



Development of a Fast Screening Method for Selecting Excipients in Formulations using MD simulations, NMR and MicroScale Thermophoresis

Indrakumar, Sowmya; Zalar, Matja; Tschammer, Nuska; Pohl, Christin Alina; Nørgaard, Allan; Streicher, Werner; Harris, Pernille; Golovanov, Alexander P; Peters, Günther H.J.

Published in:
European Journal of Pharmaceutics and Biopharmaceutics

Link to article, DOI:
[10.1016/j.ejpb.2020.10.015](https://doi.org/10.1016/j.ejpb.2020.10.015)

Publication date:
2021

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Indrakumar, S., Zalar, M., Tschammer, N., Pohl, C. A., Nørgaard, A., Streicher, W., Harris, P., Golovanov, A. P., & Peters, G. H. J. (2021). Development of a Fast Screening Method for Selecting Excipients in Formulations using MD simulations, NMR and MicroScale Thermophoresis. *European Journal of Pharmaceutics and Biopharmaceutics*, 158, 11-20. <https://doi.org/10.1016/j.ejpb.2020.10.015>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Journal Pre-proofs

Development of a Fast Screening Method for Selecting Excipients in Formulations using MD simulations, NMR and MicroScale Thermophoresis

Sowmya Indrakumar, Matja Zalar, Nuska Tschammer, Christin Pohl, Allan Nørgaard, Werner Streicher, Pernille Harris, Alexander P. Golovanov, Günther H. J. Peters

PII: S0939-6411(20)30315-5
DOI: <https://doi.org/10.1016/j.ejpb.2020.10.015>
Reference: EJPB 13439

To appear in: *European Journal of Pharmaceutics and Biopharmaceutics*

Received Date: 6 July 2020
Revised Date: 22 October 2020
Accepted Date: 24 October 2020

Please cite this article as: S. Indrakumar, M. Zalar, N. Tschammer, C. Pohl, A. Nørgaard, W. Streicher, P. Harris, A.P. Golovanov, G. H. J. Peters, Development of a Fast Screening Method for Selecting Excipients in Formulations using MD simulations, NMR and MicroScale Thermophoresis, *European Journal of Pharmaceutics and Biopharmaceutics* (2020), doi: <https://doi.org/10.1016/j.ejpb.2020.10.015>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier B.V.



Development of a Fast Screening Method for Selecting Excipients in Formulations using MD simulations, NMR and MicroScale Thermophoresis

Sowmya Indrakumar^{1,2*}, Matja Zalar^{3,4}, Nuska Tschammer^{5,6}, Christin Pohl^{1,7}, Allan Nørgaard⁷, Werner Streicher^{7,8}, Pernille Harris¹, Alexander P. Golovanov³, and Günther H. J. Peters^{1*}

¹Technical University of Denmark, Department of Chemistry, 2800 Kgs. Lyngby, Denmark;

²Current address: University of Copenhagen, Biomolecular Sciences. Ole Maaløes Vej 5. 2200 København N, Denmark; ³Manchester Institute of Biotechnology, and Department of Chemistry,

School of Natural Sciences, Faculty of Science and Engineering, The University of Manchester, Manchester M1 7DN, United Kingdom; ⁴Current address: CRUK Beatson Institute, Garscube

Estate, Switchback Road, Glasgow G61 1BD, United Kingdom; ⁵NanoTemper GmbH Technologies, Floessergasse 4 81369 Munich, Germany; ⁶Current address: CreLux GmbH - a WuXi

AppTec Company, Am Klopferspitz 19a, 82152 Planegg, Germany; ⁷Novozymes, Krogshøjvej 36, 2880 Bagsvaerd, Denmark; ⁸Current address: NanoTemper GmbH Technologies,

Floessergasse 4 81369 Munich, Germany.

ABSTRACT

Development of peptide therapeutics generally involves screening of excipients that inhibit peptide-peptide interactions, hence aggregation, and improve peptide stability. We used the therapeutic peptide plectasin to develop a fast screening method that combines microscale thermophoresis titration assays and molecular dynamics simulations to relatively rank the excipients with respect to binding affinity and to study key peptide-excipient interaction hotspots on a molecular level, respectively. Additionally, ^1H - ^{13}C -HSQC NMR titration experiments were performed to validate the fast screening approach. The NMR results are in qualitative agreement with results from the fast screening method demonstrating that this approach can be reliably applied to other peptides and proteins as a fast screening method to relatively rank excipients and predict possible excipient binding sites.

Keywords: molecular dynamics simulations; FTMap; NMR; microscale thermophoresis; chemical shift perturbations; hotspots; excipients

INTRODUCTION

Naturally occurring peptides are important biomolecules that have crucial roles in human physiology, including actions as hormones, growth factors, or host defense.[1] The peptides generally show high affinity, specificity, and efficacy for their targets, and therefore, there is an increasing interest in using peptides as therapeutics.[2] One example of such peptides is plectasin, initially isolated from the saprophytic ascomycete *Pseudoplectania nigrella* fungus found on the ground of northern European pine forests.[3] It is a small, 40 amino acid, cystine-stabilized peptide with antibiotic properties.[4] Plectasin has a tertiary structure that closely resembles those of defensins found in spiders, scorpions, and mussels, and the peptide belongs to the class of amphipathic antimicrobial peptides (AMPs) that are widely distributed as intrinsic host defense molecules produced by all multicellular organisms.[5,6] The studies have shown effective antimicrobial and antiviral effects against a broad range of bacteria, viruses, and fungi. Plectasin and its variants are especially active against Gram-positive bacteria including *S. aureus* and *S. pneumoniae*. [7] For instance, plectasin cured pneumococcal peritonitis and pneumonia with low toxicity in mice signifying the therapeutic potential of plectasin as an antimicrobial peptide.[3,8–10] While their smaller sizes make them easier to deliver across biological barriers compared to larger proteins, their formulation can be problematic since these peptides are prone to chemical and colloidal instabilities[11,12] that can lead to aggregation.[11,13,14]

The colloidal stability of peptides depends on the nature of the peptides, and their stability can be modulated by adjusting solution conditions such as pH, buffer composition, ionic strength, and by the addition of excipients.[11]–[15] The addition of appropriate excipients increases peptide stability, since the excipients interact with aggregation prone regions on the peptide surface and thereby reducing peptide-peptide interactions and subsequently aggregation.[16] Identifying appropriate excipients can involve extensive experimental work that is often time-consuming since it involves screening of excipients at different physicochemical conditions. Alternatively, *in-silico* methods can be utilized to screen different conditions to narrow the number of potential excipients used to increase colloidal peptide stability. Previously, review work by Pandya, Dalby et. al., 2018, has well documented the various biophysical methods especially NMR and modelling approaches to address protein-excipient interactions.[17] The study showed the potential use of molecular docking and simulations methods to develop a high-throughput screening method to facilitate

initial formulation. Other recent studies have shown the potential use of docking and simulation techniques in combination with experimental approaches to identify protein-excipient binding sites and rational design of excipient to aid formulation of biologics.[18–20] We have previously applied such *in-silico* screening methods in combination with small angle X-ray scattering to map protein-excipient interaction sites and study their effect on dynamics and oligomerization of human transferrin and albumin-neprilysin fusion protein.[21,22]

In this study, we introduce an alternative docking approach which is relatively fast in comparison with the already existing approaches and performs blind docking and clustering to identify predominant binding sites. Taking this further in combination with MD simulations and MST, one can predict fairly well the protein-excipient binding sites. Here, *in-silico* studies were validated with NMR studies. Furthermore, the excipient concentration for *in-silico* studies were similar to NMR experiments, that is certain number of excipient molecules corresponding to a specific concentration were randomly positioned in the solvated protein box. Such an approach can be extended further to study interactions of proteins with combination of different excipients.

In this study, the focus is on developing a fast screening methodology that utilizes microscale thermophoresis (MST) measurements and computational methods to predict peptide-excipient interaction hotspots. The initial ranking of excipients based on overall binding affinities (K_d) was obtained using MST experiments. Peptide-excipient interaction hotspots were predicted using FTMap[23] with the listed excipients in scope in combination with molecular dynamics (MD) simulations. If the predicted sites are also plausible sites for peptide-peptide interactions, excipients that bind to these sites are potentially good candidates for inclusion in the formulation.[16,24,25] Aggregation prone regions (APRs) were identified using freely available webservers for predicting APRs.[26,27] Additionally, NMR titration experiments were performed to validate the fast screening methodology. This workflow shows good agreement between the *in-silico* and experimental methods and can be applied effectively to other peptide and protein systems.

MATERIALS AND METHODS

In-silico Approaches

Structure preparation and aggregation prone regions. A library of 12 different excipients was included in the study to verify the computational screening methods. The excipients are amino acids (arginine, glycine, histidine, methionine, and proline), sugars (sucrose and trehalose), sugar glycols (mannitol, sorbitol, and glycerol), and buffering agents (Tris and succinate). The excipient structures were taken from the ZINC database (www.zinc.docking.org/)[28], and the LigPrep program in the Schrödinger suite 2016-3 (Schrödinger, LLC, New York, NY, USA)[29] was used to adjust the protonation state at pH 4 (corresponding to the pH used in the experiments). The X-ray crystal structure of plectasin was retrieved from the protein data bank, PDB ID:3E7U[30], and the PrepWiz tool in the Schrödinger suite 2016-3 was used to adjust the protonation state of the residues at pH 4. The structure was subsequently used as a starting structure for MD simulations.

