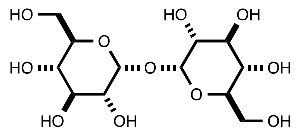
Supplementary Information

Discovery of a Kojibiose Phosphorylase in *Escherichia coli* K-12

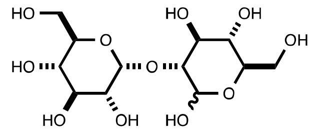
Keya Mukherjee[Φ], Tamari Narindoshvili[Ψ], and Frank M. Raushel[Φ,Ψ,*]

[Φ]Department of Biochemistry & Biophysics, Texas A&M University, College Station, Texas, 77844, USA.

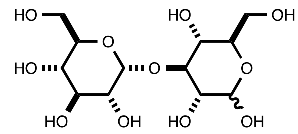
[Ψ]Department of Chemistry, Texas A&M University, College Station, Texas 77842, USA



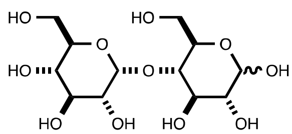
trehalose
α,α-(1,1)-D-glucose-D-glucose



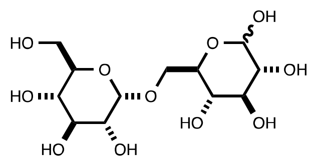
kojibiose
α-(1,2)-D-glucose-D-glucose



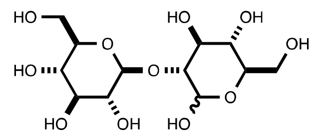
nigerose
α-(1,3)-D-glucose-D-glucose



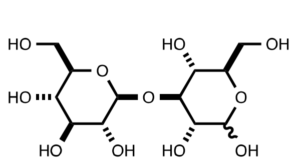
maltose
α-(1,4)-D-glucose-D-glucose



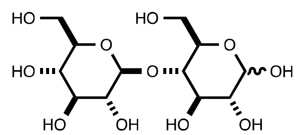
isomaltose
α-(1,6)-D-glucose-D-glucose



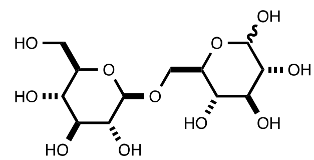
sophorose
β-(1,2)-D-glucose-D-glucose



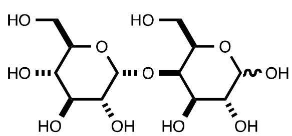
laminaribiose
β-(1,3)-D-glucose-D-glucose



