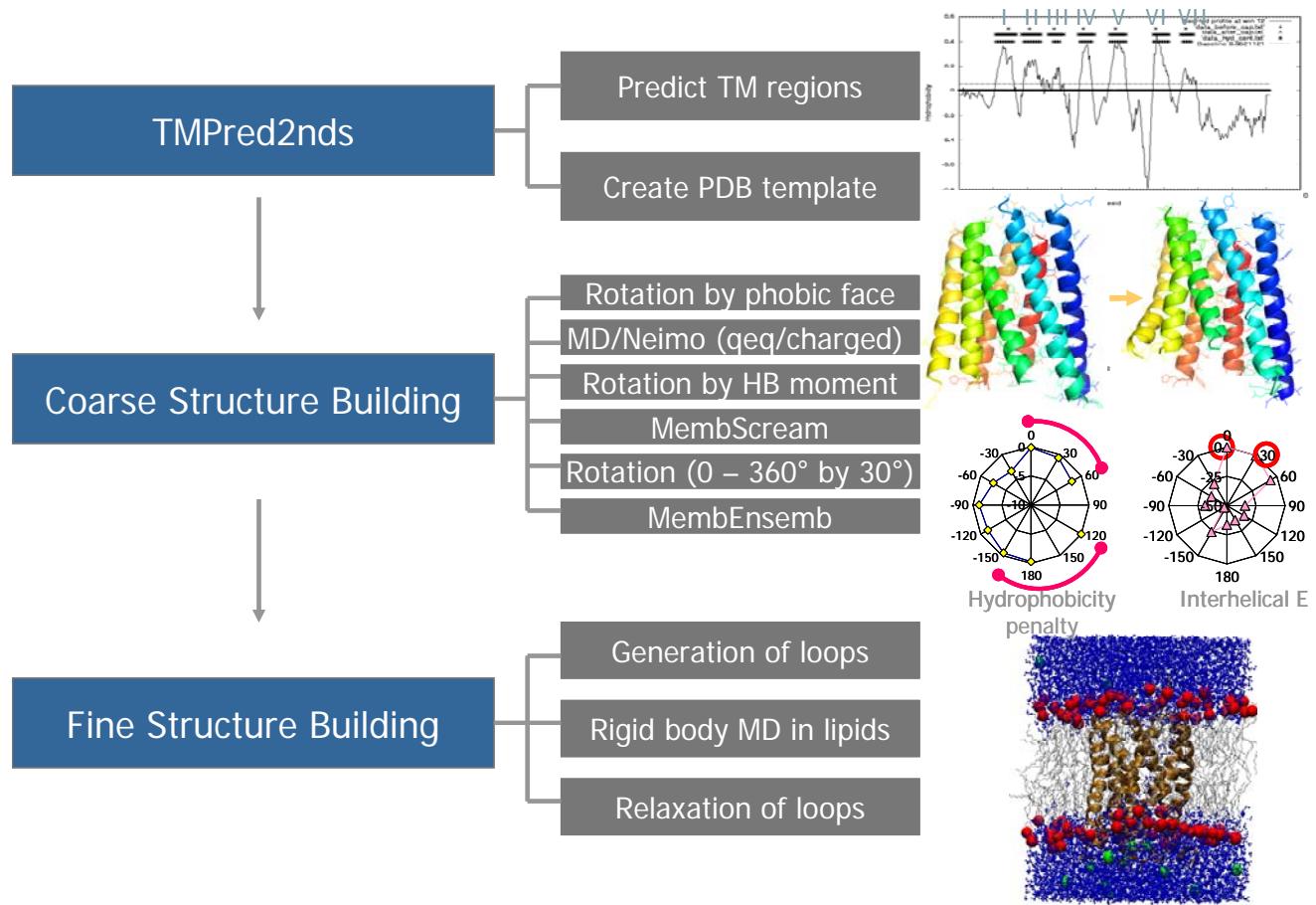
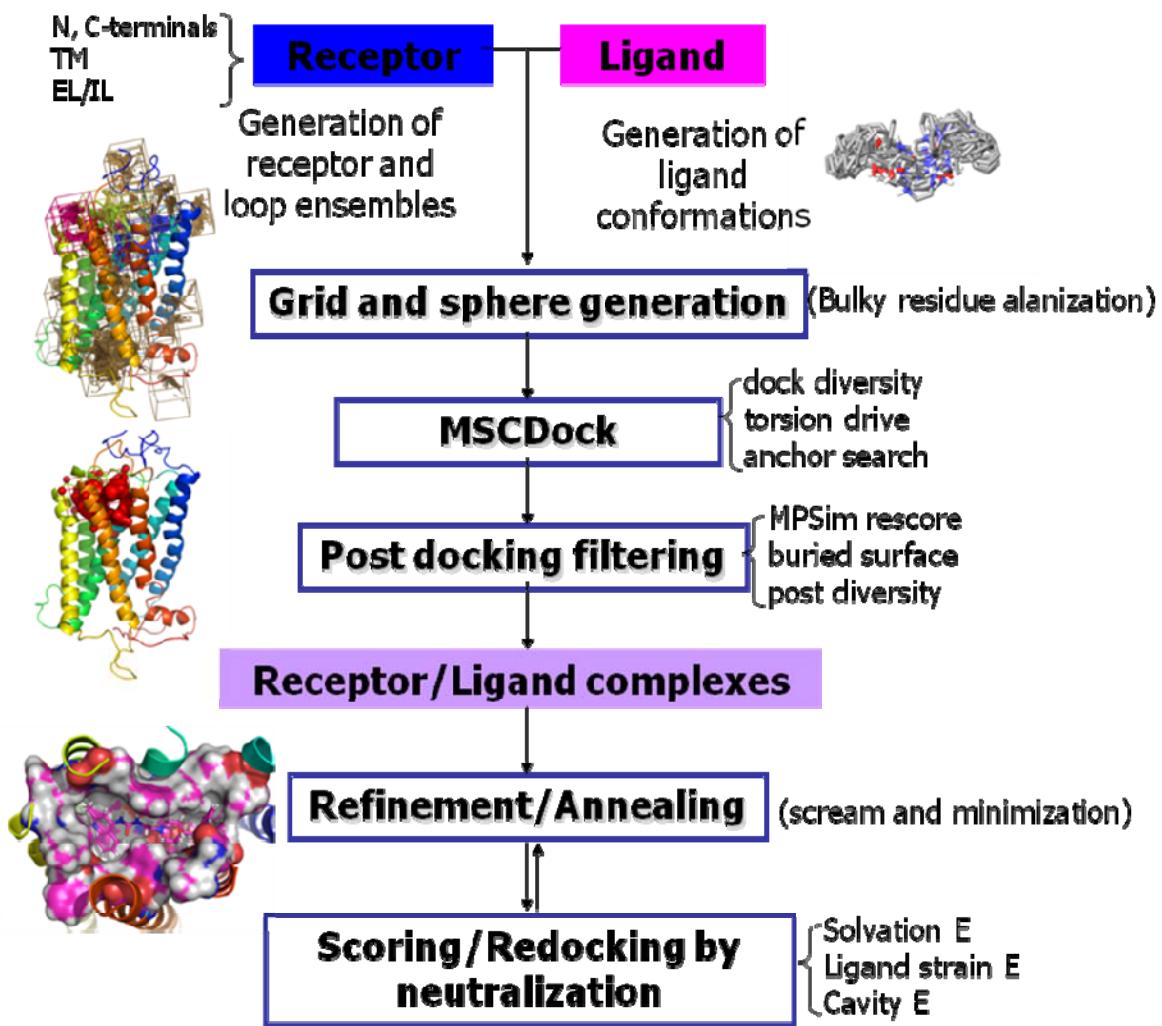


