Amino Acid Basics and H++ analysis

This work will be done during Friday's class time in your assigned group. Meet anywhere – there is no formal class session for this assignment. If you have questions, text or call Dr. Provost at 701-306-1586.

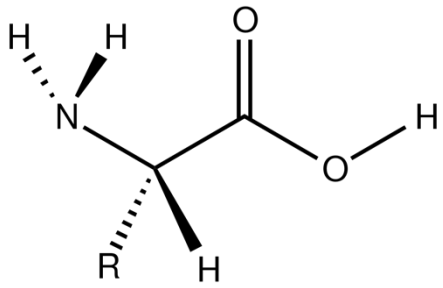
The worksheet will be done as a group on Friday then the group will submit the finished worksheet **IN CLASS TO YOUR INSTRUCTOR SEPT 21** as a group (one person submits for all members). ALL material in the worksheet may show up on the test so it is important that all members are engaged. IF one of your group members does not attend, please let me know.

Section 1: Group 1 (Beckett, Hildreth, Cordova) Group 2 (Dzotsenidze, Rich, Mouawad)
Group 3 (Leveroni, Nash, Choudhary), Group 4 (Spillman, Kobzeff, Riley),
Group 5 (Walsh, Farhoud, Goodoy, Petty), Group 6 (Johnson, Zhou, Robertson)
Group 7 (Imholte, Tancreto, Mauhay)

Section 3: Group 1 (Anawalt, Piracha, Rodriguez-Agiss) Group 2 (Eggemeyer, Holt, Burch)
Group 3 (N'Senga, Duhaney, Heras), Group 4 (Terry, Grey, Tikekar, Fagen),
Group 5 (Reimer, Marfai, Howard), Group 6 (Giles, Cowles, Pakhlevanyan, Pulido)
Group 7 (Sevilla, Bohanon, Warder, Bidwell-Astaburuaga,), Group 8 (Tarbox, Karminsky, Miller)

Part 1. Amino Acid Basics.

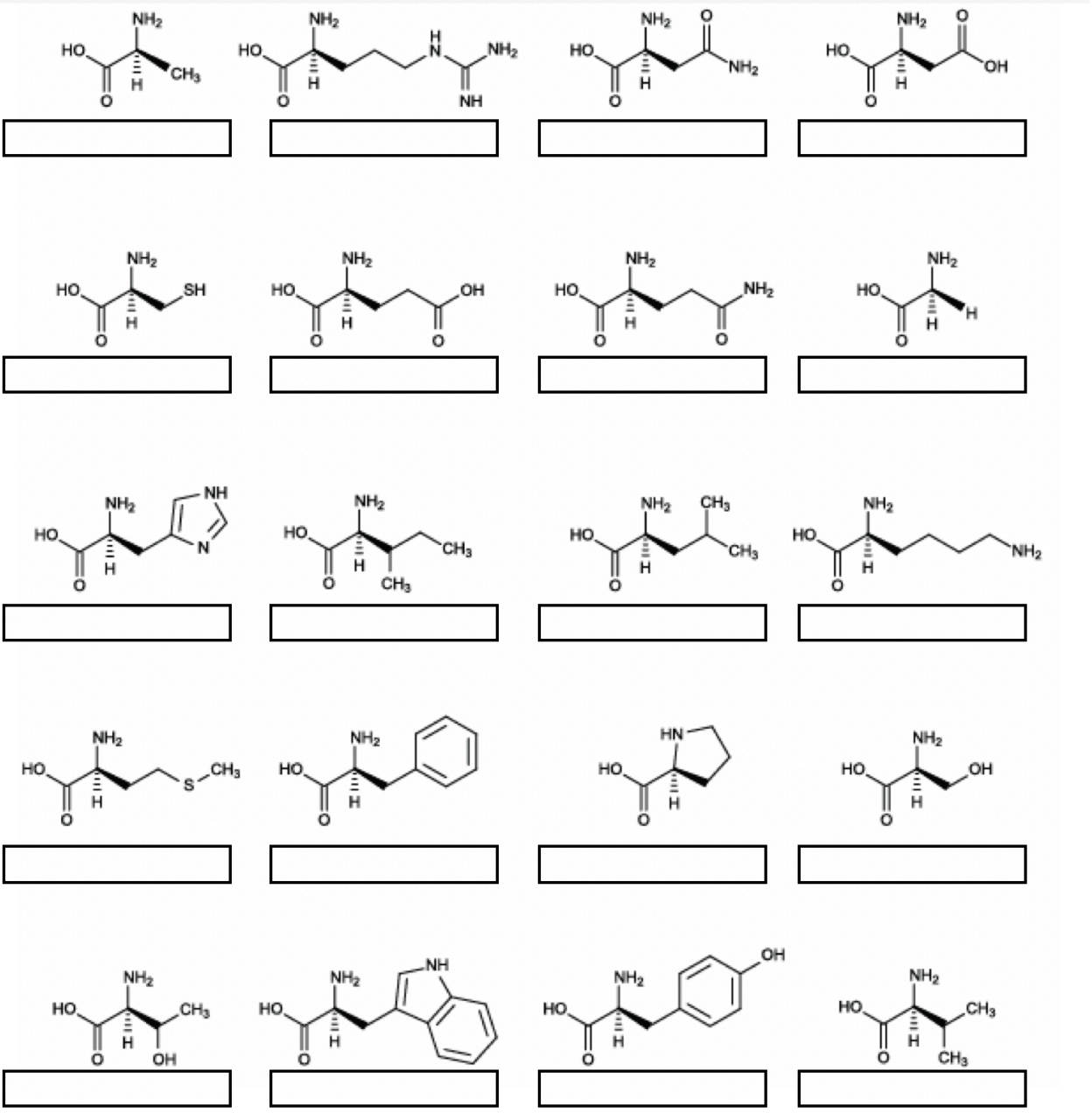
a. Inspect the structure of an generic amino acid in the figure to the right. There is one basic error in the structure as shown – discuss and report what your group finds.



