

---

# Loopholes and missing links in protein modeling

---

KAREN A. ROSSI, CAROLYN A. WEIGELT, AKBAR NAYEEM,  
AND STANLEY R. KRISTEK JR.

Computer-Assisted Drug Design, Pharmaceutical Research Institute, Bristol-Myers Squibb Company,  
Princeton, New Jersey 08543 USA

(RECEIVED March 21, 2007; FINAL REVISION June 8, 2007; ACCEPTED June 9, 2007)

## Abstract

This paper provides an unbiased comparison of four commercially available programs for loop sampling, Prime, Modeler, ICM, and Sybyl, each of which uses a different modeling protocol. The study assesses the quality of results and examines the relative strengths and weaknesses of each method. The set of loops to be modeled varied in length from 4–12 amino acids. The approaches used for loop modeling can be classified into two methodologies: ab initio loop generation (Modeler and Prime) and database searches (Sybyl and ICM). Comparison of the modeled loops to the native structures was used to determine the accuracy of each method. All of the protocols returned similar results for short loop lengths (four to six residues), but as loop length increased, the quality of the results varied among the programs. Prime generated loops with RMSDs <2.5 Å for loops up to 10 residues, while the other three methods met the 2.5 Å criteria at seven-residue loops. Additionally, the ability of the software to utilize disulfide bonds and X-ray crystal packing influenced the quality of the results. In the final analysis, the top-ranking loop from each program was rarely the loop with the lowest RMSD with respect to the native template, revealing a weakness in all programs to correctly rank the modeled loops.

**Keywords:** loop modeling; homology modeling; protein structure

The determination of the three-dimensional structure of a protein is a key component in understanding the mechanism of action, and is a fundamental part of rational drug design (Hillisch and Hilgenfeld 2003). Unfortunately, the three-dimensional structure for many therapeutically relevant targets is often not available at the initiation of a drug discovery program. In these cases, comparative modeling can often provide a useful three-dimensional model for a protein target that is related to one or more known protein structures. Templates for modeling are selected from the same protein family or proteins with similar biochemical function (Hillisch et al. 2004; Jacobson and Sali 2004; Wieman et al. 2004; Esposito et al. 2006).

Even within a family of homologous proteins, functional variations can arise as a consequence of structural

differences which are often found on the protein surface. Such structural differences within a family are often a result of amino acid substitutions, insertions, and deletions, and are typically found on exposed loop regions connecting elements of conserved secondary structure in the protein fold. Thus, loops frequently determine the functional specificity of a given protein framework, and can contribute to the active and binding sites of proteins. Accuracy of loop modeling can therefore play a critical role in structure-based design when studying interactions between a cognate protein and its ligand(s). Examples of these types of interactions include the use of models for ligand binding site recognition (Jones and Thornton 1997; Fetrow et al. 1998; Russell et al. 1998; Kasuya and Thornton 1999; Kleywegt 1999; Wei et al. 1999), virtual screening (Kairys et al. 2006), and ligand docking (Kick et al. 1997; Hobrath and Wang 2006).

Numerous examples demonstrating the successful application of comparative protein modeling in drug discovery are well documented (Hillisch et al. 2004; Jacobson and

---

Reprint requests to: Karen A. Rossi, Bristol-Myers Squibb Company, 311 Pennington-Rocky Hill Road, Hopewell, NJ 08534-5400, USA; e-mail: karen.rossi@bms.com; fax: (609) 818-3545.

Article published online ahead of print. Article and publication date are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.072887807>.

Sali 2004). In several such cases, the success of the method relies on the accuracy of predicting the conformation of one or more loops that are near the active site, and which can, therefore, influence ligand binding. Some recent examples include the modeling of the *P* and *R* loops in DP8 and DP9 in order to explain the observed lack of selectivity of some DP4 inhibitors (Rummey and Metz 2007), the modeling of the second extracellular loop in the dopamine receptor (a G protein-coupled receptor) in order to explain the interaction with substituted benzamides (Kortagere et al. 2006), and the modeling of the 10-residue-long linker domain containing the active site residues in 1-deoxy-D-xylulose-5-phosphate reductoisomerase, a potential antimalaria target (Singh et al. 2006). Furthermore, in the modeling of kinases, the conformation and orientation of the activation loop is critical in discerning the catalytically active from inactive form of the protein (Kornev et al. 2006, and references therein).

The studies cited above underscore the need for reliable methods for loop structure prediction because loops are often an essential component of the functional domain of proteins. Because of the observed variation in conformation, and thus flexibility, even within a given protein family, *ab initio* approaches to loop modeling that do not depend on homology modeling have emerged. Most *ab initio* methods for loop prediction deal with solvent-exposed loops in globular proteins (Jacobson et al. 2004; Rohl et al. 2004). However, in cases involving partially buried loops, more sophisticated protocols that involve molecular dynamics have evolved as alternatives (Mehler et al. 2006). The primary purpose of our study is to evaluate the relative performance of different algorithms for modeling a wide range of loops from a selection of globular proteins. The algorithms do not include the protocol described by Mehler et al. (2006), and transmembrane proteins have been excluded from this study.

Loop modeling can be viewed as a constrained protein folding problem. The conformation of a given segment of a polypeptide chain (i.e., the loop) is modeled from the sequence of the segment using geometric constraints imposed by the backbone atoms at the N and C termini that anchor the loop to the remainder of the protein. Studies have shown that identical peptide segments of up to nine residues can have entirely unrelated conformations in different proteins (Sander and Schneider 1991; Cohen et al. 1993; Mezei 1998). Thus, the conformation of a given loop is determined not only by the anchor regions that orient the loop, but also by the protein architecture that surrounds the loop.

The relative performance of various commercial software for building homology models has been the subject of recent studies (Wallner and Elofsson 2005; Nayeem

et al. 2006). These studies focused on comparing the relative quality of sequence alignments and the accuracy of the 3D models that are produced based on the correct sequence alignment. Since sequence alignments contain relative insertions and deletions, loop modeling was not examined exclusively. The present study focuses solely on loop modeling, assessing the relative performance of four commercial software packages. The loop-modeling programs can be classified into those that use (1) *ab initio* methods, as exemplified by Prime (Schrödinger, LLC) and Modeler (Accelrys Software Inc.), and (2) database searches, as exemplified by Sybyl (Tripos, Inc.) and ICM (Molsoft LLC). For a given loop sequence, the *ab initio* programs perform a conformational search and rank the subsequent loops using a scoring function. The database methods search protein structural databases for protein backbone segments that best fit the anchor regions of a selected loop (the parts of the main chain that precede and follow the loop); segments are then superimposed and annealed onto the anchor regions, optionally followed by refinement, and scored in order of “fitness.”

Four software packages were selected based on commercial availability and prevalence of use in the pharmaceutical industry. Historically, Modeler has been the standard for homology modeling; Prime has rapidly become a viable alternative to Modeler but has not been widely studied in the literature. Sybyl and ICM both utilize database search techniques with different methods of assessing the constructed loops.

The loop data set used in this study is a hand-picked subset of the loops used by Jacobson et al. (2004), Fiser and Sali (2000), and Xiang et al. (2002), specifically selected to span a wide range of protein architectures and structural classes. Individual program performance was evaluated based on the RMSD of the calculated loops with respect to the native protein and comparing the results thusly obtained using these different methodologies. In addition, attention was focused on the relative quality of the rank ordering of the loops produced for each program. Since the rank ordering (scoring) of the resultant loops is a major consideration when building loops for homology models, the assumption that the top-ranked loop is correct was examined in this study. Finally, a critical assessment, including visual evaluation, was performed to exemplify the strengths and weaknesses of each method.

## Methods

### *Data set selection*

The set of loops to be examined in this study was chosen from a previously compiled filtered loop database described by Jacobson et al. (2004). All protein structures in

the data set were determined by X-ray crystallography at  $\leq 2.0$  Å resolution and are diverse, sharing a maximum sequence identity of  $\leq 60\%$ . Loops selected for this study were also required to meet the following criteria: (1) Atoms in the loops have low-temperature factors; specifically, loop backbone atoms (N, C, C $\alpha$ , and O) have an average temperature factor of  $\leq 35$ ; (2) loops belong to proteins representing a wide range of protein structural families; (3) both solvent exposed and buried loops, spanning a variety of secondary structural elements (helix to helix, strand to helix, helix to strand) are represented; and (4) protein structures having incomplete residues within 10 Å of the loop to be sampled were excluded. A total of 197 loops (107 protein structures), ranging from 4–12 residues in length, were selected for this work. As shown in Table 1, each set of loops ranging from 4–10 residues had 19–27 loops per length, while the 11- and 12-residue loops, being more restricted by the criteria listed above, had 14 and 10 loops per length.