## **Supporting Information**

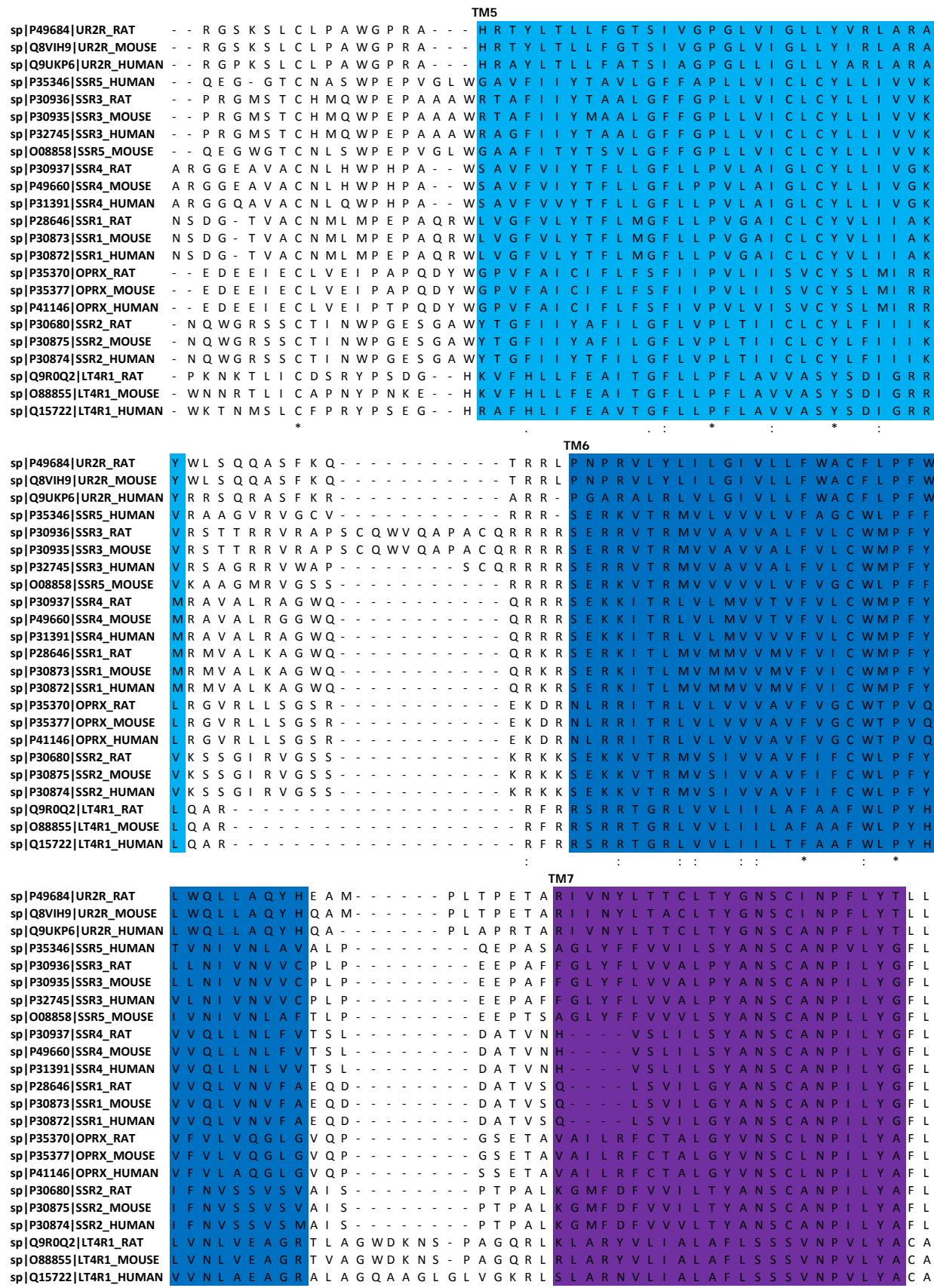


**Figure A-1.** MembScream Algorithm

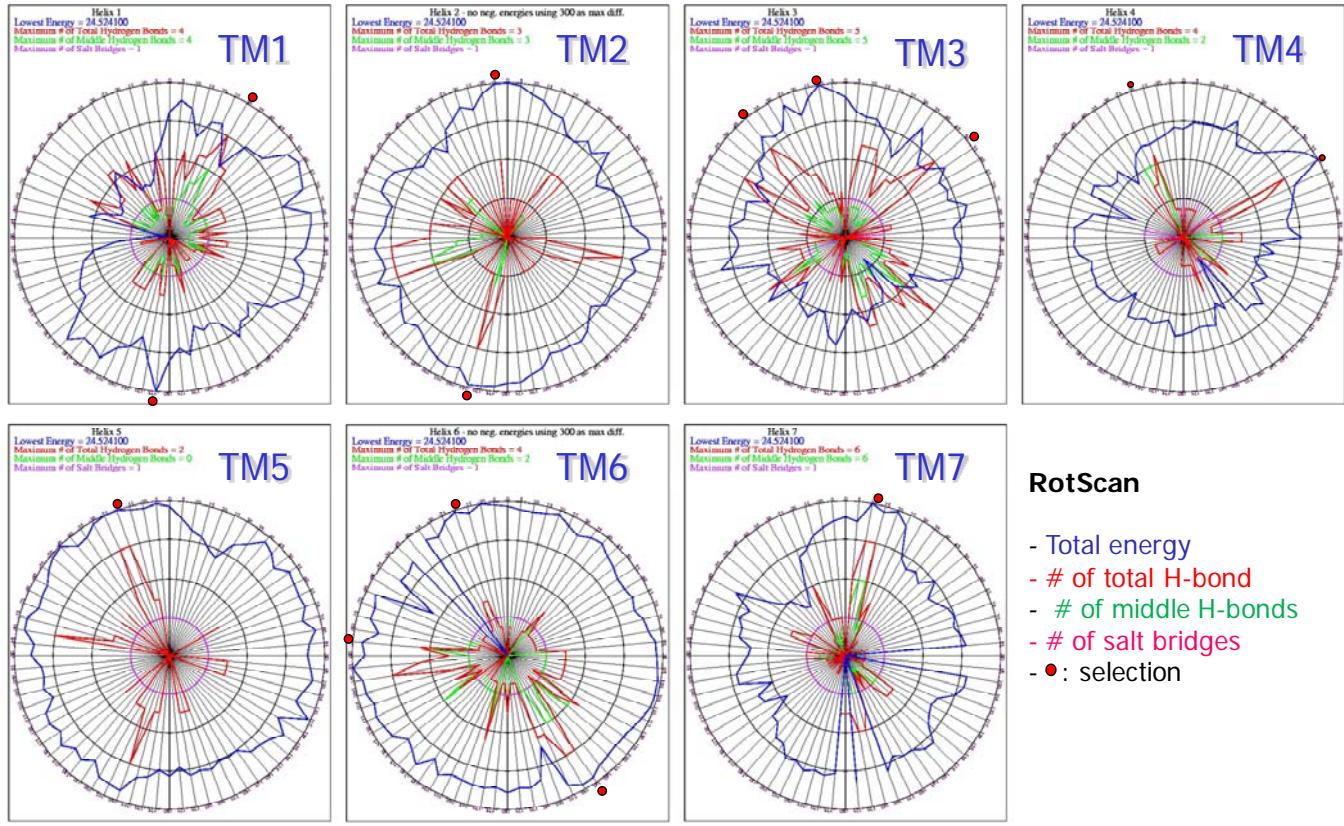


**Figure A-2.** The outline of the MSCDock procedure.

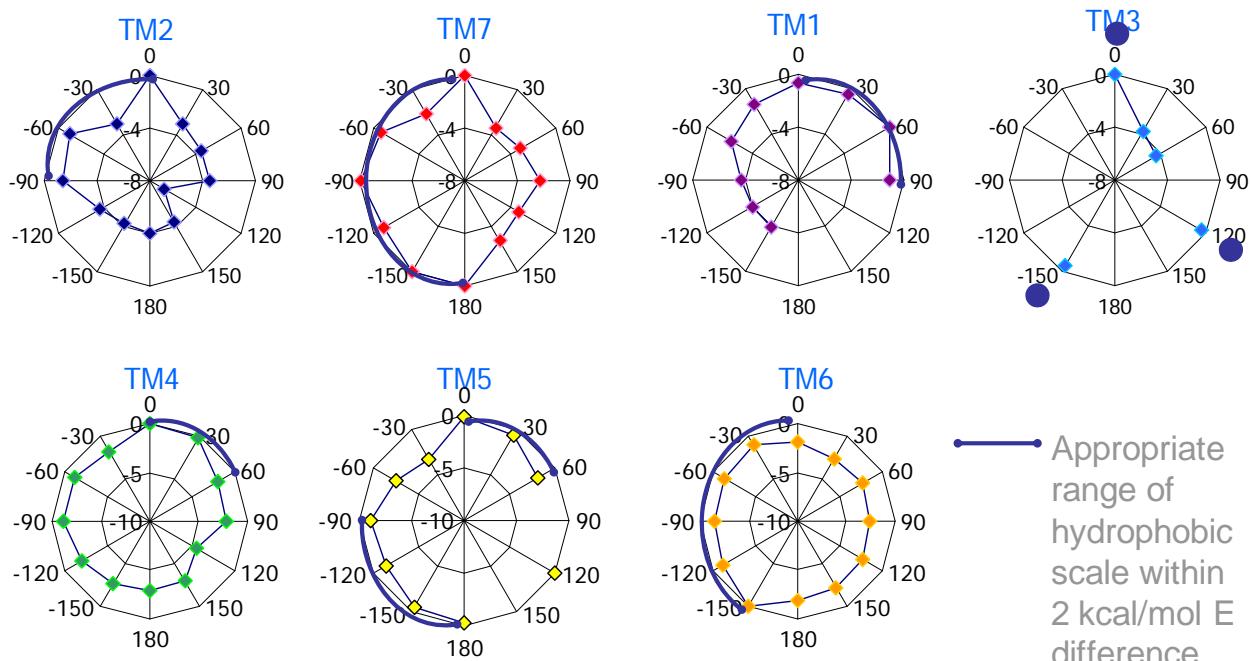
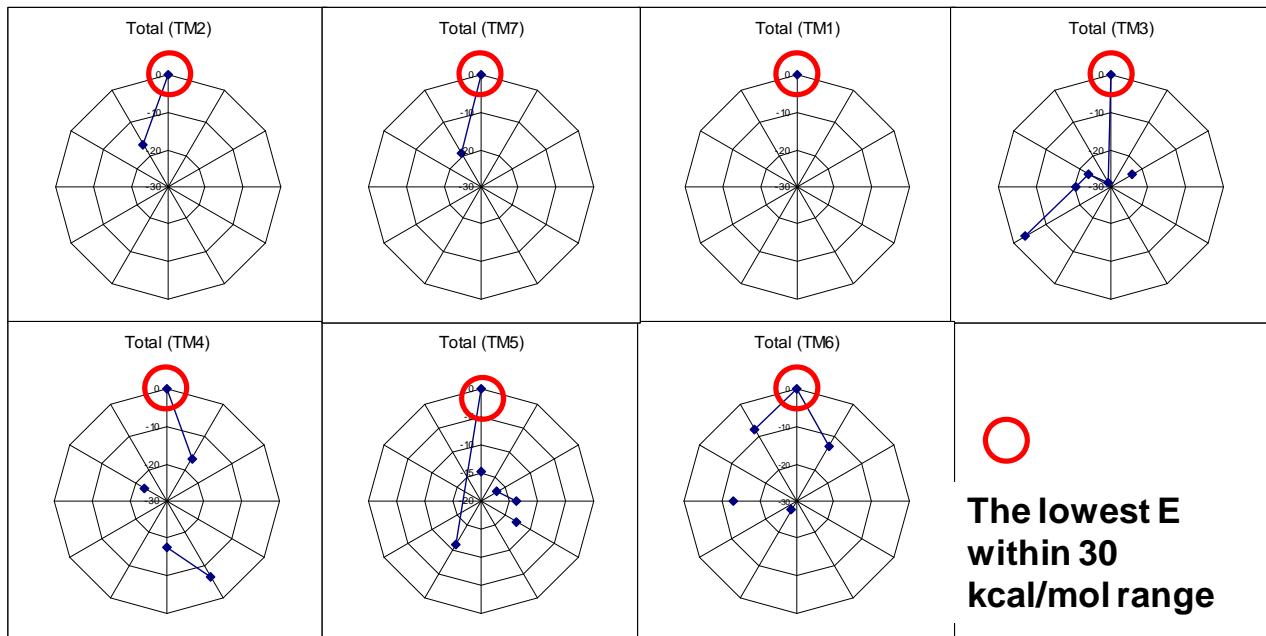
		TM1	TM2	
sp P49684 UR2R_RAT	S L K	D L V A T G V I G A V L S A M G V V G M V G N V Y T L V V M C R F	F L R A S A S M Y V Y V V N L	
sp Q8VIH9 UR2R_MOUSE	S L Q	D L V A T G V I G A V L S T M G V V G V V G N V Y T L V V M C R F	F L R A S A S M Y V Y V V N L	
sp Q9UKP6 UR2R_HUMAN	S L E	D L V A T G T I G T L L S A M G V V G V G N A Y T L V V T C R S L R A V A S M Y V Y V V N L		
sp P35346 SSR5_HUMAN	A P S	A G A R A V L V P V L Y L L V C A A G L G G N T L V I Y V V L R F A K M K	T V T N I Y L N L	
sp P30936 SSR3_RAT	L A G L	A V S G I L I S L V Y L V V C V V G L L G N S L V I Y V V L R H T S S P S V T S V Y I L N L		
sp P30935 SSR3_MOUSE	L T G L	A V S G I L I S L V Y L V V C V V G L L G N S L V I Y V V L R H T S S P S V T S V Y I L N L		
sp P32745 SSR3_HUMAN	P A G L	A V S G V L I P L V Y L V V C V V G L L G N S L V I Y V V L R H T A S P S V T N V Y I L N L		
sp O08858 SSR5_MOUSE	V S P	M G A R A V L V P V L Y L L V C T V G L G G N T L V I Y V V L R Y A K M K	T V T N V Y I L N L	
sp P30937 SSR4_RAT	G - -	- T A G M V T I Q C I Y A L V C L V G L V G N A L V I F V I L R Y A K M K	T A T N I Y L L N L	
sp P49660 SSR4_MOUSE	G - -	- T A G M V T I Q C I Y A L V C L V G L V G N A L V I F V I L R Y A K M K	T A T N I Y L L N L	
sp P31391 SSR4_HUMAN	A R -	- A A G M V A I Q C I Y A L V C L V G L V G N A L V I F V I L R Y A K M K	T A T N I Y L L N L	
sp P28646 SSR1_RAT	L S E	G Q G S A I L I S F I Y S V V C L V G L C G N S M V I Y V I L R Y A K M K	T A T N I Y I L N L	
sp P30873 SSR1_MOUSE	L S E	G Q G S A I L I S F I Y S V V C L V G L C G N S M V I Y V I L R Y A K M K	T A T N I Y I L N L	
sp P30872 SSR1_HUMAN	L S E	G Q G S A I L I S F I Y S V V C L V G L C G N S M V I Y V I L R Y A K M K	T A T N I Y I L N L	
sp P35370 OPRX_RAT	A F L	P L G L K V T I V G L Y L A V C I G G L L G N C L V M Y V I L R H T K M K	T A T N I Y I F N L	
sp P35377 OPRX_MOUSE	A F L	P L G L K V T I V G L Y L A V C V G G L L G N C L V M Y V I L R H T K M K	T A T N I Y I F N L	
sp P41146 OPRX_HUMAN	A F L	P L G L K V T I V G L Y L A V C V G G L L G N C L V M Y V I L R H T K M K	T A T N I Y I F N L	
sp P30680 SSR2_RAT	Y Y D	M T S N A V L T - F I Y F V V C V V G L C G N T L V I Y V I L R Y A K M K	T I T N I Y I L N L	
sp P30875 SSR2_MOUSE	Y Y D	M T S N A V L T - F I Y F V V C V V G L C G N T L V I Y V I L R Y A K M K	T I T N I Y I L N L	
sp P30874 SSR2_HUMAN	Y Y D	L T S N A V L T - F I Y F V V C V V G L C G N T L V I Y V I L R Y A K M K	T I T N I Y I L N L	
sp Q9RQ02 LT4R1_RAT	S P G	G M S L S L L P I V L L S V A L V V G L P G N S F V V W S I L K R M Q K R S V T A L L V L N L		
sp O88855 LT4R1_MOUSE	S P G	G M S L S L L P I V L L S V A L V V G L P G N S F V V W S I L K R M Q K R T V T A L L V L N L		
sp Q15722 LT4R1_HUMAN	S L G	V E F I S L L A I I I L L S V A L A V G L P G N S F V V W S I L K R M Q K R S V T A L M V L N L		
	:	*	*	*
		TM3		
sp P49684 UR2R_RAT	A L A D	L L Y L L S I P F I I A T Y V T K D W H F G D V G C R V L F S L D F L T M H A S I F T L T I		
sp Q8VIH9 UR2R_MOUSE	A L A D	L L Y L L S I P F I I A T Y V T K D W H F G D V G C R V L F S L D F L T M H A S I F T L T I		
sp Q9UKP6 UR2R_HUMAN	A L A D	L L Y L L S I P F I I A T Y V T K D W H F G D V G C R V L F S L D F L T M H A S I F T L T I		
sp P35346 SSR5_HUMAN	A V A D	V L Y M L G L P F L A T Q N A A S F W P F G P V L C R L V M T L D G V N Q F T S V F C L T V		
sp P30935 SSR3_RAT	A L A D E	L F M L G L P F L A A Q N A L S Y W P F G S L M C R L V M A V D G I N Q F T S I F C L T V		
sp P32745 SSR3_MOUSE	A L A D E	L F M L G L P F L A A Q N A L S Y W P F G S L M C R L V M A V D G I N Q F T S I F C L T V		
sp O08858 SSR5_MOUSE	A V A D	V L F M L G L P F L A A Q N A L S Y W P F G S L M C R L V M A V D G I N Q F T S I F C L T V		
sp P30937 SSR4_RAT	A V A D E	L F M L S V P F V A S A A A L R H W P F G A V L C R A V L S V D G L N M F T S V F C L T V		
sp P49660 SSR4_MOUSE	A V A D E	L F M L S V P F V R S A A A L R H W P F G A V L C R A V L S V D G L N M F T S V F C L T V		
sp P31391 SSR4_HUMAN	A V A D E	L F M L S V P F V A S S A A L R H W P F G S V L C R A V L S V D G L N M F T S V F C L T V		
