

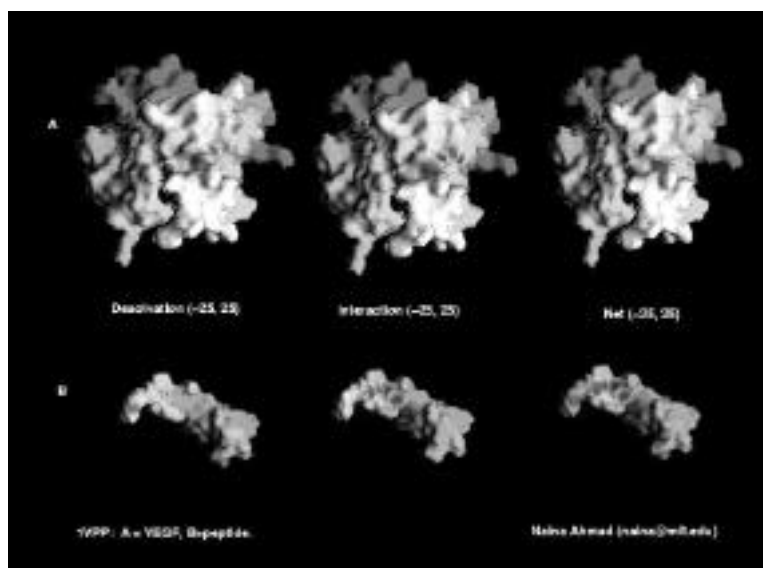
The Glue between Proteins:

An Investigation into Protein Binding

Naina Shehzeen Ahmad

Perhaps the most important aspect of unraveling the mystery behind how biological molecules dispense their function is to understand the structural ramifications that enable them to fit into their roles. This provides not only a comprehensive view of how these biologically active molecules actually work, but also valuable insight into why they may malfunction in diseases and what can be done to fix this malfunction. Of the many biologically active and crucial molecules, proteins present a complex and interesting molecule of study, mainly because of the astounding range of functions they are involved in. Proteins are responsible for catalysis of an extraordinary range of biochemical reactions, providing structural rigidity to the cells, controlling the transport of materials across membranes, regulating concentrations of metabolites, acting as sensors and switches, causing motion, and controlling gene function and expression.

Proteins perform these diverse roles within the biochemical pathways of the living system predominantly by binding either to another substrate molecule or to another protein, DNA, or receptor. Enzymatic action, activation of genes for expression, or hormone action are all caused by binding properties of proteins to a partner forming a complex. The specific nature of binding enables proteins to recognize their target molecule and to bind only to their preferred substrate. The affinity of binding, or how strongly the protein can bind to its specific target molecule, also plays a major role in the binding of these molecules, enabling the protein to pick out its target substrate and tightly bind to it despite competition for its binding sites by similar substrates. Why might one protein molecule bind more tightly to one target molecule



but not to another of similar size and shape? What factors might cause the protein to recognize its target molecule and not bind to any other molecule in the vicinity?

Clearly, the binding choices are made on energy–minimization criterion. What is less obvious is exactly what factors come to play in these energy calculations. During the binding of two protein molecules, there is a gain in energy of each molecule as it desolvates from its surrounding aqueous media in order to form bonds to each other. This destabilizing energy gain must be compensated for by the stabilizing loss in energy of each molecule, due to the interactions with each other and with the new aqueous surroundings, if a stable binding system is to be achieved.

Why should protein binding generate any interest at all? Specificity of protein binding lies at the base of its functionality, but how does this fit into the broader perspective of what we are able to do with the information from modeling such protein binding interactions? With advancements in modern age, cutting–edge technology and our ability to exploit computer programs to run computation–intensive modeling programs, we are now able to visualize the interaction at the molecular level of each protein–binding interface. This presents a novel understanding of why one protein binds to another, hence enabling scientists to design synthetic molecules either to directly disrupt pathogenic protein–binding interactions or to promote certain other interactions that will negate the pathogenic effects.

Background

In this report, various computational and graphical programs have been used to investigate the contributing factors in the binding of a Vascular Endothelial Growth Factor and a synthetic receptor blocking peptide.

