversity of the amino acids incorporated at each of the 12 positions of these 14 sequences was first analyzed by the ratio between the frequency of each different amino acid at the *n* randomized position of the 12-mer

Table 1 Phage clones sequences identified in the phage display panning

Phage clone	Sequence	# repetitions
PC1	STVRPLLMDKY	1/33
PC2	HHLRIPYALDQT	1/33
PC3	DSAPSYNYRPSY	1/33
PC6	DYHDPSPPTLRK	1/33
PC7	KVYFSPWRVPM	1/33
PC8	QVNLGERSQQM	1/33
PC9	HSNDPRLITMRK	1/33
PC10	TCFAHTHNNFGH	1/33
PC11	DYHDPSPPTLRK	4/33
PC13	GNNPLHVHDKR	2/33
PC19	RDYHPRDHTATW	1/33
PC28	IPGTAPPLARTG	1/33
PC29	KDFLPSPQTATW	1/33
PC31	VRAFSGEHSFVS	1/33

sequence over the total 20 amino acids. This value was further normalized by the value 0.7 (ratio 14/20), corresponding to maximum of 14 different amino acids possible at  $n$  position (Fig. S1b†). Fig. S1b† shows that all positions of the 14 lead peptide sequences present an amino acid probability higher than 50%, except for position 5, which indicates a preferred type of amino acid at this position. The distribution of the amino acids observed at the 12 randomized positions for all the 14 peptide sequences was then analyzed and plotted in Fig. S1c.† Proline amino acid is statistically relevant specially between the positions at the middle of the sequence at positions 5 (43%), 7 (30%) and 9 (35%).

Although there is an overabundance of proline in the phage-displayed peptide population of the Ph.D-12™ libraries, this statistic overrepresentation is typically observed at positions 3 and 12,<sup>40</sup> which contrasts with the results obtained here. As discussed later, the incorporation of a proline at the middle-sequence is important for peptide conformation,<sup>5,41,42</sup> giving rise to  $\beta$ -turns, which was found to be important for the induced binding towards supramolecular phosphorylated-based aggregates.

All sequences (except for PC10) had at least one positively charged residue (R/K) or dyad RK/KR. Cationic amino acids are preferentially located at both termini as shown in Fig. S1c,† with a higher abundance at C-terminus, suggesting electrostatic interactions with the phosphate moiety. A distribution of negatively charged residues such as aspartic acid is also observed, near the N-terminus, which can counterbalance the presence of the positive charges in the peptide sequence. According to literature, the charged residues (R/K/D) are underrepresented in these type of phage-displayed peptides<sup>40</sup> and these residues are typically associated with phosphate-binding.<sup>43</sup>

Another observation is the preferential presence of aromatic residues such as tyrosine and phenylalanine, at positions 3 and 4. This is interesting from the perspective of supramolecular aggregation interference, considering that the inner part of aggregates is enriched by aromatic fluorenyl, phenyl and phenol functional groups. Other amino acids with a frequency higher than 50% are found in a distributed manner throughout the peptide sequences, without position dependence or enrichment. These include histidine, hydrophobic residues such as leucine, alanine, valine, or polar residues such as threonine, glycine and serine. Overall, the selected peptides had several common compositional features with proline, aromaticity and charge playing a role in the selection.

The analysis of phage clones binding to supramolecular Fmoc-FpY particles involved incubation of each amplified phage clone with Fmoc-FpY at panning conditions, which was followed by buffer washing to avoid non-specific interactions and subsequent transfer of the phage/Fmoc-FpY containing solutions to a 96-microplate for ELISA assay, to quantify the supramolecular target-bound phages (Fig. S2a†).

PC3, PC7 and PC29 were found to be the top binders towards Fmoc-FpY (Fig. 1c), and interestingly, PC3 and PC7 present the highest combination of proline and aromatic residues with two and three respectively, followed by the PC29 with both two

proline and aromatic residues. In contrast, the weakest binder, PC9, does not present any aromatic residue and only one proline residue. A common feature between all the peptide sequences is the presence of positively charged residues regardless of the extent of binding. In fact, PC9 is the lowest binder and the one containing the highest number of R/K residues which suggests non-specific binding towards Fmoc-FpY.

Interactions of phages towards Fmoc-FpY aggregates could also be imaged by TEM (Fig. 1b shows binding of the tip of amplified PC7 to the supramolecular target. Several dozen TEM images were obtained of samples of the PC7 and phage control M13 with Fmoc-FpY). For PC7, we observed several examples of phage termini interacting with Fmoc-FpY particles (additional examples are shown in Fig. S3†), while for the M13 phage none showed evidence of interaction.

Although it was not selected for in our screening setup, as hydrolysis of Fmoc-FpY could not be detected, the phosphatase-like catalytic activity of the identified phage clones was also tested using *para*-nitrophenol phosphate (pNPP) with PC7 and PC8 observed to be significantly above background (Fig. 1c and S2b†).

