

Application Guide
January 6, 2009

Antibody Modeling:
Improving modeling of hypervariable loops
using unbiased physics based methods

Francisco G. Hernandez-Guzman, PhD
Lead Solutions Scientist - Accelrys

Introduction:

With the advent of high throughput methods in the areas of protein expression, production, purification, and structure determination, more and more antibody structures are available to researchers. However, antibodies are unique in their ability to recognize a vast diverse array of antigens, and therefore it is common that the structure of interest may not be readily available. The challenge in modeling antibodies is that even though the antibody fold is highly conserved, and hence highly amenable to modeling by homology, the complimentary determining regions (CDR) are often diverse. Nevertheless, we believe that high quality models may be constructed by combining homology and de novo methods to provide sensible structural information in the absence of experimentally determined structures. The goal of this application note is to present a method which takes a homology model of an antibody fragment as the starting point, and seeks to refine the challenging H3 CDR loop by using de novo physics based methods that reconstruct the loop ignoring its starting conformation, and thus breaking the bias inherited from the initial template.

Outline:

Part I: Challenges of loop refinement for Homology Models

Part II: Method for loop refinement of CDR loops

Part III: Examples of refinement of the H3 CDR loop of homology models of Fab fragments

In this application note we start our modeling work from a previously built homology model of a known Fab crystal structure. We will describe some of the challenges encountered for de novo refinement of CDR loops when the starting model is of homology origin. Next, we will propose a method for improving our results during modeling, and finally we will end with some examples to illustrate the method. This application note is the continuation of previous work illustrating how to build an initial antibody homology model, as well as some initial data for efforts done in CDR refinement. For access to the preceding work please follow the link to the Application note entitled: Antibody Modeling: A quick reference guide to building Antibody homology models, and graphical identification and refinement of CDR loops in Discovery Studio.

(http://accelrys.com/reference/cases/studies/antibody_modeling_app_guide.pdf).

Part I: Challenges of loop refinement for Homology Models

We have observed that the LOOPER algorithm is capable of achieving a high degree of success in reconstructing loops from crystal structures especially for loops of less than 10 residues in length (Spasov et al, 2008). Success has been routinely defined by reconstruction of loops to less than 2Å backbone RMS deviation from the starting crystal structure as a top solution (Fiser et al, 2000)). However, we're unable to match the same degree of success when dealing with homology model cases. Table 1 presents two test cases which illustrate this point.

PDB target:	PDB template:	Case:	Loop size	Loop sequence	Observation:	Before LOOPER	After LOOPER
2aab	2aab	H3 Loop	8	YVGYHVRW	Crystal Structure	0 Å	1.30Å
2aab	1mf2	H3 Loop	8	YVGYHVRW	Homology Model	4.11 Å	3.81 Å
1bbd	1bbd	H3 Loop	9	YYSYYDMDY	Crystal Structure	0 Å	0.69 Å
1bbd	1ifh	H3 Loop	9	YYSYYDMDY	Homology Model	7.35 Å	5.37 Å

Table 1: Initial RMSD values for backbone atoms of the H3 loop of 2aab and 1bbd before and after refinement using LOOPER. Please note that loop re-construction by LOOPER will ignore the original conformation of the loop. *(Reported RMSD's are for the top scored solution provided by LOOPER.)

In addition, we have noticed that changes in the orientation of side chains neighboring the loop of interest can have drastic effects on the scoring of the loops. To illustrate this point we have done the following analysis using the 1bbd.pdb structure as an example:

We methodically modified the loop environment as defined of all residues within a 5Å radius from the selected H3 loop. In the case of 1bbd, the following residues were selected: Light chain: 42, 52, 52, 55, 59, 60 and 61; and for the Heavy chain: 2, 3, 4, 26, 27, 28, 31, 32, 33, 35, 96, 97 and 108. Out of these residues, the side-chain angles of only a subset could be modified because in many cases changes to the side-chain conformations would lead to clear clashes. Obvious residues like glycines, alanines and prolines could not be modified. Modification of the side-chain conformation was done based on changes to a high likelihood rotamer using the definitions of the rotamer library by Richardson and Richardson (2000). Table 2 summarizes the residues that were modified in this test.

