Supporting Information

The Binding Space Concept:

A New Approach to Enhance the Reliability of Docking scores, and its Application to Predicting Butyrylcholinesterase (BChE) Hydrolytic Activity

Giulio Vistoli[1*], Angelica Mazzolari[1], Bernard Testa[2], Alessandro Pedretti[1]

1) Dipartimento di Scienze Farmaceutiche "Pietro Pratesi", Facoltà di Farmacia, Università degli Studi di Milano, Via Mangiagalli, 25, I-20133 Milano, Italy

2) Dept of Pharmacy, University Hospital Centre (CHUV), Rue du Bugnon, CH-1011 Lausanne, Switzerland

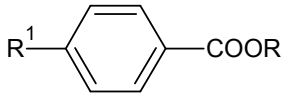
Corresponding Author

Giulio Vistoli, Giulio.vistoli@unimi.it; Phone: +390250319349; Fax: +390250319359

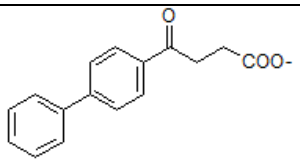
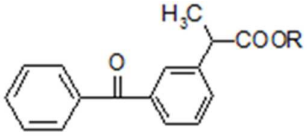
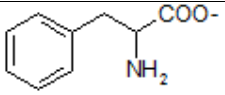
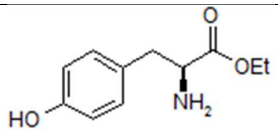
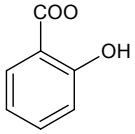
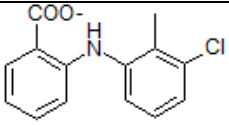
Table S1 Substrates considered in the docking studies (n = 75) as subdivided into:

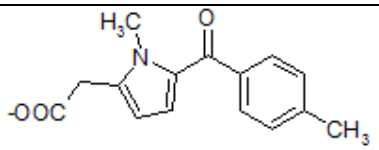
a) Benzoic esters (3 derivatives cited in refs 8 and 3)

Compound #	R	R^1	$t_{1/2}$ (m) in P80
1	Methyl	OH	3000
2	Ethyl	NH_2	6000

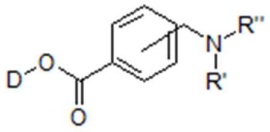


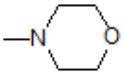
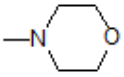
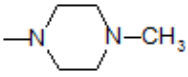
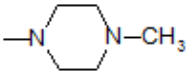
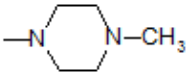
b) Methyl and ethyl esters (6 derivatives cited in ref. 2 and 3)

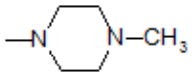
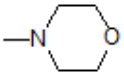
Compound	Ester moiety name	R-CH_3	$t_{1/2}$ (m) in P80
3	Fenbufen		282
4	Ketoprofen		1200
5	L-phenylanaline		29
6	L-tyrosine ethyl ester		59
27	Salicylic acid		1056
28	Tolfenamic acid		6000

29	Tolmetin		1140

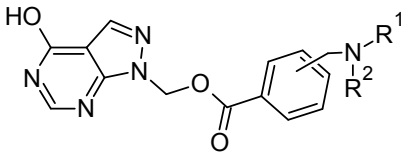
c) N-Substituted (Aminomethyl) benzoate esters of drugs (13 derivatives cited in ref. 1)

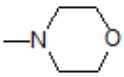
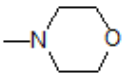
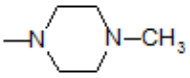
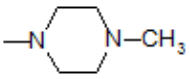
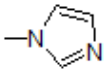


Compound #	Drug (D-OH)	R'	R''	-CH_2NR'R'' position	$t_{1/2}$ (m) in P80
7	Hydrocortisone	H	CH3	Meta	25
8	Hydrocortisone	H	C3H7	Meta	38
9	Hydrocortisone	C2H5	C2H5	Meta	34
10	Hydrocortisone	C2H5	C2H5	Para	107
11	Hydrocortisone	C2H5	CH2CH2NEt2	Meta	8.0
12	Hydrocortisone			Meta	62
13	Hydrocortisone			Para	92
14	Hydrocortisone			Meta	15
15	Hydrocortisone			Para	147
16	Prednisolone			Meta	18

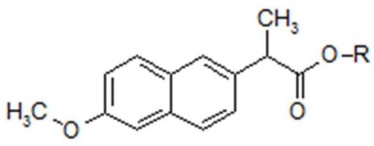
17	Methylprednisolone			Meta	29
18	Methylprednisolone	CH3	CH2CH2NEt2	Meta	18
19	Acyclovir			Para	3.7

d) Allopurinol esters (6 derivatives cited in ref. 1)



Cmpd	**-$CH_2NR'R''$ position**	**R1**	**R2**	**$t_{1/2}$ (m) in P80**
20	Para	-CH_3	-CH_3	9.40
21	Para			1.90
22	Meta			8.50
23	Para			11.00
24	Meta			0.50
25	Meta			3.60

