University of Mississippi
eGrove

Honors Theses

Honors College (Sally McDonnell Barksdale Honors College)

2015

A Computational Comparison of the Hydrophobic Pocket of Neural and Epithelial Cadherins for the Purpose of Identifying a Selective Inhibitor

Ross E. Straughan

University of Mississippi. Sally McDonnell Barksdale Honors College

Follow this and additional works at: https://egrove.olemiss.edu/hon_thesis

Part of the Chemistry Commons

A COMPUTATIONAL COMPARISON OF THE HYDROPHOBIC POCKET OF

NEURAL AND EPITHELIAL CADHERINS FOR THE PURPOSE OF IDENTIFYING

A SELECTIVE INHIBITOR

by
Ross Elliott Straughan

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of

the requirements of the Sally McDonnell Barksdale Honors College.

Oxford
May 2015

Approved by

_____

Advisor: Dr. Susan Pedigo

_____

Reader: Dr. Daniell Mattern

_____

Reader: Dr. Jonah Jurss

ACKNOWLEDGEMENTS

ABSTRACT

Classical cadherins are a subfamily of calcium-dependent cellular adhesion molecules

that play an important role in the formation of cellular junctions in many tissues. The

extracellular portion of cadherins consists of five tandem-repeated domains (EC1-EC5).

The critical first step in cadherin-mediated cellular adhesion occurs at the interface

between two adjacent EC1 domains in which the transition from monomer to dimer is

accomplished by docking the W2 residue of the N-terminal β-strand of one EC1 domain

into the hydrophobic pocket of its partner domain. Cancer and many other diseases have

been linked to the aberrant expression of Epithelial (E-cad) and Neural Cadherins (N-

cad). Due to the importance of cadherins in the study of cancer, the hydrophobic pocket

of the EC1 domain is of interest because it provides a possible site for the selective

inhibition of dimerization as an anti-cancer treatment. Furthermore, if the shape of the

hydrophobic pocket is different in E-cad and N-cad, then perhaps these differences may

be exploited to target a specific tissue or specific form of cancer. In order to study the

significance of the hydrophobic pocket, we studied the crystal structures of the EC1-EC2

domains of E-cad and N-cad using the imaging software *Chimera*. First, we compared the

position of critical hydrophobic pocket residues in two "identical" crystal structures of N-

cad and likewise for E-cad. Second, we used a specific function in *Chimera* to obtain

area and volume measurements of the hydrophobic pocket and its opening in each

iv

structure. Subsequently, a database of indole derivatives were docked into the hydrophobic pocket using the software *OpenEye* to identify potential ligands that could selectively bind a single Cadherin subtype. Results indicate that there is indeed a difference in the size and shape of the hydrophobic pockets of N-cad and E-cad that leads to differences in the optimal indole derivatives predicted to bind to N-cad and E-cad. These critical results suggest that the hydrophobic pockets of these two proteins are different and may be exploited for selective inhibition of dimerization by cadherin subtypes. Future studies will be directed toward developing these unique indole structures as possible cancer therapies.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

CASTp   Computed Atlas of Surface Topography of proteins

EC1   Extracellular Domain 1

EC2   Extracellular Domain 2

E-cad   Epithelial Cadherin

I3A   Indole-3-Acetic Acid

MS   Molecular Surface

N-cad   Neural Cadherin

pdb   Protein Data Bank

SA   Solvent Accessible

Trp-2   Tryptophan 2

INTRODUCTION

Classical cadherins are a subfamily of calcium-dependent cellular adhesion molecules that play an important role in the formation of cellular junctions in many tissues. Cadherins consist of five extracellular domains, a transmembrane domain, and a cytoplasmic domain. Each of the five extracellular domains exhibit structural similarities (each consisting of ~110 amino acids) and Ca^{2+} modulates the stability of the protein by binding between each successive domain[1]. The crucial first step in cadherin-mediated cellular adhesion involves the interface between two monomeric N-terminal (EC1) domains of Cadherin molecules emanating from adjacent cells. The transition from monomer to dimer is accomplished by the "strand swap" of the N-terminal β-strand of one EC1 domain that docks into the hydrophobic pocket of its partner and vice-versa[2, 3]. Achievement of the strand-swap interface is driven by the docking of a conserved Trp-2 residue in the βA-strand into the conserved hydrophobic pocket of the EC1 domain on the adjacent cell[4].

Epithelial cadherin (E-cad) plays an important role within epithelial tissue and is critical in establishing cellular junctions, cellular polarity, and permeability barriers[5]. A relative to E-cad, Neural cadherin (N-cad), is predominately expressed in neural, endothelial, and smooth muscle cells, as well as some carcinomas[6]. The relevance of cadherins to cancer is linked to a phenomenon termed "Cadherin Switching[7]." Cadherin Switching occurs during the normal development of tissues

and aids in the segregation of cell types[8-10]. However, Cadherin Switching also

occurs within tumor cells and enables them to metastasize[7, 11, 12].

The mode by which cadherin switching occurs can take one of several forms.

One of the defining characteristics of a transition from healthy epithelium to

carcinoma is the loss of E-cadherin expression[13]. Multiple studies have supported

this by showing that in tumors *in situ* E-cadherin expression is often lost[14-16].

Additionally, some epithelial-derived cancer cells inappropriately up-regulate the

expression of N-cadherin, which has been shown to promote the motility of

cancerous cells[17] (**Figure 1-1**).
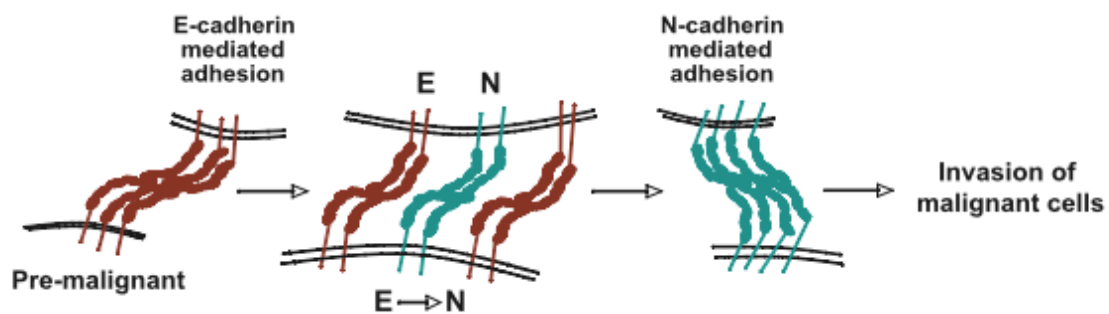
Figure 1-1. Schematic of tumors that undergo Cadherin Switching from E-cadherin to N-cadherin. The E-cad to N-cad transition allows these cells to exit the epithelial tissue and invade surrounding tissues. Once the malignant cells invade these tissues, N-cadherin is able to interact with mesenchymal cells[18].