*Protein preparation*

Hydrogen atoms were added to all proteins structures, and standard protonation states at pH 7.0 were used (Asp, Glu ionized; Arg, Lys protonated). Protein assignment was performed using Maestro (Schrödinger, LLC) to optimize the positions of Ser, Thr, and Tyr hydroxyl protons, and His protonation and tautomeric states. The positions of Asn, Gln, and His side chains were also optimized via 180 degree terminal-chi flips. Waters, metals, and cofactors were removed, while any bound ligands or substrates were retained. Ligand bond orders and formal charges were set to their proper values. Since X-ray structures are optimized to fit electron densities rather than minimum energy conformations, they often exhibit significant deviations from regularized molecular geometries. Such distorted geometries can cause force-field energies to be high. Therefore, each protein structure was subjected to relaxation using a series of restrained, partial

**Table 1.** Characteristics of loop data set<sup>a</sup>

Loop number	Loop Length								
	4	5	6	7	8	9	10	11	12
1	1aaj:82–85	153l:131–135	1ads:149–154	1a62:89–95	1a62:71–78	1aac:58–66	135l:18–27	153l:154–164	153l:98–109
2	1ads:99–102	1a2y:14–18	1ads:150–155	1bkf:64–70	1ads:274–281	1arb:168–176	1ads:170–179	1a2p:76–86	1akz:181–192
3	1cbs:21–24	1a8e:197–201	1brt:174–179	1ads:186–192	1alc:34–41	1arp:127–135	1ads:171–180	1a2y:91–101	1arb:74–85
4	1frd:59–62	1frd:83–87	1brt:253–258	1brt:226–232	1arb:136–143	1aru:36–44	1amp:181–190	1akz:211–221	1bkf:9–20
5	1gpr:123–126	1gpr:54–58	1cbs:66–71	1cvi:111–117	1cvi:148–155	1cse:95–103	1arb:41–50	1awq:1101–1111	1cex:40–51
6	1nfp:37–40	1hbg:19–23	1dim:318–323	1dad:116–122	1gof:606–613	1csh:252–260	1arp:37–46	1cvi:257–267	1dim:213–224
7	1pbe:117–120	1hbq:158–162	1dts:146–151	1dim:198–204	1hbq:31–38	1ede:257–265	1aru:128–137	1dad:42–52	1ixh:160–171
8	1pda:139–142	1kuh:37–41	1ede:180–185	1edg:309–315	1hfc:119–126	1fus:91–99	1btl:170–179	1fus:28–38	1xyz:813–824
9	1plc:74–77	1lit:131–135	1gca:100–105	1gca:196–202	1hfc:142–149	1llk:142–150	1dim:87–96	1ixh:120–130	1luc:158–169
10	1ppn:42–45	1lit:51–55	1mpr:233–238	1hbg:46–52	1nar:192–199	1mla:194–202	1fkf:63–72	1mla:9–19	2ayh:21–32
11	1ref:111–114	1lkk:186–190	1nif:211–216	1hfc:152–158	1nif:221–228	1nls:131–139	1gpr:133–142	1ref:122–132	
12	1thw:194–197	1mla:102–106	1noa:25–30	1iab:142–148	1nif:279–286	1onc:70–78	1gvp:49–58	2eng:124–134	
13	1tib:46–49	1mla:275–279	1onc:12–17	1lif:64–70	1nls:97–104	1pda:108–116	1ixh:84–93	2pth:8–18	
14	1tml:42–45	1nar:56–60	1rge:73–78	1mbd:17–23	1nwp:84–91	1pgs:117–125	1knt:35–44	3pte:91–101	
15	1xif:82–85	1nfp:95–99	1rhs:50–55	1mla:80–86	1oyc:80–87	1php:91–99	1mrj:173–182		
16	2exo:161–164	1noa:88–92	1tca:38–43	1nif:65–71	1prn:150–157	1sgp:109–117	1plc:42–51		
17	2sga:44–47	1prn:187–191	1tca:94–99	1php:135–141	1thw:18–25	1xnb:116–124	1ppn:190–199		
18	2sil:220–223	1rie:149–153	1tys:66–71	1rhs:21–27	1tml:187–194	1xnb:133–141	1scs:65–74		
19	2tgi:72–75	1sbp:181–185	1xyz:633–638	1sgp:128–134	2ayh:194–201	1xyz:795–803	1tca:23–32		
20	4enl:335–338	1tca:157–161	1xyz:711–716	1tca:132–138		2ayh:169–177	1whi:47–56		
21	4gcr:116–119	1tml:147–151	2ayh:81–86	1tml:20–26		2cpl:24–32	2cmd:57–66		
22	7rsa:47–50	1vcc:63–67	2eng:9–14	1xyz:689–695		2eng:172–180	2mnr:91–100		
23		1xyz:559–563	2mnr:308–313	2mnr:270–276		3pte:107–115	2sil:197–206		
24		2cba:168–172	2ran:40–45	2pth:95–101			3hsc:28–37		
25		2cmd:188–192	2sil:176–181	3tgi:159–165			7rsa:110–119		
26		2hbg:37–41	3pte:131–136	5p21:83–89			7rsa:33–42		
27		7rsa:75–79	3pte:256–261				7rsa:87–96		
28			5p21:104–109						
29			8abp:65–70						
Total loops	22	27	29	26	19	23	27	14	10

<sup>a</sup>The specific loops studied are listed by PDB code and separated by loop length. The total number of loops for each length is shown at the bottom. The loop number corresponds to the numbers shown in the graphs in Figure 1A–J.

minimizations using the OPLS-2001 force field (Jorgensen et al. 1996; Kaminski et al. 2001; Jacobson et al. 2002) until the average RMS deviation of the heavy atoms reached a maximum of 0.3 Å with respect to the starting structure. These minimized structures were used as the starting point for loop generation, and also served as the reference templates for evaluation of results. To ensure no bias was incurred using the minimized structure, several loops were calculated using the native, unminimized protein, and no difference in the trends were observed.

### *Prime*

The Loop Refinement module in Prime version 2.5 (Schrödinger, LLC) uses an ab initio loop prediction method. The energy function used by Prime is the OPLS-2001 all-atom protein force field with the Surface Generalized Born implicit solvent model (Ghosh et al. 1998; Gallicchio et al. 2002). Briefly, Prime uses a dihedral-angle-based buildup procedure to extensively sample conformational space, followed by iterative cycles of hierarchical screening and clustering to select representative loop candidates. Loop candidates then undergo side-chain optimization followed by a final complete energy minimization of the side-chain optimized loop structures. The residues of the loop to be sampled were specified and the Prime loop refinement sampling method set to “serial loop sampling” at an “extended high” sampling accuracy. The default setting fixed all core protein backbone atoms, but side-chain atoms within 7.5 Å of the calculated loop were allowed to relax. For loops with cysteines known to participate in disulfide bonds, the sulfurs of the cysteine residues were typed as disulfides. For each calculation, results were returned for the top six loop predictions ranked by energy. Prime also offers the opportunity to carry out loop refinement calculations within a protein’s crystal environment. In this study, the set of proteins for the 11-residue loops was selected to undergo loop refinement in the crystal environment in addition to the calculations described above.

### *Modeler*

Loops were generated in Modeler, version 7.0 (Accelrys Software Inc.) using the Refine Loop utility. Modeler uses conjugate gradients and molecular dynamics with simulated annealing to optimize the positions of all non-hydrogen atoms. The energy of the loop is derived from a combination of molecular mechanics terms from the CHARM-22 force field (MacKerell Jr. et al. 1998) and empirically derived terms based on statistical preferences for the main-chain and side-chain dihedrals and non-bonded atomic contacts derived from known protein structures. The residues of the loop to be sampled were

specified, and the loop optimization level was set to high, using the most thorough annealing protocol available in this program. The six top-ranking loops generated were returned for analysis. If disulfide bonds were present in the loop, those constraints were added to the input file to create the disulfide bonds on the resultant loops.

### *ICM*

The Loop Sampling option was used in ICM, version 3.4–8 (Molsoft LLC) to generate plausible loop conformations for the test set using a knowledge-based approach for loop searching. A nonredundant version of the PDB (proteins with >90% identity were removed from the database) was used to assemble a custom database for loop searches that excluded proteins from the loop test set. The original loop was cleaved out of the protein and the sequence of the desired loop is used for the search. Computational searches of the database select replacement loops by geometrically matching loop ends, proper sequence length, and as close a loop sequence as possible to the query loop sequence. The template loop was then inserted into the protein model and side-chain types were modified to match the desired loop sequence. The resulting loops can have additional problems such as loop attachment points, which may have strained geometry and steric clashes that can occur with the main portion of the protein. The minimize loop feature in ICM was used to remove steric violations and relax geometry at loop junctions. The loop sampling procedure available in the ICM graphical user interface was used for the model building.