Site-specific binding studies. Excipients structures were obtained from Zinc Database[28]. These excipients were prepared at pH 4 using the Ligprep tool in Schrödinger release 2016-3 (Schrödinger, LLC, New York, NY, USA)[29]. Histidine arginine and Tris at pH 4 carry an overall +1 positive charge. The prepared excipients were docked to each of the two sites using Glide program that is part of the Schrödinger suite 2016-3. The docking experiments were performed using default parameters for grid generation and docking runs, where the latter was performed with extra precision (XP).[31] From each site, the pose with the lowest energy was chosen further for MD simulations. All-atom simulations were performed for 500 ns each with the same settings as described in the FTMap section of the main manuscript, the excipient concentration was adjusted to 150 mM that corresponds to the same concentration used in NMR titration experiment. In total, 31 molecules were added to the solvated system containing approximately 11250 water molecules. The parameter file for the excipients was prepared using the antechamber[32] module in Amber 16 at pH 4 and applying the AM1-BCC[33] charge method. The number of excipients accumulating at every residue was calculated by counting the number of excipients within 4Å of the residue divided by the sum of the number of contacts for all excipients, also defined as contacts.

Molecular dynamics simulations. Classical all-atom constant pH MD simulations[34] in explicit solvent were carried out for 100 ns at pH 4 using the software Amber 16 with the amber force field

ff99SB[35] for proteins to generate an ensemble of possible peptide structure in solution. The titratable residues (Asp, Glu, His) were allowed to change their protonation state during the simulations. The peptide was solvated in a truncated octahedron water box having a 15 Å cutoff in all directions, and water molecules were represented using the TIP3P[36] water model. The system was neutralized with four chloride ions. The particle mesh Ewald method[37] was employed with a real-space cutoff of 8 Å to determine the non-bonded electrostatics energies. The system was minimized for 5000 steps using 1500 steps of steepest descent method followed by 3500 steps of conjugate gradient method. Successively, the system was heated linearly from 10 K to 300 K using the Langevin thermostat[38] with a collision frequency of 5 ps⁻¹ for 1ns. The systems were then equilibrated for 4 ns with coordinates saved every 5 ps at constant temperature (300 K) and pressure (1bar). Finally, production simulations were run for 100 ns, and coordinates were saved every 10 ps.

Experimental Methods

Relative Ranking of excipients using Microscale Thermophoresis. The experiments were performed using a Monolith NT.115 LabelFree® instrument (NanoTemper Technologies, Munich, Germany). Label-free system makes use of the intrinsic protein fluorescence caused by the aromatic amino acids phenylalanine, tyrosine, and tryptophan (Trp), with the latter being the dominant intrinsic fluorophore. Plectasin samples were provided by Novozymes A/S. Plectasin was extensively dialyzed into 10 mM acetate buffer pH 4 at 4°C using slide-a-lyzer 2000 MWCO dialysis cassettes (Thermo Scientific) for 20 hours with three times buffer exchange after 3, 6 and 12 hours. The final concentration of plectasin was 17mg/mL as determined by measuring UV absorbance at 280 nm ($A_{280}=38.08$, $\epsilon = 2.24 \text{ mL}\cdot\text{mg}^{-1}\cdot\text{cm}^{-1}$) with a Nanodrop2000 (Thermo Scientific). The excipient stocks were prepared in 10 mM acetate buffer pH 4 (Table S1). In the case of pH shift, pH was adjusted using either NaOH or HCl. Table S1 shows the stock concentrations $[C]_{\text{stock}}$ used for the measurements. The excipient was diluted in 16 serial steps in PCR tubes. Subsequently, to each of the PCR tubes, plectasin was added to a final concentration of 10 μM. The prepared samples were incubated for 1 hour before loading into MST premium coated capillaries, where the coating minimizes peptide binding to the capillary wall. Experiments were performed in triplicates at the same MST-power to determine uncertainties in the measured

values. MST power was set to 40 % for all excipients except succinate and mannitol where the power was set to 20%.

NMR titration experiment to identify peptide-excipient binding regions. The plectasin sample was prepared as mentioned previously described in MST sample preparation. The total peptide sample volume was 500 μL , which included 2 mM peptide, 1 μL trimethylsilylpropanoic acid (TSP) (used as a reference for peak identification), 25 μL D_2O and volume adjusted to 500 μL with buffer. Excipients were added stepwise to the peptide sample from the 1M stock solution, which minimized dilution effects. All NMR experiments were acquired at 25 $^\circ\text{C}$ on Bruker 800MHz Avance III spectrometer equipped with a temperature control unit and a 5 mm TCI cryoprobe. The plectasin assignments are based on those published previously.[39] ArgHCl was added to a final concentration of 1, 5, 7.5, 10, 15, 20, 30, 40, 50 and 75 mM to the prepared 500 μL plectasin sample. All other excipients were added to a final concentration of 25, 50, 75, 100 and 150 mM. At each titration point, 1D ^1H and 2D natural abundance ^1H - ^{13}C -heteronuclear single quantum correlation (HSQC) experiment with sensitivity enhancement, gradient coherence selection, and multiplicity editing was acquired. For all 1D ^1H experiments, 32k complex points were collected with spectral windows of 16.0 ppm, while 2D data matrix of 2k x 512 complex points with spectral windows 16.0 ppm for ^1H and 165 ppm for ^{13}C dimension were acquired. The total acquisition time for 2D-HSQC experiments was 150 min. All HSQC spectra were referenced against the TSP peak. The weighted chemical shift perturbations (CSP), $\Delta\delta_{CH}$, occurring due to binding of excipients to plectasin were calculated as follows:

$$\Delta\delta_{CH} = \sqrt{\frac{1}{2}[\delta_H^2 + (0.43^2 \cdot \delta_C^2)]} \quad (1)$$

where $\Delta\delta_H$ and $\Delta\delta_C$ are chemical shift changes in proton and carbon dimensions, respectively. The weight attributed to the carbon chemical shift was calculated as a ratio between the average peak width in the proton and carbon dimension. To obtain the binding affinity K_d values, we have assumed 1:1 binding mode on the individual residue basis. As this assumption might not be necessarily true when considering larger surface area (i.e. patch) of the peptide, we did not attempt the global fits even for local patches. The binding affinity (K_d) and max CSP ($\Delta\delta_{max}$) can be obtained as follows:

$$\Delta\delta_{CH} = \frac{\Delta\delta_{max}\{([P]_t + [L]_t + K_d) - [([P]_t + [L]_t + K_d)^2 - 4[P]_t[L]_t]^{0.5}\}}{2[P]_t} \quad (2)$$

where $[P]_t$ and $[L]_t$ are the total protein (2 mM) and ligand concentrations, respectively. The threshold was set to 0.025 ppm corresponding to the standard deviation when taking into account all CSPs. CSPs above the threshold reflect significant signal perturbations. The same cutoff of 0.025 ppm cutoff was used for all the excipients to account for the differences in their binding affinities. It is essential to choose an appropriate threshold value to increase specificity (to pick residues genuinely interacting with excipients).[40] Residue specific K_d values were calculated considering peaks with significant perturbations. To get site specific K_d values at site I and II (sites predicted by FTMap and SiteMap), average K_d value was calculated by taking an average of residue-specific K_d values (<500 mM) of residues belonging to a specific site. Moreover, significant curvature of $\Delta\delta_{CH}$ vs excipient concentrations is required to calculate accurate binding affinities. The excipients used in this study show weak interactions (10-500 mM range).

RESULTS

In this section, we will present the findings from experimental as well as computational studies. The rationale is to develop a fast screening approach by combining complementary methods to enhance our understanding on the protein-excipient interactions.

Relative ranking of excipients using MST. A library of 12 different excipients was included in the study (Table 1). The excipients can be broadly defined in four classes as sugars (sucrose and trehalose), sugar glycols (mannitol, sorbitol, and glycerol), amino acids (arginine, glycine, histidine, methionine, and proline), and buffering agents (Tris and succinate) that are frequently used in stabilization of therapeutics.[24,41,42] MST label-free measurements were done in 10 mM acetate buffer at pH 4 as high colloidal stability was previously reported for plectasin in acidic pH range.[12] Data were analyzed and K_d values were obtained using the MO software version 2.3 from NanoTemper Technologies. At least 10 out of 16 titration points from the titration series were taken for the analysis to improve fitting (based on χ^2). For most of the excipients, the maximum obtainable stock concentration was 5 M (Table S1 in Supporting Information (SI)). The viscosity of all excipient solutions was measured using the automated micro viscosizer (Anton Paar). All measured viscosities were within the limit of the MST instrument, negating any effect of viscosity on the thermophoresis of the molecule. The MST trace of each capillary was analyzed for the difference between the fluorescence during thermophoresis (F_{hot}) and baseline fluorescence (F_{cold}). The normalized fluorescence (F_{norm}) is defined as $F_{\text{hot}}/F_{\text{cold}}$. Binding curves can be displayed using ΔF_{norm} values (baseline-corrected normalized fluorescence) or fraction bound (ΔF_{norm} values are divided by the amplitude of the curve) against excipient concentration (Figure 1). Table 1 shows the resulting apparent K_d values from the MST measurements, which reflect overall binding of excipients to protein, possibly at multiple sites. For some excipients, saturation is starting to appear that enables accurate determination of K_d (corresponding to 50% saturation) (Figure 1). A numerical comparison of the apparent K_d values allows for the ranking of excipients based on the binding affinities (Figure 1). Saturation was reached for arginine ($K_d = 35$ mM), histidine (46 mM), Tris (56 mM), and trehalose (65 mM). In the case of other excipients, saturation was not obtained, which implies only an estimate of K_d value can be extracted from the data (Figure 1).

