Supporting Information

Functional Insight into *Cordyceps militaris* Sugar Transporters by Structure Modeling, Network Analysis and Allosteric Regulation

Xin Liu,[a,b] Hanyang Zhang,[a] Ziyun Zhou,[a,c] Pranesha Prabhakaran, [d,e] Wanwipa Vongsangnak,[*e,f] Guang Hu,[*a,c] and Fei Xiao [*a]

a.Center for Systems Biology, Department of Bioinformatics, School of Biology and Basic Medical Sciences, Suzhou Medical College of Soochow University, Suzhou 215123, China.

b.Institute of Blood and Marrow Transplantation, Medical College of Soochow University, Jiangsu Institute of Hematology, The first Affiliated Hospital of Soochow University, Collaborative Innovation Center of Hematology, National Clinical Research Center for Hematologic Diseases, Soochow University, Suzhou 215123, China.

c.Jiangsu Province Engineering Research Center of Precision Diagnostics and Therapeutics Development, Suzhou, 215123, China

d.Interdisciplinary Graduate Program in Bioscience, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.

e.Omics Center for Agriculture, Bioresources, Food, and Health, Faculty of Science, Kasetsart University (OmiKU), Bangkok 10900, Thailand.

f.Department of Zoology, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.

*† *Corresponding authors.*

Email address: wanwipa.v@ku.ac.th (W.V.); huguang@suda.edu.cn (G.H.); xiaofei@suda.edu.cn (F.X.)

