Proof-of-Concept In-Silico Design of a Small Molecule Inhibitor of the PCDH1 Surface Protein in order to Combat Andes Orthohantavirus Infection

October 19th, 2022

Sai Sahvir Bhaskaruni

Abstract

The Andes orthohantavirus (ANDV) is a strain of virus from the Hantavirus family which is known to cause Hantavirus Cardiopulmonary Syndrome, an upper respiratory infection. As there is currently no FDA-approved treatment for ANDV, an infection which has a 35 to 50% mortality rate, it is imperative that a treatment be developed to address this issue. The ANDV is known to bind to human cells via the glycoprotein c on the virus fusing with the extracellular cadherin-repeat domain one (EC1 domain) on the PCDH1 receptor. PCDH1, a type one membrane protein, is a protocadherin protein with a primary function in cell-to-cell adhesion. To combat ANDV infection, we have targeted cellular uptake by screening for a small molecule inhibitor of the PCDH1 EC1 domain. A crystal structure of the PCDH1 EC1 domain has been previously determined and herein used for in silico evaluation of its binding to compounds that mimic the binding site of the ANDV glycoprotein c. It is expected that an inhibitor with strong binding to PCDH1 can be used for further testing in vivo and in vitro.

Introduction

The recent pandemic has increased awareness of pathogens that pose a distinct threat to public health. One family of these pathogens are hantaviruses, which the WHO has classified as an emerging public health threat , largely due to the fact that they currently lack a broadly approved treatment or vaccine to combat their infection (Hangaragi et. al, 2020). Hantaviruses are a family of enveloped, negative-sense RNA viruses in the family Bunyaviridae. In humans, hantaviruses can cause two types of diseases: Hantavirus Cardiopulmonary Syndrome (HCPS), and Hemorrhagic Fever with Renal Syndrome (HFRS). HCPS has a fatality rate of around 40%, and HFRS has a fatality rate 15% (Liu et al., 2020). HCPS is native to the New World, and therefore virus strains that cause HCPS are called New World Hantavirus.

The most threatening New World hantavirus is the ANDV. With a mortality rate between 35% and 50%, it is imperative that a treatment be developed to address this issue (Vial et al., 2020).

The genome of the hantavirus has three strands: the small (S) strand, the medium (M) strand, and the large (L) strand. The S strand encodes the nucleocapsid protein, the M strand encodes the external glycoproteins, and the L strand encodes viral polymerase (Muyangwa et al., 2015). The external glycoproteins are the main factor that enables virus-membrane fusion (Mittler et al., 2017), and therefore could be the target of a small molecule inhibitor.

Before an inhibitor can be designed, the virus-membrane fusion process has to be studied. Multiple previous studies have identified the ectodomain 1 (EC1) of the protocadherin 1 (PCDH1) protein as the surface protein involved in membrane fusion of New World Hantavirus, particularly the ANDV (Hepojoki et al., 2010). PCDH1, a type one membrane protein, is a protocadherin protein with a primary function in cell-to-cell adhesion. It was found that the

n-terminus of the G_N glycoprotein was responsible for starting the process of virus-membrane fusion (Hepojoki et al., 2010).

The n-terminus of the G_N glycoprotein was used as a baseline to design a small molecule inhibitor using the precedent set by (Brogi et al., 2020). The design and editing of the molecule was done through PyMol, the docking studies were done by AutoDock Vina on the UCSF Chimera program, and further analysis was done using ChemDoodle.

Methods

The first step conducted in the method was to recreate the binding site of ANDV glycoprotein C in order to determine the rudimentary framework of the small molecule. Based on the research done by Slough et. al, the backbone would have to be either threonine-isoleucine-tyrosine (encoded TIY) or threonine-isoleucine (encoded TI) since Hepojoki et. al found that the *n*-terminus of the G_C ANDV glycoprotein interacts with PCDH1. As a small molecule inhibitor is by nature two or three amino acids long, the sequences from the *n*-terminus had to be either TI or TIY. These two compounds were constructed on PyMol and then docked to the EC1 on PCDH1 using Autodock Vina on UCSF Chimera. These results are shown in **Figures 1 and 2**.

