1. Assume that we have t motifs of length 1. The consensus of these motifs is the nucleotide that appears the most in all the motifs, so by definition, the sum of the hamming distances will be the minimum since P is the same as the most motifs, and P is the median string. Then, for any t motifs of length k, the sum of the total hamming distances will simply be the sum between each set of motifs of length 1 and the corresponding P. Since for each nucleotide in P the sum of hamming distances with the corresponding nucleotides in the motifs is the minimum, the sum for all nucleotides in P will be the minimum, and the consensus P will be the median string of all t motifs.

2. No, randomized algorithms, in the end, are still randomized. Over a large number of trials, they will tend to converge to the correct answer, but there is still a nonzero probability that they will select a consensus that is not the median string. Since it is a randomized algorithm, the selected consensus could be another consensus with a low hamming distance sum, but not the one with the lowest depending on what the original randomly selected motifs are. An example would be a set of sequences with 2 implanted k-mers. The Gibbs sampler could easily find the one with the higher total sum of hamming distance in some cases, since it tends towards a localized minimum.

3. No, as it can only output results that are k-mers that already exist in the motifs, and the correct median string of a given set of motifs may not exist in the set of motifs themselves.

4. GibbsSamplerWithComplement(DNA, k, t, N):
     Randomly select motifs$_i$ from in each string in DNA
     for j <- 1 to N:
       i <- Random(t)
       Profile <- profile matrix formed from all strings and reverse complements in Motifs except for Motif$_i$
       Motif$_i$ <- Profile-randomly generated k-mer in the ith sequence
       if Score(motifs) < Score(bestMotifs):
         bestMotifs = motifs

     Return bestMotifs

5. B, because the median string of B, AAACCGGCAC, bears less relation to any of the potential binding sites as compared to the relation between A and its potential binding sites.

6. GTATGGGTG, GTGGGTATG
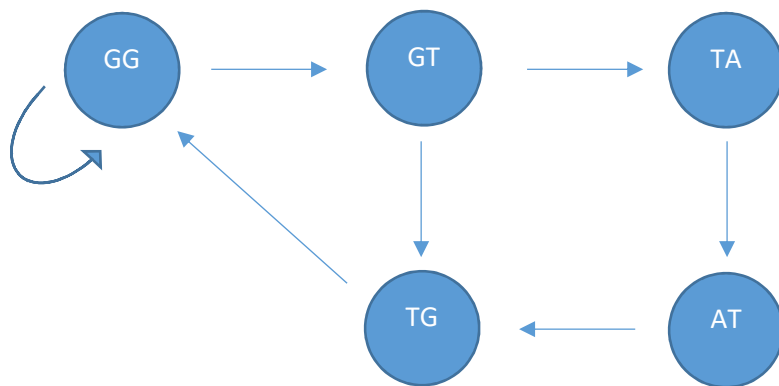
7. Yes, but it is also a cycle since all nodes are balanced. As such, there is no one path to find, but rather many. An example would be ABCFBEADEFIEHDGHIA.

8. For each node, merge each one with its complement, and use the edges to figure out the correct order of each potential assembly sequence.

9. No, since a segment with all of the same nucleotide will get reduced to the length of the largest fragment.

10. [(N, 114) (L, 113) (Y, 163) (V, 99) (NL, 227) (LY, 276) (YV, 262) (VN, 213)
(NLY, 390) (LYV, 375) (YVN, 376) (VNL, 326) (NLYV, 489)]