Cadherin switching also includes situations in which the expression of E-cadherin is

not significantly altered, yet, the concurrent over-expression of N-cadherin leads to

the development of carcinomas[19]. This altered expression of cadherins in cancerous

cells holds many similarities to the cadherin switching that occurs during regular

embryonic development. Due to the important role of cadherins in the development of cancer, it is of interest to study the differences between N- and E-cadherin for the purposes of selectively inhibiting adhesion by one relative to the other. In particular their hydrophobic pockets offer a unique opportunity as a target for chemotherapeutics since the symmetrical docking of βA-strands into the hydrophobic pocket of its partner protomer is essential for cellular adhesion.

From the literature, we know that there is significant interest in the hydrophobic pocket of classical cadherins for the purpose of selective control over dimerization[20, 21] (**Figure 1-2**). However, the importance of the hydrophobic pocket cannot be fully understood without also considering the Trp-2 residue that it binds. Several studies have addressed the requirement of Trp-2 in dimerization. Studies by Tamura et al[22] demonstrated that in cell aggregation studies, when Trp-2 was mutated to phenylalanine, adhesion was impaired. Further, when Trp-2 was mutated to alanine, no aggregation was observed. These results were supported by studies in our laboratory on adhesion between truncated constructs on the first two domains of N-cad (NCAD12) in which Trp-2 was mutated to alanine[23]. Similar studies on the Trp-2 to alanine mutant of E-cad were published by Chitaev et al[24]. Given the importance of Trp-2, it is reasonable to assume that interference with the docking of Trp-2 into the hydrophobic pocket will disrupt dimerization, thereby making the disruption of the Trp-2 docking a promising chemotherapeutic target.

We are not the first to think of this idea. Previous cellular adhesion assays were performed in the presence of tryptophan analogues, indole-3-acetic acid (I3A) and its relative, 5-methyl indole-3-acetic acid, in order to study the potential

inhibitory effects of indole derivatives on cadherin dimerization. It was found that

I3A inhibited N-cad dimerization while 5-methyl I3A had no effect upon

dimerization[22]. These studies reveal that small modifications to the 6-membered

ring of the indole have a dramatic effect on docking of the indole moiety in the

hydrophobic pocket. It is interesting to note that other Cadherin-Cadherin

interactions are also targets of chemotherapeutics. The N-cad antagonist, the

pentapeptide N-Ac-*CHAVC*-NH$_2$ (designated ADH-1), has been shown to inhibit

tumor growth and metastasis in a mouse model of pancreatic cancer[25]. This peptide

is directed toward the interaction between Cadherins emanating from the same cell

surface, an interaction initially predicted from the first x-ray crystal structure

(although the interactions were apparently misinterpreted in the publication[2]. Due

to the ability of select indole derivatives to inhibit cellular adhesion, the

identification of a novel antagonist to selectively inhibit dimerization in specific

types of Cadherins is of great importance in the development of anti-cancer

therapies.
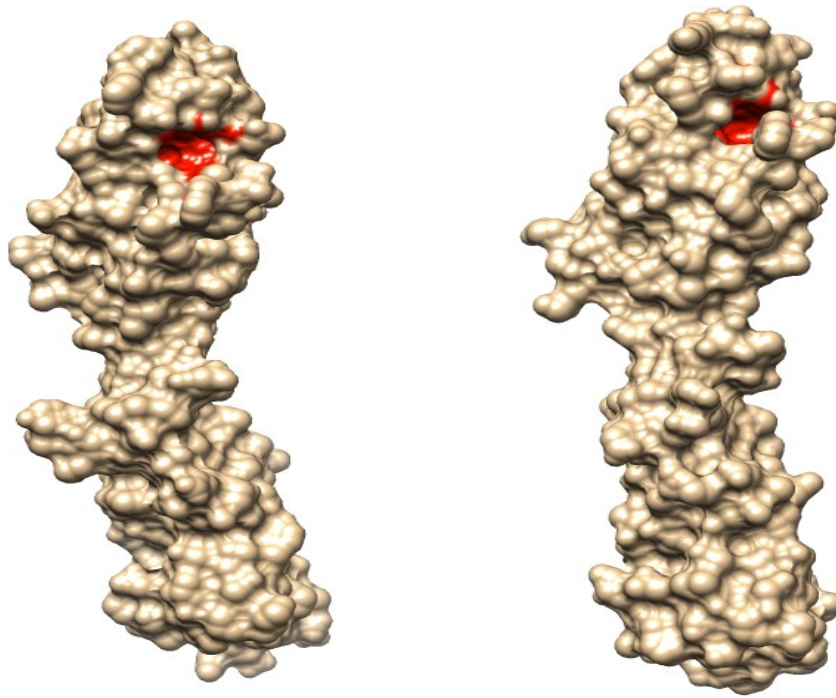
Figure 1-2. The first two extracellular domains (EC1, EC2) of N-cad (left) and E-cad (right). The hydrophobic pocket, located in EC1, is highlighted in red. The opening to the pocket is partially obstructed by the Tryptophan (Trp-2) residue of the β-strand. Both images are facing into the pocket from the mouth opening. It is interesting to note from this presentation of the two-domain constructs that there are a number of local differences in the overall shapes of the molecules. Some of these regions have been the focus of research as a source of difference in the kinetic and equilibrium of dimerization between E-cad and N-cad.

