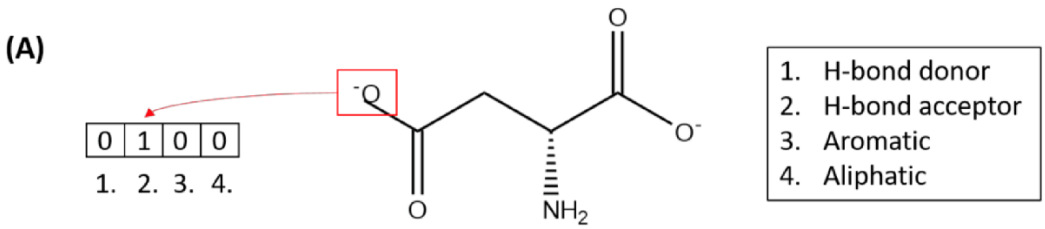
Supporting Information

Table of contents

Supporting Figures



(B)

Binding Pocket 1	F				S				E				T				Y				M			
	0	0	1	0	1	1	0	0	0	1	0	0	1	1	0	0	1	1	1	0	0	0	0	1
Binding Pocket 2	V				D				Y				M				F				K			
	0	0	0	1	0	1	0	0	1	1	1	0	0	0	0	1	0	0	1	0	1	0	0	0

$$R = \frac{c}{a+b-c} = \frac{3}{10+8-3} = 0.2$$

Figure S1. Illustrative example explaining the calculation of the residue chemical similarity (*R*) term. (A) Shown is the binary representation of the chemical features of aspartate. Each bit is position dependent and corresponds to one of four chemical features (H-bond donor, H-bond acceptor, aromatic, and aliphatic). The anionic oxygen from the carboxyl group (red box) is a H-bond acceptor, thus a bit is set in the second position. (B) In this simplified example, the two binding pockets are a binary representation of the chemical features of the aligned binding pockets produced from PPS-Align. The *R* score is calculated as a Tanimoto coefficient between the two bit-string representations of the binding pockets, where *a* is the number of bits in Binding Pocket 1, *b* is the number of bits in Binding Pocket 2, and *c* is the number of bits shared between the two. Here, the two binding pockets have a RCS of 0.2, as shown in the figure.

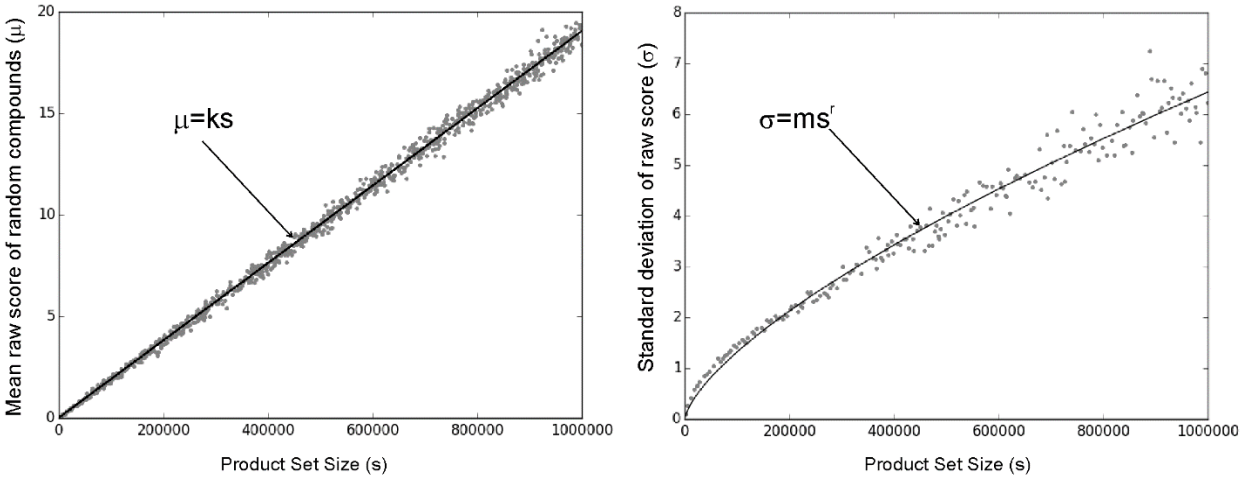


Figure S2. Mean and standard deviation of ligand similarities between random ligand sets from the GLASS database versus the product of set sizes. Only the ligand pairs with a Tanimoto coefficient above 0.84 are calculated and the data follow well the linear and power-law equation shown in Eq. (6).