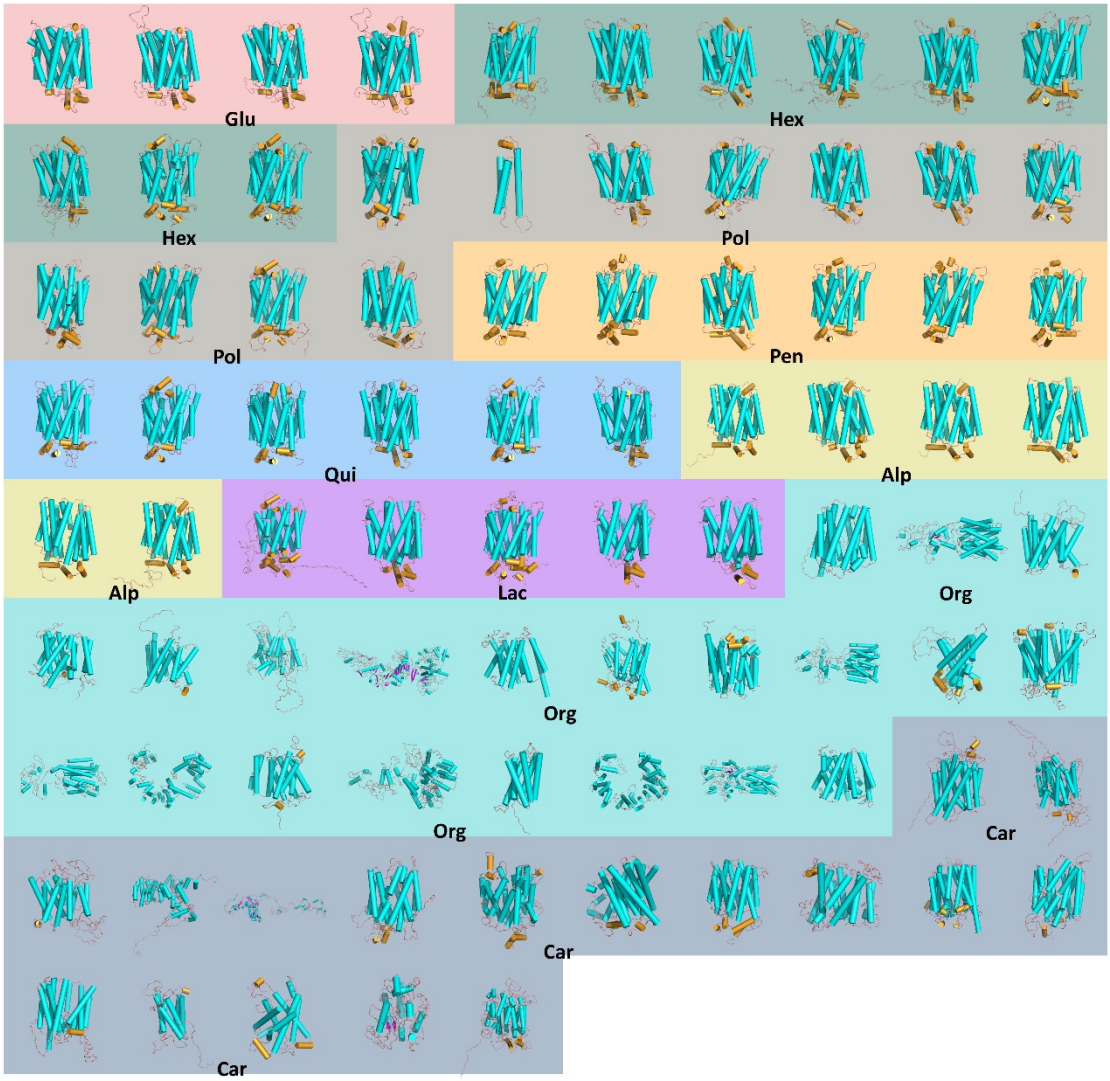

Fig. S1 Three-dimensional structural modeling results of 85 *C. militaris* sugar transport proteins. Protein structures are represented as colored cartoon, and each subfamily is marked with a different background color.

