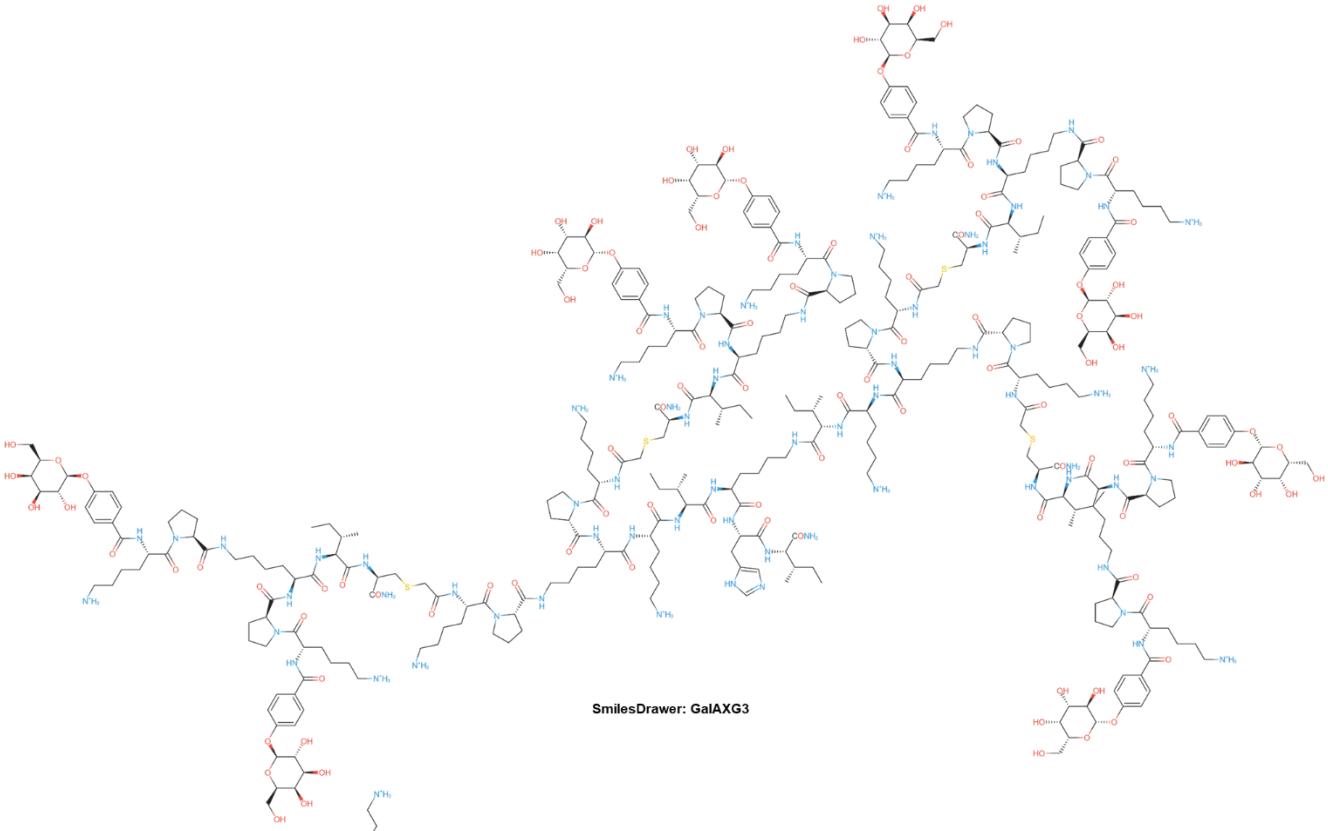
Supporting Information for:

SmilesDrawer: parsing and drawing SMILES-encoded molecular structures using client-side JavaScript

