ENHANCING PROTEIN SURFACE RECOGNITION THROUGH MULTIVALENT INTERACTIONS

Reported by Kyle Plunkett November 8, 2001

INTRODUCTION

Molecular recognition events, such as DNA-protein, protein-substrate, and protein-protein interactions, play essential roles in the specificity of many biological processes. The importance of these recognition events becomes more apparent when one considers that it is the balanced regulation of biological processes that allows life to flourish. By inhibiting the actions of proteins, researchers have been able to regain control over processes that have become unregulated. Previously, protein inhibition has focused on the protein-substrate recognition event that occurs at the active site of a protein.[1] Libraries of compounds, with the potential to interact at the protein's active site, have been synthesized and screened for biological activity.[2] A new strategy targets the multivalent interactions that occur during protein-protein binding. The result has been a rigorous study into the precise interactions that occur in protein-protein complexes,[3] as well as the design of molecules capable of binding selectively to the surface of proteins.[4] Binding of these molecules may cause a variety of responses, including the disruption of required recognition sites on the protein, the induction of allosteric conformational changes in the protein, or even the encumberment of the active site through steric hinderance (Figure 1).[4]

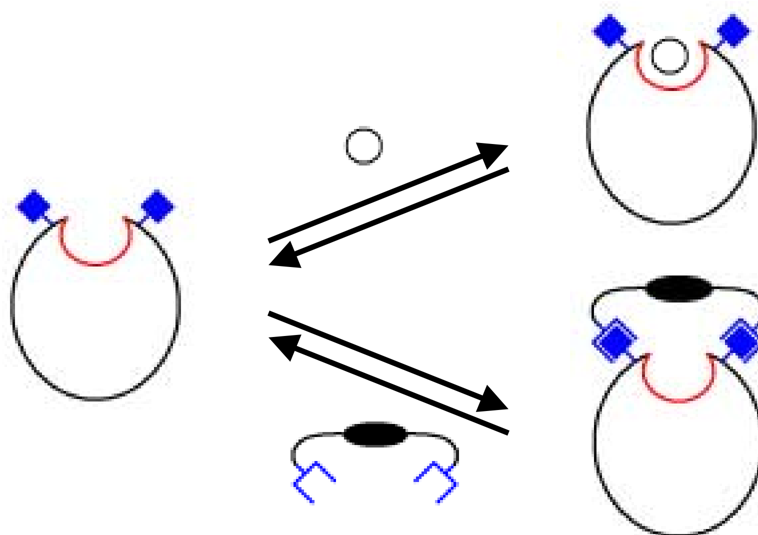


Figure 1. Steric inhibition of active site through a large, surface bound molecule.

Copyright © 2001 Kyle N. Plunkett

The surface area involved in protein-protein binding, or epitope, has generally been found to contain large hydrophobic areas that energetically drive aggregation in water.[5] Although this binding motif is most frequent, it is not unique.[6] Some protein interfaces are composed of multiple hydrophobic patches intermixed with hydrogen bonds and water molecules. Because of the divergent nature of these protein surfaces, a molecule or protein receptor that binds to another protein is typically required to interact with both hydrophobic and hydrophilic residues. Multivalent interactions are being recognized as being capable of providing both specificity and high association. This report will focus on the understanding of multivalentcy and how it may be applied to the design of protein inhibitors.

DETERMINATION OF PROTEIN – PROTEIN BINDING SITES

Protein-protein interactions occur over large surface areas (ca. 700 to 5000Å^2 per subunit) and are stabilized by a combination of the same forces (hydrophobicity, hydrogen bonding, electrostatic interactions, and van der Waals interactions) that stabilize intramolecular folding motifs in proteins.[3] Structural analysis of known protein-protein dimers in the Brookhaven Protein Databank (PDB) have found general trends in protein binding motifs, with the goal of identifying epitopes on proteins having unknown dimer binding structures.[7] By identifying structural properties such as interface residue propensity, protrusion index, hydrophobicity, and accessible surface area of the known dimer structures, and then applying these predictions to unknown homodimers, Jones and coworkers were able to correctly select greater than 70% of the interface regions.[7]