### *Sybyl*

Version 7.1 of Sybyl was used in this study. The Protein Loop Search functionality within Sybyl/Biopolymer uses a knowledge-based approach to model loops. The calculations were run by cleaving the loop from the core protein and supplying the sequence for the calculation. This method finds fragments in a database, constructed from the Protein Data Bank, that are the proper residue length whose anchor regions have a good geometric fit to the anchor regions of the modeled protein, and then melds the fragments onto the anchor regions of the protein. Application of this procedure usually generates several candidate loops that satisfy the geometrical requirements. A spreadsheet with interactive graphical tools is available for analyzing and selecting from the retrieved fragments on the basis of quality of fit to the anchor regions and sequence homology. In the event that the database search fails to identify hits, an alternative method called Random Tweak Loop Generation, which relies on a constrained conformational search, is used (Sybyl 2005). The method is based on that of Fine et al. (1986) and Shenkin et al.



(1987). For our purposes, all fragment hits that were derived from proteins showing sequence identity of 90% or higher were removed from the search.

#### *Accuracy of loop prediction*

Six loop models were calculated for every loop using the four modeling programs. The accuracy of each loop prediction was evaluated by comparing it with the refined native protein. Results are reported as “global” RMSDs. The whole protein, excluding the loop, was superimposed on the minimized native structure, and the RMSDs were calculated for the main-chain atoms (N, C, C $\alpha$ , and O) of the modeled loop with respect to the native loop. Two distinct criteria, providing different evaluation metrics, were used to assess the modeled loops; the “top-rank RMSD” and “best RMSD” are defined as follows. The top-rank RMSD is the RMSD of the loop that was ranked first of the six loops calculated from each program. This type of result is useful to assess how well each of the loop-modeling algorithms score the modeled loops, and is discussed in detail in this study. The best RMSD is the RMSD of the modeled loop that most closely matches the native loop irrespective of rank (first through sixth). While this number is only useful (and calculable) when it can be compared to the correct answer, i.e., the native structure, it provides a measure of how well the algorithms sample the loop and whether it is possible to achieve a good-quality result from a given loop-modeling program.

#### **Results**

First we assess the performance of the loop-modeling algorithms for all loop lengths. The loops are then broken down into representative short (five residue), medium (eight residue), and long (11 residue) loop lengths. Specific loops from each of these categories were used to illustrate interesting behavior from a modeling or structural viewpoint, and are discussed in detail. The loops were evaluated by top-rank RMSD and best RMSD as defined above. The best RMSD of the results was used to evaluate the strengths and weaknesses of each program. Additional observations are made that evaluate different features available in the programs.

Overall, the four programs generated reasonable quality loops (<2 Å) for at least a few loops within each set of loop lengths up to 10-residue loops (Fig. 1A–I). Figure 1J gives an overview of how well the programs performed by taking the average of the RMSDs from each loop length for each method, evaluating both best RMSD and top-rank RMSD. Prime produced loops with the lowest RMSD compared to the crystal structure; calculated using both top-ranking as well as best RMSD. When assessing best RMSD, all programs performed well in the shorter

loop lengths (four to six residues), with RMSDs usually below 1 Å, which is an unexpectedly good result. This supports the use of loop-modeling methods for modeling shorter length loops with good accuracy. The medium length loops, seven to nine residues, yielded loops that had best RMSDs >2 Å over 50% of the time, and showed more variation between the programs where Prime began to stand out from the other three programs (Fig. 1J). The accuracy of the calculated loops drops off for the longer length loops to 3–5 Å (10–12 residues) for all methods. The 27 loops in the 10-residue set had similar RMSDs (Fig. 1G) for all methods, with no single program standing out significantly. The two other sets of long loops that were modeled, 11- and 12-residue lengths, however, showed a very wide variation in the quality of results depending upon modeling algorithm (Figs. 1H–I), with Prime producing the best loops for both lengths yielding four of the 10 loops with best RMSDs under 2.5 Å, and Sybyl generating the worst by a large margin producing four loops with best RMSDs over 6 Å.

#### *Short loops: 5 residues, 27 loops*

All four programs generated loops with good accuracy (Fig. 1B) for this short length loop, yielding the majority of results with an RMSD <0.6 Å. The best RMSDs for the five-residue loops from Prime most closely matched the native structure for 23 of 27 loops; nine of these loops also scored as the top-ranking loops by Prime, yielding the best results compared to other methods for this loop length. The results from Modeler were second best, with 21 of 27 the best RMSDs under 1 Å. However, the top-ranked loops from Modeler had significantly higher RMSDs than the loops with the best RMSDs. This shows that the scoring function used by Modeler is less accurate than the one used by Prime to rank loop conformations. Modeler is capable of generating loops that are very close to the native structure, but is unable to select the loop that most closely resembles the native structure as top ranked. Despite the fact that ICM and Sybyl both use database search methods, ICM seems to outperform Sybyl in the majority of the loops sampled. This may be due in part to the database from which the loops are selected, or it could be due to the criteria used to select loops from database. Sybyl had the poorest results with significantly worse top-rank RMSDs, six loops over 3 Å and three of those >4 Å.

#### *Effect of salt bridges*

##### *1vcc, residues 63–67*

Figure 2A shows the top-ranking and best RMSD results from all four loop-modeling algorithms for the five-residue loop in 1vcc. Although Prime produced the

most accurate results for five-residue loops, loop 1vcc proved to be the greatest challenge for Prime. Close inspection of this loop revealed that Sybyl and ICM generated loops similar to the native structure with best RMSDs near 0.5 Å. Modeler has the general shape of the loop correct, but is shifted from the native by 1.5 Å. Prime gave a best RMSD of 2.6 Å, the worst result for this program in this loop length. The conformation of the Prime loop is folded against the protein. For Prime this is a typical result, because loop conformations are rejected as they are being constructed if the backbone of the loop

is too far from the protein core. There are many cases in this study where this procedure works in favor of Prime, because many loops prefer to form contacts with the rest of the protein rather than being solvent exposed. For 1vcc the Prime loop conformation is also driven by a salt bridge formed between Lys 65 in the loop and protein core residue Asp 42. Generally, including electrostatics yields good quality results. In this case, however, the combination of the salt bridge and the packing of the loop against the core of the protein caused Prime to produce loop models that had the highest best RMSD.