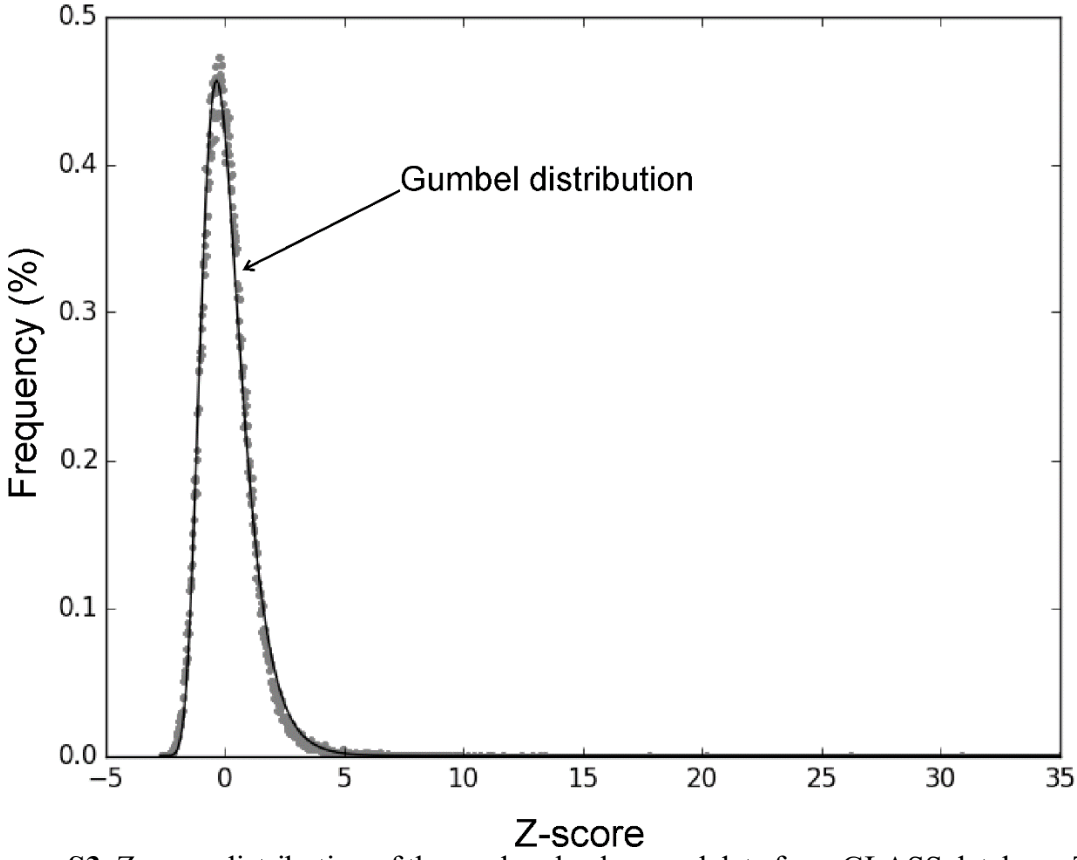


Figure S3. Z-score distribution of the random background data from GLASS database. The Tanimoto coefficient of 0.84 was found to be the best fit with the Gumbel distribution.