sp P28646 SSR1_RAT	A I A D E	L L M L S V P F L V T S T L L R H W P F G A L L C R L V L S V D A V N M F T S I Y C L T V		
sp P30873 SSR1_MOUSE	A I A D E	L L M L S V P F L V T S T L L R H W P F G A L L C R L V L S V D A V N M F T S I Y C L T V		
sp P30872 SSR1_HUMAN	A I A D E	L L M L S V P F L V T S T L L R H W P F G A L L C R L V L S V D A V N M F T S I Y C L T V		
sp P35370 OPRX_RAT	A L A D T	L V L L T L P F Q G T D I L L G F W P F G N A L C K T V I A I D Y Y N M F T S T F T L T A		
sp P35377 OPRX_MOUSE	A L A D T	L V L L T L P F Q G T D I L L G F W P F G N A L C K T V I A I D Y Y N M F T S T F T L T A		
sp P41146 OPRX_HUMAN	A L A D T	L V L L T L P F Q G T D I L L G F W P F G N A L C K T V I A I D Y Y N M F T S T F T L T A		
sp P30680 SSR2_RAT	A I A D E	F M L G L P F L A M Q V A L V H W P F G K A I C R V V M T V D G I N Q F T S I F C L T V		
sp P30875 SSR2_MOUSE	A I A D E	F M L G L P F L A M Q V A L V H W P F G K A I C R V V M T V D G I N Q F T S I F C L T V		
sp P30874 SSR2_HUMAN	A I A D E	F M L G L P F L A M Q V A L V H W P F G K A I C R V V M T V D G I N Q F T S I F C L T V		
sp Q9RQ02 LT4R1_RAT	A L A D	L A V L L T A P F F L H F L A R G T W S F E V T G C R L C H Y V C G V S M Y A S V L L I T I		
sp O88855 LT4R1_MOUSE	A L A D	L A V L L T A P F F L H F L A Q G T W S F G L A G C R L C H Y V C G V S M Y A S V L L I T A		
sp Q15722 LT4R1_HUMAN	A L A D	L A V L L T A P F F L H F L A Q G T W S F G L A G C R L C H Y V C G V S M Y A S V L L I T A		
	*	*	*	*
		TM4		
sp P49684 UR2R_RAT	M S S E	R Y A A V L R P L D T V Q - R S K G Y R K L L V L G T W L L A L L L T L P M M L A I Q L V R		
sp Q8VIH9 UR2R_MOUSE	M S S E	R Y A A V L R P L D T V Q - R S K G Y R K L L A L G T W L L A L L L T L P M M L A I R L V R		
sp Q9UKP6 UR2R_HUMAN	M S S E	R Y A A V L R P L D T V Q - R P K G Y R K L L A L G T W L L A L L L T L P V M L A M R L V R		
sp P35346 SSR5_HUMAN	M S V D	R Y L A V V H P L S S A R W R R P R V A K L A S A A A W V L S L C M S L P L L V F A D V - -		
sp P30936 SSR3_RAT	M S V D	R Y L A V V H P T R S A R W R T A P V A R M V S A A V W V A S A V V V L P V V V F S G V - -		
sp P30935 SSR3_MOUSE	M S V D	R Y L A V V H P T R S A R W R T A P V A R T V S R A V W V A S A V V V L P V V V F S G V - -		
sp P32745 SSR3_HUMAN	M S V D	R Y L A V V H P T R S A R W R T A P V A R T V S A A V W V A S A V V V L P V V V F S G V - -		
sp O08858 SSR5_MOUSE	M S V D	R Y L A V V H P L R S A R W R R P R V A K L A S A A A W V W F S L L M S L P L L V F A D V - -		
sp P30937 SSR4_RAT	L S V D	R Y V A V V H P L R A A T Y R R P S V A K L I N L G V W L A S L L V T L P I A V F A D T R P		
sp P49660 SSR4_MOUSE	L S V D	R Y V A V V H P L R T A T Y R R P S V A K L I N L G V W L A S L L V T L P I A V F A D T R P		
sp P31391 SSR4_HUMAN	L S V D	R Y V A V V H P L R A A T Y R R P S V A K L I N L G V W L A S L L V T L P I A I F A D T R P		
sp P28646 SSR1_RAT	L S V D	R Y V A V V H P I K A A R Y R R P T V A K V V N L G V W V L S L L V I L P I V V F S R T A A		
sp P30873 SSR1_MOUSE	L S V D	R Y V A V V H P I K A A R Y R R P T V A K V V N L G V W V L S L L V I L P I V V F S R T A A		
sp P30872 SSR1_HUMAN	L S V D	R Y V A V V H P I K A A R Y R R P T V A K V V N L G V W V L S L L V I L P I V V F S R T A A		
sp P35370 OPRX_RAT	M S V D	R Y V A I C H P I R A L D V R T S S K A Q A V N V A I W A L A S V V G V P V A I M G S A Q V		
sp P35377 OPRX_MOUSE	M S V D	R Y V A I C H P I R A L D V R T S S K A Q A V N V A I W A L A S V V G V P V A I M G S A Q V		
sp P41146 OPRX_HUMAN	M S V D	R Y V A I C H P I R A L D V R T S S K A Q A V N V A I W A L A S V V G V P V A I M G S A Q V		
sp P30680 SSR2_RAT	M S I D	R Y V A P F I K S A K W R R P R T A K M I N V A V W G V S L L V I L P I M I Y A G L R S		
sp P30875 SSR2_MOUSE	M S I D	R Y V A P F I K S A K W R R P R T A K M I N V A V W C V S L L V I L P I M I Y A G L R S		
sp P30874 SSR2_HUMAN	M S I D	R Y V A P F I K S A K W R R P R T A K M I N V A V W G V S L L V I L P I M I Y A G L R S		
sp Q9RQ02 LT4R1_RAT	M S L D	R S L A V A R P F V S Q K V R T K A F A R W V L A G I W V V S F L L A I P V L V Y R T V T -		
sp O88855 LT4R1_MOUSE	M S L D	R S L A V A R P F V S Q K L R T K A M A R R V L A G I W V L S F L L A T P V L A Y R T V V P		
sp Q15722 LT4R1_HUMAN	M S L D	R S L A V A R P F V S Q K L R T K A M A R R V L A G I W V L S F L L A T P V L A Y R T V V P		
	*	*	*	*