Table 1. Experimentally determined K_d of excipients to plectasin using MST label-free method in 10 mM acetate, pH 4. Here, % $[PE]/100$ corresponds to the fraction of peptide (P)-excipient (E) complex; $[P] = 17\text{mg/mL}$ and $[E]_{\text{stock}} = 150\text{ mM}$. Saturation is not achieved for the excipients having % $[PE]/100$ lower than 50%. Therefore, the calculated K_d represents the minimum K_d value.

Excipient	K_d (mM)	% $[PE]/100$	Excipient	K_d (mM)	% $[PE]/100$
Arginine	35±12	0.81	Glycine	217±204	0.41
Histidine	46±6	0.77	Sucrose	275±83	0.35
Tris	56±14	0.73	Proline	437±44	0.26
Trehalose	65±48	0.7	Mannitol	600±870	0.2
Succinate	137±77	0.52	Sorbitol	900±600	0.14
Methionine	180±140	0.45	Glycerol	4200±7000	0.03

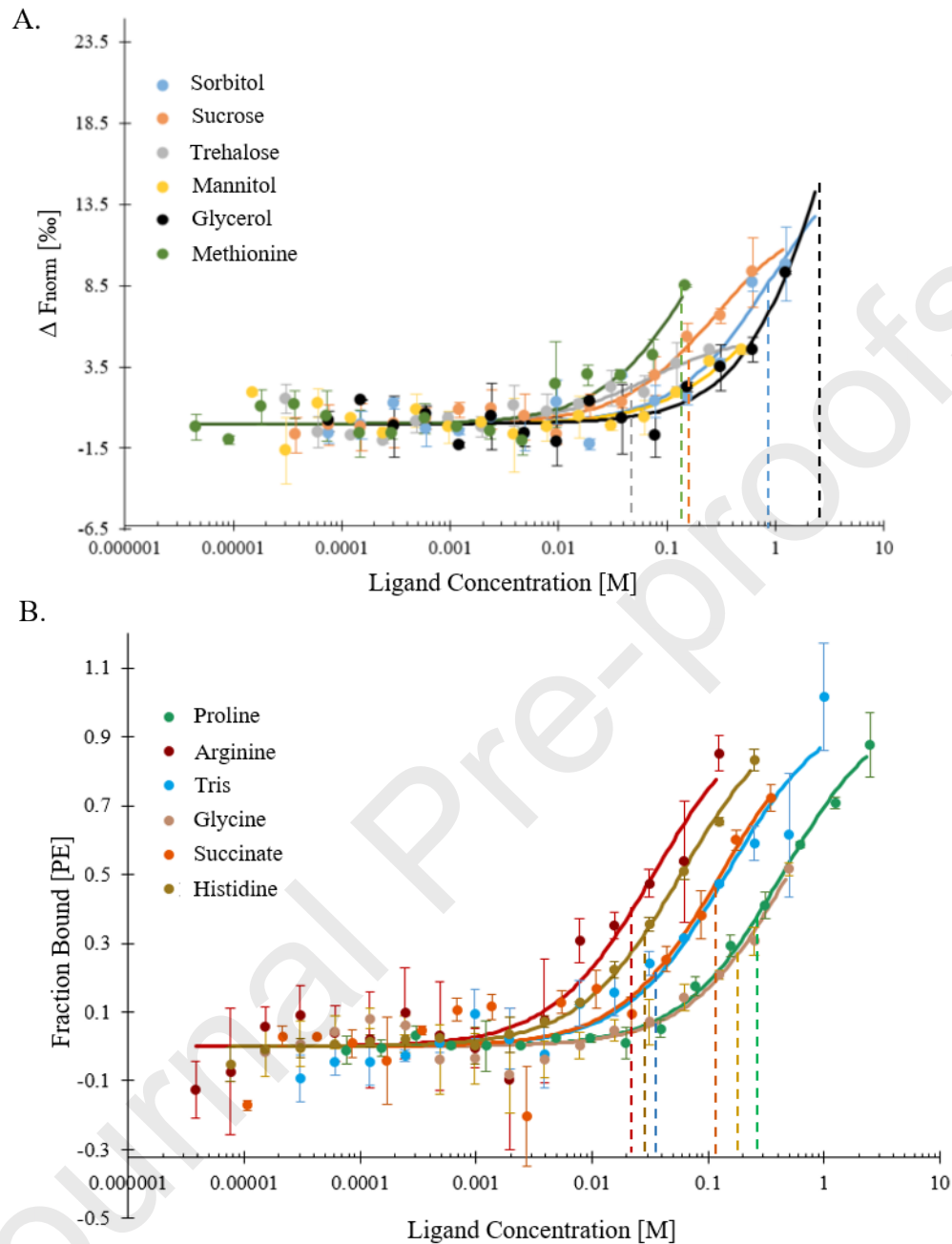


Figure 1. A. Concentration-response curves for the binding interactions between different excipients and unlabeled plectasin displayed using ΔF_{norm} against ligand (excipient) concentrations. B. Concentration-response curves for the binding interactions between different excipients and unlabeled plectasin displayed using fraction bound against ligand concentrations.

In-silico Approaches

Aggregation prone regions and peptide-excipient interaction hotspot. Aggregation prone regions were predicted using AGGRESKAN3D[26], CAMSOL[27], CSSP[43], that all are based on structure-based methods and hence consider the local environment around a given residue to estimate a solubility score. Different servers were used to see if similar regions were picked as APRs. As shown in Figure 2A, APRs are concentrated in the termini; especially, a large segment of a beta-strand is predicted to be prone to aggregation.

Molecular dynamics simulations. Classical all-atom constant pH MD simulations[34] in explicit solvent were carried out for 100 ns at pH 4. Conformations taken from the MD simulation trajectories were grouped into distinct clusters using the *average-linkage* method, which is part of Cpptraj[44] in Amber 16. Cluster representatives were analyzed using FTMap (<http://ftmap.bu.edu/param>)[23], which identifies binding hotspots of macromolecules. Using FTMap, small organic molecules (present as default) and molecules of choice are docked blindly onto the peptide surface, clustered, and ranked based on average energy.[45] Results from FTMap on different cluster representatives were analyzed further using a simple statistical method to rank sites that are predominant interaction sites for excipients in all cluster representatives. To do this, every residue, j , is assigned an interaction score, I_{score} , defined as:

$$I_{score} = \sum_{i=1}^{10} c_i (I(j)_{norm} \geq 1) \quad (3)$$

where

$$I(j)_{norm} = \frac{I(j)^{non-bonded}}{\langle I(j)^{non-bonded} \rangle} \quad (4)$$

$I(j)^{non-bonded}$ is the non-bonded interaction value at residue (j) obtained from FTMap. $I(j)_{norm}$ is the normalized non-bonded interaction value at residue j , where $I(j)^{non-bonded}$ is divided by the mean of $I(j)^{non-bonded}$ overall 40 residues. Non-bonded interactions were considered as they

define the main hotspots.[45] In equation 3, for a given residue, j , and frame, i , $c_i = 1$ if $I(j)_{norm} \geq 1$, otherwise 0. Equation 3 is applied over all 10 representative conformers obtained from clustering. The I_{score} threshold of 4 was arbitrarily chosen to highlight residues that are often visited by excipients and hence potential hotspots. Figure 2B shows the plectasin structure colored based on I_{score} . Residues Gly1, Phe2, Glu10, His16, His18, Tyr25, Tyr29, Cys30, Val36, Cys37, and Lys38 have $I_{score} > 4$. Similar segments (Gly1, Phe2, Gln5, Tyr29, Lys32, Gly33, Phe35, Val36, Lys38, and Tyr40) are also found to be APRs (Figure 2A).

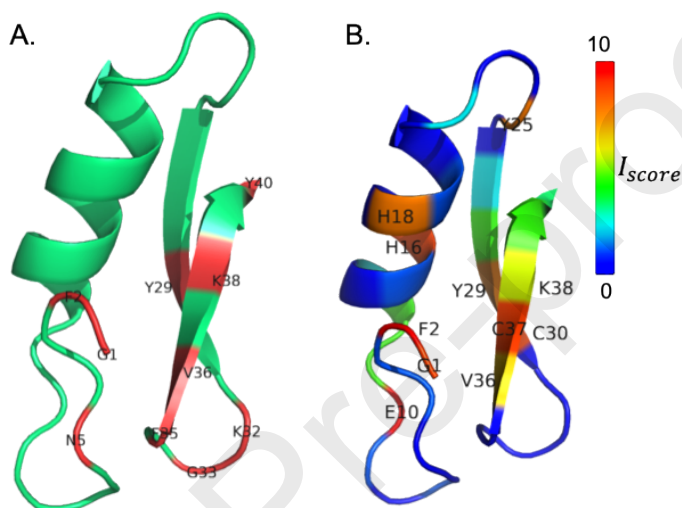


Figure 2. A. APRs are shown in red on the plectasin structure. The residue is colored red if it is predicted as aggregation prone by any one of the webservers (AGGRESCAN3D[26], CAMSOL[27], CSSP[43]). B. Structure coloring based on FTMap score. Residues with a score above an arbitrarily chosen threshold ($I_{score} > 4$) are labeled and considered as protein-excipient interaction hotspots. These residues are Gly1, Phe2, Glu10, His16, His18, Tyr25, Tyr29, Cys30, Val36, Cys37, and Lys38.

