

Site of Tagging Influences the Ochratoxin Recognition by Peptide NFO4: A Molecular Dynamics Study

Aby A. Thyparambil, Tigran M. Abramyan, Ingrid Bazin, Anthony

Guiseppi-Elie

► To cite this version:

Aby A. Thyparambil, Tigran M. Abramyan, Ingrid Bazin, Anthony Guiseppi-Elie. Site of Tagging Influences the Ochratoxin Recognition by Peptide NFO4: A Molecular Dynamics Study. Journal of Chemical Information and Modeling, 2017, 57 (8), pp.2035-2044. 10.1021/acs.jcim.7b00312 . hal-02892728

HAL Id: hal-02892728 https://hal.science/hal-02892728

Submitted on 5 Dec 2023 $\,$

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Site of Tagging Influences the Ochratoxin Recognition by Peptide NFO4: A Molecular Dynamics Study

Aby A. Thyparambil^{a,b}, Tigran M. Abramyan^c, Ingrid Bazin^d and Anthony Guiseppi Elie^{a,b,d,e,*}.

^a Center for Bioelectronics, Biosensors and Biochips (C3B), Texas A&M University, College Station, TX 77843, USA

^b Department of Biomedical Engineering, College of Engineering, Texas A&M University, College Station, TX 77843, USA

^c Computational Biophysics & Molecular Design, Center for Integrative Chemical Biology and Drug Discovery, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599-7363, USA

^d Ecole des mines d'Ales, institut Mines Telecom, 6 avenue de Clavieres, 30319 Ales cede, France

^e ABTECH Scientific, Inc., Biotechnology Research Park, 800 East Leigh Street, Richmond, VA 23219, USA

KEYWORDS. Peptides, Hapten, Ochratoxins, Affinity Tag, Binding Energy, Bias Exchange Metadynamics, Markov State Models, All-Atom Molecular Dynamics.

ABSTRACT: Molecular recognition by synthetic peptides is growing in importance in the design of biosensing elements used in the detection and monitoring of a wide variety of hapten bioanlaytes. Conferring specificity via bio-immobilization and subsequent recovery and purification of such sensing elements are aided by the use of affinity tags. However, the tag and its site of placement can potentially compromise the hapten recognition capabilities of the peptide, necessitating a detailed experimental characterization and optimization of the tagged molecular recognition entity. The objective of this study was to assess the impact of site-specific tags on a native peptide's fold and hapten recognition capabilities using advanced molecular dynamics (MD) simulation approach involving bias exchange metadynamics and Markov state models. The in-solution binding preferences of affinity tagged NFO4 (VYMNRKYYKCCK) to chlorinated (OTA) and non-chlorinated (OTB) analogues of ochratoxin were evaluated by appending hexa-histidine tags (6× His-tag) to the peptide's N-terminus (NterNFO4) or C-terminus (CterNFO4), respectively. The untagged NFO4 (NFO4), previously shown to bind with high affinity and selectivity to OTA, served as the control. Results indicate that the addition of site-specific $6 \times$ His-tags altered the peptide's native fold and the ochratoxin binding mechanism, with the influence of site-specific affinity tags being most evident on the peptide's interaction with OTA. The tags at the N-terminus of NFO4 preserved the native fold and actively contributed to the non-bonded interactions with OTA. In contrast, the tags at the C-terminus of NFO4 altered the native fold and were agnostic in its non-bonded interactions with OTA. The tags also increased the penalty associated with desolvating the peptide-OTA complex. Interestingly, the tags did not significantly influence the non-bonded interactions or the penalty associated with solvating the peptide-OTB complex. Overall, the combined contributions of non-bonded interaction and solvation penalty were responsible for the retention of the native hapten recognition capabilities in NterNFO4 and compromised native recognition capabilities in CterNFO4. Advanced MD approaches thus provide structural and energetic insights critical to evaluate the impact of site-specific tags, and may aid in the selection and optimization of the binding preferences of a specific biosensing element.

57

59

61

INTRODUCTION

34 A major percentage of the pesticides, herbicides, toxins, metals and allergens for which there is need to design and 35 36 fabricate biosensors are potential haptens.¹ Haptens are a class 37 of low molecular weight chemicals that are inherently not 38 antigenic but can elicit an immune response when conjugated 39 to serum proteins or low molecular weight peptides.¹ Prolonged exposures to some haptens have been linked to multi-40 41 ple health risks including organ failure, increased cancer risk, 42 genotoxicity, autoimmune disorders, and hypersensitive reactions.²⁻⁴ During the last decade a wide range of sensing and 43 detection platforms involving the use of synthetic peptides 44 have been developed against a variety of haptens such as 45 ochratoxin-A (OTÅ), 2, 4, 6 trinitrotoluene, and 2,4 dichloro-phenoxyacetic acid.⁵⁻¹⁰ In all these applications, site-specific 46 47

conjugation of the molecular recognition peptide to achieve 48 49 its immobilization on or within the biosensing element or 50 bioassay was enabled by the use of site-specific affinity tags. The tags of the conjugates were essential to ensure the immo-51 52 bilization, efficient hapten recognition, recovery, stability, and cost-effective operation of the detection platform. In this re-53 54 gard, affinity tags are a very popular type of tag as they allow 55 for orientation of the recognition peptide and possible reusability of the biotransducer platform.^{11, 12} 56

The introduction of affinity tags generally involves the co-58 valent conjugation of contiguous amino acids, peptides, enzymes, and protein domains to specific sites of the molecular 60 recognition entity. These tags then bind with high specificity to a target biological or a chemical ligands on the adsorbent phase. In so doing the tag facilitates the immobilization of the 62

1 molecular recognition entity on the detection platform while 62 optimally orienting it towards the solution phase for hapten 2 63 capture.¹³ Furthermore, immobilization aids in the cost-3 64 effective usage of the molecular recognition platforms, by 4 65 facilitating the recovery, stability, and possible reusability of 5 66 6 the detection platform from the bioanalyte matrix. However, a 67 7 major challenge in the design of biosensing elements pertains 68 8 to the selection and location of affinity tags, so as not to com-69 9 promise the inherent activity, specificity, and selectivity of the 70 10 molecular recognition entity. Ideally, affinity tags are ex-71 pected to be small in size to prevent or minimize direct inter-72 11 12 action with the biosensing elements or with the analyte of 73 13 interest. Direct interactions of tags with the biorecognition 74 14 peptide can also affect the orientation and the molecular 75 recognition properties of the sensing element. Tagging the 76 15 16 molecular recognition entity is particularly more challenging 77 17 when the tags and sensing elements are of comparable size 78 18 and chemistry.

19 Hexahistidine tag (6×His-tag) is one such example of an af-80 finity tag that binds with high affinity to substrates coated with metal ions, specifically zinc.^{8, 14} Because of its short 81 20 82 21 22 length and relative ease of synthesis and removal, 6×His-tags 83 23 are widely used in protein purification and immobilization. 84 24 ¹⁷ The introduction of 6×His-tag may have either positive or 85 25 negative effects on the biochemical properties of a molecular 86 recognition entity of a sensing element.^{7, 8, 18-22} For example, 26 87 the addition of 6×His-tags has been found to increase the sen-88 27 sitivity of binding assays with tagged ScFv.²³ But, in case of 28 89 29 glycine N-acyltransferase GNAT peptides from Pseudomonas 90 30 aeruginosa, the 6×His-tags were found to act as weak com-91 petitive inhibitors and interfered with the intrinsic activity of 31 the peptide.²² In addition to the type of tag, the site of tag con-32 92 jugation may also affect the intrinsic biochemical activity of 33 93 34 the molecular recognition entity of a sensing element.¹⁸ For 94 35 example, drugs conjugated to antibodies via glutamine tags 95 36 showed differences in their pharmacokinetics depending on 96 37 the location of the drug on the heavy and light chains of the 97 antibodies.²⁴ Similarly, on solid platforms, a synthetic peptide 38 98 39 (12 amino acid long) oriented with 6× His-tag on the N-99 40 terminus was found to be significantly better at hapten recog- 100 41 nition than those oriented with the affinity tags on the C- 101 terminus of the peptide.^{7, 8} Clearly, the impact of site-specific 102 42 43 tags on the biological recognition capabilities of sensing ele- 103 44 ments is non-trivial, especially when the tags are of compara- 104 45 ble size and chemistry and must be thoroughly assessed be- 105 46 fore being reliably applied for biosensing applications. 106 107

47 One approach to identifying a site-specific tag suitable for 108 48 peptide immobilization involves cyclic testing of known tags 109 at different sites using routine experimental biophysical tech- 110 49 niques, such as equilibration dialysis^{25, 26} or surface plasmon 111 50 resonance.^{27, 28} Unfortunately, these techniques lack the reso- 112 51 52 lution of sub-molecular events and the likelihood of such an 113 53 approach leading to an optimized design is extremely small. A 114 sub-molecular understanding of the impact of tags on the pep- 115 54 55 tide-hapten binding is essential because of the dynamic nature of the peptide. At equilibrium, peptides exist in an ensemble 56 of conformational states with their respective probabilities.²⁹ 57 ³² Within this ensemble of peptide conformations, haptens 58 59 bind with high affinity to conformations that produce the 60 bound state of the peptide, while weakly interacting hapten-61 peptide pairs are associated with other conformational states.

