

Supplementary Materials

Rescoring docking hit lists for model cavity sites: predictions and experimental testing

Alan P. Graves, Devleena M. Shivakumar, Sarah E. Boyce, Matthew P. Jacobson, David A.
Case, and Brian K. Shoichet

Contents:

1. PLOP side chain rotamer search and minimization algorithm.
2. AMBERDOCK parameters and optimization.
3. Tables S1-S9.
4. Figures S1-S3.

PLOP side chain rotamer search and minimization algorithm. The side chain rotamer search algorithm proceeds by optimizing one side chain at a time (keeping the others fixed), and keeps looping through the residues until they stop changing conformation. During the side chain rotamer search, the default resolution of 10° for the rotamer library was used, and the initial side chain conformations were utilized by specifying the parameter “randomize no”. The initial (ofac_init) and minimal (ofac_min) overlap factors, which define steric clashes, were set to 0.75 and 0.5, respectively. The calculation of the complex energy ($E_{complex}$) involved an initial short minimization of the ligand with a maximum number of steps (mxitn) set to 15 followed by the sidechain rotamer search, minimization of residues with a maximum number of steps (mxitn) set to 30, and finally minimization of ligand to convergence. Minimization and calculation of the free receptor energy ($E_{receptor}$) began with the initial side chain rotamer search followed by the minimization of residues with a maximum number of steps (mxitn) set to 30.

AMBERDOCK parameters and optimization. The traditional AMBER force field has been parameterized to work with biological macromolecules, and has limited parameters for organic molecules. The general AMBER force field (GAFF), which has been specifically designed to cover most pharmaceutical molecules, is compatible with the existing AMBER force fields in such a way that the two can be mixed during a simulation. It uses a simple harmonic energy function as the additive AMBER force fields, but the atom types used in GAFF are much more general such that they cover most of the organic chemical space. The current implementation of the GAFF force field consists of 33 basic atom types and 22 special atom types covering almost all of the organic chemical space that is made up of C, N, O, S, P, H, F, Cl, Br and I atoms. The input ligand files for AMBERDOCK can be generated automatically using the perl script (prepare_amber.pl) provided with the DOCK6.1 program. This perl script calls for programs such as *antechamber* to calculate the charges for the ligands and tleap to assign the parameter set for protein and ligand atoms. The newer version of the DOCK6 program

bypasses this external perl step, and the input files are generated internally when the AMBER score routine is called.

Binding free energy calculations with AMBERDOCK follows the flowchart in Figure S1. In Step 1, the input files (pdb, inpcrd, prmtop) in AMBER format are read. In Step 2, the frozen atoms are defined based on the choice of the user. The user can specify the atoms that are allowed to relax in the DOCK6 input file (dock.in). A brief minimization is carried out using the GB implicit solvent model in Step 2. In Step 3, the molecular dynamics (MD) simulation is carried out on the minimized system. In step 4, a brief minimization is carried out to minimize the structure generated from the previous MD step. Steps 2-4 are performed with the frozen atom approximation in GB implicit solvent model. When frozen atoms are specified, the NAB program calculates the energy for only those atoms that are allowed to move. This significantly speeds up the calculations and uses less memory. In Step 5, a single point energy is calculated on the structure obtained from Step 4. In this step, the energy of the full system is calculated without any frozen atoms and non-bonded cutoffs. Also, the surface area term is added to this final step to get a more accurate energy term. The Steps 1-5 were repeated for complex ($E_{complex}$), ligand (E_{ligand}) and receptor ($E_{receptor}$). To expedite the scoring process, we calculated the energy of the receptor for the first ligand and used this energy as a constant term during the energy evaluations for the rest of the ligands in the database. It is suggested to repeat these steps with various combinations of molecular mechanics (MM) options to optimize the variables.

Several scoring protocols were tested on the 58 known ligands and 17 known decoys for L99A/M102Q to find the optimal set of conditions for AMBERDOCK rescoring. The initial attempt was to use AMBERDOCK with only minimization and no molecular dynamics simulation. For example, *score1* (Figure S2b) involved scoring the test set with Gasteiger-Marselli PEOE charges for the ligands and AMBER parm 94 parameter set for protein atoms; the Hawkins-Cramer-Truhlar pairwise GB model (equivalent to $igb=1$ in AMBER); the LCPO

method to calculate the surface area term; and minimization of nine binding site residues (78, 84, 88, 91, 102, 111, 118, 121, and 153). The results clearly show that the minimization alone is not sufficient to separate out the ligands from decoys with several known decoys scoring better than known ligands. Infact, the results from DOCK (Figure S2a) is adversely affected by running *score1*. We also tried rescoring the same test set with a series of AMBERDOCK scores with different combinations of the molecular mechanics input parameters, using large or small flexible receptor regions, GB models (*igb*=1, 2, 5) with or without the surface area term, number of minimization steps, etc. We found that all of them predicted several decoys among the top 20 scored ligands with only minimization of the system, and no molecular dynamics (MD) simulations. As a logical next step, we introduced MD simulations along with minimization in the system. The initial receptor-ligand configuration obtained from docking was subjected to a few steps of minimization followed by a few picoseconds of MD simulations, and then a final minimization run to obtain the total energy. Selection of charge models is also very important. The AMBERDOCK scores obtained when using the AM1-BCC charge model provided better results compared to the Gasteiger-Marselli PEOE charge model. Using *score24G* with Gasteiger-Marselli PEOE charges with MD and minimization, 6 decoys scored as well as the top 20 ligands (Figure S2c). Using *score24* with AM1-BCC charges with MD and minimization, only 1 decoy scored as well as the top 20 scoring ligands (Figure S2d). This scoring protocol produced good results for most of our test sets of known ligands and decoys compared to the other scoring protocols that we tested. The specifications of *score24* are: GB model of 5 (corresponding to *igb*=5 in the AMBER program); surface area term calculated using the LCPO model; a non-bonded cutoff of 18 Å; 100 steps of conjugate gradient minimization; and 3000 steps of MD simulation with a 1 fs time step at a temperature of 300K, followed by 100 steps of minimization. During the minimization and MD, only the ligand and the protein residues within 5 Å of the ligand were allowed to move.