Site-specific binding. The previous results from FTMap were based on representative peptide structures obtained from 100 ns MD simulation. To access the effect of peptide flexibility on peptide-excipient binding, MD simulations of the complexes were performed for 500 ns with the same setting as described above. The excipients that had relatively low affinities (strong binding) in MST experiments were taken for independent MD simulations. These are arginine, histidine, trehalose, succinate, and Tris. Glycerol was taken as a negative control to access its weak binding. For each excipient, the concentration was adjusted to 150 mM, which corresponds to 31 excipient molecules for the solvated system containing approximately 11250 water molecules (as described in MD methodology section). Predominantly, four regions were identified as potential excipient binding sites using FTMap (red arrows on top of Figure 3), which were also predicted as the main binding sites for small molecules using the SiteMap program in the Schrödinger suite 2016-3 (data not shown). For simplicity, we have assigned four sites in this order; residues Gly1, Phe2, Val36, Cys37, and Lys38 belong to site I, residues His16, His18, Tyr29, and Cys30 belong to site II, and regions around Glu10 and Tyr25 are annotated as site III and IV, respectively. Figure 3 provides an overview of the excipient binding in different regions at the peptide surface. Relatively high contact frequency is found for the four sites. Figure 4 summarizes *in-silico* results in the sequence positions of plectasin. The residues marked in red are predicted hotspot regions from FTMap, of which some of the predicted regions are also APRs. It is observed that excipients arginine, histidine, and Tris make more contact with the peptide (Figure 3), which are also the ones with relatively strong affinities (Table 1). On contrast, glycerol makes the least contact of all excipients

during the MD simulations (Figure 3), which is in good agreement with MST experiments where we see negligible binding (Table 1).

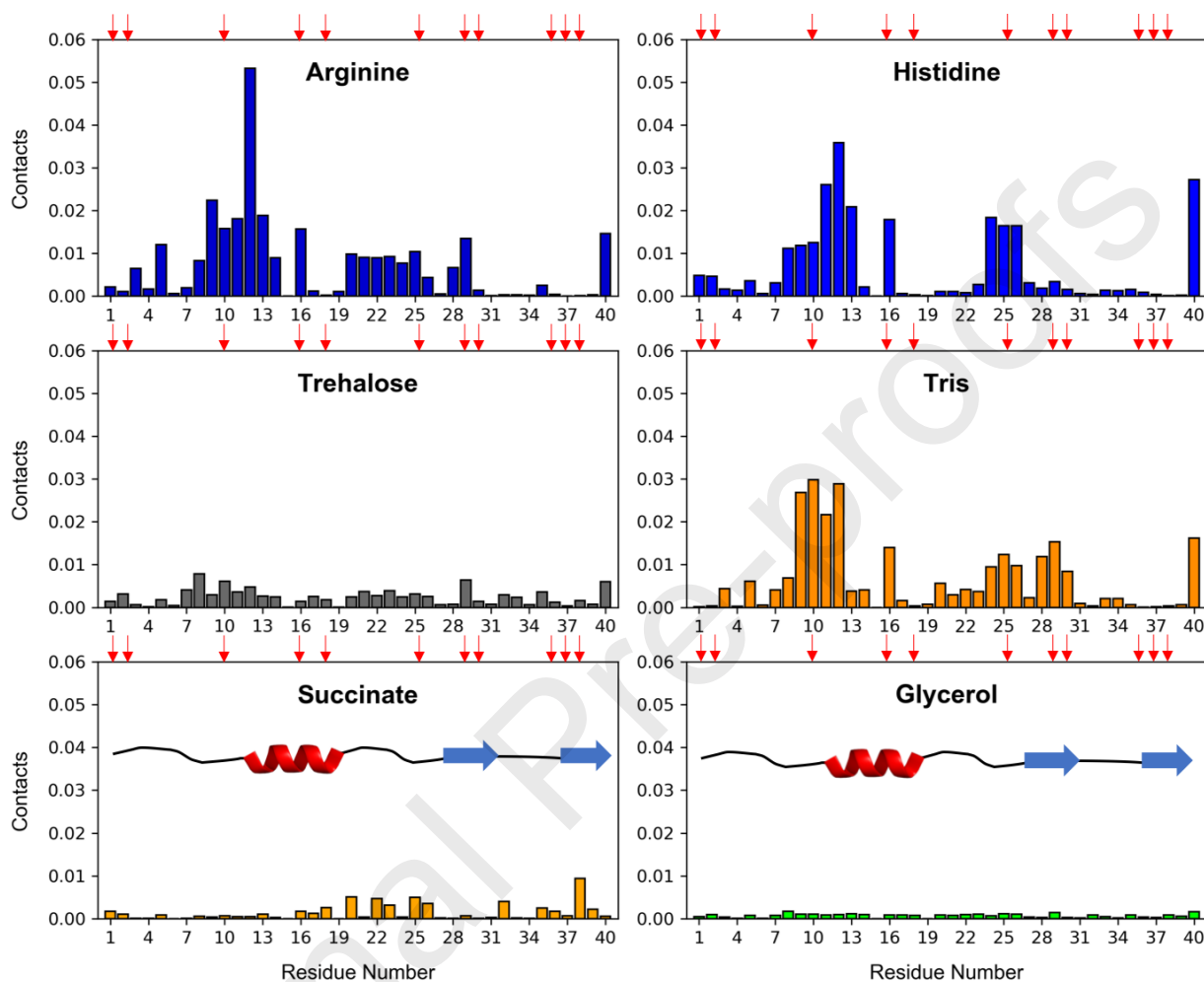


Figure 3. The contacts formed between plectasin and the different excipients. The number of contacts is normalized by the sum of contacts calculated over all excipients. The hotspot regions predicted by FTMap are marked as red arrows on top of each Figure.

NMR ^1H - ^{13}C HSQCs titration experiments. The backbone and side-chain assignments of plectasin were achieved as previously described[39] and are shown in Figure S1 and S2. To identify interaction sites for excipients on plectasin, ^1H - ^{13}C HSQCs titration experiments were performed in 10 mM acetate buffer, pH 4 for all the 12 excipients (for experimental details see Table SI). An example of ^1H - ^{13}C HSQC spectra overlays for arginine as an excipient is shown in

Figure 5. For the other excipients, HSQC overlays are given in the Supporting Information (Figures S3-13).

APR	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
FTMap	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
NMR consensus	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Arginine	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Histidine	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Methionine	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Glycine	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Proline	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Glycerol	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Mannitol	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Sorbitol	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Sucrose	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Trehalose	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Tris	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
Succinate	GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY
	1 5 10 15 20 25 30 35 40

Figure 4. A. Mapping the results of *in-silico* calculations and NMR signal perturbation mapping on sequence positions of plectasin. For APR, residues that were predicted as aggregation prone from different servers are marked in red. For FTMap, residues with $I_{score} > 4$ are marked in red. The distinct classes of excipients are colored differently. Amino acids (dark blue), sugar glycols (green), sugar (gray), and buffering agents (orange). In NMR consensus, residues that are perturbed significantly (CSP > 0.025 ppm) in the presence of at least five different excipients are marked in red. Finally, for each of the different excipients, residues that showed a significant perturbation in NMR experiments (CSP > 0.025 ppm) are marked in red.

Maximum chemical shift perturbations were calculated for each residue in the presence of different excipients, and CSPs above 0.025 ppm were considered to be significant (Figure 6). Additionally, significant chemical shift perturbations (CSP > 0.025 ppm) were mapped onto the plectasin structure (Figure 7A-I). It is observed that amino acids and buffering agents interact at many sites on the peptide, as highlighted in red, whereas sugars and sugar glycols interact in fewer places (Figure 7).

