

Molecular Modeling and Conformational Analysis of MuSK protein

Vasilis Haidinis, Georgios Dalkas, Konstantinos Poulas, Georgios Spyroulias
Department of Pharmacy, University of Patras, Rio-Patras, 26504, GREECE

Abstract. Muscle-specific kinase (MuSK) is a crucial receptor tyrosine kinase required for the development and function of neuromuscular junction. Although many protein domains have already been modeled with crystallographic techniques in various organisms, a single model for the whole human structure is not yet available. A model of the entire protein was constructed by using two parallel Homology Modeling approaches, one unsupervised and one driven by the user. In addition, by applying Molecular Dynamics (MD) simulations the present study provides further insights on the structure, and the intermolecular interactions of the protein were examined. The expected semi rigid globular form of the protein was confirmed and in addition a hydrophobic core and a hydrogen bond network that enhances the stability of the molecule were observed. Furthermore, these calculations identified an intriguing rotation of Ig domains and this finding sets the base for additional hypothesis and further investigation.

Keywords: homology modeling, molecular dynamics, MuSK, neuromuscular junction.

1 Introduction

The neuromuscular junction (NMJ) is a specific synapse whose function consists in transmitting electrical activity from the nerve endings in skeletal muscle [1]. In recent years there have been many efforts to understand the function of the NMJ through various experimental techniques, one of them being X-ray crystallography. MuSK protein is a single pass transmembrane polypeptide located in the area, considered to be an organizer in a series of interactions and signaling pathways involved in the creation and functioning of the NMJ. Mutations of MuSK are responsible for diseases and disorders of the NMJ, such as congenital myasthenic syndromes, while absence of MuSK results in non formation of NMJ [2,3].

The extracellular region of MuSK consists of four globular domains: three N-terminal immunoglobulin-like (Ig) domains, and a C-terminal cysteine - rich region similar to the cysteine-rich domain of Frizzled, the receptor for Wnt [4,5]. Agrin, is a ligand that is synthesized and released by motor neurons and is usually detected in synaptic basal lamina. MuSK phosphorylation in muscle cells is triggered by Agrin.

However, there is no evidence that the two proteins interact directly. A hypothetical molecule, MASC (myotube-associated specificity component), is proposed as a binding partner of agrin to transduce signals to MuSK [6]. Recent studies suggest that this hypothetical molecule may be the low-density lipoprotein receptor-related protein 4 (Lrp4) [7]. Lrp4 binds to neural agrin and forms a complex with MuSK in a manner that is stimulated by agrin. MuSK induces clustering of acetylcholine receptors (AChRs) via Dok-7 and Tid1. Dok-7 interacts with the cytoplasmic domain of MuSK and induces its autophosphorylation [8], while Tid1 orchestrates the cascade that leads to AChR clustering, by interactions with other regulatory proteins downstream of MuSK.

The tertiary structure prediction is a difficult process and there is not a single method that can satisfy all those characteristics needed to eventually have a successful outcome. It is a complex and technically demanding method that requires computational techniques and there have been developed three main categories of algorithms. The first one is Homology Modeling, which uses known protein structures as templates in order to create models of unknown structures and is based on the amino acid sequence homology between the two proteins. The second, known as Threading or Fold Recognition, searches for structures with similar folding, without any obvious sequence similarity, while the third, known as *Ab initio* structure prediction and *de novo* modeling, predicts folding based on physical chemistry principles [9].

In the present study homology modeling and Molecular Dynamics simulations were employed in order to gain structural and functional information about MuSK. This information can be used to determine, for the first time, a model that represents the conformational properties of the protein and the interactions between the different domains of Musk.

2 Computational Methods

2.1 Homology Modeling

Protein sequence of the human MuSK was acquired from the SWISS-PROT [10] database (O15146). Since there are three different isoforms of MuSK sequence, produced by alternative splicing, the canonical sequence or Isoform 1 was used for model building. Two different methods were used in parallel for the molecular building, and the models with the highest score were selected and analyzed. Models were subjected to energy minimization and those with the lowest energy were selected. Using the first approach, tools such as MODELLER [11] and SWISS-MODEL [12] were used, both attempting to create through an unsupervised way, based only on the protein amino acid sequence, the final model of the protein. In the second method, by using the online tool BLAST [13], a search was implemented for those structures that have the highest percentage of identity and which will be subsequently used as templates for the creation of our model. The sequences of these structures are manually aligned to the amino acid sequence of the target protein and then initial models are constructed.

