JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

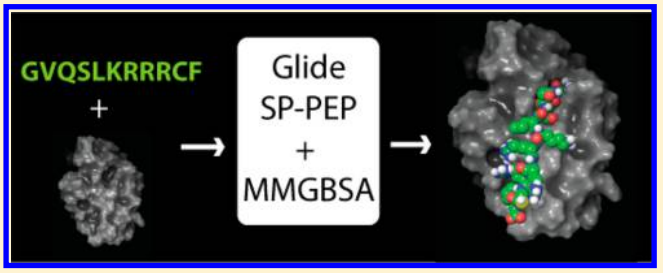
Improved Docking of Polypeptides with Glide

Ivan Tubert-Brohman, Woody Sherman, Matt Repasky, and Thijs Beuming*

Schrödinger, Inc., 120 West 45th street, New York, New York 10036, United States

Supporting Information

ABSTRACT: Predicting the binding mode of flexible polypeptides to proteins is an important task that falls outside the domain of applicability of most small molecule and protein−protein docking tools. Here, we test the small molecule flexible ligand docking program Glide on a set of 19 non-α-helical peptides and systematically improve pose prediction accuracy by enhancing Glide sampling for flexible polypeptides. In addition, scoring of the poses was improved by post-processing with physics-based implicit solvent MM-GBSA calculations. Using the best RMSD among the top 10 scoring poses as a metric, the success rate (RMSD $\leq$ 2.0 Å for the interface backbone atoms) increased from 21% with default Glide SP settings to 58% with the enhanced peptide sampling and scoring protocol in the case of redocking to the native protein structure. This approaches the accuracy of the recently developed Rosetta FlexPepDock method (63% success for these 19 peptides) while being over 100 times faster. Cross-docking was performed for a subset of cases where an unbound receptor structure was available, and in that case, 40% of peptides were docked successfully. We analyze the results and find that the optimized polypeptide protocol is most accurate for extended peptides of limited size and number of formal charges, defining a domain of applicability for this approach.



INTRODUCTION

Discovery of small-molecule therapeutics has benefited from the availability of accurate in silico docking tools that predict the binding mode of ligands to their protein targets. With the recently increased focus on the development of peptide-based therapeutics, there is a growing need to extend docking technologies to the prediction of peptide−protein binding modes. The interest in polypeptides as therapeutics has reemerged in recent years with the rising need to inhibit protein−protein interactions and other difficult drug targets coupled with advances in technologies that make peptides viable as drug candidates. The increasing interest in peptide therapeutics can be seen from the growing number of peptides entering the clinic and getting approved as drugs.[1−3] Advances in experimental approaches to reduce proteolytic cleavage and improve half-life, such as antibody Fc fusions and binding to serum albumin, have proven successful at mitigating some peptide liabilities.[4−6] While experimental approaches have made significant advances related to peptide therapeutics, less progress has been made on structure-based computational approaches.

Several studies describe computational approaches for binding mode prediction of peptide−protein complexes based on Monte Carlo or molecular dynamics sampling schemes.[7−13] However, most of these approaches are target-specific (e.g., developed specifically for PDZ[9] or MHC domains[14]) and often require the input of an approximate bound-state conformation of the peptide,[15] which may not be available in drug discovery projects. A more general tool for peptide docking, called Rosetta FlexPepDock ab initio (henceforth referred to as FlexPepDock), was recently developed and shows good performance across a range of targets and polypeptides.[12] More recently, a version of the HADDOCK protein−protein docking algorithm was shown to also perform very well on this data set.[16] Both of these methods (FlexPepDock and HADDOCK) have approached the peptide-docking problem by adapting strategies developed for protein−protein docking. While the results are encouraging, the methods are computationally intensive, thereby limiting the applicability to a small number of docking calculations run on a large number of processors. For example, FlexPepDock requires thousands of CPU hours per polypeptide docking calculation.

Here, we have approached the problem by adapting a small-molecule docking program to better treat peptides. The molecular docking program Glide[14,17,18] has proven to be reliable for predicting the binding mode of small molecules in protein complexes and identifying active compounds in virtual screening campaigns.[19,20] The Glide SP algorithm is computationally efficient, generally yielding results in 15 s for a typical drug-like molecule. This speed allows large databases of compounds to be screened for putative binders with modest computational recourses. However, to date, Glide has been developed and optimized for docking drug-like small molecules and not polypeptides.

To assess the applicability of Glide SP to the more challenging case of polypeptide docking, we predicted the binding mode of all non-α-helical peptides studied in the

Received: February 26, 2013
Published: June 26, 2013

dx.doi.org/10.1021/ci400128m | J. Chem. Inf. Model. 2013, 53, 1689−1699

Table 1. Peptides Used in the Current Docking Benchmark[a]

PDB ID Holo/Apo	Sequence	Atoms	Rotatable Bonds	Residues	Secondary Structure
1AWR/2ALF	**HAGPIA**	80	19	6	C
1ER8	**HPFHLLVY**	145	35	8	C
1N7F/1N7E	**AVTRTYSC**	124	39	8	b+C
1NLN	**QVQSLKRRRCF**	198	62	11	b+C
1NVR/2QHN	**ASVSA**	61	18	5	b+C
1QKZ	A**NGGASGQVK**	124	40	10	C
1RXZ/1RWZ	KS**TQATLERWF**	193	58	11	b+C
1SSH/1OOT	GP**PPAMPARP**T	154	36	11	C
1TW6	**AVPI**	62	12	4	C
1W9E/1RJ6	**NEFYF**	92	25	5	b+C
1Z9O	**EDEFYDALS**	137	44	9	C
2C3I/2J2I	**KRRRHPSG**	147	44	8	C
2FGR/2FGQ	**DNWQNGTS**	116	37	8	C
2FNT	**RQVNFLG**	120	34	7	C
2J6F	**PPKPRPRR**	152	38	8	C
2O9V/2O9S	V**PPPVPPPPS**	144	25	10	C
2P1K	**SATSAKATQTD**	148	50	11	b+C
2VJ0/1B9K	**FEDNFV**P	106	27	7	C
3D1E	**GQLGLF**	91	25	6	C

[a]All peptides from the FlexPepDock[12] benchmark except for those with α-helical content were included. For cases where both holo and apo structures were available, both PDB codes are given. Secondary structures are indicated as β-strand (b) or random coil (C). Residues that form the peptide−protein interface are indicated in bold font and are used in the iRMSD calculations (see Experimental Section for details).

