Structure-based virtual screening and MD refinement reveal novel inhibitors of *Escherichia coli* arginyl-tRNA synthetase

M. Bilal Tayyab[a-e] (https://orcid.org/0000-0003-0166-707X), *Jianshu Dong*[a-f]*

[a] School of Pharmaceutical Sciences, Zhengzhou University, Zhengzhou 450001, China

[b] Key Laboratory of Advanced Drug Preparation Technologies, Ministry of Education; Zhengzhou University, Zhengzhou 450001, China

[c] Key Laboratory of Henan Province for Drug Quality control and Evaluation, Zhengzhou University, Zhengzhou 450001, China

[d] Collaborative Innovation Center of New Drug Research and Safety Evaluation, Zhengzhou University, Zhengzhou 450001, China

[e] Institute of Drug Discovery and Development, Zhengzhou University, Zhengzhou 450001, China

[f] University of Chinese Academy of Sciences, Beijing 100864, China

Corresponding Author

*Prof. Jianshu Dong

School of Pharmaceutical Sciences, Zhengzhou University, Zhengzhou 450001, China

Email: dongjianshu08@mails.ucas.ac.cn

Key Words

Antibacterial resistance, multi-drug resistance (MDR), Escherichia coli, Arginyl tRNA synthetase, Anthraquinone-based ArgRS inhibitors

Supplementary information

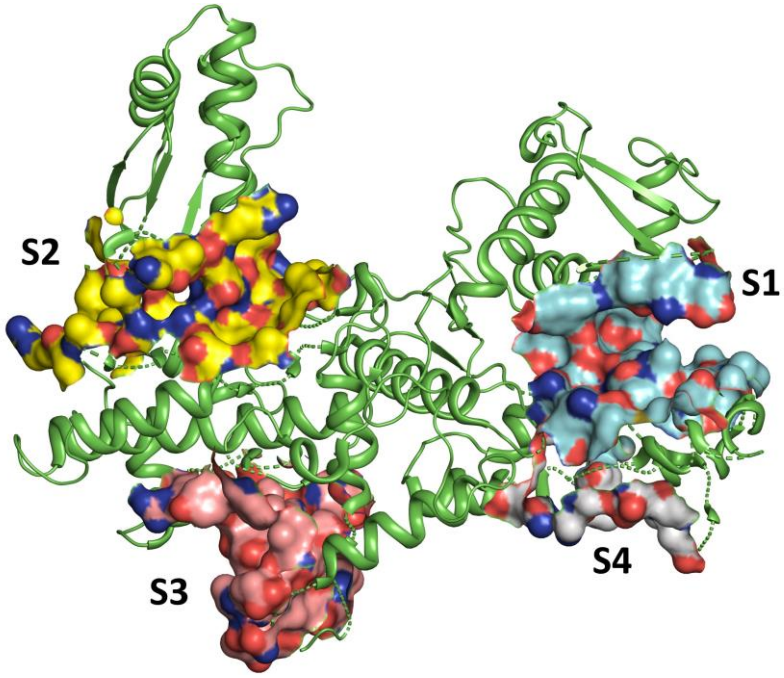


Figure S1: Predicted binding site by Sitemap

The sitemap was used to validate the binding pocket selection in our study. Four sites are predicted on the Arg-tRNA synthetase site which is marked in the figure S1. Table S1 is showing the details about these predicted binding sites. S1 is comprises of Ins1 and Ins2 domains, in which Ins1 domain is included in the catalytic pocket where arginylation process takes place. Ins1 domain of S1 site and Add1 domain of S2 site involves in cognate tRNA recognition and binding via D-loop of tRNA (PDB 4OBY, 5YYN). S1 is the catalytic site, predominantly hydrophobic in nature and facilitates the binding ligands interactions as hydrogen bond acceptor. Residues that are walling the pocket S1 is represented in the table.

Table S1: Predicted sites details obtained by Sitemap

Title	S1	S2	S3	S4
Site Score	1.019	0.976	0.979	0.977

Size	243	143	110	87
Volume	466.13	421.89	290.17	231.52
Exposure	0.52	0.63	0.69	0.55
Enclosure	0.69	0.64	0.62	0.68
Contact	0.92	0.83	0.77	0.82
Phobic/ Philic	0.90	0.23	0.73	1.85
Don/Acc	0.85	1.43	1.15	0.95
Residues	123,124,125,126,158,159,160,162,163,164,165,166,167,168,193,200,210,211,214,215,217,218,219,220,225,258,262,265,306,308,309,310,311,312,313,316	2,4,29,30,31,32,33,34,37,38,86,87,89,488,489,490,491,494,495,498,499,502,505,506,508,509,518,521,522,525	404,405,407,408,409,410,411,415,416,417,419,420,423,424,452,458,461,462,464,465,466,572,573,574,575,576	124,125,172,173,176,178,179,180,184,186,187,188,189,191,192,195,230,231,234,235,238

Table S2: Compounds selected for molecular docking after HTVS

A library of 50,000 FDA approved drug like compounds was retrieved from PubChem database. This library was used to perform virtual screening on the catalytic site S1. The following compounds were showing greater scores as compared to natural substrate L-arginine and selected to perform molecular docking.

PubChem ID	SMILES
2725	Clc1ccc([C@H](CCN(C)C)c2ncccc2)cc1
2726	Clc1cc2N(CCCN(C)C)c3c(Sc2cc1)cccc3
2802	Clc1c(C2=NCC(=O)Nc3c2cc([N+](=O)[O-])cc3)cccc1
2907	ClCCN([P@]1(=O)OCCCN1)CCCl
2950	O=C1c2c(C(=O)c3c1c(O)ccc3)cccc2O
2997	Clc1cc2C(=NCC(=O)Nc2cc1)c1ccccc1
3092	O=C(N)c1c(c([N+](=O)[O-])cc([N+](=O)[O-])c1)C
3182	O[C@@H](Cn1c2c(n(c(=O)n(c2=O)C)C)nc1)CO
3371	FC(F)(F)c1cc(Nc2c(cccc2)C(=O)O)ccc1
3374	Fc1cc2CC[C@H](n3c2c(c(=O)c(c3)C(=O)O)c1)C
3418	P(=O)(Oc1c(cccc1)C(=O)O)(O)O
3559	Clc1ccc(C2(O)CCN(CC2)CCCC(=O)c2ccc(F)cc2)cc1
3676	O=C(Nc1c(cccc1C)C)CN(CC)CC
3758	O=c1n(CC(C)C)c2nc[nH]c2c(=O)n1C
3828	o1c2c(c(OC)c3c(occ3)c2OC)c(=O)cc1C
3869	O[C@H](CN[C@H](CCc1ccccc1)C)c1cc(c(O)cc1)C(=O)N
3883	[S@](=O)(Cc1nccc(OCC(F)(F)F)c1C)c1[nH]c2c(n1)cccc2
3958	Clc1cc2C(=N[C@H](O)C(=O)Nc2cc1)c1c(Cl)cccc1
4044	OC(=O)c1c(Nc2c(c(ccc2)C)C)cccc1
4076	O(c1cc(CCN)cc(OC)c1OC)C
4201	On1c(N)cc(N2CCCCC2)nc1=N
4421	O=c1c2c(n(CC)cc1C(=O)O)nc(cc2)C
4477	Clc1c(NC(=O)c2c(O)ccc(Cl)c2)ccc([N+](=O)[O-])c1
4735	O(CCCCCOc1ccc(cc1)C(=N)N)c1ccc(cc1)C(=N)N
4828	O(C[C@@H](O)CNC(C)C)c1c2c([nH]cc2)ccc1
4839	O([C@H]1CCCN(C1)CC)C(=O)C(c1ccccc1)c1ccccc1

