Take Home Exam: Build, Energy Minimize, and Simulate a Complete Monoclonal Antibody Structure

In this exercise you will be given known coordinates for a majority of a monoclonal antibody which contains protein and carbohydrate components. In addition, you will be given the file with the sequence information of protein. With this information, your task will be: build the entire model, carry out energy minimization, conduct a torsion angle monte carlo (TAMC) simulation, and then use SasCalc to calculate SANS profiles for the ensemble.

Provided Files:

- fab.pdb
- fc.pdb
- carbohydrate.pdb
- heavy_chain_sequence_fasta.txt
- disulfides.txt
- carbohydrate_patches.txt
- charmm36.zip

Part I: Build Complete Structure

A monoclonal antibody (mAb) is a protein that typically consists of two "Fab" domains attached to a single "Fc" domain as shown in Figure 1 below.

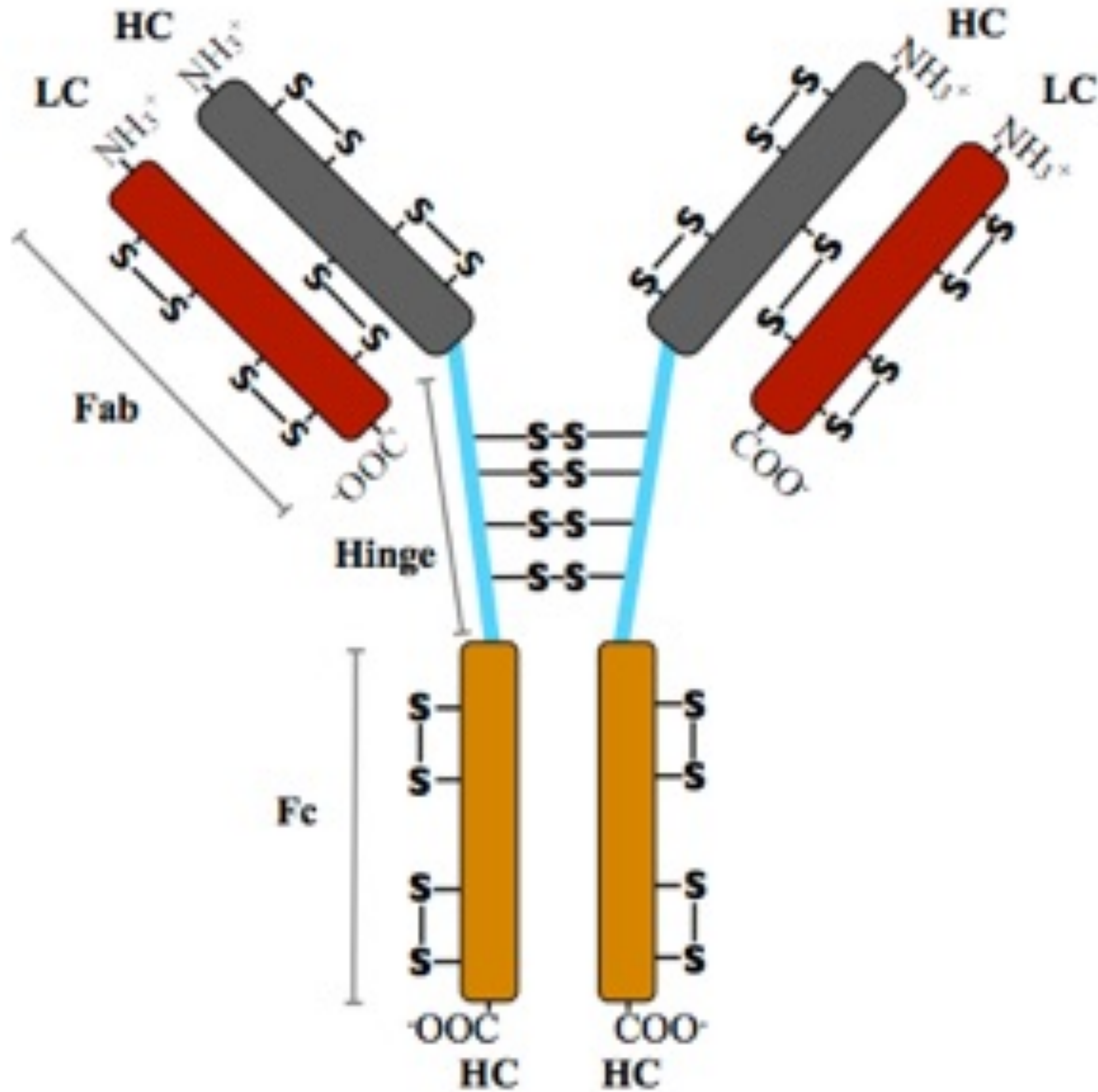

Figure 1. Schematic of the Structure of an Antibody. Note that the number of disulfide bonds in the hinge regions varies between antibody sub-types.