Selected H3 CDR loop neighboring residues			
L-chain residues	Modification	H-chain residues	Modification
51	R1	2	R1
55	R2	3	R1
61	R1	4	R1
		27	R1
		32	R1
		33	R1
		97	R1

Table 2: Summary of residues for which the side-chain conformation was modified to match a rotamer conformation. R1 = rotamer number 1. R2 = rotamer number 2. Rotamers are defined as the most common conformation that a side-chain adopts in proteins, and have decreasing statistical values. i.e. R1 is more likely to occur than R2.

After the side-chains of a select group of residues were modified, LOOPER was used to re-construct the H3 loop. A summary of results is shown in table 3.

Structure	Modifications	H3 loop RMSD to initial crystal structure for the top three LOOPER solutions:		
		Sol #1:	Sol #2:	Sol #3:
1bbd	None	0.69Å	0.70Å	0.82Å
1bbd	Selected* neighboring Side-chains modified	4.03Å	3.20Å	3.75Å

Table 3: RMSD backbone atom results from loop reconstruction of the H3 loop of 1bbd.pdb before and after systematic modifications of selected neighboring side-chains. *Selected neighboring side-chains are listed in Table 2.

From this example we show that systematic modifications to the loop environment for the H3 CDR, can have significant effects on LOOPER's ability to reconstruct the loop with a loop having good RMSD to the initial crystal structure loop conformation.

Part II: Method for loop refinement of CDR loops

Based on the observations made in Part I we formulated the following hypothesis: *if we're able to properly model the neighboring side chains around the loop of interest, then LOOPER may yield higher quality results for the loop conformation than without neighboring side chain optimization.*

In order to test our hypothesis we performed the following steps.

Step 1: We selected the H3 loop and then selected residues around this loop that were within 8Å.

Step 2: We mutated the H3 loop residues to Ala. (Pro, Ala and Gly residues remained as is).

Step 3: On the selection done in Step 1, we used CHIROTOR (Spasov, et al, 2007) to refine the side chain positions of the neighboring residues.

Step 4: We then mutated the loop residues back to their native state.

Step 5: And finally we used LOOPER on the H3 loop.

The logic for mutation of the loop residues to Ala was to allow neighboring side chains enough space to rotate to an optimal conformation, allowing side-chain to back-bone interactions to take precedence as observed by Spasov et al, 2007. The initial premise is that the loop is not at a proper conformation to begin with, and its current conformation may in turn interfere with the proper optimization of the neighboring side chains.

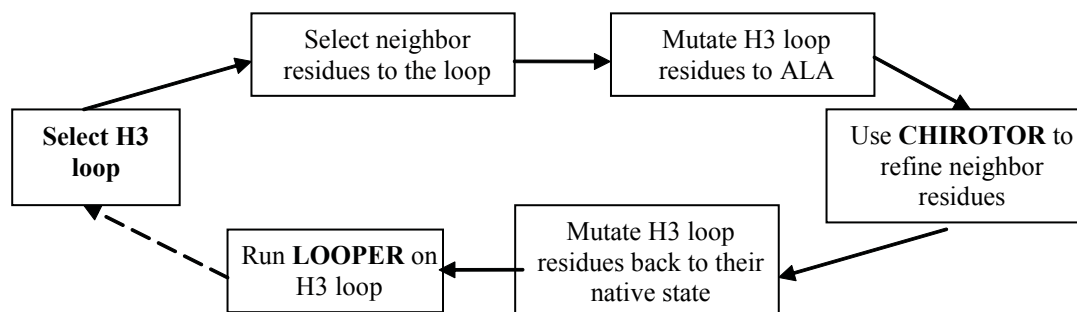


Fig 1: Graphical representation of the proposed workflow for reconstructing flexible loops.

The logic for mutation of the loop residues to Ala was to allow neighboring side chains enough space to rotate to an optimal conformation. Our initial premise is that the loop is not at a proper conformation to start with, and its current conformation might interfere with the proper optimization of the neighboring side chains.

Part III: Example of refinement of the H3 CDR loop of one Fab crystal structure and two Fab fragment homology models

In this section we present a few cases in which we assessed the success of our method proposed in Part III. Cases 1a, 1b and 1c, represent our first example in which we sought to build a homology model of the 2aab crystal structure, and refine the H3 loop. Cases 2a, 2b and 2c, our second example, looks at comparing rebuilding of the H3 loop of the native crystal 1bbd crystal structure, versus the structured prepared in Part II of this application note, as well as a re-refined neighboring side-chains of such structure using the method proposed in Part III. Cases 3 (a,b,c,d,e and f) represent our third example in which we furthered the study by including homology models for the crystal structure of the 1bbd Fab fragment as described in Part I. Table 4 summarizes the results after following the steps suggested in Part III.