4926	S1c2c(N(CCCN(C)C)c3c1cccc3)cccc2
4946	O(C[C@@H](O)CNC(C)C)c1c2c(ccc1)cccc2
5011	N1(CCNCC1)c1nc2c(cc1)cccc2
5430	s1cc(nc1)c1[nH]c2c(n1)cccc2
5733	Clc1ccc(C(=O)c2n(c(cc2C)CC(=O)O)C)cc1
5816	O[C@H](c1cc(O)c(O)cc1)CNC
6322	OC(=O)[C@@H](N)CCCN=C(N)N
9444	O1[C@@H]([C@@H](O)[C@@H](O)[C@@H]1n1c(=O)nc(nc1)N)CO
10661	Clc1[nH]c2c(n(c(=O)n(c2=O)C)C)n1
164739	O[C@@H]([C@@H](N)C)c1cc(O)c(O)cc1
439260	O[C@H](c1cc(O)c(O)cc1)CN
667468	O1Cc2c(/C(=C/CCN(C)C)/c3c1cccc3)cccc2
5281607	o1c2c(c(=O)cc1c1ccccc1)c(O)cc(O)c2
25273634	Fc1cc2CC[C@@H](Oc2cc1)[C@@H](O)CNC[C@@H](O)[C@@H]1Oc2c(CC1)cc(F)cc2
29985098	Fc1cc2CC[C@@H](Oc2cc1)[C@H](O)CNC[C@H](O)[C@H]1Oc2c(CC1)cc(F)cc2
29985107	Fc1cc2CC[C@@H](Oc2cc1)[C@@H](O)CNC[C@@H](O)[C@H]1Oc2c(CC1)cc(F)cc2
1229	Ic1c(OC)cc(C[C@@H](N)C)c(OC)c1
3127	Clc1ccc(C[C@@H](N)C)cc1
4201	On1c(N)cc(N2CCCCC2)nc1=N
5823	O=C(Nc1ccc([N+](=O)[O-])cc1)N=C(N)N
6971	OC(=O)c1c(NC(=O)C)cccc1
9362	[O-][n+]1c2c(nc3c1cccc3)cccc2
9986	Fc1ccc(C[C@@H](N)C)cc1
10316	[O-]/[N+](=N\c1ccccc1)/c1ccccc1
10777	Clc1c(NC(=O)C)cccc1
11974	Brc1c(OCCCOc2c(Br)cc(cc2)C(=N)N)ccc(c1)C(=N)N
16751	FC(F)(F)c1cc2c(n(c(=O)cc2)C)cc1
20591	Oc1c(N)c(ccc1)C(=O)C
26446	[O-][n+]1c2c(c([N+](=O)[O-])c(c1)C)cccc2
29050	Clc1c2[n+](onc2c([N+](=O)[O-])cc1)[O-]

30760	Clc1cc(NC(=O)N=C(N)N)ccc1C#N
31126	NC(=N)Cc1cc2c(cc1)cccc2
49896	Clc1c(scc1Cl)C[C@@H](N)C
50096	O[C@@H]1c2c(c3c(c(c4c(c3C)cccc4)C)cc2)C=C[C@H]1O
53653	O[C@@H]1c2c(c3nc4c(cc3cc2)cccc4)C=C[C@H]1O
62065	Brc1c(OC)cc(C[C@@H](N)C)c(OC)c1
62787	O(c1c(C[C@@H](N)C)cc(OC)cc1)C
63952	O[C@@H](CN[C@@H](Cc1ccc(OC)cc1)C)c1c2c([nH]c(=O)cc2)c(O)cc1
68754	Clc1c(NCCN=C(N)N)c(Cl)ccc1
72826	[O-][n+]1nc2c(c3c1cccc3)cccc2
74787	[O-][N+]1=C(C(=O)c2c1cccc2)c1ccccc1
74981	Clc1cc2N(CCCN)c3c(Sc2cc1)cccc3
76601	Oc1c(ccc(O)c1)C(=O)N
76632	O(c1c(CCN)cc(OC)cc1)C
77339	Clc1cc2c([n+]([O-])ccc2[N+](=O)[O-])cc1
77804	S(=O)(=O)([O-])c1cc(c2o[n+](CC)cc2)ccc1
77805	S(=O)(=O)(O)c1cc(c2o[n+](CC)cc2)ccc1
77855	Oc1c(c2cc(O)ccc2)ccc(O)c1
85345	[O-][n+]1c2c(c([N+](=O)[O-])cc1)c([N+](=O)[O-])ccc2
97270	OC(=O)c1c2c3c(CCc3ccc2C(=O)O)cc1
99573	[O-][n+]1c2c3c(ccc2ccc1)cccc3
100282	O=C1c2c(c(NCCN)ccc2NCCN)C(=O)c2c1cccc2
100371	O=C1c2c(c(NCCNC)ccc2NCCNC)C(=O)c2c1cccc2
104223	FC(F)(F)c1cc(CCN)ccc1
117475	[O-][N+](=O)c1ccc(cc1)C(=N)N
121501	Fc1cc(C[C@@H](N)C)ccc1
122886	O=C(Nc1c2ncccc2ccc1)N
123686	O(C[C@H](O)CNCCNC(=O)Nc1ccccc1)c1c(cccc1)C#N
123775	Oc1c(CN2CCCC2)cc(Nc2ncnc3c2cccc3)cc1CN1CCCC1
124846	FC(F)(F)c1ccc(NC(=O)CCCC[C@H](NCCc2[nH]cnc2)C)cc1
134019	O=C1c2c(c(NCCN)ccc2NCCN)C(=O)c2c1ccnc2
135468	O=C1c2c(c(NCCN(C)C)ccc2NCCN(C)C)C(=O)c2c1c(O)ccc2O
135740	O(c1c(CCN)cc(OC)c(c1)C)C