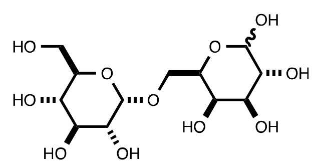
cellobiose
β-(1,4)-D-glucose-D-glucose



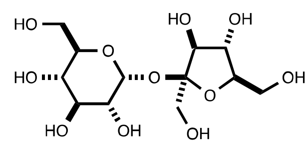
gentiobiose
β-(1,6)-D-glucose-D-glucose



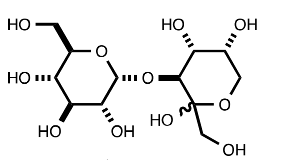
lactose
β-(1,4)-D-glucose-D-galactose



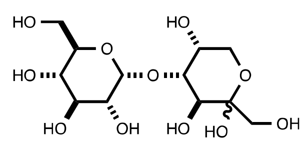
melibiose
α-(1,6)-D-glucose-D-galactose



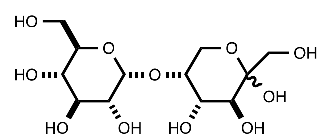
sucrose
α,β-(1,2)-D-glucose-D-fructose



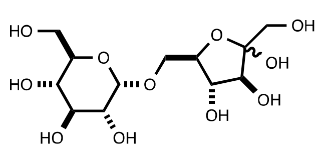
turanose
α-(1,3)-D-glucose-D-fructose



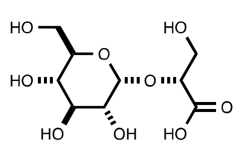
maltulose
α-(1,4)-D-glucose-D-fructose



leucrose
