Protein Design

DOI: 10.1002/ange.201305963

Computational Design of Helical Peptides Targeting TNFα**

*Changsheng Zhang, Qi Shen, Bo Tang, and Luhua Lai**

Computationally designing novel proteins that can bind to a specific target protein is promising yet challenging, as binding and folding need to be considered at the same time.[1] Regular secondary structures, especially a single helix, which is often involved in native protein–protein interactions,[2] serve as an ideal folding motif for designing novel binds from scratch. Though examples for designing single-helix peptides for a specific protein target have been reported,[3] a general applicable approach is not yet established. We have developed a general computational strategy for de novo design of helical peptides that can bind to a given target. This strategy involves two steps: automated binding site detection and sequence design.

We used this strategy to design peptide inhibitors of tumor necrosis factor-α (TNFα), an established therapeutic target for rheumatoid arthritis and other autoimmune diseases.[4] Smaller, less-expensive peptide drugs may be better alternatives to clinically used monoclonal antibodies and engineered soluble TNF receptors. Since TNFα exerts its biological effects by binding to cell-surface TNF receptors (TNFR),[5] peptides were designed to target the TNFR binding site on TNFα to block its signaling.

TNFα functions as a trimer with three TNFR binding sites (Figure 1a). All the sites between the three monomers (A, B, C) in the two TNFα structures (Protein Data Bank (PDB) code: 1TNF, 1A8M)[6] were used for binding-peptide design. Initially, a 28.6 Å line, representing the axis of a 20-residue helix, was used as the detector, which was systematically translated and rotated on the target surface (Figure 1b). The optimal positions of the detector defined by our scoring functions were clustered and integrated (see Supporting Information) into a 25-residue long helix at two positions (Figure 1c). Helical peptides were designed targeting these two positions.