Table 2. Number of Test Cases with at Least One Good Pose after Each Stage of Docking for Glide Calculations Using Different Sampling Settings (see Supporting Information Table S1 for a detailed overview of all parameter values used)[a]

	experiment	ConfGen	rough scoring	SP min	PDM	final pose
1	default settings (default funnel, 302 directions, 25 rotation steps, 1.0 grid density)	17	10	9	5	4
2	input conformation	17	11	11	7	5
3	wide funnel	17	10	10	8	3
4	wider funnel	17	10	9	8	3
5	widest funnel	17	11	9	8	3
6	widest funnel + 642 directions	17	12	10	8	5
7	**widest funnel + 1002 directions**	**17**	**11**	**10**	**8**	7
8	widest funnel + 1962 directions	17	13	9	9	6
9	1002 directions	17	9	10	5	4
10	wider funnel + 1002 directions	17	10	11	8	6
11	wider funnel + 1002 directions + 50 rotation steps	17	11	10	8	6
12	wider funnel + 1002 directions + 100 rotation steps	17	11	12	9	4
13	widest funnel + 1002 directions + 50 rotation steps	17	11	10	9	6
14	widest funnel + 1002 directions + 100 rotation steps	17	12	10	9	4
15	dense grid	17	9	9	4	3
16	widest funnel + dense grid	17	10	9	9	5
17	widest funnel + dense grid +1002 directions	17	11	9	9	4
18	widest funnel + rigid input conformation	17	12	11	10	6
19	widest funnel + input conformation + 1002 directions	17	13	12	12	11
20	rigid docking	17	6	10	10	10
21	rigid docking + widest funnel	17	6	10	10	10
22	rigid docking + widest funnel + 1002 directions	17	8	11	11	11
23	rigid docking + widest funnel + 1002 directions + dense grid	17	9	11	11	11
24	rigid docking + widest funnel + 1002 directions + 50 rotation steps	17	8	11	11	11
25	rigid docking + widest funnel + 1002 directions + 100 rotation steps	17	8	12	12	12

[a]A pose is considered accurate when it has iRMSD < 2.0 Å, where iRMSD is the RMSD of the backbone atoms for interface residues (see Table 1 for interface residues). The iRMSD was determined for a maximum of 100 poses from a single Glide run. The best result that did not include the input conformer is shown in bold (experiment 7). Only 17 out of the 19 peptides were used in this section of the benchmark, as two peptides (1NLN and 1RXZ) exceeded the maximum number of rotatable bonds (50) in ConfGen and required amide bond constraints.

FlexPepDock benchmark. Docking was performed using both the holo and, when available, the apo structure of the protein. The effect of various docking settings and post-processing steps on the prediction accuracy of Glide SP was systematically evaluated. The optimized polypeptide docking protocol, which includes enhanced sampling plus post-processing of poses with

physics-based implicit solvent MM-GBSA calculations, greatly improves the prediction accuracy compared to default Glide SP docking and approaches the accuracy of FlexPepDock at a fraction of the computational cost. Analysis of the features of peptides that are docked well with the new peptide docking workflow identify smaller size, absence of charged residues, and a high degree of extendedness correlating with accurate reproduction of native protein−ligand poses.

RESULTS AND DISCUSSION

The FlexPepDock[12] data set has 26 polypeptides with diverse secondary structure, including β-strand, random-coil, and α-helical peptides. While the ConfGen[21] conformational sampling algorithm used in Glide has been shown to perform well for small molecules,[22] it fails to generate α-helical conformations for polypeptides unless backbone torsional constraints are added. Therefore, the seven α-helical polypeptides were excluded in this work, and we focus on modifying the Glide docking algorithm to improve predictions on the remaining 19 peptides with an aim to enable ConfGen to generate helical polypeptide conformers in future work. Table 1 lists the final test set of 19 peptides used in this study along with some of their properties. The peptides range from 4 to 11 residues in length and have between 12 and 62 rotatable bonds.

Optimization of Glide SP Sampling for Large Polypeptides. The Glide SP docking algorithm has been described in detail previously.[17] Briefly, Glide uses a series of hierarchical filters to search for possible locations of a ligand in the binding site using a pre-defined grid representation of the rigid receptor. The series of filters can be thought of as a funnel, where the shape and properties of the protein are represented on a grid by different sets of fields that provide progressively more accurate scoring. Initial ligand conformations are generated by the ConfGen algorithm,[21] and a large number of candidate initial poses are generated using an approximate description of both protein and peptide (rough scoring stage). The poses passing this rough scoring stage are partially optimized using the standard OPLS_2005[23−25] molecular mechanics force field (SP minimization stage). Finally, all poses are subjected to post-docking minimization (PDM stage) in the gridded protein field. Top poses are rescored and ranked by a custom scoring function (Emodel).

In this work, the accuracy of each pose at various points in the Glide docking funnel was checked after (1) ConfGen, (2) rough scoring, (3) SP minimization, (4) post-docking minimization, and (5) the final top pose by Emodel is determined. In order to improve pose prediction accuracy for polypeptides, we first optimized sampling by systematically varying parameters that control pose generation and evaluation during the Glide sampling process. These parameters are described in detail in the Experimental Section. In brief, they control the Glide funnel width, ligand diameter orientation, ligand diameter angle, grid density, and van der Waals scaling. Results from all the tests described in methods other than those involving the van der Waals scaling are summarized in Table 2. All holo structures were considered at this stage, with the exception of structures 1NLN and 1RXZ, which were not used in this parameter exploration due to their large number of rotatable bonds. They are, however, included in the scoring optimization benchmark described in the next section.

With default Glide SP settings (referred to as experiment 1), 4 out of 17 test cases result in a good iRMSD (defined as the interface residues having a backbone RMSD under 2.0 Å). The number of cases with at least one good pose decreases at each stage of the Glide docking funnel, with the largest drop occurring during the first docking stage (rough scoring). Encouragingly, the ConfGen stage always succeeds in producing at least one conformer with a good iRMSD. However, the iRMSD metric is not sufficient for evaluating conformers produced by ConfGen, as one could have a perfect backbone conformation (iRMSD = 0.0 Å) that is undockable because of poor side-chain conformations. Fortunately, the RMSD for all heavy atoms also shows good performance of ConfGen, with all but one case (i.e., 94%) having at least one good conformer (data not shown).

To test the effects of conformation generation, we performed rigid docking using the crystallographic polypeptide conformation to give an upper bound for the docking accuracy if the ConfGen algorithm were able to perfectly reproduce the bioactive conformation of all molecules. Using otherwise default settings, rigid docking (experiment 20) succeeds in finding good final poses for 10 test cases. Interestingly, only six test cases have good poses after the rough scoring stage. In fact, in only 11 of the cases was it possible to find at least one pose that passed rough scoring, meaning that the remaining six tests did not return any pose. The reason for this failure could be that rough scoring is not accurate enough to correctly identify the good poses, rotational/translational sampling is not fine enough, or a combination of these two factors. Indeed, as ligands become longer it is necessary to sample the rotational space much more finely to account for the lever arm effect.

The remaining rigid docking experiments (experiments 21−25) attempt to answer whether using a denser grid, more ligand directions, and increased funnel width can increase the number of cases with a good pose. After rough scoring the number of good poses is increased to nine, but this is still below the number for the default flexible docking run. As a final test of the initial conformation generation, we performed flexible docking in which the input conformation is appended to the automatically generated conformational ensemble to give an indication of how an improved ConfGen algorithm could affect docking accuracy. Default settings (experiment 2), including the input conformer, only increase the success rate from four to five, indicating that other parts of the Glide docking funnel are responsible for the incorrect results.

Next, we explored various parameters associated with the docking funnel after the initial conformation generation stage. Increasing funnel width (experiments 3−5) with the other parameters at their default values degrades final results slightly (three cases produce a good pose in the top 10). The difference is not statistically significant. Therefore, it is not possible to draw conclusions, but the lack of improvements emphasizes the challenges of polypeptide docking and that simply tuning a single parameter may not lead to improved results. However, the number of cases with good poses at the post-docking minimization stage after increasing the funnel width actually increases from 5 to 8, showing that there are cases where a good pose exists in the top 100 poses but does not get selected in the final top 10 poses. Indeed, optimization of the scoring function (see below) suggests that the current form of Emodel is not an optimal scoring function for polypeptide pose selection.

