Bioorg. Med. Chem. Lett. 51 (2021) 128341

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

HSymM-guided engineering of the immunodominant p53 transactivation domain putative peptide antigen for improved binding to its anti-p53 monoclonal antibody

Zachary R. Fritz [a], Rene S. Schloss [a], Martin L. Yarmush [a], Lawrence J. Williams [b,*]

[a] Department of Biomedical Engineering, Rutgers University, Piscataway, NJ 08854, United States
[b] Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, United States

ARTICLE INFO

Keywords:
Peptides
Antibody
Peptide engineering
Epitope
Paratope
p53
Hidden symmetry model
Coarse grained models
Immunoassays
Autoantibodies

ABSTRACT

A novel engineering strategy to improve autoantibody detection with peptide fragments derived from the parent antigen is presented. The model system studied was the binding of the putative p53 TAD peptide antigen (residues 46–55) to its cognate anti-p53 antibody, ab28. Each engineered peptide contained the full decapeptide epitope and differed only in the flanking regions. Since minimal structural information was available to guide the design, a simple epitope:paratope binding model was applied. The Hidden Symmetry Model, which we recently reported, was used to guide peptide design and estimate per-residue contributions to interaction free energy as a function of added C- and *N*-terminal flanking peptides. Twenty-four peptide constructs were designed, synthesized, and assessed for binding affinity to ab28 by surface plasmon resonance, and a subset of these peptides were evaluated in a simulated immunoassay for limit of detection. Many peptides exhibited over 200-fold enhancements in binding affinity and improved limits of detection. The epitope was reevaluated and is proposed to be the undecapeptide corresponding to residues 45–55. HSymM calculated binding free energy and experimental data were found to be in good agreement ($R^2 > 0.75$).

Here we report a novel engineering strategy to enhance the binding of a peptide to a protein. The model system we studied was the binding of the p53 transactivation domain (TAD) decapeptide antigen (Ag), **1**, which corresponds to residues 46–55 of the full protein, to its specific monoclonal antibody (mAb), **ab28** (Figure 1). The p53 protein appears to be the most commonly mutated protein in cancer.[1] Many mutations compromise the important regulatory roles of p53 and lead to its overexpression, cellular accumulation, and autoimmune targeting.[2] Detection of anti-p53 autoantibodies (aAb) in the circulatory system is a potential marker of early disease onset. Recent clinical data show that detection anti-p53 and related autoantibodies in human patients represents a modality of cancer detection that can precede other methods of diagnosis by well-over two years.[3,4] Normally, an exact facsimile of an antigenic whole protein or protein segment is prepared and immobilized–*as is*–for antibody detection.[5] In recent years the cancer aAb research landscape has expanded to include linear peptide antigens. Even though such antigens lack the ability to detect aAbs specific to non-contiguous epitopes, immunoassays based on peptides offer many advantages, especially ease of synthesis, greater stability, and a high degree of customizability.[3,6] Our interest in protein interactions, molecular design, and aAb detection led us to consider the advantages of engineering peptides to have superior protein binding properties. Linear epitopes of the p53 protein, for example the TAD decamer (**1**) and others near the N- or C- termini, have been shown to be immunodominant tumor associated antigens useful for aAb detection and clinical evaluation.[6,7] We reasoned that antigens with enhanced affinities to aAbs may be able to detect a greater proportion of the low and medium affinity polyclonal aAb response, thereby increasing immunoassay sensitivity. Guided by our method of estimating per-residue contributions to interaction free energy, we designed and synthesized twenty-four peptide constructs, measured their binding affinity to **ab28** by surface plasmon resonance (SPR), and evaluated a subset of these peptides in a simulated immunoassay limit of detection (LOD) study. Each engineered peptide contained the full decapeptide epitope and differed only in the flanking regions. Several peptides exhibited over 200-fold enhancements in binding affinity and up to 200-fold improvements in their LOD. The data led us to reevaluate the putative epitope and suggest that the correct epitope is the undecapeptide corresponding to residues 45–55 in the

* Corresponding author.
E-mail address: Lwilliams@chem.rutgers.edu (L.J. Williams).

https://doi.org/10.1016/j.bmcl.2021.128341
Received 1 June 2021; Received in revised form 17 August 2021; Accepted 22 August 2021
Available online 26 August 2021