Based on the results of the previous step, additions were added to maximize hydrogen bonds. The first simulation completed in this step was the addition of oxygens on the free hydrogens in the molecule, and the next step was the addition of nitrogens on the free hydrogens in the molecule. The hydrogen bonds created by these two simulations were tracked to determine ideal nitrogen and oxygen placement for the compound library, as were the ΔG values indicating binding affinity. These results are shown in **Figures 3 and 4**.

That data was then compiled, and the most efficacious, measured by prominence in bonding, additions to the backbone were added in accordance with Lipinski's Rule of 5, the guiding principle which helps define druggable compounds. Those compounds which had a lower ΔG than the backbone, signifying a higher binding affinity, were then analyzed to record the octanol: water partition coefficient LogP, molecular weight, and conformational change in the PCDH1 protein. Significant conformational change in PCDH1 indicates a high possibility of secondary effects on the cell; those trials which showed a large amount of conformational change were eliminated. ChemDoodle was used to determine LogP values, PyMol was used to determine molecular weight, and UCSF Chimera was used to determine conformational change in the PCDH1 protein. These results are shown in **Figure 5** and **Supplementary Figures 1-10**.

Results

The TI docking and TIY docking compounds, shown in **Figures 1 and 2**, revealed that TIY had a higher ΔG, gibbs free energy change, value than TI. This was strongly indicative of TIY binding more strongly to the PCDH1 EC1 domain. For this reason it was chosen as the backbone for the small molecule inhibitor candidates designed below.

The Oxygen and Nitrogen diagnostic docking was conducted to determine which additions would be the most beneficial in small molecule inhibitors. Through the analysis of the results of the diagnostic dockings it was determined that Oxygen additions would be more beneficial as they resulted in the creation of more H-bonds. A larger abundance of H-bonds results in a stronger binding strength and therefore a better inhibitor.

Through the analysis of the H-bonds created by the Oxygen Diagnostic Docking it was found that additions in the tail of the isoleucine amino acid were most beneficial in creating H-bonds. However the tail of the threonine and the ring of the tyrosine amino acids also had an

abundance of H-bonds created when oxygen additions were made. The results of the Nitrogen Diagnostic Docking were similar, however the number of H-bonds created by the Oxygen additions outweighed the the number created by the Nitrogen additions.

Lipinski's Rule of 5 limits molecules to have 5 H-bond acceptors, and since the TIY molecule had 2 H-bond acceptors only three additions were made from construction trials 1-8. The only viable candidate from those trials was Construction Trial 8 shown in **Supplementary Figure 1**. From construction trials 9-21 the natural H-bond donors were replaced with atoms that made it so that the atoms at that position were no longer H-bond donors, and thus 5 additions were able to be made which improved the binding affinity. From this group several Construction Trials were successful, they are shown in **Supplementary Figures 2 - 7**. For Construction Trial 22 only 4 H-bond donor additions were made instead of the five in Construction Trial 9-21, however the other edits remained the same. Construction Trial 22 is shown in **Supplementary Figure 8**. Construction Trials 23-25 reverted back to the format of Construction Trials 9-21. Only Construction Trials 24 and 25 were successful and they are shown in **Figure 5** and **Supplementary Figure 9** respectively. Construction Trial 24 was the most successful as it had the highest absolute ΔG value. Construction Trials 26-30 involved the additions of nitrogens instead of oxygens, but still followed the same format as Construction Trials 9-21. This was conducted to determine if the less than optimal results of the Nitrogen Diagnostic Docking were only extrapolated to that test. Of this group, only 1 successful trial emerged, Construction Trial 28, which is shown in **Supplementary Figure 10**.

