ISSN 2348-6201
Volume 6 Number 1
Journal of Advances in Biotechnology

A graph-theoretic model of single point mutations in the cystic fibrosis transmembrane conductance regulator

Samuel Kakraba[1], Debra Knisley[2]
[1]University of Arkansas at Little Rock and the University of Arkansas Medical Sciences
sxkakraba@ualr.edu
[2]East Tennessee State University, Department of Mathematics and Statistics, Johnson City TN USA 37614
knisleyd@etsu.edu

ABSTRACT

Cystic fibrosis is one of the most prevalent inherited diseases. This disease is caused by a mutation in a membrane protein, the cystic fibrosis transmembrane conductance regulator (CFTR). CFTR is known to function as a chloride channel that regulates the viscosity of mucus that lines the ducts of a number of organs. The most prevalent mutation of CFTR is located in one of two nucleotide binding domains, namely, the nucleotide binding domain one (NBD1). However, some mutations in nucleotide binding domain two (NBD2) can equally cause cystic fibrosis. In this work, a graph-theoretic model is built for NBD2. Using this model for NBD2, we examine the consequences of single point mutations on NBD2. We collate the wildtype structure with eight of the most prevalent mutations and observe how the NBD2 is affected by each of these mutations.

Indexing terms/Keywords

Cystic fibrosis; graph theory; graph-theoretic model; mutation.

Academic Discipline And Sub-Disciplines

Computational Biology, Applied Discrete Mathematics, Protein Science

SUBJECT CLASSIFICATION

Combinatorial Mathematics, Structural Biology

TYPE (METHOD/APPROACH)

Mathematical Modeling

1. INTRODUCTION

Cystic fibrosis is caused by a single point mutation in a membrane protein and is the most common genetic disorder in the caucasion population. The cystic fibrosis transmembrance conductance regulartor (CFTR) is a large protein that is embedded in the membrane of epithelial cells and functions as a chloride channel [1–4]. The most common mutation of CFTR results in the misfolding of the protein and subsequent degradation. Consequently, people with this mutation have an absence or diminished amount of mature, functioning CFTR in the membrane of the epithelial cells of organs such as the lungs, pancreas and colon. The viscosity of the mucus that lines the ducts of these organs is altered, resulting in sticky mucus plugs that disrupt the normal function of these organs. Lung disease is the primary cause of death for cystic fibrosis (CF) patients. The affected protein, CFTR, is a large protein consisting of five domains; two transmembrane domains, two nucleotide binding domains and a regulatory domain. The deletion of phenylalanine at position 508 (ΔF508) occurs in more than 90% of the CF population [5] and occurs in one of the two nucleotide binding domains (NBD1).

It is still not fully understood how mutations of CFTR cause the mis-folding of the protein. In previous work by Knisley et al, [6], a graph-theoretic model was built to study the effects of various mutations on the protein domain NBD1. Despite the fact that most of the disease causing mutations occur in NBD1, a number of them also occur in NBD2. There are seventeen mutations in the LSGGQ sequence and Walker B motif of NBD1 which cause CF, while there are four mutations in the same respective region of NBD2. In addition, research has shown that whereas there is only one mutation in the Walker A motif of NBD1 causing cystic fibrosis, there are as many as five of these mutations taking place in NBD2. In view of the fact that mutations that results in cystic fibrosis can equally occur in NBD2, it is appropriate to make an effort to gain understanding on how mutations in NBD2 can impact significantly on NBD2 [7, 8. 9].

To address the key question, we build a graph-theoretic model of NBD2. With a vertex-weighted nested graph representation of the protein domain NBD2, we present a method to model the effect of a single point mutation of a protein. Using the vertex-weighted graph, we define novel combinatorial descriptors based on these vertex-weights. We employ these graph-theoretic measures to quantify the consequences of nine mutations of CFTR's NBD2. Each mutation results in a distinct set of graph-theoretic measures that are both local and global and capture the underlying structural network consequences of the mutation. Our graph-theoretic approach models a process by which a local change can produce a significant global change. Even though the method of this research is analogous to that used by Knisley et al. [6], the molecular descriptors that are calculated to inform the graph-theoretic model are much more extensive and particular to NBD2. A more complete understanding of the effects of a single point mutation on the structure and function of the protein molecule are necessary to guide the rational design of drugs to treat this fatal disease.

