PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

A simple contact mapping algorithm for identifying potential peptide mimetics in protein–protein interaction partners

Alex Krall,[1] Jonathan Brunn,[2] Spandana Kankanala,[2] and Michael H. Peters[2,3,4]*

[1] Department of Computer Science, Virginia Commonwealth University, Richmond, Virginia 23284

[2] Department of Chemical and Life Science Engineering, Virginia Commonwealth University, Richmond, Virginia 23284

[3] Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, Virginia 23284

[4] Massey Cancer Center, Virginia Commonwealth University, Richmond, Virginia 23298

ABSTRACT

A simple, static contact mapping algorithm has been developed as a first step at identifying potential peptide biomimetics from protein interaction partner structure files. This rapid and simple mapping algorithm, "OpenContact" provides screened or parsed protein interaction files based on specified criteria for interatomic separation distances and interatomic potential interactions. The algorithm, which uses all-atom Amber03 force field models, was blindly tested on several unrelated cases from the literature where potential peptide mimetics have been experimentally developed to varying degrees of success. In all cases, the screening algorithm efficiently predicted proposed or potential peptide biomimetics, or close variations thereof, and provided complete atom-atom interaction data necessary for further detailed analysis and drug development. In addition, we used the static parsing/mapping method to develop a peptide mimetic to the cancer protein target, epidermal growth factor receptor. In this case, secondary, loop structure for the peptide was indicated from the intra-protein mapping, and the peptide was subsequently synthesized and shown to exhibit successful binding to the target protein. The case studies, which all involved experimental peptide drug advancement, illustrate many of the challenges associated with the development of peptide biomimetics, in general

Proteins 2014; 82:2253–2262.

Key words: protein–protein interactions; peptide mimetics; static mapping.

INTRODUCTION

Protein–protein interactions (PPIs) that underlie cell signal pathways are a promising avenue for the development of inhibitory compounds aimed at disrupting cellular networks. Specifically, small peptide inhibitors that mimic protein segments associated with specific residue-residue interactions (peptidyl-biomimetics) in native, cellular processes have shown success as therapeutic agents.[1] However, there are many different facets of peptide biomimetics that must be considered before moving forward with any potential drug candidate. These facets include, for example, extracellular versus intracellular proteins, agonist versus antagonist actions, solubility, specificity, residue length, nature of binding forces and strength, and structural stability and dynamics.[2]

In general, intra-protein and inter-protein interactions are characterized by regions or segments of strong atomic charge and/or van der Waals interactions associated with specific residues of the parent proteins. (Also referred to as "hot spots" or "hot segments";[1] see Fig. 1) For example, regions or segments where relatively strong van der Waals forces dominate are typically associated with hydrophobic protein interactions, whereas regions of strong atomic charge and partial atomic charge interactions are associated with salt bridges, hydrogen bonding, and polar interactions. Now, these regions or segments can potentially function in isolation as peptide mimetics

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Michael H. Peters, Department of Chemical and Life Science Engineering, Virginia Commonwealth University, 601 West Main St, Richmond, VA 23284. E-mail: mpeters@vcu.edu

Received 15 January 2014; Revised 7 April 2014; Accepted 15 April 2014
Published online 22 April 2014 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24592

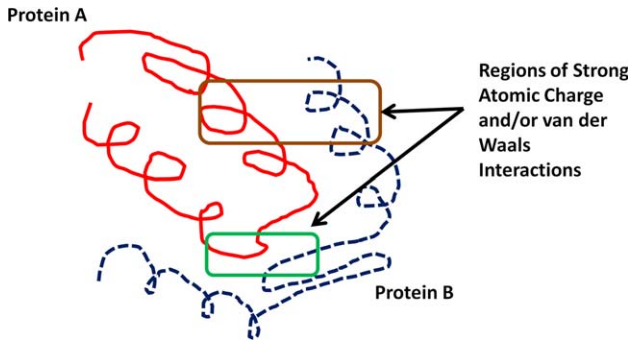

Figure 1

Overview of the contact mapping strategy for peptidyl biomimetics from protein interaction partners. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

that may block or inhibit the formation of their parent protein complexes. Thus, most peptide mimetics function as antagonistic drugs in application. Clearly, there are questions about any such region's or segment's effectiveness in isolation, including the maintenance of structure, strength of binding, solubility, and so forth, but nonetheless a simple, static mapping of the interactions could provide a rapid starting point for more detailed studies.

Recently, London *et al.*[3] have reviewed the identification of "hot spots" or "hot segments" of PPIs and potential peptide mimetics. In the majority of cases, specific residues or sequences of residues from the parent proteins could be identified. Many of the current methods of peptide identification are based on different schemes of docking calculations or Free Energy calculations.[1,3] In this study, assuming that the protein partner interaction structure information is known, we describe a very simple, transparent, and rapid static contact mapping algorithm for parsing out key binding areas along with their associated residues. These segments or regions can then be further explored as potential peptide mimetics via more lengthy dynamic algorithms, such as molecular dynamics or free energy, docking algorithms, and experimental methods. In addition, the parsing provides the specific type and magnitude of the atom–atom interaction forces of the associated residues based on all-atom interaction force field models, which has uses beyond peptidyl biomimetics including mutation analysis as illustrated in one case below. In this study, the algorithm is specifically tested against several reported, but unrelated, experimental peptide biomimetic development examples from the literature. In addition, we give our own example of experimental peptide development for anti-cancer therapy that involves the successful creation of secondary, loop peptide structure with experimentally demonstrated binding to the target protein. The examples given here also broadly illustrate some of the challenges and pitfalls associated with peptide biomimetic development in general. For completeness, we note that this method is restricted to the development of potential peptide biomimetics from known protein interaction partner structure files and does not necessarily include the identification of all "druggable hot spots"[4,5] or ligand identification through peptide data mining methods[6] or docking methods involving peptide libraries and small bioactive molecules.[4,7,8]

MATERIALS AND METHODS

Computational

The parsing and contact mapping program is diagrammed in Figure 2. The external *.pdb file or equivalent structure file is supplied by the user. Interaction partners (Protein A and Protein B) are specified by the user through identification of the Chain ID in the