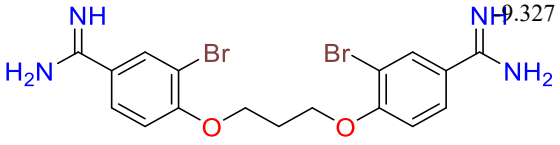
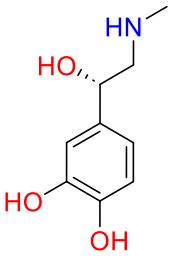
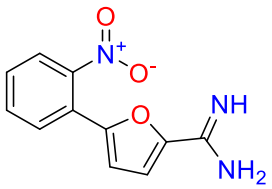
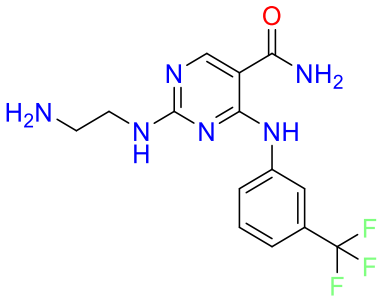
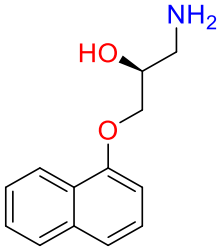
141047	O(c1c(C[C@@H](N)C)ccc(OC)c1)C
151954	O(c1c(CCN)cc(OC)c(OC)c1)C
159599	o1c2c(c(c1C)COCCN)cc1c(oc(=O)cc1C)c2C
159899	O(C[C@@H](O)CN)c1c2c(ccc1)cccc2
163304	o1c(c2c([N+](=O)[O-])cccc2)ccc1C(=N)N
190609	S=P(Oc1nc(N(CC)CC)nc(c1)C)(O)O
198496	S(Cc1cc2nsnc2cc1)C(=N)N
205668	Brc1ccc(C[C@@H](N)C)cc1
248349	Oc1c(c2cc(O)cc(O)c2)c(O)cc(O)c1
340997	n12[nH]cnc2nc(c1/N=N/c1ccccc1)C
348494	Clc1cc(CSC(=N)N)ccc1Cl
396364	Fc1c2oc(c3cc(F)c(NC(=O)[C@@H](N)CCCCN)cc3)cc(=O)c2c(N)c(F)c1C
493974	ClCc1[n+]([O-])c(c2c(n1)cccc2)C
493976	[O-][n+]1c(c2c(nc1C)cccc2)C
494279	O(n1c2c([n+]([O-])c3c(c2)cccc3)c2c1cccc2)C
521099	FC(F)(F)c1cc(cc(c1)C(F)(F)F)CN
719678	Clc1ccc(C[C@@H](N)C)cc1
1551525	Clc1nc(=N)n(O)c(N)c1
2793810	O(C[C@@H](O)CN)c1c2c(ccc1)cccc2
3016271	O=C(c1c(cccc1)C(=O)O)c1cc(N)c(O)cc1
3036381	[O-]/[N+](=C\c1ccccc1)/c1ccccc1
3038498	P(=O)(OCOc1c(C(C)C)cccc1C(C)C)(O)O
3045227	Clc1cc2n(C[C@@H](N)C)ccc2cc1F
3513127	N(=C(N)N)CCCc1ccccc1
4474062	[S@@](=O)(c1c(=O)c2c(n(c1)C)cc(F)cc2)C
6433300	N(=C(N)N)C/C=C/c1ccccc1
6452226	[O-][n+]1nc(N(C)C)nc2c1cc(cc2)C
6603801	Ic1c(OC)cc(C[C@@H](N)C)c(OC)c1
6917740	S(C1=Cc2c(Oc3c1cc(F)cc3)cccc2)CCNC
9549296	FC(F)(F)c1cc(Nc2nc(NCCN)ncc2C(=O)N)ccc1
9573442	Clc1ccc(cc1)/C=N/N=C(N)N
9799509	O=c1c2c(n(CCCN)c3c1cccc3)c(c(cc2)C)C
9800271	[S@](=O)(Cc1nccc(OCC(F)(F)F)c1C)c1[nH]c2c(n1)ccc(O)c2

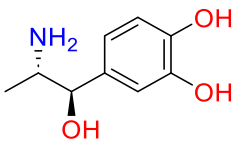
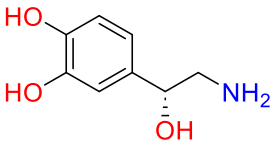
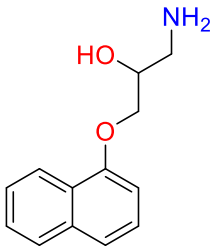
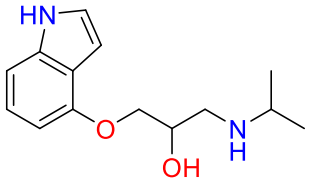
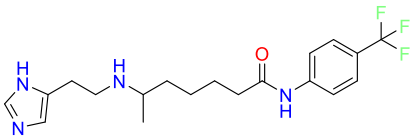
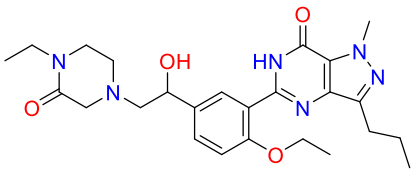
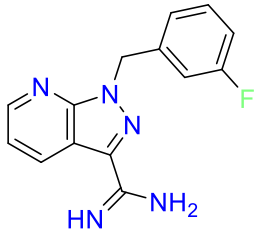
9815618	FC(F)(F)c1cc(C[C@@H](N)C)ccc1
9816106	O[C@]1([C@H](CCCC1)CNC)c1cc(O)ccc1
10036637	O(c1c(CCN)cc(OC)c([N+](=O)[O-])c1)C
10130548	o1c2c(C[C@H](N)C)cccc2cc1
10400521	FC(F)(F)c1c(OC)cc(C[C@@H](N)C)c(OC)c1
10543931	FC(F)(F)c1c2c(OCC2)c(c2c1OCC2)C[C@@H](N)C
11160133	[nH]1c2c(c3c1cccc3)cccc2CCN
11370394	Brc1c(OC)cc([C@@H](O)CN)c(OC)c1
11789033	O(c1c(C[C@@H](N)C)cc(OC)c(c1)C)C
12262505	O(c1c(C[C@H](N)C)cc(OC)cc1)C
12262506	O(c1c(C[C@@H](N)C)cc(OC)cc1)C
12315927	S(c1c(OC)cc(CCN)c(OC)c1)C
12918504	O[C@@H]([C@H]([n+]1noc(/N=C(\[O-])/Nc2ccccc2)c1)C)c1ccccc1
12918505	O[C@@H]([C@H]([n+]1noc(NC(=O)Nc2ccccc2)c1)C)c1ccccc1
12985378	O1c2c(c3n(ncc3cc2)C[C@@H](N)C)CCC1
13530900	OC(=O)Cc1c(CCN(CCC)CCC)cccc1[N+](=O)[O-]
13547416	[125I]c1c(OC)cc(C[C@@H](N)C)c(OC)c1
14320105	O=C1c2c(C(=O)c3c1cccc3)ccc(NCCN)c2
15061891	Brc1cc(C[C@@H](N)C)ccc1
15434373	o1c(=O)c2c(cc(OC)cc2O)cc1C(=O)O
15915372	[S@@](=O)(c1cc(C[C@@H](N)C)c(OC)cc1C)C
20239040	Clc1ccc(NNS(=O)(=O)O)cc1
21072314	Clc1cc(c2cc(ccc2)C(=O)O)c(O)c(N)c1
22019919	NC(c1cc(C(N)(C)C)ccc1)(C)C
23983765	Brc1c(OC)cc(C[C@@H](N)C)c(OC)c1
24729233	O(c1c(CCN)cc(OC)c(c1)CC)C
24937628	S1(=O)(=O)N(CCN2CCNCC2)c2c(N1c1c(F)cccc1F)cccc2
29979100	Clc1c(OC)cc(CCN)c(OC)c1
36689857	O(c1cc(C[C@@H](N)C)cc(OC)c1OC)C
38988468	O(c1c(C[C@@H](N)C)cc(OC)c(OC)c1)C
39235321	Fc1ccc([C@@H]2[C@@H](N)C2)cc1
40425353	O(c1c2c3c([nH]c2ccc1)cccc3)C[C@@H](O)CNCCOc1c(OC)ccc(O)c1
40463719	S(c1c(OC)cc(C[C@@H](N)C)c(OC)c1)C

