J Comput Aided Mol Des (2009) 23:411–418
DOI 10.1007/s10822-009-9274-3

Closing the side-chain gap in protein loop modeling

Karen A. Rossi · Akbar Nayeem · Carolyn A. Weigelt · Stanley R. Krystek Jr.

Received: 7 November 2008 / Accepted: 18 April 2009 / Published online: 21 May 2009

Abstract The success of structure-based drug design relies on accurate protein modeling where one of the key issues is the modeling and refinement of loops. This study takes a critical look at modeled loops, determining the effect of re-sampling side-chains after the loop conformation has been generated. The results are evaluated in terms of backbone and side-chain conformations with respect to the native loop. While models can contain loops with high quality backbone conformations, the side-chain orientations could be poor, and therefore unsuitable for ligand docking and structure-based design. In this study, we report on the ability to model loop side-chains accurately using a variety of commercially available algorithms that include rotamer libraries, systematic torsion scans and knowledge-based methods.

Keywords Homology modeling · Side-chain sampling · Structure-based drug design

Introduction

Knowledge of protein structure plays a crucial role in drug discovery, and the use of structural biology by the pharmaceutical community is widely accepted as a paradigm for structure-based drug design. Despite the progress in the field of structural biology (NMR spectroscopy and X-ray crystallography) the gap between the number of known protein sequences and the number of experimentally determined structures remains significant. Computational approaches have been developed to fill this gap, with homology modeling emerging as a commonly used procedure which is often an essential part of the structure-based design paradigm. A key component of homology modeling is the identification of one or more suitable templates. Once a template is found, homology modeling can be used to generate three dimensional models of the query with varying degrees of accuracy [1]. For example, homology models have been useful for developing kinase inhibitors as agents for cancer treatment [2, 3]. Additionally, homology models have been used as successful starting points for virtual screening [4–7]. Accumulated protein modeling studies have shown that homology modeling is the most successful method for structure prediction [8–10].

Homology modeling is used to build three-dimensional models for a protein target based upon a single or multiple templates with known structure. The accuracy of homology modeling is dependent on the fact that evolutionarily related protein sequences often have similar three-dimensional structures. The homology modeling procedure is a multi-step process that can be summarized in the following steps [11]:

1. Sequence alignment of the target and template
2. Target backbone generation
3. Target loop modeling
4. Target side-chain modeling
5. Target model refinement

Electronic supplementary material The online version of this article (doi:10.1007/s10822-009-9274-3) contains supplementary material, which is available to authorized users.

K. A. Rossi (✉) · A. Nayeem · C. A. Weigelt · S. R. Krystek Jr.
Bristol-Myers Squibb Company, Research & Development, Computer-Assisted Drug Design, P.O. Box 5400, Princeton, NJ 08543, USA
e-mail: karen.rossi@bms.com

In our previous work, we addressed issues related to protein sequence alignment and model backbone generation [1]. Those initial studies prompted us to explore the accuracy of protein loop modeling, and results of a subsequent publication showed that loops up to 12 residues in length could be modeled accurately with a backbone RMSD of <2.5 Å [12]. However, only the RMSD of the backbone atoms were scrutinized. Further examination of the results showed the importance of correct side-chain placement; therefore, in the current study we focus our attention on the ability to accurately model amino acid side-chains. In particular, we build upon our previous studies and compare different algorithms in order to examine possible trends in side-chain placement that allow for accurate modeling of protein side-chains following loop sampling.

Materials and methods

Dataset selection

The purpose of this study is to examine the accuracy of protein side-chain sampling in conjunction with loop modeling. Therefore, calculated loops from our previous study were used as the starting point for side-chain sampling. We used our previous dataset containing a total of 197 loops with lengths ranging from 4 to 12 residues [12]. The native protein structures selected for this dataset are a subset of those used by both Sali [13] and Friesner [14], that contain high quality X-ray structures with resolutions better than 2.0 Å and average temperature factors of below 35. The modeled loops used in this study were selected based upon the calculated loop with the lowest RMSD compared to the native structure, irrespective of the method that was used to calculate the loop or the rank order of the loop within the result set. This new dataset contains loops that were generated using Prime, Modeler, Sybyl and ICM. The backbone RMSDs of 171 of the 197 calculated loops were between 0.07 and 2.0 Å.

