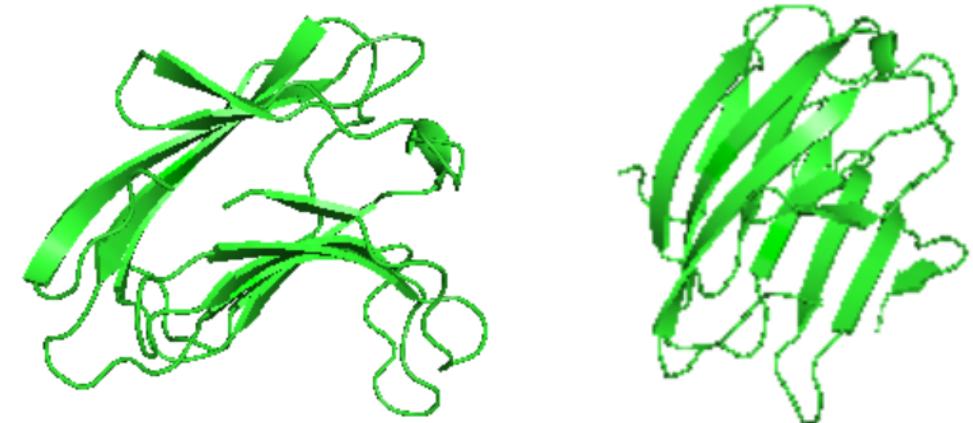
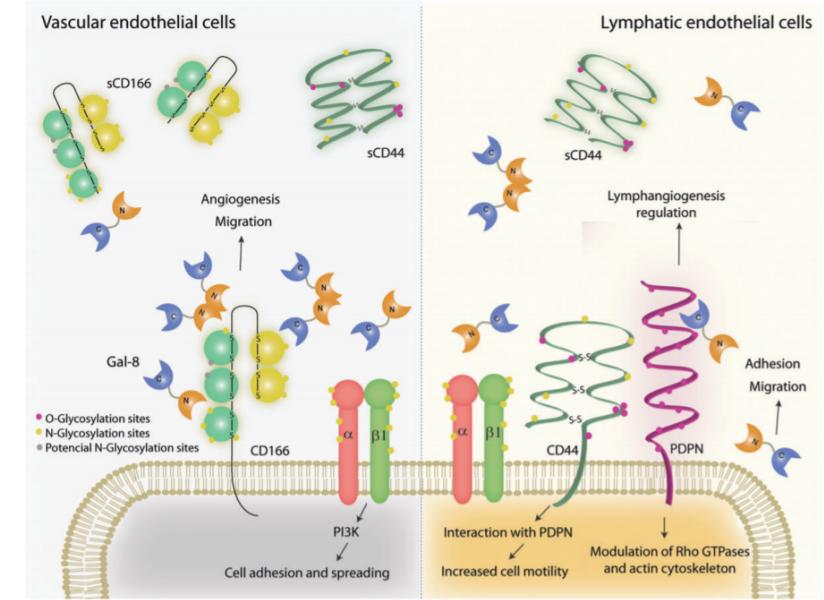


# Analysis of the I35 Mutations of Galactin-8

CBB 752 Final project

# Galectin-8

- Member of the galectin family
  - Binds  $\beta$ -galactose
  - Contains carbohydrate recognition domain (CRD)
  - Binds integrins, CD44, etc.
- Role in tumorigenesis
  - Controls endothelial cell migration and angiogenesis