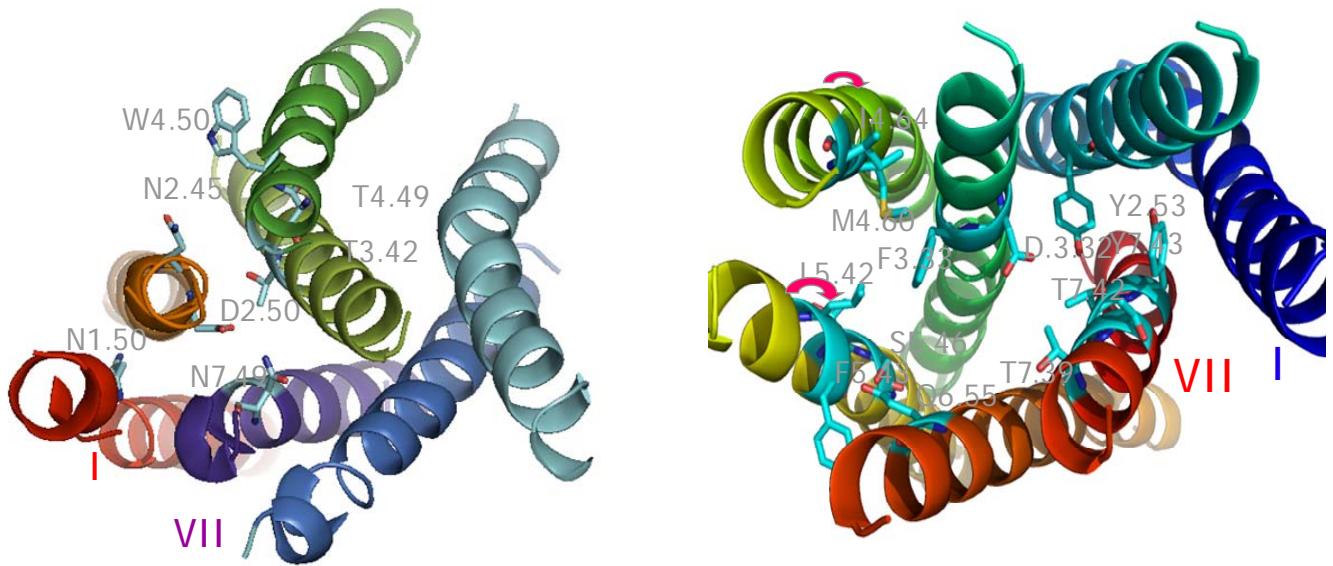
**Figure A-3.** The clustal W (1.83) multiple alignments of 23 GPCRs sequences with 20 to 90% identity.