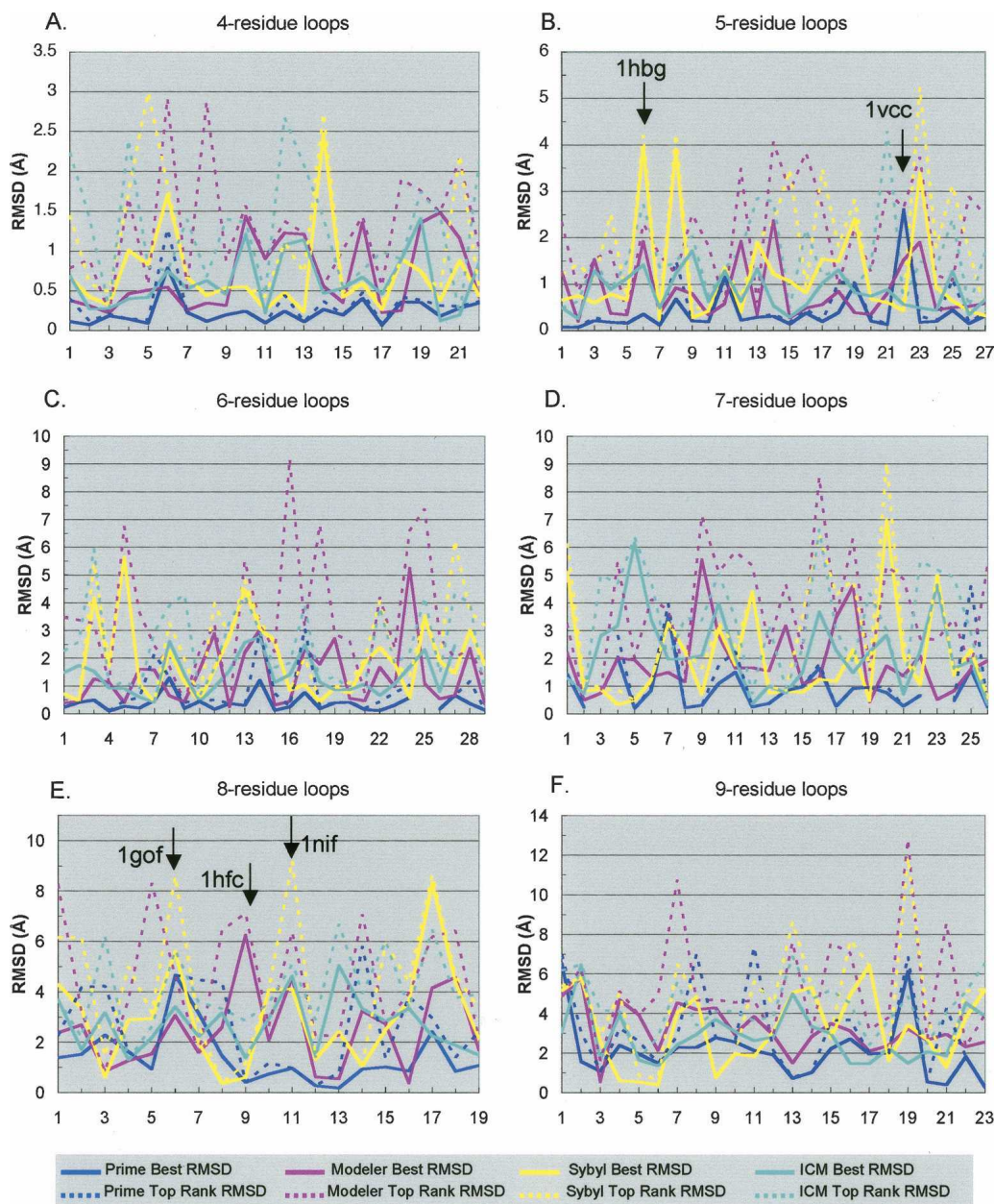


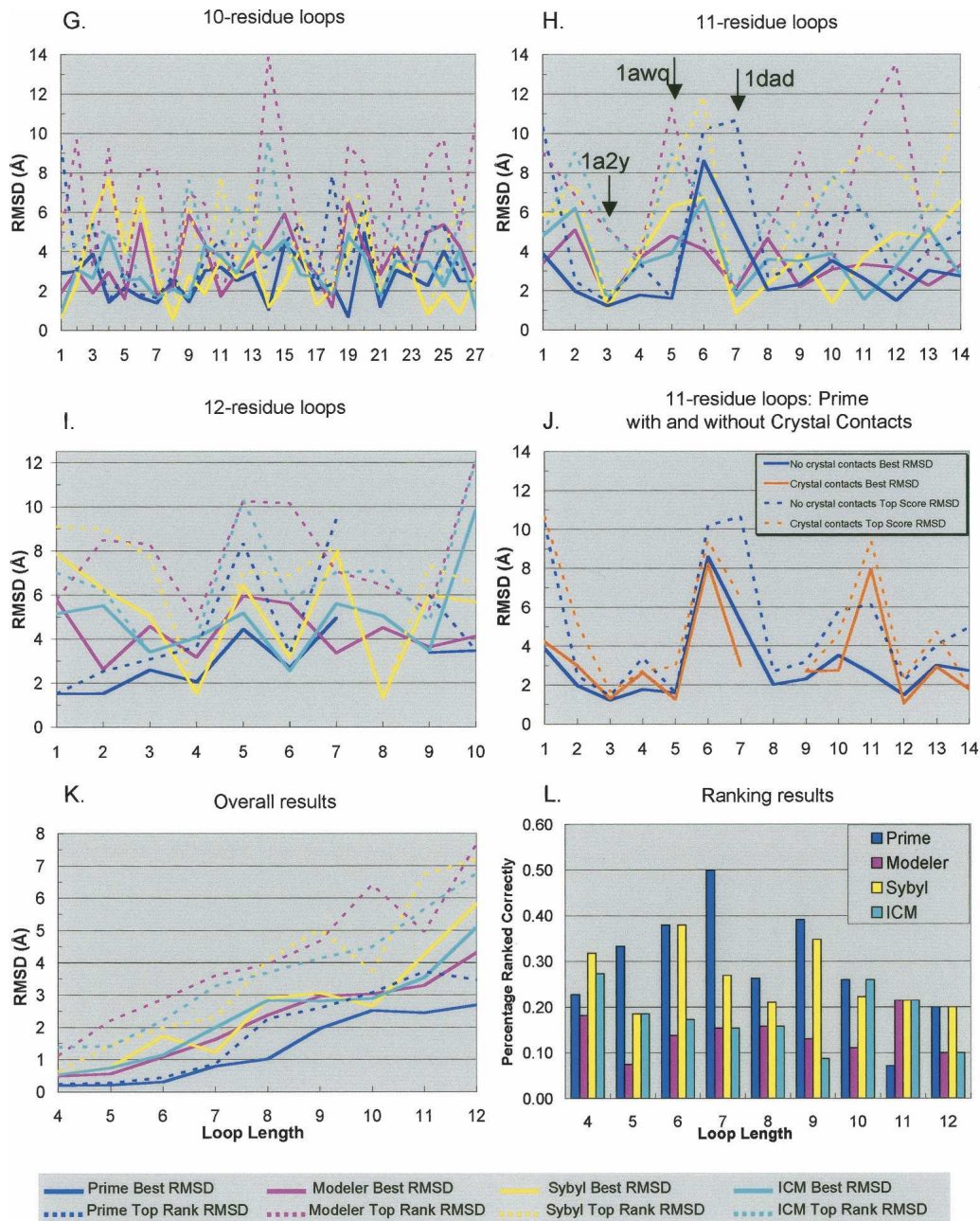
Figure 1. (Continued on next page)



*1hbg, residues 19–23*

Prime generates better loops for *1hbg* than the other three programs by a significant margin, with a top-rank RMSD of 0.3 Å, Figure 2B. In this case, correct formation of a salt bridge between loop residue Asp 21 and protein core Arg 117 proves to be advantageous for Prime. ICM

gives the second most accurate result, having a best RMSD of 1.4 Å and maintaining the overall general shape of the loop with a small shifting of the backbone. Modeler inverts the shape of the loop, yielding a best RMSD of 1.9 Å. Sybyl produces the worst result in this group with a best RMSD of 4.0 Å. It is apparent that the



**Figure 1.** (A–J) Graphs showing the best and top-ranked RMSDs from each program for the 4–12 residue loops. The X-axis is the loop number (which corresponds to the loops shown in Table 1) and the Y-axis is the RMSD in Angstroms. The colors are described in the key at the bottom of each figure. (J) The 11-residue loops generated by Prime with (orange) and without (dark blue) crystal contacts. (K) The average RMSD of both the best loop and top-ranked loop over each loop length for each program. (L) The percentage of loops scored correctly for each loop length: Prime (dark blue), Modeler (magenta), Sybyl (yellow), and ICM (cyan).

lack of a steric component in the Sybyl scoring function greatly affects the quality of this result, allowing side chains to come within van der Waals contact of one another. In this example several instances of atomic distances  $<1 \text{ \AA}$  are observed.

*Medium loops: 8 residues, 19 loops*

As expected, the top-ranking solutions for eight residue loops (Fig. 1E) were less similar to the native structure compared to top-ranking solutions for five residue loops (Fig. 1B), which rarely had best RMSDs  $>2.5 \text{ \AA}$ . The results for the eight-residue loops show many best

RMSDs above  $4 \text{ \AA}$  and the top-ranking loops produced by the modeling programs produced many loops with RMSDs over  $6 \text{ \AA}$ . Overall, Prime performed best for the eight-residue loops, but not as well as it did for the five-residue loops. Of the 19 loops sampled, Prime produced 12 solutions that had best RMSDs  $<1.5 \text{ \AA}$ , which, like the five-residue loops, are accurate results that would be acceptable for use in structure-based drug design. These 12 loops most closely matched the native structure compared to the other methods; however, only four of these were also ranked to be best. Modeler, again, performed better than ICM and Sybyl, giving four solutions most like the native out of these three programs,