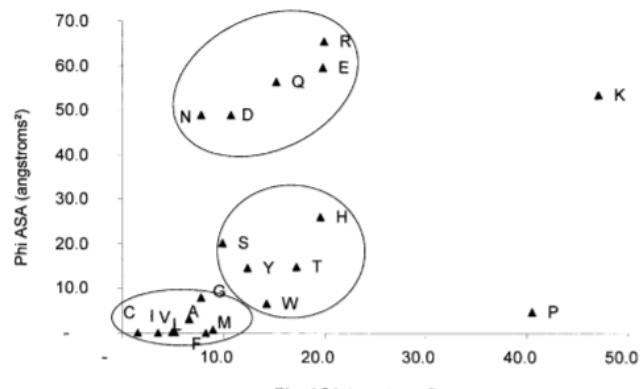


Cummings and Liu, 2009  
Troncoso et al. 2014

PDB 4BMB

# Amino Acid Mutations

- Size
- Charge
- Hydrophobic effect
- Change in potential energy



Lins et al. 2003

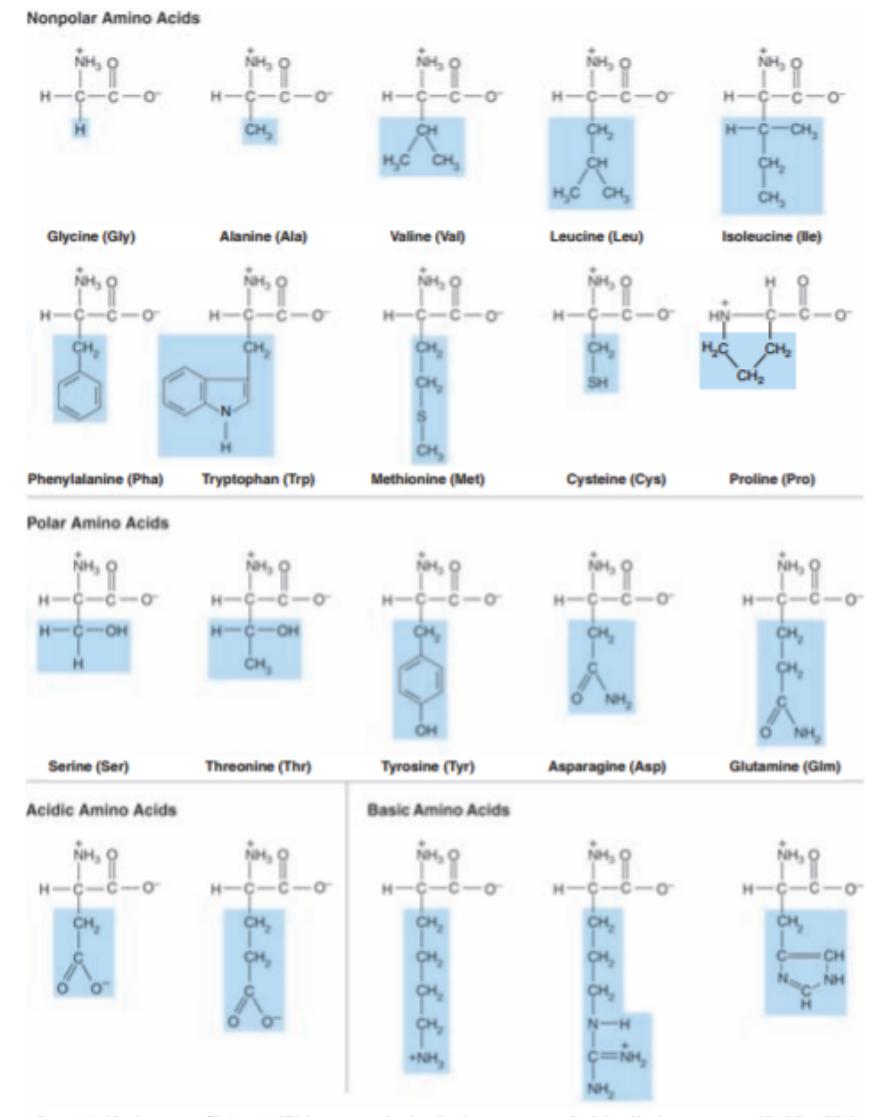


FIGURE 5.2

McKee, T., & McKee, J. R. (2012). *Biochemistry: The molecular basis of life*. Oxford: Oxford University Press.

# Common Mutations

REPLACEMENT AMINO ACID

	ORIGINAL AMINO ACID																			
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
A Ala	9867	2	9	10	3	8	17	21	2	6	4	2	6	2	22	35	32	0	2	18
R Arg	1	9913	1	0	1	10	0	0	10	3	1	19	4	1	4	6	1	8	0	1
N Asn	4	1	9822	36	0	4	5	6	21	3	1	13	0	1	2	20	9	1	4	1
D Asp	6	0	42	9859	0	6	53	6	4	1	0	3	0	0	1	5	3	0	0	1
C Cys	1	1	0	0	9973	0	0	0	1	1	0	0	0	0	1	5	1	0	3	2
Q Gln	3	9	4	5	0	9876	27	1	23	1	3	6	4	0	6	2	2	0	0	1
E Glu	10	0	7	56	0	35	9865	4	2	3	1	4	1	0	3	4	2	0	1	2
G Gly	21	1	12	11	1	3	7	9935	1	0	1	2	1	1	3	21	3	0	0	5
H His	1	2	18	3	1	20	1	0	9912	0	1	1	0	2	3	1	1	1	4	1
I Ile	2	2	3	1	2	1	2	0	0	9872	9	2	12	7	0	1	7	0	1	33
L Leu	3	1	3	0	0	6	1	1	4	22	9947	2	45	13	3	1	3	4	2	15
K Lys	2	37	25	6	0	12	7	2	2	4	1	9926	20	0	3	8	11	0	1	1
M Met	1	1	0	0	0	2	0	0	0	5	8	4	9874	1	0	1	2	0	0	4
F Phe	1	1	1	0	0	0	0	1	2	8	6	0	4	9946	0	2	1	3	28	0
P Pro	13	5	2	1	1	8	3	2	5	1	2	2	1	1	1	9926	12	4	0	2
S Ser	28	11	34	7	11	4	6	16	2	2	1	7	4	3	17	9840	38	5	2	2
T Thr	22	2	13	4	1	3	2	2	1	11	2	8	6	1	5	32	9871	0	2	9
W Trp	0	2	0	0	0	0	0	0	0	0	0	0	1	0	1	0	9976	1	0	0
Y Tyr	1	0	3	0	3	0	1	0	4	1	1	0	0	21	0	1	1	2	9945	1
V Val	13	2	1	1	3	2	2	3	3	57	11	1	17	1	3	2	10	0	2	9901

PAM 1

REPLACEMENT AMINO ACID

	ORIGINAL AMINO ACID																			
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
A Ala	13	6	9	9	5	8	9	12	6	8	6	7	7	4	11	11	11	2	4	9
R Arg	3	17	4	3	2	5	3	2	6	3	2	9	4	1	4	4	3	7	2	2
N Asn	4	4	6	7	2	5	6	4	6	3	2	5	3	2	4	5	4	2	3	3
D Asp	5	4	8	11	1	7	10	5	6	3	2	5	3	1	4	5	5	1	2	3
C Cys	2	1	1	1	52	1	1	2	2	2	1	1	1	1	1	2	3	2	1	4
Q Gln	3	5	5	6	1	10	7	3	7	2	3	5	3	1	4	3	3	1	2	3
E Glu	5	4	7	11	1	9	12	5	6	3	2	5	3	1	4	5	5	1	2	3
G Gly	12	5	10	10	4	7	9	27	5	5	4	6	5	3	8	11	9	2	3	7
H His	2	5	5	4	2	7	4	2	15	2	2	3	2	2	3	3	2	2	3	2
I Ile	3	2	2	2	2	2	2	2	2	10	6	2	6	5	2	3	4	1	3	9
L Leu	6	4	4	3	2	6	4	3	5	15	34	4	20	13	5	4	6	6	7	13
K Lys	6	18	10	8	2	10	8	5	8	5	4	24	9	2	6	8	8	4	3	5
M Met	1	1	1	1	0	1	1	1	1	2	3	2	6	2	1	1	1	1	2	2
F Phe	2	1	2	1	1	1	1	1	3	5	6	1	4	32	1	2	2	4	20	3
P Pro	7	5	5	4	3	5	4	5	5	3	2	4	3	2	20	6	5	1	2	4
S Ser	9	6	8	7	7	6	7	9	6	5	4	7	5	3	9	10	9	4	4	6
T Thr	8	5	6	6	4	5	5	6	4	6	4	6	5	3	6	8	11	2	3	6
W Trp	0	2	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0
Y Tyr	1	1	2	1	3	1	1	1	3	2	2	1	2	15	1	2	2	3	31	2
V Val	7	4	4	4	4	4	4	5	4	15	10	4	10	5	5	5	7	2	4	17