Prime

The side-chain sampling algorithm used in Prime [14] version 1.6 (Schrödinger, LLC, New York, NY) randomizes the side-chain rotamers. Starting with the first residue at the N-terminus of the loop to be sampled, a side-chain rotamer library identifies the lowest energy rotamer while keeping all other side-chains fixed. All subsequent residues are subjected to the same protocol while holding all other side-chains fixed. This process is repeated until the side-chain rotamers have reached convergence. Finally, all side-chain atoms are minimized using the OPLS2005 force field while holding the backbone atoms fixed. Prime was run in default mode, including allowance for flexibility of all protein side-chains within 7.5 Å of the loop.

ChiRotor

Discovery Studio, version 2.1 (Accelrys Software Inc., San Diego, CA) contains a side-chain sampling method, ChiRotor [15] that uses a systematic torsion scan. Loops were subjected to a constrained minimization using the CHARMm force field prior to side-chain sampling. Prior to side-chain sampling, all side-chain atoms are removed and each residue adopts 3 initial conformations with chi1 = 60°, −60° and 180°. All other chi angles are derived from the CHARMm force field. The three rotamers for each residue are optimized using the CHARMm polar H force field [16] and the two lowest energy conformations are stored. All resulting conformations are minimized using CHARMm. The first residue is then replaced by the second lowest energy side-chain conformation and minimized. If this new structure has a lower energy than the first, it replaces the first one. The subsequent side-chains are sampled in the same manner until the lowest energy structure is obtained.

Sybyl

The "Add Side-chains" functionality as implemented in the Biopolymer module of Sybyl, version 7.1 (Tripos, Inc., St. Louis, MO) was examined in this study using the standard settings for side-chain addition. For each residue side-chain, Sybyl assigns the conformation of the matching residue in the Biopolymer Protein library. Side-chains added in this manner ignore contacts with the rest of the molecule and therefore can often result in severe steric clashes (the option of using a rotamer library that uses conformations that result in the fewest bumps with the rest of the molecule is available in current versions of Sybyl/Biopolymer). Following side-chain addition to the loop residues, a local minimization over the side-chain atoms, keeping the rest of the molecule intact, was carried out with the Tripos force field.

Orchestrar

Andante [17] (Tripos, St. Louis, MO) was run using "no borrow" mode where the side-chains are built from rotamer libraries, optimizing the packing. Rotamers from the library are eliminated based upon steric clashes. If all rotamers fail the steric check, then side-chain conformations from the rotamer library are used. The side-chain with the lowest energy is built for each residue. The results are clustered and subjected to simulated annealing to identify low energy

combinations of rotamers. The lowest energy conformation side-chains are then built back onto the backbone.

ICM

Side-chain modeling as implemented in ICM, version 3.6, (Molsoft LLC, San Diego, CA) relies upon conversion of the molecule into internal coordinates. For amino acids to be modeled, the side-chain torsion angles are systematically sampled using ICM biased probability Monte Carlo optimization [18]. Biased probability Monte Carlo is a powerful global energy optimization method for biomolecular structure prediction [19]. A series of low energy conformations are evaluated for each side-chain, selecting the side-chain orientations with the lowest energy. Following side-chain sampling, the lowest energy side-chains are built onto the backbone and relaxed by local minimization.