Decreasing the grid spacing by a factor of 2 for finer sampling (experiment 15) with default Glide SP settings degrades the final results slightly, resulting in three good predictions. However, combined with increased funnel width (experiment

Table 3. Number of Test Cases Producing Good Poses (columns 6−9) for Various Values of the Ligand and Receptor vdW Scaling Factors (columns 2−5)[a]

experiment	ligand charge cutoff	ligand scaling	receptor charge cutoff	receptor scaling	rough all	rough	SP min	final
rigid	0.15	0.8	0.25	1	11	6	10	10
flexible	0.15	0.8	0.25	1	17	10	9	4
rigid/scale ligand A	0.15	0.6	0.25	1	13	7	11	10
rigid/scale ligand B	0.15	0.4	0.25	1	14	7	11	10
rigid/scale ligand C	0.15	0.2	0.25	1	14	7	12	9
rigid/scale ligand D	1	0.8	0.25	1	12	7	11	10
rigid/scale ligand E	1	0.6	0.25	1	14	8	11	10
rigid/scale ligand F	1	0.4	0.25	1	15	9	12	12
rigid/scale ligand G	1	0.2	0.25	1	16	10	12	10
rigid/scale receptor A	0.15	0.8	0.25	0.9	13	5	7	7
rigid/scale receptor B	0.15	0.8	0.25	0.8	13	5	7	6
flexible/scale ligand	1	0.4	0.25	1	17	11	9	4
flexible/scale ligand/experiment 7	1	0.4	0.25	1	17	11	11	5

[a]"Rough all" is the number of test cases that produce at least one pose (good or bad) after rough scoring.

16), the number of good predictions improved to five. It is possible that increased grid density without a corresponding increase in funnel width causes the set of rough poses to be less comprehensive, thus increasing the chances that a good pose will be "crowded out" based on the approximate rough score when in fact it could rank better with a more accurate scoring function.

Increasing the number of directions for the ligand diameter from the default value of 302 while keeping other settings at the default Glide SP values (experiment 9) does not affect the final number of good poses. However, combined with an increased funnel width (experiments 6−8), there is an increase in the number of cases with good final poses, reaching a maximum of seven cases for 1002 directions (increasing the directions further deteriorates the result, and importantly, increases the CPU/memory usage significantly). Experiment 7 shows the best results obtained from tuning the docking funnel parameters and is henceforth referred as the SP-PEP method. Including the crystallographic conformation in the initial ConfGen conformational ensemble with the settings from experiment 7 increases the number of cases with good final poses to 11, indicating the best result that can be achieved with the current Glide docking and scoring protocol if the ConfGen algorithm were able to perfectly reproduce the bioactive conformation for all polypeptides studied here.

Increasing the number of steps used to rotate the ligand around its diameter did not improve docking accuracy for flexible docking. This setting was only tested in conjunction with an increased number of directions and increased funnel width (experiments 11−14). In fact, increasing the number of steps degraded the result (i.e., 100 rotation steps was worse than 50 rotation steps, which was worse than the default 25). For rigid docking, however, the end result did improve slightly and the best result from Table 2 (experiment 25) is the one for rigid docking with "widest funnel", 1002 directions, and 100 rotation steps (12 cases resulted in a good final pose).

The effect of changing the vdW scaling factor for ligand and receptor was tested first in the rigid docking experiment. Changing the ligand scaling factor to reduce vdW radii resulted in improvements in the number of cases with good poses after rough scoring (Table 3). The best results for rigid docking were obtained from scaling all atoms instead of only the nonpolar atoms and setting the scaling value to 0.4 (experiment "rigid/scale ligand F"). This is a rather extreme reduction in vdW radii, and applying it to flexible docking indeed decreased the benefit of increasing the number of directions and funnel width (e.g., changes equivalent to experiment 7, Table 2). Changing the receptor scaling factor did not improve the overall docking outcome. The vdW scaling results are shown in Table 3.

Because the greatest loss of good poses happens at the rough scoring stage, especially for rigid docking, we performed additional vdW scaling at this stage. VdW scaling was applied to all atoms regardless of their partial charge. The results of these experiments are shown in Table 4. While the additional vdW scaling improved rigid docking somewhat, flexible docking accuracy degraded.

Table 4. Number of Test Cases Producing Good Poses (columns 3−6) for Various Values of the Rough Score-Specific vdW Scaling Factor[a]

experiment	ligand scaling	rough all	rough	SP min	final
rigid	1	11	6	10	10
rigid/scale ligand	0.8	14	7	11	10
rigid/scale ligand	0.6	14	9	11	11
rigid/scale ligand	0.4	15	10	12	12
rigid/scale ligand	0.2	16	10	12	12
flexible	0.6	17	11	10	4
flexible/scale ligand + experiment 7	0.6	17	11	12	3
flexible/scale ligand + experiment 7 + input conformation	0.6	17	14	14	10

[a]"Rough all" is the number of test cases that produce at least one pose (good or bad) after rough scoring.

Physics-Based Post-Processing and Rescoring (MM-GBSA). As stated above, the best possible combination of Glide docking settings identified here are in experiment 7 (Table 2) and referred to as SP-PEP. Given the improved sampling of this algorithm, we next tested whether application of post-processing and rescoring steps could further improve the identification of correct poses from the ensemble of predictions. Leveraging a dependence of Glide on the input bond lengths and angles of the ligand,[26] we ran the SP-PEP method 10 times for each of the holo structures with different starting conformations, collecting a total of 100 poses for each run. The 10 starting conformations were generated using a MacroModel sampling method recently developed to efficiently

Table 5. Accuracy of Docking Using Multiple Independent Default and SP-PEP Glide Runs (columns 2−4) and Accuracy of Pose Ranking Using Various Scoring Functions (columns 5−9)[a]

	number of good poses (<2 Å iRMSD) in ensemble (1000)			lowest iRMSD in ensemble	lowest iRMSD of 10 top-scoring poses			
peptide	Glide SP	Glide SP-PEP	SP-PEP + MM-GBSA	SP-PEP + MM-GBSA	GlideScore	Emodel	MM-GBSA backbone/heavy atom	FlexPepDock
1AWR	170	197	172	0.94	1.90	1.43	1.12/3.27	0.90
1ER8	0	2	2	1.77	1.22	1.78	1.78/2.33	1.40
1N7F	5	18	67	1.22	2.18	2.18	1.96/2.42	1.40
1NLN	2	138	144	0.41	0.64	0.64	0.41/1.41	2.00
1NVR	19	154	154	0.8	0.70	0.70	0.63/0.92	0.50
1QKZ	0	0	2	1.74	1.69	1.69	7.40/8.03	8.80
1RXZ	0	0	0	3.57	5.31	5.31	4.77/6.57	0.70
1SSH	26	46	125	0.48	0.52	1.26	0.75/0.74	3.10
1TW6	307	196	262	0.34	0.33	0.33	0.34/0.34	5.90
1W9E	2	5	9	0.3	2.12	1.92	1.65/2.94	1.50
1Z9O	0	0	0	3.65	5.45	5.45	5.45/5.34	1.60
2C3I	0	0	0	2.96	3.39	3.39	4.00/3.77	6.00
2FGR	0	0	0	4.52	7.78	4.91	4.91/5.11	8.60
2FNT	1	3	4	1.31	1.81	1.81	1.31/2.09	1.00
2J6F	2	10	11	1.32	10.66	3.48	10.74/12.86	5.70
2O9V	0	0	2	1.94	3.78	3.78	2.05/2.49	0.60
2P1K	0	4	3	1.08	1.73	2.62	2.43/2.72	0.60
2VJ0	1	1	4	1.79	2.54	2.54	1.79/1.40	1.50
3D1E	56	135	170	0.39	0.85	0.85	0.85/0.76	1.10
			<2 Å iRMSD	15	10	10	11/6	12
			<3 Å iRMSD	16	13	13	13/12	13

[a]Peptides were docked against the holo structure. Ten Glide runs produced up to 1000 total poses (maximum of 100 from each run), and the number of poses with iRMSD < 2.0 Å are reported. For the best sampling method (SP-PEP + MM-GBSA), we also report the lowest iRMSD in the total ensemble. Lowest iRMSD amongst the top-scoring poses from 10 runs of the SP-PEP method using different scoring functions from the Glide and Prime programs was calculated, and compared to those using FlexPepDock.