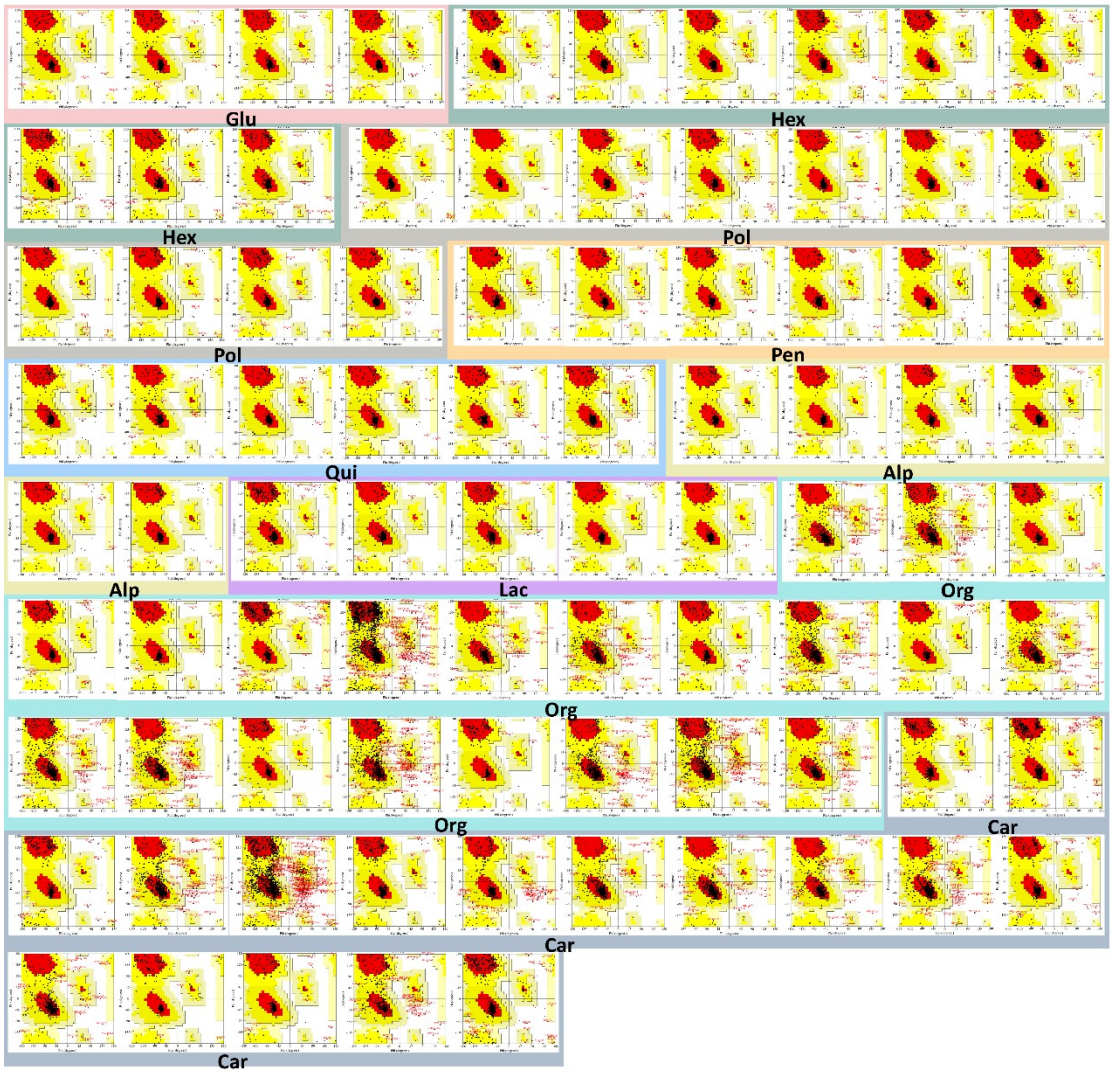

Fig. S2 Ramanchandran plots of evaluation results of all *C. militaris* sugar transport protein modeling results.

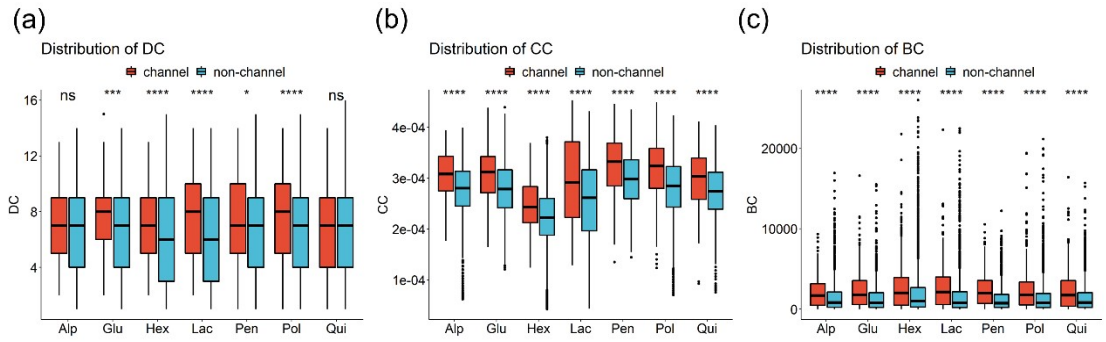

Fig. S3 Distribution of network parameters (a) DC, (b) CC, (c) BC in channel and non-channel residues in each subfamily.