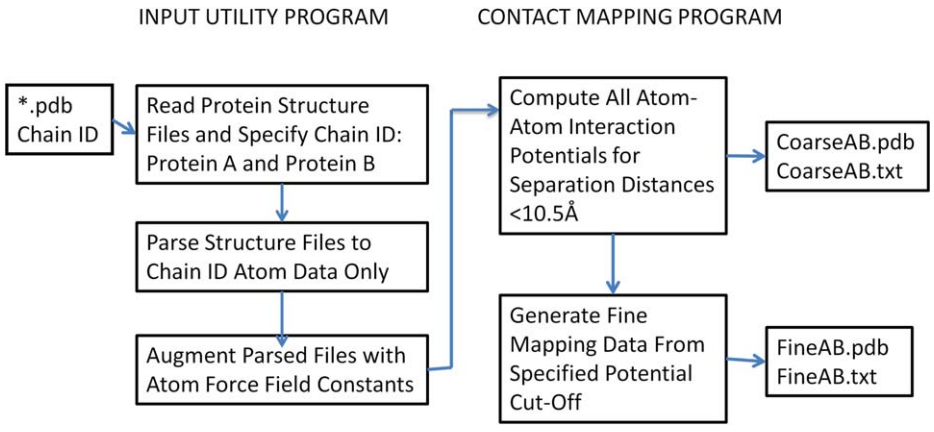

Figure 2

Flowchart for the parsing and contact mapping program: OpenContact. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

particular protein data bank file(s). As illustrated below, interactions between two chains of a single protein may be specified or two chains of a multiple protein interaction complex. Amber03[9] supplies partial atomic charges and Lennard-Jones parameters (van der Waals and Born force constants) for all protein atoms. In the Coulombic potential calculations given here, we use a distance dependent solvent dielectric of sigmoidal shape ($D = 0.2$ Å^{-1}) following the work of Ramstein and Lavery[10] as reviewed by Smith and Pettitt.[11] In this model, the dielectric constant ranges from unity at zero separation distance between atoms to its bulk value at $\sim$ 15 Å and, therefore, approximately captures solvent effects for a wide range of contact distances. Apolar implicit solvent forces have not been included in the results given here. For "coarse parsing" of the interactions, atom–atom separation distances greater than 10.5 Å are excluded. In the "fine parsing" of the interactions, an additional restriction on the atom–atom interaction potentials are specified. For all "fine parsing" results given below, we have selected upper limits to restrict Coulombic or Lennard Jones interactions as

$$U^*_{\mathrm{Coul}} \leq -0.3$$

or

$$U^*_{\mathrm{LJ}} \leq -0.1$$

respectively, where $U^* = U/kT$ with k being Boltzmann's constant and T is temperature (taken as 310.15 K in all results given here). These optimized, empirical cut-offs, the efficiency of which are demonstrated in case studies below, ensure that only the strongest attractive interaction potentials, for a given type, are included in the fine parsing output results. Because of the partial atomic charges assigned by all atom force filed models there are typically hundreds of attractive Coulombic interactions in the dimensionless range of $(-0.2,0)$; similar arguments apply to the Lennard Jones attractive term. However, we note that the fine parsing criteria given above can also be changed by the user via the coarse parsing data.

In addition, and following Amber03 modeling,[9] any atom–atom overlaps are defined as separation distances less than 0.9 times the average pair molecular diameter. In those cases, the potentials are computed at this minimally selected distance. The final coarse and fine parsing results are provided to the user in a highly manageable and friendly format in both *.pdb and *.txt (text files) for spread sheeting, plotting, and more detailed interaction analysis, as illustrated in the case studies below (see Supporting Information).

Experimental—EGFR/P28 kinetic binding measurements

The peptide mimetic, P28, was synthesized using Liberty Automated Microwave Peptide Synthesizer from CEM Corporation. Purification (>95%) was performed through High Performance Liquid Chromatography using a reverse phase column. For disulfide bond formation, the peptide sample, P28, was dissolved in 0.01*M* ammonium bicarbonate buffer (pH 8) at a concentration of 0.1 mg/mL and the solution was left to stir in open atmosphere. The progress of the reaction was monitored by analytical HPLC (peak shifted after disulfide bond formation). After the reaction was complete, the peptide solution was purified and checked for mass (loss of 2 protons) using mass spectrometry.

The binding kinetics of growth factor and peptides were measured on a Biacore 3000 instrument (GE, New Jersey) using HBS-EP as running buffer (10 m*M* HEPES, pH 7.4, 150 m*M* NaCl, 3 m*M* EDTA, and 0.15% surfactant P20). EGFR-Fc (R&D) was immobilized to CM5 sensor chip surface using standard amine-coupling chemistry by injecting the following reagents (45 μL) at a flow rate of 5 μL/min: 0.05*M* *N*-hydroxysuccinimide, 0.2*M* *N,N*-(3-dimethylaminopropyl)-*N*-ethyl-carbodiimide mixture, EGFR-Fc (10 μg/mL in 10 m*M* NaAc, pH 5.0), and 1*M* ethanolamine-HCl (pH 8.5).

RESULTS

Based on our methods, we have examined four different cases of protein interaction partners that have lead to the development of peptide mimetics.

Peptide inhibitors to gp41—an envelop glycoprotein of HIV-1

Gp41 is part of a glycoprotein complex of HIV-1 that binds to target cell receptors CD4 and CCR-5 or CXCR-4.[12] Gp41 is a three-stranded coiled-coil structure that is exposed during the viral entry process (prefusion state). Gp41, therefore, has been a target for the development of inhibitory compounds that bind to it and disrupt the viral entry process. Each subunit of gp41 consists of an N-heptad repeat unit from its N-terminal region (NHR) and C-heptad repeat unit from the C-terminal end (CHR) arranged in an antiparallel fashion. During fusion the subunits fold to form a six bundle helix with three NHR regions in the core stabilized by interactions with three ectodomain CHR regions. The NHR and CHR interacting regions were synthesized and structurally determined.[13] Peptide sequences based on the CHR region (C peptides) potentially bind to the NHR region and vice versa.[14] C-peptides have been experimentally shown to be potent inhibitors resulting in, for example, the successful drug Fuzeon (Roche) or Enfuvirtide (T-20).[15]