Case #:	PDB target:	PDB template:	Case (loop size):	Obs:	Neighboring side-chain refinement	Before LOOPER (top sol.)	After LOOPER (top sol.)
1a	2aab	2aab	H3 Loop (8)	Crystal	-	0 Å	1.30 Å
1b	2aab	1mf2		Homology Model	NO	4.11 Å	3.81 Å
1c	2aab	1mf2	H3 Loop (8)	Homology Model	YES	4.11 Å	1.79 Å
2a	1bbd	1bbd	H3 Loop (9)	Crystal	-	0 Å	0.69 Å
2b	1bbd	1bbd	H3 Loop (9)	Crystal *	NO	0 Å	4.03 Å
2c	1bbd	1bbd	H3 Loop (9)	Crystal *	YES	4.03 Å	1.10 Å
3a	1bbd	1iffh	H3 Loop (9)	Homology Model	NO	7.35 Å	5.37 Å
3b	1bbd	1iffh	H3 Loop (9)	Homology Model	YES	7.35 Å	4.92 Å
3c	1bbd	1txv	H3 Loop (9)	Homology Model	NO	8.74 Å	3.53 Å
3d	1bbd	1txv	H3 Loop (9)	Homology Model	YES	8.74 Å	3.63 Å
3e	1bbd	Chimera 1iffh(L)-1txv(H)	H3 Loop (9)	Homology Model	NO	8.69 Å	7.30 Å
3f	1bbd	Chimera 1iffh(L)-1txv(H)	H3 Loop (9)	Homology Model	YES	8.69 Å	2.54 Å

Table 4: RMSD backbone results from the proposed de novo refinement method using LOOPER for homology models built for crystal structures 2aab and 1bbd. * The 1bbd crystal structure was modified as described in Part II. ** This structure was modified according to Part II and then re-refined using CHIROTOR (Spassov, et al 2007).

Discussion:

In this application note we followed the hypothesis that even minor side chain conformational changes of surrounding residues, can have drastic effects in the successful reconstruction of loops using the de novo loop building tool LOOPER. We proposed a method for optimizing the neighboring side-chains, and we validated the

method with three examples (cases 1,2 and 3), two of which (1 and 3) where from a homology model origin. Finally, the goal was to present a reasonable loop conformation by fully reconstructing the H3 loop and assessing success via an RMS comparison to the crystal structure.

In the first test case of the homology modeling set (2aab – cases 1b and 1c), the proposed method worked fairly straight forward. The RMS changed from 3.81Å to 1.79Å (case 1c) after optimization indicating a significant improvement in the starting loop conformation. See figure 1.

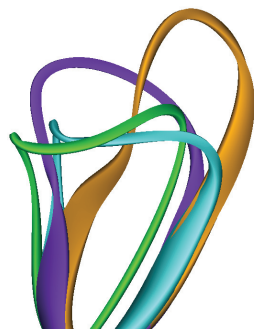


Figure 1: Green: 2aab crystal structure for the H3-loop. Orange: H3-loop of homology model of 2aab before LOOPER optimization (Backbone RMSD = 4.11Å). Purple: H3-loop of homology model of 2aab after LOOPER optimization, but without neighboring side-chain optimization (Backbone RMSD = 3.81Å). Cyan: H3-loop of homology model of 2aab after neighboring side-chain optimization and LOOPER refinement (Backbone RMSD = 1.79Å).

In the second case (cases 2a, 2b and 2c) we provided proof that in the case of crystal structures, it is possible to refine modified side chains of crystal structures using CHIROTOR and successfully retrieve a reasonable loop conformation using LOOPER (case 2c). See figure 2.

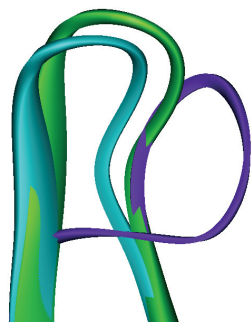


Figure 2: Green: 1bbd crystal structure for the H3-loop. Purple: H3-loop of the 1bbd crystal structure after LOOPER and select modification of neighboring side-chains according to Part II (Backbone RMSD = 3.83Å). Cyan: H3-loop of the 1dbd crystal structure after neighboring side-chain optimization using CHIROTOR (Spassov, et al. 2007) and LOOPER (Backbone RMSD = 1.10Å).