It is important to note that this may not be a single confor mation but in fact a smearing of energetically similar confor mations or an alternate minimum. Consequently, haptens that bind to the peptides with high affinity tend to shift the popula tion distribution favoring the bound states of the peptide hapten complex; and haptens that bind to the peptides with low affinity tend to skew the equilibrium towards the unbound states.^{31, 33} The foregoing requires that a clear energetic distinction be established between bound and unbound states. In essence, the impact of tags on a peptide's hapten binding properties, including its affinity, specificity and selectivity, can be thoroughly assessed, or even engineered, by considering the population distribution and redistribution of a peptide's conformational states, in tandem with the structureenergetic characteristics of the bound states. Undoubtedly, in silico techniques like molecular dynamics (MD) are the best tools for visualizing such sub-molecular events.

However, regular molecular dynamics (MD) approaches alone are inept in generating thermodynamically favored conformations of peptide-hapten complexes or in describing the kinetics of peptide-hapten binding.^{30, 33} Regular MD simulations typically capture biophysical phenomena occurring on the order of hundreds of nanoseconds, while the unbinding kinetics of a peptide-hapten system with high affinity would involve time scales often exceeding several hours. As an alternative, the authors have recently demonstrated the capabilities of advanced MD simulations involving accelerated configurational search and Markov state models (MSM) to assess the mechanism involved in the binding of synthetic peptides to a hapten such as ochratoxin.³⁴

79

The objective of the current study was to assess the impact of dimensionally and chemically comparable site-specific tags on a peptide's fold and hapten recognition. For this purpose, the impact of 6× His-tag and its site of placement on the native fold of NFO4 and its native in-solution binding behavior were assessed using an in silico approach involving enhanced MD sampling in conjunction with MSM. NFO4, a dodecamer (VYMNRKYYKCCK), was previously shown to bind with high affinity and selectivity to the chlorinated analogue (OTA) of a hapten like ochratoxin as opposed to a nonchlorinated analogue (OTB) (Figure 1).^{6, 34, 35} The affinity tags were "covalently conjugated" via peptide bond formation to the N-terminus (NterNFO4) and C-terminus (CterNFO4) of the peptide. Untagged NFO4 served as the control. All binding energy calculations were carried out in explicit solvent with the solution pH mimicking wine (pH 3-4), a matrix of concern for the presence of ochratoxins. Advanced sampling was performed using bias-exchange metadynamics (BEMD), which involves the introduction of time dependent bias potentials to sample rare events and generate experimentally relevant biomolecular conformations.³⁶ The equilibrium dynamics of the system was then predicted based on the conditional transition probability of diverse peptide conformations using MSM functionalities provided within HTMD.³⁷



2 Figure 1: Chemical structure of (a) ochratoxin-A (OTA), (b) 3 ochratoxin-B (OTB), (c) the dodecapeptide NFO4, (d) NterN-4 FO4 and (e) CterNFO4. The structural analogues of ochratox-5 ins (i.e. OTA and OTB) differ in the 'Cl' moiety highlighted in red. Similarly, the peptides differ in the location of 6× His-6 7 tags with the NFO4 lacking a tag, NterNFO4 having a tag at 8 the N-terminus of NFO4 (blue dotted box), and CterNFO4 9 having a tag at the C-terminus of NFO4 (red dotted box).

EXPERIMENTAL PROCEDURES 10

11 **Model Setup:** Linear structures of VYMNRKYYKCCK (NFO4), HHHHHH-VYMNRKYYKCCK (NterNFO4) and 12 13 VYMNR KYYKCCK-HHHHHHH (CterNFO4) were generated from the primary sequence of the respective peptides using 14 Visual Molecular Dynamics (v 1.9.2).³⁸ The initial coordi-15 nates of ochratoxins (OTA and OTB) were obtained from the 16 ZINC database; and the topology and parameters for the 17 18 ochratoxins were generated through the automated topology builder (ATB) and repository 2.1.39, 40 The amino acids and 19 20 the hapten were modeled with the likely protonation state in a 21 wine-like matrix of pH 3-4. Consequently, the N-termini, histidines, arginines, and lysines within the peptide were pro-22 23 tonated, and the C-termini, aspartic and glutamic acids, and 24 the ochratoxins were modeled as uncharged or neutral.

25 The simulation procedure was previously described in de-26 tail.³⁴ Briefly, MD simulations were carried out in a GROMACS (v 5.0.4) simulator.⁴¹ GROMOS 54A7 force field 27 and SPC water model was used to describe the molecular mechanics of the modelled system.^{42, 43} Long-range electro-28 29 30 static interactions (real-space truncation at 1.4 nm and grid 31 spacing of 0.12 nm) was handled using periodic boundary 32 conditions with particle mesh Ewald summation, and was 33 updated every 10 fs, together with the pair list generation. The 34 Lennard-Jones 6-12 potential was used to evaluate the van 35 der Waals interactions within a cutoff distance of 1.4 nm and 36 was updated at every step. The LINCS algorithm was used to 37 constrain the lengths of covalent bonds and the geometry of 38 the water molecules. A coupling scheme using velocity rescal-39 ing with a stochastic term was applied to maintain the temper-40 ature at 298 K with the modified Berendsen coupling method 41 and a relaxation time constant of 0.1 ps. The pressure of the 42 system was maintained at 1 bar using Parrinello-Rahman cou- 100 43 pling method with a time step of 2 fs and a relaxation time 101 constant of 2.0 ps. The starting atomic velocities were gener- 102 44 103

45 ated randomly using Maxwell-Boltzmann distribution at 298 46 K.

47 Sampling of Kinetically Relevant Peptide and Pep-48 tide-Hapten Configurations. Prior to binding simulations, the peptide configurations were sampled using BEMD.44, 45 49 All BEMD simulations were performed in the NPT ensemble 50 51 at 298K and 1 bar using GROMACS (v 5.0.4) compiled with the PLUMED (v 2.2.2) plugin.⁴⁶ Three independent BEMD 52 53 simulations were initialized from random configurations of 54 the peptide. For each BEMD simulation of peptide folding, 55 three replicas were used. Each of the first two replicas was 56 biased by a different collective variable (CV), namely, main 57 chain radius of gyration (R_{e}) , and number of hydrogen bonds 58 (N_{hb}) within the peptide main chain. The third replica was 59 simulated without any bias. From the peptide folding simula 60 tions, four macrostates were identified and were used in the 61 binding simulations. For each of the sampled peptide configu 62 rations, binding simulation was initiated by biasing five dif ferent CVs. The first four replicas were biased by main chain 63 R_{g} , N_{hb} between the peptide and haptens, degree of similarity 64 between the torsional angles traversed by the peptide and hap 65 66 ten (Φ_{corr}), and the distance between the center of mass of the 67 hapten and center of mass of the peptide (d_1) . The fifth replica 68 was simulated without any bias.

The convergence and equilibrium state of the BEMD trajec tories were verified by calculating the potential of mean force (PMF) along different CVs (Figure S1). The evolution of each replica in the free energy landscape was further monitored and traced using the replica index. For the peptide folding and binding simulations, all CVs could reconstruct similar PMF profile. However, there were differences in the convergence time. In general, the simulation using CVs such as Rg, Nhb, and d₁ reached convergence in relatively similar simulation time (<30 ns). PMF took longer to converge (>40 ns) when biased along Φ_{cor} . When all four CVs were used, the average replica exchange probability was ~ 0.26 . After equilibration and convergence, the simulations were extended for an addi tional 100 ns.