In the application of OpenContact for this system, the CHR peptide (Chain ID = C of 1AIK.pdb)was assigned as Protein A and the NHR peptide (Chain ID = N of 1AIK.pdb) was assigned as Protein B. The coarse parsing

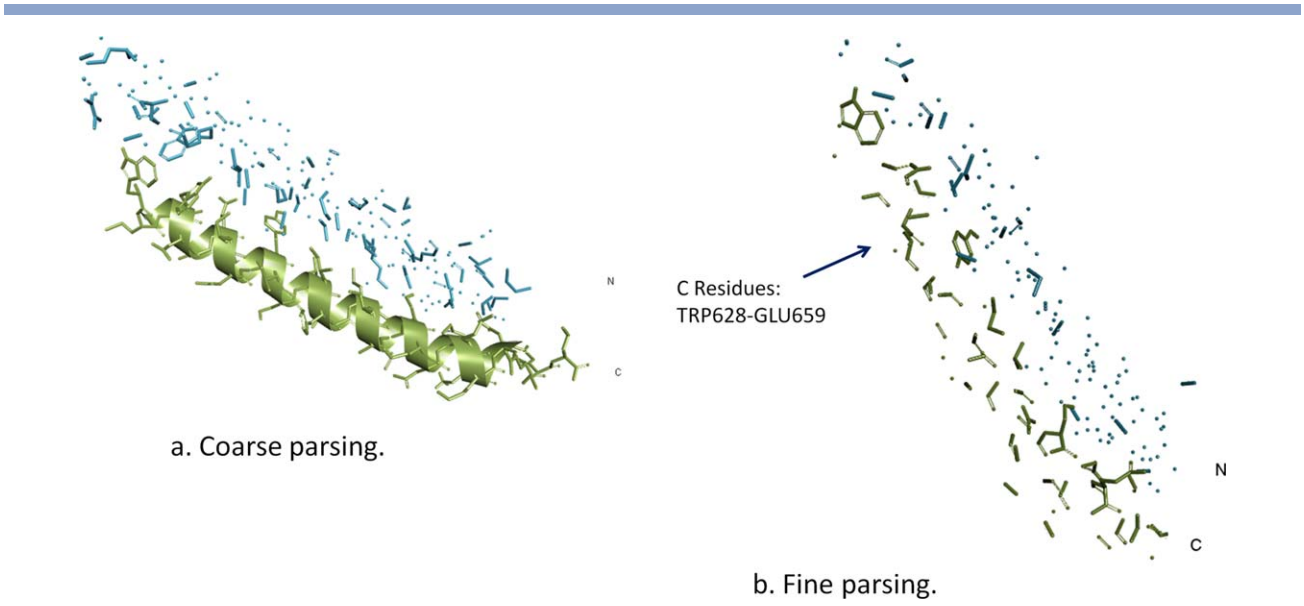

Figure 3

a: Coarse parsing or CoarseAB.pdb file for gp41. Due to multiple, close atom contacts, the CHR domain retains its helical structure. **b**: Fine parsing or FineAB.pdb file for the gp41 system. The CHR domain sequence predicted (C628–C659) has been experimentally shown to be one of the most potent inhibitors of gp41. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

*.pdb file (CoarseAB.pdb), for contacts less than 10.5 Å, is shown in Figure 3(a).

As seen in Figure 3(a) and in the Supporting Information, there are many more contact atoms for the CHR unit than for NHR unit. This result is consistent with experimental studies that show peptide inhibitors based on CHR domain are generally more potent than NHR domain peptides. In addition, as shown in Figure 3(b), the fine parsing based on the Coulombic and LJ criteria yielded a 32 residue peptide sequence (C Residues TRP628-GLU659) almost exactly overlying a number of known nanomolar binding peptides (C34's and SJ-2176) from the CHR region:[14]

C34: TRP628-LEU661
SJ2176: GLU630-GLU659

Figure 4(a) (Lennard Jones Potential plot) and 4(b) (Coulombic Potential Plot) show examples using the FineAB.txt output file for this system, where the parsing correctly predicted the dominance of apolar C-peptide binding to the NHR hydrophobic pocket region (N656-N573) [Fig. 4(a)], as well as several important salt bridge and polar interactions [Fig. 4(b)] (Complete output files and spreadsheets are given in the Supporting Information.).

As noted above, OpenContact is simply a first-step at possible peptidyl-biomimetics, and C34 peptides have been further experimentally studied in terms of their binding affinities, solubilities, residue variations, viral strain variations, and so forth.[14] For example, C34 peptides exhibited solubility issues in application.[14] Note that Fuzeon or DP-178 (C Residues TRY638-PHE673) demonstrated improved solubility and equally strong binding, when compared with C34 peptides.[14,15] Thus, simple contact mapping will not rectify potential solubility issues of isolated peptide segments. It is also unclear as to the role played by the additional residues of DP-178 outside of the contact map, that is, residues 662–673, although they could be associated with maintaining the helical structure/stability of Fuzeon or a binding segment not evident from the static map (see Discussion below). The next example demonstrates how simple contact mapping yields a small linear peptide from a larger protein partner file.

Peptide mimetics for EphB4-EphrinB2 interactions

The PPI between membrane EphrinB2 ligand and membrane tyrosine kinase receptor EphB4 represents a critical signaling pathway in embryonic and tumor angiogenesis.[16,17] Inhibiting this interaction could potentially slow angiogenesis in growing tumors and, thus, small molecule, peptide inhibitors based on phage display libraries[16,17] or peptide mimetics[1] have been proposed. The structure of this interaction complex has been determined[17] (pdb file: 2HLE) and is used to investigate potential peptide mimetics as given below.