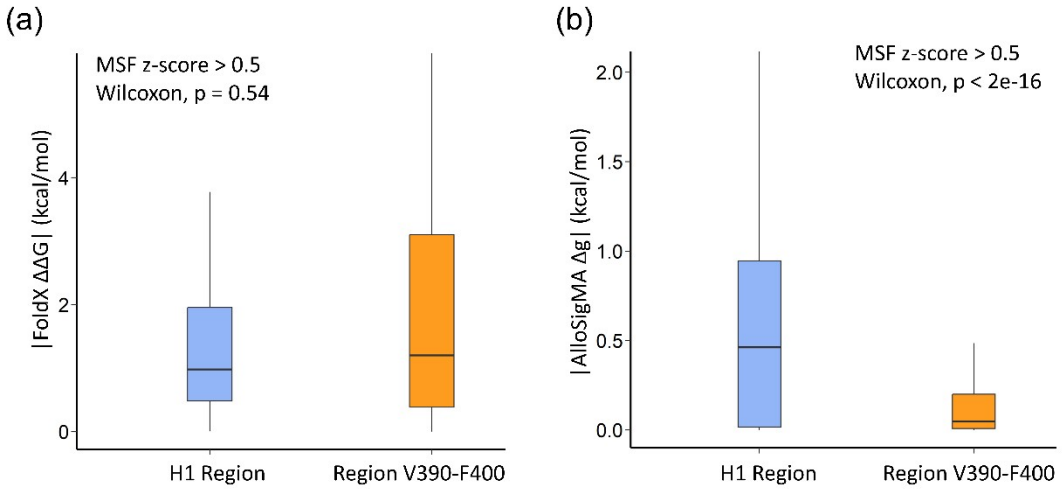

Fig. S4 The effect of saturation mutation on (a) the folding free energy ΔΔG for all residues with MSF z-score > 0.5 in the H1 region and V390-F400 region, as well as (b) the allosteric Δg, was compared by Wilcoxon signed ranking test.

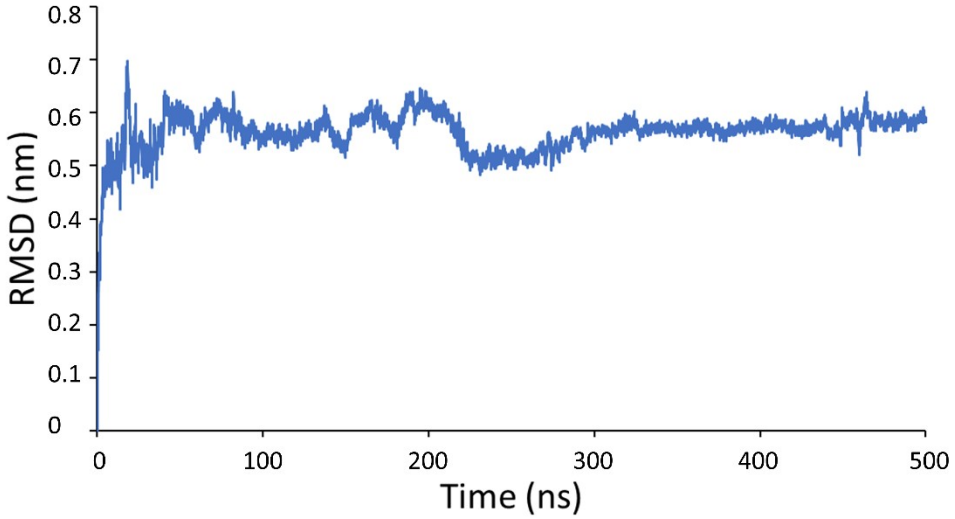

Fig. S5 RMSD of Cα atoms of Hex06741 during MD simulation.

Table S1. Subfamily classification information of *C. militaris* sugar transport proteins.

Subfamily	Abbreviation	Protein id	Protein abbreviation
alpha-glucoside	Alp	CCM_00728	Alp 00728
alpha-glucoside	Alp	CCM_02083	Alp 02083
alpha-glucoside	Alp	CCM_07255	Alp 07255
alpha-glucoside	Alp	CCM_07490	Alp 07490
alpha-glucoside	Alp	CCM_07588	Alp 07588
alpha-glucoside	Alp	CCM_08696	Alp 08696
carboxylate	Car	CCM_00851	Car 00851
carboxylate	Car	CCM_01120	Car 01120
carboxylate	Car	CCM_01719	Car 01719
carboxylate	Car	CCM_03592	Car 03592
carboxylate	Car	CCM_04459	Car 04459
carboxylate	Car	CCM_04581	Car 04581
carboxylate	Car	CCM_04643	Car 04643
carboxylate	Car	CCM_05263	Car 05263
carboxylate	Car	CCM_05651	Car 05651
carboxylate	Car	CCM_05803	Car 05803
carboxylate	Car	CCM_06052	Car 06052
carboxylate	Car	CCM_06456	Car 06456
carboxylate	Car	CCM_07432	Car 07432
carboxylate	Car	CCM_07990	Car 07990
carboxylate	Car	CCM_08224	Car 08224
carboxylate	Car	CCM_09352	Car 09352
carboxylate	Car	CCM_09518	Car 09518
glucose	Glu	CCM_00222	Glu 00222
glucose	Glu	CCM_02140	Glu 02140
glucose	Glu	CCM_03623	Glu 03623
glucose	Glu	CCM_04896	Glu 04896
other hexose sugars	Hex	CCM_01166	Hex 01166
other hexose sugars	Hex	CCM_02729	Hex 02729
other hexose sugars	Hex	CCM_03583	Hex 03583
other hexose sugars	Hex	CCM_04946	Hex 04946
other hexose sugars	Hex	CCM_05726	Hex 05726
other hexose sugars	Hex	CCM_06741	Hex 06741
other hexose sugars	Hex	CCM_07674	Hex 07674
other hexose sugars	Hex	CCM_08090	Hex 08090
other hexose sugars	Hex	CCM_09544	Hex 09544
lactose	Lac	CCM_03819	Lac 03819
lactose	Lac	CCM_05297	Lac 05297
lactose	Lac	CCM_07980	Lac 07980
lactose	Lac	CCM_08600	Lac 08600
lactose	Lac	CCM_09094	Lac 09094