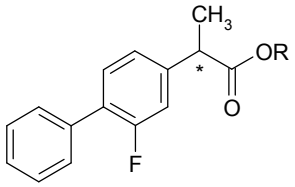
e) Naproxen esters (1 derivative cited in ref. 2)

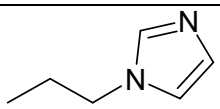
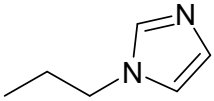
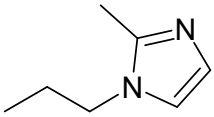
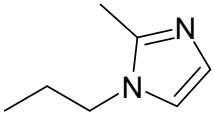


Cmpd #	**Ester moiety name**	**R**	**$t_{1/2}$ (m) in S80**

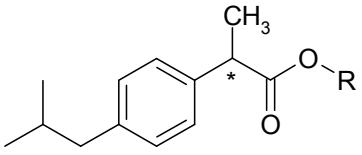
26	Ethyl	$-CH_2CH_3$	1206 (ref. 7)

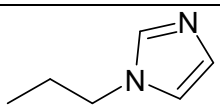
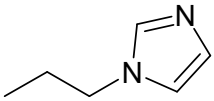
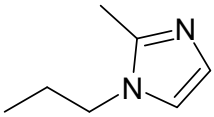
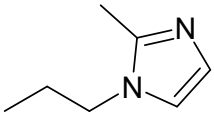
f) Flurbiprofen esters (4 derivatives cited in ref. 7)



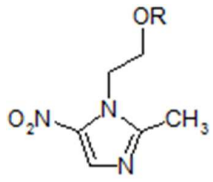
Compound #	Configuration	R	$t_{1/2}$ (m) in P80
30	R		1080
31	S		26700
32	R		2280
33	S		21300

g) Ibuprofen esters (8 derivatives cited in ref. 7)

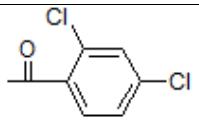
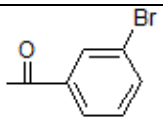
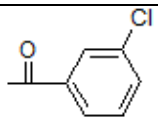
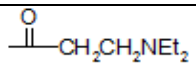
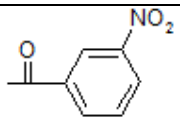
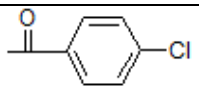
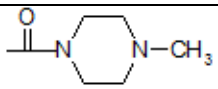
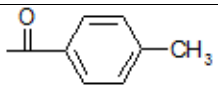
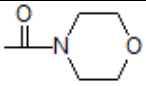
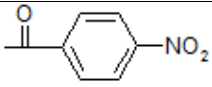
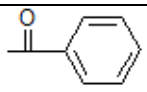
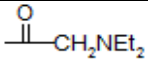


Compound #	Configuration	R	$t_{1/2}$ (m) in P80
34	R	$-CH_2CH_3$	5580
35	S	$-CH_2CH_3$	15900
36	R	$-CH_2CH_2N(CH_3)_2$	102
37	S	$-CH_2CH_2N(CH_3)_2$	156
38	R		1140
39	S		57600
40	R		312
41	S		8100

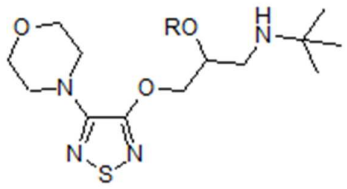
h) Metronidazole esters (19 derivatives cited in ref. 5 and ref. 6)



Compound #	R	$t_{1/2}$ (m) in P80

42		4.8
43		5.4
44		5.4
45		46.2
46		13.8
47		4.2
48		522
49		7.8
50		30
51		4.2
52		78
53		2.4
54		12

i) Timolol esters (21 derivatives cited in ref. 4)



Compound #	Ester moiety name	R'	$t_{1/2}$ (m) in P80
55	*O*-1'-Methylcyclopropanoyl	-CO*c*C_3H_4-1'-CH_3	690
56	*O*-2-Aminobenzoyl	-COC_6H_4-*o*-NH_2	4200
57	*O*-2'-Methylcyclopropanoyl	-CO*c*C_3H_4-2'-CH_3	792
58	*O*-2-Methylbenzoyl	-COC_6H_4-*o*-CH_3	3120
59	*O*-2-Methoxybenzoyl	-COC_6H_4-*o*-OCH_3	438
60	*O*-3,3-Dimethylbutyryl	-$COCH_2C(CH_3)_3$	6300
61	*O*-3-Thienyl	-COC_4H_3S	432
62	*O*-4-Methoxybenzoyl	-COC_6H_4-*p*-OCH_3	1200
63	*O*-4-Methylbenzoyl	-COC_6H_4-*p*-CH_3	1200
64	*O*-Acetyl	-$COCH_3$	35
65	*O*-Benzoyl	-COC_6H_{11}	264
66	*O*-Butyryl	-$CO(CH_2)_2CH_3$	106
67	*O*-Cyclobutanoyl	-CO*c*C_4H_7	39
68	*O*-Cyclohexanoyl	-CO*c*C_6H_{11}	270
69	*O*-Cyclopentanoyl	-CO*c*C_5H_9	132
70	*O*-Cyclopropanoyl	-CO*c*C_3H_5	294
71	*O*-Hexanoyl	-$CO(CH_2)_4CH_3$	186