2. METHODS

2.1 Graph-theoretic Model

Each of the twenty most common amino acids is modeled as a vertex-weighted graph. The backbone and central carbon atom are represented by a single vertex and each of the atoms in the corresponding amino acid residue's chain structure is represented by a vertex. The hydrogen atoms are suppressed. Vertices in the residue chain are weighted by the nearest integer value of the molecular mass of the corresponding atom and edges are determined by molecular bonds. Using each of these hydrogen suppressed models of the twenty most common amino acids, we obtain twenty corresponding vectors of descriptors of length twenty two that are based on the graph-theoretic measures. The definitions of the measures used can are found below.

2.1.1 Molecular Descriptors

The molecular descriptors in the Tables 1A, 1B and 1C are computed from the graph-theoretic model based on weighted degree and the assignment of atomic numbers as to each vertex. **Keys**: *d*1 = number of vertices (order of the graph), *d*2 = number of edges (size of the graph), *d*3 = total weighted degree of the graph (obtained by adding all the weights of all the vertices), *d*4 = weighted domination number, *d*5 = average eccentricity, *d*6 = radius (minimum eccentricity), *d*7 = diameter (maximum eccentricity), *d*8 = average weighted degree (total degree, divided by the number of vertices), *d*9 = maximum eigenvalue of the weighted Laplacian matrix of the graph, *d*10 = minimum eigenvalue of the weighted Laplacian matrix of the graph, *d*11 = Average eigenvalue of the Laplacian matrix of the the graph, *d*12 = second smallest eigenvalue of the Laplacian matrix of the graph. Using the atomic numbers as weights of vertices in the graph theoretic model of each of the amino acids, we obtain the following descriptors in Tables 2-3: *d*13 = weighted domination number using the atomic number, *d*14 = average weighted eccentricity based on the the atomic number, *d*15 = weighted radius based on the atomic number (minimum eccentricity), *d*16 = weighted diameter based on the atomic number (maximum eccentricity), *d*17 = total weighted atomic number of the graph (obtained by summing all the atomic number of each of the vertices in the graph), *d*18 = average weighted atomic number or degree based on atomic number in the graph. Descriptors *d*19 through *d*22 in the Tables 1A, 1B and 1C were obtained from the weighted Laplacian matrix, *d*19 = weighted maximum eigenvalue based on atomic number, *d*20 = weighted min- imum eigenvalue based on the atomic numbers, *d*21 = weighted average eigenvalue based on the atomic numbers, and *d*22 = weighted second smallest eigenvalue of the weighted Laplacian matrix. The resulting tables are below.

Molecule	Symbol	d1	d2	d3	d4	d5	d6	d7	d8
Arginine	R	8.00	7.00	12.00	6.00	8.120	6.00	12.00	1.50
Histidine	H	7.00	6.00	14.00	6.000	6.71	6.000	9.00	2.00
Lysine	K	6.00	5.00	10.00	4.00	7.00	5.00	9.00	1.667
Aspartic Acid	D	5.00	4.00	8.00	4.00	5.17	3.00	6.00	1.60
Glutamic Acid	E	6.00	5.00	10.00	5.00	6.00	4.00	8.00	1.667
Serine	S	3.00	2.00	4.00	2.00	1.670	2.00	3.00	1.333
Threonine	T	4.00	3.00	6.00	3.00	3.250	1.00	4.00	1.50
Asparagine	N	5.00	4.00	8.00	4.00	5.00	3.00	6.00	1.60
Glutamine	Q	6.00	5.00	10.00	4.00	5.860	4.00	8.00	1.667
Cysteine	C	3.00	2.00	4.00	2.00	2.33	1.00	3.00	1.333
Glycine	G	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Proline	P	4.00	4.00	8.00	4.00	4.00	4.00	4.000	2.00
Alanine	A	2.00	1.00	2.00	1.00	1.00	1.00	1.00	1.00
Isoleucine	I	5.00	4.00	8.00	4.00	3.25	3.00	6.00	1.600
Valine	V	4.00	3.00	6.00	3.00	3.25	1.00	4.00	1.50
Leucine	L	5.00	4.00	8.00	4.00	5.00	3.00	6.00	1.60
Methionine	M	5.00	4.00	8.00	4.00	5.40	3.00	7.00	1.60
Phenylalaine	F	8.00	8.00	14.00	6.00	7.00	6.000	11.000	1.750
Tyrosine	Y	9.00	9.00	18.00	7.00	8.88	6.000	13.000	2.000
Tryptophan	W	11.00	12.00	24.00	8.00	11.10	9.000	14.000	2.182

