Integration Solutions

(1)

a) With no active glycosyltransferase of either type, an ii individual would not be able to add any sugars to the O form of the lipopolysaccharide. Thus, the only lipopolysaccharide present would be the O type.

b) The enzyme encoded by the i allele is essentially inactive. Having one copy of the enzyme (encoded by the I^A or I^B allele) that adds sugars onto the O antigen must be sufficient to produce enough of the A or B antigen on the surface of red cells to provoke an immune response.

c) Here, both enzymes encoded by the I^A and I^B alleles are active, and both reaction 1 and reaction 2 will occur simultaneously. Thus, the cells will be producing both A and B antigen, and the red cells will have both molecules on their surfaces. Consequently, both antigens will be available for immune response.

d)

```
                          .Leu Val Val Thr Pro Trp...
Sequence of I^A allele:   ...C GTG GTG ACC CCT T...

                          .Leu Val Val Pro Leu...
Sequence of i allele:     ...C GTG GTA CCC CTT....
```

e) Note that the reading frame must be the same in both sequences, starting from the left.

```
                                266 267 268
                          ..Tyr Tyr Leu Gly Gly Phe Phe...
Sequence of I^A allele:   ...AC TAC CTG GGG GGG TTC TT...

                                266 267 268
                          ..Tyr Tyr Met Gly Ala Phe Phe...
Sequence of I^B allele:   ...AC TAC ATG GGG GCG TTC TT...
```

f) The difference between the type-A and type-B sugars is mostly size; both groups are capable of forming hydrogen bonds with the enzyme. The type-A sugar is larger than the type-B sugar. Thus, the side chains in the binding site of the type-A enzyme are smaller (Leu and Gly) than the corresponding side chains in the type-B enzyme (Ala and Met). Thus, the larger substrate can only fit into the larger binding site while the smaller can fit into the smaller binding site.

g)

i) This would change one of the amino acids at the binding site from very small (Ala or Gly) to very large (Arg). This would make the binding site too small to accommodate the substrate, resulting in an inactive enzyme.

ii) This would result in a substantially shorter glycosyltransferase enzyme which would likely be missing critical parts and be inactive.

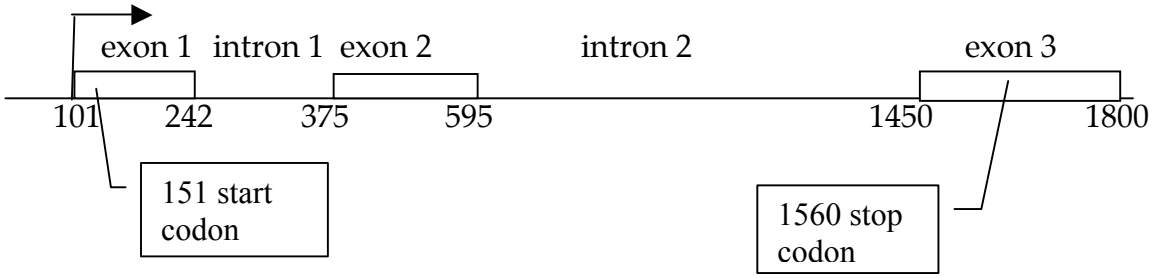
h)

i) The differences in sequence must eliminate the UDP-X substrate specificity of the substrate-binding pocket, so that it now binds either UDP-Gal or UDP-GalNAc.

ii) If you had a patient with type AB who had parents of types AB and O, this would indicate the presence of a *cis*-AB allele. Without the *cis*-AB allele, this cross would be: $I^AI^B \times ii$, which cannot give an AB child. The only explanation (barring parental infidelity) is $I^{cisAB} \times ii$, giving a I^{cisAB} i child with type AB blood.

(2)

a) Your map should look roughly like this:



b)

A1: UGG ⇒ UAG nonsense mutation.
This would result in a protein only ~15 amino acids long. The vast majority of the protein would be missing, so it would be nonfunctional.