72	*O*-Isobutyryl	-$COCH(CH_3)_2$	78
73	*O*-Octanoyl	-$CO(CH_2)_6CH_3$	444
74	*O*-Pivaloyl	-$COC(CH_3)_3$	525
75	*O*-Propionyl	-$COCH_2CH_3$	45

Table S2: Complete list of the ligand-based and docking-based descriptors utilized in the correlative studies

Descriptor	**Description**
Rotors	Number of flexible torsions
Atoms	Number of atoms
HeavyAtoms	Number of heavy atoms
Bonds	Number of Bonds
Unsaturations	Number of Unsaturations
HbA	Number of H-bond acceptors
HbD	Number of H-bond donors
HbTot	Number of H-bond groups
Dipole	Dipole moment
Lipole	Lipole moment
Mass	Molecular mass
SAS	Solvent accessible surface
PSA	Polar surface area
Volume	Molecular volume
R_giration	Radius of gyration
Ovality	Molecular Ovality
$logP_{MLP}$	Log P as computed by MLP approach
Docking Score	**Description**
CHARMM	LJ energy computed by CHARMM force-field
CVFF	LJ energy computed by CVFF force-field
Elect	Electrostatic energy as computed by Coulomb's law with and without distance-dependent dielectric function
$Elect_{DD}$	
MLP_{InS}	MLP_{InS} hydrophobic score as computed using different distance functions (linear, quadratic, cubic, Fermi-like)
MLP_{InS}^{2}	
MLP_{InS}^{3}	
MLP_{InS}^{F}	
Contacts	Contacts score based on the number of surrounding residues and normalized per heavy atoms (H) and molecular weight (W)
$Contacts_{H}$	
$Contacts_{W}$	
APBS	Electrostatic energy as computed by APBS
CHEMPLP	Docking score functions as computed by PLANTS and normalized per heavy atoms (H) and molecular weight (W)
$CHEMPLP_{H}$	
$CHEMPLP_{W}$	
PLP	
PLP_{H}	
PLP_{W}	
PLP95	
$PLP95_{H}$	
$PLP95_{W}$	
$XScore_{HP}$	XScore function and its components

$XScore_{HM}$	
$XSScore_{HS}$	
$XScore_{A}$	
$XScore_{B}$	

Table S3 Correlative equations as generated by introducing the score sensitivity values. In all equations n = 75 and $p < 0.0001$ (Min: Y means minimized complexes, N non-minimized complexes).

N.	PDB	scores	Min	Equation, pK =	R^2	Q^2	SE	F
S1	2wij	Eq. 30 + $sens_U$	Y	1.92 – 0.21 $LogP_{MLP}$ – 74.06 $Contacts_W$ + 0.64 PLP95_sens – 15.23 $Xscore_{HS}$_sens – 0.56 PLP_sens	0.68	0.63	0.66	29.89
S2	2wij	Eq. 31 + $sens_U$	N	-0.34 – 99.01 $Contacts_W$ – 0.062 MLP_{InS} – 0.49 $PLP95_H$ – 14.51 $Xscore_{HS}$_sens + 0.41 PLP95_sens	0.70	0.65	0.65	32.73
S3	3o9m	Eq. 32 + $sens_U$	Y	8.86 – 0.32 $LogP_{MLP}$ – 87.38 $Contacts_W$ – 53.48 $ChemPLP_W$_range – 0.76 Vdiam + 12.30 $ChemPLP_H$_sens	0.61	0.55	0.73	21.86
S4	3o9m	Eq. 33 + $sens_U$	N	-0.99 + 0.4123 HBA – 29.15 PLP_W_range + 0.22 MLP_{InS}_range + 0.16 PLP95_sens – 0.28 Contact	0.67	0.62	0.67	28.38
S5	4bds	Eq. 34 + $sens_U$	Y	0.52 + 0.025 PSA – 0.36 Mobility – 2.31 MLP_{InS} – 0.0976 CVFF_sens – 0.5105 $Xscore_{HS}$	0.68	0.62	0.67	29.16
S6	4bds	Eq. 35 + $sens_U$	N	5.97 – 119.41 $Contacts_W$ – 0.5788 MLP_{InS} – 0.61 Mobility + 2.9088 $PLP95_H$_range – 0.7234 PLP95_sens	0.65	0.58	0.73	25.13
S7	4tpk	Eq. 36 + $sens_U$	Y	7.14 – 7.59 $Contacts_H$ – 0.23 Lipole – 1.89 MLP_{InS} + 0.35 MLP_{InS}_sens – 0.86 $Xscore_{HS}$	0.65	0.59	0.72	25.42
S8	4tpk	Eq. 37 + $sens_U$	N	-3.69 + 0.32 HBA – 1.36 MLP_{InS} + 0.59 MLP_{InS}_sens – 0.43 Contacts – 0.041 PLP95	0.69	0.63	0.66	30.76
S9	4xii	Eq. 38 + $sens_U$	Y	-3.05 + 0.51 HBA + 0.86 $RMSD_S$_mean + 0.62 MLP_{InS}_sens – 0.33 Contacts + 26.51 $Contacts_W$_range	0.72	0.67	0.63	34.87
S10	4xii	Eq. 39 +	N	-0.39 + 0.42 HBA + 0.23 ElectDD_sens – 1.01 $Xscore_{HS}$ + 0.012 Elect + 0.76 $RMSD_S$ _ mean	0.65	0.58	0.71	26.31