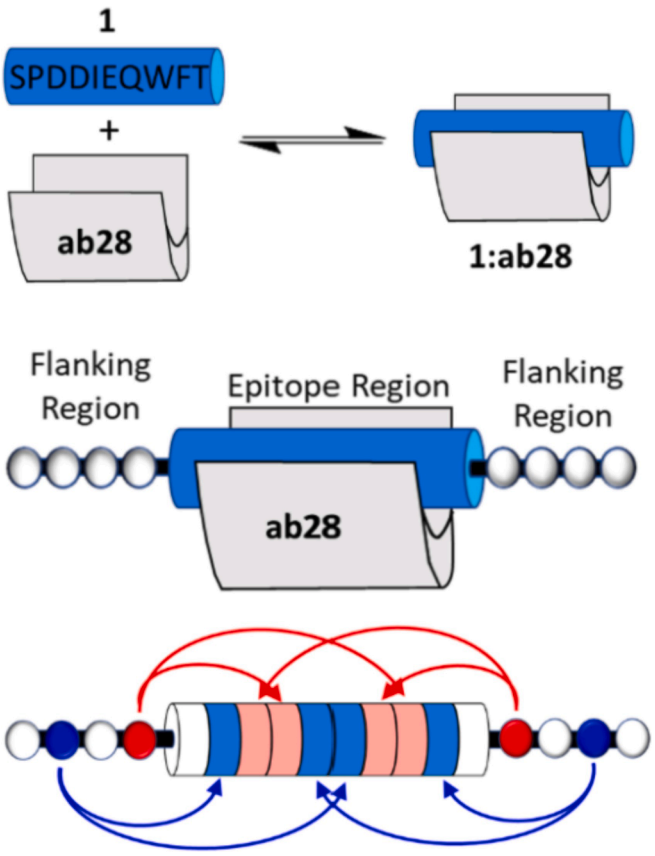

Figure 1. Peptide Engineering Strategy. The monoclonal antibody (**ab28**) is specific to the decapeptide sequence corresponding to residues 46–55 of human p53 (**1**). Addition of flanking residues is proposed to modulate the binding affinity of the peptide by increasing (red arrows) or attenuating (blue arrows) the interaction energies of the interior residues. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

protein. Comparison of our predicted binding free energy to the experimental data gave a good correlation ($R^2 > 0.75$).

We used the Hidden Symmetry model (HSymM) to guide our engineering effort. This model was recently described and used to rationalize mutational effects on protein thermal stability.[8] Even though per-residue interaction energy predictions are expected to be approximate only, in our benchmark study HSymM provided good free energy estimates of a set of 28 T4 lysozyme mutants ($R^2 > 0.7$) and suggested that this approximation method may be useful for many applications, including peptide-protein binding. The full equations and parameters are discussed in reference[8] and are included in the Supplementary Information (SI). Briefly, the model assumes extensive, small dynamical fluctuations and describes the effective contribution to interaction energy of each residue (μ_i) as a simple function of the intrinsic contribution to interaction energy (γ_i), sequence, and coarse residue conformation (σ_i). The γ-scale was reported in 2007[9] and appears to be the only demonstrably context-independent and scale-invariant amino acid descriptor set that reflects compaction forces and excluded volume effects [see also reference[9] and Table S1, SI]. The polypeptide conformation is indexed as one of four structural arrangements and is primarily dependent on whether the residue makes a contact with the first, second, third, or fourth residue in the sequence (σ = 1, 2, 3, or 4). The effective contribution (μ) is determined by normalizing the intrinsic contributions (γ) of the residue with a subset of residues nearby in the sequence depending on σ. In this way, residues nearby in the sequence may bear on the μ-values of other residues. γ-values range from about 0.250 to 0.070, typical μ-values are near 0.145, and values near 0.200 are very high whereas values near 0.100 are very low. A scaling factor (λ = 12.5 kcal/mol) allows the interaction energy between residues *i* and *j* to be approximated ($G - \lambda \mu_i \mu_j$). For example, two residues of μ = 0.200 would interact more strongly than two residues of μ = 0.100 by an estimated –0.375 kcal/mol. Conveniently, the model allows fast (<1 *sec*), single-state calculations of proteins and protein complexes with minimal computing ability, and the output can be viewed as a heat map, where midrange μ-values are depicted as white, above midrange values are shown as increasingly red (hot) and below midrange values are shown as increasingly blue (cold).

As depicted in Figure 1, we aimed to improve the binding of **1** to **ab28** without modifying the peptide epitope. Changing residues outside the binding interface, according to our model, would change one or more μ-values of the ligand epitope and thereby contribute to the binding equilibrium (**1** + **ab28** → **1:ab28**). Since the binding epitope is constant, the sidechain and backbone entropy of the binding interface should be constant and can be ignored; the binding of **ab28** to an engineered variant relative to **1** can be approximated as $\Delta\Delta G = -\lambda(\Sigma\mu_k\mu_{Ab} - \Sigma\mu_j\mu_{Ab})$, where μ_j corresponds to epitope values calculated for **1**, μ_k corresponds to those of the engineered peptide, and μ_{Ab} to the paratope residues of **ab28**. Accordingly, flanking sequences that were expected to increase the μ-values of the decapeptide epitope and increase binding affinity were added to either the C-, N-, or both termini of the epitope.

We adopted a minimalist model for the **1:ab28** structure (Figure 2). Precise modeling of Ag:aAb binding is essentially impossible, because immunogenicity of autoantigens differs across individuals, as does antigenicity and affinity of the polyclonal antibody population. It is perhaps appropriately inconvenient then that no structural information is available for **ab28** or for the **1:ab28** complex. A crystallographic structure of the p53 transactivation domain, however, has been reported (**I**, Figure 2, PDB ID: 2L14).[10] In this structure, the residues corresponding to the peptide epitope are approximately **α**-helical (**II**). We used this structure to model peptide **1** with HSymM (depicted as **III**, σ = 4 for all residues 46–55). As shown in **III** and in helical wheel format **V**, the epitope in the context of the protein is relatively cool, but in the isolated decapeptide (**IV** and **VI**) the differences are more extreme. The warmer residues in **V** tend to be hotter in **VI** (S46, D49, I50, W53, F54) and the cooler residues colder (P47, D48, E51, Q52, T55). The epitope was assumed to be helical upon binding. Each residue of the epitope was assigned four or five contacts to the antibody (**VII**), since for helical peptide:Ab complexes the paratope tends to form four to five contacts to each residue of the epitope (c.f. PDBIDs: 1MVU, 4HPO, 2AP2; see: SI).[11,12] We assumed that hotter residues were more important for binding. Hence, four contacts were assigned to residues calculated to have μ-values lower than the average in the WT protein and five antibody contacts were assigned for those calculated μ-values greater than the average (**VII**). Since no structural data are available for the Ab, we assigned the paratope contacts of residue *i* the average μ-value of residue i and nearby epitope residues (i.e. i, i + 3, i + 4, i + 7, i − 3, i − 4, i – 7, where relevant). This rationale is consistent with the view that Ab binding is a balance of maximizing both binding affinity and selectivity. Admittedly, it is unlikely that all residues in a helical peptide would make contacts with the antibody paratope. However, this model avoids the risk of neglecting critical epitope residues inherent to a partial binding surface (c.f. **VIII**). The assumptions that the epitope for all peptides in this study are approximately helical (σ = 4), that the whole-antigen binding surface depicted in **VII** is suitable, and that the μ-values of the paratope mirror the epitope seemed certain to add noise to the estimated binding energies and weaken the correlation of predicted and observed binding affinity. Nevertheless, our findings show that the assumptions of this minimalist model appear to be reasonable.