The purpose of this paper is to establish a computational basis for the differences between the hydrophobic pockets of E-cad and N-cad; and to investigate the subtype-dependence of predicted indole derivative antagonists that will compete with Trp-2 of the β strand and selectively inhibit dimerization. Through the analysis of E-cad and N-cad crystal structures in *Chimera*, we have determined that, indeed, a difference does exist between the hydrophobic pockets of the two cadherins. From this finding, the crystal structures of both N-cad and E-cad were screened against a database of indole derivatives to model the binding of the hydrophobic pocket in order to identify a ligand that will bind selectively to a single cadherin subtype. Finding an indole derivative that can selectively modulate the activity of specific cadherins may prove useful for anti-cancer treatments by inhibiting metastasis.

MATERIALS AND METHODS

Much of the data presented in this paper has been gathered by application of

the computer program, *Chimera*[26]. *Chimera* is developed by the Resource for

Biocomputing, Visualization, and Informatics at the University of California, San

Francisco, which was supported by NIGMS P41-GM103311. *Chimera* provides an

array of tools for the visualization and analysis of biomolecules. *Chimera* was first

used to access from the Protein Data Bank (pdb;http://www.rcsb.org/pdb/home/

home.do) the crystal structures of E-cad and N-cad. Subsequently, a series of

manipulations were performed within the program in order to study the

hydrophobic pockets of these proteins.

The visualization of E-cad and N-cad would not have been possible without

the use of crystal structures of these proteins from the Protein Data Bank (pdb).

From previous experiments conducted by x-ray crystallography, the pdb files

consist primarily of coordinates for biological molecules. These files contain the

atoms of each protein in their three-dimensional structure in a structure that is

thought to represent their functional form. Additionally, these files contain citation

information, structure solution details, and text describing experimental

observations used in determining the protein structure. Specifically, for use in the

experiments conducted here, the pdb files for 1NCG.pdb[2] and 2QVI.pdb[1] were used

for murine N-cad and 2QVF.pdb[1] and 1EDH.pdb[27] were used for murine E-cad.

These structures were the basis for the analysis of bond distances, structural

similarities, and all other comparisons between E-cad and N-cad.

To begin the comparison between the hydrophobic pockets of N-cad and E-

cad, the pbd files were uploaded into *Chimera* in order to superimpose the two

structures upon one another and then to create a sequence and structural alignment

from the superimposed image. The superposition was carried out through the

"MatchMaker" module within *Chimera*. MatchMaker does this by first matching the

two sequences in a pairwise manner and then fitting the α-carbons of paired

residues from the sequence alignment. Next, a structural-based alignment was

carried out by using the "Match -> Align" module. "Match -> Align" only uses the

distances between α-carbons when creating the alignment and does not take

residue type into account. These alignments were then visualized in the "Multialign

Viewer."

The next step in comparing the hydrophobic pockets of N-cad and E-cad was

accomplished by uploading the pdb files into *Chimera* in order to measure the

volumes and areas of each hydrophobic pocket. *Chimera* cannot directly measure

the volume of a pocket that opens to the outside surface because it does not know

where to mark the plane between inside and outside. However, the pbd files can be

uploaded into *Chimera* with reference to the Computed Atlas of Surface Topography

---

[1] The x-ray crystal structures for 2QVI.pdb and 2QVF.pdb are posted on the rcsb.org website without a citation in a peer reviewed journal.

of proteins (CASTp) Database[28], which utilizes Delaunay triangulation and alpha

shapes in order to determine pocket boundaries. The CASTp calculations are carried

out using a solvent probe with a 1.4 angstrom radius, which is the radius of a water

molecule that is considered spherical (**Figure 2-1**). The results of this surface probe

provide analytical measurements of the area and volume of every pocket and cavity

within the protein. The data from this analysis is presented by two values: the

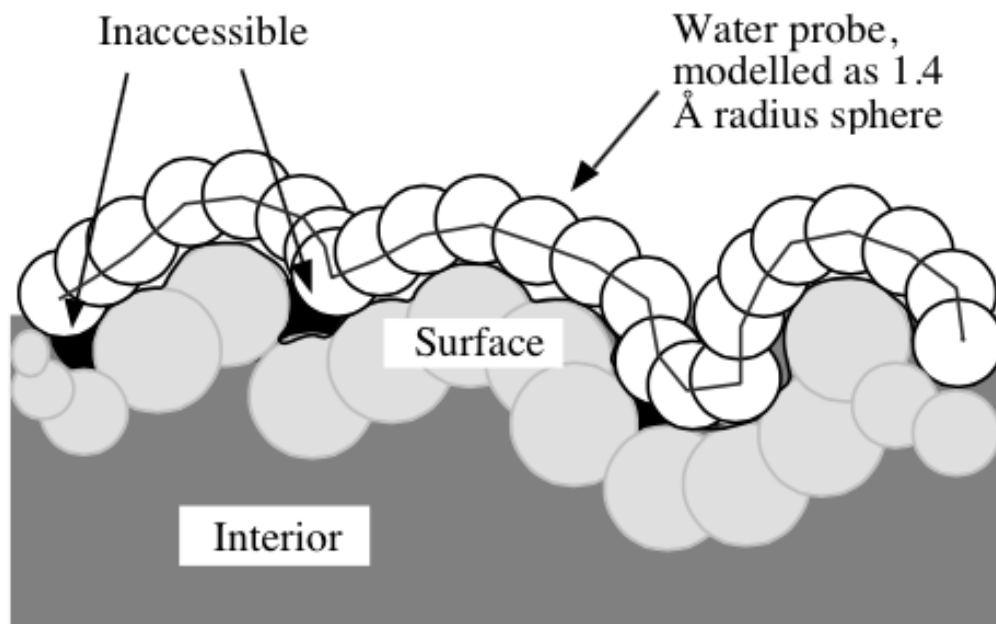solvent accessible surface (SA) and the molecular surface (MS).

Figure 2-1. A schematic representation of a water probe as it scans the surface of a
protein. Due to the spherical shape of the probe, a small fraction of the surface area is
inaccessible to the probe. The surface area that the probe can access is used to calculate
the Solvent Accessible (SA) area and volume. (Figure is reproduced from *Biophysics
Textbook Online* by Victor Bloomfield at biophysics.org.)

The MS is always larger than the SA because the MS includes all of the spaces that

are too small to accommodate a water molecule. In addition to this, the CASTp

output provides information about the circumference and the area of an opening. All measurements of area and volume are presented in square angstroms and cubic angstroms, respectively.