Peptides P3, P7, and P29 were synthesized by solid-phase peptide synthesis and used in their free form in the following studies aiming to quantify binding and understand the interaction with the target. The P9 peptide sequence corresponding to the weakest binding phage was also studied as a negative control. The mode of binding of the free peptides P3, P7, P29 and P9 towards assembled Fmoc-FpY was carried out in the same manner by using 1D<sup>1</sup>H Nuclear Magnetic Resonance (NMR). A solution of 20 mM Fmoc-FpY was titrated with increasing concentrations of the respective peptide in the range 50–1000  $\mu$ M and the chemical shift changes of the Fmoc-FpY proton signals were monitored (as exemplified for P7 and Fmoc-FpY in Fig. S4–S6†). The signals of the peptide could not be resolved because of the difference in concentrations between Fmoc-FpY and peptides, as exemplified for P7 in Fig. S4.†

Analysis of the Fmoc-FpY chemical shifts perturbation as a function of the peptide's concentration yielded binding curves that suggest a level of cooperativity in the interactions (Fig. 2 and S7†) for all the peptides except for P9 (Fig. S7†). The data were fitted to the Hill equation  $\Delta\delta = [\text{peptide}_{\text{free}}]^n / (K_D^n + [\text{peptide}_{\text{free}}]^n)$ , where  $K_D$  is the dissociation constant and  $n$  the Hill coefficient that describes peptide cooperativity (Fig. 2). The  $K_D$  values that were obtained from the fitting of the peptide-induced chemical shift perturbation curves (Fig. 2a, S7† and Table 2) suggest the formation of a supramolecular-binding complex between Fmoc-FpY and the peptides P3, P7 and P29 but with different  $K_D$  as shown in Table 2.

P7 was revealed to have the highest affinity towards Fmoc-FpY with a  $K_D$  of  $74 \pm 3 \mu$ M, which is in the same order of magnitude compared to the values obtained for the peptides with affinity for previously reported self-assembling structures.<sup>33,34</sup>

The data obtained for P3 and P29 showed comparable cooperative binding profiles as for P7 but with  $K_D$  values that were an order magnitude higher than P7 (Fig. 2, S7† and Table