b. As a group draw a tri-peptide. Pick three amino acids, choose the order and draw the structure using the line formula. There is no right or wrong choice of amino acids, just pick three. Consult your textbook on the orientation of the R groups before you draw your tri-peptide.
   - Label the following items on your line drawing: amino terminal, carboxy terminal, peptide bond, alpha carbon and R groups.

c. As a group, select one of the amino acids with an ionizable side chain (R group). Draw the structure of the amino acid in its most acidic form (every group protonated). Using your book, circle each ionizable functional group and indicate the pKa for each group. Then show the structure of the amino acid as the pH rises. Recall your understanding of how pKa values relate to the ionization of a weak acid or base to titrate the appropriate ionizable group. Calculate and record the overall charge of the amino acid in each state.

d. Below is a drawing of all 20 common amino acids.
   - Record the full name, three and single letter abbreviation.
   - Circle the side chain of each amino acid and write the name of the functional group. i.e. guanidino, imidazole, phenyl, ect…



e. In the figure above of the 20 common amino acids, identify which amino acids are post translationally modified by the following: phosphorylation, acetylation, methylation, carboxylation. Indicate below how the modification changes the chemistry of the side chain.

f. In the space below, group the sidechains (using three letter abbreviations) into as many groups as you like. Explain the basis of each grouping. Some amino acids may be in multiple groups.

g. Create a list of all amino acids (using the three letter abbreviation) that contain titratable side groups. Indicate the approximate pKa values of the side chains of the free amino acids (free means in solution not in a peptide bond/protein). In words, describe what would be the structure of these sidechains at pH 2, pH 5, pH 8 and pH 10. Only worry about the sidechains: ignore the amino and carboxyl groups of the alpha carbon.

h. The pKa of the sidechain of glutamate, tyrosine and serine are very different. Why might that be? Consider the rest of the molecule as electrons are pulled or donated from the rest of the R group.

Part 2. Amino Acid in Protein. H++ and local environment effect on pKa.

In the first part of this assignment, you characterized the side chains of free amino acids and examined the chemistry and pKas of the side chains of the 20 common amino acids. However, when in a protein, the local environment of an amino acid can shift the equilibria of ionizable side chains. Carefully read chapter 4 part D "the pK values of ionizable groups depend on nearby groups" and apply this information to interpret and understand the work in this section.

In this part, you will analyze the pKa of protonatable side groups in the local environment of a folded protein. Click on the link (https://www.rcsb.org) of the protein database and ender in the ID number: 7QID. Poke around and see the kind of information stored on this page. No need to record anything.

Now click on the website (http://biophysics.cs.vt.edu/). This links to the program H++.