Analysis

The results were assessed by comparing the RMSDs of the re-sampled side-chains with the RMSDs of the side-chains from the calculated loops that were used as the starting point for the sampling. Therefore, no change in RMSD was expected if the starting loop originated from the same method as the re-sampled loop. In each case, the RMSDs were calculated with respect to the heavy atoms of the native side-chains. Since the backbone atoms were unmodified during the sampling process, they were excluded from the calculation. The comparison of the two RMSDs with respect to the native structure allowed for a quantitative assessment of the need for side-chain re-sampling with respect to the original results.

The results were categorized as better, worse or equal, as shown in Table 1. When the RMSDs of the re-sampled side-chains are lower than the RMSDs of the starting side-chains by more than 0.15 Å, then the re-sampled RMSDs were considered better. On the other hand, if the re-sampled side-chains RMSDs were greater than the starting side-chain RMSDs by more than 0.15 Å, the re-sampled side-chain RMSDs were considered worse. Differences between the re-sampled and starting side-chain RMSDs in the range of ±0.15 Å were considered to be equal. The quality of the re-sampling was assessed an alternative way by using a ratio of side-chains with improved RMSDs

Table 1 Complete assessment of dataset

	Prime	ChiRotor	Sybyl	Orchestrar	ICM	% Improved	Prime (excluded dataset)
Overall results (197)							(71)
Better	45	44	50	83	57		33
Worse	18	90	95	69	35		14
Equal	134	63	52	45	105		24
Improvement ratio	2.5:1	0.5:1	0.5:1	1.2:1	1.6:1	65	2.4:1
Short (78)							(15)
Better	13	12	19	27	15		6
Worse	5	43	44	35	13		5
Equal	60	23	15	16	50		4
Improvement ratio	2.6:1	0.3:1	0.4:1	0.8:1	1.2:1	59	1.2:1
Medium (68)							(28)
Better	15	19	21	35	25		13
Worse	6	26	28	18	9		4
Equal	47	23	19	15	34		22
Improvement ratio	2.5:1	0.7:1	0.7:1	1.9:1	2.8:1	75	3.3:1
Long (51)							(28)
Better	17	13	10	21	17		14
Worse	7	21	23	16	13		5
Equal	27	17	18	14	21		9
Improvement ratio	2.4:1	0.6:1	0.4:1	1.3:1	1.3:1	63	2.8:1

The complete results are shown for each method, split by overall dataset; short (4–6 aa), medium (7–9 aa) and long (10–12 aa). The number of loops represented by each loop length is shown in parenthesis. The RMSDs (Å) of the starting and re-sampled side-chains are compared to the native PDB side-chains. Categories were determined by the following: $RMSD_{starting} - RMSD_{re\text{-}sampled} \geq 0.15$ Å = better; $RMSD_{re\text{-}sampled} - RMSD_{starting} \geq 0.15$ Å = worse; $| RMSD_{re\text{-}sampled} - RMSD_{starting} | = < 0.15$ Å = equal. The % Improved column represents the percentage of structures that resulted in at least one better side-chain RMSD after re-sampling. The last column shows the Prime re-sampled side-chains on a dataset that excludes Prime starting loops. The numbers in the parenthesis represent the number of loops for each length

compared to worse RMSDs. A ratio higher than 1 is more desirable, where the number of better side-chain RMSDs is greater than the number of worse side-chain RMSDs. These analyses were used to determine which method yielded superior results.

Additionally, a percent improved value was calculated that combined the results and identified whether side-chain re-sampling by any of the different methods showed improvement. Using the criteria described above for categorizing the results, if at least one method improved the RMSD the side-chain re-sampling was considered to have been an improvement. If no improvements were made, but at least one method had equivalent results, the re-sampling was considered equivalent. If all methods resulted in worse RMSDs for the sampled loops, the re-sampling was considered worse. The results for each loop were calculated and normalized over the total number of loops. This type of analysis provided a way to evaluate the need for side-chain re-sampling showing examples where at least one of the methods was able to improve the side-chain RMSD of the calculated loop.