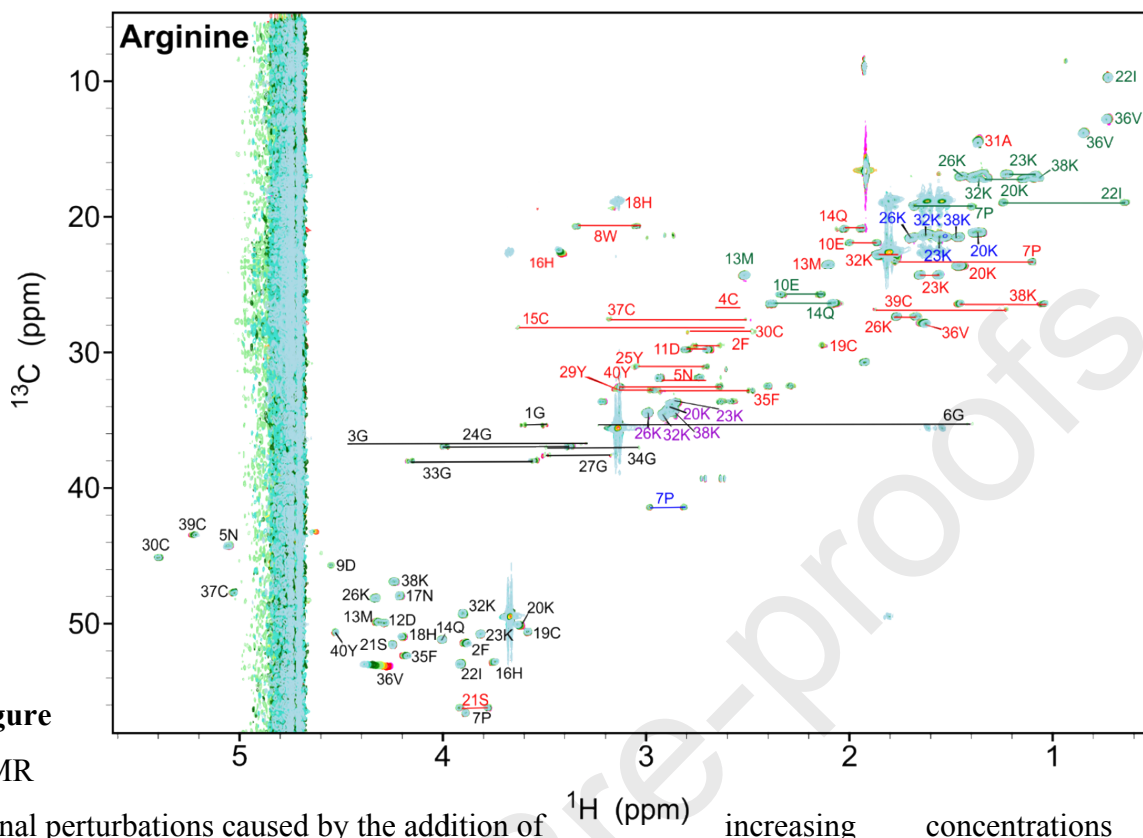


Figure
NMR

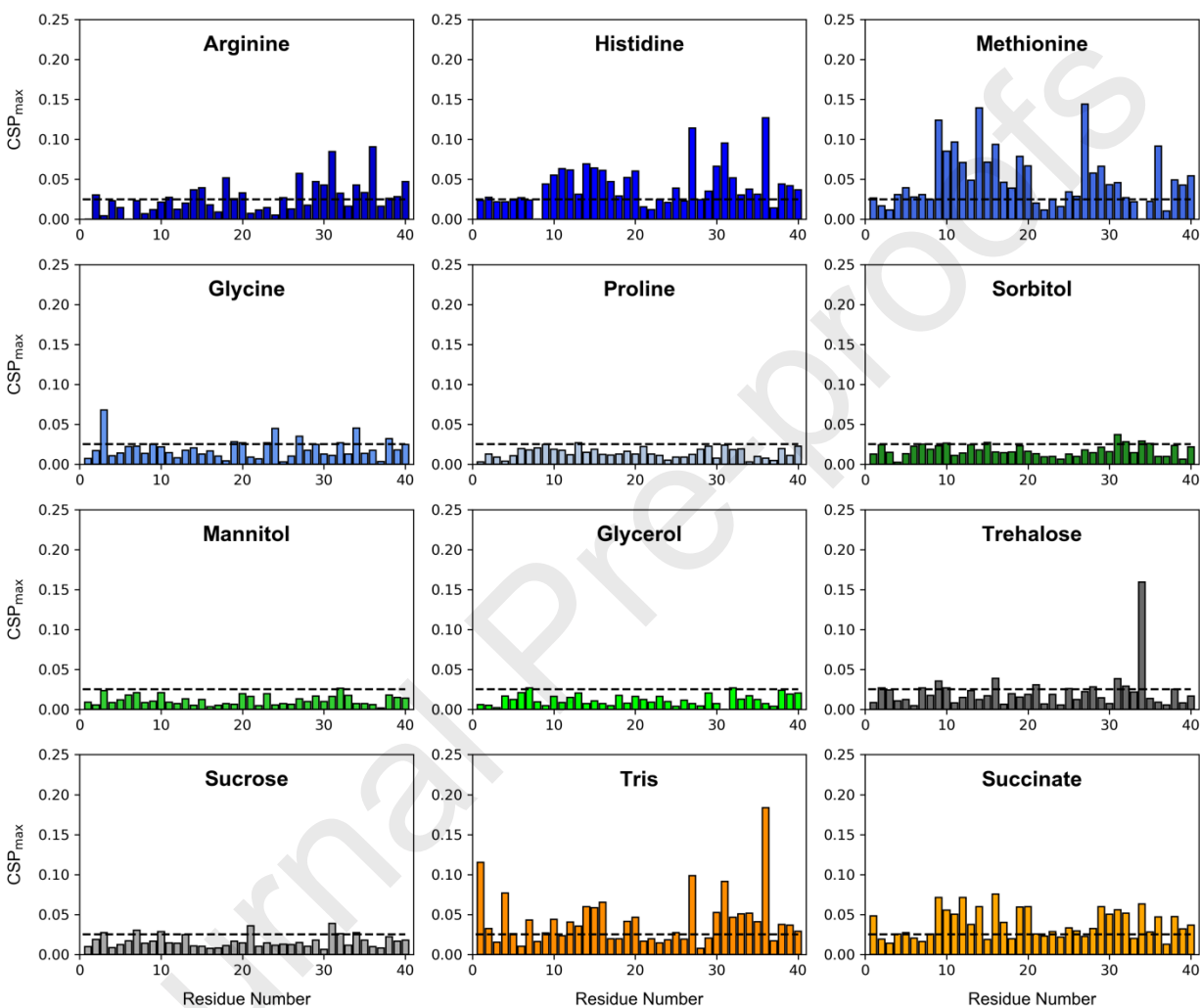
5.

signal perturbations caused by the addition of ^1H (ppm) increasing concentrations of excipient. ^1H - ^{13}C HSQC spectra overlay of plectasin in the absence (magenta) and in the presence of 1 mM (red), 5 mM (orange), 7.5 mM (yellow), 10 mM (light green), 15 mM (green), 20 mM (dark green), 30 mM (cyan) and 40 mM (light blue) arginine. Assignments are noted next to the corresponding peaks. Peaks labeled in black, red, green, purple, and blue belong to $\text{C}_\alpha\text{-H}_\alpha$, $\text{C}_\beta\text{-H}_\beta$, $\text{C}_\gamma\text{-H}_\gamma$, $\text{C}_\epsilon\text{-H}_\epsilon$, and $\text{C}_\delta\text{-H}_\delta$ groups. Lines in the spectra connect peaks from two protons that are connected to the same C atom.

Similar to I_{score} in the FTMap analysis, an NMR_{score} was assigned to each residue. NMR_{score} is defined as the number of times a given residue is perturbed significantly due to interactions with the different excipients. A threshold of five ($NMR_{score} > 5$) was arbitrarily chosen to define a residue as a protein-excipient interaction hotspot based on NMR experiments (Figure 7M). Perturbations with $\text{CSP} > 0.025$ ppm were used to estimate the K_d values for local binding (Figure S14). Furthermore, site-specific K_d values at site I and II were calculated (see NMR section in SI). Previous studies have shown that affinities as weak as 10 mM can be determined using NMR, which is the upper limit that can be calculated.[40] The excipients used in this study show weak interactions (K_d in the 10-500 mM range), which make it challenging to determine K_d values

precisely using NMR (Figure S14). Figure 4 maps the NMR results onto the sequence positions of plectasin, where residues that are significantly perturbed are marked in red.

Figure 6. The maximum CSP mapping of plectasin upon the addition of different excipients. The dotted line represents the threshold (CSP = 0.025 ppm) above which the perturbations are



considered to be significant.