Table 6. Accuracy of Docking against apo Structures Using SP-PEP Glide with MM-GBSA Scoring[a]

peptide		number of good poses in ensemble (1000)	smallest iRMSD in top 10			
apo ID	holo ID	SP-PEP + MM-GBSA	SP-PEP + MM-GBSA	FlexPepDock	unbound receptor Cα-iRMSD	native docking success
2ALF	1AWR	105	0.86	1.4	0.3	x
1N7E	1N7F	19	1.64	1.0	0.4	x
2QHN	1NVR	74	0.54	0.4	0.3	x
1RZW	1RXZ	0	3.94	4.3	1.5	
1OOT	1SSH	0	4.45	3.2	0.7	x
1R6S	1W9E	2	1.99	1.9	0.6	x
2J2I	2C3I	0	3.62	3.7	0.2	
2FGQ	2FGR	0	5.02	10.1	0.3	
2O9S	2O9V	0	11.23	0.6	0.3	
1B9K	2VJ0	0	5.22	1.4	0.3	x
		<2 Å iRMSD	4	6		

[a]In 4 out of 10 cases, accurate poses are found among the top-scoring poses from 10 independent runs, compared to six for FlexPepDock docking to the apo structures and six for native SP-PEP docking to the holo structures. Cα-iRMSD is the RMSD of Cα atoms for binding site residues, as has been described previously.[12]

sample macrocycles, similar to the LowModeMD approach.[27] The total number of correct poses in the ensemble of 1000 is shown in column 3 of Table 5 and compared to values for default Glide SP (column 2). SP-PEP poses were subjected to MM-GBSA post-processing. MM-GBSA uses Prime[28] with the OPLS_2005 force field and the VSGB2 implicit solvent model. For the MM-GBSA work here, the receptor was treated rigidly and the peptide was minimized. This combinations is henceforth referred to as SP-PEP + MM-GBSA.

As shown in column 5 of Table 5, near-native solutions were found in the ensemble for 15 out of 19 cases (79%) using the SP-PEP algorithm with MM-GBSA post-processing. Poses from this SP-PEP + MM-GBSA sampling strategy were ranked using a variety of scoring functions. These included the scoring functions GlideScore and Emodel from Glide as well as the MM-GBSA free energy from Prime, which is obtained by a post-processing of the Glide poses. The metric used to determine prediction accuracy of FlexPepDock was the lowest RMSD of the interface residues (iRMSD) among the top 10 poses. Therefore, the best (i.e., lowest) iRMSD values among the top pose from the 10 independent SP-PEP calculations is shown in Table 5, columns 6−9.

The best overall performance is obtained using the MM-GBSA score, with 11 cases producing a pose with iRMSD < 2.0

dx.doi.org/10.1021/ci400128m | J. Chem. Inf. Model. 2013, 53, 1689−1699

Table 7. Classification of Complexes Based on Pose Reproduction[a]

		GlideScore		Emodel		MM-GBSA		peptide structural properties			
		number of accurate poses in top 10	highest ranking of accurate pose	number of accurate poses in top 10	highest ranking of accurate pose	number of accurate poses in top 10	highest ranking of accurate pose	residues	normalized extendedness	charged side chains	solvated charged side chains
A	1N7F	10	1	10	1	5	2	**8**	0.91	1	**1**
	1NLN	9	1	9	1	10	1	**11**	0.92	**4**	0
	1NVR	7	2	5	2	9	1	5	0.94	0	0
	1TW6	10	1	10	1	10	1	4	0.79	0	0
	2FNT	2	1	2	1	4	1	7	0.9	1	0
	3D1E	7	2	5	1	10	1	6	0.86	0	0
B	1AWR	3	7	3	3	10	1	6	0.85	0	0
	1ER8	0	67	0	50	1	5	**8**	0.78	0	0
	1SSH	1	7	0	16	8	1	**11**	0.91	1	0
	1W9E	1	9	2	5	0	18	5	0.87	1	**1**
	2P1K	0	57	0	80	1	9	**11**	0.9	**2**	**1**
	2VJ0	0	62	0	92	0	14	7	**0.63**	1	0
C	2J6F	0	67	0	101	0	321	**8**	0.88	**4**	**2**
	2O9V	0	164	0	246	0	13	**10**	0.88	0	0
D	1QKZ	1	4	1	8	0	92	**10**	**0.41**	1	**1**
	1RXZ	0	–	0	–	0	–	**11**	**0.75**	**3**	**3**
	1Z9O	0	–	0	–	0	–	**9**	0.91	**4**	**4**
	2C3I	0	–	0	–	0	–	**8**	0.87	**5**	**1**
	2FGR	0	–	0	–	0	–	**8**	**0.54**	1	**1**

[a]For each scoring function, both the number of poses in the top 10 of the ensemble (1000 poses) and the ranking of the highest pose is reported. Category A contains complexes that produce many good results with SP-PEP. Category B complexes have accurate predictions for a small subsegment of the peptide. Category C is difficult due to internal symmetry. Category D peptides have significant internal folding and are hard to predict. Several structural criteria associated with accurate and inaccurate docking of polypeptides are reported in the last four columns. Large number of residues, limited extendedness, large number of charged side chains, and large number of solvated charged side chains correlate with low docking accuracy (categories C and D).

Å. GlideScore and Emodel have similar performance, with 10 out of 19 cases predicted under 2 Å. The best results obtained here are slightly worse than FlexPepDock (12 out of 19) but come at a fraction of the computational cost. When the iRMSD threshold is raised to 3.0 Å, which still yields reasonable-looking poses given the large size of the polypeptides studied here, all methods reach the same performance (13/19 cases). Interestingly, the correlation between iRMSD results for SP-SEP and the FlexPepDock is relatively low (r^2 value of 0.31), indicating the orthogonality of these two methods. Indeed, several cases that perform poorly with SP-PEP (1RXZ, 1Z9O, and 2O9V) yield good results with FlexPepDock and vice versa (1QKZ, 1SSH, and 1TW6). Cases that present problems for all methods include 2C3I, 2FGR, and 2J6F. For the SP-PEP + MM-GBSA protocol we also measured the iRMSD for all heavy atoms of the peptide (Table 5). As expected, iRMSD values are higher than for backbone alone, but in 12 out of 19 cases the iRMSD values of the best pose in the top 10 predictions remains under 3 Å.

Finally, to apply the optimized docking method to a more realistic problem, we docked peptides to a subset of cases where an apo structure was available. The results are shown in Table 6. As expected, the accuracy of rigid-receptor docking SP-PEP docking against apo structures is worse than docking against holo structures due to the suboptimal geometry of the apo binding site. However, for a number of cases, correct solutions are found, in particular for the subset where induced fit effects are small (Cα iRMSD < 0.6 Å), and docking to the holo structure results in a good solution. A more appropriate strategy for dealing with induced-fit effect would require implementation of the SP-PEP algorithm in the Induced Fit Docking protocol,[29] which will be the focus of future work.

