UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Merck Sharp & Dohme LLC,
Petitioner,

v.

Halozyme Inc.,
Patent Owner.

Case No. PGR2025-00017
U.S. Patent No. 12,110,520

Declaration of Dr. Sheldon Park

Table of Contents

I. Introduction

A. Background and Qualifications

1. My educational background, career history, and other relevant qualifications are summarized below. I attach to this Declaration my *curriculum vitae* (Appendix B) which provides a full and accurate description of my educational background, professional experience, and qualifications.

2. I received my Ph.D. in Biophysics from Harvard University in 2000. I completed my postdoctoral training at the University of Pennsylvania. I also received an M.S. in Physics from the Massachusetts Institute of Technology ("MIT") in 1994. I received a B.S. in Physics and Math from the University of California, Berkeley, in 1991.

3. I currently serve as an associate professor in the Department of Chemical and Biological Engineering and previously served as Director of Graduate Studies for the department. I am also affiliated with the University's Genetics, Genomics and Bioinformatics graduate program. I have been a professor at the University of Buffalo since completing my postdoctoral research position at the University of Pennsylvania.

4. During my career I have taught courses in, among other things, Protein Engineering, and Biotechnology Principles for Chemical Engineers. I began teaching these courses in 2007.

5. I have nearly two decades of experience in the field of protein engineering. In my lab, we use computational and experimental tools to characterize and design protein molecules with novel physical and biological properties. For example, my lab engineered the first functional monomeric streptavidin that is used in *in vivo* imaging and biomolecular detection. My lab has also developed and engineered many proteins based on rational design and directed evolution techniques.

6. In 2011, I was named a recipient of a CAREER award from the National Science Foundation, which is an award given to early-career faculty who have the potential to serve as academic role models in research and education and to lead advances in the mission of their department or organization.

7. I have authored 30 peer-reviewed publications, many of which are directed to topics in proteins, protein characterization, and protein design and engineering. Along with Jennifer Cochran, I was an editor of the textbook "Protein Engineering and Design," which was published in 2009 by CRC Press.

B. Compensation

8. I am being compensated for my time at the rate of $350 per hour for my work in connection with this matter. I am being reimbursed for reasonable and customary expenses associated with my work in this investigation. This compensation is not dependent in any way on the contents of this Declaration, the

substance of any further opinions or testimony that I may provide, or the ultimate outcome of this matter.

C. Information Considered

9. My opinions are based on my years of education, research, and experience, as well as my investigation and study of relevant materials. I also have relied upon the materials listed in Appendix A in forming these opinions.

D. Person of Ordinary Skill in the Art

10. I understand that my analysis and opinions are to be provided using the perspective of a person of ordinary skill in the art in the December 2011 time frame.

11. I have been informed that a "person of ordinary skill" is a hypothetical person who has certain educational qualifications and experience, and possesses an ordinary level of insights and skill.

12. Counsel for Merck provided the following description of a person of ordinary skill in the art in the 2011 time frame for me to evaluate:

> A person of ordinary skill in the art in the 2011 time-frame would have had an undergraduate degree, a Ph.D., and post-doctoral experience in scientific fields relevant to study of protein structure and function (*e.g.*, chemistry, biochemistry, biology, biophysics). From training and experience, the person would have been familiar with factors influencing protein structure, folding and activity, production of

> modified proteins using recombinant DNA techniques, and use of biological assays to characterize protein function, as well with techniques and tools used to analyze protein structure-structure relationship (*i.e.*, sequence searching and alignments, protein modeling software, etc.).

13. I believe the description of a person of ordinary skill in ¶ 12 is accurate, and that I had those qualifications by December of 2011.

14. In preparing this declaration, I have used the perspective of a person of ordinary skill in the art in the 2011 time frame as it is described in ¶ 12.

II. Scope of My Declaration

15. I was asked if a person of ordinary skill in the art in 2011 would have been able to identify the single amino acid substitutions within non-essential regions of $PH20_{1-447}$ that would be tolerated by the protein (*i.e.*, would not cause a substantial loss of hyaluronidase activity) without making and testing each possible single-substituted $PH20_{1-447}$ protein.

16. As I demonstrate below, a skilled person, using sequence analysis and protein modeling techniques known in 2011, would have readily identified many specific amino acid substitutions, some of which are not necessarily conservative substitutions, that would be tolerated by the $PH20_{1-447}$ protein structure and would therefore enable the protein to retain its hyaluronidase activity. I base this conclusion on an analysis I performed, which is described in § IV.C.

17. I identified single amino acid substitutions that would be tolerated at position 324 in $PH20_{1-447}$, including E324D, E324N, E324R, E324A, E324H, and E324S. I summarize this analysis in § V.

18. Finally, I was asked to determine the number of distinct polypeptides that share 91% or 95% sequence identity with human PH20 sequences of varying length, assuming certain conditions. I provide my calculations below in § VII.

III. Terminology Used in this Declaration

19. I will use the following abbreviations in this declaration:

(a) I use "PH20" to refer to the human PH20 protein. The full-length sequence of the human PH20 protein has 509 amino acids and was first published in 1993.[1] The sequence is reported as SEQ ID NO:1 in U.S. Patent No. 7,767,429 (the '429 Patent), which is identical to the sequence deposited with the Universal Protein Resource website ("www.uniprot.org") with ID P38567.

(b) The full-length human wild-type PH20 sequence includes a 35 amino acid signal sequence (positions 1-35 of Uniprot ID: P38567). The mature amino acid sequence PH20 is found at

[1] EX1029 (Gmachl), 546, Fig. 1.

positions 36-509 of Uniprot ID: P38567, and is 1-474 in sequences that omit the signal sequence.

(c) "$PH20_{1\text{-}n}$" refers to the human wild-type PH20 polypeptide sequence starting at position 1 and terminating at position "n" of the mature protein (*i.e.*, lacking the signal sequence). For example, $PH20_{1\text{-}447}$ means the polypeptide starting at position 1 and ending at position 447 of the mature human wild-type PH20 sequence. The corresponding positions in the human wild-type PH20 sequence having the signal sequence are 36-482 (*e.g.*, SEQ ID NO: 1 of U.S. 7,767,429, Uniprot: P38567).

(d) "AxxxB" refers to an amino acid substitution at position xxx, where the wild-type residue is A and the residue after the substitution is B.

(e) The use of a position number in this declaration is referring to the position in the mature form of human PH20 (*i.e.*, omitting the 1-35 amino acid signal sequence). I also will occasionally use both the full-length and the mature numbers.

IV. Overview of Methodologies Used for Identification of Tolerated Amino Acid Substitutions in $PH20_{1\text{-}447}$

20. Proteins often can tolerate a single amino acid substitution in non-essential regions of the protein's structure. I use tolerate here to mean that the

presence of a single amino acid at a particular position of the protein's amino acid sequence that is different than the naturally occurring ("wild-type") amino acid at that position of the protein does not materially alter the local structure around that position in the protein and thus does not meaningfully alter the biological activity of the protein. Of course, there are exceptions to this general point, and any particular substitution will need to be assessed to determine if that particular substitution will be tolerated.

21. The concept of tolerance can be readily appreciated by comparing evolutionarily related proteins that share structure homology. While there are amino acids at many positions within a set of evolutionarily related proteins that are conserved (*i.e.*, amino acids that do not vary or vary only rarely), the amino acids at many other positions can vary extensively.[2] The variability in the amino acids at these non-conserved positions indicates that the protein structure(s) is not

[2] EX1014 (Brandon), 351 ("The underlying assumption is that secondary and tertiary structure has been more conserved during evolution than amino acid sequence; in other words only such changes have been retained during evolution that conserve the structure. Consequently, the pattern of residue changes within homologous proteins contains specific information about the structure.").

dependent on the exact identity of the amino acid at those positions and is able to accommodate different amino acids without causing a loss of protein function.

22. To identify single amino acid substitutions that would be tolerated in human $PH20_{1-447}$, I used procedures that were widely used in rational design methods of protein engineering before December 2011. Generally, these include the following steps:

(a) identify homologous sequences based on sequence comparison;

(b) perform a multiple sequence alignment of retrieved sequences;

(c) identify the frequency of occurrence ("profile") of amino acids at each position of the protein across the set of sequences;

(d) generate a protein model using SWISS-MODEL;

(e) use the model to visually assess individual substitutions within the local environment of the protein at a particular position.[3]

23. The techniques I describe above of finding homologous sequences, aligning them, and using the sequence identity information to identify single amino

[3] As I describe in further detail, below, the model would have been reliable for evaluating certain single amino acid substitutions (depending on the position of the substitution), but not for modeling multiple concurrent substitutions, which would quickly make the model unreliable. *See* ¶¶ 160-162.

acid substitutions that would be tolerated were well-known and widely used by skilled artisans in 2011.[4]

24. I believe a person of ordinary skill in the art in 2011 would have been familiar with the tools and techniques discussed in Green (EX1017).[5] These include: (i) BLAST, for performing protein sequence searching, (ii) sequence alignment tools like CLUSTAL (*e.g.*, CLUSTAL-Omega), (iii) the protein structure modeling tool SWISS-MODEL, and (iv) software to view protein structures, like PyMol. I discuss these tools further in § VI, below.

A. Identification of Non-Essential Residues From Homologous Sequences

25. In general, amino acids that do not vary at the same aligned positions within a set of structurally-related proteins are essential to the structure and functions of the protein.[6] When a residue does not vary in so many naturally

---

[4] EX1017 (Green), 223-230, 236 (discussing rational design techniques); EX1016 (Steipe), 181-186.

[5] These tools and techniques are also discussed in Steipe (EX1016).

[6] EX1017 (Green), 224 ("By considering the common features of the sequences of these proteins, it is possible to deduce the key elements that determine protein structure and function…").

occurring proteins, it is strong evidence that it is essential to the protein's structure and function and needs to be preserved.[7]

26. I investigated essential residues in hyaluronidase enzymes in two ways. First, I identified invariant residues by analyzing a multiple sequence alignment of a set of published hyaluronidase sequences that was available in December of 2011. Second, I reviewed scientific literature that identified important residues in hyaluronidase proteins or which reported experimental results showing that modifying single residues impaired or eliminated activity of the enzymes.[8] These two overlapped significantly and many residues that were experimentally shown to be essential for activity proved to be highly conserved based on sequence alignment. Below I address the first investigation.

1. BLAST Results and the Multiple Sequence Alignment

27. I generated a dataset of sequences that were homologous to PH20 and that were publicly available by the end of 2011. To do that I performed a BLAST

[7] EX1017 (Green), 224 ("Evolution provides a tremendously useful model for protein design.").

[8] I reviewed the scientific literature in detail below. *See* § IV.C.3.

search and culled the results to yield a set of 88 nonredundant homologous sequences that were publicly available by 2011.[9]

28. I then aligned the 88 homologous sequences in a multiple sequence alignment ("MSA").[10] The homologous sequences "must be aligned such that conserved positions are in register with one another."[11] Since the sequences are all different, the MSA attempts to align the sequences, but this can result in gaps in the MSA. MSA algorithms seek to "optimize a global score across an alignment," which is "typically based on sequence similarity alone and does not take structural or functional information into account."[12]

---

[9] I explain this process in further detail below. *See* § VI.A.

[10] EX1017 (Green), 224 ("One of the most straightforward applications of primary sequence data in protein engineering is the use of multiple-sequence alignments to define consensus motifs for a particular structure or function."); EX1016 (Steipe), 184 ("multiple sequence alignments are required").

[11] EX1017 (Green), 224.

[12] EX1016 (Steipe), 184.

29. I used the Clustal Omega tool to produce an MSA of the 88 homologous sequences.[13] The Clustal Omega program computes and reports whether a residue is "conserved" or "semi-conserved," which takes into account how similar a residue is across all of the sequences. A "conserved residue" is one which has the same amino acid in all of the sequences. A "semi-conserved" residue is one where all amino acids in that position have similar properties. The multiple sequence alignment I produced is in Exhibit 1058.[14] Each row shows 60 positions of the aligned sequences. Below each set of 60 positions is a consensus report using 70%, 80%, 90% and 100% identity thresholds for consensus residues at that position (*i.e.*, the amino acid that appears in 70%, 80%, 90% or 100% of the aligned sequences, if any). The consensus report gives a quick overview of where

[13] *See, e.g.*, EX1043 (Sievers), 1-2, 4-5. I explain this process in further detail below. *See* § VI.B.

[14] The MSA depicted in Exhibit 1058 was produced by the MSA viewer available from the European Bioinformatics Institute of the European Molecular Biology Laboratory organization (EMBL-EBI) (https://www.ebi.ac.uk/jdispatcher/msa/mview?stype=protein). The text-based output of the CLUSTAL-Omega MSA is also provided in Exhibit 1057.

conserved residues appear. I performed a more detailed analysis of the MSA using custom scripts I prepared, as described below.

2. Essential Residues in PH20

30. Using the alignment, I identified 68 largely invariant residues that a skilled artisan would have deemed "essential residues" in $PH20_{1-447}$ based on the sequence alignment (table below).[15] These are positions where the amino acid was conserved 95% of the time or more and non-identical amino acids appear in less than ~5% of the proteins in the data set.[16]

[15] The essential residues are also listed in Appendix D-3 (EX1004, 171).

[16] The frequency at which an amino acid appears in the MSA is compiled in Appendix D-1 (EX1004, 133).

Residue #	Mature Residue #	PH20 Residue	Residue %		Residue #	Mature Residue #	PH20 Residue	Residue %
49	14	F	95.5		222	187	P	100
53	18	W	100		224	189	C	100
60	25	C	100		226	191	N	98.9
91	56	F	98.9		234	199	Y	98.9
92	57	Y	98.9		236	201	G	98.9
97	62	G	100		238	203	C	100
99	64	Y	97.7		246	211	N	100
100	65	P	100		249	214	L	98.9
112	77	G	98.9		251	216	W	98.9
113	78	G	100		253	218	W	98.9
115	80	P	100		256	221	S	100
116	81	Q	100		258	223	A	97.7
123	88	H	98.9		259	224	L	100
141	106	G	98.9		261	226	P	98.9
144	109	V	95.5		281	246	R	97.7
146	111	D	98.9		284	249	E	100
147	112	W	100		287	252	R	98.9
148	113	E	100		299	264	P	98.9
150	115	W	100		316	281	L	98.9
152	117	P	100		321	286	L	97.7
154	119	W	97.7		326	291	G	98.9
157	122	N	98.9		332	297	G	98.9
158	123	W	96.6		335	300	G	98.9
164	129	Y	100		339	304	W	98.9
168	133	S	100		351	316	C	100
188	153	A	100		362	327	L	96.6
192	157	F	100		368	333	N	97.7
203	168	T	98.9		369	334	V	96.6
211	176	R	95.5		376	341	C	100
212	177	P	97.7		377	342	S	96.6
215	180	L	95.5		381	346	C	100
216	181	W	100		385	350	G	100
217	182	G	100		387	352	C	100
219	184	Y	100		398	363	L	96.6

3. Non-Essential Residues in PH20

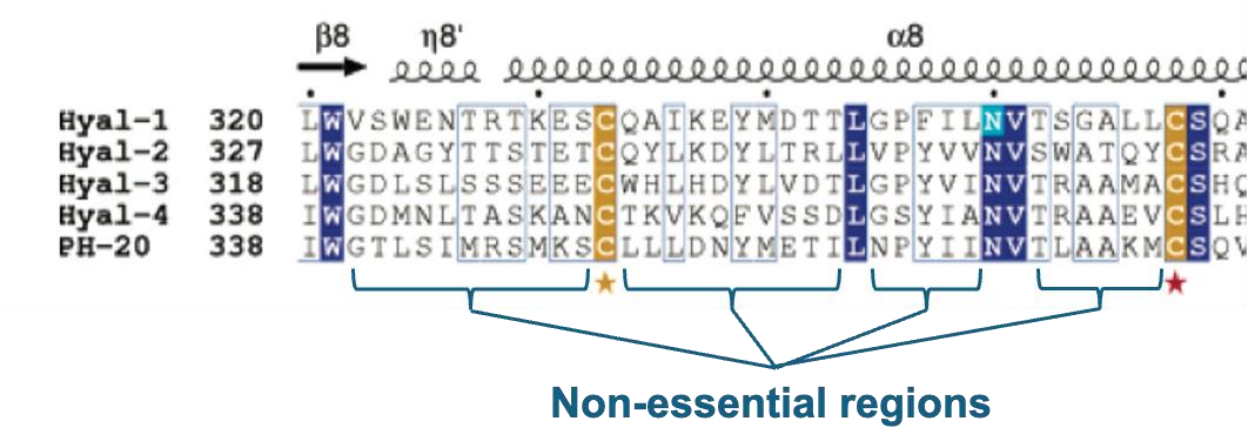
31. The 379 positions in $PH20_{1-447}$ other than the 68 essential residues are positions at which there is evolutionary variation among the 88 protein sequences that I analyzed. The existence of evolutionary variation at these positions indicates that the homologous proteins have tolerated different amino acids at those positions. Amino acids at these positions in the different hyaluronidase proteins would be considered "non-essential" residues because the proteins presumably still exhibit hyaluronidase activity when they are mutated. The 379 non-essential positions in $PH20_{1-447}$ are compiled in Appendix D-2.

32. I also reviewed observations in the '429 Patent concerning making modifications to PH20 proteins. As it explains:

> Suitable conservative substitutions of amino acids are known to those of skill in this art and can be made generally without altering the biological activity, for example enzymatic activity, of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity …[17]

[17] EX1005 ('429 Patent), 16:14-22; *also* 9:47-49.

I believe the positions that I identified in Appendix D-2, including position 324, align with what I consider to be the "non-essential regions" referred to by the '429 Patent. An illustration of some non-essential regions in PH20 is provided below.[18]



B. Modeling the PH20 Structure for Visual Inspection

1. Visualization of the Protein Structure

33. In 2011, to assess whether a potential amino acid substitution at a particular position in a protein would be tolerated, a person of ordinary skill would have visualized the mutation by building a structural model of the protein. Visualization allows one to assess the interactions between the wild-type amino acid being changed and its neighboring amino acids at that position, and thereby compare the interactions of different amino acids at that position with neighboring

[18] I annotated Figure 3 in Chao (EX1006) to illustrate some of the non-essential regions I identified in my analysis.

amino acids. Comparing the set of interactions seen with the wild-type residue to those observed for the substituted amino acid at that position will allow prediction of whether the substitution will be favorable, neutral, or unfavorable to the protein's structure.

34. Assessing amino acid substitutions by visual inspection of the site of the substitution in a protein structural model was a commonly used technique in 2011. For example, as one group reported:

> Many protein-engineering applications involve the creation of a small number of mutations to a naturally occurring protein so as to enhance its function in a well-defined manner. In these cases, a structural biologist's intuition is often an important tool in the design of the desired variants, an approach that may be termed structure-based protein design to borrow a term from the drug design field. Visualization of the known reference structure is a key component of this. For example, visualization can identify unsatisfied hydrogen bond donors or acceptors that may be mutated to increase stability or affinity. Similarly, visualizing steric interactions can help engineer interactions to discriminate among several potential binding targets.[19]

[19] EX1017 (Green), 228-29.

35. Another example is work done by a group led by Dr. Moult at UMBI. His group performed visual inspections of protein models to assess the impact of single amino acid substitutions in proteins that were caused by single nucleotide polymorphisms. The results of those visual inspections were published in peer review journals.[20] I view that as a validation of this technique of visually assessing the effects of an amino acid substitution that I used in my analysis here.

36. The local environment of the protein where an amino acid substitution is being made can be visualized using a structural model of the protein. In December of 2011 (and even today), the structure of human PH20 was not solved. The absence of an experimentally determined structure for a protein does not preclude the use of structure-based modeling methods.[21] That is because the

[20] EX1031 (Yue), 459 (reporting effects of single amino acid substitutions "relied primarily on visual inspection of an amino acid substitution on protein structure and function."); EX1032 (Wang), 265-266 (analyzing homology models for single amino acid substitutions to study mechanisms of hereditary disease).

[21] EX1017 (Green), 229 ("In many cases, protein engineering targets a protein whose structure has not been solved"), ("This does not preclude the use of structure-based methods, as the known structures of related proteins can be used to create model structures through the process of homology modeling").

structure of a protein can be modeled in a process called "homology modeling."[22] Homology modeling was routinely used by 2011 and provided an "accurate computational method to generate reliable structural models."[23]

2. Homology Models for Proteins with Unknown Structures

37. Generally, a protein homology model is built using an experimentally determined reference structure with a highly homologous sequence.[24] That is

---

[22] EX1017 (Green), 229-230 ("This does not preclude the use of structure-based methods, as the known structures of related proteins can be used to create model structures through the process of homology modeling"), ("Homology building can allow a model structure to be built from the structure of a related sequence").

[23] EX1012 (Bordoli), 1 ("Homology modeling is currently the most accurate computational method to generate reliable structural models and is routinely used in many biological applications"), (Homology modeling is "the method of choice to build reliable" models); EX1014 (Brandon), 348 ("This model can serve as an excellent basis for identifying amino acid residues involved in the active site…").

[24] EX1017 (Green), 229; EX1012 (Bordoli), 1 ("Homology modeling aims to build three-dimensional protein structure models using experimentally

possible because "[h]omologous proteins have similar three-dimensional structures."[25] To build a homology model, the "backbone of homologous residues from a protein of unknown structure" is mapped "onto a known structure."[26] In many cases, however, "there are regions of nonhomologous sequence, even in highly homologous proteins" in which a "new backbone…must be constructed for these regions."[27]

38. Numerous homology modeling programs were available by 2011, including the SWISS-MODEL program.[28] This program searches a library of experimentally determined protein structures to identify suitable templates to use

---

determined structures of related family members as templates"); EX1014 (Brandon), 348 ("If significant amino acid sequence identity is found with a protein of known crystal structure, a three-dimensional model of the novel protein can be constructed, using computer modeling, on the basis of the sequence alignment and the known three-dimensional structure").

[25] EX1014 (Brandon), 370.

[26] EX1017 (Green), 229.

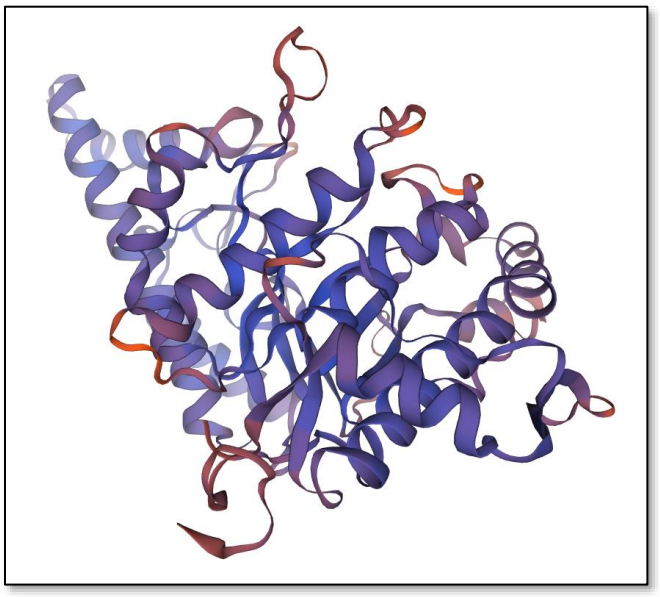
[27] EX1017 (Green), 229.

[28] EX1017 (Green), 229, Table 10.2; *also* EX1012 (Bordoli), 2. I explain this in further detail below. *See* § VI.C.

for the protein of interest, and then generates a model for that protein based on the sequence alignment of the target protein and the template structure.[29]

3. Generation of the Model of PH20 Structure

39. I used the SWISS-MODEL program to generate a model (shown below) of the PH20 structure.[30]



[29] EX1012 (Bordoli), 1; EX1066 (SWISS-MODEL), 3 ("Building a homology model comprises four main steps: identification of structural template(s), alignment of target sequence and template structure(s), model building, and model quality evaluation").

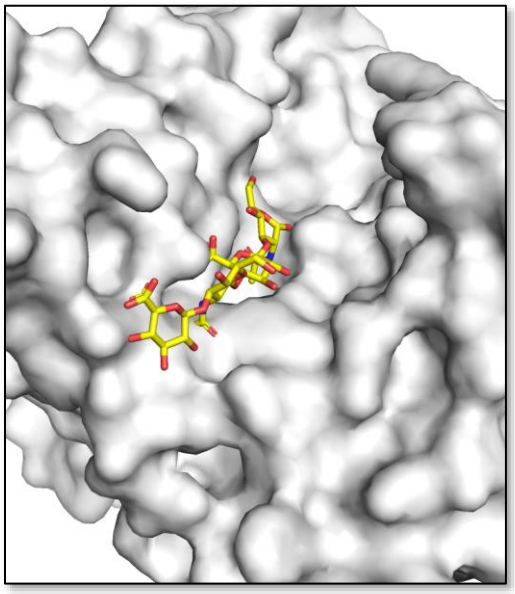
[30] I explain this process and the generation of the displayed model in further detail below. *See* § D.C.

40. The PH20 homology model does not include its carbohydrate ligand (hyaluronic acid) because the HYAL1 structure is not based on a complex of that protein with its ligand. Another structure provided an assessment of the ligand interaction, however, which was the hyaluronidase protein in bee venom (ID: 1FCV).[31] The bee venom hyaluronidase is also highly homologous to human PH20 and HYAL1, sharing about 30% sequence identity.[32] Therefore, using PyMol, I also mapped the bee venom hyaluronidase's ligand interaction to the human PH20 model to get a sense of the residues in human PH20 that are likely to

[31] EX1033 (Markovic-Housley), 1028-1029, 1035 (determined the x-ray structure of hyaluronidase from bee venom in complex with an HA oligomer); *see also* EX1006 (Chao), 6912 (citing Markovic-Housley to support that "[t]he N-terminal 325 residues [of HYAL1] exhibit 31% sequence identity with bee venom hyaluronidase (bvHyal), whose structure has been determined in complex with a HA tetrasaccharide").

[32] EX1006 (Chao), 6912 (reporting that "[t]he N-terminal 325 residues [of HYAL1] exhibit 31% sequence identity with bee venom hyaluronidase (bvHyal), whose structure has been determined in complex with a HA tetrasaccharide").

interact with the ligand.[33] An illustration of the PH20 structural model with the ligand is shown below.



C. Evaluation of Substitutions at Non-Essential Residues

41. Generally, a skilled artisan would expect that amino acids found in non-essential positions in naturally occurring, homologous hyaluronidase proteins would be tolerated as single amino acid substitutions at the corresponding position in $PH20_{1-447}$.[34] That is because those amino acids are tolerated in the

[33] I explain this process in further detail below. *See* § D.D.

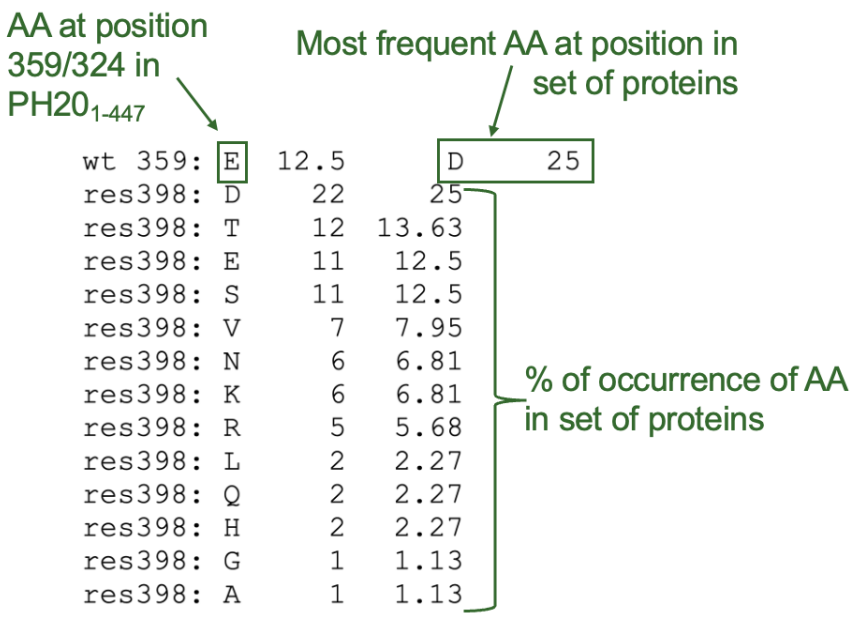
[34] EX1005 ('429 Patent), 9:46-52 ("Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity []"), 16:14-20.

corresponding position in homologous proteins, implying that they are compatible with the hyaluronidase structure and function.

42. Within the 88 homologous hyaluronidase proteins that I analyzed, there were varying numbers of amino acids that occurred at each non-essential position.[35] A higher frequency for one amino acid within a set of amino acids at a particular position indicates that a larger number of naturally occurring hyaluronidase proteins contain that amino acid at that position. For example, a 5% frequency of occurrence of an amino acid in a set of 88 related proteins means that ~4 naturally occurring hyaluronidase proteins have that amino acid at that position. In essence, evolution has "sampled" those substitutions and found they are tolerated in at least 4 isoforms of PH20.

43. An example of a set of amino acids from the MSA based on 88 sequences for position 324 (359 with the signal sequence) of $PH20_{1-447}$ is provided below.

[35] Appendices D-1 and D-2.



44. I evaluated several single amino acid substitutions in non-essential positions in $PH20_{1-447}$ using the PH20 structure model I prepared.[36] For positions that I evaluated, I assessed possible interactions between the wild-type residue at that position and its neighboring amino acids.[37]

45. An amino acid at a particular position in a protein may interact with solvent and/or with neighboring amino acids in the protein. The set of those interactions involving that residue may contribute positively, negatively, or may be

[36] As I explained in § IV.A.3, I refer to the "non-essential residues" as the residues which were conserved in less than 95% of the sequences in the MSA.

[37] I consider neighboring amino acids to be those which are within 5 Å of the side chain of the residue of interest. *See* ¶ 56.

inconsequential to the stability of the protein structure at the location of the residue.

46. There are a range of possible interactions that can occur between a target residue and its neighboring amino acids depending on the size and the chemical attributes of the amino acids involved. Amino acids may be classified as: (i) hydrophobic, (ii) charged (ionic), or (iii) polar (hydrophilic but without a permanent charge and capable of hydrogen bonding).[38] Amino acid side chains also vary in size and rigidity, ranging from the smallest (*i.e.*, glycine, no side chain) to large side chains (*e.g.*, phenylalanine, tryptophan). Both factors, *i.e.*, hydrophobicity and sterics, contribute to the nature of possible interactions that an amino acid may have with its neighbors in a protein structure.

47. The varying characteristics of side chains gives rise to a large number of possible types of interactions. For example, two charged side chains can have a favorable interaction if the charges are opposed (+/-) or an unfavorable interaction if the charges are the same (+/+). Hydrophobic residues likewise can have favorable interactions with other hydrophobic residues, and unfavorable interactions with hydrophilic residues. Still, other forces can be observed based on how close the residues are to each other. For example, van der Waals attractions

[38] EX1014 (Brandon), 4-5.

occur between residues that are within a few angstroms of each other, but if the atoms get too close, there are strong repulsions. Likewise, polar residues can exhibit hydrogen bonding interactions, again depending on how far apart the involved residues are and whether the geometry is permissible. An illustration of some of these interactions is provided below:[39]

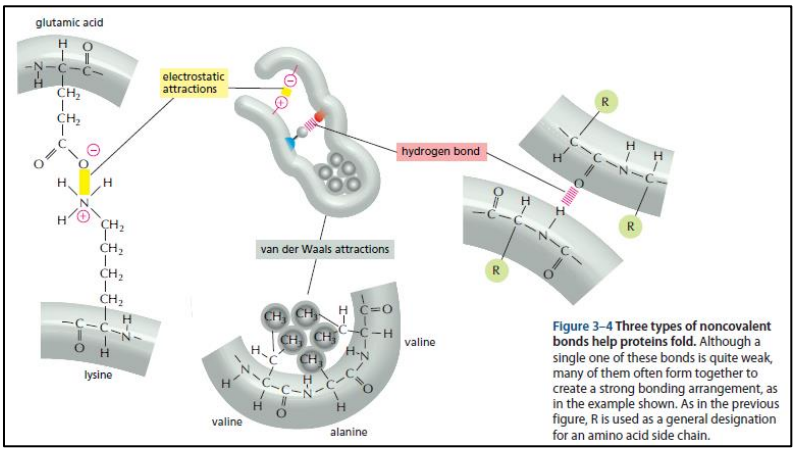


Figure 3–4 Three types of noncovalent bonds help proteins fold. Although a single one of these bonds is quite weak, many of them often form together to create a strong bonding arrangement, as in the example shown. As in the previous figure, R is used as a general designation for an amino acid side chain.

1. Published Assessments of Single Amino Acid Substitutions

48. Two papers from Moult's lab in the early 2000s illustrate how assessments of a residue's interactions with its neighbors may be carried out to assess single amino acid substitutions in proteins. The first (Wang) provided a survey of types of interactions that would suggest a substitution would be destabilizing to the protein's structure.[40] These included:

[39] EX1039 (Alberts), 130, Fig. 3-4.

[40] EX1032 (Wang), 266.

1) Loss of one or more hydrogen bonds: A hydrogen bond is defined as a donor to acceptor distance of ≤2.5 Ångström ("Å") and an angle at the acceptor ≥90.0°.

2) Reduced hydrophobic interaction: Loss of burial of 50 $Å^2$ or more of non-polar area on folding.

3) Loss of a salt bridge: A salt bridge is defined as at least one pair of atoms on oppositely charged groups within 4.5 Å.

4) Buried charged residue: Introduction of a zero-accessibility electrostatically isolated charge.

5) Over-packing: Introduction of a larger side chain, making unavoidable short atomic contacts (<2.5 Å in the interior of a protein or <2.0 Å on the surface).

6) Internal cavity: Replacement of one or more zero-solvent accessibility groups with a smaller side chain.

7) Electrostatic repulsion: Introduction of a charged group, such that at least one pair of atoms on like-charged groups are unavoidably within 4.5 Å.

8) Buried polar residue: One or more polar groups unavoidably have zero solvent accessibility and no hydrogen bonds.

9) Disruption of metal binding: A metal liganding group is replaced by a non-liganding group.

10) Breakage of a disulfide bond: A cysteine residue in an S-S bond is replaced by a non-cysteine residue.

11) Backbone strain: Gly is changed to any other residue, with phi/psi values not in the Ramachandran general sterically allowed region, or any other residue is changed to Pro, where phi is unfavorable (permitted phi for Pro = –60 ± 15°); or a cis Pro (omega = 0 ± 60°) is changed to any other residue.

12) Destabilization of a protein multimer: Any of the above rules, involving atoms on a neighboring subunit.

49. The second paper in 2005 from Moult's lab (Yue) also summarized the types of destabilizing interactions that can occur between amino acids, and how they are assessed when evaluating the effects of a single substitution in a protein.[41] The Yue paper listed 11 types of interactions:

[41] EX1031 (Yue). Yue reported on an analysis of human genes containing a single nucleotide polymorphism, which cause production of a human protein having a single amino acid change relative to the wild-type form of the protein.

> Eleven contributions to the energy and entropy of protein stability are considered. There are four classes of electrostatic interaction: reduction of charge–charge, charge–polar or polar–polar energy, or introduction of electrostatic repulsion; three solvation effects: burying of charge or polar groups, and reduction in non-polar area buried on folding; and two terms representing steric strain: backbone strain and overpacking. The other two contributions considered are cavity formation (affecting van der Waals energy), and loss of a disulfide bridge.[42]

50. The Yue paper then came up with 15 factors based on these known types of interactions that can affect protein stability, which also considered the location of the interaction in the protein.[43]

Table 3. The 15 factors included in the model of protein stability

Type	Factors
Continuous factors	Electrostatic interaction: polar–polar, polar–charge, charge–charge
	Over-packing
	Hydrophobic burial
	Surface accessibility
	Structural rigidity: crystallographic *B*-factor, *Z* score and standard deviation
Binary factors	Cavity
	Electrostatic repulsion
	Backbone strain
	Buried charge
	Buried polar
	Breakage of a disulfide bond

The effect of each single residue mutant on stability is expressed in terms of the value of one or more of these contributions to the energy and entropy. Continuous factors are represented by a continuous variable, binary factors are two state, either significantly or not significantly affecting stability.

[42] EX1031 (Yue), 462; *also* EX1031 (Yue), 470-471.

[43] EX1031 (Yue), 469, Table 3.

51. The Yue paper also used a classification scale for the effects of a substitution. The first two classifications were that the changes "stabilize or mildly destabilize the folded state" of the protein.[44] In other words, the substitution made the protein more stable, or only mildly destabilized the protein structure. These first two categories were differentiated from substitutions that more substantially destabilized the protein structure.[45] The Yue paper also assessed the severity of the destabilizing changes by assessing changes in free energy, explaining that changes "in the 3-4 kcal/mol range" or higher correlated to more the substantially destabilized proteins, which were those with "disease causing mutations."[46]

52. The Yue paper also pointed out that there was a better correlation between substitutions and destabilizing changes when the nature of an interaction was combined with its location. For example, proteins were more consistently destabilized by a single amino acid change that introduced a large side chain that caused a steric clash within a buried position of the protein, or which introduced electrostatic repulsion within an overpacked region of the protein.[47]

---

[44] EX1031 (Yue), 464.

[45] EX1031 (Yue), 464.

[46] EX1031 (Yue), 464.

[47] EX1031 (Yue), 462-463, Figure 2.

2. My Assessment of Factors Influencing Single Amino Acid Substitutions

53. In my assessment of single amino acid substitutions that I address in this declaration, I considered factors within the context of the specific location of the substitution. The location of the amino acid being changed within the protein structure is an important consideration in assessing whether a single amino acid substitution will be tolerated or not.

54. Generally, residues on the surface of the protein's structure, where they are more accessible to solvent, have more freedom of movement and thus tend to tolerate a single amino acid substitution with a greater number of alternative amino acids. That is because the original and new amino acid are both primarily interacting with solvent, and interactions with other parts of the protein aren't nearly as important. By contrast, amino acids that are buried within the protein's structure tend to be more sensitive to (and tolerate fewer) substitutions, because the new and original amino acids are interacting with other residues in the protein and there are more constraints on the shape and biochemical properties of the residue. Changing the amino acid at that location can disrupt interactions that are important to protein structure.

55. I used these established principles to define a classification system for assessing single amino acid substitutions in PH20. Similar to the Yue paper, I used three classifications: (i) changes that were likely to stabilize the protein, (ii)

changes that were either neutral or mildly positive or negative for the local protein structure, and (iii) changes that were likely to be significantly destabilizing.

56. I first inspected the PH20 structure I produced using the PyMol viewer to identify the neighbors of the residues that I assessed in the wild-type PH20 sequence. I noted "neighbors" that were within ~ 5 Å of the side chain of the amino acid being evaluated. A distance of 5 Å is an appropriate cutoff for neighbors, as most direct interactions of consequence between residues will occur within that distance.[48]

57. A greater number of neighbors at a position (*e.g.*, ~ 10 or higher) is generally indicative of the residue being buried, while fewer neighbors (*e.g.*, less than ~5) is generally indicative of a solvent accessible surface residue. The neighbor count alone, however, does not determine if a residue is solvent accessible or inaccessible. Inspection of the residue within the PH20 model helps to identify a residue as being a surface residue or a buried residue, although this classification is qualitative and is used more for discussion than for strict classification purposes.

[48] For example, Mihel et al. 2008 uses 4.7 – 6 Å as cutoff to define interaction. *See* EX1043 (Mihel), 2 (Table 1).

58. Fractional solvent accessible surface area ("SASA") is a quantitative measurement which can be used to determine whether a residue is solvent accessible. Fractional SASA is computed by calculating the SASA of each amino acid and dividing it by the maximum SASA achievable for the same amino acid.[49] Although I prefer to analyze solvent accessibility by visualizing each amino acid in the structure, I have also considered the fractional SASA for amino acids in Appendix D-5 which I can use to assess how buried or solvent accessible particular residues are. For example, the information may tell me that the amino acid is more or less solvent exposed at a certain location in the structure than it is on average. For several substitutions, I also computed the fractional SASA for the substituted residues in the mutated PH20 structure.

59. I also inspected the distance between neighbors and a particular residue to determine whether there was likely a van der Waals interaction, which

[49] EX1035 (Lins), 1408, Table 2 (reporting "[m]edian values of total hydrophobic (pho), hydrophilic (phi) accessible surface (ASA) of whole residues from folded proteins").

generally occurs when the residues are within 3-4.5 Å (with 3 Å being strong van der Waals interaction).[50]

60. I also periodically used a custom script that I wrote which runs within the PyMol environment. My python script shows all of the neighboring amino acids encompassed in a shell, allowing one to visualize the chemical moieties surrounding an amino acid at a given position.[51] By visualizing the neighbors as a surface, it can be determined whether the residue of interest is buried (with neighbors on all sides) or solvent exposed (missing neighbors on some sides). This feature is especially useful when considering the hydrophobicity factor for each position, as described below.

61. PyMol also includes a "mutagenesis" feature that replaces an amino acid at a defined position with another amino acid.[52] For several substitutions, I used this feature to evaluate whether a mutation would likely be tolerated. It was especially helpful when evaluating how a mutation may disrupt or introduce interactions with other residues.

---

[50] This is also displayed in the GUI window along with the number of neighbors. I explain this process in further detail below. *See* § VI.D.

[51] I describe this script in further detail below. *See* § VI.D.

[52] I describe this function in further detail below. *See* § VI.D.

62. In my assessments whether a mutation would likely be tolerated, I also considered the chemical similarity of the substitution to the wild-type amino acid at the position being evaluated. As I illustrate below, chemical similarity alone often is not sufficient to predict whether a particular mutation would be tolerated or not.

63. Certain types of substitutions create interactions with neighboring residues that are more impactful than others. Examples include: (i) changes that introduce a hydrophobic residue into a hydrophilic environment (or vice versa), (ii) changes that affect secondary structures in the protein, and (iii) changes that create steric clashes due to increased size of the amino acid side chain, or which create holes within the protein structure due to a smaller side chain.

a. **Hydrophobicity**

64. Hydrophobic amino acids are generally nonpolar and tend to repel water. Hydrophilic amino acids are generally polar and mix well with water. Amino acids are classified as hydrophobic or hydrophilic based on their side chains (summarized in the table below):

Hydrophobic	Hydrophilic
Alanine (A)	Aspartic acid (D)
Cysteine (C)	Glutamic acid (E)
Phenylalanine (F)	Histidine (H)
Glycine (G)	Lysine (K)
Isoleucine (I)	Asparagine (N)
Leucine (L)	Glutamine (Q)
Methionine (M)	Arginine (R)
Proline (P)	Serine (S)
Valine (V)	Threonine (T)
Tryptophan (W)	
Tyrosine (Y)	

65. An amino acid substitution that "matches" the hydrophobicity of the environment of the wild-type residue will generally be tolerated by the protein, assuming other structural constraints are satisfied. For example, a hydrophobic amino acid, such as leucine (L), in a hydrophobic environment often can be substituted with valine (V), another hydrophobic amino acid.

66. Many environments have mixed hydrophobic and hydrophilic characteristics. Introducing a mutation that alters the hydrophobicity of the residue in a mixed environment may cause more modest changes in interaction.

67. Positions in a protein can be solvent exposed even if the environment is overall hydrophobic. An amino acid in such an environment may be able to interact with solvent (water) based on its characteristics. For example, a residue like lysine may be long enough to cause its hydrophilic portion (the NH3 group at the end of the side chain) to extend past a hydrophobic pocket and interact with water. Likewise a position that is solvent exposed may also be hydrophobic if it is

surrounded by hydrophobic amino acids or the hydrophobic groups of polar amino acids. For example, the side chains of lysine, proline, asparagine may create a localized hydrophobic pocket that may be occupied by a nonpolar amino acid, such as leucine.

68. When the wildtype amino acid does not match the hydrophobicity of its environment, mutations that would correct the mismatch would be favored substitutions and may improve the protein's properties (*e.g.*, stability, activity, etc.). As an example, increasing the number of hydrophobic residues in a buried position, such as the protein's interior core, may result in a more thermostable protein.[53]

b. **Secondary Structure**

69. For each substitution I considered, I also evaluated its compatibility with the predicted secondary structure at that position. Amino acids have varying propensities for influencing a secondary structure. In my evaluation of substitutions, I considered whether the new amino acid would reduce or improve the propensity for the secondary structure at each position.

70. The propensity of different amino acids to support helix or beta sheet formation was generally known. For example, some amino acids are considered

[53] EX1014 (Brandon), 354.

good helix formers while others are poor helix formers. Similarly, some amino acids are good beta sheet formers while others are not. I have categorized the amino acids below based on my general understanding.

Alpha Helix		**Beta Sheet**	
Good Former	Poor Former	Good Former	Poor Former
Alanine (A)	Proline (P)	Threonine (T)	Alanine (A)
Arginine (R)	Glycine (G)	Valine (V)	Cysteine (C)
Lysine (K)	Aspartic acid (D)	Phenylalanine (F)	Proline (P)
Leucine (L)	Asparagine (N)	Isoleucine (I)	Glycine (G)
Methionine (M)	Serine (S)	Tryptophan (W)	Asparagine (N)
Glutamine (Q)	Threonine (T)	Tyrosine (Y)	Serine (S)
Glutamic acid (E)	Valine (V)		Histidine (H)

71. There are exceptions to these general categories. For example, proline (P) and asparagine (N) are listed as poor alpha helix formers but are favored at the beginning of an alpha helix.

72. Additionally, some amino acids are generally considered to be good at forming turns and loops, including asparagine (N), glycine (G), aspartic acid (D), serine (S), and proline (P). In loop positions, a mutation that reduces the flexibility of the backbone may be favored because reduced flexibility reduces the entropy of folding. As a result, the mutation is likely to have a stabilizing effect. Thus, proline (P) may be favored at loop positions over other amino acids which tend to be more flexible.

73. Similarly, another substitution that would likely increase stability is the replacement of a residue with high conformational freedom (*e.g.*, glycine (G))

with a proline (P), as long as the replacement does not violate other constraints. Because proline is more rigid than most amino acids, it will generally reduce the conformational freedom of the protein.[54]

c. Steric and Tertiary Interactions

74. I considered how a substitution would alter steric interactions compared to the wild-type residue. Steric interactions can cause physical disruptions due to their size and/or physical orientation of the side chain of the newly introduced amino acid.

75. Certain substitutions may create or disrupt interactions with nearby amino acids. One example would be a substitution that removes favorable van der Waals contacts, which could occur if the substituted amino acid replaces a large hydrophobic amino acid with a smaller hydrophobic amino acid.

76. A substitution may also form a cavity within the protein structure if it causes defective atomic packing of the hydrophobic core. For instance, if a large buried amino acid is replaced with a small amino acid, a cavity may form. Such a

[54] EX1014 (Brandon), 356 ("Another way to decrease the number of possible unfolded structures of a protein, and hence stabilize the native structure, is, therefore, to mutate glycine residues to any other residue and to increase the number of proline residues.").

mutation would likely be disfavored energetically even if the overall structure is maintained. The disfavored effects of a substitution that creates a cavity may be minimized if the cavity is filled with structural water molecules. I did not investigate this issue in my evaluation of potential mutations.

77. On the other hand, "mutations designed to fill existing cavities may be effective in some cases," and will likely be tolerated. However, naturally occurring proteins in general do not have cavities that can be easily filled perfectly through a single mutation, and a mutation that increases the volume of the residue usually requires the rest of the protein to make structural adjustments.[55] I considered this possibility in my evaluation of potential mutations, although this was done qualitatively. Cavity-causing mutations are more straightforward and can be more definitively determined based on visual inspection of a static structure.

78. Generally, hydrophobic clusters should have a high packing density—as high as possible without causing steric clashes. This means that hydrophobic clusters should have as many carbon groups in the region as possible while maintaining appropriate distances. An appropriate distance is generally around 3.5 to 4.5 Å between carbons.

[55] EX1014 (Brandon), 358.

79. It was also known by 2011 that a protein can adjust its overall structure to accommodate a mutation that causes modest increase in a side chain volume. I, however, did not explicitly consider this in my evaluation of potential mutations because these adjustments are difficult to predict accurately.