Table 1A

Molecule	Symbol	*d*9	*d*10	*d*11	*d*12	*d*13	*d*14	*d*15	*d*16
Arginine	R	12.499	-4.307	3.500	-2.590	19.00	31.444	20.00	38.00
Histidine	H	12.876	-3.721	4.286	-1.185	15.00	23.10	18.00	31.00
Lysine	K	10.363	-3.151	3.00	-0.536	12.00	24.50	18.00	31.00
Aspartic Acid	D	11.539	-4.178	3.20	0.528	12.00	16.40	12.00	20.00
Glutamic Acid	E	11.530	-3.425	3.333	-0.538	12.00	21.00	14.00	26.00
Serine	S	5.00	1.00	2.00	2.00	6.00	13.33	8.00	20.00
Threonine	T	9.928	-3.928	3.00	3.00	6.00	12.40	8.00	14.00
Asparagine	N	11.539	-4.178	3.20	0.528	12.00	16.50	14.00	20.00
Glutamine	Q	12.207	-4.255	3.333	-1.043	12.00	21.167	15.00	24.00
Cysteine	C	6.243	-2.243	2.00	2.00	6.00	16.670	12.00	22.00
Glycine	G	0.00	0.00	0.00	0.00	1.00	3.50	1.00	6.00
Proline	P	12.00	-4.00	4.00	4.00	12.00	12.00	12.00	12.00
Alanine	A	2.00	0.00	1.00	2.00	6.00	6.00	6.00	6.00
Isoleucine	I	10.851	-6.085	1.80	-1.517	12.00	15.60	12.00	18.00
Valine	V	9.928	-3.928	3.00	3.00	6.00	10.50	6.00	12.00
Leucine	L	11.029	-4.729	3.20	1.052	12.00	15.60	12.00	18.00
Methionine	M	9.49	-2.812	2.80	0.678	18.00	27.20	18.00	34.00
Phenylalaine	F	14.851	-4.801	4.25	-1.672	18.00	23.25	18.00	24.00
Tyrosine	Y	12.868	-4.793	4.333	-2.054	18.00	27.78	20.00	38.00
Tryptophan	W	13.511	-6.324	4.00	-2.576	24.00	27.50	18.00	36.00

Table 1B

Molecule	Symbol	*d*17	*d*18	*d*19	*d*20	*d*21	*d*22
Arginine	R	45.00	5.00	23.343	0.00	10.667	4.20
Hisitidine	H	47.00	4.70	24.243	-1.734	10.400	1.605
Lysine	K	37.00	6.17	22.739	-0.179	10.167	1.372
Aspartic Acid	D	34.00	6.80	28.634	0.00	10.40	2.969
Glutamic Acid	E	40.00	6.67	28.731	0.00	10.667	1.822
Serine	S	22.00	7.33	20.00	0.00	8.667	6.00
Threonine	T	27.00	5.40	23.819	-4.227	9.00	6.00
Asparagine	N	33.007	6.60	27.708	0.00	10.00	3.00
Glutamine	Q	39.00	6.50	27.831	0.00	10.50	1.849
Cysteine	C	28.00	9.33	28.00	0.00	11.333	6.00
Glycine	G	7.00	3.50	7.00	0.00	3.50	0.00
Proline	P	24.00	6.00	24.00	0.00	12.00	12.00
Alanine	A	12.00	6.00	12.00	0.00	6.00	0.00
Isoleucine	I	30.00	6.00	24.841	-1.641	9.60	3.373
Valine	V	24.007	6.00	24.00	0.00	9.00	6.00
Leucine	L	30.00	6.00	25.021	0.00	9.60	3.113
Methionine	M	40.00	8.00	31.344	0.00	13.60	2.656
Phenylalaine	F	48.00	6.00	26.993	0.00	12.00	2.026
Tyrosine	Y	56.00	6.22	28.252	-0.96	12.222	1.599
Tryptophan	W	68.00	5.667	29.778	0.211	12.75	2.044

Table 1C

2.1.2 Subdomains of the Protein

We next partition the sequence of CFTR that corresponds to the NBD2 domain into nine subsequences using 3GD7 from the Protein Data Bank [10]. In particular, we obtain the following subsequences denoted by S_i, i = 1…9 as shown in the table below. In determining these subsequences, we are guided by the secondary structures of the protein. Each subsequence contains one and only one type of secondary structure, either a beta strand, an alpha helix, or a loop. The loop regions may contain turns, a 3/10-helix or an alpha helix with no more than 6 residues. The corresponding subsequences of each subdomain are provided in Table 2.