α-(1,5)-D-glucose-D-fructose

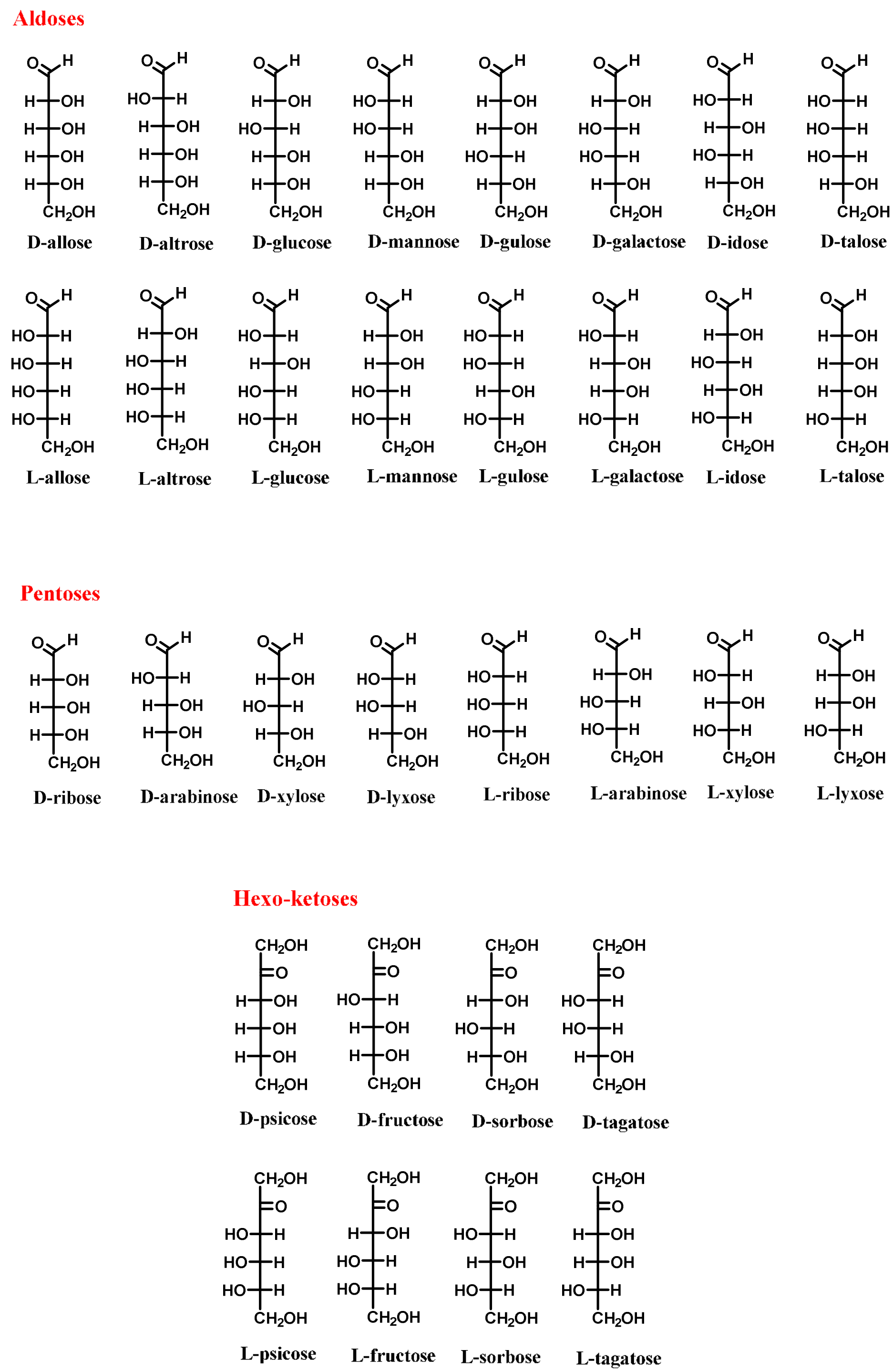


palatinose
α-(1,6)-D-glucose-D-fructose

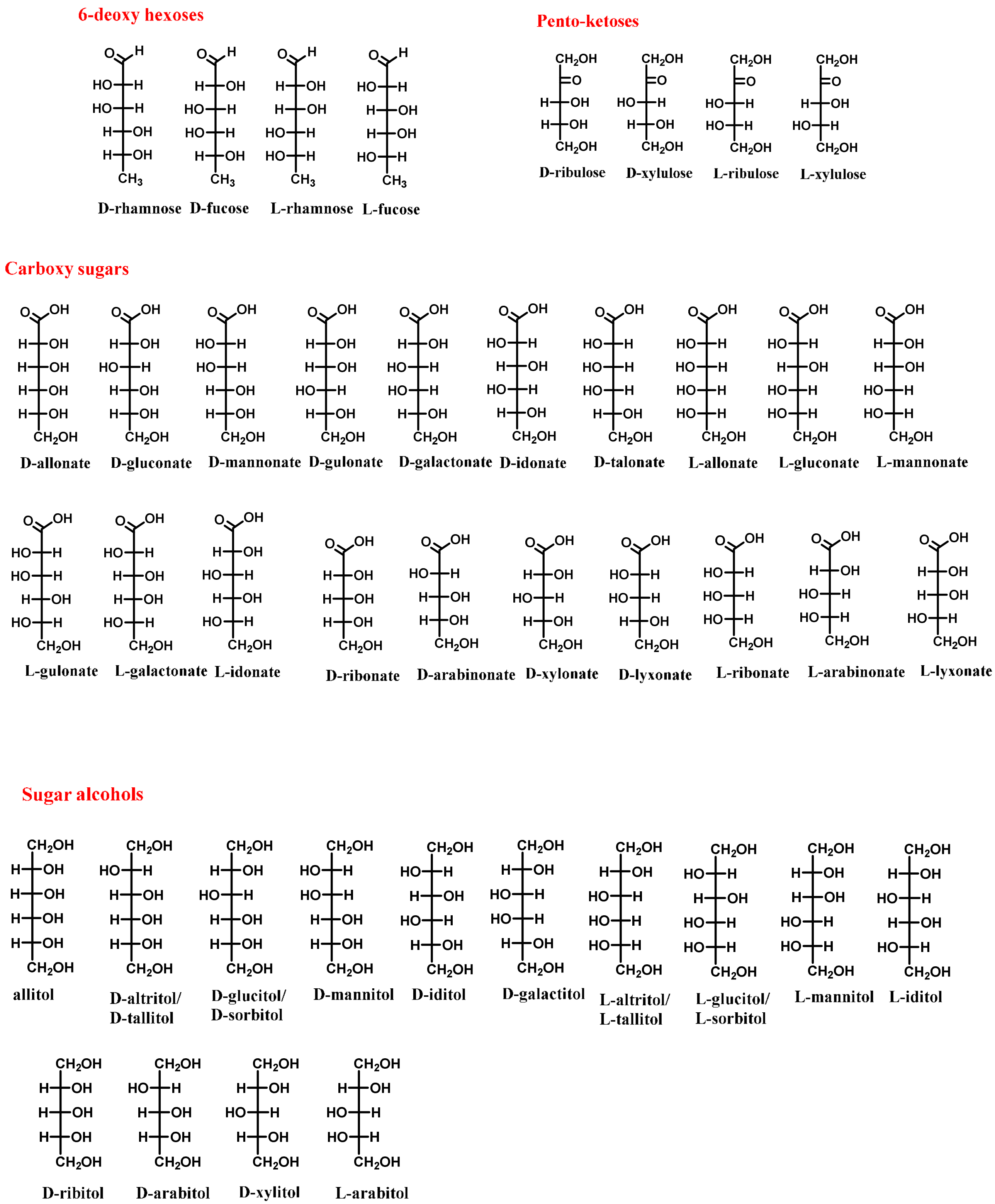


α-(1,2)-D-glucose-D-glycerate

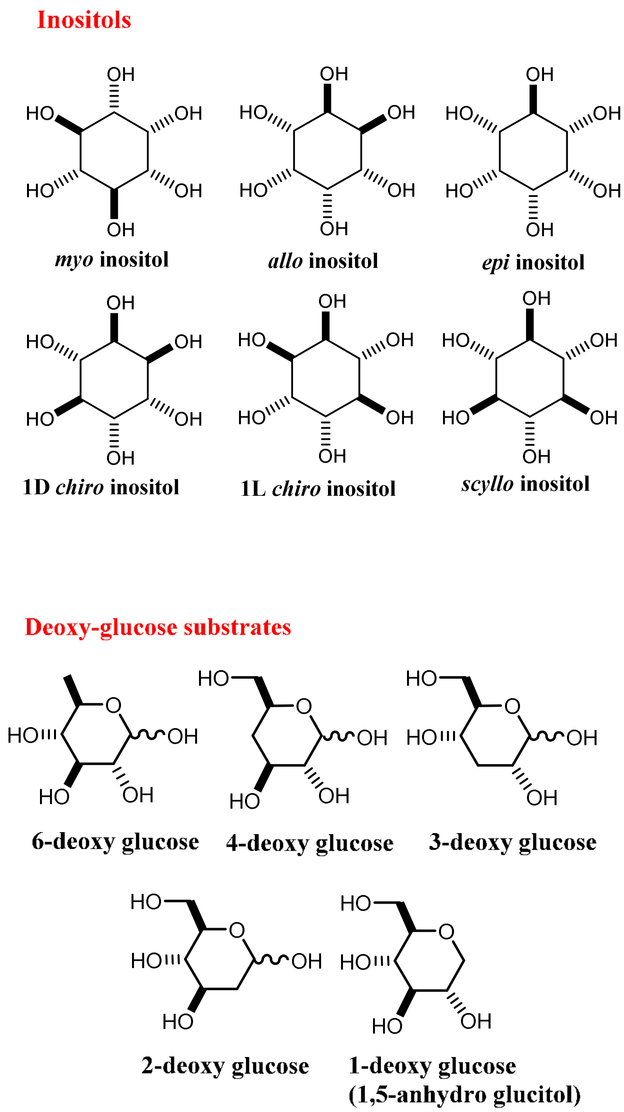
Scheme S1: Structures of disaccharides tested as substrates for YcjM and YcjT.