Table 2 shows a third method of analysis; the development of a scoring function that quantified the relative performance of the different methods. The scoring function assigned a value of 1 to an RMSD that decreased by more than 0.15 Å relative to the starting side-chain RMSD, a value of 0.5 was applied to RMSDs that remain within a ±0.15 Å RMSD, and finally a value of 0 was given for RMSDs that were higher than 0.15 Å compared to the starting side-chain orientation. These numbers were added together for each method then divided by the number of loops sampled and multiplied by 100 to generate a score with a range of 0–100, with 100 being the best.

Results and discussion

Overall results

A table containing the complete dataset results is available as supplementary material.

Table 2 Scoring the dataset

	Prime	ChiRotor	Sybyl	Orchestrar	ICM
Overall score	57	38	39	54	56
Short loop score	55	30	34	45	51
Medium loop score	57	45	45	63	62
Long loop score	60	42	37	55	54

Each loop is assigned a score of 0, 0.5 or 1 dependent on the RSMD of the re-sampled side-chain. The normalized score, shown here, is calculated as the weighted average score over the number of loops, expressed as a percentage

A summary of the results of the side-chain re-sampling is shown in Table 1. As mentioned above, the initial conformations of the loops originated from our previous study [12] where loops with the lowest backbone RMSD relative to the native form were selected regardless of which algorithm was used to calculate the loop. The previous study concluded that Prime generated the most accurate loops in the test set, therefore, it is not surprising that the majority of starting loop and side-chain orientations (126 out of 197) originated from Prime calculations.

The overall analysis seen in Table 1 shows a variety of results depending on which method was used for re-sampling. Prime had the largest number of equal side-chain RMSDs which is in-part; due to the fact that the majority of loops re-sampled were based on loops generated using Prime. The Prime loop sampling algorithm uses a side-chain sampling component as part of the protocol; therefore, most of the Prime-generated loops had no change (134/197). A small number of the 126 Prime-generated loops that were re-sampled using Prime returned differing results, which is due to using different versions of the software containing minor changes in the force field. Prime improved approximately 25% (45/197) of the side-chain RMSDs, while only 18 out of 197 were rated worse. This gave the best improvement ratio where 2.5 times as many side-chains were improved compared to those with worse RMSDs. This study suggests additional re-sampling using Prime would likely be beneficial to the model, and is therefore recommended for use in homology modeling protocols.

The large number of Prime-generated starting loops skewed the number of Prime equal side-chain RMSDs. Therefore, an additional assessment of the Prime side-chain sampling algorithm was done which excluded loops that originated from Prime, removing the bias towards equal side-chain RMSDs. These results are shown in the last column of Table 1. This subset of the dataset contains 71 proteins, and the overall results for Prime had a much smaller number of equal RMSDs. The improvement ratios were similar to those of the complete dataset, although the short loops had a considerably lower ratio. This is likely due to the small number of loops in the dataset, containing only 15 examples. The medium and long length loops had slightly higher improvement ratios compared to the complete dataset, with significantly lower numbers of results with equal RMSDs. Overall, for each of the three loop lengths, Prime gave consistently good results. These results reinforce that Prime yields very high quality results for side-chain sampling.

Side-chains re-sampled using ICM had slightly worse results compared to Prime. ICM improved 57/197 side-chains, a larger number than Prime. However, a larger number (35/197) of side-chain orientations were worse

than the starting RMSD compared to Prime. About half of all side-chain RMSDs remained equal after sampling (105/197); therefore a good improvement ratio was observed with a value of 1.6:1. ICM results were particularly good for the medium length loops. Based on our analysis, it was the best algorithm for this loop length, with an improvement ratio of 2.8:1, better than that of Prime (2.5:1). ICM showed fair results for the short (1.2:1) and long (1.3:1) loops.

