

In-Silico Discovery and Validation of Neuropeptide Y Binding Peptides for Sensors

Xingqing Xiao,¹ Zhifeng Kuang,² B.J. Burke,² Yaroslav Chushak,² Barry L. Farmer,² Peter A. Mirau*,² Rajesh R. Naik,² and Carol K. Hall*¹

¹ Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27695, United States

² Materials and Manufacturing Directorate and 711th Human Performance Wing, Air Force Research Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio 45433, United States

KEYWORDS: *Biorecognition Elements, Peptide-based biosensor, Computational Peptide Design, Phage Display Library, Neuropeptide Y*

ABSTRACT: Wearable sensors for human health, performance and state monitoring which have a linear response to the binding of biomarkers found in sweat, saliva or urine are of current interest for many applications. A critical part of any device is a *biological recognition element* (BRE) that is able to bind a biomarker at the surface of a sensor with a high affinity and selectivity to produce a measurable signal response. In this study, we discover and compare 12-mer peptides that bind to neuropeptide Y (NPY), a stress and human health biomarker, using independent and complimentary experimental and computational approaches. The affinities of the NPY-binding peptides discovered by both methods are equivalent and below the μM level, which makes them suitable for application in sensors. The *in-silico* design protocol for peptide-based BREs is low cost, highly efficient, and simple, suggesting its utility for discovering peptide binders to a variety of biomarker targets.

INTRODUCTION

Sensors that measure the concentrations of biomarkers in biofluids are of great interest for assessing health and cognitive state.^{1,4} Among the fundamental challenges limiting the development of these sensors are the discovery of biological recognition elements (BREs) that are able to recognize and bind to biomarker targets with high affinity and specificity to detect low biomarker concentrations. Antibodies are currently the BRE of choice, but they often function poorly in devices because of their high cost, weak thermostability, short shelf life and issues with reproducibility.⁵

Human neuropeptide Y (NPY) is a highly conserved 36-mer peptide (YPSKPDNPGEDAPAEDMARYY-SALRHYINLITRQRY) that functions as a neurotransmitter that is widely distributed in the human central and peripheral nervous systems.⁶ Recent clinical studies have shown that NPY plays an essential role in the regulation of basic physiological effects, including mood disorders, stress responses, and memory processing.^{3,7} The NPY concentration in human biological fluids is an important indicator in the diagnosis of depression, anxiety, and stress-related disorders, including post-traumatic stress disorder (PTSD).⁸ The detection of NPY levels remains a fundamental challenge for incorporation into wearable devices. Traditional methods for NPY detection rely on complex, time consuming, and expensive assays using antibodies.^{1,9} Antibodies against NPY can exhibit cross-reactivity

when used as sensing elements in microarrays, and the production process is laborious and cost-ineffective. Moreover, antibodies are large proteins with a short shelf-life as a result of protein denaturation, which limits their scope in biosensor development.^{5,10}

Short peptides^{3,4} are appealing as BREs because of their high thermostability, strong selectivity, and low cost. In addition, their small size places them in close proximity to the active electronic or plasmonic material surface in sensors, greatly decreasing the limit of detection.¹¹ To date, most peptide-based BREs are identified from natural sources¹² or from combinatorial phage display libraries¹³ which determine binders experimentally from large pools of sequences. The identification of peptides from combinatorial libraries can be a difficult experiment as it is subject to artifacts related to parasitic sequences¹³ and sometimes yields peptides with low binding affinities that are difficult to use in biofluid sensors.

We have been working to develop fast and automated methods to design peptides with exceptional binding affinities for protein or RNA targets.¹⁴⁻¹⁹ Our computational algorithm uses atomistic force fields rather than knowledge-based information to design peptide sequences; this enables us to design high-affinity binding peptides to targets that have no known binders available in the protein data bank. In recent work, we used the computational algorithm to successfully evolve a 12-mer peptide-based BRE for the detection of cardiac event biomarker protein troponin I (cTnI).¹⁷ The results showed that

the *in-silico* evolved peptide binds to cTnI with a high affinity (0.27 nM), which is comparable to that of the natural antibody (0.12 nM) for cTnI. Using plasmonic paper detection, we found a detection limit of 10 fM (0.23 pg/mL), a significant improvement over the commercial assays that require a clinical lab and nearly 24 hours to develop.

In this work, we use two different approaches, viz. combinatorial phage display library and computational peptide design algorithm, to identify high affinity peptide-based BREs for detection of NPY in sensors. The binding affinities of the phage-display and *in-silico* discovered peptides were characterized experimentally using bilayer interferometry (BLI),^{20–21} surface plasmon resonance (SPR),²² and circular dichroism (CD).^{23–24} Our results show that the *in-silico* peptide RNPQPMWQMNW (N16) has an affinity equivalent to that of the phage displayed peptide FPNWSLRPMNQM (N3) in binding with NPY. The binding kinetic properties of peptides to NPY were evaluated by experimentally measuring the association and dissociation coefficients, k_a and k_d . The N16 peptide has more rapid NPY binding kinetics, so the *in-silico* designed peptide might perform better in sensors.

MATERIALS AND METHODS

The reagents for buffers were purchased from Sigma-Aldrich and used without purification. Peptides with a CGGG linker for SPR and biotinylated peptides with a GGG linker were purchased from Peptide 2.0 and HPLC purified to 95%.

Circular dichroism spectra were collected on a Jasco J-815 circular dichroism (CD) spectrometer using a 750 μ L quartz cuvette from 180 to 260 nm with a data pitch of 0.1 nm, bandwidth of 1 nm, scan rate of 50 nm/min, and averaged over 3 scans.

A BioNavis multiparametric surface plasmon resonance (MP-SPR) Navi 210A instrument integrated with a degasser and automated for six samples was used to measure NPY binding kinetics. Gold coated SPR sensors purchased from BioNavis were cleaned via UV-ozone treatment for 10 min, heated in a 7.5 : 1 : 1 solution of water–30% H₂O₂–NH₄OH at 80 °C for 10 min, thoroughly rinsed with double deionized water, and dried with N₂. The clean gold coated SPR sensors were mounted in a sensor holder and inserted into MP-SPR instrument. For immobilization of NPY-binding peptides on gold coated SPR sensors, 500 μ L of peptide at a concentration of 50 μ g/mL in deionized and filtered water was injected at a flow rate of 30 μ L/min using a pre-wait delay time of 2 min, 10 min injection period of peptide or target, and 2–40 min post-wait time to allow for peptide dissociation.