The third case was much more challenging, and several homology models had to be considered (cases 3a through 3f). Homology models using 1lff.pdb as a template had high homology to the VL domain but not the VH. Slight but significant structural changes to the Heavy chain were observed and it was to no surprise that the there was no significant improvement in loop conformation. (See Figure 3a).

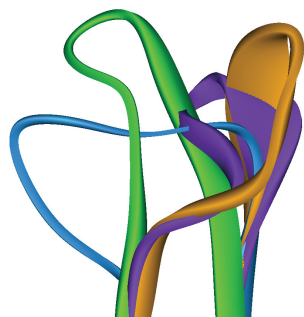


Figure 3a: Green: 1bbd crystal structure. Orange: Homology model of 1bbd using 1lff as a template with no refinement (Backbone RMSD = 7.34Å). Purple: Homology model of 1bbd using 1lff as a template with LOOPER refinement (Backbone RMSD = 5.37Å). Blue: Homology model of 1bbd using 1lff as a template with neighbor side-chain optimization followed by LOOPER refinement. (Backbone RMSD = 4.92Å)

The 1txv.pdb molecule was also used a template because it had the highest sequence homology to the Heavy chain. In this case we saw improvement from 8.74Å to 3.63Å, but this is still considered unacceptable. (see Figure 3b).

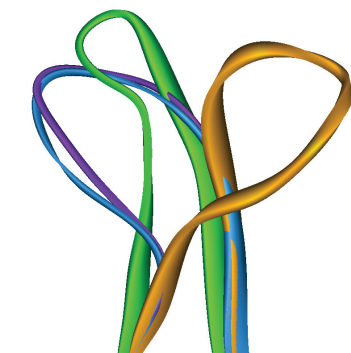


Figure 3b: Green: 1bbd crystal structure. Orange: Homology model of 1bbd using 1txv as a template with no refinement (Backbone RMSD = 8.78Å). Purple: Homology model of 1bbd using 1txv as a template with LOOPER refinement (Backbone RMSD = 3.53Å). Blue: Homology model of 1bbd using 1txv as a template with neighbor side-chain optimization followed by LOOPER refinement. (Backbone RMSD = 3.63Å)

Lastly we constructed a chimera model using the Light chain from 1ifh.pdb and the Heavy chain from 1txv.pdb. In this case we were able to significantly improve the loop conformation from 8.69Å down to 2.54Å which was almost within the success limit of 2.0Å RMSD. (See Figure 3c).

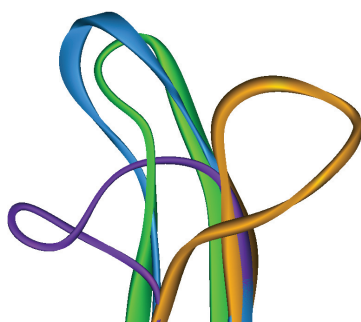


Figure 3c: Green: 1bbd crystal structure. Orange: Homology model of 1bbd using a chimera model of 1ifh for the light chain, and 1txv for the heavy chain as a template with no loop refinement (Backbone RMSD = 8.67Å). Purple: Homology model of 1bbd using the chimera template followed by LOOPER refinement (Backbone RMSD = 7.36Å). Blue: Homology model of 1bbd using the chimera template with neighboring side-chain optimization followed by LOOPER refinement. (Backbone RMSD = 2.51Å)

Further inspection by superposition of the 1bbd.pdb crystal structure to the templates used for homology modeling, revealed that part of the reason for having limited success was that none of the templates were able to account for a significant structural shift of the N-terminus (a 3.5 Å shift for the C_α atom position) of the heavy chain residues which shared common contacts with the H3 loop. For this reason, most loop predictions done in this area were shifted and gave less accurate conformations. (See Figure 4).

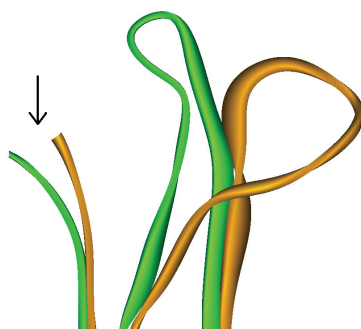


Figure 4: Green: 1bbd crystal structure. Orange: Template structure used to build homology models of 1bbd. On the right is the H3-loop, and on the left the N-terminus residues of the Heavy chain. A significant displacement of the N-terminus is required to fully accommodate the H3 loop in the conformation found in the crystal structure (N-terminus pointed by the black arrow).