The Vascular Endothelial Growth Factor (VEGF) is a highly specific extracellular growth factor, or mitogen, for endothelial cells. It plays a major role in the regulation of the formation of new vessels, termed vasculogenesis, and in the formation of blood vessels, a process termed angiogenesis. The basic mechanism by which mitogens act is by binding to a membrane receptor, altering its conformation and causing a cascade of serial bindings to occur, which in turn results in regulation of the cell cycle. At the molecular level, this growth factor will bind to the Flt-1 (fms-like tyrosine kinase) and KDR (kinase domain receptor). These receptor tyrosine kinases are proteins that are able to add a phosphate group to other proteins in order to change their activity. They are uniquely

expressed on endothelial cell surfaces and have Ig-like domains in their extracellular portions. On binding, vasculogenesis and angiogenesis are stimulated, causing new blood vessels to form.^{1,2}

This natural process may have adverse manifestations in pathogenic circumstances, for example, in the case of cancer or diabetic retinopathy, in which the spread of disease is based on the proliferation of diseased cells. By disrupting the VEGF action, angiogenesis can be effectively slowed down, perhaps even stopped, at target sites, causing regression of the disease. A neutralizing murine anti-VEGF antibody was found to inhibit cancer proliferation in mice, and in recent years, a humanized version of the antibody was developed for testing in clinical trials.³

In order to search for potential candidates that disrupt the binding between the VEGF and the receptors to which it binds natively, it is crucial to accurately determine the interactions between the growth factor and its native receptors. From results of X-ray crystallography, the interactions have been visualized and a synthetic peptide sequenced that can bind to the VEGF competitively with the receptor, hence blocking off the receptor binding sites on VEGF. Naïve peptide phage display techniques have been used to identify disulphide bond constrained 16–20-mer peptides that bind to VEGF and block the interaction with the VEGF receptors.⁴

From the calculations in this report, it has been shown that the peptide studied is not an ideal candidate for a receptor-blocking angiogenesis inhibitor because of the extensive electrostatic imbalance in its bound state with the VEGF molecule.

Method

The exact spatial location of atoms within a complex crystal structure can now be fairly accurately determined by X-ray crystallography. A beam of X-ray is passed through a crystal of the sample, either a pure compound or, as in this case, a bound complex of a protein with its ligand. From the diffraction patterns, using various tools, a final structure can be predicted within acceptable error values. In this final structure, the three-dimensional location of almost all atoms can be predicted and recorded. The standard for presentation of such results is often through pdf files that can be viewed using Molecular Modeling software.

The crystal structure of the VEGF-peptide complex was obtained from the Protein Data Bank and analyzed for the project using Quanta and Grasp as the visualizing softwares.

The crystal structure of the complex had been

refined to an R-value of 19.0 percent (from a free R-value of 27.3 percent) using all the reflections between 20 and 1.9 angstroms. The structure was visualized by Quanta, and the homodimeric VEGF molecule with the two 20-mer peptides bound at either poles of the dimer was identified.

Often there are entire amino acid residues missing from X-ray crystallography structures, and often there is ambiguity in the results that cannot conclusively place a residue in any one out of two equally probable spatial positions. Furthermore, hydrogen atoms are too low in electron density to actually be visualized by X-ray crystallography, and this inability to confirm the location of the hydrogen atoms often leaves ambiguity regarding the correct native titration states of some amino acid residues.

These ambiguities in the determined crystal structure were dealt with in the project by using a program called Charmm. Charmm is able to compare each residue in the given pdf input file to a database of possible residues and is able to build in any reasonable parts of the residues that might be missing, including the missing hydrogen atoms that were undetectable by the X-ray crystallography. It is also able to build in missing residues given the sequence of the protein and the peptide it is taking as an input.

From visualization in Quanta, the peptide binds to the homodimer and interacts with both monomers, though 95 percent of the interfacial interactions come from one monomer and only 5 percent from the other.

The peptide under scrutiny binds with its residues 3 to 8, which have a slightly irregular Beta strand conformation, and residues 89 to 95 of the strand Beta-6 of the first monomer (VEGF-1). Its residues 8 to 10 form a hairpin loop and bring the remaining peptide, residues 11 to 20, into close proximity to bind to residues 79 to 82 of the strand Beta-5 of VEGF-1, residue 48 on Beta-2, residues 38 to 42 of the Alpha-2 region of VEGF-1, and residue Phe 17 on the N terminus of the Alpha-1 region of the VEGF-2.0.1

The coordinate files were visualized by Quanta and then fed into Charmm with appropriate modifications for computation. The original file from the crystallographers contained the homodimer bound to two peptide molecules at the two poles, and my modified file deleted one of the two peptides in order to focus on one binding face. Since both chains of the dimer contributed to binding at the interface, both were retained and named A and B chains, which were to bind with the peptide, or the C chain.