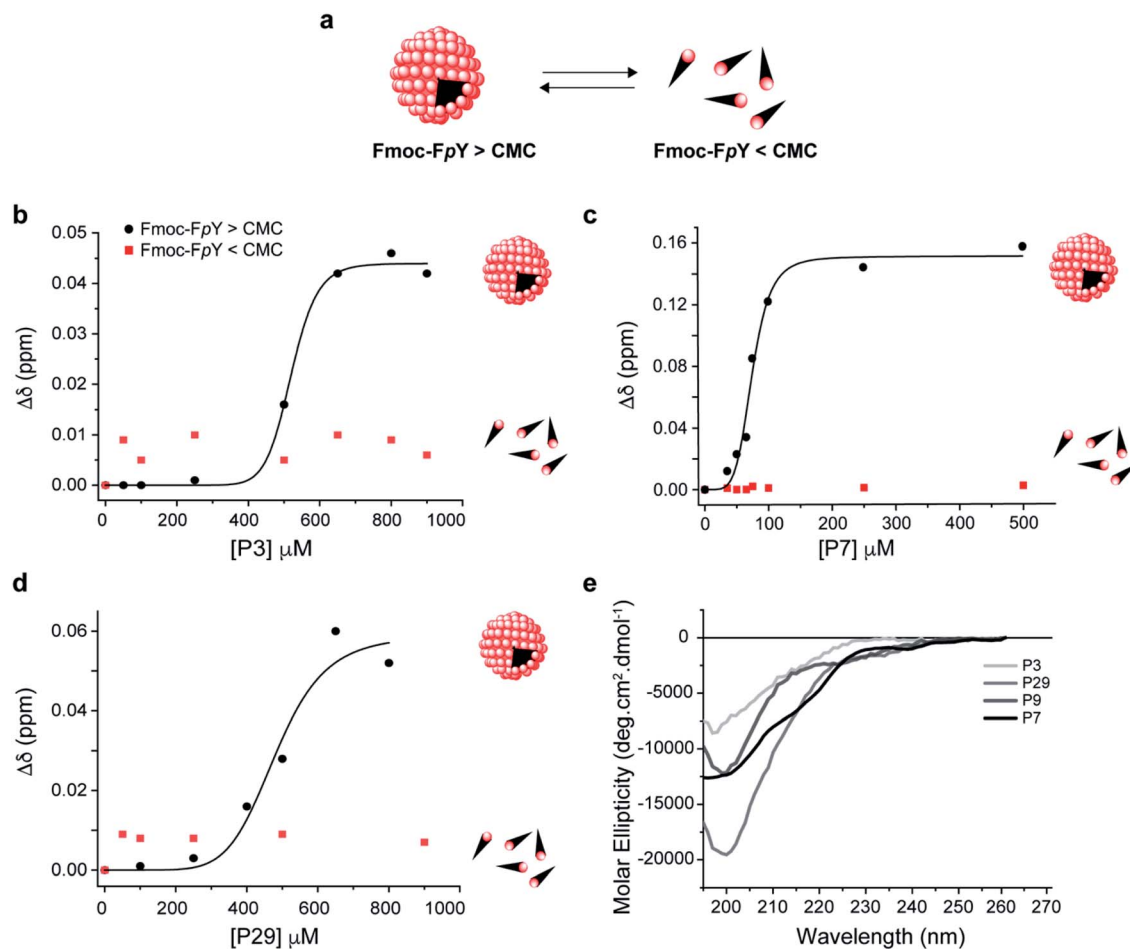


Fig. 2 Supramolecular recognition of the top binding peptide (P3, P7 and P29) towards (a) Fmoc-FpY and conformation analysis of the peptides. Binding curve for Fmoc-FpY (1 mM, 20 mM) and (b) P3 (c) P7, and (d) P29. Interaction from 1D <sup>1</sup>H NMR Chemical shift changes of the most downfield Fmoc-FpY aromatic proton signal are shown. CD spectra of (e) P3, P7, P29 and P9 in 10 mM sodium phosphate pH 8.0. The fitting details can be found in the ESI.†

Table 2 The binding affinity for the various dodecapeptides for Fmoc-FpY determined by NMR titration experiments

Peptide	Sequence	$K_D$ (μM)	$r^2$
P3	DSAPSYNYRPSY-NH <sub>2</sub>	520 ± 7	0.99
P7	KVYFISIPWRVPM-NH <sub>2</sub>	74 ± 3	0.96
P29	KDFLPSPQTATW-NH <sub>2</sub>	488 ± 33	0.95
P9	HSNDPRLITMRK-NH <sub>2</sub>	n.d.	—

2), revealing lower affinity of the P3 and P29 for the assembled target. In contrast, the P9 concentration range used was not sufficient to reach saturation of the complex, and therefore a value of  $K_D$  was not possible to be determined (Fig. S8†). These results are in accordance with the panning results. Effectively, the peptide with the highest binding affinity towards the supramolecular Fmoc-FpY aggregates was P7, which corresponds to the lead phage clone binder.

To test whether the peptides P3, P7 and P29 for Fmoc-FpY differentiate between the self-assembled target and the free molecules, the titration was repeated at a concentration where

Fmoc-FpY is not assembled (1 mM) (Fig. 2 and S9†). At this concentration, no chemical shift perturbations were observed, confirming that P3, P7 and P29 displays supramolecular target selectivity (Fig. 2). The shape of the curves for all peptides also indicates a mechanism that is similar to positive allosteric cooperativity ( $n > 1$ ) (Fig. 2 and S7†), suggesting that binding of one peptide molecule facilitates the binding of additional molecules. For the three peptides, the CH and aromatic protons of the Fmoc group (4.1 and 7.2–7.9 ppm, respectively) showed the highest chemical shift perturbation, followed by the H $\alpha$  of the phenylalanine (3.5 ppm), and the H $\delta$  of phosphorylated tyrosine (7 ppm) (Fig. S5 and S7†). These observations reveal that the Fmoc group experiences a more dramatic change in its chemical environment upon binding when compared with the more exposed aromatic protons of the tyrosine (Fig. S7†), suggesting that peptide binding interferes with the Fmoc stacking within the micellar aggregates. This structural disruption is expected to play a role in the observed allosteric cooperativity.

To understand the impact of the peptides' conformation on the mode of interaction, circular dichroism (CD) studies were performed for all the peptide sequences (Fig. 2d). The CD



spectra revealed that all peptides appear to adopt a random coil conformation with a minimum around 200 nm, except for P7 that revealed a  $\beta$ -hairpin-like conformation,<sup>44</sup> with a minimum at 196 nm, a small shoulder at 210 nm, and a weak positive band around 230 nm (Fig. 2d).

P7 peptide is the lead binder with a  $\beta$ -hairpin-like conformation, which can indicate that the binding interaction can benefit from a defined conformation.

Inspection of the P7 sequence KVFYSPWRPM-NH<sub>2</sub> reveals that the optimal combination between the proline positioning, aromaticity, and the N-terminal sequence (KVYF) that contains both cationic and aromatic residue can provide a complement of the Fmoc-FpY structure. Therefore, the respective mode of interaction of P7 with Fmoc-FpY was then studied in more detail.

