**Homology modeling of antibodies for antibody engineering**

**Maya Sharma , Udhaya Arimpur and Arun Kaniyamattom**

**Abstract**

The immune systems of living organism have defensive mechanism where immunoglobin play important roles by which they protect the organism from the foreign invaders or molecules, so-called antigens. Over the time, the foreign substances encounter inside the immune system, thereafter, they somatically mutate their genome sequences and generate high affinity defense system. Antibody humanization is the process to convert any non-human antibody sequence into humanized antibodies to reduce immunogenicity during drug administration in human.

Amino acid sequence-based homology modeling of any murine antibody protein sequence is used to evaluate antibody structural properties such as binding loop predication, structural stability over antibody humanization and affinity maturation. Energy minimization software is applied to mouse gene antibody and antibody models to describe residues which will experience steric clashes as a result of modification in tertiary conformation. Therefore, a model guided homology antibody provides an application guide note to researchers in antibody and protein engineering such as humanization and affinity maturation process.

**Key words** Antibody-engineering, Sampling, Computer-aided design, Molecular simulations, Scoring, Affinity maturation

**Introduction:**

Homology modeling/comparative modeling is a technique based on biological fact that similar sequences are identical and hence they are similar in structure too (more or less). Sequence of amino acid target protein as well as homologous protein of experimental 3D structure protein of it are used and an unknown atomic resolution of the target protein is created.

Homology models are helpful to induce a rough plan wherever the alpha carbons of key residues in the protein folds helps to guide mutagenesis or gives idea about structure and function relationships of proteins. However, these models are unreliable in predicting the conformations of insertions or deletions, and the details of sidechain positions. These models are unlikely to be helpful in modeling ligand docking used in drug-design experiment unless the sequence identity with the template is >70%, and even then, less reliable than associate empirical crystallographic or NMR .

During the course of the antibody drug development process, conventional mouse monoclonal antibody was administered to humans. Conversely, mouse antibody which could be used in therapeutic applications, may generate human anti-mouse antibody(HAMA) response, resulting in severe allergic reactions in humans. Homology modelling is a vital technique in the field of structural biology which helps narrowing down the gap between the already known sequences of proteins and experimentally derived antibody structures. This also allows researchers to generate reliable protein models without being a specific computational expert. Modelling gives easy access to visualize and comprehend the results obtained.

Humanization of mouse antibodies defines the substitute of mouse framework regions with comparable human germline framework areas based on similar sequence of amino acid. The mouse scaffold is changed with a human scaffold to lessen immunogenicity of the mouse antibody on administration to humans. Further optimization is done by generating the antibody model of mouse, and non-conserved residues exist in this model are modified which reflect germline sequences of human.

non-conserved residues in mouse antibody model are mutated to reflect human germline sequences. After that optimized mouse model’s CDRs and human frameworks are then subjected to energy minimization done with the use of one of the force fields.

This approach is well implemented by researchers using homology model guided antibody engineering process.

