**Nanobody Engineering – Based on Framework of the anti-RR6 Nanobody**

Laurence Chen, Dr. Xiaohua Huang

Genomics and Systems Biotechnology Laboratory

Created: July 8, 2015

**Nanobody**: anti-RR6

**PDB ID:** 1QD0

http://www.rcsb.org/pdb/explore/explore.do?structureId=1QD0

http://www.ncbi.nlm.nih.gov/pubmed/10684599?dopt=Abstract

1. **BACKGROUND**

Nanobodies, also known as camelid antibodies, are distinct in that nanobodies do not have light chains as well as CH1 domains, and their heavy-chain variable domains (VHH) can be harnessed for a wide variety of protein engineering purposes. The physical size of a nanobody VHH fragment is, compared to a conventional antibody, relatively small since it is only comprised of one immunoglobulin (IgG) domain. It has also been reported, “[nanobody] levels of expression and solubility are significantly higher than those of classical Fab or Fvs” (Spinelli, 2000).

Under Dr. Xiaohua Huang’s guidance, a senior design team at UC San Diego developed a system to produce and select for a library of nanobodies binding to specific antigens. The senior design team based their framework off of NbSyn2, PDB ID 2X6M, a nanobody in which a binding pocket, using two of its CDRs, is coupled with its target antigen.

1. **RATIONALE**

The success of the senior design team and Dr. Huang’s nanobody production system can be utilized as a platform to produce different nanobodies with distinct desired characteristics. Through this, nanobody libraries that have the ability to bind to distinct groups of antigens can be created using protein engineering technologies.

Using the senior design team’s project as a model, a new nanobody with similar characteristics was discovered. The anti-RR6 Nanobody, PDB ID 1QD0, was selected after looking at the structures of numerous nanobodies available on the protein data bank (PDB). The criteria in combing through literature was to find a nanobody with a concave binding site on it, in which protruding antigens can insert itself and bind to. I will define this class of nanobodies as *pocket nanobodies*. The anti-RR6 was selected because it had a deep binding pocket, as well as a secondary binding pocket that can be useful if mutated well.

1. **AIMS**

In light of the goal to produce a pocket nanobody that binds to any amino acid group, a series of design steps will be explored and outlined as follows:

1. Residues to mutate
2. Mutations to be introduced & optimization
3. Gene construction
4. **DESIGN**

**IV.A. Residues to mutate**

The amino acid residues on the anti-RR6 nanobody were selected after analysis using both protein alignments with other nanobodies as well as 3D structures using ChemBio3D Ultra 2014.

Protein alignments were carried out using COBALT from the NCBI. A broad alignment scope comparing 5 different nanobodies was input into COBALT to look for conserved regions as well as hyper variable regions. Alignment results can be viewed under the “Cobalt” tab of [PocketNanobody\_ProteinAlignment.xlsx] or recalled on COBALT by inputting its alignment ID PA0ZHE9821S.

A narrow alignment scope was also applied to compare the senior design team’s NbSyn2 (2X6M) nanobody and the anti-RR6 nanobody (1QD0). The alignment results were used to examine differences between these pocket nanobody candidates. Again, alignment results can be viewed in the attached excel document or by recalling its alignment ID PD7HUH6B21S.

The following are several iterations and thought process in determining which amino acids were ideal candidates for mutation:

**First iteration: Residues on Nanobodies to Mutate**

(H) His32(addition)

(G) Gly33(addition)

(H) His34 (addition)

(Y) Tyr35

(G) Gly36 (in combination with glycine, sulfate may be stabilized by Tyr35)

(R) Arg55 **corresponds to 51N mutation of SDT**

(W) Trp56

(S) Ser57

(K) Lys59

(R) Arg102

(P) Pro103 **corresponds to F100 mutation of SDT**

(V) Val104 **corresponds to S101 mutation of SDT**

(R) Arg105

(V) Val106

(A) Ala107 **corresponds to Y103 mutation of SDT**

(D) Asp108 **corresponds to C104 mutation of SDT**

(I) Ile109

SWSN (108-111) corresponds to LPVG (111-114)

All the way to Ser110 (insertion of the 3rd loop)

**Remarks:** The residues I chose to mutate were solely based on modifying the binding pocket using Chem3D, residues that appeared to be in contact with the dye peptide of 1QD0. What was failed to be accounted for was that the goal was to create a nanobody platform in which there is a lot of diversity. Only mutating residues at the pocket is not sufficient for diversity, so amino acid regions further away from the pocket was analyzed.