Subdomain	Subsequence	Amino Acid Seqennce
S1	1209..1224	QMTVKDLTAKYTEGGN
S2	1225..1238	AILENISFSISPGQ
S3	1239..1261	RVGLLGRTGSGKSTLLSAFLRLL
S4	1262..1277	NTEGEIQIDGVSWDSI
S5	1278..1305	TLEQWRKAFGVIPQKVFIFSGTFRKNLD
S6	1306..1324	PNAAHSDQEIWKVADEVGL
S7	1325..1340	RSVIEQFPGKLDFVLV
S8	1341..1364	DGGCVLSHGHKQLMCLARSVLSKA
S9	1365..1391	KILLLDEPSAHLDPVTYQIIRRTLKQA

Table 2

The secondary structures that guide the cutoff points of each of the subdomains are given in the Table 3 below.

Subdomain	Subsequence	Reason
S1	1209..1224	beta strand, binding site, turn, bend
S2	1225..1238	binding site, beta strand, turn
S3	1239..1261	beta strand, binding site, turn, bend, alpha-helix
S4	1262..1277	bend, beta strand, turn
S5	1278..1305	binding site, alpha helix, bend, turn, beta strand
S6	1306..1324	turn, bend, alpha helix
S7	1325..1340	bend, turn, alpha helix, beta strand, 3/10 -alpha helix
S8	1341..1364	turn, bend, alpha helix, beta strand
S9	1365..1391	alpha helix, beta strand, bend

Table 3

We model the structure corresponding to each subsequence as a graph to obtain nine graphs which we denote by G_i where i = 1…9. The vertices(nodes) of each graph represent an amino acid in the sequence and edges are determined by a proximity measure of eight angstroms. The distance endpoints are determined by the center of mass of each amino acid residue in the 3D structure provided in 3GD7. Figure 3 is the graphs of the substructure S4 and S5.

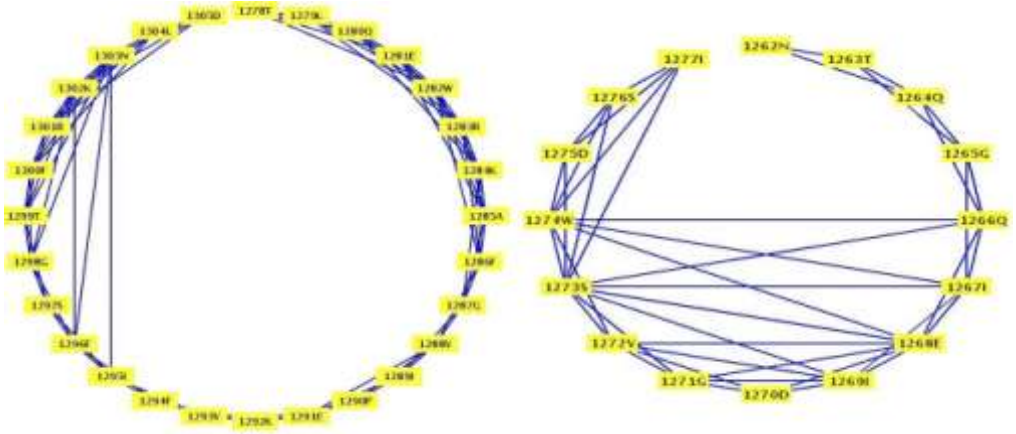

Figure 1: Subdomain Graphs for corresponding to S4 and S5

Finally, each subdomain graph such as the two above is represented by a vertex in the NBD2 CFTR domain graph. The edges of the NBD2 CFTR domain graph, or simply the domain graph, are based on a proximity measure where the distance endpoints are determined by a threshold distance of 8 angstroms between any two residues of each subdomain. Summarizing the method of construction, the nested graph has three layers. At the lowest level we have a collection of twenty small vertex-weighted graphs, one for each of the twenty most common amino acids. At the middle level, we have a collection of nine vertex-weighted graphs, in which each vertex represents an amino acid and the weights of the vertices are the combinatorial descriptors of the amino acid graphs at the lower level. At the highest level, (third in this case) we have a vertex-weighted graph that represents the nucleotide binding domain NBD2. The vertices in the domain graph each represent one of the subdomain graphs and the weights assigned to these vertices are derived from the vertex-weighted graph descriptors of each. Using these measures we obtain a final set of graph-theoretic-based measures for the domain graph of NBD2. The graph-theoretic model for the NDB2 is depicted in Figure 2.