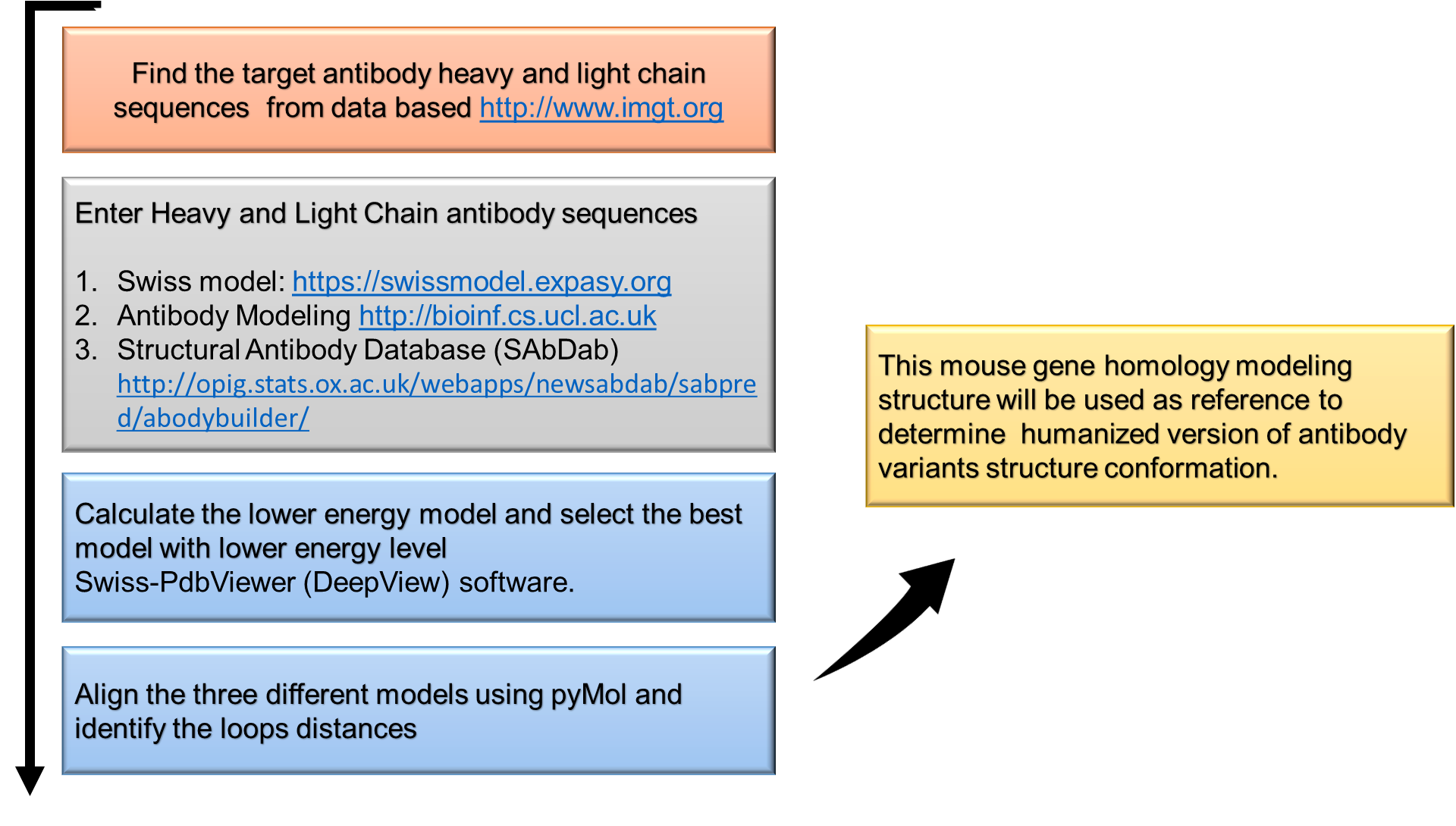
**Goal of this project:**

To construct humanized antibody homology model from primary sequence of mouse antibody.

Screening of surface accessibility to find conformationally exposed residues, and mutate to eliminate potential immunogenicity.

Energy minimization and stereochemical analysis to obtain optimum model.

**Workflow:**



**Antibody structure**

Antibody is Y shape molecule along with heavy chain and light chain. The antibody 1s 150 kDa molecular weight. The antibody has foreign substances binding sites is called commentary determine regions (CDRs). The three CDR regions of heavy chain [

(H-CDR1, H-CDR2, and H-CDR3) and the three CDR regions of light chain (L- CDR1, L-CDR2, and L-CDR3) are denoted. The contact CDR definition of antibody sequence to position the CDRs of the mouse antibody light chain and heavy chains locations are “H-CDR1 26-35, H-CDR2 47-65, H-CDR3 93-101, L-CDR124-36, L-CDR2 46-55, and L-CDR3 89-96”.

