Molecular tweezers target a protein–protein interface and thereby modulate complex formation†

F. Trusch,‡a K. Kowski,a K. Bravo-Rodriguez,c C. Beuck,a A. Sowislok,b B. Wettig,b A. Matena,a E. Sanchez-Garcia,c H. Meyer,d T. Schraderb and P. Bayer*a

Molecular tweezers for lysine and arginine select a few residues on a protein surface and by their unique complexation mode disrupt a critical protein–protein interaction. Detailed structural information was gained by NMR experiments, strongly supported by QM/MM calculations and further substantiated by ITC, fluorescence anisotropy, ELISA and bio-layer-interference studies.

Modulation of protein–protein interactions (PPI) with small molecules is one of the most promising new approaches in chemical biology and drug discovery.¹ To date, most discoveries result from structure-based design, natural product-inspired PPI modulation or compound library generation.² However, supramolecular ligands are less frequently used for PPI modulation, since the design of compounds which target a predictable well-defined area on a protein is still challenging.³ A promising approach is the construction of compounds which bind single exposed amino acids. In this respect, cyclodextrins, cucurbiturils, resorcinarenes and calixarenes have been used to successfully block critical residues on protein surfaces.⁴ Recently, we introduced water-soluble supramolecular tweezers highly selective for lysines and arginines (Scheme 1). By a unique threading mechanism, the positively charged side chains of these amino acids are drawn into the tweezers’ tailored cavity.⁵,⁶ Applications of the tweezers have been so far limited to enzyme inhibition and the disruption of protein–peptide complexes.⁵,⁷,⁸ Here, we report on the inhibition of the protein–protein interaction between the domains of p97 and its cofactor UBXD1 which we use as a model case (for interaction details see SI2, ESI†).

Specifically, we reveal the affinity profile of the tweezers to preferred lysine and arginine residues on the p97-N surface with NMR spectroscopy in combination with Molecular Dynamics (MD) and Quantum Mechanics/Molecular Mechanics (QM/MM) calculations. Subsequent biophysical assays show how binding of the tweezers to p97-N disrupts the interaction with UBXD1-N.

p97 (also called VCP) is an abundant AAA-ATPase involved in cell cycle regulation, DNA repair and protein degradation.⁹ Accordingly, missense mutations in p97 cause degenerative diseases.¹⁰ Its N-terminal domain p97-N1-199 with 32 basic residues represents an attractive target for the molecular tweezers (Fig. S1, ESI†). Hence, we performed 2D NMR experiments to elucidate whether binding to the mainly positively charged protein surface is unspecific or if the tweezers occupy preferred well-defined interaction sites. Sub-stoichiometric addition of tweezers to ¹⁵N-labeled p97-N up to a molar ratio of 0.5 : 1 produced shifting as well as disappearing N¹H-signals in a series of ¹H–¹⁵N-HSQC spectra (Fig. S2, ESI†). Especially decreasing signal intensities indicate specific interactions,⁸ while chemical shift changes may also occur as averaged values when ligands explore the protein surface mainly guided by electrostatic interactions in a process termed “protein surface camouflage”.¹¹ When mapping the
Due to the extensive aromatic system of the tweezers, several N_H-signals of residues around a bound lysine or arginine also show changes. Analysis and subsequent mapping of resonances are in some cases ambiguous since some patches contain more than one positively charged amino acid. With NMR data we could assign host residues for tweezers binding as K20, K60/K63, R113/K148, R86/R144/R155, R159/K164 and K190. Chemical shifts and intensity changes of N_H-signals of all involved amino acids within the preferred binding regions produced titration curves suitable for nonlinear regression with dissociation constants in the lower to medium µM-range (Table 1 and Fig. 1a).

To retrieve thermodynamic parameters, binding of tweezers to p97-N was also monitored by isothermal titration calorimetry (ITC) (Fig. 2).

The binding process is enthalpically driven as indicated by a strong negative $\Delta H = -64.1$ kcal mol$^{-1}$ that overcompensates the unfavorable entropic changes of $- T \Delta S = 56.9$ kcal mol$^{-1}$. This phenomenon is typical for the specific inclusion inside the tweezers’ cavity at the cost of reduced mobility of the amino acid side chain. The analysis of calorimetric data regarding the molar ratio between tweezers and p97-N results in a stoichiometry of 6 tweezers per p97-N molecule, which is in remarkable agreement with the above discussed $^1$H–$^{15}$N-HSQC-titration experiment (Fig. 1a). From the averaged Gibbs free energy of $-7.2$ kcal mol$^{-1}$ for each single binding event, the average $K_D$ value was calculated to 6 µM, which is slightly lower than the estimated affinities from NMR titration experiments (Table 1).