The input file was programmed and modified

accordingly to take into account these three chains. The two chains of the homodimer are disulphide bonded to each other, and this was considered during the modification of the input file. The disulphide bonds and their locations are summarized in Table 1.

Charmm was used to build in the missing residues and the missing hydrogen atoms. The

Table 1: Summarizing Disulphide bonds:

Disulphide bonds are formed between Chains 1 and Chains 2.

#	Chain-1	Residue #	Chain-2	Residue #
1	VEGF-1	26	VEGF-2	68
2	VEGF-1	51	VEGF-2	60
3	VEGF-1	57	VEGF-1	102
4	VEGF-1	60	VEGF-2	51
5	VEGF-1	61	VEGF-1	104
6	VEGF-2	26	VEGF-2	68
7	VEGF-2	57	VEGF-2	102
8	VEGF-2	61	VEGF-2	104
9	Pep-1	7	Pep-1	15

Key: VEGF-1 and VEGF-2 are the monomer chains of the component A in the report. Pep-1 is the 20-mer peptide that is denoted by B in my project.

Table 2: Interfacial residues investigated for their titration states

No.	Residue name-Chain-Residue No.	Interactions.
1	Glu-C5	OE1 is H-bonded to HE1 of Trp C3 OE2 H-bonded to HD1 of His A90 Apparently correct titration state
2	Asp-C10	OD1 is H-bonded to H Tyr C12 and H of Arg C14 OD2 is H-bonded to HE of Arg C14 Best left deprotonated.
3	Asp-C11	OD1 is 2.713angstroms from HE21 of Gln A89 OD2 is 7.924angstroms from H21 of A 48 Even though there are no H-bonds, these spatial positionings confer stability and this titration state seems to be the best.
4	His-A 90	HD1 is H-bonded to OE2 of Glu C5 NE2 is 5.268 angstroms away from H of Ala C8 Hence the titration state is appropriate at HIS and neither HSD nor HSC seems to be supported.
5	Asp A41	OD1 is H-bonded to HH12 of Arg C1 OD2 is H-bonded to HH22 of Arg C1 Hence Asp A41 must remain deprotonated
6	Tyr A39	OH is H-bonded to HT2 of Arg C1 OH is H-bonded to H of Ser A95 Titration state is appropriate
7	Glu-A38	O is 3.167 angstrom away from HH11 of Arg C1 OE1 is H-bonded to Leu A97 OE2 is H-bonded to HH22 Arg A56 Stable interaction at the current titration state.
8	Glu A42	H is H-bonded to H of Tyr of A 45 Titration state seems stable.

Key:

VEGF-1 is chain A
VEGF-2 is chain B
Pep-1 is chain C

result was visualized in Quanta again in order to decide whether the titration states of the amino acids, particularly those at the interface, were accurate or not.

A summary of the residues investigated and the results are appended in Table 2.

The software, Delphi, was then used to run standard electrostatic calculations. Now, the entire homodimer was taken to be part A, and the peptide was part B, so that the interface between A and B is the actual binding interface between the entire homodimer (not taking each monomer chain separately) and the peptide.

Another software, Grasp, was used to map out these results from the electrostatic calculations. From the results, it appears there are some residues that have a net electrostatic potential, which are not entirely self-compensatory after desolvation and protein-peptide interactions. These results are visualized in Figure 1.

Appropriate changes are summarized in List 1 and 2, and visualized in Figures 2 and 3.

Finally, the proposed structures were once again mapped out on Grasp after a Delphi calculation and studied. Figures of the images shown by Grasp were visualized and printed out using Snapshot.

Results

Tables A through D summarize the results of the project.

The electrostatic calculations run on Delphi indicate the net potential on the surface of the molecule both in its bound and unbound states. From this, we can calculate the net surface potentials, due to the desolvation process as well as the interaction between the two molecules that take part in binding.

The red regions on the Grasp image reflect the regions with a net negative charge in each case, and the blue regions indicate the regions with a net positive charge. The degree of intensity of the colors reflects the magnitude of the surface potential.