In the application of OpenContact to this system, the input files were established as follows:

Protein A: 2HLE.pdb Chain ID = A (EphB4 receptor)
Protein B: 2HLE.pdb Chain ID = B (EphrinB2 ligand)

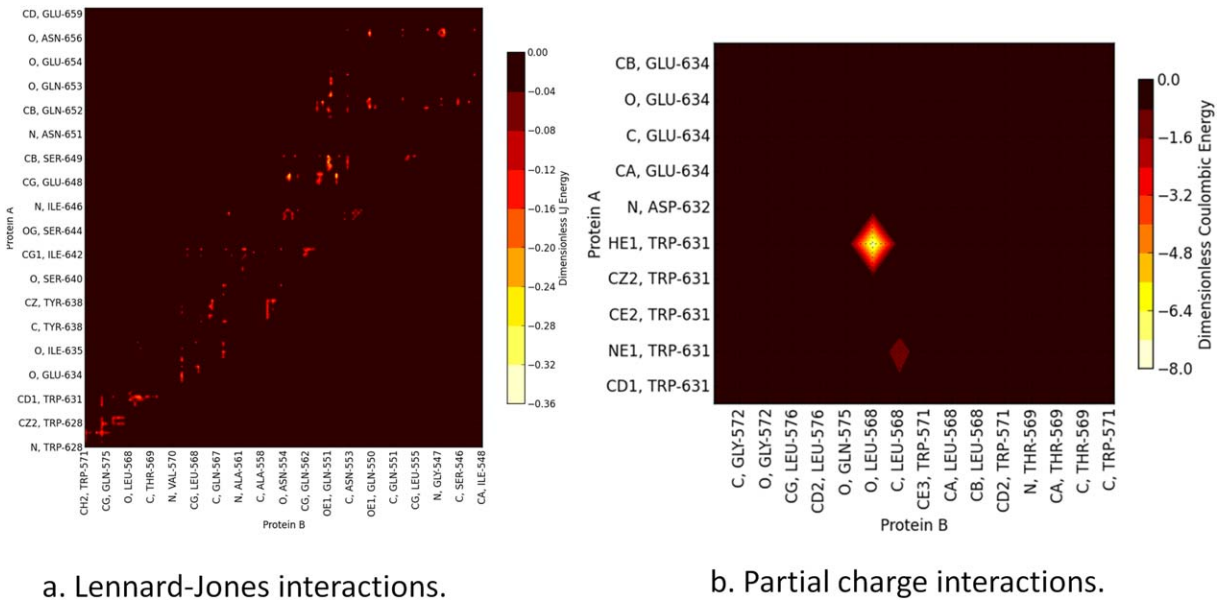

Figure 4

a: Lennard Jones potential plot for atom-atom interactions predicted from FineAB.txt showing a significant number of strong van der Waals interactions for this system. The dimensionless potential values are given by the colorbar. **b**: Coulombic potential plot for atom–atom interactions predicted from FineAB.txt showing a limited number of partial charge interactions for this system as compared to van der Waals potential interactions. The open source Matplotlib feature of OpenContact allows the user to zoom into specific atom-atom interactions. Herein, we show the partial atomic charge interaction predicted between O, LEU-568 of NHR and HE1, TRP-631 of CHR. All naming conventions follow the Protein Data Bank format.

The coarse parsing, shown in Figure 5, indicates a significant number of residue–residue interactions. However, many of these interactions are associated with secondary, beta sheet structures of the EphB4 receptor that peptide segments may not mimic in isolation. However, smaller linear segments associated with loops and other less structured segments could be potential inhibitor candidates.[1] In particular, the coarse parsing demonstrates that previously proposed[1] residues 116-128 of the EphrinB2 G-H loop (Chain ID B shown) make intimate contact with the hydrophobic pocket of the EphB4 receptor. This is shown more clearly in the fine parsing *.pdb file given in Figure 6 and in the Lennard Jones plot created from FineAB.txt file (Supporting Information). In addition, the spreadsheet created from the FineAB.txt file (Supporting Information) concurs with London *et al.*[1] that these specific receptor residues (116B-128B) confer a majority of the interaction energy via a number of strong van der Waals interactions associated with the hydrophobic pocket. In addition, however, another potential small linear peptide mimetic is identified from a loop of the ligand (A chain or EphB4 protein) as LYS24A-TRP32A, which interestingly retains its structure in the coarse parsing (Fig. 5). Ligand (membrane) based inhibitors have not been explored for this system to our knowledge, and further analysis, such as dynamic docking, is beyond the scope of this article; OpenContact simply provides a rapid assessment of potential peptide mimetics for further detailed studies. As a final note, the fine parsing from OpenContact also correctly listed LEU95A as a significant "player" in the interactions (with PHE120B and PRO122B); this ligand residue is believed responsible for its selectivity to EphB4[16] (Supporting Information).

This example demonstrates the initial design/discovery strategy often used for identifying potential peptide biomimetics, that is, looking for consecutively numbered contact residues from the parent protein chain that make strong contact with partner protein chains. Consecutive residues are necessary to help maintain the native structure of the isolated peptide mimetic. Although there are a number of residues identified in the coarse and fine parsing results, as shown for example in Figure 6, only consecutively numbered residues greater than approximately 8 are generally considered as potential candidates.[2] Thus, a significant number of residues are required for strength in binding and those residues should be in sequence order from the parent protein for structure considerations. Contact mapping only provides a starting point for such an identification and further dynamic computational/experimental studies of any potential peptide candidate identified are, of course, still required.

Peptide inhibitors to the proto-oncogenic transcription factor Myc-Max

The heterodimer protein complex Myc-Max (pdb ID: 1NKP) is a transcription factor in cell proliferation and