**Additional steps:** Take a look at the senior design team’s residue mutations to see if a rationale can be understood.

**Second iteration: Residues on Nanobodies to Mutate**

**Remarks:** Compared the residues the senior design team mutated and aligned it with 1QD0. The mutation residues were not all clustered around the binding pocket, which would contribute to greater diversity than the first iteration. However, what seemed to be missing between the first and this iteration was establishing a framework – a framework that would hold the core structure for even more nanobodies to be produced.

**Additional steps:** Align nanobodies that are more dissimilar to see which regions are highly conserved, and which contribute to diversity. Also make note that when making a turn, there needs to be a proline, glycine, or serine. Establish a rationale behind each mutation

**Third iteration: Residues on Nanobodies to Mutate**

(A) Alanine – 29

(H) Histidine – 32 (added loop)

(G) Glycine – 33 (added loop)

(H) Histidine – 34 (added loop)

(G) Glycine – 36 (does not contribute to turn)

~~(V) Valine – 43~~

(R) Arginine – 55

(W) Tryptophan – 56

~~(K) Lysine – 68~~

(P) Proline – 103

(V) Valine – 104

(V) Valine – 106

(A) Alanine – 107

**Remarks**: This iteration takes into consideration the conserved regions and the hyper variable regions as aligned from 5 different nanobodies that function as pockets. The main characteristic difference in 1QD0 that attributes to its deeper pocket is its 3 additional amino acids on its CDR1 region, creating an extended loop structure. In addition, hyper variable amino acids located relatively further from the binding pocket were also taken into account.

**Fourth iteration: Residues on Nanobodies to Mutate (total of 18)**

(R) Arginine – 27

(A) Alanine – 28

(A) Alanine – 29

(H) Histidine – 32 (added loop)

(G) Glycine – 33 (added loop)

(H) Histidine – 34 (added loop)

(G) Glycine – 36 (does not contribute to turn)