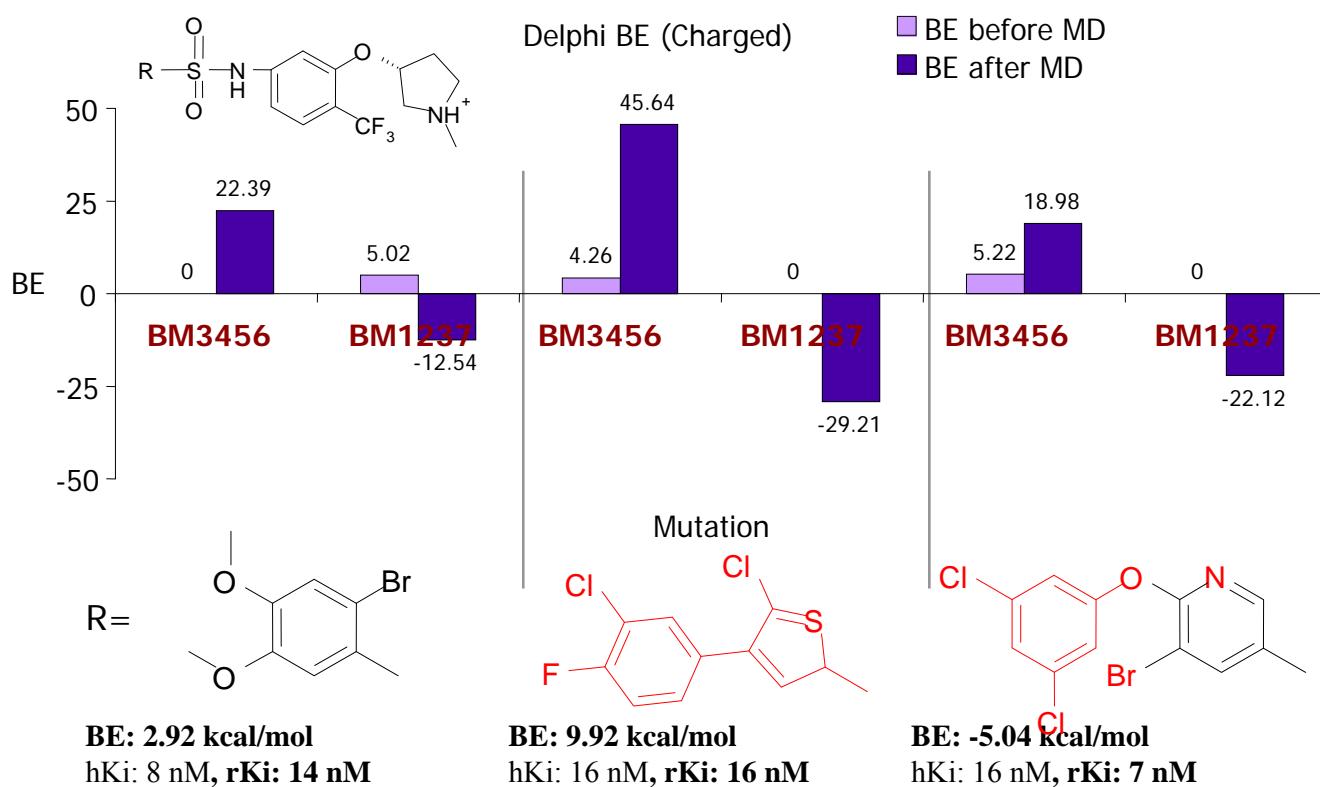
**Figure A-4.** Interhelical interaction energies of RotScan. Total energy and the maximum number of total H-bonds, middle H-bonds, and salt bridges for each helix are radically plotted in outward. Red circles were selected angles for combinatorial scan.



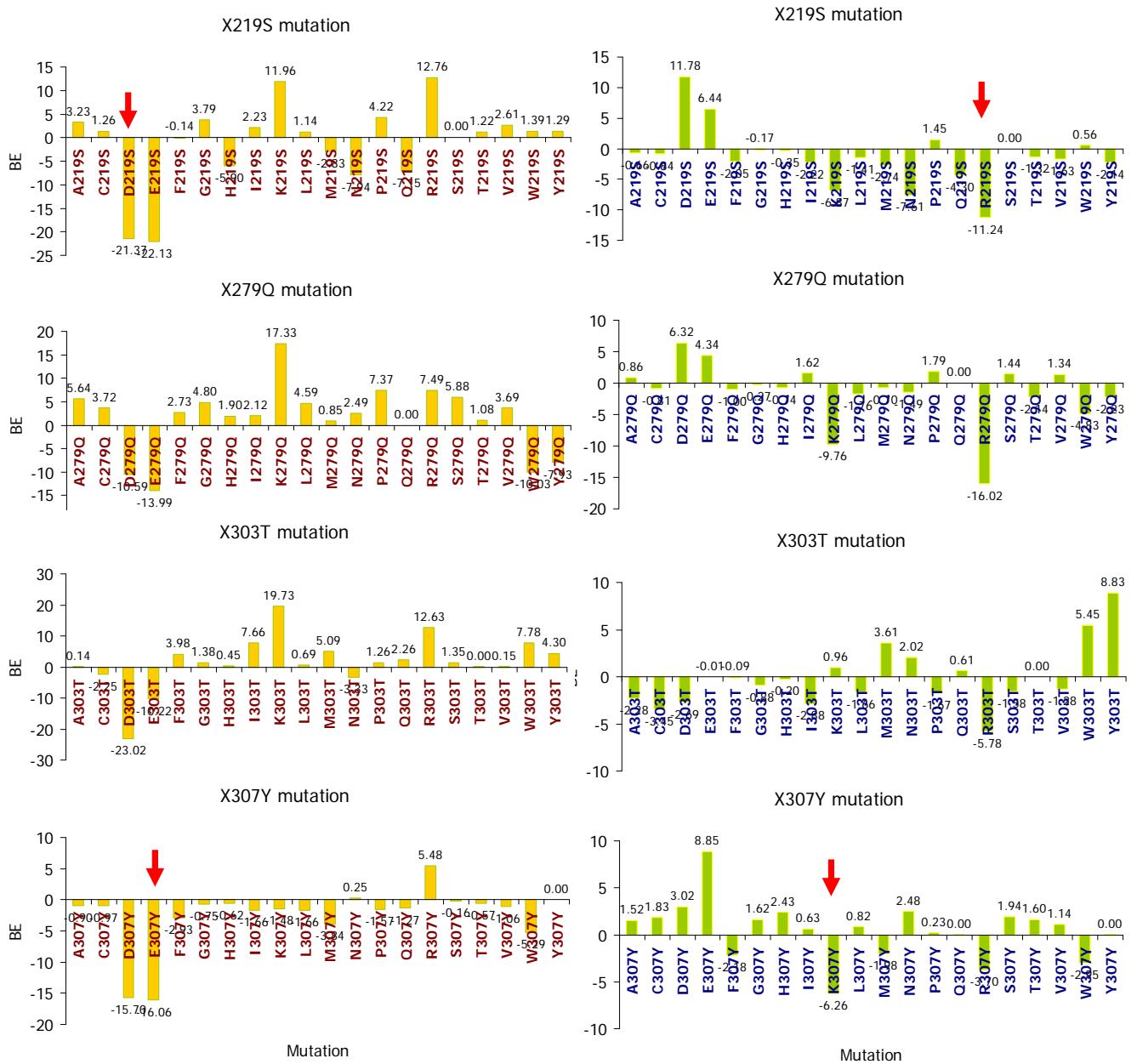
**Figure A-5.** SCREAM E of the human (top) and rat (bottom) Urotensin-II receptor. Blue line is an appropriate range of hydrophobic scale within 2 kcal/mol energy difference for further interhelical scan.