The final step in comparing the hydrophobic pockets of N-cad and E-cad was carried out through the utilization of *in silico* virtual screening. Virtual screening is a widely used method in computational chemistry in which protein crystal structures are screened against a large ligand database in order to determine potential drug interactions. For the purposes of this paper, the indole-derivative ligands were selected from the ZINC database, a free database of over 35 million commercially-available compounds[29]. Next, the crystal structures for N-cad (2QVI.pdb) and E-cad (2QVF.pdb), along with the indole derivatives, were loaded into the virtual screening program *OpenEye* (eyesopen.com). Subsequently, *OpenEye* performed an examination of all protein-ligand interactions and scored the ligands using the chemgauss4 scoring function. This scoring function takes into account shape, hydrogen bonding between ligand and protein, hydrogen bonding with solvent, and hydrogen bonding network effects. From this, Gaussian potentials are used to measure the complementarity of ligands within the active site. The results of the virtual screening were visualized using *PyMol*[30].

RESULTS AND DISCUSSION

STRUCTURAL AND SEQUENCE ALIGNMENT OF N- AND E-CADHERINS

The comparison of protein sequence and structure between subtypes is important toward understanding protein function. Sequence comparison provides information about residue conservation between subtypes, while structural comparison may indicate the importance of divergent residues in relation to protein function in similar but different proteins. More specifically for the purposes of this paper, these tools were utilized to gain greater insight into the differences in residues that define the hydrophobic pockets of N-cad and E-cad. The sequence-based structural alignment performed here was conducted with focus on the residues that were indicated, by CASTp data, as those that define the shape of the hydrophobic pocket.

The results of this alignment indicate that there are indeed several differences in the residues that define the hydrophobic pockets of E- and N-cadherin (**Figure 3-1**). These differences can be seen at 4 residues. For E-cad, the residues of interest are Lys 25, Asn 27, Ser 78 and Asp 90. The corresponding residues in N-cad are Arg 25, Asp 27, Ala 78, and Asn 90. It is particularly interesting to note the difference at residue 78 between Alanine and Serine. This residue defines much of the bottom of the pocket and the extended arm of Serine, relative to the smaller Alanine residue, may contribute significantly to the difference in shape of the two hydrophobic pockets. Additionally, the hydrophobic pocket of E-cad is
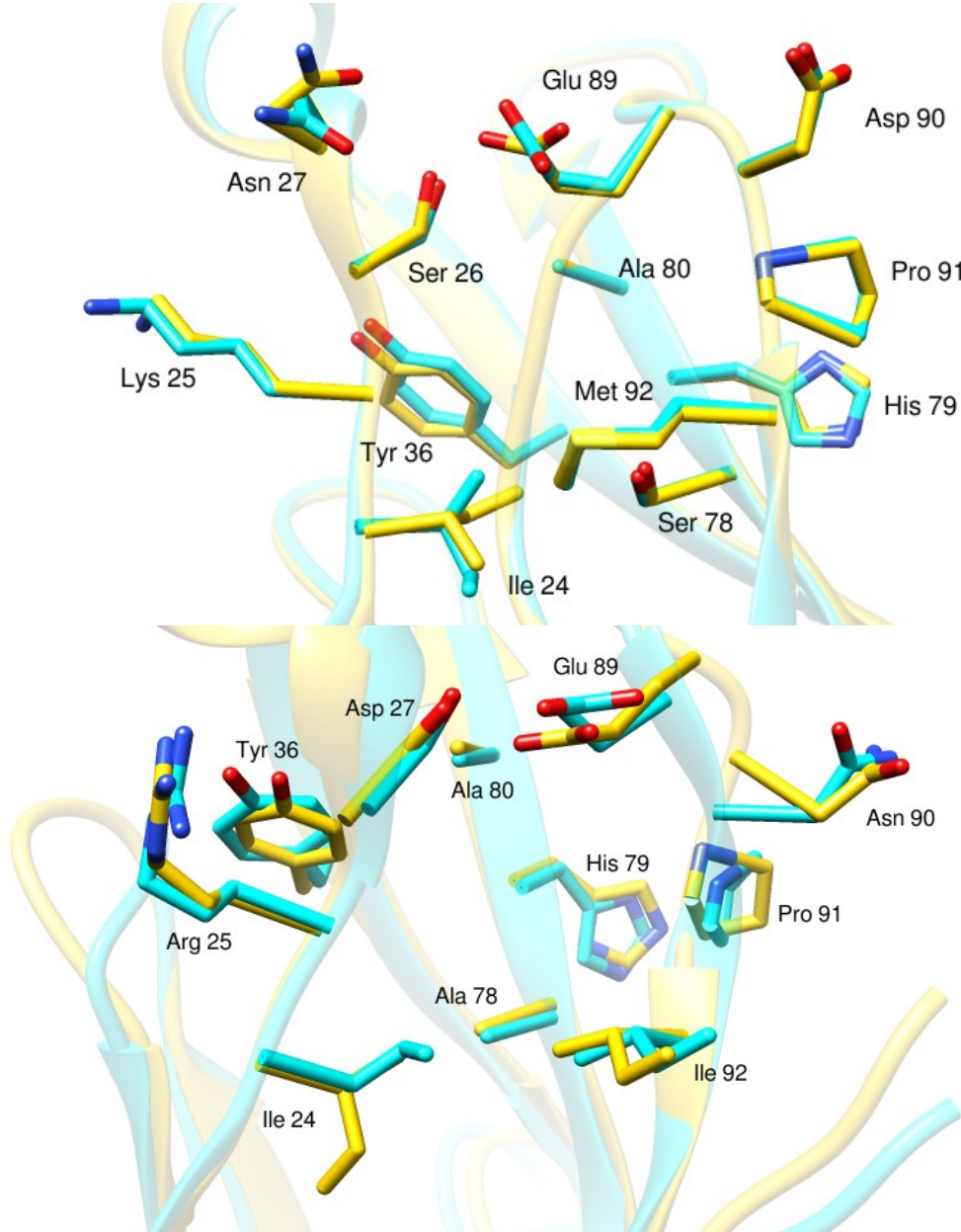
11

Figure 3-1. Importance of the residues that define the hydrophobic pocket. Two aligned structures of E-cad (top) and two aligned structures of N-cad (bottom) are shown. Inspection reveals differences in position and identity of several residues that define the hydrophobic pockets of each Cadherin type.

also defined by an additional residue that does not contribute to the pocket in N-cad, Ser 26. Such differences in the residues that define each hydrophobic pocket indicates that a structural difference between the two cadherin types exists, which could be potentially exploited as a therapeutic target.