To understand the importance of the phosphate-moiety in mediating the binding mechanism towards Fmoc-FpY, a control experiment was performed to assess binding between P7 and Fmoc-FY. Considering the propensity of Fmoc-FY to undergo gelation,<sup>39</sup> the samples of Fmoc-FY were prepared at lower concentrations, specifically at below and above the critical assembly concentration. Concentrations were chosen according to the emission spectra of Fmoc-FY where fibers form without gelation (1 mM), and in absence of the fibers formed (0.1 mM) (Fig. S10a†). Fiber formation was also confirmed by TEM in Fig. S10.† Considering these results, the study of the binding of P7 towards Fmoc-FY was conducted by NMR below (0.1 mM) and above (1 mM) the critical assembly concentration. 1D <sup>1</sup>H NMR chemical shift changes of the most downfield Fmoc-FpY aromatic proton signal were monitored to understand the mode of interactions of the lead peptide P7 and Fmoc-FY (Fig. S11†). The results shown in Fig. S12† revealed that in the absence of the phosphate-moiety, the binding event between P7 and Fmoc-FpY does not occur for both Fmoc-FY concentrations tested, which implies a pivotal role for the phosphate-moiety at mediating the binding with P7, possibly through the positively charged residues R and K.

The proposed hairpin-like P7 conformation in solution was also observed in MD simulations. As may be expected for a short peptide, P7 samples multiple conformations throughout the simulations (3 × 250 ns). However, for a significant (30%) part of the simulations the peptide adopts a hairpin-like conformation (Fig. 3a) which corroborates the CD spectra (Fig. 2b). This P7 conformation is stabilized by a cooperative exchange of backbone hydrogen bonds (Fig. 3b, S13 and S14†), and the key interactions are summarized in (Fig. S13 and S15†). Root mean square fluctuations (RMSF) of the C $\alpha$  atoms from each residue shows that the core of the hairpin (residues 4–9, FSIPWR) is less dynamic compared to the termini (Fig. S16†).

Mutational analysis of the leading P7 peptide sequence (Fig. S15†) were then studied by 1D<sup>1</sup>H NMR to assess the contribution of each residue to the dynamic hairpin conformation of P7 (Fig. 3c) and to understand in more detail the mechanism of interaction of the supramolecular complex Fmoc-FpY/P7 (Fig. 3d). Due to the precipitation of certain mutants at high concentrations (K1A, P7A, R9A and, P11A), this comparison was carried out at 75  $\mu$ M where all mutants were

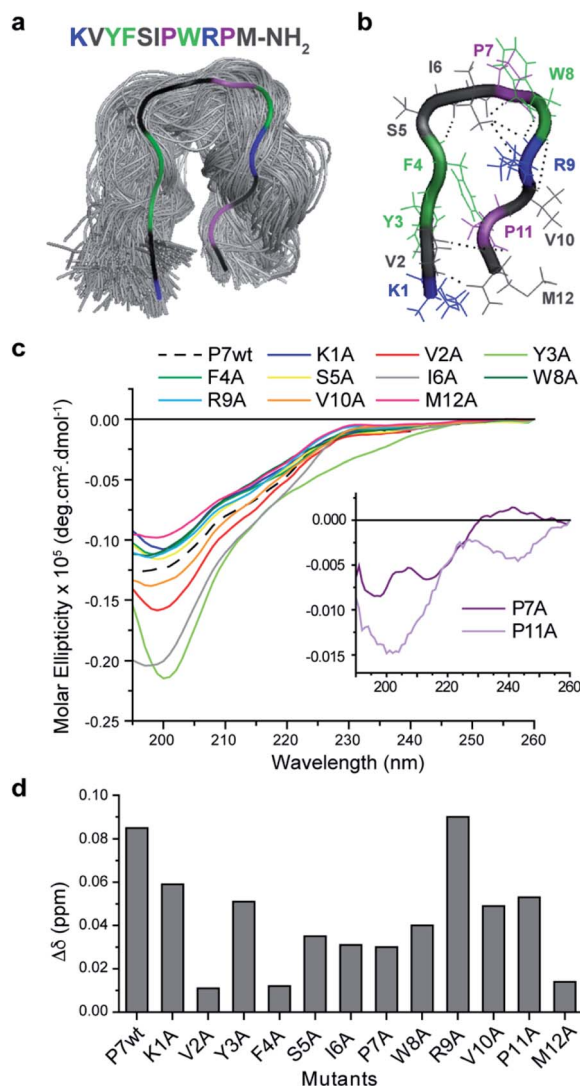


Fig. 3 Conformational analysis of P7 and the respective alanine mutants (a) top cluster (30%) from MD simulations of P7 showing a  $\beta$ -hairpin-like structure maintained by specific (b) H-bonds. (c) CD spectra of P7 and Ala mutants in 10 mM sodium phosphate pH 8.0 and their (d) influence on the supramolecular binding complex of P7/Fmoc-FpY evaluated by chemical shifts changes at the aromatic proton of the Fmoc group (7.9 ppm) of 20 mM Fmoc-FpY.

soluble. Lower chemical shift differences are related to a decrease in the binding efficiency, and the results indicate that all residues contribute to binding, except for R9 (Fig. 3d). Overall, the KVYF terminus (in particular V and F) of the P7 has a stronger impact on binding compared to the C-terminal of P7, further confirming a role of aforementioned complementarity with Fmoc-FpY. The core of the hairpin (4–8, FSIPW) is also critical for binding, thought to provide structural stability to the hairpin.