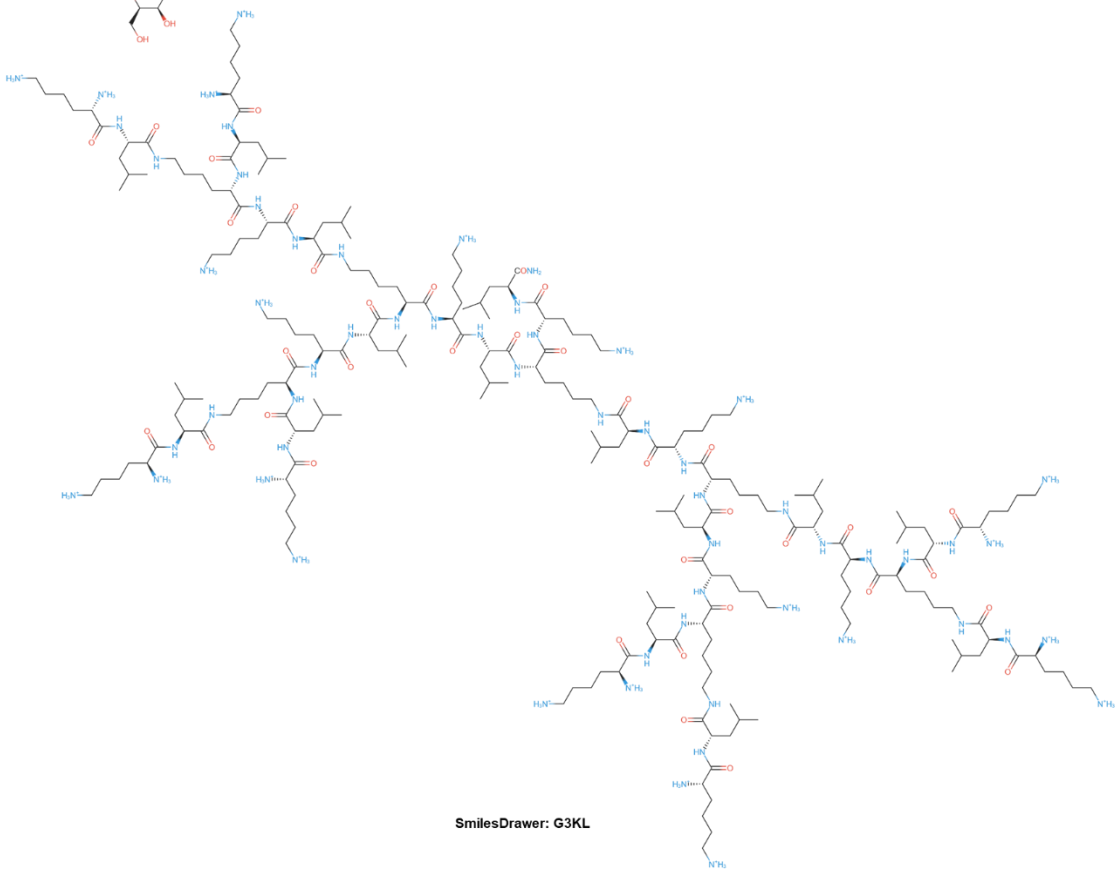
Daniel Probst[1,*] and Jean-Louis Reymond[1,*]

[1]*Department of Chemistry and Biochemistry, National Center for Competence in Research NCCR TransCure, University of Berne, Freiestrasse 3, 3012 Berne, Switzerland*