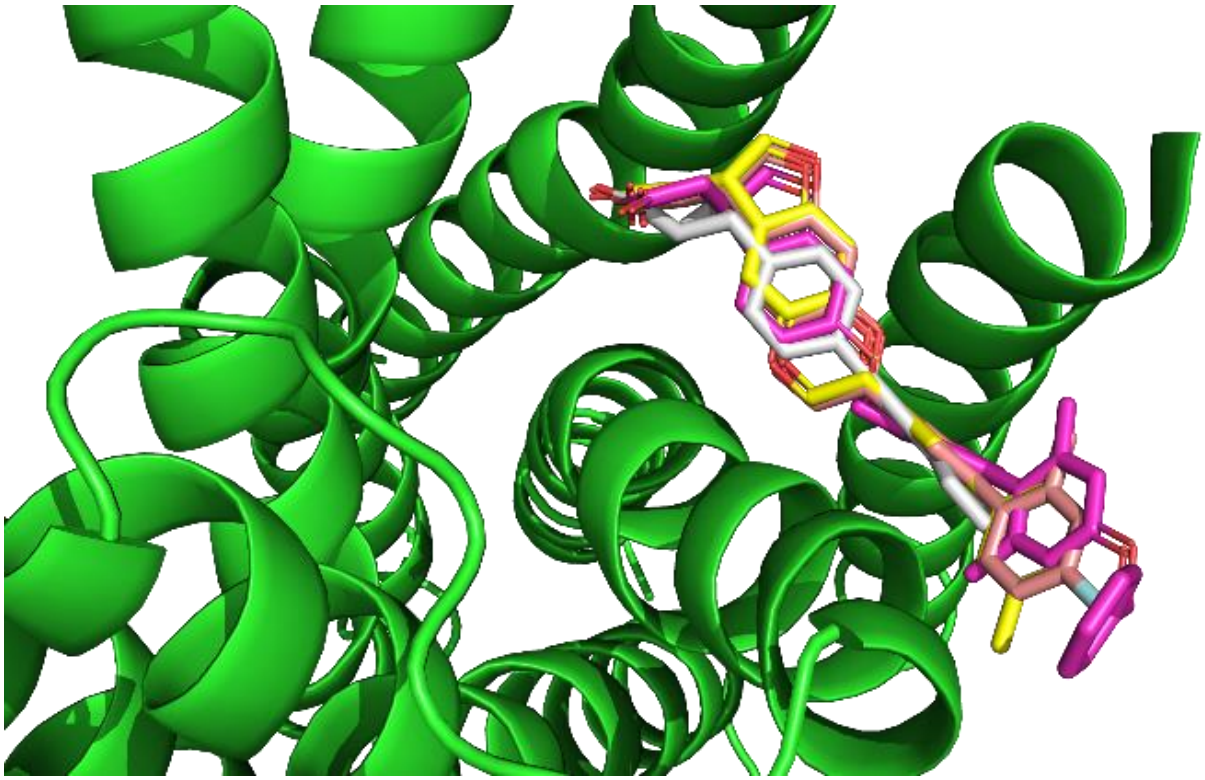


Figure S4. Top 5 active compound results for free fatty acid Receptor 1 using AutoDock Vina. Note that the binding pocket resides between two of its transmembrane domains, extending from the lipid bilayer to the inner cavity of the receptor. As is evident from the structure of the docked ligands, they share an obvious chemotype that also extends to the rest of the active set.