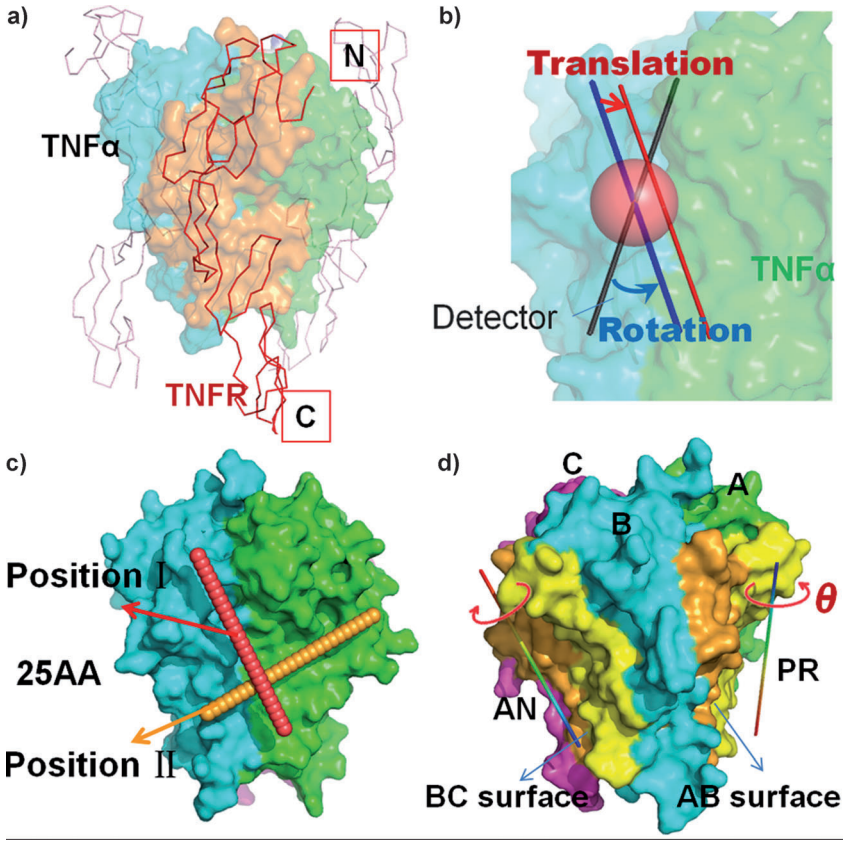

Figure 1. Identification of helix-binding positions on the TNFα surface. a) Model of the TNFα-TNFR1 structure based on the TNFβ-TNFR1 structure (PDB code: 1TNR).[5] The interaction surface of TNFα is shown in orange. b) Illustration of how the detector was systematically translated and rotated on the target surface. c) Two positions for binding 25-residue helical peptides. d) The optimized detector positions and the two opposing directions, PR (pro the binding direction of TNFR, see panel (a)) and AN (anti), adopted by the helix for position I. N-termini blue and C-termini red. θ is the rotation angle for sampling.

Polyalanine sequence was used first to refine the location of the 25-residue long helix. Figure 1d shows the optimization result for position I. The two reversed directions with alternate N- and C-termini and rotation angles of 0°, 60°, 120°, 180°, 240°, or 300° were sampled. The sequences and binding conformations of the 25-residue helical peptide were then computationally generated from the initial polyalanine

[*] Dr. C. Zhang,[+] Prof. L. Lai
BNLMS, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, and Peking-Tsinghua Center for Life Sciences at College of Chemistry and Molecular Engineering, Peking University Beijing 100871 (China)
E-mail: lhlai@pku.edu.cn
Q. Shen,[+] B. Tang, Prof. L. Lai
Center for Quantitative Biology, Peking University (China)
[+] These authors contributed equally to this work.

[**] This work was supported in part by the Ministry of Science and Technology of China (2009CB918500) and the National Natural Science Foundation of China (11021463, 90913021). TNFα = tumor necrosis factor-α.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201305963.

sequence. To consider backbone flexibility, the loop design module of Rosetta (version 2.3.0)[7] was used. The helical peptide was disguised as the C-terminal "loop" of TNFα. The TNFα backbone was fixed while the side chain conformations at the interface were optimized. The best sequences and binding conformations of the helix were generated using the Monte Carlo sampling strategy. One hundred trajectories were performed from each initial binding conformation. In total, we obtained 2 (positions) × 6 (target surfaces) × 2 (directions) × 6 (rotations) × 100 = 14400 solutions. The resultant helices were usually shifted or bent slightly relative to the initial conformation (Supporting Information, Figure S2). Information entropy[8] for sequences generated from a specific initial conformation indicated that Rosetta can produce sequences with appropriate diversity (Figure S3).

The folding probability of a designed sequence was estimated using FoldIndex.[9] AggreScan[10] was used to estimate the aggregation probability. The average residue helical propensity was calculated to estimate whether a designed sequence can fold into a helix. Rosetta-calculated energy was used to estimate the stability of the complex and the binding affinity. Peptide sequences that demonstrated poor Rosetta energy, low folding probability, high aggregation probability, low helix propensity, poor hydrophobic packing, or small contact size were eliminated according to the respective scoring parameter cutoff (see Supporting Information). The remaining peptides were sorted using a hybrid scoring function containing all the above parameters. The top 50 position-I and top 20 position-II sequences (Tables S2 and S3) were manually selected using three additional criteria. First, sequences derived from the same initial binding conformation were aligned and analyzed. Conserved residues were considered important for binding. Second, selected peptides should have maximum interaction with hot spots on the TNFα surface for TNFR binding, including R31, R32, Y87, Y115, D143, and F144.[11] Third, complementarities of electrostatic potentials were necessary.