A2: AAG ⇒ UAG nonsense mutation.
This would result in a protein only ~16 amino acids long. The vast majority of the protein would be missing, so it would be nonfunctional.

A3: CAG ⇒ UAG nonsense mutation. This would result in a protein only ~35 amino acids long. The vast majority of the protein would be missing, so it would be nonfunctional.

A4: Frameshift. Hint: Use the translation below the DNA instead of the codon table. The protein is now N-Met-Val-His-Leu-Thr-Pro-GLY-Arg-Ser-Leu; the reading frame is shifted to frame b starting with the Arg (the GLY comes right at the deletion). The remaining amino acids (until the early stop at about 205) will all be different than in the normal protein. This protein will therefore be nonfunctional.

A5: Frameshift to frame c. N...-Glu-Glu-Val-Cys-.... (frame c starting from Leu). The amino acids will all be different from those in the normal protein up to a premature termination at 217 or so. This protein will therefore be nonfunctional.

A6: Frameshift to frame c. N...-Glu-Glu-LYS-Val-Cys.... (frame c starting from Leu). The amino acids will all be different than those in the normal protein up to a premature termination at 217 or so. This protein will therefore be nonfunctional.

Each of these mutations will result in a recessive phenotype because in each case the mutant gene does not produce functional protein. One functional copy of the β-globin gene produces enough protein to make healthy red blood cells. Thus, the heterozygote is normal.

c)
B1: GGC (Gly, small and hydrophobic) ⇒ CGC (Arg, big and hydrophilic). The side chain is buried in the hydrophobic core, an unfriendly place for a hydrophilic amino acid. This substitution disrupts the β-globin structure and makes it inactive.

B2: GCC (Ala, small and hydrophobic) ⇒ GAC (Asp, small and hydrophilic). The side chain is buried in the hydrophobic core, an unfriendly place for a hydrophilic amino acid. This substitution disrupts the β-globin structure and makes it inactive.

B3: AAG (Lys, big, hydrophilic, and positively charged) ⇒ GAG (Glu, big, hydrophilic, and negatively charged). If the side chain of this amino acid was involved in an interaction with another, this substitution might be expected to disrupt the function of the protein. Normal function is observed, so the side chain is likely on the surface due to its hydrophilicity, but the charge may not be important.

B4: GUC (Leu, medium-sized and hydrophobic) ⇒ UUC (Phe, big and hydrophobic). The side chain is in the interior of the protein and the larger size likely alters the folding of the protein.

B5: CAU (His, medium and hydrophilic) ⇒ CGU (Arg, big and hydrophilic). The His side chain binds to the heme iron, so size alters the binding.

B6: AAA (Lys, big, hydrophilic, and positively charged) ⇒ GAA (Glu, big, hydrophilic, and negatively charged). The side chain is near COO^- of the heme. Changing the amino acid to a negatively charged one will repel heme. Without heme, the structure is really disrupted and the proteins stick together.

Each of these mutations will result in a dominant phenotype because one mutant copy of the gene produces enough mutant protein to aggregate even in the presence of normal protein.

(3)
a) Four of the DNA codons for arginine (CGA, CGG, CGC, CGT) contain the highly mutable CG sequence. If the CG in the CGT or CGC sequences were mutated to TG as described, they would become TGT or TGC, which code for cysteine. Thus, the first factor would make these mutations likely. Also, changing the highly hydrophilic arginine to the hydrophobic cysteine would likely have a big effect on protein structure. Therefore, individuals with such a mutation would likely show a phenotype.

b) This is similar to (a). The required mutation, CGG to TGG, is likely. Also, changing the highly hydrophilic Arg (hydrophilic) to Trp (hydrophobic) would have a big effect on protein structure.

c) The required mutation, CGT or CGC to CAT or CAC, is very likely, but both Arg and His have hydrophilic and positively charged side chains. It is not surprising that such a mutation occurs, but more that the resulting amino acid change has a noticeable effect on protein structure.