The models created for the target protein through the automated methods, as described above, failed to provide a structure for the entire extracellular domain of MuSK, since the proposed structures were not in accordance with the available published research data. For this reason, we decided to construct the model based on the second method.

The crystal structure of immunoglobulin-like domains 1 and 2 of the mouse MuSK protein (*Rattus norvegicus* PDB ID: 2IEP) [5] was used as a template to create the homology model of the immunoglobulin-like domains 1 and 2 of human MuSK. The sequence alignment revealed a 99% identity. Additionally the crystal structure of N-terminal fragment of human TAX1 (PDB ID: 2OM5) [14] was used as a template to create the homology model of the immunoglobulin-like domain 3 while the crystal structure of the frizzled-like cysteine-rich domain of mouse MuSK (*Rattus norvegicus* PDB ID: 3HKL) was used to create the corresponding homology model of the CRD domain of human MuSK. The sequence alignment identity between 2OM5 and our target sequence is 35%, while the corresponding identity for 3HKL is 99%.

A ten-amino acids linker peptide is between the three Ig-like domains and the cysteine-rich domain (CRD). This small part of the molecule was modeled using as templates the crystal structures of proteins PDB ID: 2A3N and PDB ID: 1JW1 [15] which show 60% identity with 100% sequence coverage and 75% with 80% coverage, respectively. Having all these partial models of the domains of MuSK we assembled the final model using Pymol [16].

All sequence alignments were performed using ClustalW2 [17]. All conformers were submitted to energy minimization and the quality of the models was assessed using computational metrics and suggestions from the literature. Validation of the model was performed using the Structural Analysis and Verification Server [18]. SAVES server allows us to check the protein structure with different quality assessment software (Procheck, What_Check, Errat, Prove and Verify_3d). Procheck returned a satisfactory result indicating that 84,8% of residues exhibit backbone dihedral angles (ϕ/ψ) in most favored regions, while only five residues bear dihedral angles in disallowed regions 1,3%. Those five residues were inherited to our model from the original structures that were used as templates.

2.2 Preparation of the model

All crystallographic water molecules were discarded from the initial models. Missing heavy and hydrogen atoms were added using the XLEaP module of AMBER 9 [19]. All the basic residues were protonated and all the acidic were deprotonated. Histidine protonation states were manually set by examining their potential for hydrogen bonding with surrounding residues, and by taking into account the pKa values calculated using the H++ web server [20]. The improved protein backbone AMBER force field, denoted as ff99SB, was employed for all protein atoms [21,22]. A set of minimization steps was carried out using the SANDER module of AMBER 9 for the relaxation of the models and for the optimization of the positions of the new atoms. In particular, 1000 steps were performed in order to allow relaxation of the hydrogen atoms only, using the steepest descent algorithm. Then, 1000 steps of conjugate gradient minimization were carried out by restraining all C α atoms using

positional harmonic restraints of 50 kcal mol⁻¹ Å⁻² force constant. A third round of 3000 unrestrained minimization steps was carried out. The generalized Born implicit salvation model (GB^{HCT}) [23] was employed throughout the energy minimization steps, with a 16- Å cutoff for the non bonded interactions.

2.3 Molecular dynamics (MD)

All the MD calculations were carried out on Intel Xeon workstations using AMBER 9. Numerical integration was performed with a 2-fs time step and all bonds involving hydrogen atoms were constrained with SHAKE [24]. Periodic boundary conditions were imposed using the particle mesh Ewald method [25] with 8.0 Å limit for the direct space sum. Temperature and pressure controls were imposed using a Berendsen-type algorithm [26] both with 1 ps coupling constants. The model was immersed in isometric truncated octahedron TIP3P-water boxes [27] and the appropriate number of counter ions was added to neutralize the total charge. In order to optimize the position of the solvent molecules and equilibrate the temperature and pressure of the systems, we executed the following preliminary stages of simulation: (a) energy minimization for 1000 steps using the steepest descent method with harmonic restraints of 50 kcal mol⁻¹ Å⁻² force constant on all solute atoms; (b) restrained constant volume dynamics (NVT ensemble) at 300 °K for 30 ps; (c) a second energy minimization for 1000 steps with 10 kcal mol⁻¹ Å⁻² restraints on all protein Ca⁺² atoms; (d) temperature was then gradually increased to 300 °K within six rounds of 5-ps constant volume dynamics (NVT), while solute atoms were restrained with kcal mol⁻¹ Å⁻²; (e) restraints were then gradually released within 20 ps in the NVT ensemble at 300 °K; (f) the density of the systems was increased to 1.0 g cm⁻³ during 150 ps of constant pressure dynamics (NPT ensemble). Subsequently, production runs were carried out under physiological conditions (300 °K, 1 atm) in the NPT ensemble for a total time of 10 ns.