PAM 250

Dayhoff et al. 1978

# Common Mutations

ORIGINAL AMINO ACID

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
REPLACEMENT AMINO ACID	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
A Ala	9867	2	9	10	3	8	17	21	2	6	4	2	6	2	22	35	32	0	2	18
R Arg	1	9913	1	0	1	10	0	0	10	3	1	19	4	1	4	6	1	8	0	1
N Asn	4	1	9822	36	0	4	6	6	21	3	1	13	0	1	2	20	9	1	4	1
D Asp	6	0	42	9859	0	6	53	6	4	1	0	3	0	0	0	1	5	3	0	0
C Cys	1	1	0	0	9973	0	0	0	1	1	0	0	0	0	0	1	5	1	0	2
Q Gln	3	9	4	5	0	9876	27	1	23	1	3	6	4	0	6	2	2	0	0	1
E Glu	10	0	7	56	0	35	9865	4	2	3	1	4	1	0	3	4	2	0	1	2
G Gly	21	1	12	11	1	3	7	9935	1	0	1	2	1	1	3	21	3	0	0	5
H His	9	18	3	1	20	1	0	9912	0	1	1	0	2	3	1	1	1	4	1	
I Ile	2	3	1	2	1	2	0	0	9872	9	2	12	7	0	1	7	0	1	33	
L Leu	3	1	3	0	0	6	1	1	4	22	9947	2	45	13	3	1	3	4	2	15
K Lys	2	37	25	6	0	12	7	2	2	4	1	9926	20	0	3	8	11	0	1	1
M Met	1	1	0	0	0	2	0	0	0	5	8	4	9874	1	0	1	2	0	0	4
F Phe	1	1	1	0	0	0	0	1	2	8	6	0	4	9946	0	2	1	3	28	0
P Pro	13	5	2	1	1	8	3	2	5	1	2	2	1	1	9926	12	4	0	0	2
S Ser	28	11	34	7	11	4	6	16	2	2	1	7	4	3	17	9840	38	5	2	2
T Thr	22	2	13	4	1	3	2	2	1	11	2	8	6	1	5	32	9871	0	2	9
W Trp	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	9976	1	0	0
Y Tyr	1	0	3	0	3	0	1	0	4	1	1	0	0	21	0	1	1	2	9945	1
V Val	13	2	1	1	3	2	2	3	3	57	11	1	17	1	3	2	10	0	2	9901

PAM 1

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
REPLACEMENT AMINO ACID	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
A Ala	13	6	9	9	5	8	9	12	6	8	6	7	7	4	11	11	11	2	4	9
R Arg	3	17	4	3	2	5	3	2	6	3	2	9	4	1	4	4	3	7	2	2
N Asn	4	4	6	7	2	5	6	4	6	3	2	5	3	2	4	5	4	2	3	3
D Asp	5	4	8	11	1	7	10	5	6	3	2	5	3	1	4	5	5	1	2	3
C Cys	2	1	1	1	52	1	1	2	2	2	1	1	1	1	1	2	3	2	1	4
Q Gln	3	5	5	6	1	10	7	3	7	2	3	5	3	1	4	3	3	1	2	3
E Glu	5	4	7	11	1	9	12	5	6	3	2	5	3	1	4	5	5	1	2	3
G Gly	12	5	10	10	4	7	9	27	5	5	4	6	5	3	8	11	9	2	3	7
H His	2	5	5	4	2	7	4	2	15	2	2	3	2	2	3	3	2	2	2	2
I Ile	3	2	2	2	2	2	2	2	2	2	10	6	2	6	5	2	3	4	1	9
L Leu	6	4	4	3	2	6	4	3	5	15	34	4	20	13	5	4	6	6	7	13
K Lys	6	18	10	8	2	10	8	5	8	5	4	24	9	2	6	8	8	4	3	5
M Met	1	1	1	1	0	1	1	1	1	2	3	2	6	2	1	1	1	1	2	2
F Phe	2	1	2	1	1	1	1	1	3	5	6	1	4	32	1	2	2	4	20	3
P Pro	7	5	5	4	3	5	4	5	5	3	2	4	3	2	20	6	5	1	2	4
S Ser	9	6	8	7	7	6	7	9	6	5	4	7	5	3	9	10	9	4	4	6
T Thr	8	5	6	6	4	5	5	6	4	6	4	6	5	3	6	8	11	2	3	6
W Trp	0	2	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0
Y Tyr	1	1	2	1	3	1	1	1	3	2	2	1	2	15	1	2	2	3	31	2
V Val	7	4	4	4	4	4	4	5	4	15	10	4	10	5	5	5	7	2	4	17

PAM 250

# Common Mutations

REPLACEMENT AMINO ACID

	ORIGINAL AMINO ACID																			
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
	Ala	Arg	Asn	Asp	Cys	Gln	Glut	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
A Ala	9867	2	9	10	3	8	17	21	2	6	4	2	6	2	22	35	32	0	2	18
R Arg	1	9913	1	0	1	10	0	0	10	3	1	19	4	1	4	6	1	8	0	1
N Asn	4	1	9822	36	0	4	6	6	21	3	1	13	0	1	2	20	9	1	4	1
D Asp	6	0	42	9859	0	6	53	6	4	1	0	3	0	0	1	5	3	0	0	1
C Cys	1	1	0	0	9973	0	0	0	1	1	0	0	0	0	1	5	1	0	3	2
Q Gln	3	9	4	5	0	9876	27	1	23	1	3	6	4	0	6	2	2	0	0	1
E Glu	10	0	7	56	0	35	9865	4	2	3	1	4	1	0	3	4	2	0	1	2
G Gly	21	1	12	11	1	3	7	9935	1	0	1	2	1	1	3	21	3	0	0	5
H His	9	18	3	1	20	1	0	9912	0	1	1	0	2	3	1	1	1	4	1	
I Ile	2	3	1	2	1	2	0	0	9872	9	2	2	7	0	1	7	0	1	33	
L Leu	3	1	3	0	0	6	1	1	4	22	9947	2	45	13	3	1	3	4	2	15
K Lys	2	37	25	6	0	12	7	2	2	4	1	9926	20	0	3	8	11	0	1	1
M Met	1	1	0	0	0	2	0	0	0	5	8	4	9874	1	0	1	2	0	0	4
F Phe	1	1	1	0	0	0	0	1	2	8	6	0	4	9946	0	2	1	3	28	0
P Pro	13	5	2	1	1	8	3	2	5	1	2	2	1	1	9926	12	4	0	0	2
S Ser	28	11	34	7	11	4	6	16	2	2	1	7	4	3	17	9840	38	5	2	2
T Thr	22	2	13	4	1	3	2	2	1	11	2	8	6	1	5	32	9871	0	2	9
W Trp	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	9976	1	0	
Y Tyr	1	0	3	0	3	0	1	0	4	1	1	0	0	21	0	1	1	2	9945	1
V Val	13	2	1	1	3	2	2	3	3	57	11	1	17	1	3	2	10	0	2	9901