80. I also considered other tertiary interactions that might be influenced by a substitution. Certain substitutions were known to produce a more stable tertiary interaction. One example is a substitution that replaces a hydrogen bond in the wild-type protein with a salt bridge. Because a salt bridge between two groups containing permanent charges is more stable than a hydrogen bond, such a change would likely increase the stability of the structure. Similarly, replacing a hydrogen bond between two polar groups (*e.g.*, asparagine-glutamine) to introduce a permanent charge in one (*e.g.*, asparagine-glutamate) should be stabilizing.

81. A mutation that maintains stabilizing tertiary interactions that exist in the wild-type sequence will likely be favored. For instance, a hydrogen bond can form between polar amino acids due to the presence of matching functional groups in their side chains (*e.g.*, hydrogen bond donor and hydrogen bond acceptor). Amino acids that contain a polar group include serine (S), threonine (T), histidine (H), asparagine (N), glutamine (Q), tyrosine (Y), and tryptophan (Y). If there is a hydrogen bond, these amino acids may be replaced with a similar chemical moiety without losing the hydrogen bond. Replacing these polar amino acids with a

charged amino acid such as aspartic acid (D), glutamic acid (E), lysine (K), and arginine (R), may also be acceptable so long as the substitution does not result in a repulsive interaction.

82. On the other hand, a mutation that reduces or destroys a stabilizing interaction would be disfavored and, depending on the level of reduction, the mutation may not be tolerated. For instance, a mutation may remove a hydrogen bond, which would result in the loss of a stabilizing interaction. A mutation may also downgrade a salt bridge to a hydrogen bond, which would reduce the magnitude of the stabilizing interaction. Common examples of substitutions which may reduce stability include changing serine (S) to alanine (A), threonine (T) to valine (V), and tyrosine (Y) to phenylalanine (F). Even though these amino acids are roughly the same size, the tertiary interactions that the different amino acids have can vary. By the same token, introducing a hydrogen bonding moiety may lead to creation of a favorable interaction, which may be stabilizing. For example, replacing glutamine (Q) to glutamate (E) may result in a stronger hydrogen bond that may be stabilizing.

83. There are a few other discrete factors that are relevant to tertiary interactions.

1. Replacing a polar amino acid that is long with a shorter polar amino acid may disrupt a van der Waals interaction and may be

disfavored. Examples of this include changing glutamic acid (E) to aspartic acid (D) or changing glutamine (Q) to asparagine (N).

2. Substituting a long amino acid (*e.g.*, lysine (K), arginine (R), glutamic acid (E), or glutamine (Q)) may disrupt tertiary interactions if the long amino acid is packed against an aromatic group to provide hydrophobic contacts.

3. Substitutions that disrupt cation-aromatic interactions may be disfavored, as these types of interactions are important to protein structure. Cation-aromatic interactions may form between cationic amino acids (*e.g.*, lysine (K) and arginine (R)) and aromatic amino acids (*e.g.*, tryptophan (W), phenylalanine (F), and tyrosine (Y)).

4. Substitutions that disrupt aromatic-aromatic interactions may be disfavored as such interactions are generally important to protein structure. Aromatic-aromatic interactions may form between histidine (H), tryptophan (W), phenylalanine (F), and tyrosine (Y).

d. **Structures Observed in Proteins from Other Species**

84. For some substitutions, I consulted the structure of human HYAL1 and/or bee venom hyaluronidase.[56] I did that to consider how a particular substitution might influence the structure of PH20. This type of comparative assessment is useful because it reflects evolutionary influences of the protein's structure. Alternatively, if the structure around that position is similar to the human PH20 model, it may provide guidance that the mutation would similarly be tolerated in the human PH20 protein.

e. **Balancing the Factors to Assess Tolerability**

85. In my evaluation of whether single amino acid substitutions may be tolerated at a position, I considered all of the factors discussed above.[57] I balanced the type of impact of the substitutions based on the magnitude of each interaction may have on the protein's structure.

86. Based on my assessment, I assigned each substitution a score of 1, 2, or 3.

[56] I mapped the PH20 positions to the positions in HYAL1 to assist with my analysis and have included this as Appendix D-4 (EX1004, 174).

[57] I discuss these factors above. *See* §§ IV.C.2.a-c.

- A score of 3 indicates that I would expect the substitution to improve the overall stability of the protein.
- A score of 1 indicates that the mutation would likely reduce the protein stability and may even result in some loss of protein function.
- A score of 2 indicates that the substitution would have either no effect on the structure of the protein, or a slightly positive or slightly negative effect on the structure and function.

87. Substitutions that I scored as a 2 or 3 would be expected to be tolerated by the PH20 protein. These types of changes would either be inconsequential to the protein structure or would cause a change that would likely be beneficial to the stability of the protein. The substitutions that I scored as a 1 would likely not be tolerated by the PH20 protein. These types of changes often significantly disrupt the protein structure at the location of the substitution, which significantly increases the probability that the substitution will adversely affect folding of the protein or structural features of the PH20 protein that are important to its biological activity (*e.g.*, catalysis, ligand binding). My scoring table is provided below.

Score	Expected Impact	Expected Toleration
1	Significantly Destabilized	Likely Not Tolerated
2	Neutral or Minor Impacts	Tolerated
3	Improved Stability	Tolerated

3. Analysis of Published Results from Mutations of Hyaluronidase Proteins

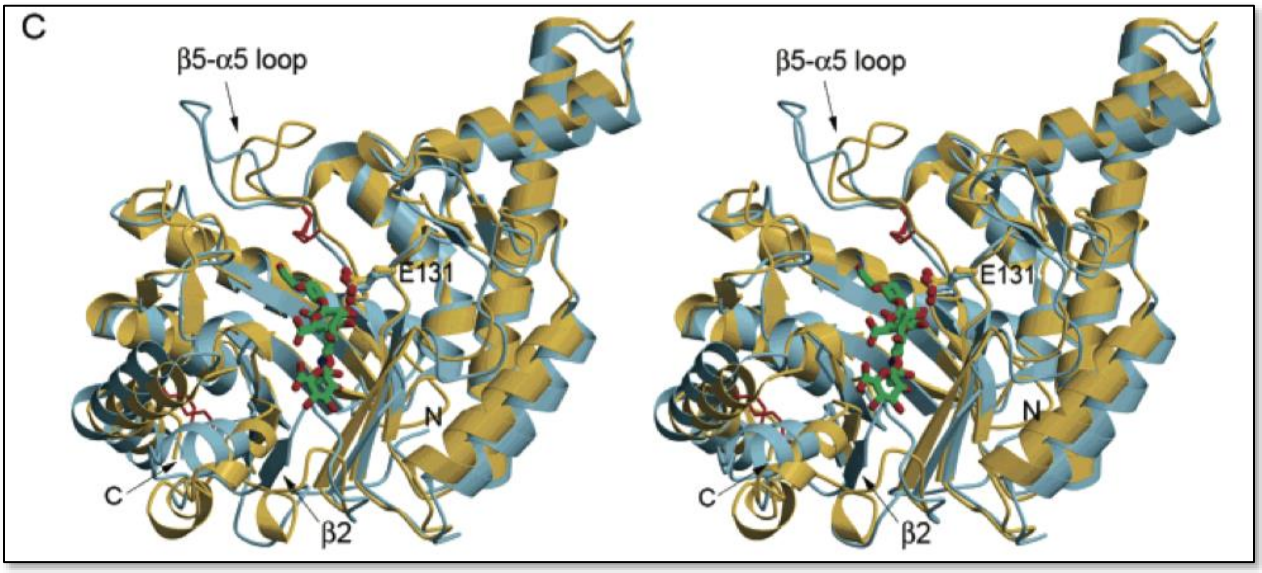
88. In my evaluation of single amino acid substitutions, I considered biochemical and structural data reported in the scientific literature before December of 2011, particularly those reported in Chao (EX1006), Zhang (EX1010), Stern (EX1008), and Arming (EX1011).

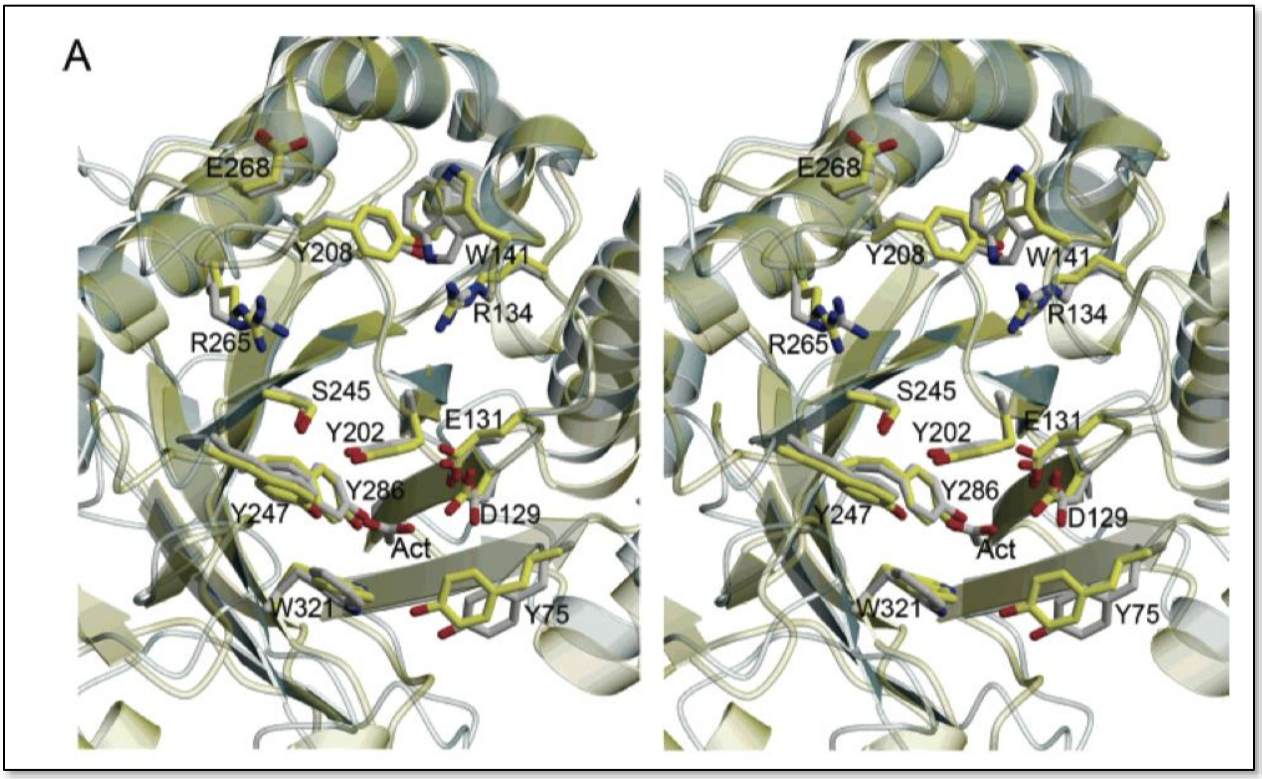
89. Chao reported an experimentally-determined structure of the human HYAL1 protein, and used it to characterize the HA binding site and other regions important to activity.[58] Chao also showed the active site of HYAL1 was structurally very similar to that in the earlier-characterized bee venom hyaluronidase protein. The top figure shows the complete structures of the two proteins overlaid (Figure 2C) while the bottom figure shows the region of the structures containing the active site overlaid (Figure 4A).[59] The images are

[58] EX1006 (Chao), 6912-13.

[59] EX1006 (Chao), 6915, 6917.

presented as "stereoscopic" images to enable 3D viewing. Either image alone is sufficient to assess the model image.





90. Chao explained that the active site of the two enzymes shared a conserved structure:

> Superposition of the structures of hHyal-1 and the bvHyal-$(GlcUA\text{-}GlcNAc)_2$ complex (22) shows that the active site clefts are similar in size and shape (Figure 2C). Many of the active site residues are conserved (Figure 4A), indicating that the mode of tetrasaccharide binding seen in bvHyal is likely to be similar in hHyal-1.[60]

91. Chao identified the residues in the active site of HYAL1 that are involved in catalysis as Asp129, Glu131, and Tyr202, which correspond to Asp146, Glu148 and Tyr219 in PH20.[61] It also identified a number of other residues in a cleft where the ligand binds (Tyr75, Trp141, Tyr202, Tyr208, Tyr210, Tyr247, Tyr261, Tyr286, Trp321).[62]

92. Chao also provided a multiple sequence alignment of the 5 human hyaluronidase enzymes. Chao's alignment shows 90 positions in PH20 that are 100% conserved among the five human hyaluronidases.[63] (Figure 3, below).

---

[60] EX1006 (Chao), 6914-15.

[61] EX1006 (Chao), 6914.

[62] EX1006 (Chao), 6914-15.

[63] EX1006 (Chao), 6916.

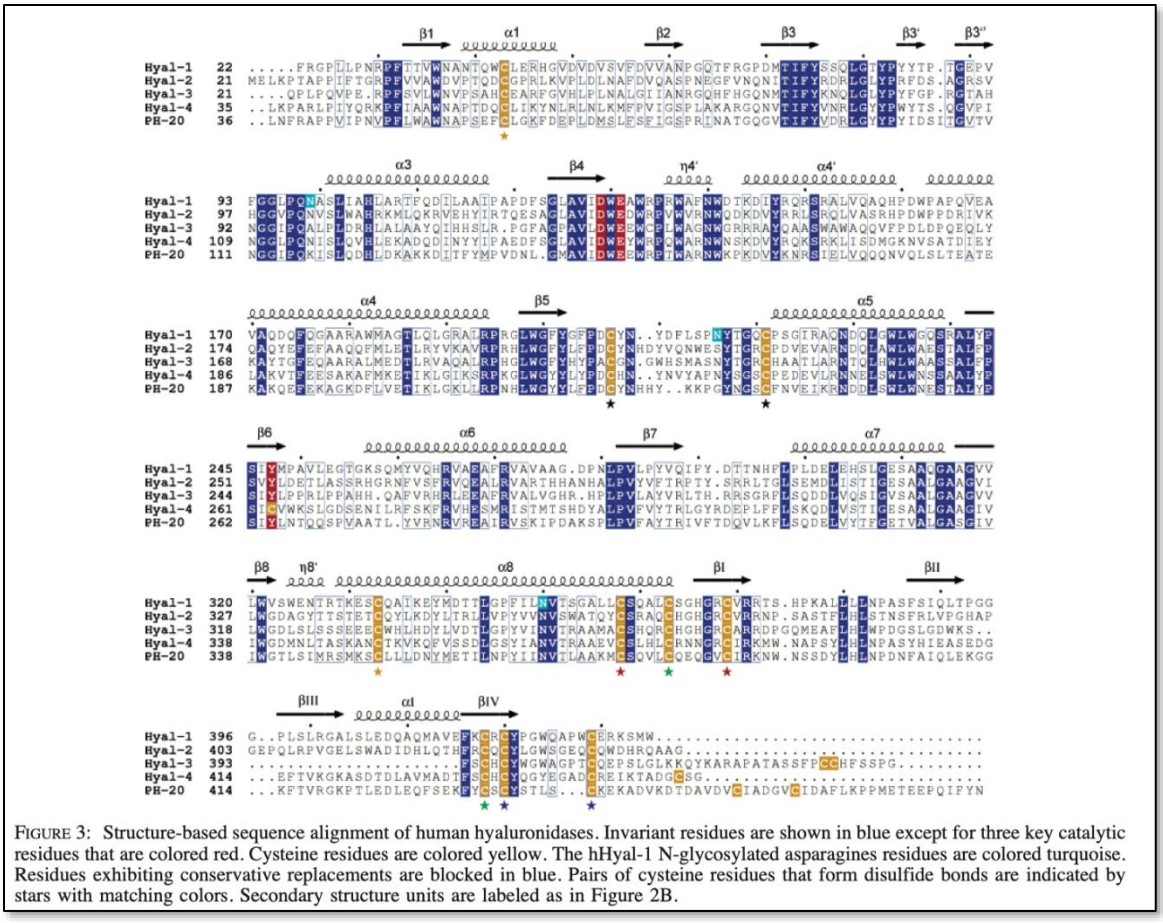


FIGURE 3: Structure-based sequence alignment of human hyaluronidases. Invariant residues are shown in blue except for three key catalytic residues that are colored red. Cysteine residues are colored yellow. The hHyal-1 N-glycosylated asparagines residues are colored turquoise. Residues exhibiting conservative replacements are blocked in blue. Pairs of cysteine residues that form disulfide bonds are indicated by stars with matching colors. Secondary structure units are labeled as in Figure 2B.

93. Chao reports more residues being conserved than the positions that I identified as being conserved within the 88 protein set I analyzed. This is not surprising because Chao only analyzed 5 human homologous proteins whereas I analyzed 88 homologous proteins from human and non-human species. Chao's alignment suggests that some of positions are conserved because of the small sample size used in their study but in fact there is variation in some of these positions. As such, I concluded that several of the positions that were reported as conserved by Chao may actually tolerate some substitutions.

94. The Zhang paper also identified residues expected to be important in HYAL1's active site using the Chao HYAL1 structure. Zhang overlaid the bee venom hyaluronidase structure in complex with HA onto the HYAL1 structure, as had been done by Chao.[64] Zhang then identified residues in HYAL1 that (i) were within 5 Å of the HA molecule and "likely to be important for catalytic activity," and (ii) residues that "lined a groove capable of accommodating a full length-HA polymer prior to cleavage." [65] (Figure 1, below).

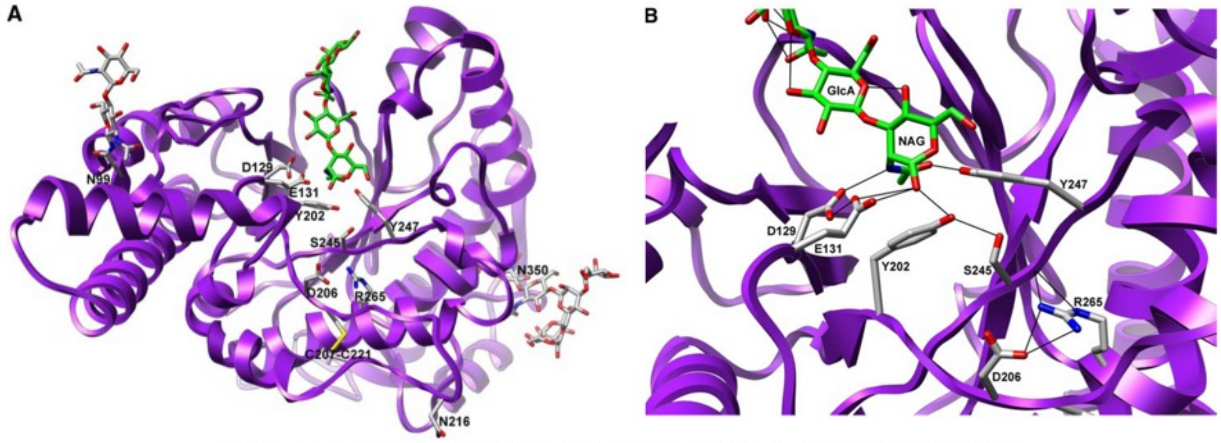


FIGURE 1. **Ribbon representation of the human Hyal1 crystal structure with bound HA tetrasaccharide.** Crystal coordinates from the HA substrate crystallized in the bee venom hyaluronidase structure (PDB code 1FCV) were used to place an HA tetrasaccharide in stick representation at the active site groove of human Hyal1 (PDB code 2PE4). *A*, side chains of residues examined in this study for potential catalytic importance are labeled and shown in stick representations. One active site proximal disulfide is indicated (Cys^{207}–Cys^{221}), as well as the three asparagines that are sites for *N*-glycosylation (Asn^{99}, Asn^{216}, and Asn^{350}), shown modified with short carbohydrate chains as determined in the crystal structure. The molecular interactions occurring between bound HA and putative catalytic residues at the active site are depicted in an active site magnification (*B*), with hydrogen bonds indicated by *thin black lines*. Carbon atoms of the protein residues and modifications are represented in *gray*, whereas those of the substrate are in *green*. In both cases, nitrogen atoms are *blue*, oxygen are *red*, and sulfurs are *yellow*. *GlcA*, glucuronic acid; *NAG*, *N*-acetylglucosamine.

64 EX1010 (Zhang), 9435-9436.

65 EX1010 (Zhang), 9435-9436.

95. Zhang then made and tested single-mutant HYAL1 proteins with mutations at the positions of the residues that were proximate to the ligand in its model. Many of them had little or no activity.[66] (Figure 2, below).

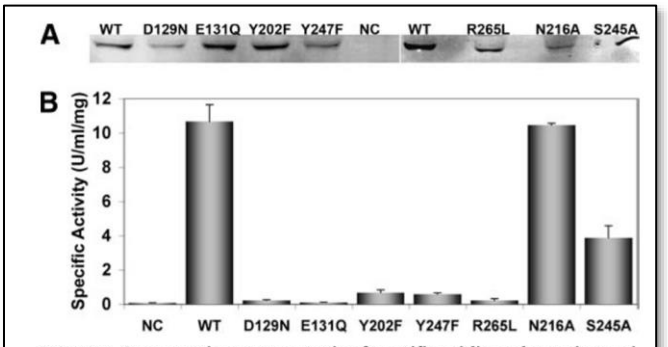


FIGURE 2. **Conservative mutagenesis of specific acidic and tyrosine residues significantly reduces Hyal1 activity.** The indicated mutations were introduced to the wild-type pFlag-Hyal1 construct as described under "Experimental Procedures." Conditioned media were prepared from 22Rv1 human prostate tumor cells transfected with the respective constructs. *A*, equivalent protein amounts, as determined by Bio-Rad assay, were analyzed by Western blot probed with anti-FLAG antibody to compare wild-type and mutant Hyal1 expression. *B*, hyaluronidase activity of the conditioned media was assayed in a quantitative microplate assay. Densitometric measurements from the Western blot in *A* were used to normalize wild-type and mutant Hyal1 content. Equal amounts of the indicated Hyal1 species were serially diluted and applied in triplicate to HA precoated microwell plates. Residual HA was quantified to obtain specific activities. Values represent the mean ± S.E.

96. Zhang also investigated mutations that eliminated two of "three canonical N-glycosylation sites" in human PH20. It found a mutation at Asn350 in the "c-terminal EGF-like domain" abolished hyaluronidase activity but one at Asn216 did not.[67]

97. Finally, Zhang expressed a form of the HYAL1 protein that removed the sequence containing residues that form the Hyal-EGF domain, which starts at

[66] EX1010 (Zhang), 9437.

[67] EX1010 (Zhang), 9438-39.

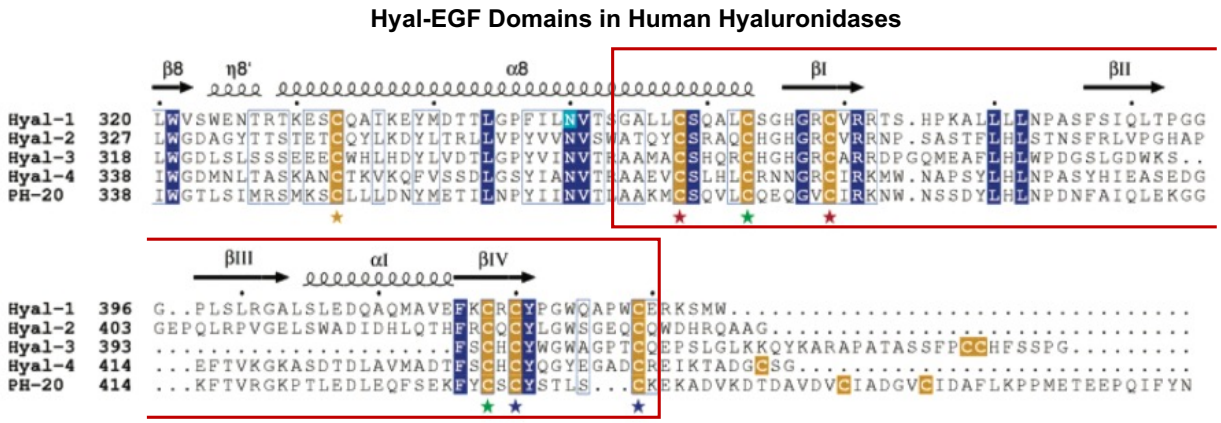
position 356 of HYAL1 and at position 374 in PH20. The Hyal-EGF domain is a unique domain found in mammalian hyaluronidases. It is characterized by a particular pattern of cysteine and glycine residues that was explained in Chao (below).[68]

> as that of bvHyal, comprising a distorted $(\beta/\alpha)_8$ barrel. No sequence homology has been reported in the scientific literature for the C-terminal domains of the mammalian hyaluronidases. However, this region contains a cysteine-rich pattern, $x_4Cx_{0-48}Cx_{3-12}Cx_{1-70}Cx_{1-6}Cx_2Gax_{0-21}Gx_2C$, where "a" denotes a hydrophobic residue, "x" denotes any residue, and the gaps between cysteine residues vary in length as indicated by the subscripts. This pattern is identified in the SMART (*23*) and PROSITE (*24*) databases as an epidermal growth factor (EGF)-like motif.

98. I have annotated the alignment in Figure 3 of Chao to show the location of the Hyal-EGF domain in the five proteins.[69] In PH20, it is located from at 337-409, with the six cysteines being at positions 341, 346, 352, 400, 402, and 408 and the four glycines being at positions 350, 377, 378, and 384.

[68] EX1006 (Chao), 6912.

[69] EX1006 (Chao), 6916.



99. When the HYAL1 mutant lacking the Hyal-EGF domain was tested, its activity was largely eliminated (~6.2%). Zhang compiled that result and the other results from testing mutants in Table 1 (below).[70]

TABLE 1
Summary of Hyal1 wild-type (WT) and mutant kinetic constants

Enzyme	K_m	V_{max}	% WT activity at 50 μM HA
	μM	μmol/min/mg	
Hyal1 wild-type	38.1 ± 4.8	12.5 ± 0.7	100.0 ± 1.8
Catalytic mutants			
E131Q	NA[a]	NA	0.08 ± 0.01
Y247F	NA	NA	0.04 ± 0.01
D129N	181 ± 19[b]	1.9 ± 0.1[b]	5.10 ± 0.09
Substrate binding mutants			
Y202F	367 ± 37[b]	6.7 ± 0.5[b]	11.1 ± 0.0
S245A	110 ± 19	10.7 ± 1.1	41.1 ± 0.4
Putative structural mutants			
R265L	NS[c]	NS	4.18 ± 0.09
N216A	103 ± 14	10.0 ± 0.8	44.6 ± 1.3
N350A	NS	NS	0.12 ± 0.01
N350tr	NS	NS	2.49 ± 0.17[d]
L356tr	NS	NS	6.29 ± 1.50[d]

[a] NA indicates no measurable activity at any HA concentration.
[b] Indicates extrapolated value from saturable curve fit.
[c] NS, not saturable, indicates data do not fit a saturation curve.
[d] Values measured at 125 μM HA.

[70] EX1010 (Zhang), 9438.

100. Stern also identified residues involved in the active site of PH20 by comparing a model of bovine PH20 complexed with bee venom hyaluronidase. Stern identified the same residues were present in comparable positions in other hyaluronidases (*i.e.*, human PH20 and HYAL1-HYAL4). Stern summarized these positions in Table 1.[71]

Table 1. Numbering Scheme for Conserved Residues among the Vertebrate Hyal Hydrolases Involved in the Catalytic Process and in Essential Positioning of the Substrate's Carbonyl of the Acetamido Group[a]

BVHyal	BPH-20	Hyal-1	Hyal-2	Hyal-3	Hyal-4	HPH-20
catalytic residue:						
Glu113	149	131	135	129	147	148
positioning residues:						
Asp111	147	129	133	127	145	146
Tyr184	220	202	206	202	218	219
Tyr227	265	247	253	246	Cys263	264
Trp301	341	321	327	319	339	339

[a] The residues were divided into the catalytic Glu and supporting residues that position the HA's carbonyl of the acetamido group for catalysis, as reported by Jedrzejas and Stern.[12] The Cys264 residue of Hyal-4 interrupts the conserved scheme and likely reflects this Hyal's specificity for chondroitin and its chondroitinase function. All other residues are strictly conserved (Figure 3).

101. Finally, in an earlier paper, Arming identified: (i) 5 single mutations that greatly reduced PH20 activity (Gln113, Gln249, Asn111, Thr252, Gly176), (ii)

[71] EX1008 (Stern), 825 (numbers include 35 residue signal sequence); EX1009 (Jedrzejas), 6912-17.

four conserved cysteine residues that form disulfide bonds and (ii) a PH20 truncated at position 341 exhibited no activity.[72]

4. Review of the Methodology Demonstrates Unbiased Evaluation

102. To develop an unbiased scoring system, it is necessary to evaluate a lot of different types of substitutions. Therefore, it was important to me to evaluate and assign scores to many different substitutions because, through this process, I developed a consistent methodology to evaluate potential substitutions. As I went through the analysis, I would revisit the scores for each position multiple times to make sure that the way I assigned scores was consistent. Because I considered a broad range of situations, I believe my methodology provided an objective and unbiased evaluation of substitutions throughout the protein.

103. I conducted my analysis in a manner that did not focus on any particular position. For instance, I considered some substitutions involving similar amino acids (*e.g.*, both amino acids were hydrophobic, charged or polar) as well as those with dissimilar amino acids (*e.g.*, hydrophobic vs. charged, polar vs. hydrophobic, large vs. small). I also considered substitutions at buried positions as

[72] EX1011 (Arming), 811, 812-813 (mutations at 113, 249, and 252 were inactive and mutations at 111 and 176 were 1-3% active).

well as substitutions at solvent exposed positions. In this declaration, I have been asked by counsel to report my conclusions with respect to position 324.

V. Analysis of Position 324 / 359

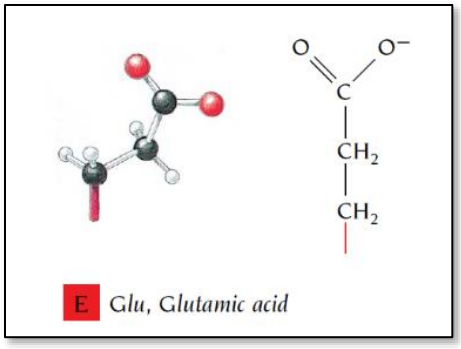
104. I evaluated potential substitutions for non-essential residues in PH20 using the methodology I described above in § IV.C.2-3. Observations from my analysis of the substitutions suggested by evolutionary variation at position 324 and a brief explanation of the basis for the score I assigned for that substitution are compiled in Appendix C.[73] I provide below a more detailed explanation for the substitutions at position 324 (position 359 with the signal sequence).

A. Description of the Structure Near Position 324

105. I assessed the local environment near position 324 in PH20, which is glutamic acid (E) in the wild-type form of human PH20. Glutamic acid is a negatively charged amino acid that is typically found on the protein surface.[74] It has an aliphatic side chain that can contribute to van der Waals contacts as well as a polar terminal carboxy group that can participate in hydrogen bonds or salt bridges.

[73] Appendix C (EX1004, 132).

[74] EX1014 (Brandon), 6.



106. There are 12 other amino acids found at position 324 across the 88 proteins analyzed in the MSA.[75] The frequency that each amino acid appeared in nature is shown in the table below. The types of amino acids that appear at position 324 vary significantly, and include polar and non-polar amino acids, charged residues, as well as residues with large and small side chains. Therefore, this position is not well conserved, suggesting that substitutions of many different amino acids are likely tolerated at position 324 in the human PH20 protein.

[75] Appendix D-1 (EX1004, 153). I also reviewed the amino acids found at position 324 in homologous proteins available by December 2012 and found that the set of amino acids remained constant.

Pos (w/s)	Pos (wo/s)	WT	Alt	Frequency (%)
359	324	E	-	12.50
359	324	-	D	25.00
359	324	-	T	13.63
359	324	-	S	12.50
359	324	-	V	7.95
359	324	-	N	6.81
359	324	-	K	6.81
359	324	-	R	5.68
359	324	-	L	2.27
359	324	-	Q	2.27
359	324	-	H	2.27
359	324	-	G	1.13
359	324	-	A	1.13

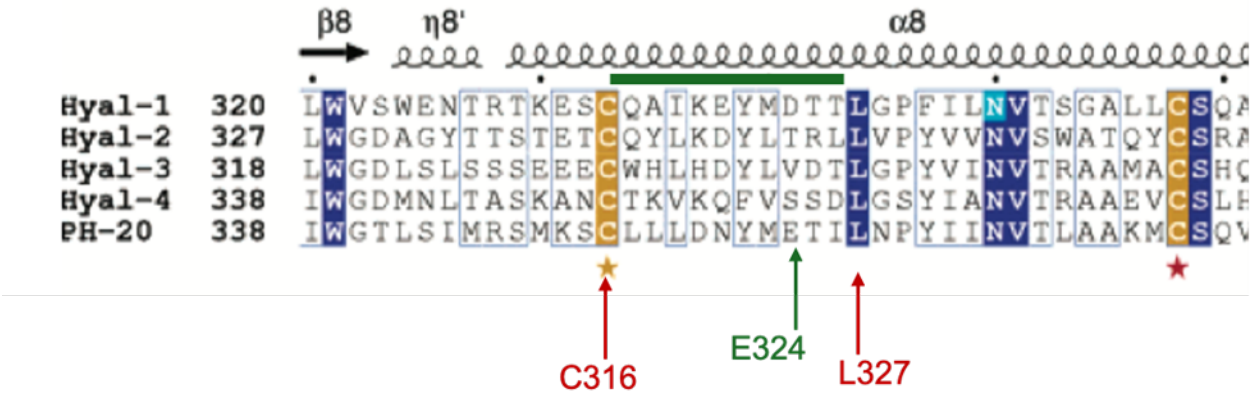
107. I visualized position 324 within the PH20 model using PyMol.[76] First, I confirmed that position 324 was not near the active site. I also noted that position 324 has 6 neighbors and inspected the placement of the neighbors and how they interacted with the glutamic acid residue at position 324.[77] Some of the neighbors are local (*i.e.*, close to S324 in sequence), including D320 and T325. Two neighbors are more distal in sequence, L374 and F380, but spatially close. I also

[76] I explain how I used PyMol in greater detail below. *See* § VI.D.

[77] Appendix E-3 (EX1004, 199).

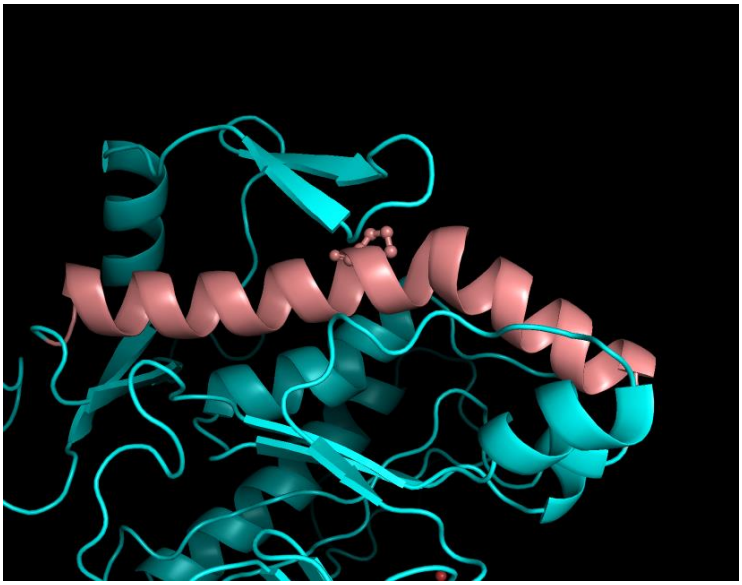
reviewed other amino acids ("non-neighbors") to confirm that more distant positions are not interacting with E324. I did not identify any such interactions.

108. Residue E324 is in the middle of the α8 helix as shown in Figure 3 of Chao (below).



109. The amino acid that occurs naturally in PH20 at position 324 is glutamic acid (E), which has a high helix propensity. Position 324 is located a few residues before a proline at position 329. When prolines occur within an alpha-helix structure, they cause a "kink" in the alpha-helical structure. The kink caused by the proline at position 329 is shown below (α8 is colored in pink). Since the secondary helical structure is partially unwound around the position of the kink, amino acids that do not have high helix propensity can be tolerated at positions near the location of the kink. For instance, glycine, which has very low helix propensity and usually does not appear in a helix, occurs in at least one

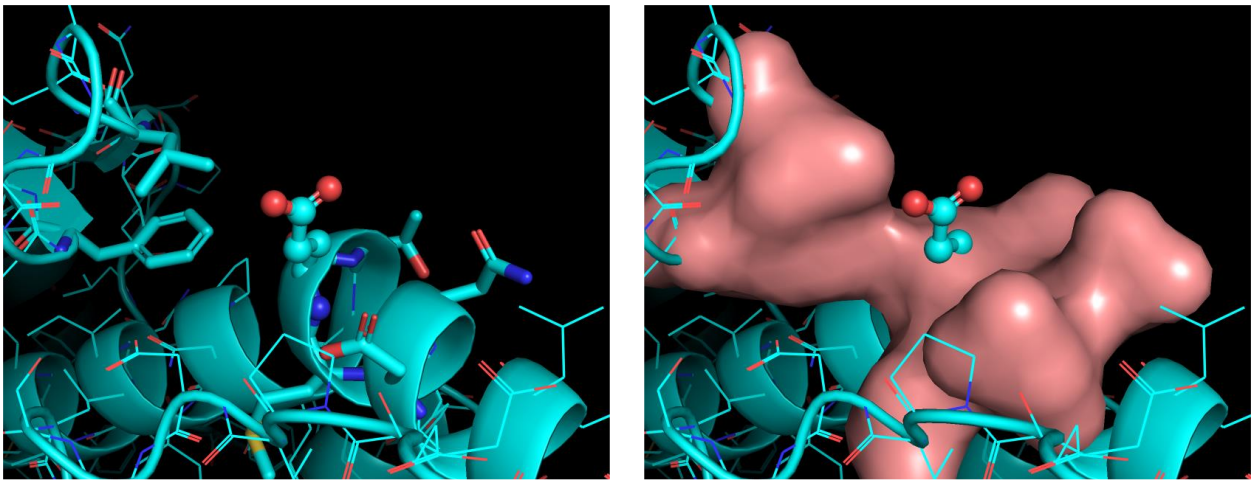
homologous hyaluronidase enzyme at a position corresponding to position 324 in PH20.[78]



110. Residue E324 is located at a solvent-exposed position. E324 has a fSASA of 0.48 which is similar to the median fSASA value for glutamic acid, which is 0.45.[79] Glutamic acid is a hydrophilic amino acid and its most important interactions involve the polar carboxyl group in its side chain. Since glutamic acid's side chain points toward the solvent, its side chain movement is not restricted by E324's six neighbors.

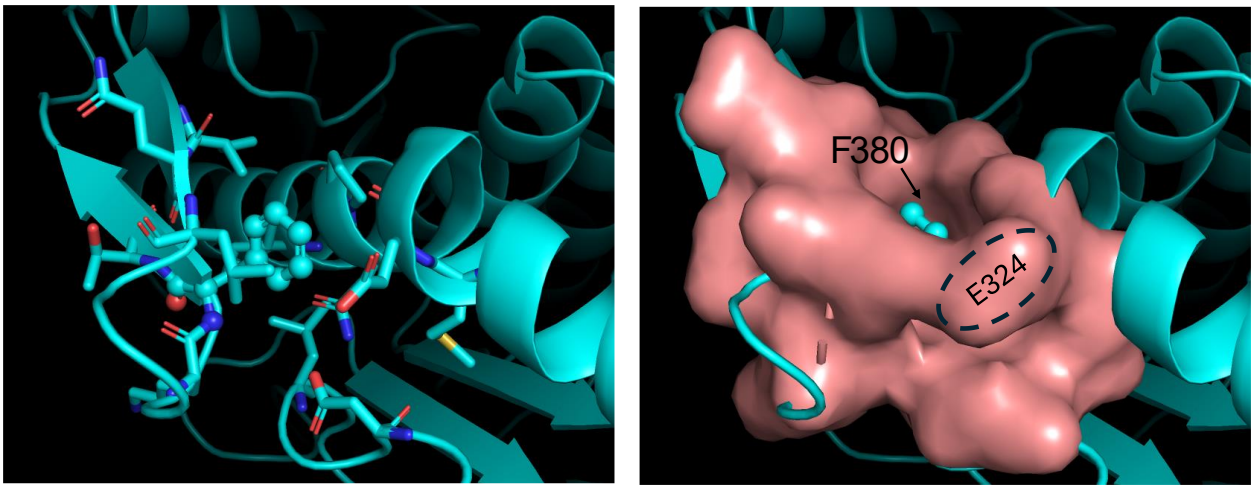
[78] Appendix D-1 (EX1004, 153).

[79] Appendix C (EX1004, 132) (fSASA of E324); Appendix D-5 (EX1004, 186).



111. The positioning of the glutamic acid at position 324 suggests that it may play a role in sterically impeding the movement of solvent molecules around F380 and creating a hydrophobic space around F380 where solvent accessibility is reduced (see figure below). In this regard, the size of the amino acid being substituted at position 324 may be an important factor that influences whether a substitution at this position will be tolerated. Many different amino acids can serve as a "gatekeeper" that restricts solvent movement, and this would explain why several amino acids with a wide range of biochemical characteristics are found at this position in homologous hyaluronidase proteins, including positively and negatively charged residues, large and small residues, and residues that have both high and low helical propensities.[80]

[80] The frequency of each amino acid are reported in Appendix D-1 (EX1004, 133). I described helix propensities above in ¶ 70.



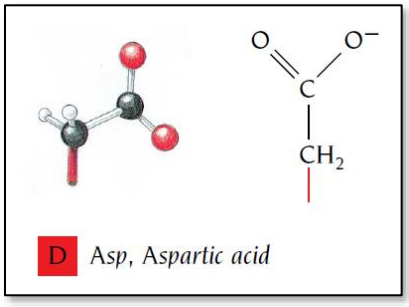
112. I evaluated the substitutions of glutamic acid (E) at position 324 with other amino acids, including aspartic acid (D), asparagine (N), arginine (R), alanine (A), histidine (H), and serine (S). For each substitution, I assessed the substituted residue within the structure using the PyMol "mutagenesis" function. I also generated mutant structures of the PH20 sequence with each substitution in it using SWISS-MODEL and assessed those models, which are also what I used for images. Based on my analysis, I believe several substitutions would be tolerated within PH20, including E324D, E324N, E324R, E324A, E324H, and E324S.

B. Assessment of E324D Substitution

113. I assessed whether aspartic acid (D) would be tolerated at position 324. Aspartic acid is found at the equivalent of position 324 in PH20 in about 25% of the 88 proteins I reviewed. It is the most frequently observed amino acid at position 324, but is not the wild-type amino acid in PH20 at that position. The

relatively high frequency of occurrence suggests that aspartic acid would be tolerated at position 324 in PH20.

114. Aspartic acid is a medium size amino acid with a negatively charged carboxyl group at the end (shown below).[81] Aspartic acid is a hydrophilic amino acid, which means it can tolerate a solvent accessible environment.



115. The PyMol protein mutagenesis feature suggested rotamer 3 as the best fit for aspartic acid at position 324, which is one of the 8 different rotamers available for aspartic acid at this position. I assessed all 8 rotamers myself and concluded that rotamer 3 would be the best fit because it avoids steric clashes with nearby residues.[82]

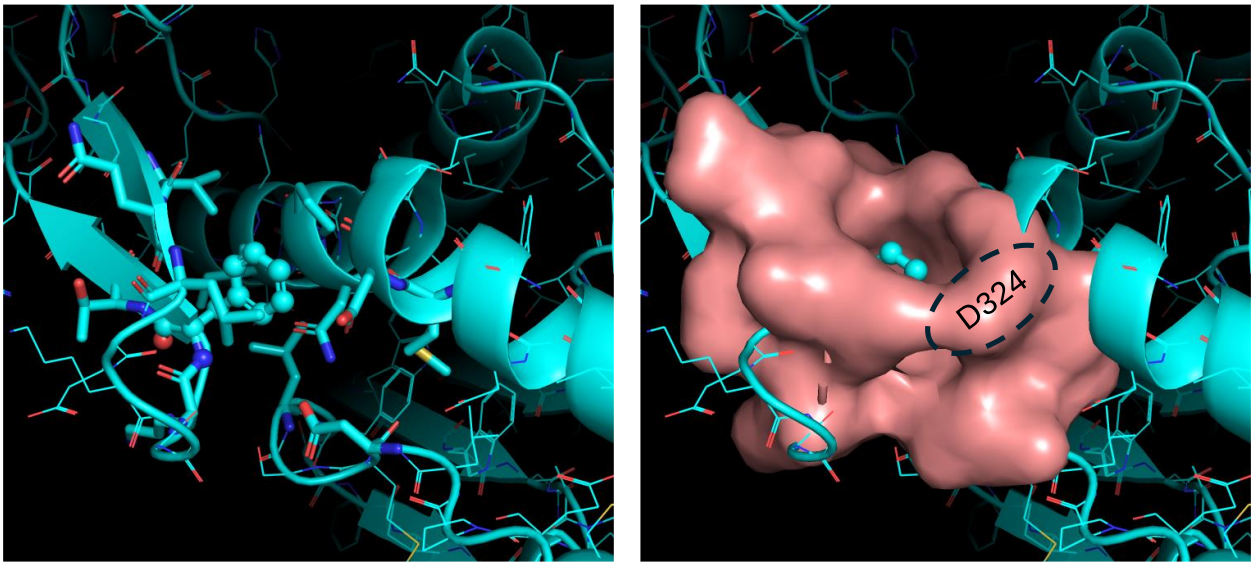
116. Similar to glutamic acid at position 324, the terminal carboxyl group of D324 will point out toward and interact with solvent. This is favorable for both

[81] EX1014 (Brandon), 6.

[82] An energy minimized form of the protein would also adjust to reduce steric clashes that might occur with other rotamers.

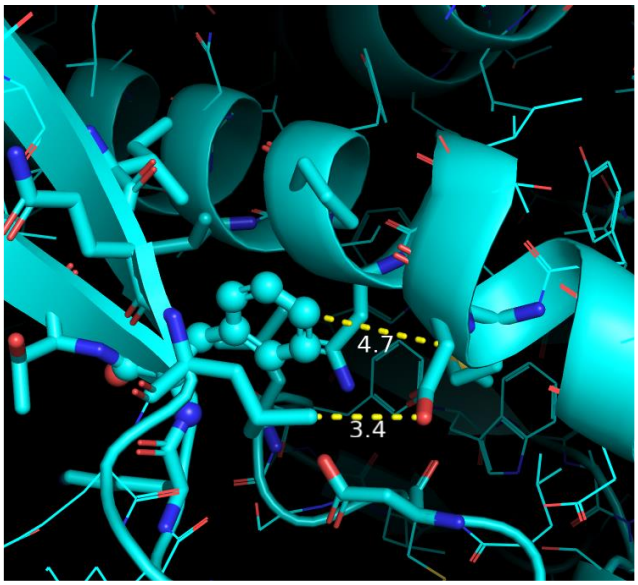
aspartic acid and glutamic acid because they are both hydrophilic residues and the terminal carboxyl groups are generally well hydrated.

117. Aspartic acid at position 324 would function similarly as glutamic acid at position 324 because it would be expected to contribute to the creation of a solvent shielded hydrophobic environment around F380. Like glutamic acid, the positioning of aspartic acid at position 324 will sterically impede the movement of solvent molecules around F380. The surface of the environment created by the neighbors around F380 with the substitution of aspartic acid at position 324 supports this conclusion (shown right below). The fractional solvent accessibility of F380 with D324 also supports this conclusion, being 0.08 with D324 compared to 0.09 with E324.



118. Substituting glutamic acid with aspartic acid at position 324 may reduce the van der Waals and/or hydrophobic interactions with its neighbors since

aspartic acid has one fewer methylene in its side chain. Aspartic acid's negative carboxyl group may also interfere with the adjacent hydrophobic interaction involving the side chains of F380 (labeled below as "4.7" Å). Additionally, the close proximity of a polar carboxyl group of D324 to a hydrophobic side chain of L374 (labeled "3.4" Å) could be somewhat destabilizing. However, I would not expect the magnitude of any reduction of van der Waals and/or hydrophobic interactions that may result from the E324D substitution to significantly impact the stability of the protein.