Also of note is the presence of a rotamer at Isoleucine-24. This is brought to attention to assess the fact that the hydrophobic pocket is not a single static shape or size. Additionally, we know from literature that there is an interest in residues 78 and 92 that define the floor of the pocket[31]. The effects of these residues on the volume and area of the hydrophobic pocket will be discussed in the next section. Overall, while these findings are more qualitative than quantitative in nature, the clear differences in the residues that define the hydrophobic pocket further indicate the potential for this site to be explored as a therapeutic target.

COMPARISON OF HYDROPHOBIC POCKET VOLUME DATA

Examination of the CASTp surface area and volume data for the hydrophobic pocket in the crystal structures of N-cad and E-cad reveals that there is a significant difference in the size and area of each protein's respective hydrophobic pocket (**Table 3-1**). The volume of the hydrophobic pocket of N-cad was revealed to be larger than E-cad by an average of ~93 angstroms for the molecular surface (MS) volume and an average of ~59 angstroms for the solvent accessible (SA) volume. Similarly, the areas of both the pocket and mouth are larger in N-cadherin.

Table 3-1. Area and Volume Data for Various Hydrophobic Pockets

Cadherin Type	pdb file used	MS Volume	SA Volume	Pocket MS Area	Pocket SA Area	Mouth MS Area	Mouth SA Area
N-cad	2qvi	551.5	206.2	302.1	199.8	169.9	83.5
N-cad	1ncg	421.3	146.1	241.9	163.2	152.4	58.9
E-cad	2qvf	393.9	117.4	240.7	152.4	134.0	47.5
E-cad	1edh	259.4	53.6	196.0	101.5	70.3	20.4

One question that arises from this data is what causes the significant difference in area and volume within each cadherin subtype. The cause for this difference becomes even more troubling when visualizing the superimposed structures of N-cad and E-cad in (**Figure 3-1**). Based on visual inspection of these images, it would be predicted that the volume and area within each cadherin subtype would be very similar. However, such similarity is not reflected in the CASTp data presented here. During the course of this research, we realized that there are significant inconsistencies in the surface area that the CASTp calculation defines as being part of the mouth of the pocket. This discrepancy can be easily seen when viewing the maps of the hydrophobic pockets given from the CASTp output files. These figures are shown in the Appendix. In particular, 2QVI.pdb includes a large portion of surface area below the mouth of the pocket that is not included in 1NCG.pdb, the other structure for N-cad. Similarly in E-cad, there are obvious differences in the shape of what the map defines as the mouth of the mapped hydrophobic pocket. This would indeed explain why such a significant difference would arise in these calculations. While this surface area may be important for ligand binding, at this time we are unable to standardize the unreliable and complex process by which the CASTp

database defines the mouth of the hydrophobic pocket. While the quantitative data presented here might not be an accurate representation of the absolute area and volume of the hydrophobic pocket, it at the very least suggests that a relative difference indeed exists between N-cad and E-cad.

The relative inconsistencies in the sizes of the structures presented could be attributed to differences in specific residues that define the hydrophobic pocket. More specifically, one possible difference in the size and shape of the hydrophobic pocket that we want to mention was the rotamers of Isoleucine-24. Its ethylene group can rotate into and out of the pocket in both E-cad and N-cad. However, the relative volume of the side chain (~4.3 cubic angstroms) is not enough to account for such a large difference in the size of the hydrophobic pockets, which is accounted for as mentioned above.

Another point of interest in determining the cause for the difference in the size and shape of the pocket are the subtype-specific residues 78 and 92 that define much of the floor of the hydrophobic pocket[31]. Residue 78 is Alanine in N-cad and Serine in E-cad, while residue 92 is Isoleucine in N-cad and Methionine in E-cad. The Serine and Methionine residues present in E-cad are much larger and more polar than the relatively hydrophobic Alanine and Isoleucine residues that are present in N-cad. This observation is consistent with the findings that the hydrophobic pocket of E-cad is indeed smaller. The importance of these residues in the hydrophobic pockets of E-cad and N-cad was demonstrated in an experiment by Vendome, et al[31] in which residues 78 and 92 in E-cad were given the mutations S78A, M92I in order to make an "N-like" hydrophobic pocket in E-cad. Interestingly, when these residues in E-cad were mutated to possess the residues present in N-cad, the N-like mutant exhibited a nearly identical dissociation constant

(23.8 μM) to wild type N-cadherin (25.8 μM) as determined by Analytical

Ultracentrifugation. These findings confirm the importance of residues 78 and 92 both in

defining cadherin subtype and also in defining the shape and size of the hydrophobic

pocket, and that the shape of the hydrophobic pocket has a role in determining the affinity

of adhesive dimer formation. Interestingly, in Vendome's study, they could not convert

N-cad into E-cad by creating the reciprocal double mutant, indicating that the

determinants for the affinity of the adhesive dimer are more complex than just

hydrophobic pocket geometry.

IN SILICO DOCKING OF INDOLE DERIVATIVES WITH N- AND E-CADHERINS

After conducting virtual screening of the protein crystal structures in *OpenEye*,

the indole derivatives tested were scored and ranked based on predicted binding (**Table

3-2**). While the program indicated over 400 possible ligands that could potentially bind

the pocket, only the top five for each cadherin subtype are shown here. Interestingly,

different indole derivatives scored better for N-cad and E-cad, further indicating that a

substantial difference exists between the hydrophobic pockets that could be targeted in

drug design.