**Figure A-6.** H-bonding networks among highly conserved residues in each transmembrane (TM) from the cytoplasmic view (Left) and putative binding sites from the extracellular view (Right). After MembScream, important residues in TMs 4 and 5 which are known as the involvement of the ligand binding directed more toward the binding site and final structure shows better H-bonding networks between N2.45 and W4.50.

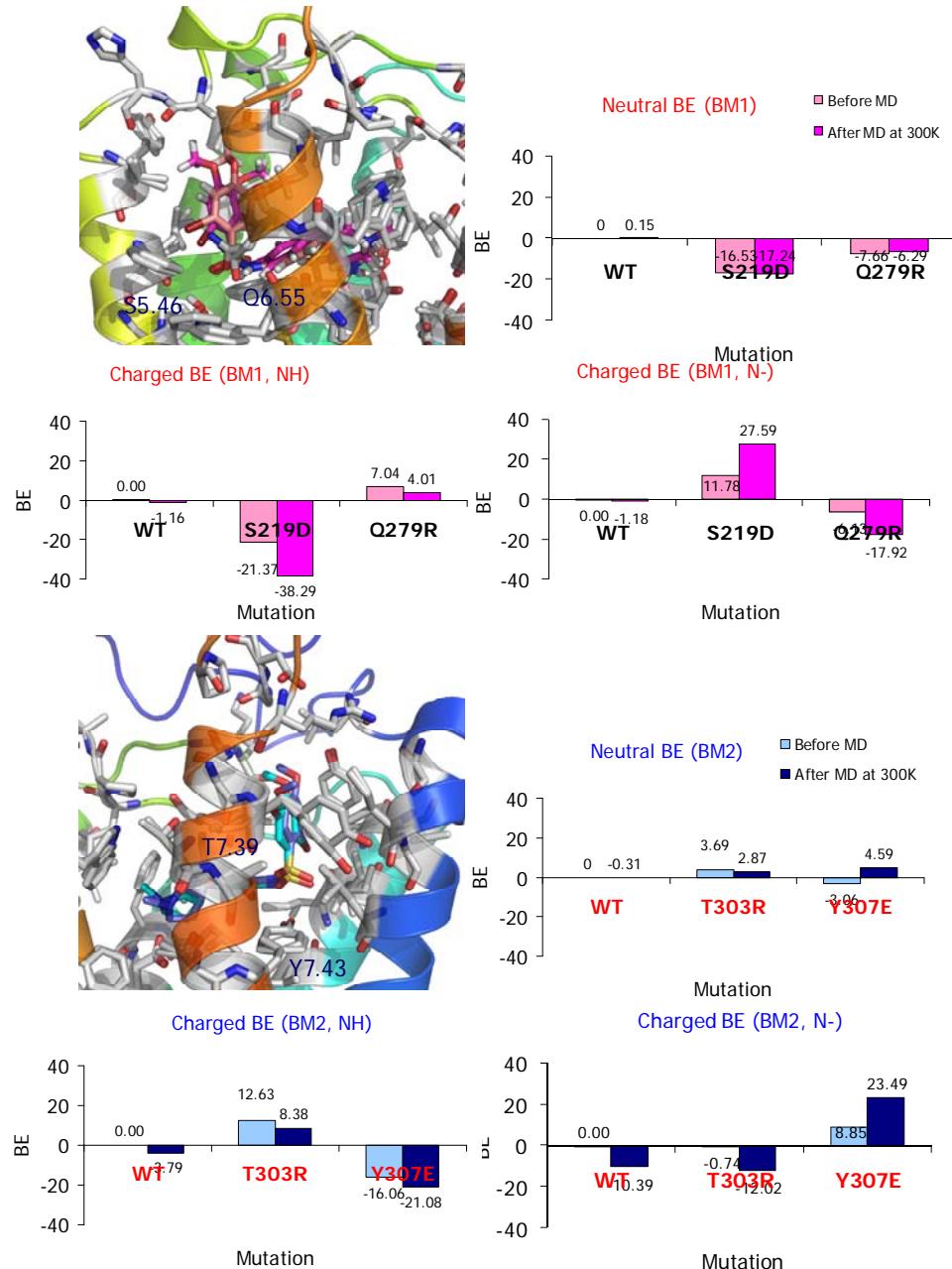


**Figure A-7.** The comparison of binding energies of three SB706375 analogues before (in light color) and after (in dark color) molecular dynamics (MD). Relative binding energies before and after 10 ps quench annealing from 50 to 600 K were calculated and compared with the binding energy of the best case in each compound before MD which was set to 0.