Table 1: Summary of Construction Trial

Molecule Name	Highest ΔG	Molecular Weight	LogP
Construction Trial 8	-5.9	381.397	2.601
Construction Trial 9	-6	413.3958	1.558
Construction Trial 10	-6.4	413.3958	1.248
Construction Trial 12	-6.3	413.3958	0.744
Construction Trial 13	-6.1	413.3958	1.285
Construction Trial 14	-6	413.3958	1.054
Construction Trial 17	-5.9	413.3958	0.397
Construction Trial 22	-6.2	397.3964	1.099
Construction Trial 24	-6.5	413.3958	0.069
Construction Trial 25	-6.3	413.3958	0.152
Construction Trial 28	-5.8	408.472	-0.733

Highest ΔG	-5.3
Molecular Weight	214.2615
LogP	0.513

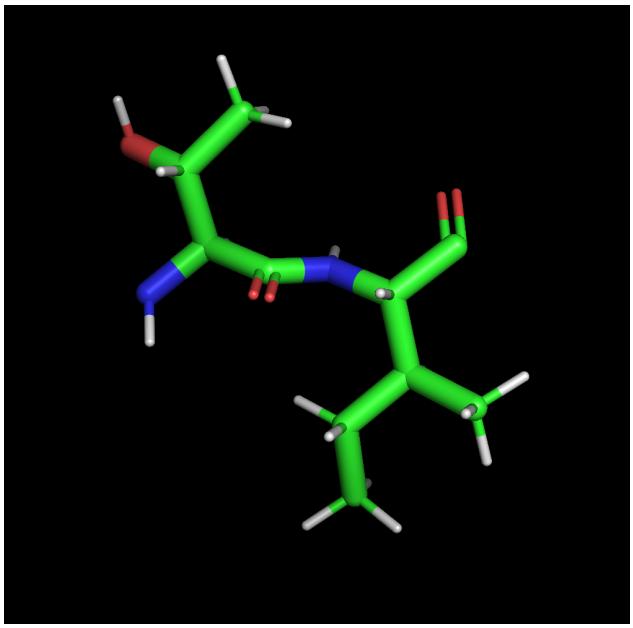


Figure 1: TI Docking
G_C ANDV glycoprotein n-terminus threonine-isoleucine (TI) peptide analog used to verify rudimentary framework of antagonist compound.

Highest ΔG	-5.7
Molecular Weight	377.4348
LogP	1.996

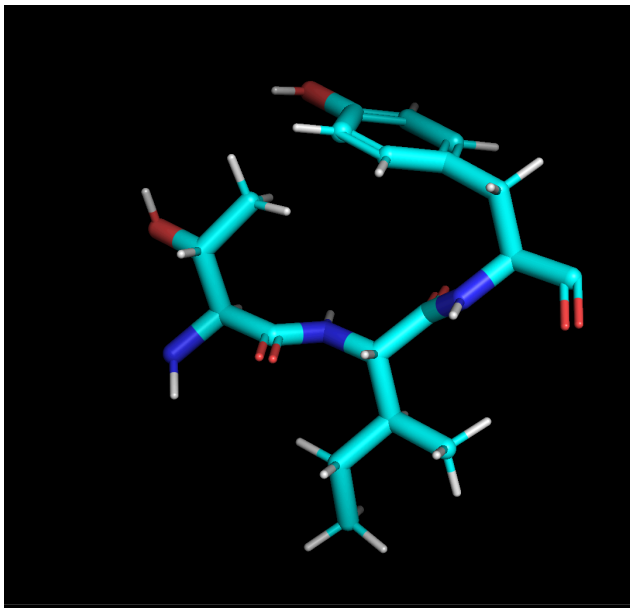


Figure 2: TIY Docking
G_C ANDV glycoprotein n-terminus threonine-isoleucine-tyrosine (TIY) peptide analog used to verify rudimentary framework of antagonist compound.