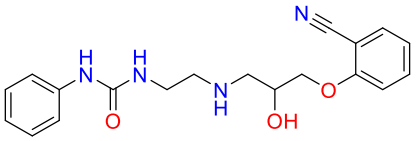
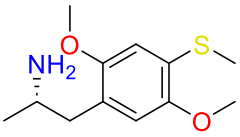
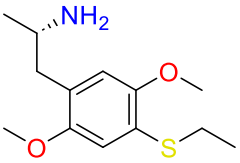
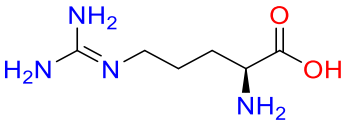
40578305	Fc1cc(C[C@@H](N)C)ccc1
44276763	S(c1c(OC)c(OC)c(CCN)cc1)C
44276771	S(c1c(OC)cc(CCN)cc1OC)C
44276881	S(c1cc(CCN)cc(OC)c1OC)C
44349945	S(c1c(OCC)cc(CCN)cc1OC)CC
44349972	S(c1c(OC)cc(CCN)c(OC)c1)CCF
44350009	S(c1cc(CCN)cc(OC)c1OC)CC
44350108	S(c1c(OC)cc(CCN)c(OC)c1)CCOC
44719499	Fc1c(OC)cc(CCN)c(OC)c1
44719510	O(c1c(OC)cc(CCN)c(OC)c1)C(C)C
45358717	O(C(=O)N1[C@@H](CCCC1)C(=O)Nc1c(cccc1C)C)C(C)(C)C
46188928	Fc1c([C@@H]2N(CCC2)c2nc3n(ncc3NC(=O)N3C[C@@H](O)CC3)cc2)cc(F)cc1
46863664	Clc1ccc(Oc2cc(CN3CCN(CC3)C(=O)Nc3cccnc3)ccc2)cc1
51888193	Fc1cc([C@@H]2[C@@H](N)C2)ccc1F
53491674	Clc1c(cc(NC(=O)Nc2c(F)cc(Oc3cc([n+]([O-])cc3)C(=O)NC)cc2)cc1)C(F)(F)F
57474247	O(c1c(CCN)cc(OC)c(c1)C#N)C
66720197	Fc1ccc(c2nn(c3ncccc3C(=O)N[C@@H](Cc3ccccc3)C(=O)C(=O)N)cc2)cc1
76957261	Clc1cc2c3c4n(c2cc1)C(=O)CC[C@@H]4NCC3
83874760	O1c2c(C[C@H](N)C)cccc2CC1
86312063	Clc1c(OC)cc(C[C@@H](N)C)c(OC)c1
92287301	O1c2cc(C[C@@H](N)C)c(OC)cc2OC1
92950747	O(c1c([C@@H]2[C@@H](N)C2)cc(OC)c(c1)C)C
93954688	Clc1c(F)cc([C@@H]2[C@@H](N)C2)cc1
96656418	Clc1cc([C@@H]2[C@@H](N)C2)ccc1F
101064799	Brc1c(OC)cc(C[C@@H](N)C)cc1OC
101990829	S(c1c(OC)cc(C[C@@H](N)C)c(OC)c1)CC
129601417	Fc1ccc(Cn2c3c(c(C(=O)N[C@H](C(C)(C)C)C(=O)N)c2)cccc3)cc1
130203787	O=C(N)C[C@H](n1ncc(c1)c1ncnc2[nH]ccc12)C1CCCC1
133082522	FC1(F)Oc2cc(C[C@@H](N)C)ccc2O1
137255451	O[C@H](CN1CCN(C(=O)C1)CC)c1cc(c2[nH]c(=O)c3n(nc(c3n2)CCC)C)c(OCC)cc1
145945450	Fc1cc(Cn2nc(c3c2nccc3)C(=N)N)ccc1

146014900	Brc1cc(c(nc1)[C@@H](O)CCCNc1[nH]c(=O)c(Cc2c[n+]([O-])c(cc2)C)cn1)C
154584829	O(CCc1cc2c(n(c3ncccc3[nH]c2=O)CC)nc1)c1c2c([n+]([O-])cc1)cccc2
156596608	O(c1c(OCCOC)cc2ncnc(/N=C\3/C=C(C(=O)C=C3)C#C)c2c1)CCOC
171390172	Fc1cc([C@@H]2[C@@H](N)C2)cc(F)c1
1880	o1c2c(c(=O)cc1c1ccccc1)ccc(O)c2O
1892	O=c1n(c(=O)n(c2ncn(c12)CCO)C)C
2000	OC(=O)[C@H](NC(=O)C)Cc1ccccc1
2016	[n+]1(c2c(cc3c1cc(N)cc3)ccc(N)c2)C
2201	OC(=O)c1c2c(cc3c1cccc3)cccc2
2474	O=C(Nc1c(cccc1C)C)[C@H]1N(CCCC1)CCCC
2554	O=C(N1c2c(C=Cc3c1cccc3)cccc2)N

Table S3: Details of chemical structures and docking scores of compounds other than anthraquinone derivatives, that are showing greater binding than natural substrate L-arginine

PubChem ID	Compound Name	IUPAC Name	Compound Structure	Docking Score (kcal mol^{-1})
11974	Dibrompropamidine	3-bromo-4-[3-(2-bromo-4-carbamimidoylphenoxy)propoxy]benzenecarboximidamide		-9.327
5816	Epinephrine	4-[(1R)-1-hydroxy-2-(methylamino)ethyl]benzene-1,2-diol		-8.639
163304	Nitrafudam	5-(2-nitrophenyl)furan-2-carboximidamide		-8.564
9549296	Syk Inhibitor II	2-(2-aminoethylamino)-4-[3-(trifluoromethyl)anilino]pyrimidine-5-carboxamide		-8.557
2793810	Norpropranolol	(2S)-1-amino-3-naphthalen-1-yloxypropan-2-ol		-8.484

164739	Levonordefrin	4-[(1R,2S)-2-amino-1-hydroxypropyl]benzene-1,2-diol		-8.289
439260	Norepinephrine	4-[(1*R*)-2-amino-1-hydroxyethyl]benzene-1,2-diol		-8.205
159899	N-Desisopropylpropranolol	1-amino-3-naphthalen-1-yloxypropan-2-ol		-8.104
4828	Pindolol	1-(1*H*-indol-4-yloxy)-3-(propan-2-ylamino)propan-2-ol		-8.027
124846	HTFMT	6-[2-(1H-imidazol-5-yl)ethylamino]-N-[4-(trifluoromethyl)phenyl]heptanamide		-7.975
137255451	Piperazinonafil	5-[2-ethoxy-5-[2-(4-ethyl-3-oxopiperazin-1-yl)-1-hydroxyethyl]phenyl]-1-methyl-3-propyl-6*H*-pyrazolo[4,3-d]pyrimidin-7-one		-7.953
145945450	1-(3-Fluorobenzyl)-1H-pyrazolo[3,4-b]pyridine-3-carboximidamide	1-[(3-fluorophenyl)methyl]pyrazolo[3,4-b]pyridine-3-carboximidamide		-7.943