PAM 1

REPLACEMENT AMINO ACID

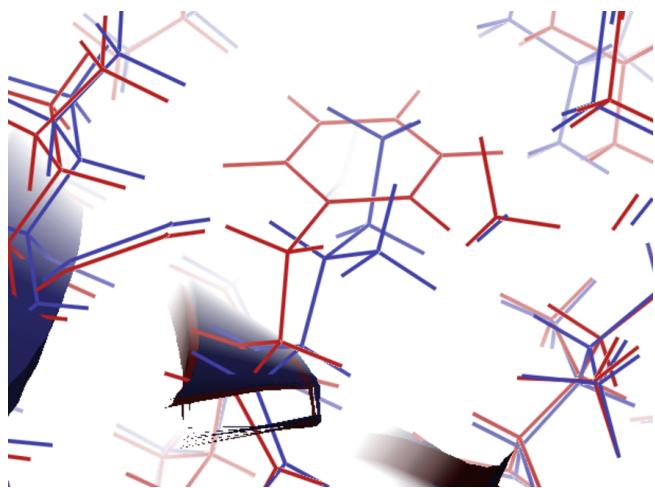
	ORIGINAL AMINO ACID																			
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
	Ala	Arg	Asn	Asp	Cys	Gln	Glut	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
A Ala	13	6	9	9	5	8	9	12	6	8	6	7	7	4	11	11	11	2	4	9
R Arg	3	17	4	3	2	5	3	2	6	3	2	9	4	1	4	4	3	7	2	2
N Asn	4	4	6	7	2	5	6	4	6	3	2	5	3	2	4	5	4	2	3	3
D Asp	5	4	8	11	1	7	10	5	6	3	2	5	3	1	4	5	5	1	2	3
C Cys	2	1	1	1	52	1	1	2	2	2	1	1	1	1	1	2	3	2	1	4
Q Gln	3	5	5	6	1	10	7	3	7	2	3	5	3	1	4	3	3	1	2	3
E Glu	5	4	7	11	1	9	12	5	6	3	2	5	3	1	4	5	5	1	2	3
G Gly	12	5	10	10	4	7	9	27	5	5	4	6	5	3	8	11	9	2	3	7
H His	2	5	5	4	2	7	4	2	15	2	2	3	2	2	3	3	2	2	2	2
I Ile	3	2	2	2	2	2	2	2	2	10	6	2	5	5	3	4	1	3	9	9
L Leu	6	4	4	3	2	6	4	3	5	15	34	4	20	11	5	4	6	6	7	13
K Lys	6	18	10	8	2	10	8	5	8	5	4	24	9	2	6	8	8	4	3	5
M Met	1	1	1	1	0	1	1	1	1	2	3	2	6	2	1	1	1	1	2	
F Phe	2	1	2	1	1	1	1	1	1	3	5	6	1	4	32	1	2	2	4	20
P Pro	7	5	5	4	3	5	4	5	5	3	3	4	2	2	20	6	5	1	2	4
S Ser	9	6	8	7	7	6	7	9	6	5	4	7	5	3	9	10	9	4	4	6
T Thr	8	5	6	6	4	5	5	6	4	6	4	6	5	3	6	8	11	2	3	6
W Trp	0	2	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	55	1
Y Tyr	1	1	2	1	3	1	1	1	3	2	2	1	2	15	1	2	2	3	31	2
V Val	7	4	4	4	4	4	4	5	4	15	10	4	10	5	5	5	7	2	4	17