Processing of the MD trajectories was performed using the PTRAJ module of AMBER 9, while VMD 1.8.7 [28] was used for their examination. MD trajectories were analyzed to identify important intermolecular interactions by extracting their geometric features (distances and angles) as a function of simulation time. Hydrogen bonding interactions were monitored using a 3.4 Å distance cutoff and 120° as the angle cutoff. Hydrophobic interactions were included for a pair of carbon atoms separated by a distance less than 4.0 Å. Only interactions present for more than half of the simulation time were considered.

3 Results

Before describing the results of MD simulations of MuSK protein, we have to mention the points that were taken under consideration in order to rationalize the results obtained so far. First, we have to take into account that the final MuSK model was created from the conjunction of models of different domains which were

constructed separately. The molecule, when located into the cell, is very close to the transmembrane helix, which possibly enhances both the stability of the molecule through interactions with the cell surface and its tendency of dimerization. In this study the molecule is not anchored to the cell surface or any transmembrane helix, which can lead to different behavior due to greater degree of freedom in its movements and its interactions.

Results of MD simulations confirmed that although protein retains its semi linear conformation almost throughout the simulation, the Ig domains, in comparison to their original position, appear to be rotated approximately by 90 degrees [Figure 1]. Furthermore within any Ig domain, this spatial change is followed by a 45 degrees bend. The motion mentioned above, could be due to efforts of the molecule to occupy a more favorable energy state in order to be able to make interactions with either another protein or another molecule of MuSK. The Ig1 domain is considered to be involved in the creation of a MuSK dimer and is the part of the protein that interacts with Agrin and Lrp4. Therefore, an increased mobility is expected.

The Ig domains have almost identical secondary structure. They consist of 8 b-strands (only the Ig2 region has 10), which create a globular structure of beta-sheets. Hydrogen bonds are formed between parallel and antiparallel beta-strands and contribute to the stability of the domain. The residue Asn71 of Ig2 domain forms hydrogen bonds with residues Met96, Pro125 and Gln174 [Figure 2]. The side chain of this amino acid acts as a link that holds the upper regions of the Ig domain. A similar bond is observed between residues Asn80, Lys3 and Val82 in Ig1 domain and between Asn263, Gly266 and Pro216 in Ig3. Additional important hydrogen bonds are formed between residues Trp131 and Gly150 in Ig2 domain and Trp38 and Gly58 in Ig1 respectively. Within Ig2 domain, a favorable salt bridge was formed between the guanidinium group of Arg143 and the carboxylic group of Asp161.

Molecular modeling results suggest that the core of the Ig domains is highly hydrophobic. During the simulation, conformational changes of the molecule seem to allow residues located on the surface to participate in hydrophobic interactions and van der Waals bonds. At the starting point of simulation these residues have their side chains oriented toward solvent. As the simulation progresses and in correlation with the Ig domains rotation, some of these residues turn their side chains toward the interior of the molecule. In many cases they interact with other hydrophobic residues by forming hydrophobic and van der Waals interactions. Specifically, potential hydrophobic and van der Waals contacts could be formed between Val18 and Met96 as well as between Leu17 and Leu173 [Figure 3]. This behavior may be due to the lack of another molecule which would provide residues for interaction e.g. another molecule of MuSK in case of dimerization, or due to the effort of the molecule to stabilize its conformation. Similarly, the lack of another molecule could explain the conformational change of the CRD domain during the course of the MD simulations. In its initial state the CRD domain displays a compact structure, with the helices a1, a2 and a3 very close to each other. The other residues of CRD, consisted mainly loops, are solvent-exposed and they participated in water-mediated interactions. As MD progresses, distance between helices increase and as a result more residues are solvent-exposed.