While viewing these results, it is interesting to note the hydrogen bonding

interactions that anchor the ligand within the hydrophobic pocket (**Figure 3-2**). From

this, it is evident that the shape of the hydrophobic pocket is different for N-cad and E-

cad. For the top scoring ligand for E-cad, Ser-78, Glu-89, and Asp-90 donate hydrogen

bonds to stabilize the ligand. Conversely, for the top scoring ligand for N-cad, Asn-90 is

the only pocket residue that is involved in hydrogen bonding with the ligand while Asp-1

of the β-strand interacts with the hydrogen of a tertiary amine group. The lesser ability of N-cadherin to hydrogen bond with ligands in the hydrophobic pocket is consistent with the differences between N-cad and E-cad at residues 78 and 92 discussed in the previous section. This data indicates that the differential binding of ligands to cadherin subtypes may be driven by both pocket size as well as non-covalent interactions within the pocket.
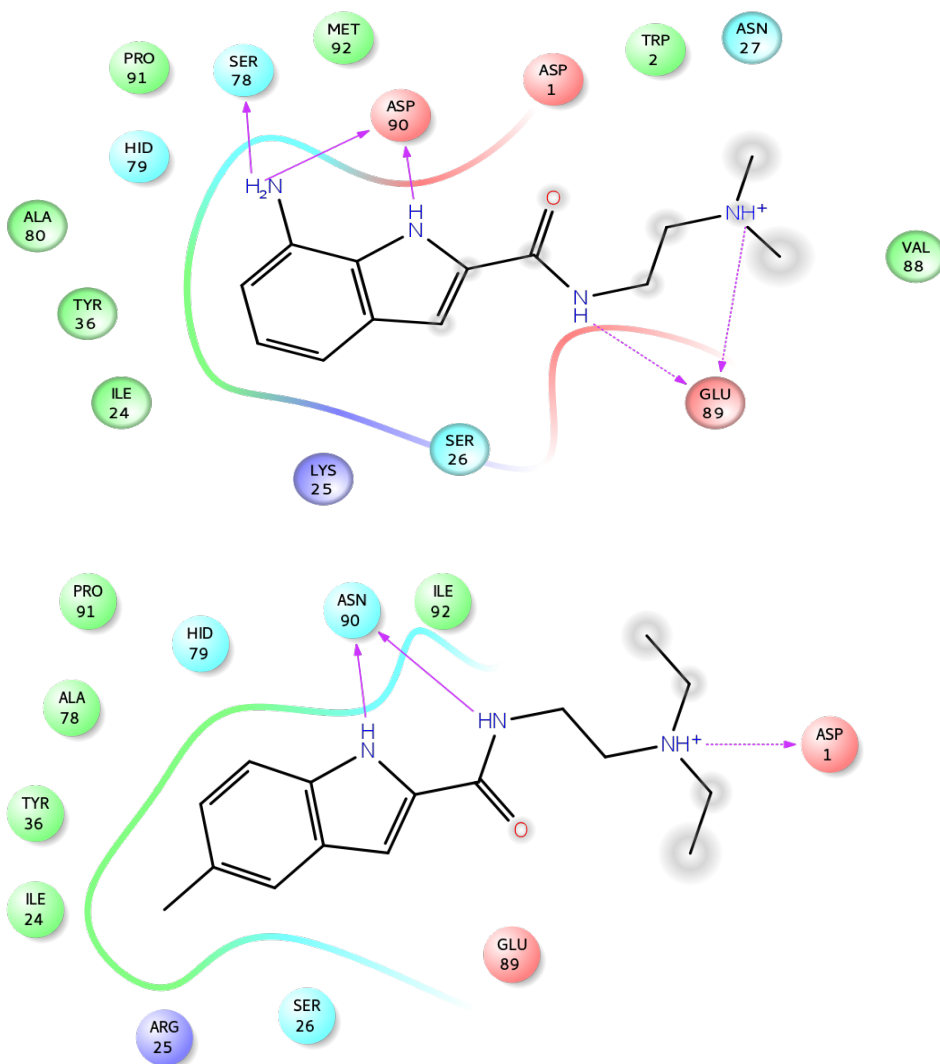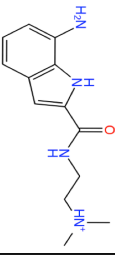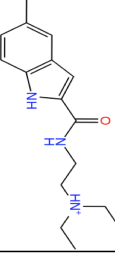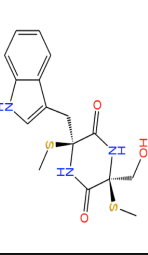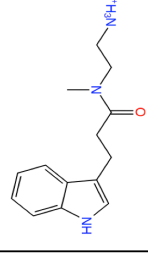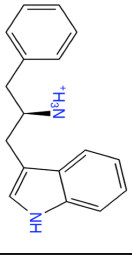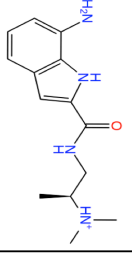

Figure 3-2. Two-dimensional view of ligand interactions within the hydrophobic pocket. E-cad (top) is docked with the top scoring ligand from virtual screening. Five potential interactions can be seen in the protein-ligand complex. N-cad (bottom), also shown docked with its top scoring ligand, has only three potential interactions with only two of those occurring in the pocket itself.
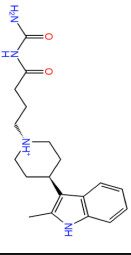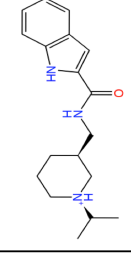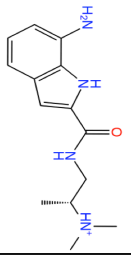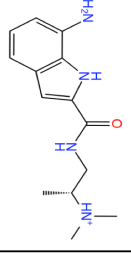
17

Table 3-2. Top Scoring Ligands from Virtual Screening Against N- and E-Cadherin

E-Cadherin				N-Cadherin			
Structure	ZINC ID	Name	Chemgauss4 score	Structure	ZINC ID	Name	Chemgauss4 score
	ZINC62833581	7-amino-N-(2-dimethylaminoethyl)-1H-indole-2-carboxamide	-14.188612		ZINC14008861	N-(2-diethylaminoethyl)-5-methyl-1H-indole-2-carboxamide	-13.966022
	ZINC13430269	(3S,6S)-3-(hydroxymethyl)-6-(1H-indol-3-ylmethyl)-3,6-bis(methylsulfanyl)pip erazine-2,5-dione	-13.554291		ZINC54417529	N-(2-aminoethyl)-3-(1H-indol-3-yl)-N-methyl-propanamide	-13.878401
	ZINC54685433	(2R)-1-(1H-indol-3-yl)-3-phenyl-propan-2-amine	-13.444638		ZINC62833841	7-amino-N-[(2S)-2-(dimethylamino)propy l]-1H-indole-2-carboxamide	-13.810176
	ZINC58344048	N-carbamoyl-4-[4-(2-methyl-1H-indol-3-yl)-1-piperidyl]butanamide	-13.251917		ZINC58330517	N-[[(3S)-1-isopropyl-3-piperidyl]methyl]-1H-indole-2-carboxamide	-13.56624
	ZINC62833842	7-amino-N-[(2R)-2-(dimethylamino)propy l]-1H-indole-2-carboxamide	-13.214027		ZINC62833842	7-amino-N-[(2R)-2-(dimethylamino)propy l]-1H-indole-2-carboxamide	-13.522629