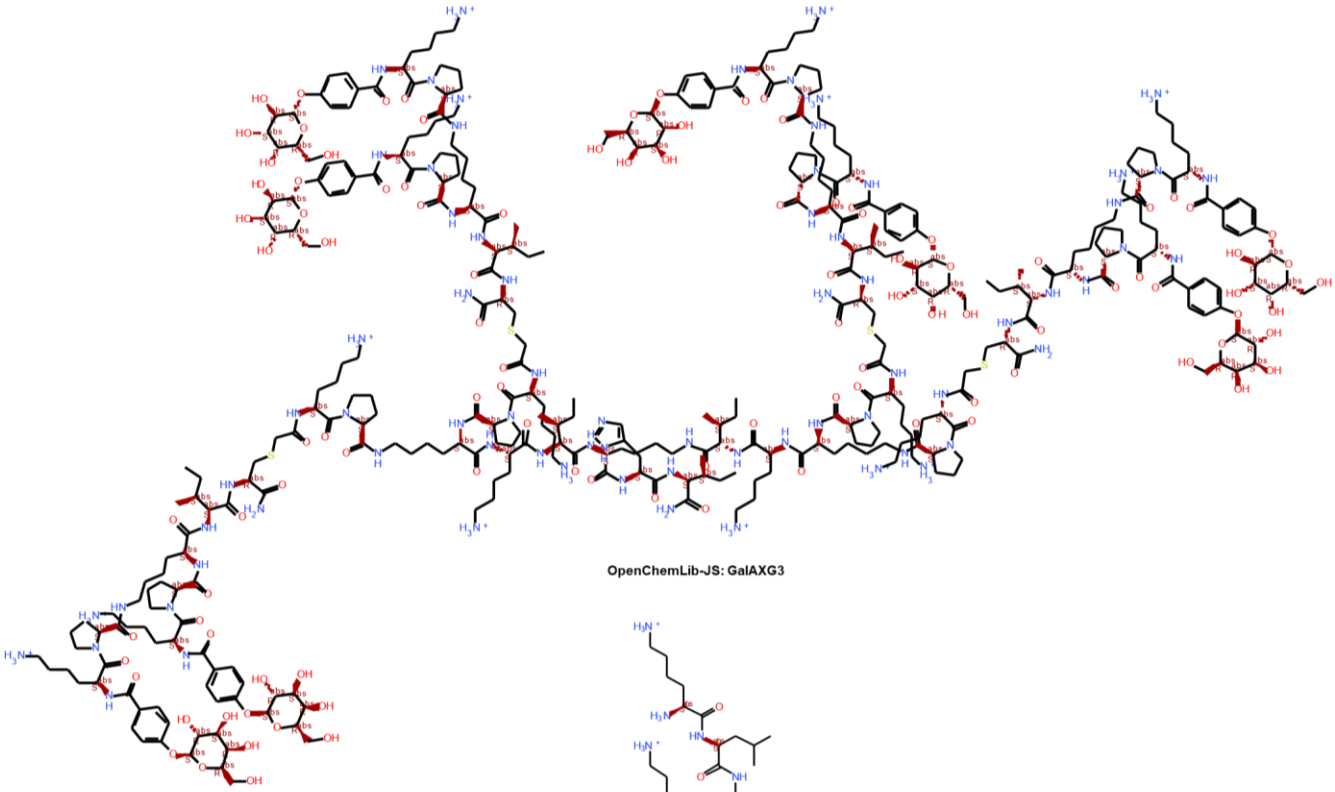
[*]e-mail: daniel.probst@dcb.unibe.ch, jean-louis.reymond@dcb.unibe.ch