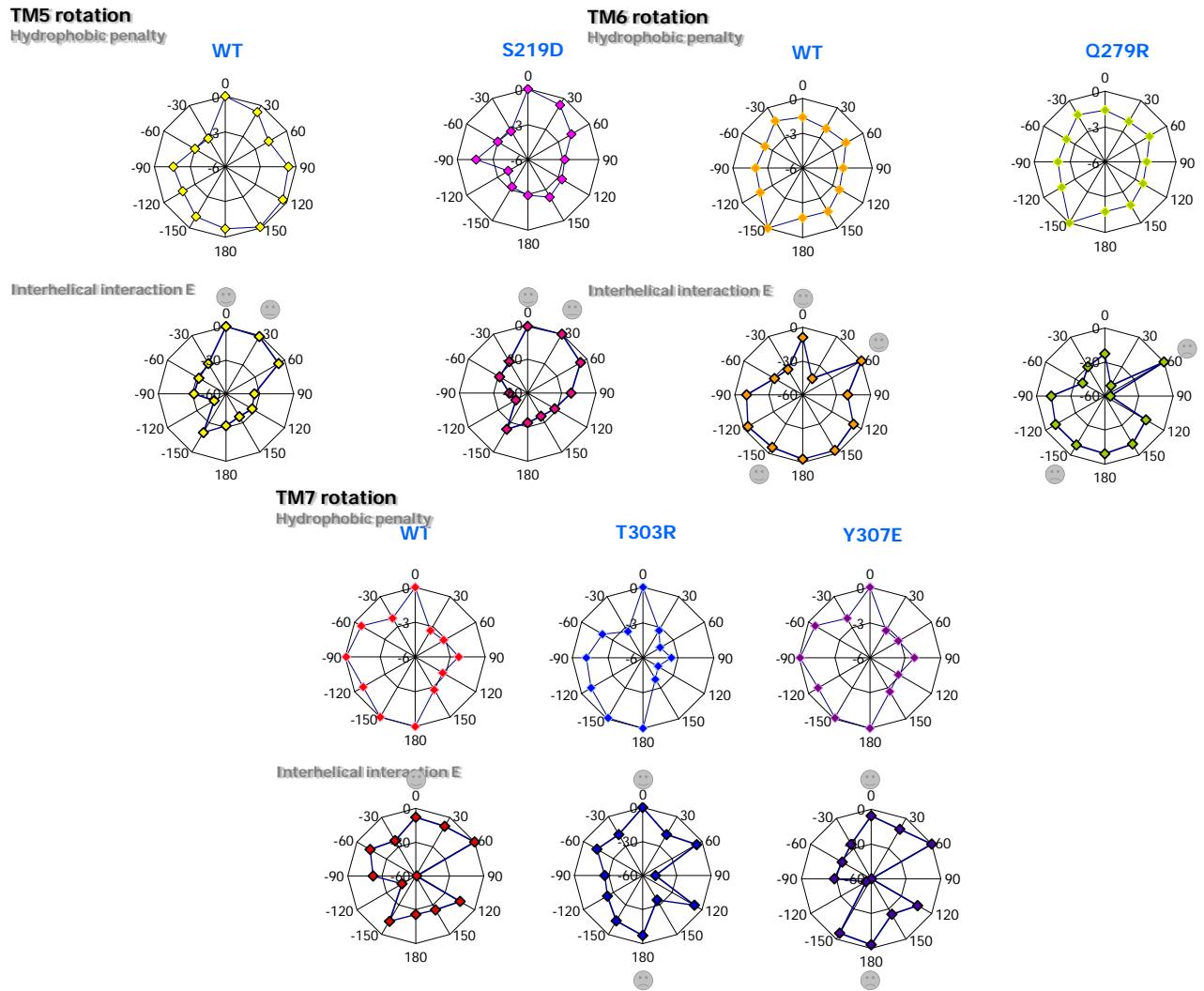


**Figure A-8.** The relative binding energies from Delphi method among 20 amino acids mutant receptors

of rat Uroteinsin II receptor complex with SB-706375 in binding mode 1 and 2 compared with the binding energy of wild type set to 0. Energetically favorable mutation is shown by red arrow.

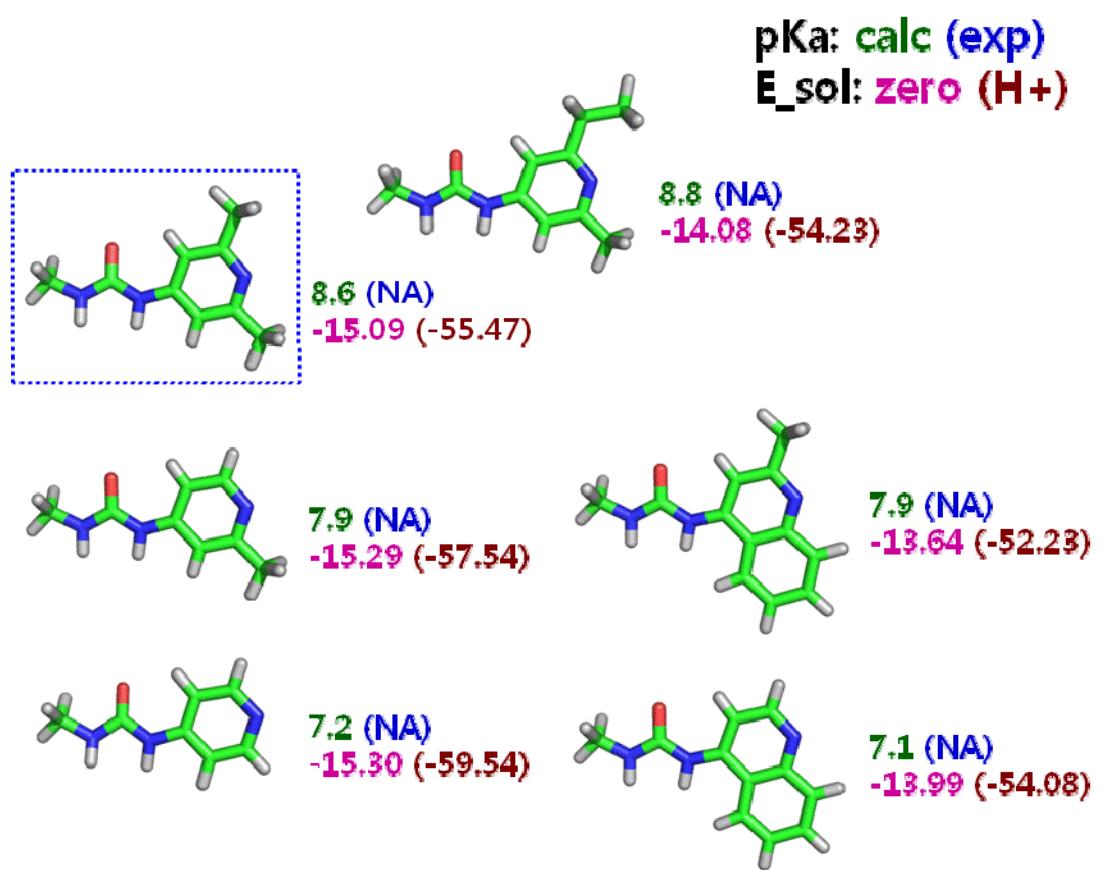


**Figure A-9.** The relative binding energies from Delphi method among mutant receptors of rat Uroteinsin II receptor complex with SB-706375 in Binding mode 1 (top) and 2 (bottom). Relative binding energies before and after 10 ps quench annealing for charged (50 to 600 K) and neutral system (50 to 300 K) were calculated and compared with the binding energy of wild type set to 0.

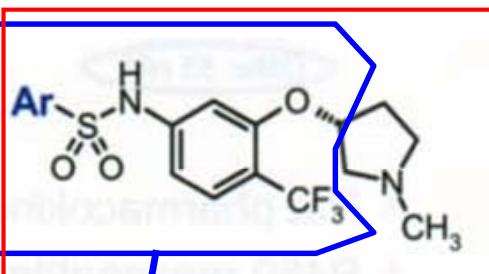


**Figure A-10.** Hydrophobicity penalty and Scream E of the wild type and mutant rat Urotensin-II receptor.

The hydrophobicity penalty and Epolar energy of transmembrane (TM) 5, 6, and 7 was calculated and plotted radically outward in kcal/mol. In the E plot, 0 is the lowest E, and others are the relative E compared with the lowest one. Energetically preferred angles of each TM are marked with circle.



**Figure A-11.** The calculated pKa (pyrimidine N) and solvation energies for various ACT structures.



The figure shows a table of hUT Ki values for various aryl groups (Ar) and three models for pKa calculations: CH3, CF3, and All Parts.

**Models used for pKa calculations**

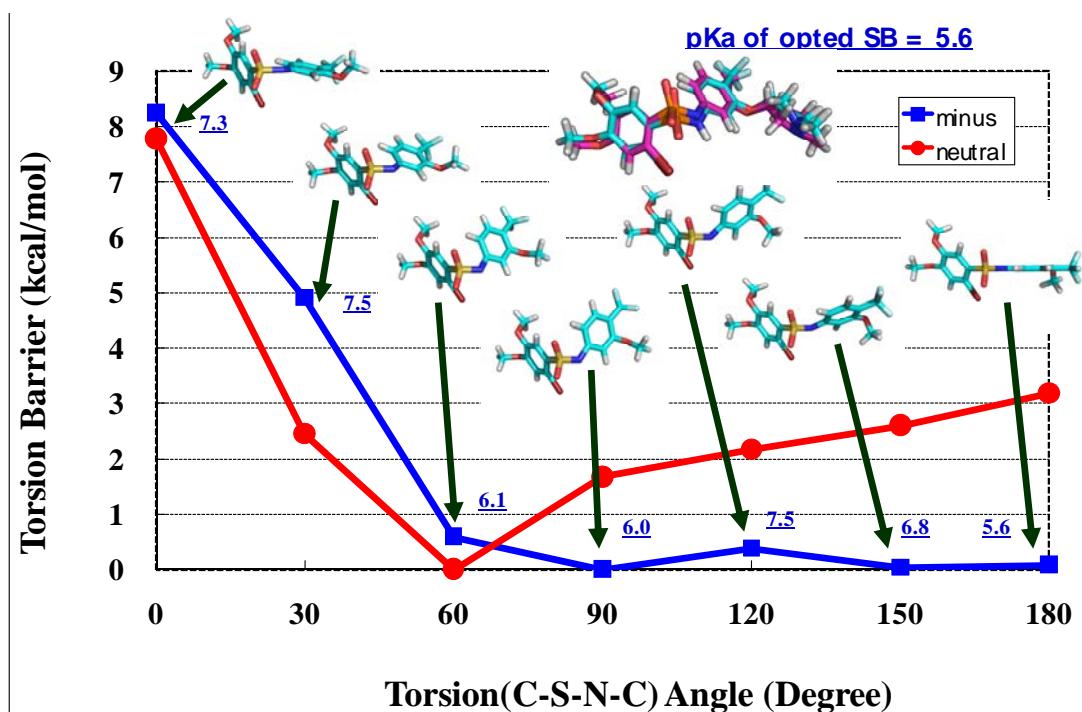
**Using Structure of SB-5-minus**

#	Ar	hUT Ki (nM)
1		1784
2		227
3		367
4		86
5		46
6		15
7		8

Ar	Ki	with CH3		with CF3		With All Parts	
		pKa	E_sol(w/H)	pKa	E_sol(w/H)	pKa	
1	1784	8.8	-11.06	-57.05	5.8	-10.72	-52.30
2	227	10.3	-12.55	-57.69	6.7	-12.00	-54.43
3	367	9.1	-11.56	-58.22	6.1	-11.15	-53.59
4	86	8.5	-9.95	-53.03	4.3	-10.26	-50.95
5	46	7.6	-9.28	-50.31	4.7	-9.29	-46.33