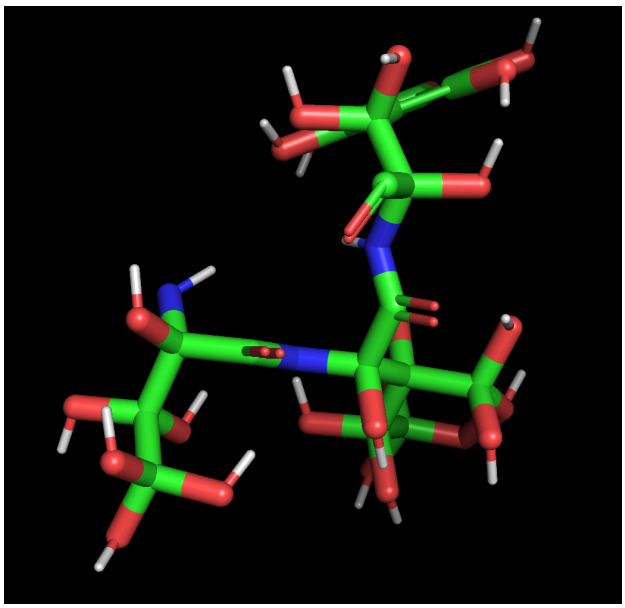
Highest ΔG	-6.3	
Molecular Weight	729.4216	
LogP	-17.045	

Figure 3: Oxygen Diagnostic Docking
A TIY peptide analog with all available atoms as oxygens.

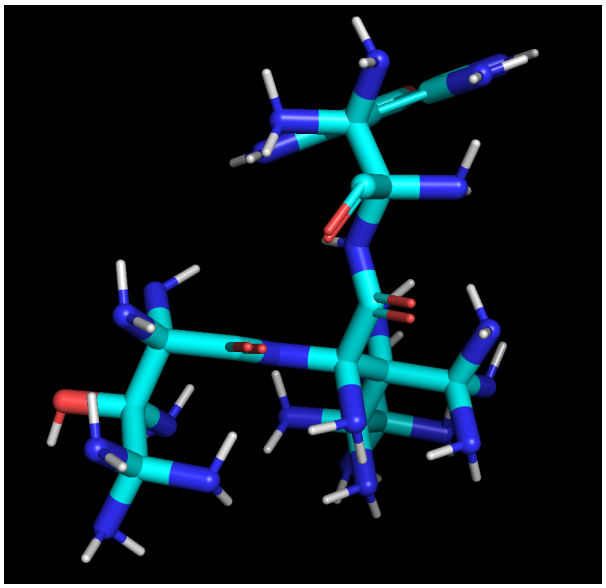
Highest ΔG	-6.2	
Molecular Weight	707.7569	
LogP	-17.755	

Figure 4: Nitrogen Diagnostic Docking
A TIY peptide analog with all available atoms as nitrogens.

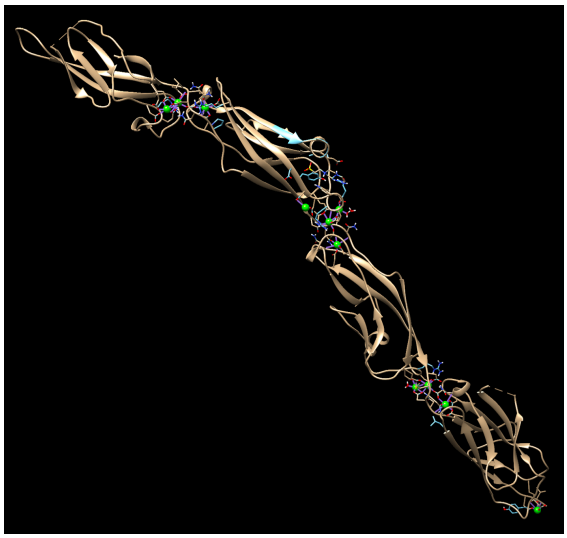
Highest ΔG	-6.5	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	0.069	

Figure 5: Construction Trial 24
Candidate compound with strongest *in silico* affinity to PCDH1