Table S1. Top 100 hits predicted by DOCK3.5.54 for L99A.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S1 (continued). Top 100 hits predicted by DOCK3.5.54 for L99A.

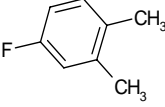
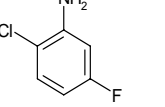
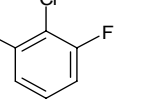
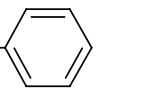
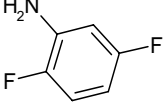
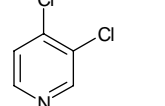
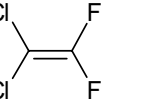
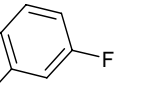
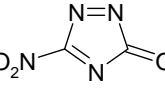
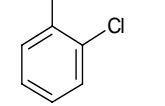
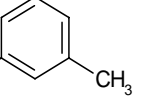
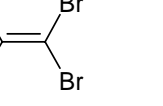
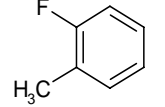
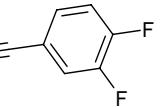
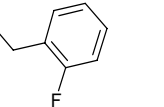
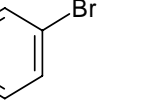
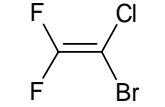
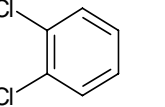
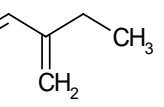
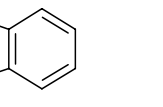
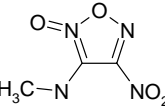
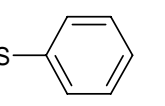
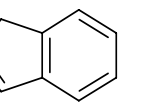
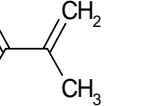
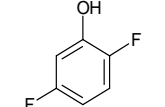
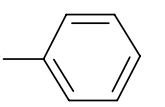
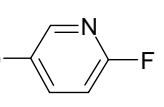
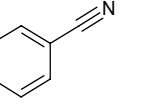
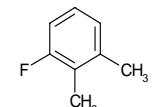
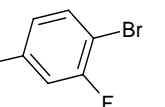
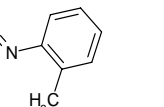
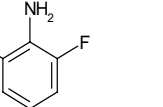
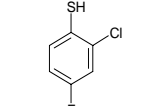
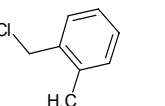
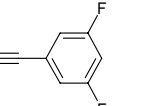
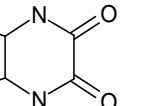
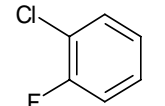
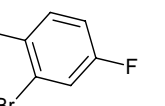
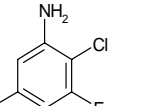
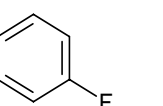
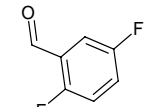
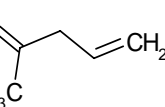
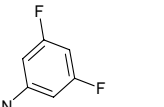
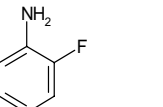
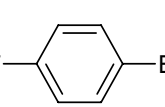
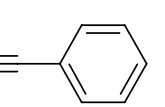
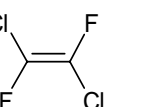
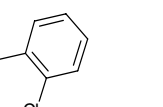
#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S2. Top 100 hits predicted by PLOP for L99A.