CPU Time. The Glide SP-PEP docking protocol increases the average time to dock each ligand from a few minutes with default Glide SP to an average of 14 min for the data set studied here. Post-processing of 100 poses per Glide-SP PEP run using MM-GBSA took on average 160 min on a single CPU. The full workflow producing the best results requires 10 independent docking calculations followed by Prime post-processing of the top 100 poses, resulting in a total computational time of approximately 24 CPU hours. As a reference, FlexPepDock takes months of CPU time per polypeptide. Both the multiconformer Glide/MM-GBSA approach and FlexPepDock are highly parallelizable, so wall clock times can be reduced considerably by using more processors (e.g., an average of 2.5 wall-clock hours per peptide when run across 10 CPUs using the protocol described here).

Structural Analysis of Successes and Failures. What follows is a structural analysis of the predictions for each of the protein−peptide complexes to assess what causes predictions to fail or succeed. While the top 10 iRMSD metric is an effective way to compare among different methods, for the present analysis it is useful to also look at how many of the top 10 predictions are accurate and to what extent native poses in the top 10 are outranked by the non-native poses. These metrics

dx.doi.org/10.1021/ci400128m | J. Chem. Inf. Model. 2013, 53, 1689−1699

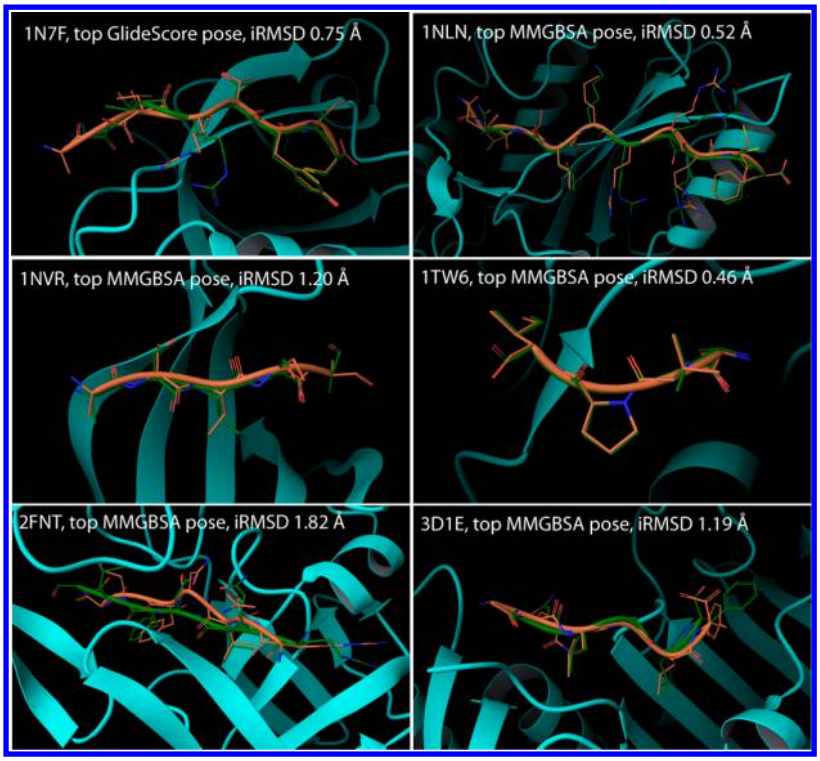


Figure 1. Protein−peptide complexes, predicted accurately using all possible versions of Glide (category A, Table 7). Proteins are shown in cyan, X-ray peptides in orange, and predictions in green. In all cases, the pose with the best MM-GBSA energy is shown, except for 1N7F, where the top GlideScore pose is shown.

are shown in Table 7, and they allow for a classification of complexes into four categories: A, B, C, and D.

Category A includes complexes that can be predicted reliably with iRMSDs of less than 2 Å using most of the scoring functions. Typically, the top-ranked pose across all experiments is accurate, and accurate poses are found multiple times in the ensemble. These complexes are 1N7F, 1NLN, 1NVR, 1TW6, 2FNT, and 3D1E (Figure 1). These complexes included both very short (1TW6, four residues) and long (1NLN, 11 residues) peptides.

Category B includes complexes that have moderately accurate predictions for the full-length peptide (albeit not necessarily using all scoring functions) but have a significantly better iRMSD for the nonterminal part of the peptide. This group includes complexes 1AWR, 1ER8, 1W9E, 1SSH, 2VJO, and 2P1K. These complexes typically have an incorrectly predicted N-terminus, C-terminus, or both. The incorrect prediction of these termini is partially due to an overestimation of the energy of interaction of the charged amino or carboxylate groups (see the analysis below), which are more solvent exposed in the X-ray structure than in the predicted top-scoring poses. Another contributing factor is the intrinsic flexibility of these residues as revealed by their relatively high B-factors. This is illustrated in Figure 2, in which for each case one or two predicted poses are compared with the X-ray structures (left), and per-residue prediction accuracy is compared with residue B-factors (right). Thus, in these cases, the conformation predicted by the docking program may represent an accessible state of the peptide that is not in the X-ray structure. A similar conclusion can be reached by comparing the iRMSD values for shorter sections of the peptide reported in the FlexPepDock benchmark with the full-length iRMSD values obtained there.

The final two categories contain the most challenging cases. Category C contains complexes 2J6F and 2O9V, for which all methods predict poses far from the native pose (Figure 3). Complexes 2J6F and 2O9V both contain poly proline motifs, which are hard cases for Glide to predict due to the challenges associated with sampling of ring conformations. In addition, these peptides have a significant degree of N→C/C→N symmetry. For example, 2J6F has a pseudosymmetric basic-Pro-basic-Pro-basic motif, while 2O9V has a pseudosymmetric Pro-Pro-Pro-Val-Pro-Pro-Pro motif. Indeed, the highest scoring peptides typically have the peptide direction reversed. In the rare cases that the correctly aligned peptide is identified, the score is substantially worse than the reversed peptides.

Finally, complexes 1QKZ, 1RXZ, 1Z9O, 2C3I, and 2FGR (category D) are the most difficult to predict with the SP-PEP approach. For 1QKZ, the algorithm identifies two poses in the entire ensemble that are close to native, while for the others not a single native-like conformation is identified. Moreover, ignoring one or more of the termini in the RMSD calculation does not improve the detection of near native poses, in contrast to the peptides in category B. The majority of these cases result in poor results with FlexPepDock as well, but 1RXZ and 1Z9O can be accurately predicted using that method, suggesting that improvements to our sampling and scoring strategy might increase accuracy.