Discussion

The results of this study indicate a significant possibility to create an anti-ANDV drug candidate. Eleven qualified candidate compounds were identified through thirty trials. This indicates a high possibility for more candidate compounds to come from the modification and optimization of the eleven candidates represented by **Supplementary Figures 1-10.** Understanding that candidate compound libraries are tens of thousands large, the purpose of the methodologies described serves as a pilot for further development of compounds. Further, success within *in silico* testing does not justify development of a compound, as many factors play a role in success of a drug. This includes first and foremost drug efficacy and potency. Although drug affinity can be measured using UCSF Chimera, the elucidated effect of the drug is difficult to quantify without *in vitro* testing. I have approximated the elucidated effect qualitatively by

visualizing receptor conformational change, but even this is inaccurate as other factors, such as cytoplasmic-side anchoring or signaling proteins, play a role in the conformational change of receptors. Secondly, success of a drug is defined by its pharmacokinetic profile: absorption, distribution, metabolism, and clearance. Without knowing the status of the drug within the body, it cannot progress clinically. Finally, off-target effects have a major role in determining progression of a drug candidate, as many target proteins play multi-faceted roles beyond what is attempting to be antagonized. Namely, PCDH1 is involved in neural cell adhesion and has been linked with neural adhesion. Therefore, introducing an antagonist may present off-target neurological effects. However, in acknowledging such obstacles, the findings of this study initiate the discovery pathway for the full design of anti-ANDV drug candidates.

References

Brogi, S., Castro-Ramalho, T., Kuca, K., Medina-Franco, J.L., & Valko, M. (2020, August 7) *In silico* Methods for Drug Design and Discovery. *Frontiers in Chemistry*, https://doi.org/10.3389/fchem.2020.00612

Hangaragi, P. S. (2020, June 14) Hantavirus: An emerging global threat, *Asian Journal of Oral Health and Allied Sciences, 10(4),* 1-5. https://doi.org/10.25259/AJOHAS_6_2020

Hepojoki, J., Strandin T., Vaheri, A., Laniken, H. (2020, January 1) Interactions and Oligomerization of Hantavirus Glycoprotein, *Journal of Virology, 84(1)*. https://doi.org/10.1128/JVI.00481-09

Liu, R., Ma, H., Shu, J., Zhang, Q., Han, M., Liu, Z., Jin, X., Zhang, F., & Wu, X. (2020, January 30) Vaccine and Therapeutics against Hantaviruses. Frontiers in Microbiology, 10, 2989. https://doi.org/10.3389/fmicb.2019.02989

Mittler, E., Dieterle, M.E., Kleinfelter, L.M., Slough, M.M., Chandran, K., & Jangra, R.K. (2019, August 17) Hantavirus entry: Perspectives and recent advances. Advances in Virus Research, 104(6), 185-224. https://doi.org/10.1016/bs.aivir.2019.07.002

Muyangwa, M., Martynova, E. V., Khaiboullina, S. F., Morzunov, S. P., & Rizvanov, A. A. (2015, November 27). Hantaviral Proteins: Structure, Functions, and Role in Hantavirus Infection. Frontiers in Microbiology, 6, 1326. https://doi.org/10.3389/fmicb.2015.01326

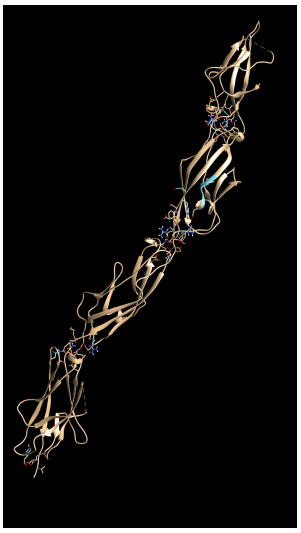
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry, 25(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Shaker, B., Ahmad, S., Lee, J., Jung, C., & Na, D. (2021 October) In silico methods and tools for drug discovery. *Computers in Biology and Medicine, 137(104851)* https://doi.org/10.1016/j.compbiomed.2021.104851

Slough, M.M., Herbert, A.S., Kuehne, A.I., Dye, J.M., Chandran, K., & Jangra, R.K. (2020 June 3) Mapping the Interface between New World Hantaviruses and Their Receptor, PCDH1. *Proceedings 2020, 50*(1), 8. https://doi.org/10.3990/proceedings2020050008

Todsen, W.L. (2014 July 29) ChemDoodle 6.0 *Journal of Chemical Information and Modeling, 54*(8), 2391-2393. https://doi.org/10.1021/ci500438j