According to the CD spectra shown in Fig. 3c, the mutations that mostly affect P7 conformation are V2A, Y3A, I6A, P7A, and P11A, revealing their importance in maintaining the stability of the hairpin framework, V2, Y3, and I6 are involved in hydrogen bonding (Fig. S13 and S14†) and CH $\alpha$  interactions (Fig. S15†)



stabilizing the hairpin-like conformation. Moreover, these residues present lower RMSF values, indicating their contribution to lower flexibility of the hairpin conformation. Also, mutants P7A and P11A revealed a complete loss of hairpin-like structure, confirming the role of these Pro residues in stabilizing the conformation (Fig. 3c).

The NMR results, together with the evaluation of the mutants' conformation by CD, suggest that P7's hairpin-like conformation is essential for the formation of the supramolecular Fmoc-FpY/P7 complex. We propose a mechanism involving selective binding towards the assembled state through interaction with the KVFY terminus which is facilitated by the hairpin-conformation and leads to partial disruption of the supramolecular aggregates. A similar induced-binding mechanism has been previously shown for particle-bound UTP-ligands<sup>14</sup> and has been suggested for a heptapeptide bound to self-assembled peptide nanofibers.<sup>34</sup>

Finally, the catalytic activity that was observed during screening, and in the phage clone PC7 was investigated for the free oligopeptide P7 using the model substrate pNPP (Fig. 4a). The dependence of the initial rate of the reaction on substrate

concentration follows Michaelis–Menten model, with  $k_{\text{cat}} = 10^{-5} \text{ s}^{-1}$ ,  $K_{\text{M}} = 14 \pm 4 \text{ mM}$ , and catalytic efficiency of  $k_{\text{cat}}/K_{\text{M}} = 4 \pm 0.3 \times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$ . We note that the catalytic efficiency of P7 compares favourably to previously reported systems for phosphate ester hydrolysis, including a designed  $\beta$ -hairpin loop based on peptide–nucleic acid conjugate,<sup>45</sup> a 42-residue helix-loop-helix peptide motif<sup>46</sup> and peptide amphiphiles<sup>47</sup> that showed catalytic activity when assembled into peptide nanofibers. A common feature between all is the presence of defined structural elements, reinforcing the importance of folding for chemical catalysis. An additional observation was that P7 concentration influences its catalytic rate, with a dramatic decrease in the catalytic rates observed as the concentration of P7 increases (Fig. 4b).

Coinciding with this reduction in catalytic activity is a conformational change of the peptide (Fig. S18†). The CD signature of the hairpin-like peptide changes as P7 concentration increases, with the minimum band blue-shifted to 202 nm and the weak positive band slightly red-shifted showing the maximum of the positive band at 230 nm and the disappearance of the shoulder at 210 nm (Fig. S16†). At concentrations higher than 100  $\mu\text{M}$ , P7 revealed a CD signature of PP-II conformation, which is characterized by a negative band around 200–205 nm and by a weaker positive band around 225 nm (ref. 48 and 49) and expected to be enabled by the two proline residues. Also, at higher concentrations, P7 self-assembles into spherical aggregates, which were further confirmed using diffusion-ordered spectroscopy (DOSY) and imaged by TEM and AFM, shown in S19.†

This P7 self-assembly is related to a loss of  $\beta$ -hairpin-like structure and a conformational change of P7 to PP-II conformation, acting also as a conformation switch in the catalytic mechanism. The results suggest that the secondary framework of P7 of hairpin-like is the catalytic active conformation, as reported also in other peptide-based catalysts.<sup>45,46,50</sup> We note that the hairpin to PP-II conformational switch observed at higher concentration was an unexpected result that was unrelated to the panning conditions, but an unintended consequence of the proline-rich sequence that was selected. The catalytic mechanism of P7 is currently not understood and will be the subject of future studies. We note that the catalytic activity for the supramolecular target Fmoc-FpY (instead of the *para*-nitrophenol substrate) could not be conclusively demonstrated because the  $k_{\text{cat}}$  of P7 was not sufficient to hydrolyse Fmoc-FpY.

While molecular recognition is the main driving during selection, the phosphatase activity was also used during the screening process. Although the catalytic activity was not sufficient to hydrolyse the Fmoc-FpY phosphate ester bond, as initially intended and previously observed for an amide condensation target,<sup>15</sup> the modest activity observed, and supramolecular regulation of this activity are interesting features to be exploited for further applications.

The peptide combines a remarkable combination of features in one sequence: binding to a supramolecular target, modest catalytic activity, and concentration-dependent supramolecular reconfiguration.