		$sens_U$						
S11	2wij	Eq. 30 + $sens_S$	Y	7.73 – 87.36 $Contacts_W$ – 0.061 MLP_{InS} + 0.088 PLP95_range – 3.45 $Xscore_{HS}$_sens – 0.78 Vdiam	0.68	0.61	0.68	28.28
S12	2wij	Eq. 31 + $sens_S$	N	-0.05 – 0.56 MLP_{InS} + 0.33 HBA – 39.14 $Contacts_W$ – 0.22 Contacts + 0.0070 Elect	0.71	0.66	0.65	34.21
S13	3o9m	Eq. 32 + $sens_S$	Y	4.98 + 0.22 HBA – 0.25 $LogP_{MLP}$ – 63.10 $Contacts_W$ – 0.12 PLP_sensS – 0.43 Sdiam	0.62	0.56	0.72	22.77
S14	3o9m	Eq. 33 + $sens_S$	N	-2.65 + 0.55 HBA – 0.12 PLP_sens + 0.37 MLP_{InS}_sens – 0.17 Contacts + 0.0065 Elect	0.69	0.63	0.66	30.32
S15	4bds	Eq. 34 + $sens_R$	Y	-2.55 + 0.031 PSA – 3.18 MLP_{InS} – 0.10 CHARMM_sens – 0.021 CVFF_range – 0.021 Bonds	0.67	0.61	0.68	27.49
S16	4bds	Eq. 35 + $sens_S$	N	0.14 + 0.32 HBA – 0.45 MLP_{InS} – 0.16 Mobility – 0.24 Contacts – 12.68 $ChemPLP_W$_sens	0.64	0.57	0.71	24.92
S17	4tpk	Eq. 36 + $sens_S$	Y	6.79 – 6.36 $Contacts_H$ – 0.18 Lipole – 0.41 MLP_{InS} – 0.099 Contacts_sens – 0.81 $Xscore_{HS}$	0.65	0.59	0.70	25.61
S18	4tpk	Eq. 37 + $sens_S$	N	7.24 – 0.17 Lipole – 0.91 $Xscore_{HS}$ - 104.50 $Contacts_W$ – 2.3 MLP_{InS} – 2.92 $Contacts_W$_sensS	0.67	0.65	0.69	27.07
S19	4xii	Eq. 38 + $sens_S$	Y	-1.39 + 0.53 HBA – 0.010 Elect_sens + 0.12 MLP_{InS}_range – 0.27 Contacts + 1.61 $Contacts_H$ _range	0.74	0.69	0.60	39.28
S20	4xii	Eq. 39 + $sens_S$	N	-0.45 + 0.47 HBA – 1.41 $Xscore_{HS}$_sens + 0.036 ElectDD_range – 0.75 $Xscore_{HS}$ + 0.058 APBS	0.63	0.54	0.73	24.25
S21	2wij	Eq. 31 + $sens_M$	Y	1.27 – 0.24 $LogP_{MLP}$ – 70.8459 $Contacts_W$ – 81.28 $ChemPLP_W$_sens – 5.31 $Xscore_{HS}$_sens + 7.99 $PLP95_H$_sens	0.69	0.64	0.66	30.40
S22	2wij	Eq. 31	N	-0.49 – 100.78 $Contacts_W$ – 0.061 MLP_{InS} – 0.52 $PLP95_H$ –	0.70	0.65	0.66	32.20