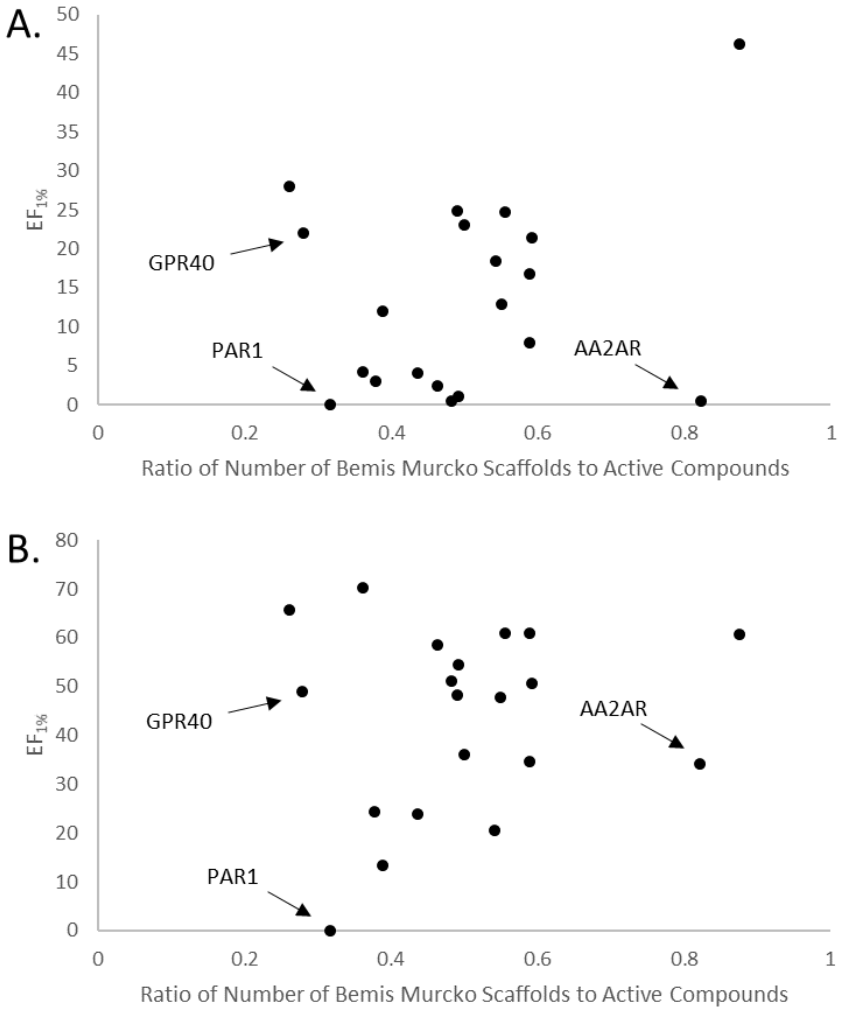


Figure S5. Comparison of Retrospective Virtual Screening Performance of MAGELLAN with Chemical Diversity for the GPCR-Bench Dataset. Results are shown (A.) with and (B.) without the 30% sequence identity cutoff. The number of Bemis Murcko scaffolds derived from the active compounds was divided by the total number of active compounds for each GPCR to obtain a ratio; a higher value indicates greater chemical scaffold diversity.