Within the HSymM framework, the 46–55 decapeptide segment stood out as a good candidate for engineered binding affinity enhancements. The key observation was that the average μ-values for the putative epitope in the context of the whole-protein (**III**/**V**) and as an

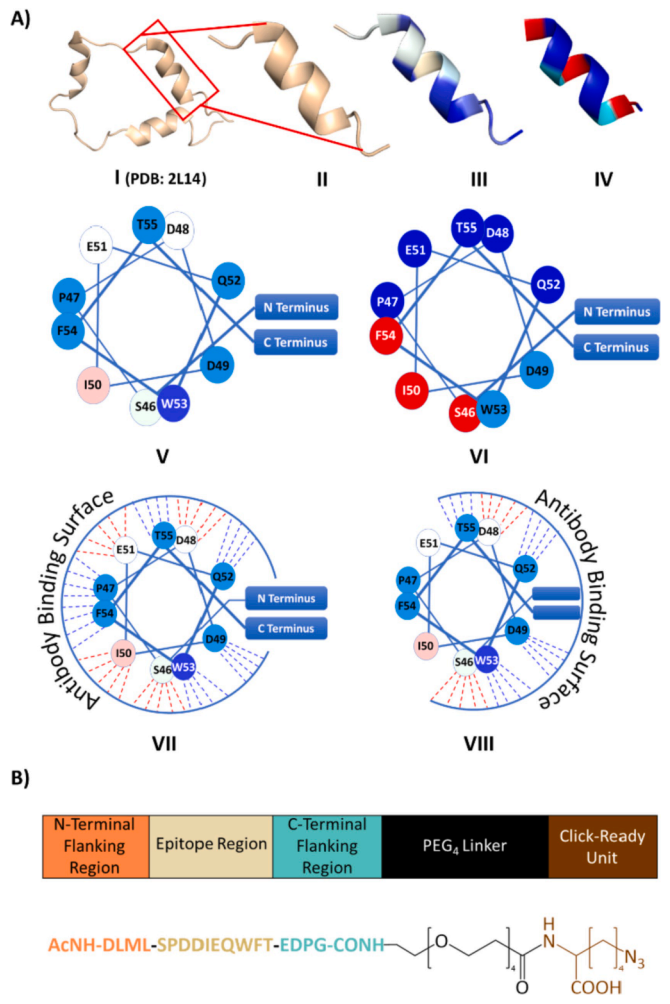

Figure 2. Peptide:Antibody Model and Construct Design. A. The PDB entry 2L14 (**I**) was used as the WT p53 reference structure for the target epitope, residues 46–55 (**II**), to generate heatmaps of the WT (**III**) and isolated peptide (**IV**). Helical wheel representations of **III** and **IV** are shown as **V** and **VI**, respectively. Paratope:epitope binding was modeled as **VII** and **VIII**. Heatmap colors indicate high (red), medium (white), and low (blue) per-residue contributions to interaction energy. B. Peptide constructs, composed of the putative epitope and C- and/or *N*-terminal flanking sequences, were attached to a polyethylene glycol (PEG_4) linker fused to an azido-lysine, click-ready unit. The construct shown matches **16**, which has the WT flanking sequences. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

isolated peptide (**IV**/**VI**) were both low (average μ = 0.135). We reasoned that careful selection of flanking residues could increase the μ-profile, and hence binding affinity, without compromising solubility or promoting aggregation. The μ-profile is an important consideration, because there is a ceiling effect inherent within the model that stems from the effective contribution (μ) being determined by normalizing intrinsic contributions (γ); the larger the number of residues averaged the lower the impact a flanking residue can have on the μ-value of a residue in the epitope. For this peptide, each μ-value is the average of at least three and as many as five γ-values. For example, an epitope residue with μ = 0.135 that makes 4 or 5 contacts to paratope residues of equal μ-value would contribute a calculated binding energy of about –1.0 kcal/mol. For a variant to increase the residue μ-value to 0.150, a flanking residue with a γ-value of about 0.200 would have to be used. This is below the 0.25 maximum of known γ-values and corresponds to leucine (see SI, Table S1). This variant would increase the epitope μ-value – but not the paratope μ-values – and would therefore be expected to enhance binding by about –0.15 kcal/mol. However, an increase in the epitope μ-value to μ = 0.165 would be impossible, since this would require a flanking reside with γ-value of at least 0.265. The low average μ-value of the WT antigen and peptide **1** suggested that significant improvements in binding affinity could be realized in variants of **1** that had short C- and the *N*-terminal flanking regions that included high γ-value amino acid residues.