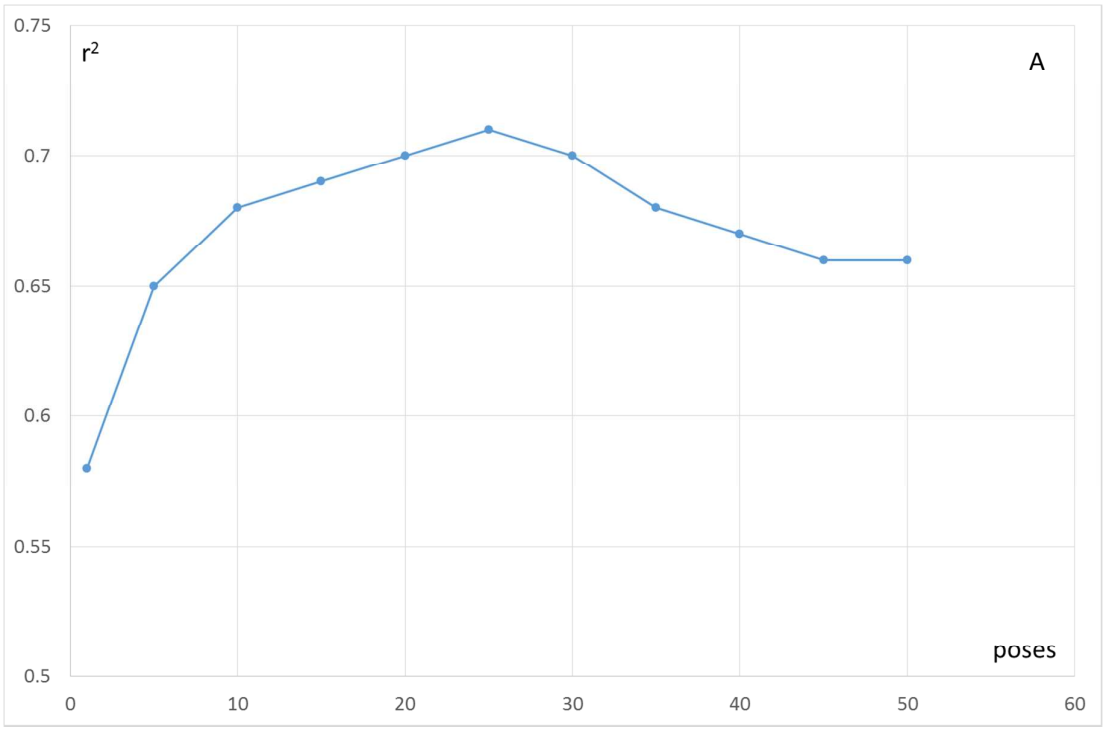
		+ $sens_M$		6.43 $Xscore_{HS}$_sens + 0.19 PLP95_sens				
S23	3o9m	Eq. 32 + $sens_M$	Y	8.98 – 0.31 $LogP_{MLP}$ – 91.11 $Contacts_W$ – 27.79 $ChemPLP_W$_range – 0.77 Vdiam + 57.73 $Contacts_W$_sens	0.61	0.53	0.74	21.20
S24	3o9m	Eq. 33 + $sens_M$	N	-0.82 + 0.42 HBA – 1.9930 PLP_H_range + 0.21 MLP_{InS}_range – 0.27 Contacts + 0.032 PLP95_sens	0.66	0.59	0.69	26.41
S25	4bds	Eq. 34 + $sens_M$	Y	-3.69 + 0.46 HBA + 0.018 Elect_sens – 2.49 MLP_{InS} – 0.025 Bonds – 0.042 CVFF_sens	0.66	0.59	0.69	26.49
S26	4bds	Eq. 35 + $sens_M$	N	3.69 – 120.04 $Contacts_W$ + 0.21 HBA – 0.69 MLP_{InS} – 0.12 Torsions + 1.42 $Xscore_{HS}$_sens	0.64	0.58	0.71	24.91
S27	4tpk	Eq. 36 + $sens_M$	Y	4.98 + 0.20 HBA – 76.86 $Contacts_W$ – 2.50 MLP_{InS} – 0.87 $Xscore_{HS}$ – 0.012 CVFF_sens	0.65	0.59	0.70	25.94
S28	4tpk	Eq. 37 + $sens_M$	N	7.07 – 100.80 $Contacts_W$ – 0.20 Lipole – 2.51 MLP_{InS} – 0.98 $Xscore_{HS}$ + 2.46 $Xscore_{HS}$_sens	0.69	0.63	0.66	30.24
S29	4xii	Eq. 38 + $sens_M$	Y	-6.39 + 0.51 HBA + 0.29 MLP_{InS}_sens – 0.36 Contacts + 0.657 Vdiam + 2.83 $Contacts_H$_range	0.71	0.65	0.64	33.32
S30	4xii	Eq. 39 + $sens_M$	N	-0.45 + 0.44 HBA – 2.30 MLP_{InS} + 0.086 ElectDD_sens – 0.74 $Xscore_{HS}$ + 0.011 Elect	0.65	0.58	0.71	26.06
S31	3o9m	Eq. 32 + sens all	Y	6.68 – 0.33 $LogP_{MLP}$ - 77.56 $Contacts_W$ - 0.57 PLP_$sens_R$ + 0.21 PLP95_$sens_U$ – 0.44 Sdiam	0.64	0.57	0.71	24.05