119. Aspartic acid has low helix propensity. Despite this, I would not expect substitution of aspartic acid at position 324 to significantly impact the

secondary helical structure because the helical structure around position 324 has been disrupted due to the proline at position 329.[83]

120. I confirmed that the modeled structure with E324D, which incorporates energy minimization, supported my evaluation based on PyMol's protein mutagenesis feature.[84]

121. Overall, I gave the E324D mutation a score of 2 as I expect for the substitution to be neutral.[85] Thus, I concluded that aspartic acid at position 324 would be tolerated at this position.

C. Assessment of E324N Substitution

122. I assessed whether asparagine (N) would be tolerated at position 324. Asparagine has a medium size side chain terminating in an amide group.[86] Asparagine is a polar amino acid, which means it can tolerate a solvent accessible environment.

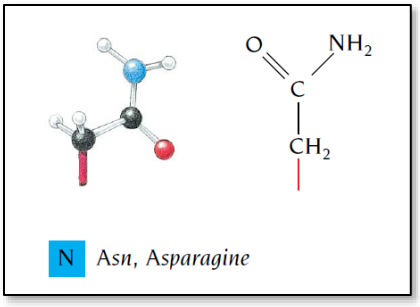
---

[83] This is also referred to as a "kink" in the helix, which is commonly introduced by a proline.

[84] The Swiss Model printout for E324D reports a QMEAN score of -3.2 (EX1070, 3) compared to PH20's QMEAN score of -2.83 (EX1069, 3).

[85] Appendix C (EX1004, 132).

[86] EX1014 (Brandon), 7.



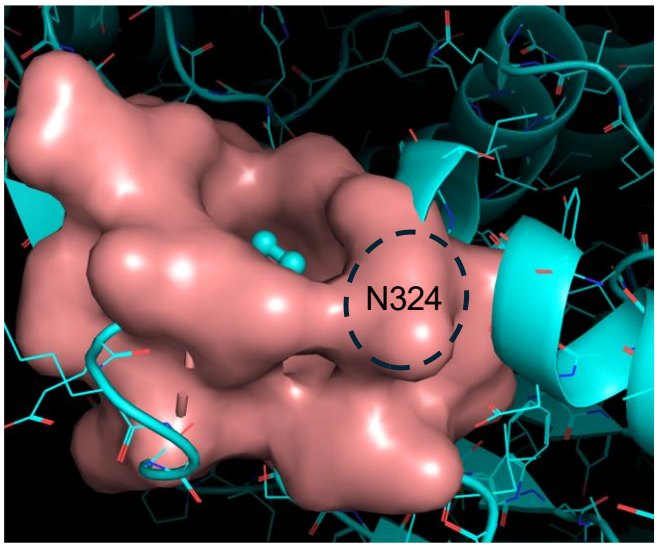
123. The PyMol protein mutagenesis feature suggested rotamer 11 as the best fit for asparagine at position 324, which is one of 11 different rotamers available for asparagine at this position. I assessed the 11 rotamers myself and concluded that rotamer 11 was a good fit as it avoids steric clashes and ensures that the side chain is well hydrated."[87] Rotamer 1 may also be an acceptable choice.

124. The terminal amide group of N324 points out toward the solvent and is therefore solvent exposed. The fSASA of N324 is 0.31, which is consistent with the median fSASA of 0.36 for asparagine. Solvent exposure is desirable for both asparagine and wild type glutamic acid because they are hydrophilic residues.

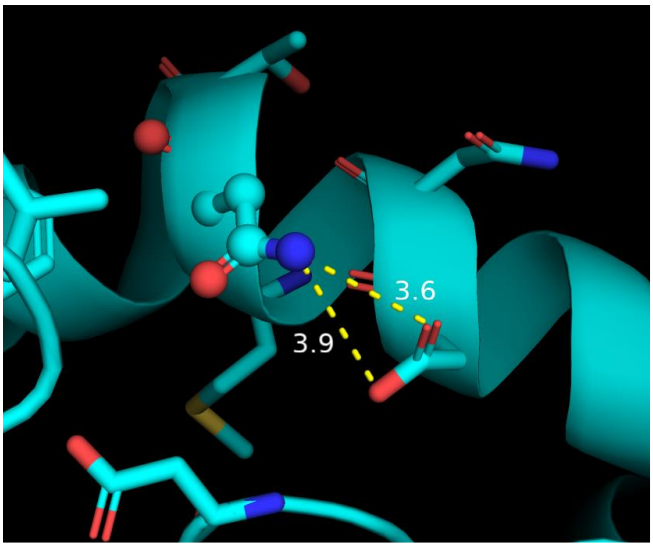
125. Asparagine at position 324 would function similarly as glutamic acid at position 324 because N324 is also expected to contribute to the creation of a solvent shielded hydrophobic environment around F380. Compared to glutamic acid's side chain, the size of asparagine's side chain is slightly smaller, but still

[87] An energy minimized form of the protein would also adjust to reduce steric clashes that might occur with other rotamers.

large enough to sterically impede the movement of solvent molecules around F380, in a way that is similar to how the side chain of glutamic acid in the wildtype PH20 at position 324 does. The surface of the environment created by the neighbors around F380 with the substitution of asparagine at position 324 supports this conclusion (shown below).



126. Additionally, the E324N substitution would substitute a negatively charged amino acid (*i.e.*, glutamic acid) with a polar, uncharged amino acid (*i.e.*, asparagine). From this perspective, the E324N substitution may be favorable because it would avoid the repulsion of charges between two negatively charged residues: E324 and D320. Asparagine at position 324 could also participate in favorable interactions with neighboring residues, such as potentially forming hydrogen bonds with aspartic acid at position 320 (shown below).



127. Asparagine has low helix propensity. Despite this, I would not expect substitution of asparagine at position 324 to significantly impact the secondary helical structure because the helix is already bent from the proline at position 329.[88]

128. I confirmed that the modeled structure with E324N supported my evaluation using PyMol's protein mutagenesis feature.[89]

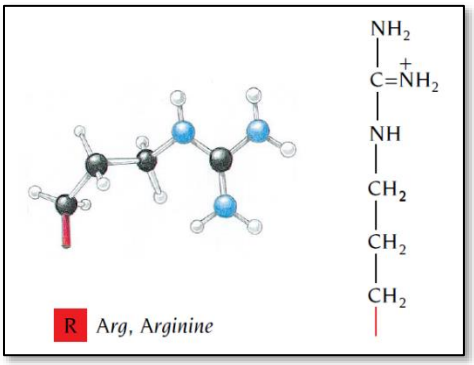
[88] This is also referred to as a "kink" in the helix, which is commonly introduced by a proline.

[89] The Swiss Model printout for E324N reports a QMEAN score of -2.85 (EX1071, 3) compared to PH20's QMEAN score of -2.83 (EX1069, 3).

129. Overall, I gave the E324N mutation a score of 2 as I expect for the substitution to be neutral.[90] Thus, I concluded that asparagine at position 324 would be tolerated in PH20.

D. Assessment of E324R Substitution

130. I assessed whether arginine (R) would be tolerated at position 324. Arginine has a long aliphatic side chain.[91] Arginine is a hydrophilic amino acid which has high helix propensity, which means it can tolerate a solvent accessible environment and is compatible with position 324's location on the α8 helix.



131. The PyMol protein mutagenesis feature suggested rotamer 1 as the best fit for arginine at position 324, which is one of 22 different rotamers available for arginine at this position. I assessed the 22 rotamers myself and concluded that rotamer 1 was a good fit as it avoids steric clashes and ensures that the side chain is

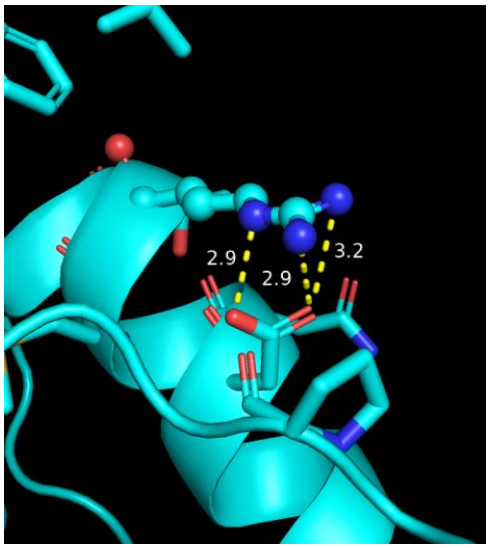
[90] Appendix C (EX1004, 132).

[91] EX1014 (Brandon), 7.

well hydrated.[92] Other rotamers, including 4, 9, and 11 may also be acceptable choices and may be preferred in some cases based on potentially favorable interactions.

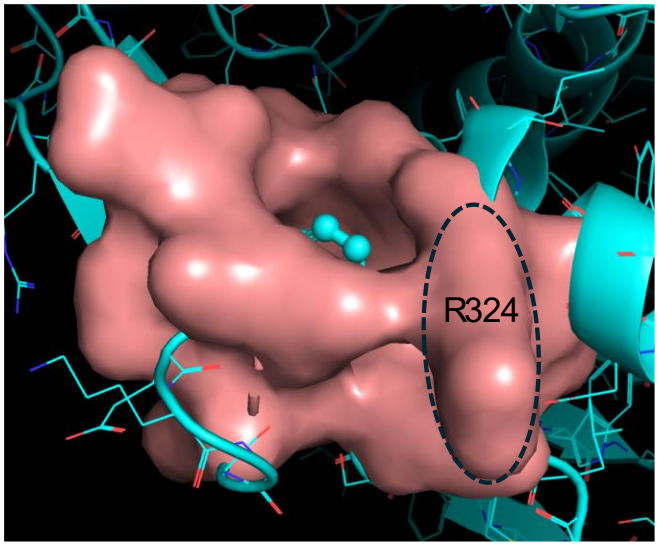
132. The terminal guanidino group of R324 points out toward the solvent and is therefore solvent exposed. This is favorable for both arginine and glutamic acid because they are both hydrophilic residues.

133. Additionally, arginine at position 324 could form a salt bridge with D320, which would stabilize the helical structure around position 324. The potential hydrogen bonds in the salt bridge between R324 and D320 are labeled below as "2.9" and "3.2" Å.



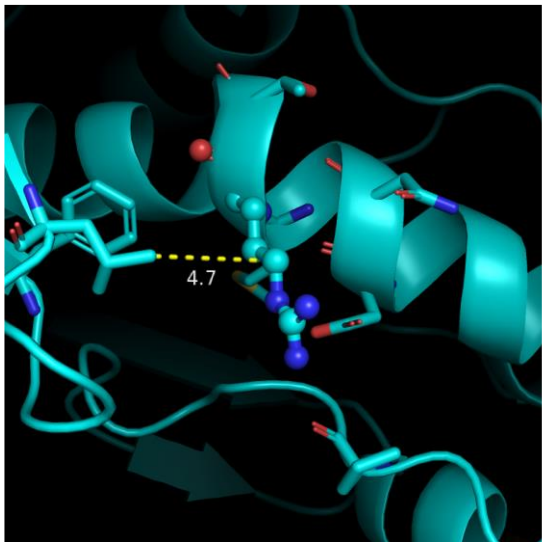
[92] An energy minimized form of the protein would also adjust to reduce steric clashes that might occur with other rotamers.

134. Arginine at position 324 would function similarly as glutamic acid at position 324 because it would contribute to the creation of a solvent shielded hydrophobic environment around F380 in a way that is similar to how glutamic acid does at position 324. Due to arginine's long side chain and its positioning at position 324, it will sterically impede the movement of solvent molecules around F380. The surface of the environment created by the neighbors around F380 with the substitution of arginine at position 324 supports this conclusion (shown below).



135. Arginine at position 324 would also participate in hydrophobic interactions similar to the interactions observed with glutamic acid at position 324. Even though arginine is a hydrophilic amino acid, it has a long aliphatic side chain that is relatively hydrophobic, similar to the aliphatic side chain of glutamic acid. Therefore, arginine's side chain can participate in hydrophobic interactions while the charged terminal guanidino group is involved in hydrophilic interactions, such

as a salt bridge. For instance, arginine at position 324 will participate in hydrophobic interactions with leucine at position 374. The hydrophobic contact between R324 and L374 is labeled below as "4.7" Å.



136. I confirmed that the modeled structure with E324R supported my evaluation using PyMol's protein mutagenesis feature.[93]

137. Overall, I gave the E324R mutation a score of 3 since the arginine would likely introduce several potentially stabilizing interactions relative to the wildtype glutamic acid residue at position 324.[94] Thus, I concluded that arginine at

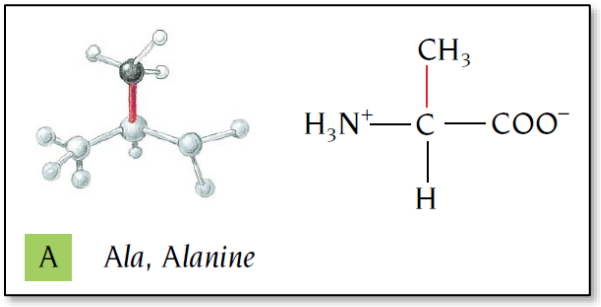
[93] The Swiss Model printout for E324R reports a QMEAN score of -2.76 (EX1072, 3) compared to PH20's QMEAN score of -2.83 (EX1069, 3).

[94] Appendix C (EX1004, 132).

position 324 may increase the stability and activity of the PH20 protein, but in any case would be tolerated at this position in PH20.

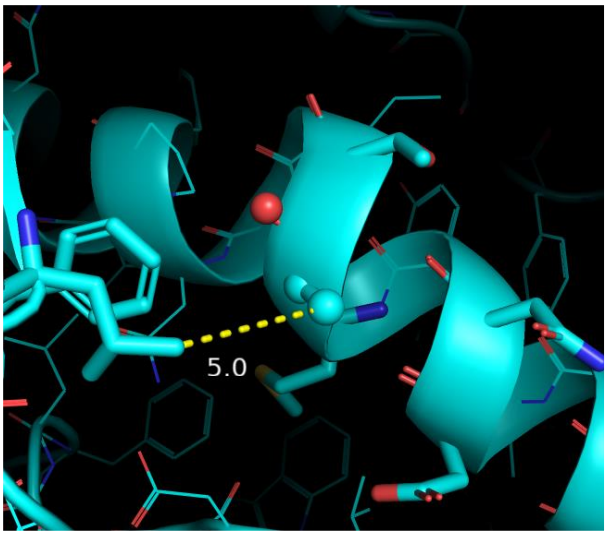
E. Assessment of E324A Substitution

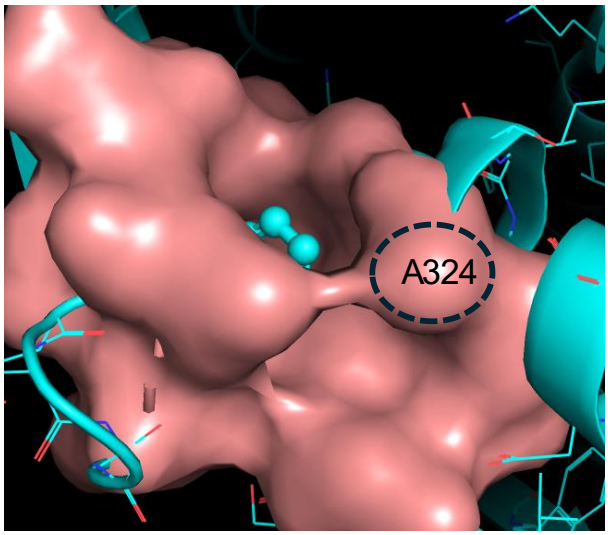
138. I assessed whether alanine (A) would be tolerated at position 324. Alanine is a small amino acid with a single carbon (Cβ) in its side chain.[95]



139. Given alanine's small size, some of the interactions involving glutamic acid at position 324 may be lost or reduced (including hydrophobic contacts to L374) upon the substitution with alanine (see below left). Alanine's side chain can still sterically impede the movement of some solvent molecules around F380. However, given its small size, alanine at position 324 may not be as effective as glutamic acid in shielding F380 from the solvent, which may have a slightly negative impact on the stability of structure in that region (see below right).

[95] EX1014 (Brandon), 6.





140. On the other hand, alanine has high helix propensity, which will help stabilize the helical secondary structure in the region around position 324. The E324A substitution also would change a negatively charged amino acid at position 324 (*i.e.*, glutamic acid) to an uncharged amino acid (*i.e.*, alanine). That change would remove a potentially repulsive interaction that occurs in the wild-type structure between two negatively charged residues (*i.e.*, the glutamic acid at position 324 and the aspartic acid at position 320).

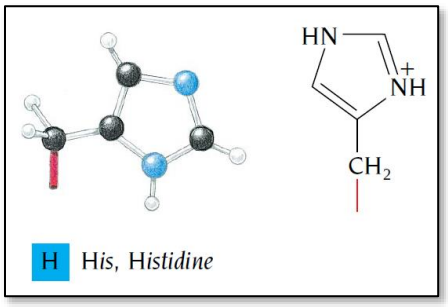
141. I confirmed that the modeled structure with E324A supported my evaluation using PyMol's protein mutagenesis feature.[96]

[96] The Swiss Model printout for E324A reports a QMEAN score of -2.99 (EX1073, 3) compared to PH20's QMEAN score of -2.83 (EX1069, 3).

142. Overall, I gave the E324A mutation a score of 2 as I expect for the substitution to be neutral.[97] Thus, I concluded that alanine at position 324 would be tolerated in PH20.

F. Assessment of E324H Substitution

143. I assessed whether histidine (H) would be tolerated at position 324. Histidine is a medium size amino acid with an aromatic side chain.[98] Histidine is a hydrophilic amino acid and is thus appropriate for position 324, which is solvent exposed.

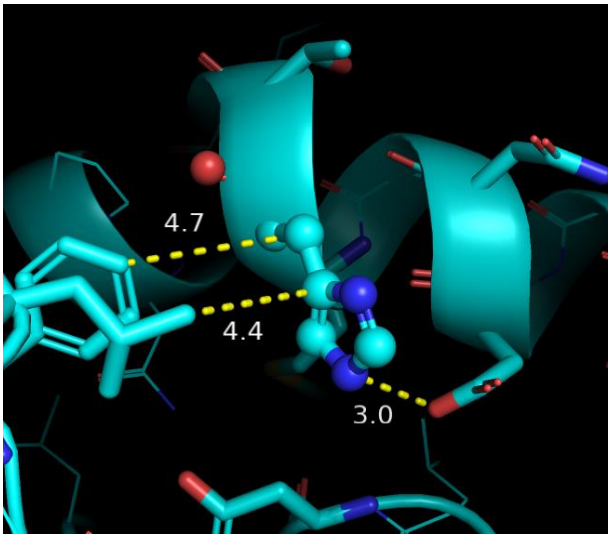


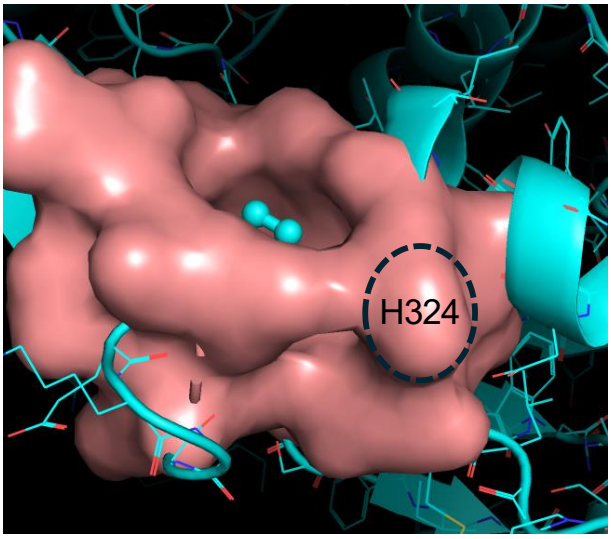
144. Histidine's side chain contains an imidazole ring with two nitrogen atoms that may participate in hydrogen bonds with D320 given their proximity (below, labeled "3.0" Å). The nitrogen atoms in histidine's imidazole ring may also participate in van der Waals interactions with L374 and F380 (below, labeled "4.4" and "4.7" Å, respectively). The histidine side chain also will contribute to

[97] Appendix C (EX1004, 132).

[98] EX1014 (Brandon), 7.

shielding the hydrophobic environment around F380, similar to the way that the wild type E324 does (below, right).





145. Even though histidine has low helix propensity, histidine at position 324 may not significantly impact the secondary helical structure because the helix is already bent from the proline at position 329.[99] Therefore, I would not expect that histidine at position 324 would significantly disrupt the secondary structure in that region.

[99] This is also referred to as a "kink" in the helix, which is commonly introduced by a proline.

146. I confirmed that the modeled structure with E324H supported my evaluation using PyMol's protein mutagenesis feature.[100]

147. Overall, I gave the E324H mutation a score of 2 with the potential to have a slightly positive impact.[101] Thus, I concluded that histidine at position 324 would be tolerated.

G. Assessment of E324S Substitution

148. I assessed whether serine (S) would be tolerated at position 324. Serine is found at the equivalent of position 324 in PH20 with the same frequency as the wild-type amino acid in PH20 at that position (*i.e.*, glutamic acid)—both appear in 12.5% of the 88 proteins I reviewed. The relatively high frequency of occurrence suggests that serine would be tolerated at position 324 in PH20.

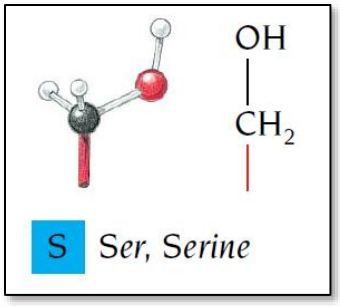
149. Serine is a small polar amino acid with a hydroxyl side chain.[102] Since serine is a hydrophilic amino acid, it is an appropriate amino acid to introduce at position 324, which is a solvent exposed position.

---

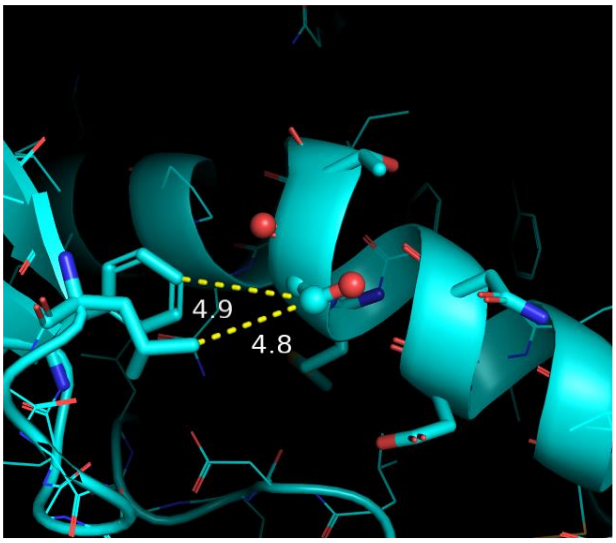
[100] The Swiss Model printout for E324H reports a QMEAN score of -2.79 (EX1074, 3) compared to PH20's QMEAN score of -2.83 (EX1069, 3).

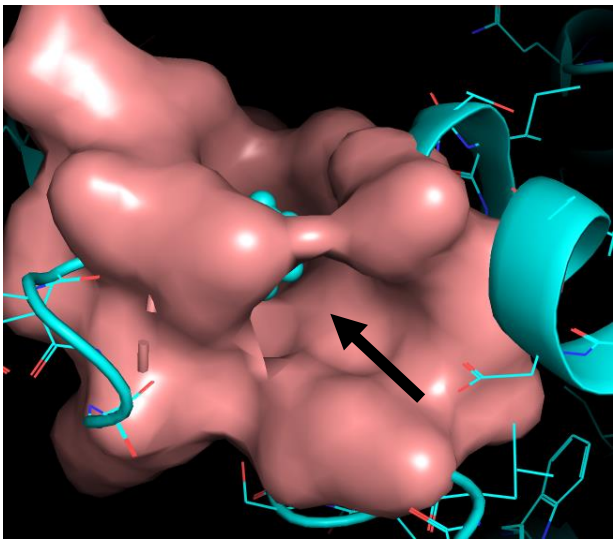
[101] Appendix C (EX1004, 132).

[102] EX1014 (Brandon), 7.



150. Given serine's short side chain, some of the interactions involving glutamic acid at position 324 may be lost (including hydrophobic contacts to L374) upon the substitution with serine (see below left). Serine's short side chain will also not be as effective as glutamic acid in shielding F380 from the solvent, which may have a slight negative impact on the stability of structure in that region (see below right). The arrow indicates the region near F380 which has increased solvent exposure as a result of serine at position 324.





151. Serine has low helix propensity. Despite this, I would not expect substitution of serine at position 324 to significantly impact the secondary helical

structure because the helical structure around position 324 has been disrupted due to the proline at position 329.[103]

152. I confirmed that the modeled structure with E324S supported my evaluation using PyMol's protein mutagenesis feature.[104]

153. Overall, I gave the E324S mutation a score of 2 as I expect for the substitution to be neutral.[105] Thus, I concluded that serine at position 324 would be tolerated in PH20.

VI. Tools Used in My Analysis

154. To perform my analysis, I used a set of tools, including BLAST, Clustal-Omega, SWISS-MODEL, and PyMol. Each of these tools was available by December 2011 and would have been used by a person of ordinary skill in the art. Although I have used these tools as they are available today, I believe the features of them that I was using in my analysis functioned equivalently to how they would have functioned in 2011. I explain how I used each tool in my analysis

[103] This is also referred to as a "kink" in the helix, which is commonly introduced by a proline.

[104] The Swiss Model printout for E324S reports a QMEAN score of -3.18 (EX1075, 3) compared to PH20's QMEAN score of -2.83 (EX1069, 3).

[105] Appendix C (EX1004, 132).

and why I believe these tools provide outputs which reliably portray what would have been available by December 2011.

A. BLAST Search and Narrowing of Returned Sequences

155. I performed a protein–protein BLAST search using the "BLASTP" program from the NIH's National Library of Medicine.[106] The "BLASTP" program allows a person to enter a FASTA sequence and search protein databases for homologous sequences.[107] This program was widely used by December 2011.[108]

---

[106] The BLASTP program that I used is available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome

[107] EX1017 (Green), 224 (BLAST is an "[o]nline service to search for related sequences"); EX1016 (Steipe), 183.

[108] *See, e.g.*, EX1016 (Steipe), 16 ("Typically, a protein-protein BLAST search would be the starting point for the collection of homologous sequences."); EX1064 (BLAST), 1, 4 ("NCBI has provided BLAST sequence analysis services for over a decade"). I reviewed webpages from BLAST which were captured by December 2011 to confirm that the program's functions in 2011 are consistent with how it functions today. EX1064 (BLAST), 10 ("Standard

156. I performed a BLAST search using the human PH20 sequence in FASTA format (Uniprot P38567)[109], and searched against the "reference proteins" database (shown below), as a skilled person would have done in 2011.[110] The general parameters were chosen to yield up to 5000 sequences with the Expect threshold of 0.001. All other algorithm parameters were left the same as default. The search retrieved 5,000 homologous sequences, which I downloaded as a text file ("hom5000.txt").

157. The 5,000 sequences returned by the BLAST search included sequences that were not available by December of 2011. To limit my analysis to only information that was available prior to December 2011, I eliminated sequences published after December 2011 by taking the following steps:

---

protein-protein BLAST (blastp) is used for both identifying a query amino acid sequence and for finding similar sequences in protein databases"), 4, 23-28.

[109] The Uniprot sequence I used in the search contains the 35 amino acid signal sequence.

[110] *See* EX1016 (Steipe), 183 ("A standard BLAST search (http://www.ncbi.nlm.nih.gov/BLAST/) with default parameters should be sufficient to retrieve the sequences of interest.").

(a) I copied the header section of the text file into a separate file ("hom5000_header.txt") to decrease the amount of text so the data would be easier to manipulate. Then, I extracted the accession numbers of the retrieved sequences from the last column and saved them as a list of 5,000 alphanumeric codes in a temporary file.

(b) I wrote and ran a perl script to determine the accession history of each sequence (date when the sequences were first available ("first seen") in the database).[111] The script produced a list with alternating lines corresponding to the accession number and the date the sequences were first available. I filtered this list by the date the sequences were first available. I saved the sequences that were available prior to December 29, 2011 to a new file "hom_pre2011.txt". I organized this information into two columns, the first with the accession number and the

[111] For example, the accession history of sequence NP_003108 is available at https://www.ncbi.nlm.nih.gov/protein/**NP_003108.2**?report=girevhist. The script switches the bold text out for each sequence so that the accession history can be similarly determined.

second with the year the sequences were first available. There were 134 rows, where each row corresponded to an accession number. I re-saved this file as "hom_pre2011.txt".[112]

(c) I then compared each row in this file ("hom_pre2011.txt") to each row in the first file ("hom5000_header.txt") and removed entries from the first file that were not available by December 29, 2011. I saved the remaining entries as a new file ("hom_pre2011_header.txt").[113] I confirmed that the number of entries in the new file ("hom_pre2011_header.txt") was the same as the number of rows in the "hom_pre2011.txt" file, which was 95.

(h) I then wrote and ran another perl script which compared the 95 entries in the "hom_pre2011_header.txt" file to identify any duplicates. In the case of duplicates, I kept only the longest isoform of each unique enzyme from each organism. This process removed 7 isoform sequences, yielding a final set of 88 unique sequences which were homologous to human PH20 and

[112] EX1053.

[113] EX1054.

available by December 2011. I saved this list as "hom_pre2011_header_clean.txt".[114]

158. To generate a file with these 88 sequences in FASTA format, I wrote and ran another perl script that retrieved the FASTA format for each sequence from the file with the original BLASTP results ("hom5000.txt"). I saved the results containing the 88 sequences in FASTA format as a new file ("hom_pre2011.fasta").[115] The saved FASTA sequences were results of BLASTP alignment and thus contained dashes indicating deletions.

B. Clustal Omega

159. To generate an MSA, I used the Clustal Omega program, which was available by December 2011.[116] I uploaded the 88 sequences in FASTA format

[114] EX1055.

[115] EX1056.

[116] EX1043 (Sievers), 1. I also reviewed webpages from Clustal Omega which were captured around December 2011 to confirm my belief that the program's functions in 2011 are consistent with how it functions today. EX1065 (Clustal Omega), 1 (Clustal Omega is "suitable for aligning protein sequences"), 4 ("Clustal-Omega is a general purpose multiple sequence alignment (MSA) program for proteins" that "produces high quality MSAs").

("hom_pre2011.fasta") to the Clustal Omega website. The MSA that was produced has gaps and therefore the total number of aligned positions in the MSA are greater than the number of amino acid positions in human PH20. Through this analysis, I kept track of the position number in the MSA as well as the residue number that the position corresponds to in the human PH20 sequence. The human PH20 sequence appears as the first row in the MSA so that it can be easily identified and referenced.

160. I saved the MSA generated from the Clustal Omega program to a local file ("ph20_pre2011.align-clustal_num"). The output of the Clustal Omega alignment of the set of 88 sequences I submitted is provided both as a text file and as an HTML file that displays coloring of the residues. The text file output is provided in Exhibit 1057, while the HTML display of the results is provided in Exhibit 1058.

C. SWISS-MODEL

161. I used the SWISS-MODEL program to generate a model of the PH20 structure.[117] The workspace that is available today, although it appears slightly different, still provides the "Automated Mode" option that was available in

[117] I used the program available at: http://swissmodel.expasy.org/

2011.[118] The "Automated Mode" option analyzes the protein sequence and identifies suitable templates based on BLAST.[119]

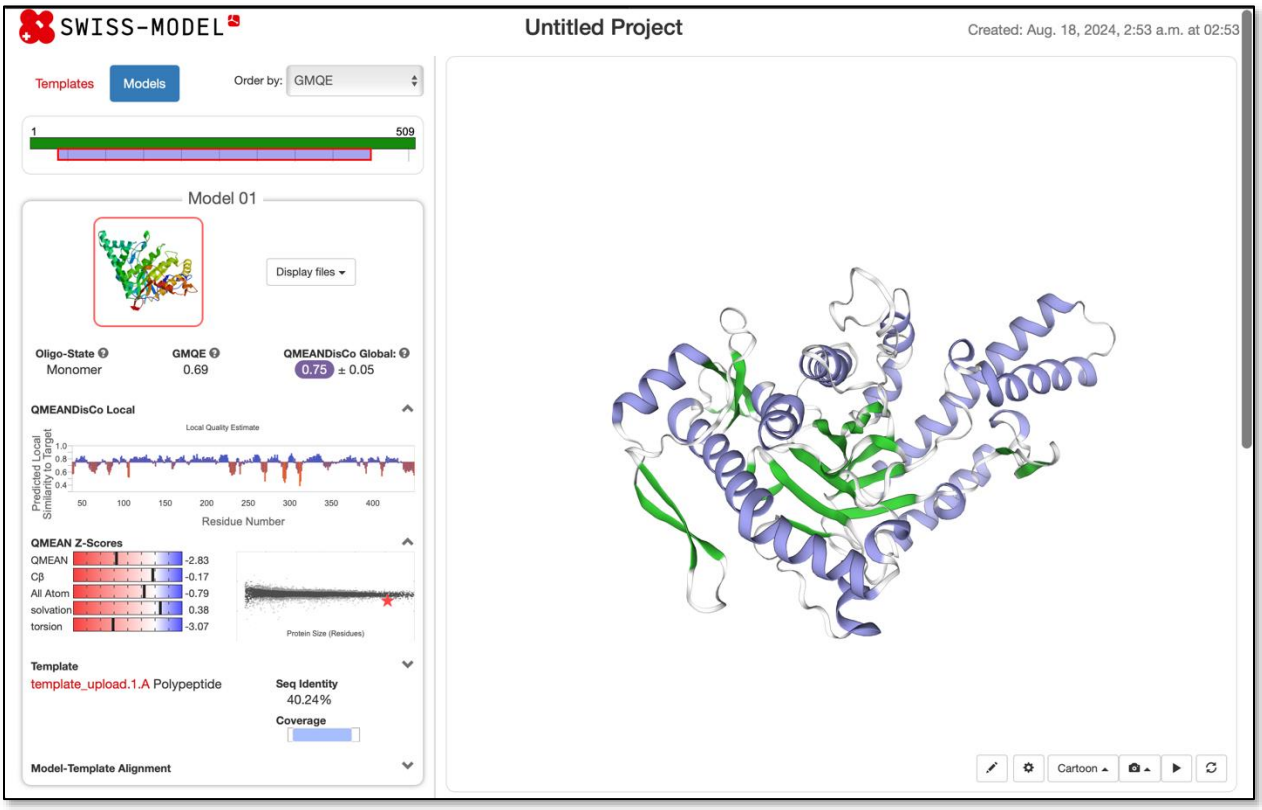
162. When I submitted the PH20 sequence, SWISS-MODEL identified as a suitable template the structure of HYAL1, which had been experimentally determined before 2011.[120] Several prior art publications also had identified and

[118] I reviewed webpages from SWISS-MODEL which were captured by December 2011 to confirm that the program's functions in 2011 are consistent with how it functions today. EX1066 (SWISS-MODEL), 1, ("Automated Mode"), 17 ("This submission requires only the amino acid sequence (FASTA format or single letter raw sequence) or the UniProt accession code of the target protein as input data."), 25 ("SwissModel Automatic Modelling Mode").

[119] EX1066 (SWISS-MODEL), 7 ("The pipeline will automatically select suitable templates based on a Blast (*Altschul et al.*) E-value limit (which can be adjusted upon submission), experimental quality, bound substrate molecules, or different conformational states of the template."), 17.

[120] This structure was determined by Chao and deposited in the RCSB Protein Data Bank on April 2, 2007. It is available online at: https://www.rcsb.org/structure/2pe4 (citing EX1006 (Chao)).

discussed the homology of the HYAL1 and PH20 proteins by 2011.[121] Of course, I could have also manually selected HYAL1 as the template by entering its PBD-ID: 2PE4, which would have also been an option in 2011.[122] An example of an output of a SWISS-MODEL session is provided below.



[121] I discuss these publications in further detail above. *See* § IV.C.3.

[122] The PDB-ID could have been entered in the "Automatic Modelling Mode" in 2011. EX1066 (SWISS-MODEL), 25 ("Use a specific template").

163. I expected that a PH20 structure modeled on the HYAL1 template would be a reliable PH20 model given the high degree of sequence identity.[123] I then confirmed that the PH20 model was reliable by assessing the QMEAN value for the model, which is a "global indicator of the quality of a given model" that is available in the SWISS-MODEL workspace.[124] The relative QMEAN assessment

[123] EX1012 (Bordoli), 4 ("The percentage of sequence identity between target and template is generally accepted as a reasonable first estimate of the quality of a model"); EX1012 (Bordoli), 2 ("*Automated mode*: If the alignment between the target and the template sequence displays a sufficiently high similarity, a fully automated homology modeling approach can be applied.").

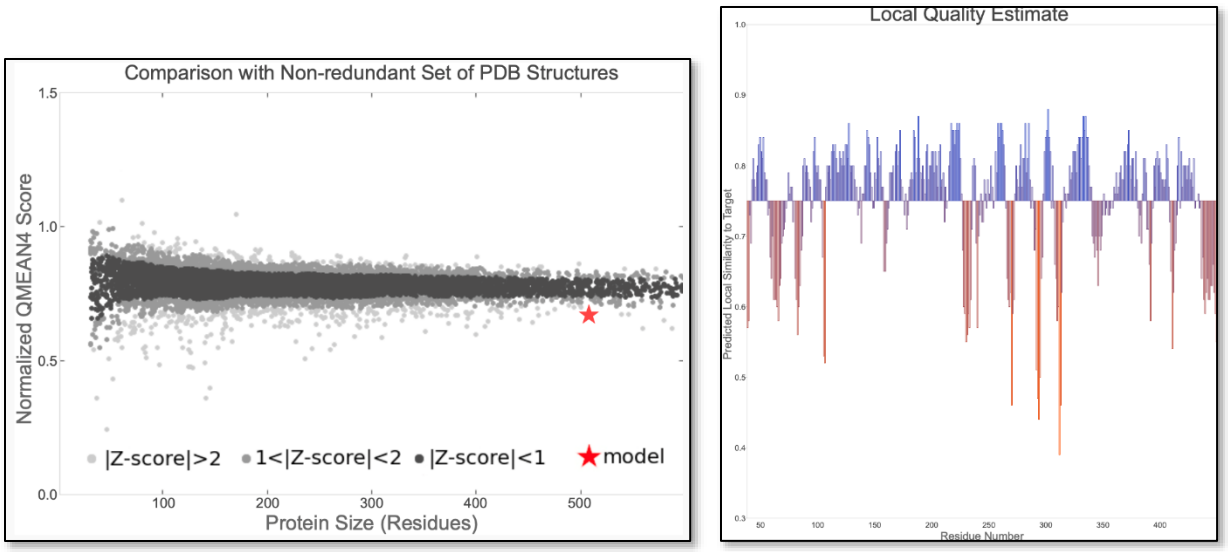
[124] EX1012 (Bordoli), 4 ("As a global indicator of the quality of a given model, the results of QMEAN, a composite scoring model,…are provided in the SWISS-MODEL workspace"). The QMEAN score was available in SWISS-MODEL workspaces by 2011. EX1066 (SWISS-MODEL), 27 ("Protein Structure & Model Assessment Tools"), 35 ("QMEAN is a composite scoring function for both the estimation of the global quality of the entire model as well as for the local per-residue analysis of different regions within a model"), 39 ("Estimating the quality of protein structure models is a vital step in protein structure prediction" and the "QMEAN server provides access to two scoring functions

for the entire protein (left) and the local QMEAN quality estimate are shown below (right). Both quality metrics indicate that the region of the model I was using for my assessments had high quality and was reliable. For example, the whole protein had an overall QMEAN score of -2.83, which is below the threshold of -4 indicative of acceptable quality.[125] Similarly, the local quality assessment remained above 0.6 for the regions that I was evaluating, particularly position 324.[126]

for the quality estimation of protein structure models which allow to rank a set of models and to identify potentially unreliable region(s)…"), 41 (showing global and local scores reported in 2011).

[125] EX1069, 3. This is the threshold I typically consider as the threshold of acceptable quality. *See also* EX1037 (Benkert), 346-47 (characterizing a QMEAN score greater than -3 as "unusually high" and a QMEAN score greater than -5 as one which indicates "something wrong with th[e] structure").

[126] EX1069, 3.



164. Portions of a structural model become less reliable in regions where there are differences in the target-template alignment.[127] The HYAL1 and PH20 sequences have different lengths: PH20 has 474 residues in its mature sequence, while HYAL1 has only 414. Also, the HYAL1 template structure cannot provide structural information for PH20 amino acids past position 403, because the sequence is no longer homologous after position 421 in HYAL.

165. To determine the reliability of the model as it approaches the C-terminus, I also analyzed that region's local QMEAN reliability (reported as B-

[127] EX1012 (Bordoli), 4 ("the accuracy of individual models may vary significantly form the expected average quality due to suboptimal target-template alignments, low template quality, structural flexibility or inaccuracies introduced by the modeling program").

factor scores in the modeled structure) of particular residues.[128] As shown above, the B-factor scores after position 403 in the PH20 model frequently begin to approach 0.6, which suggests that this region in the model is not reliable. Additionally, there is an insertion in PH20 compared to HYAL1 at the N-terminus, which means the first 3 residues of the mature PH20 protein are modeled de novo and cannot be trusted. I therefore limited my assessments based on the PH20 model in my report to positions between 7 and 403 of PH20.

166. The PH20 structural model I generated would be very similar if not identical to the model that SWISS-MODEL would have generated in 2011. As I explained above, SWISS-MODEL works by first identifying a template structure from a library of templates, based on sequence comparisons. If there is more than one suitable template structure, the template structures are superimposed and the atom positions in the structural backbone are averaged to construct the core of the

[128] EX1012 (Bordoli), 4 ("However a good global score does not guarantee that important functional sites of a protein have been modeled correcting. Therefore, tools for local model quality estimates are included."); EX1012 (Bordoli), 8 ("Local reliability ('B factor') for analyzing the local (per-residue) model reliability that can help identifying potentially incorrect regions in the model.").

model.[129] If there is only one suitable template structure, like there is here, then no averaging is done and the backbone of the template structure is directly adopted.

167. HYAL1 is the only suitable template structure for PH20 and was used as the template of the PH20 model I prepared. To make the PH20 model, SWISS-MODEL first models the conserved residues (*i.e.*, the residues in PH20 which also appear in HYAL1) by adopting their conformation from the HYAL1 template structure. This means that where PH20 and HYAL1 sequences match, the HYAL1 amino acid conformations are directly imported from the HYAL1 template structure into the modeled PH20 structure. This is the case for 165 positions in PH20. As a result, the modeled PH20 structure used in this analysis shows an identical backbone conformation as the HYAL1 structure and identical side chain conformation at conserved positions, as would be expected. SWISS-MODEL would have generated the same backbone structure for the PH20 in 2011.

168. SWISS-MODEL models positions in PH20 which are not conserved. For instance, there are 3 insertions and 3 deletions between HYAL1 and PH20, meaning one has more or fewer amino acids than the other. Insertions and deletions can introduce variation in the main chain conformation, but these

[129] EX1038 (Schwede), 3382 ("To generate the core of the model, the backbone atom positions of the template structure are averaged.").

constitute only minor perturbations because every insertion and deletion between HYAL1 and PH20 only involve one amino acid.

169. An example of an insertion with only minor perturbation is position 277 in PH20, where the loop near this position is modeled. This position contacts the residue at position 309. Although this introduces some uncertainty, precise modeling of position 277 is not critical for the evaluation of position 309 because 277 is located in a solvent exposed loop and its conformation can easily change in response to a new amino acid at position 309.

170. Side chain modeling also was well established by 2011 and SWISS-MODEL modeled non-conserved side chains by searching through rotamer libraries while minimizing energy function.

171. Therefore, overall, the PH20 model from SWISS-MODEL that I used in my analysis should be very similar to what a skilled artisan would have obtained in 2011. Changes in the modeled structures between today and in 2011, if any, would be very subtle and would not impact my analysis, which was based on my visual analysis of the structures at the positions where substitutions were being assessed. Even though homology modeling inherently involves some uncertainties, the modeled PH20 structure provides a structure that can be visualized and manipulated by the skilled person to assess the interactions between individual amino acids and their environment. My opinions reflect my visual

inspection and analysis of the modeled structure but are based on my insights as a skilled artisan.

172. Protein structure models generated by SWISS-MODEL and similar tools in the 2011 time frame would have been reliable for evaluating single amino acid substitutions that are in regions of the PH20 model where there is a sufficiently reliable local QMEAN score. This is possible because only one residue is being changed and the remainder of the protein is still based on a reliable experimentally-determined structure (here, HYAL1).

173. In 2011, a second or third substitution could possibly be considered, but modeling multiple substitutions is more challenging and predictions can quickly become unreliable. This is especially true if the mutated residues are close to each other, either in sequence or in space, so that they interact with each other. That is because one needs to model interaction between substituted residues as well as interaction between a substituted residue with constant neighbors. Because this is difficult to visualize, it becomes necessary to model individual combinations through homology modeling, which becomes combinatorial intractable.

174. Consequently, a skilled artisan in 2011 could not have used SWISS-MODEL to reliably predict the effects on the structure of a modified PH20 polypeptide of many substitutions (*e.g.*, 5 or more) in the PH20 sequence, if each position is allowed to vary independently to one of several amino acids. In

addition, a skilled artisan could not have used a SWISS-MODEL of PH20 to reliably predict the effects of substitutions within portions of the PH20 sequence where the model is not reliable (*i.e.*, outside of the PH20 sequence between positions 7 and 403).

D. PyMol

175. As I explained above, I used a program called PyMol in my analysis, which is a protein data base (pdb) file viewer that was available by 2011.[130]

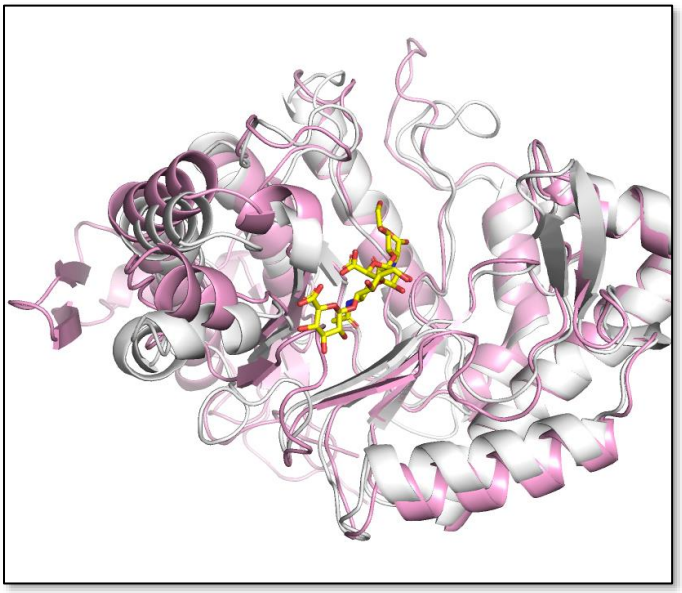
176. I first used PyMol to visualize the interaction between human PH20 and its ligand, hyaluronic acid. To do this, I opened the human PH20 structure that was modeled by SWISS-MODEL in PyMol.[131] I also opened the bee venom hyaluronidase structure with the ligand (PDB: 1FCV) that was published in 2000.[132] Within PyMol, I superimposed the bee venom hyaluronidase structure

[130] I used a recent version of PyMol: version 3.0.3.

[131] Models generated in SWISS-MODEL could be viewed in PyMol by 2011. EX1066 (SWISS-MODEL), 43.

[132] The "1FCV" structure is the trigonal form in complex with the HA tetramer. EX1033 (Markovic-Housley), 1035. The authors of Markovic-Housley deposited the structure in the RCSB Protein Data Bank by 2001. It is available online at: https://www.rcsb.org/structure/1fcv

with the ligand to overlap with the human PH20 model, as shown below.[133] Then, I hid the bee venom hyaluronidase structure so that only the human PH20 model and the ligand were visible.



177. To assist in my evaluation of positions in PH20, I created scripts that run within the PyMol environment. The scripts use built-in functions in PyMol, and are provided as Appendices E-1 and E-2 E. When I invoke the scripts, I specify the position I am inspecting (*e.g.*, 359 for position 324 in the mature protein).[134] Using my scripts, I can have PyMol identify each residue, its

[133] The bee venom hyaluronidase structure is in white, the human PH20 model is in pink, and the ligand is in yellow.