4 Discussion

Based on crystal structures, we have generated *in silico* the model of human MuSK protein. Two different approaches of Homology Modeling were followed and the conformational dynamics properties of MuSK were studied through molecular dynamics simulations. In the first approach all protein models were constructed automatically providing only the target amino acid sequence. This approach failed to deliver a satisfactory result mostly due to contradictions with published scientific literature. In the second approach we provided the structures that would be used as templates based on sequence alignment identity. Due to the inability of finding a template for the whole extracellular region of MuSK, we created partial models of the protein domains. Assembly of these models led to the construction of the complete model.

Simulation data analysis confirmed the preservation of a linear semi-rigid conformation of Ig domains. We have observed an interesting rotation of Ig domains, followed by a bend of Ig1. Although the above motion is of great interest, it is not easy to confirm if this movement is due to the effort of the molecule to occupy a most favorable energy position (in order to be able to interact with other neighboring molecules), or due to initial disposition of the model.

The model exhibits enhanced structural stability, due to a network of hydrophobic interactions, hydrogen bonds and salt bridges within the structure. In addition a number of hydrogen bonds between parallel and antiparallel b-strands on its surface were observed.

The creation of alternative MuSK models, based on different orientation of its domains, and the study of interactions between MuSK and other proteins such as Agrin and Lrp4 are under investigation.

References

1. Hughes BW, Kusner LL, Kaminski HJ.: Molecular architecture of the neuromuscular junction, *Muscle Nerve*, 33, 445-61 (2006).
2. DeChiara TM, Bowen DC, Valenzuela DM, Simmons MV, Poueymirou WT, Thomas S, Kinetz E, Compton DL, Rojas E, Park JS, Smith C, DiStefano PS, Glass DJ, Burden SJ, Yancopoulos GD.: The receptor tyrosine kinase MuSK is required for neuromuscular junction formation in vivo, *Cell*, 85, 501-512 (1996).
3. Werner Hoch, John McConville, Sigrun Helms, John Newsom-Davis, Arthur Melms & Angela Vincent: Auto-antibodies to the receptor tyrosine kinase MuSK in patients with myasthenia gravis without acetylcholine receptor antibodies *Nature Medicine* 7, 365 – 368 (2001).
4. Amy L. Stiegler, Steven J. Burden and Stevan R. Hubbard.: Crystal Structure of the Frizzled-Like Cysteine-Rich Domain of the Receptor Tyrosine Kinase MuSK, *J. Mol. Biol.* 393, 1–9 (2009).
5. Amy L. Stiegler, Steven J. Burden and Stevan R. Hubbard.: Crystal Structure of the Agrin-responsive Immunoglobulin-like Domains 1 and 2 of the Receptor Tyrosine Kinase MuSK, *J. Mol. Biol.* 364, 424–433 (2006).