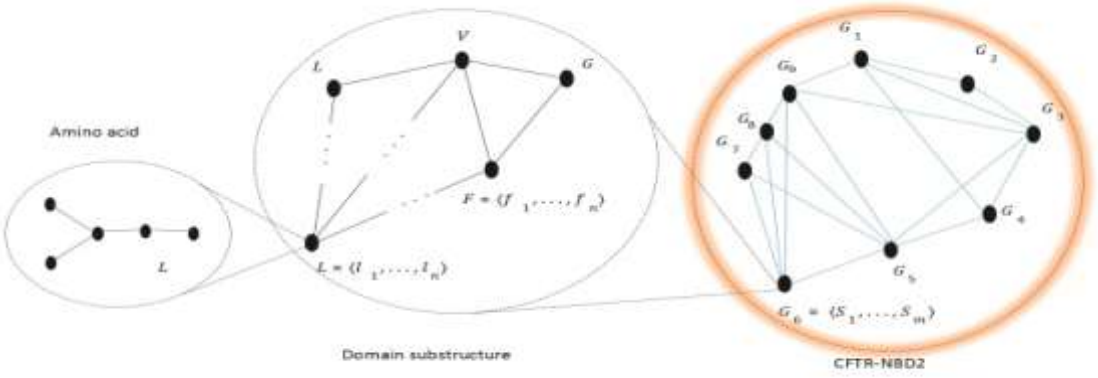

Figure 2: Nested Graph Model for NBD2

2.2 Virtual Mutations

Having obtained a set of measures for the wildtype NBD2 domain, we select eight disease causing mutations found in the Cystic Fibrosis Mutation Databank [11] that occur in NBD2. Given that we have selected eight mutations to model, we now obtain a set of graph-theoretic measures for each mutation in the following way. To measure the global structural effect of a single point mutation, we first make the corresponding change at the residue level. This change affects one of the subdomains . We obtain a new set of subdomain graph measures of the effected subdomain. Note that an amino acid switch (virtual mutation) produces a graph that contains a different residue together with different set of combinatorial descriptors. Since both the structure of, and possibly a vertex-weight for one vertex may have changed, this changes the corresponding vector of vertex-weights for the vertex in that substructure . In this way, we incorporate the graph-theoretic changes of a single point mutation with the predicted structural change by using the vertex-weights at each level. The edge set of the domain graph remains unchanged, but the weights of the vertices are adjusted according to the structural (both vertex and edge) changes of the underlying subdomain graphs. Since our measures are based on vertex-weights, we obtain a new set of values associated with each mutation. These, together with the wildtype, provide a set of measures for nine distinct graphs. The eight mutations that were selected and that occur most frequently in the population together with the clinical manisfestations are given in Table 4.

Mutation	CF
Wildtype	No
Y1212G	Mild
G1271E	Mild
S1347R	Mild
I1234V	Mild
D1270N	Mild
V1212W	Mild
S1235R	Mild
N1303K	Severe

Table 4 : Some CF-causing Mutations in NBD2

The resulting change in the descriptors for the main domain graph for each mutation is calculated and given below. The definitions of the t values for Table 7 follow: t1 = average of non-zero numbers in column 3 of the weighted adjacency matrix, t2 = average of non-zero numbers in column 4 of the weighted adjacency matrix, t3 = average of non-zero numbers in column 9 of the weighted adjacency matrix, t4 = average weighted degree of the top level graph (divided by thousand), t5 = total weighted degree of the top level graph (divided by thousand), t6 = weighted connectivity for row 1 of the weighted adjacency matrix (obtained by summing all the numbers on row 1), t7 = weighted connectivity for row 2 of the weighted adjacency matrix, t8 =weighted connectivity for row 3 of the weighted adjacency matrix, t9 = weighted connectivity for row 4 of the weighted adjacency matrix, and t10 = weighted connectivity for row 5 of the weighted adjacency matrix.