Six position-I sequences and two position-II sequences were selected (Tables S2 and S3). The binding models of these peptides are shown in Figure S4. To experimentally screen for sequences that can bind TNFα and to minimize the effect of possible peptide aggregation, each of the eight selected sequences was expressed in *E. coli* as GST fusion proteins and then subjected to a surface plasmon resonance (SPR) binding assay. In brief, TNFα was immobilized onto a CM-5 sensor chip, and GST fusion proteins were passed through the flow cells as analytes. One position-I sequence TBHa31 and one position-II sequence TBHp03 showed significant binding with TNFα (Figure 2).

Next, label-free TBHa31 and TBHp03 peptides were chemically synthesized for further study. Their circular dichroism (CD) spectra indicated typical α-helical secondary structure (Figure 2c). An luciferase assay was performed to measure the biological activity within cellular environment. In transfected cells, TNFα induced *NF-κB* activation through TNFR, which resulted in luciferase activity. The luminescence signal decreased with increasing concentrations of TBHa31 or TBHp03 (Figure 2d), demonstrating that both peptides inhibited TNFα function in the cells. By fitting the signal-

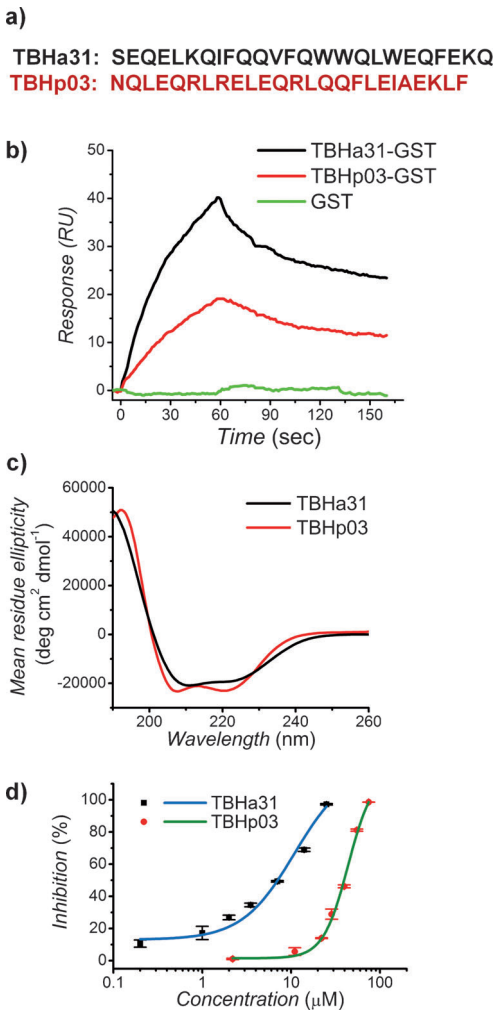

Figure 2. Properties of the two helical peptides, TBHa31 and TBHp03, designed to bind to position I and II in TNFα, respectively. a) Peptide sequences. b) SPR binding curve of GST fusion proteins with immobilized TNFα. c) CD spectra. d) Dose-response curves of cellular luciferase assays. Experimental data were fitted using the Hill model.

concentration curve using the Hill model, the half maximal inhibitory concentration (IC_{50}) values of TBHa31 and TBHp03 were 10.9 ± 0.5 μM and 43.9 ± 2.2 μM, respectively. We selected the peptide with high activity, TBHa31, for further experimental characterization.

The binding affinity of TBHa31 with TNFα was measured by SPR with immobilized TNFα and TBHa31 as analytes. The binding K_d was measured to be 1.97 ± 0.22 μM, fitted using the steady state model (Figure 3a and b). The in vitro activity of TBHa31 was further confirmed by a SPR competitive binding assay. A pre-incubated mixture of TNFα and TBHa31 was injected over the sensor chip on which the extracellular domain of TNFR1 was immobilized. TNFα and TNFR1 binding was effectively blocked by incubating TBHa31 with TNFα (Figure 3c). The IC_{50} value measured from the dose-response curve (Figure 3d) was 0.65 ± 0.01 μM.