organic anion	Org	CCM_00002	Org 00002
organic anion	Org	CCM_00013	Org 00013
organic anion	Org	CCM_00156	Org 00156
organic anion	Org	CCM_00266	Org 00266
organic anion	Org	CCM_00449	Org 00449
organic anion	Org	CCM_01038	Org 01038
organic anion	Org	CCM_02068	Org 02068
organic anion	Org	CCM_02414	Org 02414
organic anion	Org	CCM_03338	Org 03338
organic anion	Org	CCM_03480	Org 03480
organic anion	Org	CCM_04743	Org 04743
organic anion	Org	CCM_05810	Org 05810
organic anion	Org	CCM_05926	Org 05926
organic anion	Org	CCM_07068	Org 07068
organic anion	Org	CCM_07086	Org 07086
organic anion	Org	CCM_07408	Org 07408
organic anion	Org	CCM_07420	Org 07420
organic anion	Org	CCM_07625	Org 07625
organic anion	Org	CCM_07673	Org 07673
organic anion	Org	CCM_07725	Org 07725
organic anion	Org	CCM_09563	Org 09563
pentose	Pen	CCM_01139	Pen 01139
pentose	Pen	CCM_02098	Pen 02098
pentose	Pen	CCM_05984	Pen 05984
pentose	Pen	CCM_06358	Pen 06358
pentose	Pen	CCM_08290	Pen 08290
pentose	Pen	CCM_09215	Pen 09215
polyol	Pol	CCM_00501	Pol 00501
polyol	Pol	CCM_01080	Pol 01080
polyol	Pol	CCM_01532	Pol 01532
polyol	Pol	CCM_03115	Pol 03115
polyol	Pol	CCM_04797	Pol 4797
polyol	Pol	CCM_05298	Pol 05298
polyol	Pol	CCM_06039	Pol 06039
polyol	Pol	CCM_06864	Pol 06864
polyol	Pol	CCM_08999	Pol 08999
polyol	Pol	CCM_09026	Pol 09026
polyol	Pol	CCM_09671	Pol 09671
quinate	Qui	CCM_00186	Qui 00186
quinate	Qui	CCM_03527	Qui 03527
quinate	Qui	CCM_04942	Qui 04942
quinate	Qui	CCM_06002	Qui 06002
quinate	Qui	CCM_07760	Qui 07760

quinate	Qui	CCM_08598	Qui 08598

Table S2. Channel properties for Alp00728 transporter.