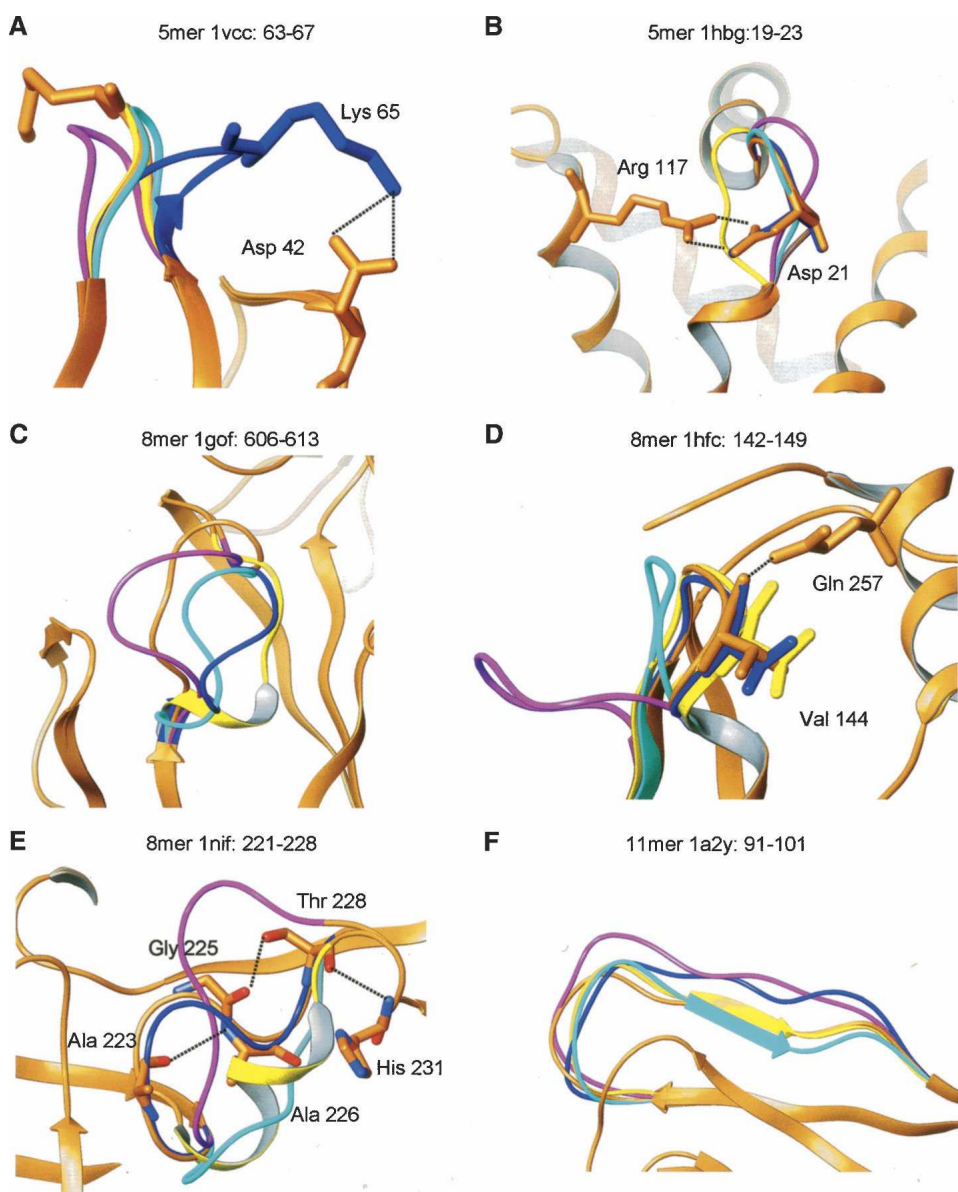


Figure 2. (Continued on next page)



with a best RMSDs  $<3.0$  Å, one of which was also top ranked. ICM had only one loop, 1thw, with the overall best result, best RMSD = 2.2 Å. This loop was predicted nearly as well by Prime, best RMSD = 2.4 Å. Sybyl performed similar to ICM, predicting one loop, best RMSD = 0.6 Å, most accurately, but gave the greatest number of solutions with high, yielding six loops with a high best RMSD of  $>4.0$  Å.

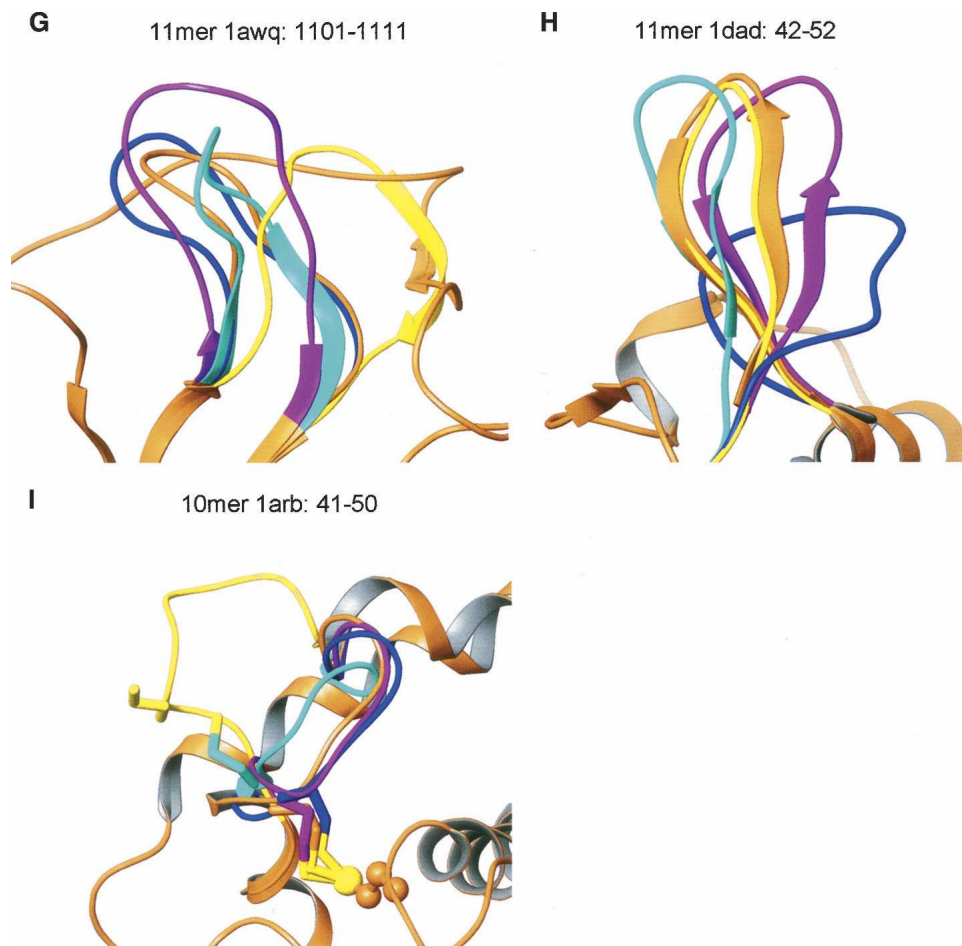
*Igof, residues 606–613*

None of the programs were able to accurately predict the conformation of Igof within  $<3.0$  Å of the native structure (Fig. 2C). Modeler most closely matched the native with a best RMSD of 3.1 Å, followed closely by ICM with a best RMSD of 3.4 Å. Both of these programs generated loops that had turn-like moieties at the end of the loop, somewhat mimicking the crystallographic loop that has a turn in that region. The best results from Sybyl

and Prime were significantly different from the native, with RMSDs of 5.6 Å and 4.7 Å, respectively. Prime packed the loop against the protein, but twisted the loop to the wrong side. Sybyl placed a helical turn at the beginning of the loop instead of the end and had steric clashes with the protein core.

*Ihfc, residues 142–149*

All programs, with the exception of Modeler, did well at modeling the loop conformation of Ihfc (Fig. 2D). The best solution given by Prime, also the top-ranked loop, was within 0.4 Å of the native structure. Prime correctly predicted a hydrogen bond between the backbone carbonyl of Val 144 in the loop and the side chain of Gln 257 in the native protein. The Prime modeled loop correctly packed against the core of the protein, decreasing the solvent accessible surface area and burying of the lipophilic residues of the loop. For this loop, Sybyl



**Figure 2.** Loop conformations that have the lowest RMSDs with respect to the native structure for the proteins and loop lengths as shown. The color scheme is as follows: native protein (brown), Prime (dark blue), Modeler (magenta), Sybyl (yellow), and ICM (cyan). (E) Shows some side chains of the native protein colored by atom, with carbons colored brown, nitrogens blue, and oxygens red. (I) Shows the cysteines of the loop that form a disulfide bridge in the native structure and the cysteine of the protein core rendered in ball and stick, and sulfurs colored yellow.

returned a best RMSD within 0.6 Å of the native, and even positioned the side chains quite well, including placing Val 144 within hydrogen bonding distance of Gln 257; however, the vector is not optimal. The best result from ICM had an RMSD of 1.4 Å. The loop from ICM was reasonably well positioned with respect to the native loop, although a steric clash was observed between the C terminus of the protein and Val 144 in the loop. This appears to be an artifact of the grid-based algorithm for steric contacts used in the ICM loop search protocol, and subsequent minimization should remove the steric interaction without perturbing the loop conformation. Modeler performed very poorly for this loop, giving a best RMSD of 6.3 Å. The loop predicted by Modeler is solvent exposed, forming almost no contacts with the native protein. In the native structure, the terminal residues of this loop form part of a beta-strand. Modeler failed to reproduce the hydrogen-bonding network required for these residues to participate in a beta-strand, allowing the loop to adopt a completely different conformation from that of the native protein.