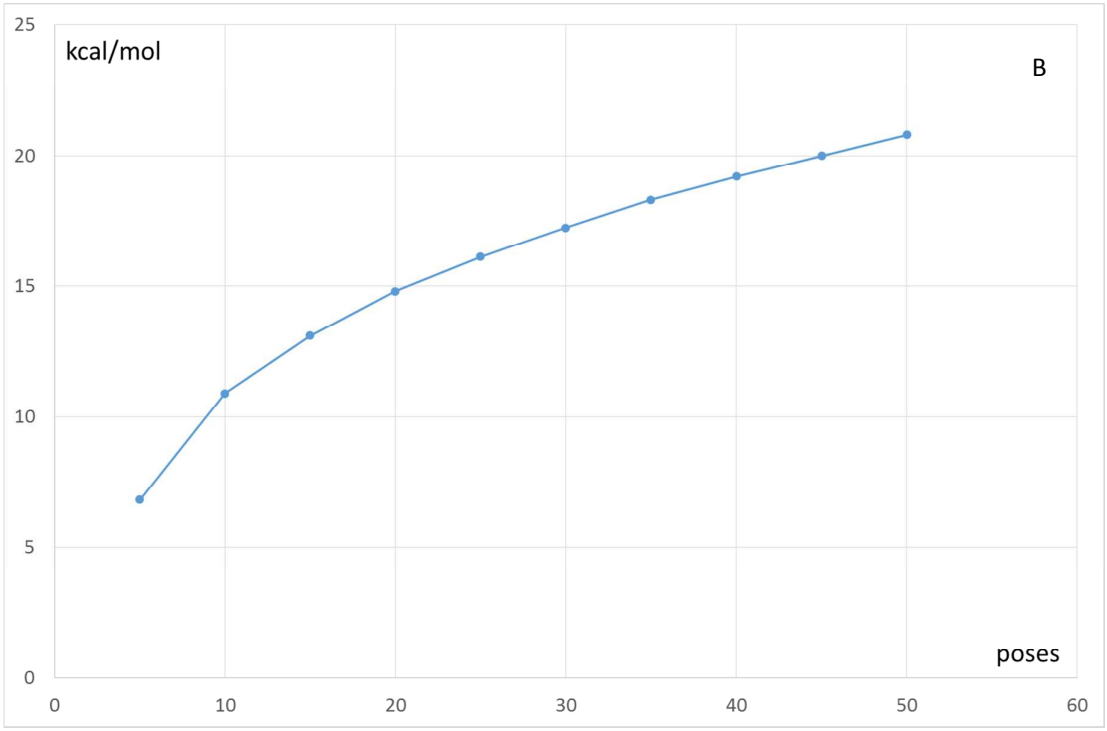
Table S4 Comparison of the coefficients of determination (r^2) as obtained in the different tested cases (Min: Y means minimized complexes, N non-minimized complexes, Δ indicates the differences between minimized and non-minimized complexes).

PDB	Min	best	mean	range	$sens_R$	$sens_U$	$Sens_S$	$Sens_M$	best → mean	mean → range	range → $sens_R$	range → $sens_U$	range → $sens_S$	range → $sens_M$	Total best-sensR	mean
2wij	Y	0.61	0.66	0.68	0.70	0.68	0.68	0.69	0.05	0.02	0.02	0.01	0.01	0.02	0.09	0.67
2wij	N	0.63	0.67	0.70	0.72	0.70	0.71	0.70	0.04	0.03	0.02	0.00	0.01	0.00	0.09	0.69
3o9m	Y	0.60	0.60	0.62	0.64	0.61	0.62	0.61	0.00	0.02	0.02	0.00	0.01	0.00	0.04	0.61
3o9m	N	0.61	0.65	0.67	0.71	0.67	0.69	0.66	0.04	0.02	0.04	0.00	0.02	-0.01	0.10	0.67
4bds	Y	0.57	0.63	0.69	0.68	0.68	0.67	0.66	0.06	0.06	-0.01	0.00	-0.01	-0.02	0.11	0.65
4bds	N	0.57	0.65	0.65	0.66	0.65	0.64	0.64	0.08	0.00	0.01	0.01	0.00	0.00	0.09	0.64
4tpk	Y	0.61	0.66	0.66	0.68	0.65	0.65	0.65	0.05	0.00	0.02	-0.01	-0.01	-0.01	0.07	0.65
4tpk	N	0.65	0.67	0.68	0.71	0.69	0.67	0.69	0.02	0.01	0.03	0.01	-0.01	0.01	0.06	0.68
4xii	Y	0.58	0.68	0.71	0.74	0.72	0.74	0.71	0.10	0.03	0.03	0.01	0.03	0.00	0.16	0.70
4xii	N	0.59	0.63	0.64	0.69	0.65	0.63	0.65	0.04	0.01	0.05	0.02	0.00	0.02	0.10	0.64
mean	Y	0.59	0.65	0.67	0.69	0.67	0.67	0.66	0.05	0.03	0.02	0.00	0.01	0.00	0.09	0.66
mean	N	0.61	0.65	0.67	0.70	0.67	0.67	0.67	0.04	0.01	0.03	0.01	0.00	0.00	0.09	0.66
mean	all	0.60	0.65	0.67	0.69	0.67	0.67	0.67	0.05	0.02	0.02	0.01	0.01	0.00	0.09	0.66

Table S5 Correlative equations as generated to evaluate the effect of docking parameters and reproducibility. In detail, the models were produced by including the mean scores as computed by repeating 5 times the docking simulations on the 4xii structure and minimizing the so generated complexes (Eqs. S32-S35 plus Eq. 28, Table 2). The obtained results emphasize the satisfactory reproducibility of the developed equations which, while including slightly different descriptors, show almost constant statistics as exemplified by the r^2 values which are comprised in a very narrow range equal to 0.02 (i.e. $r^2 = 0.68 \pm 0.01$). Again, the simulations performed by varying the rmsd value (Eqs. S36-S37, Table S5) also produce almost constant results suggesting that the rmsd value, while influencing the clustering of the generated poses, has a limited impact on the calculations of the binding space parameters.