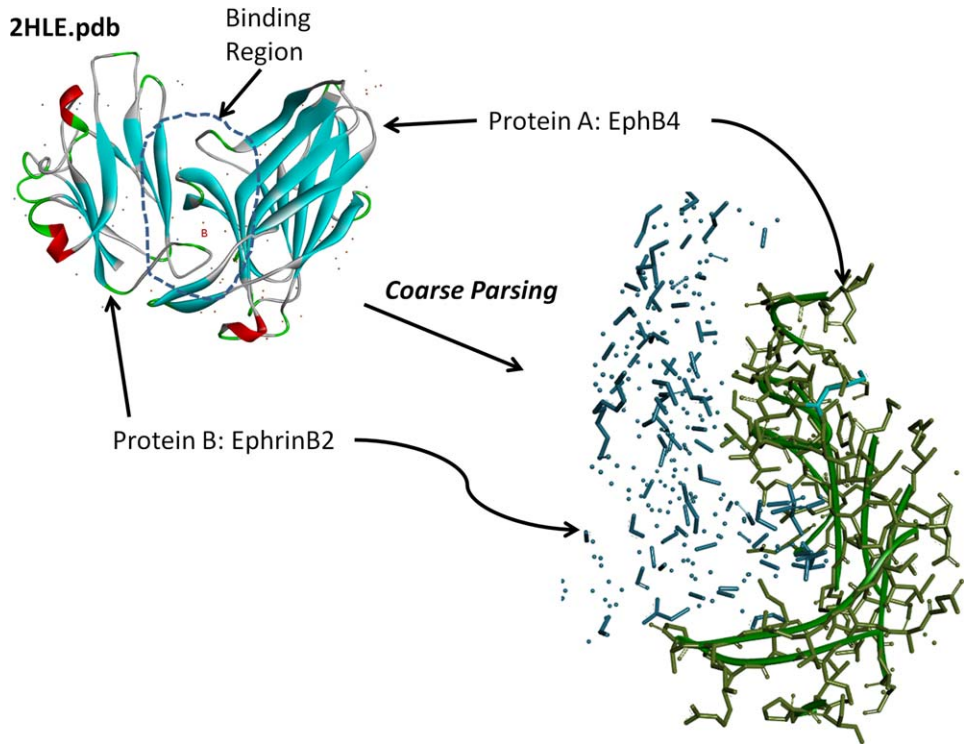

Figure 5

Coarse parsing or CoarseAB.pdb file for EphB4-EphrinB2 system. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

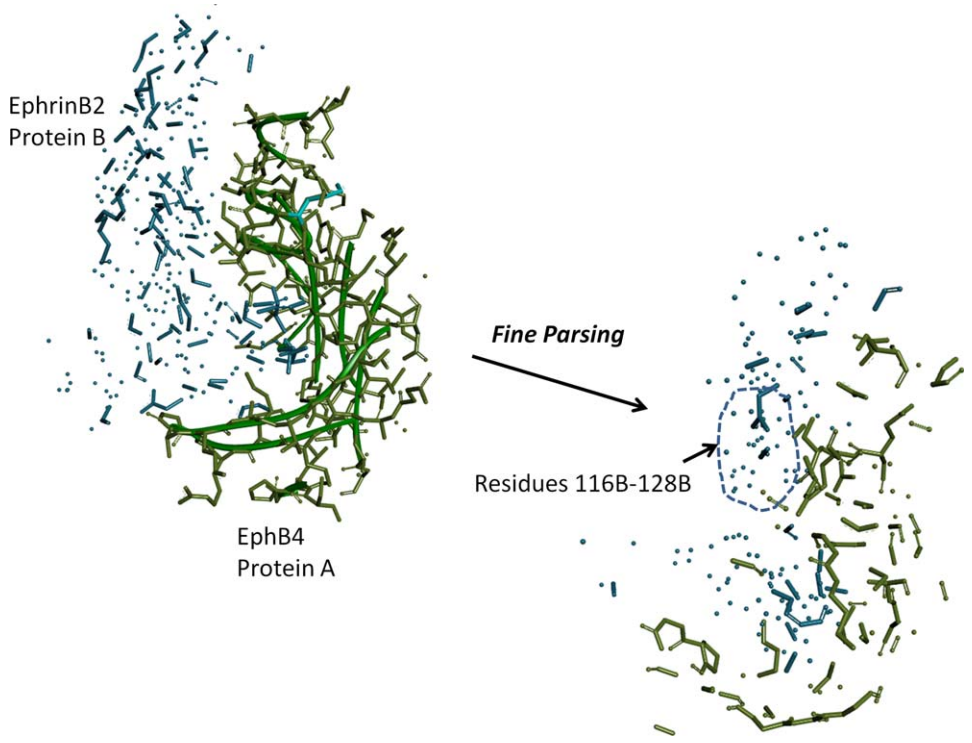

Figure 6

Fine parsing or FineAB.pdb file for EphB4-EphrinB2 system. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

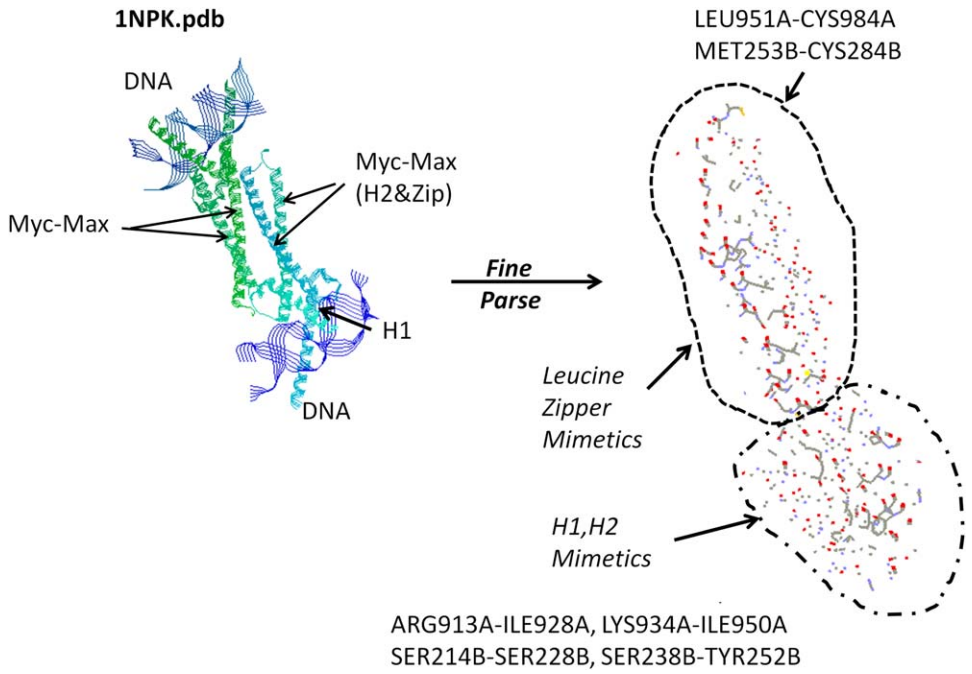

Figure 7

Myc-Max tetramer system: pdb ID: 1NKPand fine parsing or FineAB.pdb file for the Myc-Max system. Not all residue labels appeared in the fine parsing plot due to insufficient number of residue atoms. The complete output files are given in the Supporting Information. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

is upregulated in many different cancer types.[18] The Myc-Max heterodimer is therefore an ideal candidate for attempted development of inhibitory compounds based on peptide biomimetic strategies.[19,20] Herein, we again use OpenContact by choosing:

Protein A: (Myc) 1NKP.pdb Chain ID = A
Protein B: (Max) 1NKP.pdb Chain ID = B

Figure 7 shows the 1NKP.pdb structure file from the Protein Data Bank of a tetramer complex involving two Myc-Max heterodimers.[18] The Myc protein has a so-called (N-term) Basic Region that binds to DNA promoter, followed by the secondary structures: {Helix (1)-Loop – Helix(2) – Leucine zipper} (C-term). The Max DNA binding protein has the same secondary structural motif (b-HLH-zip) as shown. Note that the H2 region is contiguous with the Leucine zipper region in these structural motifs.[18,19]

The fine structure interaction (FineAB.pdb) for this system is also shown in Figure 7, and FineAB.txt gives all atom-atom interactions under fine parsing (Supporting Information). The fine parsing from OpenContact identifies the following potential peptide biomimetics from either the Myc or Max protein:

Myc Protein: *H1 Region*: ARG913A-ILE928A; *H2 Region*: LYS939A-ILE950A; *Zipper*: LEU951A-CYS984A
Max Protein: *H1 Region*: SER214B-SER228B; *H2 Region*: SER238B-TYR252B; *Zipper*: MET253B-CYS284B

Our H1 mimetic results almost exactly overlay previously proposed, synthesized, and studied[19] mimetics c-Myc H1: ASN915A-ILE928A and Max H1 mimetic: ASP216B-VAL229B. In addition, c-Myc H1 mimetics have been developed with amino acid substitutions (peptide mimetic mutants) to improve their helical content.[19] These mutants also exhibited complimentary binding to c-Myc and have been further studied including complexing with cell penetrating polypeptides.[20–22] Note that the Max peptide synthesized from the H1 region was experimentally shown to lack helical structure with no detectable interaction with c-Myc.[19] No studies on H2 or Zipper mimetics for this system have been reported to our knowledge. This system nicely illustrates that not all potential helical peptide mimetics will retain their structure in isolation from their parent protein chain or parent protein chain partner.

EGFR-EGF system

The epidermal growth factor receptor (EGFR) family, also called ErbB family, consists of four kinds of receptors, EGFR (ErbB1), Neu (ErbB2), ErbB3, and ErbB4. Each receptor consists of a cytoplasmic tyrosine kinase domain, a single transmembrane spanning region, and an extracellular region (sEGFR), the latter of which contains approximately 620 amino acids. The extracellular region of each receptor has four domains, L1 and L2 (or

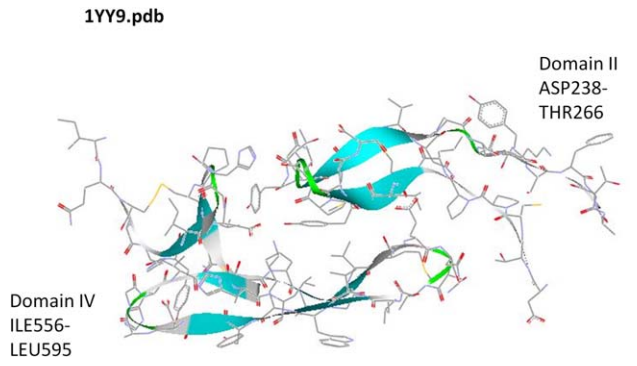

Figure 8

Domains II and IV of sEGFR in the unextended, inactive state. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

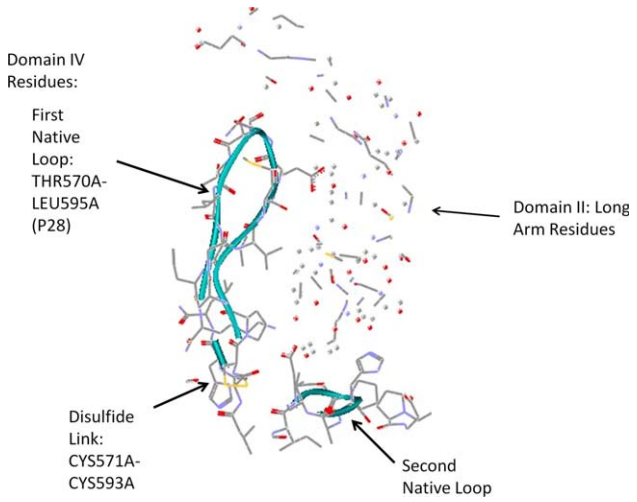

Figure 9

CoarseAB.pdb Output Results from OpenContact for the sEGFR (Domains II and IV) System. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

I and III), which are leucine-rich regions, and CR1 and CR2 (or II and IV), which are cysteine-rich.[23] The ligand EGF binds to the extracellular domain of the receptor and leads to receptor dimerization and activation of tyrosine kinase enzyme located in the intracellular domain of the receptor. Kinase activation leads to transphosphorylation of the receptor intracellular domains and initiates multiple signal transduction pathways.[24] The over expression, mutation or truncation of the EGF receptor leads to constant activity of the receptor which results in excess cell growth and, in turn, cancer. EGFR is implicated in the development of wide range of epithelial cancers, including those of breast, colon, head and neck, kidney, lung, pancreas and prostate.[25]

Unliganded EGF receptors are predominantly in the unextended form (inactive state) with Domain II buried by an intramolecular interaction with Domain IV, as shown in Figure 8.[26] Structural studies on EGF-EGFR complex have shown that the growth factor EGF binds to the Domains I and III of extracellular-EGFR simultaneously. This binding alters the spatial arrangement of all extracellular domains (now called the "active, extended state") and exposes a critical region, known as the dimerization arm, of Domain II. The unexposed arm intimately assists in the dimerization of the receptor, which subsequently leads to activation of EGFs intracellular growth signaling pathway.[23,26] We also note that approximately 10% of unliganded EGFR is in its extended or "active" state at any given time.[23,26]