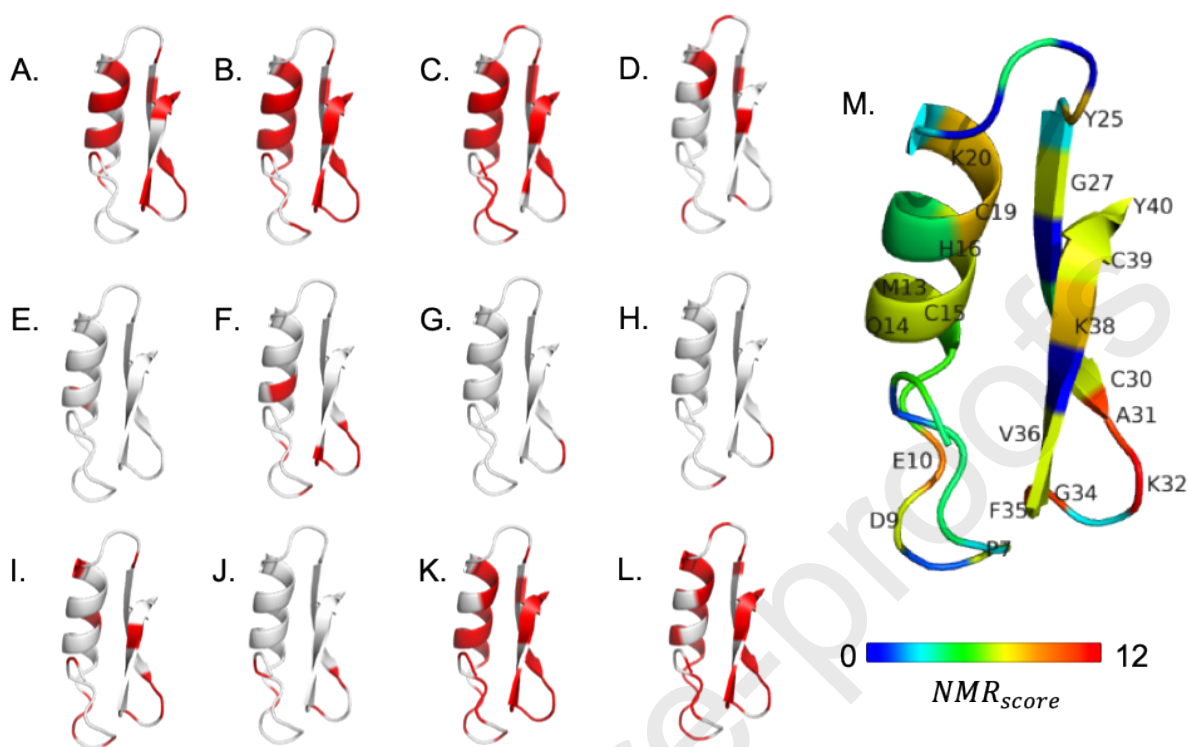


Figure 7. Structure coloring of plectasin based on CSP. Regions with significant perturbations (CSP > 0.025 ppm) are colored red. A. Arginine, B. Histidine, C. Methionine, D. Glycine, E. Proline, F. Sorbitol, G. Mannitol, H. Glycerol, I. Trehalose, J. Sucrose, K. Tris, and L. Succinate. M. Structure coloring based on NMR_{score} . Residues with a score above an arbitrarily chosen threshold ($NMR_{score} > 5$) are labelled and considered as main interaction hotspots for excipient interactions.

DISCUSSION

Microscale thermophoresis is a robust and fast method used to characterize peptide-excipient interactions and their binding affinities. As the sample is free to move during the measurement, it creates conditions close to the native state.[46] We used a label-free system, which takes into account the intrinsic tryptophan fluorescence of plectasin. It measures the change in thermophoresis upon direct binding of small molecules to the target peptide, thus allowing quantitative measurement of direct binding affinity. Based on the MST approach, an initial ranking of excipients based on binding affinities to plectasin was obtained. Effective overall K_d values were predicted in the case of arginine (35 mM), histidine (46 mM), Tris (56 mM), and trehalose (65 mM), which reached saturation in the dose-response curve (Figure 1) resulting in K_d values with fairly low standard deviations (Table 1).

In order to study the excipient interaction sites on the peptide surface on a molecular level, FTMap, a docking method, was used.[23,45] FTMAP is a direct computational equivalent of the experimental screening approaches, and have been previously used in various studies to identify potential binding hotspots of small molecules on protein surfaces.[47–49] In this study, FTMap was able to pick a few prominent hotspots, which we classified as sites I-IV, which consists of residues: Gly1, Phe2, Glu10, His16, His18, Tyr25, Tyr29, Cys30, Val36, Cys37, and Lys38 (Figure 2B and 4). Hotspot is defined as a residue with $I_{score} > 4$, which implies any listed excipient is certainly going to interact with the identified hotspots. Although the FTMap predicted hotspots are unspecific, the hotspots provide an overall picture about the possible binding sites irrespective of the excipient type. Some of the hotspot regions are also identified as aggregation prone regions (residues: 1-2, 29, 36-40) (Figure 2B and 4).

Independently, constant pH MD simulations were performed considering arginine, histidine, Tris, trehalose, succinate and glycerol as excipients. The five excipients except glycerol showed relatively strong binding to plectasin, whereas glycerol did not show any binding. In Figure 3, it is observed that all excipients make more contacts compared to glycerol with the peptide surface. Comparing Tris and succinate, it is observed that succinate makes fewer contacts with plectasin than Tris, which correlates well with the MST results indicating that succinate binds only weakly to plectasin. In case of arginine, histidine and Tris even though the apparent affinities discussed

here are in the millimolar range, specific binding of excipients at certain regions on the peptide surface has the potential to affect peptide stability – either in a positive or in a negative way.[11,42,50] Each of the three excipients carry an overall net charge of +1 at the pH 4, which leads to stronger interactions with the tetrapeptide segment (9-12) mainly due to the negatively charged patch (Site III) and around Tyr25 (Site IV); both of which are predicted hotspot regions from FTMap (Figure 3). Rather weak interactions are observed at site I that constitutes of a flexible N-terminus and a beta strand, and it is the flexibility arising from the free end that causes weak binding. Likewise, excipients interact favorable with His16 of site II as discussed in the NMR section below. The *in-silico* method in combination with experimental methods is therefore useful for fast screening of potential excipients.

Additionally, series of NMR spectra of the peptide were recorded in the absence and presence of varying concentrations of excipients to validate the *in-silico* predictions. For all the 12 excipients, ^1H - ^{13}C HSQCs were recorded and processed considering the peaks that could be unambiguously assigned. Overall, the addition of excipient can lead to direct binding of excipient to the peptide that may be highlighted as ‘perturbations’ i.e. CSP. Significant CSP (CSP > 0.025 ppm) were mapped on to the structure (Figure 7A-L). Figure 7M shows the secondary structure of plectasin, where the coloring is based on the NMR_{score} , which is defined as the number of time a given residue is perturbed significantly in the presence of different excipient. NMR_{score} highlight regions that are often perturbed due to interactions with the different excipients and it is similar to I_{score} (from FTMap). It is observed that NMR and FTMap pick similar regions for peptide-excipient interactions. Residues in the C-terminal region (36-40) that constitute site I as well as APRs are identified by both techniques (Figure 3 and 7M). Furthermore, residues around His16 (Site II), regions around Glu10 (site III) and Tyr25 (site IV) have been identified by both *in-silico* and NMR approaches. For each excipient, interaction sites on the peptide are discussed below.

In general, regions having CSP below 0.025 ppm do not interact with excipients, or interactions are too weak to be detected by NMR. However, there are certain regions (CSP > 0.025 ppm) that interact relatively strongly with excipients (Figure 6 and 7). Although there were many peaks with noticeable perturbations, K_d values could be calculated only for some peaks as significant curvature of the CSPs against excipient concentration is required to determine accurate K_d values (Figure S14). In all cases, excipients binding to site II interact favorably with the side chains atoms

as compared to when they bind close to site I, where they interact with the backbone atoms (Figure 5, Figures S3-S13). Figure 5 shows ^1H - ^{13}C HSQC spectra overlay for arginine as an excipient, and peaks corresponding to $\text{C}_\alpha\text{-H}_\alpha$ of Gly34, and Val36 have noticeable perturbations. The average binding affinities at the two main binding sites are equivalent (site I: $K_d = 30$ mM and site II: $K_d = 18$ mM). It is the same for His as excipient (site I $K_d = 61$ mM as compared to site II $K_d = 52$ mM). His and Arg at pH 4 carry an overall +1 charge. Jungwirth and coworkers[51,52] showed in a computational study that His-His, Arg-Arg, His-Arg pairing makes favorable interactions in spite of alike charges. This could explain the favorable binding of His and Arg in site II (His16 and His18). Moreover, favorable electrostatic interactions of positively charged excipients (His, Arg and Tris) with the negatively charged tetrapeptide stretch leads to greater accumulation of excipients in this region. Further, it is observed that Arginine and Histidine as excipients show similar interaction patterns with the peptide, the respective side chains make less contact with the tetrapeptide stretch as opposed to the backbone atoms (Figure S15). Here, one can notice favorable electrostatic interactions of the charged excipients with the peptide backbone and certain amino acids on the peptide. Methionine as excipient has preference for site II (18 mM). Moreover, these excipients also bind to FTMap predicted hotspot at Tyr25 (site IV) with affinity in the range of 1-100 mM.

In the case of neutral or multi-OH containing molecules (sugars, sugar glycols, glycine, and proline), binding is weak (Figure 6D-J, Figure 7D-J). However, glycine and trehalose interact weakly (10-30 mM) with plectasin around site I and II (Figure 7, Figure S14). Previously, a study showed that aromatic residues on peptide surface frequently interact with sugars in the order Trp >> Tyr > His.[53] Hydroxyl groups on the sugars have the tendency to replace a water molecule close to the peptide during dehydration making itself available to interact with the polar patches on the peptide leading to peptide stabilization.[54] Alternatively, buffering reagents, Tris and succinate, interact with many segments on the peptide as arginine, histidine, and methionine. Tris and succinate prefer binding to site I with affinities around 19 mM and 56 mM, respectively. Overall, Tris (30 mM) has favorable interactions as compared to succinate (58 mM). This is in good agreement with MD simulations where Tris makes more contact with the peptide surface compared to succinate (Figure 3, Figure S15). Tris makes more contact with the peptide using the amino group as opposed to the hydroxyl groups (Figure S15). It should be cautiously noted that lysines form clustered peaks, for example, peaks corresponding to $\text{C}_\epsilon\text{-H}_\epsilon$, $\text{C}_\gamma\text{-H}_\gamma$, and $\text{C}_\delta\text{-H}_\delta$ groups

of lysines form clusters (Figure 5, Figures S3-13). Therefore, the peaks for these groups might be placed vaguely and may not imply true binding in these regions. On the whole, most of the excipients certainly bind in the predicted hotspot regions from FTMap, which also constitute the two predicted binding sites with higher binding affinity, for instance, Val36 and His16 that are part of site I and site II, respectively, are perturbed significantly (Figures S3-S13). The observed favorable peptide-excipient interactions could explain the improved peptide stability.[16,24]

Overall, there is a good correlation between the total sum of normalized contacts deduced from simulations and the overall K_d values measured using MST. For instance, Arginine that has high affinity towards plectasin and in comparison, to the other excipients makes more contacts with the peptide as observed in the MD simulations. The methodology presented here can provide insight into the molecular origin of the relative ranking of the excipients based on measured binding affinities to the protein and identify protein-excipients interactions sites deduced from *in-silico* modelling, which can certainly accelerate formulation procedure.