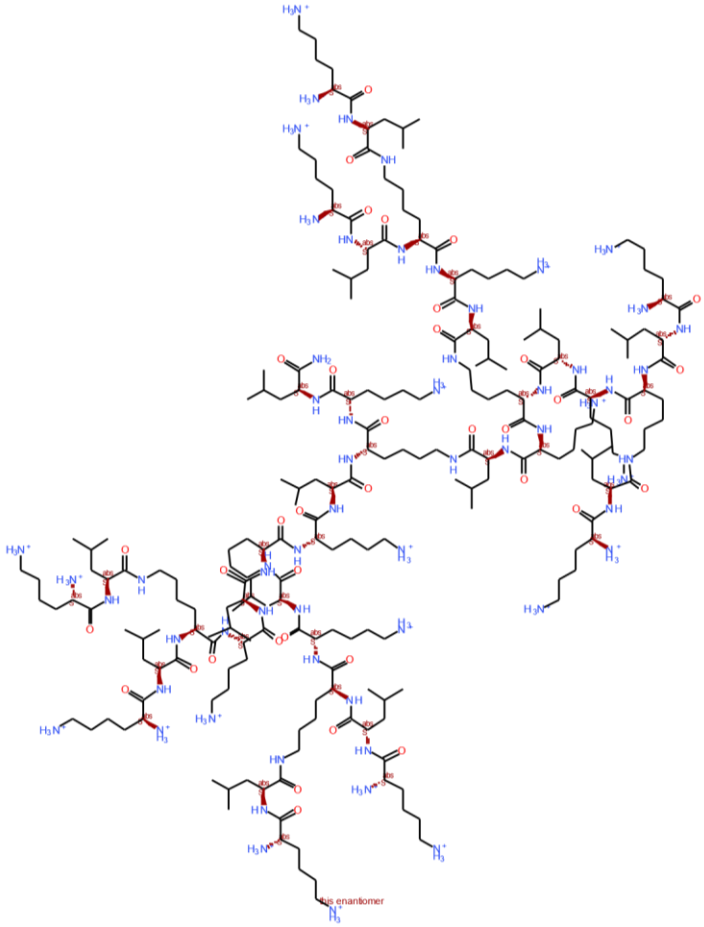
SmilesDrawer: GalAXG3



SmilesDrawer: G3KL



OpenChemLib-JS: GalAXG3



OpenChemLib-JS: G3KL

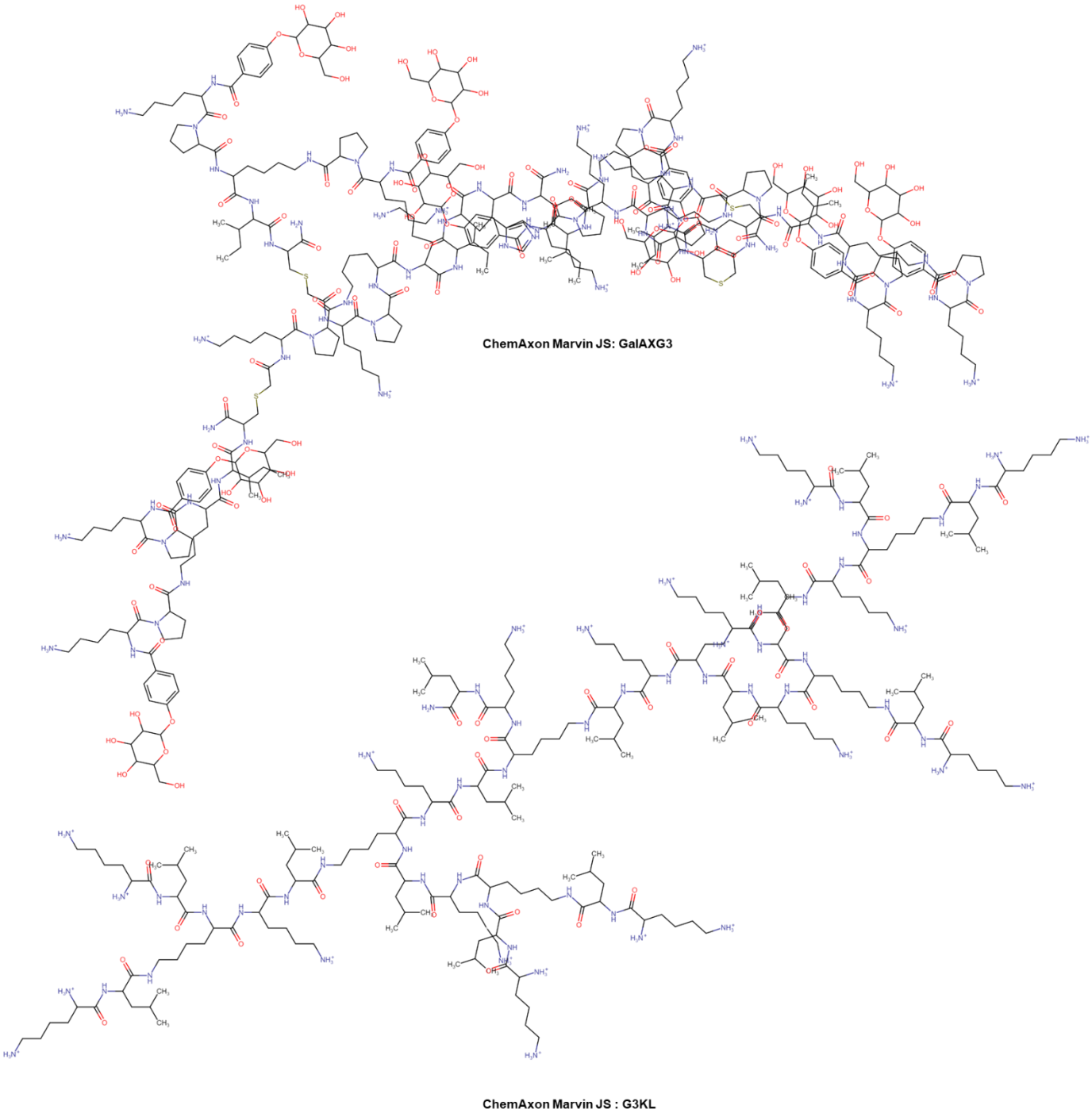


Fig. S1 Drawing of large peptides using SmilesDrawer, OpenChemLib-JS and ChemAxon Marvin JS. Comparisons of large peptide depictions show that SmilesDrawer excels at both keeping the main chain of a peptide straight and avoiding overlaps. The top drawings each depict glycopeptide dendrimer **GalAXG3** rendered with SmilesDrawer, OpenChemLib-JS and ChemAxon Marvin JS. Bottom drawings show the antimicrobial peptide dendrimer **G3KL** rendered with SmilesDrawer, OpenChemLib-JS and ChemAxon Marvin JS.

Table. S1 SMILES strings used in this article. (1-16) SMILES used in Figure 1. (17, 18) SMILES used in Figure S1.