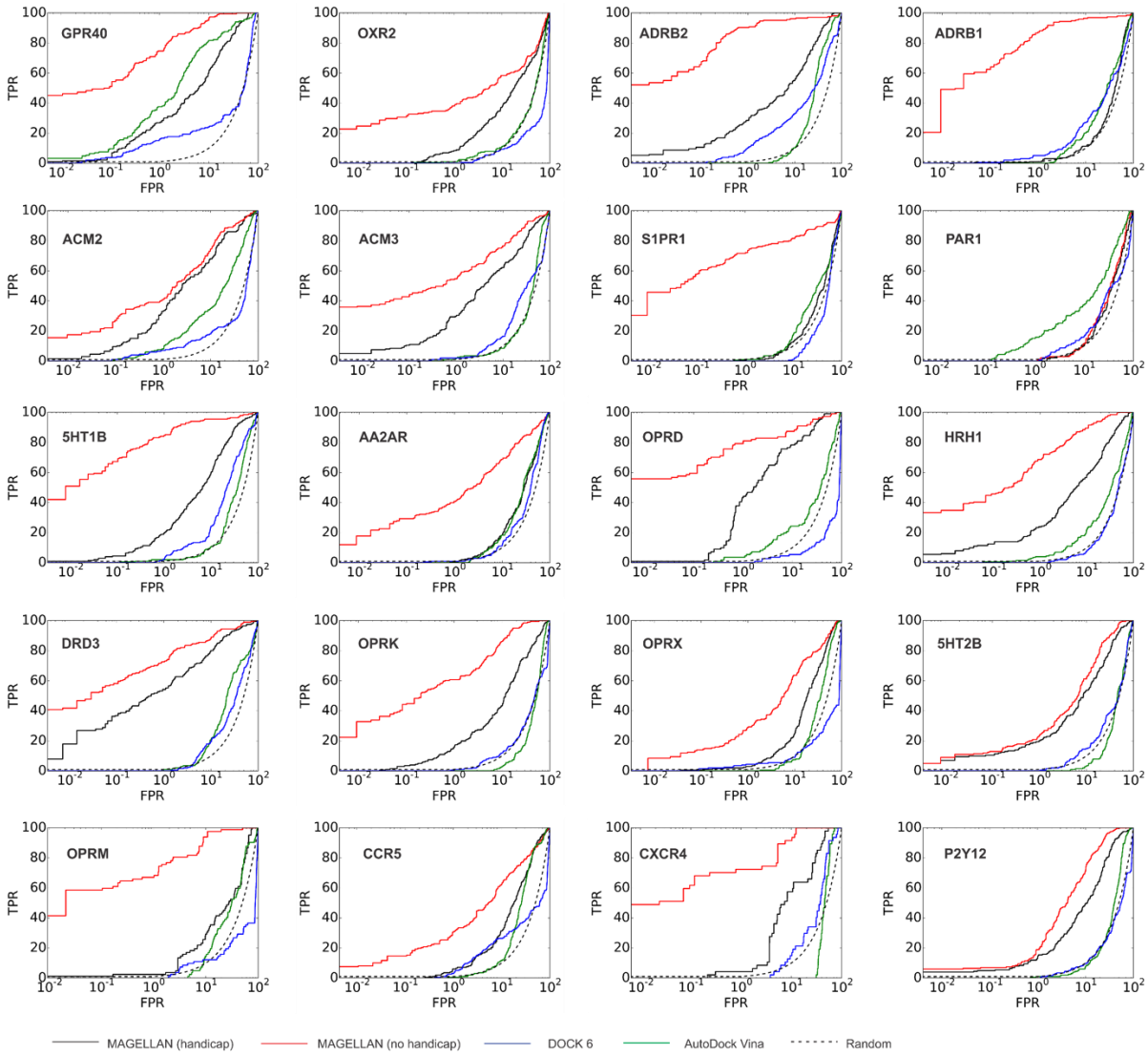


Figure S6. Log ROC Curves for Retrospective Virtual Screen Results of MAGELLAN, autodock Vina, and DOCK 6 with GPCR-Bench Dataset. The following GPCRs were tested: free fatty acid receptor 1 (GPR40), orexin receptor 2 (OXR2), beta-2 adrenergic receptor (ADRB2), be ta-1 adrenergic receptor (ADRB1), muscarinic acetylcholine receptor 2 (ACM2), muscarinic acetylcholine receptor 3 (ACM3), sphingosine 1-phosphate receptor (S1PR1), proteinase-activated receptor 1 (PAR1), 5-hydroxytryptamine receptor 1B (5HT1B), adenosine receptor A2A (AA2AR), delta opioid receptor (OPRD), histamine H1 receptor (HRH1), dopamine D3 receptor (DRD3), kappa opioid receptor (OPRK), nociception receptor (OPRX), 5-hydroxytryptamine receptor 2B (5HT2B), mu opioid receptor (OPRM), C-C chemokine receptor type 5 (CCR5), C-X-C chemokine receptor type 4 (CXCR4), and purinergic receptor (P2Y12).

BW:	1.35	1.39	1.42	1.46	2.53	2.57	2.58	2.61	2.62
Query:	V	C	L	G	I	L	P	L	Y
	:	:	:	:	:		:	:	
Target:	V	C	L	G	I	M	P	L	F
BW:	2.64	2.65	3.28	3.29	3.32	3.33	3.36	3.37	3.39
Query:	L	W	S	L	G	E	T	Y	T
	:	:				:	:	:	:
Target:	L	W	F	Q	S	E	T	Y	T
BW:	3.40	4.56	4.57	4.60	4.61	5.38	5.39	5.42	5.43
Query:	L	S	A	F	L	V	M	V	T
		:	:			:	:	:	
Target:	V	S	A	I	F	V	M	V	S
BW:	5.46	5.47	6.44	6.47	6.48	6.51	6.52	6.55	6.58
Query:	Y	F	F	C	W	F	H	R	Y
		:	:	:	:	:	:	:	
Target:	F	F	F	C	W	F	H	R	F
BW:	6.59	7.35	7.39	7.42	7.43	7.45	7.46		
Query:	I	N	L	F	Y	S	A		
		:		:	:	:	:		
Target:	S	N	F	F	Y	S	A		

Figure S7. BindRes alignment for human motilin receptor. BW refers to the Ballesteros-Weinstein numbering scheme for the binding site residues, while the query and target are the human motilin receptor (UniProt: O43193) and pig growth hormone secretagogue receptor type 1 (UniProt: Q95254), respectively. The colons depict residues that are identical.