Several polypeptide properties have been identified to correlate with low iRMSD pose predictions (Table 7, last four columns). As expected, there is an inverse correlation of iRMSD with peptide length, where longer peptides generally produce less accurate pose predictions. Thus, seven cases with poor SP-PEP results have a peptide length of eight residues or longer, compared to five out of 12 cases where SP-PEP is successful. Peptide extendedness, however, is positively

dx.doi.org/10.1021/ci400128m | J. Chem. Inf. Model. 2013, 53, 1689−1699

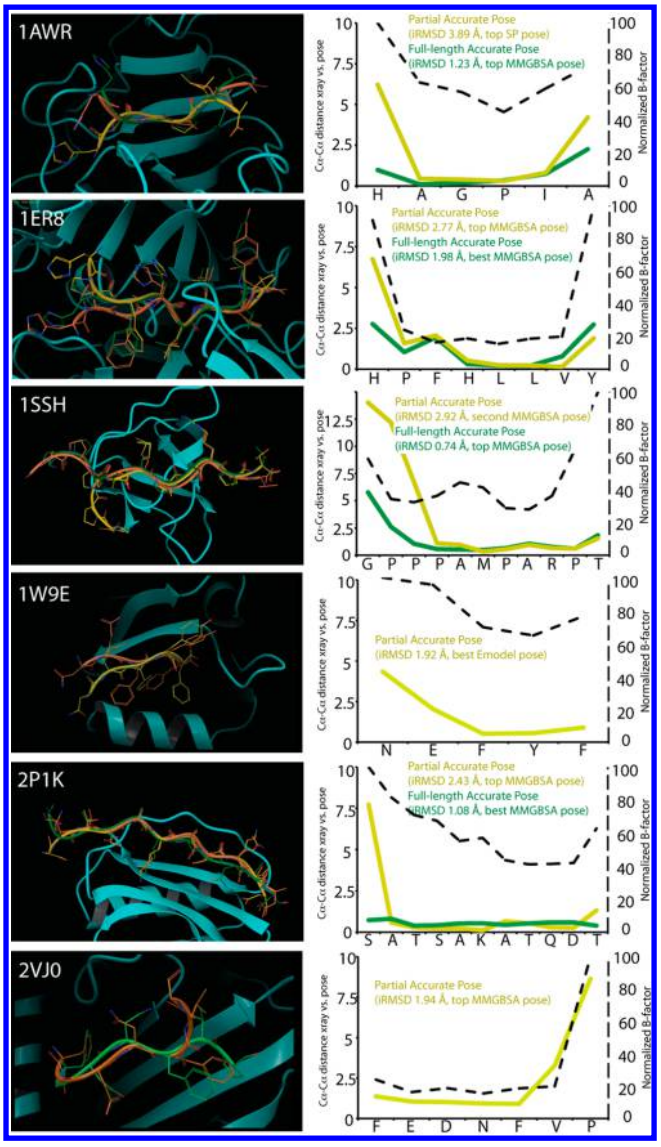


Figure 2. Protein−peptide complexes where inaccurately predicted N- and/or C-termini lead to high RMSD values (category B, Table 7). In the panels on the left, proteins are shown in cyan, X-ray peptides in orange, and predictions in green and yellow. In cases where all predicted poses have inaccurate termini, only one pose is shown (1W9E and 2VJ0). For all other cases, two poses are shown that have the termini either correctly (green) or incorrectly predicted (yellow). Panels on the right compare the per-residue prediction accuracy (as measured by Cα−Cα distances between predictions and structures) with whole-residue B-factors (dashed line). In all of these systems, a subset of the peptide can be predicted with very high accuracy (<1 Å) while N- and C-termini typically have much higher errors. In all cases, the incorrectly predicted segments have high B-factors.

correlated with the quality of pose prediction, and three out of seven cases with poor SP-PEP results form more compact conformations in the crystal structure (i.e., low extendedness) compared to one out of 12 cases where SP-PEP is successful (setting the cutoff to 75% of the maximum extension of a β-strand). It is possible that a modification of the Glide algorithm to explicitly reward internal hydrogen bond formation could improve predictions of these nonextended cases and will be explored in future work.

Finally, the number of charged side chains and the number of solvent exposed charged side chains that do not form a salt

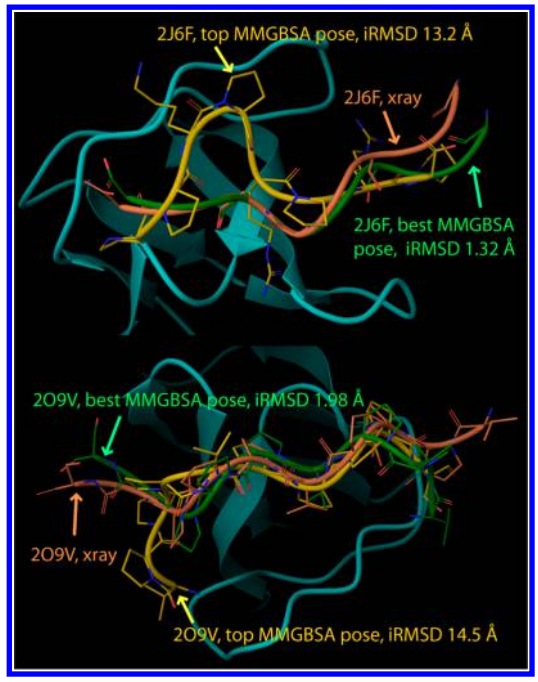


Figure 3. Incorrectly predicted symmetrical polyproline containing peptides (category C, Table 7). Proteins are shown in cyan, X-ray peptides in orange, predictions in green (best pose) and yellow (top MM-GBSA pose). In both cases, the peptide direction is reversed for the top MM-GBSA pose (yellow).

bridge with the protein are inversely correlated with good results. In particular, four out of seven cases with poor SP-PEP results have more than one charged side chain, compared to two out of 12 cases where SP-PEP is successful. Along these same lines, six out of seven cases with poor SP-PEP results have a charged side chain that is not involved in a salt bridge with the protein, compared to three out of 12 cases where SP-PEP is successful. Notably, the only highly charged peptide that produces good results (1NLN) has all four charged side chains forming salt bridges with the protein. In summary, using the SP-PEP docking protocol, small, extended, uncharged peptides tend to produce good results, whereas large, internally folded, and charged peptides with several solvated charged side chains are more challenging. The size and charge of peptides can be directly obtained from the amino acid sequence of the peptide. For the prediction of peptide secondary structure, several accurate algorithms exist.[30,31] Thus, is should be possible to prospectively determine the difficulty of docking peptides with the SP-PEP protocol in Glide.

Identification of the Peptide Binding Site on the Protein Surface. In the results presented above, it was assumed that the approximate location of the peptide binding site on the protein surface was known. While this is a reasonable assumption for a large number of conserved peptide binding domains for which sufficient structural information is available (e.g., SH2, PDZ, and WW domains), there are cases where this information is lacking. The binding site detection algorithm SiteMap[32,33] has shown success in identifying small molecule binding sites on protein surfaces, but due to the more shallow and increased polar nature of many protein and peptide binding sites, it is not necessarily suitable for the type of binding sites studied here. To increase performance of SiteMap recognition of peptide and protein binding sites, we found that modification of two key parameters in the SiteMap algorithm

dx.doi.org/10.1021/ci400128m | J. Chem. Inf. Model. 2013, 53, 1689−1699

("enclosure" and "maxvdw", which control the enclosure and hydrophobicity cutoffs used in detecting binding sites, respectively) improved the coverage of binding site detection with an acceptable loss in precision. As shown in Table 8, using

Table 8. Prediction of the Peptide Binding Site Using SiteMap, Which Predicts and Ranks Putative Binding Sites on the Surface of Proteins[a]

protein surface	default sitemap ranking	peptide-optimized sitemap ranking	class (from Table 7)
1AWR	1	1	B
1ER8	1	1	B
1N7F	2	1	A
1NLN	1	1	A
1NVR	not detected	2	A
1QKZ	not detected	not detected	D
1RXZ	2	not detected	D
1SSH	not detected	2	B
1TW6	1	1	A
1W9E	3	2	B
1Z9O	not detected	not detected	D
2C3I	1	3	D
2FGR	not detected	not detected	D
2FNT	1	1	A
2J6F	not detected	3	C
2O9V	not detected	2	C
2P1K	not detected	1	B
2VJ0	not detected	not detected	B
3D1E	3	1	A

[a]Modification of several parameters significantly increases the performance for protein−peptide complexes.

these modified settings the peptide binding site is correctly predicted for 14 out of 19 cases (compared to 10 out of 19 for the original SiteMap settings). In eight of the cases, the correct binding site is the top ranked prediction with the new SiteMap settings (six using the original SiteMap settings). Finally, accuracy increases significantly when considering only class A and B peptides (Table 7). Here, 11 out of 12 sites are identified, with eight of these being the top prediction.