CONCLUSION

We show that computational approaches in combination with MST correlate well with the NMR titration experiments utilized here. Binding affinities of distinct excipients to plectasin were determined using the fast experimental screening approach, MST. For some excipients (arginine, histidine, trehalose, and Tris), apparent K_d values were predicted accurately due to 50% saturation of peptide with the excipient. In other cases, K_d values were calculated with relatively large uncertainties. Furthermore, these overall binding affinities were traced to a molecular level using *in-silico* methods. Using the FTMap approach, a quick blind docking of excipients were performed on different conformers of peptide obtained from MD trajectory clustering. Subsequently, simple statistical analyses were performed to give an interaction score per residue to identify regions that are prone to interaction with excipients of interest. Additionally, constant pH MD simulations were performed, and the results indicated that the five tested excipients, arginine, histidine, trehalose, Tris and succinate, bind favorably to regions corresponding to the FTMap predicted hotspot regions. Glycerol did not show any preference, which is in good agreement with the MST and NMR results presented in this study. These findings were further investigated using ^1H - ^{13}C HSQC titration experiments. Most of the peaks could be assigned explicitly. *In silico* approaches and NMR consensus regions picked analogous segments on the peptide as the most frequently

interacting regions with excipients (regions around Val36-Tyr40) that are also APRs. In principle, weak binding of excipients enhances colloidal stability at pH 4, and favorable electrostatic interactions of positively charged excipients (His, Arg and Tris) with the negatively charged tetrapeptide stretch leads to greater accumulation of excipients in this region.

We have shown MST works for peptides, but also for bigger proteins or more complex systems.[55,56] Our approach provides a methodology that can be extended to other biotherapeutics, such as mAbs and fusion proteins, in order to capture protein dynamics at desired physicochemical conditions. Such an approach in combination with coarse grained simulation can be performed to simulate on longer time scales in short duration. Previously, we have applied such an *in-silico* screening method in combination with small angle X-ray scattering to understand protein-excipient interactions and their effect on oligomerization and dynamics.[21,22] Thus, the discussed fast screening methodology can effectively be carried forward to study the effect of combination of different excipients and how it alters proteins dynamics and thus aggregation.

ASSOCIATED CONTENT**Supporting Information**

Methods and materials; Table S1, excipient stock solution concentration for the different excipients; Figure S1-S2, ^1H - ^{13}C NMR-HSQC spectra showing assignment for plectasin; Figure S3-S13, ^1H - ^{13}C NMR-HSQC spectra of plectasin at different concentration of 12 different excipients; Figure S14, mapping of binding affinities on plectasin; Figure S15, electrostatics surface coloring of plectasin and normalized contacts formed between excipients and plectasin.

AUTHOR INFORMATION**Corresponding Author**

E-mail: soemya@kemi.dtu.dk

Phone number: +45(45252486)

E-mail: ghp@kemi.dtu.dk

Phone number: +45(45252486)

AUTHOR CONTRIBUTIONS

S.I. performed the computational work, the NMR experiments with the help of M.Z. and A.P.G, the MST with the help of N.T, and drafted the manuscript. S.I. and G.H.J.P. contributed to the design of the computational study and discussion of the results. S.I performed the majority of the NMR data analyses such as sequence-specific signal assignment, peak picking, and CSP and K_d analyses. M.Z. and A.P.G. contributed to the design of the NMR study, peak picking, and CSP analyses. C.P. helped in the early design of the buffer conditions for MST. P.H., A.N., and W.S contributed to the interpretation of the results. All authors contributed to the writing of the manuscript.

ACKNOWLEDGMENTS

This study was funded by a project part of the EU Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie grant agreement No 675074 – “Protein-excipient Interactions and Protein-Protein Interactions in formulation” (PIPPi); www.pippi.kemi.dtu.dk. Simulations were performed on the CPU/GPU cluster at DTU Chemistry and the High Performance Computing cluster at DTU. NMR experiments were performed in the NMR Facility at the Manchester Institute of Biotechnology, The University of Manchester. MST experiments were performed at Nanotemper GmbH, Munich. Viscosities were measured at Ludwig Maximilian University of Munich, Germany with help from Dr. Hristo Svilenov and Dr. Andreas Tosstorff from the group of Professor Gerhard Winter.

ABBREVIATIONS

AMP, antimicrobial peptide; APR, aggregation prone region; CSP, chemical shift perturbation; MD, molecular dynamics; MST, microscale thermophoresis

REFERENCES

- [1] A. Nijnik, R. Hancock, Host defence peptides: antimicrobial and immunomodulatory activity and potential applications for tackling antibiotic-resistant infections, *Emerg. Health Threats J.* 2 (2009) e1.
- [2] J.L. Lau, M.K. Dunn, Therapeutic peptides: Historical perspectives, current development trends, and future directions, *Bioorg. Med. Chem.* 26 (2018) 2700–2707.
- [3] P.H. Mygind, R.L. Fischer, K.M. Schnorr, M.T. Hansen, C.P. Sönksen, S. Ludvigsen, D. Raventós, S. Buskov, B. Christensen, L. De Maria, O. Taboureau, D. Yaver, S.G. Elvig-Jørgensen, M. V. Sørensen, B.E. Christensen, S. Kjærulff, N. Frimodt-Møller, R.I. Lehrer, M. Zasloff, H.-H. Kristensen, Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus, *Nature.* 437 (2005) 975–980.
- [4] T. Schneider, T. Kruse, R. Wimmer, I. Wiedemann, V. Sass, U. Pag, A. Jansen, A.K. Nielsen, P.H. Mygind, D.S. Raventos, S. Neve, B. Ravn, A.M.J.J. Bonvin, L. De Maria, A.S. Andersen, L.K. Gammelgaard, H.-G. Sahl, H.-H. Kristensen, Plectasin, a Fungal Defensin, Targets the Bacterial Cell Wall Precursor Lipid II, *Science* (80-.). 328 (2010) 1168–1172.
- [5] J.A. Hoffmann, Phylogenetic Perspectives in Innate Immunity, *Science* (80-.). 284 (1999) 1313–1318.
- [6] P. Schmitt, R.D. Rosa, D. Destoumieux-Garzón, An intimate link between antimicrobial peptide sequence diversity and binding to essential components of bacterial membranes, *Biochim. Biophys. Acta - Biomembr.* 1858 (2016) 958–970.
- [7] H. Chen, R. Mao, D. Teng, X. Wang, Y. Hao, X. Feng, J. Wang, Design and pharmacodynamics of recombinant NZ2114 histidine mutants with improved activity against methicillin-resistant *Staphylococcus aureus*, *AMB Express.* 7 (2017) 46.
- [8] Z. Li, X. Wang, X. Wang, D. Teng, R. Mao, Y. Hao, J. Wang, Research advances on plectasin and its derivatives as new potential antimicrobial candidates, *Process Biochem.* 56 (2017) 62–70.

- [9] F. Xiang, Z. Xie, J. Feng, W. Yang, Z. Cao, W. Li, Z. Chen, Y. Wu, Plectasin, First Animal Toxin-Like Fungal Defensin Blocking Potassium Channels through Recognizing Channel Pore Region, *Toxins (Basel)*. 7 (2015) 34–42.
- [10] K. Klein, R.B. Grønnemose, M. Alm, K.S. Brinch, H.J. Kolmos, T.E. Andersen, Controlled Release of Plectasin NZ2114 from a Hybrid Silicone-Hydrogel Material for Inhibition of *Staphylococcus aureus* Biofilm, *Antimicrob. Agents Chemother.* 61 (2017).
- [11] K.L. Zapadka, F.J. Becher, A.L. Gomes dos Santos, S.E. Jackson, Factors affecting the physical stability (aggregation) of peptide therapeutics, *Interface Focus*. 7 (2017) 20170030.
- [12] C. Pohl, M. Zalar, I. El Bialy, S. Indrakumar, G.H.J. Peters, W. Friess, A.P. Golovanov, W.W. Streicher, A. Noergaard, P. Harris, The Effect of Point Mutations on the Biophysical Properties of an Antimicrobial Peptide: Development of a Screening Protocol for Peptide Stability Screening, *Mol. Pharm.* 17 (2020) 3298–3313.
- [13] R. Riek, D.S. Eisenberg, The activities of amyloids from a structural perspective, *Nature*. 539 (2016) 227–235.
- [14] K. Trainor, A. Broom, E.M. Meiering, Exploring the relationships between protein sequence, structure and solubility, *Curr. Opin. Struct. Biol.* 42 (2017) 136–146.
- [15] J. Valente, R. Payne, M. Manning, W. Wilson, C. Henry, Colloidal Behavior of Proteins: Effects of the Second Virial Coefficient on Solubility, Crystallization and Aggregation of Proteins in Aqueous Solution, *Curr. Pharm. Biotechnol.* 6 (2005) 427–436.
- [16] E.T. Maggio, Use of excipients to control aggregation in peptide and protein formulations, *J. Excipients Food Chem.* 1 (2010) 40–49.
- [17] A. Pandya, M. Howard, M. Zloh, P. Dalby, An Evaluation of the Potential of NMR Spectroscopy and Computational Modelling Methods to Inform Biopharmaceutical Formulations, *Pharmaceutics*. 10 (2018) 165.
- [18] T. Barata, C. Zhang, P. Dalby, S. Brocchini, M. Zloh, Identification of Protein–Excipient