With these considerations in mind, we designed a construct of the type shown in Figure 2B. A short polyethylene unit, which should help maintain solubility and serve as a spacer, was inserted between a click-ready portion and the peptide. The peptides were composed of the constant epitope region and the variable C- and *N*-terminal flanking regions. Accordingly, peptide **1** and 23 related constructs (**2**–**24**) were designed, prepared, assessed for binding to **ab28** by surface plasmon resonance (SPR), and the experimental binding data were compared to the calculated HSymM binding free energies (Tables 1-2). Other than in specifically designed cases, we avoided peptides with flanking regions dominated by non-polar residues, since these might otherwise render the construct prone to non-specific binding, aggregation, and low solubility. Polar flanking residues with low γ-values, such as K, R, and D, were placed in positions that would cool the hottest residues in the epitope, e.g. residues 46, 50, and 54, since these were calculated to be significantly hotter than the WT sequence. Non-polar residues with high γ-values, such as L, F and V, were positioned to warm the coolest residues in the epitope, e.g. residues 48 and 52, since these were calculated to be significantly colder than the WT sequence. In this way, flanking residues were combined in order to match or exceed the μ-values calculated for the WT protein. Flanking residues that matched the WT protein sequence and a limited number of permutations of the various flanking sequences were also studied. (Table S2 provides the full list of μ-values for the peptides in this study.)

The parent construct **1** bound **ab28** with good affinity (Table 1) and was compared to C-terminal modified variants **2**–**9**. Construct **2** contained the WT C-terminal flanking sequence, -EDPG, and bound more strongly than **1**. The calculated binding was expected to be weaker than **1**. This is a rare example in this set where prediction and experiment diverged. The permuted variant, -PGED (**3**), had a reduced binding affinity, as expected. Consistent with experiment, constructs **4**–**9** were expected to bind with greater affinity than **1**. These sequences were selected in part to be distinct from the WT flanking sequence, containing

Table 1
Binding of C-terminal peptide variants of **1**. Constructs with the peptide sequence shown were assessed by SPR to give the disassociation constants (K_D) from which the binding affinity relative to **1** and the corresponding change in binding free energy ($\Delta\Delta G_{Exp}$) were determined and compared to calculated free energy ($\Delta\Delta G_{Calc}$).

No.	Sequence	K_D [M]	Relative Affinity (25 °C)	$\Delta\Delta G_{Exp}$ (kcal/mol)	$\Delta\Delta G_{Calc}$ (kcal/mol)
1	SPDDIEQWFT	3.3E-06	–	–	–
2	SPDDIEQWFT-**EDPG**	2.1E-06	2	−0.3	+0.3
3	SPDDIEQWFT-**PGED**	3.5E-06	1	+0.1	+0.2
4	SPDDIEQWFT-**LLNR**	7.6E-07	4	−0.9	−0.4
5	SPDDIEQWFT-**LNRL**	1.4E-06	2	−0.5	−0.3
6	SPDDIEQWFT-**RNLL**	1.8E-06	2	−0.3	−0.2
7	SPDDIEQWFT-**VFNK**	7.4E-07	4	−0.9	−0.6
8	SPDDIEQWFT-**NVFK**	6.9E-07	5	−0.9	−0.5
9	SPDDIEQWFT-**KNFV**	1.7E-06	2	−0.4	−0.4

Table 2
Binding of N- and/or C-terminal peptide variants of **1**. Constructs with the peptide sequence shown were studied by SPR to give the disassociation constants (K_D) from which the binding affinity relative to **1** and corresponding change in binding free energy ($\Delta\Delta G_{Exp}$) were determined and compared to the calculated free energy based on the decapeptide (Decamer $\Delta\Delta G_{Calc}$) and to the calculated free energy based on the undecapeptide (Undecamer $\Delta\Delta G_{Calc}$) (see Experimental Methods in SI).

No.	Sequence	K_D [M]	Relative Affinity (25 °C)	$\Delta\Delta G_{Exp}$ (kcal/mol)	Decamer $\Delta\Delta G_{Calc}$ (kcal/mol)	Undecamer $\Delta\Delta G_{Calc}$ (kcal/mol)
1	SPDDIEQWFT	3.3E-06	–	–	–	–
10	**VFNK**-SPDDIEQWFT	6.8E-08	48	−2.3	−0.4	−1.3
11	**KFVN**-SPDDIEQWFT	7.2E-08	45	−2.2	−0.5	−1.6
12	**NKVF**-SPDDIEQWFT	1.6E-08	200	−3.1	−0.6	−2.0
13	**LLDM**-SPDDIEQWFT	2.7E-08	120	−2.8	−0.7	−2.1
14	**LMLD**-SPDDIEQWFT	3.5E-07	9	−1.3	−0.8	−1.7
15	**DLML**-SPDDIEQWFT	9.1E-09	360	−3.5	−0.8	−2.1
16	**DLML**-SPDDIEQWFT-**EDPG**	1.4E-08	240	−3.2	−0.4	−1.7
17	**KFVN**-SPDDIEQWFT-**ANRA**	1.0E-07	33	−2.1	−0.3	−1.4
18	**ARNA**-SPDDIEQWFT-**NVFK**	7.0E-08	47	−2.3	−0.3	−1.5
19	**RFKV**-SPDDIEQWFT-**AANN**	5.3E-08	62	−2.4	−0.3	−1.7
20	**L**SPDDIEQWFT	6.3E-09	520	−3.7	−0.3	−1.6
21	**KKNN-L**SPDDIEQWFT	8.7E-09	380	−3.5	+0.3	−0.8
22	**DDLM-L**SPDDIEQWFT	7.0E-09	470	−3.6	−0.7	−1.9
23	**YYEY-L**SPDDIEQWFT	7.7E-09	430	−3.6	−1.1	−2.5
24	**YYEY-L**SPDDIEQWFT-**YYEE**	1.3E-08	260	−3.3	−1.2	−2.7

