Supplementary Information

Tables

Supplementary Table 1: Disulfide bond features. Table showing the C_α-C_α distance of the cystines in CDR H1 and H3 loops that form a disulfide bond. Details on the position of the cystines are also shown. [s]stretched twist, [t]twist.

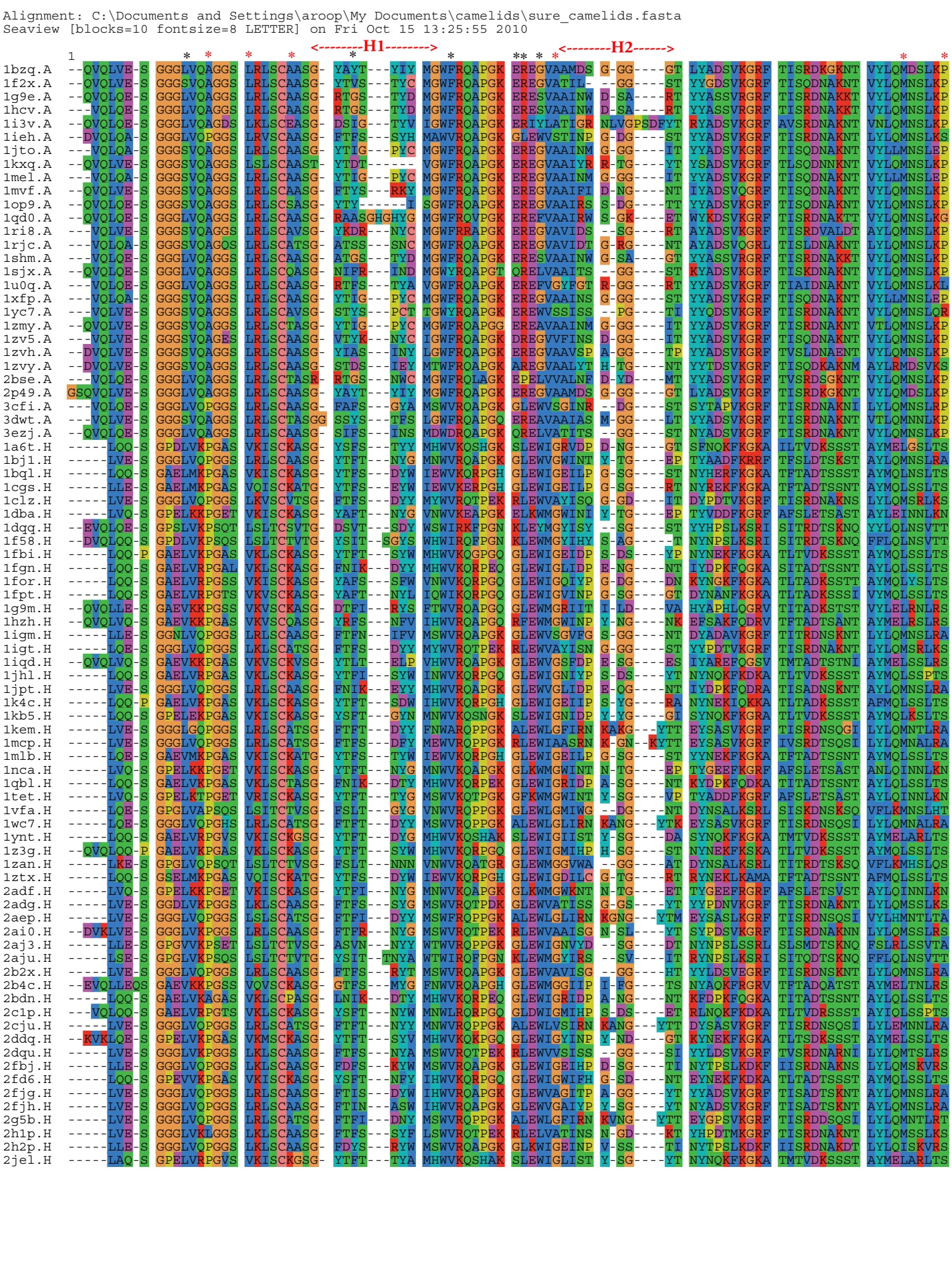
PDB ID	H3 Length	Cys(C_α)-Cys(C_α) (Å)	Cys Position in H1	Cys Position in H3	Relative Position of Cys in H3		H3 Midpoint
					End	Start	
1YC7	10	4.0	32	96	*n*-8	2	5.0
1ZV5[t]	12	6.1	33	100B	*n*-4	8	6.0
1RI8[s]	17	5.7	33	100B	*n*-9	8	8.5
1F2X[t]	19	5.7	33	100C	*n*-10	9	9.5
1RJC[s]	19	5.9	33	100D	*n*-9	10	9.5
1JTO[s]	24	5.7	33	100E	*n*-13	11	12.0
1MEL[s]	24	5.7	33	100E	*n*-13	11	12.0
1XFP[s]	24	5.7	33	100E	*n*-13	11	12.0
1ZMY[s]	24	5.8	33	100E	*n*-13	11	12.0
Mean Distance		5.6					
Standard Deviation		0.6					