[134] This is the case for the "highlight.py" script. The "highlight_mt.py" script functions similarly by identifying the mutant residue within the structure. This

neighbors, the distance between the residue and each neighbor, as well as the surface of the environment.[135] I can also display the pockets of a particular location in the structure using a script that invokes display options built into PyMol that creates an image that visualizes the surface of the neighboring amino acids. This is helpful to quickly gain information about solvent accessibility. My scripts use functionality that has been present in PyMol since before December 2011.[136]

178. I also used a built-in function in PyMol that replaces the amino acid in the structure at a position with a different amino acid ("mutagenesis").[137] This

---

script to inspect the PH20 models with single amino acid substitutions, such as E324D.

[135] The neighbors and the distances between the residue and each neighbor populate in an External GUI Window. For my analysis of position 324, I have copied the contents from the GUI Windows into a document, attached as Appendix E-3 (EX1004, 198).

[136] I reviewed a collection of webpages from PyMol Wiki which were captured by December 2011 to confirm this. EX1067 (PyMol), 6-51, 53-57.

[137] This feature is available in PyMol by clicking on the magic wand icon in the workspace. The "Wizards" window opens and allows the user to select

function makes it possible to visualize a particular amino acid substitution at any position in the protein. It also allows the user to evaluate different rotamers of the amino acid to assess which provides the best fit in the environment being studied. I confirmed that this protein mutagenesis function was also available in PyMol by December 2011.[138]

179. The PyMol commands I used in my scripts and in my analyses of structures would have been available within versions of PyMol available in 2011 and would have been used by a skilled person in 2011. Changes in PyMol's capabilities between today and in 2011 would not impact my analysis. My opinions reflect my visual inspection of the modeled structure but are based on my insights as a skilled artisan.

VII. Determination of Numbers of Distinct Polypeptides

180. I was asked to determine the number of distinct polypeptides that have (i) one substitution at a defined position to either one or one of seven alternative amino acids, and (ii) varying percentages of sequence identity to human PH20

"Protein Mutagenesis." A user can visualize a substitution by inputting an amino acid and then clicking on the residue to be substituted.

[138] EX1067 (PyMol), 61 ("PyMol has a Mutagenesis Wizard to make mutagenesis very easy for the end user"), ("To mutate a residue follow these easy steps…").

proteins having varying lengths of C-terminal truncations. I assumed changes were due to replacements of amino acids in the sequence. I did not include a variable accounting for the possibility of changes being terminal deletions of residues, which would increase the numbers I provide below.

181. The parameters of sequence identity that I assessed were as follows:

(a) 91% sequence identity to human PH20, positions 1 to 465;

(b) 91% sequence identity to human PH20, positions 1 to 447;

(c) 91% sequence identity to human PH20, positions 1 to 474;

(d) 91% sequence identity to human PH20, positions 1 to 433; and

(e) 91% sequence identity to human PH20, positions 1 to 430.

182. I performed the above calculations for 95% sequence identity:

(a) 95% sequence identity to human PH20, positions 1 to 465;

(b) 95% sequence identity to human PH20, positions 1 to 447;

(c) 95% sequence identity to human PH20, positions 1 to 474;

(d) 95% sequence identity to human PH20, positions 1 to 433; and

(e) 95% sequence identity to human PH20, positions 1 to 430.

183. I developed a python script to calculate the number of distinct polypeptides for each set.[139] The number of replacements is calculated by multiplying the length of the protein (*e.g.*, 1-465) by 100 minus the specified percentage sequence identity (*i.e.*, 100-95%) and then rounding down to the nearest whole number. So, for example, 95% sequence identity to $PH20_{1\text{-}465}$ means that the protein can have 23 total changes. Accounting for the one required change, that permits 22 additional changes (*e.g.*, replacements).

184. I calculated the number of distinct polypeptides that exist that meet the specified criteria.[140] The number of distinct polypeptides is extremely large by all accounts, ranging from 10^{59} to 10^{112}.

[139] In it, I used "n1" to represent the total number of residues in the sequence, "n2" to be the number of required mutations at particular positions, "n3" to be the number of choices for the required mutations, "n4" to be the total number of mutations that are permitted, and "n5" to be the number of choices for the optional mutations. Appendix F-1 (EX1004, 206).

[140] Appendix F-2 (EX1004, 208).

PH20 length	*Sequence Identity %*	*# Changes*	*Pos. 324 Choices*	*Add'l Changes*	*# of Distinct Polypeptides*
430	91	38	7	37	5.36×10^{101}
447	91	40	7	39	9.77×10^{106}
474	91	42	7	41	4.42×10^{112}
430	95	21	7	20	3.08×10^{60}
447	95	22	7	21	2.63×10^{63}
474	95	23	7	22	3.63×10^{66}
430	91	38	1	37	7.66×10^{100}
447	91	40	1	39	1.40×10^{106}
474	91	42	1	41	6.32×10^{111}
430	91	38	2	37	1.53×10^{101}
447	91	40	2	39	2.79×10^{106}
474	91	42	2	41	1.26×10^{112}
465	91	41	7	40	9.88×10^{109}
465	91	41	2	40	2.83×10^{109}
465	91	41	1	40	1.41×10^{109}
433	91	38	7	37	7.02×10^{101}
433	91	38	1	37	1.00×10^{101}
430	95	21	1	20	4.40×10^{59}
433	95	21	1	20	5.08×10^{59}

I, Sheldon Park, do hereby declare and state, that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, under Section 1001 of Title 18 of the United States Code.

Executed on: Jan 15, 2015

APPENDIX A

Exhibit List

The list of exhibits that I relied upon in forming my opinions is below.

No.	Exhibit Description
1005	U.S. Patent No. 7,767,429
1006	Chao et al., "Structure of Human Huyaluronidase-1, a Hyaluronan Hydrolyzing Enzyme Involved in Tumor Growth and Angiogenesis," Biochemistry, 46:6911-6920 (2007)
1008	Stern et al., "Hyaluronidases: Their Genomics, Structures, and Mechanisms of Action," Chem. Rev. 106:818-839 (2006)
1009	Jedzrejas et al., "Structures of Vertebrate Hyaluronidases and Their Unique Enzymatic Mechanism of Hydrolysis," PROTEINS: Structure, Function and Bioinformatics, 61:227-238 (2005)
1010	Zhang et al., "Hyaluronidase Activity of Human Hyal1 Requires Active Site Acidic and Tyrosine Residues," J. Biol. Chem., 284(14):9433-9442 (2009).
1011	Arming et al., "In vitro mutagenesis of PH-20 hyaluronidase from human sperm," Eur. J. Biochem., 247:810-814 (1997).
1012	Bordoli et al., "Protein structure homology modeling using SWISS-MODEL workspace," Nature Protocols, 4(1):1-13 (2008)
1014	Brandon & Tooze, "Introduction to Protein Structure," Second Ed., Chapters 1-6, 11-12, 17-18 (1999).
1016	Steipe, "Consensus-Based Engineering of Protein Stability: From Intrabodies to Thermostable Enzymes," Methods in Enzymology, 388:176-186 (2004).
1017	Green, "Computer Graphics, Homology Modeling, and Bioinformatics," Protein Eng'g & Design, Ch. 10 (2009).
1018	Chica et al., "Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design," Curr. Opin. Biotechnol., (4):378-84 (2005).
1029	Gmachl et al., "The human sperm protein PH-20 has hyaluronidase activity," FEBS Letters, 3:545-548 (1993).

No.	Exhibit Description
1031	Yue et al., "Loss of Protein Structure Stability as a Major Causative Factor in Monogenic Disease," J. Mol. Biol., 353:459-473 (2005).
1032	Wang & Moult, "SNPs, Protein Structure, and Disease," Hum. Mutation, 17:263-270 (2001).
1033	Marković-Housley et al., "Crystal Structure of Hyaluronidase, a Major Allergen of Bee Venom," Structure, 8:1025-1035 (2000)
1035	Lins et al., "Analysis of Accessible Surface of Residues in Proteins," Protein Sci., 12:1406-1417 (2003).
1037	Benkert et al., "Toward the Estimation of the Absolute Quality of Individual Protein Structure Models," Bioinformatics, 27:343-350 (2010).
1038	Schwede et al., "SWISS-MODEL: An Automated Protein Homology-Modeling Server," Nucleic Acids Rsch., 31:3381-3385 (2003).
1039	Alberts, "Molecular Biology of the Cell," Fifth Edition, Chapter 3 (2007).
1043	Sievers et al., "Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega," Molecular Sys. Biology, 7.1 (2011)
1044	Mihel, "PSAIA – Protein Structure and Interaction Analyzer," BMC Structural Biology, 8:21 (2008)
1053	Hom_pre2011
1054	Hom_pre2011_header
1055	Hom_pre2011_header_clean
1056	Hom_pre2011.fasta
1057	Ph20_pre2011.aln-clustal_num
1058	Ph20_pre2011 Alignment html
1064	Collection of BLAST Webpages from the Internet Archive, navigable from: https://web.archive.org/web/20111022151531/http://www.clustal.org/omega/

No.	Exhibit Description
1065	Collection of Clustal Omega Webpages from the Internet Archive, navigable from: https://web.archive.org/web/20111022151531/http://www.clustal.org/omega/
1066	Collection of SWISS-MODEL Webpages from the Internet Archive, navigable from: https://web.archive.org/web/20110519141121/http://swissmodel.expasy.org/?pid=smh01&uid=&token=
1067	Collection of PyMol Webpages from the Internet Archive, navigable from: https://web.archive.org/web/20110701072314/http://pymol.org/
1069	Swiss Model Printout of PH20 Model
1070	Swiss Model Printout of PH20 Model with E324D Mutation
1071	Swiss Model Printout of PH20 Model with E324N Mutation
1072	Swiss Model Printout of PH20 Model with E324R Mutation
1073	Swiss Model Printout of PH20 Model with E324A Mutation
1074	Swiss Model Printout of PH20 Model with E324H Mutation
1075	Swiss Model Printout of PH20 Model with E324S Mutation

APPENDIX B

Dr. Park's CV

SHELDON J. PARK, PHD

Department of Chemical and Biological Engineering,
905 Furnas Hall, University at Buffalo
Email: sjpark6@buffalo.edu, Tel: 716-645-1199

EDUCATION

1994-2000	**PhD, Biophysics**, Harvard University
1991-1994	**Master of Science, Physics,** Massachusetts Institute of Technology
1987-1991	**Bachelor of Arts, Physics and Mathematics,** University of California, Berkeley

EMPLOYMENT HISTORY

2016-2021	**Director of Graduate Studies**, Department of Chemical and Biological Engineering
2014-present	**Associate Professor**, Department of Chemical and Biological Engineering, University at Buffalo
2013-present	Affiliated with the Genetics, Genomics and Bioinformatics Graduate program
2006-2014	**Assistant Professor,** Department of Chemical and Biological Engineering, University at Buffalo
2002-2006	**Postdoctoral Research Associate**, Department of Chemistry & Department of Chemical and Biomolecular Engineering, University of Pennsylvania, Philadelphia, PA
1995-2000	**Research Associate**, Department of Chemistry, Harvard University, Cambridge, MA
1991-1994	**Research Associate**, Department of Physics, Massachusetts Institute of Technology, Cambridge, MA

AWARD

National Science Foundation: CAREER Award (2011)

RESEARCH

Intellectual Properties

- Sheldon Park and Kok Hong Lim, Compositions Comprising Monomeric Streptavidin and Methods for Using Same, Technology disclosed (2011, 2013)
- Sheldon Park, Monomeric streptavidin mutants, methods of using the same and processes of manufacturing proteins, US Patent 10,759,835 (Issued 9/2020)
- Monomeric streptavidin licensed to EMD Millipore (Aug 2017)

- pH dependent trastuzumab (technology disclosure, Feb 2019)

Funding

Ongoing Support

None

Completed Support

1. R21 GM139160 ($422,167), Park (PI), Neelamegham (co-PI) 9/15/2020 – 9/14/2022
Engineering of glycosyltransferases to obtain glycan binding proteins
The grant supports the engineering and characterization of novel glycan binding proteins capable of recognizing sialic acid in a context dependent manner.

2. R21 AI138195-02 ($425,383), Park (PI), Shah, D (Co-I) 1/21/2018 – 12/31/2021 (NCE)
Modulation of antigen pharmacokinetics with pH dependent antibody
The grant supports the engineering and characterization of pH dependent antibody to neutralize and eliminate pathogenic molecules in vivo.

3. NSF CBET 1264051 **(**$300,000) Park (PI) 8/1/2013 – 7/31/2017
Temperature dependent subcellular localization
Project description: The grant supports the engineering and characterization of novel temperature sensitive intein to control the subcellular localization of proteins using the growth temperature.

4. NSF CBET 1053608 ($400,000) Park (PI) 2/1/2011 – 1/31/2016
CAREER: Yeast-based disulfide trapping for engineering selective inhibitors of a protein kinase
Project description: The grant supports the engineering and characterization of novel protein inhibitors of protein-protein interaction using yeast surface display.

5. NSF CBET 1134371 ($99,000) Tsianou, M (PI), Park, S., Alexandridis, P. (Co-PI) 9/1/2011 – 8/31/2014
Isothermal Titration Calorimeter for Bio/Nano-Materials Research and Education
Project description: The grant supported the purchase of ITC to support research in biotechnology and materials and to support education.

6. TeraGrid (20,000 hours computing time) Park (PI) 9/1/2009 – 8/31/2010
Design of high affinity monomeric streptavidin by MD simulation
Project description: Perform molecular dynamics simulations to design and test different monomeric streptavidin variants to enable high affinity biotin interaction.

7. University at Buffalo—Interdepartmental Research Development Fund (IRDF) $33,000
Park (PI), Daniels, D. (UB, Psychology, Co-I) 11/1/2008 – 10/30/2009
Modulating Physiological Response with Novel GPCR Ligands

Project description: Use yeast display to engineer novel GPCR ligands to control body fluid homeostasis.

8. American Chemical Society: Travel Award ($500) Park (PI) 8/2005
Title: Ab initio study of the phosphodiester bond hydrolysis by Cre recombinase.

Peer-Reviewed Publications

h-index (17), i10-index (22), Google citations (1458 total)
(Underline: UB undergraduate)

1. Hombu, R., Beatty, L., Tomaszewski, J.E., **Park, S.**, Neelamegham, S., Engineering glycosyltransferases into glycan binding proteins using a novel mammalian surface display platform, *Under review*.
2. Hombu, R., Neelamegham, S., **Park, S.**, Cellular and molecular engineering of glycan sialylation in heterologous systems, *Molecules* 26:5950 (2021)
3. Yang, Q., Hughes, T. A., Kelkar, A., Yu, X., Cheng, K., **Park, S.**, Huang, W. C., Lovell, J. F., Neelamegham, S., Inhibition of SARS-CoV-2 viral entry upon blocking N- and O-glycan elaboration, *eLife*, 2020;9:e61552 (doi: 10.7554/eLife.61552)
4. Le, Q., Nguyen, V., **Park, S.**, Recent advances in the engineering and application of streptavidin-like molecules, *Applied Microbio & Biotech* 103: 7355-7365 (2019) (doi: 10.1007/s00253-019-10036-5)
5. Kroetsch, A., Qiao, C., Heavey, M., Guo, L., Shah, D. K., **Park, S.**, Engineered pH dependent recycling antibodies enhance elimination of Staphylococcal enterotoxin B superantigen in Mice, *mAbs* 11:411-421 (2019)
6. Zhou, F., Kroetsch, A., Nguyen, V.P., Huang, X., Ogoke, O., Parashurama, N., **Park, S.**, High-Affinity Antibody Detection with a Bivalent Circularized Peptide Containing Antibody-Binding Domains, *Biotechnol J* 14: 1800647 (2019) (doi: 10.1002/biot.201800647)
7. Kroetsch, A., Chin, B., Nguyen, V., Gao, J., **Park, S.** Functional expression of monomeric streptavidin and fusion proteins in Escherichia coli: applications in flow cytometry and ELISA, *Applied Microbio & Biotech* 102, 10079-10089 (2018)
8. Lee, S.H., Jin, C., Cai, E., Ge, P., Ishitsuka, Y., Teng, K.W., de Thomaz, A. A., Nall, D., Baday, M., Jeyifous O., Demonte, D., Dundas, C.M., **Park, S.**, Delgado, J.Y., Green, W.N., Selvin, P.R., Super-resolution Imaging of Synaptic and Extra-synaptic Pools of AMPA Receptors with Different-sized Fluorescent Probes, *eLife* e27744 (2017)
9. Mann, J.K., Shen, J., **Park, S.** Enhancement of muramyl dipeptide-dependent NOD2 activity by a self-derived peptide, *J Cell Biochem* 118, 1127-1238 (2017)
10. Mann, J.K.(*), Demonte, D.(*), Dundas, C.M., **Park, S.**, Cell labeling and proximity dependent biotinylation with engineered monomeric streptavidin, *Technology* 4, 1-7 (2016)
11. Chamma, I., Letellier, M., Butler, C., Lim, K. H., Gauthereau, I., Choquet, D., Sibarita, J.B., **Park, S.**, Sainlos, M., Thoumine, O. Mapping the dynamics and nanoscale organization of synaptic adhesion proteins using monomeric streptavidin, *Nature Commun* 7:10773 (2016)

12. Demonte, D., Li, N., **Park, S.**, Postsynthetic domain assembly with NpuDnaE and SspDnaB split inteins, *App Biochem Biotech* 177, 1137-51 (2015)
13. Kroetsch, A. and **Park, S.,** More than one way to skin a cat: in-situ engineering of an antibody through photo-conjugated C2 domain, *Biotechnol J* 10, 508-509 (2015)
14. Wang, H., Carrier, S., **Park, S.**, Schultz, Z.D., Selective TERS Detection and Imaging through Controlled Plasmonics, *Faraday Discussions* 178:221-35 (2015)
15. Mann, J., **Park, S.**, Epitope specific binder design by yeast surface display *Meth Mol Biol* 1319:143-54 (2015)
16. Demonte, D., Dundas, C.M., **Park, S.**, Expression and purification of soluble monomeric streptavidin in E. coli *App Microbiol Biotech* 98: 6285-95 (2014)
17. Dundas, C.M., Demonte, D., **Park, S.**, Streptavidin-biotin technology: improvements and innovations in chemical and biological applications. *App Microbiol Biotech* 97:9343-9353 (2013)
18. **Park, S.**, Mann, J.K., Li, N., Targeted inhibitor design: lessons from small molecule drug design, directed evolution, and vaccine research. *J SciMed Central Chem Eng Proc Tech* 1:1004 (2013)
19. Demonte, D., Drake, E., Lim, K.H., Gulick, A., **Park, S.**, Structure based engineering of streptavidin monomer with a reduced biotin dissociation rate. *Proteins: Structure, Function, and Bioinformatics* 81, 1621-1633 (2013)
20. Mann, J.K., Wood, J.F., Stephan, A.F., Tzanakakis, E.S., Ferkey, D.M., **Park, S.** Epitope guided engineering of monobody binders for in vivo inhibition of Erk-2 signaling. *ACS Chem Biol* 8, 608-616 (2013)
21. Lim, K.H., Huang, H., Pralle, A., **Park, S.** Stable, High-Affinity Streptavidin Monomer for Protein Labeling and Monovalent Biotin Detection, *Biotechnol Bioeng* 110, 57-67 (2013)
22. Lim, K.H., Hwang, I., **Park, S.** Biotin-assisted folding of streptavidin on the yeast surface, *Biotechnol Prog* 28, 276-283 (2011)
23. Lim, K.H., Huang, H., Pralle, A., **Park, S.** Engineered streptavidin monomer and dimer with improved stability and function, *Biochemistry* 50, 8682-8691 (2011)
24. Hsu, C.K. and **Park, S.** Computational and mutagenesis studies of the streptavidin native dimer interface, *J Mol Graph Model* 29, 295-308 (2010)
25. Lim, K.H., Hsu, C.K., **Park, S.** Flow cytometric analysis of genetic FRET detectors containing variable substrate sequences, *Biotechnol Prog* 26, 1765-1771 (2010)
26. Lim, K.H., Madabhushi, S. (*), Mann, J. (*), Neelamegham, S., **Park, S.** Disulfide trapping of protein complexes on the yeast surface, *Biotechnol Bioeng* 106, 27-41 (2010)
27. Hwang, I. and **Park, S.** Computational design of protein therapeutics. *Drug Disc Today: Tech* 5:e43-8 (2009)
28. Szep, S.*, **Park, S.***†, Boder, E. T., Van Duyne, G., Saven, J. G . Structural coupling between FKBP12 and buried water. *Proteins* 74, 603 (2009) (*) Equal contribution. (†) Corresponding author.
29. **Park, S.**, Xu, Y., Stowell, X. F., Gai, F., Saven, J. G, Boder, E. T. Limitations of yeast surface display in engineering proteins of high thermostability. *Protein Eng Des Sel (PEDS)* 19, 211-217 (2006).

30. **Park, S.** and Saven, J. G. Simulation of pH-dependent edge strand rearrangement in human β-2 microglobulin. *Protein Sci* 15, 200-207 (2006).
31. **Park, S.** and Saven, J. G. Statistical and molecular dynamics studies of buried waters in globular proteins. *Proteins* 60, 450-463 (2005).
32. **Park, S.**, Boder, E. T., Saven, J. G. Modulating the DNA affinity of Elk-1 with computationally selected mutations. *J Mol Biol* 348, 75-83 (2005).
33. **Park, S.**, Kono, H., Wang, W., Boder, E. T., Saven, J. G. Progress in the development and application of computational methods for probabilistic protein design. *Comp Chem Eng* 29, 407-421 (2005).
34. **Park, S.** and Saven, J. G. Computationally assisted protein design. *Annu Rep Comp Chem*, Vol 1, Chapter 18, 245-253 (2005).
35. **Park, S.**, Fu, X., Wang, W., Yang, X., Saven, J. G. Computational protein design and discovery. *Annu Rep Prog Chem* Sect C 100, 195-236 (2004).
36. **Park, S.**, Yang, X., Saven, J. G. Advances in computational protein design. *Curr Opin Struct Biol* 14, 487-494 (2004).
37. **Park, S.**, Uesugi, M., Verdine, G. L. A second calcineurin binding site on NFAT regulatory domain. *Proc Natl Acad Sci U S A* 97, 7130-5 (2000).

Edited Book

Sheldon J. Park and Jennifer R. Cochran (Eds). Protein Engineering and Design. Taylor & Francis/CRC Press. (2009)

PhD Student Dissertations

1. "Optimized Engineering Platform for the Generation of Recycling Therapeutic Antibodies", Andrew Kroetsch, February 2019
2. "Engineering novel proteins for biotechnology application", Daniel Demonte, August 2015
3. "Selective Targeting of Cell Signaling and Inflammatory Proteins using Biologics", Jasdeep Mann, January 2014
4. "Development of Novel Screens and Molecular Tools and High Throughput Screening for Protein Engineering", Kok Hong Lim, April 2012

MS Student Theses

1. "Engineered Lasso for Stable and Stoichiometric Binding to Antibody", Fangyu "Amy" Zhou, July 2017
2. "Sequence robust loop modeling using PyRosetta", Aparajita Dasgupta, May 2016
3. "Engineered peptide to restore NOD2 response to bacterial peptidoglycan", Jiaochen Shen, August, 2015
4. "High Throughput Engineering Of Split Intein with Improved Trans-Splicing Kinetics and a Novel Split Intein Mediated Modular Synthesis", Naiyi Li, May 2014
5. "Engineering a temperature-sensitive self-cleaving intein", Daniel Demonte, May 2012
6. "Investigation on the Stability and Functions of Engineered Streptavidin", Jia Yen Leong, April 2012

7. “GPCR expression and activation in yeast”, Francis Perez, August 2011
8. “Mutational studies of monomeric streptavidin”, Taela Durst, December 2010
9. “Heterologously expressed G-protein coupled receptor in yeast”, Hyung Joon Cho, February 2009

MS Student Projects

1. “Engineering circularized peptide for high affinity antibody binding”, Xiao Huang, May 2018
2. “Expression of Functional Monomeric Streptavidin and Fluorescent Protein Fusions”, Brandon Chin, May 2018
3. “Application of Engineered Monomeric Streptavidin in ELISA”, Vyncent Nguyen, May 2017

Oral Presentations

(*) Underlined corresponds to the presenter.

1. “Expanding the molecular toolbox with engineered proteins”, Sheldon Park, Department of Chemistry, UB, December 2015
2. “Engineered and natural proteins in biotechnology”, Sheldon Park, Department of Chemical Engineering, NYU-Polytech, April 2015
3. “Engineered molecular recognition in biomedicine”, Sheldon Park, Department of Pharmaceutical Sciences, UB, September 2014
4. “Engineered molecular recognition in biotechnology and medicine”, Sheldon Park, WPI, April 2014
5. “Structure Guided Engineering of Monomeric Streptavidin With Improved Ligand Binding”, Daniel Demonte, Eric Drake, Kok Hong Lim, Andrew Gulick, Sheldon Park, AIChE annual Meeting, San Francisco, CA, November 2013
6. “Pathobiology of Crohn's Disease Risk Nod2 Mutation”, Jasdeep Mann, Sheldon Park, AIChE annual Meeting, San Francisco, CA, November 2013
7. “Targeted inhibitor design: bridging rational and directed evolution studies” Sheldon Park, Indiana University of Pennsylvania, November 2013
8. “Engineered molecular recognition in biotechnology and medicine”, Sheldon Park, Ohio State University, October 2013
9. “Engineered monobody inhibitors of Erk-2 dependent signaling”, Jasdeep Mann, Jordan Wood, Anne-Fleur Stephan, Emmanuel Tzanakakis, Denise Ferkey, Sheldon Park, AIChE Annual Meeting, Pittsburg, PA, October 2012
10. “Rational and Irrational Engineering of Novel Protein Reagents”, Sheldon Park, Department of Biology, University at Buffalo, April 2012
11. “Stable High Affinity Streptavidin Protomers for Cell Biology Application.” Kok Hong Lim,* Heng Hu, Arnd Pralle, Sheldon Park, AIChE Annual Meeting, Minneapolis, MN, October 2011.
12. “Rational Protein Engineering on the Yeast Surface.” Sheldon Park, Cheng Kuo Hsu, Kok Hong Lim, Department of Chemical and Biomolecular Engineering, University of Notre Dame, November 2010

13. "Protein complementation assay for detecting homodimeric association on the yeast surface." Sheldon Park, Jasdeep Mann, Cheng Kuo Hsu, AIChE Annual Meeting, Salt Lake City, UT, November 2010.
14. "Molecular Dynamics Simulation as a Protein Engineering Tool." Sheldon Park, Kok Hong Lim, Cheng Kuo Hsu, American Association of Pharmaceutical Scientists 2010 National Biotechnology Conference, San Francisco, CA
15. "Computational design of high affinity monomeric streptavidin." Sheldon Park, Kok Hong Lim, Cheng Kuo Hsu, AIChE Annual Meeting, Nashville, TN, November 2009.
16. "Disulfide trapping of protein complexes on the yeast surface." Sheldon Park, Kok Hong Lim, Jasdeep Mann, Sri Madabhushi, Sriram Neelamegham, AIChE Annual Meeting, Nashville, TN, November 2009.
17. "Computational design of high affinity monomeric streptavidin." Sheldon Park, BMES Annual Meeting, Pittsburgh, PA, October 2009.
18. "Structure based protein engineering." Sheldon Park, Department of Biochemistry and Hauptmann Woodward Institute, University at Buffalo. April 28, 2009
19. "Modulating the DNA affinity of Elk-1 with computationally selected mutations: relationship between protein stability and DNA binding." Sheldon Park, Eric Boder, Jeffery G. Saven, AIChE Annual Meeting, San Francisco, CA November 2006.
20. "Simulation of pH-dependent edge strand rearrangement in human β-2 microglobulin." Sheldon Park, Jeffery G. Saven, AIChE Annual Meeting, San Francisco, CA November 2006.
21. "Essential structural roles of a buried water in FKBP12: a combined approach using database search molecular dynamics, and structure determination." Sheldon Park, Jeffery G. Saven, AIChE Annual Meeting, San Francisco, CA, November 2006.
22. "Simulation of pH-dependent edge strand rearrangement in human β-2 microglobulin." Sheldon Park, Jeffery G. Saven, Chemical and Biophysics Mini-Symposium, University of Pennsylvania, Philadelphia, PA, April 2006.
23. "Using computation and experiment to study structure-function relationship in protein", Sheldon Park, University at Buffalo, March 2006
24. "Using computation and experiment to study structure-function relationship in protein", Sheldon Park, Duke University, NC, February 2006
25. "Using computation and experiment to study structure-function relationship in protein", Sheldon Park, Caltech, CA, January 2006
26. "Biophysical studies and computational design of a DNA binding protein." Sheldon Park, Eric Boder, Jeffery G. Saven, Chemical and Biophysics Mini-Symposium, University of Pennsylvania, Philadelphia, PA, October 2005
27. "Computational and experimental approaches to protein design." Sheldon Park, Virtual Institute Colloquia in Protein Modeling and Bio-nanotechnology, Drexel University, Philadelphia, PA, July 2005
28. "Database and molecular dynamics studies of buried water molecules in globular proteins", Sheldon Park and Jeffery G. Saven, Chemical Biophysical Mini-Symposium, Department of Biochemistry and Biophysics, University of Pennsylvania, PA 2005

29. "Application of molecular dynamics simulation to the study of protein-DNA interaction." Sheldon Park, Eric Boder, Jeffery G. Saven, Biomedical Engineering Society, Philadelphia, PA, October 2004
30. "Computation-assisted design of a transcription factor." Sheldon Park, Institute for Medicine and Engineering, University of Pennsylvania, Philadelphia, PA, May 2004.

Poster Presentations

1. "Disulfide trapping on the yeast surface", Jasdeep Mann, Kok Hong Lim, Sheldon Park, International Conference on Biomolecular Engineering (ICBE) 2011, San Francisco, January 2011
2. "Engineering of a stable domain swapped streptavidin dimer", Cheng-Kuo Hsu, Kok Hong Lim, Sheldon Park, ICBE 2011, San Francisco, January 2011
3. "Flow cytometric analysis of generic FRET detectors", Kok Hong Lim, Cheng-Kuo Hsu, Sheldon Park, ICBE 2011, San Francisco, January 2011
4. "Design of high affinity monomeric streptavidin by MD simulation" Sheldon Park, Proteins Gordon Conference, Holderness School, New Hampshire, June 2009
5. "Computational design of monomeric streptavidin." Sheldon Park, Cheng-Kuo Hsu, Kok Hong Lim. 2nd symposium on Nanotechnology in Biology and Medicine, sponsored by Integrated Nanostructured Systems, University at Buffalo, May 13, 2009
6. "Functional expression of streptavidin on the yeast surface." Cheng-Kuo Hsu, Inseong Hwang, Sheldon Park, AIChE, Philadelphia, PA 2008
7. "Atomic resolution structures of FKBP12 wild type and mutants show the existence of a coupled network of amino acids and a structural water in the protein core", Sheldon Park, Szilvia Szep, Eric T. Boder, Gregory D. Van Duyne, Jeffery G. Saven, San Diego, CA 2006
8. "Ab initio study of the phosphodiester bond hydrolysis by Cre recombinase", Sheldon Park and Jeffery Saven, American Chemical Society, Washington, DC 2005
9. "Statistical and Molecular Dynamics Studies of Buried Waters in Globular Proteins", Sheldon Park and Jeffery G. Saven, Protein Society, Boston, MA 2005
10. "Modulating the DNA Affinity of Elk-1 with Computationally Selected Mutations", Sheldon Park, Eric T. Boder, and Jeffery G. Saven, Protein Society, Boston, MA 2005
11. "Simulation of pH-dependent edge strand rearrangement in human β-2 microglobulin." Sheldon Park, Jeffery Saven, US-Japan Protein Engineering Symposium, University of Pennsylvania, PA, 2005.
12. "Computationally selected mutations in the transcription factor Elk-1 and its DNA binding", Sheldon Park, Eric Boder, Jeffery Saven, Biochemistry and Molecular Biophysics Symposium, University of Pennsylvania, PA 2004
13. "Synthetic protein folding and self-assembly: Computational library design and yeast expression screening", Sheldon Park, Magdalena Jonikas, Eric Boder, Jeffery Saven, American Chemical Society, New York, NY, 2003
14. "Existence of a second calcineurin binding site on select isoforms of NFAT", Sheldon Park, Gregory Verdine, American Society of Biochemistry and Molecular Biology, San Francisco, CA, 1999

Intradepartmental Student Presentations (Oral)

Includes the Annual Graduate Student Research Symposium and CBE seminars
(Underline: presenter)

1. "A Platform Engineering Approach for the Design of Recycling Therapeutic Antibodies", Andrew Kroetsch, October 2018 (Oral presentation at the Annual Research Symposium)
2. "Engineered monobody inhibitors of Erk-2 dependent signaling", Jasdeep Mann, Jordan Wood, Anne-Fleur Stephan, Emmanuel Tzanakakis, Denise Ferkey, Sheldon Park, October 2012 (Oral presentation at the Annual Research Symposium)
3. "Designed Streptavidin with Improved Affinity and Stability", Kok Hong Lim and Sheldon Park, April 2011 (Departmental Seminar)
4. "Engineering Stable High Affinity Streptavidin Monomers for Cell Biology Applications", Kok Hong Lim, Heng Huang, Arnd Pralle, Sheldon J. Park, Annual Research Symposium, September 2011 (Oral presentation at the Annual Research Symposium)

Intradepartmental Student Presentations (Poster)

(Underline: presenter)

1. "Engineering high affinity sialic acid binding protein from sialyltransferase ST3Gal1", Julia Diehl, Sriram Tendulkar, Uma Subramanian, Sheldon Park, October 2022
2. "Self-Assembling Protein Nanoparticle for Drug Delivery", Sriram Tendulkar, Uma Subramanian, Casey Flatt, Sivaker Srithar, Sheldon Park, October 2022
3. "Engineered single chain Interleukin 17 for inhibition of IL-17-driven inflammatory signaling", Quan Le, Vyncent Nguyen, Sheldon Park, October 2019
4. "Novel diagnostic tool to monitor early-stage fibrosis", Xiao Huang, Fangyu Zhou, Ogechi Ogoke, Natesh Parashurama, Sheldon Park, September 2017
5. "Expression and Purification of Functional Monomeric Streptavidin and Green Fluorescent Protein Fusion Protein", Brandon Chin, Sheldon Park, September 2017
6. "Engineered pH dependent antibody against Staphylococcal enterotoxin B", Andrew Kroetsch, Chunxia Qiao, Dhaval K. Shah, Sheldon Park, September 2017
7. "Antibody Lasso for stable and stoichiometric binding of antibody", Fangyu Zhou*, Vyncent Nguyen*, Andrew Kroetsch, Sheldon Park", October 2016
8. "Monomeric Streptavidin in Order to Promote Antibody Fluorescence Expression", Jingyuan Gao, Andrew M. Kroetsch, Sheldon Park, October 2016
9. "Engineering Z Domain of Staphylococcal Protein A as Novel Antibody Lasso and its application", Vyncent Nguyen, Fangyu Zhou, Andrew Kroetsch, Sheldon Park, October 2016
10. "Optimized Engineering Platform for Generation of pH Sensitive Antibodies", Andrew M. Kroetsch, Hsueh Yuan Chang, Dhavalkumar Shah, Sheldon Park, October 2016
11. "Sequence-robust loop modelling with PyRosetta", Aparajita Dasgupta, Sheldon Park, October 2015
12. "Engineering high affinity monomeric streptavidin", Daniel Demonte and Sheldon Park, 2012

13. "Engineered monobody to target Erk-2 dependent signaling", Jasdeep K. Mann and Sheldon J. Park, Annual Research Symposium, University at Buffalo, NY, 2011
14. "Construction and analysis of novel domain swapped streptavidin dimer", Cheng-Kuo Hsu, Kok Hong Lim, Sheldon Park, 2010
15. "Engineering High-Affinity Monomeric Streptavidin", Kok Hong Lim, Sheldon Park, 2010 – Best Poster award for Kok Hong Lim
16. "Stable streptavidin heterodimer containing complementary interfacial mutations", Jasdeep Mann, Sheldon Park, 2010
17. "Engineering of Monomeric Streptavidin", Kok Hong Lim and Sheldon Park, 2009
18. "Engineering Specificity at the Dimer Interface of Streptavidin", Cheng-Kuo Hsu and Sheldon Park, 2009
19. "Disulfide Trapping of Transient Protein Complexes on the Yeast Surface", Kok Hong Lim, Jasdeep Mann, Sri Ranganayaki Madabhushi, Sriram Neelamegham, and Sheldon Park, 2009
20. "Computational design of monomeric streptavidin" Sheldon Park, Cheng-Kuo Hsu, Kok Hong Lim. 2nd symposium on Nanotechnology in Biology and Medicine, sponsored by Integrated Nanostructured Systems, University at Buffalo, May 13, 2009
21. "Heterologously expressed GPCR in yeast", Hyung Joon Cho and Sheldon Park, 2008
22. "Functional expression of streptavidin on yeast surface", Cheng-Kuo Hsu and Sheldon Park, 2008
23. "Substrate structure and the phosphorylation kinetics", Kok Hong Lim and Sheldon Park, 2008
24. "High-throughput screening of a peptide library to modulate the activity of heterologously expressed GPCR in yeast", Hyung Joon Cho and Sheldon Park, 2007
25. "Development of a high throughput screen for protein-protein interface based on conformation-dependent proteolytic susceptibility", Cheng-Kuo Hsu and Sheldon Park, 2007
26. "Engineering of Monomeric Streptavidin", Inseong Hwang and Sheldon Park, 2007

Other Presentations

1. Guest lecturer: Pharmaceutical Sciences 539, University at Buffalo, Jan 2011
2. Guest lecturer: Biomolecular Engineering 101, University at Buffalo, Feb 2011
3. Guest lecturer: Biomolecular Engineering 101, University at Buffalo, Nov 2009

Research Collaborators

Former

- Arnd Pralle, Department of Physics, University at Buffalo
- Denise Ferkey, Department of Biological Sciences, University at Buffalo
- Manolis Tzanakakis, Department of Chemical Engineering, University at Buffalo
- Andrew Gulick, Hauptman Woodward Institute
- Olivier Thoumine, Institut Interdisciplinaire de Neurosciences, Université Bordeaux
- Paul Selvin, Department of Physics, University of Illinois, Urbana-Champagne
- Natesh Parashurama, Department of Chemical and Biological Engineering, UB
- Dhaval Shah, Department of Pharmaceutical Sciences, UB

Current

- Sriram Neelamegham, Department of Chemical and Biological Engineering, UB

TEACHING

Courses Taught

EAS 230 Higher Level Programming Language with Prof. H. Stenger—Spring 2008 (324 students), Spring 2009 (315 students)

CE 220 Biotechnology Principles—Spring 2023 (45 students)

CE 327 Chemical Engineering Lab I with Prof T. Kofke—Fall 2006 (32 students), Fall 2007 (38 students), Fall 2008 (43 students), Fall 2009 (48 students)

CE 427 Chemical Engineering Lab III with Prof. T. Kofke—Fall 2006 (29 students), Fall 2007 (30 students), Fall 2008 (35 students), Fall 2009 (42 students)

CE 434 Process Dynamics and Control—Fall 2010 (52 students), Fall 2011 (50 students), Fall 2012 (61 students), Fall 2013 (70 students), Fall 2014 (81 students), Fall 2015 (66 students), Fall 2016 (77 students), Spring 2017 (85 students), Spring 2018 (98 students), Spring 2019 (81 students)

CE 450/550 Protein Engineering—Spring 2007 (14 students), Spring 2008 (16 students), Fall 2009 (14 students), Spring 2011 (18 students), Spring 2012 (16 students), Spring 2013 (19 students), Spring 2014 (13 students), Spring 2016 (21 students), Fall 2017 (28 students), Fall 2018 (35 students), Fall 2019 (14 students), Fall 2020 (29 students), Fall 2021 (18 students)

CE 498 Undergraduate Research—Fall 2019 (9 students), Spring 2020 (11 students), Fall 2020 (2 students), Spring 2021 (8 students), Fall 2021 (5 students)

CE 517 Bioengineering Principles with Profs. S. Neelamegham and E. Tzanakakis—Spring 2010 (11 students)

CE 599 Supervised Teaching—Fall 2016 (20 students), Spring 2017 (17 students), Fall 2017 (15 students), Spring 2018 (13 students), Fall 2018 (13 students), Spring 2019 (12 students), Fall 2019 (13 students), Spring 2020 (14 students), Fall 2020 (11 students), Spring 2021 (16 students)

Current Students

None

Former graduate students and postdocs

PhD Students

- Andrew Kroetsch (May 2019)
  Senior Manager, Takeda, MA
  Dissertation: "*Optimized Engineering Platform for the Generation of Recycling Therapeutic Antibodies*"

- Daniel Demonte (August 2015)

Scientist, Alpine Immune Biosciences, Seattle, WA
Dissertation: "*Engineering novel proteins for biotechnology application*"

- Jasdeep Mann (February 2014)
  Senior Scientist, Bluebird Biotech, Seattle, WA
  Dissertation: "*Selective Targeting of Cell Signaling and Inflammatory Proteins Using Biologics*"

- Kok Hong Lim (May 2012)
  Senior Principal Scientist, Pfizer, Cambridge, MA
  Dissertation: "*Development of Novel Screens and Molecular Tools in Protein Engineering"*

Master Students (Thesis and Project)

- Sathvik Kumar Reddy Toppireddy (May 2024)
  "Design, Production and Applications of Novel Protein-based Nanoparticle Drug Delivery"

- Ly Tran (December 2023)
  "Targeted Degradation of an Intracellular Protein using Proteolysis Targeting Chimera (PROTACs) and Endosomolytic Peptides"

- Sriram Tendulkar (May 2023)
  University at Delaware, PhD student
  "Self-Assembling Protein Based Chassis For Drug Delivery, Immunology, Vaccine, And Research Assay Applications"

- Xiao Huang (May 2018)
  GeneWiz, South Plainfield, New Jersey
  Project: "*Engineering circularized peptide for high affinity antibody binding*"

- Brandon Lee Chin (May 2018)
  Compass Therapeutics, Cambridge, MA
  Project: "*Expression of Functional Monomeric Streptavidin and Fluorescent Protein Fusion Protein*"

- Fangyu Amy Zhou (August 2017)
  HarkerBIO, Buffalo, NY
  Thesis: "*Engineered Lasso for Stable and Stoichiometric Binding to Antibody*"

- Vyncent Nguyen (May 2017)
  Albany Molecular Research Inc, Buffalo, NY
  Project: "*Application of Engineered Monomeric Streptavidin in ELISA*"

- Aparajita Dasgupta (May 2016)
  PhD student in MDI, UB
  Thesis: "*Sequence-robust loop modeling using PyRosetta*"

- Jiaochen Shen (August 2015)
  Biogen IDEC, Cambridge, MA
  Thesis: "*Engineered peptide to restore NOD2 response to bacterial peptidoglycan*"

- Naiyi Li (May 2014)
  Bristol Myers Squibb, Devens, MA
  Thesis: "*High Throughput Engineering Of Split Intein with Improved Trans-Splicing Kinetics and Novel Split Intein Mediated Modular Synthesis*"

- Daniel Demonte (May 2012)
  PhD student in CBE, UB
  Project: "*Engineering a temperature-sensitive self-cleaving intein*"

- Jia Yen Leong (May 2012)
  Thesis: "*Investigation on the Stability and Functions of Engineered Streptavidin*"

- Francis Perez (August 2011):
  Research Scientist at NYC Department of Health and Mental Hygiene
  Project: "*GPCR Expression and Activation In Yeast*"

- Taela Durst (January 2010)
  Thesis: "*Mutational studies of monomeric streptavidin*"

- Hyung Joon Cho (June 2008)
  PhD student at Virginia Tech
  Thesis: "*Heterologously Expressed G-Protein Coupled Receptor in Yeast*"

Master of Engineering

- Shreya Iyer (May 2024)
- Aabha Satish Tengeri(May 2024)
- Anton Jayakodiarachchige (May 2024)
- Julia Diehl (May 2023)
- Uma Subramanian (May 2023)
- Sivaker Srithar (May 2022)
- Krish Subramanian (Dec 2022)
- Casey Flatt (May 2022)
- Jiapeng Zou (May 2022)
- Jacquekia Jefferson (May 2022)
- William Erdman (Dec 2021)

- Shaista Shabbir (Nov 2019): ThermoFisher Scientific
- Jingyuan Gao (December 2016)
- Cheng-Kuo Hsu (June 2011): Production Manager, Hsinchu City, China

Master of Science Track 3 course-based (purview of DGS)

- Krysta Clark (May 2021)
- Ngoc Nguyen (May 2021)
- Taylor Congdon (April 2021)
- Zhixin Zong (January 2021)
- Ming Yang (January 2021)
- Alex Rakfeldt (May 2020)
- Christian Ferger (July 2020)
- Jacob Brooks (July 2020)
- Timothy Chewens (May 2020)
- Vishal Tuli (July 2020)
- Ziming Wang (May 2020)
- Kiruthika Narayani Santhanam (Jan 2020)
- Kate Chen (August 2019)
- Bharath Vaidyanathan (July 2019)
- Jinhao Zheng (July 2019)
- Jing Xue (July 2019)
- Shin Yan Zheng (July 2019)
- Shubhankar Kapoor (May 2019)
- Abubaker Khalafalla (May 2019)
- Jayne Beckmann (Jan 2019)
- Gabriel Guzman (Jan 2019)
- Juntao Gao (Dec 2018)

Postdoc

- Dr. Jasdeep Mann: Postdoc (Feb - May 2014). Senior Scientist, Bluebird Biotech, Seattle, WA
- Dr. Inseong Hwang: Postdoc (Sept 2006 - Nov 2008). Head of R&D at DOCSMEDI, South Korea

Others

- Ryoma Hombu: PhD (December 2020 – Sept 2021). Currently a PhD student with Dr. Sriram Neelamegham, UB
- Vyncent Nguyen: PhD candidate (Jan - July 2019). Withdrew from the program. Scientist III, Thermo Fisher Scientific, Buffalo, NY)
- Pooja Chakrabarty: MS candidate (2009 – 2010)
- Dylan Servos: MS candidate (2019 – 2020)
- Nivethitha Srinivaas: ME candidate (2019 – 2020)

Former Undergraduates and High School Students

- Quan Le: CBE undergraduate (Feb 2018 – May 2020).
  PhD student at Boston University.
- Alex Mack: CBE undergraduate (Sept 2019 – May 2020)
- Thalia Taylor: CBE undergraduate (June 2019 – Dec 2019)
- Callie Bailey-Wickens: CBE undergraduate (Summer and Fall 2018)
- Benjamin Carlson: CBE undergraduate (Fall 2015 - 2017)
- Thomas Straubinger: Pharmaceutical Sciences undergraduate (Fall 2015)
- Christopher Dundas: CBE undergraduate (Fall 2011 – Spring 2015).
  PhD student at U Texas, Austin.
- Laura Saunders: CBE undergraduate (Spring 2015)
- Darwin Chen: CBE undergraduate (Fall 2014)
- Ming Chen: CBE undergraduate (Spring 2014)
- Karl Barber: CBE undergraduate (Fall 2009 – Spring 2012)
  PhD student at Yale University
- Shannon Griffin: Kenmore West High School (Summer 2011)
- Yimin Tang: CBE undergraduate (2009 – 2010)
- Lye Lin Lock: CBE undergraduate (Spring and Summer 2009)
- Oluwatobi Busari: Mechanical and Aerospace Engineering (Summer 2009)
- Brett Van Groenewould: CBE Undergraduate (Fall 2008)
- Ryan Tomko: CBE undergraduate (Spring 2008)
- Mary Stottele: Sacred Heart Academy, Amherst, NY (Summer 2008)
- Mary Brummond: CBE undergraduate (Summer 2007)
- Kristen Marra: CBE undergraduate (Summer 2007)

Student Honors

- Andrew Kroetsch: Mark Diamond Research Fund 2017
- Andrew Kroetsch: University at Buffalo Engineering Alumni Association Scholarship 2016
- Christopher Dundas: NSF-REU 2014
- Christopher Dundas: Goldwater Scholarship, 2013
- Karl Barber: Fulbright Scholarship, 2012
- Kok Hong Lim: Best poster award, Annual Graduate Student Research Symposium 2010

SERVICES

Departmental, School of Engineering, and University Services

- Participated in the School of Engineering Open House (2006 – 2010)
- Served as an organizer of the Annual Graduate Research Symposium (2006 – 2013)
- Served as a judge at the Annual Departmental Poster Competition (2006, 2008, 2012 - present)
- Served on the Safety Committee (2006 – 2010, 2012 - 2014)
- Served as the Fire Evacuation Warden for the 9th floor of Furnas (2008 - 2013)
- Served on the Junior Faculty Panel for "Future Faculty Workshop" organized by Dr. Batta (2008, 2010)
- Served on the UB Graduate School Grievance Committee (2009 – present)
- Served on the Department of Chemical and Biological Graduate Committee (2007 – present)
- Served as a freshman mentor for EAS 202 (2012 – 2014)
- Faculty advisor to the UB student club of the Society for Biological Engineers (2011 – 2013)
- Served on the search committee for the Dean of the School of Engineering and Applied Sciences (2012)
- Faculty advisor to the student chapter of the American Institute of Chemical Engineers (2012 – 2013)
- Furnas Hall, UB, Fire Evacuation Coordinator (2013 – present)
- Served on the Program for Genetics, Genomics and Biophysics Grievance Committee (2013 – present)
- Served as a judge at the Annual School of Engineering Poster Competition (2014 – present)
- **Director of Graduate Studies** (2016 – 2021)
- Graduate Fellowship and Scholarship Committee (2019 – 2023)

Research Committees

PhD Defense Committee

Student Name	Year	Advisor
Indrajeet Singh	2007	Sriram Neelamegham
Yajun Yan	2008	Mattheos Koffas
Chin Giaw Lim	2011	Mattheos Koffas
Kok Hong Lim	2012	Sheldon Park
Jasdeep Mann	2014	Sheldon Park
Sri Madabhushi	2014	Sriram Neelamegham
Nandini Mondal	2014	Sriram Neelamegham
Shobhit Gogia	2014	Sriram Neelamegham
Daniel Demonte	2015	Sheldon Park
Mahmoud Ahmadi	2016	Blaine Pfeifer
Shuen Shiuan Wang	2016	Sriram Neelamegham

Yi Li	2016	Sriram Neelamegham
Lei Fang	2017	Blaine Pfeifer
Marie Beitelshees	2018	Blaine Pfeifer
Nicholas Moscatello	2018	Blaine Pfeifer
Kai Cheng	2018	Sriram Neelamegham
Changjie Chen	2018	Sriram Neelamegham
Yuqi Zhu	2018	Sriram Neelamegham
Andrew Kroetsch	2019	Sheldon Park
James Tran	2020	Haiqing Lin
Ogechi Ogoke	2020	Natesh Parashurama
Sheida Jamalzadeh	2020	Paul Cullen
Yuqi Zhu	2020	Sriram Neelamegham
Dongwon Park	2021	Blaine Pfeifer
Girish Swayambhu	2021	Blaine Pfeifer
Panagiotis Chrysinas	TBA	Rudiyanto Guanawan
Ashritha Mandava	TBA	Sriram Neelamegham
Qi Yang	TBA	Sriram Neelamegham
Arezoo Momeni	2021	Sriram Neelamegham
Kai Cheng	TBA	Sriram Neelamegham
Nika Rajabian	2022	Stelios Andreadis
Ashish Kumar Podder	TBA	Stelios Andreadis
Masoud Zamani	TBA	Stelios Andreadis
Na Rong	2021	Stelios Andreadis
Ryoma Hombh	TBA	Sriram Neelamegham
Thomas Hughes	TBA	Sriram Neelamegham
Debanik Choudury	2024	Stelios Andreadis
Arun Singh	TBA	Sriram Neelamegham

Masters Defense Committee

Student Name	**Year**	**Advisor**
Savatha Prakash	2007	Sriram Neelamegham
Amruta Bedekar	2008	Mattheos Koffas
Namita Bhan	2011	Mattheos Koffas
Lynn Wong	2011	Mattheos Koffas
Sankaranarayanan Venkiteswaran	2011	Mattheos Koffas
Nagarajan Krishna	2012	Emmanuel Tzanakakis
Anne Fleur Stephans	2012	Emmanuel Tzanakakis
Aishwarya Ranganathan	2012	Stelios Andreadis
Randall Smith	2013	Stelios Andreadis

Seoyoung Son	2013	Stelios Andreadis
Jingyuan Gao	2016	Sheldon Park
Aparajita Dasgupta	2016	Sheldon Park
Xiaoyan Wang	2016	Stelios Andreadis
Fangyu Zhou	2017	Sheldon Park
Vyncent Nguyen	2017	Sheldon Park
Brandon Chin	2018	Sheldon Park
Xiao Huang	2018	Sheldon Park
Iyan Warren	2018	Natesh Parashurama
Zach Dickman	2019	Natesh Parashurama
Jill Annis	2019	Rudiyanto Guanawan
Xiaoyang Shu	2020	Natesh Parashurama
Tram Hoang Anh Nguyen	2020	Natesh Parashurama
Ashritha Mandava	2020	Sriram Neelamegham
Samaneh Moghadasi Boroujeni	2020	Stelios Andreadis
William Erdman	2021	Sheldon Park

Undergraduate Academic Advising

Kristen Marra, Theodore Einar Olsen-Tan, Julio Cesar Valenzuela Roc, John Seeley Vallone, Andrew John Westcot, Michael Joseph Wolny, Alexander Philip Anneling, Kathryn Rose Callan, Joshua Urbach, Kevin Carter, Christopher Dundas, Dylan Keefe, James Knox, Hayley Riesel, Keelan Schwab, Lee Terry, Stephen Vitello, Leong Chee Yap, Eunji Youn, Hunter Jones, Zachary Valenza, Yi Chen, Jaehyug Choi, Jermel Griffin, Min Gi Jin, Vivian Luo, Thomas Prato, Zehey Wehbe, Qinqin Xiao, Muhammad Azmi, Justine Cappiello, Nur Hazirah, Ali Jamali, Gino Rosati, Benjamin Skoog, Jinhao Zheng, Yiyang Zhou, David Tran, Andy Lam, Michael Ganbarg, Suting Huang, Xiao Ci Lin, Dylan Servos, Zeen Xing, Vinh Bui, Matan Abittan, Adam Ballard, Jacob Carbone, Fran Duffy, Jason Gomez, Derrick Huang, Daniel Krohn, Scott Reese, Alex Mack, Quan Le, Adam Ballard, Alexander Deangelis, Lamon Friday, Daniel Krohn, Rylee MacDonald, Gianna Pezzino

Professional Services

- Editor for "Protein Engineering and Design", together with Dr. Jennifer Cochran, CRC Press (2009)
- AIChE Meeting, Session Chair, 2006—Systems Biology
- BMES Meeting, Session Chair, 2007—Biomolecular Engineering
- AIChE Meeting, Session Chair, 2007-2010—Food, Pharmaceutical, and Bioengineering Division (15c). Advances in Protein Structure, Function, and Stability
- AIChE Meeting, Session Chair, 2011—Food, Pharmaceutical, and Bioengineering Division (15c). Molecular Modeling of Biophysical Processes I - Molecular Binding and Protein Structure and Dynamics
- AIChE Meeting, Session Chair, 2012—Food, Pharmaceutical, and Bioengineering Division (15c). Molecular Modeling of Biophysical Processes I - Molecular Binding and Protein Structure and Dynamics. Biomaterials.