positively charged residues K and R, instead of the negatively charged D and E, and canonical non-polar sidechains L, V, F instead of the more polar WT sequence. The inclusion of permutations was designed to reveal the potential spurious effects of non-epitope related binding to **ab28**. Indeed, the improved binding of **2** over **1**, in contrast to expectation, could be due to a stabilizing contact outside the epitope. These simple modifications changed the observed binding free energy across the range of +0.1 to –0.9 kcal/mol, which closely matched the range expected based on the model (calculated range: +0.3 to – 0.6 kcal/mol). Despite the minimalistic model and other simple approximations regarding the structure of these peptides, the antibody, and the Ab:Ag complex, the overall agreement between the calculated and experimentally observed binding for this set was very good ($R^2 = 0.78$).

As shown in Table 2, the parent construct (**1**) was also compared to *N*-terminal modified variants (**10**–**15**), combined N- and C-terminal variants (**16**–**19**), the undecapeptide corresponding to WT residues 45–55 (**20**), and four additional variants (**21**–**24**). Disappointingly, there was no meaningful agreement between HSymM predictions and experimental observation for the putative decapeptide epitope (compare $\Delta\Delta G_{Exp}$ with *Decamer* $\Delta\Delta G_{calc}$). The flanking residues of constructs **10**–**12** were permutations of the C-flanking variants **7**–**9**. The four residues used in the flanking sequence in constructs **13**–**15** were the same, in **13** and **14** the order was scrambled and in **15** the order matched the WT protein. We also examined **16**, which contains N- and C- terminal flanking residues that matched the WT sequence, as well as **17**–**19**, which were designed to serve as interesting permutations of the N- and C-flanking variants. Although one goal of the study was realized – use of a non-native sequence, e.g. **12**, markedly improved the binding affinity of this antigen over the putative epitope sequence (>200-fold) – closer examination of the data suggested epitope-specific binding in the *N*-terminus flanking sequence and that the putative epitope was wrongly assigned.

Three considerations support revision of the epitope. Firstly, the experimental improvement in affinity for peptides **10**–**15** corresponded to a binding energy of –1.3 to –3.5 kcal/mol, which were large changes consistent with new specific interactions between the protein and the peptide. Most of the observed values were significantly greater than our HSymM predicted flanking residue effects (*Decamer* $\Delta\Delta G_{calc}$, Table 2). We expected the stabilization attributable to a single epitope residue interacting with 4 or 5 residues of the paratope to be on the order of –1 to –2 kcal/mol, depending on whether the interacting residues are cool ($\mu = 0.135$) or warm ($\mu = 0.180$). The observed changes were more consistent with the combination of specific interactions and designed flanking effects. Secondly, the changes in binding qualitatively correspond to a non-polar hydrocarbon functionality being favored in the sidechain of residue 45. This residue position is underlined in constructs **10**–**19**. The long hydrocarbon in lysine in **10** gave an excellent improvement in binding compared to **1**, whereas shortening the methylene chain with retention of charge, as in **14**, reduced the binding improvement significantly. When this charge was removed but size and polar functionality was maintained, as in **11**, the affinity returned; the binding affinity of **10** and **11** were virtually identical. When nonpolar groups were placed in this position, the binding improved drastically (compare **12**, **13**, and **15**). In light of this, it is not surprising that the affinity of **16**, which corresponds to the WT sequence, far surpasses that of the putative epitope. Similarly, constructs **17**–**19**, which were designed to test ideas regarding the decapeptide epitope, exhibited binding affinity strongly dependent on the identity and nature of the residue in position 45 (underlined). Hence, for polar residues (**17**, c.f. **11**) the affinity was significantly increased compared to **1** but was much lower than that of **16**; for non-polar truncated or β-branched residues, as in **18** and **19**, the affinity was much greater than expected based on the decapeptide but was also inferior to **16**. Thirdly, and most compellingly, undecapeptide construct **20**, which corresponds to residues 45–55 of p53 TAD, had an affinity for **ab28** that was over 500 times that of the decapeptide.

The last group of peptide constructs we examined were variants of **20**. Constructs **21**–**23** had *N*-terminal flanking residues of different types. The KKNN sequence of **21** effected the fusion of polar sidechains, whereas the DDLM group of **22** corresponded to an exact facsimile of the WT p53 TAD. The final two peptide constructs **23** and **24**, were dominated by tyrosine residues. We return to the behavior of these constructs below.

The monoclonal antibody immunoassays on select peptide constructs corroborated the SPR measurements (Table 3). Our click-ready peptide constructs were conjugated to carboxylated magnetic microbeads (Luminex) that had been amidated with dibenzocyclooctyne (DBCO)-

Table 3
Immunoassays. LODs for selected immobilized constructs were determined in sandwich assays. The > 250-fold improvement in the LOD corresponds well with SPR binding affinity data.