123686	ICI 89406	1-[2-[[3-(2-cyanophenoxy)-2-hydroxypropyl]amino]ethyl]-3-phenylurea		-7.89
40463719	V894NUX7WX	(2*S*)-1-(2,5-dimethoxy-4-methylsulfanylphenyl)propan-2-amine		-7.86
101990829	M8RD23GZ1J	(2*S*)-1-(4-ethylsulfanyl-2,5-dimethoxyphenyl)propan-2-amine		-7.838
6322	L-arginine **(reference)**	(2S)-2-amino-5-(diaminomethylideneamino)pentanoic acid		-7.836

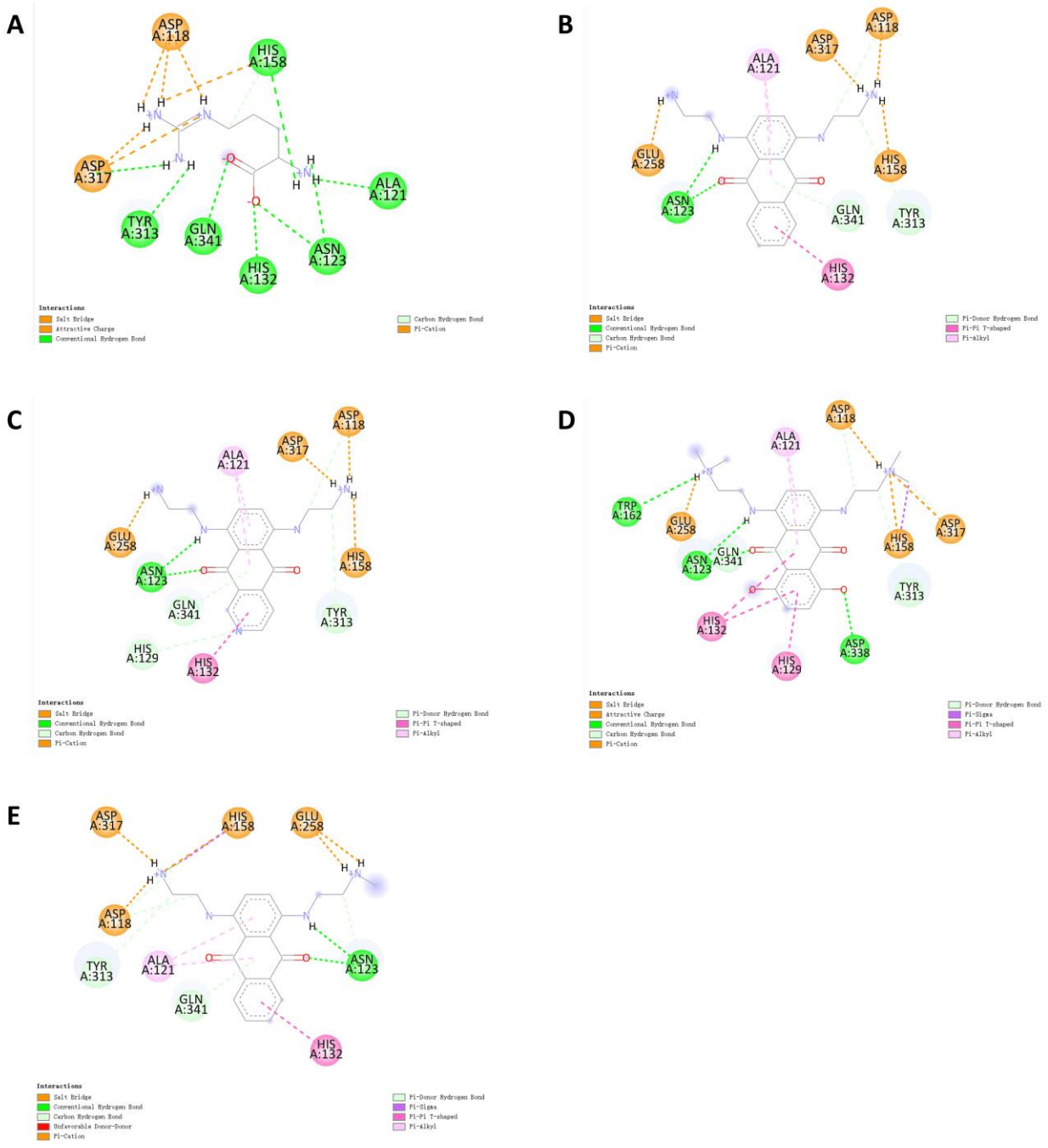


Figure S2: 2D representation of molecular interactions between 4OBY and (A) L-arginine, (B) 100282, (C) 134019, (D) 135468, (E) 100371 after molecular docking. The 2D plots were generated using BIOVIA Discovery Studio.