Mutation	**t1**	**t2**	**t3**	**t4**	**t5**	**t6**	**t7**	**t8**	**t9**	**t10**
Wildtype	4.42	2.99	4.93	0.90	8.12	13.95	7.66	18.31	13.20	24.60
Y1219G	4.33	2.90	4.49	0.82	7.37	13.65	6.67	18.00	13.11	24.43
V1212W	3.15	3.60	5.14	0.92	8.23	13.61	6.39	19.24	11.93	24.15
I1234V	4.67	2.99	4.93	0.90	8.10	14.20	7.90	18.30	13.44	24.85
S1235R	3.09	2.99	4.93	0.91	8.18	12.62	6.40	18.38	11.94	23.27
G1271E	4.04	2.99	4.93	0.91	8.23	13.57	7.40	18.55	13.06	23.99
S1347R	4.42	2.99	5.19	0.91	8.18	14.21	7.66	18.64	13.27	24.98
N1303K	4.44	3.01	4.95	0.91	8.14	14.01	7.68	18.35	13.22	27.95
D1270N	4.53	2.99	4.93	0.91	8.23	14.06	7.89	18.54	13.54	24.71

Table 5: Mutation-specific molecular descriptors for NBD2

3. RESULTS AND DISCUSSION

The R statistical Software [12] was used to cluster the mutations using the molecular descriptors for the top level graph when the single point mutations were performed. The single linkage function in R was used for our hierarchical clustering as it was ascertained to be the best to reduce biased results. The dendrogram clusters (shown in Figure 13) the wildtype mutation and other mutations using the molecular descriptors from Table 5 and the results are shown below.

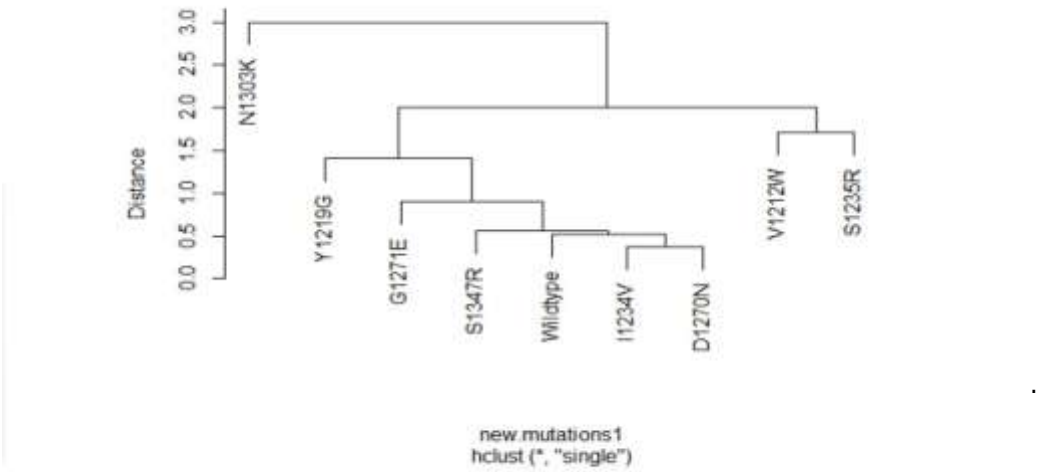

Figure 3 : Heirachical Clustering of CF-causing Mutations in NBD2

The N1303K mutation is one of the known mutations in NBD2 that causes severe cystic fibrosis. The evidence that the substitution of N with K at position 1303 leads to variations in the arrangement of the molecule when folded in the lab has not explained why the molecule does not fold appropriately in the cell. N1303K is said to be linked to pancreatic insufficiency cystic fibrosis [13,14] . Our results in Figure 2 show that the resulting structural effects of N1303K are expressively distinct from the wildtype. Also, it is obvious from our results that the difference between the wildtype and domain graphs caused by mutations like I1234V, S1345R and D1270N are less significant. More so, our results lead to a conclusion that Y1219G, G1271E, V1212W and S1235R are also considerably distinct from wildtype, even though they all belong to one larger cluster. Our findings call for the need for further investigations. For instance, questions such as, under what circumstance would N1303K match up to or mirror the wildtype in our nested graph model? In other words, what graph-theoretic or combinatorial descriptors of the graph containing N1303K would result in a graph that is very similar to

wildtype or will cause the clustering of N1303K along the wildtype or other mild mutations? Answers to such questions are of paramount importance since they might ultimately lead to a useful discernment into a line of action for the design of a molecule that can correct this specific mutation associated with cystic fibrosis.