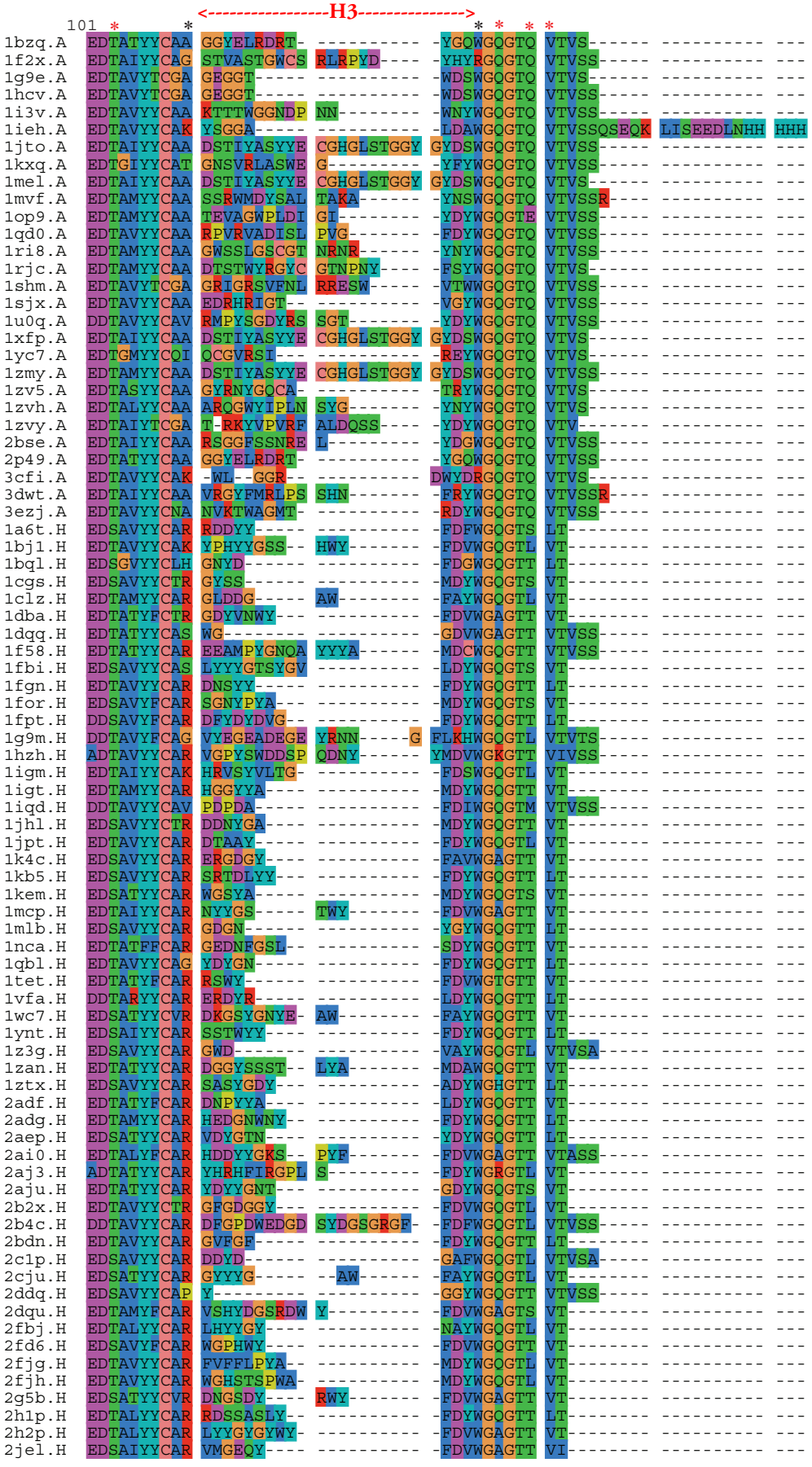
Supplementary Table 2: Comparison of different loop building techniques for long CDR H3 loops (n_{H3} ≥ 16 residues) using global rmsds. *LowE*, *LowRMS* and *LowALL* indicates the lowest-scoring CDR H3 loop model, the CDR H3 with the lowest global rmsd in the ten-lowest scoring models and the model with the lowest global-rmsd observed respectively. [s] and [t] indicates structures whose CDR H3 loops adopt the stretched-twist and twist conformations respectively. [c] indicates that the native contains a disulfide bond between CDRs H1 and H3.