- Interaction Hotspots Using Computational Approaches, *Int. J. Mol. Sci.* 17 (2016) 853.
- [19] S. Jo, A. Xu, J.E. Curtis, S. Somani, A.D. MacKerell, Computational Characterization of Antibody–Excipient Interactions for Rational Excipient Selection Using the Site Identification by Ligand Competitive Saturation-Biologics Approach, *Mol. Pharm.* (2020) acs.molpharmaceut.0c00775.
- [20] A. Tosstorff, G.H.J. Peters, G. Winter, Study of the interaction between a novel, protein-stabilizing dipeptide and Interferon-alpha-2a by construction of a Markov state model from molecular dynamics simulations, *Eur. J. Pharm. Biopharm.* 149 (2020) 105–112.
- [21] A. Kulakova, S. Indrakumar, P. Sønderby, L. Gentiluomo, W. Streicher, D. Roessner, W. Frieß, G.H.J. Peters, P. Harris, Small angle X-ray scattering and molecular dynamic simulations provide molecular insight for stability of recombinant human transferrin, *J. Struct. Biol.* X. (2019) 100017.
- [22] A. Kulakova, S. Indrakumar, P. Sønderby Tuelung, S. Mahapatra, W.W. Streicher, G.H.J. Peters, P. Harris, Albumin-nepilysin fusion protein: understanding stability using small angle X-ray scattering and molecular dynamic simulations, *Sci. Rep.* 10 (2020) 10089.
- [23] C.H. Ngan, T. Bohnuud, S.E. Mottarella, D. Beglov, E.A. Villar, D.R. Hall, D. Kozakov, S. Vajda, FTMAP: extended protein mapping with user-selected probe molecules, *Nucleic Acids Res.* 40 (2012) W271–W275.
- [24] T.J. Kamerzell, R. Esfandiary, S.B. Joshi, C.R. Middaugh, D.B. Volkin, Protein–excipient interactions: Mechanisms and biophysical characterization applied to protein formulation development, *Adv. Drug Deliv. Rev.* 63 (2011) 1118–1159.
- [25] M. Christie, D. Peritt, R.M. Torres, T.W. Randolph, J.F. Carpenter, The Role of Protein Excipient in Driving Antibody Responses to Erythropoietin, *J. Pharm. Sci.* 104 (2015) 4041–4055.
- [26] R. Zambrano, M. Jamroz, A. Szczasiuk, J. Pujols, S. Kmiecik, S. Ventura, AGGRESAN3D (A3D): server for prediction of aggregation properties of protein structures, *Nucleic Acids Res.* 43 (2015) W306–W313.

- [27] P. Sormanni, F.A. Aprile, M. Vendruscolo, The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility, *J. Mol. Biol.* 427 (2015) 478–490.
- [28] J.J. Irwin, T. Sterling, M.M. Mysinger, E.S. Bolstad, R.G. Coleman, ZINC: A Free Tool to Discover Chemistry for Biology, *J. Chem. Inf. Model.* 52 (2012) 1757–1768.
- [29] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, *J. Comput. Aided. Mol. Des.* 27 (2013) 221–234.
- [30] K. Mandal, B.L. Pentelute, V. Tereshko, V. Thammavongsa, O. Schneewind, A.A. Kossiakoff, S.B.H. Kent, Racemic crystallography of synthetic protein enantiomers used to determine the X-ray structure of plectasin by direct methods., *Protein Sci.* 18 (2009) 1146–1154.
- [31] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, *J. Med. Chem.* 47 (2004) 1739–1749.
- [32] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174.
- [33] A. Jakalian, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, *J. Comput. Chem.* 23 (2002) 1623–1641.
- [34] J.M. Swails, D.M. York, A.E. Roitberg, Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation, *J. Chem. Theory Comput.* 10 (2014) 1341–1352.
- [35] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D.E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field., *Proteins.* 78 (2010) 1950–8.

- [36] M.W. Mahoney, W.L. Jorgensen, A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions, *J. Chem. Phys.* 112 (2000) 8910–8922.
- [37] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, *J. Chem. Phys.* 98 (1993) 10089–10092.
- [38] R.W. Pastor, B.R. Brooks, A. Szabo, An analysis of the accuracy of Langevin and molecular dynamics algorithms, *Mol. Phys.* 65 (1988) 1409–1419.
- [39] S. Indrakumar, M. Zalar, C. Pohl, A. Nørgaard, W. Streicher, P. Harris, A.P. Golovanov, G.H.J. Peters, Conformational Stability Study of a Therapeutic Peptide Plectasin Using Molecular Dynamics Simulations in Combination with NMR, *J. Phys. Chem. B.* 123 (2019) 4867–4877.
- [40] M.P. Williamson, Using chemical shift perturbation to characterise ligand binding, *Prog. Nucl. Magn. Reson. Spectrosc.* 73 (2013) 1–16.
- [41] S. Ohtake, Y.J. Wang, Trehalose: Current Use and Future Applications, *J. Pharm. Sci.* 100 (2011) 2020–2053.
- [42] E.T. Maggio, Use of excipients to control aggregation in peptide and protein formulations, *J. Excipients Food Chem.* 1 (2010) 40–49.
- [43] C. Kim, J. Choi, S.J. Lee, W.J. Welsh, S. Yoon, NetCSSP: web application for predicting chameleon sequences and amyloid fibril formation, *Nucleic Acids Res.* 37 (2009) W469–W473.
- [44] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, *J. Chem. Theory Comput.* 9 (2013) 3084–3095.
- [45] D. Kozakov, L.E. Grove, D.R. Hall, T. Bohnuud, S.E. Mottarella, L. Luo, B. Xia, D. Beglov, S. Vajda, The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins, *Nat. Protoc.* 10 (2015) 733–755.

- [46] M. Jerabek-Willemsen, T. André, R. Wanner, H.M. Roth, S. Duhr, P. Baaske, D. Breitsprecher, *MicroScale Thermophoresis: Interaction analysis and beyond*, *J. Mol. Struct.* 1077 (2014) 101–113.
- [47] M.-S. Kim, L.A. Buisson, D.A. Heathcote, H. Hu, D.C. Braddock, A.G.M. Barrett, P.G. Ashton-Rickardt, J.P. Snyder, *Approaches to design non-covalent inhibitors for human granzyme B (hGrB)*, *Org. Biomol. Chem.* 12 (2014) 8952–8965.
- [48] D.H. Hall, L.E. Grove, C. Yueh, C.H. Ngan, D. Kozakov, S. Vajda, *Robust Identification of Binding Hot Spots Using Continuum Electrostatics: Application to Hen Egg-White Lysozyme*, *J. Am. Chem. Soc.* 133 (2011) 20668–20671.
- [49] D. Beglov, D.R. Hall, A.E. Wakefield, L. Luo, K.N. Allen, D. Kozakov, A. Whitty, S. Vajda, *Exploring the structural origins of cryptic sites on proteins*, *Proc. Natl. Acad. Sci.* 115 (2018) E3416–E3425.
- [50] S. Ohtake, Y. Kita, T. Arakawa, *Interactions of formulation excipients with proteins in solution and in the dried state*, *Adv. Drug Deliv. Rev.* 63 (2011) 1053–1073.
- [51] J. Vondrášek, P.E. Mason, J. Heyda, K.D. Collins, P. Jungwirth, *The Molecular Origin of Like-Charge Arginine–Arginine Pairing in Water*, *J. Phys. Chem. B.* 113 (2009) 9041–9045.
- [52] J. Heyda, P.E. Mason, P. Jungwirth, *Attractive Interactions between Side Chains of Histidine-Histidine and Histidine-Arginine-Based Cationic Dipeptides in Water*, *J. Phys. Chem. B.* 114 (2010) 8744–8749.
- [53] K.L. Hudson, G.J. Bartlett, R.C. Diehl, J. Agirre, T. Gallagher, L.L. Kiessling, D.N. Woolfson, *Carbohydrate–Aromatic Interactions in Proteins*, *J. Am. Chem. Soc.* 137 (2015) 15152–15160.
- [54] N. Soltanizadeh, L. Mirmoghtadaie, F. Nejati, L.I. Najafabadi, M.K. Heshmati, M. Jafari, *Solid-State Protein-Carbohydrate Interactions and Their Application in the Food Industry*, *Compr. Rev. Food Sci. Food Saf.* 13 (2014) 860–870.

- [55] T. Zhang, L. Nagel-Steger, D. Willbold, Solution-Based Determination of Dissociation Constants for the Binding of A β 42 to Antibodies, *ChemistryOpen*. 8 (2019) 989–994.
- [56] M. Ramakrishnan, F. Alves De Melo, B.M. Kinsey, J.E. Ladbury, T.R. Kosten, F.M. Orson, Probing Cocaine-Antibody Interactions in Buffer and Human Serum, *PLoS One*. 7 (2012) e40518.

Journal Pre-proofs

Graphical abstract