**Materials and required software:**

**Mouse antibody sequences and structures**

At first VH and VL chains are defined, sequences of which will be used as the starting material to build models. Antibody gene in the form of mouse fragment antigen-binding sequence of antibody was then obtained from antibody database IMGT [as](http://www.imgt.org/%20%20as) FASTA format file and also antibody sequences with variable heavy and light chain sequences in Fv(fragment variable format). The mouse antibody variable and heavy and mild sequences are used as an input to create a homology version by using any publicly available servers like Swiss model, SAbPred, or Kotai etc. homology modeling software. Finally, visualization of this antibody systems are done through PyMOL.

Swiss model: <https://swissmodel.expasy.org><http://opig.stats.ox.ac.uk/webapps/sabdabsabpred/WelcomeSAbPred.php> ,<https://sysimm.org/tools>

molecular visualization software are PyMOL, DeepView\_SwissPDBViewer are used from <http://spdbv.vital-it.ch/disclaim.html>.

**Mouse antibody sequence**

**>variable heavy chain VH**

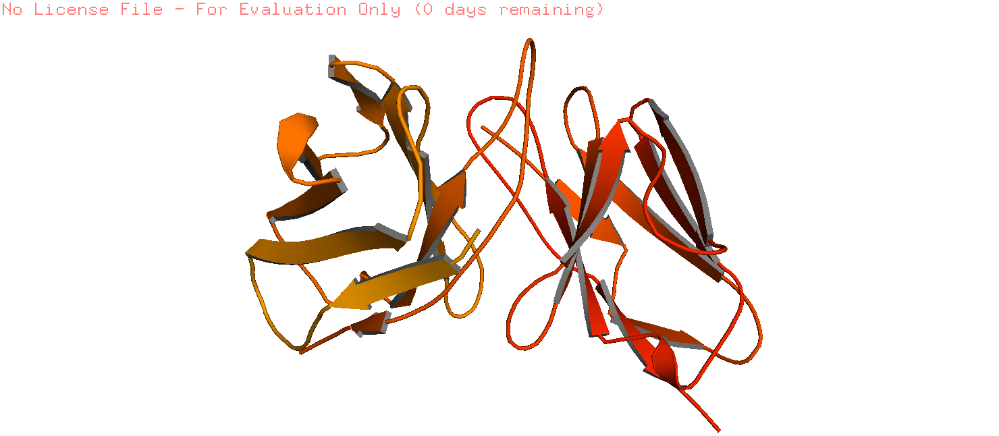
**DVQLQQSGPGLVAPSQSLSITCTVSGFSLTNYGIHWVRQPPGKGLEWLGVIWAGGYTKYNSALMSRLSMSKDNSKSQVFLKMNSLQTDDTAMYYCARDEVRRDYYAMDHWGQGTTVTVSA**

**>Variable light chain (VL)**

**DIQMTQSPAILSASPGEKVTMTCWASSGVSYMHWYQQKPGSSPKPWIFATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSFNPLTFGAGTKLEIKR**

**A close up of a flower

Description automatically generated**

**A B**

C

Fig.1.Antibody generated using these software

**A**. Swiss model: <https://swissmodel.expasy.org> **B.** Structural Antibody Database: [http://opig.stats.ox.ac.uk/webapps/sabdabsabpred/](http://opig.stats.ox.ac.uk/webapps/sabdabsabpred/WelcomeSAbPred.php) **C.** Kotai: <https://sysimm.org/tools>.