We checked whether TBHa31 interacts with TNFα as designed. From the TNFα-TBHa31 complex model, key interaction sites were identified as: salt bridges (E4-R138,

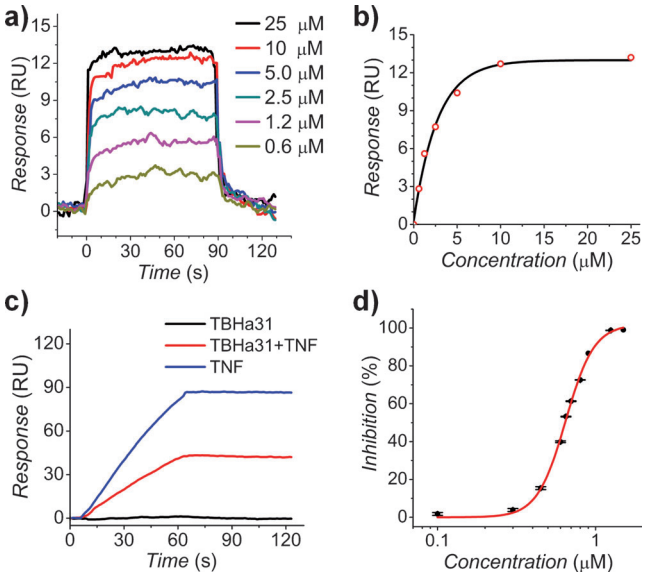

Figure 3. SPR direct and competitive binding assay results of TBHa31. a) SPR direct binding curves with increasing TBHa31 concentrations with TNFα immobilized on the sensor chip for measuring the TBHa31-TNFα binding affinity. b) Dose-response curve of TBHa31 binding with immobilized TNFα. c) SPR competitive binding curves. TNFR1 (extracellular domain) was immobilized on the sensor chip. Pre-incubation of TBHa31 with TNFα effectively blocked the binding between TNF-α and TNFR1. No binding signal was observed when only TBHa31 was applied. d) Dose-response curve of SPR competitive binding.