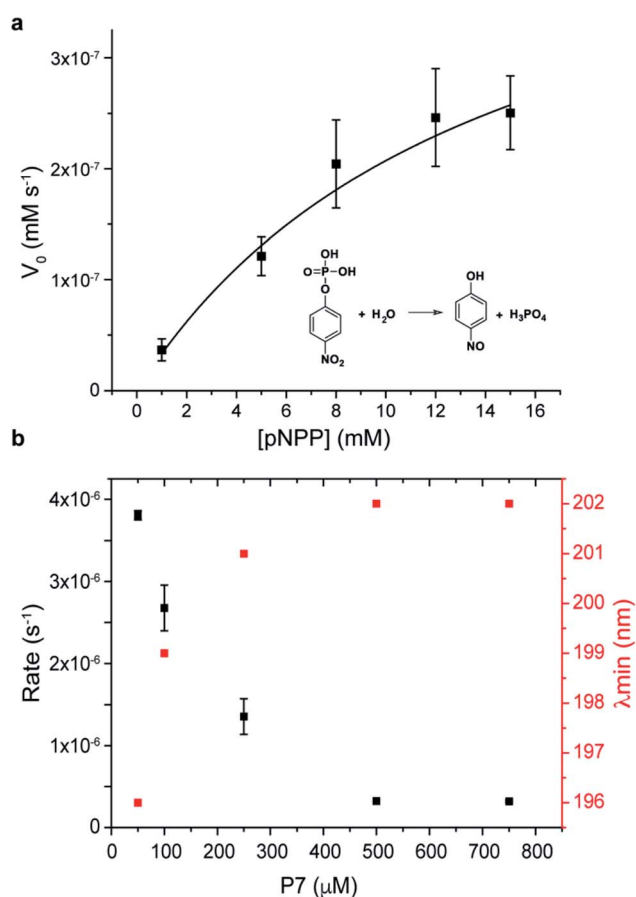


Fig. 4 Kinetics of P7 peptide (a) towards the model phosphatase substrate pNPP. (a) Dependence of the initial rate of the reaction on pNPP substrate concentration. (b) Phosphatase activity of P7 was influenced by conformation as a function of P7 concentration. Error bars represent standard deviations of at least three independent measurements.



## Conclusions

A biopanning strategy based on supramolecular recognition led to the discovery of oligopeptides' sequences with chemical diversity and with sequence-dependent patterns. In particular, there is a clear preference of charged residues (R/K) located at sequence terminals, the presence of proline residues in the middle of the sequence, and aromatic residues (Y/F), preferentially positioned in the first half of the peptide. Although the chemical diversity was key to generate binders with different properties, it is envisaged that in future studies, a consensus sequence can be obtained by carrying out more rounds of biopanning.

The detailed binding mechanism of the lead oligopeptide-based sequences revealed micromolar range  $K_{Ds}$  (74–520  $\mu\text{M}$ ). This indicates reversible binding, and an ability to differentiate tyrosine phosphorylation between monomeric and aggregation states. The P3, P7 and P29 oligopeptides revealed selectivity for supramolecular over non-supramolecular states of phosphate molecules, enabling the creation of micromolar-affinity supramolecular complexes. The lead peptide P7 presented an optimal combination of chemical composition with a defined hairpin-like structure that allowed for an adaptive behavior towards supramolecular structures, *via* destabilization of the tyrosine-phosphorylated micellar aggregates. In future, we envision that the panning strategy developed in this work can also be applied to find binders to other dynamic targets such as the phase-separated biomolecular condensates.

Although there is no counterpart of the phosphorylated self-assembled Fmoc-FpY in biology, high-density tyrosine phosphorylation is often seen in different diseases (*e.g.*, metabolic, neurodegenerative diseases). Therefore, the development of peptide modalities and functional materials that selectively target and perturb high-density phosphorylation can be incorporated in future diagnostic tools and therapeutics.

In addition, we also foresee that the chemical information dictating the ability of dodecapeptides at catalyzing phosphate ester hydrolysis phosphate expands the peptide field in aqueous media.

## Data availability

The data supporting the findings of this study are included in the main text and in the ESI file.†

## Author contributions

A. S. P., K. L. D, L. M., A. C. A. R., and R. V. U. conceived and designed the experiments. A. S. P., K. D. and S. C. carried out the phage display experiments. A. S. P, K. D., S. C. and B. P. M carried out the kinetic measurements and experimental conformational analysis. A. S. P. and L. M. conducted the binding NMR studies and DOSY NMR spectroscopy experiments. A. S. P, B. M. M. F., D. K. and V. N. performed the AFM imaging. I. F. and C. P. contributed for the self-assembly analysis. A. S. P., C. P. and T. W. performed the TEM experiments and contributed for sample preparation and data analysis. H. F.