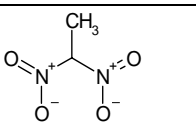
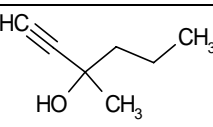
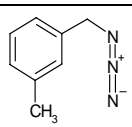
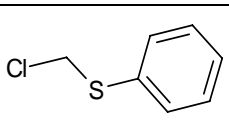
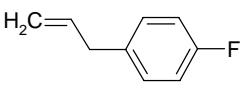
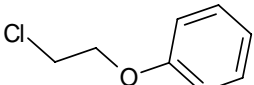
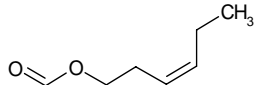
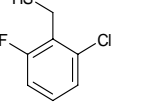
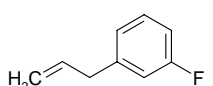
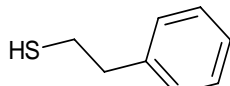
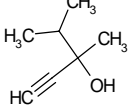
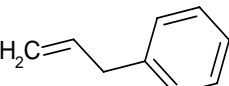
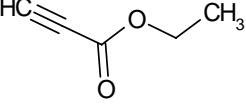
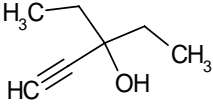
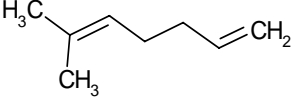
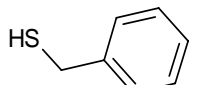
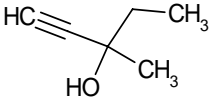
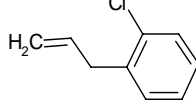
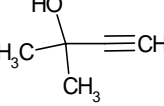
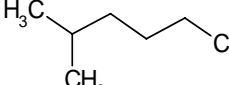
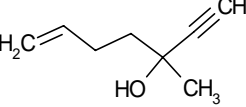
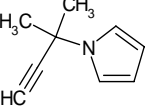
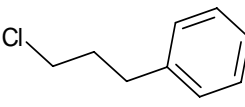
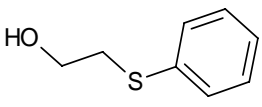
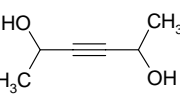
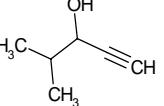
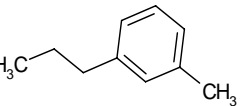
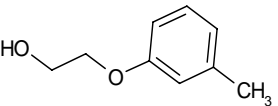
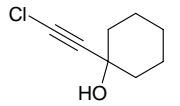
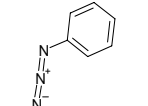
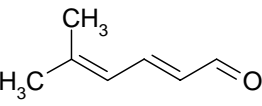
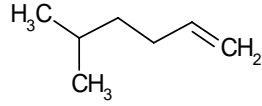
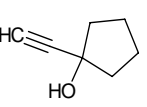
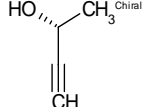
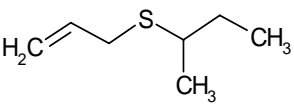
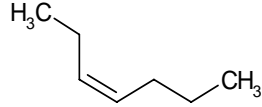
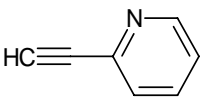
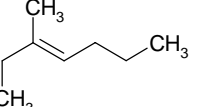
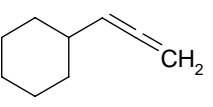
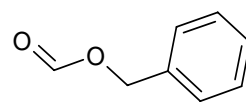
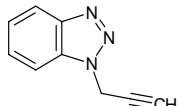
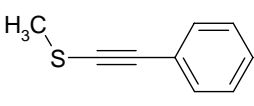
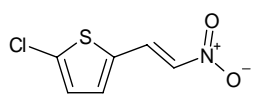
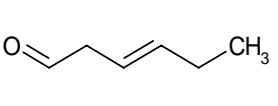
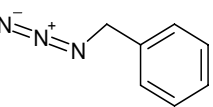
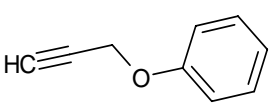
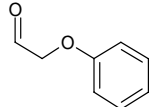
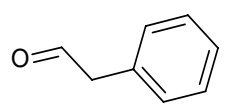
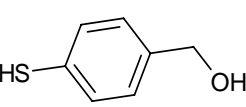
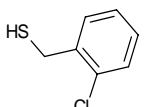
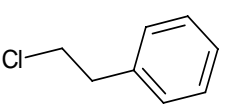
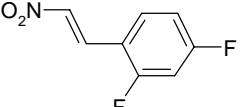
#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S2 (continued). Top 100 hits predicted by PLOP for L99A.