Alanine-scanning mutagenesis has also been employed to probe protein epitopes and has identified “hot spots” that are responsible for the majority of the binding energy.[8] By sequentially mutating individual amino acids at a proteins surface, side chains that affect the binding were identified. Alanine was chosen as the replacement residue in the mutated protein, because it eliminates the side chain beyond the β-carbon, but does not significantly alter the main-chain conformation. The knowledge of specific residues that play key roles in binding may lead to an improved strategy for multivalent inhibition of a protein.

MULTIVALENTCY IN BIOLOGY

Although a few examples of small molecules that inhibit protein-protein interactions have been discovered,[9-11] the search for small molecules with high association constants that disrupt these large surface area interactions has been difficult.[4] Multivalent interactions between a few small molecules (or multiple functionalities on a single molecule) and a large protein epitope have been investigated.[12]

Multivalentcy is an important element of many recognition processes, including protein-sugar recognition, cellular adhesion, and the blocking of pathogens from adhesion to target cells.[13] Multiple ligands tethered together minimize the degree of motion at the binding site and increases binding between the ligand and its receptor. For example, if molecule A binds to the receptor with an exothermicity of X, and molecule B binds to the same receptor, at a different location, with an exothermicity of Y, then the connection of A and B through a strain-free tether should result in an exothermicity of binding that is greater than X + Y. This effect has been described as the enthalpic chelate effect.[14] Entropy loss is also minimized with the binding of a multi-ligand molecule. The first contact between the receptor and the multivalent-ligand is made with a large loss of entropy, but further penalties are not accrued with additional binding, because of the proximity of the secondary ligands.

A good example of this type of binding is illustrated with the multivalent version of the drug vancomycin. Vancomycin has been found to bind to a short peptide sequence (D-ala-D-ala) on the exterior of a bacterial cell wall through both hydrophobic interactions and hydrogen bonding. A trimeric multivalent vancomycin molecule (Figure 2) was prepared by covalently attaching three vancomycin molecules to a rigid scaffold.[15] Similarly, a trimer of the D-ala-D-ala recognition sequence was synthesized and tested for binding to the vancomycin trimer. A dissociation constant (K_d) of 4×10^{-17} M was found for the multivalent interaction, whereas the attachment of three monomeric D-ala-D-ala molecules to the vancomycin trimer had a K_d of 1×10^{-6} M (Figure 2). Thus, an effective 10^{11} increase in the binding affinity between vancomycin and D-ala-D-ala was observed by making this interaction trivalent. Although protein binding to ligands does not follow this model strictly, some studies have hinted at the added value of utilizing multivalent interactions when designing molecules for protein inhibition.

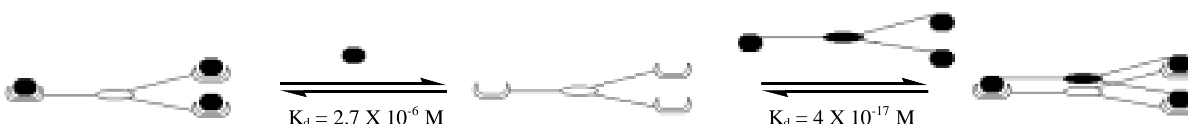


Figure 2. Multivalent interaction model using vancomycin ( ) and D-ala-D-ala ( ).