4. CONCLUSION

A single point mutation that results in either the absence of a single residue or a switch from one residue to another in a protein structure can have a profound local effect, but how this local perturbation manifests to a global one remains unclear. In this work, using the nested graph model, we quantify the network of interacting amino acids by viewing the molecule as a system with levels of interaction. A systems biology approach such as this one may aid in helping researchers understand the consequences of a small mutation on a large protein molecule and thereby direct the design of a drug to address the faulty protein. Whereas we are focused on cystic fibrosis, the method easily extends to the study of any disease that is the consequence of a mutation, especially single point mutations such as muscular dystrophy and sickle cell anemia.

REFERENCES

1. K. Roberts, P. Cushing, P. Boisguerin, D. Madden, and B. Donald, "Computational Design of a PDZ domain peptide inhibitor that rescues CFTR activity," PLOS Computational Biology, vol. 8, no. 4, Article ID e1002477, 2012.
2. A. Aleksandrov, P. Kota, L. Cui et al., "Allosteric modulation balances thermodynamic stability and restores function of ΔF508 CFTR," Journal of Molecular Biology, vol. 419, pp. 41–60, 2012.
3. A. W. R. Serohijos, T. Hegedus, A. A. Aleksandrov et al., "Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D structure crucial to assembly and channel function,"Proceedings of the National Academy of Sciences of the United States of America, vol. 105, no. 9, pp. 3256–3261, 2008.
4. L. He, A. A. Aleksandrov, A. W. R. Serohijos et al., "Multiple membrane-cytoplasmic domain contacts in the cystic fibrosis transmembrane conductance regulator (CFTR) mediate regulation of channel gating,"Journal of Biological Chemistry, vol. 283, no. 39, pp. 26383–26390, 2008.
5. A. W. R. Serohijos, T. Hegedus, J. R. Riordan, and N. V. Dokholyan, "Diminished self-chaperoning activity of the ΔF508 mutant of CFTR results in protein misfolding," PLoS Computational Biology, vol. 4, no. 2, Article ID e1000008, 2008.
6. D. J. Knisley, J.R. Knisley and A. C. Herron, "Graph-Theoretic Models of Mutations in the Nucleotide Binding Domain 1 of the Cystic Fibrosis Transmembrane Conductance Regulator," Computational Biology Journal, vol. 2013, Article ID 938169, 9 pages, 2013. doi:10.1155/2013/938169
7. J.M. Chen, C. Cutler, C. Jacques, G. Bœuf, E. Denamur, G. Lecointre, B. Mercier, G. Cramb, and C. F´erec. , "A combined analysis of the cystic fibrosis transmembrane conductance regulator: Implications for structure and disease models, . Molecular Biology and Evolution, 18(9):1771–1788, 2001.
8. D. Knisley and J Knisley, "Predicting protein-protein interactions using graph invariants and a neural network " Computationa lBiology and Chemistry, 35(2):108 -– 113, 2011.
9. L. Osborne, R. Knight, G. Santis, and M. Hodson. " A mutation in the second nucleotide binding fold of the cystic fibrosis gene", 48(PMC1682979), 1991/03/.
10. The Protein Data Bank, www.pdb.org (Crystal Structure of Human Nbd2 Complexed With N6- Phenylethyl-atp (P-atp).
11. The Clinical Functional and Translation Website for Cystic Fibrosis, www.cftr2.org
12. 12. R: A language and environment for statistical computing, http://www.R-project.org/, Retrieved on 3/12/2015.
13. About cystic fibrosis. http://www.cff.org/AboutCF, Retrieved on 3/20/2015.
14. D. Rapino, I. Sabirzhanova, M. Lopes-Pacheco, R. Grover, W.B. Guggino, and L. Cebotaru, "Rescue of NBD2 mutants of N1303K and S1235R of CFTR by small-molecule correctors and transcomplementation" ,.PLOS ONE, 10(3):e0119796. doi: 10.1371/journal.pone.0119796. eCollection 2015.