H++ is an automated system that computes pK values of ionizable groups in macromolecules in a folded protein. Why H++ ? Structure and function of macromolecules depend critically on the ionization (protonation) states (pK) of their acidic and basic groups. For example, affinities of proteins for ligands depend on the pKs of the groups in or near the binding sites; groups with unusual pKs are often found in active sites of enzymes. Atomistic simulations of macromolecules require specification of the protonation state of the titratable groups. Whether or not a given group is protonated depends, in a non-trivial way, on the position of the group within the molecule as well on the characteristics of the surrounding solvent such as its pH and ionic strength. Experimental determination (usually by NMR) of protonation equilibria is expensive and often cannot be performed for every group of interest. The most common source of high resolution structures -- X-ray crystallography -- normally does not provide positions of the hydrogen atoms. The gap is bridged by theoretical methods that predict protonation states (pK) of ionizable groups within the macromolecule based on its atomic resolution structure. This web site provides access to a set of tools that automate this process. The H++ server allows both experts and novices to quickly obtain estimates of pKs as well as other related characteristics of bio-molecules such as isoelectric points, titration curves, and energies of protonation microstates. It also automates the tedious process of preparing the input files for typical molecular dynamics simulations.

Simply put – the program estimates the pKa of titratable sidechains based on the local environment of a folded protein.

Table I shows the results of entering 7QID into H++. Use this information to answer the rest of your questions…

a. Review the resultant data of the computed pKa values of a select set of pKa shown in table I.
   - Record the highest and lowest value you obtain for each shown amino acids shown in table I.
   - Are these values different from the pKa of a free amino acid?
   - Discuss why you think the results are the way they are. Refer to the section of your textbook for insight.

Table I H++ results of 7QID analysis of select amino acids

Amino Acid	pKa
ARG-114	11.299
ARG-118	8.624
ARG-135	11.461
ARG-14	11.848
ARG-181	9.907
ARG-19	11.704
ARG-229	12.325
ARG-236	11.741
ARG-252	11.151
ARG-270	11.275
ARG-271	12.271
ARG-331	12.231
ARG-345	11.010
ARG-371	11.544
ARG-372	11.413
ARG-383	11.633
ARG-386	12.112
ARG-389	10.856
ARG-409	11.430
ARG-42	11.686
ARG-451	12.306
ARG-454	10.916
ARG-47	10.631
ARG-479	12.926
ARG-488	12.145
ARG-498	11.605
ARG-539	10.782
ARG-554	11.302
ARG-576	11.550
ARG-577	11.072
ARG-634	10.965
ARG-65	10.066
ARG-656	10.355
ARG-702	11.794
ARG-717	12.079

Amino Acid	pKa
ASP-12	4.710
ASP-132	5.273
ASP-138	6.071
ASP-142	4.966
ASP-150	3.653
ASP-151	3.894
ASP-157	4.230
ASP-220	4.979
ASP-221	4.942
ASP-234	4.071
ASP-250	4.410
ASP-261	4.999
ASP-322	5.863
ASP-404	4.469
ASP-413	3.541
ASP-45	3.750
ASP-456	4.608
ASP-464	6.014
ASP-48	4.602
ASP-483	4.142
ASP-496	3.895
ASP-499	3.917
ASP-519	5.844
ASP-522	4.468
ASP-533	4.740
ASP-535	5.282
ASP-542	4.243
ASP-574	3.953
ASP-584	4.275
ASP-59	4.737
ASP-591	4.534
ASP-600	4.432
ASP-620	4.143
ASP-638	3.624

Amino Acid	pKa
HIS-185	5.579
HIS-199	7.269
HIS-209	5.590
HIS-21	6.054
HIS-264	6.567
HIS-280	5.192
HIS-313	5.742
HIS-32	4.717
HIS-440	4.117
HIS-548	5.314
HIS-627	6.396
HIS-710	6.375

Amino Acid	pKa
TYR-107	10.774
TYR-127	10.415
TYR-144	9.826
TYR-232	10.869
TYR-245	9.803
TYR-246	10.747
TYR-277	10.449
TYR-292	9.928
TYR-3	13.241
TYR-367	9.829
TYR-374	10.238
TYR-398	12.311
TYR-401	10.718
TYR-430	9.641
TYR-477	10.393
TYR-492	10.172
TYR-507	10.231
TYR-512	9.559
TYR-562	10.982
TYR-579	10.698
TYR-587	13.272
TYR-60	11.160
TYR-628	12.814
TYR-646	10.482
TYR-67	10.380
TYR-675	12.043
TYR-708	9.812
TYR-91	13.129