PROTEIN BINDING THROUGH MULTIVALENT INTERACTIONS

Divalent Small Molecule Recognition

Shuker and coworkers have prepared multivalent molecules using combinatorial library searching, followed by the linking of two protein-binding molecules to form one high affinity molecule.[16] They have entitled their approach 'Structure-Affinity Relationships by Nuclear Magnetic Resonance' (SAR by NMR). In the first step (Figure 3B), a small library of molecules was screened via

^{15}N NMR for protein binding affinity. Binding events were detected by changes in the chemical shift of the protein amide backbone in the 2D ^{15}N-HSQC spectra. Once a lead molecule was identified, analogs were developed to maximize binding to the surface of the protein (Figure 3C). Binding of a second (different) compound to another pocket on the same protein, in the presence of the first molecule, was identified (Figure 3D) and optimized (Figure 3E) in a similar manner. Once the two lead molecules were found, X-ray crystallography was utilized to find their locations and relative orientations when complexed to the protein. This facilitated the design of the appropriate inert tether (Figure 3F), with the goal of producing one high-affinity, divalent molecule.

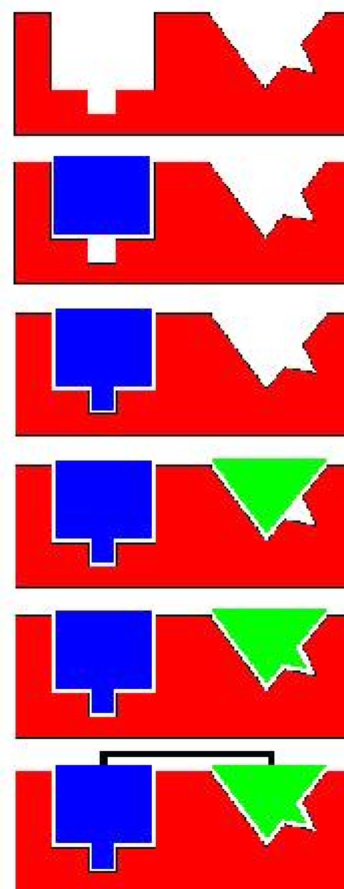
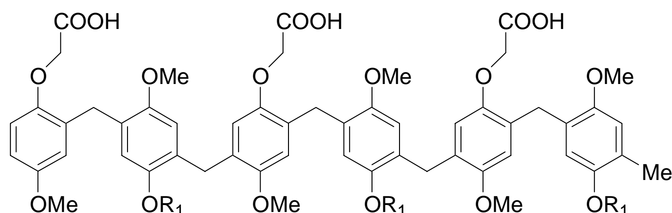


Figure 3. SAR by NMR method.

The SAR by NMR strategy was applied to the FK506 binding protein (FKBP), which inhibits calcineurin (a serine-threonine phosphatase) and blocks T cell activation.[17] After screening 10,000 small molecules by NMR concentration experiments, trimethoxyphenyl pipecolinic acid (**1**) was found to have the highest affinity for FKBP (K_d = 2.0 μM). Given the relatively high binding affinity of ligand **1,** no further optimization was pursued. The library was screened again in the presence of excess amounts of **1** to block binding at the first site. NMR chemical shift data indicated that a second ligand bound to a pocket near **1**. Optimization of the substituents on the aromatic rings of this second molecule (Figure 4) led to compound **2**, with a K_d = 100 μM. The X-ray structure of FKBP complexed to ligand **1** and **2** showed that the methyl ester of ligand **1** was in close proximity to the hydroxyl group on the benzoyl ring of ligand **2**. The methyl ester of **1** and hydroxyl of **2** were linked together by a hydrocarbon chain of varied length (n = 0 - 6) and subsequently tested for binding by fluorescence microscopy (Scheme 1). The tethered compound with n = 3 showed the highest