PAM 250

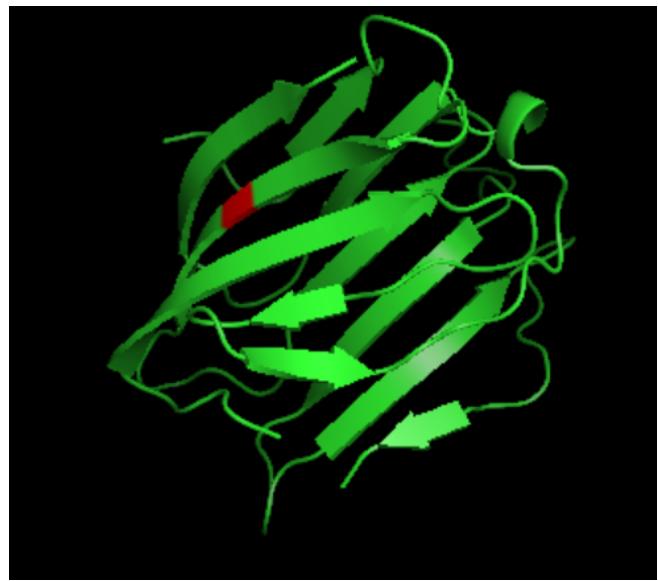
Dayhoff et al. 1978

# I35F

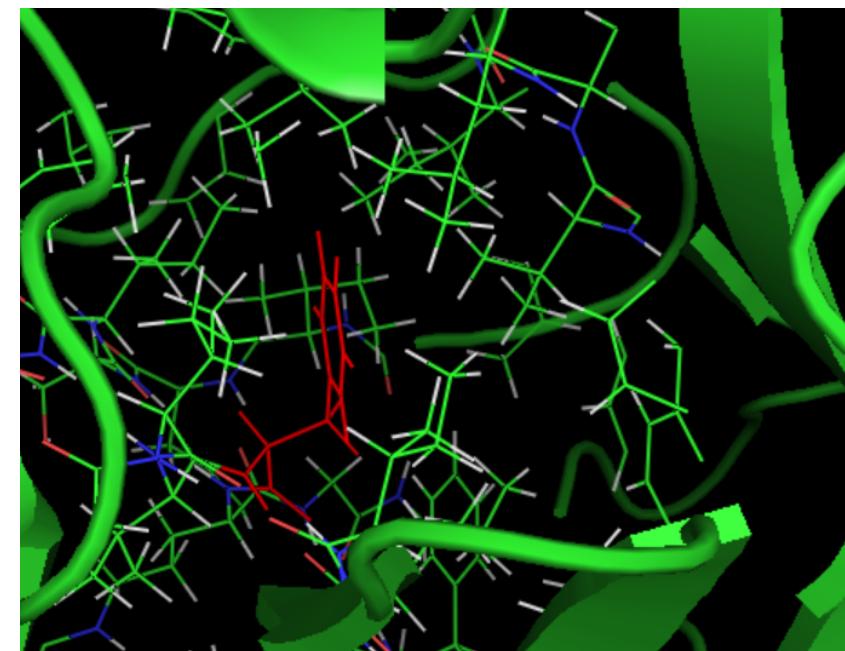
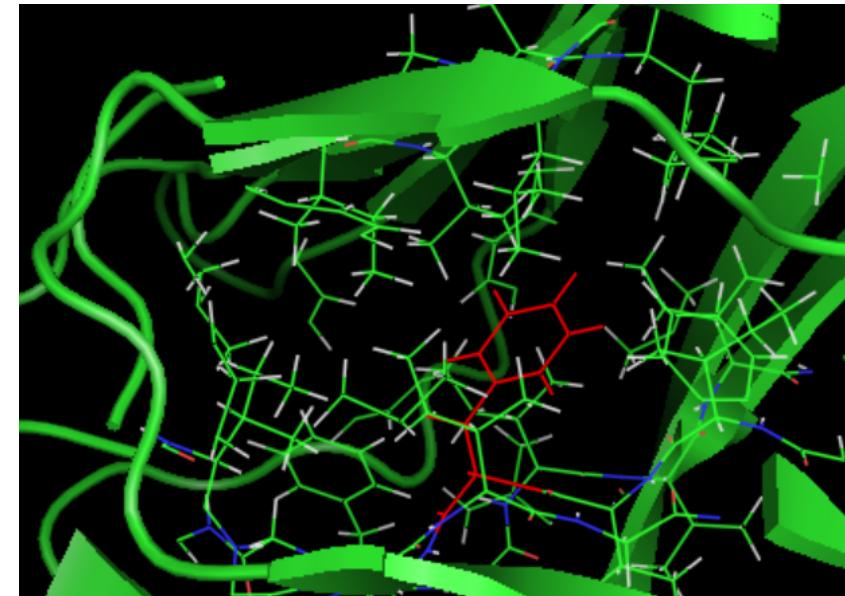
Rosetta Energy: +4.821  
→ Damaging mutation



RMSD



I35

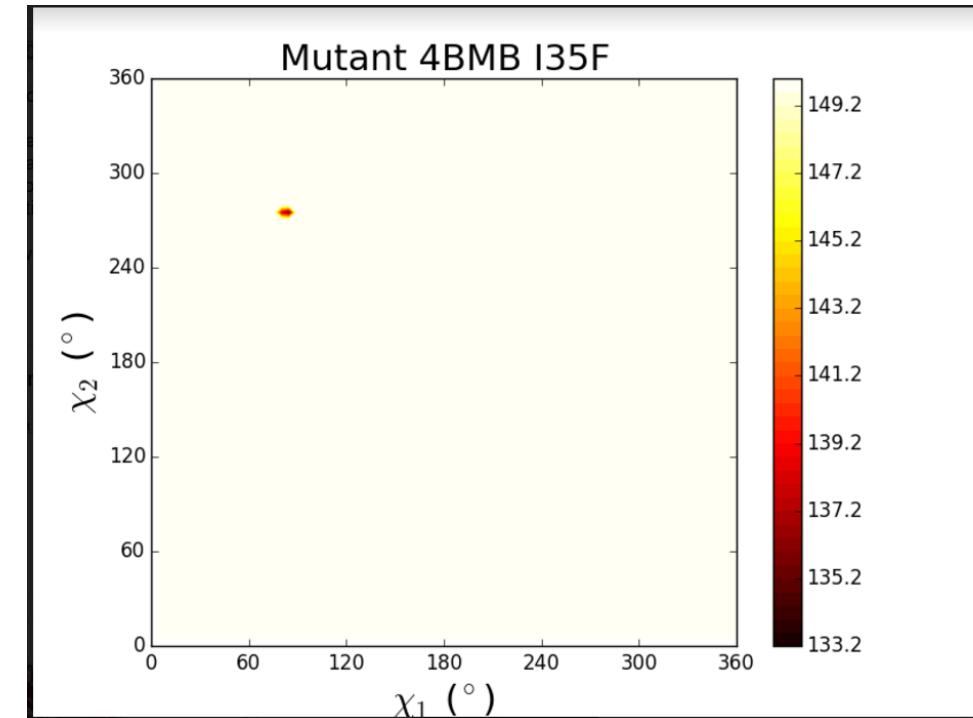
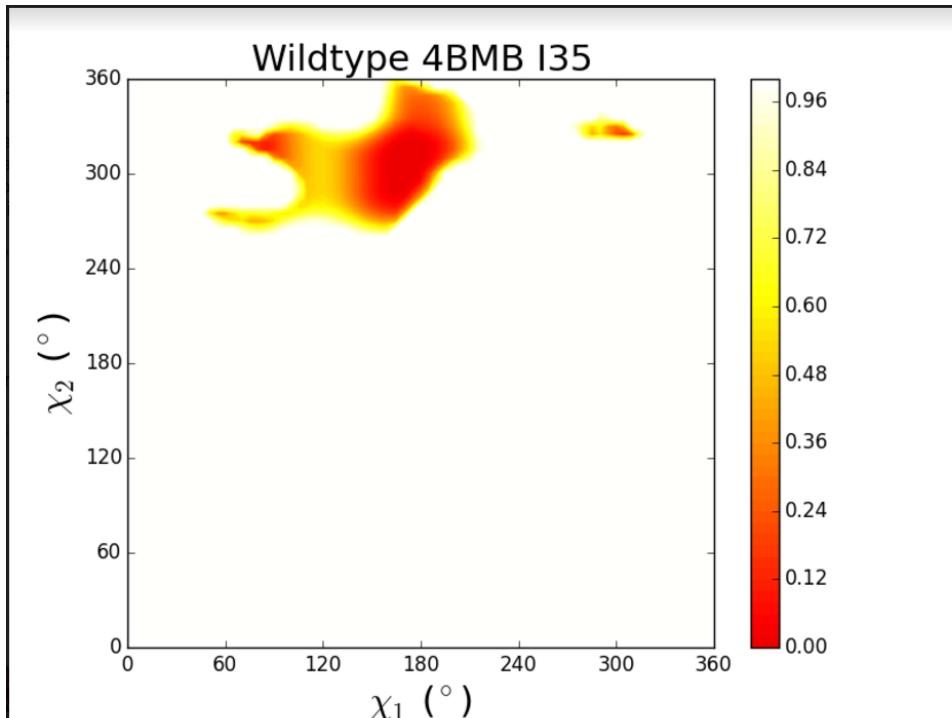