C., A. J. C and I. P. M. performed the computational studies. A. S. P, L. M., A. C. A. R. and R. V. U. co-wrote the paper.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

The authors thank Dr Ana Sofia Ferreira for the help with DOSY NMR spectroscopy and Professor Carlos Salgueiro for helpful discussions. The authors thank the financial support of Advanced Science Research Center – City University of New York and Applied Molecular Biosciences Unit – UCIBIO which is financed by national funds from FCT (UIDP/04378/2020 and UIDB/04378/2020). The authors thank Fundação para a Ciência e Tecnologia, Portugal, for the research fellowships SFRH//BD/90644/2012 (H. F. C.), SFRH/BPD/97585/2013 (A. S. P.), SFRH/BPD/114848/2016 (L. M.) and SFRH/BPD/112543/2015 (A. J. M. B.) and the grant PTDC/BII-BIO/28878/2017 (A. C. A. R.). The NMR spectrometer is part of the National NMR Network (PTNMR) and are supported by Infrastructure Project No. 022161 (co-financed by FEDER through COMPETE 2020, POCI, and PORL and FCT through PIDDAC). R. V. U. acknowledges the Air Force Office of Scientific Research (FA9550-19-1-0111) for financial support. A. S. P. acknowledges funding from the Research Fund RF19-4862 of Royal Society of Chemistry.

## References

- 1 Y. Bai, A. Chotera, O. Taran, C. Liang, G. Ashkenasy and D. G. Lynn, *Chem. Soc. Rev.*, 2018, **47**, 5444–5456.
- 2 E. Mattia and S. Otto, *Nat. Nanotechnol.*, 2015, **10**, 111–119.
- 3 J. M. Slocik and R. R. Naik, *Curr. Opin. Biotechnol.*, 2017, **46**, 7–13.
- 4 S. Zhang, *Nat. Biotechnol.*, 2003, **21**, 1171–1178.
- 5 O. Zozulia, M. A. Dolan and I. V. Korendovych, *Chem. Soc. Rev.*, 2018, **47**, 3621–3639.
- 6 F. Sheehan, D. Sementa, A. Jain, M. Kumar, M. Tayarani-Najjaran, D. Kroiss and R. V. Uljin, *Chem. Rev.*, 2021, **121**(22), 13869–13914.
- 7 P. S. Huang, S. E. Boyken and D. Baker, *Nature*, 2016, **537**, 320–327.
- 8 I. V. Korendovych and W. F. DeGrado, *Q. Rev. Biophys.*, 2020, **53**, e3.
- 9 D. N. Woolfson, G. J. Bartlett, A. J. Burton, J. W. Heal, A. Niitsu, A. R. Thomson and C. W. Wood, *Curr. Opin. Struct. Biol.*, 2015, **33**, 16–26.
- 10 K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski and R. J. Knapp, *Nature*, 1991, **354**, 82–84.
- 11 H. B. Albada and R. M. J. Liskamp, *J. Comb. Chem.*, 2008, **10**, 814–824.
- 12 C. G. Pappas, R. Shafi, I. R. Sasselli, H. Siccardi, T. Wang, V. Narang, R. Abzalimov, N. Wijerathne and R. V. Uljin, *Nat. Nanotechnol.*, 2016, **11**, 960–967.
- 13 K. A. Günay and H.-A. Klok, *Bioconjugate Chem.*, 2015, **26**, 2002–2015.