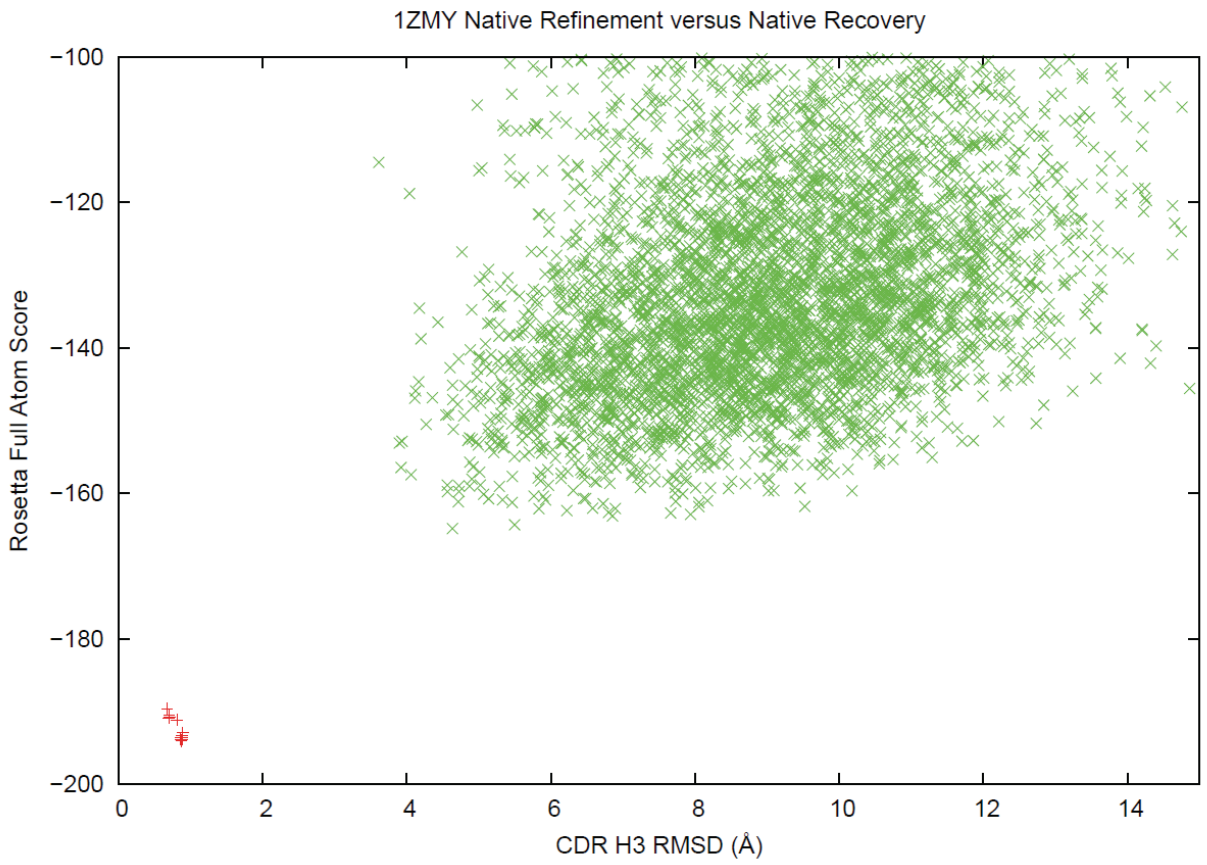
No. of models		No constraints with nine residue fragments only (Å) 5,000			Constraints with nine residue fragments only (Å) 5,000			Constraints with three residue fragments following nine residue fragments (Å) 5,000			Constraints with three residue fragments following nine residue fragments (Å) 20,000		
PDB ID	CDR H3 length	LowE	LowRMS	LowALL	LowE	LowRMS	LowALL	LowE	LowRMS	LowALL	LowE	LowRMS	LowALL
1ZVH[s]	16	7.3	3.3	2.9	5.3	2.3	2.2	4.5	3.4	2.3	4.9	3.2	2.3
1U0Q[s,c]	16	2.6	2.6	2.5	3.0	2.8	2.8	2.8	2.8	2.8	3.1	2.8	2.1
3DWT[s]	16	8.0	3.3	2.6	4.2	3.1	2.5	3.3	3.3	2.6	3.2	3.2	2.6