#### *Inif, residues 221–228*

Loop Inif (221–228) is a very unusual loop that is a distorted  $3_{10}$  helix (Fig. 2E). Hydrogen bonds are formed between backbone atoms of residues Ala 223 and Ala 226 and Thr 228 and His 231 and backbone–side-chain interaction between Gly225 and Thr 228. Prime produces a loop with 0.9 Å for the best RMSD that mimics the backbone interactions quite well, but fails to replicate the Gly 225/Thr 228 interaction and rotates the threonine side chain away by  $\sim 120^\circ$ . The loop conformation emphasizes the value of incorporating a term in the loop scoring function that favors burying hydrophobic groups. The best loop generated by Modeler has a best RMSD of 4.6 Å, and fully extended the loop instead of forming the helical-like structure. The first residue fails to make the Ala 223/Ala 226 hydrogen bond, as seen in the native structure, and makes two backbone/side-chain hydrogen bonds to residues in the protein core that are not seen in the native structure; additionally, no intra-loop hydrogen bonds are made that are seen in the native structure. The best solution from Sybyl has a poor best RMSD of 4.1 Å; however, the generated loop is purely alpha-helical in nature. It is interesting that an alpha-helical structure was found for a loop that in the native protein is a distorted  $3_{10}$  helix, but unfortunately, the lack of a steric term (used in the Sybyl loop scoring function) greatly reduces the quality of this loop. There are several instances of close steric contacts ( $<1.0$  Å) and even one example of a loop backbone coming into direct contact with a  $C_\alpha$  atom of a neighboring residue. ICM generates a loop with a best RMSD of 4.6 Å that is similar to Sybyl with respect to residues 226–228, but significantly different for residues

221–225. Again, steric clashes occur between many of the loop residues and the protein core. It is clear that for this example, with steric clashes, a minimization protocol and rescoring would likely improve the quality of the ranking.

#### *Long loops: 11 residues, 14 loops*

As to be expected, the longer loop lengths have significantly higher RMSDs, averaging 5.3 Å for the top-ranking loops and 3.4 Å for the loops with the best RMSD (Fig. 1K). Many solutions have top-ranking RMSDs above 8 Å, and some that are as high as 12 Å or more depending upon the specific loop and loop-modeling algorithm (Fig. 1H). Obviously, due to the high degree of conformational flexibility in these loops, these results are not surprising. It is interesting to note that Prime produces the loop with the closest RMSD to the native structure in seven of the 14 loops, a lower percentage of best RMSDs compared to the short and medium length loops. This is the worst showing for Prime; however, Prime still produced more loops closest to the native loop than the other three methods, with six loops having an RMSD of  $<2$  Å. Sybyl has three cases where loops were the best generated of all four programs. These three loops were further investigated to determine what proteins the loop segments were derived from. It was found that even though the proteins contained  $<50\%$  sequence identity to the native protein, they likely belonged to the same structural family of proteins. ICM also had two loops that scored closest to the native PDB in comparison to the other programs. This highlights the value of using a database searching method, which calculates results very quickly even for the longer loops. When crystal structures from homologous proteins in the same family are available, a good-quality homology model may be built quickly, especially for longer length loops. All programs had at least some examples where each produced the best overall RMSD for a given loop; however, the RMSDs are so large for most of the best solutions (with the exception of Prime), that these models would not generally be acceptable for structure-based design.

#### *1a2y, residues 91–101*

All of the loop-modeling protocols perform quite well on this loop, despite the long length, with RMSDs  $<1.6$  Å (Fig. 2F). The results for this loop suggest that longer length loops may be predicted with reasonable accuracy when the loop is in a fairly confined region of the protein. Sybyl produced a loop with the lowest best RMSD, 1.1 Å, which validates the use of database search methods when a common protein fold is being examined, in this case, a beta sandwich. Prime compares favorably with Sybyl, generating a loop with a best RMSD of 1.2 Å that places the side chains similarly to the native crystal structure. ICM generated a loop with a best RMSD of 1.6 Å, but had

significant steric contact with neighboring side chains. Modeler also has a good best RMSD for this loop length, 1.5 Å, but upon closer inspection, the backbone is positioned close to the backbone of the native crystal, but rotated 180°. Thus, all of the residues are pointing in the opposite direction from the native crystal.

#### *lawq, residues 1101–1111*

Prime generates the loop with the lowest best RMSD for lawq (1.6 Å), which is significantly better than the loops from ICM (3.8 Å), Modeler (4.8 Å), and Sybyl (6.3 Å) (Fig. 2G). Again, a long loop length is sampled in this case, but Prime generated a loop conformation that had a low RMSD compared to the native crystal structure. This partially buried, partially exposed loop was modeled well by Prime. The resulting model even contained a series of hydrogen bonds constraining the loop in a conformation in a manner similar to the native loop. The loop from ICM makes hydrogen bonds to a different neighboring loop, and this draws the loop into the wrong conformation. Modeler fails to produce loops that interact with neighboring loops; rather, it extends this loop toward solvent producing an inaccurate model. Sybyl generates a loop that not only shows a high RMSD for the best loop modeled, but the modeled loop weaves in and out of the protein, making significant steric clashes in numerous places. This is yet another example of the value of including a steric term in the loop-modeling scoring function, which would reject this type of loop conformation.

#### *ldad, residues 42–52*

For the loop in ldad, Modeler, Sybyl, and ICM all generate loops with low best RMSDs (Fig. 2H). In this example, Sybyl outperforms the other three loop-modeling algorithms with a best RMSD of 0.8 Å, reproducing the hairpin turn of the native loop. Modeler (2.1 Å) and ICM (1.7 Å) both generate reasonable loops similar in shape to the native loop. The best RMSD for Prime (5.3 Å) is one of the higher RMSDs of all Prime loops. In this case, the modeled loop folds against the protein core, instead of extending toward the solvent, as it is in the native loop. Further investigation of the loop revealed that the crystallographic conformation is induced by contact from a neighboring protein in the unit cell; therefore, modeling of a single protein may produce incorrect loop conformations.

#### *Disulfide bonds: Iarp, residues 41–50 (10 mer)*

Modeler and Prime have the ability to create disulfide linkages between cysteines when generating loops. As one would expect, the ability to utilize disulfide constraints produced much better loops, as evidenced by the 10-residue loop Iarp (Fig. 1I). Both Modeler and Prime

generated loops with a best RMSD <2.0 Å, and are clearly anchored by the disulfide linkages between the modeled loop and the protein core. Without identifying the disulfide linkages, Modeler and Prime generate significantly worse results; the best RMSD for Prime was 4.2 Å, while the Modeler best was RMSD 5.4 Å. ICM does fairly well at reproducing the loop conformation with an RMSD of 2.5 Å, but the cysteine faces away from the partner disulfide. Sybyl produces an extremely poor best RMSD (6.5 Å) loop which occupies a significantly different region of the protein. In both Modeler and Prime, prior to calculating the loops, the disulfides need to be specified and the atom typing fixed so this constraint can be utilized by the loop-sampling algorithms. Sybyl and ICM both have advanced options, not part of the standard protocol, that allow the user to define constraints between cysteines. However, the constraints are utilized in a post-processing minimization step that was not considered for purposes of this study.

#### *Crystal contacts*

Of the four programs tested, Prime is the only one that is capable of including crystal contact information during loop generation. The complete set of 11-residue loops were retested (Fig. 1J) using the crystal packing option in Prime. All 14 loops had at least one atom that was within 10 Å of a neighboring protein. One of the loops (I fus, loop 28–38) did not yield any solutions. Unfortunately, efforts to determine the source of this error were unsuccessful. Seven of the remaining 13 loops sampled improved using crystal contacts, with ldad showing the most significant improvement, decreasing the original best RMSD from 5.3 Å to 3.0 Å. The final six loops had higher RMSDs when utilizing the crystal packing option, but were <1 Å for all but one of the loops. Even with the additional information regarding crystal packing, this loop generation fails to deliver a loop within the 2 Å RMSD cutoff identifying good-quality loops. Four of the seven loops that had smaller RMSDs using crystal contacts were considerably small improvements of  $\leq 0.4$  Å. Since only half of the loops sampled had improved RMSDs, and not particularly large improvements, this suggests that crystal packing is not necessary or even beneficial in most cases. The true value of crystal packing in Prime is that it allows the solvent-exposed loops to extend away from the protein core, to interact with a neighboring protein. Several loops that appear to be solvent exposed when viewing a single asymmetric unit, are in fact, interacting with a symmetry related protein. Since the Prime loop-modeling algorithm was designed to increase steric contact between the loop and protein core, the addition of the crystal unit cell information allows Prime to generate loops that would not be produced without the neighboring atoms of the symmetry related protein being present.