- 14 D. Kroiss, J. M. Aramini, S. A. McPhee, T. Tuttle and R. V. Ulijn, *ChemSystemsChem*, 2019, **1**, e1900013.
- 15 Y. Maeda, N. Javid, K. Duncan, L. Birchall, K. F. Gibson, D. Cannon, Y. Kanetsuki, C. Knapp, T. Tuttle, R. V. Ulijn and H. Matsui, *J. Am. Chem. Soc.*, 2014, **136**, 15893–15896.
- 16 F. Tanaka and C. F. Barbas III, *Chem. Comm.*, 2001, 769–770, DOI: 10.1039/B100479B.
- 17 P. W. J. M. Frederix, G. G. Scott, Y. M. Abul-Haija, D. Kalafatovic, C. G. Pappas, N. Javid, N. T. Hunt, R. V. Ulijn and T. Tuttle, *Nat. Chem.*, 2014, **7**, 30.
- 18 X. Xiao, Z. Kuang, J. M. Slocik, S. Tadepalli, M. Brothers, S. Kim, P. A. Mirau, C. Butkus, B. L. Farmer, S. Singamaneni, C. K. Hall and R. R. Naik, *ACS Sensors*, 2018, **3**, 1024–1031.
- 19 L. Cronin, S. H. M. Mehr and J. M. Granda, *Chem*, 2018, **4**, 1759–1761.
- 20 L. Tallorin, J. Wang, W. E. Kim, S. Sahu, N. M. Kosa, P. Yang, M. Thompson, M. K. Gilson, P. I. Frazier, M. D. Burkart and N. C. Gianneschi, *Nat. Comm.*, 2018, **9**, 5253.
- 21 A. van Teijlingen and T. Tuttle, *J. Chem. Theory Comput.*, 2021, **17**, 3221–3232.
- 22 T. Clackson, H. R. Hoogenboom, A. D. Griffiths and G. Winter, *Nature*, 1991, **352**, 624–628.
- 23 C. S. M. Fernandes, I. Barbosa, R. Castro, A. S. Pina, A. S. Coroadinha, A. Barbas and A. C. A. Roque, *Biotechnol. J.*, 2016, **11**, 1513–1524.
- 24 J. McCafferty, A. D. Griffiths, G. Winter and D. J. Chiswell, *Nature*, 1990, **348**, 552–554.
- 25 I. M. Martins, R. L. Reis and H. S. Azevedo, *ACS Chem. Biol.*, 2016, **11**, 2962–2980.
- 26 S. Sonzini, A. Marcozzi, R. J. Gubeli, C. F. vanderWalle, P. Ravn, A. Herrmann and O. A. Scherman, *Angew. Chem., Int. Ed.*, 2016, **55**, 14000–14004.
- 27 C. A. Stevens, F. Bachtiger, X.-D. Kong, L. A. Abriata, G. C. Sosso, M. I. Gibson and H.-A. Klok, *Nat. Comm.*, 2021, **12**, 2675.
- 28 S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara and A. M. Belcher, *Nature*, 2000, **405**, 665–668.
- 29 D. B. Pacardo, M. Sethi, S. E. Jones, R. R. Naik and M. R. Knecht, *ACS Nano*, 2009, **3**, 1288–1296.
- 30 J. P. Casey, R. J. Barbero, N. Heldman and A. M. Belcher, *J. Am. Chem. Soc.*, 2014, **136**, 16508–16514.
- 31 L. Giger, S. Caner, R. Obexer, P. Kast, D. Baker, N. Ban and D. Hilvert, *Nat. Chem. Biol.*, 2013, **9**, 494.
- 32 J. Wagner, R. A. Lerner and C. F. Barbas, *Science*, 1995, **270**, 1797.
- 33 B. P. Orner, L. Liu, R. M. Murphy and L. L. Kiessling, *J. Am. Chem. Soc.*, 2006, **128**, 11882–11889.
- 34 T. Sawada, T. Takahashi and H. Mihara, *J. Am. Chem. Soc.*, 2009, **131**, 14434–14441.
- 35 M.-M. Zhou, J. E. Harlan, W. S. Wade, S. Crosby, K. S. Ravichandran, S. J. Burakoff and S. W. Fesik, *J. Biol. Chem.*, 1995, **270**, 31119–31123.
- 36 J. A. Jadwin, T. G. Curran, A. T. Lafontaine, F. M. White and B. J. Mayer, *J. Biol. Chem.*, 2018, **293**, 623–637.
- 37 A. H. Ross, D. Baltimore and H. N. Eisen, *Nature*, 1981, **294**, 654–656.
- 38 A. C. Nairn, J. A. Detre, J. E. Casnellie and P. Greengard, *Nature*, 1982, **299**, 734–736.
- 39 J. W. Sadownik, J. Leckie and R. V. Ulijn, *Chem. Comm.*, 2011, **47**, 728–730.
- 40 D. J. Rodi, A. S. Soares and L. Makowski, *J. Mol. Biol.*, 2002, **322**, 1039–1052.
- 41 K. Akagawa and K. Kudo, *Acc. Chem. Res.*, 2017, **50**, 2429–2439.
- 42 A. J. Metrano, N. C. Abascal, B. Q. Mercado, E. K. Paulson, A. E. Hurlley and S. J. Miller, *J. Am. Chem. Soc.*, 2017, **139**, 492–516.
- 43 A. E. Hargrove, S. Nieto, T. Zhang, J. L. Sessler and E. V. Anslyn, *Chem. Rev.*, 2011, **111**, 6603–6782.
- 44 R. Mahalakshmi, G. Shanmugam, P. L. Polavarapu and P. Balaram, *ChemBioChem*, 2005, **6**, 2152–2158.
- 45 T. Machida, S. Dutt and N. Winssinger, *Angew. Chem., Int. Ed.*, 2016, **55**, 8595–8598.
- 46 J. Razkin, H. Nilsson and L. Baltzer, *J. Am. Chem. Soc.*, 2007, **129**, 14752–14758.
- 47 G. Gulseren, I. C. Yasa, O. Ustahuseyin, E. D. Tekin, A. B. Tekinay and M. O. Guler, *Biomacromolecules*, 2015, **16**, 2198–2208.
- 48 A. K. Pandey, K. M. Thomas, C. R. Forbes and N. J. Zondlo, *Biochemistry*, 2014, **53**, 5307–5314.
- 49 N. Zanna, L. Milli, B. Del Secco and C. Tomasini, *Org. Lett.*, 2016, **18**, 1662–1665.
- 50 M. L. Romero Romero, F. Yang, Y.-R. Lin, A. Toth-Petroczy, I. N. Berezovsky, A. Goncarencu, W. Yang, A. Wellner, F. Kumar-Deshmukh, M. Sharon, D. Baker, G. Varani and D. S. Tawfik, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E11943.