Conclusion:

In this application note we have presented a workflow for reconstructing highly flexible loops such as the H3 loop in Antibodies. The workflow is applied on difficult cases such as rebuilding of loops from homology model origin, and we propose that proper orientation of neighboring side chains to the loop may be critical to the proper modeling of such loop. We presented an example to illustrate that relatively small changes can drastically change proper prediction of loops, and therefore modeling of loops from a homology model origin may be an even greater challenge given the inherent inaccuracies throughout the model. Using de novo loop reconstruction methods such as LOOPER, and side chain reconstruction technology such as CHIROTOR, can be used to tackle with a high degree of success the difficult problem of loop refinement. The examples presented indicate that it is possible to achieve a reasonable degree of accuracy in predicting proper loop orientations when considering a homology model as a starting point. Our examples have focused on Antibodies, but the methodology can be used for any family of proteins requiring loop refinement.

Modules used:

DS MODELER
DS CHARMm
DS Biopolymer
DS Protein Refine (LOOPER & ChiRotor)

References

1. Al-Lazikani, B., Chothia, C., and Lesk, A. M., "Standard conformations for the canonical structures of immunoglobulins," *J. Mol. Biol.*, **1997**, 273(4), 927.
2. Berman, et al. [The Protein Data Bank](#). Nucleic Acids Research, **2000**, 28, 235-242
3. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M., "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations," *J. Comput. Chem.*, **1983**, 4, 187-217.
4. Chang, C., Hernandez-Guzman, F.G., Luo, W., Wang, X., Ferrone, S., Ghosh, D., "Structural basis of antigen mimicry in a clinical relevant melanoma antigen system." **2005**, *J. Biol. Chem.* 280, 41546 – 41552
5. Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., et al., "Conformations of immunoglobulin hypervariable regions," *Nature*, **1989**, 342, 877.
6. Edwards, D.J., Hubbard, R.E., and Brady, R.L., "Homology Modeling of Antibody Combining Sites," *ImmunoMethods*, **1992**, 1, 71.
7. Eswar, N., Eramian, D., Webb, B., Shen, M., Sali, A. Protein Structure Modeling With MODELLER. *Current Protocols in Bioinformatics John Wiley & Sons, Inc.*, **Supplement 15**, 5.6.1-5.6.30, 2006
8. Fiser, A., Gian Do, R.K. and Sali, A. "Modeling of loops in protein structures" **2000**, *Protein Science* 9: 1753 – 1773.
9. Imai, K. and Takaoka, A., "Comparing antibody and small-molecule therapies for cancer," *Nature Reviews Cancer*, **2006**, 6, 714 – 727.
10. Lescar, J., Stouracova, R., Riottot, M.M., Chitarra, V., Brynda, J., Fabry, M., Horejsi, M., Sedlacek, J., Bentley, G.A. "Three-dimensional structure of an Fab-peptide complex: structural basis of HIV-1 protease inhibition by a monoclonal antibody," *J. Mol. Biol.* **1997**, 267, 1207-1222
11. Lowell, S., Word, J.M., Richardson, J., Richardson, D., *Proteins: Structure, 2000, Function and Genetics*, 40, 389 – 408.
12. Morea, V., Tramontano, A., Rustici, M., Chothia, C., and Lesk, A. M., "Conformation of the third hypervariable region in the Vh domain of antibodies," *J. Mol. Biol.*, **1998**, 275, 269.
13. Morea, V., Lesk, A., and Tramontano, A. "Antibody Modeling: Implications for Engineering and Design," *METHODS*, **2000**, 20, 267.
14. Sali, A., Pottertone, L., Yuan, F., van Vlijmen, H., and Karplus, M., "Evaluation of comparative protein modeling by MODELLER," *Proteins*, **1995**, 23, 318.
15. Spassov, V., Yan, L. and Flook, P. "The Dominant Role of Side-chain Backbone Interactions in Structural Realization of Amino-acid Code. ChiRotor: a Side-chain Prediction Algorithm Based on Side-chain Backbone Interactions," *Protein Science*, **2007**
16. Spassov, V., Flook, P. and Yan, L. "LOOPER: a molecular mechanics-based algorithm for protein loop prediction." *Protein Eng. Des. Sel.* **2008** **21(2):91-100**
17. Tormo, et al. Three-dimensional structure of the Fab fragment of a neutralizing antibody to human rhinovirus serotype 2. *Protein Sci.* **1992**, 1, 1154-1161