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

MSM Construction for Predicting Folded Conformations and Binding Kinetics. MSM construction, analysis and validation were achieved using the functionalities provided with HTMD (v 1.7.0).^{37, 47, 48} It should be pointed out that the main goal of the application of the MSM analysis in the current study was to obtain the population distribution of the macrostates of each molecular system rather than the transition rates among these macrostates. Also, the terminology 'macrostates' was adopted from the HTMD documentation to indicate the metastable states of the peptide. Distinct folded conformations of the peptide were identified from MSM by featurizing the intramolecular contacts made by the alpha carbon atoms of the peptide. For the binding simulations, all the distances between the heavy atoms of the hapten and the peptide were used as the metric to resolve the conformational space. For the current study, the unbiased replica from three independent simulations was used for the MSM model construction .Two atoms were in contact if their distance was less than 5 Å. Time-lag independent component analysis (IC) was performed on the featurized trajectories to find the slow linear sub-space of the input features. IC components that accounted

1 for > 90% of the total kinetic variance in the configurations 2 were retained for projection and analysis. For binding simula-3 tions, dimensionality reduction was achieved by projecting 4 the IC components onto five dimensions. Each IC component 5 was subsequently scaled according to its corresponding ei-6 genvalue to obtain a kinetic map in which Euclidean distances 7 were proportional to kinetic distances, providing an optimal 8 space to perform clustering. Mini batch k-means clustering 9 method was employed to obtain 100 microstates.

⁹ method was employed to obtain 100 microstates.

10 Markov model was constructed by the analysis of the implied time scales (ITS) (Figure S2). The validity of the Mar-11 kov model was verified using Chapman-Kolmogorov. The 12 robustness of MSM was also assessed by increasing the num-13 14 ber of clusters and projected dimensions. The number of 15 macrostates was determined by the separation of the transi-16 tions in the ITS plot. Macrostates were identified by lumping 17 kinetically close microstates using the Perron cluster cluster analysis (PCCA+) lumping algorithm. For each macrostate, 18 19 the binding free energy (equation 1) and the equilibrium dis-20 tribution of the peptides were computed from the reversible transition matrix by comparing the probabilities of bound to 21 22 unbound states. The reversible transition matrix was estimated 23 using the maximum likelihood estimator. The binding free 24 energies of the peptide-hapten complex at equilibrium were 25 calculated with the following equation

$$26 \qquad \Delta G^{\circ}_{binding} = -kT \ln\left(\frac{p_{bound}}{p_{unbound}}\right), \tag{1}$$

27 where p_{bound} is the equilibrium probability of the complex in 28 the bound state, p_{unbound} is the equilibrium probability of the 29 unbound state of the complex with no peptide-hapten con-30 tacts.

31 Trajectory Analysis. Trajectories were visualized using Visual Molecular Dynamics (VMD v 1.9.8). Analysis tools 32 33 provided with GROMACS were used for estimating the radial 34 distribution function of water (RDF) around the peptide and the hapten, root-mean-squared fluctuation (RMSF), the H-35 bond profile, and for generating the contact maps.⁴¹ DSSP 36 analysis provided with MDtraj (v 1.7.2.) was used for second-37 ary structure analysis.49 38

92

93

94

95

96

97

109

110

111

39 The binding free energy of the individual macrostate was 98 40 further decomposed as a cumulative sum of three main ener-99 41 getic components: non-bonded energetic contributions, solva- 100 42 tion penalty, and conformational entropy. The non-bonded 101energetic contributions and solvation penalties of the peptide- 102 43 hapten complex (associated with solvating the polar and non-103 44 polar groups within the bound complex) were estimated using 104 45 g mmpbsa, a GROMACS-based plugin for high-throughput 105 46 47 MM/PBSA calculation.⁵⁰ Penalties involved in conformation- 106 48 al entropy were obtained by quasi-harmonic mode analysis, 107 49 using the functionalities provided in GROMACS. 108

50 • RESULTS AND DISCUSSION

51 Site-Specific 6× His Tag Influence NFO4 Folding. 112

52 BEMD was applied to fully explore the conformational space

53 of NFO4 to identify the conformations relevant to hapten 113

54 binding. For this purpose, all simulations were initiated from $\frac{114}{114}$

55 fully extended conformations of the peptides, and were car-

57 ure S1). To assure equilibrium, the simulations were extended 58 for an additional 100 ns. Once the peptide conformations were 59 sampled, the lowest energy conformations were identified 60 using MSM. Figure 2 represents the free energy surface (FES) 61 and the population distribution of each conformation follow 62 ing the projection of the two major IC components (represent 63 ed by IC1 and IC2). Within the FES, the four low free energy basins were designated as '1', '2', '3', and '4' and the equilib 64 65 rium distribution of each of these peptide conformations were 66 also provided. The peptide structure that was representative of 67 each basin is shown in the adjacent panel. Table 1 summarizes structural and energetic attributes of these four 68 the 69 macrostates of the peptides. Figures S3-S6 provide infor mation on the RMSF, secondary structure preference, contact 70 71 map, and the average number of intramolecular H bonds for 72 each state of the peptides, respectively. In general, discussion 73 will be limited to the peptide conformation represented by 74 basin '4', the dominant conformation in each of the peptide 75 system.

76 (a) NFO4: Structural characteristics of the most dominant 77 conformation in NFO4 (NFO4 '4', 99%, Figure 2a) indicated 78 a bent structure, with the constituent residues being rigidly 79 held (RMSF < 0.2 nm) (Figure S3a). In the bent structure, 80 residues 1 4 formed the trailing segment, residues 5 8 formed 81 the loop structure and residues 9 12 formed the leading seg 82 ment of the peptide. At least 4 of the 12 residues (Met3, Asn4, 83 Lys8 and Cys9) within the bent were involved in intramolecu 84 lar H bonding (Figure S4a). Secondary structure analysis of 85 the sampled NFO4 conformations indicated a general prefer 86 ence towards β turn (33%) and coiled like (67%) structures. 87 The proximity of the neighboring residues and the potential 88 intramolecular contacts might be responsible for the reduced 89 flexibility of the residues (Figure S5a). Overall, the folded 90 conformation of the peptide was relatively stable compared to 91 its unfolded conformation by $4.5 k_{\rm b}T$.

(b) NterNFO4: Of the 18 residues in NterNFO4, residues 1 6 represent, the $6 \times$ His tag appended to the N terminus, and residues 7 18 represent the NFO4 peptide. Within the domi nant peptide population (NterNFO4 '4', 99%, Figure 2b), residues 1 5 were flexible (RMSF > 0.3 nm) and unstructured while the remaining residues were involved in a compact bent structure formation that was relatively less flexible than the N terminus end of the peptide (Figure S3b and S4b.). In the bent structure, the trailing and the leading segments in the β turn were composed of residues 6-10 and residues 13-18 respectively. The residues 11-12 formed the peptide's loop. The proximity of several residues (distance < 0.5 nm, Figure S5b), especially within the bent structure resulted in at least 9 intramolecular H-bonds between Val7, Tyr8, Met9, Asn10, Tyr13, Tyr14, Lys15 and Cys10. The peptide's structures were additionally stabilized by H-bonding with the solvent (Figure S6b). Overall, the introduction of $6 \times$ His-tags at the N-termini of NFO4 preserved the native peptide fold and did not shift its overall stability ($<0.5 \text{ k}_{b}\text{T}$), despite altering some of the intrinsic contact network within the native NFO4 peptide.

(c) CterNFO4: CterNFO4 represented C-termini modified NFO4 with residues 1-12 representing the native NFO4 pep-

1 tide and residues 13-18 representing the $6 \times$ His-tag appended 14 to the C-terminus of the peptide. Unlike NFO4 and NterN-2 15 FO4, CterNFO4 favored a broader distribution of the peptide 3 16 4 conformations (Figure 2c)). With the C-terminus tag in place 17 5 many of the peptide segments were too far apart (> 0.5 nm) to 18 make intra-molecular contacts. The drop in ordered structure 19 6 7 was an additional indicator of the disruptive influence of 6× 8 His-tags on the intrinsic contact network of NFO4. The dis-20 9 ruption in the contact network also explains the peptides' 21 enhanced flexibility (> 0.6 nm) (Figure S3c). Interestingly, 10 22 the dominant conformation of CterNFO4 was relatively more 11 23 12 stable (> 1.0 k_bT) than NFO4 and NterNFO4 and favored 24 13 ordered structure, with 17% β -turn and 83% coil. This is an 25

interesting dichotomy, a broader distribution of states but with the dominant state being more stable. Overall, the addition of $6 \times$ His tag at the C termini of NFO4 reduced the secondary structure content, disrupted the native contacts and increased the residue flexibility, all of which compromised the native fold.