Despite Prime having the best improvement ratio compared to other methods, re-sampling side-chains with Orchestrar improved the largest number of side-chains of all methods, 83 (Orchestrar) compared to 57 for ICM and 45 for Prime. This shows that re-sampling side-chains using Orchestrar can yield excellent results. However, Orchestrar also had a large number (69) of side-chains with worse RMSDs which is similar to the number of improved side-chain RMSDs. This result is reflected in the improvement ratio of 1.2:1 for Orchestrar. The re-sampling nets a positive ratio, suggesting this method will have slightly better odds of returning improved results over worse ones. The ratio could be greatly improved if Orchestrar was able to reduce the number of worse RMSDs, since this algorithm produced the highest number of improved side-chain RMSDs.

ChiRotor and Sybyl had significantly worse results compared to the three previous methods. These two algorithms had similar results to one another, despite using very different methodology. While the number of improved side-chain RMSDs were similar to Prime and ICM; ChiRotor (44/197) and Sybyl (50/197), the number of worse RMSDs were significantly higher for these two methods than for the other three methods. Both methods had higher RMSDs for the re-sampled side-chains for almost half of the dataset. Sybyl was used in the original study to generate loop conformations, some of which were applied to this study. Since Sybyl uses a database search to determine the side-chain orientations, the results based on Sybyl-generated loops were identical when re-sampled with Sybyl, much like the Prime-originated loops, slightly elevating the number of equal results for Sybyl. The improvement ratio was significantly poorer for ChiRotor and Sybyl compared to the other methods, yielding ratios of 0.5:1 for both methods, suggesting that these methods have a higher probability of returning worse results than better ones. Ideally, one would want to use an algorithm that produces a larger number of better side-chain orientations than worse. These data do not support using ChiRotor and Sybyl for re-sampling side-chains, whereas re-sampling with Prime, ICM and Orchestrar is recommended.

Finally, the overall results can be described by the percentage improved. This value describes the possibility of calculating a better side-chain orientation using any of the methods examined in this study. The value for the complete dataset shows that over 2/3rds of the side-chains sampled were improved by one or more of the methods tested. This result is significant, as it suggests that using an ensemble of side-chain sampling algorithms will yield better results than using any one single algorithm.

Short loops

Additional analysis separates the loops by length: short (4–6 aa), medium (7–9 aa) and long (10–12 aa). When examining the short loops (4–6 amino acids), the starting side-chain RMSDs of the original models for the majority of the examples were <2.0 Å. These results suggest that for those models it would be very difficult to improve upon the placement of side-chains. This is exemplified in Table 1 where out of 78 short loops, three of the methods (Prime, ChiRotor and ICM) improved fewer than 20% of the examples, approximately 13 side-chain RMSDs. Sybyl performed slightly better than the previous three methods, improving side-chain RMSDs by 24% (19/78). The results from Orchestrar are noteworthy, where 35% of the side-chains (27/78) were improved compared to about half as many from the other methods. Another notable result is that only 5 out of 78 side-chain RMSDs were worse using Prime re-sampling. This is associated with the fact that the majority of the starting side-chain orientations came from Prime loop sampling, therefore, a very large number (60/78) short loops did not change when re-sampled using Prime. Compared to the complete dataset, the improvement ratio was lower for all of the methods for re-sampling short loop side-chains, with the exception of Prime. Since the RMSDs of the starting side-chains were initially quite small it is possible that the software algorithms were simply unable to improve upon such good results. The current results support the published data, including our own, which suggest that modeling software can accurately model short loops in proteins [12–14, 20]. Table 1 shows that while the individual results from each method demonstrated little improvement, examination of the percent improved value tells a different story. Although individual sampling methods improved 15–35% of the results, the ensemble of methods improved 59% of the side-chain RMSDs. This can be compared to 65% for the overall results.