Interpretations

The figures show snapshots of both the protein molecule (A) and the peptide (B) in three distinct stages of calculations. The images for the desolvation indicate the surface potentials resulting from the desolvation penalty. Hence, regions that are red are the residues, or their parts thereof, that have acquired a net positive charge by losing their interactions with the solvent water molecules, while the residues in the blue regions have acquired a net positive charge as their desolvation penalty.

Whether two molecules will tightly bind to each other depends largely on whether these net surface potentials can be well balanced by the results of interaction between the two molecules. The second stage shows the images of the surfaces of both the protein as well as the peptide due to their interaction with each other. The red regions show the residues that have gained a positive charge as a result of this interaction, while the residues in the blue regions have gained a net positive charge.

For the tightest binding, residues that gained a negative charge due to desolvation should acquire an equal but negative charge during intermolecular interaction to compensate for the desolvation penalty. The net surface potential should be zero in order to achieve the most stable electrostatic configuration.

Discussion

Studying the interactions of this peptide binding to the VEGF molecule is in essence a drug design problem. The VEGF, when bound to its native receptor on the endothelial surfaces, will regulate angiogenesis. In cancerous cells in which inhibiting angiogenesis has been clearly linked to regression of the disease, it is desirable to try to block the receptor binding sites of this VEGF molecule so that it is unable to recognize its native receptor.

In order for the drug to be an effective competitive inhibitor of VEGF-receptor binding, it must have a high specificity as well as a high affinity for VEGF. Ideally, the affinity of VEGF to this peptide must exceed the affinity that the VEGF has for its own native receptor. Only then will it actually bind to this inhibitory molecule instead of the native receptor.

In running the electrostatic calculations for the peptide with the VEGF protein molecule, it is rather obvious that the peptide is far from optimized binding circumstances. In its unaltered form, as put forward by the crystallographers (Figure 1), there is extensive, uncompensated surface potentials during binding. It is unlikely that there are such imbalances in the VEGF-receptor binding complex, in which case, VEGF will preferentially bind to its receptor and the peptide would be rather ineffective as an angiogenesis-inhibitory drug.

However, it is possible to optimize binding conditions from simulations run on these programs, which narrows down our search for the ideal blocker and determines exactly what sort of structure is required for an effective inhibitor.

After extensive changes to the peptide (Figure 3) it appears there are two main residues, Cys B7

and Val B4, which remain largely uncompensated for.

Notably, Cys B7 is involved in a disulphide bond with Cys B15 within the peptide. This interaction appears critical in shaping the peptide so that it can bend over in its hairpin loop and completely bind the entire face of the receptor binding site. It might be well worth the while to replace this Cys B7 with Ser, perhaps, or another similar-size residue to see whether the disruption in disulphide bonding affects the binding properties of the molecules.

In the protein itself, changing the titration state of Glu 93 by protonating it has improved the net surface potentials. Perhaps binding is optimized at a pH where the Glu residue remains protonated. This seems a rather key point to consider during drug design.

Similarly, His 90 is better off electrostatically when it is a charged molecule. This is partly due to the interactions it has with its binding partners of the peptide.

Ile 90 has been changed to give it a more negative charge in order to balance the dark blue regions caused by its electrostatic imbalance. However, it appears that Ile90 is still quite positive and remains unbalanced by interactions. It might be interesting to investigate the residue on the binding peptide and to try to change its charge to make it more negative, thus balancing the net charge.

Therefore, although the peptide being studied might not offer the best angiogenesis inhibitor possible, it does offer a starting point from which we can gain valuable understanding of the characteristics needed in potential drug candidates. Further work may be undertaken to make the necessary alterations to the synthetic peptide molecule and to investigate its effectiveness as an angiogenesis inhibitor through in vitro characterizations.

List 1: Alterations made to the peptide molecule

Arg B1 Appears intensely red in both desolvation as well as interaction. Thus, this residue gains a net positive charge as a desolvation penalty, and there are positive groups interacting when it binds to the protein. This is a very unstable condition and leads to poor binding due to electrostatic repulsions.

Changes: I have changed the charges on the residue in the complex.crd file to see whether it alleviates the situation. In Figures 2 and 3, the difference can be

clearly seen as the net effect is less intensely positive. The charge imbalance has been largely corrected. All atomic charges were changed to zero, and HT2 was given a net negative charge of -0.33. The HH11 was given a charge of -0.33 instead of 0.0000.