The Fab and Fc domains are linked together via amino acids in a "hinge" region. Another way to describe a mAb is two identical heavy chains (HC) that are covalently bonded to two identical light chains (LC). The HC and LC have several intra-chain and inter-chain disulfide bonds in the Fab and hinge domains, and additional intra-chain disulfide bonds exist in the Fc domains. Typically, mAb have post-translational modifications which include heterogenous carbohydrate chains covalently attached to amino-acid side chains in the Fc domain.

Due to the flexibility of mAb proteins, obtaining the coordinates using X-ray crystallography is

usually restricted to partial structures of Fab and Fc domains obtained without the presence of the remainder of the mAb protein.

In this exercise, you are given coordinates for Fab & Fc domains in separate files, and the full-sequence of a single HC domain is provided. The amino-acid numbering in these files should be used as provided to simplify the model building process. **This is often not the case, and careful inspection is required to provide the correct numbering scheme for general structure building.**

To guide you through the process, the following tasks should be completed. One can use the "Build Utilities" module to rename segments and create sequence pdb files from the provided FASTA sequence file.

1) Open the fab.pdb and fc.pdb files in the same VMD session and manually move and orient the fab region so that the carboxyl terminus the HC portion of the fab domain is reasonably close to one of the HC chains amino terminus of the fc domain. Save the modified coordinates of the fab molecule as fab1.pdb. Repeat the process to move the fab to the other HC chain of the fc domain in a similar manner. Save the modified coordinates as fab2.pdb. Make sure that there is no apparent overlap of fab1.pdb, fab2.pdb and fc.pdb. The fab and fc domains should be within 10-20 angstroms apart.

2) Using the modified fab1.pdb, fab2.pdb & fc.pdb files generated in step 1, create four new pdb files (two for each LC and HC). Name the segments in the files LC1 and LC2 for the LC files and HC1 and HC2 for the HC files. These files will contain the known structural information from the experimentally determined X-ray coordinates.

3) Using full-sequence FASTA file, create two sequence pdb files (one each for HC1 and HC2).

4) Using the provided disulfide bonding information, write a psfgen build script and use psfgen to create a protein-only pdb / psf files. **Note that you will need to use the topology and parameter files provided in the charmm36.zip file.** Read the output of the psfgen build and verify by inspection in VMD that the structure is reasonable.

5) Create constraint PDB files and carry out a series of energy minimization runs to relax the structure:

- constrain all atoms in non-hinge regions
- constrain all heavy atoms
- constrain all backbone atoms

6) Energy minimize the protein-only structure without constraints. Inspect the final structure and make sure that secondary structure elements in the Fab and Fc domains have not been comprimised.

7) With the final coordinates of the relaxed protein-only mAb model, use the provided

carbohydrate.pdb to create two new pdb files with modified coordinates of the carbohydrates only that are roughly oriented so that atom 1 of each sugar is close to the side chain of residue 300 of HC1. Separately, do the same thing for residue 300 of HC2. In both cases, you should be moving and orienting the coordinates in VMD. Save the modified coordinates of the sugars to carbohydrate_1.pdb and carbohydrate_2.pdb. Use build utilities to name the segments CAR1 and CAR2 respectively.

8) Modify the psfgen build script you generated in step (4) to include carbohydrate_1.pdb and carbohydrate_2.pdb files. Use the information in "carbohydrate_patches.txt" that will build the bonding within the carbohydrate molecules and between the carbohydrate molecules and HC1 and HC2.

8) Create constraint PDB files and carry out a series of energy minimization runs to relax the structure:

- constrain all protein atoms
- constrain all heavy atoms
- constrain all backbone atoms

9) Energy minimize the complete structure (protein and carbohydrate) without constraints. Inspect the final structure.

10) Run a 10 ps vacuum MD simulation of the complete system without constraints. Save the final coordinates to a new pdb file.

Part II: TAMC & SasCalc

Using the final coordinates of the full mAb model created in Part I, carry out a TAMC simulation to generate a trajectory with > 20,000 accepted structures by sampling the backbone dihedral angles in the appropriate residues in the hinge regions. Calculate the SANS profiles for your ensemble using SasCalc.