#	**Compound Name**	**SMILES**
1	cubane	C12C3C4C1C5C2C3C45
2	trinorbornane	C1CC23CCC4CC2C1CC34
3	heptacyclo[6.4.0.02,7.03,6.04,11.05,10.09,12]dodecane	C12C3C4C1C1C4C4C5C3C2C5C14
4	dodecahedrane	C12C3C4C5C1C6C7C2C8C3C9C4C1C5C6C2C7C8C9C12
5	buckminsterfullerene	C12=C3C4=C5C6=C1C7=C8C9=C1C%10=C%11C(=C29)C3=C2C3=C4C4=C5C5=C9C6=C7C6=C7C8=C1C1=C8C%10=C%10C%11=C2C2=C3C3=C4C4=C5C5=C%11C%12=C(C6=C95)C7=C1C1=C%12C5=C%11C4=C3C3=C5C(=C81)C%10=C23
6	vitamin B12	C[C@H](CNC(=O)CC[C@]1(C)[C@@H](CC(N)=O)[C@H]2[N]3=C1C(C)=C1[C@@H](CCC(N)=O)C(C)(C)C4=[N]1[Co+]31(C#N)N3C([C@@H](CCC(N)=O)[C@](C)(CC(N)=O)[C@]23C)=C(C)C2=[N]1C(=C4)[C@@H](CCC(N)=O)[C@]2(C)CC(N)=O)OP([O-])(=O)O[C@@H]1[C@@H](CO)O[C@@H]([C@@H]1O)N1C=NC2=C1C=C(C)C(C)=C2
7	hydromethylglutaryl-coenzyme A	CN[C@H](CC(C)C)C(=O)N[C@@H]1[C@H](O)C2=CC=C(OC3=C(O[C@@H]4O[C@H](CO)[C@@H](O)[C@H](O)[C@H]4O[C@H]4C[C@](C)(N)[C@H](O)[C@H](C)O4)C4=CC(=C3)[C@@H](NC(=O)[C@H](CC(N)=O)NC1=O)C(=O)N[C@@H]1C3=CC(=C(O)C=C3)C3=C(C=C(O)C=C3O)[C@H](NC(=O)[C@@H](NC1=O)[C@H](O)C1=CC=C(O4)C(Cl)=C1)C(O)=O)C(Cl)=C2
8	cholesterol	C[C@H](CCCC(C)C)[C@H]1CC[C@@H]2[C@@]1(CC[C@H]3[C@H]2CC=C4[C@@]3(CC[C@@H](C4)O)C)C
9	arachidonic acid	CCCCC/C=C\C/C=C\C/C=C\C/C=C\CCCC(=O)O
10	prostaglandin E1	CCCCC[C@@H](/C=C/[C@H]1[C@@H](CC(=O)[C@@H]1CCCCCCC(=O)O)O)O
11	strychnine	O=C1C[C@@H]2OCC=C3CN4CC[C@@]56[C@@H]4C[C@@H]3[C@@H]2[C@@H]5N1C1=CC=CC=C61
12	quinine	OC1=CC2=C(C=C1)N=CC=C2[C@@H](O)[C@@H]1C[C@@H]2CCN1C[C@@H]2C=C