The combined assessment of the above results suggests that the site specific placement of the affinity tags can affect the native fold of the NFO4 peptide. While $6 \times$ His tags at the N terminus of NFO4 showed a general tendency to preserve the native peptide fold, the tags at the C terminus of the NFO4 had a disruptive influence on the peptide's native fold.

Table 1: Structural and Energetic Preferences of the Folded Peptides

Recognition Molecule	Number of Residues ¹	Pref. Macro. ²	Basin Depth ³	Secondary Structure Preference ⁴	PrPr. H-bonds ⁵	PrSolv. H-bonds ⁶
NFO4	12	'4'	4.5 k _b T	33% β-turn, 67 % Coil	4	38
NterNFO4	18	'4'	4.7 k _b T	44% β-turn, 56 % Coil	9	58
CterNFO4	18	'4'	5.4 k _b T	17% β-turn, 83% Coil	8	60

¹ Number of Residues refers to the total number of amino acids constituting the peptide.

² Pref. Macro. refers to the preferred macrostate which was most populated at equilibrium as determined by the MSM analysis..

³ Basin depths refer to the free energy associated with the most preferred model, as determined from the MSM analysis.

⁴ Secondary structure preference refers to the average secondary structure content of the most preferred model.

⁵ Pr.-Pr. H-bonds refer to the average number of peptide-peptide hydrogen bonds (H-bonds) within the most preferred macrostate.

⁶ Pr.-Solv. H-bonds refer to the average number of hydrogen bonds (H-bonds) between the peptide and the solvent within the most preferred macrostate.



Figure 2. Illustration of the free energy landscape, the representative peptide configurations for each macrostate, and the equilibrium distributions associated with (A) NFO4, (B) NterNFO4, and (C) CterNFO4 as identified from MSM analysis of the trajectories obtained in BEMD accelerated sampling simulations. IC1 (x-axis) and IC2 (y-axis) represent the dimensional projection along the

top two components of the IC analysis. Peptide conformations corresponding to the four low free energy basins were identified as '1', '2', '3', and '4', and its equilibrium population distribution were designated as ' π '. $\pi < 1\%$ were considered statistically insig nificant peptide conformations. The scale bar indicates the free energy surface (FES, units of k_b T) associated with the peptide fold ing, with blue (higher number) indicating the least favorable configuration and red (lower number) indicating the most favorable configuration. Peptide segment similar toNFO4 were represented as *ribbons* (colored in grey), while the His residue side chains were displayed in *licorice* (colored by atom name).

62

Site-Specific 6× His Tag Influenced the Hapten Bind- 63 8 9 ing Properties of the NFO4 Peptide. Binding affinities 64 10 and the equilibrium population distributions of the peptides 65 11 bound to OTA and OTB were predicted using BEMD and 66 MSM analysis. The FES associated with ochratoxin binding to 67 12 the peptides are provided in Figure S7. The regions that are 68 13 14 more intensely blue within the FES correspond to the less en- 69 15 ergetically favorable peptide-hapten configurations while the 70 16 regions that were more intensely red correspond to the more 71 17 energetically favorable configurations. The macrostates identi-72 18 fied for each peptide-hapten configurations were represented 73 19 by an individually colored marker. The placement of each 74 20 colored marker within the FES represents the microstates from 75 21 where the specific peptide-hapten configurations were sam- 76 22 pled. For this study, peptides were considered bound to the 77 23 haptens when the haptens were within 0.5 nm of the peptide. 78 24 Based on the peptide-hapten contacts, at least two macrostates 79 25 representing the bound and unbound states of the peptide were 80 26 identified. The most populated bound and unbound states are 81 27 shown in Figure 3. Table 2 summarizes the structural and 82 thermodynamic preferences for the untagged and tagged 83 28 29 NFO4 when bound to OTA and OTB. For each of the discrete 84 macrostates identified from the FES maps, residue flexibility 85 30 31 (Figures S8 and S9), water density surrounding the complex 86 32 (Figures S10-S15), secondary structure preference (Figures 87 S16 and S17), peptide-hapten contact maps (Figures S18 and 88 33 34 S19), H-bonding profiles (Figures S20 and S21), energetic 89 contribution to peptide-hapten binding (Tables S1-S2), and 90 35 residue-level energetic breakup (Tables S3-S5) were evaluated 91 36 to better understand the molecular-level processes involved in 92 37 peptide binding to ochratoxins. As the binding energy esti- 93 38 mates determined in this study were based on the definition 94 39 provided in equation (1), most of the ensuing discussion is 95 40 41 limited to the structural and energetic characteristics of the 96 42 most dominant states.

(a) Site-Specific 6× His Tag Significantly Influenced the 98 43 44 Native Peptide's Interaction with OTA. Figures 3a, 3c and 99 3e represent the most dominant equilibrium distributions of100 45 46 NFO4, NterNFO4 and CterNFO4 when bound and unbound to101 47 OTA, respectively. Table S1 provides an overview of the en-102 48 ergetic breakdown for the peptide interactions with OTA, and103 49 Tables S3a, S4a, and S5a provide a residue-level energetic104 breakdown for peptide-OTA interaction. The combined data105 50 51 suggests that the presence of the tag significantly influenced106 52 the predicted in-solution binding affinities of the peptide to107 53 OTA. 108 Previous studies using experimental and simulation tech-109 54

⁵⁴ Previous studies using experimental and simulation tech-¹¹⁰
⁵⁵ niques had established that NFO4 binds with high affinity and ¹¹⁰
⁵⁶ selectivity to OTA. ³⁴ In line with these results, NFO4 shows111
⁵⁷ higher affinity and selectivity to OTA. Three macrostates were112
⁵⁸ identified in NFO4-OTA system (Figure S7a): NFO4-OTA-'1'113
⁵⁹ represented the unbound state of the peptide and NFO4-OTA-114
⁶⁰ and NFO4-OTA-'2' represented the bound state of the115
⁶¹ peptide-hapten complex. The bound configurations represent-116

ed by the NFO4-OTA-'2' were energetically favored (Δ PMF \approx -10.4 kcal/mol) and dominated (> 99%) the sampled configurations. The unbound NFO4 peptides were relatively ordered (33% β-turn, 67% Coil) (Figure S16a), with a dense water layer surrounding the peptide and OTA (Figure S11). However, post-binding to OTA, both the NFO4 and the OTA underwent desolvation (Figure S12) and broke most of its intrinsic internal contacts (Figures S16, S18, and S19) to form new contacts with OTA. Contact map indicates that at least seven of the 12 residues (2-4, 6-8, and 10) were within 0.35 nm of OTA, among which at least three of the contacts were likely mediated via H-bonding. The disruption in the peptide's internal contacts resulted in a loss of its secondary structure. The loss in internal contacts, however, did not increase the average fluctuations of the individual residues but rather decreased them, probably due to the tight interactions between the residues in the peptide and OTA. Consistent with our previous reports, the higher affinity and selectivity of NFO4 to OTA stems from the lower solvation penalty associated with the NFO4-OTA complex.³²

The addition of 6x His-tags at the N-terminus generally preserved the affinity and selectivity of NFO4 to OTA (Table 2). Two distinct macrostates (Figure. S7c) were identified:-NterNFO4-OTA-'0' (16% of the sampled conformations) and NterNFO4-OTA-'1' (84% of the sampled conformations) which represented the unbound and the bound states of the peptide respectively (Figure 3c). Among these two states, the bound state was energetically favored ($\Delta PMF \approx -9.8$ kcal/mol). The addition of affinity tags at the N-terminus of NFO4 reduced the number of water molecules surrounding the peptide. But the NterNFO4 peptide interaction with OTA, induced more extensive desolvation in the hapten than the peptide (Figure S11). The peptide interaction with OTA disrupted many of the native contacts within the peptide (Figure. S18.b) to form new intra-molecular and peptide-hapten contacts (Figure. S20.b) all of which disrupted the peptide's ordered structure. Within the bound complex, OTA was found to be in contact with 12 of the 18 residues (residue numbers 2-6, 8, 11, 13-17) of which at least four of the contacts were mediated by H-bond interaction (Figure S19b). The intra-molecular and the peptide-hapten contacts were additionally responsible for the lowered residue fluctuations within the bound complex. In fact, the extensive increase in the non-bonded interactions with OTA was accompanied by an increase in the solvation penalty of NterNFO4-OTA complex. Despite the increase in solvation penalty for NterNFO4-OTA complex, the favorable interaction of histidine residues at the N-terminus with OTA (Table S4a), was the primary factor responsible for preserving the binding preference of the native peptide.