Medium loops

Compared to the short loops, the medium length loops (7–9 amino acids) had improved results for all methods except Prime, which remained approximately the same. The medium length loops represent about 1/3rd of the dataset (68/197). Compared to the starting side-chain RMSDs,

side-chain re-sampling for all methods improved about equally for the five methods tested, with Orchestrar again having the largest number of improved RMSDs. The number of worse RMSDs dropped significantly for ChiRotor, Sybyl, and Orchestrar for this loop length. ICM had the largest increase with respect to the improvement ratio, even compared to Prime, with a ratio of 2.8:1 compared to Prime's 2.5:1. The boost to the ICM ratio came from a significant increase in the improved side-chain RMSDs. The results show that re-sampling with ICM produced accurate side-chain conformations for loops with 7–9 amino acids. While ChiRotor and Sybyl both had the least impressive results, the number of better RMSDs, and the improvement ratio increased significantly when compared to the short loops. In contrast to the shorter loop lengths, more than half of the medium length loops contained side-chain RMSDs >2.0 Å; therefore, it is not surprising that the re-sampling algorithms were able to provide better side-chain orientations than the original side-chain orientations. Again, Prime contained the largest number of unchanged results, because a large portion of the dataset originated from loops modeled with Prime. The percent improved results for the medium length loops increased by a large margin compared to the short loop lengths, showing that 75% of the examples had at least one method that yielded better side-chain RMSDs when re-sampled.

Long loops

The long length loops were comprised of a smaller dataset than the short and medium length loops, representing only 51 out of the 197 loop dataset, as fewer long length loops met the criteria for high quality structures for use in validation studies. The improvement ratios for the long loops were similar to those for the entire data set; all methods generated results with a quality between the medium and short length loops. Again, Orchestrar produced the largest number of improved side-chain RMSDs with 21/51 of the examples categorized as better, while Prime and ICM both yielded 17/51 improved results. ChiRotor and Sybyl had results significantly worse than the other methods with improvement ratios well under 1, though ChiRotor had better results than Sybyl. Again, Prime had a very small number of worse RMSDs, giving way to a very high improvement ratio of 2.4:1; almost double that of ICM or Orchestrar. The percent improved for the long loops was also similar to the overall dataset, with 63% of the loops having improved RMSDs using at least one of the methods described. Reports in the literature, which show that high quality loops can be calculated for even longer loops [21], generally focus on the backbone conformation of the loops and overlook the side-chain orientation. The re-sampled side-chain RMSDs of the medium and long loops exemplify the value of applying additional side-chain refinement to these calculated loops. While Prime generally performed the best for these loop lengths, Orchestrar and ICM both improved an equal or higher number of side-chains. If the number of worse RMSDs were reduced, these methods could have greatly improved results.

The percent improved results showed that using a side-chain re-sampling protocol after loop refinement resulted in improvement of side-chain RMSDs in approximately 2/3rds of the proteins examined. This study shows that after loop refinement is completed and side-chains are initially placed, at least one additional method for side-chain sampling should be performed, regardless of the initial algorithm. The results from Table 1 also suggest that for the medium to long length loops, additional side-chain re-sampling, using a second algorithm such as Prime, ICM or Orchestrar, could be used to help improve the placement of side-chains in the refined loops of homology models.

Results of the accuracy score

While the previous analysis stresses the merits of using an ensemble of side-chain sampling algorithms for use in structure-based drug design, it is also important to identify the method that produces the most accurate results. For this purpose, a scoring function was developed to assess and compare more quantitatively, the relative performance of the different re-sampling methods. Table 2 shows the overall score for each of the methods and clearly shows that Prime (57), ICM (56) and Orchestrar (54) generated significantly higher quality results than ChiRotor or Sybyl, which had scores of 38 and 39, respectively.

When scores were broken down into short, medium and long loop lengths, some differences were observed. The scores for the short loops showed that Prime was superior, with the highest score of 55. ICM had a modest score of 51, while poorer results were seen for Orchestrar, 45. Relative to these three methods, Sybyl and ChiRotor had significantly lower scores, 34 and 30, respectively. Additionally, the short loop results for ChiRotor, Sybyl and Orchestrar showed significantly lower scores than the overall score for each of these methods. This is due to the large number of worse RMSDs generated by these algorithms. As described earlier, the side-chain RMSDs for the shorter loops were quite good prior to side-chain re-sampling, and these methods may not be able to improve upon these results.