CONCLUSIONS

The increased interest in exploring polypeptides and peptidomimetics as potential therapeutics has prompted the development and adaptation of docking tools specifically toward these molecules. In this study, we explored the ability of Glide SP to predict the structure of peptide−protein complexes. We implemented a new peptide-specific sampling and scoring protocol, termed SP-PEP, to improve the docking accuracy level. Using default Glide SP settings, only 24% of the cases examined here produced an accurate pose (iRMSD < 2 Å) in the ensemble (10 poses) of a single Glide run. Through systematically modifying settings in Glide, the accuracy of sampling in polypeptide prediction was increased. In particular, changing the parameters that increase the number of initial ligand placements (so-called ligand diameter directions) with parameters that control the "funnel width" of Glide had the largest effect on the accuracy. Interestingly, the ConfGen algorithm that is used for generating receptor-independent conformations of ligands performed exceptionally well, with 94% of the non-helical peptides having a heavy-atom RMSD < 2.0 Å. This is particularly noteworthy given that ConfGen was designed for small drug-like molecules and has not been parametrized for polypeptides.

In addition to modifying docking settings, it was found to be advantageous to run Glide multiple times using different random starting conformations, leveraging the fact that there is a dependence of the docking result on the starting bond and angle values. Thus, we ran Glide 10 times saving the top 100 poses from each calculation and consequently analyzed a total of 1000 poses for docking accuracy. Finally, we refined each of these docked poses by a minimization using a physics-based scoring method with an implicit solvent model, as implemented in Prime. Using this approach, in 79% of the cases, a good pose was identified in the ensemble, compared to 53% for 10 default Glide runs.

Having increased the efficiency of the sampling algorithm, several scoring functions were tested to identify native-like solutions in the ensemble. While all tested scoring functions (GlideScore, Emodel, and MM-GBSA free energy) showed improvements, the poses scored with MM-GBSA were the most accurate, with 58% of the cases having a good pose in the top 10 ranked poses, compared to 63% for the FlexPepDock benchmark. An important point is that in several cases an incorrectly predicted N- or C-termini of the peptide leads to a relatively high RMSD, while the majority of residues are predicted at very high accuracy (Figure 2). These errors might reflect intrinsic disorder of that part of the peptide, as typically the residues with the largest error have elevated B-factors as well (see the black dashed line in the right panels of Figure 2). In addition, poorly docked peptides tended to have more solvated charged groups, suggesting errors in termini prediction results in part from a tendency to over-reward buried charge−charge interaction using the current scoring functions.

Glide is a rigid receptor docking tool, and as such we have ignored here any treatment of flexibility of the receptor. Induced Fit Docking[29] and similar methods exist to treat protein flexibility during docking, although that was not explored in this work. However, it appears from a structural analysis of known protein−peptide complexes that in general most peptides do not induce large-scale conformational changes in their protein binding partner.[13] Indeed, when applying the SP-PEP protocol against the available apo structures for a subset of the peptides studied here, accurate poses were found in a significant number of cases (40%). Another important limitation of the current implementation of the ConfGen algorithm in Glide is that it is unable to handle α-helical peptides without imposing torsional constraints, which is contrast to the good results obtained by the FlexPepDock method on α-helical peptides. However, the previously carried out structural analysis shows that the majority of peptides in the peptide−protein complex in the PDB have a non-helical coil or strand character,[13] which is predicted well with the current methodology. In addition, if polypeptides are expected to be helical based on sequence or other information, that can be added as a constraint to the conformation generation protocol; that was not explored here. Finally, the force field and empirical scoring parameters in Glide have not been optimized specifically for peptides. This has the benefit that molecules with features from both peptides and small molecules (e.g., peptidomimetics) can be easily handled by the current SP-PEP algorithm without any further parametrization. However, optimization of Glide scoring with a focus on large compounds binding to flat protein interfaces is likely to lead to further

improvements in the docking performance and will be the focus of future work.

EXPERIMENTAL SECTION

General Calculations and Measurements. All peptide complexes and corresponding apo structures listed in Table 1 were downloaded from the Protein Data Bank.[34] Structures were prepared with the Protein Preparation Wizard in Maestro 9.3.[35] Glide docking grids were prepared with default settings using the peptide ligand to define the location of the center of the grid. Peptides were docked using either default Glide SP or using optimized parameters for peptides (see below). Peptides were docked starting from multiple random conformations, which were generated using a custom MacroModel sampling method developed to efficiently sample macrocycles. Ten representative peptide conformations were selected after clustering of conformers. Poses were refined in Prime[28] using a simple default minimization with the VSGB2.0 implicit solvent model.[36] B-factors (Figure 2) were obtained by averaging over all heavy atoms in each amino acid and normalizing by setting the highest value in the peptide to 100. Normalized extendedness of peptides (Table 7) was determined by taking the Cα−Cα distance of the first and last residue in the peptide and dividing it by the number of residues and by the maximum value observed in fully extended β-strands (3.0 Å)

Metric of Success. For comparison with FlexPepDock, we use the iRMSD metric defined previously,[11] which is the RMSD of the interface backbone atoms. A peptide residue belongs to the interface if its Cβ atom is less than 8 Å from any Cβ atom on the receptor (Cα atom for glycine). All residues in the peptide that are not at the interface and excluded from the RMSD calculation are shown in gray in Table 1. In this work, a pose is considered "accurate" when it has iRMSD < 2.0 Å. To track the performance of the conformation search algorithm, the accuracy of conformers was established along various points in the Glide funnel. For conformers in the Glide funnel, the iRMSD is calculated after superimposing it optimally on the native pose to assess whether a good conformation exists. However, for the intermediate and final docking results, the iRMSD is calculated without the superposition step.

Docking Parameters Optimization for the SP-PEP Algorithm. In order to improve the ability of Glide SP to dock peptides, the following parameters explored, either alone or in combination.

Funnel Width. Four different "funnel widths" were considered. They are defined by the MAXKEEP, MAXREF, and POSTDOCK_NPOSE keywords in Glide. In these tests, the POSTDOCK_NPOSE keyword controls the number of poses that go through post-docking minimization. The default Glide SP funnel employs parameters MAXKEEP, MAXREF, and POSTDOCK_NPOSE with default values of 5000, 400, and 5, respectively. For the "wide", "wider", and "widest" funnels, MAXREF was set to 1000 and POSTDOCK_NPOSE to 100. The MAXKEEP parameter was set to 10000, 20000, and 100000 for the three funnels, respectively. For analysis purposes, in all cases we set POSES_PER_LIG equal to POSTDOCK_NPOSE to get all the post-minimized poses to be reported.

Number of Directions. One of the early steps in docking the ligand consists of orienting the "ligand diameter" along a number of predetermined directions. The default number of directions is 302, and to use a larger number of directions the value of this constant had to be increased. The following numbers of directions were used: 302 (default), 642, 1002, and 1962.