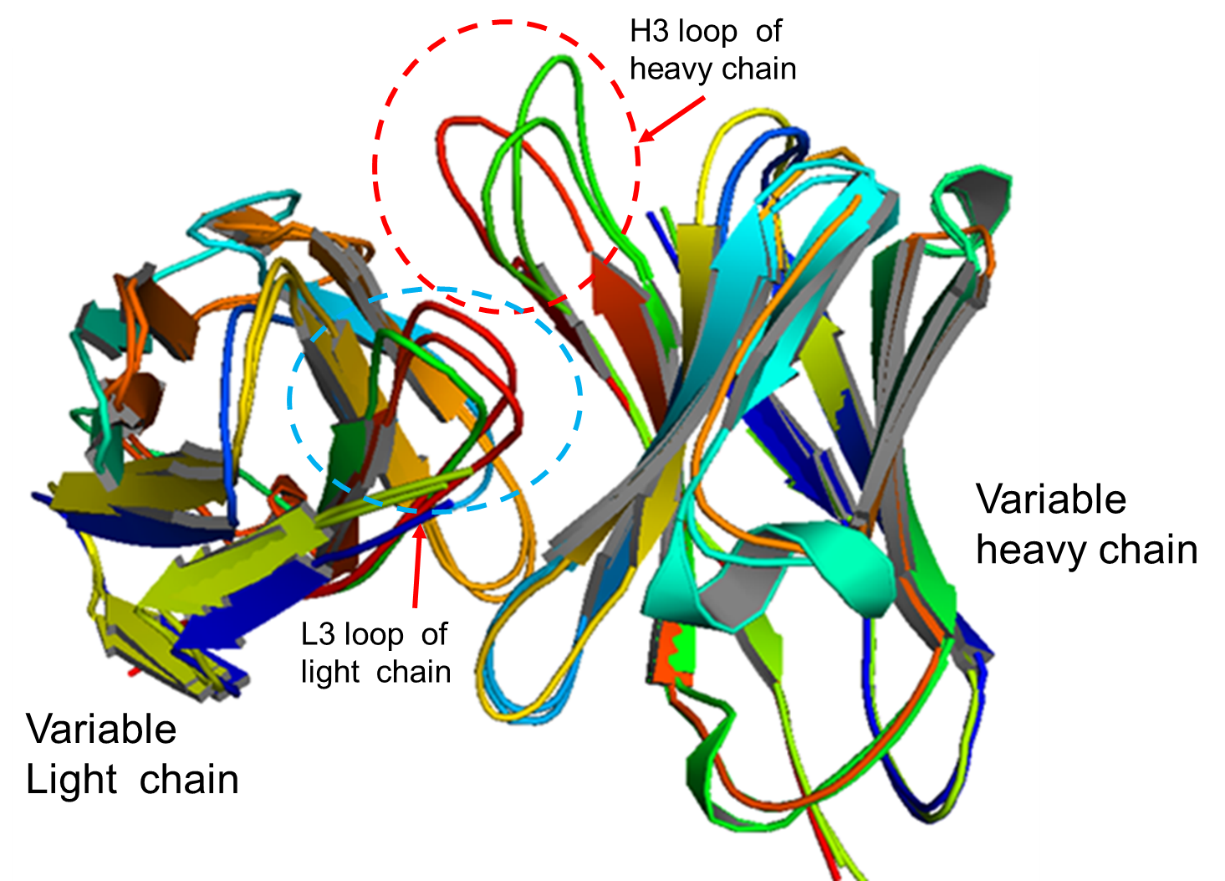
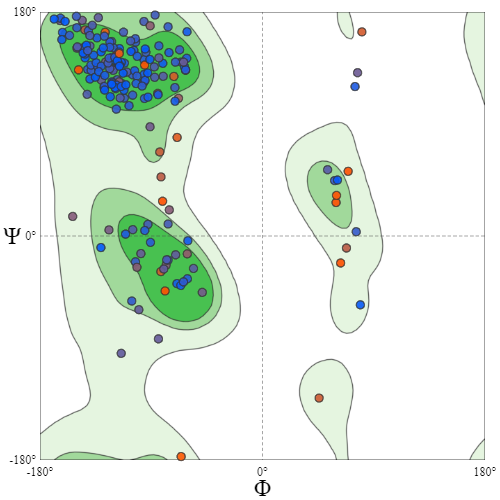


Fig 2. Overlapping of these structures after energy minimization using DeepView-Swiss-PDBViewer. This structure s visualized using pymol.