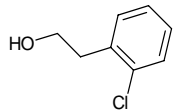
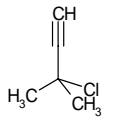
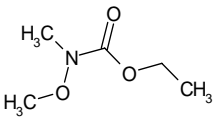
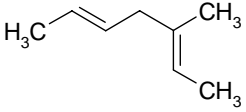
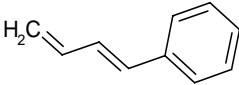
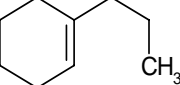
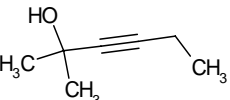
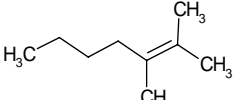
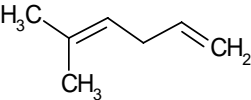
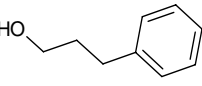
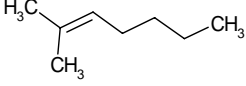
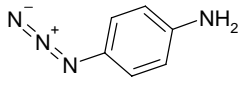
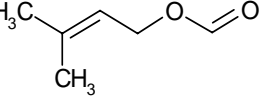
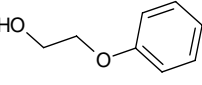
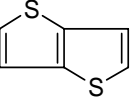
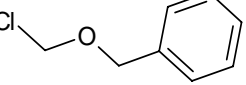
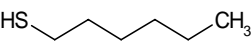
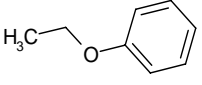
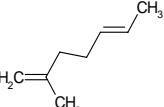
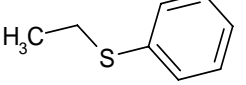
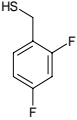
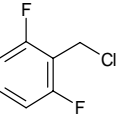
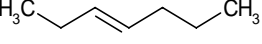
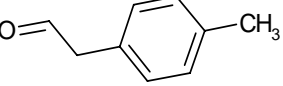
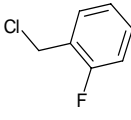
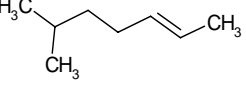
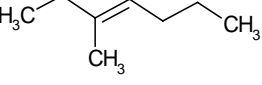
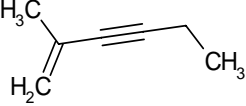
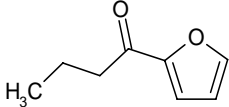
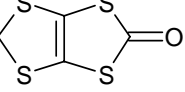
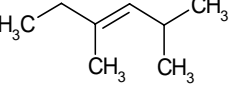
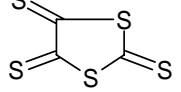
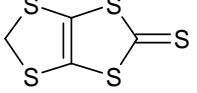
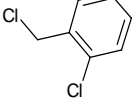
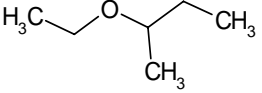
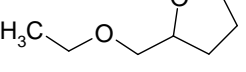
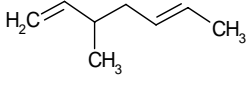
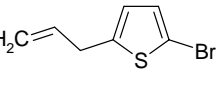
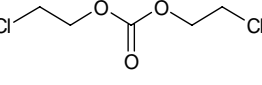
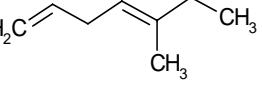
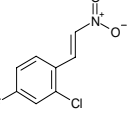
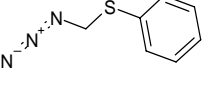
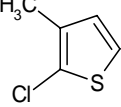
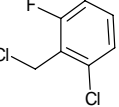
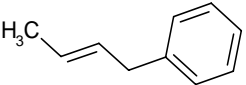
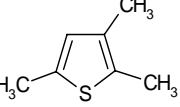
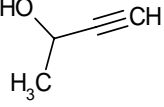
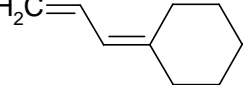
#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S3. Top 100 hits predicted by AMBERDOCK for L99A.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table 3S (continued). Top 100 hits predicted by AMBERDOCK for L99A.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S4. Top 100 hits predicted by DOCK3.5.45 for L99A/M102Q.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S4 (continued). Top 100 hits predicted by DOCK3.5.45 for L99A/M102Q.

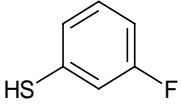
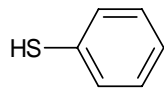
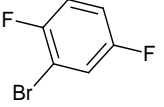
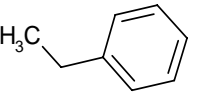
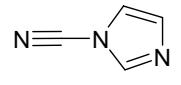
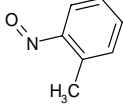

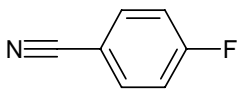
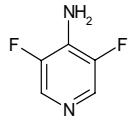
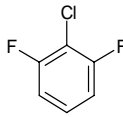
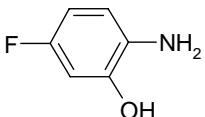
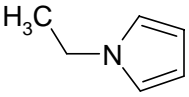
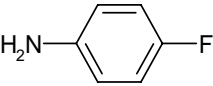
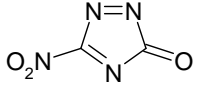
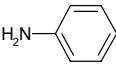
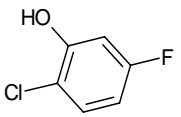
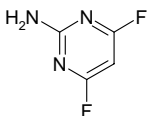
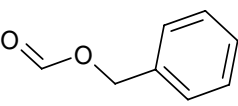
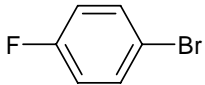
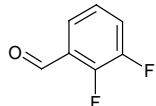
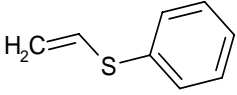
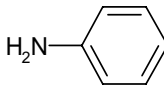
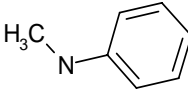
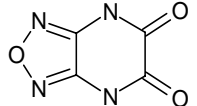
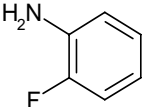
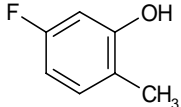
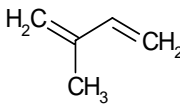
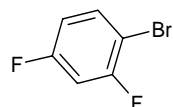
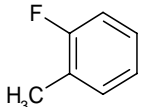
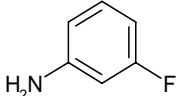
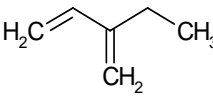
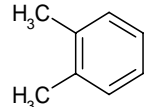
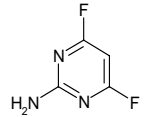
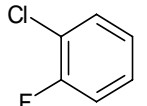
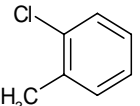
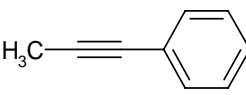
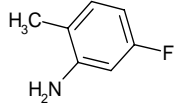
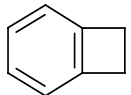
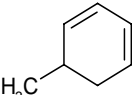
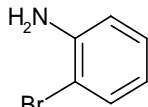
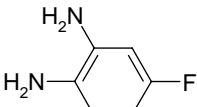
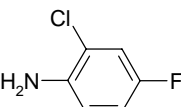
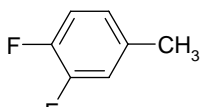
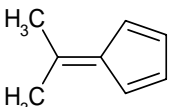
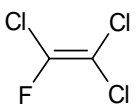
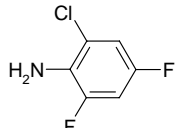
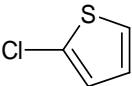
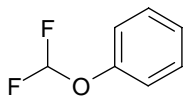
#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S5. Top 100 hits predicted by PLOP for L99A/M102Q.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S5 (continued). Top 100 hits predicted by PLOP for L99A/M102Q.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S6. Top 100 hits predicted by AMBERDOCK for L99A/M102Q.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S6 (continued). Top 100 hits predicted by AMBERDOCK for L99A/M102Q.