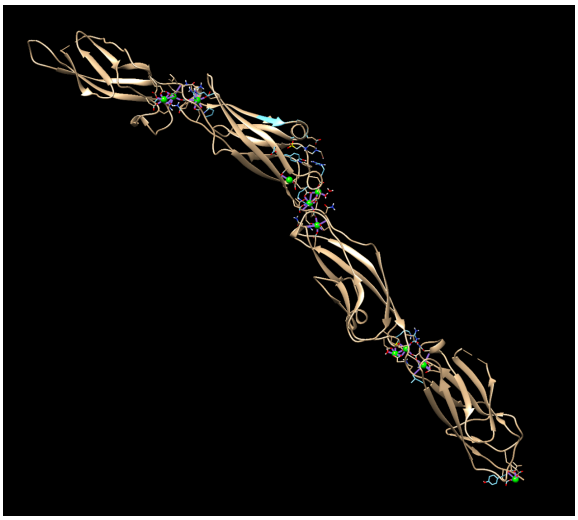
Supplementary Figures

Highest ΔG	-5.9	PCDH1 Conformational Change 
Molecular Weight	381.397	
LogP	2.601	

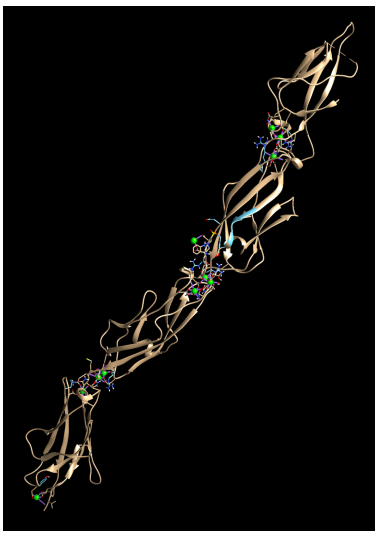
S1: Construction Trial 8

Highest ΔG	-6.0	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	1.558	

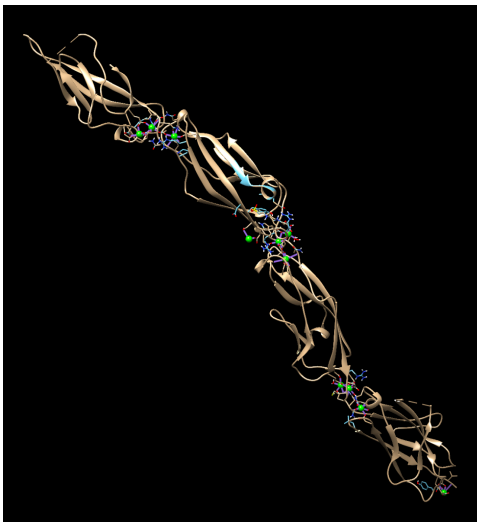
S2: Construction Trial 9

Highest ΔG	-6.4	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	1.248	

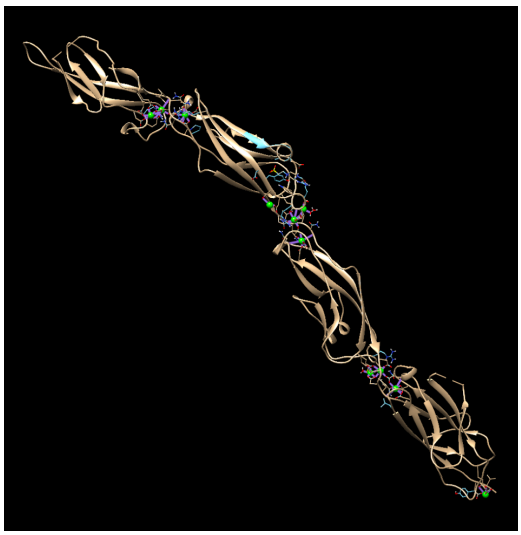
S3: Construction Trial 10

Highest ΔG	-6.3	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	0.744	

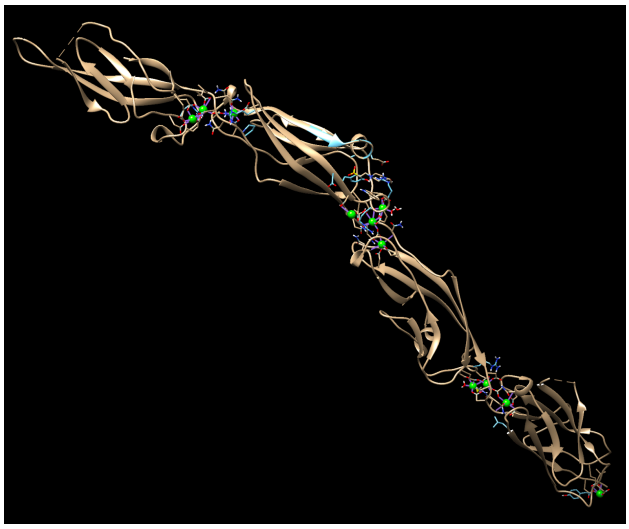
S4: Construction Trial 12

Highest ΔG	-6.1	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	1.285	