The addition of 6x His-tags at the C-terminus radically altered the binding behavior of the peptide, by effectively eliminating the native peptide's affinity and selectivity to OTA. Unlike other peptide systems, the peptide configurations corresponding to the unbound state (75%) were sampled more than the bound state (Figure 3e). The remaining configurations

1 represented the bound configurations, with CterNFO4-OTA- 39 2 '2' representing the dominant configuration (Figure S7e). Both 40 the unbound and bound states were sampled from the redder 41 3 4 regions of the FES, with the bound state being slightly more 42 5 energetically favored than the unbound state ($\Delta PMF \approx -2.0$ 43 6 kcal/mol). When the Cter NFO4 were bound to OTA, the pep- 44 7 tide and OTA underwent desolvation with the hapten desolva- 45 8 tion being more prominent than the peptide desolvation. Also, 46 9 the OTA interaction did not induce any significant secondary 47 structural shift or significantly alter the native internal network 48 10 within the peptide. At least one of the contacts between the 49 11 CterNFO4 (residue numbers 5 and 6) and OTA was likely 50 12 13 mediated via H-bond interaction. However, the contribution of 51 non-bonded interaction within the CterNFO4-OTA system to 52 14 the overall binding energetics was the least among all the OTA 15 53 systems (Table S1). The residue-level assessment of the bind-16 54 ing energetics, further indicated that the 6x His-tag in the 17 55 CterNFO4 barely interacted with OTA. The positioning of the 18 56 19 6x His-tags significantly influenced the non-bonded contribu-57 20 tion to the overall binding energetics (compare Tables S4a and 58 21 Table S5a). 59 22 (b) Site-Specific 6× His Tag Did Not Significantly Influence 60 23 the Native Peptide's Interaction with OTB: Figures 3b, 3d 61

24 and 3f represent the most dominant equilibrium distributions 62 25 of NFO4, NterNFO4 and CterNFO4 when bound and unbound 63 26 to OTB, respectively. In general, diverse and broader distribu- 64 tions of conformational states were identified in the peptide 65 27 interaction with-OTB system (Figures 3b, 3d, and 3f) as op- 66 28 29 posed to the peptide's interaction with OTA. Table S2 provide 67 an overview of the energetic breakdown for the peptide inter- $\frac{1}{68}$ 30 31 actions with OTB, and Tables S3b, S4b, and S5b provide a 6932 residue-level energetic breakdown for peptide-OTB interac- 70 33 tion. As evident, the combined data suggests that the tagging 71 process did not significantly affect the predicted in-solution 72 34 35 affinities to OTB. 73

36 Similar to OTA, the binding interaction of NFO4 with OTB 74
37 also involved desolvation of the peptide and OTB (Figures 75
38 S13b-S15b), and disruption in the peptide's inherent contacts

(Figures S17, S19, and S21) to form new contacts with OTB. The loss in intrinsic contacts compromised the 2° structural integrity of the NFO4 peptide, and was expected to increase the residue fluctuation. However, the increase in residue fluc tuation was primarily limited to the terminals of the peptide, with the other segments of the NFO4 peptide remaining rela tively intact. A probable reason for the lack in residue fluctua tion could be due to the new contacts formed between the res idues 3 9 in NFO4 and OTB. Of these new contacts, at least three contacts were mediated via H bonding. As it could be seen from the Tables S1 and S2, NFO4 shows a higher *insolution* affinity and selectivity to OTA than OTB, despite the non bonded contributions for both systems being similar.

The addition of 6x His tags at the N terminal end of NFO4 (Figure 3d) also involved the formation of new peptide hapten contacts via the desolvation of the peptide and the hapten, and the disruption in the native contact network (Figures S14 and S19b). Although the disruption in the native contact network induced a loss in ordered secondary structure, the residue flex ibility in the bound state was lower than the unbound state, probably due to the peptide hapten contacts and the more ex tensive intra molecular contact network (Figures S17b, S19b, and S21b). The desolvation process however, promoted the non bonded interactions with OTB while increasing the solva tion penalty associated with the NterNFO4 OTB system.

In case of the CterNFO4 OTB system (Figure 3f), the desolvation process was accompanied by less extensive pep tide hapten contacts (Figures S14, S16c, and S18c). Further more, the relatively minor change in the internal contacts did not change the secondary structure of the peptide or alter the flexibility of the residues. In fact, the contribution of non bonded interaction and the solvation penalty within the CterNFO4 OTB system to the overall binding energetics was the least among all the OTB systems. (Tables S1 and S2). The residue level assessment of the binding energetics, further indicated that the 6x His tag in the CterNFO4 barely interact ed with OTB.



Figure 3. The representative states for unbound and bound peptide, equilibrium probabilities of these states (π), and free energies of binding (designated as ΔG° (kJ/mol)) (a) NFO-OTA, (b) NFO4-OTB, (c) NterNFO4-OTA, (d) NterNFO4-OTB, (e) CterNFO4-OTA, and (f) CterNFO4-OTB, identified from MSM analysis of the trajectories generated in BEMD enhanced sampling. ' π ' represent the distribution of the most populated bound or unbound macrostates (the sums of probabilities of all bound, p_{bound} , and all un-

bound states, $p_{unbound}$, were used for binding free energy calculations in Equation 1). Peptides were depicted with the translucent *molecular surface* (in grey) and the ochratoxins were shown in *licorice* (colored by the default atom color except for carbon which was colored as magenta (for OTA) and orange (OTB)).

Recognition Molecule	Complex Dist. ¹ (%)	Str. Shifs ²	Cont. Freq. ³	RMSF Shift ⁴	ΔG° (kJ/mol) ⁵
NFO4-OTA	> 99	Yes	7/12	Decreased	-42.34 (6.2)
NFO4-OTB	> 99	Yes	6/12	Increased	-22.76 (5.52)
NterNFO4-OTA	~84	Yes	12/18	Decreased	-36.92 (7.92)
NterNFO4-OTB	~94	Yes	6/18	Decreased	-19.07 (1.69)
CterNFO4-OTA	~25	No	2/18	Decreased	-9.16 (2.42)
CterNFO4-OTB	~93	Yes	5/18	Decreased	-17.74 (2.34)

Table 2. Structural and thermodynamic characteristics of the peptide-hapten complex

¹ Complex Dist. refers to the overall bound fraction (%) of the peptide-hapten complex.

 2 Str. Shifts refers to the secondary structure shifts within the Pref. State relative to the unbound macrostate.

³ Cont. Freq. refers to the average number of residues within 0.35 nm of the hapten within the Pref. Macro.

⁴ RMSF Shift refers to the structural fluctuation within the Pref. State relative to the unbound macrostate.

 $^{5}\Delta G^{\circ}$ refers to the average equilibrium binding free energy. The values in the parenthesis provides the statistical error associated with ΔG° estimates by block averaging the data over 20 ns.