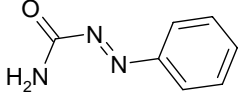
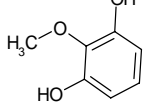
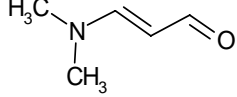
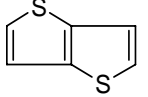
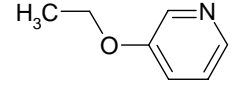
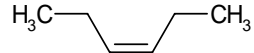
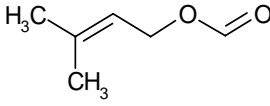
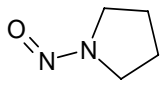
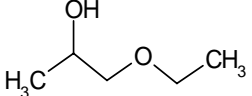
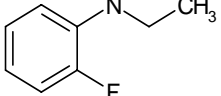
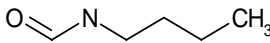
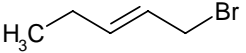
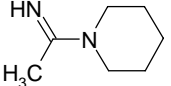
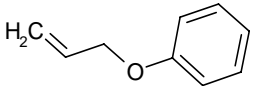
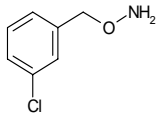
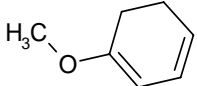
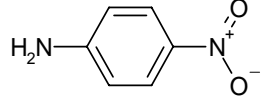
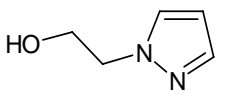
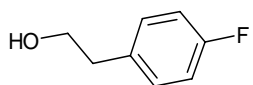
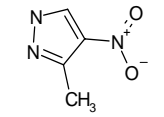
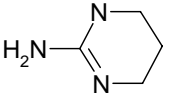
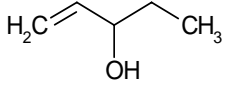
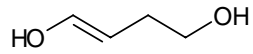
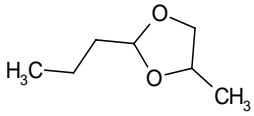
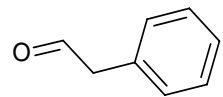
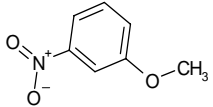
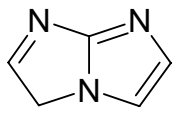
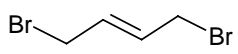
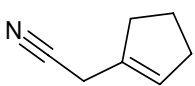
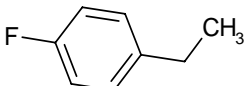
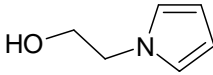
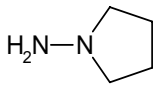
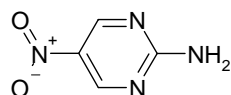
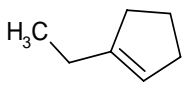
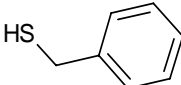

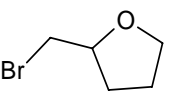
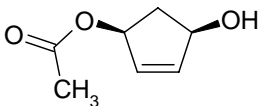
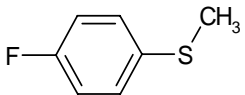
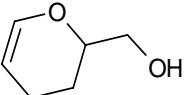
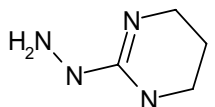
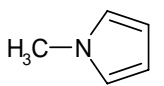
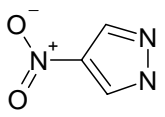
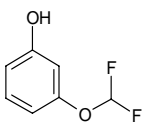
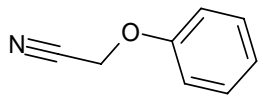
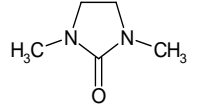
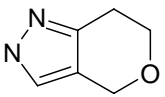
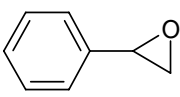
#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S7. Top 100 hits predicted by DOCK3.5.54 for CCP.