Gly B2 The O of Gly B2 was more blue than when it recovered in interaction due to desolvation. I changed the charge on the oxygen to -1.0000 from -0.5500.

Cys B7 Gains a more positive charge from interaction than what it needs to compensate for its desolvation penalty. Hence, making this residue more negatively charged will make the interaction tighter. I changed the charge on O of Cys B7 to -1 from -0.55. However, the rest of the molecule is deep red from the interaction.

Ala B9 Red. This residue is too negative in its desolvation and does not get enough compensatory interactions from binding. By making the charges more positive, N of Ala B9 from -0.4 to 1 and O of Ala B9 from -0.55 to 0.55, I have tried to reduce this effect.

Asp B11 The OD1 of Asp B11 is intensely red, indicating a high negative charge from desolvation that does not interact with the protein to compensate. Thus, it appears that protonating the residue will improve this condition. The original files were altered to take in AS1 instead of ASP while building missing residues in Charmm, and the entire calculations were redone (Figure 2). Indeed, this dramatically improved the results.

Glu B5 This residue was intensely red in interaction, possibly indicating many negative charges surrounding it during interactions that were not needed to compensate for the desolvation penalty. Hence, the imbalance can only be corrected by protonating the B5 Glu, altering the files that were fed into Charmm, and running the entire program over. Indeed, results show that this improves the surface potentials in such a way so as to stabilize the net interactions (Figure 2).

Hence, the two residues, Asp B11 and Glu B5, were altered to accommodate a different titration

References

1. Wiesmann C, Christinger HW, Cochran AG, et al. "Crystal structure of the complex between VEGF and a receptor-blocking peptide." *Biochemistry*, 1998;37:17765-17772.
2. Muller YA, Li B, Christinger HW. "Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site." *Proc. Natl. Acad. Sci. U.S.A.*, 1997;94:7192-7197.
3. Shaheen RM, Tsent WW, Vellagas R, et al. "Effects of an antibody to vascular endothelial growth factor receptor-2 on survival, tumor vascularity, and apoptosis in a murine model of colon carcinomatosis." *International Journal of Oncology*, 2001;18(2): 221-226.
4. Fairbrother WJ, Christinger HW, Cochran AG, et al. "Novel peptides selected to bind vascular endothelial growth factor target the receptor binding site." *Biochemistry*, 1998; 37(51):17754-64.

Acknowledgments

Special thanks to Professor Bruce Tidor, 5.67 IAP, 2001.

state that would optimize binding. However, these residues may not be in this titration state in their natural state. If we are using the peptide molecule as a model for drug candidates for blocking the VEGF from binding to its native receptor and in order to inhibit angiogenesis in cancerous cells, it would be wise to investigate small molecules that exhibit the certain electrostatic characteristics required at these residues in order to maximize strength of binding.

List 2: Alterations made to the protein molecule

In general, most of the residues of the protein seem to be rather well compensated for in the desolvation and interactions. Those residues that gain a negative charge during desolvation are able to neutralize their charges during interactions, and vice versa.

For example, at the interface, Glu A38, Tyr A39, Pro 40, and Asp 41 are all negatively charged due to desolvation, but during their interactions, they are surrounded by positively charged residues from the binding peptide to compensate for this penalty.

Similarly, parts of Glu 93 and Tyr 39 that are blue from desolvation are surrounded by negative charges during interaction to compensate for their net charges.

Residue Interactions

- Ile 91 The O, H, and CB of Ile 90 were more blue than red after interaction and desolvation, and I changed the charges on these atoms to try to reduce the disparity in their interactions.
- Arg 82 Parts of this molecule are more red than blue, while the rest are more blue than red. Hence, the charges on NE and HE were made more negative, while those on CZ and NH1 were made positive to optimize binding conditions.
- Glu 93 OE1 of Glu A93 was red on desolvation, but did not have any corresponding interactions. I have changed its titration state to G11, so that it does not gain a large negative charge on desolvation. The net effect is to reduce the imbalance in the surface potential (Figure 3).
- His 90 The net charge on His 90 was negative, and in order to counter this effect, its titration state was converted to HSC to impart a positive charge, to compensate for the negative charge that it acquires upon interaction with the peptide molecule. 