6. Glass DJ, Bowen DC, Stitt TN, Radziejewski C, Bruno J, Ryan TE, Gies DR, Shah S, Mattsson K, Burden SJ, DiStefano PS, Valenzuela DM, DeChiara TM, Yancopoulos GD: Agrin acts via a MuSK receptor complex. *Cell* 85: 513–523 (1996).
7. Kim N., Stiegler, A. L., Cameron, T. O., Hallock, P. T., Gomez, A. M., Huang, J. H., Hubbard, S. R., Dustin, M. L. and Burden, S. J.: Lrp4 is a receptor for Agrin and forms a complex with MuSK. *Cell* 135, 334-342 (2008).
8. Okada, K., Inoue, A., Okada, M., Murata, Y., Kakuta, S., Jigami, T., Kubo, S., Shiraishi, H., Eguchi, K., Motomura, M., et al.: The muscle protein Dok-7 is essential for neuromuscular synaptogenesis. *Science* 312, 1802–1805. (2006).
9. Xiang, Z., : Advances in Homology Protein Structure Modeling, *Curr Protein Sci.* 7, 217-227 (2006).
10. Bairoch A, Boeckmann B, Ferro S, Gasteiger E.: Swiss-Prot: juggling between evolution and stability. *Brief Bioinform* 5:39–55 (2004).
11. Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, Pieper U, Stuart AC, Marti-Renom MA, Madhusudhan MS, Yerkovich B, Sali A.: Tools for comparative protein structure modeling and analysis. *Nucleic Acids Res* 31:3375–3380 (2003).
12. Arnold, K., Bordoli, L., Kopp, J., Schwede, T.: The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22 (2), 195-201 (2006).
13. Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J.: Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410 (1990).
14. Mortl M., Sonderegger P., Diederichs K., Welte W.: The crystal structure of the ligand-binding module of human TAG-1 suggests a new mode of homophilic interaction *Protein Sci.* 16: 2174-2183 (2007).
15. Kumar P., Yadav S., Singh T.P: Crystallization and structure determination of goat lactoferrin at 4.0Å resolution: A new form of packing in lactoferrins with a high solvent content in crystals *Indian J. Biochem. Biophys.* 39: 16-21 (2002)
16. DeLano, W.L. The PyMOL Molecular Graphics System DeLano Scientific, San Carlos, CA, USA (2002).
17. Thompson, J. D.; Higgins, D. G.; Gibson, T. J.; CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22 (22), 4673-4680 (1994).
18. Structural Analysis and Verification Server. <http://nihserver.mbi.ucla.edu/SAVES/>.
19. Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., Woods, R. J.: The Amber biomolecular simulation programs. *Journal of Computational Chemistry* 26 (16), 1668 -1688 (2005).
20. Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res* 33:W368–W371 (2005).
21. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M.; Ferguson, D. M., Spellmeyer, D. C. Fox, T., Caldwell, J. W., Kollman, P. A., A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *Journal of the American Chemical Society*, 117 (19), 5179-5197 (1995).
22. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C., Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins Structure Function and Bioinformatics*, 65 (3), 712-725 (2006).
23. Tsui, V., Case, D. A., Theory and applications of the generalized Born solvation model in macromolecular Simulations. *Biopolymers*, 56 (4), 275-291 (2000).
24. Miyamoto, S., Kollman, P. A., Settle an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *Journal of Computational Chemistry*, 13 (8), 952-962 (1992).

25. Cheatham, T. E., Miller, J. L., Fox, T.; Darden, T. A., Kollman, P. A., Molecular Dynamics Simulations on Solvated Biomolecular Systems - the Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, Rna, and Proteins. *Journal of the American Chemical Society*, 117 (14), 4193-4194 (1995).
26. Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., Dinola, A.; Haak, J. R., Molecular Dynamics with Coupling to an External Bath. *Journal of Chemical Physics*, 81 (8), 3684-3690 (1984).
27. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics*, 79 (2), 926-935 (1983).
28. Humphrey, W., Dalke, A., Schulten, K., VMD: visual molecular dynamics. *J Mol Graph*, 14 (1), 33-38, 27-38 (1996).