Entry	Construct No.	Sequence	Antibody LOD (ng/mL)
1	1	SPDDIEQWFT	2.52
2	10	VFNK-SPDDIEQWFT	0.31
3	11	KFVN-SPDDIEQWFT	0.30
4	12	NKVF-SPDDIEQWFT	0.04
5	16	DLM-LSPDDIEQWFT-EDPG	0.06
6	20	LSPDDIEQWFT	0.01
7	22	DDLM-LSPDDIEQWFT	0.02
8	23	YYEY-LSPDDIEQWFT	0.02

PEG_4-amine (not shown, see SI for experimental details). Diluted fetal bovine sera were spiked with **ab28** standards (0–100 ng/ml). The functionalized beads were plated and assessed using a sandwich assay protocol with a BioRad immunoassay kit and BioPlex 200 system (Bio-Rad). Although designed constructs **10** and **11** constituted significant improvements in LOD over **1** (compare entries 1–3), construct **12** (entry 4) showed marked improvement. Similarly, constructs **16**, **20**, **22**, and **23**, each of which included L45 and showed very high affinity for **ab28** had the lowest LODs (entries 5–8). Thus, our non-native peptide constructs showed over 100-fold improvement in detection compared to **1**, and peptides that housed the WT leucine in position 45 had over 200-fold improvement over putative epitope **1**.

Despite our disappointment in finding that the epitope of the **ab28** antigen was misassigned, we examined the correlation of the experimental data with our calculations. Figure 3 shows the fit between the calculated binding energies of the putative decapeptide epitope and the binding data presented in Tables 1 and 2. The C-terminal flanking set (Table 1), as mentioned above, was well approximated by our minimalist model (Figure 3A, red line, $R^2 = 0.71$), but the full data set gave a poor fit ($R^2 = 0.24$). Fitting the data to the undecapeptide epitope (residues 45–55), however, gave a good fit between the experimental and our HSymM calculated data (Figure 3B, $R^2 = 0.76$). The outliers are readily understood, as well (e.g. **14** and **21**; Figure 3B, circled). Our model did not account for the destabilization that would be expected from replacing the neutral epitope residue, L45, with aspartate (**14**, see top circle in Figure 3B). The unexpected improvement in binding observed for **21** may reflect a serendipitous specific binding interaction between this construct and the antibody, but the degree to which it is an outlier also depends on our minimalistic model (see bottom circle in Figure 3B). For example, if a more realistic binding model is used (**VIII** in Figure 2, with the inclusion of residue L45) the fit is still very good ($R^2 = 0.74$) and **21** is not an extreme outlier (see Figure S3C in the SI).

In contrast to the 46–55 decapeptide segment, which stood out as a good candidate for engineered binding affinity enhancements because of its low average μ-value ($\mu = 135$), the undecapeptide epitope, i.e. **20**, has a significantly higher μ-value (average $\mu = 140$, Table S2). Increasing this average without potentially compromising the construct is challenging. For example, the flanking sequence YYEY, which represents an average γ-value of >185, only raises the average μ-value of the undecapeptide epitope to 152 (c.g. **23**). Moreover, these sidechains introduce serious potential liabilities in terms of solubility, aggregation, and non-specific binding. Although the SPR analysis gave well-behaved sensograms for all peptides (see SI for steady state and kinetic curve fitting data), the presence of the YYEY flanking residues or the *N*-flanking LLM triad, in any order, caused the constructs to deviate from a 1:1 peptide: Ab kinetic binding model at higher concentrations. This suggests that aggregation may occur at higher concentrations for some of the high affinity peptides. Indeed, potential aggregation may be reflected in the increased scatter in Figure 3B for the higher affinity peptides.

The correlation between our calculations and experiment is remarkable given the minimal structural information available to guide the design. The goodness of the fit enabled us to reexamine our all-residue antibody binding model (**VII**, Figure 2). We considered partial helix surfaces as potential epitopes similar to the mode depicted in **VIII**. Evaluation of epitopes of 4–7 residues instead of the full undecapeptide were evaluated (see Figure S3A-R). Surfaces that did not include residue L45 failed to provide a meaningful correlation ($R^2 < 0.14$), and although many surfaces that included L45 were able to adequately correlate the data, the mode that corresponded to **VIII** with L45 was found to be realistic and gave a good fit ($R^2 = 0.74$, Figure S3C). Other considerations, for example modeling the epitope as a single turn and excluding a certain residue from the fit (data not shown), or correlating the fit to helix-based considerations such as estimated percent helicity (Figure S5), did not provide meaningful correlations or improvements.[13–20] Taken together, the data strongly suggest that residues 45–55, as shown in **22**, constitute the correct epitope for **ab28**,

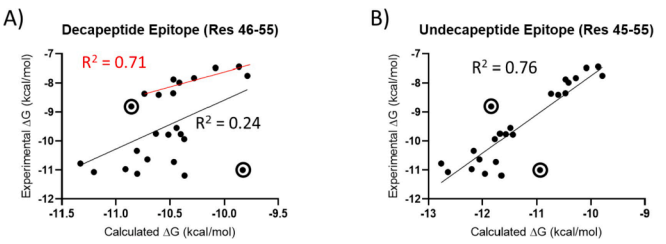

Figure 3. Comparison of Experimental and Calculated Affinities. **(A)** Calculated binding free energies based on the putative decapeptide epitope (residues 46–55) compared to experimentally determined values (peptides **1**–**9:** red line, $y = 0.95x + 1.80$, $R^2 = 0.71$; peptides **1**–**24**: black line, $y = 1.69x + 8.30$, $R^2 = 0.24$). **(B)** Calculated binding free energies based on the proposed undecapeptide epitope (residues 45–55) compared to experimentally determined values (peptides **1**–**24**: $y = 1.33x + 5.57$, $R^2 = 0.76$). See text comments on outliers **14** and **21** (circled). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

that HSymM modeling of flanking peptides, and the minimalist model served well to guide this peptide-protein binding study.