All methods performed significantly better for the medium length loops compared to the short loops, with the exception of Prime, which was marginally better. Orchestrar and ICM (63 and 62, respectively) had the highest scores for all methods for the medium loop length. While Prime (57) did not have the highest score for this length, it was consistent with the scores for the short and long length

loops. Although ChiRotor (45) and Sybyl (45) had significantly improved results for medium loops compared to short loops, both scores fell below 50, showing that these methods each yielded a higher number of worse RMSDs than improved RMSDs.

The results for the long length loops were similar to the overall results for each method. Prime performed the best for long loops, with a score of 60. In our previous study, we confirmed that Prime produced high quality loop conformations for longer length loops; this study validates the use of Prime for refining side-chains orientations of long loops. Orchestrar and ICM both performed well, with scores of 55 and 54 respectively, while ChiRotor and Sybyl fell well below with scores of 42 and 37.

Assessment of the scores indicated that all methods appeared to work best for re-sampling medium length loops of 7-9 amino acids, with the exception of Prime which performed best for the long loops. Prime had the most consistent scores between loops lengths (short = 55, medium = 57, long = 60) which suggests that Prime works equally well for a range of loop lengths. Additionally, an ensemble, comprised of side-chain orientations generated by Prime, ICM and Orchestrar, is more likely to produce high quality results than one including ChiRotor and Sybyl.

Conclusions

Homology modeling remains an important cornerstone of the structure-based design paradigm due to the fact that structures of many therapeutically relevant targets are unavailable. Our previous studies have detailed the ability to generate quality homology models, even when templates have low sequence identity [1]. We have also shown that high quality loops can be calculated, with backbone RMSDs <1 Å for short and medium length loops and <2.5 Å for longer length loops. While a measure of the backbone RMSD gives a good indication of the quality of the loop conformation, the side-chain orientation is not accounted for. However, side-chain positions are critical in structure-based design; if incorrect side-chain orientations for residues in the binding site were used in molecular docking, the correct pose of the ligand would be difficult to obtain, and the results would be questionable. We are therefore, concerned with the accurate placement of side-chains following backbone modeling. It has become clear from cross-docking experiments and ensemble docking methodology that accurate and/or multiple side-chain orientations are often required to explore the conformational variability of protein-ligand complexes. Calculated loops from our previous study were selected for side-chain re-sampling using five algorithms from commercially-available software that include rotamer libraries, systematic torsion scanning and knowledge-based methods to assess the accuracy of side-chain modeling.

The overall results from this study show the importance of side-chain re-sampling after loop conformation generation. The side-chains re-sampled using Prime had the best overall results, with the highest improvement ratio as well as the best score. While Prime rates highest overall, ICM is a close second improving more side-chain RMSDs than Prime, but also yielding a greater number of worse RMSDs, resulting in a lower ratio and score. While Orchestrar did not receive scores and ratios as high as ICM, Orchestrar improved the largest number of side-chain orientations overall. Since Orchestrar was not used in the previous study to generate loop conformations, it appears that using an additional side-chain sampling algorithm that is different from the algorithm used to calculate the loop will yield better results than using a single algorithm for the complete loop modeling protocol. While Prime generally performed the best, it is clear these results can be improved upon by using an alternative method for side-chain re-sampling such as Orchestrar or ICM. ChiRotor and Sybyl produced inferior results, far below the other methods, with poor improvement ratios that fell well below 1 and scores that were significantly less than 50%. The percent improved analysis clearly indicates that irrespective of the software used to calculate a loop conformation, alternative algorithms could be used for side-chain sampling. While each method returns a single model for the side-chain placement, our study shows that an ensemble of side-chain orientations generated by the collective use of multiple methods can often lead to better models, this is particularly true for Prime, ICM and Orchestrar. These results are consistent with results from ensemble docking experiments [22–24]. However, given multiple side-chain orientations for a loop, current scoring functions are unable to distinguish which side-chain orientation is best, therefore this work supports further development of scoring functions for side-chain optimization.