The purpose of this experiment was to identify a ligand that has the potential to bind a single cadherin subtype in order to selectively inhibit dimerization. Through examining the results in search of a selective inhibitor of E-cad, it was noted that the highest ranked ligand for E-cad is ranked third for N-cad. Due to the fact that this ligand ranks so highly for both cadherin subtypes, it is not reasonable to believe that this ligand would be of use in selectively binding E-cad. However, the search for a selective inhibitor of N-cad indicated a much more promising result. Interestingly, the highest ranking ligand for N-cad is ranked 251st on the list of ligands for E-cad. Such a difference gives legitimate reason to believe that we have identified a ligand that will selectively inhibit N-cad dimerization. This finding is exciting because if indeed this compound selectively binds to N-cad and inhibits dimerization, it could potentially serve as a cancer therapy. More specifically, it could be a possible means to suppress metastasis in the particular carcinomas that show up-regulated expression of N-cad, the hallmark of metastasis in particular forms of prostate, breast and pancreatic cancer[17, 19]. Due to the exciting potential of this ligand as a cancer therapy, the future direction of this research will be to test this ligand's ability to decrease dimer formation *in vitro* adhesion assays.

There are obvious imperfections in the *in silico* screening procedure that warrant further testing. As stated previously, the next step in this research will be to test the highest scoring ligand for N-cad to determine its true efficacy in selectively binding the hydrophobic pocket. However, the results indicated here are very exciting and give promise for the future development of new anti-cancer therapies by exploiting the differences in the hydrophobic pockets of N-cad and E-cad.

BIBLIOGRAPHY

1. Chothia, C., and Jones, E. Y. (1997) The molecular structure of cell adhesion molecules. *Annu Rev Biochem 66*, 823-862.

2. Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grubel, G., Legrand, J. F., Als-Nielsen, J., Colman, D. R., and Hendrickson, W. A. (1995) Structural basis of cell-cell adhesion by cadherins [see comments]. *Nature 374*, 327-337.

3. Harrison, O. J., Corps, E. M., Berge, T., and Kilshaw, P. J. (2005) The mechanism of cell adhesion by classical cadherins: the role of domain 1. *J Cell Sci 118*, 711-721.

4. Klingelhofer, J., Troyanovsky, R. B., Laur, O. Y., and Troyanovsky, S. (2000) Amino-terminal domain of classic cadherins determines the specificity of the adhesive interactions. *J Cell Sci 113 (Pt 16)*, 2829-2836.

5. Baum, B., and Georgiou, M. (2011) Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling. *The Journal of cell biology 192*, 907-917.

6. Hatta, K., and Takeichi, M. (1986) Expression of N-cadherin adhesion molecules associated with early morphogenetic events in chick development. *Nature 320*, 447-449.

7.   Gravdal, K., Halvorsen, O. J., Haukaas, S. A., and Akslen, L. A. (2007) A switch from E-cadherin to N-cadherin expression indicates epithelial to mesenchymal transition and is of strong and independent importance for the progress of prostate cancer. *Clinical Cancer Research 13*, 7003-7011.

8.   Niessen, C. M., and Gumbiner, B. M. (2002) Cadherin-mediated cell sorting not determined by binding or adhesion specificity. *The Journal of cell biology 156*, 389-399.

9.   Tepass, U., Truong, K., Godt, D., Ikura, M., and Peifer, M. (2000) Cadherins in embryonic and neural morphogenesis. *Nat Rev Mol Cell Biol 1*, 91-100.

10.   Miyatani, S., Shimamura, K., Hatta, M., Nagafuchi, A., Nose, A., Matsunaga, M., Hatta, K., and Takeichi, M. (1989) Neural cadherin: role in selective cell-cell adhesion. *Science 245*, 631-635.

11.   Hazan, R. B., Kang, L., Whooley, B. P., and Borgen, P. I. (1997) N-cadherin promotes adhesion between invasive breast cancer cells and the stroma. *Cell Adhes Commun 4*, 399-411.

12.   Nieman, M. T., Prudoff, R. S., Johnson, K. R., and Wheelock, M. J. (1999) N-cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression. *The Journal of cell biology 147*, 631-644.

13.   Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., and Garcia De Herreros, A. (2000) The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nature cell biology 2*, 84-89.

14. Al-Aynati, M. M., Radulovich, N., Riddell, R. H., and Tsao, M. S. (2004) Epithelial-cadherin and beta-catenin expression changes in pancreatic intraepithelial neoplasia. *Clin Cancer Res 10*, 1235-1240.

15. Berx, G., Cleton-Jansen, A. M., Nollet, F., de Leeuw, W. J., van de Vijver, M., Cornelisse, C., and van Roy, F. (1995) E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. *Embo J 14*, 6107-6115.

16. Berx, G., and van Roy, F. (2009) Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harbor perspectives in biology 1*, a003129.

17. Hazan, R. B., Phillips, G. R., Qiao, R. F., Norton, L., and Aaronson, S. A. (2000) Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis. *The Journal of cell biology 148*, 779-790.

18. Maret, D., Gruzglin, E., Sadr, M. S., Siu, V., Shan, W., Koch, A. W., Seidah, N. G., Del Maestro, R. F., and Colman, D. R. (2010) Surface expression of precursor N-cadherin promotes tumor cell invasion. *Neoplasia 12*, 1066-1080.

19. Wheelock, M. J., Shintani, Y., Maeda, M., Fukumoto, Y., and Johnson, K. R. (2008) Cadherin switching. *Journal of cell science 121*, 727-735.