*No solutions found: 2sil: 176–181 (6 mer), 1ads: 186–192 (7 mer), 2 mnr: 270–276 (7 mer), 1xyz: 813–824 (12 mer)*

There were four examples of loops for which Prime was unable to determine a solution. When investigating these cases, it was noted that all four loops were buried deep within the protein. Because Prime builds the loop residues simultaneously from each end, the program was unable to find a path that was free of steric clashes. This may be a result of poor sampling, but could also be due to steric clashes between the generated loop and the core protein where a force field-based program would yield very high energies for such poses.

### Ranking loops

As frequently noted, the resultant loops were ranked very poorly with respect to the rank ordering of the RMSDs with respect to the native crystal. Figure 1K represents the average best RMSD from each loop length as well as the average top-ranking RMSD. This allows for an assessment of both the average quality of the best RMSD results as well as the average quality of the scoring function. For short loops (four to six residues), all programs generated loops with best RMSDs averaging <1 Å, with the exception of the Sybyl six-residue loops, confirming that all four algorithms yield satisfactory results for use in modeling studies. Although Modeler, Sybyl, and ICM had low best RMSDs, the methods did not rank the results well, and show a 1–1.5 Å difference between the top-ranked and best RMSD loops, while the results from Prime were much better, with only 0.1 Å difference between the two RMSDs. For medium-length loops (seven to nine residues) Prime stands out as the best method, with RMSDs averaging <2 Å while the other three methods approach RMSDs of 3 Å. At these lengths, the best RMSDs deviate from the top-ranked RMSDs for all methods, but Prime has a smaller difference of ~1 Å vs. ~1.5 Å for the others. Even for long loops (10–12 residues) Prime was able to produce loops with RMSDs averaging ~2.5 Å. All methods had similar results for the 10-residue loops, but the 11- and 12-residue loops break down for Modeler, Sybyl, and ICM with RMSDs ranging from 3–6 Å. The deviation between top-ranked vs. best RMSDs remains consistent with the medium-length loops, showing that the ranking does not correlate with loop length. The high RMSDs observed in the longer loops would not be acceptable for use in homology modeling.

Another way to assess the results is by comparing the number of times each program selected the loop with the best RMSD as the top-ranking loop (Fig. 1L). By this criterion, it is evident that Prime has the best method for ranking loops, since it scored 30% of the loops properly compared to 14% for Modeler, 25% for Sybyl, and 18%

for ICM. Prime performed equal to or better than all methods for all loop lengths except the four-residue loops and 11-residue loops. This representation does overlook the fact that Prime often clusters the loop conformations together closely; therefore, the best RMSD and top-ranking RMSD may only be 0.1 Å apart. Sybyl performs slightly worse than Prime when rating the resulting loops, with an average of 25% scored correctly. The scores in Sybyl would be greatly improved if appropriate steric components were added for selection of loop models, since the scoring is currently assessed simply by RMSD of the terminal ends of the loop compared to the linkage trajectory to the protein core. It is striking to note that despite the lack of an energy-based scoring function, Sybyl ranked the loops almost as well as Prime for all loop lengths except the five- and seven-residue loops. The performance for both Modeler and ICM is less encouraging, universally, averaging 14% and 18% for loops correctly ranked for each program, respectively. Modeler samples conformational space relatively well, and generates structures that are closer to the crystallographic conformation than ICM, but does not rank the generated loops well. For loop lengths up to 10 residues, the methods deliver results that have RMSDs generally <2–2.5 Å, which would be acceptable levels of resolution for modeling studies. Unfortunately all methods are unable to rank the resultant loops with good accuracy; it is clear that a user must visually inspect all loop modeling results and not rely on the ability of a modeling algorithm.

### Discussion

Since loops near the binding site can influence the manner of ligand binding, accurate loop modeling can often be an important prerequisite to structure-based design. Various examples have already been cited in the introduction. Therefore, the results of this study are intended to elucidate the strengths and weaknesses of four loop-modeling algorithms.

Overall, all four algorithms generated loops that were within 2 Å of the crystallographic conformation for at least one loop at each loop length except the 12-residue loops, where Prime and Sybyl had a few examples of accurate loops. Prime, however, performed most consistently and produced loops with the lowest RMSD for the majority of the loops modeled at all lengths. Prime is the only program that uses force field-based energy terms that includes electrostatics and sterics to score the loops with a continuum solvent model, and therefore, appears to be a superior method of selecting and ranking loops. However, in the best case (seven-residue loops) Prime is only capable of ranking the best RMSD loop as the top loop 50% of the time, and averaging only 30% correct over all loop lengths. As stated before, this representation

does overlook the fact that Prime often clusters the loop conformations together closely; therefore, the best RMSD and top-ranking RMSD may only be 0.1 Å apart. The benefit of incorporating electrostatics into the calculation is evident, seen in several examples where salt bridges and hydrogen bonds were driving the conformation of the loop, resulting in loops with very low RMSDs compared to the loops generated by the other programs. Additionally, the hydrophobic term that Prime uses to model loops yields loops that tend to be packed against the protein. This packing is often observed in crystal structures, and frequently seen in this study. When loops are solvent exposed, Prime has an option to use crystal packing to reproduce the extended conformation of the loop. The crystal packing option, however, shows mixed results in this study, improving some of the loops, but not others, and no clear circumstances as to when it should be utilized.

When examining the best RMSDs, Modeller, Sybyl, and ICM reported similar results, but the rank ordering of the loops varied greatly between the programs. The loops with the best RMSDs from Modeller were generally closer to the native loop than those from Sybyl or ICM, and it is clear that conformational space is better sampled by Modeller than the other two methods. The random positioning of the starting loop followed by minimization, molecular dynamics, and simulated annealing result in a wide range of conformations sampled by Modeller, yielding loops with relatively good best RMSDs compared to Sybyl or ICM. However, in loops longer than seven residues, even the best RMSDs are over 2 Å. The energy function used to score and rank the loop models is the main weakness of the Modeller program, evident by the large disparity between the best and top-ranking RMSD. The use of statistical preferences for nonbonded interactions and dihedrals does not score the resultant loops as well as using electrostatic interactions, which Prime successfully uses.

Sybyl and ICM are both database search-based methods and report similar results, with ICM performing slightly better than Sybyl, possibly because the database of protein templates used by ICM was more current and larger than that used by Sybyl. The results for Sybyl are erratic, generating very good results for some loop lengths and very poor results for others. Database search methods perform well on common loop conformations, such as beta hairpin turns, and type I turns are useful for quickly generating models of loops, including long length loops. If proteins with a reasonable level of sequence identity and common three-dimensional fold family members are available, good quality loops can be generated, as was observed in the 12-residue loops (Fig. 11). Sybyl, surprisingly, ranked the resultant loops second best of the four programs, but only ranking 25% of the loops

correctly. Both Sybyl and ICM do not have scoring functions based on energy; rather, they rank the loops based on the RMSD of the spliced loop to the protein core. The absence of a bump check results in a large number of loops that have severe steric clashes with both the protein core side chains as well as backbone atoms. Incorporating a simple steric term into the loop modeling and ranking would greatly improve the quality of the results from both Sybyl and ICM.

Modeller, Sybyl, and ICM could all benefit from a post-processing set that involves an energy calculation to determine the rank ordering of the resultant loops, allowing high energy loops to be rejected. Although Prime scored the results the best of the four programs, the top-ranking loops were only predicted correctly 30% of the time. This underscores the need to improve the scoring functions used to rank loop models for all loop-modeling programs in order to assist the user in selecting the best loop to use in a homology model. Currently, the user cannot rely solely on the energy or rank order of the loop from any one method. As a result of this study, the user is encouraged to visually inspect the loops to determine which is most reasonable. The following observations were made regarding what constitutes a “reasonable” loop conformation; most importantly, one must inspect for steric clashes of both the backbone and side-chain atoms. Sybyl and ICM generated numerous structures with unacceptable steric clashes which need to be eliminated through visual inspection. Second, identify possible salt bridges and hydrogen bonds that may rigidify or drive the conformation of the loop. Compactness of the loop to the protein core was often observed in the native structure. The algorithm used in Prime includes a term that takes this into account and, as a result, yields loops more similar to the native crystal than the other methods. The ability to form disulfides is a key component to loop modeling. When opportunity arises to form disulfides, the conformational flexibility of the loop is significantly reduced. When a disulfide is not specified when one exists, the resultant loop will adopt a conformation that keeps the cysteines far from one another. Finally, it is prudent to examine the characteristics of the residues in the loop. Hydrophobic residues generally pack against the protein while hydrophilic residues tend to be solvent exposed. Looking for these types of criteria can help identify the best loop from the set generated by the different methods.