NPY binding to peptides was also measured using bilayer interference (BLI) on a Fortebio Octet 96 spectrometer in 96-well plates under constant agitation. The streptavidin coated sensors were equilibrated with biotinylated peptides, rinsed and dipped into wells containing different concentrations of NPY. All BLI experiments were performed in PBS buffer with 1% bovine serum albumin and 0.002% Tween 20.

Computational peptide design algorithm. Our peptide-design algorithm is an iterative procedure that searches for high-affinity peptide binders to a target biomolecule.^{14–15,19} Figure S1 shows a flow sheet for the computational peptide-de-

sign algorithm. The algorithm is initialized by choosing a reference peptide binder to the target and then determining the initial structure of the peptide-target complex from the PDB, crystallography or atomistic molecular dynamics simulation. The initial rotamers (side-chain conformations) for amino acid repacking along the chain are taken from Lovell's rotamer library.²⁵ Two types of trial moves, sequence change and conformation change, are included in the algorithm to generate new target-binding peptides. (i) Sequence change move: There are two types of trial moves to change the peptide sequence. The first is a random substitution of a new residue for an old one. The new residue should be of the same residue type as the old one to maintain the peptide's hydration properties. The second type of trial move is a random exchange of two chosen residues, regardless of their residue type. Each trial peptide is subjected to Broyden-Fletcher-Goldfarb-Shanno (BFGS) energy minimization to determine optimal side-chain configurations for the amino acids along the chain. (ii) Conformation change move: There are three types of trial "moves" to change the peptide backbone conformation. The first uses the extended concerted rotation method to displace a series of consecutive residues in the middle of the peptide chain, leaving the other residues fixed. The second rotates a peptide fragment on one of the two ends (N- and C- terminus) and the third translates the entire peptide backbone conformation. Two parameters, $\delta_{\max}$ and $kT_{\text{conformation}}$, are used to control the magnitude of the conformation change moves. The root mean-square deviation (RMSD), δ_{rmsd} , of the new trial conformer from its original peptide conformation is evaluated to make sure that it is not too big ($<\delta_{\max}$) and not too small ($>\delta_{\min}$). The parameter $kT_{\text{conformation}}$ controls the likelihood that a new peptide conformer will be accepted, with higher values making acceptance easier. All attempts to generate new peptide backbone conformers are considered as long as (1) the torsion angles (ϕ and ψ) satisfy the Ramachandran plot²⁶ and (2) there are no atomic overlaps between the peptide's backbone and the target. Once trial backbone conformers are generated, side chains are repacked on the trial backbone conformers and BFGS energy minimization is conducted to optimize their configurations. The score Γ_{score} of each trial peptide sequence or conformer is evaluated, and the Monte Carlo (MC) Metropolis algorithm is used to accept or reject the new trial peptide by calculating the acceptance probability

$$P = \min \{1, \exp[(\Gamma_{\text{score}}^{\text{old}} - \Gamma_{\text{score}}^{\text{new}})/kT]\}.$$

The score function that we use to evaluate the merits of each trial peptide in the computational algorithm is given in equation (1), which takes into account the binding affinity of the peptide to the target (first term) and the conformational stability of the peptide when bound to the target (second term):¹⁶

$$\Gamma_{\text{score}} = \Delta U_{\text{binding}} + \lambda \cdot (U_{\text{peptide-VDW}}^{\text{bound}} + U_{\text{peptide-ELE}}^{\text{bound}} + U_{\text{peptide-EGS}}^{\text{bound}}) \quad (1)$$

The binding energy $\Delta U_{\text{binding}}$ is defined to be the difference between the energy of the complex and the energies of the peptide and the target prior to binding. The evaluation of energies (U) of the complex, the peptide and the target involves the calculation of the internal energy (INT), van der Waals energy (VDW), electrostatic energy (ELE), the polar solvation energy (EGS) and the non-polar solvation energy (GBSUR). More detailed descriptions about the calculation of the energy and score function can be seen in Supporting Information and our

previous paper.¹⁵ Lower scores (more negative values) mean better binders. All the force field parameters are taken from the Amber 14SB force field.²⁷

The input for the computational designs is an initial binding structure of the peptide-target complex, and several controllable parameter settings, such as the pH value, the value of ($\delta_{\max}$, $kT_{\text{conformation}}$), the number of residues of each of the six residue types (see below), and an initial random seed that generates random numbers for the sequence and conformation change moves. The computational algorithm cannot be used to predict the active site of a target molecule. Our peptide design is limited to cases in which (i) the active site of a target molecule is known from the protein data bank or can be predetermined using computational approaches, and (ii) an initial peptide sequence is known.

Hydration properties of *in-silico* peptides. Since NPY shows poor solubility at neutral pH, the peptide-binding experiments are performed at pH=5.0. Below pH=6.0, the polar histidine (His) is protonated, becoming a positively-charged amino acid (Hip) (Table S1). To be consistent with the experimental data, our peptide designs are conducted at pH=5.0. If not specified, the letter “H” stands for the positively-charged histidine (Hip). The twenty natural amino acids are classified into six residue types according to their hydrophobicity, polarity, charge and size (Table S1). The length of the *in-silico* evolved peptides is set to be the same as the phage-display-discovered peptide FPNWSLRPMNQM which has twelve residues: five hydrophobic residues ($N_{\text{hydrophobic}}=5$), no negatively-charged residue ($N_{\text{negative charge}}=0$), one positively-charged residue ($N_{\text{positive charge}}=1$), four hydrophilic residues ($N_{\text{hydrophilic}}=4$), two other residues ($N_{\text{other}}=2$), and no glycine ($N_{\text{glycine}}=0$).