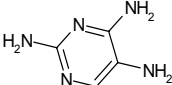
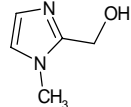
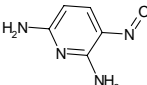
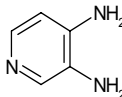
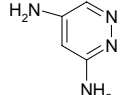
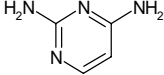
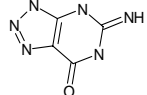
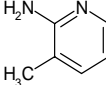
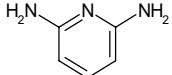
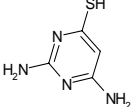
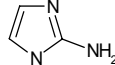
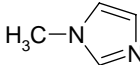
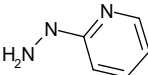
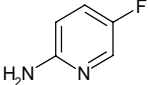
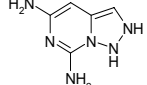
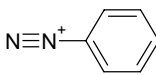
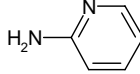
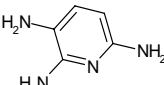
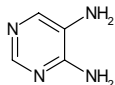
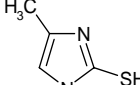
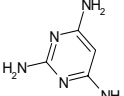
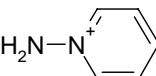
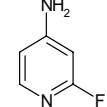
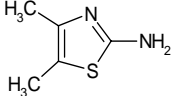
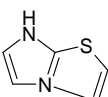
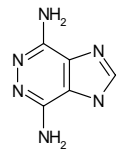
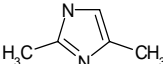
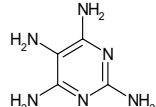
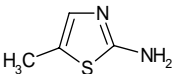
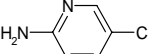
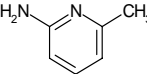
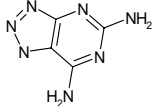
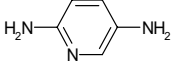
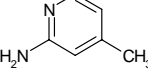
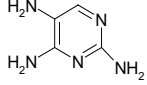
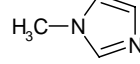
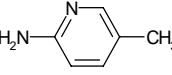
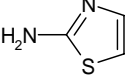
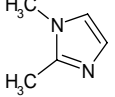
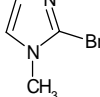
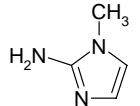
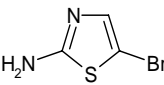
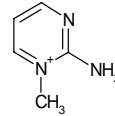
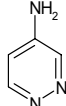
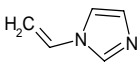
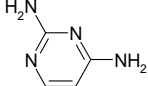
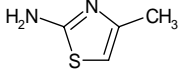
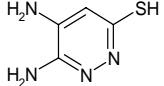
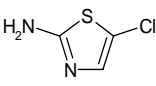
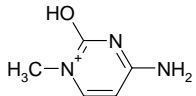
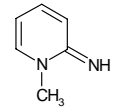
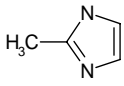
#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S7 (continued). Top 100 hits predicted by DOCK3.5.54 for CCP.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S8. Top 100 hits predicted by PLOP for CCP.

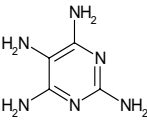
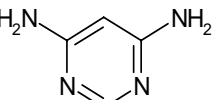
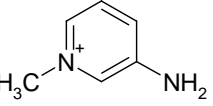
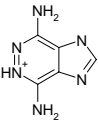
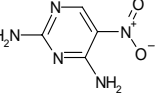
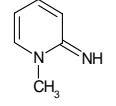
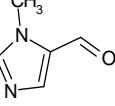
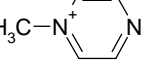
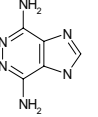
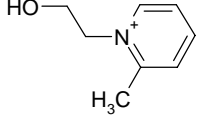
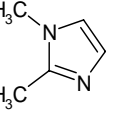
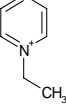
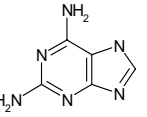
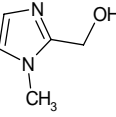
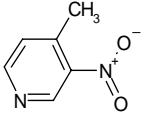
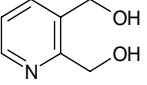
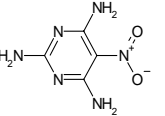
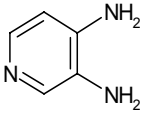
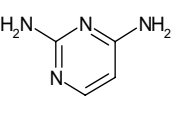
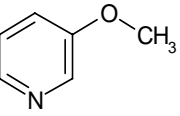
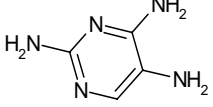
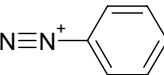
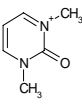
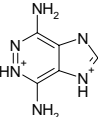
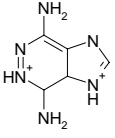
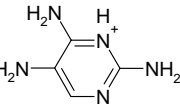
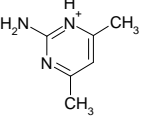
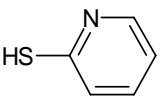
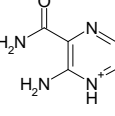
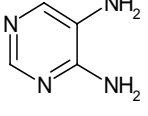
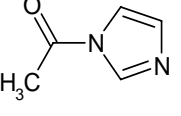
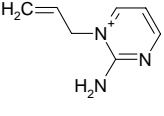
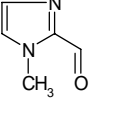
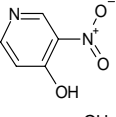
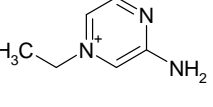
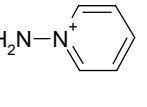
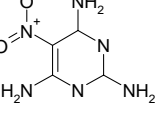
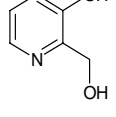
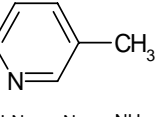
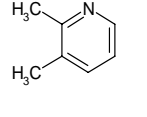
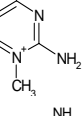
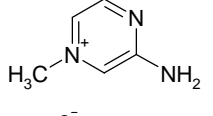
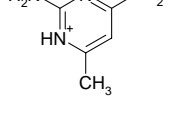
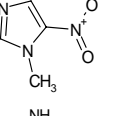
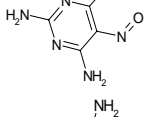
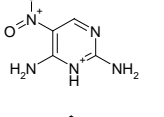
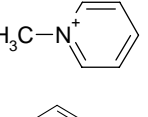
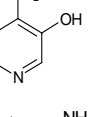
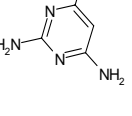
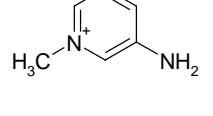
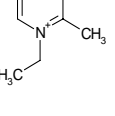
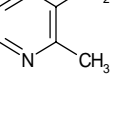
#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S8 (continued). Top 100 hits predicted by PLOP for CCP.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Table S9. Top 100 hits predicted by AMBERDOCK for CCP.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
1		14		27		40	
2		15		28		41	
3		16		29		42	
4		17		30		43	
5		18		31		44	
6		19		32		45	
7		20		33		46	
8		21		34		47	
9		22		35		48	
10		23		36		49	
11		24		37		50	
12		25		38		51	
13		26		39		52	