**Figure A-12.** The calculated pKa (sulfonamide N) and solvation energies for various SB structures.

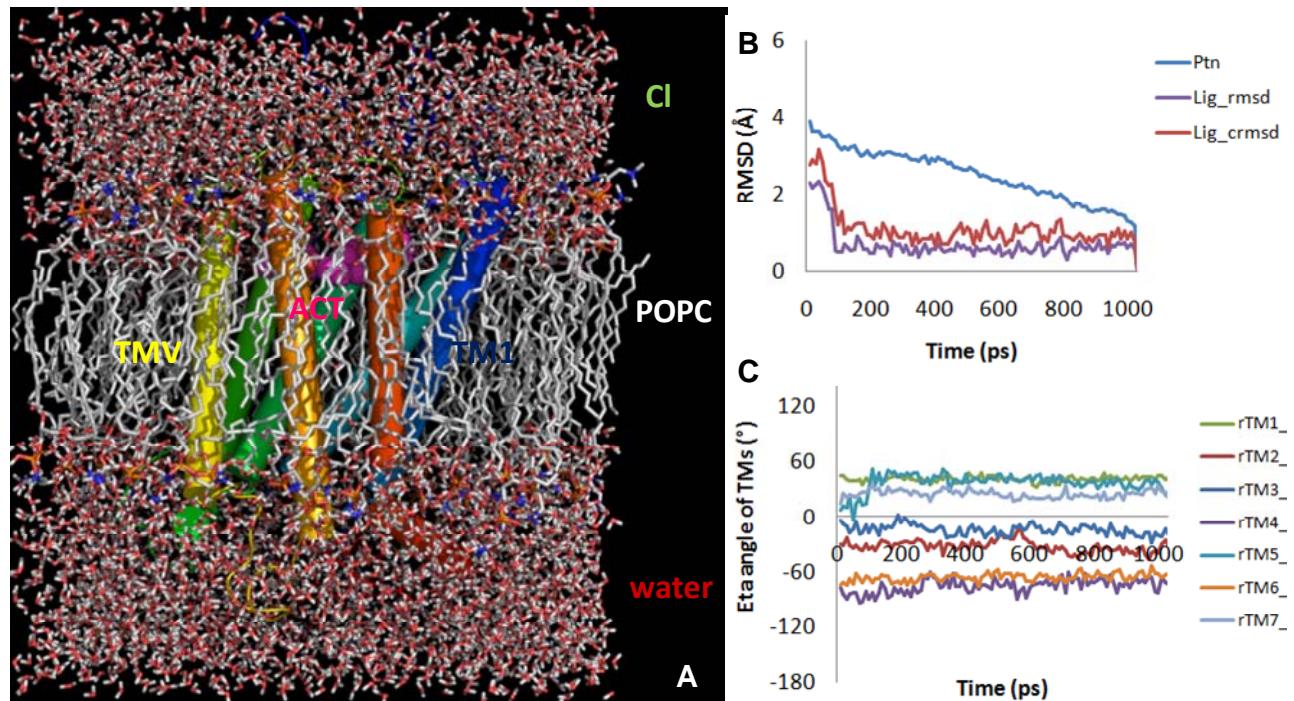


**Figure A-13.** The potential energy barrier of SB-sulfonamide calculated by B3LYP/LACVP\*

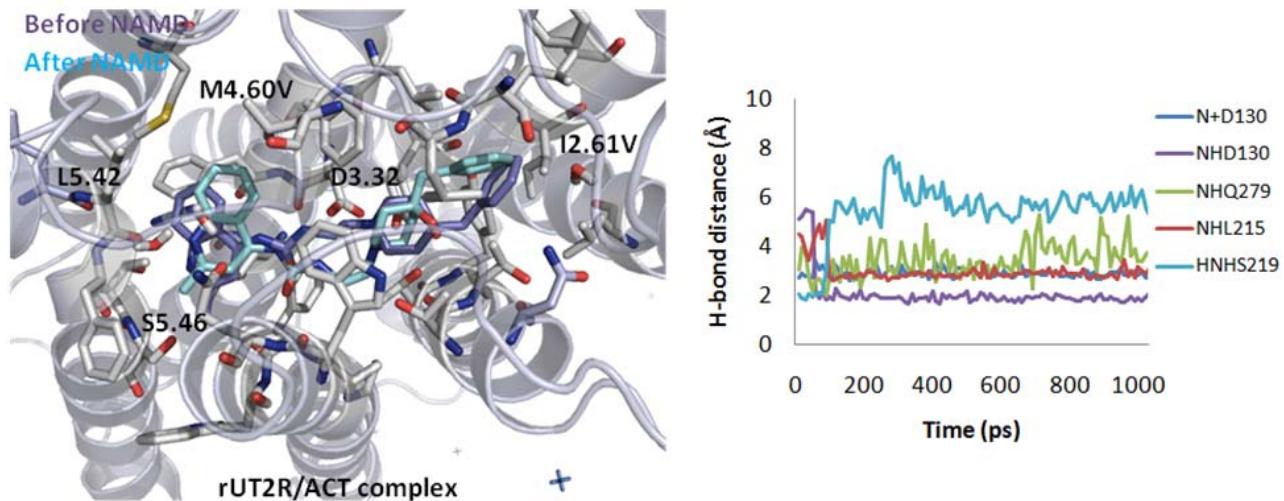
\* Full membrane full solvent simulations

The loops and N and C termini were added, and we carried out 1 ns of MD simulations on the apo protein, SB706375-bound rUT2R, ACT058362-bound rat and human UT2R in explicit membrane and water, as shown in Figure A-14A. The root mean square deviation (RMSD) trajectory of protein and ligand through 1ns dynamics is given in Figure 5B, indicating that the ligand was stabilized after 100ps in the binding site. The reason for ligand fluctuation until 100ps is that the original H-bond of NH is in a quinoline ring and S219<sup>5.46</sup> destabilized and formed an alternative H-bond with the backbone carbonyl atom of L215<sup>5.42</sup>. Thus, the  $\eta$  angle of TM5 did not converge until 100ps, as shown in Figure A-14C. In particular, Figure A-15 shows the important receptor-ligand interactions in the binding site. The starting structure showed several H-bonds: D130<sup>3.32</sup> and a protonated amine in piperidine ring, S219<sup>5.46</sup> and the NH in the quinoline ring, and Q279<sup>6.55</sup> and the CO atom of the urea group. During these dynamics, water entered into the binding site, bridging between water and the NH of the quinoline ring thus replacing the H-bond of the NH quinoline with the side chain OH of S219<sup>5.46</sup> into the backbone CO atom of L219<sup>5.42</sup>.

Our conclusion from these full MD studies is that the MembScream predicted structure is stable. All classical H-bonding networks among TMs 1-2-7 (1.50-2.50-7.49) and TMs 2-3-4 (2.45-3.42-4.50) in rUT2R structures were stabilized through 1ns simulations.



**Figure A-14.** Molecular dynamics simulation of ACT058362-bound rat-Urotensin II receptor in explicit lipid and water. A) Final structure embedded into POPC and water, B) root mean square deviation (RMSD) trajectory of protein and ligand (Lig\_rmsd: absolute RMSD, Lig\_crmsd: relative RMSD), C)  $\eta$  angle variation of each transmembrane helix through 1ns simulations.



**Figure A-15.** Important interaction of ACT058362 at Urotensin II receptor. A) The superimposition of ACT058362 at rat Urotensin II receptor before and after 1ns molecular dynamics, B) Several hydrogen bonding distances through 1ns dynamics.