Numerous therapeutic, antagonistic compounds, primarily antibodies, have been developed that bind to extracellular domains I and III (L1 and L2) of EGFR.[27,28] Herein, however, we looked for a peptide biomimetic analog of Domain IV, using OpenContact, that may bind to the exposed arm of Domain II in the active state of EGFR thereby potentially inhibiting dimerization. Thus, we selected the following input data for exploration via OpenContact:

Protein A: 1YY9.pdb Chain A SER501-THR614 (Domain IV residues)
Protein B: 1YY9.pdb Chain A CYS207-CYS309 (Domain II residues)

As shown in Figure 9 from CoarseAB.pdb, one complete (native) loop structure and one linear segment from Domain IV residues were identified:

First Native Loop: THR570A - LEU595A; (Disulfide Bridge: CYS571A-CYS593A)
Linear Chain from Part of Second Native Loop: TYR561A - VAL568A; (Disulfide Bridge: CYS558-CYS567)

The first, 26 residue loop structure is natively stabilized by a disulfide bridge between CYS571-CYS593A. The second linear peptide overlaps with part of another disulfide stabilized loop (CYS558-CYS567). Atom–atom interaction potentials from FineAB.txt show a mixture of significant van der Waals and partial charge interactions from both residue sequences (Supporting Information).

The disulfide bond stabilization of these identified loop structures provided a unique opportunity to investigate the inclusion of secondary structure in peptide mimetics in addition to the helices studied in previous cases. Thus, we synthesized a 28 residue segment of the larger loop (VAL568-LEU595), simply called P28, and added the disulfide bridge between Cysteine residues.

For kinetic measurements, EGF or P28 at concentrations ranging from 1 μM to 30 nM and 10 μM to 1 μM, respectively, in running buffer were injected over EGFR at a flow rate of 30 mL/min for 3 min with dissociation time of 10 min. After each cycle, surfaces were regenerated with 10 mM glycine, pH 1.5. The data was reference subtracted and fitted to 1:1 Langmuir binding model to obtain the

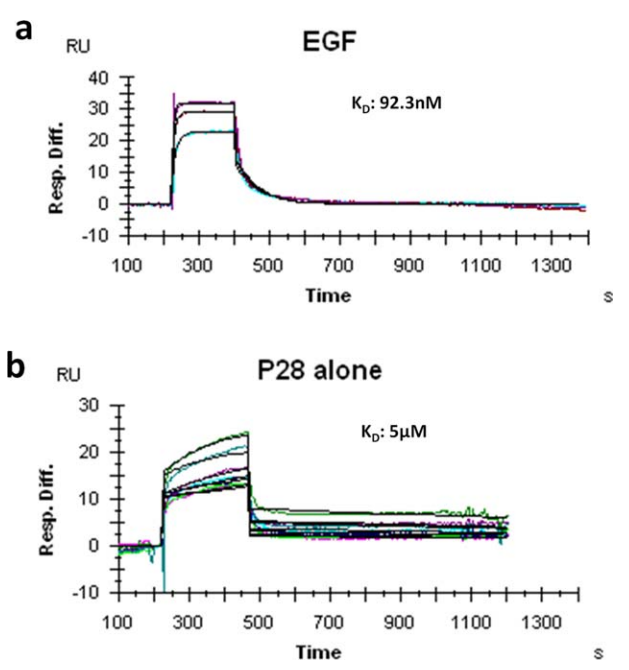

Figure 10

Response curves from Biacore kinetic binding assays: **a**: EGF as analyte; **b**: P28 as analyte. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

association rate constant (k_a), dissociation rate constant (k_d), and the equilibrium dissociation rate constant (K_D).

Figure 10(a,b) shows the initial results of the kinetic binding assays. For EGF binding [Fig. 10(a)], the calculated equilibrium dissociation constant K_D was 92.3 nM with association rate constant (k_a) of 1.75×10^5 M^{-1} s^{-1} and dissociation rate constant (k_d) of 0.016 s^{-1}. For P28 [Fig. 10(b)], the calculated equilibrium dissociation constant K_D was 5 μM with association rate constant (k_a) of 67.6 M^{-1} s^{-1} and the dissociation rate constant (k_d) of 3.82×10^{-4} s^{-1}. The equilibrium disassociation constant, K_D, for EGF-EGFR-Fc system shown is consistent with previously published values[23] and verifies the activity of the EGFR-Fc ligand. With only P28 as analyte, as shown in Figure 10(b), micromolar binding was predicted which is consistent with only 10% of the ligand in the active state. Other possibilities of the observed P28 binding include conformational changes of EGFR-Fc upon surface coupling. More experimental studies are currently underway in order to pinpoint the binding of P28 with EGFR domain II. Nonetheless, OpenContact was successful at identifying a potential peptide mimetic for this system that included secondary structural components.

DISCUSSION

The simple, static parsing algorithm, OpenContact, yielded results almost exactly equivalent to proposed peptide mimetics from four unrelated cases that were further experimentally studied. The output results included insights into static force interactions, contact numbers, secondary structure considerations, recognition of key residues, and a comprehensive listing of all potential mimetics. The empirical distance and potential cutoffs employed were shown to lead to a manageable, yet lucrative, numbers of potential peptide mimetics for any given system. The general format of the output results allow for more detailed analysis of specific residue-residue and atom–atom interactions by the user for any given system. All results given here were obtained on an off-the-shelf, desktop computer with less than a few minutes of computational time maximum; output files for each case were less than ~1 MB.

We note that potential peptide mimetics identified through the simple, static contact mapping of protein interaction partner files can be further screened via more complex and time-consuming dynamic algorithms, such as molecular dynamics and docking methods, which are outside the scope of this work. We also note that the literature experimental case studies on peptidyl biomimetics given here were limited in number and more experimental studies based on static mapping identification are needed. In addition, future studies are needed to examine the possibility of strong binding segments of protein interactions that may only arise through dynamic studies that are not evident in the static structure analysis given here. Nonetheless, the simple static structure analysis was shown to quickly and easily uncover potential peptide mimetics that could be "pipe-lined" for further study.