Amino Acid	pKa
LYS-102	9.576
LYS-121	9.124
LYS-149	9.876
LYS-164	10.038
LYS-166	10.322
LYS-190	9.726
LYS-197	9.482
LYS-224	10.007
LYS-265	9.870
LYS-267	9.289
LYS-283	9.877
LYS-310	11.065
LYS-319	10.142
LYS-369	9.434
LYS-384	10.126
LYS-40	10.061
LYS-416	10.299
LYS-425	9.567
LYS-433	8.096
LYS-441	8.865
LYS-449	9.339
LYS-460	10.138
LYS-474	9.061
LYS-484	8.967
LYS-508	5.588
LYS-53	9.963
LYS-544	10.125
LYS-557	10.181
LYS-567	8.033
LYS-582	8.998
LYS-614	10.124
LYS-616	10.360
LYS-649	9.796
LYS-652	10.541

b. The insulin receptor is expressed on the surface of cells where the pH of the extracellular matrix ranges from 7.0 - 6.7 in metabolically active cells. With this information and the information found in Table 1, predict if His 548 will be charged or uncharged and protonated or unprotonated. What about His 199? What would the charge state be for the same two His residues when cells are "resting" and the pH ranges from 7.2 – 7.4?

c. If insulin were to bind ionically to either His 199, would the pH impact how the receptor might bind and respond to insulin?

Next let's look at the structure of one of the titratable sidechains in the protein. Click back to the PDB database (https://www.rcsb.org) and re-enter 7QID. On the PDB page, click on the link circled in the figure to open up the primary sequence and the three dimensional rendering of the protein. That will open up to a view as shown in the figure to the right. Click on an amino acid from the sequence, the image will zoom in and show some of the bonds between that amino acid and other side chains in the folded protein. Take a few min to become acquainted with some of the visualization options and controls.

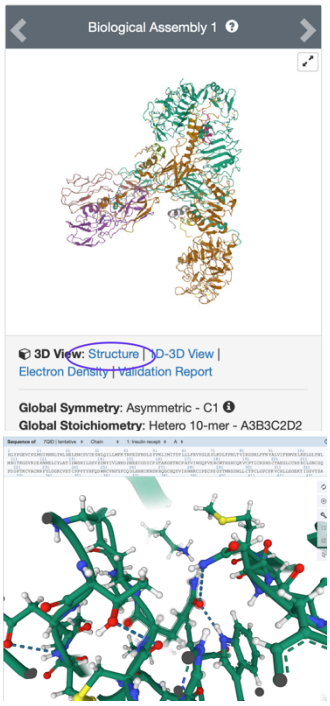


Table II shows a set of amino acids (bolded) whose pKa has shifted as determined by H++. The other amino acids in each box is a list of nearby amino acids that have a strong contribution to the pKa of the "bolded" amino acid. The first number after these amino acid sis the background contribution from non-titratable groups; The second number is the interaction with other titratable groups.

d. Pick one of the amino acids in a box from table II and record the pKa from table I and the pKa for the free amino acid found in your book. Predict the impact of 3-4 of the residues that influence the pKa of the bolded amino acid.

Table II Selected amino acids and the surrounding amino acids that influence the pKa of the "site" residue.

HIS- 21	SITE: **ARG- 135**	**SITE: ASP- 74**
18 THR 0.157794 0	132 ASP -0.0520526 0.545004	43 PRO 0.115013 0
19 ARG -0.080854 -0.308314	134 SER -0.533691 0	44 GLU 0.0062117 0.175896
22 GLU -0.562851 0.420378	136 ILE -0.27413 0	45 ASP -0.0784181 0.124871
24 GLU -0.0148476 0.422492	138 ASP 0.0733634 0.294151	47 ARG -0.233836 -0.513633
45 ASP 0.0507442 0.103738	141 GLU 0.107448 0.231986	48 ASP -0.0492348 0.159272
48 ASP -0.155628 0.289686	173 VAL 0.21646 0	49 LEU 0.196024 0
22 ARG 0.0341623 -0.142832	174 ILE -0.214934 0	50 SER -0.3212 0
26 TYR 0.0136532 0.105182	181 ARG 0.211053 -1.31017	53 LYS -0.0210506 -0.108624
	197 LYS -0.00495648 -0.359997	69 LEU -0.196035 0
	229 ARG 0.0284102 -0.10202	70 GLU -0.227632 0.416457
		71 SER 0.384642 0
		73 LYS 0.0162653 -1.6708
		75 LEU -0.256956 0
		77 PRO -0.0420102 0
		LYS -0.0811147 -0.288629

e. (BONUS ROUND – Extra Credit): Using the structure from RCSB, zoom in and see where a few of these contributing amino acid side chains are in proximity to the bolded site amino acid you've selected. Describe the local environment around each of the illustrate residues. How do you think that local environment impacts the pKa of the labeled sidechains.