Explicit-solvent atomistic molecular dynamics simulation. Explicit-solvent atomistic MD simulations are carried out in the canonical (NVT) ensemble using the AMBER15 package with the force field ff14SB²⁷ to examine the dynamics of the binding process of peptides to NPY. The peptides examined include one phage-display-discovered peptide N3 and four *in silico*-discovered peptides. The starting conformations of NPY complexed with the four evolved peptides for the atomistic MD simulations are obtained from our peptide-design algorithm. Each peptide-NPY complex is solvated in a periodic truncated octahedral box containing an 8 Ångstrom buffer of TIP3P water²⁸ (~8000 water molecules) surrounding the complex in each direction. Chloride counterions (Cl^-) are added to neutralize the system. Three independent simulations are conducted for the peptide-NPY complexes in 150 ns to ensure that our systems reach an equilibrated state. K-means clustering analysis²⁹⁻³⁰ is performed on the last 10 ns of the simulation trajectories to obtain representative structures for these complexes in solution. Using the implicit-solvent molecular mechanics/generalized born surface area (MM/GBSA)³¹ approach with the variable internal dielectric constant model, we post-analyze the last 10 ns simulation trajectories of all the peptide-NPY complexes to calculate their binding free energies. Details of the computational procedures can be found in our previous work.¹⁴⁻¹⁹

Phage display discovery of NPY-binding peptide. The deployment of biochemical sensors for biofluids has been limited by the discovery of BREs for biomarkers of interest. The phage display techniques for the peptide-based BRE discovery and the challenges as well as the limitations of this approach have been reviewed.³² In this study, we firstly utilized the combinatorial phage display library to identify a promising NPY-binding peptide FPNWSLRPMNQM termed as “N3” for convenience. A measurement of the binding affinity of peptide N3 to NPY was obtained by quartz crystal microbalance (QCM); the experimentally-measured disassociation constant was $K_{d(\text{QCM})}=23.9 \mu\text{M}$.³³

Modeling of NPY-peptide structure. As noted in the section on the computational peptide design algorithm above, an initial structure of NPY complexed with a model peptide is required input for the *in silico* discovery of peptide-based BREs. Since the initial structure of the NPY-N3 complex is not in the Protein Data Bank, we used a molecular modeling approach to identify the domains within NPY where peptide N3 binds. The NMR solution structure of neuropeptide Y (NPY) was obtained from the Protein Data Bank (PDB code: 1RON).³⁴ The conformation of (isolated) peptide N3 was initially generated using the Rosetta *ab initio* fragment assembly package.³⁵ An ensemble of 200 folded structures was then generated using replica exchange molecular dynamics simulations in 200-ns at eight temperatures (277.15K, 284.53K, 292.11K, 299.89K, 307.88K, 316.09K, 324.51K and 333.15K). The conformation for the peptide-NPY complex was generated by docking the folded peptide N3 with the NMR structure of NPY in two steps. Rigid docking was performed using the ZDOCK³⁶ package with the default force field parameters, and the generated structures were processed in Rosetta with the flexible docking package. The refined complexes were scored and ranked according to the Rosetta energy function. Table S2 shows the scoring energy for the 5 best conformations of NPY with the peptide N3. Figure 1(a) shows the low-energy structure for NPY bound to the peptide N3 that was used as input for the *in silico* peptide evolution.

RESULTS

***In-silico* evolution of NPY binding peptides.** We performed four independent evolutions to ensure that the computational algorithm samples peptides in a broad sequence and conformation space. The four independent evolutions start from random sequences and proceed along different searching pathways that are controlled by setting distinct input parameters, such as $\delta_{\max}$ and $kT_{\text{conformation}}$. The first two evolutions include sequence change moves only, while the last two evolutions include both sequence and conformation change moves. Figure 1(b) shows an example of the score vs. number of steps when only sequence changes are attempted. It is clear that the score profile fluctuates considerably with the number of evolution steps, indicating that this procedure examines a significant range in sequence space. By examining the score profile over the course of the evolution steps, we can identify the lowest scores which correspond to the best peptide sequences for this search. Figure 1(c) shows an example of the score vs. number of steps when both sequence and conformation changes are attempted at ($\delta_{\max}$, $kT_{\text{conformation}}$) = (4.0, 4.0). It is clear that the peptide conformations experience a sizeable variation in the RMSD profile, implying that the

search is accessing a broad conformational space. The score profile associated with newly-generated peptide conformations fluctuates considerably, and eventually gets down to an even lower value than in the absence of conformation changes (Figure 1b).

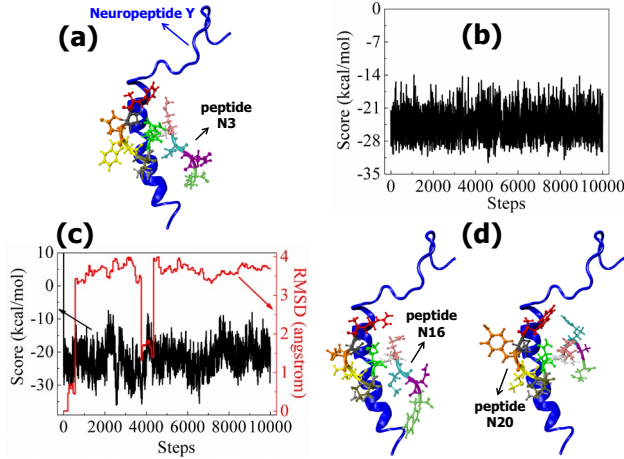


Figure 1. Computational peptide-design algorithm is used to discover high-affinity NPY-binding peptides. (a) Docking pose of the phage-display-discovered peptide N3 on the 36-mer NPY is obtained using the ZDOCK package. This docking pose is used as a starting conformation in the algorithm to evolve other peptide binders. (b) Sequence evolution proceeds with only sequence change attempts, resulting in the best-scoring peptide N16. The fluctuation in the score vs. number of evolution step indicates the extent of variation of the peptide sequence. (c) Sequence evolution proceeds with both sequence and conformation change attempts at $(\delta_{\max}, kT_{\text{conformation}}) = (4.0, 4.0)$, resulting in the best-scoring peptide N20. Profile of RMSD vs. evolution step (red trace) indicates the extent of variation of the peptide conformation. (d) The structures of the complexes N16-NPY and N20-NPY are obtained via the computational peptide-design algorithm.