I35F

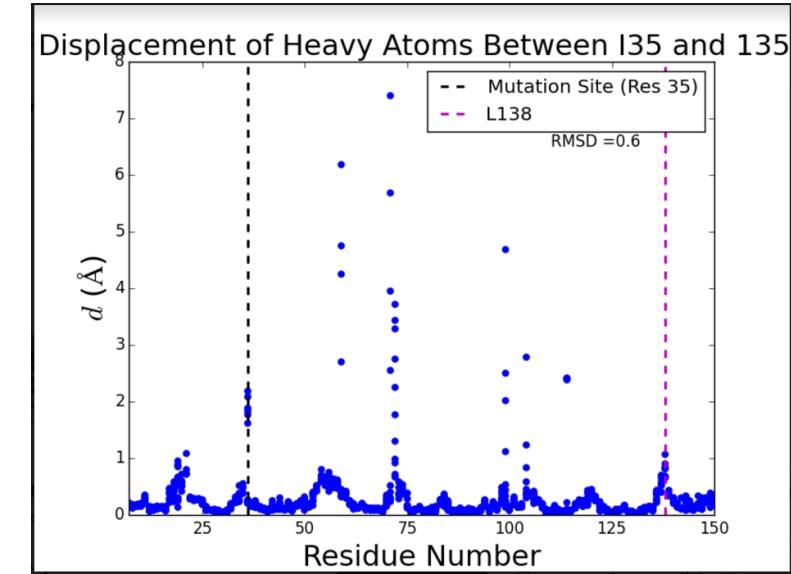
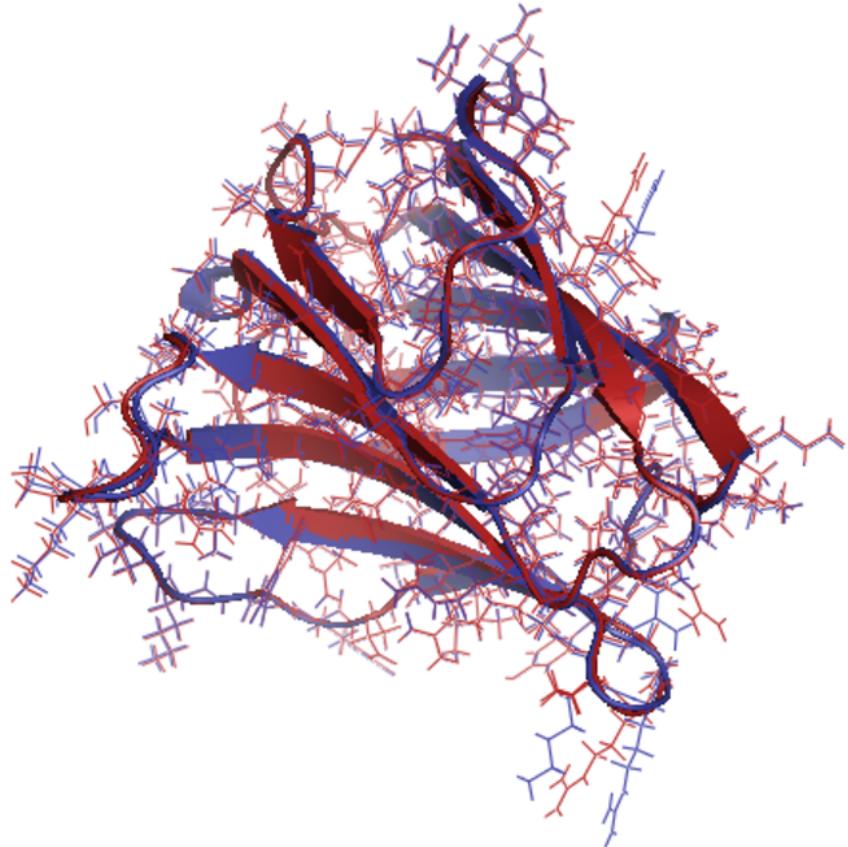
# Coding Problem:

- Calculate total potential energy of interactions of side chains
  - $U_{RLJ}(\chi_1, \chi_2) = \sum_{i>j} U_{RLJ}(r_{ij})$
  - Plot as function of  $\chi_1$  and  $\chi_2$
- RMSD of heavy chain atoms

# I35F Energy Plots



# I35F RMSD Calculations



Overall heavy atom  
RMSD: 0.591

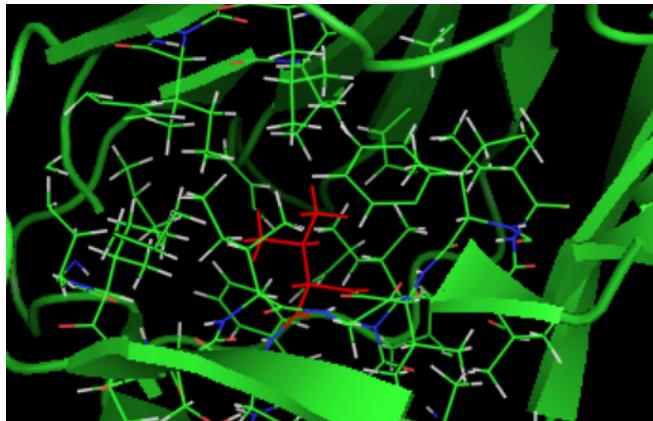
# Pipeline: Rosetta

- Lennard-Jones attractive
- Internal energy of sidechain rotamers (Dunbrack)
- Electrostatic
- Lennard-Jones repulsive (intraresidue and interresidue)
- Lazaridis-Jarplus solvation energy
- Sidechain-backbone hydrogen bond energy
- Distant backbone-backbone hydrogen bond
- Sidechain-sidechain hydrogen bond energy
- Omega backbone dihedral
- Overlap clashes
- Probability at phi-psi angle
- Ramachandran preferences