(R) Arginine – 55

(W) Tryptophan – 56

(S) Serine – 57

(A) Alanine – 78

(K) Lysine – 79

(R) Arginine – 102

(P) Proline – 103

(V) Valine – 104

(R) Arginine – 105

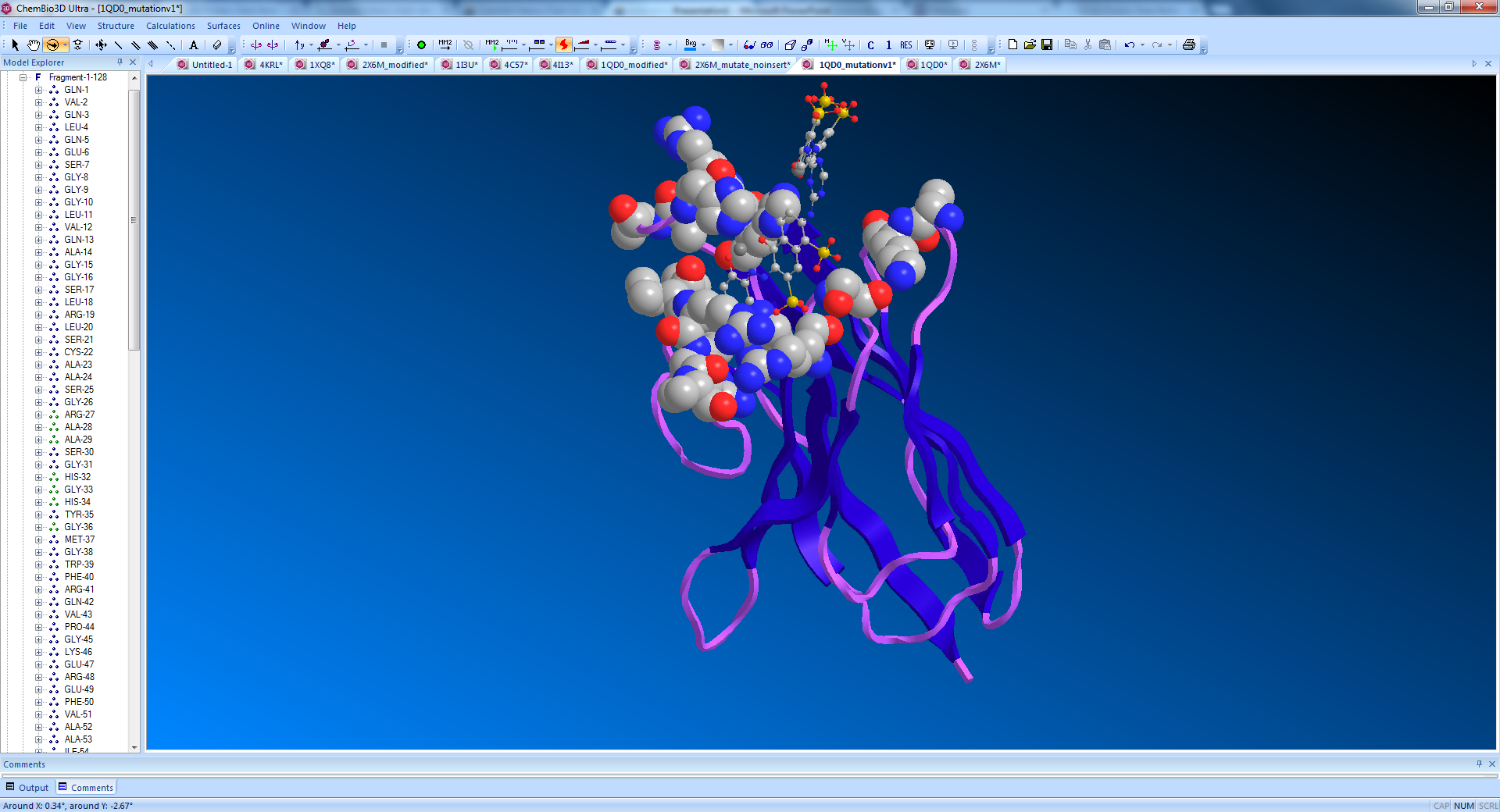
(V) Valine – 106

(A) Alanine – 107

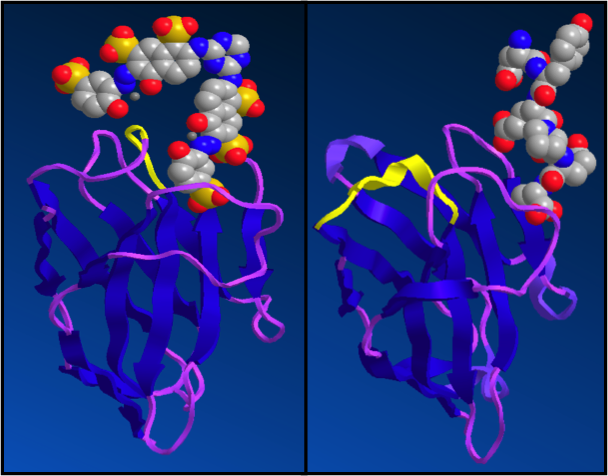
**Remarks**: Arginine 102 and 105 were added because it was noted that the 3 Arginines (as well as 102) contribute to stabilization of the sulphate in the deep pocket. However, if the arginines are not mutated, then it would not support non-negatively charged groups at the end and would not work well. As a result, the mutations further away from the framework were trimmed off.

Also, there is potential for 1QD0 to have a second pocket. The antigen in mind will have several benzene rings and 2 main side chains. 1QD0 has the potential to recognize both, thus creating even more diversity and allows for better amino acid discrimination (20\*20). Thus, Ser57, Arg28, Ala29, Ala 78, and Lys 79 were chosen.