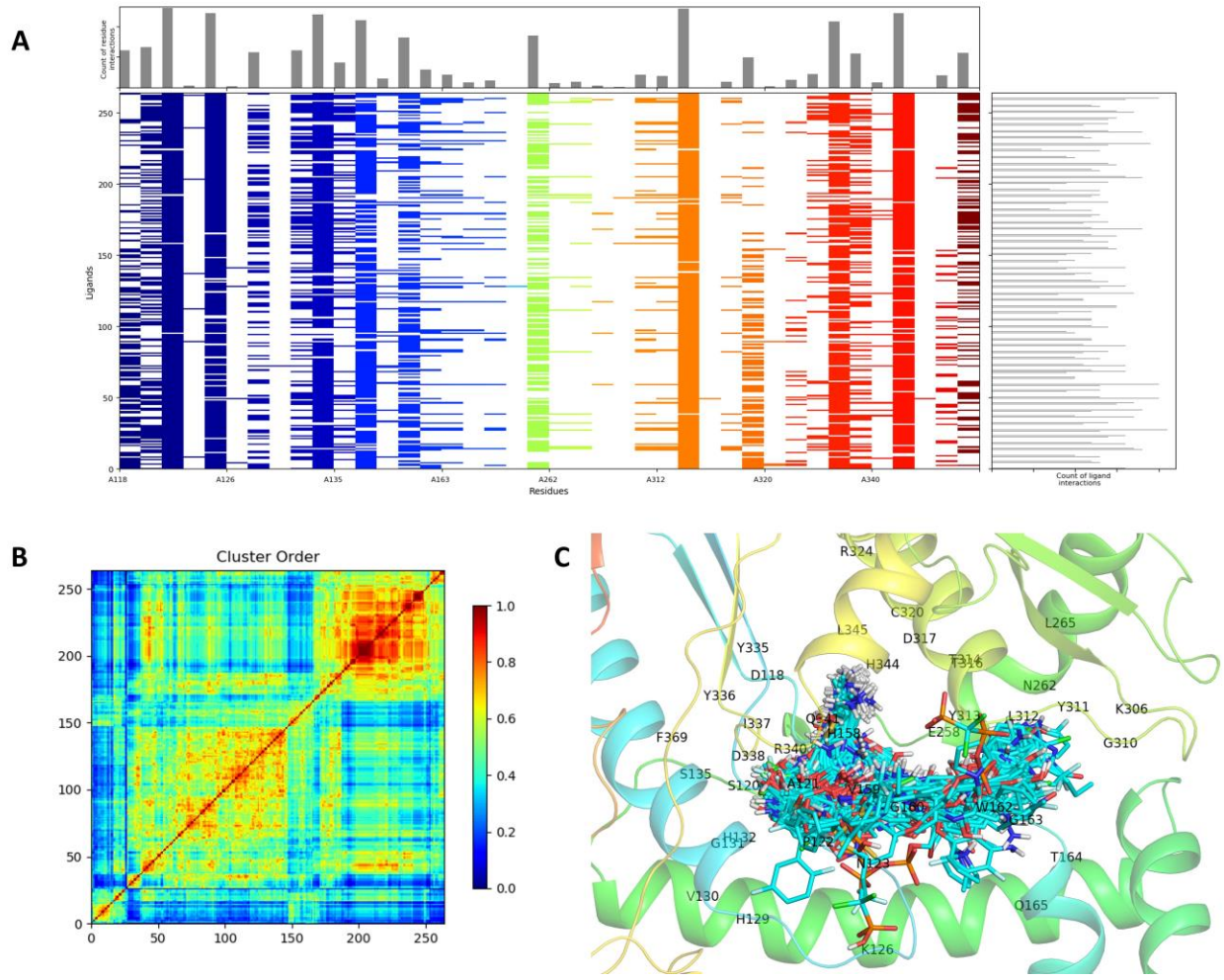


Figure S3: Structural interaction fingerprinting (SIFt) analysis of all compounds selected for molecular docking after virtual screening. (A) Quantitative representation of SIFt data, blue blocks representing polar, hydrogen-bond contacts while red blocks representing non-polar, hydrophobic contacts. **(B)** Cluster analysis of hits based on binding pattern. **(C)** All interactions made by hits during XP docking are represented by superposing them in the center.

DFT Methods details and results explanation

Computed descriptors for L-arginine and key anthraquinone hits are summarized in Table 3: ΔE_gap values were ~0.067 eV (CID100282), ~0.066 eV (CID134019), ~0.071 eV (CID100371), and ~0.077 eV (CID135468). For CID100282, extensive orbital delocalization lowered the ionization potential (≈1.69 eV) and increased chemical softness (≈1.09 eV), yielding deep-blue

MESP lobes at terminal ethylenediamines (cationic donors to Asp118/Asp317) and carbonyl-centered negative zones (acceptors for Asn123/Glu258). In CID134019, LUMO density shifted toward the pyridyl nitrogen, increasing electronegativity ($\chi \approx 3.06$ eV) and electrophilicity ($\omega \approx 0.55$ eV), which oriented the ring nitrogen toward Glu258 in docked poses. In CID100371, removal of the ethylenediamine arms concentrated both frontier orbitals on the anthraquinone core, shrinking positive MESP regions and limiting robust dyad engagement. In CID135468, para-dihydroxyls biased negative potential toward solvent-exposed sites (e.g., Asp338), and tertiary amines diminished terminal basicity, collectively reducing stable Asp118/Asp317 contacts and shifting stabilization toward π-stacking with His129/His132 (see Fig. 4 for representative surfaces and poses).

Post-MD Simulation Results analysis

Quantitative traces are provided in Fig. 5 and Table 4: global Cα RMSD (arginine: rapid rise to ~0.35 nm by 5 ns, then 0.30–0.45 nm; CID100282: smooth increase to ~0.45 nm, plateau by ~45 ns; CID135468: ~0.40 nm by ~25 ns), pocket RMSD (~0.15–0.25 nm for all systems), and ligand RMSD (arginine ~0.45 nm plateau; inhibitors convergent plateaus after the above times). RMSF showed low-mobility helices (<0.10 nm) and mobile loops (190–210, 270–290, 390–410 at ~0.35–0.60 nm), with peak loop fluctuations reduced by ~0.05 nm for CID100282 and slightly more for CID135468 near residues 115–140. Rg values were 2.88 ± 0.03 nm (arginine) and 2.90 ± 0.04 nm (both leads); CID100282 displayed a transient Rg increase to ~3.02 nm at 45–60 ns concurrent with a ligand-RMSD spike, followed by recontraction. Hydrogen-bond counts over 100 ns: arginine 5–8 total; CID100282 3–5; CID135468 2–4, with >90% persistence for a single Asp118 contact in both leads. PCA/FEL: arginine showed a single deep basin from 1–10 ns to 90–100 ns; CID100282 early phase featured multiple shallow minima across PC1 ± 7, PC2 ± 5 $Å^2$ that coalesced into a deeper basin shifted ~1.5 Å along PC1 at 90–100 ns; CID135468 contracted from a $\leqslant$12 kJ mol^{-1} plateau into wells <6 kJ mol^{-1} with stable centroids. Representative late-phase hydrogen-bond triads involved Asp118–Ser120–Asp317 (CID100282) and long-lived Asp118 anchoring with intermittent Asp317 support (CID135468).