N.	PDB	Scores	Min	Equation, pK =	R^2	Q^2	SE	F
S32	4xii	Mean 2 run	Y	2.30 - 78.49 $Contacts_W$ + 0.32 HbA - 0.19 $LogP_{MLP}$ - 0.068 Torsions - 0.073 Contacts	0.67	0.61	0.68	27.54
S33	4xii	Mean 3 run	Y	5.39 - 86.10 $Contacts_W$ - 0.23 Lipole - 0.94 MLP_{InS} + 0.45 Vdiam - 1.23 $XScore_{HS}$	0.68	0.63	0.67	29.69
S34	4xii	Mean 4 run	Y	0.73 - 114.08 $Contacts_W$ + 0.37 HbA - 0.19 $LogP_{MLP}$ - 0.071 Torsions - 6.78 $ChemPLP_W$	0.69	0.64	0.66	30.49
S35	4xxi	Mean 5 run	Y	2.70 - 7.53 $Contacts_H$ + 0.34 HbA - 0.18 $LogP_{MLP}$ - 0.098 Torsions + 0.0002 APBS	0.68	0.64	0.66	29.27
S36	4xii	Mean rmsd = 1 Å	Y	-0.75 + 9.85 $Contacts_H$ - 0.047 MLP_{InS} + 27.46 PLP_W - 0.71 Contacts - 0.084 PLP95	0.66	0.60	0.69	26.47
S37	4xii	Mean rmsd = 3 Å	Y	1.15 + 0.4654 HbAcc - 5.79 $Contacts_H$ - 0.11 Torsions - 2.83 MLP_{InS} + 0.0056 Elect	0.65	0.59	0.70	25.87