Journal Reviews

- ACS Chemical Biology
- ACS Medicinal Chemistry
- ACS Nano
- ACS Omega
- ACS Synthetic Biology
- AIChE Journal
- Analytical Chemistry
- Annals of Biomedical Engineering
- Applied Biochemistry and Biotechnology
- Applied Microbiology and Biotechnology
- Artificial Intelligence in Medicine
- Biochemical Engineering Journal
- Biochemistry
- Biochimica et Biophysica Acta
- Bioconjugate Chemistry
- Biophysical Journal
- Biotechnology and Bioengineering
- Biotechnology Journal
- Biotechnology Advances
- Biotechnology Progress
- BMC Bioinformatics
- ChemBioChem
- Chemistry & Biology
- Chemical Biology and Drug Design
- Computational and Structural Biotechnology Journal
- Critical Reviews in Biotechnology
- Electrophoresis
- Enzyme and Microbial Technology

- Integrative Biology
- International Journal of Biological Macromolecules
- Journal of the American Chemical Society
- Journal of Biotechnology
- Journal of Molecular Biology
- Journal of Molecular Modeling
- Journal of Physical Chemistry
- Letters in Drug Design and Discovery
- Micro Cell Factory
- Molecular Biotechnology
- Nature Communications
- Nature Protocol
- New Biotechnology
- Oncotarget
- PNAS
- Protein Science
- Proteins
- PLoS One
- Science Report

Also a reviewer for the following books or book chapters: Wiley Encyclopedia of Chemical Biology, Dictionary of Chemical Engineering, Process Dynamics Modeling and Control.

Grant reviews

- Czech Science Foundation (2013)
- National Science Foundation, CBET (2013, Jan 2015, Feb 2015)
- Kentucky Science and Engineering Foundation (2017)
- Worldwide Cancer Research (2019)
- National Institutes of Health, SBIR, AIDC-C(12) (2021)
- National Institutes of Health, SBIR, AIDC (2022)
- National Institutes of Health, SBIR, ZRG1 MBBC-G (2022)

Professional Development Activities

Workshop

- Conversations on Career Advancement for Tenured Associate Faculty (Apr 2016)
- SEAS Faculty Workshop: Cultural Competence for Faculty in the Academy (Nov 2015)

Professional Conferences

- Annual AIChE meeting (2006 – present)
- Annual Protein Society meeting (2006, 2007, 2008)
- Annual BMES meeting (2007, 2009)

Teaching Related Activities

- A workshop on the use of UBlearns (August 2006)
- Summer School for Chemical Engineering Faculty—organized by the American Society for Engineering Education (ASEE), Chemical Engineering Division (August 2007)

APPENDIX C

Position 324

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	**Hyal1 Residue#**	**Hyal1 Residue**	**PH20 Residue #**	**PH20 Mature Residue #**	**PH20 Residue**	**Alternative Residue**	**Residue %**	**Rating (1-3)**	**Comments**	**#neigh**	**fSASA**	**Factors**		
2												**Hydrophobicity**	**Secondary Structure**	**Interactions**
3	341	D	359	**324**	E	-	12.5			6	0.48			
4	341		359	**324**	-	D	25	2	neutral: lower H propensty, unfavorable hydrophobic interaction, seen in hyal1	7	-	1	1	1
5	341		359	**324**	-	T	13.63	2	gain: increased hydrophobic interaction	7	-	1	1	1
6	341		359	**324**	-	S	12.5	2	neutral: low H propensity, limited interaction	6	-		1	1
7	341		359	**324**	-	V	7.95	2	gain: increased hydrophobic interaction	6	-			1
8	341		359	**324**	-	N	6.81	2	neutral	7	-			
9	341		359	**324**	-	K	6.81	3	gain: increased hydrophobic interaction, charge interaction with D355	6	-			1
10	341		359	**324**	-	R	5.68	3	gain: increased hydrophobic interaction, salt bridge to D355	7	-			1
11	341		359	**324**	-	L	2.27	3	gain: improved hydrophobic contacts	7	-	1	1	1
12	341		359	**324**	-	Q	2.27	2	gain: improved van der Waals	7	-			1
13	341		359	**324**	-	H	2.27	2	gain: improved van der Waals	7	-			1
14	341		359	**324**	-	G	1.13	2	loss: low H propensity	7	-		1	1
15	341		359	**324**	-	A	1.13	2	neutral: avoid charge repulsion	6	-		1	

APPENDIX D

Part 1: Amino Acid Profile

wt 1: M	18.18	-	81.81
res0: -	72	81.81	
res0: M	16	18.18	
wt 2: G	10.22	-	81.81
res1: -	72	81.81	
res1: G	9	10.22	
res1: R	3	3.4	
res1: K	2	2.27	
res1: A	1	1.13	
res1: E	1	1.13	
wt 3: V	12.5	-	80.68
res2: -	71	80.68	
res2: V	11	12.5	
res2: M	2	2.27	
res2: R	1	1.13	
res2: A	1	1.13	
res2: G	1	1.13	
res2: E	1	1.13	
wt 4: L	15.9	-	80.68
res3: -	71	80.68	
res3: L	14	15.9	
res3: V	1	1.13	
res3: F	1	1.13	
res3: Q	1	1.13	
wt 5: K	5.68	-	79.54
res4: -	70	79.54	
res4: R	6	6.81	
res4: K	5	5.68	
res4: S	2	2.27	
res4: C	2	2.27	
res4: P	1	1.13	
res4: N	1	1.13	
res4: T	1	1.13	
wt 6: F	11.36	-	78.4
res5: -	69	78.4	
res5: F	10	11.36	
res5: L	4	4.54	
res5: E	2	2.27	
res5: R	1	1.13	
res5: H	1	1.13	
res5: C	1	1.13	
wt 7: K	11.36	-	78.4
res6: -	69	78.4	
res6: K	10	11.36	
res6: Q	4	4.54	
res6: G	2	2.27	
res6: E	1	1.13	
res6: H	1	1.13	
res6: P	1	1.13	
wt 8: H	19.31	-	78.4
res7: -	69	78.4	
res7: H	17	19.31	
res7: Q	1	1.13	
res7: R	1	1.13	

wt 9: I	11.36	-	78.4
res8: -	69	78.4	
res8: I	10	11.36	
res8: L	5	5.68	
res8: S	3	3.4	
res8: V	1	1.13	
wt 10: F	12.5	-	77.27
res9: -	68	77.27	
res9: F	11	12.5	
res9: S	4	4.54	
res9: R	2	2.27	
res9: L	2	2.27	
res9: H	1	1.13	
wt 11: F	13.63	-	76.13
res10: -	67	76.13	
res10: F	12	13.63	
res10: L	2	2.27	
res10: S	2	2.27	
res10: W	2	2.27	
res10: Y	1	1.13	
res10: R	1	1.13	
res10: C	1	1.13	
wt 12: R	9.09	-	75
res11: -	66	75	
res11: R	8	9.09	
res11: G	7	7.95	
res11: C	3	3.4	
res11: K	2	2.27	
res11: S	1	1.13	
res11: A	1	1.13	
wt 13: S	15.9	-	75
res12: -	66	75	
res12: S	14	15.9	
res12: V	3	3.4	
res12: K	2	2.27	
res12: G	2	2.27	
res12: N	1	1.13	
wt 14: F	13.63	-	73.86
res13: -	65	73.86	
res13: F	12	13.63	
res13: V	3	3.4	
res13: L	3	3.4	
res13: T	1	1.13	
res13: I	1	1.13	
res13: A	1	1.13	
res13: W	1	1.13	
res13: C	1	1.13	

wt 15: V	11.36	-	73.86
res14: -	65	73.86	
res14: V	10	11.36	
res14: Q	4	4.54	
res14: L	3	3.4	
res14: A	2	2.27	
res14: T	1	1.13	
res14: I	1	1.13	
res14: G	1	1.13	
res14: F	1	1.13	
wt 16: K	4.54	-	76.13
res15: -	67	76.13	
res15: E	5	5.68	
res15: K	4	4.54	
res15: P	4	4.54	
res15: G	3	3.4	
res15: V	3	3.4	
res15: S	1	1.13	
res15: R	1	1.13	
wt 17: S	12.5	-	75
res16: -	66	75	
res16: S	11	12.5	
res16: I	4	4.54	
res16: C	2	2.27	
res16: V	2	2.27	
res16: F	1	1.13	
res16: P	1	1.13	
res16: T	1	1.13	
wt 18: S	7.95	-	76.13
res17: -	67	76.13	
res17: S	7	7.95	
res17: N	4	4.54	
res17: H	4	4.54	
res17: R	2	2.27	
res17: G	1	1.13	
res17: C	1	1.13	
res17: T	1	1.13	
res17: Q	1	1.13	
wt 19: G	13.63	-	76.13
res18: -	67	76.13	
res18: G	12	13.63	
res18: L	4	4.54	
res18: R	2	2.27	
res18: Q	2	2.27	
res18: A	1	1.13	
wt 20: V	4.54	-	71.59
res19: -	63	71.59	
res19: A	8	9.09	
res19: T	7	7.95	
res19: V	4	4.54	
res19: L	3	3.4	
res19: M	1	1.13	
res19: I	1	1.13	
res19: S	1	1.13	

wt 21: S	14.77	-	71.59
res20: -	63	71.59	
res20: S	13	14.77	
res20: P	6	6.81	
res20: F	3	3.4	
res20: T	1	1.13	
res20: G	1	1.13	
res20: L	1	1.13	

wt 22: Q	14.77	-	70.45
res21: -	62	70.45	
res21: Q	13	14.77	
res21: W	6	6.81	
res21: G	2	2.27	
res21: C	1	1.13	
res21: L	1	1.13	
res21: F	1	1.13	
res21: P	1	1.13	
res21: V	1	1.13	

wt 23: I	5.68	-	81.81
res22: -	72	81.81	
res22: T	5	5.68	
res22: I	5	5.68	
res22: V	4	4.54	
res22: A	2	2.27	

wt 24: V	14.77	-	79.54
res23: -	70	79.54	
res23: V	13	14.77	
res23: C	2	2.27	
res23: L	2	2.27	
res23: A	1	1.13	

wt 25: F	10.22	-	64.77
res24: -	57	64.77	
res24: L	15	17.04	
res24: F	9	10.22	
res24: I	2	2.27	
res24: T	2	2.27	
res24: S	1	1.13	
res24: V	1	1.13	
res24: Y	1	1.13	

wt 26: T	9.09	-	62.5
res26: -	55	62.5	
res26: I	12	13.63	
res26: V	8	9.09	
res26: T	8	9.09	
res26: A	2	2.27	
res26: L	1	1.13	
res26: S	1	1.13	
res26: C	1	1.13	

wt 27: F	27.27	-	55.68
res27: -	49	55.68	
res27: F	24	27.27	
res27: L	11	12.5	
res27: S	2	2.27	
res27: V	1	1.13	
res27: T	1	1.13	

wt 28: L	23.86	-	56.81
res44: -	50	56.81	
res44: L	21	23.86	
res44: F	9	10.22	
res44: I	2	2.27	
res44: S	1	1.13	
res44: C	1	1.13	
res44: G	1	1.13	
res44: V	1	1.13	
res44: A	1	1.13	
res44: Q	1	1.13	

wt 29: L	23.86	-	57.95
res45: -	51	57.95	
res45: L	21	23.86	
res45: I	7	7.95	
res45: F	5	5.68	
res45: S	1	1.13	
res45: Q	1	1.13	
res45: M	1	1.13	
res45: V	1	1.13	

wt 30: I	15.9	-	54.54
res46: -	48	54.54	
res46: L	20	22.72	
res46: I	14	15.9	
res46: G	2	2.27	
res46: V	2	2.27	
res46: Y	1	1.13	
res46: S	1	1.13	

wt 31: P	15.9	-	54.54
res47: -	48	54.54	
res47: P	14	15.9	
res47: K	8	9.09	
res47: V	3	3.4	
res47: T	3	3.4	
res47: S	3	3.4	
res47: N	2	2.27	
res47: L	2	2.27	
res47: R	1	1.13	
res47: H	1	1.13	
res47: A	1	1.13	
res47: G	1	1.13	
res47: E	1	1.13	

wt 32: C	15.9	-	52.27
res48: -	46	52.27	
res48: C	14	15.9	
res48: S	9	10.22	
res48: L	7	7.95	
res48: T	3	3.4	
res48: D	2	2.27	
res48: A	2	2.27	
res48: R	2	2.27	
res48: F	1	1.13	
res48: W	1	1.13	
res48: K	1	1.13	

wt 33: C	18.18	-	42.04
res49: -	37	42.04	
res49: C	16	18.18	
res49: A	11	12.5	
res49: I	7	7.95	
res49: L	7	7.95	
res49: S	3	3.4	
res49: T	2	2.27	
res49: V	2	2.27	
res49: Q	1	1.13	
res49: P	1	1.13	
res49: G	1	1.13	

wt 34: L	17.04	-	42.04
res50: -	37	42.04	
res50: S	15	17.04	
res50: L	15	17.04	
res50: M	7	7.95	
res50: T	3	3.4	
res50: C	3	3.4	
res50: Q	3	3.4	
res50: D	2	2.27	
res50: G	1	1.13	
res50: A	1	1.13	
res50: W	1	1.13	

wt 35: T	12.5	-	42.04
res51: -	37	42.04	
res51: T	11	12.5	
res51: S	11	12.5	
res51: E	9	10.22	
res51: A	6	6.81	
res51: L	4	4.54	
res51: M	3	3.4	
res51: G	2	2.27	
res51: V	1	1.13	
res51: Q	1	1.13	
res51: W	1	1.13	
res51: C	1	1.13	
res51: P	1	1.13	

wt 36: L	36.36	-	43.18
res52: -	38	43.18	
res52: L	32	36.36	
res52: A	5	5.68	
res52: V	4	4.54	
res52: M	2	2.27	
res52: G	2	2.27	
res52: Q	1	1.13	
res52: N	1	1.13	
res52: K	1	1.13	
res52: T	1	1.13	
res52: E	1	1.13	

wt 37: N	7.95	-	42.04
res53: -	37	42.04	
res53: K	23	26.13	
res53: D	11	12.5	
res53: N	7	7.95	
res53: Q	4	4.54	
res53: H	2	2.27	
res53: E	1	1.13	
res53: C	1	1.13	
res53: M	1	1.13	
res53: A	1	1.13	
wt 38: F	12.5	-	42.04
res54: -	37	42.04	
res54: P	23	26.13	
res54: F	11	12.5	
res54: G	5	5.68	
res54: Y	3	3.4	
res54: I	3	3.4	
res54: N	2	2.27	
res54: V	1	1.13	
res54: A	1	1.13	
res54: K	1	1.13	
res54: M	1	1.13	
wt 39: R	15.9	-	42.04
res55: -	37	42.04	
res55: R	14	15.9	
res55: T	14	15.9	
res55: A	10	11.36	
res55: Q	3	3.4	
res55: S	3	3.4	
res55: F	2	2.27	
res55: K	2	2.27	
res55: I	1	1.13	
res55: G	1	1.13	
res55: P	1	1.13	
wt 40: A	34.09	-	38.63
res56: -	34	38.63	
res56: A	30	34.09	
res56: R	14	15.9	
res56: Q	2	2.27	
res56: G	2	2.27	
res56: M	1	1.13	
res56: T	1	1.13	
res56: P	1	1.13	
res56: N	1	1.13	
res56: W	1	1.13	
res56: F	1	1.13	
wt 41: P	34.09	-	34.09
res57: -	30	34.09	
res57: P	30	34.09	
res57: L	10	11.36	
res57: A	4	4.54	
res57: G	3	3.4	
res57: D	3	3.4	
res57: R	2	2.27	
res57: W	2	2.27	
res57: S	2	2.27	
res57: Q	1	1.13	
res57: H	1	1.13	

wt 42: P	84.09	P	84.09
res58: P	74	84.09	
res58: -	11	12.5	
res58: T	2	2.27	
res58: L	1	1.13	
wt 43: V	25	I	27.27
res59: I	24	27.27	
res59: V	22	25	
res59: L	22	25	
res59: -	8	9.09	
res59: M	3	3.4	
res59: T	2	2.27	
res59: Q	2	2.27	
res59: N	2	2.27	
res59: F	2	2.27	
res59: P	1	1.13	
wt 44: I	23.86	I	23.86
res60: I	21	23.86	
res60: F	18	20.45	
res60: L	15	17.04	
res60: Y	14	15.9	
res60: V	12	13.63	
res60: -	5	5.68	
res60: G	2	2.27	
res60: S	1	1.13	
wt 45: P	39.77	P	39.77
res61: P	35	39.77	
res61: T	16	18.18	
res61: Q	14	15.9	
res61: S	9	10.22	
res61: K	5	5.68	
res61: G	2	2.27	
res61: -	2	2.27	
res61: L	2	2.27	
res61: R	1	1.13	
res61: I	1	1.13	
res61: V	1	1.13	
wt 46: N	36.36	N	36.36
res62: N	32	36.36	
res62: R	22	25	
res62: G	12	13.63	
res62: E	5	5.68	
res62: -	4	4.54	
res62: K	4	4.54	
res62: D	3	3.4	
res62: S	2	2.27	
res62: H	2	2.27	
res62: P	1	1.13	
res62: A	1	1.13	

wt 47: V	11.36	R	36.36
res63: R	32	36.36	
res63: K	23	26.13	
res63: V	10	11.36	
res63: Q	6	6.81	
res63: H	3	3.4	
res63: T	3	3.4	
res63: L	3	3.4	
res63: -	2	2.27	
res63: S	2	2.27	
res63: A	1	1.13	
res63: I	1	1.13	
res63: C	1	1.13	
res63: E	1	1.13	
wt 48: P	88.63	P	88.63
res64: P	78	88.63	
res64: A	2	2.27	
res64: T	2	2.27	
res64: N	2	2.27	
res64: S	2	2.27	
res64: I	1	1.13	
res64: G	1	1.13	
wt 49: F	95.45	F	95.45
res65: F	84	95.45	
res65: L	4	4.54	
wt 50: L	25	I	27.27
res66: I	24	27.27	
res66: L	22	25	
res66: V	17	19.31	
res66: T	9	10.22	
res66: S	5	5.68	
res66: N	4	4.54	
res66: A	3	3.4	
res66: M	2	2.27	
res66: Q	1	1.13	
res66: C	1	1.13	
wt 51: W	14.77	V	30.68
res67: V	27	30.68	
res67: A	18	20.45	
res67: T	16	18.18	
res67: W	13	14.77	
res67: I	5	5.68	
res67: S	4	4.54	
res67: L	3	3.4	
res67: G	1	1.13	
res67: C	1	1.13	
wt 52: A	46.59	A	46.59
res68: A	41	46.59	
res68: V	23	26.13	
res68: I	12	13.63	
res68: L	7	7.95	
res68: F	4	4.54	
res68: G	1	1.13	
wt 53: W	100	W	100
res69: W	88	100	

wt 54: N	85.22		N	85.22
res70: N	75	85.22		
res70: D	9	10.22		
res70: A	4	4.54		
wt 55: A	71.59		A	71.59
res71: A	63	71.59		
res71: V	21	23.86		
res71: I	2	2.27		
res71: G	1	1.13		
res71: T	1	1.13		
wt 56: P	82.95		P	82.95
res72: P	73	82.95		
res72: N	8	9.09		
res72: D	7	7.95		
wt 57: S	15.9		T	79.54
res73: T	70	79.54		
res73: S	14	15.9		
res73: I	2	2.27		
res73: L	1	1.13		
res73: V	1	1.13		
wt 58: E	30.68		E	30.68
res74: E	27	30.68		
res74: Q	25	28.4		
res74: D	15	17.04		
res74: A	8	9.09		
res74: H	3	3.4		
res74: S	3	3.4		
res74: N	2	2.27		
res74: L	2	2.27		
res74: M	1	1.13		
res74: Y	1	1.13		
res74: F	1	1.13		
wt 59: F	9.09		D	19.31
res75: D	17	19.31		
res75: Q	14	15.9		
res75: W	12	13.63		
res75: R	12	13.63		
res75: L	9	10.22		
res75: F	8	9.09		
res75: S	5	5.68		
res75: H	4	4.54		
res75: N	1	1.13		
res75: Y	1	1.13		
res75: P	1	1.13		
res75: K	1	1.13		
res75: T	1	1.13		
res75: A	1	1.13		
res75: E	1	1.13		
wt 60: C	100		C	100
res76: C	88	100		

wt 61: L	40.9		L	40.9
res77: L	36	40.9		
res77: K	11	12.5		
res77: G	9	10.22		
res77: A	8	9.09		
res77: R	5	5.68		
res77: Q	4	4.54		
res77: M	3	3.4		
res77: P	2	2.27		
res77: T	2	2.27		
res77: V	2	2.27		
res77: I	2	2.27		
res77: H	1	1.13		
res77: S	1	1.13		
res77: E	1	1.13		
res77: W	1	1.13		
wt 62: G	9.09		P	21.59
res78: P	19	21.59		
res78: E	12	13.63		
res78: I	11	12.5		
res78: G	8	9.09		
res78: K	6	6.81		
res78: T	5	5.68		
res78: Q	5	5.68		
res78: A	4	4.54		
res78: L	4	4.54		
res78: R	3	3.4		
res78: D	3	3.4		
res78: M	2	2.27		
res78: S	2	2.27		
res78: N	2	2.27		
res78: F	2	2.27		
wt 63: K	31.81		R	50
res80: R	44	50		
res80: K	28	31.81		
res80: N	6	6.81		
res80: Q	2	2.27		
res80: M	1	1.13		
res80: I	1	1.13		
res80: -	1	1.13		
res80: H	1	1.13		
res80: E	1	1.13		
res80: G	1	1.13		
res80: F	1	1.13		
res80: D	1	1.13		
wt 64: F	39.77		F	39.77
res81: F	35	39.77		
res81: Y	26	29.54		
res81: H	17	19.31		
res81: L	6	6.81		
res81: T	2	2.27		
res81: V	1	1.13		
res81: S	1	1.13		

wt 65: D	7.95		N	29.54
res82: N	26	29.54		
res82: G	25	28.4		
res82: K	19	21.59		
res82: D	7	7.95		
res82: S	5	5.68		
res82: Q	3	3.4		
res82: Y	1	1.13		
res82: H	1	1.13		
res82: R	1	1.13		
wt 66: E	6.81		V	65.9
res83: V	58	65.9		
res83: L	8	9.09		
res83: I	6	6.81		
res83: E	6	6.81		
res83: M	5	5.68		
res83: Q	2	2.27		
res83: T	1	1.13		
res83: H	1	1.13		
res83: D	1	1.13		
wt 67: P	27.27		P	27.27
res84: P	24	27.27		
res84: D	20	22.72		
res84: R	8	9.09		
res84: H	7	7.95		
res84: T	5	5.68		
res84: S	5	5.68		
res84: A	4	4.54		
res84: G	4	4.54		
res84: N	4	4.54		
res84: Q	3	3.4		
res84: K	2	2.27		
res84: E	2	2.27		
wt 68: L	63.63		L	63.63
res85: L	56	63.63		
res85: V	15	17.04		
res85: I	8	9.09		
res85: P	3	3.4		
res85: M	3	3.4		
res85: F	2	2.27		
res85: T	1	1.13		
wt 69: D	64.77		D	64.77
res86: D	57	64.77		
res86: N	20	22.72		
res86: P	8	9.09		
res86: Q	2	2.27		
res86: S	1	1.13		
wt 70: M	11.36		L	64.77
res90: L	57	64.77		
res90: V	14	15.9		
res90: M	10	11.36		
res90: F	4	4.54		
res90: I	2	2.27		
res90: T	1	1.13		

wt 71: S	43.18		S	43.18
res91: S	38	43.18		
res91: K	19	21.59		
res91: N	10	11.36		
res91: E	5	5.68		
res91: D	5	5.68		
res91: Q	5	5.68		
res91: R	3	3.4		
res91: G	2	2.27		
res91: A	1	1.13		
wt 72: L	21.59		V	26.13
res93: V	23	26.13		
res93: L	19	21.59		
res93: A	18	20.45		
res93: M	11	12.5		
res93: I	8	9.09		
res93: F	3	3.4		
res93: Q	3	3.4		
res93: S	1	1.13		
res93: D	1	1.13		
res93: T	1	1.13		
wt 73: F	90.9		F	90.9
res94: F	80	90.9		
res94: L	6	6.81		
res94: Y	2	2.27		
wt 74: S	14.77		D	40.9
res95: D	36	40.9		
res95: Q	14	15.9		
res95: S	13	14.77		
res95: G	8	9.09		
res95: N	7	7.95		
res95: P	5	5.68		
res95: H	3	3.4		
res95: T	1	1.13		
res95: E	1	1.13		
wt 75: F	3.4		V	44.31
res96: V	39	44.31		
res96: I	28	31.81		
res96: L	15	17.04		
res96: F	3	3.4		
res96: M	2	2.27		
res96: A	1	1.13		
wt 76: I	26.13		V	39.77
res97: V	35	39.77		
res97: I	23	26.13		
res97: Q	12	13.63		
res97: T	5	5.68		
res97: L	3	3.4		
res97: N	2	2.27		
res97: K	2	2.27		
res97: F	1	1.13		
res97: H	1	1.13		
res97: A	1	1.13		
res97: M	1	1.13		
res97: S	1	1.13		
res97: E	1	1.13		

wt 77: G	36.36		A	43.18
res98: A	38	43.18		
res98: G	32	36.36		
res98: V	3	3.4		
res98: P	3	3.4		
res98: H	3	3.4		
res98: T	3	3.4		
res98: S	3	3.4		
res98: Q	1	1.13		
res98: L	1	1.13		
res98: E	1	1.13		
wt 78: S	53.4		S	53.4
res99: S	47	53.4		
res99: N	36	40.9		
res99: T	5	5.68		
wt 79: P	77.27		P	77.27
res100: P	68	77.27		
res100: T	4	4.54		
res100: H	4	4.54		
res100: K	3	3.4		
res100: R	3	3.4		
res100: Q	2	2.27		
res100: C	1	1.13		
res100: D	1	1.13		
res100: L	1	1.13		
res100: S	1	1.13		
wt 80: R	15.9		N	30.68
res101: N	27	30.68		
res101: G	17	19.31		
res101: L	14	15.9		
res101: R	14	15.9		
res101: F	4	4.54		
res101: E	4	4.54		
res101: Q	4	4.54		
res101: S	1	1.13		
res101: H	1	1.13		
res101: K	1	1.13		
res101: A	1	1.13		
wt 81: I	4.54		E	32.95
res102: E	29	32.95		
res102: Q	24	27.27		
res102: A	13	14.77		
res102: K	10	11.36		
res102: I	4	4.54		
res102: L	2	2.27		
res102: G	1	1.13		
res102: S	1	1.13		
res102: T	1	1.13		
res102: R	1	1.13		
res102: H	1	1.13		
res102: N	1	1.13		

wt 82: N	6.81		T	27.27
res103: T	24	27.27		
res103: G	18	20.45		
res103: K	12	13.63		
res103: S	10	11.36		
res103: R	7	7.95		
res103: N	6	6.81		
res103: D	3	3.4		
res103: H	3	3.4		
res103: E	2	2.27		
res103: A	2	2.27		
res103: -	1	1.13		
wt 83: A	21.59		F	50
res104: F	44	50		
res104: A	19	21.59		
res104: L	5	5.68		
res104: V	5	5.68		
res104: Q	4	4.54		
res104: I	3	3.4		
res104: D	2	2.27		
res104: S	2	2.27		
res104: W	1	1.13		
res104: Y	1	1.13		
res104: K	1	1.13		
res104: T	1	1.13		
wt 84: T	18.18		R	26.13
res105: R	23	26.13		
res105: V	18	20.45		
res105: T	16	18.18		
res105: S	7	7.95		
res105: H	7	7.95		
res105: Q	5	5.68		
res105: I	4	4.54		
res105: M	3	3.4		
res105: F	2	2.27		
res105: A	1	1.13		
res105: K	1	1.13		
res105: N	1	1.13		
wt 85: G	75		G	75
res106: G	66	75		
res106: N	13	14.77		
res106: D	5	5.68		
res106: A	2	2.27		
res106: R	1	1.13		
res106: K	1	1.13		
wt 86: Q	69.31		Q	69.31
res107: Q	61	69.31		
res107: P	16	18.18		
res107: S	7	7.95		
res107: R	1	1.13		
res107: N	1	1.13		
res107: H	1	1.13		
res107: D	1	1.13		

wt 87: G	5.68	N	62.5
res108: N	55	62.5	
res108: D	8	9.09	
res108: G	5	5.68	
res108: P	5	5.68	
res108: T	4	4.54	
res108: S	3	3.4	
res108: E	2	2.27	
res108: F	2	2.27	
res108: K	2	2.27	
res108: I	1	1.13	
res108: Q	1	1.13	
wt 88: V	34.09	I	37.5
res109: I	33	37.5	
res109: V	30	34.09	
res109: M	15	17.04	
res109: L	10	11.36	
wt 89: T	88.63	T	88.63
res110: T	78	88.63	
res110: V	4	4.54	
res110: S	3	3.4	
res110: A	3	3.4	
wt 90: I	84.09	I	84.09
res111: I	74	84.09	
res111: L	11	12.5	
res111: V	2	2.27	
res111: T	1	1.13	
wt 91: F	98.86	F	98.86
res112: F	87	98.86	
res112: Y	1	1.13	
wt 92: Y	98.86	Y	98.86
res113: Y	87	98.86	
res113: F	1	1.13	
wt 93: V	21.59	V	21.59
res114: V	19	21.59	
res114: S	16	18.18	
res114: K	14	15.9	
res114: R	11	12.5	
res114: A	10	11.36	
res114: P	6	6.81	
res114: T	4	4.54	
res114: H	4	4.54	
res114: Q	2	2.27	
res114: Y	2	2.27	
wt 94: D	35.22	D	35.22
res115: D	31	35.22	
res115: N	25	28.4	
res115: S	13	14.77	
res115: E	9	10.22	
res115: T	3	3.4	
res115: R	2	2.27	
res115: K	2	2.27	
res115: Y	1	1.13	
res115: A	1	1.13	
res115: W	1	1.13	

wt 95: R	62.5	R	62.5
res116: R	55	62.5	
res116: Q	16	18.18	
res116: K	7	7.95	
res116: E	6	6.81	
res116: H	3	3.4	
res116: L	1	1.13	
wt 96: L	94.31	L	94.31
res117: L	83	94.31	
res117: F	3	3.4	
res117: I	2	2.27	
wt 97: G	100	G	100
res118: G	88	100	
wt 98: Y	40.9	Y	40.9
res119: Y	36	40.9	
res119: L	26	29.54	
res119: T	13	14.77	
res119: M	5	5.68	
res119: F	4	4.54	
res119: H	2	2.27	
res119: S	1	1.13	
res119: K	1	1.13	
wt 99: Y	97.72	Y	97.72
res120: Y	86	97.72	
res120: F	2	2.27	
wt 100: P	100	P	100
res121: P	88	100	
wt 101: Y	53.4	Y	53.4
res122: Y	47	53.4	
res122: H	15	17.04	
res122: W	14	15.9	
res122: R	7	7.95	
res122: S	3	3.4	
res122: K	2	2.27	
wt 102: I	19.31	Y	46.59
res123: Y	41	46.59	
res123: F	22	25	
res123: I	17	19.31	
res123: L	4	4.54	
res123: V	2	2.27	
res123: H	1	1.13	
res123: K	1	1.13	
wt 103: D	29.54	T	35.22
res124: T	31	35.22	
res124: D	26	29.54	
res124: N	13	14.77	
res124: G	8	9.09	
res124: S	8	9.09	
res124: E	1	1.13	
res124: K	1	1.13	

wt 104: S	35.22	S	35.22
res125: S	31	35.22	
res125: P	20	22.72	
res125: E	16	18.18	
res125: A	5	5.68	
res125: D	4	4.54	
res125: T	3	3.4	
res125: K	2	2.27	
res125: Q	2	2.27	
res125: L	2	2.27	
res125: R	1	1.13	
res125: H	1	1.13	
res125: N	1	1.13	
wt 105: I	4.54	Q	17.04
res126: Q	15	17.04	
res126: T	11	12.5	
res126: A	11	12.5	
res126: R	8	9.09	
res126: D	7	7.95	
res126: S	7	7.95	
res126: H	6	6.81	
res126: E	5	5.68	
res126: V	4	4.54	
res126: N	4	4.54	
res126: I	4	4.54	
res126: K	4	4.54	
res126: G	1	1.13	
res126: M	1	1.13	
wt 106: T	12.5	-	84.09
res127: -	74	84.09	
res127: T	11	12.5	
res127: N	2	2.27	
res127: S	1	1.13	
wt 107: G	79.54	G	79.54
res128: G	70	79.54	
res128: L	4	4.54	
res128: E	3	3.4	
res128: N	2	2.27	
res128: D	2	2.27	
res128: Q	1	1.13	
res128: W	1	1.13	
res128: M	1	1.13	
res128: K	1	1.13	
res128: H	1	1.13	
res128: R	1	1.13	
res128: S	1	1.13	
wt 108: V	26.13	V	26.13
res129: V	23	26.13	
res129: E	18	20.45	
res129: K	11	12.5	
res129: T	11	12.5	
res129: R	10	11.36	
res129: A	5	5.68	
res129: D	3	3.4	
res129: M	2	2.27	
res129: I	2	2.27	
res129: S	1	1.13	
res129: Q	1	1.13	
res129: L	1	1.13	

wt 109: T	5.68		P	47.72
res130: P	42	47.72		
res130: S	19	21.59		
res130: A	11	12.5		
res130: T	5	5.68		
res130: I	3	3.4		
res130: E	3	3.4		
res130: N	2	2.27		
res130: H	1	1.13		
res130: V	1	1.13		
res130: K	1	1.13		
wt 110: V	54.54		V	54.54
res131: V	48	54.54		
res131: I	21	23.86		
res131: H	9	10.22		
res131: F	5	5.68		
res131: L	1	1.13		
res131: S	1	1.13		
res131: Y	1	1.13		
res131: K	1	1.13		
res131: T	1	1.13		
wt 111: N	52.27		N	52.27
res132: N	46	52.27		
res132: H	18	20.45		
res132: F	15	17.04		
res132: Y	6	6.81		
res132: S	1	1.13		
res132: A	1	1.13		
res132: K	1	1.13		
wt 112: G	98.86		G	98.86
res133: G	87	98.86		
res133: T	1	1.13		
wt 113: G	100		G	100
res134: G	88	100		
wt 114: I	32.95		L	46.59
res135: L	41	46.59		
res135: I	29	32.95		
res135: V	18	20.45		
wt 115: P	100		P	100
res136: P	88	100		
wt 116: Q	100		Q	100
res137: Q	88	100		
wt 117: K	9.09		N	61.36
res138: N	54	61.36		
res138: L	11	12.5		
res138: K	8	9.09		
res138: A	7	7.95		
res138: R	3	3.4		
res138: V	2	2.27		
res138: Q	1	1.13		
res138: S	1	1.13		
res138: E	1	1.13		
wt 118: I	13.63		A	26.13
res139: A	23	26.13		
res139: G	18	20.45		
res139: V	13	14.77		
res139: I	12	13.63		
res139: S	6	6.81		
res139: T	5	5.68		
res139: Q	3	3.4		
res139: C	2	2.27		
res139: E	2	2.27		
res139: N	1	1.13		
res139: L	1	1.13		
res139: F	1	1.13		
res139: M	1	1.13		
wt 119: S	77.27		S	77.27
res140: S	68	77.27		
res140: N	11	12.5		
res140: P	6	6.81		
res140: Y	1	1.13		
res140: L	1	1.13		
res140: D	1	1.13		
wt 120: L	90.9		L	90.9
res141: L	80	90.9		
res141: I	4	4.54		
res141: M	2	2.27		
res141: R	1	1.13		
res141: Y	1	1.13		
wt 121: Q	25		Q	25
res142: Q	22	25		
res142: D	11	12.5		
res142: K	9	10.22		
res142: S	8	9.09		
res142: W	8	9.09		
res142: I	7	7.95		
res142: A	5	5.68		
res142: R	4	4.54		
res142: E	3	3.4		
res142: V	3	3.4		
res142: T	2	2.27		
res142: G	2	2.27		
res142: N	2	2.27		
res142: M	1	1.13		
res142: C	1	1.13		
wt 122: D	5.68		A	23.86
res143: A	21	23.86		
res143: V	15	17.04		
res143: E	11	12.5		
res143: T	8	9.09		
res143: K	7	7.95		
res143: D	5	5.68		
res143: R	4	4.54		
res143: Q	3	3.4		
res143: S	3	3.4		
res143: H	3	3.4		
res143: C	2	2.27		
res143: F	2	2.27		
res143: N	2	2.27		
res143: L	1	1.13		
res143: M	1	1.13		
wt 123: H	98.86		H	98.86
res144: H	87	98.86		
res144: S	1	1.13		
wt 124: L	93.18		L	93.18
res145: L	82	93.18		
res145: R	4	4.54		
res145: I	1	1.13		
res145: Y	1	1.13		
wt 125: D	12.5		E	26.13
res146: E	23	26.13		
res146: A	23	26.13		
res146: K	13	14.77		
res146: D	11	12.5		
res146: R	5	5.68		
res146: N	3	3.4		
res146: G	2	2.27		
res146: S	2	2.27		
res146: I	2	2.27		
res146: T	1	1.13		
res146: P	1	1.13		
res146: F	1	1.13		
res146: Q	1	1.13		
wt 126: K	55.68		K	55.68
res147: K	49	55.68		
res147: R	17	19.31		
res147: M	10	11.36		
res147: H	3	3.4		
res147: Q	3	3.4		
res147: L	2	2.27		
res147: C	1	1.13		
res147: V	1	1.13		
res147: I	1	1.13		
res147: Y	1	1.13		
wt 127: A	53.4		A	53.4
res148: A	47	53.4		
res148: L	16	18.18		
res148: T	11	12.5		
res148: S	8	9.09		
res148: V	3	3.4		
res148: M	2	2.27		
res148: C	1	1.13		
wt 128: K	21.59		K	21.59
res149: K	19	21.59		
res149: F	14	15.9		
res149: A	13	14.77		
res149: D	11	12.5		
res149: R	8	9.09		
res149: Q	7	7.95		
res149: P	5	5.68		
res149: E	3	3.4		
res149: Y	2	2.27		
res149: S	2	2.27		
res149: N	1	1.13		
res149: V	1	1.13		
res149: G	1	1.13		
res149: T	1	1.13		

wt 129: K	13.63	Q	31.81	
res150: Q	28	31.81		
res150: E	14	15.9		
res150: K	12	13.63		
res150: D	6	6.81		
res150: G	6	6.81		
res150: H	6	6.81		
res150: Y	3	3.4		
res150: S	3	3.4		
res150: L	2	2.27		
res150: V	2	2.27		
res150: N	2	2.27		
res150: T	2	2.27		
res150: M	1	1.13		
res150: A	1	1.13		
wt 130: D	70.45	D	70.45	
res151: D	62	70.45		
res151: Q	6	6.81		
res151: G	5	5.68		
res151: R	4	4.54		
res151: H	4	4.54		
res151: E	2	2.27		
res151: N	2	2.27		
res151: S	1	1.13		
res151: A	1	1.13		
res151: P	1	1.13		
wt 131: I	79.54	I	79.54	
res152: I	70	79.54		
res152: V	14	15.9		
res152: L	3	3.4		
res152: F	1	1.13		
wt 132: T	3.4	L	19.31	
res153: L	17	19.31		
res153: E	14	15.9		
res153: N	14	15.9		
res153: R	7	7.95		
res153: K	7	7.95		
res153: A	6	6.81		
res153: Q	6	6.81		
res153: H	4	4.54		
res153: D	4	4.54		
res153: T	3	3.4		
res153: G	3	3.4		
res153: M	1	1.13		
res153: S	1	1.13		
res153: V	1	1.13		

wt 133: F	6.81	H	26.13	
res154: H	23	26.13		
res154: Y	20	22.72		
res154: A	15	17.04		
res154: K	9	10.22		
res154: F	6	6.81		
res154: R	4	4.54		
res154: E	3	3.4		
res154: D	2	2.27		
res154: N	2	2.27		
res154: T	1	1.13		
res154: L	1	1.13		
res154: Q	1	1.13		
res154: V	1	1.13		
wt 134: Y	57.95	Y	57.95	
res155: Y	51	57.95		
res155: A	14	15.9		
res155: S	6	6.81		
res155: T	4	4.54		
res155: L	3	3.4		
res155: F	3	3.4		
res155: N	2	2.27		
res155: R	2	2.27		
res155: V	1	1.13		
res155: K	1	1.13		
res155: G	1	1.13		
wt 135: M	20.45	I	62.5	
res156: I	55	62.5		
res156: M	18	20.45		
res156: L	11	12.5		
res156: V	4	4.54		
wt 136: P	59.09	P	59.09	
res157: P	52	59.09		
res157: R	21	23.86		
res157: T	4	4.54		
res157: S	3	3.4		
res157: Q	3	3.4		
res157: G	2	2.27		
res157: W	1	1.13		
res157: K	1	1.13		
res157: H	1	1.13		
wt 137: V	6.81	S	18.18	
res158: S	16	18.18		
res158: A	16	18.18		
res158: T	13	14.77		
res158: -	8	9.09		
res158: D	7	7.95		
res158: V	6	6.81		
res158: I	5	5.68		
res158: E	5	5.68		
res158: L	2	2.27		
res158: W	2	2.27		
res158: N	2	2.27		
res158: M	2	2.27		
res158: H	2	2.27		
res158: Y	1	1.13		
res158: K	1	1.13		