1QD0[s]	16	7.4	3.1	2.5	1.5	1.5	1.5	0.9	0.9	0.9	0.9	0.9	0.9
1MVF[s]	17	9.3	4.0	3.5	4.0	3.9	3.2	3.4	3.4	3.4	2.3	2.3	1.8
1RI8[s,c]	17	10.0	4.6	4.0	4.8	3.3	2.7	5.1	4.0	2.6	3.6	2.6	2.5
1ZVY[s]	18	6.6	6.6	3.0	3.9	3.1	2.8	4.3	3.3	3.0	4.3	3.4	2.2
1F2X[t,c]	19	14.7	9.0	4.5	5.4	4.6	3.5	3.9	3.9	3.3	5.7	3.3	3.0
1RJC[s,c]	19	7.1	6.3	4.4	2.5	2.5	2.5	3.6	3.6	2.8	3.6	2.0	2.0
1JTO[s,c]	24	10.8	5.2	4.8	4.5	4.1	3.7	6.7	4.3	3.4	5.4	4.3	3.4
1MEL[s,c]	24	18.6	6.2	3.6	5.1	4.6	3.5	4.0	4.0	3.8	6.3	4.0	3.3
1XFP[s,c]	24	12.9	9.4	5.2	5.0	5.0	3.2	4.2	4.2	3.9	4.5	4.2	3.4
1ZMY[s,c]	24	7.8	7.6	4.8	4.6	4.6	3.6	3.9	3.9	3.8	4.0	3.9	3.3
Median		8.0	5.2	3.6	4.5	3.3	2.8	3.9	3.6	3.0	4.0	3.2	2.5