Table S9 (continued). Top 100 hits predicted by AMBERDOCK for CCP.

#	STRUCTURE	#	STRUCTURE	#	STRUCTURE	#	STRUCTURE
53		65		77		89	
54		66		78		90	
55		67		79		91	
56		68		80		92	
57		69		81		93	
58		70		82		94	
59		71		83		95	
60		72		84		96	
61		73		85		97	
62		74		86		98	
63		75		87		99	
64		76		88		100	

Figure S1. Distribution of pair-wise Tanimoto similarities among the top-ranked docked and rescored molecules, calculated using ECFP4 fingerprints.

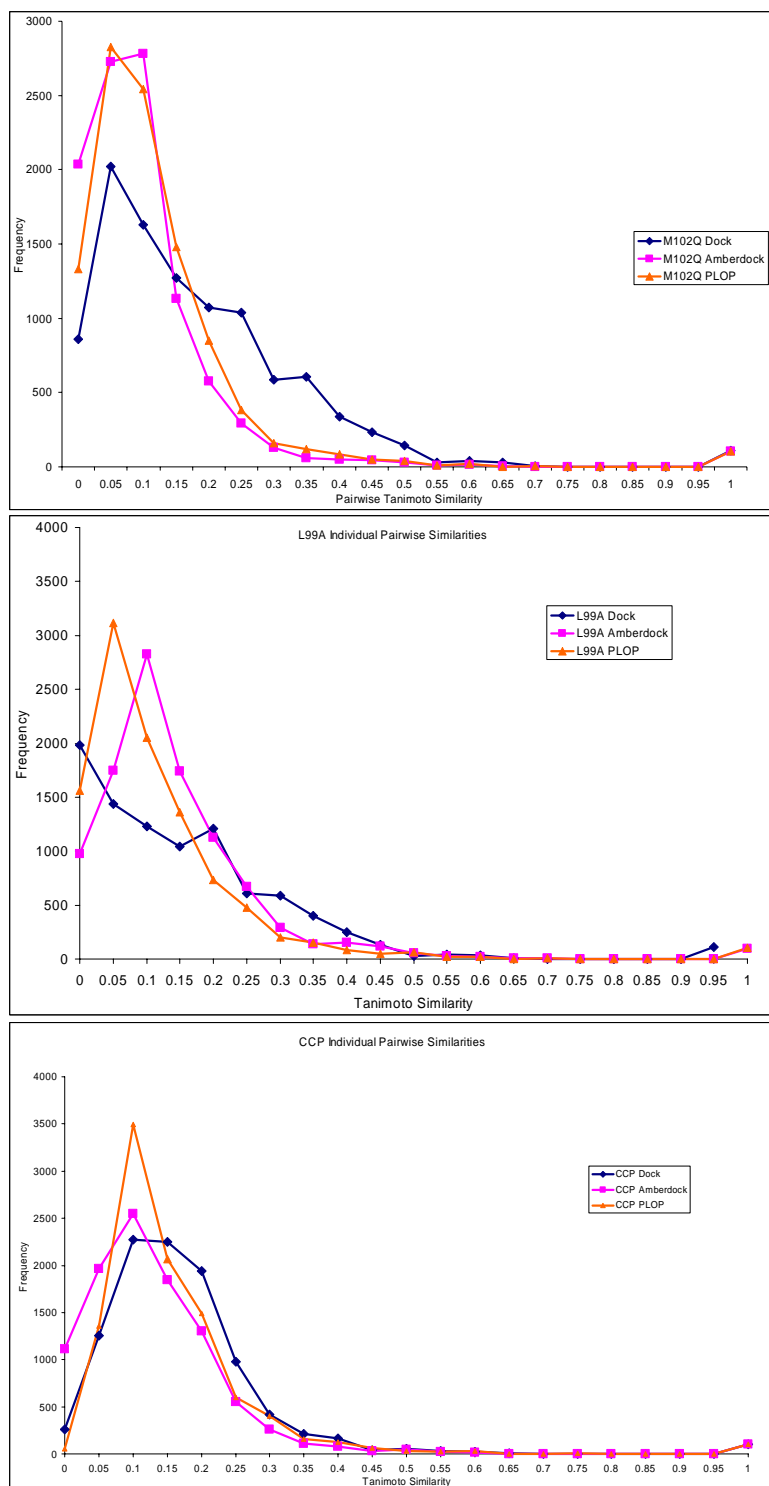
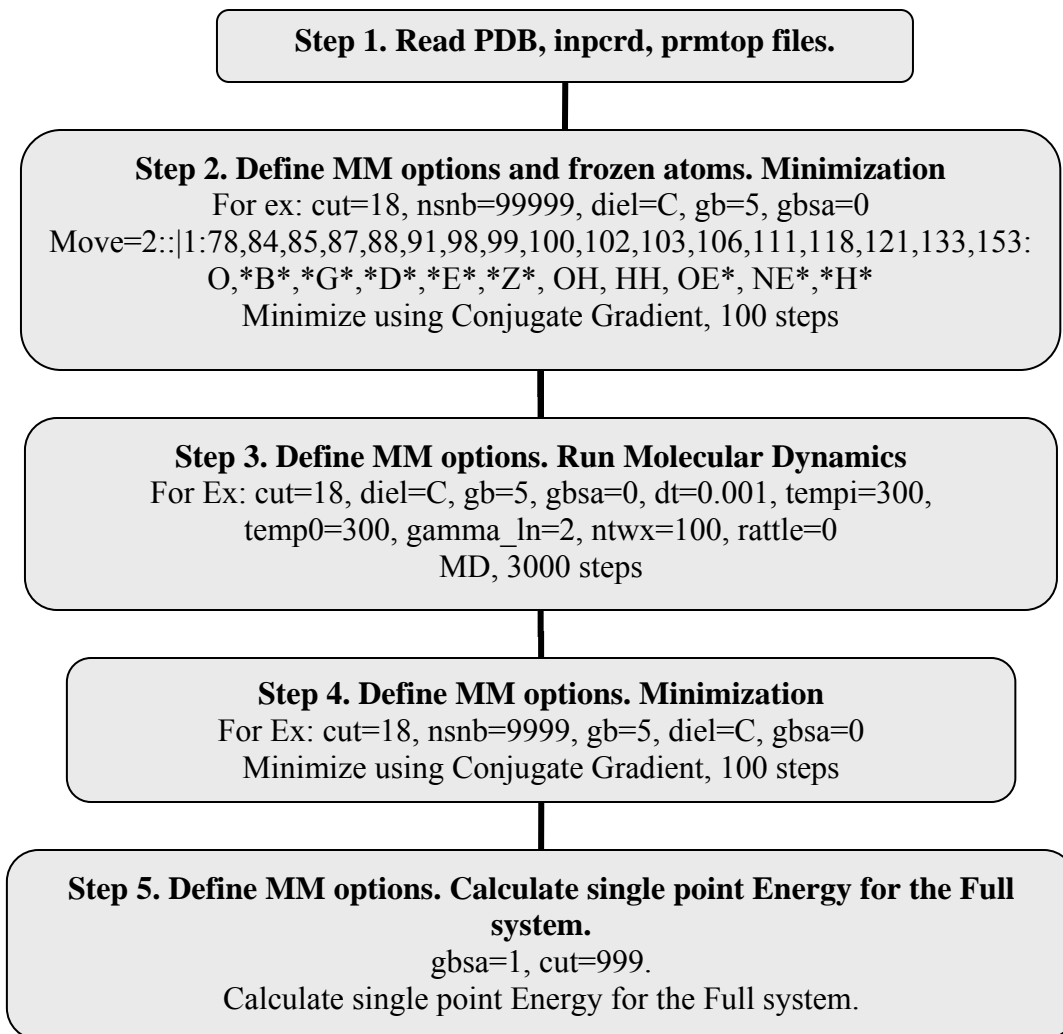


Figure S2. Flowchart and parameters for AMBERDOCK rescoring.



Description of the keywords: gb=2, Onufriev, Bashford, Case (OBC) variant of GB, with $\alpha=0.8$, $\beta=0.0$, $\gamma=2.909$. Similar to igb=2 in AMBER.; gb=5, with $\alpha=1.0$, $\beta=0.8$, $\gamma=4.85$; diel=C, uses constant dielectric; dt, time step, ps; tautp, temperature coupling parameter, in ps; gamma_ln, collision frequency for Langevin dynamics, in ps⁻¹; temp0, target temperature; gbsa, add a surface-area dependent energy.

Figure S2. Preliminary scoring protocols for AMBERDOCK. The known L99A/M102Q ligands (blue bars) and decoys (red bars) are plotted versus the respective score (x -axis). A. DOCK score. B. AMBERDOCK *score1* with minimization only. This uses PEOE charges, $gb=1$, $gbsa=1$, and moveable residues=(78, 84, 88, 91, 102, 111, 118, 121, 153). C. AMBERDOCK *score24G* with MD and minimization. The scoring protocol includes 3000 steps of MD with a moveable region of 5 Å from the ligand, and $gb=5$. D. AMBERDOCK *score24* with MD and minimization. This is the same as *score24G*, except ligand charges were calculated at the AM1-BCC level of theory.