The sequences of the four best-scoring peptides, viz. N16, N17, N20 and N21, identified by *in-silico* evolution are given in Table 1. The peptides N16 and N17 result from sequence-change-move-only searches at two distinct sets of random number seeds, while the peptides N20 and N21 result from cases where both sequence and conformation change moves were attempted for $(\delta_{\max}, kT_{\text{conformation}}) = (4.0, 4.0)$ and $(5.0, 4.0)$, respectively. Figure 1(d) shows the structures of the *in-silico* peptides N16 and N20 bound to NPY that are obtained via the computational algorithm. By comparing Figures 1(a) and 1(d), we can see that peptide N16 retains the same backbone conformation as the original N3, but with a different residue sequence draped on its backbone motif. The conformation and sequence of Peptide N20 are both different from that of the original N3.

Explicit-solvent atomistic MD simulations are carried out in 150 ns simulations to examine the dynamic properties of NPY when bound to the phage-display-discovered peptide N3 and to the four *in-silico*-discovered peptides N16, N17, N20, and N21. The motivation for these MD simulations is that the target NPY is not allowed to move in response to the changes of peptide sequence, so we cannot guarantee that the *in-silico*

peptides with the lowest scores have higher affinity for NPY than the original peptide N3. The last 10-ns simulation trajectories of all the peptide-NPY complexes are analyzed to calculate their respective binding free energies (ΔG_{cal}) using the implicit-solvent MM/GBSA approach with the variable internal dielectric constant model, as listed in Table 1. Our simulation results reveal that the original peptide N3 exhibits a good affinity to NPY with a low computed binding free energy -14.43 kcal/mol. (Note that the lower the value of ΔG_{cal} , the higher the binding affinity.) By comparison, three out of the four *in-silico* peptides, viz. N16, N17 and N21, are found to have even lower binding free energies: -18.86 kcal/mol, -16.83 kcal/mol, and -18.90 kcal/mol, respectively, suggesting that the three *in-silico* peptides may bind to the target NPY with higher affinity than the original N3.

Table 1. Sequences of the original peptide N3 and the four evolved peptides as well as their corresponding scores and binding free energies obtained from the computational algorithm and the atomistic MD simulations, respectively.

Peptide	Sequence	Γ_{score} (kcal/mol)	ΔG_{cal} (kcal/mol)
N3 ^a	FPNWSLRPMNQM	-21.86	-14.43
N16 ^b	RNPQPMMWQMNW	-32.58	-18.86
N17	RNPQPWTWWLTW	-32.86	-16.83
N20 ^c	WQYMPMQWQRAQ	-37.34	-3.34
N21	YNPQPMTMRYNW	-36.12	-18.90

(a) The peptide N3 is discovered using phage display technique; (b) the peptides N16 and N17 result from the *in-silico* evolution with only sequence change attempts; (c) the peptides N20 and N21 result from the *in-silico* evolution with both sequence and conformation change attempts.

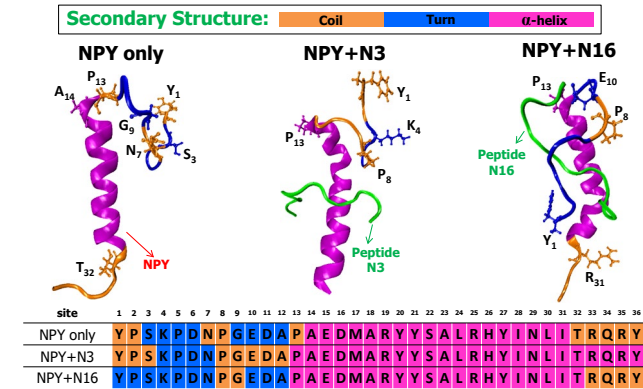


Figure 2. The last 10-ns MD simulation trajectories of NPY only, NPY+N3 and NPY+N16 complexes are analyzed to observe the secondary structures of the 36-mer NPY in solution. The random coils are colored in orange, the turns are in blue, and the α -helices are in pink. The secondary structure content along the sequence is shown to facilitate comparison of the structural changes of NPY when bound to the three peptides.

The effect of binding on the secondary structures of NPY was examined using the VMD program suite.³⁷ In order to compare the structures of the bound peptides, we analyzed all the residues on NPY in the last 10-ns MD simulation trajectories

to obtain a probability distribution for the secondary structure content along the 36 chain sites (Figure S2). K-means clustering analysis²⁹ is performed to obtain representative structures of NPY alone in TIP3P water and when bound by the peptides N3 and N16 in TIP3P water. Three primary secondary structures: coil, turn, and α -helix, were observed for NPY in both the free and bound states. In comparison with free NPY (Figure 2), we found that when bound to NPY, the peptides N3 and N16 do not cause a significant change in the folded structure of NPY; instead NPY (more-or-less) maintains the same α -helix in the middle (A₁₄-I₃₁).

Experimental Validation of *In-silico* Binding Peptides.

The binding affinities for the phage-display and *in silico* discovered peptides were measured using bilayer interferometry (BLI)²¹ and surface plasmon resonance (SPR).²² The BLI experiments were performed using biotin-labeled peptides with GGG linkers on the amino terminus. The signals were measured following exposure to concentrations of NPY between 0.074 and 2.38 μ M. The SPR experiments used thiolated peptides with a CGGG linker on a gold surface.

Both BLI and SPR are sensitive to the changes in refractive index when the target molecules bind to the BRE attached to the surface. In BLI the reflected light from a probe dipped in target solutions of varying concentrations will change as the target binds to the surface. Under optimal conditions the reflected intensity will rise as the target binds and the time-dependent signal change can be directly related to the association constant k_a . The signal then saturates as the system reaches equilibrium. Finally, the dissociation coefficient k_d can be calculated from the time-dependent change in intensity as the surface is washed with buffer.