0.75
r²
A
0.7
0.65
0.6
0.55
0.5
poses
0
10
20
30
40
50
60


25
kcal/mol
B
20
15
10
5
0
poses
0
10
20
30
40
50
60

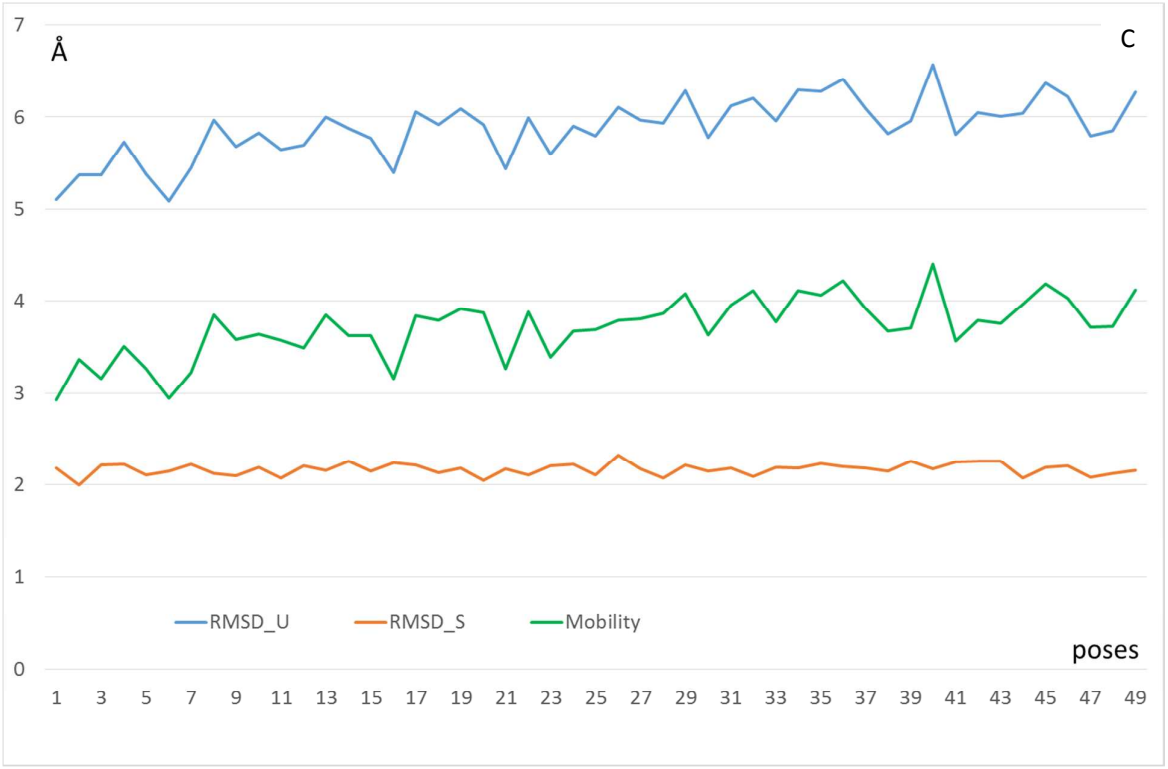


Figure S1 Relationships between the number of poses considered during the calculation of the binding space parameters and: (A) the reliability of the obtained correlative model as encoded by the corresponding r^2 value; (B) the extent of the score range as computed for the primary score function ChemPLP (C), the average values of the three computed rmsd functions. The reported results are obtained by docking simulations on the 4xii structure and by generating 50 poses for each ligand. The poses were rescored as detailed under Methods after complex minimization. Figure S1A shows that the best r^2 value is obtained when averaging 25 poses even though the most marked improvements are obtained when shifting from 1 to 5 poses and from 5 to 10 poses. When considering more than 25 poses the corresponding r^2 value gradually decreases although the beneficial effect of averaging more poses compared to the best pose only is clearly appreciable even when considering all 50 poses. Notably, the equation developed including both ranges and sensitivity values as computed considering the optimum of 25 poses does not show a correspondingly enhanced r^2 value ($r^2 = 0.73$, Eq. not shown) thus suggesting that considering more poses has a beneficial effect on the reliability of the score averages but too wide ranges can negatively impact on the statistics of the obtained models. As depicted in Figure S1B, the score ranges (computed for the ChemPLP function) increase in a logarithmic way when increasing the number of considered poses, thus suggesting that the number of considered poses has to represent a compromise between the exhaustiveness of the average scores and the reliability of the corresponding ranges. In this context, the here used number of 10 poses appears to be a reasonable compromise also balancing the computational cost of the corresponding docking and rescoring calculations. Finally, Figure S1C shows that the rmsd profiles are characterized by a markedly less pronounced variability. In detail, the superimposed rmsd_S values show quite constant average values suggesting that they are influenced by the intrinsic flexibility of the simulated ligands and not by the number of considered poses. Again, the unsuperimposed rmsd_S values and

consequently the mobility values show averages which marginally increase with the number of considered poses and remain in a quite narrow range (5.0 to 6.5 Å).

Table S6 including the data based on the best poses utilized to develop Eq. 1-19 (XLS)

Table S7 including the data based on the multiple poses utilized to develop Eq. 19-59 plus S1-S31 (XLS)

Table S8 reporting the docking scores as computed by ReScore+ for the minimized complexes of the first ten substrates as generated for 4xii (as an example) and describing how binding space parameters can be easily calculated using the address functions in excel (XLS)

Additional references

[1] Bundgaard, H.; Falch, E.; Jensen, E. A Novel Solution-Stable, Water-Soluble Prodrug Type for Drugs Containing a Hydroxyl or an NH-acidic Group. *J Med Chem*. **1989**, 32, 2503-7

[2] Rautio, J.; Taipale, H.; Gynther, J.; Vepsalainen, J.; Nevalainen, T.; Jarvinen T. In Vitro Evaluation of Acyloxyalkyl Esters as Dermal Prodrugs of Ketoprofen and Naproxen. *J Pharm Sci*. **1998**, 87, 1622-8.

[3] Bundgaard, H.; Buur, A.; Shih-Chieh, C.; Lee, V.H. Timolol Prodrugs: Synthesis, Stability and Lipophilicity of Various Alkyl, Cycloalkyl and Aromatic Esters of Timolol *Int J Pharmaceut* **1988**, 46, 77-88

[4] Johansen, M.; Larsen C. A Comparison of the Chemical Stability and the Enzymatic Hydrolysis of a Series of Aliphatic and Aromatic Ester Derivatives of Metronidazole *Int J Pharmaceut* **1985**, 26, 227-241

[5] Larsen, C.; Kurtzhals, P.; Johansen M. Kinetics of Regeneration of Metronidazole from Hemiesters of Maleic Acid, Succinic Acid and Glutaric Acid in Aqueous Buffer, Human Plasma and Pig Liver Homogenate, *Int J Pharmaceut* **1988**, 41, 121-9

[6] Mørk, N.; Bundgaard, H. Stereoselective Enzymatic Hydrolysis of Various Ester Prodrugs of Ibuprofen and Flurbiprofen in Human Plasma. *Pharm Res*. **1992**, 9, 492-6.

[7] Biggadike, K.; Angell, R.M.; Burgess, C.M.; Farrell, R.M.; Hancock, A.P.; Harker, A.J.; Irving W.R.; Ioannou, C.; Procopiou, P.A.; Shaw, R.E.; Solanke, Y.E.; Singh, O.M.; Snowden, M.A.; Stubbs, R.J.; Walton, S.; Weston, H.E. Selective Plasma Hydrolysis of Glucocorticoid Gamma-Lactones and Cyclic Carbonates by the Enzyme Paraoxonase: an Ideal Plasma Inactivation Mechanism *J Med Chem*. **2000**, 43, 19-21.

[8] Nielsen, N.M.; Bundgaard, H. Glycolamide Esters as Biolabile Prodrugs of Carboxylic Acid Agents: Synthesis, Stability, Bioconversion, and Physicochemical Properties. *J Pharm Sci.* **1988**, 77, 285-98.