Ramchandran plot

Mol Probity Results

|  |  |  |  |
| --- | --- | --- | --- |
| Mol Probity Score |  | 2.25 |  |
| Clash Score |  | 4.38 | (A60 ARG-A78 GLU) |
| Ramachandran Favored |  | 90.99% |  |
| Ramachandran Outliers |  | 1.8% | A31 TYR, B4 LEU, B39 GLN, B40 PRO |
| Rotamer Outliers |  | 4.74% | A91 SER, B2 VAL, A46 TRP, A89 GLN, B90 THR, B24 VAL, A29 VAL, A18 LYS, A90 TRP |
| C-Beta Deviations |  | 8 | A31 TYR, B90 THR, B39 GLN, A95 LEU, B96 ALA, B32 TYR, A93 ASN, B92 MET |
| Bad Bonds |  | 0 / 1792 |  |
| Bad Angles |  | 12 / 2437 | B103 ASP, B108 ASP, B90 THR, A4 MET, B96 ALA, A33 HIS, A90 TRP, (B103 ASP-B104 TYR), (B40 PRO-B41 PRO), B109 HIS, B35 HIS |
| Cis Prolines |  | 1 / 11 | (A7 SER-A8 PRO) |
| Twisted Prolines |  | 3 / 11 | (B39 GLN-B40 PRO), (B40 PRO-B41 PRO), (A93 ASN-A94 PRO) |

****

**Fig.3. Ramachandra plot of antibody sequences.**

**Results**

Here 3 mouse antibody homology models were derived using above three servers and were then visualized in the SPDviewer and Pymol as shown in Fig. 1A, B and C. These models were overlapped to evaluate the model errors, especially H3 and L3 loop modeling as shown in Fig.2. The Ramachandra plot as shown in Fig.3, each of residue amino acid visualized in the form of dots in of φ vs. ψ graph, is known as a Ramachandran plot. Residues are shown as blue dots are residue and red dots are selected one. Green lines refer high resolution set proteins which are probability contours. These dots provided information for evaluation of heavy chain and light chain amino acid property, that provide an ideal to select mutant residues at desirable positions during antibody engineering process.

**Conclusion**

Computational based homologues modelling methods are very important experimental structural biology tools used to increase knowledge of the protein world as well as properties. There had been numerous techniques evolved for antibody engineering such as humanization. Mostly, antibody engineering is exertions in depth, additionally it's miles design and cycle approach, mostly expected results shown noticeable drop in affinity towards antigen because of structural and sequence knowledgebase design fashions to optimization of molecules that added values in protein engineering research vicinity.

**Reference**

1.Yoonjoo Choi, Casey Hua, Charles L Sentman, Margaret E Ackerman & Chris Bailey-Kellogg (2015), Counter acting agent humanization by structure-based computational protein plan, m Abs,1045-1057, DOI: 10.1080/19420862.2015.1076600

2. Andrew Waterhouse et.al; SWISS-MODEL: homology demonstrating of protein structures and complexes, Nucleic Acids Inquire about, 2018, Vol. 46

3.Vinodh B Kurella & Reddy Gali; Structure guided homology model-based plan and designing of mouse antibodies for humanization, Bioinformation 180-186 (2014)

4. Paolo Marcatili et, al; Nat. Protoc. 9, 2771–2783 (2014); doi:10.1038/nprot.2014.189; 5. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Proteins; Structure approval by Cα geometry: φ,ψ and Cβ deviation, 2003 15;437-450

6. Rosalba Lepore et, al; PIGSPro: forecast of immunoglobulin structures v2, Nucleic Acids Inquire about, 2017, Vol. 45, Web Server issue