Supporting Tables

Table S1. Comparison of Boltzmann-enhanced receiver operating characteristic (BEDROC) values by MAGELLAN, AutoDock Vina and Dock 6 on 20 Class A GPCRs in GPCR-Bench. Values out and in parentheses are the results for MAGELLAN with or without handicap cutoffs.

Gene Name	**UniProt ID**	**MAGELLAN**	**AutoDock Vina**	**Dock 6**
GPR40	O14842	0.46 (0.87)	0.59	0.35
OX2R	O43614	0.26 (0.51)	0.12	0.03
ADRB2	P07550	0.45 (0.94)	0.17	0.23
ADRB1	P07700	0.07 (0.92)	0.07	0.23
ACM2	P08172	0.54 (0.62)	0.11	0.21
ACM3	P08483	0.49 (0.69)	0.07	0.13
S1PR1	P21453	0.06 (0.80)	0.22	0.13
PAR1	P25116	0.06 (0.06)	0.34	0.26
5HT1B	P28222	0.42 (0.92)	0.07	0.02
AA2AR	P29274	0.10 (0.58)	0.29	0.13
OPRD	P32300	0.60 (0.84)	0.08	0.19
HRH1	P35367	0.43 (0.80)	0.11	0.18
DRD3	P35462	0.69 (0.83)	0.09	0.07
OPRK	P41145	0.33 (0.75)	0.13	0.08
OPRX	P41146	0.15 (0.46)	0.18	0.04
5HT2B	P41595	0.38 (0.44)	0.16	0.06
OPRM	P42866	0.16 (0.82)	0.12	0.12
CCR5	P51681	0.20 (0.49)	0.04	0.05
CXCR4	P61073	0.28 (0.81)	0.06	0.09
P2Y12	Q9H244	0.33 (0.49)	0.09	0.14
Average		**0.32 (0.68)**	**0.16**	**0.14**

Supporting Text

Parameters for Molecular Docking

For AutoDock Vina (version 1.1.2), all compounds were converted from Mol2 to PDBQT format. Additionally, the experimental GPCR PDB files were converted into the PDBQT format, whereby hydrogens and partial charges were added to all PDBQT files. A $30\times30\times30$ $Å^3$ search space was defined on the receptor so that it centered upon the crystal ligand. Default settings were used for the docking runs: 1.) exhaustivity was set to 8, 2.) a single core was assigned to the job, 3.) nine binding modes were generated after each run, and 4.) the maximum energy difference between the best and worst binding modes displayed was set to 3 kcal/mol. Compounds were ultimately ranked according to their docking scores based on AutoDock Vina's scoring function.

To examine the flexible rescoring and refinement, the top-scoring poses generated from AutoDock Vina were converted from PDBQT to SDF format with OpenBabel (version 2.4.0). All compounds for each receptor were imported into a Molecular Operating Environment (MOE, version 2019.0101) database. Poses were rescored with induced refinement using the GBVI/WSA dG scoring methodology. Side chains of the receptor were tethered with a weight of 10. Minimization was set to terminate after 50 iterations with an RMS gradient of 0.01. The refined poses were then ranked according to their scores.

For DOCK 6 (version 6.7), all GPCR PDB files were converted to the Mol2 format, where hydrogens and partial charges were added. Molecular surfaces were generated with the DMS tool in Chimera with a probe radius of 1.4 Å. Spheres were selected within 5 Å of the crystal ligand, while the scoring grid enclosed the spheres with a 5 Å margin. A grid was constructed around the selected spheres with a padding of 5 Å in all 6 directions; a grid spacing of 0.3 Å was used, while van der Waals overlap was set to 0.75. The simplex minimizer was set to the following settings: 1.) one minimization cycle, 2.) exit cycle when converging at a cutoff of 0.1 kcal/mol, 3.) exit minimization when cycles converge to cutoff of 1 kcal/mol, 4.) translation step size of 0.1 Å, 5.) rotation step size of 0.1 radian, 6.) torsion angle step size of 10 degrees, 7.) 500 iterations per cycle per anchor, 8.) 500 iterations per cycle per growth step, and 9.) random number generator seed of 0. All other parameters (e.g. Lennard-Jones attractive/repulsive components, dielectric factor, etc.) were set to the default recommended values. Docking was performed with the recommended settings for which ligand flexibility was accounted but receptor flexibility was neglected, with compounds ranked according to their grid-based score.