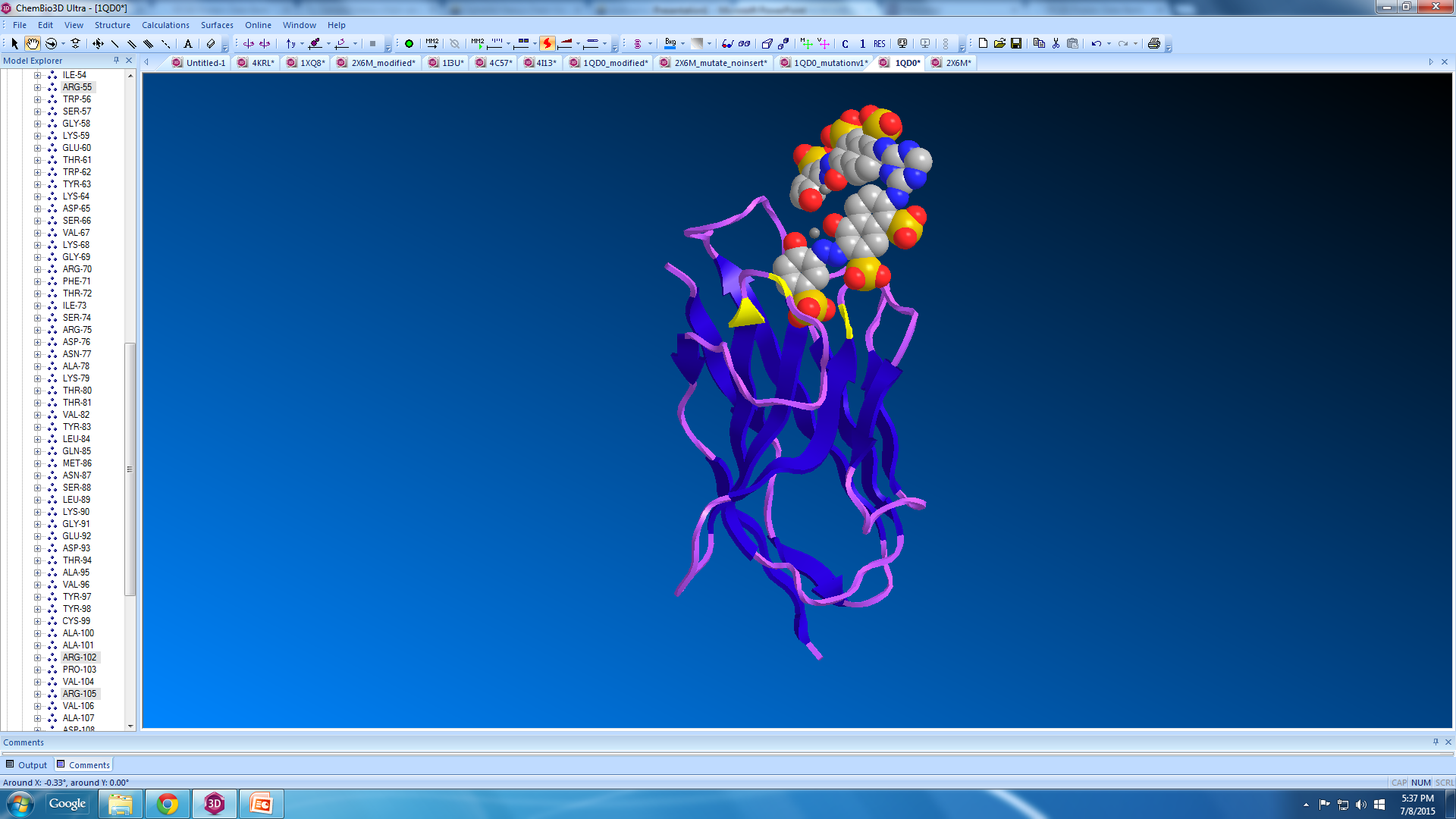
**We can visually summarize the mutations into categories:**

****

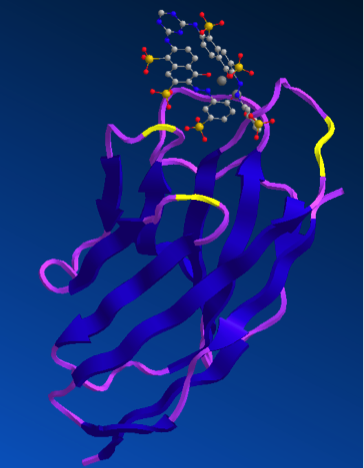
**Figure 1:** Amino acid residues to be mutated. Amino acid residues that to be mutated are visualized in space filling mode. The peptide antigen is visualized in ball and stick. The blue arrow points towards the deepest area of the binding pocket, and the red arrow points towards where the desired variable R group will be located on the designed antigen.