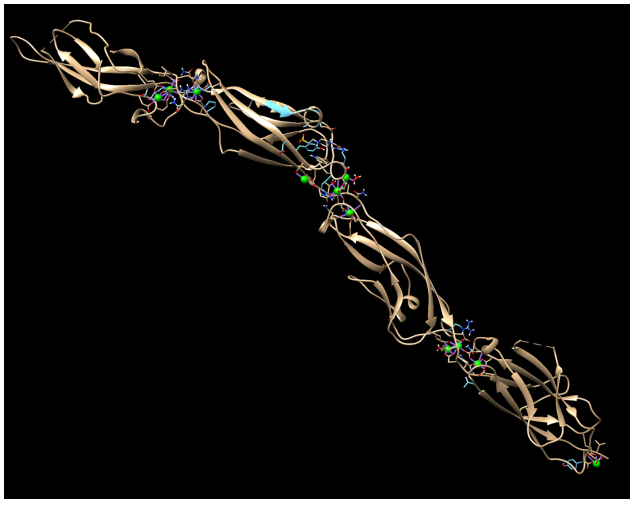
S5: Construction Trial 13

Highest ΔG	-6.0	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	1.054	

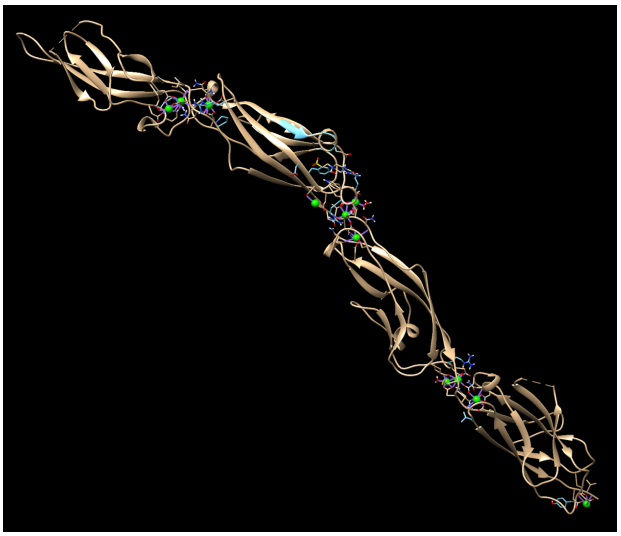
S6: Construction Trial 14

Highest ΔG	-5.9	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	0.397	

S7: Construction Trial 17

Highest ΔG	-6.2	PCDH1 Conformational Change 
Molecular Weight	397.3964	
LogP	1.099	

S8: Construction Trial 22

Highest ΔG	-6.3	PCDH1 Conformational Change 
Molecular Weight	413.3958	
LogP	0.152	

S9: Construction Trial 25

Highest ΔG	-5.8	PCDH1 Conformational Change 
Molecular Weight	408.472	
LogP	-0.733	

S10: Construction Trial 28