Snapshot	Channel	Throughput[1]	Cost[2]	Bottleneck radius[3]	Length[4]	Curvature[5]	Bottleneck Residue[6]
Alp00728	18	0.26	1.34	0.89	69.41	1.44	Y283;G282;
Alp00728	19	0.25	1.39	0.89	62.41	1.3	I281;T24;
Alp00728	27	0.17	1.78	0.89	76.65	1.32	I23;L59;W55

Note: [1]Probability that the pathway is used as a route for transporter of the substances using the formula. [2]The lower the cost value, the higher the order of the channels. In a biological sense, the energy spent by small molecules passing through a channel with a large bottleneck and a short path is lower. [3]Maximal probe size which can fit in the narrowest part of the channel. [4]Length of the channel from the starting point to the protein surface. [5] Shape of the channel as the ratio between the length of the channel and the shortest possible distance between the starting point and the channel ending point. [6]The narrowest part of the channel (bottleneck) including a list of surrounding residues and a static picture of the bottleneck with the channel visualized as spheres and surrounding residues as sticks.

Table S3. Basic information on protein transporter channels of different subfamilies.

Group	Mean of bottleneck radius (Å)	Standard deviation of bottleneck radius (Å)	Mean of channel length (Å)	Standard deviation of channel length (Å)
Alp	0.926	0.074	69.926	5.688
Glu	0.898	0.084	80.255	11.092
Hex	0.859	0.034	73.538	10.671
Lac	0.937	0.12	68.188	9.222
Pen	0.869	0.05	70.638	6.712
Pol	0.955	0.118	58.613	19.853
Qui	0.935	0.12	72.526	5.666

Table S4. Modeling templates for *C. militaris* sugar transporter proteins.

Subfamily	Sequence id	Modeller templates	I-TASSER templates
alpha-glucoside	CCM_02083	4YBQ_A	
alpha-glucoside	CCM_07490	4ZW9_A	
alpha-glucoside	CCM_00728	4LDS_A	
alpha-glucoside	CCM_07255	5EQG_A	
alpha-glucoside	CCM_07588	6N3I_A	
alpha-glucoside	CCM_08696	4ZW9_A	
carboxylate	CCM_06456	4J05_A	
carboxylate	CCM_04581	4J05_A	
carboxylate	CCM_01120	4J05_A	
carboxylate	CCM_03592		4m64A,5gm6G,1pv6A,6cc4A,6t1zA, 3ayfA,6g9x,6af0A,6w4s,6n1zA
carboxylate	CCM_05651		4j05A,4j05A,4j05A,6h7d,6h7d,4j05A, 5c65,4j05A,6h7dA,4j05A
carboxylate	CCM_09518	4LDS_A	
carboxylate	CCM_04643		5ifeC,4j05A,5gm6G,4j05A,4xriA, 6rw3A,6m2l,6m2l,4j05A,5c65
carboxylate	CCM_06052		4zowA,6d4hA,4zp0A,1qgrA,4zowA, 1pw4,4zow,6kki,4zowA,4zp0A
carboxylate	CCM_07432		1pw4A,3o7pA,3o7pA,6g9x,3o7p, 1pw4A,1pw4,3o7pA,1pw4A,3o7pA
carboxylate	CCM_05263		3m71A,3ayfA,3m71A,3m73A,6qq5A, 3m71,5z1wA,3m71S,5khnB,3m71A
carboxylate	CCM_08224	2LCK_A	
carboxylate	CCM_01719	6OH4_A	
carboxylate	CCM_05803		6i1rA,6i1rA,6oh2A,5y78,6ukj, 6i1rA,6ukj,6i1rA,6t15A,6i1rA
carboxylate	CCM_07990	5YS3_A	
carboxylate	CCM_00851	4J05_A	
carboxylate	CCM_04459		2qshA,2qshA,2qshA,2qsh,2qsh, 2qshA,2qsh,2qsfA,2qshA,2qsfA
carboxylate	CCM_09352		5tqqA,5m5iA,5y4gA,1gvhA, 3t38,5olnA,3t38A,4jo0A
glucose	CCM_04896	4GBY_A	
glucose	CCM_02140	4LDS_A	
glucose	CCM_03623	6H7D_A	
glucose	CCM_00222	6H7D_A	
other hexose sugars	CCM_06741	6H7D_A	
other hexose sugars	CCM_08090	6H7D_A	
other hexose sugars	CCM_03583	6H7D_A	
other hexose sugars	CCM_07674	6H7D_A	
other hexose sugars	CCM_04946	6H7D_A	