47 13 Validation of Results and Implications for Computer- 48 Aided Design of Sensing Elements. Table 3 compares the 49 14 simulated in-solution affinity and selectivity with the immobi- 50 15 16 lized state binding properties of the tagged and untagged pep- 51 17 tides Previous studies using experimental and simulation tech- 52 niques have established that NFO4 binds with high affinity 53 18 and selectivity to OTA.³⁴ In line with these results, NFO4 54 19 showed higher affinity and selectivity to OTA than OTB. 55 20 21 Though at least three different structural variants of ochratoxin 56 22 are known to naturally occur, only OTA and OTB are found 57 prevalently in food, feed and beverages.^{3, 4} Of these two ochra- 58 23 toxins, OTA is more toxic³ and has been known to bio-59 24 accumulate in the consumer, necessitating strict regulatory 60 oversight.^{3, 4, 51, 52} In this regard, biosensing platforms based 61 25 27 on immobilized 6x His-tag conjugated NFO4 provide a cost- 62 28 effective alternative to antibody based bioassays for the im- 63 29 munocapture of OTA in the detection and monitoring of OTA 64 30 over OTB. But, a main challenge in the design of biosensing 65 elements for OTA pertains to the selection and location of 66 31 32 affinity tags, so as not to compromise the inherent activity of 67 33 the molecular recognition entity. The choice of tag is particu- 68 34 larly more challenging when the tags and sensing elements are 69 35 of comparable size and chemistry. 70 71 In the current study, only the *in-solution* affinities of the 6×72 36 His tagged NFO4 were assessed. Our study indicated that the 73 37

addition of affinity tags at either ends of the NFO4 favored 74 38 39 solvation more than the untagged peptide Also, OTB is more 75 hydrophilic than OTA and the improved solvation from the 76 40 41 affinity tag placement at either ends of the peptide minimized 77 42 the peptide interaction with OTB (Tables S1 and S2). The site 78 43 of 6x His-tag placement influenced the native fold, with the 79 44 tags at the Nter preserving the native fold and those at the Cter 80 45 altering the fold. The alteration in the fold further affected the 81

46 NFO4's interactions with OTA, with the NterNFO4 retaining 82

the general affinity and selectivity to OTA while the CterN-FO4 lacking any ochratoxin-specific recognition capabilities (by comparing the $\Delta G^{\circ}_{binding}$ in Table 2). Infact, among the tagged peptides, the affinities of NFO4 and NterNFO4 towards OTA were not significantly different but the affinity of CterNFO4 was ~4x lower than that of the native peptides. In a recent experimental study, involving the immobilization of 6× His tagged NFO4, Soleri et al. had demonstrated that the dense layers of immobilized NFO4 peptides were significantly better(2x) at OTA recognition when oriented on Zn^{++} -laden chitosan foams via N-terminus modification as opposed to Cterminus modification.^{7, 8} While it is certainly encouraging to observe that the in-solution predictions from simulations generally followed the experimental data on immobilized system, a direct comparison of these results is challenging. For example, experimental studies on the immobilized hexamer (SNLHPK) on a wide range of substrates indicate that the peptide's affinity to OTA on solid phase ($K_D \sim 0.01 \ \mu M - 1.00$ μ M) were drastically different from those *in-solution* (K_D ~ 29 μ M).^{53, 54} These observed differences in affinity would suggest that in addition to the molecular recognition entity, the surface may also have an active role in the hapten recognition, even though the exact mechanisms are currently unknown. Future studies could also explore and validate strategies that would minimize the direct interaction between the peptide and the tags by introducing more flexible spacers like polyethylene glycol (PEG) so as to assess if such modification could recover the native binding properties of the NFO4 irrespective of the site of modification. Nevertheless, the results presented in this study are in line with those and other studies that demonstrated the negative impact of 6x His tags on the binding properties of peptides especially when the tags were placed at the C-terminus of peptides/proteins.^{18, 20, 22, 55} Based on these rationale, the current simulations were considered representative of the real-world binding interactions within an affinity tagged NFO4.

Table 3. Binding properties	(mean ±95% C	L.) of sy	nthetic po	eptides to	ochratoxins	in a win	e-like solution	pH.
-----------------------------	--------------	-----------	------------	------------	-------------	----------	-----------------	-----

Recognition Molecule	ΔG° _{OTA-PRE} (kJ/mol)	ΔG° _{OTB-PRE} (kJ/mol)	$K_{D^{-}OTA}(\mu M)$ Pred. ¹	Selectivity Pred. ²	Comments
NFO4	-42.3 (6.2)	-22.8 (5.5)	0.038	~1.9	Immobilized NFO4 - $3x$ more selective to OTA than OTB ³
NterNFO4	-36.9 (7.9)	-19.1 (1.7)	0.338	~1.9	Immobilized NterNFO4 were 2.5x more effective at OTA
CterNFO4	-9.2 (2.4)	-17.7 (2.3)	24353	~0.5	detection than CterNFO4 ⁴

¹ The dissociation constant (K_D) for ochratoxins was estimated from the average ΔG° estimate using the equation $K_D = \frac{1}{\exp(-\frac{\Delta G}{2\pi})}$ where R

 $= 0.00831 \text{ kJ mol}^{-1} \text{ K}^{-1} \text{ and } \text{ T} = 298 \text{ K}.$

² Selectivity-Pred was based on the relative ratios of $\Delta G^{\circ}_{\text{OTA-PRE}}$ to $\Delta G^{\circ}_{\text{OTB-PRE}}$

³ See Ref 6 Estimate using HPLC-FLD.

⁴ See Ref 8 Estimate using fluorescence.

CONCLUSIONS

51 in the current study can be extended to virtually any peptide-52 based system so as to assess the impact of tags or site-specific tags on the peptide's intrinsic activity.

ASSOCIATED CONTENT

The following supplementary materials are available online. Figure S1: Convergence plot in peptide folding simulation. Figure S2: Implied time scale plot of the folded peptide for (a) NFO4, (b) NterNFO4, and (c) CterNFO4, Figure S3: RMSF of the folded peptide. Figure S4: Secondary structural preferences of the folded peptides, Figure S5: Contact map of the folded peptides, Figure S6: Average H-bond profile in folded peptides. Figure S7: FES map of the peptide systems involved in ochratoxin binding. Figure S8-S9: RMSF of the peptide bound to OTA and OTB. Figure S10-S12: RDF of the peptides and OTA. Figure S13-S15: RDF of the peptides and OTB. Figure S16-S17: Secondary structural preferences of the peptides when bound to OTA and OTB. Figure **S18-S19**: Contact map of the residues within a peptide-hapten complex. Figure S20-S21: H-bond profile within the peptidehapten complex. Tables S1-S2: Breakdown of the energetic components contributing to the overall binding energetics. Tables S3-S5: Residue-level contribution the overall binding energetics of each peptide system.

AUTHOR INFORMATION

Corresponding Author

*Anthony Guiseppi-Elie.

E-mail: guiseppi@tamu.edu. Phone: +1(979) 458 1239. Fax: +1(979) 458 8219.

ORCID

Aby A. Thyparambil: 0000-0003-1760-1454 Tigran M. Abramyan: 0000-0002-7224-6072 Ingrid Bazin: 0000-0002-8293-5592 Anthony Guiseppi-Elie: 0000-0003-3218-9285

Author Contributions

AG-E and AAT designed the study, AAT and TMA executed the study, TMA and IB contributed insights to the study, and AAT and AG-E wrote the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors acknowledge the support of the computing and storage facility of Texas A & M University's High Performance Research Computing Resources.

53 8 Advanced sampling and MSM were used to assess the im-9 pact of site-specific 6x His tags on the in-solution binding 54 properties of NFO4 to ochratoxins in a wine-like matrix envi- 55 10 ronment. The in-silico framework outlined in the current study 56 11 facilitated structural and thermodynamic characterization of 57 12 13 the sensing elements with tags of comparable size and chemis- 58 14 try in detail that were otherwise not obtainable using routine 59 15 experimental biophysics techniques or regular MD simula- 60 tions. Results showed that the 6x His tag placement influenced 61 16 the folding and binding behavior of the peptides. While the $6x^{62}$ 17 63 His tags placed on the N-terminus of the NFO4 generally pre-18 64 served the native fold of the peptide the tags altered the intrin-19 65 sic contact network and actively interacted with OTA. The 20 66 21 overall affinity and selectivity of NterNFO4 was driven by the 67 combined influence of the non-bonded interactions involving 22 23 OTA with the His residues and the higher solvation penalty of 69 peptide-OTA complex. In contrast, the 6x His tags placed on 24 70 25 the C-terminus of the NFO4 improved the stability of the native peptide by altering the native fold and the intrinsic contact 26 72 27 network within the peptide. These His residues did not con-28 tribute significantly to the overall non-bonded energetics, and 73 29 also introduced a higher solvation penalty for solvating the peptide-OTA complex as opposed to peptide-OTB complex. 74 30 The non-bonded interactions and the solvation penalties asso- 75 31 32 ciated with the NFO4 interaction with OTB were relatively 76 unaffected by the site-specific placement of the tags. Conse-77 33 quently, the NterNFO4 preserved the generally affinity and 34 78 35 selectivity of NFO4 while the CterNFO4 eliminated the OTAspecific binding capabilities of NFO4. However, the binding 79 36 mechanisms of the NFO4 and NterNFO4 to OTA were differ- 80 37 ent. The OTA-specific binding behavior of NFO4 was entirely 81 38 driven by the lower solvation penalty of the NFO4-OTA, as 39 opposed the OTA-specific binding by NterNFO4 that was 83 40 driven by the interplay of non-bonded interactions and solva-41 84 tion penalty associated with the peptide-OTA complex. 42 85 86 43 While further work is certainly needed before the in-silico