These test cases also nicely demonstrate some of the challenges associated with the development and application of peptide mimetics. In Case A, the gp41 system, solubility issues for C34 peptides were encountered despite their strong binding ability. Cases B, EphB4-EphrinB2 and C, Myc-Max, demonstrated that binding regions associated with particular secondary structural motifs, such as beta sheets and helices, cannot necessarily be maintained in isolation. In Case B, only small linear peptide segments were experimentally possible as mimetics and, in Case C., helical content was improved through synthetic mutation. Whether or not peptide segments of native protein structure can be maintained in isolation or to what degree peptide segments can maintain native binding ability will always be an outstanding issue in the development of peptide mimetics. Case C also demonstrated the need for complexing peptides for intracellular drug delivery applications. Case D, our own experimental peptide development based on static mapping, demonstrates that more sophisticated, native secondary structural maintenance may be possible for some peptide mimetics. Future improvements to OpenContact, currently underway, involve the inclusion in the mapping algorithm of the implicit apolar solvent potentials via the

solvent accessible surface area and volume and the addition of conformational flexibility of protein interactions.

ACKNOWLEDGMENTS

The authors thank Dr. Matthew Hartman at the VCU Massey Cancer Center for his helpful guidance on synthesis and purification of P28 and his generosity in providing reagents and equipment needed for the process. OpenContact is freely downloadable at www.engineering.vcu.edu/proteinengineering under the SleepyCat, OSI compliant license, or by contacting the corresponding author.

REFERENCES

1. London N, Raveh B, Movshovitz-Attias D, Schueler-Furman O. Can self-inhibitory peptides be derived from the interfaces of globular protein–protein interactions? Proteins 2010;78:3140–3149.
2. Groner B, editor. Peptides as drugs. Weinheim: Wiley-VCH; 2009.
3. London N, Raveh B, Schueler-Furman O. Druggable protein–protein interactions—from hot spots to hot segments, Curr Opin Chem Biol 2013;17:1–8.
4. Ulucan O, Eyrisch S, Helms V. Druggability of dynamic protein–protein interfaces. Curr Pharm Des May 29 2012.
5. Fry DC. Small-molecule inhibitors of protein–protein interactions: how to mimic a protein partner. Curr Pharm Des May 29 2012.
6. Roberts KE, Cushing PR, Boisguerin P, Madden DR, Donald BR. Computational design of a PDZ domain peptide inhibitor that rescues CFTR activity. PLoS Comp Biol 2012;8:1–12.
7. Tang YT, Marshall GR. Virtual screening for lead discovery. Methods Mol Biol 2011;716:1.
8. Pande J, Szewczyk M, Grover AK. Phage display: concept, innovations, applications, and future. Biotechnol Adv 2010;28:849–858.
9. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. A point-charge force field for molecular mechanics simulations of proteins based on condensed phase quantum mechanical calculations. J Comp Chem 2003;24:1999–2012.
10. Ramstein J. Lavery, R. Energetic coupling between DNA bending and base pair opening. Proc Natl Acad Sci USA 1988;85:7231.
11. Smith PE, Pettitt BM. Modeling solvent in biomolecular systems. J Phys Chem 1994;98:9700.
12. Chan DC, Kim PS. HIV entry and its inhibition. Cell 1998;93:681–684.
13. Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. Cell 1997;89:263–273.
14. Debnath AK. Progress in identifying peptides and small molecule inhibitor targeted to gp41 of HIV-1. Expert Opin Investig Drugs 2006;15:465–478.
15. Wild C, Dubay JW, Greenwell T. Propensity for a leucine zipper-like domain of the human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex, Proc Natl Acad Sci USA 1994;91:12676–12680.
16. Chrencik JE, Brooun A, Kraus ML, Recht MI, Kolatkar AR, Han GW, Seifert JM, Widmer H, Auer M, Kuhn P. Structural and biophysical characterization of the EphB4-EphrinB2 protein–protein interaction and receptor selectivity. J Biol Chem 2006;281:28185–28192.
17. Noberini R, Mitra S, Salvucci O, Valencia F, Duggineni S, Prigozhina N, Wei K, Tosato G, Huang Z, Pasquale EB. PEGylation potentiates the effectiveness of an antagonistic peptide that targets the EphB4 receptor with nanomolar affinity. PLoS One 2011;6: e2861(1–13).
18. Nair SK, Burley SK. X-ray structures of Myc-Max and Mad-Max recognizing DNA: molecular bases of regulation by proto-oncogenic transcription factors. Cell 2003;112:193–205.
19. Draeger LJ, Mullen GP. Interaction of the bHLH-zip domain of c-Myc with H1-type peptides. J Biol Chem 1994;269:1785–1793.
20. Giorella J, Clerico L, Pescarolo MP, Vikhanskaya F, Salmona M, Colella G, Bruno S, Mancuso T, Bagnasco I, Russo P, Parodi S. Inhibition of cancer cell growth and c-Myc transcriptional activity by a c-Myc helix type 1 peptide fused to an internalization sequence. Cancer Res 1998;58:3654–3659.
21. Bidwell GL, III, Raucher D. Application of thermally responsive polypeptides directed against c-Myc transcriptional function for cancer therapy. Mol Cancer Ther 2005;4:1076–1085.
22. Bidwell GL, III, Raucher D. Cell penetrating elastin-like polypeptides for therapeutic peptide delivery. Adv Drug Deliv 2010;62: 1486–1496.
23. Ferguson KM, Berger MB, Menddrola JM, Cho H-S, Leahy DJ, Lemmon MA. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell 2003;11:501–517.
24. Oda K, Matsuoka Y, Funahashi A, Kitano H. A comprehensive pathway map of epidermal growth factor receptor signaling. Mol Syst Biol 2005;1:1–17.
25. Scaltriti M, Baselga J. The epidermal growth factor receptor pathway: a model for targeted therapy. Clin Cancer Res 2006;12:5268–5272.
26. Li S, Schmitz KR, Jeffrey PD, Wiltzius JJW, Kussie P, Ferguson KM. Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell 2005;7:301–311.
27. Kamat V, Donaldson JM, Kari C, Quadros MRD, Lelkes PI, Chaiken I, Cocklin S, Williams JC, Papazoglou E, Rodeck U. Enhanced EGFR inhibition and distinct epitope recognition by EGFR antagonistic mAbs C225 and 425. Cancer Biol Ther 2008;7:726–733.
28. Li S, Kussie P, Ferguson KM. Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure 2008;16: 216–227.