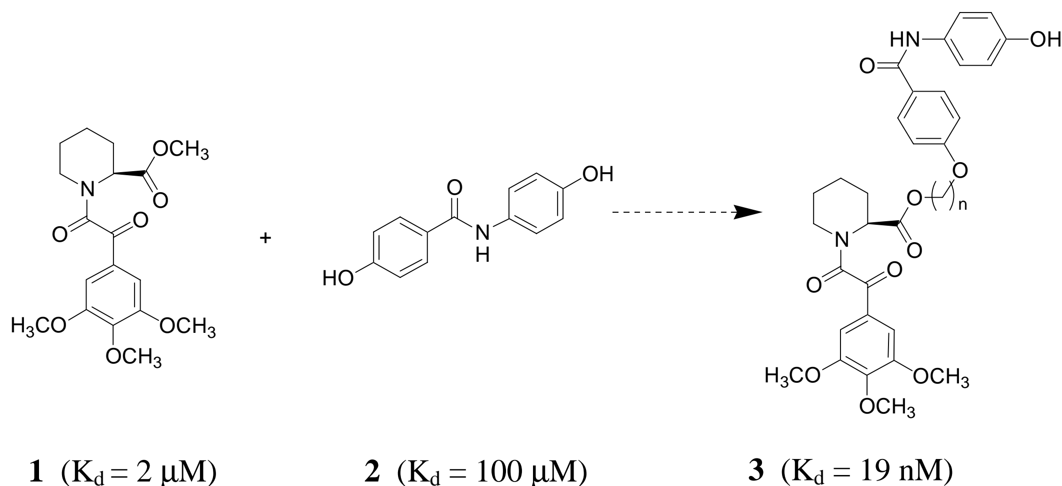


R_1	K_i (nM)
-H	1300
$-CH_2O(CH_2)_2CH_3$	1000
$-(CH_2)_3Ph$	100
$-(CH_2)_5CH_3$	17
$-CH_2$-(cyclo)-C_6H_{11}	20

Figure 4. Oligomer hydrophobic substituents and their inhibition of HLE.

binding affinity, with a K_d = 19 nM. To determine the location of the tethered molecule, NOEs were compared to the unlinked ligands, **1** and **2**. The designed molecule was observed to bind at the same location, but with a small shift of 1 to 2 Å. In contrast to combinatorial libraries of multivalent compounds, relatively few compounds must be synthesized for the SAR by NMR method, because the un-tethered ligands are optimized separately prior to linking. Major limitations of this technique, however, include the requirement that the biomacromolecule be sufficiently small for NMR analysis (<30 kDa) and available in large quantities (>200 mg).

Scheme 1. Covalent linkage of two protein binding molecules to form one multivalent molecule.

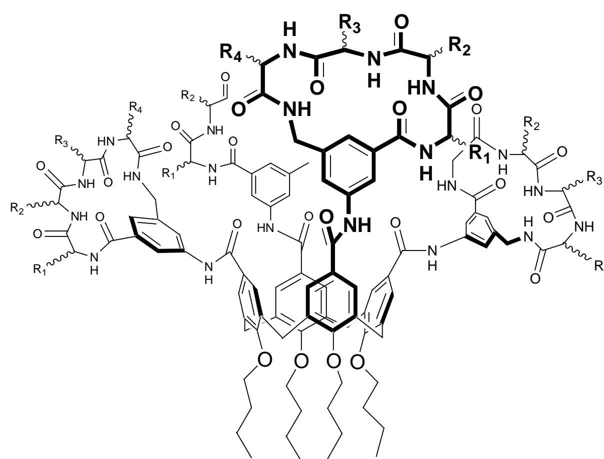


Oligomer / Polymer Multivalent Inhibitors

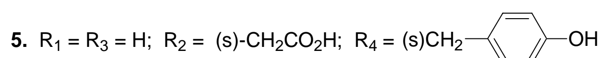
Using a surface recognizing oligomer, Regan and coworkers identified molecules that inhibit human leukocyte elastase (HLE).[18] HLE is a 30 kDa serine protease found to have a highly basic surface that contains 18 arginine residues and only 6 acidic residues. Adjacent to the active site are four arginine residues surrounded by hydrophobic amino acids that are located within 40 Å of each other. Taking advantage of the proximity of these residues, a diphenylmethane-based oligomer containing anionic and lipophilic functionalities was designed and synthesized so that it would interact with multiple sites and consequently sterically block the active site of the protein. The strategy involved positioning the carboxylic acid groups of the oligomer close to the guanidinium groups on the protein, and hydrophobic groups close to hydrophobic domains of the protein. The investigators found that a minimum size (five phenoxyacetic acid units) was necessary for inhibition, but extending the series farther had little additive effect. The addition of hydrophobic functionalities to the aromatic rings of the oligomer resulted in even greater inhibitory effects (Figure 4). The specific inhibitory effects of the

oligomer on other proteins were measured. Cathepsin G, which is more basic than HLE, was not found to be inhibited significantly by the oligomer, indicating that the multivalent interactions are specific for HLE.