Results from this study point out the benefits of *ab initio* and database methods for loop modeling. Given the ever-increasing computational speed available and the increase in conformational sampling resulting from this, it is fair to expect that protein models, and the loops generated by computational methods, should be held to rigorous standards. This is especially true if they are to be

used for structure-based drug design where multiple hypotheses are derived from inspection of the protein models. These studies provide new ways to impact drug design. For example, given the expected increase in loop flexibility with increasing loop length, an approach to structure-based ligand docking that considers an ensemble of conformations rather than a single conformational solution for a given loop is more reasonable. Independent studies of loops in G protein-coupled receptors (Kortagere et al. 2006; Mehler et al. 2006), as well as a study showing the relative benefits of ligand docking to an ensemble of protein structures rather than a single structure (Polgar and Keserue 2006) also appear to support this finding. Furthermore, using an ensemble of loop conformations rather than a single conformation reduces uncertainties that can arise from inadequacies of the scoring function. Finally, the loop-modeling data presented herein suggest that the incorporation of implicit solvent models results in higher accuracy of ranking and improved selection of loop structures.

## Acknowledgment

We thank Ramy Farid for supplying the unfiltered data set by Jacobson et al. (2004) used in this study.

## References

- Cohen, B.I., Presnell, S.R., and Cohen, F.E. 1993. Origins of structural diversity within sequentially identical hexapeptides. *Protein Sci.* **2**: 2134–2145.
- Esposito, E.X., Tobi, D., and Madura, J.D. 2006. Comparative protein modeling. In *Reviews in Computational Chemistry* (eds. K.B. Lipkowitz et al.) Vol. 22, pp. 57–167. Wiley-VCH, Weinheim, Germany.
- Fetrow, J.S., Godzik, A., and Skolnick, J.S. 1998. Functional analysis of the *Escherichia coli* genome using sequence–structure–function paradigm: Identification of proteins exhibiting the glutaredoxin/thioredoxin disulfide oxidoreductase activity. *J. Mol. Biol.* **282**: 703–711.
- Fine, R.M., Wang, H., Shenkin, P.S., Yarmush, D.L., and Levinthal, C. 1986. Predicting antibody hypervariable loop conformations II: Minimization and molecular dynamics studies of MCPC603 from many randomly generated loop conformations. *Proteins* **1**: 342–362.
- Fiser, A., Do, R.K., and Sali, A. 2000. Modeling of loops in protein structures. *Protein Sci.* **9**: 1753–1773.
- Gallicchio, E., Zhang, L.Y., and Levy, R.M. 2002. The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators. *J. Comput. Chem.* **23**: 517–529.
- Ghosh, A., Rapp, C.S., and Friesner, R.A. 1998. Generalized born model based on surface integral formulation. *J. Phys. Chem. B* **102**: 10983–10990.
- Hillisch, A. and Hilgenfeld, R. 2003. The role of protein 3D structures in the drug discovery process. In *Modern methods of drug discovery*. (eds. A. Hillisch and R. Hilgenfeld), pp. 157–218. Birkhäuser Verlag, Basel, Switzerland.
- Hillisch, A., Pineda, L.F., and Hilgenfeld, R. 2004. Utility of homology models in the drug discovery process. *Drug Discov. Today* **9**: 659–669.
- Hobrath, J.V. and Wang, S. 2006. Computational elucidation of the structural basis of ligand binding to the Dopamine 3 receptor through docking and homology modeling. *J. Med. Chem.* **49**: 4470–4476.
- Jacobson, M.P. and Sali, A. 2004. Comparative protein structure modeling and its applications to drug discovery. *Annu. Rep. Med. Chem.* **39**: 259–276.
- Jacobson, M.P., Kaminski, G.A., Friesner, R.A., and Rapp, C.S. 2002. Force field validation using protein side chain prediction. *J. Phys. Chem. B* **106**: 11673–11680.
- Jacobson, M.P., Pincus, D.L., Rapp, C.S., Day, T.J.F., Honig, B., Shaw, D.E., and Friesner, R.A. 2004. A hierarchical approach to all-atom protein loop prediction. *Proteins* **55**: 351–367.
- Jones, S. and Thornton, J.M. 1997. Prediction of protein–protein interaction sites using patch analysis. *J. Mol. Biol.* **272**: 133–143.
- Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**: 11225–11236.
- Kairys, V., Fernandes, M.X., and Gilson, M.K. 2006. Screening drug-like compounds by docking to homology models: A systematic study. *J. Chem. Inf. Model.* **46**: 365–379.
- Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., and Jorgensen, W.L. 2001. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105**: 6474–6487.
- Kasuya, A. and Thornton, J.M. 1999. Three-dimensional structure analysis of PROSITE patterns. *J. Mol. Biol.* **286**: 1673–1691.
- Kick, E.K., Roe, D.C., Skillman, A.G., Liu, G., Ewing, T.J., Sun, Y., Kuntz, I.D., and Ellman, J.A. 1997. Structure-based design and combinatorial chemistry yield low nano-molar inhibitors of cathepsin D. *Chem. Biol.* **4**: 297–307.
- Kleywegt, G.J. 1999. Recognition of spatial motifs in protein structures. *J. Mol. Biol.* **285**: 1887–1897.
- Kornev, A.P., Haste, N.M., Taylor, S.S., and Ten Eyck, L.F. 2006. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc. Natl. Acad. Sci.* **103**: 17783–17788.
- Kortagere, S., Roy, A., and Mehler, E.L. 2006. Ab initio computational modeling of long loops in G-protein coupled receptors. *J. Comput. Aided Mol. Des.* **20**: 427–436.
- MacKerell Jr., A.D., Bashford, D., Bellott, M., Dunbrack Jr., R., Evanseck, J., Field, M., Fischer, S., Gao, J., Guo, H., Ha, S., et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**: 3586–3616.
- Mehler, E.L., Hassan, S.A., Kortagere, S., and Weinstein, H. 2006. Ab initio computational modeling of loops in G-protein coupled receptors: Lessons from the crystal structure of rhodopsin. *Proteins* **64**: 673–690.
- Mezei, M. 1998. Chameleon sequences in the PDB. *Protein Eng.* **11**: 411–414.
- Nayeem, A., Sitkoff, D., and Krystek Jr., S.R. 2006. A comparative study of available software for high accuracy homology modeling: From sequence alignments to structural models. *Protein Sci.* **15**: 808–824.
- Polgar, T. and Keserue, G.M. 2006. Ensemble docking into flexible active sites. Critical evaluation of FlexE against JNK-3 and  $\beta$ -Secretase. *J. Chem. Inf. Model.* **46**: 1795–1805.
- Rohl, C.A., Strauss, C.E.M., Chivian, D., and Baker, D. 2004. Modeling structurally variable regions in homologous proteins with Rosetta. *Proteins* **55**: 656–677.
- Rummey, C. and Metz, G. 2007. Homology models of dipeptidyl peptidases 8 and 9 with a focus on loop predictions near the active site. *Proteins* **66**: 160–171.
- Russell, R.B., Sasieni, P.D., and Sternberg, M.J.E. 1998. Supersites within superfolds. Binding site similarity in the absence of homology. *J. Mol. Biol.* **282**: 903–918.
- Sander, C. and Schneider, R. 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9**: 56–68.
- Shenkin, P.S., Yarmush, D.L., Fine, R.M., Want, H., and Levinthal, C. 1987. Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for ring like structures. *Biopolymers* **26**: 2053–2085.
- Singh, N., Cheve, G., Avery, M.A., and McCurdy, C.R. 2006. Comparative protein modeling of 1-deoxy-D-xylulose-5-phosphate reductoisomerase enzyme from plasmodium falciparum: A potential target for antimalarial drug discovery. *J. Chem. Inf. Model.* **46**: 1360–1370.
- Sybyl. 2005. Tripos, Inc. St. Louis, MO.
- Wallner, B. and Elofsson, A. 2005. All are not equal: A benchmark of different homology modeling programs. *Protein Sci.* **14**: 1315–1327.
- Wei, L., Huang, E.S., and Altman, R.B. 1999. Are protein structures good enough to preserve protein sites? *Structure* **7**: 643–650.
- Wiemann, H., Tondel, K., Anderssen, E., and Drablos, F. 2004. Homology-based modelling of targets for rational drug design. *Mini Rev. Med. Chem.* **4**: 793–804.
- Xiang, Z.X., Soto, C.S., and Honig, B. 2002. Evaluating conformational free energies; The colony energy and its application to the problem of loop prediction. *Proc. Natl. Acad. Sci.* **99**: 7432–7437.