Diameter Rotation Angle. After the ligand diameter is oriented, the angle of rotation around the diameter needs to be determined. Glide does this by rotating the ligand 360° in discrete increments. The following values were used: 25 (the default, 14.4° increments), 50 (7.2°), and 100 (3.6°).

Grid Density. Glide places the center of the ligand on grid points that are normally spaced 1.0 Å apart. In addition to the default value of 1.0, we also use a finer grid density at 0.5 Å. The latter value is referred to as "dense grid".

van der Waals Scaling. The default vdW ligand atom scaling factor in Glide is 0.8 for atoms with an absolute value of partial charge smaller than 0.15, and no scaling for other atoms. We tested ligand scaling factor values of 0.6, 0.4, and 0.2 for the default partial charge cutoff of 0.15 and also for a modified partial-charge cutoff of 1.0. In addition to the ligand scaling described above, it is also possible to scale the vdW radius of the receptor atoms during grid generation. By default, no scaling is used, and we explored the effects of using scaling factors of 0.9 and 0.8. Finally, in a separate experiment, the scaling factor to the ligand−atom vdW radius was applied only during the rough scoring stage. The values for this new scaling factor were 0.2, 0.4, 0.6, and 0.8.

Rigid Docking. As a control experiment, the crystallographic conformation of the native pose was docked to the receptor rigidly. This allows separating the influence of the conformational generation of the ligand (ConfGen) from the actual docking—that is, positioning and orientation—of the ligand.

Inclusion of Crystallographic Conformer. A second control experiment was to include the crystallographic conformation in the calculation along with all of the other conformations that are generated by ConfGen. This allows testing how well rough scoring and subsequent steps can pick the native conformation among non-native conformations.

Binding Site Prediction. We have previously developed a protein−protein interaction (PPI) specific SiteMap mode by optimizing several parameters that control the size, degree of enclosure, and hydrophobic/hydrophilic character of the predicted site.[18] We aimed to optimize ranking and binding site coverage, while minimizing the total size of the predicted pocket, for a set of six well-characterized PPI interfaces.[37] Optimal results were obtained by changing the parameters "enclosure" from 0.5 to 0.4 and "maxvdw" from 1.1 to 0.55. These settings were then used to predict the peptide binding site on the surface of the 19 proteins studied here.

ASSOCIATED CONTENT

Supporting Information

Detailed overview of all parameter values used in the optimization of the SP-PEP algorithm. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: thijs.beuming@schrodinger.com.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Saladin, P.; Zhang, B.; Reichert, J. Current trends in the clinical development of peptide therapeutics. *IDrugs* **2009**, *12*, 779−784.
(2) Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M. Synthetic therapeutic peptides: Science and market. *Drug Discovery Today* **2010**, *15*, 40−56.
(3) Mocellin, S.; Pilati, P.; Nitti, D. Peptide-based anticancer vaccines: Recent advances and future perspectives. *Curr. Med. Chem.* **2009**, *16*, 4779−4796.
(4) McGregor, D. P. Discovering and improving novel peptide therapeutics. *Curr. Opin. Pharmacol.* **2008**, *8*, 616−619.
(5) Huang, C. Receptor-Fc fusion therapeutics, traps, and MIMETIBODY technology. *Curr. Opin. Biotechnol.* **2009**, *20*, 692−699.
(6) Andersen, J. T.; Pehrson, R.; Tolmachev, V.; Daba, M. B.; Abrahmsén, L.; Ekblad, C. Extending half-life by indirect targeting of the neonatal Fc receptor (FcRn) using a minimal albumin binding domain. *J. Biol. Chem.* **2011**, *286*, 5234−5241.
(7) Staneva, I.; Wallin, S. All-atom Monte Carlo approach to protein-peptide binding. *J. Mol. Biol.* **2009**, *393*, 1118−1128.
(8) Bordner, A. J.; Abagyan, R. Ab initio prediction of peptide-MHC binding geometry for diverse class I MHC allotypes. *Proteins* **2006**, *63*, 512−526.
(9) Niv, M. Y.; Weinstein, H. A flexible docking procedure for the exploration of peptide binding selectivity to known structures and homology models of PDZ domains. *J. Am. Chem. Soc.* **2005**, *127*, 14072−14079.
(10) Antes, I. DynaDock: A new molecular dynamics-based algorithm for protein-peptide docking including receptor flexibility. *Proteins* **2010**, *78*, 1084−1104.
(11) Raveh, B.; London, N.; Schueler-Furman, O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins* **2010**, *78*, 2029−2040.
(12) Raveh, B.; London, N.; Zimmerman, L.; Schueler-Furman, O. Rosetta FlexPepDock ab-initio: Simultaneous folding, docking and refinement of peptides onto their receptors. *PloS One* **2011**, *6*, e18934.
(13) London, N.; Movshovitz-Attias, D.; Schueler-Furman, O. The structural basis of peptide−protein binding strategies. *Structure* **2010**, *18*, 188−199.
(14) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(15) Donsky, E.; Wolfson, H. J. PepCrawler: A fast RRT-based algorithm for high-resolution refinement and binding affinity estimation of peptide inhibitors. *Bioinformatics* **2011**, *27*, 2836−2842.
(16) Trellet, M.; Melquiond, A. S.; Bonvin, A. M. A unified conformational selection and induced fit approach to protein−peptide docking. *PloS One* **2013**, *8*, e58769.
(17) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739−1749.
(18) *SiteMap*, version 2.7; Schrödinger, LLC: New York, 2012.
(19) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 235−249.
(20) Zhou, Z.; Felts, A. K.; Friesner, R. A.; Levy, R. M. Comparative performance of several flexible docking programs and scoring functions: enrichment studies for a diverse set of pharmaceutically relevant targets. *J. Chem. Inf. Model.* **2007**, *47*, 1599−1608.
(21) Watts, K. S.; Dalal, P.; Murphy, R. B.; Sherman, W.; Friesner, R. A.; Shelley, J. C. ConfGen: A conformational search method for efficient generation of bioactive conformers. *J. Chem. Inf. Model.* **2010**, *50*, 534−546.
(22) Chen, I. J.; Foloppe, N. Drug-like bioactive structures and conformational coverage with the LigPrep/ConfGen suite: comparison to programs MOE and catalyst. *J. Chem. Inf. Model.* **2010**, *50*, 822−839.
(23) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory and Comput.* **2010**, *6*, 1509−1519.
(24) Jorgensen, W. L.; Tirado-Rives, J. The OPLS potential function for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(25) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development andtesting of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(26) Feher, M.; Williams, C. I. Numerical errors and chaotic behavior in docking simulations. *J. Chem. Inf. Model.* **2012**, *52*, 724−738.
(27) Labute, P. LowModeMD: Implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops. *J. Chem. Inf. Model.* **2010**, *50*, 792−800.
(28) *Prime*, version 3.1; Schrödinger, LLC: New York, 2012.
(29) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534−553.
(30) Thevenet, P.; Shen, Y.; Maupetit, J.; Guyon, F.; Derreumaux, P.; Tuffery, P. PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Res.* **2012**, *40*, W288−293.
(31) Kaur, H.; Garg, A.; Raghava, G. P. PEPstr: A de novo method for tertiary structure prediction of small bioactive peptides. *Protein Peptide Lett.* **2007**, *14*, 626−631.
(32) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377−389.
(33) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146−148.
(34) Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **2003**, *10*, 980.
(35) *Maestro*, version 9.3; Schrödinger, LLC: New York: 2012.
(36) Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A. The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins* **2011**, *79*, 2794−2812.
(37) Wells, J. A.; McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450*, 1001−1009.