other hexose sugars	CCM_01166	6H7D_A	
other hexose sugars	CCM_05726	6H7D_A	
other hexose sugars	CCM_09544	6H7D_A	
other hexose sugars	CCM_02729	6H7D_A	
lactose	CCM_08600	4LDS_A	
lactose	CCM_05297	6H7D_A	
lactose	CCM_09094	6H7D_A	
lactose	CCM_03819	6H7D_A	
lactose	CCM_07980	6H7D_A	
organic anion	CCM_03480	7BP3_A	
organic anion	CCM_00002		3o7pA,3o7pA,3o7pA,6kki,3o7p, 3o7pA,6kki,3o7pA,3o7pA,3o7pA,
organic anion	CCM_00013		6d4hA,3aqpA,6d4hA,3jac,3jac, 5voxB,5tj5,3rkoC,3ayfA,2gb5A
organic anion	CCM_04743		6d4hA,3aqpA,6d4hA,3jac,3jac, 6d4hA,5tj5,1kplA,2vc9A,5n9yA
organic anion	CCM_05926		4m64A,6t1zA,6t1zA,6g9x,4m64, 4m64A,4m64,1pw4A,4m64A,1b3uA
organic anion	CCM_07673		6bugC,1qgrA,6bugA,1qgrA,6bugC, 5oc9A,6bugC,3jd8A,6bug,4fgvA
organic anion	CCM_01038	5Y78_A	
organic anion	CCM_00266	5Y78_A	
organic anion	CCM_09563		5y78A,5y78A,5y78A,5y78,5y78, 5y78A,5y78,5y78A,6yleC,5y78A
organic anion	CCM_00156	5OGE_A	
organic anion	CCM_00449	5OGE_A	
organic anion	CCM_07068		5y78A,6d4hA,5y78A,5voxB,5y78A, 6d4hA,5y78,5y78,5y78A,5y78
organic anion	CCM_02414		5y78A,5y78A,5y78A,5y78,6ukj, 5y78A,6ukj,5y78A,6cptA,5y78A
organic anion	CCM_07625	5ZUG_A	
organic anion	CCM_07408	5Y78_A	
organic anion	CCM_07420		5y78A,4cr2Z,5y78A,6d4hA, 5y78,5yfpC,5y78,2qizA
organic anion	CCM_07725		6d4hA,5y78A,5y78A,5y78,5y78, 5y81B,5y78,5y78A,1u6gC,5y78A
organic anion	CCM_05810	4C9G_A	
organic anion	CCM_02068		6ar6A,6c9dA,5jcsS,6c9dA,5z33, 5jcsS,4tmb,5byzA,6emkA,2y94A
organic anion	CCM_03338		3wdoA,6e8jA,6e8jA,6g9x,6e8j, 6e8jA,6e8j,3wdoA,1qgrA,3wdoA
organic anion	CCM_07086		4zowA,4c0oA,3wdoA,4xriA,6t1zA, 5m87A,6exs,1qgrA,1pw4,3wdo

pentose	CCM_05984	4LDS_A	
pentose	CCM_08290	6H7D_A	
pentose	CCM_02098	4LDS_A	
pentose	CCM_01139	4LDS_A	
pentose	CCM_06358	6H7D_A	
pentose	CCM_09215	6H7D_A	
polyol	CCM_03115	6H7D_A	
polyol	CCM_09026	6H7D_A	
polyol	CCM_01080	6H7D_A	
polyol	CCM_01532	4LDS_A	
polyol	CCM_05298	4LDS_A	
polyol	CCM_00501	4LDS_A	
polyol	CCM_04797	4LDS_A	
polyol	CCM_06864	4LDS_A	
polyol	CCM_08999	4LDS_A	
polyol	CCM_06039	4PYP_A	
polyol	CCM_09671	5C65_A	
quinate	CCM_00186	6H7D_A	
quinate	CCM_03527	6H7D_A	
quinate	CCM_04942	4GBY_A	
quinate	CCM_06002	6H7D_A	
quinate	CCM_07760	6H7D_A	
quinate	CCM_08598	4QIQ_A	