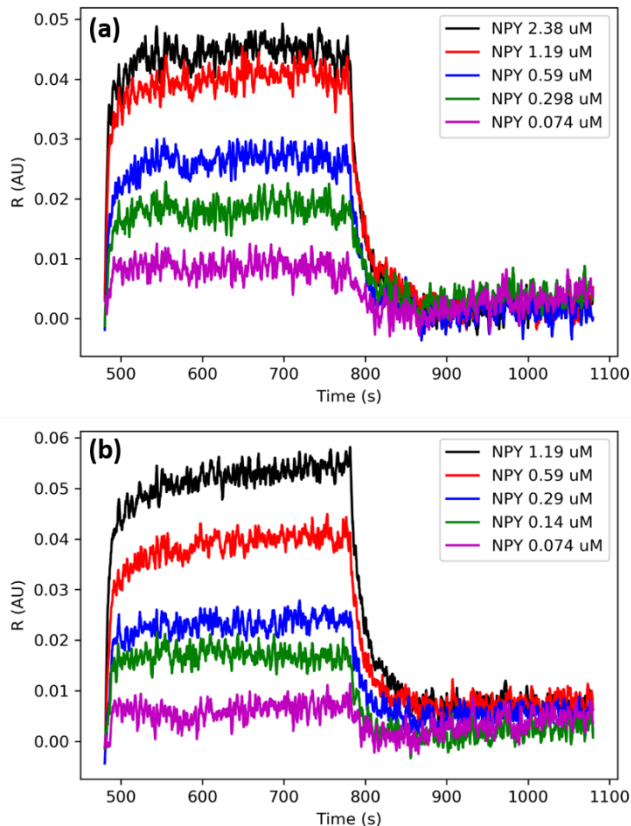


Figure 3. The BLI data (signal intensity R) for the peptides (a) N3 and (b) N16 with increasing concentrations ([NPY]) of NPY. Changes in signal intensity are due to changes in refractive index as NPY binds the surface-attached peptide. The change in signal is proportional to the amount bound and changes to nonspecific binding have been subtracted. After surface coating with the peptides, NPY was added at 480 s and buffer washed starting at 780 s.

Figure 3 shows the baseline-subtracted BLI data for the peptides N3 and N16. In both cases, a rapid rise in signal was observed upon exposure to NPY followed by a plateau as the system comes to equilibrium. A return to the baseline signal is observed after washing with buffer at 780 s due to dissociation of the NPY. No rise in the BLI signal was observed for the *in-silico* peptides N17 and N21 peptides under these conditions. It is possible that the N17 and N21 are weaker peptide binders and that signals could have been observed at higher NPY concentrations, but we did not pursue these experiments.

The binding affinities can be determined in principle from either the rate of change in signal intensity upon exposure to NPY or the plateau signal intensity in either the BLI or SPR experiments. The initial change in signal intensity can be fit to a rising exponential given by

$$R(t) = R_{eq}(1 - e^{-\frac{t}{k_{obs}}}),$$

where $R(t)$ is the observed signal, R_{eq} is the plateau value of the observed intensity, and the observed association coefficient k_{obs} is related to the association and dissociation coefficients by

$$k_{obs} = k_a[NPY] + k_d \quad (2)$$

The slope of a plot of k_{obs} vs. [NPY] (not shown) gives k_a , the intercept gives k_d , and the binding affinity K_d is given by the ratio of k_d/k_a . The k_d value can independently be determined from a fit of the decrease in signal intensity when the sample equilibrated with NPY is washed with buffer. The buffer wash occurs at 780 s in the data shown in Figure 3. The binding affinity can also be determined from the equilibrium plateau value of the signal intensity R_{eq} in the BLI or SPR experiments as a function of NPY concentration. In this case the plateau value R_{eq} is related to the maximum change in R (R_{max}) and the K_d as a function of NPY concentration by

$$R_{eq} = \frac{[NPY]R_{max}}{(K_d + [NPY])} \quad (3)$$

and the values for K_d and R_{max} are obtained from a least squares fit of R_{eq} vs [NPY].

The BLI signal intensity for probes with the surface-attached N3 and N16 peptides in Figure 3 shows large changes in the BLI response with changing NPY concentrations, demonstrating that both peptides bind to NPY. The initial rise in intensity is too rapid to be accurately fit by the kinetic equation, but the data give an accurate measure of the equilibrium binding plateau and the disassociation coefficient k_d . Figure 4 shows that the BLI plateau values gives a good fit to the equilibrium model, allowing us to determine the value for K_d . Although the association kinetics are too fast to accurately measure by BLI at these concentrations, we can directly measure the dissociation constant k_d . Given that K_d and k_d can be accurately measured, the value for the association coefficient k_a can be inferred from the ratio of k_d/K_d .

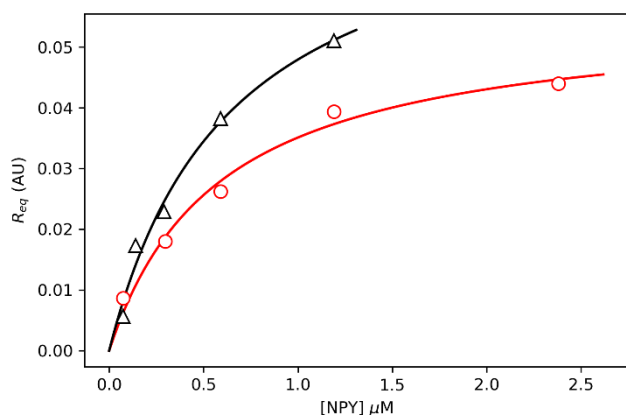


Figure 4. The equilibrium analysis of the BLI data (signal intensity R) for the peptides N3 (red) and N16 (black). The value of R_{eq} was determined from the plateau binding intensity as a function of NPY concentration as shown in Figure 3.

Table 2. The NPY kinetics and binding affinities for surface attached biotinylated N3 and N16.