wt 138: D	50	D	50	
res160: D	44	50		
res160: E	14	15.9		
res160: G	6	6.81		
res160: R	5	5.68		
res160: N	5	5.68		
res160: S	4	4.54		
res160: K	3	3.4		
res160: T	3	3.4		
res160: A	2	2.27		
res160: W	1	1.13		
res160: I	1	1.13		
wt 139: N	9.09	F	50	
res161: F	44	50		
res161: S	14	15.9		
res161: N	8	9.09		
res161: K	7	7.95		
res161: P	6	6.81		
res161: Y	3	3.4		
res161: Q	3	3.4		
res161: V	1	1.13		
res161: R	1	1.13		
res161: A	1	1.13		
wt 140: L	6.81	S	31.81	
res162: S	28	31.81		
res162: A	13	14.77		
res162: E	8	9.09		
res162: V	8	9.09		
res162: L	6	6.81		
res162: R	4	4.54		
res162: T	4	4.54		
res162: N	4	4.54		
res162: Q	3	3.4		
res162: H	2	2.27		
res162: K	2	2.27		
res162: D	2	2.27		
res162: P	1	1.13		
res162: W	1	1.13		
res162: F	1	1.13		
res162: I	1	1.13		
wt 141: G	98.86	G	98.86	
res163: G	87	98.86		
res163: E	1	1.13		
wt 142: M	4.54	L	92.04	
res164: L	81	92.04		
res164: M	4	4.54		
res164: I	1	1.13		
res164: V	1	1.13		
res164: P	1	1.13		
wt 143: A	88.63	A	88.63	
res165: A	78	88.63		
res165: G	6	6.81		
res165: V	3	3.4		
res165: S	1	1.13		

wt 144: V	95.45		V	95.45
res166: V	84	95.45		
res166: I	3	3.4		
res166: M	1	1.13		
wt 145: I	86.36		I	86.36
res167: I	76	86.36		
res167: L	7	7.95		
res167: V	4	4.54		
res167: M	1	1.13		
wt 146: D	98.86		D	98.86
res168: D	87	98.86		
res168: E	1	1.13		
wt 147: W	100		W	100
res169: W	88	100		
wt 148: E	100		E	100
res170: E	88	100		
wt 149: E	27.27		E	27.27
res171: E	24	27.27		
res171: A	15	17.04		
res171: Y	13	14.77		
res171: D	13	14.77		
res171: N	11	12.5		
res171: S	4	4.54		
res171: G	3	3.4		
res171: K	2	2.27		
res171: H	2	2.27		
res171: F	1	1.13		
wt 150: W	100		W	100
res172: W	88	100		
wt 151: R	88.63		R	88.63
res173: R	78	88.63		
res173: C	3	3.4		
res173: K	2	2.27		
res173: Y	2	2.27		
res173: E	1	1.13		
res173: S	1	1.13		
res173: L	1	1.13		
wt 152: P	100		P	100
res174: P	88	100		
wt 153: T	14.77		Q	27.27
res175: Q	24	27.27		
res175: L	15	17.04		
res175: T	13	14.77		
res175: V	13	14.77		
res175: R	13	14.77		
res175: I	8	9.09		
res175: E	1	1.13		
res175: K	1	1.13		

wt 154: W	97.72		W	97.72
res176: W	86	97.72		
res176: F	1	1.13		
res176: Y	1	1.13		
wt 155: A	43.18		A	43.18
res177: A	38	43.18		
res177: V	16	18.18		
res177: D	9	10.22		
res177: I	8	9.09		
res177: E	5	5.68		
res177: S	4	4.54		
res177: M	3	3.4		
res177: L	2	2.27		
res177: T	2	2.27		
res177: K	1	1.13		
wt 156: R	75		R	75
res178: R	66	75		
res178: F	13	14.77		
res178: G	6	6.81		
res178: W	1	1.13		
res178: L	1	1.13		
res178: C	1	1.13		
wt 157: N	98.86		N	98.86
res179: N	87	98.86		
res179: A	1	1.13		
wt 158: W	96.59		W	96.59
res181: W	85	96.59		
res181: F	2	2.27		
res181: R	1	1.13		
wt 159: K	18.18		G	22.72
res182: G	20	22.72		
res182: D	19	21.59		
res182: K	16	18.18		
res182: Q	14	15.9		
res182: N	14	15.9		
res182: A	2	2.27		
res182: R	1	1.13		
res182: H	1	1.13		
res182: E	1	1.13		
wt 160: P	21.59		P	21.59
res183: P	19	21.59		
res183: T	19	21.59		
res183: S	18	20.45		
res183: D	8	9.09		
res183: A	7	7.95		
res183: R	6	6.81		
res183: N	3	3.4		
res183: E	3	3.4		
res183: K	2	2.27		
res183: L	1	1.13		
res183: F	1	1.13		
res183: G	1	1.13		

wt 161: K	84.09		K	84.09
res185: K	74	84.09		
res185: R	11	12.5		
res185: M	2	2.27		
res185: G	1	1.13		
wt 162: D	67.04		D	67.04
res186: D	59	67.04		
res186: R	4	4.54		
res186: N	4	4.54		
res186: I	4	4.54		
res186: Q	4	4.54		
res186: A	3	3.4		
res186: L	3	3.4		
res186: K	2	2.27		
res186: E	2	2.27		
res186: H	1	1.13		
res186: M	1	1.13		
res186: T	1	1.13		
wt 163: V	47.72		V	47.72
res187: V	42	47.72		
res187: I	42	47.72		
res187: A	3	3.4		
res187: N	1	1.13		
wt 164: Y	100		Y	100
res188: Y	88	100		
wt 165: K	15.9		R	68.18
res189: R	60	68.18		
res189: K	14	15.9		
res189: Q	11	12.5		
res189: W	1	1.13		
res189: L	1	1.13		
res189: M	1	1.13		
wt 166: N	20.45		Q	34.09
res190: Q	30	34.09		
res190: N	18	20.45		
res190: R	14	15.9		
res190: K	6	6.81		
res190: E	5	5.68		
res190: A	5	5.68		
res190: D	3	3.4		
res190: I	2	2.27		
res190: H	2	2.27		
res190: L	1	1.13		
res190: T	1	1.13		
res190: M	1	1.13		

wt 167: R	23.86		K	27.27
res191: K	24	27.27		
res191: R	21	23.86		
res191: A	11	12.5		
res191: L	7	7.95		
res191: S	5	5.68		
res191: H	5	5.68		
res191: Q	4	4.54		
res191: N	4	4.54		
res191: V	3	3.4		
res191: E	2	2.27		
res191: G	1	1.13		
res191: W	1	1.13		
wt 168: S	100		S	100
res192: S	88	100		
wt 169: I	22.72		R	46.59
res193: R	41	46.59		
res193: I	20	22.72		
res193: L	6	6.81		
res193: K	5	5.68		
res193: W	5	5.68		
res193: Q	4	4.54		
res193: M	2	2.27		
res193: V	2	2.27		
res193: E	2	2.27		
res193: C	1	1.13		
wt 170: E	35.22		E	35.22
res194: E	31	35.22		
res194: A	23	26.13		
res194: Q	12	13.63		
res194: K	11	12.5		
res194: T	3	3.4		
res194: N	3	3.4		
res194: R	2	2.27		
res194: H	1	1.13		
res194: V	1	1.13		
res194: D	1	1.13		
wt 171: L	79.54		L	79.54
res195: L	70	79.54		
res195: W	7	7.95		
res195: F	7	7.95		
res195: Q	1	1.13		
res195: I	1	1.13		
res195: K	1	1.13		
res195: Y	1	1.13		
wt 172: V	63.63		V	63.63
res196: V	56	63.63		
res196: I	17	19.31		
res196: A	8	9.09		
res196: T	7	7.95		

wt 173: Q	36.36		Q	36.36
res197: Q	32	36.36		
res197: R	15	17.04		
res197: S	15	17.04		
res197: A	12	13.63		
res197: K	8	9.09		
res197: L	4	4.54		
res197: E	1	1.13		
res197: F	1	1.13		
wt 174: Q	25		Q	25
res198: Q	22	25		
res198: S	15	17.04		
res198: A	13	14.77		
res198: D	9	10.22		
res198: E	9	10.22		
res198: K	8	9.09		
res198: H	3	3.4		
res198: N	2	2.27		
res198: W	2	2.27		
res198: R	2	2.27		
res198: I	1	1.13		
res198: L	1	1.13		
res198: G	1	1.13		
wt 175: Q	22.72		Q	22.72
res199: Q	20	22.72		
res199: R	19	21.59		
res199: M	13	14.77		
res199: K	8	9.09		
res199: H	7	7.95		
res199: E	6	6.81		
res199: T	4	4.54		
res199: V	4	4.54		
res199: L	4	4.54		
res199: A	1	1.13		
res199: N	1	1.13		
res199: F	1	1.13		
wt 176: N	17.04		H	50
res200: H	44	50		
res200: N	15	17.04		
res200: Q	10	11.36		
res200: F	5	5.68		
res200: G	4	4.54		
res200: D	4	4.54		
res200: Y	3	3.4		
res200: K	2	2.27		
res200: L	1	1.13		
wt 177: V	6.81		P	63.63
res201: P	56	63.63		
res201: E	7	7.95		
res201: V	6	6.81		
res201: I	5	5.68		
res201: R	4	4.54		
res201: K	3	3.4		
res201: Q	2	2.27		
res201: A	2	2.27		
res201: S	1	1.13		
res201: L	1	1.13		
res201: T	1	1.13		

wt 178: Q	11.36		D	35.22
res202: D	31	35.22		
res202: N	18	20.45		
res202: Q	10	11.36		
res202: T	9	10.22		
res202: H	4	4.54		
res202: G	4	4.54		
res202: S	4	4.54		
res202: Y	2	2.27		
res202: E	2	2.27		
res202: F	2	2.27		
res202: -	1	1.13		
res202: K	1	1.13		
wt 179: L	23.86		W	50
res203: W	44	50		
res203: L	21	23.86		
res203: V	10	11.36		
res203: I	6	6.81		
res203: M	3	3.4		
res203: F	1	1.13		
res203: -	1	1.13		
res203: A	1	1.13		
res203: Y	1	1.13		
wt 180: S	37.5		S	37.5
res204: S	33	37.5		
res204: P	31	35.22		
res204: N	7	7.95		
res204: T	6	6.81		
res204: D	5	5.68		
res204: A	2	2.27		
res204: -	1	1.13		
res204: Q	1	1.13		
res204: K	1	1.13		
res204: L	1	1.13		
wt 181: L	10.22		P	23.86
res205: P	21	23.86		
res205: A	20	22.72		
res205: E	13	14.77		
res205: L	9	10.22		
res205: I	3	3.4		
res205: H	3	3.4		
res205: S	3	3.4		
res205: V	2	2.27		
res205: Q	2	2.27		
res205: D	2	2.27		
res205: F	2	2.27		
res205: R	2	2.27		
res205: T	2	2.27		
res205: K	2	2.27		
res205: -	1	1.13		
res205: M	1	1.13		

wt 182: T	22.72		T	22.72
res206: T	20	22.72		
res206: D	14	15.9		
res206: P	10	11.36		
res206: E	10	11.36		
res206: A	9	10.22		
res206: Q	7	7.95		
res206: S	6	6.81		
res206: K	5	5.68		
res206: N	4	4.54		
res206: M	1	1.13		
res206: -	1	1.13		
res206: W	1	1.13		
wt 183: E	23.86		E	23.86
res207: E	21	23.86		
res207: Q	18	20.45		
res207: D	13	14.77		
res207: R	12	13.63		
res207: K	12	13.63		
res207: A	4	4.54		
res207: L	3	3.4		
res207: W	2	2.27		
res207: S	1	1.13		
res207: Y	1	1.13		
res207: -	1	1.13		
wt 184: A	20.45		V	43.18
res208: V	38	43.18		
res208: I	21	23.86		
res208: A	18	20.45		
res208: Q	6	6.81		
res208: L	3	3.4		
res208: S	1	1.13		
res208: G	1	1.13		
wt 185: T	13.63		E	37.5
res209: E	33	37.5		
res209: V	12	13.63		
res209: T	12	13.63		
res209: K	7	7.95		
res209: L	6	6.81		
res209: R	5	5.68		
res209: D	3	3.4		
res209: A	3	3.4		
res209: N	2	2.27		
res209: I	1	1.13		
res209: W	1	1.13		
res209: H	1	1.13		
res209: G	1	1.13		
res209: S	1	1.13		

wt 186: E	7.95		K	30.68
res210: K	27	30.68		
res210: A	13	14.77		
res210: Y	12	13.63		
res210: T	8	9.09		
res210: E	7	7.95		
res210: H	6	6.81		
res210: R	3	3.4		
res210: Q	3	3.4		
res210: D	3	3.4		
res210: N	2	2.27		
res210: S	1	1.13		
res210: G	1	1.13		
res210: L	1	1.13		
res210: C	1	1.13		
wt 187: K	17.04		L	19.31
res211: L	17	19.31		
res211: K	15	17.04		
res211: V	13	14.77		
res211: Q	11	12.5		
res211: E	11	12.5		
res211: A	8	9.09		
res211: I	6	6.81		
res211: R	3	3.4		
res211: S	2	2.27		
res211: M	1	1.13		
res211: T	1	1.13		
wt 188: A	100		A	100
res212: A	88	100		
wt 189: K	45.45		K	45.45
res213: K	40	45.45		
res213: Q	30	34.09		
res213: R	7	7.95		
res213: V	4	4.54		
res213: L	2	2.27		
res213: I	1	1.13		
res213: Y	1	1.13		
res213: A	1	1.13		
res213: H	1	1.13		
res213: E	1	1.13		
wt 190: Q	13.63		Y	15.9
res214: Y	14	15.9		
res214: D	13	14.77		
res214: A	13	14.77		
res214: Q	12	13.63		
res214: V	8	9.09		
res214: E	6	6.81		
res214: T	5	5.68		
res214: K	4	4.54		
res214: R	4	4.54		
res214: I	3	3.4		
res214: F	3	3.4		
res214: N	1	1.13		
res214: L	1	1.13		
res214: M	1	1.13		

wt 191: E	42.04		E	42.04
res215: E	37	42.04		
res215: Q	19	21.59		
res215: T	12	13.63		
res215: D	8	9.09		
res215: G	6	6.81		
res215: S	3	3.4		
res215: A	1	1.13		
res215: V	1	1.13		
res215: R	1	1.13		
wt 192: F	100		F	100
res216: F	88	100		
wt 193: E	81.81		E	81.81
res217: E	72	81.81		
res217: Q	14	15.9		
res217: N	1	1.13		
res217: D	1	1.13		
wt 194: K	18.18		E	22.72
res218: E	20	22.72		
res218: K	16	18.18		
res218: F	11	12.5		
res218: N	9	10.22		
res218: Q	7	7.95		
res218: G	7	7.95		
res218: S	5	5.68		
res218: T	4	4.54		
res218: R	4	4.54		
res218: L	2	2.27		
res218: C	1	1.13		
res218: A	1	1.13		
res218: M	1	1.13		
wt 195: A	75		A	75
res219: A	66	75		
res219: S	19	21.59		
res219: C	1	1.13		
res219: E	1	1.13		
res219: T	1	1.13		
wt 196: G	21.59		A	77.27
res220: A	68	77.27		
res220: G	19	21.59		
res220: S	1	1.13		
wt 197: K	36.36		R	44.31
res221: R	39	44.31		
res221: K	32	36.36		
res221: Q	12	13.63		
res221: E	2	2.27		
res221: M	1	1.13		
res221: S	1	1.13		
res221: C	1	1.13		

wt 198: D	5.68		A	40.9
res222: A	36	40.9		
res222: Q	13	14.77		
res222: S	9	10.22		
res222: N	8	9.09		
res222: D	5	5.68		
res222: E	4	4.54		
res222: K	3	3.4		
res222: H	2	2.27		
res222: T	2	2.27		
res222: C	2	2.27		
res222: R	1	1.13		
res222: L	1	1.13		
res222: V	1	1.13		
res222: I	1	1.13		
wt 199: F	71.59		F	71.59
res223: F	63	71.59		
res223: W	14	15.9		
res223: L	10	11.36		
res223: Y	1	1.13		
wt 200: L	2.27		M	97.72
res224: M	86	97.72		
res224: L	2	2.27		
wt 201: V	6.81		L	22.72
res225: L	20	22.72		
res225: E	17	19.31		
res225: K	12	13.63		
res225: A	12	13.63		
res225: N	8	9.09		
res225: V	6	6.81		
res225: Q	6	6.81		
res225: T	4	4.54		
res225: M	1	1.13		
res225: I	1	1.13		
res225: S	1	1.13		
wt 202: E	50		E	50
res226: E	44	50		
res226: G	15	17.04		
res226: D	6	6.81		
res226: Q	5	5.68		
res226: K	4	4.54		
res226: S	3	3.4		
res226: T	2	2.27		
res226: V	2	2.27		
res226: I	2	2.27		
res226: H	1	1.13		
res226: R	1	1.13		
res226: L	1	1.13		
res226: N	1	1.13		
res226: Y	1	1.13		
wt 203: T	98.86		T	98.86
res227: T	87	98.86		
res227: S	1	1.13		
wt 204: I	21.59		L	78.4
res228: L	69	78.4		
res228: I	19	21.59		
wt 205: K	35.22		K	35.22
res229: K	31	35.22		
res229: R	22	25		
res229: Q	13	14.77		
res229: E	7	7.95		
res229: T	4	4.54		
res229: S	2	2.27		
res229: A	2	2.27		
res229: H	2	2.27		
res229: W	2	2.27		
res229: L	2	2.27		
res229: N	1	1.13		
wt 206: L	78.4		L	78.4
res230: L	69	78.4		
res230: Y	9	10.22		
res230: F	3	3.4		
res230: V	2	2.27		
res230: H	2	2.27		
res230: K	1	1.13		
res230: Q	1	1.13		
res230: W	1	1.13		
wt 207: G	67.04		G	67.04
res231: G	59	67.04		
res231: A	18	20.45		
res231: V	11	12.5		
wt 208: K	39.77		K	39.77
res232: K	35	39.77		
res232: R	16	18.18		
res232: Q	14	15.9		
res232: I	13	14.77		
res232: L	5	5.68		
res232: E	2	2.27		
res232: M	2	2.27		
res232: S	1	1.13		
wt 209: L	10.22		A	26.13
res233: A	23	26.13		
res233: K	14	15.9		
res233: S	13	14.77		
res233: L	9	10.22		
res233: E	7	7.95		
res233: N	5	5.68		
res233: T	5	5.68		
res233: R	3	3.4		
res233: V	2	2.27		
res233: F	2	2.27		
res233: M	2	2.27		
res233: I	1	1.13		
res233: Q	1	1.13		
res233: G	1	1.13		
wt 210: L	47.72		L	47.72
res234: L	42	47.72		
res234: S	12	13.63		
res234: M	8	9.09		
res234: V	8	9.09		
res234: F	7	7.95		
res234: R	3	3.4		
res234: T	2	2.27		
res234: Y	2	2.27		
res234: K	1	1.13		
res234: A	1	1.13		
res234: E	1	1.13		
res234: N	1	1.13		
wt 211: R	95.45		R	95.45
res235: R	84	95.45		
res235: Q	2	2.27		
res235: C	1	1.13		
res235: K	1	1.13		
wt 212: P	97.72		P	97.72
res236: P	86	97.72		
res236: S	2	2.27		
wt 213: N	20.45		R	26.13
res237: R	23	26.13		
res237: K	22	25		
res237: N	18	20.45		
res237: Q	7	7.95		
res237: H	6	6.81		
res237: S	5	5.68		
res237: G	3	3.4		
res237: Y	1	1.13		
res237: D	1	1.13		
res237: A	1	1.13		
res237: E	1	1.13		
wt 214: H	23.86		G	48.86
res238: G	43	48.86		
res238: H	21	23.86		
res238: Q	9	10.22		
res238: R	5	5.68		
res238: Y	4	4.54		
res238: A	4	4.54		
res238: C	1	1.13		
res238: S	1	1.13		
wt 215: L	95.45		L	95.45
res239: L	84	95.45		
res239: F	2	2.27		
res239: H	1	1.13		
res239: Y	1	1.13		
wt 216: W	100		W	100
res240: W	88	100		
wt 217: G	100		G	100
res241: G	88	100		
wt 218: Y	44.31		F	54.54
res242: F	48	54.54		
res242: Y	39	44.31		
res242: L	1	1.13		

wt 219: Y	100		Y	100
res243: Y	88	100		
wt 220: L	70.45		L	70.45
res244: L	62	70.45		
res244: G	18	20.45		
res244: R	5	5.68		
res244: H	2	2.27		
res244: N	1	1.13		
wt 221: F	70.45		F	70.45
res245: F	62	70.45		
res245: Y	26	29.54		
wt 222: P	100		P	100
res246: P	88	100		
wt 223: D	84.09		D	84.09
res247: D	74	84.09		
res247: A	6	6.81		
res247: N	3	3.4		
res247: S	2	2.27		
res247: C	1	1.13		
res247: E	1	1.13		
res247: V	1	1.13		
wt 224: C	100		C	100
res248: C	88	100		
wt 225: Y	75		Y	75
res249: Y	66	75		
res249: H	15	17.04		
res249: G	6	6.81		
res249: F	1	1.13		
wt 226: N	98.86		N	98.86
res250: N	87	98.86		
res250: S	1	1.13		
wt 227: H	32.95		Y	48.86
res251: Y	43	48.86		
res251: H	29	32.95		
res251: N	7	7.95		
res251: G	6	6.81		
res251: T	1	1.13		
res251: S	1	1.13		
res251: D	1	1.13		
wt 228: H	9.09		D	44.31
res252: D	39	44.31		
res252: N	26	29.54		
res252: H	8	9.09		
res252: W	6	6.81		
res252: G	4	4.54		
res252: K	3	3.4		
res252: Y	1	1.13		
res252: Q	1	1.13		

wt 229: Y	46.59		Y	46.59
res253: Y	41	46.59		
res253: F	24	27.27		
res253: V	9	10.22		
res253: H	5	5.68		
res253: R	2	2.27		
res253: I	1	1.13		
res253: L	1	1.13		
res253: W	1	1.13		
res253: K	1	1.13		
res253: M	1	1.13		
res253: S	1	1.13		
res253: G	1	1.13		
wt 230: K	21.59		K	21.59
res254: K	19	21.59		
res254: V	12	13.63		
res254: Y	11	12.5		
res254: L	10	11.36		
res254: Q	6	6.81		
res254: N	5	5.68		
res254: R	5	5.68		
res254: H	5	5.68		
res254: D	4	4.54		
res254: S	4	4.54		
res254: G	3	3.4		
res254: T	2	2.27		
res254: M	1	1.13		
res254: A	1	1.13		
wt 231: K	19.31		S	19.31
res255: S	17	19.31		
res255: K	17	19.31		
res255: Q	16	18.18		
res255: A	10	11.36		
res255: D	7	7.95		
res255: I	5	5.68		
res255: M	4	4.54		
res255: T	4	4.54		
res255: G	3	3.4		
res255: E	3	3.4		
res255: H	1	1.13		
res255: N	1	1.13		
wt 232: P	13.63		-	35.22
res260: -	31	35.22		
res260: E	16	18.18		
res260: P	12	13.63		
res260: S	9	10.22		
res260: D	4	4.54		
res260: Q	4	4.54		
res260: K	4	4.54		
res260: A	3	3.4		
res260: L	1	1.13		
res260: F	1	1.13		
res260: N	1	1.13		
res260: T	1	1.13		
res260: H	1	1.13		

wt 233: G	7.95		N	51.13
res261: N	45	51.13		
res261: S	16	18.18		
res261: T	7	7.95		
res261: G	7	7.95		
res261: F	3	3.4		
res261: K	2	2.27		
res261: Q	2	2.27		
res261: R	1	1.13		
res261: H	1	1.13		
res261: P	1	1.13		
res261: I	1	1.13		
res261: -	1	1.13		
res261: D	1	1.13		
wt 234: Y	98.86		Y	98.86
res262: Y	87	98.86		
res262: -	1	1.13		
wt 235: N	12.5		T	78.4
res263: T	69	78.4		
res263: N	11	12.5		
res263: K	3	3.4		
res263: D	3	3.4		
res263: R	1	1.13		
res263: S	1	1.13		
wt 236: G	98.86		G	98.86
res264: G	87	98.86		
res264: M	1	1.13		
wt 237: S	27.27		S	27.27
res265: S	24	27.27		
res265: R	17	19.31		
res265: Q	17	19.31		
res265: H	10	11.36		
res265: E	6	6.81		
res265: T	3	3.4		
res265: K	3	3.4		
res265: N	3	3.4		
res265: A	2	2.27		
res265: F	1	1.13		
res265: D	1	1.13		
res265: I	1	1.13		
wt 238: C	100		C	100
res266: C	88	100		
wt 239: F	7.95		P	73.86
res267: P	65	73.86		
res267: F	7	7.95		
res267: H	7	7.95		
res267: S	4	4.54		
res267: L	2	2.27		
res267: R	1	1.13		
res267: T	1	1.13		
res267: A	1	1.13		

wt 240: N	6.81	D	40.9
res268: D	36	40.9	
res268: E	16	18.18	
res268: P	8	9.09	
res268: A	7	7.95	
res268: N	6	6.81	
res268: L	5	5.68	
res268: S	5	5.68	
res268: K	2	2.27	
res268: I	1	1.13	
res268: V	1	1.13	
res268: T	1	1.13	
wt 241: V	31.81	V	31.81
res269: V	28	31.81	
res269: E	12	13.63	
res269: G	10	11.36	
res269: I	10	11.36	
res269: D	8	9.09	
res269: A	7	7.95	
res269: L	5	5.68	
res269: S	2	2.27	
res269: N	2	2.27	
res269: K	1	1.13	
res269: T	1	1.13	
res269: F	1	1.13	
res269: Y	1	1.13	
wt 242: E	72.72	E	72.72
res270: E	64	72.72	
res270: I	13	14.77	
res270: T	6	6.81	
res270: A	3	3.4	
res270: V	2	2.27	
wt 243: I	17.04	V	32.95
res271: V	29	32.95	
res271: I	15	17.04	
res271: K	14	15.9	
res271: R	8	9.09	
res271: L	5	5.68	
res271: S	4	4.54	
res271: H	3	3.4	
res271: E	2	2.27	
res271: M	2	2.27	
res271: Q	2	2.27	
res271: P	1	1.13	
res271: T	1	1.13	
res271: C	1	1.13	
res271: Y	1	1.13	

wt 244: K	10.22	A	22.72
res272: A	20	22.72	
res272: L	16	18.18	
res272: S	10	11.36	
res272: K	9	10.22	
res272: Q	8	9.09	
res272: T	7	7.95	
res272: R	4	4.54	
res272: F	3	3.4	
res272: D	3	3.4	
res272: V	2	2.27	
res272: E	2	2.27	
res272: I	1	1.13	
res272: N	1	1.13	
res272: H	1	1.13	
res272: P	1	1.13	
wt 245: R	78.4	R	78.4
res273: R	69	78.4	
res273: Q	12	13.63	
res273: L	3	3.4	
res273: H	1	1.13	
res273: K	1	1.13	
res273: M	1	1.13	
res273: E	1	1.13	
wt 246: N	100	N	100
res274: N	88	100	
wt 247: D	68.18	D	68.18
res275: D	60	68.18	
res275: N	20	22.72	
res275: T	6	6.81	
res275: Q	2	2.27	
wt 248: D	11.36	Q	44.31
res276: Q	39	44.31	
res276: E	19	21.59	
res276: D	10	11.36	
res276: A	4	4.54	
res276: K	4	4.54	
res276: H	3	3.4	
res276: N	3	3.4	
res276: L	3	3.4	
res276: R	2	2.27	
res276: G	1	1.13	
wt 249: L	98.86	L	98.86
res277: L	87	98.86	
res277: I	1	1.13	

wt 250: S	25	S	25
res278: S	22	25	
res278: A	14	15.9	
res278: G	13	14.77	
res278: L	5	5.68	
res278: H	5	5.68	
res278: Q	4	4.54	
res278: F	4	4.54	
res278: K	4	4.54	
res278: N	3	3.4	
res278: D	3	3.4	
res278: R	3	3.4	
res278: M	3	3.4	
res278: V	1	1.13	
res278: I	1	1.13	
res278: Y	1	1.13	
res278: T	1	1.13	
res278: E	1	1.13	
wt 251: W	98.86	W	98.86
res279: W	87	98.86	
res279: F	1	1.13	
wt 252: L	94.31	L	94.31
res280: L	83	94.31	
res280: M	3	3.4	
res280: I	2	2.27	
wt 253: W	98.86	W	98.86
res281: W	87	98.86	
res281: F	1	1.13	
wt 254: N	31.81	N	31.81
res282: N	28	31.81	
res282: A	18	20.45	
res282: K	18	20.45	
res282: E	9	10.22	
res282: G	8	9.09	
res282: D	2	2.27	
res282: T	2	2.27	
res282: S	1	1.13	
res282: R	1	1.13	
res282: Q	1	1.13	
wt 255: E	39.77	E	39.77
res283: E	35	39.77	
res283: Q	15	17.04	
res283: S	14	15.9	
res283: A	11	12.5	
res283: K	5	5.68	
res283: V	2	2.27	
res283: D	1	1.13	
res283: N	1	1.13	
res283: I	1	1.13	
res283: G	1	1.13	
res283: H	1	1.13	
res283: T	1	1.13	
wt 256: S	100	S	100
res284: S	88	100	

wt 257: T	48.86		T	48.86
res285: T	43	48.86		
res285: A	14	15.9		
res285: R	10	11.36		
res285: S	10	11.36		
res285: Y	3	3.4		
res285: Q	2	2.27		
res285: M	2	2.27		
res285: N	2	2.27		
res285: C	1	1.13		
res285: L	1	1.13		
wt 258: A	97.72		A	97.72
res286: A	86	97.72		
res286: G	2	2.27		
wt 259: L	100		L	100
res287: L	88	100		
wt 260: Y	71.59		Y	71.59
res288: Y	63	71.59		
res288: F	25	28.4		
wt 261: P	98.86		P	98.86
res289: P	87	98.86		
res289: S	1	1.13		
wt 262: S	92.04		S	92.04
res290: S	81	92.04		
res290: D	3	3.4		
res290: N	2	2.27		
res290: A	2	2.27		
wt 263: I	71.59		I	71.59
res291: I	63	71.59		
res291: V	23	26.13		
res291: A	1	1.13		
res291: T	1	1.13		
wt 264: Y	81.81		Y	81.81
res292: Y	72	81.81		
res292: G	9	10.22		
res292: S	4	4.54		
res292: T	1	1.13		
res292: V	1	1.13		
res292: H	1	1.13		
wt 265: L	71.59		L	71.59
res293: L	63	71.59		
res293: V	10	11.36		
res293: M	9	10.22		
res293: I	6	6.81		

wt 266: N	9.09		P	26.13
res294: P	23	26.13		
res294: D	18	20.45		
res294: E	10	11.36		
res294: R	9	10.22		
res294: N	8	9.09		
res294: K	7	7.95		
res294: S	5	5.68		
res294: W	4	4.54		
res294: G	2	2.27		
res294: Q	1	1.13		
res294: T	1	1.13		
wt 267: T	11.36		K	25
res295: K	22	25		
res295: A	12	13.63		
res295: E	11	12.5		
res295: P	10	11.36		
res295: T	10	11.36		
res295: S	6	6.81		
res295: L	3	3.4		
res295: V	3	3.4		
res295: Q	3	3.4		
res295: Y	3	3.4		
res295: R	2	2.27		
res295: I	2	2.27		
res295: C	1	1.13		
wt 268: Q	5.68		S	15.9
res296: S	14	15.9		
res296: A	13	14.77		
res296: T	10	11.36		
res296: R	9	10.22		
res296: V	7	7.95		
res296: I	7	7.95		
res296: Q	5	5.68		
res296: E	4	4.54		
res296: G	4	4.54		
res296: L	4	4.54		
res296: K	3	3.4		
res296: M	2	2.27		
res296: D	2	2.27		
res296: H	1	1.13		
res296: F	1	1.13		
res296: W	1	1.13		
res296: C	1	1.13		
wt 269: Q	4.54		L	87.5
res297: L	77	87.5		
res297: Q	4	4.54		
res297: F	3	3.4		
res297: V	1	1.13		
res297: H	1	1.13		
res297: M	1	1.13		
res297: K	1	1.13		

wt 270: S	6.81		K	26.13
res298: K	23	26.13		
res298: A	18	20.45		
res298: E	11	12.5		
res298: G	10	11.36		
res298: P	7	7.95		
res298: R	6	6.81		
res298: S	6	6.81		
res298: M	3	3.4		
res298: Q	2	2.27		
res298: V	1	1.13		
res298: N	1	1.13		
wt 271: P	11.36		S	43.18
res299: S	38	43.18		
res299: G	18	20.45		
res299: D	12	13.63		
res299: P	10	11.36		
res299: N	7	7.95		
res299: V	1	1.13		
res299: M	1	1.13		
res299: L	1	1.13		
wt 272: V	4.54		S	54.54
res301: S	48	54.54		
res301: -	14	15.9		
res301: N	7	7.95		
res301: A	6	6.81		
res301: T	5	5.68		
res301: V	4	4.54		
res301: G	2	2.27		
res301: H	1	1.13		
res301: D	1	1.13		
wt 273: A	4.54		N	38.63
res303: N	34	38.63		
res303: K	16	18.18		
res303: H	14	15.9		
res303: Y	6	6.81		
res303: Q	5	5.68		
res303: A	4	4.54		
res303: S	3	3.4		
res303: L	2	2.27		
res303: E	1	1.13		
res303: -	1	1.13		
res303: V	1	1.13		
res303: D	1	1.13		
wt 274: A	27.27		A	27.27
res304: A	24	27.27		
res304: G	16	18.18		
res304: S	12	13.63		
res304: I	9	10.22		
res304: T	9	10.22		
res304: V	7	7.95		
res304: H	5	5.68		
res304: R	2	2.27		
res304: L	1	1.13		
res304: C	1	1.13		
res304: N	1	1.13		
res304: E	1	1.13		

wt 275: T	5.68		L	27.27
res305: L	24	27.27		
res305: R	20	22.72		
res305: Q	20	22.72		
res305: A	9	10.22		
res305: T	5	5.68		
res305: V	4	4.54		
res305: W	3	3.4		
res305: P	1	1.13		
res305: M	1	1.13		
res305: G	1	1.13		
wt 276: L	27.27		L	27.27
res306: L	24	27.27		
res306: K	15	17.04		
res306: M	11	12.5		
res306: R	11	12.5		
res306: N	10	11.36		
res306: A	6	6.81		
res306: H	5	5.68		
res306: Q	2	2.27		
res306: P	2	2.27		
res306: T	1	1.13		
res306: E	1	1.13		
wt 277: Y	32.95		F	65.9
res307: F	58	65.9		
res307: Y	29	32.95		
res307: M	1	1.13		
wt 278: V	78.4		V	78.4
res308: V	69	78.4		
res308: S	15	17.04		
res308: A	2	2.27		
res308: T	1	1.13		
res308: C	1	1.13		
wt 279: R	51.13		R	51.13
res309: R	45	51.13		
res309: Q	15	17.04		
res309: H	11	12.5		
res309: S	10	11.36		
res309: K	5	5.68		
res309: W	1	1.13		
res309: A	1	1.13		
wt 280: N	19.31		F	30.68
res310: F	27	30.68		
res310: H	25	28.4		
res310: N	17	19.31		
res310: Y	12	13.63		
res310: S	4	4.54		
res310: A	2	2.27		
res310: E	1	1.13		
wt 281: R	97.72		R	97.72
res311: R	86	97.72		
res311: Q	2	2.27		
wt 282: V	85.22		V	85.22
res312: V	75	85.22		
res312: L	7	7.95		
res312: I	6	6.81		

wt 283: R	14.77		Q	29.54
res313: Q	26	29.54		
res313: R	13	14.77		
res313: H	10	11.36		
res313: K	8	9.09		
res313: A	8	9.09		
res313: E	8	9.09		
res313: L	4	4.54		
res313: N	3	3.4		
res313: M	2	2.27		
res313: V	2	2.27		
res313: T	1	1.13		
res313: G	1	1.13		
res313: W	1	1.13		
res313: C	1	1.13		
wt 284: E	100		E	100
res314: E	88	100		
wt 285: A	77.27		A	77.27
res315: A	68	77.27		
res315: S	17	19.31		
res315: G	3	3.4		
wt 286: I	20.45		L	28.4
res316: L	25	28.4		
res316: M	22	25		
res316: F	20	22.72		
res316: I	18	20.45		
res316: R	2	2.27		
res316: V	1	1.13		
wt 287: R	98.86		R	98.86
res317: R	87	98.86		
res317: A	1	1.13		
wt 288: V	60.22		V	60.22
res318: V	53	60.22		
res318: I	23	26.13		
res318: L	8	9.09		
res318: M	2	2.27		
res318: A	1	1.13		
res318: T	1	1.13		
wt 289: S	42.04		A	54.54
res319: A	48	54.54		
res319: S	37	42.04		
res319: R	1	1.13		
res319: D	1	1.13		
res319: Q	1	1.13		

wt 290: K	15.9		T	15.9
res320: T	14	15.9		
res320: K	14	15.9		
res320: S	11	12.5		
res320: R	7	7.95		
res320: L	7	7.95		
res320: E	6	6.81		
res320: V	5	5.68		
res320: Q	5	5.68		
res320: A	4	4.54		
res320: H	3	3.4		
res320: D	3	3.4		
res320: I	3	3.4		
res320: Y	3	3.4		
res320: F	2	2.27		
res320: M	1	1.13		
wt 291: I	10.22		M	19.31
res321: M	17	19.31		
res321: V	17	19.31		
res321: A	10	11.36		
res321: T	10	11.36		
res321: I	9	10.22		
res321: Q	5	5.68		
res321: L	5	5.68		
res321: G	4	4.54		
res321: R	2	2.27		
res321: K	2	2.27		
res321: S	2	2.27		
res321: -	2	2.27		
res321: E	1	1.13		
res321: D	1	1.13		
res321: F	1	1.13		
wt 292: P	6.81		A	26.13
res322: A	23	26.13		
res322: H	17	19.31		
res322: T	16	18.18		
res322: G	6	6.81		
res322: S	6	6.81		
res322: P	6	6.81		
res322: -	4	4.54		
res322: V	2	2.27		
res322: K	2	2.27		
res322: M	2	2.27		
res322: R	2	2.27		
res322: I	1	1.13		
res322: E	1	1.13		
wt 293: D	9.09		H	25
res323: H	22	25		
res323: -	21	23.86		
res323: S	19	21.59		
res323: R	9	10.22		
res323: D	8	9.09		
res323: N	5	5.68		
res323: K	2	2.27		
res323: G	2	2.27		

wt 294: A	20.45		A	20.45
res324: A	18	20.45		
res324: H	13	14.77		
res324: G	10	11.36		
res324: -	10	11.36		
res324: K	9	10.22		
res324: R	5	5.68		
res324: P	4	4.54		
res324: D	4	4.54		
res324: E	3	3.4		
res324: T	3	3.4		
res324: V	3	3.4		
res324: N	2	2.27		
res324: S	1	1.13		
res324: C	1	1.13		
res324: L	1	1.13		
res324: Q	1	1.13		
wt 295: K	13.63		D	42.04
res326: D	37	42.04		
res326: N	13	14.77		
res326: K	12	13.63		
res326: G	6	6.81		
res326: S	4	4.54		
res326: P	3	3.4		
res326: E	2	2.27		
res326: H	2	2.27		
res326: Q	2	2.27		
res326: -	2	2.27		
res326: R	1	1.13		
res326: A	1	1.13		
res326: L	1	1.13		
res326: V	1	1.13		
res326: I	1	1.13		
wt 296: S	11.36		Y	31.81
res327: Y	28	31.81		
res327: H	20	22.72		
res327: P	13	14.77		
res327: S	10	11.36		
res327: N	4	4.54		
res327: F	3	3.4		
res327: D	2	2.27		
res327: E	1	1.13		
res327: G	1	1.13		
res327: A	1	1.13		
res327: C	1	1.13		
res327: L	1	1.13		
res327: -	1	1.13		
res327: I	1	1.13		
res327: Q	1	1.13		

wt 297: P	26.13		A	32.95
res328: A	29	32.95		
res328: P	23	26.13		
res328: S	10	11.36		
res328: N	7	7.95		
res328: T	6	6.81		
res328: D	3	3.4		
res328: V	3	3.4		
res328: H	2	2.27		
res328: K	2	2.27		
res328: G	1	1.13		
res328: R	1	1.13		
res328: M	1	1.13		
wt 298: L	84.09		L	84.09
res329: L	74	84.09		
res329: V	7	7.95		
res329: R	2	2.27		
res329: A	2	2.27		
res329: P	1	1.13		
res329: T	1	1.13		
res329: I	1	1.13		
wt 299: P	98.86		P	98.86
res330: P	87	98.86		
res330: S	1	1.13		
wt 300: V	84.09		V	84.09
res331: V	74	84.09		
res331: I	13	14.77		
res331: T	1	1.13		
wt 301: F	54.54		F	54.54
res332: F	48	54.54		
res332: L	20	22.72		
res332: Y	17	19.31		
res332: M	2	2.27		
res332: V	1	1.13		
wt 302: A	10.22		V	65.9
res333: V	58	65.9		
res333: P	18	20.45		
res333: A	9	10.22		
res333: I	3	3.4		
wt 303: Y	84.09		Y	84.09
res334: Y	74	84.09		
res334: F	14	15.9		
wt 304: T	48.86		T	48.86
res335: T	43	48.86		
res335: A	14	15.9		
res335: S	13	14.77		
res335: V	10	11.36		
res335: I	3	3.4		
res335: L	2	2.27		
res335: M	1	1.13		
res335: N	1	1.13		
res335: F	1	1.13		

wt 305: R	80.68		R	80.68
res336: R	71	80.68		
res336: Q	16	18.18		
res336: N	1	1.13		
wt 306: I	25		P	42.04
res337: P	37	42.04		
res337: L	28	31.81		
res337: I	22	25		
res337: V	1	1.13		
wt 307: V	22.72		T	27.27
res338: T	24	27.27		
res338: F	23	26.13		
res338: V	20	22.72		
res338: G	14	15.9		
res338: A	3	3.4		
res338: Y	1	1.13		
res338: D	1	1.13		
res338: L	1	1.13		
res338: S	1	1.13		
wt 308: F	23.86		Y	67.04
res339: Y	59	67.04		
res339: F	21	23.86		
res339: H	7	7.95		
res339: L	1	1.13		
wt 309: T	25		T	25
res340: T	22	25		
res340: R	21	23.86		
res340: S	13	14.77		
res340: D	9	10.22		
res340: E	5	5.68		
res340: A	5	5.68		
res340: L	5	5.68		
res340: I	3	3.4		
res340: K	2	2.27		
res340: G	1	1.13		
res340: M	1	1.13		
res340: N	1	1.13		
wt 310: D	28.4		D	28.4
res341: D	25	28.4		
res341: R	17	19.31		
res341: Y	8	9.09		
res341: N	7	7.95		
res341: M	7	7.95		
res341: S	5	5.68		
res341: Q	4	4.54		
res341: K	4	4.54		
res341: E	3	3.4		
res341: T	3	3.4		
res341: H	3	3.4		
res341: L	1	1.13		
res341: G	1	1.13		

wt 311: Q	11.36	T	25	
res342: T	22	25		
res342: S	15	17.04		
res342: E	10	11.36		
res342: -	10	11.36		
res342: Q	10	11.36		
res342: K	5	5.68		
res342: R	4	4.54		
res342: A	3	3.4		
res342: V	2	2.27		
res342: N	2	2.27		
res342: L	2	2.27		
res342: M	1	1.13		
res342: G	1	1.13		
res342: D	1	1.13		
wt 312: V	4.54	-	46.59	
res350: -	41	46.59		
res350: P	17	19.31		
res350: S	5	5.68		
res350: T	5	5.68		
res350: R	5	5.68		
res350: V	4	4.54		
res350: G	4	4.54		
res350: K	2	2.27		
res350: I	1	1.13		
res350: Y	1	1.13		
res350: N	1	1.13		
res350: E	1	1.13		
res350: A	1	1.13		
wt 313: L	43.18	L	43.18	
res351: L	38	43.18		
res351: N	11	12.5		
res351: F	8	9.09		
res351: S	8	9.09		
res351: G	6	6.81		
res351: D	4	4.54		
res351: M	4	4.54		
res351: I	2	2.27		
res351: T	2	2.27		
res351: Y	1	1.13		
res351: -	1	1.13		
res351: E	1	1.13		
res351: P	1	1.13		
res351: K	1	1.13		
wt 314: K	7.95	T	17.04	
res352: T	15	17.04		
res352: R	12	13.63		
res352: E	12	13.63		
res352: H	10	11.36		
res352: F	9	10.22		
res352: K	7	7.95		
res352: L	6	6.81		
res352: D	6	6.81		
res352: Y	3	3.4		
res352: Q	3	3.4		
res352: N	2	2.27		
res352: V	2	2.27		
res352: -	1	1.13		

wt 315: F	55.68	F	55.68	
res353: F	49	55.68		
res353: Y	10	11.36		
res353: G	9	10.22		
res353: L	6	6.81		
res353: A	5	5.68		
res353: V	5	5.68		
res353: E	1	1.13		
res353: Q	1	1.13		
res353: T	1	1.13		
res353: P	1	1.13		
wt 316: L	98.86	L	98.86	
res354: L	87	98.86		
res354: M	1	1.13		
wt 317: S	70.45	S	70.45	
res355: S	62	70.45		
res355: P	13	14.77		
res355: T	6	6.81		
res355: E	3	3.4		
res355: N	2	2.27		
res355: L	1	1.13		
res355: Q	1	1.13		
wt 318: Q	28.4	E	31.81	
res356: E	28	31.81		
res356: Q	25	28.4		
res356: L	15	17.04		
res356: K	12	13.63		
res356: R	3	3.4		
res356: T	2	2.27		
res356: P	1	1.13		
res356: A	1	1.13		
res356: M	1	1.13		
wt 319: D	27.27	D	27.27	
res357: D	24	27.27		
res357: E	16	18.18		
res357: M	15	17.04		
res357: Q	13	14.77		
res357: I	4	4.54		
res357: V	3	3.4		
res357: S	2	2.27		
res357: T	2	2.27		
res357: K	2	2.27		
res357: A	2	2.27		
res357: P	2	2.27		
res357: G	1	1.13		
res357: H	1	1.13		
res357: -	1	1.13		
wt 320: E	21.59	D	76.13	
res358: D	67	76.13		
res358: E	19	21.59		
res358: H	1	1.13		
res358: -	1	1.13		
wt 321: L	97.72	L	97.72	
res359: L	86	97.72		
res359: I	1	1.13		
res359: -	1	1.13		

wt 322: V	42.04	V	42.04	
res360: V	37	42.04		
res360: I	30	34.09		
res360: E	15	17.04		
res360: M	3	3.4		
res360: A	2	2.27		
res360: -	1	1.13		
wt 323: Y	5.68	S	38.63	
res361: S	34	38.63		
res361: H	26	29.54		
res361: N	15	17.04		
res361: Q	7	7.95		
res361: Y	5	5.68		
res361: -	1	1.13		
wt 324: T	77.27	T	77.27	
res362: T	68	77.27		
res362: S	17	19.31		
res362: V	2	2.27		
res362: -	1	1.13		
wt 325: F	2.27	I	77.27	
res363: I	68	77.27		
res363: L	14	15.9		
res363: F	2	2.27		
res363: V	2	2.27		
res363: S	1	1.13		
res363: -	1	1.13		
wt 326: G	98.86	G	98.86	
res364: G	87	98.86		
res364: -	1	1.13		
wt 327: E	89.77	E	89.77	
res365: E	79	89.77		
res365: V	6	6.81		
res365: Q	1	1.13		
res365: T	1	1.13		
res365: -	1	1.13		
wt 328: T	9.09	S	79.54	
res366: S	70	79.54		
res366: T	8	9.09		
res366: I	7	7.95		
res366: A	1	1.13		
res366: -	1	1.13		
res366: C	1	1.13		
wt 329: V	19.31	A	77.27	
res367: A	68	77.27		
res367: V	17	19.31		
res367: -	1	1.13		
res367: M	1	1.13		
res367: I	1	1.13		
wt 330: A	94.31	A	94.31	
res368: A	83	94.31		
res368: S	2	2.27		
res368: V	1	1.13		
res368: P	1	1.13		
res368: -	1	1.13		