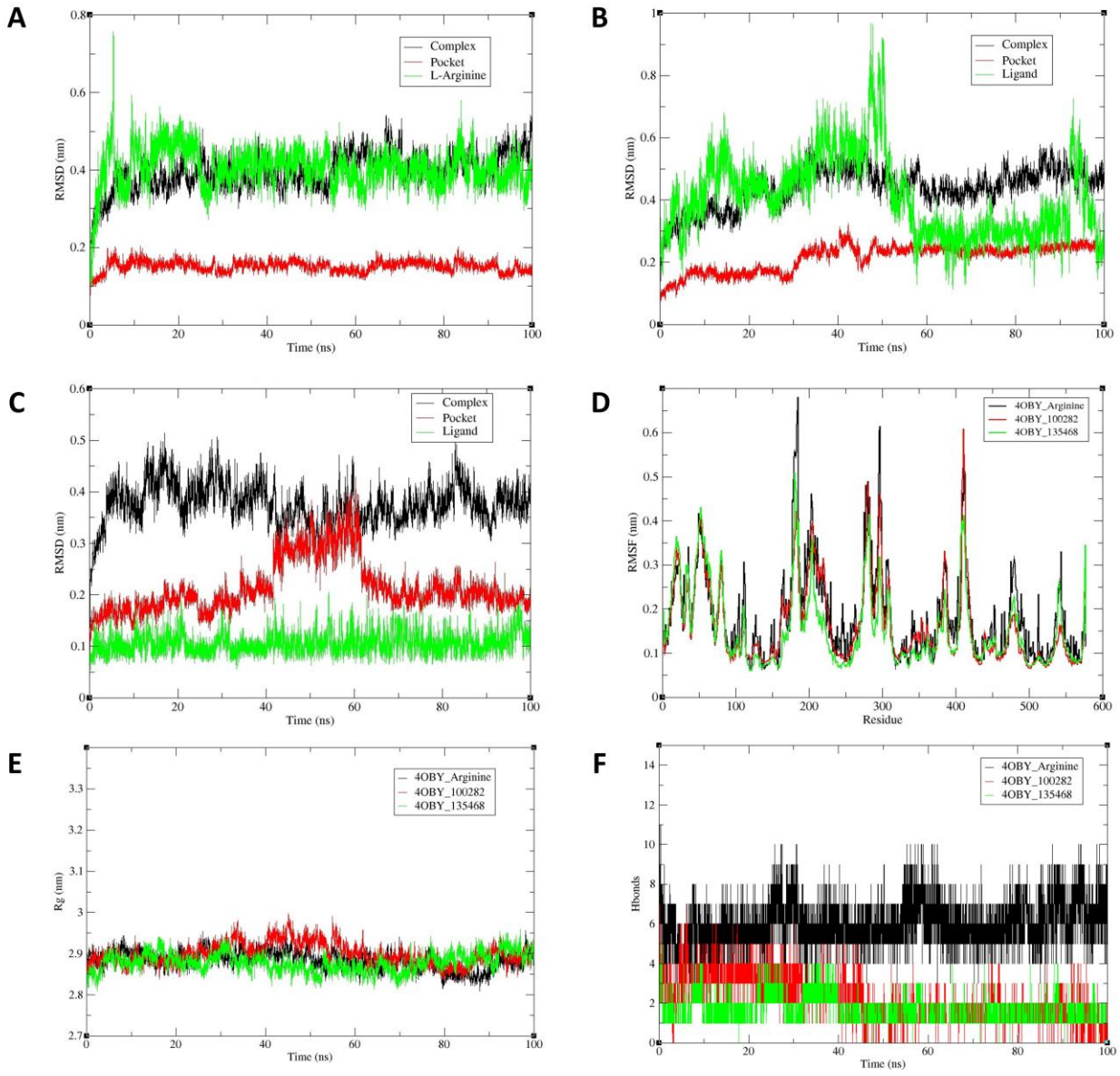


Figure S4: The RMSD results of L-arginine (A), CID100282 (B), and CID135468 (C), RMSF (D), R_g (E), and hydrogen bonds count (F) is shown in the diagram. These results are for the first run. The simulations were re-performed upon the request of the reviewers and the updated results are presented in the original manuscript.

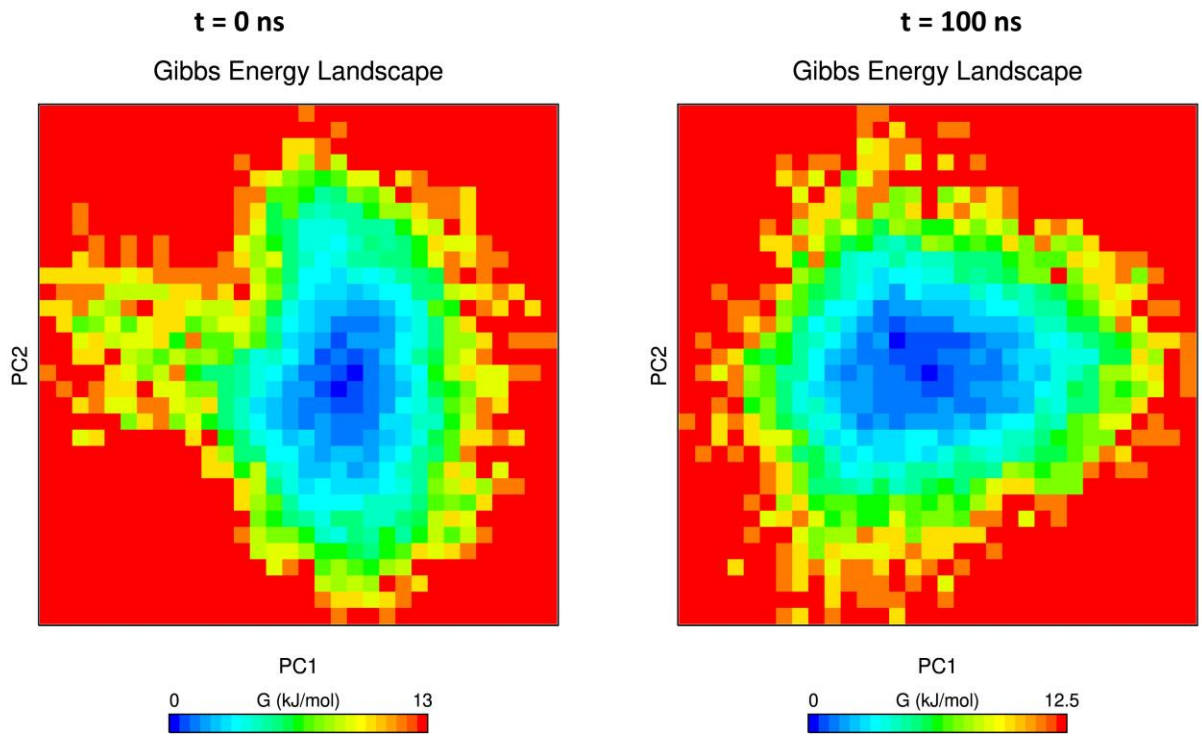


Figure S5: 2D representation of PCA analysis for 4OBY-Arginine complex before and after MD simulation. The PCA analysis was performed by using "gmx covar" and "gmx anaeig" in GROMACS.

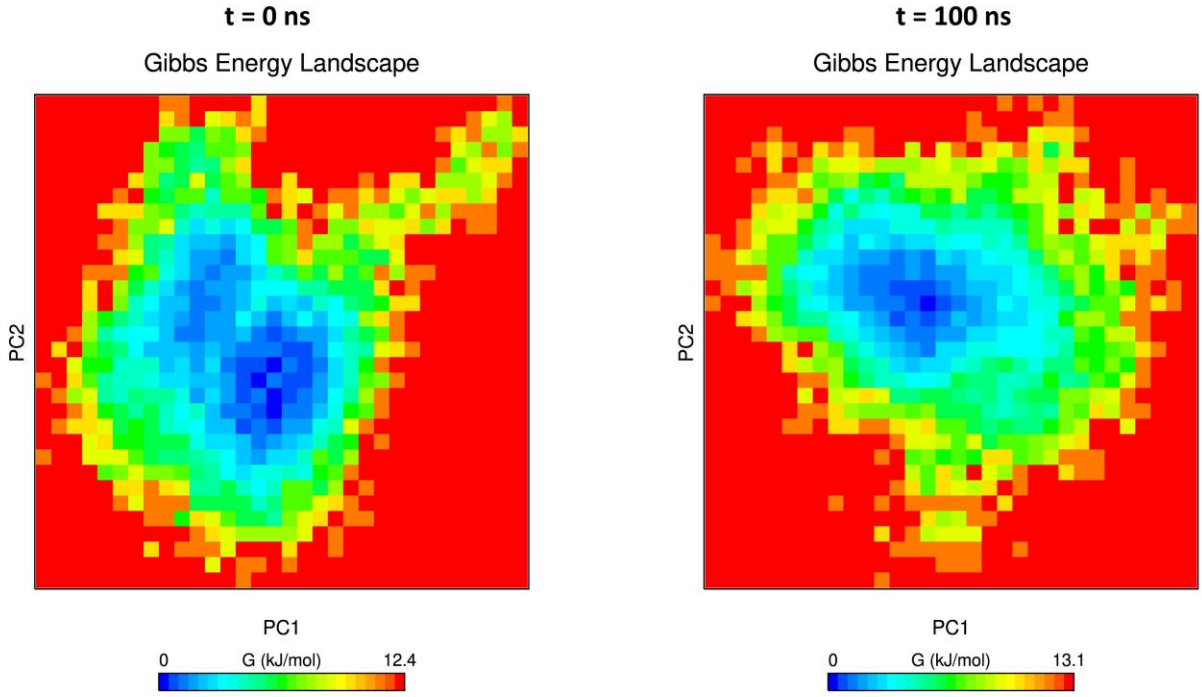


Figure S6: 2D representation of PCA analysis for 4OBY-100282 complex before and after MD simulation. The PCA analysis was performed by using "gmx covar" and "gmx anaeig" in GROMACS.