	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_d (μM)	ΔG_{exp} (kcal/mol)
bN3	3.2×10^3	1.9×10^{-3}	0.59 ± 0.1	-8.49
bN16	7.1×10^3	4.6×10^{-3}	0.64 ± 0.1	-8.44

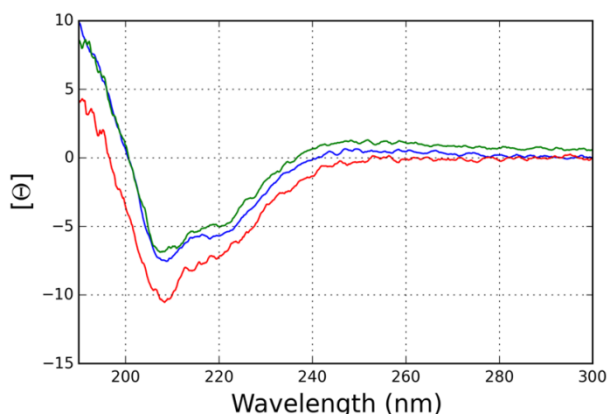


Figure 5. The CD spectra of NPY (blue) in solution and in the presence of the 3-fold excess of the peptides N3 (red) and N16 (green). The spectra of the free peptide have been subtracted for the N3 and N16.

Table 2 shows the results of the kinetic and equilibrium binding analysis for biotinylated N3 and N16 binding to NPY. The binding affinities for both peptides are below the μM limit (N3: $0.59 \pm 0.1 \mu M$ and N16: $0.64 \pm 0.1 \mu M$), and are very similar to each other. According to the equation $\Delta G = -RT \ln(K_d)$, we calculated the experimental binding free energies, -8.49 and -8.44 kcal/mol, of the peptides N3 and N6, respectively. By comparing the calculated and experimentally-measured binding free energies in Tables 1 and 2, we found that the values of ΔG_{cal} (-14.43 and -18.86 kcal/mol) are lower than the values of ΔG_{exp} (-8.49 and -8.44 kcal/mol). The reason for this inconsistency is that we neglect the enthalpic and entropic contributions from water. This is a consequence of our use of the

implicit-solvent MM/GBSA approach with the variable internal dielectric constant model to calculate the binding free energy of peptide and NPY. While the affinities are similar, the results show that the N16 peptide has both faster association and dissociation with NPY.

The secondary structure of the peptide-NPY complexes was also evaluated by Circular Dichroism (CD) spectroscopy²³ for comparison with our atomistic MD simulation. Previous studies have shown NPY to be partially folded in an α -helix, with the helical content depending on the solvent, pH, temperature and the presence of lipid micelles.³⁴⁻³⁸ Figure 5 shows the CD spectra of NPY in the absence and presence of the N3 and N16 peptides. For direct comparison with the NPY conformation, the spectra from the free peptides have been subtracted from the NPY-peptide mixture. The results for NPY show the typical features expected from an α -helical peptide, with negative peaks at 208 and 222 nm. No significant changes are observed in the CD spectra of NPY in the presence of the N16, suggesting that the α -helix is maintained during the binding event. The CD intensity increases in the presence of N3, but there is no significant change in the intensity ratio of 208 to 222 peaks that are a signature of α -helix formation. Taken together these data suggest that the binding of N3 or N16 does not lead to loss in α -helical content, which is consistent with our molecular dynamics simulations and the structures shown above (Figure 2).

DISCUSSION

Our experimental and *in-silico* results show that it is feasible to identify peptide binders with high affinity to NPY using both computational peptide design algorithms and phage-display combinatorial libraries. As noted in the literature, phage display discovery can be experimentally challenging and time consuming.³² Thus, it is reasonable to believe that our computational protocols represent a significant advance in the state-of-the-art since they are applicable to a wide variety of experimental systems. The *in-silico* discovery of target-binding peptides will also be advanced further by improving our computational resources in the future.

The initial binding affinity measurements using QCM suggested that the phage display binding peptide N3 had a low binding affinity ($K_{d(QCM)} = 23.9 \mu M$) for NPY and therefore would not have been useful in sensors measuring biologically relevant NPY concentrations.³³ However, using the high sensitivity BLI device, we measured a much stronger affinity for the N3 peptide ($0.59 \pm 0.1 \mu M$). Our hypothesis is that this discrepancy is a consequence of the lower sensitivity of QCM³⁹ compared to BLI. The resonant frequency of the quartz crystal changes with adsorbed mass as NPY binds to a surface-attached peptide and the concentration ranges for the BLI experiments ($[NPY] > 0.07 \mu M$) are well below those that can be measured by QCM. We believe that the more sensitive BLI allows us to measure a binding mode that is not accessible to QCM. To confirm this hypothesis, we measured a binding affinity of N16 for NPY using an independent method. The results from an SPR analysis (Figure S3) also shows a sub- μM affinity for NPY ($K_d = 0.25 \mu M$).

While equivalent binding affinities were observed for the N3 and N16, we note that a significant difference is observed in the association and dissociation kinetics. These differences could have an important impact on sensor response since the

association rate determines the time required for signal measurements and the dissociation rate is critical for sensor regeneration.

Although the values of K_d for the peptides N3 and N16 are below the μM range, they are significantly higher than the concentration of NPY in human plasma (nM) and sweat (pM).⁴⁰ Recent studies have shown that such peptide-based BREs can be useful for sensors. It has been reported that peptide BREs with μM affinities can be used to create sensors based on graphene transistors with a lower limit of detection of 10 pM.³³ We would therefore expect that the *in-silico* peptide N16 could be used as a drop-in replacement for the N3. The more rapid association and dissociation kinetics would be useful for generation of the signal responses in continuous monitoring sensors.