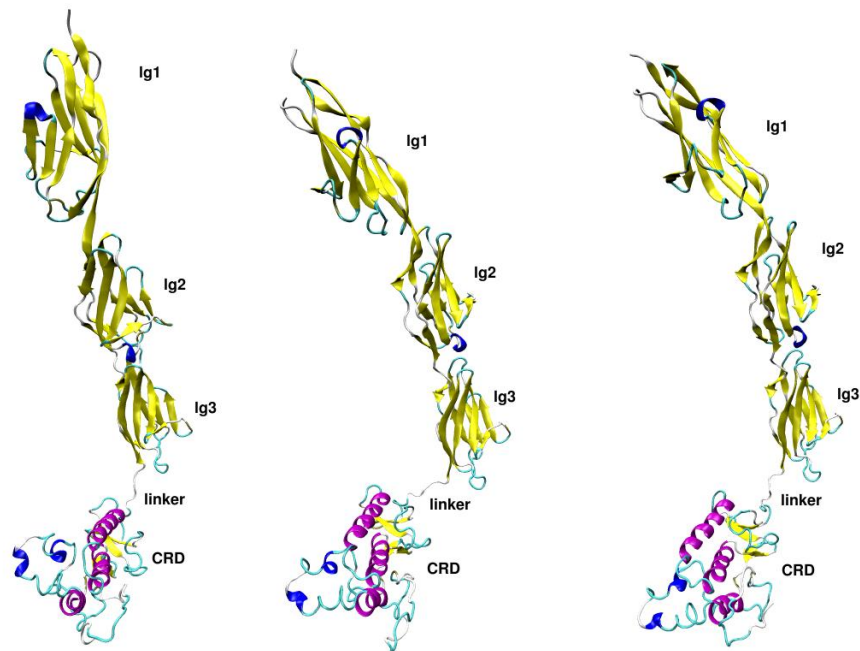


Fig. 1. Ribbon representation of MuSK protein. The image above shows the gradual “unpacking” of the CRD domain and the rotation of the Ig domains. The image was created using the VMD [28] analysis software.

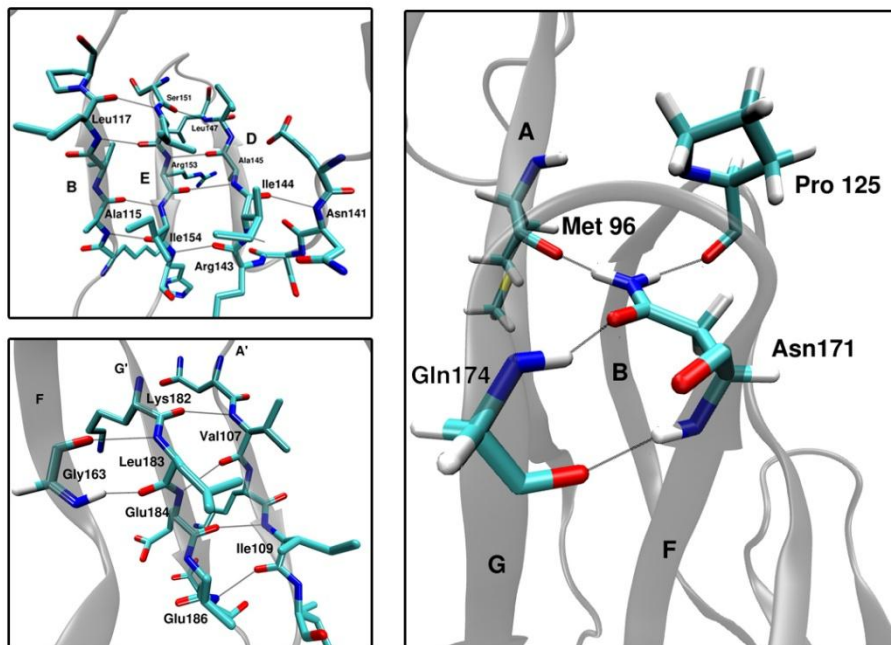


Fig. 2. Hydrogen bond formation in Ig domains of MuSK model. The image was created using the VMD [28] analysis software.

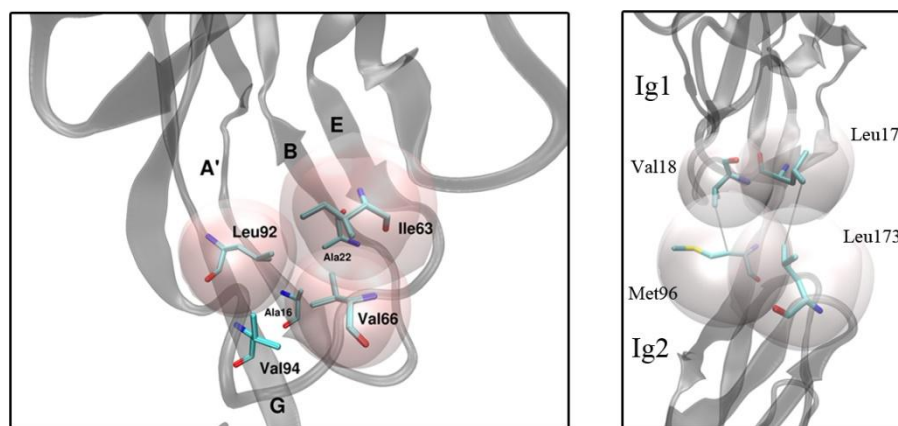


Fig. 3. Possible van der Waals interactions. The image was created using the VMD [28] analysis software.