Acknowledgments We would like to thank the following people for helpful discussions as well as assistance with software: Hongwei Huang (Accelrys), Tamsin Mansley (Tripos), Andrew Orry (Molsoft), and Woody Sherman (Schrödinger).

References

1. Nayeem A, Sitkoff D, Krystek S Jr (2006) Protein Sci 15(4):808. doi:10.1110/ps.051892906
2. Diller DJ, Li R (2003) J Med Chem 46(22):4638. doi:10.1021/jm020503a
3. Rockey WM, Elcock AH (2006) Curr Protein Pept Sci 7(5):437. doi:10.2174/138920306778559368

4. Engel S, Skoumbourdis AP, Childress J, Neumann S, Deschamps JR, Thomas CJ, Colson AO, Costanzi S, Gershengorn MC (2008) J Am Chem Soc 130(15):5115. doi:10.1021/ja077620l
5. Evers A, Klabunde T (2005) J Med Chem 48(4):1088. doi:10.1021/jm0491804
6. Evers A, Klebe G (2004) J Med Chem 47(22):5381. doi:10.1021/jm0311487
7. Radestock S, Weil T, Renner S (2008) J Chem Inf Model 48(5):1104. doi:10.1021/ci8000265
8. Cozzetto D, Kryshtafovych A, Ceriani M, Tramontano A (2007) Proteins 69(Suppl. 8):175. doi:10.1002/prot.21669
9. Moult J, Fidelis K, Kryshtafovych A, Rost B, Hubbard T, Tramontano A (2007) Proteins 69(Suppl. 8):3. doi:10.1002/prot.21767
10. Tramontano A, Morea V (2003) Proteins 53(Suppl. 6):352. doi:10.1002/prot.10543
11. Greer J (1991) Methods Enzymol 202:239. doi:10.1016/0076-6879(91)02014-Z
12. Rossi KA, Weigelt CA, Nayeem A, Krystek SR Jr (2007) Protein Sci 16(9):1999. doi:10.1110/ps.072887807
13. Fiser A, Do RKG, Sali A (2000) Protein Sci 9(9):1753. doi:10.1110/ps.9.9.1753
14. Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, Friesner RA (2004) Proteins 55(2):351. doi:10.1002/prot.10613
15. Spassov VZ, Yan L, Flook PK (2007) Protein Sci 16(3):494. doi:10.1110/ps.062447107
16. Momany FA, Rone R (1992) J Comput Chem 13:888. doi:10.1002/jcc.540130714
17. Smith RE, Lovell SC, Burke DF, Montalvao RW, Blundell TL (2007) Bioinformatics 23(9):1099. doi:10.1093/bioinformatics/btm073
18. Abagyan R, Totrov M (1994) J Mol Biol 235(3):983. doi:10.1006/jmbi.1994.1052
19. Totrov M, Abagyan R (1996) J Struct Biol 116(1):138. doi:10.1006/jsbi.1996.0022
20. Xiang Z, Soto CS, Honig B (2002) Proc Natl Acad Sci USA 99(11):7432. doi:10.1073/pnas.102179699
21. Zhu K, Shirts, Friesner RA (2007) J Chem Theory Comput 3(6):2108–2119. doi:10.1021/ct700166f
22. Rao S, Sanschagrin PC, Greenwood JR, Repasky MP, Sherman W, Farid R (2008) J Comput Aided Mol Des 22(9):621. doi:10.1007/s10822-008-9182-y
23. Kim J, Park JG, Chong Y (2007) Mol Simul 33(8):667. doi:10.1080/08927020701297401
24. Huang SY, Zou X (2007) Proteins 66(2):399. doi:10.1002/prot.21214