As with the phage-display peptide, the *in-silico* peptides discovered by the computational algorithm must be validated in experimental binding studies. Of the three *in-silico* peptides considered here, N16, N17 and N21, only the peptide N16 shows a good binding affinity to NPY. The *in-silico* peptides N17 and N21 are false positives in our computational designs because they exhibit better affinities than the phage-display peptide N3 in the atomistic MD simulations but poorer affinities in the experimental measurements. The peptide N17 is similar to N16 but contains two extra hydrophobic residues tryptophan (W), giving N17 a low solubility in solution and therefore making it difficult to test experimentally. The peptide N21 does not have the solubility issue, but it also fails to show a proper response in the BLI NPY-binding experiments over the μM concentration range. Thus, we consider N21 to be either a weak binder or a non-binder. One of the challenges in our current algorithm for peptide-based BRE discovery is that the *in-silico* peptides with the best scores may not represent the best experimental binders. Based on the feedback from experiments, we will further improve the score function in future work by introducing a peptide hydration term to address the solubility issues, thus enhancing the performance of the computational algorithm in designing good binding peptides. An expression that describes a peptide's hydration properties given in Pawar's work⁴¹, will be introduced and modified in our work to avoid the designed peptides being over-hydrophobic.

CONCLUSION

The discovery of peptide-based BREs for targets of interest for monitoring human health remains a significant challenge. In this study, we have combined the combinatorial phage display technique with an *in-silico* peptide design method to identify high affinity peptide binders for neuropeptide Y (NPY), a biomarker for stress and cognitive performance.³⁷ The experimental results show that the *in-silico* evolved peptide exhibits an affinity equivalent to the phage display discovered peptide in binding with NPY but with more rapid kinetics. We have

REFERENCES

1. Huang, W.; Diallo, A. K.; Dailey, J. L.; Besar, K.; Katz, H. E., Electrochemical processes and mechanistic aspects of field-effect sensors for biomolecules. *Journal of Materials Chemistry C* **2015**, *3*, 6445-6470.
2. Kim, S. S.; Kuang, Z.; Ngo, Y. H.; Farmer, B. L.; Naik, R. R., Biotic-abiotic interactions: Factors that influence peptide graphene interactions. *ACS Applied Materials and Interfaces* **2015**, *7*, 20447-20453.

shown previously with cardiac troponin I that the high-affinity peptide-based BREs can be incorporated into optically detected devices, including plasmonic paper, to enable very low limits of detection.^{14,17} The results presented here show a computational strategy to optimize peptide-based BREs for a wide variety of biomarkers that could be incorporated into the next generation of wearable sensors.

ASSOCIATED CONTENT

Supporting Information. Computational and experimental methods are included in Supporting Information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

* Email: hall@ncsu.edu

* Email: peter.mirau@us.af.mil

Author Contributions

The manuscript was written through contributions of all authors and all authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

CKH and XX acknowledge support from the Air Force Office of Scientific Research Award No. FA9550-16-1-0078 and from the National Science Foundation, CBET-1830272 and CBET 1743432. We would like to thank ForteBio Systems for assistance in acquiring BLI data. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. XX and CKH thanks San Diego Supercomputer Center (SDSC) for providing computing time.

ABBREVIATIONS

BRE, Biological Recognition Element; SPR Surface Plasmon Resonance; BLI, Bilayer Interference, MD, Molecular Dynamics, NPY, Neuropeptide Y; HPLC, High Performance Liquid Chromatography; RMSD, Root Mean Square Deviation; QCM, Quartz Crystal Microbalance; PBS, Phosphate Buffered Saline

3. Kormos, V.; Gaszner, B., Role of neuropeptides in anxiety, stress, and depression: From animals to humans. *Neuropeptide* **2013**, *47*, 401-419.
4. Kuang, Z.; Kim, S. N.; Crookes, W. J.; Farmer, B. L.; Naik, R. R., Biomimetic chemosensor: Designing peptide recognition elements for surface functionalization of carbon nanotube field effect transistors. *ACS Nano* **2010**, *4*, 452-458.