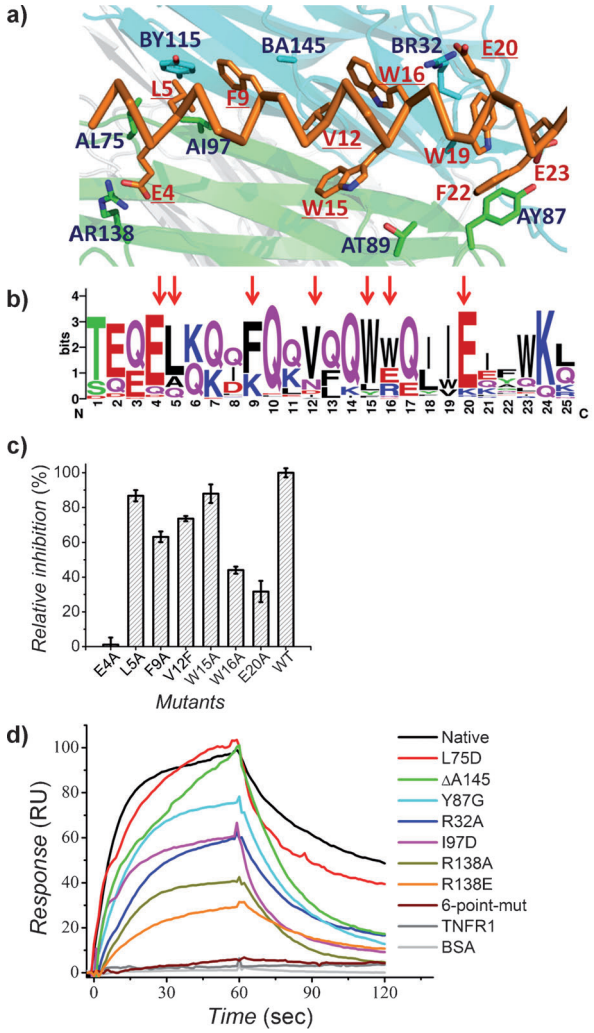

Figure 4. Analysis of TBHa31-TNFα interaction and binding site verification. a) Interaction analysis. TNFα chain A green, chain B cyan, TBHa31 orange. Conserved residues of TBHa31 are underlined. b) Multiple sequence alignment of TBHa31 sequences from the same original is represented using weblogo.[12] Conserved interface residues are indicated by red arrows. c) Inhibition of TNFα activity, as measured by luciferase assay, following site-specific mutagenesis of TBHa31 relative to wild-type TBHa31. d) SPR direct binding curves of TNFα mutants and those of native TNFα. The GST-TBHa31 fusion peptide was immobilized on the sensor chip. Δ145A was a deletion mutant. The 6-point mutant was R32A-L75A-Y87A-I97A-Y115A-A145D. The proteins were injected over the chip at a fixed concentration of 5 μM. TNFR1 and BSA were used as references.

E20-R32), hydrogen bonds (E23-Y87, W15-T89), and hydrophobic packing ([F9 L5]-[L75,I97, Y115], F22-Y87; Figure 4a). Mutation studies both on the helix and TNFα binding surface were conducted. The most highly conserved sites of TBHa31 (indicated in Figure 4b), were selected for mutagenesis studies. V12 was mutated to phenylalanine and the other sites were all mutated to alanine. All the seven mutants, showed diminished inhibitory activity relative to TBHa31 in luciferase assays (Figure 4c). In the SPR competitive binding assays, mutants E4A, V12F, F9A almost lost activity completely (Figure S5). Then, we measured the direct binding between TBHa31 and TNFα mutants (shown in Figure 4d legend) using a sensor chip on which the GST-TBHa31 fusion peptide was immobilized. Compared to native TNFα, all its mutants displayed reduced binding to TBHa31 (Figure 4d) and the 6-point mutant showed no response. The behavior of these mutants implies that TBHa31 interacts with TNFα as designed.

As helical peptides may aggregate in solution, we also checked the aggregation state of TBHa31. Gel filtration and analytical ultracentrifugation experiments indicated that TBHa31 exists mainly as an octamer with a small amount of monomer in solution (Figure S6). The binding of monomeric TBHa31 with TNFα may shift the equilibrium toward the monomer. Consistent with this, pre-incubation of TBHa31 with TNFα significantly enhances its inhibitory effect.

Using our de novo helical-binder design method, we successfully obtained an helical peptide TBHa31 that binds to TNFα and inhibits its activity. Though the binding affinity of TBHa31 was moderate (micromol), it is comparable to or better than other de novo designed protein-protein interactions at the computational design stage.[1,13] In fact, many therapeutic peptides, for example, the HIV gp41 targeting peptide, enfurvitide,[14] function with micromol activity. The binding affinity can be further optimized using various methods, such as homolog shotgun scanning by phage display.[15]

The de novo helical-binder design method described herein can be generally applied to identify helix binding sites on target proteins and for designing binding peptide sequences de novo. It can also be used as a general approach to

design peptide inhibitors or allosteric regulators targeting the novel sites discovered.

Received: July 9, 2013
Published online: August 26, 2013

Keywords: computational design · helical peptides · protein/peptide drug · protein–protein interactions · TNFα

[1] a) S. J. Fleishman, T. A. Whitehead, D. C. Ekiert, C. Dreyfus, J. E. Corn, E. M. Strauch, I. A. Wilson, D. Baker, *Science* **2011**, *332*, 816 – 821; b) C. Zhang, L. Lai, *Biochem. Soc. Trans.* **2011**, *39*, 1382 – 1386.

[2] B. N. Bullock, A. L. Jochim, P. S. Arora, *J. Am. Chem. Soc.* **2011**, *133*, 14220 – 14223.