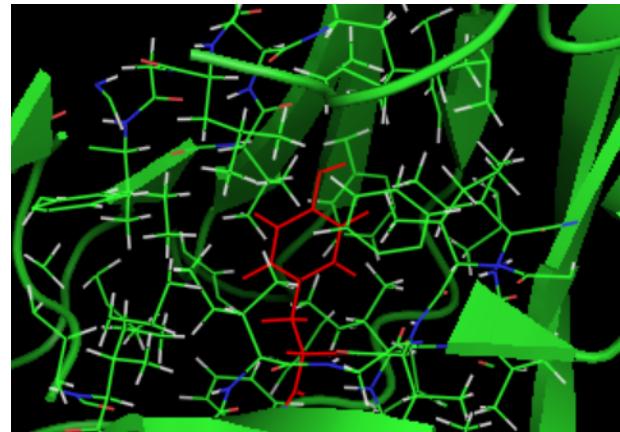
Mutation	Change in Energy (in REU)
I35V	-2.775
I35Y	-2.660
I35E	-2.652
I35L	-1.712
I35A	-1.097
I35P	-0.115
I35G	1.013
I35D	1.887
I35M	2.087
I35F	4.821
I35R	11.913

SCORE:	description	total_score	score	dslf_fa13	fa_atr	fa_dun	fa_ele	fa_intra	fa_rep	fa_sol	hbond_t	hbond_lr	hbond_s	hbond_ss	linear_cf	omega	overlap_p_aa_pp	pro_clos	rama	ref	time	gihh_planar	
SCORE: 4bmb.A-I32V_0001_0001	-307.053	-2.775	-307.05	0	-674.93	166.02	-80.646	1664	64.254	338.62	-10.005	-66.075	-13.718	-10.077	0	16.374	0	-31.176	0.57	-10.052	2.109	0	0.02
SCORE: 4bmb.A-I32Y_0001_0001	-306.938	-2.660	-306.94	0	-697.17	166.06	-79.877	1679	68.8	354.53	-11.504	-65.213	-11.26	-9.901	0	16.274	0	-31.557	0.702	-10.01	1.292	0	0.22
SCORE: 4bmb.A-I32E_0001_0001	-306.93	-2.852	-306.93	0	-683.55	166.54	-79.941	1639	64.049	347.2	-10.292	-66.535	-9.798	-9.962	0	15.374	0	-30.947	0.354	-10.24	-0.832	0	0.015
SCORE: 4bmb.A-I32L_0001_0001	-305.99	-1.712	-305.99	0	-687.24	165.87	-81.77	1605	68.258	347.22	-10.688	-65.678	-9.322	-9.435	0	15.546	0	-31.859	0.707	-11.118	1.89	0	0.025
SCORE: 4bmb.A-I32A_0001_0001	-305.375	-1.097	-305.38	0	-683.98	165.38	-82.126	1662	66.976	345.79	-10.552	-66.36	-11.951	-9.959	0	16.903	0	-30.575	0.583	-9.078	1.903	0	0.012
SCORE: 4bmb.A-I32P_0001_0001	-304.393	-0.115	-304.39	0	-670.94	163.13	-81.853	1689	65.125	337.56	-10.372	-64.641	-12.697	-10.175	0	18.059	0	-30.839	0.932	-10.257	0.879	0	0.008
SCORE: 4bmb.A_0001_0001	-304.278	-0.103	-304.28	0	-677.05	166.62	-78.964	1662	63.431	340.38	-9.247	-66.104	-10.908	-10.3	0	15.197	0	-30.793	0.298	-10.724	2.21	0	0.013
SCORE: 4bmb.A-I32G_0001_0001	-303.265	1.013	-303.27	0	-680.61	164.61	-84.174	1652	65.345	348.48	-10.425	-65.981	-11.989	-10.167	0	17.247	0	-29.897	0.55	-9.212	1.302	0	0.007
SCORE: 4bmb.A-I32D_0001_0001	-302.391	1.887	-302.39	0	-682.15	168.81	-85.316	1663	64.98	351.52	-10.047	-66.828	-11.008	-10.101	0	16.938	0	-31.094	0.587	-9.862	-0.501	0	0.021
SCORE: 4bmb.A-I32M_0001_0001	-302.191	2.087	-302.19	0	-675.96	165.69	-80.897	165	65.919	342.51	-9.519	-68.068	-9.471	-9.668	0	15.679	0	-31.275	0.35	-10.514	1.379	0	0.006
SCORE: 4bmb.A-I32F_0001_0001	-299.457	4.821	-299.46	0	-677.39	169.18	-80.914	1693	65.655	342.25	-11.349	-66.035	-12.652	-10.139	0	18.864	0	-30.667	0.407	-10.329	1.748	0	0.216
SCORE: 4bmb.A-I32R_0001_0001	-292.365	11.913	-292.37	0	-668.8	163.84	-78.859	1649	63.435	344.89	-8.89	-65.75	-10.822	-10.214	0	17.684	0	-30.813	0.505	-11.037	0.805	0	0.003