Hamilton and coworkers extended the oligomer design into a rigid antibody like structure for protein surface recognition. Antibodies use multiple equivalent receptor sites (i.e., two in IgGs and four in IgAs) to recognize antigens or other ligands multivalently.[13] A calix[4]arenes scaffold was used to attach multiple peptide loops in stable hairpin-turn conformations that served as binding sites. Binding of this rigid scaffold was tested on α-chymotrypsin, which has a positively charged surface consisting of several lysine and arginine groups.[19] Peptide loops were synthesized containing four natural amino acids and the 3-aminomethylbenzoyl dipeptide analogue, which induces β-turn conformations.[20] The surface areas of these structures were estimated to be around 450-500 Å^2, sufficient for the ligand to interact with large areas of the protein. In this study, three peptide sequences were tested with properties ranging from anionic to cationic (Figure 5). The peptide sequence with multiple aspartic acid residues was the most potent inhibitor of α-chymotrypsin, having an inhibition constant (K_i) of 0.81 μM. Competitive binding experiments showed that compound **4** was able to displace trypsin inhibitor, suggesting that **4** is binding in a similar location. Even though molecule **4** has a greater number of anionic groups than **5** or **6** (K_i = 15 and 36 μM respectively), it is not believed that electrostatic interactions are the sole contributor to the protein-ligand interaction. A calix[4]arene scaffold with carboxylic acid groups in place of the peptide loops showed lower inhibition (K_i = 26 μM) than the aspartic acid containing peptide loops. A monomeric peptide loop without the scaffold and three other peptide loops showed no inhibition, thus signifying the importance of these multiple interactions.



4. $R_1 = R_3 = H$; $R_2 = R_4 =$ (s)-CH_2CO_2H



6. $R_1 = R_3 = H$; $R_2 = R_4 =$ (s)-$(CH_2)_4NH3^+$

Figure 5. Calix[4]arene scaffold with 4 peptide loops that mimic an antibody.

Collier and coworkers have recently reported a polyvalent polyacrylamide polymer that inhibits anthrax toxin.[21] *Bacillus anthracis,* the causative agent of anthrax, contains three protein monomers: a receptor-binding moiety, termed protective antigen (PA), and two enzymatic moieties, termed edema

factor (EF) and lethal factor (LF). Once the protein monomers are released from the bacteria, PA binds to the surface of a mammalian cell and fragments into subunits (PA63) that assemble into a protein heptamer. EF and LF bind to the heptamer and form a complex that penetrates the cell and causes the death of macrophages and ultimately the host. Using phage display,[22] peptides that inhibit the binding of EF and LF to the PA63 heptamer were screened. A dodecapeptide (HTSTYWWLDGAP) was found to disrupt binding best with an IC_{50} = 150 μM. The peptide was hypothesized to be interacting with the hydrophobic surface of a PA63 heptamer and thus inhibiting complexation to EF or LF. To increase interactions with the heptamer complex, a polyacrylamide polymer with multiple copies of the dodecapeptide was synthesized. The polymer was found to inhibit LF binding with an IC_{50} = 20 nM and was later tested *in vivo* and found to eliminate toxicity of the anthrax toxin in mice.

CONCLUSIONS

Multivalent interactions between designed molecules and protein surfaces can, in general, raise the potency of protein inhibition, far exceeding that of monovalent interactions. These surface-binding molecules can affect the function of proteins by sterically blocking the substrate or physically disrupting key binding amino acid interactions. The rational design of these multivalent molecules may provide alternative routes to protein regulation useful for pharmaceutical applications.