Along with proteins and small molecules, peptides have many biomedical applications, including as vaccines,[21,22] therapeutics,[23,24] scaffolds for tissue engineering,[25,26] drug-delivery carrier and targeting systems,[27,28] as well as in clinical and investigatory immunoassays.[3,29] The advantages that peptides offer over proteins in terms of ease of synthesis, potentially better properties, and lower risk of unwanted side effects make them ideal, provided they can be designed accordlingly.[23,30–32] Since most applications depend on the peptide binding to a protein target, the methods of phage display, microarray libraries, high-throughput screening,[33–35] as well as in silico docking[36,37] and all-atom simulations[38–40] can be used to design peptides with specific properties for a wide range of applications.[32,41] The coarse-grained protein energetics model we used complements these powerful strategies. HSymM is fast and computationally inexpensive. The findings of this study further suggest that HSymM is a useful tool for visualizing and approximating per-residue contributions to interaction energy.[42,43]

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Partial funding support was provided by the NIH grants T32 GM135141 (“Rutgers Biotechnology Training Program”) and 1S10OD026750-01 (“A Biacore 8 K System.”). The authors would also like to acknowledge Drs. Nilgun Tumer and Xiao-ping Li for providing the Biacore system and assisting with the SPR measurements, as well as Parinita Jain for her immunoassay work.

Appendix A. Supplementary data

See the SI for General Experimental Methods, explanations of HSymM parameters and equations, alternative peptide:antibody binding mode affinity correlations, and peptide characterization data (SPR sensograms, HPLC, MS).

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bmcl.2021.128341.