13	vancomycin	CN[C@H](CC(C)C)C(=O)N[C@@H]1[C@H](O)C2=CC=C(OC3=C(O[C@@H]4O[C@H](CO)[C@@H](O)[C@H](O)[C@H]4O[C@H]4C[C@](C)(N)[C@H](O)[C@H](C)O4)C4=CC(=C3)[C@@H](NC(=O)[C@H](CC(N)=O)NC1=O)C(=O)N[C@@H]1C3=CC(=C(O)C=C3)C3=C(C=C(O)C=C3O)[C@H](NC(=O)[C@@H](NC1=O)[C@H](O)C1=CC=C(O4)C(Cl)=C1)C(O)=O)C(Cl)=C2
14	calicheamycin γ1	CCN[C@H]1CO[C@H](C[C@@H]1OC)O[C@@H]1[C@@H](O)[C@H](NO[C@H]2C[C@H](O)[C@H](SC(=O)C3=C(OC)C(OC)=C(O[C@@H]4O[C@@H](C)[C@H](O)[C@@H](OC)[C@H]4O)C(I)=C3C)[C@@H](C)O2)[C@@H](C)O[C@H]1O[C@H]1C#C\C=C/C#C[C@@]2(O)CC(=O)C(NC(=O)OC)=C1\C2=C/CSSSC
15	cyclosporin A	CC[C@H]1C(=O)N(CC(=O)N([C@H](C(=O)N[C@H](C(=O)N([C@H](C(=O)N[C@H](C(=O)N[C@@H](C(=O)N([C@H](C(=O)N([C@H](C(=O)N([C@H](C(=O)N1)[C@@H]([C@H](C)C/C=C/C)O)C)C(C)C)CC(C)C)C)CC(C)C)C)C)CC(C)C)C)C(C)C)CC(C)C)C)
16	maitotoxin	C[C@H](CC[C@@H]([C@@H]([C@H](C)C[C@H](C(=C)/C(=C/CO)/C)O)O)OS(=O)(=O)[O-])[C@H]([C@@H](C)[C@H]1[C@@H]([C@@H]([C@H]2[C@H](O1)[C@@H](C[C@]3([C@H](O2)C[C@H]4[C@H](O3)C[C@]5([C@H](O4)[C@H]([C@H]6[C@H](O5)C[C@H]([C@H](O6)[C@@H]([C@H](C[C@H]7[C@@H]([C@@H]([C@H]8[C@H](O7)C[C@H]9[C@H](O8)C[C@H]1[C@H](O9)[C@H]([C@@H]2[C@@H](O1)[C@@H]([C@H]([C@@H](O2)[C@H]1[C@@H]([C@H]([C@H]2[C@@H](O1)C[C@H]([C@@H](O2)[C@@H](C[C@H](C[C@H]1[C@@H]([C@H]([C@H]2[C@@H](O1)C[C@H]([C@@H](O2)[C@H]1[C@@H](C[C@]2([C@H](O1)[C@@H]([C@]1([C@H](O2)C[C@]2([C@H](O1)CC[C@]1([C@H](O2)C[C@]2([C@H](O1)C[C@H]1[C@H](O2)CC[C@H](O1)[C@]1([C@@H](C[C@H]2[C@](O1)(C[C@H]1[C@](O2)(CC[C@]2([C@H](O1)C[C@H]1[C@](O2)(C[C@H]2[C@H](O1)C/C=C\[C@H]1[C@H](O2)C[C@H]2[C@](O1)(C[C@]1([C@H](O2)C[C@H]2[C@](O1)(CC[C@H](O2)[C@H]([C@@H](C[C@@H](C)[C@@H](C)CC=C)O)O)C)C)C)C)C)C)C)O)C)C)C)C)O)C)O)O)O)O)O)O)O)O)O)O)O)O)O)OS(=O)(=O)[O-])O)O)O)O)C)C)O)O)O)O