d) The required mutation GTX (where X is A, G, C, or T) to CCX requires two base changes, neither of which involves the highly mutable CG sequence. Thus, the mutation is unlikely. Also, the mutation, from one hydrophobic amino acid to another, is unlikely to have a dramatic effect on protein structure.

e) The required mutation (CTT, CTC, or CTA to ATT, ATC, or ATA) requires only one base change, so it is reasonably likely. However, both amino acids have very similar side chains, so this mutation would not be expected to have a large effect on a protein's structure. It is surprising that individuals with such a mutation show a phenotype at all.

f) The required mutation (GGX to UUX) requires two different mutations, neither of which involves a CG sequence. Therefore, this would not be expected to be very frequent. Also, the mutation makes a dramatic change in the amino acid. Replacing the amino acid with the smallest side chain with an amino acid with a large side chain would have a big effect on protein structure.

(4)

a)

5′-	G	A	A	U	C	G	C	U	A	C	A	A	- 3′

b)

```
5′ TGCCA TCCGA TTGGT GTTCC TTCCA TGAAG GATGC ACAAC GCAAA 3′
3′ ACGGT AGGCT AACCA CAAGG AAGGT ACTTC CTACG TGTTG CGTTT 5′

5′ TACAC GCTTA GCTGA CTATA AGGAC GAATC GCTAC AACGA TGCGA 3′
3′ ATGTG CGAAT CGACT GATAT TCCTG CTTAG CGATG TTGCT ACGCT 5′

5′ TGCCA TCCGA TTGGT GTTCC TTCCA TGAAG GATGC ACAAC GCAAA 3′
3′ ACGGT AGGCT AACCA CAAGG AAGGT ACTTC CTACG TGTTG CGTTT 5′
```

c)

d) You must add ribosomes and aminoacyl-tRNA molecules (charged tRNA molecules).

e) Puromycin mimics a charged tRNA. It would not directly affect transcription.

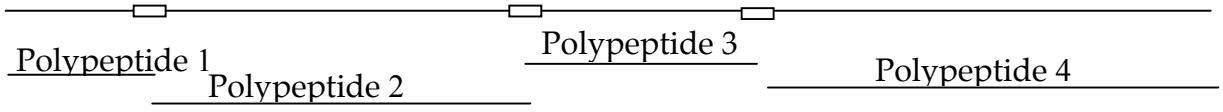
f) Puromycin mimics a charged tRNA. When incorporated into the growing polypeptide, it interrupts translation and results in truncated proteins.

g)

i) In test tube 1 (no puromycin) you get a polypeptide that is 100 amino acids long. Each amino acid represents a codon on the mRNA. A codon is three nucleotides long, so the mRNA was at least 300 nucleotides.

ii) If puromycin does **NOT** recognize a specific codon and is in limiting amounts, then one expects that some translation events will not be interrupted and those that are will be interrupted in a random fashion. The population of polypeptides in tube 2 will include polypeptides of all sizes, i.e., dipeptides, tripeptides, etc.

iii) If puromycin recognizes a specific codon that occurs three times in the mRNA and these codons are not evenly spaced, then you might expect to find four types of polypeptides that are each different lengths.



h)

i)

ii) The alternative sequence has an insertion.

iii) You would not expect this DNA sequence to encode a protein that binds serotonin because all amino acids after the insertion are altered. In this case three of the five amino acids involved in the binding, Ser, Val, and Glu, are after the site of the insertion.

i)

i)

```
5′ ACC AAT GGA CCA GCA GGA AGC GGG GTA GCT GAT TAC 3′
3′ TGG TTA CCT GGT CGT CCT TCG CCC CAT CGA CTA ATG 5′
```

ii) This sequence has a substitution mutation.

iii) You would expect this DNA sequence to encode a protein that binds serotonin. The substitution replaces a Glu with Asp. These two amino acids are very similar.