References

1 Soussi T, Wiman KG. TP53: an oncogene in disguise. *Cell Death Differ*. 2015;22(8):1239–1249.

2 Suppiah A, Greenman J. Clinical utility of anti-p53 auto-antibody: systematic review and focus on colorectal cancer. *World J Gastroenterol*. 2013;19(29):4651–4670.
3 Pedersen JW, Gentry-Maharaj A, Fourkala E-O, et al. Early detection of cancer in the general population: a blinded case-control study of p53 autoantibodies in colorectal cancer. *Br J Cancer*. 2013;108(1):107–114.
4 Macdonald IK, Parsy-Kowalska CB, Chapman CJ. Autoantibodies: Opportunities for Early Cancer Detection. *Trends in Cancer*. 2017;3(3):198–213.
5 Qin J, Zeng Ni, Yang T, et al. Diagnostic Value of Autoantibodies in Lung Cancer: a Systematic Review and Meta-Analysis. *Cell Physiol Biochem*. 2018;51(6):2631–2646.
6 Katchman BA, Barderas R, Alam R, et al. Proteomic mapping of p53 immunogenicity in pancreatic, ovarian, and breast cancers. *Proteomics. Clinical applications*. 2016;10 (7):720–731.
7 Soussi T. p53 Antibodies in the Sera of Patients with Various Types of Cancer: A Review. *Cancer Res*. 2000;60(7):1777.
8 Williams LJ, Schendt BJ, Fritz ZR, Attali Y, Lavroff RH, Yarmush ML. A protein interaction free energy model based on amino acid residue contributions: Assessment of point mutation stability of T4 lysozyme. *Technology (Singap World Sci)*. 2019;07 (01n02):12–39.
9 Moret MA, Zebende GF. Amino acid hydrophobicity and accessible surface area. *Phys Rev E*. 2007;75(1), 011920.
10 Lee CW, Martinez-Yamout MA, Dyson HJ, Wright PE. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry*. 2010;49(46):9964–9971. https://doi.org/10.1021/bi1012996.
11 Liao H-X, Bonsignori M, Alam SM, et al. Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2. *Immunity*. 2013;38(1):176–186.
12 van den Elsen JMH, Kuntz DA, Hoedemaeker FJ, Rose DR. Antibody C219 recognizes an alpha-helical epitope on P-glycoprotein. *PNAS*. 1999;96(24):13679–13684.
13 Singh, H.; Singh, S.; Singh Raghava, G. P., Peptide Secondary Structure Prediction using Evolutionary Information. bioRxiv 2019, 558791.
14 Muñoz V, Serrano L. Elucidating the folding problem of helical peptides using empirical parameters. *Nat Struct Biol*. 1994;1(6):399–409.
15 Muñoz V, Serrano L. Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J Mol Biol*. 1995;245(3):275–296.
16 Muñoz V, Serrano L. Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *J Mol Biol*. 1995;245(3): 297–308.
17 Lamiable A, Thévenet P, Rey J, Vavrusa M, Derreumaux P, Tufféry P. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res*. 2016;44(W1):W449–W454.
18 Shen Y, Maupetit J, Derreumaux P, Tufféry P. Improved PEP-FOLD Approach for Peptide and Miniprotein Structure Prediction. *J Chem Theory Comput*. 2014;10(10): 4745–4758.
19 Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic acids research*. 2012;40(Web Server issue):W288–W293.
20 Yan R, Xu D, Yang J, Walker S, Zhang Y. A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction. *Sci Rep*. 2013;3(1):2619.
21 Malonis RJ, Lai JR, Vergnolle O. Peptide-Based Vaccines: Current Progress and Future Challenges. *Chem Rev*. 2020;120(6):3210–3229.
22 Hos BJ, Tondini E, van Kasteren SI, Ossendorp F. Approaches to Improve Chemically Defined Synthetic Peptide Vaccines. *Front Immunol*. 2018;9. https://doi.org/10.3389/fimmu.2018.00884.
23 Marqus S, Pirogova E, Piva TJ. Evaluation of the use of therapeutic peptides for cancer treatment. *J Biomed Sci*. 2017;24(1):21.
24 Lau JL, Dunn MK. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg Med Chem*. 2018;26(10): 2700–2707.
25 Chow D, et al. Peptide-based Biopolymers in Biomedicine and Biotechnology. Materials science & engineering. *R, Reports : a review journal*. 2008;62(4):125–155.
26 Altunbas A, Pochan DJ, Peptide-Based and Polypeptide-Based Hydrogels for Drug Delivery and Tissue Engineering, in Peptide-Based Materials, T. Deming, Editor. 2012, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 135-167.
27 Eskandari S, Guerin T, Toth I, Stephenson RJ. Recent advances in self-assembled peptides: Implications for targeted drug delivery and vaccine engineering. *Adv Drug Deliv Rev*. 2017;110-111:169–187.
28 Liu T, Qiao JX, Poss MA, Yu J-Q. Palladium(II)-Catalyzed Site-Selective C(sp3)−H Alkynylation of Oligopeptides: A Linchpin Approach for Oligopeptide-Drug Conjugation. *Angew Chem Int Ed*. 2017;56(36):10924–10927.
29 Sanchez AB, Nguyen T, Dema-Ala R, Kummel AC, Kipps TJ, Messmer BT. A general process for the development of peptide-based immunoassays for monoclonal antibodies. *Cancer Chemother Pharmacol*. 2010;66(5):919–925.
30 Guidotti G, Brambilla L, Rossi D. Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends Pharmacol Sci*. 2017;38(4):406–424.
31 Wuo MG, Hong SH, Singh A, Arora PS. Synthetic Control of Tertiary Helical Structures in Short Peptides. *J Am Chem Soc*. 2018;140(47):16284–16290.
32 Kheirabadi M, Creech GS, Qiao JX, et al. Leveraging a "Catch–Release" Logic Gate Process for the Synthesis and Nonchromatographic Purification of Thioether- or Amine-Bridged Macrocyclic Peptides. *The Journal of Organic Chemistry*. 2018;83(8): 4323–4335.
33 Vanhee P, van der Sloot AM, Verschueren E, Serrano L, Rousseau F, Schymkowitz J. Computational design of peptide ligands. *Trends Biotechnol*. 2011;29(5):231–239.
34 Wu C-H, Liu I-J, Lu R-M, Wu H-C. Advancement and applications of peptide phage display technology in biomedical science. *J Biomed Sci*. 2016;23(1). https://doi.org/10.1186/s12929-016-0223-x.
35 Ozgul S, von Daake S, Kakehi S, et al. An ELISA-Based Screening Platform for Ligand-Receptor Discovery. *Methods Enzymol*. 2019;615:453–475.
36 Pirkhezranian Z, Tahmoorespur M, Monhemi H, Sekhavati MH. Computational Peptide Engineering Approach for Selection the Best Engendered Camel Lactoferrin-Derive Peptide with Potency to Interact with DNA. *Int J Pept Res Ther*. 2020;26(4): 2203–2212.
37 Zanuy D, Sayago FJ, Revilla-López G, et al. Engineering strategy to improve peptide analogs: from structure-based computational design to tumor homing. *J Comput Aided Mol Des*. 2013;27(1):31–43.
38 Hospital A, et al. Molecular dynamics simulations: advances and applications. *AABC*. 2015;8:37–47.
39 Kmiecik S, Kouza M, Badaczewska-Dawid A, Kloczkowski A, Kolinski A. Modeling of Protein Structural Flexibility and Large-Scale Dynamics: Coarse-Grained Simulations and Elastic Network Models. *Int J Mol Sci*. 2018;19(11):3496. https://doi.org/10.3390/ijms19113496.
40 Introduction to molecular simulation in Quantitative Biology: Computational Methods and Examples, ed. B. Munsky, W.S. Hlavacek and L.S. Tsimring, MIT Press (2018) P. Šulc, J.P.K. Doye and A. A. Louis.
41 Abe K, Kobayashi N, Sode K, Ikebukuro K. Peptide ligand screening of α-synuclein aggregation modulators by in silico panning. *BMC Bioinf*. 2007;8(1):451. https://doi.org/10.1186/1471-2105-8-451.
42 Gee MH, Sibener LV, Birnbaum ME, et al. Stress-testing the relationship between T cell receptor/peptide-MHC affinity and cross-reactivity using peptide velcro. *PNAS*. 2018;115(31):E7369–E7378.
43 Gorelik M, Davidson AR. Distinct peptide binding specificities of Src homology 3 (SH3) protein domains can be determined by modulation of local energetics across the binding interface. *J Biological Chem*. 2012;287(12):9168–9177.