87 44 platforms presented in the current study could be reliably used

to engineer peptides or direct peptide modifications with de- 88 45 46 sired sensing efficacies to a target hapten, the mechanistic

- insights gained from the approach presented in this work can 89 47
- facilitate the selection and placement of affinity tags or the 48
- 90 49 modification of the molecular recognition entity at least in a
- solution environment. In a broader sense, the methods outlined $\frac{1}{92}$ 50

REFERENCES

1

2

3

4

5

6

7

8

- Singh, M. K.; Srivastava, S.; Raghava, G. P.; Varshney, G. C., 77 HaptenDB: A Comprehensive Database of Haptens, Carrier Proteins 78 and Anti-hapten Antibodies. *Bioinformatics* 2006, 22, (2), 253-5. 79
- Erkes, D. A.; Selvan, S. R., Hapten-Induced Contact 80 Hypersensitivity, Autoimmune Reactions, and Tumor Regression: 81 Plausibility of Mediating Antitumor Immunity. *J Immunol Res* 2014, 82 2014, 175265.
- 9 3. Malir, F.; Ostry, V.; Pfohl-Leszkowicz, A.; Malir, J.; Toman, J., 84
 10 Ochratoxin A: 50 Years of Research. *Toxins (Basel)* 2016, 8, (7), 85
 11 191.
- Heussner, A. H.; Bingle, L. E., Comparative Ochratoxin Toxicity: A 87
 Review of the Available Data. *Toxins (Basel)* 2015, 7, (10), 4253-82.
- 145.Ha, T. H., Recent Advances for the Detection of Ochratoxin A.15Toxins (Basel) 2015, 7, (12), 5276-300.90
- Bazin, I.; Tria, S. A.; Hayat, A.; Marty, J. L., New Biorecognition 91
 Molecules in Biosensors for the Detection of Toxins. *Biosens* 92 *Bioelectron* 2017, 87, 285-298.
- Tria, S. A.; Lopez-Ferber, D.; Gonzalez, C.; Bazin, I.; Guiseppi-Elie, 94
 A., Microfabricated Biosensor for the Simultaneous Amperometric 95
 and Luminescence Detection and Monitoring of Ochratoxin A.
 Biosens Bioelectron 2016, 79, 835-42.
 Soleri, R.; Demey, H.; Tria, S. A.; Guiseppi-Elie, A.; Hassine, A. I.; 00
- Soleri, R.; Demey, H.; Tria, S. A.; Guiseppi-Elie, A.; Hassine, A. I.; 9'8
 Gonzalez, C.; Bazin, I., Peptide Conjugated Chitosan Foam as a 99
 Novel Approach for Capture-Purification and Rapid Detection of 100
 Hapten--Example of Ochratoxin A. *Biosens Bioelectron* 2015, 67,101
 634-41.
- Goldman, E. R.; Cohill, T. J.; Patterson, C. H.; Anderson, G. P.;103
 Kusterbeck, A. W.; Mauro, J. M., Detection of 2,4,6-Trinitrotoluene104
 in Environmental Samples Using a Homogeneous105
 Fluoroimmunoassay. *Environmental Science & Technology* 2003,106
 37, (20), 4733-4736.
- 31 37, (20), 4735-4736.
 32 10. Bade, S.; Rockendorf, N.; Franek, M.; Gorris, H. H.; Lindner, B.;108
 34 Olivier, V.; Schaper, K. J.; Frey, A., Biolabeling with 2,4-109
 35 dichlorophenoxyacetic acid derivatives: the 2,4-D tag. *Anal Chem*110
 36 2009, 81, (23), 9695-702.
- Grieshaber, D.; MacKenzie, R.; Voros, J.; Reimhult, E.,111
 Electrochemical Biosensors Sensor Principles and Architectures.113
 Sensors (Basel) 2008, 8, (3), 1400-1458.
- 40 12. Madrid, R. E.; Chehin, R.; Chen, T.-H.; Guiseppi-Elie, A.,115
 41 Biosensors and Nanobiosensors. In *Further Understanding of the*,116
 42 *Human Machine: The Road to Bioengineering*, Valentinuzzi, M. E.,117
 43 Ed. World Scientific Publishing Co.: Singapore, 2017.
- Arnau, J.; Lauritzen, C.; Petersen, G. E.; Pedersen, J., Reprint of: 119
 Current Strategies for the Use of Affinity Tags and Tag Removal for 120
 the Purification of Recombinant Proteins. *Protein Expr Purif* 2011. 121
- 14. Kimple, M. E.; Brill, A. L.; Pasker, R. L., Overview of Affinity Tags 121
 for Protein Purification. In *Current Protocols in Protein Science*, 123
 John Wiley & Sons, Inc.: 2001.
- Bornhorst, J. A.; Falke, J. J., Purification of Proteins using125
 Polyhistidine Affinity Tags. *Methods in Enzymology* 2000, 326, 245-126
 254.
 Kang, E.; Park, J. W.; McClellan, S. J.; Kim, J. M.; Holland, D. P.;128
- 16. Kang, E.; Park, J. W.; McClellan, S. J.; Kim, J. M.; Holland, D. P.; 128
 Lee, G. U.; Franses, E. I.; Park, K.; Thompson, D. H., Specific129
 Adsorption of Histidine-Tagged Proteins on Silica Surfaces130
 Modified with Ni2+/NTA-Derivatized Poly(ethylene glycol).131
 Langmuir 2007, 23, (11), 6281-8.
- Samanta, D.; Sarkar, A., Immobilization of Bio-Macromolecules on 133
 Self-Assembled Monolayers: Methods and Sensor Applications. 134 *Chem Soc Rev* 2011, 40, (5), 2567-92.
 135
- 61 18. Goel, A.; Colcher, D.; Koo, J. S.; Booth, B. J.; Pavlinkova, G.;136
 62 Batra, S. K., Relative Position of the Hexahistidine Tag Effects137
 63 Binding Properties of a Tumor-Associated Single-Chain Fv138
 64 Construct. *Biochim Biophys Acta* 2000, 1523, (1), 13-20.
 65 10
- Chant, A.; Kraemer-Pecore, C. M.; Watkin, R.; Kneale, G. G. 140
 Attachment of a Histidine Tag to the Minimal Zinc Finger Protein of 141
 the Aspergillus nidulans Gene Regulatory Protein AreA Causes a142
 Conformational Change at the DNA-Binding Site. Protein Expr143 *Purif* 2005, 39, (2), 152-9.
- 7020.Carson, M.; Johnson, D. H.; McDonald, H.; Brouillette, C.; Delucas, 14571L. J., His-Tag Impact on Structure. Acta Crystallogr D Biol 14672Crystallogr 2007, 63, (Pt 3), 295-301.
- Panek, A.; Pietrow, O.; Filipkowski, P.; Synowiecki, J., Effects of 148
 the Polyhistidine Tag on Kinetics and Other Properties of Trehalose 149

Synthase from *Deinococcus geothermalis*. Acta Biochim Pol **2013**, 60, (2), 163 6.