17	G3KL	O=C(N[C@@H](CC(C)C)C(N)=O)[C@@H](NC([C@H](CCCCNC([C@H](CC(C)C)NC([C@@H](NC([C@H](CCCCNC([C@H](CC(C)C)NC([C@@H](NC([C@H](CCCCNC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)NC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)=O)CCCC[NH3+])=O)=O)NC([C@H](CC(C)C)NC([C@@H](NC([C@H](CCCCNC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)NC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)=O)CCCC[NH3+])=O)=O)=O)CCCC[NH3+])=O)=O)NC([C@H](CC(C)C)NC([C@@H](NC([C@H](CCCCNC([C@H](CC(C)C)NC([C@@H](NC([C@H](CCCCNC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)NC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)=O)CCCC[NH3+])=O)=O)NC([C@H](CC(C)C)NC([C@@H](NC([C@H](CCCCNC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)NC([C@H](CC(C)C)NC([C@@H]([NH3+])CCCC[NH3+])=O)=O)=O)CCCC[NH3+])=O)=O)=O)CCCC[NH3+]
18	GalAXG3	O=C([C@H](CCCCNC([C@@H](NC([C@@H](NC([C@H](CCCCNC([C@H]1N(C([C@@H](NC(CSC[C@H](NC([C@H]([C@@H](C)CC)NC([C@H](CCCCNC([C@@H]2CCCN2C([C@@H](NC(C3=CC=C(O[C@H]4[C@H](O)[C@@H](O)[C@@H](O)[C@@H](CO)O4)C=C3)=O)CCCC[NH3+])=O)=O)NC([C@@H]5CCCN5C([C@@H](NC(C6=CC=C(O[C@H]7[C@H](O)[C@@H](O)[C@@H](O)[C@@H](CO)O7)C=C6)=O)CCCC[NH3+])=O)=O)=O)=O)C(N)=O)=O)CCCC[NH3+])=O)CCC1)=O)NC([C@H]8N(C([C@@H](NC(CSC[C@H](NC([C@H]([C@@H](C)CC)NC([C@H](CCCCNC([C@@H]9CCCN9C([C@@H](NC(C%10=CC=C(O[C@@H]%11O[C@H](CO)[C@H](O)[C@H](O)[C@H]%11O)C=C%10)=O)CCCC[NH3+])=O)=O)NC([C@@H]%12CCCN%12C([C@@H](NC(C%13=CC=C(O[C@@H]%14O[C@H](CO)[C@H](O)[C@H](O)[C@H]%14O)C=C%13)=O)CCCC[NH3+])=O)=O)=O)=O)C(N)=O)=O)CCCC[NH3+])=O)CCC8)=O)=O)CCCC[NH3+])=O)[C@@H](C)CC)=O)NC([C@@H](NC([C@@H](NC([C@H](CCCCNC([C@H]%15N(C([C@@H](NC(CSC[C@H](NC([C@H]([C@@H](C)CC)NC([C@H](CCCCNC([C@@H]%16CCCN%16C([C@@H](NC(C%17=CC=C(O[C@@H]%18O[C@H](CO)[C@H](O)[C@H](O)[C@H]%18O)C=C%17)=O)CCCC[NH3+])=O)=O)NC([C@@H]%19CCCN%19C([C@@H](NC(C%20=CC=C(O[C@@H]%21O[C@H](CO)[C@H](O)[C@H](O)[C@H]%21O)C=C%20)=O)CCCC[NH3+])=O)=O)=O)=O)C(N)=O)=O)CCCC[NH3+])=O)CCC%15)=O)NC([C@H]%22N(C([C@@H](NC(CSC[C@H](NC([C@H]([C@@H](C)CC)NC([C@H](CCCCNC([C@@H]%23CCCN%23C([C@@H](NC(C%24=CC=C(O[C@H]%25[C@H](O)[C@@H](O)[C@@H](O)[C@@H](CO)O%25)C=C%24)=O)CCCC[NH3+])=O)=O)NC([C@@H]%26CCCN%26C([C@@H](NC(C%27=CC=C(O[C@H]%28[C@H](O)[C@@H](O)[C@@H](O)[C@@H](CO)O%28)C=C%27)=O)CCCC[NH3+])=O)=O)=O)=O)C(N)=O)=O)CCCC[NH3+])=O)CCC%22)=O)=O)CCCC[NH3+])=O)[C@@H](C)CC)=O)N[C@@H](CC%29=CN=CN%29)C(N[C@@H]([C@H](CC)C)C(N)=O)=O