**Figure 2:** Length extension of CDR1. NbSyn2, to the right, interacts with its antigen by only using two of its CDR regions. Anti-RR6 nanobody, to the left, on the other hand interacts with its antigen by forming a deep pocket using all three CDRs. This can be attributed to its extension of its CDR1 as it has an additional 32H, 33G, and 34H compared to NbSyn2. The highlights in yellow indicate where the extension on anti-RR6 occurs versus where it would have been on NbSyn2. Thus mutating the 32-34 extension residues can give rise to diversity in how the pocket is formed.

******

**Figure 3:** Sulfate group stabilizers. The provided antigen contains a sulfate group at the end, burying itself into the pocket of the nanobody. There are three arginine residues, highlighted in yellow, surrounding the deep end of the pocket that stabilizes the negative charge. Thus, it is important to mutate the arginine residues, as not all antigens will be negatively charged.



**Figure 4:** Second binding pocket. There is a pocket region, next to the main antigen-binding site, that has potential to serve as a second binding pocket. By mutating the highlighted residues, diversity in how a secondary binding pocket can be created. Creating a second pocket would enhance the nanobody’s ability to recognize a wide variety of antigens (20 x 20 AAs) as well as enhance its specificity.

**Nanobody Sequence (antiRR6, PDB ID: 1QD0)**

001 010 020

Q V Q L Q E S G G G L V Q A G G S L R L

CAG GTT CAG CTG CAA GAA AGC GGT GGT GGT CTG GTT CAG GCA GGC GGT AGC CTG CGT CTG

021 **027 028 029** 030 **032 033 034** **036** 040

S C A A S G **R** **A** **A** S G **H G H** Y **G** M G W F

AGC TGT GCA AGC AGC GGT CGT GCA GCA TCA GGT CAT GGT CAT TAT GGT ATG GGT TGG TTT

041 050  **055** **056**  060

R Q V P G K E R E F V A A I **R W** S G K E T

CGT CAG GTT CCG GGT AAA GAA CGT GAA TTT GTT GCA GCA ATT CGT TGG AGC GGT AAA GAA

061 070 080

T W Y K D S V K G R F T I S R D N A K T

ACC TGG TAT AAA GAT AGC GTG AAA GGT CGT TTT ACC ATC AGC CGT GAT AAT GCA AAA ACC

081 090 100

T V Y L Q M N S L K G E D T A V Y Y C A

ACC GTT TAC CTG CAG ATG AAT AGT CTG AAA GGT GAA GAT ACG GCA GTG TAT TAT TGT GCA

101 **103 104 106 107** 110 120

A R **P V** R **V A** D ISLP V G F D Y W G Q

GCA CGT CCG GTT CGT GTT GCA GAT ATT AGC CTG CCG GTT GGT TTT GAT TAT TGG GGC CAG

121 128

G T Q V T V S S

GGG ACC CAG GTT ACC GTT AGC AGC

**Codon Table**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **2nd Base** | | | |  | |
| **T** | **C** | **A** | **G** |  | |
| **1st**  **base** | **T** | **Phe** | **Ser** | **Tyr** | **Cys** | **T**  C  **A**  **G** | **3rd**  **base** |
| Phe | Ser | Tyr | Cys |
| Leu | Ser | Stop | Stop |
| **Leu** | **Ser** | **Stop** | **Trp** |
| C | **Leu** | ***Pro*** | ***His*** | **Arg** | **T**  C  **A**  **G** |
| Leu | Pro | His | Arg |
| Leu | Pro | Gln | Arg |
| **Leu** | ***Pro*** | **Gln** | **Arg** |
| **A** | **Ile** | **Thr** | **Asn** | **Arg** | **T**  C  **A**  **G** |
| Ile | Thr | Asn | Ser |
| Ile | Thr | Lys | Arg |
| **Met** | **Thr** | **Lys** | **Arg** |
| **G** | **Val** | **Ala** | **Asp** | **Gly** | **T**  C  **A**  **G** |
| Val | Ala | Asp | Gly |
| Val | Ala | Glu | Gly |
| **Val** | **Ala** | **Glu** | **Gly** |

**A: Small polar; B: Large polar/charged; C: Small nonpolar; D: Large non-polar; E: Aromatic**

**Mixed bases:**

**D: T, A and G**

**K: T and G**

**M: A and C**

**W: A and T**

**S: G and C**

**R: G and A**