20. Devemy, E., and Blaschuk, O. W. (2008) Identification of a novel N-cadherin antagonist. *Peptides 29*, 1853-1861.

21. Devemy, E., and Blaschuk, O. W. (2009) Identification of a novel dual E- and N-cadherin antagonist. *Peptides 30*, 1539-1547.

22.   Tamura, K., Shan, W. S., Hendrickson, W. A., Colman, D. R., and Shapiro, L. (1998) Structure-function analysis of cell adhesion by neural (N-) cadherin. *Neuron 20*, 1153-1163.

23.   Vunnam, N., and Pedigo, S. (2011) Calcium-induced strain in the monomer promotes dimerization in neural cadherin. *Biochemistry 50*, 8437-8444.

24.   Chitaev, N. A., and Troyanovsky, S. M. (1998) Adhesive but not lateral E-cadherin complexes require calcium and catenins for their formation. *The Journal of cell biology 142*, 837-846.

25.   Blaschuk, O. W. (2012) Discovery and development of N-cadherin antagonists. *Cell Tissue Res.*

26.   Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) USCF Chimera - A visulaization System for Exploratory Research and Analysis. *J Comput Chem 25*, 1605-1612.

27.   Nagar, B., Overduin, M., Ikura, M., and Rini, J. M. (1996) Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature 380*, 360-364.

28.   Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic acids research 34*, W116-118.

29.   Irwin, J. J., and Shoichet, B. K. (2005) ZINC--a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling 45*, 177-182.

30. Schrodinger, L. (2010) The PyMOL Molecular Graphics System, Version 1.3r1. *www.pymol.org*.

31. Vendome, J., Felsovalyi, K., Song, H., Yang, Z., Jin, X., Brasch, J., Harrison, O. J., Ahlsen, G., Bahna, F., Kaczynska, A., Katsamba, P. S., Edmond, D., Hubbell, W. L., Shapiro, L., and Honig, B. (2014) Structural and energetic determinants of adhesive binding specificity in type I cadherins. *Proceedings of the National Academy of Sciences of the United States of America 111*, E4175-4184.

A
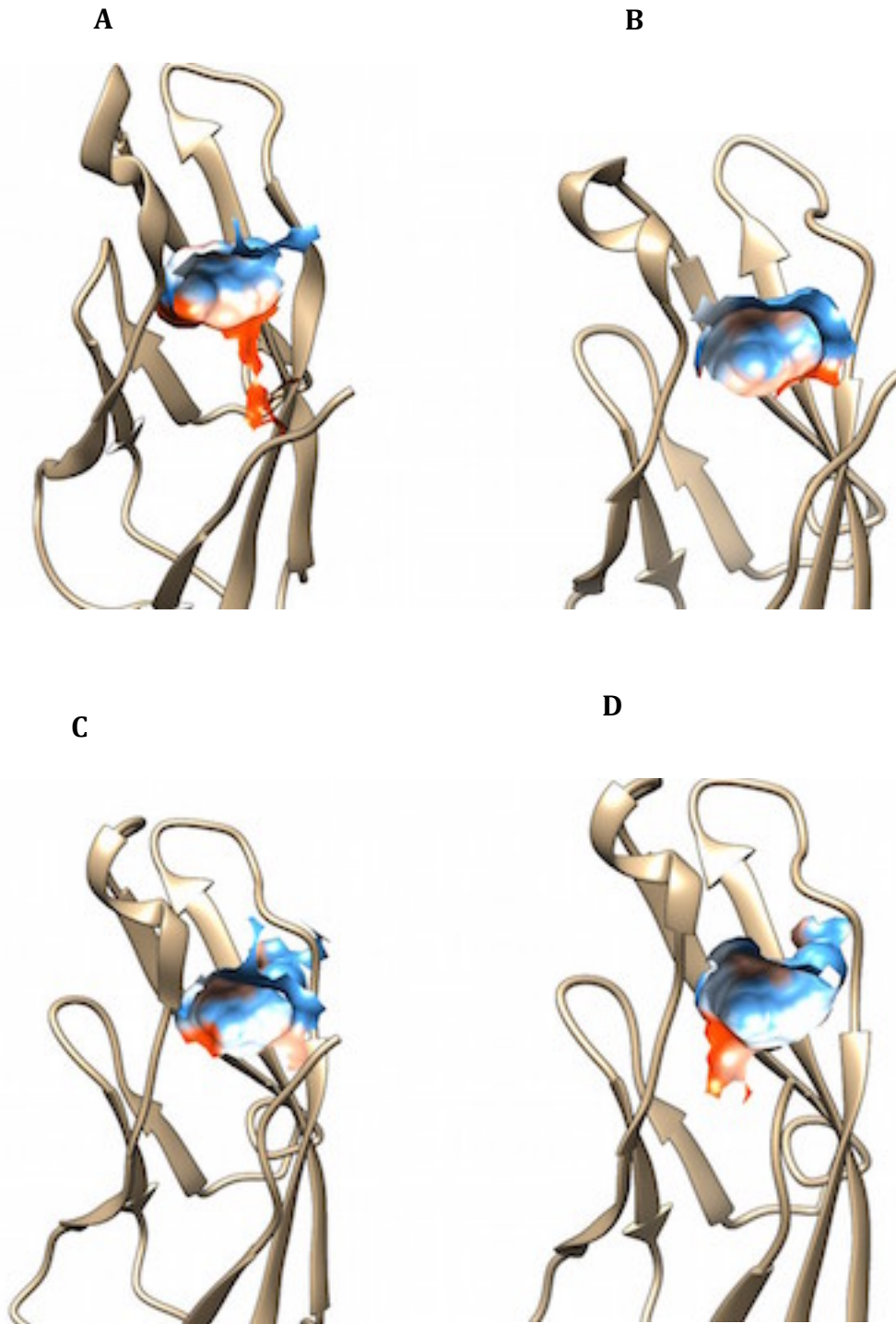
B

C

D

Figure A-1. Mapped hydrophobic pockets of several cadherins based on CASTp output data. (A) The mapped pocket of 2qvi.pdb with visible surface area that extends below the

hydrophobic pocket. Notice that this is not present in the map of the other E-cad crystal structure1ncg.pdb (B). The mapped pockets of 2qvf.pdb (C) and 1edh.pdb (D) also exhibit noticeable differences in the defined area of the pocket.