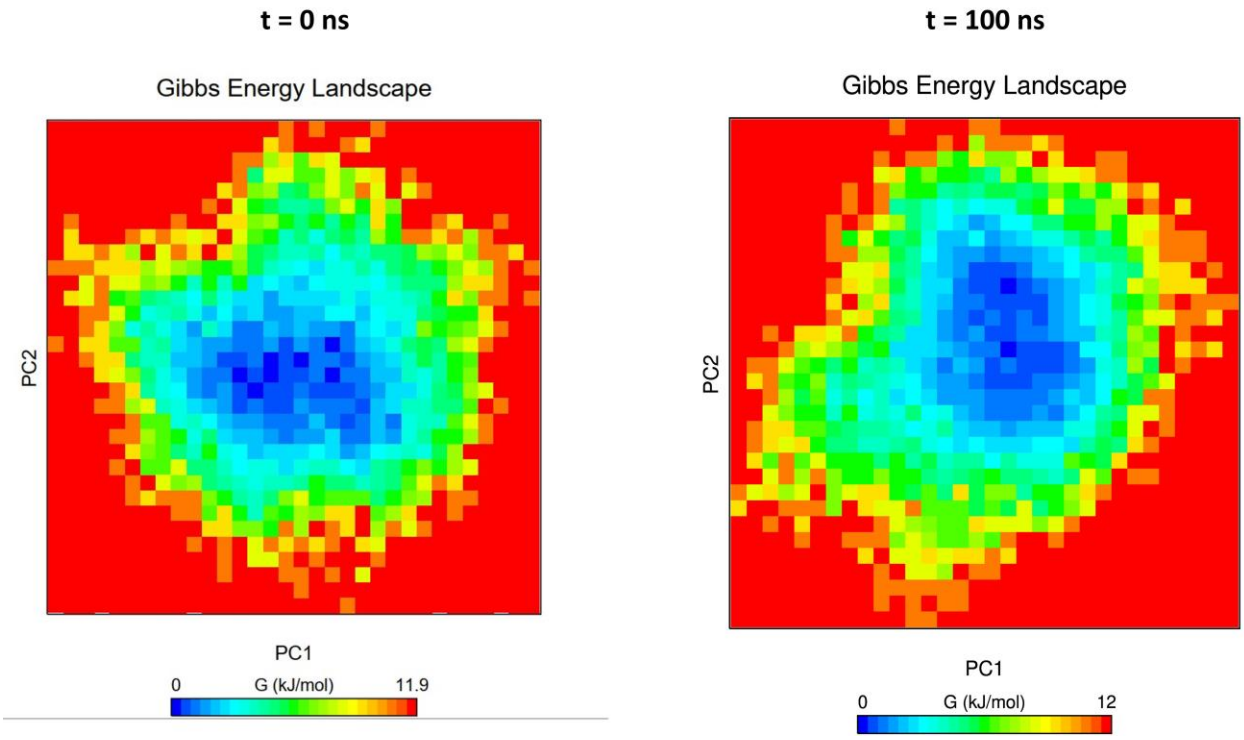


Figure S7: 2D representation of PCA analysis for 4OBY-135468 complex before and after MD simulation. The PCA analysis was performed by using "gmx covar" and "gmx anaeig" in GROMACS.

Table S4: Energy contributions of all residues of S1 pocket eArgRS to the binding of natural arginine and anthraquinone-based inhibitors.

Residue	Arginine	CID100282	CID135468
Asp118	-4.16	-4.62	-1.7
Tyr119	0.77	-0.24	-0.19
Ser120	-1.01	-0.24	-0.27
Ala121	-0.99	-0.79	-0.68
Pro122	-0.27	-0.08	0.09
Asn123	0.02	-0.17	-0.14
Val124	-----	-0.002	-----
Lys126	-----	0.17	0.17
His129	-0.04	0.04	0.03
Val130	-----	-----	-0.07
Gly131	-----	-0.26	-0.29
His132	-0.11	-1.34	-2.31
Arg134	-----	-----	-0.09
Ser135	-2.63	-0.2	-0.65
Ala156	-0.02	-0.09	-----
Asn157	-----	0.16	-----
His158	0.03	-0.16	-0.61
Val159	-0.04	-0.1	-0.04
Gly160	-0.0003	0.02	-0.05
Asp161	-0.02	-0.02	0.0006
Trp162	-----	0.02	0.02
Gly163	-----	0.006	0.02
Thr164	-----	0.005	0.01
Gln165	-----	0.01	0.01
Phe166	-----	0.005	0.005
Gly257	-----	-0.02	-0.03
Glu258	-0.01	0.42	0.52
Leu312	-----	0.007	-0.002
Tyr313	-1.15	-0.93	-0.83
Thr314	-----	-----	-0.02
Thr316	0.06	0.06	-0.03
Asp317	-4.05	-0.16	0.41
Cys320	-0.01	-0.44	-0.36
Arg324	0.74	1.91	0.81
Tyr335	-0.1	-1.95	-0.65
Tyr336	-1.11	-0.15	-0.06
Ile337	-3.67	-1.39	-2.57

Asp338	-1.77	-0.19	0.16
Arg340	-----	0.03	-0.01
Gln341	-1.21	-0.54	-1.09
His344	0.22	0.02	0.12
Leu345	-0.89	-0.15	-0.56
Trp349	-0.01	0.006	0.01
Phe369	-0.19	-----	-0.36

Table S5: All third-party software or online tools used, their versions and download or accessible links.

Software / Online tool	**Version**	**Download or public access link**
PDB		https://www.rcsb.org/
PubChem		https://pubchem.ncbi.nlm.nih.gov/
SwissADME		http://www.swissadme.ch/
ADMET lab	3.0	https://admetlab3.scbdd.com/server/evaluation
PyMOL	Molecular graphics and figure preparation were performed with PyMOL 2.5.4 (academic build; Schrödinger, LLC).	The binary installer for non-commercial research use is freely available after academic registration at https://pymol.org/2/
Schrodinger	Schrodinger 2020-3	https://www.schrodinger.com/
Gaussian 09	Version 9.5	https://gaussian.com/
GROMACS	2024.4 version	https://manual.gromacs.org/2024.4/install-guide/index.html
gmx_MMPBSA		https://valdes-tresanco-ms.github.io/gmx_MMPBSA/dev/