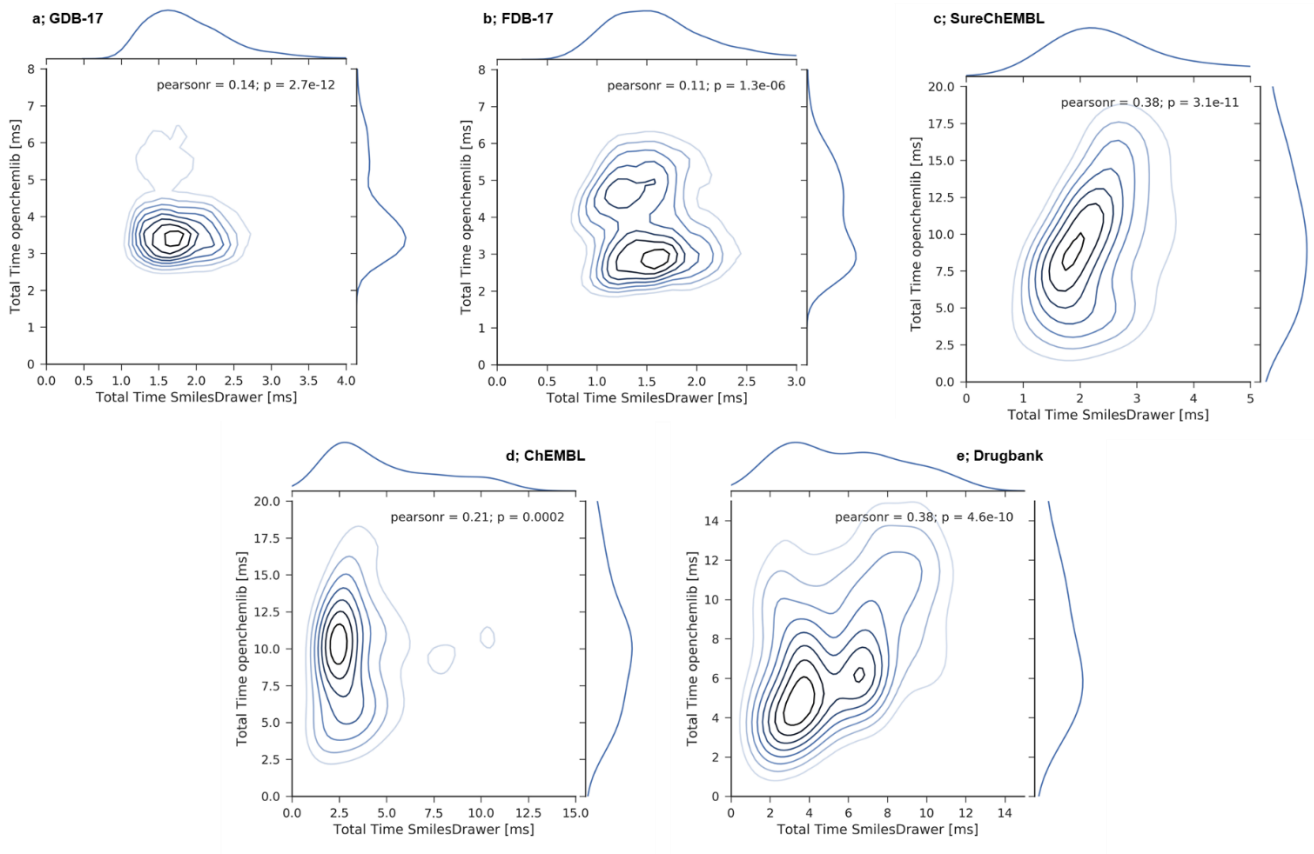


Fig. S2 Comparison to OpenChemLib-JS. Comparing the performance of SmilesDrawer to OpenChemLib-JS in regard to compounds containing bridged ring systems for the test sets (**a**) GDB-17, (**b**) FDB-17, (**c**) SureChEMBL, (**d**) ChEMBL and (**e**) Drugbank. The data shows that our implementation of a dynamic systems simulation based on the algorithm by Kamada and Kawai for the placement of atoms in bridged rings does not slow down drawing performance compared to OpenChemLib-JS. 456 (1.262%) and 509 (1.409%) compounds were removed from the SmilesDrawer and OpenChemLib-JS set respectively, as they interfered with the readability of these plots.