wt 331: L	80.68		L	80.68
res369: L	71	80.68		
res369: Q	16	18.18		
res369: -	1	1.13		
wt 332: G	98.86		G	98.86
res370: G	87	98.86		
res370: -	1	1.13		
wt 333: A	86.36		A	86.36
res371: A	76	86.36		
res371: V	6	6.81		
res371: S	2	2.27		
res371: T	2	2.27		
res371: P	1	1.13		
res371: -	1	1.13		
wt 334: S	22.72		A	73.86
res372: A	65	73.86		
res372: S	20	22.72		
res372: D	2	2.27		
res372: -	1	1.13		
wt 335: G	98.86		G	98.86
res373: G	87	98.86		
res373: -	1	1.13		
wt 336: I	39.77		V	47.72
res374: V	42	47.72		
res374: I	35	39.77		
res374: F	4	4.54		
res374: A	4	4.54		
res374: T	1	1.13		
res374: M	1	1.13		
res374: -	1	1.13		
wt 337: V	63.63		V	63.63
res375: V	56	63.63		
res375: I	31	35.22		
res375: -	1	1.13		
wt 338: I	23.86		L	48.86
res376: L	43	48.86		
res376: I	21	23.86		
res376: M	10	11.36		
res376: F	9	10.22		
res376: V	4	4.54		
res376: -	1	1.13		
wt 339: W	98.86		W	98.86
res377: W	87	98.86		
res377: -	1	1.13		
wt 340: G	81.81		G	81.81
res378: G	72	81.81		
res378: V	11	12.5		
res378: D	2	2.27		
res378: L	1	1.13		
res378: -	1	1.13		
res378: I	1	1.13		

wt 341: T	4.54		D	51.13
res379: D	45	51.13		
res379: S	28	31.81		
res379: G	5	5.68		
res379: T	4	4.54		
res379: E	2	2.27		
res379: A	2	2.27		
res379: N	1	1.13		
res379: -	1	1.13		
wt 342: L	23.86		L	23.86
res380: L	21	23.86		
res380: M	19	21.59		
res380: A	13	14.77		
res380: S	11	12.5		
res380: W	8	9.09		
res380: H	3	3.4		
res380: V	3	3.4		
res380: Y	3	3.4		
res380: I	3	3.4		
res380: G	1	1.13		
res380: T	1	1.13		
res380: F	1	1.13		
res380: -	1	1.13		
wt 343: S	17.04		N	28.4
res382: N	25	28.4		
res382: E	22	25		
res382: S	15	17.04		
res382: G	9	10.22		
res382: Q	3	3.4		
res382: D	3	3.4		
res382: A	3	3.4		
res382: -	2	2.27		
res382: T	2	2.27		
res382: Y	1	1.13		
res382: I	1	1.13		
res382: V	1	1.13		
res382: L	1	1.13		
wt 344: I	4.54		L	32.95
res383: L	29	32.95		
res383: Y	26	29.54		
res383: F	8	9.09		
res383: N	8	9.09		
res383: D	4	4.54		
res383: I	4	4.54		
res383: K	2	2.27		
res383: M	2	2.27		
res383: -	2	2.27		
res383: S	1	1.13		
res383: C	1	1.13		
res383: V	1	1.13		
wt 345: M	2.27		T	54.54
res384: T	48	54.54		
res384: S	21	23.86		
res384: A	12	13.63		
res384: V	2	2.27		
res384: M	2	2.27		
res384: -	2	2.27		
res384: Y	1	1.13		

wt 346: R	20.45		S	30.68
res385: S	27	30.68		
res385: R	18	20.45		
res385: K	11	12.5		
res385: T	10	11.36		
res385: Q	7	7.95		
res385: L	4	4.54		
res385: N	3	3.4		
res385: H	2	2.27		
res385: A	2	2.27		
res385: D	2	2.27		
res385: -	1	1.13		
res385: E	1	1.13		
wt 347: S	64.77		S	64.77
res386: S	57	64.77		
res386: T	22	25		
res386: N	5	5.68		
res386: R	2	2.27		
res386: D	1	1.13		
res386: -	1	1.13		
wt 348: M	13.63		K	39.77
res387: K	35	39.77		
res387: E	15	17.04		
res387: M	12	13.63		
res387: T	5	5.68		
res387: A	4	4.54		
res387: R	4	4.54		
res387: Q	4	4.54		
res387: Y	2	2.27		
res387: V	2	2.27		
res387: N	2	2.27		
res387: P	1	1.13		
res387: L	1	1.13		
res387: -	1	1.13		
wt 349: K	12.5		E	54.54
res388: E	48	54.54		
res388: K	11	12.5		
res388: G	7	7.95		
res388: A	6	6.81		
res388: Q	4	4.54		
res388: R	3	3.4		
res388: H	2	2.27		
res388: T	2	2.27		
res388: V	2	2.27		
res388: P	1	1.13		
res388: D	1	1.13		
res388: N	1	1.13		
wt 350: S	36.36		S	36.36
res389: S	32	36.36		
res389: T	21	23.86		
res389: N	21	23.86		
res389: E	5	5.68		
res389: A	3	3.4		
res389: M	2	2.27		
res389: G	2	2.27		
res389: K	1	1.13		
res389: Q	1	1.13		

wt 351: C	100	C	100
res390: C	88	100		
wt 352: L	19.31	Q	29.54
res391: Q	26	29.54		
res391: T	19	21.59		
res391: L	17	19.31		
res391: W	6	6.81		
res391: E	4	4.54		
res391: R	4	4.54		
res391: M	4	4.54		
res391: S	3	3.4		
res391: P	2	2.27		
res391: I	2	2.27		
res391: A	1	1.13		
wt 353: L	5.68	A	17.04
res392: A	15	17.04		
res392: K	15	17.04		
res392: T	9	10.22		
res392: N	9	10.22		
res392: S	7	7.95		
res392: Y	7	7.95		
res392: R	6	6.81		
res392: L	5	5.68		
res392: H	4	4.54		
res392: I	3	3.4		
res392: D	3	3.4		
res392: M	2	2.27		
res392: G	1	1.13		
res392: E	1	1.13		
res392: V	1	1.13		
wt 354: L	48.86	L	48.86
res393: L	43	48.86		
res393: V	26	29.54		
res393: I	19	21.59		
wt 355: D	10.22	K	57.95
res394: K	51	57.95		
res394: D	9	10.22		
res394: H	9	10.22		
res394: R	5	5.68		
res394: N	5	5.68		
res394: Q	4	4.54		
res394: S	2	2.27		
res394: G	2	2.27		
res394: E	1	1.13		
wt 356: N	10.22	D	25
res395: D	22	25		
res395: Q	16	18.18		
res395: E	12	13.63		
res395: K	11	12.5		
res395: N	9	10.22		
res395: S	6	6.81		
res395: R	3	3.4		
res395: T	3	3.4		
res395: A	3	3.4		
res395: Y	1	1.13		
res395: H	1	1.13		
res395: G	1	1.13		
wt 357: Y	80.68	Y	80.68
res396: Y	71	80.68		
res396: F	13	14.77		
res396: T	3	3.4		
res396: S	1	1.13		
wt 358: M	23.86	L	39.77
res397: L	35	39.77		
res397: V	22	25		
res397: M	21	23.86		
res397: I	10	11.36		
wt 359: E	12.5	D	25
res398: D	22	25		
res398: T	12	13.63		
res398: E	11	12.5		
res398: S	11	12.5		
res398: V	7	7.95		
res398: N	6	6.81		
res398: K	6	6.81		
res398: R	5	5.68		
res398: L	2	2.27		
res398: Q	2	2.27		
res398: H	2	2.27		
res398: G	1	1.13		
res398: A	1	1.13		
wt 360: T	22.72	S	23.86
res399: S	21	23.86		
res399: G	20	22.72		
res399: T	20	22.72		
res399: R	9	10.22		
res399: N	5	5.68		
res399: E	4	4.54		
res399: D	3	3.4		
res399: K	2	2.27		
res399: Q	1	1.13		
res399: V	1	1.13		
res399: M	1	1.13		
res399: H	1	1.13		
wt 361: I	10.22	T	39.77
res400: T	35	39.77		
res400: D	16	18.18		
res400: L	10	11.36		
res400: I	9	10.22		
res400: P	8	9.09		
res400: R	2	2.27		
res400: S	2	2.27		
res400: V	2	2.27		
res400: K	1	1.13		
res400: H	1	1.13		
res400: E	1	1.13		
res400: N	1	1.13		
wt 362: L	96.59	L	96.59
res401: L	85	96.59		
res401: F	3	3.4		
wt 363: N	20.45	G	64.77
res402: G	57	64.77		
res402: N	18	20.45		
res402: V	10	11.36		
res402: L	1	1.13		
res402: T	1	1.13		
res402: A	1	1.13		
wt 364: P	63.63	P	63.63
res403: P	56	63.63		
res403: H	10	11.36		
res403: S	8	9.09		
res403: R	8	9.09		
res403: E	1	1.13		
res403: L	1	1.13		
res403: Y	1	1.13		
res403: Q	1	1.13		
res403: N	1	1.13		
res403: V	1	1.13		
wt 365: Y	84.09	Y	84.09
res404: Y	74	84.09		
res404: F	13	14.77		
res404: S	1	1.13		
wt 366: I	57.95	I	57.95
res405: I	51	57.95		
res405: V	22	25		
res405: L	13	14.77		
res405: A	2	2.27		
wt 367: I	31.81	V	37.5
res406: V	33	37.5		
res406: I	28	31.81		
res406: L	18	20.45		
res406: A	4	4.54		
res406: T	2	2.27		
res406: M	1	1.13		
res406: K	1	1.13		
res406: F	1	1.13		
wt 368: N	97.72	N	97.72
res407: N	86	97.72		
res407: S	1	1.13		
res407: A	1	1.13		
wt 369: V	96.59	V	96.59
res408: V	85	96.59		
res408: L	2	2.27		
res408: A	1	1.13		
wt 370: T	81.81	T	81.81
res409: T	72	81.81		
res409: S	15	17.04		
res409: R	1	1.13		

wt 371: L	19.31		S	27.27
res410: S	24	27.27		
res410: L	17	19.31		
res410: R	13	14.77		
res410: T	11	12.5		
res410: W	10	11.36		
res410: K	4	4.54		
res410: A	4	4.54		
res410: E	2	2.27		
res410: D	1	1.13		
res410: M	1	1.13		
res410: G	1	1.13		
wt 372: A	84.09		A	84.09
res411: A	74	84.09		
res411: G	10	11.36		
res411: S	4	4.54		
wt 373: A	89.77		A	89.77
res412: A	79	89.77		
res412: T	8	9.09		
res412: V	1	1.13		
wt 374: K	28.4		K	28.4
res413: K	25	28.4		
res413: E	19	21.59		
res413: L	13	14.77		
res413: Q	13	14.77		
res413: T	4	4.54		
res413: M	4	4.54		
res413: R	3	3.4		
res413: D	2	2.27		
res413: F	1	1.13		
res413: V	1	1.13		
res413: N	1	1.13		
res413: H	1	1.13		
res413: A	1	1.13		
wt 375: M	19.31		L	37.5
res414: L	33	37.5		
res414: M	17	19.31		
res414: V	12	13.63		
res414: Y	10	11.36		
res414: A	7	7.95		
res414: I	3	3.4		
res414: H	2	2.27		
res414: R	1	1.13		
res414: E	1	1.13		
res414: Q	1	1.13		
res414: N	1	1.13		
wt 376: C	100		C	100
res415: C	88	100		
wt 377: S	96.59		S	96.59
res416: S	85	96.59		
res416: N	2	2.27		
res416: G	1	1.13		

wt 378: Q	44.31		Q	44.31
res417: Q	39	44.31		
res417: L	10	11.36		
res417: R	9	10.22		
res417: W	8	9.09		
res417: H	7	7.95		
res417: E	6	6.81		
res417: K	6	6.81		
res417: I	2	2.27		
res417: T	1	1.13		
wt 379: V	20.45		A	28.4
res418: A	25	28.4		
res418: V	18	20.45		
res418: H	14	15.9		
res418: T	9	10.22		
res418: S	6	6.81		
res418: Q	6	6.81		
res418: N	4	4.54		
res418: R	1	1.13		
res418: K	1	1.13		
res418: E	1	1.13		
res418: I	1	1.13		
res418: D	1	1.13		
res418: F	1	1.13		
wt 380: L	71.59		L	71.59
res419: L	63	71.59		
res419: Q	12	13.63		
res419: R	6	6.81		
res419: V	4	4.54		
res419: H	1	1.13		
res419: M	1	1.13		
res419: T	1	1.13		
wt 381: C	100		C	100
res420: C	88	100		
wt 382: Q	18.18		H	23.86
res421: H	21	23.86		
res421: Q	16	18.18		
res421: S	16	18.18		
res421: N	10	11.36		
res421: R	9	10.22		
res421: K	7	7.95		
res421: G	4	4.54		
res421: T	2	2.27		
res421: E	2	2.27		
res421: F	1	1.13		
wt 383: E	12.5		G	45.45
res422: G	40	45.45		
res422: N	15	17.04		
res422: E	11	12.5		
res422: S	10	11.36		
res422: K	5	5.68		
res422: D	4	4.54		
res422: R	1	1.13		
res422: A	1	1.13		
res422: F	1	1.13		

wt 384: Q	15.9		H	44.31
res423: H	39	44.31		
res423: N	25	28.4		
res423: Q	14	15.9		
res423: R	4	4.54		
res423: E	3	3.4		
res423: K	2	2.27		
res423: S	1	1.13		
wt 385: G	100		G	100
res424: G	88	100		
wt 386: V	11.36		R	78.4
res425: R	69	78.4		
res425: V	10	11.36		
res425: I	3	3.4		
res425: L	2	2.27		
res425: A	2	2.27		
res425: M	1	1.13		
res425: K	1	1.13		
wt 387: C	100		C	100
res426: C	88	100		
wt 388: I	19.31		V	37.5
res427: V	33	37.5		
res427: I	17	19.31		
res427: A	12	13.63		
res427: T	8	9.09		
res427: L	7	7.95		
res427: S	6	6.81		
res427: R	2	2.27		
res427: Q	1	1.13		
res427: K	1	1.13		
res427: Y	1	1.13		
wt 389: R	94.31		R	94.31
res428: R	83	94.31		
res428: W	3	3.4		
res428: -	1	1.13		
res428: K	1	1.13		
wt 390: K	50		K	50
res429: K	44	50		
res429: R	33	37.5		
res429: Q	5	5.68		
res429: H	3	3.4		
res429: W	1	1.13		
res429: N	1	1.13		
res429: -	1	1.13		

wt 391: N 22.72 D 23.86
res430: D 21 23.86
res430: N 20 22.72
res430: P 11 12.5
res430: T 8 9.09
res430: M 5 5.68
res430: H 5 5.68
res430: Q 4 4.54
res430: V 4 4.54
res430: A 2 2.27
res430: K 2 2.27
res430: R 2 2.27
res430: S 2 2.27
res430: E 1 1.13
res430: - 1 1.13

wt 392: W 34.09 W 34.09
res431: W 30 34.09
res431: P 28 31.81
res431: S 17 19.31
res431: N 4 4.54
res431: L 2 2.27
res431: E 2 2.27
res431: V 1 1.13
res431: - 1 1.13
res431: I 1 1.13
res431: K 1 1.13
res431: D 1 1.13

wt 393: N 28.4 N 28.4
res432: N 25 28.4
res432: H 12 13.63
res432: S 11 12.5
res432: K 8 9.09
res432: G 7 7.95
res432: D 7 7.95
res432: E 5 5.68
res432: Y 4 4.54
res432: T 3 3.4
res432: - 2 2.27
res432: R 2 2.27
res432: Q 1 1.13
res432: I 1 1.13

wt 394: S 35.22 S 35.22
res433: S 31 35.22
res433: A 27 30.68
res433: P 13 14.77
res433: Q 7 7.95
res433: T 7 7.95
res433: E 1 1.13
res433: - 1 1.13
res433: M 1 1.13

wt 395: S 22.72 S 22.72
res435: S 20 22.72
res435: N 13 14.77
res435: D 11 12.5
res435: E 11 12.5
res435: P 10 11.36
res435: K 8 9.09
res435: H 5 5.68
res435: F 2 2.27
res435: A 2 2.27
res435: L 1 1.13
res435: Q 1 1.13
res435: - 1 1.13
res435: G 1 1.13
res435: R 1 1.13
res435: T 1 1.13

wt 396: D 27.27 D 27.27
res436: D 24 27.27
res436: A 23 26.13
res436: T 14 15.9
res436: V 9 10.22
res436: F 4 4.54
res436: S 4 4.54
res436: H 3 3.4
res436: I 2 2.27
res436: N 2 2.27
res436: - 1 1.13
res436: M 1 1.13
res436: G 1 1.13

wt 397: Y 53.4 Y 53.4
res437: Y 47 53.4
res437: F 24 27.27
res437: L 13 14.77
res437: - 1 1.13
res437: I 1 1.13
res437: H 1 1.13
res437: R 1 1.13

wt 398: L 96.59 L 96.59
res438: L 85 96.59
res438: P 2 2.27
res438: - 1 1.13

wt 399: H 81.81 H 81.81
res439: H 72 81.81
res439: L 5 5.68
res439: I 5 5.68
res439: T 3 3.4
res439: V 1 1.13
res439: Y 1 1.13
res439: - 1 1.13

wt 400: L 93.18 L 93.18
res440: L 82 93.18
res440: M 4 4.54
res440: I 1 1.13
res440: - 1 1.13

wt 401: N 60.22 N 60.22
res441: N 53 60.22
res441: S 18 20.45
res441: Q 4 4.54
res441: P 4 4.54
res441: H 3 3.4
res441: G 2 2.27
res441: D 2 2.27
res441: W 1 1.13
res441: - 1 1.13

wt 402: P 69.31 P 69.31
res442: P 61 69.31
res442: A 8 9.09
res442: S 7 7.95
res442: T 5 5.68
res442: E 4 4.54
res442: L 1 1.13
res442: N 1 1.13
res442: - 1 1.13

wt 403: D 17.04 A 30.68
res443: A 27 30.68
res443: D 15 17.04
res443: S 12 13.63
res443: N 7 7.95
res443: R 5 5.68
res443: T 5 5.68
res443: G 3 3.4
res443: E 3 3.4
res443: K 3 3.4
res443: Q 3 3.4
res443: M 2 2.27
res443: H 1 1.13
res443: L 1 1.13
res443: - 1 1.13

wt 404: N 19.31 S 55.68
res444: S 49 55.68
res444: N 17 19.31
res444: - 9 10.22
res444: T 6 6.81
res444: H 2 2.27
res444: F 2 2.27
res444: Q 1 1.13
res444: M 1 1.13
res444: R 1 1.13

wt 405: F 61.36 F 61.36
res445: F 54 61.36
res445: Y 13 14.77
res445: - 9 10.22
res445: H 5 5.68
res445: W 2 2.27
res445: L 2 2.27
res445: I 1 1.13
res445: Q 1 1.13
res445: G 1 1.13

wt 406: A	12.5	R	18.18
res446: R	16	18.18		
res446: S	15	17.04		
res446: H	12	13.63		
res446: A	11	12.5		
res446: -	9	10.22		
res446: Q	8	9.09		
res446: K	8	9.09		
res446: D	3	3.4		
res446: Y	2	2.27		
res446: N	1	1.13		
res446: V	1	1.13		
res446: G	1	1.13		
res446: T	1	1.13		
wt 407: I	69.31	I	69.31
res447: I	61	69.31		
res447: L	13	14.77		
res447: -	9	10.22		
res447: N	2	2.27		
res447: S	1	1.13		
res447: V	1	1.13		
res447: E	1	1.13		
wt 408: Q	22.72	Q	22.72
res448: Q	20	22.72		
res448: E	19	21.59		
res448: V	14	15.9		
res448: -	9	10.22		
res448: R	7	7.95		
res448: H	5	5.68		
res448: K	4	4.54		
res448: Y	3	3.4		
res448: I	1	1.13		
res448: L	1	1.13		
res448: D	1	1.13		
res448: A	1	1.13		
res448: M	1	1.13		
res448: G	1	1.13		
res448: S	1	1.13		
wt 409: L	20.45	L	20.45
res449: L	18	20.45		
res449: P	15	17.04		
res449: A	13	14.77		
res449: -	12	13.63		
res449: R	5	5.68		
res449: F	4	4.54		
res449: V	4	4.54		
res449: T	4	4.54		
res449: I	3	3.4		
res449: S	3	3.4		
res449: H	3	3.4		
res449: Y	1	1.13		
res449: E	1	1.13		
res449: K	1	1.13		
res449: M	1	1.13		

wt 410: E	5.68	-	34.09
res450: -	30	34.09		
res450: S	19	21.59		
res450: G	10	11.36		
res450: R	5	5.68		
res450: E	5	5.68		
res450: A	4	4.54		
res450: T	4	4.54		
res450: Q	4	4.54		
res450: N	3	3.4		
res450: K	2	2.27		
res450: I	1	1.13		
res450: V	1	1.13		
wt 411: K	19.31	G	21.59
res456: G	19	21.59		
res456: K	17	19.31		
res456: E	16	18.18		
res456: -	13	14.77		
res456: D	7	7.95		
res456: Q	6	6.81		
res456: N	3	3.4		
res456: S	3	3.4		
res456: R	2	2.27		
res456: P	1	1.13		
res456: A	1	1.13		
wt 412: G	38.63	G	38.63
res457: G	34	38.63		
res457: D	18	20.45		
res457: E	10	11.36		
res457: N	7	7.95		
res457: -	5	5.68		
res457: S	4	4.54		
res457: A	3	3.4		
res457: T	2	2.27		
res457: R	2	2.27		
res457: K	1	1.13		
res457: H	1	1.13		
res457: Y	1	1.13		
wt 413: G	45.45	G	45.45
res463: G	40	45.45		
res463: P	19	21.59		
res463: -	9	10.22		
res463: R	7	7.95		
res463: F	3	3.4		
res463: K	3	3.4		
res463: N	3	3.4		
res463: E	1	1.13		
res463: L	1	1.13		
res463: Q	1	1.13		
res463: V	1	1.13		

wt 414: K	22.72	K	22.72
res464: K	20	22.72		
res464: P	13	14.77		
res464: E	11	12.5		
res464: -	9	10.22		
res464: Q	8	9.09		
res464: R	7	7.95		
res464: L	4	4.54		
res464: G	4	4.54		
res464: S	3	3.4		
res464: T	3	3.4		
res464: D	2	2.27		
res464: C	1	1.13		
res464: W	1	1.13		
res464: N	1	1.13		
res464: H	1	1.13		
wt 415: F	31.81	L	31.81
res465: L	28	31.81		
res465: F	28	31.81		
res465: Y	16	18.18		
res465: -	9	10.22		
res465: I	3	3.4		
res465: P	2	2.27		
res465: V	2	2.27		
wt 416: T	28.4	T	28.4
res466: T	25	28.4		
res466: R	12	13.63		
res466: V	10	11.36		
res466: S	9	10.22		
res466: -	9	10.22		
res466: E	5	5.68		
res466: I	5	5.68		
res466: K	3	3.4		
res466: W	2	2.27		
res466: Y	2	2.27		
res466: H	1	1.13		
res466: M	1	1.13		
res466: A	1	1.13		
res466: F	1	1.13		
res466: L	1	1.13		
res466: Q	1	1.13		
wt 417: V	38.63	V	38.63
res467: V	34	38.63		
res467: L	16	18.18		
res467: P	10	11.36		
res467: A	10	11.36		
res467: -	9	10.22		
res467: I	8	9.09		
res467: T	1	1.13		

wt 418: R	11.36		K	19.31
res468: K	17	19.31		
res468: R	10	11.36		
res468: Q	10	11.36		
res468: E	9	10.22		
res468: -	9	10.22		
res468: S	6	6.81		
res468: T	6	6.81		
res468: H	6	6.81		
res468: V	6	6.81		
res468: P	2	2.27		
res468: L	2	2.27		
res468: N	2	2.27		
res468: A	1	1.13		
res468: D	1	1.13		
res468: I	1	1.13		

wt 419: G	82.95		G	82.95
res469: G	73	82.95		
res469: -	9	10.22		
res469: S	2	2.27		
res469: A	1	1.13		
res469: P	1	1.13		
res469: R	1	1.13		
res469: E	1	1.13		

wt 420: K	27.27		K	27.27
res470: K	24	27.27		
res470: A	11	12.5		
res470: E	11	12.5		
res470: -	9	10.22		
res470: R	7	7.95		
res470: N	6	6.81		
res470: H	5	5.68		
res470: T	5	5.68		
res470: Q	4	4.54		
res470: S	2	2.27		
res470: P	1	1.13		
res470: Y	1	1.13		
res470: L	1	1.13		
res470: W	1	1.13		

wt 421: P	20.45		L	38.63
res471: L	34	38.63		
res471: P	18	20.45		
res471: A	14	15.9		
res471: -	9	10.22		
res471: V	3	3.4		
res471: M	3	3.4		
res471: S	3	3.4		
res471: N	1	1.13		
res471: H	1	1.13		
res471: F	1	1.13		
res471: E	1	1.13		

wt 422: T	15.9		S	55.68
res472: S	49	55.68		
res472: T	14	15.9		
res472: -	9	10.22		
res472: K	6	6.81		
res472: P	2	2.27		
res472: N	2	2.27		
res472: R	2	2.27		
res472: G	2	2.27		
res472: Q	1	1.13		
res472: W	1	1.13		

wt 423: L	28.4		L	28.4
res475: L	25	28.4		
res475: D	12	13.63		
res475: W	8	9.09		
res475: -	6	6.81		
res475: A	4	4.54		
res475: P	4	4.54		
res475: H	4	4.54		
res475: S	4	4.54		
res475: K	3	3.4		
res475: R	3	3.4		
res475: Q	3	3.4		
res475: Y	3	3.4		
res475: V	3	3.4		
res475: G	2	2.27		
res475: E	2	2.27		
res475: N	2	2.27		

wt 424: E	30.68		E	30.68
res476: E	27	30.68		
res476: A	15	17.04		
res476: T	12	13.63		
res476: K	10	11.36		
res476: -	6	6.81		
res476: L	3	3.4		
res476: D	3	3.4		
res476: H	2	2.27		
res476: R	2	2.27		
res476: G	2	2.27		
res476: Q	1	1.13		
res476: N	1	1.13		
res476: W	1	1.13		
res476: M	1	1.13		
res476: P	1	1.13		
res476: S	1	1.13		

wt 425: D	79.54		D	79.54
res477: D	70	79.54		
res477: -	6	6.81		
res477: G	3	3.4		
res477: E	3	3.4		
res477: T	3	3.4		
res477: A	1	1.13		
res477: S	1	1.13		
res477: H	1	1.13		

wt 426: L	45.45		L	45.45
res478: L	40	45.45		
res478: I	14	15.9		
res478: R	10	11.36		
res478: Q	9	10.22		
res478: -	8	9.09		
res478: K	2	2.27		
res478: V	2	2.27		
res478: G	1	1.13		
res478: S	1	1.13		
res478: T	1	1.13		

wt 427: E	11.36		A	25
res479: A	22	25		
res479: E	10	11.36		
res479: Q	9	10.22		
res479: -	8	9.09		
res479: K	6	6.81		
res479: L	6	6.81		
res479: D	5	5.68		
res479: S	4	4.54		
res479: H	4	4.54		
res479: N	4	4.54		
res479: M	3	3.4		
res479: V	3	3.4		
res479: T	2	2.27		
res479: G	1	1.13		
res479: I	1	1.13		

wt 428: Q	21.59		Q	21.59
res480: Q	19	21.59		
res480: V	11	12.5		
res480: Y	10	11.36		
res480: -	8	9.09		
res480: H	8	9.09		
res480: R	6	6.81		
res480: D	5	5.68		
res480: A	5	5.68		
res480: N	3	3.4		
res480: F	3	3.4		
res480: E	3	3.4		
res480: L	2	2.27		
res480: T	2	2.27		
res480: I	1	1.13		
res480: K	1	1.13		
res480: G	1	1.13		

wt 429: F	23.86		M	40.9
res481: M	36	40.9		
res481: L	22	25		
res481: F	21	23.86		
res481: -	5	5.68		
res481: S	1	1.13		
res481: W	1	1.13		
res481: D	1	1.13		
res481: A	1	1.13		

wt 430: S	19.31		A	26.13
res482: A	23	26.13		
res482: S	17	19.31		
res482: Q	12	13.63		
res482: K	9	10.22		
res482: E	7	7.95		
res482: R	6	6.81		
res482: -	5	5.68		
res482: W	3	3.4		
res482: M	2	2.27		
res482: V	1	1.13		
res482: I	1	1.13		
res482: T	1	1.13		
res482: G	1	1.13		
wt 431: E	28.4		E	28.4
res483: E	25	28.4		
res483: K	11	12.5		
res483: T	10	11.36		
res483: V	8	9.09		
res483: D	6	6.81		
res483: -	5	5.68		
res483: M	5	5.68		
res483: S	5	5.68		
res483: N	4	4.54		
res483: Q	4	4.54		
res483: I	2	2.27		
res483: A	2	2.27		
res483: H	1	1.13		
wt 432: K	23.86		K	23.86
res484: K	21	23.86		
res484: H	17	19.31		
res484: E	13	14.77		
res484: N	11	12.5		
res484: -	6	6.81		
res484: G	4	4.54		
res484: S	4	4.54		
res484: T	4	4.54		
res484: R	4	4.54		
res484: D	2	2.27		
res484: Q	2	2.27		
wt 433: F	93.18		F	93.18
res485: F	82	93.18		
res485: -	6	6.81		
wt 434: Y	10.22		R	22.72
res486: R	20	22.72		
res486: S	13	14.77		
res486: Q	11	12.5		
res486: K	10	11.36		
res486: Y	9	10.22		
res486: -	6	6.81		
res486: T	5	5.68		
res486: V	4	4.54		
res486: M	3	3.4		
res486: E	2	2.27		
res486: L	2	2.27		
res486: H	1	1.13		
res486: F	1	1.13		
res486: D	1	1.13		
wt 435: C	93.18		C	93.18
res487: C	82	93.18		
res487: -	6	6.81		
wt 436: S	15.9		H	31.81
res488: H	28	31.81		
res488: Q	23	26.13		
res488: R	15	17.04		
res488: S	14	15.9		
res488: -	6	6.81		
res488: L	1	1.13		
res488: V	1	1.13		
wt 437: C	93.18		C	93.18
res489: C	82	93.18		
res489: -	6	6.81		
wt 438: Y	85.22		Y	85.22
res490: Y	75	85.22		
res490: -	6	6.81		
res490: F	6	6.81		
res490: N	1	1.13		
wt 439: S	9.09		-	37.5
res491: -	33	37.5		
res491: Q	16	18.18		
res491: A	9	10.22		
res491: S	8	9.09		
res491: E	5	5.68		
res491: P	4	4.54		
res491: T	4	4.54		
res491: L	4	4.54		
res491: R	1	1.13		
res491: H	1	1.13		
res491: N	1	1.13		
res491: Y	1	1.13		
res491: W	1	1.13		
wt 440: T	3.4		G	43.18
res492: G	38	43.18		
res492: -	35	39.77		
res492: N	10	11.36		
res492: T	3	3.4		
res492: D	1	1.13		
res492: R	1	1.13		
wt 441: L	7.95		-	39.77
res493: -	35	39.77		
res493: W	19	21.59		
res493: Y	12	13.63		
res493: L	7	7.95		
res493: R	3	3.4		
res493: H	3	3.4		
res493: M	2	2.27		
res493: F	2	2.27		
res493: I	2	2.27		
res493: V	2	2.27		
res493: S	1	1.13		
wt 442: S	11.36		-	43.18
res497: -	38	43.18		
res497: D	14	15.9		
res497: S	10	11.36		
res497: N	6	6.81		
res497: W	5	5.68		
res497: Q	4	4.54		
res497: T	4	4.54		
res497: H	3	3.4		
res497: G	1	1.13		
res497: R	1	1.13		
res497: A	1	1.13		
res497: K	1	1.13		
wt 443: C	52.27		C	52.27
res498: C	46	52.27		
res498: -	42	47.72		
wt 444: K	20.45		-	48.86
res499: -	43	48.86		
res499: K	18	20.45		
res499: R	12	13.63		
res499: E	6	6.81		
res499: Q	6	6.81		
res499: S	2	2.27		
res499: D	1	1.13		
wt 445: E	28.4		-	48.86
res500: -	43	48.86		
res500: E	25	28.4		
res500: Q	6	6.81		
res500: K	4	4.54		
res500: D	2	2.27		
res500: G	2	2.27		
res500: S	2	2.27		
res500: I	1	1.13		
res500: A	1	1.13		
res500: T	1	1.13		
res500: P	1	1.13		
wt 446: K	10.22		-	53.4
res501: -	47	53.4		
res501: R	10	11.36		
res501: K	9	10.22		
res501: T	5	5.68		
res501: M	5	5.68		
res501: I	4	4.54		
res501: Q	2	2.27		
res501: P	2	2.27		
res501: L	1	1.13		
res501: V	1	1.13		
res501: S	1	1.13		
res501: H	1	1.13		

wt 447: A	9.09	-	56.81
res502: -	50	56.81	
res502: K	10	11.36	
res502: A	8	9.09	
res502: S	5	5.68	
res502: V	4	4.54	
res502: F	2	2.27	
res502: D	2	2.27	
res502: L	1	1.13	
res502: I	1	1.13	
res502: Q	1	1.13	
res502: N	1	1.13	
res502: T	1	1.13	
res502: R	1	1.13	
res502: P	1	1.13	

wt 448: D	17.04	-	63.63
res503: -	56	63.63	
res503: D	15	17.04	
res503: T	9	10.22	
res503: N	2	2.27	
res503: P	2	2.27	
res503: R	1	1.13	
res503: H	1	1.13	
res503: K	1	1.13	
res503: E	1	1.13	

wt 449: V	12.5	-	63.63
res504: -	56	63.63	
res504: V	11	12.5	
res504: A	7	7.95	
res504: I	6	6.81	
res504: L	3	3.4	
res504: M	2	2.27	
res504: P	1	1.13	
res504: N	1	1.13	
res504: E	1	1.13	

wt 450: K	9.09	-	63.63
res505: -	56	63.63	
res505: D	11	12.5	
res505: K	8	9.09	
res505: N	3	3.4	
res505: E	3	3.4	
res505: T	2	2.27	
res505: L	2	2.27	
res505: H	1	1.13	
res505: R	1	1.13	
res505: Q	1	1.13	

wt 451: D	6.81	-	63.63
res506: -	56	63.63	
res506: G	11	12.5	
res506: N	7	7.95	
res506: D	6	6.81	
res506: E	2	2.27	
res506: T	1	1.13	
res506: H	1	1.13	
res506: R	1	1.13	
res506: S	1	1.13	
res506: I	1	1.13	
res506: L	1	1.13	

wt 452: T	5.68	-	65.9
res507: -	58	65.9	
res507: C	8	9.09	
res507: I	7	7.95	
res507: V	6	6.81	
res507: T	5	5.68	
res507: S	2	2.27	
res507: P	1	1.13	
res507: D	1	1.13	

wt 453: D	4.54	-	75
res508: -	66	75	
res508: H	5	5.68	
res508: D	4	4.54	
res508: S	3	3.4	
res508: Q	2	2.27	
res508: G	2	2.27	
res508: R	2	2.27	
res508: E	1	1.13	
res508: T	1	1.13	
res508: K	1	1.13	
res508: M	1	1.13	

wt 454: A	5.68	-	75
res509: -	66	75	
res509: A	5	5.68	
res509: T	4	4.54	
res509: D	2	2.27	
res509: L	2	2.27	
res509: G	2	2.27	
res509: S	2	2.27	
res509: F	1	1.13	
res509: Y	1	1.13	
res509: E	1	1.13	
res509: K	1	1.13	
res509: R	1	1.13	

wt 455: V	15.9	-	77.27
res518: -	68	77.27	
res518: V	14	15.9	
res518: I	3	3.4	
res518: N	1	1.13	
res518: E	1	1.13	
res518: L	1	1.13	

wt 456: D	4.54	-	78.4
res519: -	69	78.4	
res519: N	10	11.36	
res519: D	4	4.54	
res519: S	2	2.27	
res519: H	1	1.13	
res519: Y	1	1.13	
res519: K	1	1.13	

wt 457: V	19.31	-	78.4
res520: -	69	78.4	
res520: V	17	19.31	
res520: A	1	1.13	
res520: I	1	1.13	

wt 458: C	20.45	-	78.4
res521: -	69	78.4	
res521: C	18	20.45	
res521: V	1	1.13	

wt 459: I	9.09	-	78.4
res522: -	69	78.4	
res522: I	8	9.09	
res522: A	2	2.27	
res522: T	2	2.27	
res522: V	2	2.27	
res522: L	2	2.27	
res522: F	2	2.27	
res522: M	1	1.13	

wt 460: A	10.22	-	78.4
res523: -	69	78.4	
res523: A	9	10.22	
res523: G	2	2.27	
res523: L	2	2.27	
res523: V	2	2.27	
res523: P	2	2.27	
res523: S	1	1.13	
res523: T	1	1.13	

wt 461: D	7.95	-	78.4
res525: -	69	78.4	
res525: D	7	7.95	
res525: E	7	7.95	
res525: S	2	2.27	
res525: N	1	1.13	
res525: M	1	1.13	
res525: A	1	1.13	

wt 462: G	4.54	-	78.4
res526: -	69	78.4	
res526: D	7	7.95	
res526: N	6	6.81	
res526: G	4	4.54	
res526: T	2	2.27	

wt 463: V	10.22	-	78.4
res527: -	69	78.4	
res527: V	9	10.22	
res527: I	8	9.09	
res527: L	1	1.13	
res527: A	1	1.13	

wt 464: C	21.59	-	78.4
res528: -	69	78.4	
res528: C	19	21.59	

wt 465: I	18.18	-	78.4
res529: -	69	78.4	
res529: I	16	18.18	
res529: V	3	3.4	

wt 466: D	9.09	-	80.68
res530: -	71	80.68	
res530: D	8	9.09	
res530: K	4	4.54	
res530: N	3	3.4	
res530: E	2	2.27	

wt 467: A	13.63	-	80.68
res531: -	71	80.68	
res531: A	12	13.63	
res531: S	4	4.54	
res531: T	1	1.13	
wt 468: F	6.81	-	82.95
res532: -	73	82.95	
res532: F	6	6.81	
res532: S	2	2.27	
res532: V	1	1.13	
res532: N	1	1.13	
res532: P	1	1.13	
res532: K	1	1.13	
res532: L	1	1.13	
res532: D	1	1.13	
res532: E	1	1.13	
wt 469: L	10.22	-	82.95
res533: -	73	82.95	
res533: L	9	10.22	
res533: V	6	6.81	
wt 470: K	5.68	-	88.63
res534: -	78	88.63	
res534: K	5	5.68	
res534: E	2	2.27	
res534: N	2	2.27	
res534: G	1	1.13	
wt 471: P	7.95	-	88.63
res535: -	78	88.63	
res535: P	7	7.95	
res535: L	1	1.13	
res535: F	1	1.13	
res535: S	1	1.13	
wt 472: P	5.68	-	90.9
res536: -	80	90.9	
res536: P	5	5.68	
res536: Q	2	2.27	
res536: N	1	1.13	
wt 473: M	2.27	-	90.9
res537: -	80	90.9	
res537: M	2	2.27	
res537: L	1	1.13	
res537: S	1	1.13	
res537: P	1	1.13	
res537: V	1	1.13	
res537: A	1	1.13	
res537: K	1	1.13	
wt 474: E	4.54	-	90.9
res538: -	80	90.9	
res538: E	4	4.54	
res538: S	2	2.27	
res538: L	1	1.13	
res538: V	1	1.13	

wt 475: T	5.68	-	90.9
res539: -	80	90.9	
res539: T	5	5.68	
res539: D	2	2.27	
res539: E	1	1.13	
wt 476: E	5.68	-	90.9
res540: -	80	90.9	
res540: E	5	5.68	
res540: Y	1	1.13	
res540: H	1	1.13	
res540: D	1	1.13	
wt 477: E	3.4	-	90.9
res541: -	80	90.9	
res541: E	3	3.4	
res541: D	2	2.27	
res541: G	1	1.13	
res541: A	1	1.13	
res541: S	1	1.13	
wt 478: P	2.27	-	90.9
res542: -	80	90.9	
res542: S	4	4.54	
res542: P	2	2.27	
res542: D	1	1.13	
res542: F	1	1.13	
wt 479: Q	3.4	-	93.18
res550: -	82	93.18	
res550: Q	3	3.4	
res550: H	1	1.13	
res550: P	1	1.13	
res550: T	1	1.13	
wt 480: I	5.68	-	93.18
res551: -	82	93.18	
res551: I	5	5.68	
res551: T	1	1.13	
wt 481: F	5.68	-	93.18
res552: -	82	93.18	
res552: F	5	5.68	
res552: L	1	1.13	
wt 482: Y	4.54	-	93.18
res553: -	82	93.18	
res553: Y	4	4.54	
res553: S	2	2.27	
wt 483: N	5.68	-	93.18
res554: -	82	93.18	
res554: N	5	5.68	
res554: S	1	1.13	
wt 484: A	3.4	-	93.18
res555: -	82	93.18	
res555: A	3	3.4	
res555: T	2	2.27	
res555: I	1	1.13	

wt 485: S	5.68	-	93.18
res556: -	82	93.18	
res556: S	5	5.68	
res556: T	1	1.13	
wt 486: P	4.54	-	93.18
res557: -	82	93.18	
res557: P	4	4.54	
res557: S	2	2.27	
wt 487: S	5.68	-	93.18
res558: -	82	93.18	
res558: S	5	5.68	
res558: I	1	1.13	
wt 488: T	5.68	-	93.18
res559: -	82	93.18	
res559: T	5	5.68	
res559: N	1	1.13	
wt 489: L	3.4	-	93.18
res560: -	82	93.18	
res560: L	3	3.4	
res560: T	1	1.13	
res560: S	1	1.13	
res560: V	1	1.13	
wt 490: S	5.68	-	93.18
res561: -	82	93.18	
res561: S	5	5.68	
res561: T	1	1.13	
wt 491: A	4.54	-	93.18
res562: -	82	93.18	
res562: A	4	4.54	
res562: T	2	2.27	
wt 492: T	6.81	-	93.18
res563: -	82	93.18	
res563: T	6	6.81	
wt 493: M	5.68	-	93.18
res564: -	82	93.18	
res564: M	5	5.68	
res564: V	1	1.13	
wt 494: F	4.54	-	95.45
res565: -	84	95.45	
res565: F	4	4.54	
wt 495: I	4.54	-	95.45
res566: -	84	95.45	
res566: I	4	4.54	

APPENDIX D

Part 2: Non-Essential Residues

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
2	36	**1**	L
3	37	**2**	N
4	38	**3**	F
5	39	**4**	R
6	40	**5**	A
7	41	**6**	P
8	42	**7**	P
9	43	**8**	V
10	44	**9**	I
11	45	**10**	P
12	46	**11**	N
13	47	**12**	V
14	48	**13**	P
15	50	**15**	L
16	51	**16**	W
17	52	**17**	A
18	54	**19**	N
19	55	**20**	A
20	56	**21**	P
21	57	**22**	S
22	58	**23**	E
23	59	**24**	F
24	61	**26**	L
25	62	**27**	G
26	63	**28**	K
27	64	**29**	F
28	65	**30**	D
29	66	**31**	E
30	67	**32**	P
31	68	**33**	L
32	69	**34**	D
33	70	**35**	M
34	71	**36**	S
35	72	**37**	L
36	73	**38**	F
37	74	**39**	S
38	75	**40**	F
39	76	**41**	I
40	77	**42**	G
41	78	**43**	S
42	79	**44**	P
43	80	**45**	R
44	81	**46**	I
45	82	**47**	N

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
46	83	**48**	A
47	84	**49**	T
48	85	**50**	G
49	86	**51**	Q
50	87	**52**	G
51	88	**53**	V
52	89	**54**	T
53	90	**55**	I
54	93	**58**	V
55	94	**59**	D
56	95	**60**	R
57	96	**61**	L
58	98	**63**	Y
59	101	**66**	Y
60	102	**67**	I
61	103	**68**	D
62	104	**69**	S
63	105	**70**	I
64	106	**71**	T
65	107	**72**	G
66	108	**73**	V
67	109	**74**	T
68	110	**75**	V
69	111	**76**	N
70	114	**79**	I
71	117	**82**	K
72	118	**83**	I
73	119	**84**	S
74	120	**85**	L
75	121	**86**	Q
76	122	**87**	D
77	124	**89**	L
78	125	**90**	D
79	126	**91**	K
80	127	**92**	A
81	128	**93**	K
82	129	**94**	K
83	130	**95**	D
84	131	**96**	I
85	132	**97**	T
86	133	**98**	F
87	134	**99**	Y
88	135	**100**	M
89	136	**101**	P

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
90	137	**102**	V
91	138	**103**	D
92	139	**104**	N
93	140	**105**	L
94	142	**107**	M
95	143	**108**	A
96	145	**110**	I
97	149	**114**	E
98	151	**116**	R
99	153	**118**	T
100	155	**120**	A
101	156	**121**	R
102	159	**124**	K
103	160	**125**	P
104	161	**126**	K
105	162	**127**	D
106	163	**128**	V
107	165	**130**	K
108	166	**131**	N
109	167	**132**	R
110	169	**134**	I
111	170	**135**	E
112	171	**136**	L
113	172	**137**	V
114	173	**138**	Q
115	174	**139**	Q
116	175	**140**	Q
117	176	**141**	N
118	177	**142**	V
119	178	**143**	Q
120	179	**144**	L
121	180	**145**	S
122	181	**146**	L
123	182	**147**	T
124	183	**148**	E
125	184	**149**	A
126	185	**150**	T
127	186	**151**	E
128	187	**152**	K
129	189	**154**	K
130	190	**155**	Q
131	191	**156**	E
132	193	**158**	E
133	194	**159**	K

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
134	195	**160**	A
135	196	**161**	G
136	197	**162**	K
137	198	**163**	D
138	199	**164**	F
139	200	**165**	L
140	201	**166**	V
141	202	**167**	E
142	204	**169**	I
143	205	**170**	K
144	206	**171**	L
145	207	**172**	G
146	208	**173**	K
147	209	**174**	L
148	210	**175**	L
149	213	**178**	N
150	214	**179**	H
151	218	**183**	Y
152	220	**185**	L
153	221	**186**	F
154	223	**188**	D
155	225	**190**	Y
156	227	**192**	H
157	228	**193**	H
158	229	**194**	Y
159	230	**195**	K
160	231	**196**	K
161	232	**197**	P
162	233	**198**	G
163	235	**200**	N
164	237	**202**	S
165	239	**204**	F
166	240	**205**	N
167	241	**206**	V
168	242	**207**	E
169	243	**208**	I
170	244	**209**	K
171	245	**210**	R
172	247	**212**	D
173	248	**213**	D
174	250	**215**	S
175	252	**217**	L
176	254	**219**	N
177	255	**220**	E

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
178	257	**222**	T
179	260	**225**	Y
180	262	**227**	S
181	263	**228**	I
182	264	**229**	Y
183	265	**230**	L
184	266	**231**	N
185	267	**232**	T
186	268	**233**	Q
187	269	**234**	Q
188	270	**235**	S
189	271	**236**	P
190	272	**237**	V
191	273	**238**	A
192	274	**239**	A
193	275	**240**	T
194	276	**241**	L
195	277	**242**	Y
196	278	**243**	V
197	279	**244**	R
198	280	**245**	N
199	282	**247**	V
200	283	**248**	R
201	285	**250**	A
202	286	**251**	I
203	288	**253**	V
204	289	**254**	S
205	290	**255**	K
206	291	**256**	I
207	292	**257**	P
208	293	**258**	D
209	294	**259**	A
210	295	**260**	K
211	296	**261**	S
212	297	**262**	P
213	298	**263**	L
214	300	**265**	V
215	301	**266**	F
216	302	**267**	A
217	303	**268**	Y
218	304	**269**	T
219	305	**270**	R
220	306	**271**	I
221	307	**272**	V