To rule out the remaining ambiguities between two potential host residues in the same patch and to reappraise the tweezers binding sites on p97-N, MD simulations were performed.
followed by QM/MM calculations. Of the 26 lysines and arginines included in the crystal structure (PDB: 3QQ7), only 13 formed conserved inclusion complexes during MD simulations. In all these cases, the host residue side chain was well threaded into the cavity of the tweezers (Fig. S3 and Table S1, ESI†). For these residues, QM/MM calculations were performed (for details see SI1, ESI†) whereby the electronic energies of the QM regions reflect the relative stability of the tweezers–amino acid complexes depending of the molecular environment. The subsequent analysis allowed assigning of the preferred tweezers binding residue in each patch (Table S2, ESI†). In patch 1, K20 is determined as the host residue by NMR experiments and has also one of the lowest QM energies (Fig. 1b and Table S2, ESI†). The QM/MM calculations also qualify K63 over K60 as the second most likely binding site in patch 2. Unlike K60, the side chain of K63 is solvent exposed and well accessible. In addition, positively charged neighbouring amino acids are able to interact with the negatively charged tweezers’ free hydrogen phosphate group (Fig. 1c). NMR data point towards K148 or R113 as a third putative tweezers host residue (patch 3). Although both complexes are not conserved during the MD simulations, the positive electrostatic surface potential of p97-N formed by several positively charged residues in close proximity. For patch 5, QM/MM calculations clearly favour R159 over K164 (Fig. 1d, ESI†) to form one of the most stable of all arginine–tweezers complexes (Table S2, ESI†). This arginine is well solvent-exposed without steric impediment and is supported by secondary electrostatic interactions between the tweezers’ hydrogen phosphate groups and positively charged residues in close proximity. For patch 5, QM/MM calculations clearly favour R159 over K164, because it forms a conserved inclusion complex with the tweezers, despite the lack of secondary interactions with surrounding residues. Taken together, we propose K20, K63, K148, R86, R159 and K190 as the preferred targets of the tweezers on p97-N (Table 1, patches 1–6) with binding to K148 taking place via non-inclusion electrostatic interactions. To rule out steric hindrance of the tweezers while simultaneous binding, we performed MD simulations with preformed inclusion complexes of the tweezers on K20, K63, R86 and R159 and a non-inclusion complex with K148 (K190 is not included in the crystal structure of p97-N used for the simulations) (Fig. 3 and Fig. S4, ESI†). All inclusion complexes were conserved and the tweezers established secondary interactions as described for the 1:1 complexes. These results confirm that there is sufficient space at the surface of p97-N to host five tweezers simultaneously and, most likely, a sixth one in the flexible C-terminus around K190.

Since residues of both the VIM groove and the ND1-linker of p97-N are targeted by the tweezers (Fig. 1a and S12, ESI†), but also mediate adaptor binding of UBXD1-N,12 we investigated the potential disrupting effect of the tweezers on the complex formation between p97-N and UBXD1-N by fluorescence anisotropy, ELISA and BLI (Fig. 4). First, complex formation of p97-N with labelled UBXD1-N was monitored by fluorescence anisotropy changes in the presence of up to 50 mol% tweezers. Protein and tweezers ratios were carefully optimized to avoid the formation of protein aggregates. The apparent K_D value ascends from 9 μM (no tweezers) to 216 μM already at low tweezers concentrations (Fig. 4a and S5, ESI†). Importantly, conformational changes on p97-N during tweezers binding as an alternative explanation for the weakened protein–protein interaction with UBXD1-N were excluded by CD spectroscopy (Fig. S7, ESI†). The inhibiting character of the tweezers on the UBXD1-N/p97-N complex formation was also confirmed by ELISA experiments. Immobilized UBXD1-N (30–210 nM) was incubated with combinations of p97-N and varying tweezers concentrations (Fig. 4b). Increasing tweezers concentrations substantially impeded complex formation between UBXD1-N and p97-N.

In bio-layer interference (BLI) experiments, titrations of immobilized p97-N with increasing concentrations of UBXD1-N (0–5 μM) were performed in the presence of different tweezers concentrations. The apparent K_D value obtained by steady-state data analysis increased from 2.1 μM to 21 μM (Fig. 4c and Fig. S6, ESI†). The effective concentration of the tweezers in ELISA and BLI experiments is likely overestimated because a fraction is also bound by the blocking agents. Nevertheless, the system reaches saturation at large tweezers excess, indicating specific binding to the protein–protein interface (Fig. 4b and c).
In total, the three different experimental approaches confirm that the formation of tweezers–lysine/arginine inclusion complexes can significantly impair the interaction between two protein domains.

We also tested the ability of our supramolecular ligand to target the binding partner UBXD1-N by titrating $^{15}$N-labeled UBXD1-N with increasing amounts of tweezers (Fig. S8, ESI†). Since this moiety of UBXD1 is disordered, most lysine residues are bound by the tweezers. Although the critical VIM motif in UBXD1 which binds to the VIM groove of p97-N (SI2, ESI†) does not contain any lysine or arginine residues, especially the adjacent K66 may also contribute to weaken the important p97-N/UBXD1-N interaction.

In summary, hydrogen phosphate tweezers are binding at sub-stoichiometric concentrations with low micromolar $K_D$ values preferably to basic lysines and arginines restricted to 6 patches on the p97-N protein surface. The tweezers binding is driven most likely by a combination of electrostatic interactions, hydrogen bonds, dispersive attraction and hydrophobic effects. Furthermore, QM/MM calculations confirmed preferred complexation sites but also allowed to discriminate between ambiguous host residues derived from NMR data. The inhibition of the protein–protein interaction between UBXD1-N and p97-N by the tweezers was confirmed by fluorescence anisotropy. ELISA and BLI experiments further support the weakening of the protein–protein interaction. In contrast to other supramolecular ligands like calixarenes and PTS (pyrenetetrasulfonic acid) the tweezers do not promote unspecific formation of protein assemblies. Thus, they may be amenable for specific inhibition of selected binding sites upon further chemical variation.

### Notes and references