[3] a) M. L. Bellows, M. S. Taylor, P. A. Cole, L. Shen, R. F. Siliciano, H. K. Fung, C. A. Floudas, *Biophys. J.* **2010**, *99*, 3445 – 3453; b) G. Grigoryan, A. W. Reinke, A. E. Keating, *Nature* **2010**, *468*, 859 – 864; c) H. Yin, J. S. Slusky, B. W. Berger, R. S. Walters, G. Vilaire, R. I. Litvinov, J. D. Lear, G. A. Caputo, J. S. Bennett, W. F. DeGrado, *Science* **2007**, *315*, 1817 – 1822; d) D. W. Sammond, D. E. Bosch, G. L. Butterfoss, C. Purbeck, M. Machius, D. P. Siderovski, B. Kuhlman, *J. Am. Chem. Soc.* **2011**, *133*, 4190 – 4192.

[4] a) M. Feldmann, L. Steinman, *Nature* **2005**, *435*, 612 – 619; b) M. Feldmann, R. N. Maini, *Nat. Med.* **2003**, *9*, 1245 – 1250.

[5] D. W. Banner, A. D'Arcy, W. Janes, R. Gentz, H.-J. Schoenfeld, C. Broger, H. Loetsche, W. Lesslauer, *Cell* **1993**, *73*, 431 – 445.

[6] a) M. J. Eck, S. R. Sprang, *J. Biol. Chem.* **1989**, *264*, 17595 – 17605; b) C. Reed, Z.-Q. Fu, J. Wu, Y.-N. Xue, R. W. Harrison, M.-J. Chen, I. T. Weber, *Protein Eng.* **1997**, *10*, 1101 – 1107.

[7] C. A. Rohl, C. E. M. Strauss, D. Chivian, D. Baker, *Proteins Struct. Funct. Bioinf.* **2004**, *55*, 656 – 677.

[8] P. S. Shenkin, B. Erman, L. D. Mastrandrea, *Proteins Struct. Funct. Genet.* **1991**, *11*, 297 – 313.

[9] J. Prilusky, C. E. Felder, T. Zeev-Ben-Mordehai, E. H. Rydberg, O. Man, J. S. Beckmann, I. Silman, J. L. Sussman, *Bioinformatics* **2005**, *21*, 3435 – 3438.

[10] O. Conchillo-Solé, N. S. de Groot, F. X. Avilés, J. Vendrell, X. Daura, S. Ventura, *BMC Bioinformatics* **2007**, DOI: 10.1186/1471-2105-8-65.

[11] a) X.-M. Zhang, I. Weber, M.-J. Chen, *J. Biol. Chem.* **1992**, *267*, 24069 – 24075; b) H. Loetscher, D. Stueber, D. Banner, F. Mackay, W. Lesslauer, *J. Biol. Chem.* **1993**, *268*, 26350 – 26357.

[12] G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, *Genome Res.* **2004**, *14*, 1188 – 1190.

[13] J. Karanicolas, J. E. Corn, I. Chen, L. A. Joachimiak, O. Dym, S. H. Peck, S. Albeck, T. Unger, W. Hu, G. Liu, S. Delbecq, G. T. Montelione, C. P. Spiegel, D. R. Liu, D. Baker, *Mol. Cell* **2011**, *42*, 250.

[14] M. Kilby, S. Hopkins, T. M. Venetta, B. DiMassimo, G. A. Cloud, J. Y. Lee, L. Alldredge, E. Hunter, D. Lambert, D. Bolognesi, T. Mathews, M. R. Johnson, M. A. Nowak, G. M. Shaw, M. S. Saag, *Nat. Med.* **1998**, *4*, 1302.

[15] K. Murase, K. L. Morrison, P. Y. Tam, R. L. Stafford, F. Jurnak, G. A. Weiss, *Chem. Biol.* **2003**, *10*, 161 – 168.