- Majorek, K. A.; Kuhn, M. L.; Chruszcz, M.; Anderson, W. F.; Minor, W., Double Trouble-Buffer Selection and His-tag Presence may be Responsible for Nonreproducibility of Biomedical Experiments. *Protein Sci* 2014, 23, (10), 1359-68.
- Wang, X.; Campoli, M.; Ko, E.; Luo, W.; Ferrone, S., Enhancement of ScFv Fragment Reactivity with Target Antigens in Binding Assays following Mixing with Anti-tag Monoclonal Antibodies. J Immunol Methods 2004, 294, (1-2), 23-35.
- Strop, P.; Liu, S. H.; Dorywalska, M.; Delaria, K.; Dushin, R. G.; Tran, T. T.; Ho, W. H.; Farias, S.; Casas, M. G.; Abdiche, Y.; Zhou, D.; Chandrasekaran, R.; Samain, C.; Loo, C.; Rossi, A.; Rickert, M.; Krimm, S.; Wong, T.; Chin, S. M.; Yu, J.; Dilley, J.; Chaparro-Riggers, J.; Filzen, G. F.; O'Donnell, C. J.; Wang, F.; Myers, J. S.; Pons, J.; Shelton, D. L.; Rajpal, A., Location Matters: Site of Conjugation Modulates Stability and Pharmacokinetics of Antibody Drug Conjugates. *Chem Biol* 2013, 20, (2), 161-7.
- Plum, A.; Jensen, L. B.; Kristensen, J. B., *In vitro* Protein Binding of Liraglutide in Human Plasma Determined by Reiterated Stepwise Equilibrium Dialysis. *J Pharm Sci* 2013, 102, (8), 2882-8.
- Waters, N. J.; Jones, R.; Williams, G.; Sohal, B., Validation of a Rapid Equilibrium Dialysis Approach for the Measurement of Plasma Protein Binding. *J Pharm Sci* 2008, 97, (10), 4586-95.
- 27. Naik, R. R.; Hagen, J. A.; Slocik, J. M.; Schmucker, A. L.; Singamaneni, S., Localized Surface Plasmon Resonance Sensing of Human Performance Biomarkers using Short Peptide Recognition Elements on Optically Active Metal Nanostructures. In Google Patents: 2017.
- Karlsson, R.; Pol, E.; Frostell, A., Comparison of Surface Plasmon Resonance Binding Curves for Characterization of Protein Interactions and Analysis of Screening Data. *Anal Biochem* 2016, 502, 53-63.
- Oliveberg, M.; Wolynes, P. G., The Experimental Survey of Protein-Folding Energy Landscapes. *Q Rev Biophys* 2005, 38, (3), 245-88.
- Lane, T. J.; Shukla, D.; Beauchamp, K. A.; Pande, V. S., To Milliseconds and Beyond: Challenges in the Simulation of Protein Folding. *Curr Opin Struct Biol* **2013**, 23, (1), 58-65.
- Boehr, D. D.; Nussinov, R.; Wright, P. E., The Role of Dynamic Conformational Ensembles in Biomolecular Recognition. *Nat Chem Biol* 2009, 5, (11), 789-96.
- Mobley, D. L.; Dill, K. A., Binding of Small-Molecule Ligands to Proteins: "What You See" is Not Always "What You Get". *Structure* 2009, 17, (4), 489-98.
- Plattner, N.; Noe, F., Protein Conformational Plasticity and Complex Ligand-Binding Kinetics Explored by Atomistic Simulations and Markov Models. *Nat Commun* 2015, 6, 7653.
- Thyparambil, A. A.; Bazin, I.; Guiseppi-Elie, A., Evaluation of Ochratoxin Recognition by Peptides Using Explicit Solvent Molecular Dynamics. *Toxins* 2017, 9, (5), 164.
- Bazin, I.; Andreotti, N.; Hassine, A. I.; De Waard, M.; Sabatier, J. M.; Gonzalez, C., Peptide binding to ochratoxin A mycotoxin: a new approach in conception of biosensors. *Biosens Bioelectron* 2013, 40, (1), 240-6.
- Cossio, P.; Trovato, A.; Pietrucci, F.; Seno, F.; Maritan, A.; Laio, A., Exploring the Universe of Protein Structures beyond the Protein Data Bank. *PLoS Comput. Biol.* **2010**, 6, (11), e1000957.
- Doerr, S.; Harvey, M. J.; Noe, F.; De Fabritiis, G., HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. *J Chem Theory Comput* 2016, 12, (4), 1845-52.
- Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual Molecular Dynamics. J Mol Graph 1996, 14, (1), 33-8, 27-8.
- Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., ZINC: A Free Tool to Discover Chemistry for Biology. J Chem Inf Model 2012, 52, (7), 1757-68.
- Koziara, K. B.; Stroet, M.; Malde, A. K.; Mark, A. E., Testing and Validation of the Automated Topology Builder (ATB) Version 2.0: Prediction of Hydration Free Enthalpies. *J Comput Aided Mol Des* 2014, 28, (3), 221-33.
- Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* 2015, 1-2, 19-25.
- 42. Wang, D.; Freitag, F.; Gattin, Z.; Haberkern, H.; Jaun, B.; Siwko, M.; Vyas, R.; van Gunsteren, W. F.; Dolenc, J., Validation of the

75 76 1 2 GROMOS 54A7 Force Field Regarding Mixed α/β Peptide 25 Molecules. *Helvetica Chimica Acta* **2012**, 95, (12), 2562-2577. 26

- 3 43. Huang, W.; Lin, Z.; van Gunsteren, W. F., Validation of the 27
 4 GROMOS 54A7 Force Field with Respect to beta-Peptide Folding. J 28
 5 Chem Theory Comput 2011, 7, (5), 1237-43. 29
- 6 44. Cossio, P.; Marinelli, F.; Laio, A.; Pietrucci, F., Optimizing the 30
 7 Performance of Bias-Exchange Metadynamics: Folding a 48-residue 31
 8 LysM Domain using a Coarse-Grained Model. *J Phys Chem B* 2010, 32
 9 114, (9), 3259-65. 33
- 10 45. Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A., Substrate Binding 34
 11 Mechanism of HIV-1 Protease from Explicit-Solvent Atomistic 35
 12 Simulations. J Am Chem Soc 2009, 131, (33), 11811-8.
- 46. Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; 37
 Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; 38
 Parrinello, M., PLUMED: A Portable Plugin for Free-Energy 39
 Calculations with Molecular Dynamics. *Computer Physics* 40 *Communications* 2009, 180, (10), 1961-1972.
- 18 47. Chodera, J. D.; Noe, F., Markov State Models of Biomolecular 42
 Conformational Dynamics. *Curr Opin Struct Biol* 2014, 25, 135-44. 43
- 48. Pande, V. S.; Beauchamp, K.; Bowman, G. R., Everything You 44
 Wanted to Know about Markov State Models but were Afraid to 45
 Ask. *Methods* 2010, 52, (1), 99-105.
 46
 49. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; 47
- 49. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; 47
 24 Swails, J. M.; Hernandez, C. X.; Schwantes, C. R.; Wang, L. P.; 48
 49

Lane, T. J.; Pande, V. S., MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys J* **2015**, 109, (8), 1528-32.

- Kumari, R.; Kumar, R.; Open Source Drug Discovery, C.; Lynn, A., <u>g_mmpbsa--A</u> GROMACS Tool for High-Throughput MM-PBSA Calculations. J Chem Inf Model 2014, 54, (7), 1951-62.
- Malir, F.; Ostry, V.; Pfohl-Leszkowicz, A.; Novotna, E., Ochratoxin A: Developmental and Reproductive Toxicity-An Overview. *Birth Defects Res B Dev Reprod Toxicol* 2013, 98, (6), 493-502.
- Kőszegi, T.; Poór, M., Ochratoxin A: Molecular Interactions, Mechanisms of Toxicity and Prevention at the Molecular Level. *Toxins* 2016, 8, (4), 111.
- 53. Giraudi, G.; Ferrero, V. E.; Anfossi, L.; Baggiani, C.; Giovannoli, C.; Tozzi, C., Solid-Phase Extraction of Ochratoxin A from Wine based on a Binding Hexapeptide Prepared by Combinatorial Synthesis. J Chromatogr A 2007, 1175, (2), 174-80.
- Giovannoli, C.; Passini, C.; Volpi, G.; Di Nardo, F.; Anfossi, L.; Baggiani, C., Peptide-Based Affinity Media for Solid-Phase Extraction of Ochratoxin A from Wine Samples: Effect of the Solid Support on Binding Properties. *Talanta* 2015, 144, 496-501.
- Thielges, M. C.; Chung, J. K.; Axup, J. Y.; Fayer, M. D., Influence of Histidine Tag Attachment on Picosecond Protein Dynamics. *Biochemistry* 2011, 50, (25), 5799-805.



IC1