5. Baker, M., BLAME IT ON THE ANTIBODIES. *Nature* **2015**, *521*, 274-276.
6. Wettstein, J. G.; Earley, B.; Junien, J. L., Central nervous system pharmacology of neuropeptide Y. *Pharmacology and Therapeutics* **1995**, *65*, 397-414.
7. MoralesMedina, J. C.; Dumont, Y.; Quirion, R., A possible role of neuropeptide Y in depression and stress. *Brain Research* **2010**, *1314*, 194-205.
8. Rasmusson, A. M.; Schnurr, P. P.; Zukowska, Z.; Sciolli, E.; Forman, D. E., Adaptation to extreme stress: Posttraumatic stress disorder, neuropeptide Y and metabolic syndrome. *Experimental Biology and Medicine* **2010**, *235*, 1150-1162.
9. Baranowska, B.; Wolińska-Witort, E.; Martyńska, M.; Chmielowska, M.; Baranowska-Bik, A., Plasma orexin A, orexin B, leptin, neuropeptide Y (NPY) and insulin in obese women. *Neuroendocrinology Letters* **2005**, *26*, 293-296.
10. Ronkainen, N. J.; Halsall, H. B.; Heineman, W. R., Electrochemical biosensors. *Chemical Society Reviews* **2010**, *39*, 1747-1763.
11. Tadepalli, S.; Kuang, Z.; Jiang, Q.; Liu, K. K.; Fisher, M. A.; Morrissey, J. J.; Kharasch, E. D.; Slocik, J. M.; Naik, R. R.; Singamaneni, S., Peptide Functionalized Gold Nanorods for the Sensitive Detection of a Cardiac Biomarker Using Plasmonic Paper Devices. *Scientific Reports* **2015**, *5*, 16206-11.
12. Kröger, N.; Deutzmann, R.; Sumper, M., Polycationic peptides from diatom biosilica that direct silica nanosphere formation. *Science* **1999**, *286*, 1129-1132.
13. Molek, P.; Strukelj, B.; Bratkovic, T., Peptide phage display as a tool for drug discovery: Targeting membrane receptors. *Molecules* **2011**, *16*, 857-887.
14. Xiao, X.; Agris, P. F.; Hall, C. K., Designing peptide sequences in flexible chain conformations to bind RNA: A search algorithm combining Monte Carlo, self-consistent mean field and concerted rotation techniques. *Journal of Chemical Theory and Computation* **2015**, *11*, 740-752.
15. Xiao, X.; Agris, P. F.; Hall, C. K., Introducing folding stability into the score function for computational design of RNA-binding peptides boosts the probability of success. *Proteins: Structure, Function and Bioinformatics* **2016**, *84*, 700-711.
16. Xiao, X.; Hall, C. K.; Agris, P. F., The design of a peptide sequence to inhibit HIV replication: A search algorithm combining Monte Carlo and self-consistent mean field techniques. *Journal of Biomolecular Structure and Dynamics* **2014**, *32*, 1523-1536.
17. Xiao, X.; Kuang, Z.; Slocik, J. M.; Tadepalli, S.; Brothers, M.; Kim, S.; Mirau, P. A.; Butkus, C.; Farmer, B. L.; Singamaneni, S.; Hall, C. K.; Naik, R. R., Advancing Peptide-Based Biorecognition Elements for Biosensors Using in-Silico Evolution. *ACS Sensors* **2018**, *3*, 1024-1031.
18. Xiao, X.; Wang, Y.; Leonard, J. N.; Hall, C. K., Extended Concerted Rotation Technique Enhances the Sampling Efficiency of the Computational Peptide-Design Algorithm. *Journal of Chemical Theory and Computation* **2017**, *13*, 5709-5720.
19. Xiao, X.; Zhao, B.; Agris, P. F.; Hall, C. K., Simulation study of the ability of a computationally-designed peptide to recognize target tRNA(Lys3) and other decoy tRNAs. *Protein Science* **2016**, *25*, 2243-2255.
20. Concepcion, J.; Witte, K.; Wartchow, C.; Choo, S.; Yao, D. F.; Persson, H.; Wei, J.; Li, P.; Heidecker, B.; Ma, W. L.; Varma, R.; Zhao, L. S.; Perillat, D.; Carricato, G.; Recknor, M.; Du, K.; Ho, H.; Ellis, T.; Gamez, J.; Howes, M.; Phi-Wilson, J.; Lockard, S.; Zuk, R.; Tan, H., Label-Free Detection of Biomolecular Interactions Using BioLayer Interferometry for Kinetic Characterization. *Combinatorial Chemistry & High Throughput Screening* **2009**, *12*, 791-800.
21. Wartchow, C. A.; Podlaski, F.; Li, S. L.; Rowan, K.; Zhang, X. L.; Mark, D.; Huang, K. S., Biosensor-based small molecule fragment screening with biolayer interferometry. *Journal of Computer Aided Molecular Design* **2011**, *25*, 669-676.
22. Homola, J., Surface plasmon resonance sensors for detection of chemical and biological species. *Chemical Reviews* **2008**, *108*, 462-493.
23. Greenfield, N. J., Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **2006**, *1*, 2876-2890.
24. Whitmore, L.; Wallace, B. A., Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **2008**, *89*, 392-400.
25. Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C., The penultimate rotamer library. *Proteins Structure Function and Genetics* **2000**, *40*, 389-408.
26. Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V., Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **1963**, *7*, 95-99.
27. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins Structure Function and Bioinformatics* **2006**, *65*, 712-725.
28. Mark, P.; Nilsson, L., Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *Journal of Physical Chemistry* **2001**, *105*, 9954-9960.
29. Kanungo, T.; Mount, D. M.; Netanyahu, N. S.; Piatko, C. D.; Silverman, R.; Wu, A. Y., An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2002**, *24*, 881-892.
30. Shao, J. Y.; Tanner, S. W.; Thompson, N.; Cheatham, T. E., Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *Journal of Chemical Theory and Computation* **2007**, *3*, 2312-2334.
31. Hou, T. J.; Wang, J. M.; Li, Y. Y.; Wang, W., Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling* **2011**, *51*, 69-82.
32. Matochko, W. L.; Chu, K. K.; Jin, B. J.; Lee, S. W.; Whitesides, G. M.; Derda, R., Deep sequencing analysis of phage libraries using Illumina platform. *Methods* **2012**, *58*, 47-55.
33. Kim, S.; Xing, L.; Hsiao, M.-s.; Ngo, Y.; Pavlyuk, O. M.; Martineau, R. L.; Bedford, N.; Islam, A.; Kadakia, M. P.; Drummy, L. F.; Rajesh, R., In operando observation of neuropeptide capture and release on GFET biosensors with pico-Molar sensitivity. *ACS Applied Materials & Interface* **2019**, *11*, 13927-13934.
34. Monks, S. A.; Karagianis, G.; Howlett, G. J.; Norton, R. S., Solution structure of human neuropeptide Y. *Journal of Biomolecular NMR* **1996**, *8*, 379-390.
35. Das, R.; Baker, D., Macromolecular modeling with Rosetta. *Annual Review of Biochemistry* **2008**, *77*, 363-382.
36. Chen, R.; Weng, Z. P., Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins Structure Function and Bioinformatics* **2002**, *47*, 281-294.
37. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics & Modelling* **1996**, *14*, 33-38.
38. Bader, R.; Bettio, A.; Beck-Sickinger, A. G.; Zerbe, O., Structure and dynamics of micelle-bound neuropeptide Y: Comparison with unligated NPY and implications for receptor selection. *Journal of Molecular Biology* **2001**, *305*, 307-329.
39. Marx, K. A., Quartz crystal microbalance: A useful tool for studying thin polymer films and complex biomolecular systems at the solution-surface interface. *Biomacromolecules* **2003**, *4*, 1099-1120.
40. Morgan, C. A.; Wang, S.; Southwick, S. M.; Rasmusson, A.; Hazlett, G.; Hauger, R. L.; Charney, D. S., Plasma neuropeptide-Y concentrations in humans exposed to military survival training. *Biological Psychiatry* **2000**, *47*, 902-909.

41. Pawar, A. P.; DuBay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M., Prediction of "Aggregation-prone" and "Aggregation-susceptible" Regions in Proteins Associated with Neurodegenerative Diseases. *Journal of Molecular Biology* **2005**, *350*, 379-392.
-