REFERENCES:

(1) Cochran, A. G. *Chemistry & Biology* **2000**, *7*, R85-R94.
(2) Babine, R. E., Bender, S.L. *Chem. Rev.* **1997**, *97*, 1359-1472.
(3) Stites, W. E. *Chem. Rev.* **1997**, *97*, 1233-1250.
(4) Peczuh, M. W., Hamilton, A.D. *Chem. Rev.* **2000**, *100*, 2479-2494.
(5) Chothia, C., Janin, J. *Nature* **1975**, *256*, 705-708.
(6) Larsen, T. A., Olson, A.J., Goodsell, D.S. *Structure* **1998**, *6*, 421-427.
(7) Jones, S., Thornton, J.M. *Proc. Nat. Acad. Sci. USA* **1996**, *93*, 13-20.
(8) Cunningham, B. C., Wells, J.A. *Science* **1989**, *244*, 1081-1085.
(9) Tilley, J. W., Chen, L., Fry, D.C., Emerson, S.D., Powers, G.D., Biondi, D., Varnell, T., Trilles, R., Guthrie, R., Mennona, F., Kaplan, G., LeMahieu, R.A., Carson, M., Han, R.-J., Liu, C.-M., Palermo, R., Ju, G. *J. Am. Chem. Soc.* **1997**, *119*, 7589-7590.
(10) Li, S., Gao, J., Satoh, T., Friedman, T.M., Edling, A.E., Koch, U., Choksi, S., Han, X., Korngold, R., Huang, Z. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 73-78.
(11) Li, R., Dowd, V., Stewart, D.J., Burton, S.J., Lowe, C.R. *Nat. Biotechnol.* **1998**, *16*, 190-195.
(12) Fan, E., Zhang, Z., Minke, W.E., Hou, Z., Verlinde, C.L.M.J., Hol, W.G.J. *J. Am. Chem. Soc.* **2000**, *122*, 2663-2664.
(13) Mammen, M., Choi, S.-K., Whitesides, G.M. *Angew. Chem. Int. Ed.* **1998**, *37*, 2754-2794.
(14) Searle, M. S., Sharman, G.J., Groves, P., Benhamu, B., Beauregard, D.A., Westwell, M.S., Dancer, R.J., Maguire, A.J., Try, A.C., Williams, D.H. *J. Chem. Soc., Perkin Trans.* **1996**, *1*, 2781-2786.
(15) Rao, J., Lahiri, J., Isaacs, L., Weis, R.M., Whitesides, G.M. *Science* **1998**, *280*, 708-711.

(16) Shuker, S. B., Hajduk, P.J., Meadows, R.P., Fesik, S.V. *Science* **1996**, *274*, 1531-1534.
(17) Liu, J., Farmer, J.D., Lane, W.S., Friedman, J., Weissman, I., Schreiber, S.L. *Cell* **1991**, *66*, 807-815.
(18) Regan, J., McGarry, D., Bruno, J., Green, D., Newman, J., Hsu, Chin-Yi, Kline, J., Barton, J., Travis, J., Choi, Y.M., Volz, F., Pauls, H., Harrison, R., Zilberstein, A., Ben-Sasson, S.A, Chang, M. *J. Med. Chem.* **1997**, *40*, 3408-3422.
(19) Capasso, C., Rizzi, M., Menegatti, E., Ascenzi, P., Bolognesi, M. *J. Mol. Recognit.* **1997**, *10*, 26-35.
(20) Smythe, M. L., von Itzstein, M. *J. Am. Chem. Soc.* **1994**, *116*, 2725-2733.
(21) Mourez, M., Kane, R.S., Mogridge, J., Metallo, S., Deschatelets, P., Sellman, B.R., Whitesides, G.M., Collier, R.J. *Nat. Biotech.* **2001**, *19*, 958-961.
(22) Zwick, M. B., Shen, J., Scott, J.K. *Curr. Opin. Biotechnol.* **1998**, *9*, 427-435.

Copyright © 2001 by Melissa J. Witmer