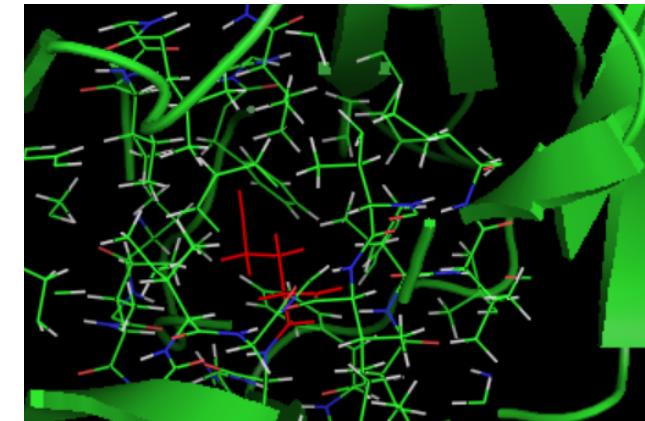
# Favorable Mutations



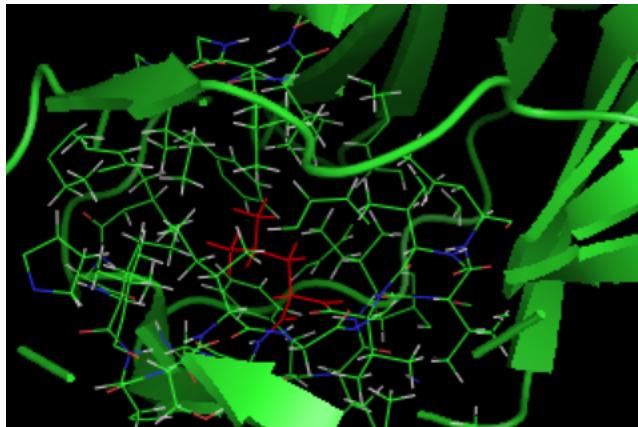
I35V  
-2.775



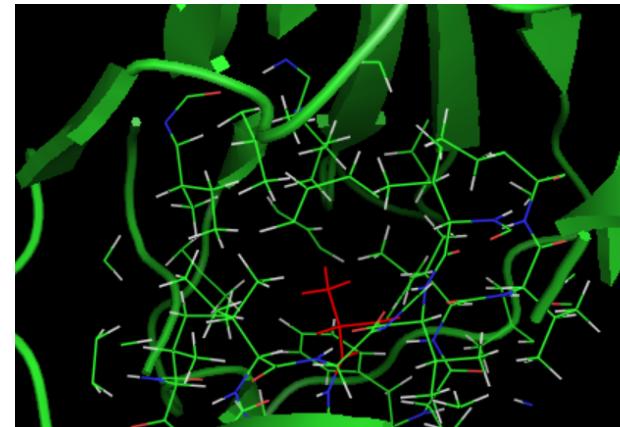
I35Y  
-2.660



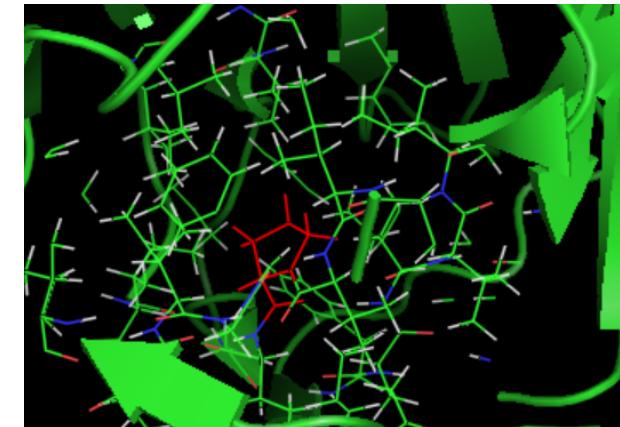
I35E  
-2.652



I35L  
-1.712

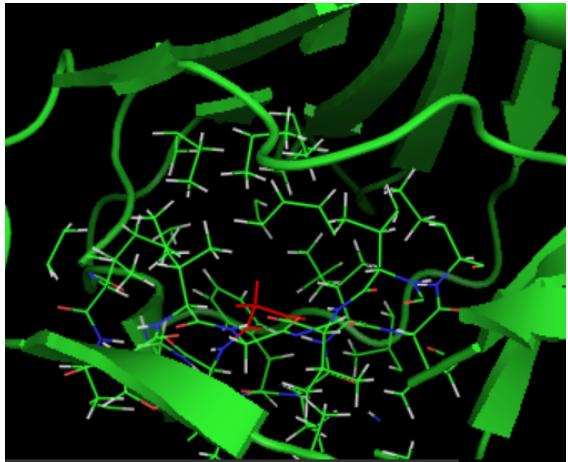


I35A  
-1.097

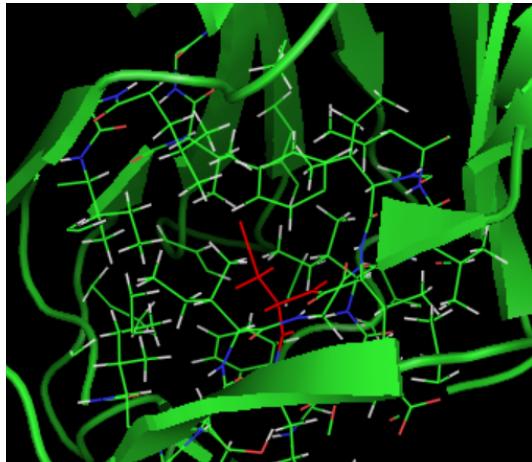


I35P  
-0.115

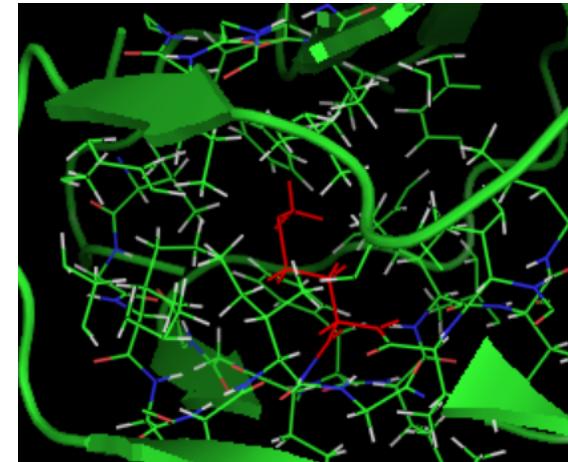
# Unfavorable Mutations



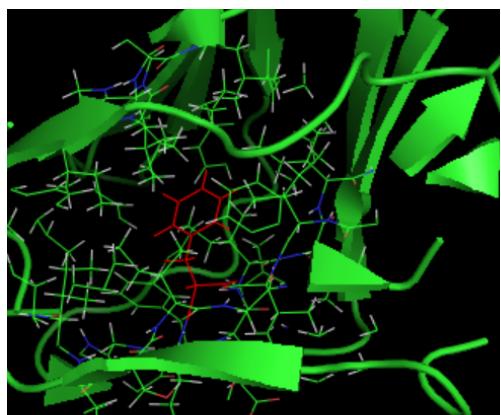
I35G  
+1.013



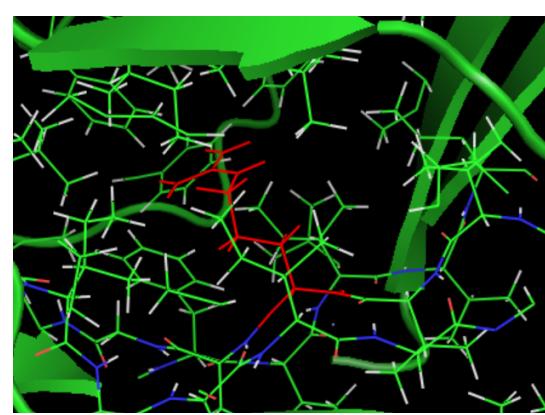
I35D  
+1.887



I35M  
+2.087



I35F  
+4.821



I35R  
+11.913

# Conclusions

- Size and charge important in predicting mutation effects
- Side chain angle potential energies increased by I35F mutation
- RMSD increased with I35F mutation
- Unexpected mutation results (D/E, F/Y)
- Interacting residues critical