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
222	308	**273**	F
223	309	**274**	T
224	310	**275**	D
225	311	**276**	Q
226	312	**277**	V
227	313	**278**	L
228	314	**279**	K
229	315	**280**	F
230	317	**282**	S
231	318	**283**	Q
232	319	**284**	D
233	320	**285**	E
234	322	**287**	V
235	323	**288**	Y
236	324	**289**	T
237	325	**290**	F
238	327	**292**	E
239	328	**293**	T
240	329	**294**	V
241	330	**295**	A
242	331	**296**	L
243	333	**298**	A
244	334	**299**	S
245	336	**301**	I
246	337	**302**	V
247	338	**303**	I
248	340	**305**	G
249	341	**306**	T
250	342	**307**	L
251	343	**308**	S
252	344	**309**	I
253	345	**310**	M
254	346	**311**	R
255	347	**312**	S
256	348	**313**	M
257	349	**314**	K
258	350	**315**	S
259	352	**317**	L
260	353	**318**	L
261	354	**319**	L
262	355	**320**	D
263	356	**321**	N
264	357	**322**	Y
265	358	**323**	M

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
266	359	**324**	E
267	360	**325**	T
268	361	**326**	I
269	363	**328**	N
270	364	**329**	P
271	365	**330**	Y
272	366	**331**	I
273	367	**332**	I
274	370	**335**	T
275	371	**336**	L
276	372	**337**	A
277	373	**338**	A
278	374	**339**	K
279	375	**340**	M
280	378	**343**	Q
281	379	**344**	V
282	380	**345**	L
283	382	**347**	Q
284	383	**348**	E
285	384	**349**	Q
286	386	**351**	V
287	388	**353**	I
288	389	**354**	R
289	390	**355**	K
290	391	**356**	N
291	392	**357**	W
292	393	**358**	N
293	394	**359**	S
294	395	**360**	S
295	396	**361**	D
296	397	**362**	Y
297	399	**364**	H
298	400	**365**	L
299	401	**366**	N
300	402	**367**	P
301	403	**368**	D
302	404	**369**	N
303	405	**370**	F
304	406	**371**	A
305	407	**372**	I
306	408	**373**	Q
307	409	**374**	L
308	410	**375**	E
309	411	**376**	K

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
310	412	**377**	G
311	413	**378**	G
312	414	**379**	K
313	415	**380**	F
314	416	**381**	T
315	417	**382**	V
316	418	**383**	R
317	419	**384**	G
318	420	**385**	K
319	421	**386**	P
320	422	**387**	T
321	423	**388**	L
322	424	**389**	E
323	425	**390**	D
324	426	**391**	L
325	427	**392**	E
326	428	**393**	Q
327	429	**394**	F
328	430	**395**	S
329	431	**396**	E
330	432	**397**	K
331	433	**398**	F
332	434	**399**	Y
333	435	**400**	C
334	436	**401**	S
335	437	**402**	C
336	438	**403**	Y
337	439	**404**	S
338	440	**405**	T
339	441	**406**	L
340	442	**407**	S
341	443	**408**	C
342	444	**409**	K
343	445	**410**	E
344	446	**411**	K
345	447	**412**	A
346	448	**413**	D
347	449	**414**	V
348	450	**415**	K
349	451	**416**	D
350	452	**417**	T
351	453	**418**	D
352	454	**419**	A
353	455	**420**	V

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
354	456	**421**	D
355	457	**422**	V
356	458	**423**	C
357	459	**424**	I
358	460	**425**	A
359	461	**426**	D
360	462	**427**	G
361	463	**428**	V
362	464	**429**	C
363	465	**430**	I
364	466	**431**	D
365	467	**432**	A
366	468	**433**	F
367	469	**434**	L
368	470	**435**	K
369	471	**436**	P
370	472	**437**	P
371	473	**438**	M
372	474	**439**	E
373	475	**440**	T
374	476	**441**	E
375	477	**442**	E
376	478	**443**	P
377	479	**444**	Q
378	480	**445**	I
379	481	**446**	F
380	482	**447**	Y

APPENDIX D

Part 3: Essential Residues

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
2	49	**14**	F
3	53	**18**	W
4	60	**25**	C
5	91	**56**	F
6	92	**57**	Y
7	97	**62**	G
8	99	**64**	Y
9	100	**65**	P
10	112	**77**	G
11	113	**78**	G
12	115	**80**	P
13	116	**81**	Q
14	123	**88**	H
15	141	**106**	G
16	144	**109**	V
17	146	**111**	D
18	147	**112**	W
19	148	**113**	E
20	150	**115**	W
21	152	**117**	P
22	154	**119**	W
23	157	**122**	N
24	158	**123**	W
25	164	**129**	Y
26	168	**133**	S
27	188	**153**	A
28	192	**157**	F
29	203	**168**	T
30	211	**176**	R
31	212	**177**	P
32	215	**180**	L
33	216	**181**	W
34	217	**182**	G
35	219	**184**	Y
36	222	**187**	P
37	224	**189**	C
38	226	**191**	N
39	234	**199**	Y
40	236	**201**	G
41	238	**203**	C
42	246	**211**	N
43	249	**214**	L
44	251	**216**	W
45	253	**218**	W

	A	B	C
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**
46	256	**221**	S
47	258	**223**	A
48	259	**224**	L
49	261	**226**	P
50	281	**246**	R
51	284	**249**	E
52	287	**252**	R
53	299	**264**	P
54	316	**281**	L
55	321	**286**	L
56	326	**291**	G
57	332	**297**	G
58	335	**300**	G
59	339	**304**	W
60	351	**316**	C
61	362	**327**	L
62	368	**333**	N
63	369	**334**	V
64	376	**341**	C
65	377	**342**	S
66	381	**346**	C
67	385	**350**	G
68	387	**352**	C
69	398	**363**	L

APPENDIX D

Part 4: PH20 & HYAL1

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
2	39	**4**	T	22	F
3	40	**5**	A	23	R
4	41	**6**	P	24	G
5	42	**7**	P	25	P
6	43	**8**	V	26	L
7	44	**9**	I	27	L
8	45	**10**	P	28	P
9	46	**11**	N	29	N
10	47	**12**	V	30	R
11	48	**13**	P	31	P
12	49	**14**	F	32	F
13	50	**15**	L	33	T
14	51	**16**	W	34	T
15	52	**17**	A	35	V
16	53	**18**	W	36	W
17	54	**19**	N	37	N
18	55	**20**	A	38	A
19	56	**21**	P	39	N
20	57	**22**	S	40	T
21	58	**23**	E	41	Q
22	59	**24**	F	42	W
23	60	**25**	C	43	C
24	61	**26**	L	44	L
25	62	**27**	G	45	E
26	63	**28**	K	46	R
27	64	**29**	F	47	H
28	65	**30**	D	48	G
29	66	**31**	E	49	V
30	67	**32**	P	50	D
31	68	**33**	L	51	V
32	69	**34**	D	52	D
33	70	**35**	M	53	V
34	71	**36**	S	54	S
35	72	**37**	L	55	V
36	73	**38**	F	56	F
37	74	**39**	S	57	D
38	75	**40**	F	58	V
39	76	**41**	I	59	V
40	77	**42**	G	60	A
41	78	**43**	S	61	N
42	79	**44**	P	62	P
43	80	**45**	R	63	G
44	81	**46**	I	64	Q
45	82	**47**	N	65	T

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
46	83	**48**	A	66	F
47	84	**49**	T	67	R
48	85	**50**	G	68	G
49	86	**51**	Q	69	P
50	87	**52**	G	70	D
51	88	**53**	V	71	M
52	89	**54**	T	72	T
53	90	**55**	I	73	I
54	91	**56**	F	74	F
55	92	**57**	Y	75	Y
56	93	**58**	V	76	S
57	94	**59**	D	77	S
58	95	**60**	R	78	Q
59	96	**61**	L	79	L
60	97	**62**	G	80	G
61	98	**63**	Y	81	T
62	99	**64**	Y	82	Y
63	100	**65**	P	83	P
64	101	**66**	Y	84	Y
65	102	**67**	I	85	Y
66	103	**68**	D	86	T
67	104	**69**	S	87	P
68	105	**70**	I	87	P
69	106	**71**	T	88	T
70	107	**72**	G	89	G
71	108	**73**	V	90	E
72	109	**74**	T	91	P
73	110	**75**	V	92	V
74	111	**76**	N	93	F
75	112	**77**	G	94	G
76	113	**78**	G	95	G
77	114	**79**	I	96	L
78	115	**80**	P	97	P
79	116	**81**	Q	98	Q
80	117	**82**	K	99	N
81	118	**83**	I	100	A
82	119	**84**	S	101	S
83	120	**85**	L	102	L
84	121	**86**	Q	103	I
85	122	**87**	D	104	A
86	123	**88**	H	105	H
87	124	**89**	L	106	L
88	125	**90**	D	107	A
89	126	**91**	K	108	R

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
90	127	**92**	A	109	T
91	128	**93**	K	110	F
92	129	**94**	K	111	Q
93	130	**95**	D	112	D
94	131	**96**	I	113	I
95	132	**97**	T	114	L
96	133	**98**	F	115	A
97	134	**99**	Y	116	A
98	135	**100**	M	117	I
99	136	**101**	P	118	P
100	137	**102**	V	119	A
101	138	**103**	D	120	P
102	139	**104**	N	122	F
103	140	**105**	L	123	S
104	141	**106**	G	124	G
105	142	**107**	M	125	L
106	143	**108**	A	126	A
107	144	**109**	V	127	V
108	145	**110**	I	128	I
109	146	**111**	D	129	D
110	147	**112**	W	130	W
111	148	**113**	E	131	E
112	149	**114**	E	132	A
113	150	**115**	W	133	W
114	151	**116**	R	134	R
115	152	**117**	P	135	P
116	153	**118**	T	136	R
117	154	**119**	W	137	W
118	155	**120**	A	138	A
119	156	**121**	R	139	F
120	157	**122**	N	140	N
121	158	**123**	W	141	W
122	159	**124**	K	142	D
123	160	**125**	P	143	T
124	161	**126**	K	144	K
125	162	**127**	D	145	D
126	163	**128**	V	146	I
127	164	**129**	Y	147	Y
128	165	**130**	K	148	R
129	166	**131**	N	149	Q
130	167	**132**	R	150	R
131	168	**133**	S	151	S
132	169	**134**	I	152	R
133	170	**135**	E	153	A

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
134	171	**136**	L	154	L
135	172	**137**	V	155	V
136	173	**138**	Q	156	Q
137	174	**139**	Q	157	A
138	175	**140**	Q	158	Q
139	176	**141**	N	159	H
140	177	**142**	V	160	P
141	178	**143**	Q	161	D
142	179	**144**	L	162	W
143	180	**145**	S	163	P
144	181	**146**	L	164	A
145	182	**147**	T	165	P
146	183	**148**	E	166	Q
147	184	**149**	A	167	V
148	185	**150**	T	168	E
149	186	**151**	E	169	A
150	187	**152**	K	170	V
151	188	**153**	A	171	A
152	189	**154**	K	172	Q
153	190	**155**	Q	173	D
154	191	**156**	E	174	Q
155	192	**157**	F	175	F
156	193	**158**	E	176	Q
157	194	**159**	K	177	G
158	195	**160**	A	178	A
159	196	**161**	G	179	A
160	197	**162**	K	180	R
161	198	**163**	D	181	A
162	199	**164**	F	182	W
163	200	**165**	L	183	M
164	201	**166**	V	184	A
165	202	**167**	E	185	G
166	203	**168**	T	186	T
167	204	**169**	I	187	L
168	205	**170**	K	188	Q
169	206	**171**	L	189	L
170	207	**172**	G	190	G
171	208	**173**	K	191	R
172	209	**174**	L	192	A
173	210	**175**	L	193	L
174	211	**176**	R	194	R
175	212	**177**	P	195	P
176	213	**178**	N	196	R
177	214	**179**	H	197	G

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
178	215	**180**	L	198	L
179	216	**181**	W	199	W
180	217	**182**	G	200	G
181	218	**183**	Y	201	F
182	219	**184**	Y	202	Y
183	220	**185**	L	203	G
184	221	**186**	F	204	F
185	222	**187**	P	205	P
186	223	**188**	D	206	D
187	224	**189**	C	207	C
188	225	**190**	Y	208	Y
189	226	**191**	N	209	N
190	227	**192**	H	210	Y
191	228	**193**	H	211	D
192	229	**194**	Y	212	F
193	230	**195**	K	213	L
194	231	**196**	K	214	S
195	232	**197**	P	215	P
196	233	**198**	G	216	N
197	234	**199**	Y	217	Y
198	235	**200**	N	218	T
199	236	**201**	G	219	G
200	237	**202**	S	220	Q
201	238	**203**	C	221	C
202	239	**204**	F	222	P
203	240	**205**	N	223	S
204	241	**206**	V	224	G
205	242	**207**	E	225	I
206	243	**208**	I	226	R
207	244	**209**	K	227	A
208	245	**210**	R	228	Q
209	246	**211**	N	229	N
210	247	**212**	D	230	D
211	248	**213**	D	231	Q
212	249	**214**	L	232	L
213	250	**215**	S	233	G
214	251	**216**	W	234	W
215	252	**217**	L	235	L
216	253	**218**	W	236	W
217	254	**219**	N	237	G
218	255	**220**	E	238	Q
219	256	**221**	S	239	S
220	257	**222**	T	240	R
221	258	**223**	A	241	A

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
222	259	**224**	L	242	L
223	260	**225**	Y	243	Y
224	261	**226**	P	244	P
225	262	**227**	S	245	S
226	263	**228**	I	246	I
227	264	**229**	Y	247	Y
228	265	**230**	L	248	M
229	266	**231**	N	249	P
230	267	**232**	T	250	A
231	268	**233**	Q	251	V
232	269	**234**	Q	252	L
233	270	**235**	S	254	G
234	271	**236**	P	255	T
235	272	**237**	V	256	G
236	273	**238**	A	257	K
237	274	**239**	A	258	S
238	275	**240**	T	259	Q
239	276	**241**	L	260	M
240	277	**242**	Y	261	Y
241	278	**243**	V	262	V
242	279	**244**	R	263	Q
243	280	**245**	N	264	H
244	281	**246**	R	265	R
245	282	**247**	V	266	V
246	283	**248**	R	267	A
247	284	**249**	E	268	E
248	285	**250**	A	269	A
249	286	**251**	I	270	F
250	287	**252**	R	271	R
251	288	**253**	V	272	V
252	289	**254**	S	273	A
253	290	**255**	K	274	V
254	291	**256**	I	275	A
255	292	**257**	P	275	A
256	293	**258**	D	276	A
257	294	**259**	A	277	G
258	295	**260**	K	278	D
259	296	**261**	S	279	P
260	297	**262**	P	280	N
261	298	**263**	L	281	L
262	299	**264**	P	282	P
263	300	**265**	V	283	V
264	301	**266**	F	284	L
265	302	**267**	A	285	P

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
266	303	**268**	Y	286	Y
267	304	**269**	T	287	V
268	305	**270**	R	288	Q
269	306	**271**	I	289	I
270	307	**272**	V	290	F
271	308	**273**	F	291	Y
272	309	**274**	T	292	D
273	310	**275**	D	293	T
274	311	**276**	Q	294	T
275	312	**277**	V	294	T
276	313	**278**	L	295	N
277	314	**279**	K	296	H
278	315	**280**	F	297	F
279	316	**281**	L	298	L
280	317	**282**	S	299	P
281	318	**283**	Q	300	L
282	319	**284**	D	301	D
283	320	**285**	E	302	E
284	321	**286**	L	303	L
285	322	**287**	V	304	E
286	323	**288**	Y	305	H
287	324	**289**	T	306	S
288	325	**290**	F	307	L
289	326	**291**	G	308	G
290	327	**292**	E	309	E
291	328	**293**	T	310	S
292	329	**294**	V	311	A
293	330	**295**	A	312	A
294	331	**296**	L	313	Q
295	332	**297**	G	314	G
296	333	**298**	A	315	A
297	334	**299**	S	316	A
298	335	**300**	G	317	G
299	336	**301**	I	318	V
300	337	**302**	V	319	V
301	338	**303**	I	320	L
302	339	**304**	W	321	W
303	340	**305**	G	322	V
304	341	**306**	T	323	S
305	342	**307**	L	324	W
306	343	**308**	S	325	E
307	344	**309**	I	326	N
308	345	**310**	M	327	T
309	346	**311**	R	328	R

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
310	347	**312**	S	329	T
311	348	**313**	M	330	K
312	349	**314**	K	331	E
313	350	**315**	S	332	S
314	351	**316**	C	333	C
315	352	**317**	L	334	Q
316	353	**318**	L	335	A
317	354	**319**	L	336	I
318	355	**320**	D	337	K
319	356	**321**	N	338	E
320	357	**322**	Y	339	Y
321	358	**323**	M	340	M
322	359	**324**	E	341	D
323	360	**325**	T	342	T
324	361	**326**	I	343	T
325	362	**327**	L	344	L
326	363	**328**	N	345	G
327	364	**329**	P	346	P
328	365	**330**	Y	347	F
329	366	**331**	I	348	I
330	367	**332**	I	349	L
331	368	**333**	N	350	N
332	369	**334**	V	351	V
333	370	**335**	T	352	T
334	371	**336**	L	353	S
335	372	**337**	A	354	G
336	373	**338**	A	355	A
337	374	**339**	K	356	L
338	375	**340**	M	357	L
339	376	**341**	C	358	C
340	377	**342**	S	359	S
341	378	**343**	Q	360	Q
342	379	**344**	V	361	A
343	380	**345**	L	362	L
344	381	**346**	C	363	C
345	382	**347**	Q	364	S
346	383	**348**	E	365	G
347	384	**349**	Q	366	H
348	385	**350**	G	367	G
349	386	**351**	V	368	R
350	387	**352**	C	369	C
351	388	**353**	I	370	V
352	389	**354**	R	371	R
353	390	**355**	K	372	R

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
354	391	**356**	N	373	T
355	392	**357**	W	374	S
356	393	**358**	N	375	H
357	394	**359**	S	376	P
358	395	**360**	S	377	K
359	396	**361**	D	378	A
360	397	**362**	Y	379	L
361	398	**363**	L	380	L
362	399	**364**	H	381	L
363	400	**365**	L	382	L
364	401	**366**	N	383	N
365	402	**367**	P	384	P
366	403	**368**	D	385	A
367	404	**369**	N	386	S
368	405	**370**	F	387	F
369	406	**371**	A	388	S
370	407	**372**	I	389	I
371	408	**373**	Q	390	Q
372	409	**374**	L	391	L
373	410	**375**	E	392	T
374	411	**376**	K	394	G
375	412	**377**	G	395	G
376	413	**378**	G	396	G
377	414	**379**	K	397	P
378	415	**380**	F	398	L
379	416	**381**	T	399	S
380	417	**382**	V	400	L
381	418	**383**	R	401	R
382	419	**384**	G	402	G
383	420	**385**	K	403	A
384	421	**386**	P	404	L
385	422	**387**	T	405	S
386	423	**388**	L	406	L
387	424	**389**	E	407	E
388	425	**390**	D	408	D
389	426	**391**	L	409	Q
390	427	**392**	E	410	A
391	428	**393**	Q	411	Q
392	429	**394**	F	412	M
393	430	**395**	S	413	A
394	431	**396**	E	414	V
395	432	**397**	K	415	E
396	433	**398**	F	416	F
397	434	**399**	Y	417	K

	A	B	C	D	E
1	**PH20 Residue #**	**Mature PH20 Residue #**	**PH20 Residue**	**Hyal1 Residue #**	**Hyal1 Residue**
398	435	**400**	C	418	C
399	436	**401**	S	419	R
400	437	**402**	C	420	C
401	438	**403**	Y	421	Y
402	439	**404**	S	422	P
403	440	**405**	T	423	G
404	441	**406**	L	424	W
405	442	**407**	S	425	Q
406	443	**408**	C	426	A
407	444	**409**	K	427	P
408	445	**410**	E	428	W
409	446	**411**	K	429	C
410	447	**412**	A	430	E
411	448	**413**	D	431	R
412	449	**414**	V	432	K
413	450	**415**	K	433	S
414	451	**416**	D	434	M
415	452	**417**	T	435	W

APPENDIX D

Part 5: SASA

Residue	Max SASA	Med SASA	Med SASA fraction
A	111	14	0.126
C	157	5	0.032
D	160	62	0.388
E	187	83	0.444
F	208	13	0.063
G	86	19	0.221
H	191	46	0.241
I	173	6	0.035
K	212	102	0.481
L	179	9	0.050
M	201	13	0.065
N	166	59	0.355
P	135	49	0.363
Q	194	74	0.381
R	250	87	0.348
S	125	35	0.280
T	144	37	0.257
V	149	8	0.054
W	249	25	0.100
Y	227	31	0.137

Analysis of accessible surface of residues in proteins

LAURENCE LINS, ANNICK THOMAS, AND ROBERT BRASSEUR
Centre de Biophysique Moléculaire Numérique (CBMN), Faculté des Sciences Agronomiques de Gembloux (FSAGx), 5030 Gembloux, Belgium

(RECEIVED January 31, 2003; FINAL REVISION March 26, 2003; ACCEPTED March 28, 2003)

based on 587 structures

APPENDIX E

Part 1: Highlight.py Script

```
################################################
# Given a resi, it zooms to the residue and show its neighbors
################################################

import numpy as np
from math import sqrt

#from __future__ import print_function
#from pymol import cmd

# mapping 3 letter code to 1 letter code
aa3 =
('ALA','CYS','ASP','GLU','PHE','GLY','HIS','ILE','LYS','LEU','MET','ASN','PRO','GLN
','ARG','SER','THR','VAL','TRP','TYR')
aa1 =
('A','C','D','E','F','G','H','I','K','L','M','N','P','Q','R','S','T','V','W','Y')
aa31 = zip(aa3,aa1)
a321 = dict()
for x in aa31:
    a321[x[0]] = x[1]

# max surface area
aa_max =
[111,157,160,187,208,86,191,173,212,179,201,166,135,194,250,125,144,149,249,227]
aa_mean =
[0.126,0.032,0.388,0.444,0.063,0.221,0.241,0.035,0.481,0.05,0.065,0.355,0.363,0.381
,0.348,0.28,0.257,0.054,0.1,0.137]
aa_area = zip(aa3,aa_max,aa_mean)
aa2area = dict()
aa2mean = dict()
for x in aa_area:
    aa2area[x[0]] = x[1]
    aa2mean[x[0]] = x[2]


try:
    cmd.set_name('ph20_with_ha4', 'ph20')
    cmd.set_name('2pe4','hyal')
    cmd.set_name('1fcv','bvh')
except:
    print('nothing to change')

map2hyal = dict()
f = open('ph20_to_hyal1.v4.map','r')
for line in f.readlines():
    ph20 = line.strip().split()[0]
    hyal = line.strip().split()[2]
    map2hyal[ph20] = hyal

map2bvh = dict()
```

```
f = open('ph20_to_bvh.map','r')
for line in f.readlines():
    ph20 = line.strip().split()[0]
    bvh = line.strip().split()[2]
    map2bvh[ph20] = bvh


def highlight(ph20resi, model=0, show_line=False):
#    cmd.select(f'br. * within 5 of resi {resi}')
    modelname = ''
    model = int(model)
    if model == 0: # ph20
        print('model: ph20')
        modelname = 'ph20'
        cmd.disable('*')
        cmd.enable('ph20')
        resi = ph20resi
    elif model == 1: # hyal1
        print('model: hyal1')
        modelname = 'hyal'
        if not ph20resi in map2hyal.keys():
            print('No matching residue\n')
            return
        resi = map2hyal[ph20resi]
        print(f'Showing ph20 {ph20resi} matching {modelname} {resi}\n')
        cmd.disable('*')
        cmd.enable('hyal')
    elif model == 2: # bvh
        print('model: bvh')
        modelname = 'bvh'
        if not ph20resi in map2bvh.keys():
            print('No matching residue\n')
            return
        resi = map2bvh[ph20resi]
        print(f'Showing ph20 {ph20resi} matching {modelname} {resi}\n')
        cmd.disable('*')
        cmd.enable('bvh')
    else: #default is ph20
        print('model: ph20')
        modelname = 'ph20'
        cmd.disable('*')
        cmd.enable('ph20')
        resi = ph20resi
        print(f'Showing ph20 {ph20resi} matching {modelname} {resi}\n')

    cmd.set('dot_solvent',1)
    area = cmd.get_area(f'{modelname} and resi {resi}')
    aaname = cmd.get_model(f'{modelname} and resi {resi}').atom[0].resn
    print('Surface area: {:.2f} (out of {}), relative area: {:.2f} (mean
{})'.format(area,aa2area[aaname],area/aa2area[aaname],aa2mean[aaname]))
```

```
    cmd.select(f'{modelname} and resi {resi} and not name c+ca+n+o')
    a = [i for i in cmd.get_model('sele').atom]
    if len(a) == 0: # gly
        cmd.select(f'{modelname} and resi {resi} and name ca')

    cmd.select('neigh',f'(br. {modelname} near_to 5 of sele) and not sele')
    cmd.create('pock','neigh')
    cmd.color('salmon','pock')

    cmd.select(f'br. {modelname} near_to 5 of sele')
    cmd.set('sphere_scale','0.3')

    cmd.hide('line','ph20')
    cmd.hide('line','hyal')
    cmd.hide('line','bvh')
    cmd.hide('sphere','ph20')
    cmd.hide('sphere','hyal')
    cmd.hide('sphere','bvh')
    cmd.hide('stick','polymer')

    cmd.show('cartoon')
    cmd.hide('cartoon','pock')
    cmd.show('surface','pock')
    cmd.disable('pock')

    cmd.show('stick','sele')
    cmd.center(f'{modelname} and resi {resi}')
    cmd.show('stick','not polymer')

    if show_line:
        cmd.show('line')
    cmd.show('sphere',f'resi {resi}')

    a = []
    a = [[x.resi, x.resn] for x in cmd.get_model('sele and n. CA').atom]
    print('# neigh: {}'.format(len(a)-1))
    for i,x in enumerate(a):
        print(f'{i+1}:\t{x[0]}\t{a321[x[1]]}')

    # measure distance to neighbors
    cutoff = 5
    mc = 0 # do not include main chain
    vdw_check(resi,cutoff,modelname)


def vdw_check(resi, cutoff, modelname, mc=0):

    """
    if modelname == 'ph20':
```

```
        cmd.disable('*')
        cmd.enable('ph20')
    elif modelname == 'hyal':
        cmd.disable('*')
        cmd.enable('hyal')
    elif modelname == 'bvh':
        cmd.disable('*')
        cmd.enable('bvh')
    else:
        print('unable to choose')
        return
    """

    cmd.select(f'{modelname} and resi {resi} and not name c+ca+n+o')
    a = [i for i in cmd.get_model('sele').atom]
    if len(a) == 0: # gly
        cmd.select(f'{modelname} and resi {resi} and name ca')

    pick = [i for i in cmd.get_model('sele').atom]
    if len(pick) == 0: # gly
        cmd.select(f'{modelname} and resi {resi} and name ca')

    neigh = [i for i in cmd.get_model(f'{modelname} near_to 5 of sele').atom]

    dist = np.empty((0,5))
    for a in pick:
        for b in neigh:
            if a.resi == b.resi:
                continue
            if mc != 0:
                next
            elif b.name == 'CA' or b.name == 'N' or b.name == 'C' or b.name =='O':
                continue
            d = distance(a,b)
            if d < cutoff:
                try:
                    dist = np.vstack([dist,[a.name, b.resi, a321[b.resn], b.name,
d]])
                except:
                    next
    sorted_dist = dist[dist[:,4].argsort()]
    for i,r in enumerate(sorted_dist):
        print(f'{i}. ', r[0], '-->', '/'.join(r[1:4]), ': ', r[4])



def distance(a1, a2):
    x1 = a1.coord[0]
    y1 = a1.coord[1]
    z1 = a1.coord[2]
```

```
    x2 = a2.coord[0]
    y2 = a2.coord[1]
    z2 = a2.coord[2]

    return sqrt((x1-x2)**2 + (y1-y2)**2 + (z1-z2)**2)


cmd.extend('highlight', highlight)
cmd.extend('vdw',vdw_check)
```

APPENDIX E

Part 2: Highlight_mt.py Script

```
################################################
# Given a resi, it zooms to the residue and show its neighbors
################################################

import numpy as np
from math import sqrt

#from __future__ import print_function
#from pymol import cmd

# mapping 3 letter code to 1 letter code
aa3 =
('ALA','CYS','ASP','GLU','PHE','GLY','HIS','ILE','LYS','LEU','MET','ASN','PRO','GLN
','ARG','SER','THR','VAL','TRP','TYR')
aa1 =
('A','C','D','E','F','G','H','I','K','L','M','N','P','Q','R','S','T','V','W','Y')
aa31 = zip(aa3,aa1)
a321 = dict()
for x in aa31:
    a321[x[0]] = x[1]

# max surface area
aa_max =
[111,157,160,187,208,86,191,173,212,179,201,166,135,194,250,125,144,149,249,227]
aa_mean =
[0.126,0.032,0.388,0.444,0.063,0.221,0.241,0.035,0.481,0.05,0.065,0.355,0.363,0.381
,0.348,0.28,0.257,0.054,0.1,0.137]
aa_area = zip(aa3,aa_max,aa_mean)
aa2area = dict()
aa2mean = dict()
for x in aa_area:
    aa2area[x[0]] = x[1]
    aa2mean[x[0]] = x[2]

def highlight_mt(ph20resi,show_line=False):
    modelname = 'ph20mt'
    print('model: ph20mt')
    resi = ph20resi
    print(f'Showing ph20 {ph20resi}\n')

    cmd.set('dot_solvent',1)
    area = cmd.get_area(f'{modelname} and resi {resi}')
    aaname = cmd.get_model(f'{modelname} and resi {resi}').atom[0].resn
    print('Surface area: {:.2f} (out of {}), relative area: {:.2f} (mean
{})'.format(area,aa2area[aaname],area/aa2area[aaname],aa2mean[aaname]))

    cmd.select(f'{modelname} and resi {resi} and not name c+ca+n+o')
    a = [i for i in cmd.get_model('sele').atom]
    if len(a) == 0: # gly
        cmd.select(f'{modelname} and resi {resi} and name ca')
```

```
    cmd.select('neigh',f'(br. {modelname} near_to 5 of sele) and not sele')
    cmd.create('pock','neigh')

    cmd.select(f'br. {modelname} near_to 5 of sele')
    cmd.set('sphere_scale','0.3')

    cmd.hide('line','ph20mt')
    cmd.hide('sphere','ph20mt')
    cmd.hide('stick','polymer')

    cmd.show('surface','pock')

    cmd.show('cartoon','ph20mt')
    cmd.show('stick','sele')
    cmd.show('stick','not polymer')
    cmd.color('cyan','ph20mt')
    cmd.color('yellow','HA4')

    cmd.center(f'{modelname} and resi {resi}')

    if show_line:
        cmd.show('line','ph20mt')
    cmd.show('sphere',f'resi {resi}')

    cmd.disable('*')
    cmd.enable('ph20mt')
    cmd.enable('HA4')
    pymol.util.cnc()
    cmd.color('salmon','pock')

    a = []
    a = [[x.resi, x.resn] for x in cmd.get_model('sele and n. CA').atom]
    print('# neigh: {}'.format(len(a)-1))
    for i,x in enumerate(a):
        print(f'{i+1}:\t{x[0]}\t{a321[x[1]]}')

    # measure distance to neighbors
    cutoff = 5
    mc = 0 # do not include main chain
    vdw_check(resi,cutoff,modelname,0)

    cmd.set('label_size',28)

def vdw_check(resi, cutoff, modelname, mc=0):

    print('Computing neighbors')
    cmd.select(f'{modelname} and resi {resi} and not name c+ca+n+o')
    a = [i for i in cmd.get_model('sele').atom]
    if len(a) == 0: # gly
```

```
        cmd.select(f'{modelname} and resi {resi} and name ca')

    pick = [i for i in cmd.get_model('sele').atom]
    if len(pick) == 0: # gly
        cmd.select(f'{modelname} and resi {resi} and name ca')

    neigh = [i for i in cmd.get_model(f'{modelname} near_to 5 of sele').atom]

    dist = np.empty((0,5))
    for a in pick:
        for b in neigh:
            if a.resi == b.resi:
                continue
            if mc != 0:
                next
            elif b.name == 'CA' or b.name == 'N' or b.name == 'C' or b.name =='O':
                continue
#               print('including mc')
            d = distance(a,b)
            if d < cutoff:
                try:
                    dist = np.vstack([dist,[a.name, b.resi, a321[b.resn], b.name,
d]])
                except:
                    next
    sorted_dist = dist[dist[:,4].argsort()]
    for i, r in enumerate(sorted_dist):
        print(f'{i}. ', r[0], '-->', '/'.join(r[1:4]), ': ', r[4])

def distance(a1, a2):
    x1 = a1.coord[0]
    y1 = a1.coord[1]
    z1 = a1.coord[2]

    x2 = a2.coord[0]
    y2 = a2.coord[1]
    z2 = a2.coord[2]

    return sqrt((x1-x2)**2 + (y1-y2)**2 + (z1-z2)**2)

def ph20mt(pid):
    cmd.reinitialize()
    cmd.delete('*')
    fname = pid + '.pdb'
    import time
    time.sleep(0.5)
    cmd.load('ha4.pdb','HA4')
    cmd.load(fname,'ph20mt')
    cmd.color('cyan','ph20mt')
    cmd.color('yellow','HA4')
```

```
    pymol.util.cnc()


cmd.extend('ph20mt',ph20mt)
cmd.extend('highlight_mt', highlight_mt)
cmd.extend('vdw',vdw_check)
```

APPENDIX E

Part 3: PyMol GUI Window

Loading E359
Showing PH20 359

Surface area: 91.16 (out of 187), relative area:
0.49 (mean 0.444)
neigh: 6
1: 355 D
2: 356 N
3: 358 M
4: 359 E
5: 360 T
6: 409 L
7: 415 F
Computing neighbors
0. OE2 --> 409/L/CD2 : 3.3896130435875564
1. CD --> 409/L/CD2 : 3.8418346285759677
2. CG --> 409/L/CD2 : 4.316818131519594
3. OE1 --> 409/L/CD2 : 4.408816958495084
4. CB --> 360/T/CG2 : 4.797433601143536
5. CB --> 415/F/CZ : 4.812685347454735
6. OE2 --> 409/L/CG : 4.82135412998977
7. CB --> 409/L/CD2 : 4.848773467450278
8. CG --> 355/D/OD2 : 4.908903802486195

Loading D359
Showing PH20 359

Surface area: 60.97 (out of 160), relative area:
0.38 (mean 0.388)
neigh: 7
1: 69 D
2: 355 D
3: 356 N
4: 358 M
5: 359 D
6: 360 T
7: 409 L
8: 415 F
Computing neighbors
0. OD2 --> 409/L/CD2 : 3.408547994082272
1. OD1 --> 415/F/CE1 : 3.7963055647968225
2. CG --> 409/L/CD2 : 3.8878097503943883
3. OD1 --> 409/L/CD2 : 4.126379866044171
4. OD1 --> 415/F/CZ : 4.266899435092037
5. CG --> 415/F/CE1 : 4.470863286668419
6. OD1 --> 69/D/CB : 4.572196128552577
7. CG --> 415/F/CZ : 4.593386909208035
8. OD1 --> 358/M/CG : 4.611814184030998
9. OD1 --> 358/M/CB : 4.613461018139933
10. CB --> 415/F/CZ : 4.677500443681096
11. CB --> 409/L/CD2 : 4.779204879832188
12. CB --> 360/T/CG2 : 4.8687616198429575
13. OD2 --> 409/L/CG : 4.89050080064918
14. CB --> 415/F/CE1 : 4.910648655327244
15. OD1 --> 415/F/CD1 : 4.94547223406323

Loading N359
Showing PH20 359

Surface area: 51.32 (out of 166), relative area:
0.31 (mean 0.355)
neigh: 7
1: 69 D
2: 355 D
3: 356 N
4: 358 M
5: 359 N
6: 360 T
7: 409 L
8: 415 F
Computing neighbors
0. ND2 --> 355/D/OD1 : 3.6424538833212696
1. ND2 --> 355/D/OD2 : 3.9347358217608677
2. ND2 --> 355/D/CG : 3.9868658910147405
3. OD1 --> 409/L/CD2 : 4.099298518525556
4. OD1 --> 415/F/CE1 : 4.126581268435471
5. CG --> 409/L/CD2 : 4.4701175327793
6. OD1 --> 69/D/CB : 4.487811260777864
7. OD1 --> 415/F/CZ : 4.554434631603517
8. CG --> 355/D/OD2 : 4.688946077067471
9. OD1 --> 358/M/CB : 4.807627735139873
10. CG --> 355/D/OD1 : 4.831846752916848
11. OD1 --> 358/M/CG : 4.8343202358421555
12. OD1 --> 355/D/OD2 : 4.860220552746158
13. CB --> 360/T/CG2 : 4.8701661552686595
14. CB --> 409/L/CD2 : 4.908677272477356
15. CB --> 415/F/CZ : 4.913758043576114
16. CG --> 355/D/CG : 4.9484620070518055
17. CG --> 358/M/CB : 4.982092746456763

Loading R359
Showing PH20 359

Surface area: 103.72 (out of 250), relative area:
0.41 (mean 0.348)
neigh: 7
1: 67 P
2: 355 D
3: 356 N
4: 358 M
5: 359 R
6: 360 T
7: 409 L
8: 415 F
Computing neighbors
0. CZ --> 355/D/OD1 : 2.829747181920966
1. NH2 --> 355/D/OD1 : 2.8571173486886723
2. NE --> 355/D/OD2 : 2.900127483211178
3. NH2 --> 355/D/OD2 : 3.109750663339618
4. NE --> 355/D/OD1 : 3.1142520111710685
5. NH1 --> 355/D/OD1 : 3.246837988295189
6. CZ --> 355/D/OD2 : 3.2907744117494593
7. NE --> 355/D/CG : 3.3213467717040146
8. NH2 --> 355/D/CG : 3.3585856434629475
9. CZ --> 355/D/CG : 3.4413857210489662
10. CD --> 355/D/OD2 : 3.8310488347045712
11. CD --> 355/D/OD1 : 3.8415588690422697
12. CD --> 355/D/CG : 4.060910078443919
13. CG --> 355/D/OD2 : 4.069022267658191
14. NH1 --> 355/D/CG : 4.177220633885001
15. NH1 --> 355/D/OD2 : 4.345261995739807
16. CG --> 409/L/CD2 : 4.546282484915213
17. CG --> 355/D/CG : 4.575240590434766
18. NH2 --> 67/P/CG : 4.69824680366515
19. NE --> 355/D/CB : 4.719231034507272
20. CD --> 409/L/CD2 : 4.735440772613497
21. CG --> 355/D/OD1 : 4.748301477707716
22. NH2 --> 355/D/CB : 4.849932246360019
23. CB --> 360/T/CG2 : 4.853961476299509
24. CB --> 415/F/CZ : 4.859386346931186
25. CZ --> 355/D/CB : 4.913214390792612
26. CB --> 409/L/CD2 : 4.917073094201744
27. CG --> 358/M/CB : 4.98199577466375

Loading A359
Showing PH20 359

Surface area: 46.38 (out of 111), relative area:
0.42 (mean 0.126)
neigh: 6
1: 355 D
2: 356 N
3: 358 M
4: 359 A
5: 360 T
6: 409 L
7: 415 F
Computing neighbors
0. CB --> 360/T/CG2 : 4.836946510585213
1. CB --> 415/F/CZ : 4.962347806333552
2. CB --> 409/L/CD2 : 4.972731569505419

Loading H359
Showing PH20 359

Surface area: 67.44 (out of 191), relative area: 0.35 (mean 0.241)
neigh: 7
1: 69 D
2: 355 D
3: 356 N
4: 358 M
5: 359 H
6: 360 T
7: 409 L
8: 415 F
Computing neighbors
0. NE2 --> 355/D/OD2 : 2.964179106079704
1. CE1 --> 355/D/OD1 : 3.2208203683656
2. NE2 --> 355/D/CG : 3.5194931627997
3. NE2 --> 355/D/OD1 : 3.536366929968038
4. CE1 --> 355/D/OD2 : 3.5373360746830014
5. CE1 --> 355/D/CG : 3.628781706384634
6. CD2 --> 355/D/OD2 : 3.7685161959103834
7. NE2 --> 69/D/CB : 3.968033118859661
8. ND1 --> 355/D/OD1 : 4.124728944858461
9. CD2 --> 69/D/CB : 4.246297308414403
10. CD2 --> 358/M/CB : 4.323931034239226
11. CD2 --> 355/D/CG : 4.371268198668278
12. CG --> 409/L/CD2 : 4.401493369002809
13. CD2 --> 358/M/CG : 4.443748661833109
14. ND1 --> 355/D/CG : 4.480615120339378
15. ND1 --> 355/D/OD2 : 4.481460093401531
16. ND1 --> 409/L/CD2 : 4.494693753804128
17. CG --> 355/D/OD2 : 4.568808601357066
18. CD2 --> 355/D/OD1 : 4.576086481868833
19. CB --> 415/F/CZ : 4.684482848016547
20. CD2 --> 415/F/CE1 : 4.743984746674277
21. CB --> 409/L/CD2 : 4.749729192790314
22. CD2 --> 409/L/CD2 : 4.784866231501209
23. NE2 --> 355/D/CB : 4.8093862645449565
24. CG --> 355/D/OD1 : 4.834246784991305
25. CG --> 355/D/CG : 4.848855495374266
26. CE1 --> 355/D/CB : 4.887693851217632
27. CB --> 415/F/CE1 : 4.897027705694935
28. CE1 --> 409/L/CD2 : 4.915692330822778
29. CB --> 360/T/CG2 : 4.926906279801616
30. CG --> 415/F/CE1 : 4.968468232755743

Loading S359
Showing PH20 359

Surface area: 50.91 (out of 125), relative area:
0.41 (mean 0.28)
neigh: 6
1: 355 D
2: 356 N
3: 358 M
4: 359 S
5: 360 T
6: 409 L
7: 415 F
Computing neighbors
0. OG --> 360/T/CG2 : 4.003260070922278
1. OG --> 356/N/OD1 : 4.513075798349715
2. OG --> 360/T/CB : 4.7411695205532185
3. OG --> 356/N/CB : 4.79180074052218
4. CB --> 409/L/CD2 : 4.825442295163705
5. CB --> 415/F/CZ : 4.882357475188843
6. CB --> 360/T/CG2 : 4.920937924506039

APPENDIX F

Part 1: Counting Script

```
import sys
def n_pick_m(n,m):
    p = 1
    try:
        if n>=m:
            for i,x in enumerate(range(n,n-m,-1)):
                p *= x
                p /= i+1
            return int(p)
        else:
            return 1
    except:
        return 1

n1 = int(sys.argv[1]) # total number of residues
n2 = int(sys.argv[2]) # number of required mutations, fixed positions
n3 = int(sys.argv[3]) # number of choices for required substitions
n4 = int(sys.argv[4]) # total number of mutations
n5 = int(sys.argv[5]) # number of choices for optional mutations

# n1 = 447
# n2 = 1
# n3 = 4
# n4 = 21
# n5 = 19

# n1 = 11
# n2 = 1
# n3 = 4
# n4 = 3
# n5 = 2

tot = 0

print(n1,n2,n3,n4,n5)
for i in range(n4):
    ni = n3**n2*n_pick_m(n1-n2,i)*n5**i
    tot += ni
#    print(f'# additional substitutions: {i}')
#    print(f'# ways of choosing sub positions: {n_pick_m(n1-n2,i):,}')
#    print(f'subtotal: {ni:,}')
#    print(f'running total: {tot:,}')
#    print(f'running total: {tot:.2e}')
#    print()

print(f'total: {tot:,}')
print(f'or: {tot:.2e}')
print()
```

APPENDIX F

Part 2: Counting Results

```
430 1 7 38 19
total:
535,948,817,140,501,380,729,667,810,825,613,171,261,187,686,897,150,562,833,753,132
,671,165,933,153,326,862,429,044,235,266,950,174,372
or: 5.36e+101

447 1 7 40 19
total:
97,699,784,943,563,842,059,154,468,679,821,249,783,874,062,660,841,100,304,251,309,
300,051,597,995,991,110,167,037,140,174,632,306,120,660,860
or: 9.77e+106

474 1 7 42 19
total:
44,226,731,338,441,319,922,842,951,626,004,477,544,622,455,863,434,756,420,643,749,
554,106,724,187,886,054,476,395,520,940,765,116,828,256,835,627,884
or: 4.42e+112

430 1 7 21 19
total:
3,083,048,321,379,051,586,868,455,325,947,044,889,122,994,035,255,818,750,793,380
or: 3.08e+60

447 1 7 22 19
total:
2,630,536,901,739,384,000,911,309,134,697,651,071,574,748,933,339,737,302,399,784,8
28
or: 2.63e+63

474 1 7 23 19
total:
3,627,301,414,276,170,284,775,262,023,044,500,321,441,769,419,340,926,622,867,532,9
66,764
or: 3.63e+66

430 1 1 38 19
total:
76,564,116,734,357,340,104,238,258,689,373,310,180,169,669,556,735,794,690,536,161,
810,166,561,879,046,694,632,720,605,038,135,739,196
or: 7.66e+100

447 1 1 40 19
total:
13,957,112,134,794,834,579,879,209,811,403,035,683,410,580,380,120,157,186,321,615,
614,293,085,427,998,730,023,862,448,596,376,043,731,522,980
or: 1.40e+106

474 1 1 42 19
total:
6,318,104,476,920,188,560,406,135,946,572,068,220,660,350,837,633,536,631,520,535,6
```

50,586,674,883,983,722,068,056,502,991,537,873,832,608,119,375,412
or: 6.32e+111

430 1 2 38 19
total:
153,128,233,468,714,680,208,476,517,378,746,620,360,339,339,113,471,589,381,072,323
,620,333,123,758,093,389,265,441,210,076,271,478,392
or: 1.53e+101

447 1 2 40 19
total:
27,914,224,269,589,669,159,758,419,622,806,071,366,821,160,760,240,314,372,643,231,
228,586,170,855,997,460,047,724,897,192,752,087,463,045,960
or: 2.79e+106

474 1 2 42 19
total:
12,636,208,953,840,377,120,812,271,893,144,136,441,320,701,675,267,073,263,041,071,
301,173,349,767,967,444,136,113,005,983,075,747,665,216,238,750,824
or: 1.26e+112

465 1 7 41 19
total:
98,801,281,111,853,038,170,032,247,261,011,559,411,452,759,177,641,859,464,242,597,
193,855,754,754,031,857,735,967,529,891,482,020,811,355,668,331
or: 9.88e+109

465 1 2 41 19
total:
28,228,937,460,529,439,477,152,070,646,003,302,688,986,502,622,183,388,418,355,027,
769,673,072,786,866,245,067,419,294,254,709,148,803,244,476,666
or: 2.82e+109

465 1 1 41 19
total:
14,114,468,730,264,719,738,576,035,323,001,651,344,493,251,311,091,694,209,177,513,
884,836,536,393,433,122,533,709,647,127,354,574,401,622,238,333
or: 1.41e+109

433 1 7 38 19
total:
701,536,937,709,980,502,014,969,318,455,861,764,372,583,984,548,006,349,395,726,965
,902,393,519,340,117,383,612,686,253,279,782,224,611
or: 7.02e+101

433 1 1 38 19
total:
100,219,562,529,997,214,573,567,045,493,694,537,767,511,997,792,572,335,627,960,995
,128,913,359,905,731,054,801,812,321,897,111,746,373
or: 1.00e+101

```
430 1 1 21 19
total:
440,435,474,482,721,655,266,922,189,421,006,412,731,856,290,750,831,250,113,340
or: 4.40e+59

433 1 1 21 19
total:
507,902,969,111,610,088,221,427,205,232,382,081,206,240,353,887,310,581,987,141
or: 5.08e+59
```