Figure

Supplementary Figure 1: Sequence alignment of cAb V_HH and classical V_HS. The camelid V_HH are listed first with the letter "A" following the four letter PDB ID. Red (*) and black stars (*) indicate alignment positions of newly and previously discovered amino acid mutations. The classical V_HS have the letter "H" following their PDB IDs. The CDR regions are annotated in red. Basic: K,R (red); Hydrophobic: A,F,I,L,M,V,W (blue); Polar: N,Q,S,T (green); H,Y (cyan); C (salmon); Acidic: D,E (magenta); Proline (yellow); Glycine (orange).


Alignment: C:\Documents and Settings\aroop\My Documents\camelids\sure_camelids.fasta
Seaview [blocks=10 fontsize=8 LETTER] on Fri Oct 15 13:25:55 2010
<--------H1-------->
<--------H2------>
1bzq.A
1f2x.A
1g9e.A
1hcv.A
1i3v.A
1ieh.A
1jto.A
1kxq.A
1mel.A
1mvf.A
1op9.A
1qd0.A
1ri8.A
1rjc.A
1shm.A
1sjx.A
1u0q.A
1xfp.A
1yc7.A
1zmy.A
1zv5.A
1zvh.A
1zvy.A
2bse.A
2p49.A
3cfi.A
3dwt.A
3ezj.A
1a6t.H
1bj1.H
1bql.H
1cgs.H
1clz.H
1dba.H
1dqq.H
1f58.H
1fbi.H
1fgn.H
1for.H
1fpt.H
1g9m.H
1hzh.H
1igm.H
1igt.H
1iqd.H
1jhl.H
1jpt.H
1k4c.H
1kb5.H
1kem.H
1mcp.H
1mlb.H
1nca.H
1qbl.H
1tet.H
1vfa.H
1wc7.H
1ynt.H
1z3g.H
1zan.H
1ztx.H
2adf.H
2adg.H
2aep.H
2ai0.H
2aj3.H
2aju.H
2b2x.H
2b4c.H
2bdn.H
2c1p.H
2cju.H
2ddq.H
2dqu.H
2fbj.H
2fd6.H
2fjg.H
2fjh.H
2g5b.H
2h1p.H
2h2p.H
2jel.H


H3
101



Supplementary Figure 2: Plot of Rosetta energies of models versus RMSD from the native crystal structure in Ångstroms, comparing refinement of crystal conformation (red) and *ab initio* loop building (green) of the CDR H3 loop in the native crystal environment.

Rosetta Command Lines

The protocols used in this paper are available in Rosetta 3.2 (Fall 2010) using the following command lines, where `$ROSETTA3` and `$ROSETTA3DB` are environment variables pointing to the Rosetta 3.0 source and database paths. `$RosettaAntibody` is the environment variable pointing to the RosettaAntibody directory, typically `$ROSETTA3/antibody`.

CDR Grafting

```
$ROSETTA3/antibody_mode.default.linuxgccrelease
     -database $ROSETTA3DB -nstruct 1
     -s FR02.pdb  -native 1mvf.pdb -in::file::fullatom -out::file::fullatom -
     out::prefix aa
     -graft_h1 -graft_h2 -graft_h3 -camelid
     -ex1aro -ex1 -ex2
     -detect_disulf true -run::rebuild_disulf true -run::find_disulf true
     -mute core -unmute protocols.AntibodyModeler -out::level 550
```

CDR H3 Loop Modeling

```
$ROSETTA3/antibody_mode.default.linuxgccrelease
     -database $ROSETTA3DB -nstruct 5000
     -s FR02.pdb -in::file::native 1mvf.pdb -in::file::fullatom -out::file::fullatom -
     out::prefix aa
     -model_h3 -camelid -constraints::cst_file camelid.cst
     -loops::strict_loops -loops::frag_sizes 9 3 -loops::frag_files
     aaFR02_09_05.200_v1_3 aaFR02_03_05.200_v1_3
     -ex1aro -ex1 -ex2
     -mute core -unmute protocols.AntibodyModeler -out::level 550
     -in::fix_disulf FR02.disulf -run::rebuild_disulf [optional flags: only if H1 and
     H3 both have CYS]
```

Master Script

The cAb V_HH sequence should be provided in a file, e.g. `query_h.fasta`. A master script executed from the directory containing the FASTA sequence will carry out all steps necessary for camelid antibody modeling:

```
perl -P $RosettaAntibody/scripts/ram_buildloop_wrapper.pl pdb1xyz_chothia.pdb 1 1 1 1
     1 1 5000 query_h.fasta camelid
```

Fragments

Protein fragments must be built for loop building. H3 sequences plus one N-terminal and two C-terminal step residues can be submitted to the Robetta server (*http://www.robetta.org*) or generated with the nn_make.pl script included in the Rosetta2.x distribution.