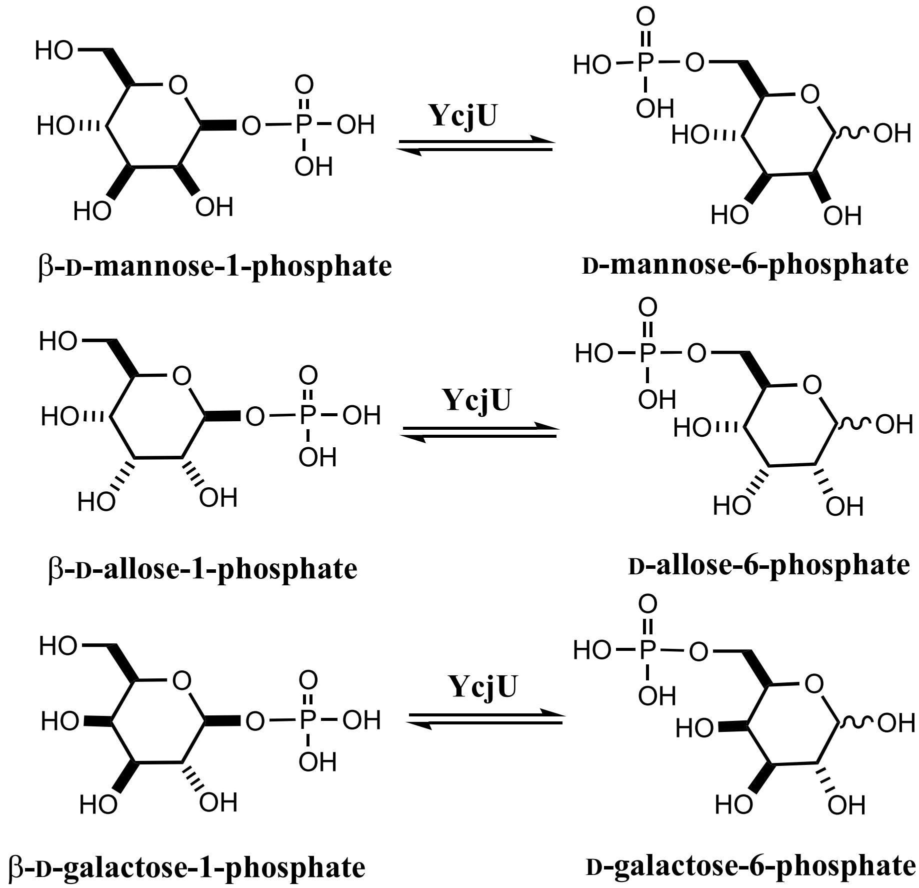
Scheme S2: Structure of aldohexoses, aldopentoses, and ketohexoses tested as substrates for YcjM and YcjT.



Scheme S3: Structures of 6-deoxyaldoses, ketopentoses, sugar carboxylates and sugar alcohols tested as substrates for YcjM and YcjT.



Scheme S4: Structures of deoxy-D-glucose variants and inositols tested as substrates for YcjM and YcjT.



Scheme S5: Reactions catalyzed by YcjU.

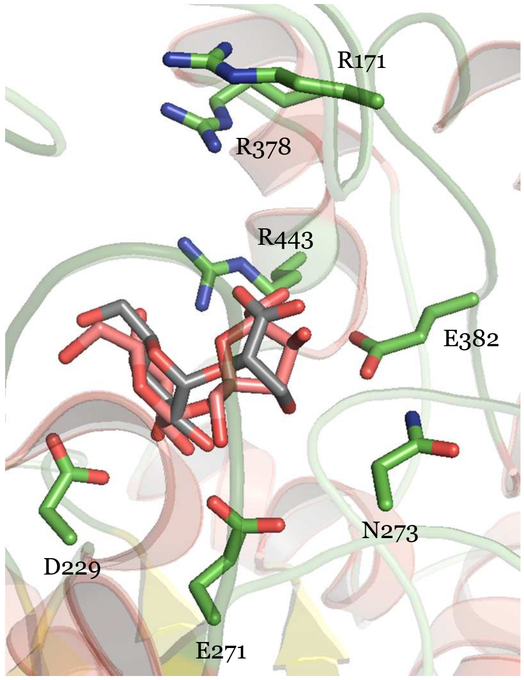


Figure S1: Homology model for active site of YcjM built using Phyre2. The α-(1,2)-D-glucose-D-glycerate was docked in the active site using AutoDock Vina. The YcjM model is superimposed with sucrose phosphorylase structure from *B. adolescentis* (PDB id: 2GDU). The bound sucrose is shown in red. In the YcjM model, α-(1,2)-D-glucose-D-glycerate is shown in grey and the active site residues of YcjM are shown in green.

```
YcjT    1 -------------MTRPVTLSEPHFSQHTLNKYASLMAQGNGYLGLRASHEEDYTRQTRGMYLAGLYHRAGKGEIN  63
PhKP    1 ------------------MRFQFGFSKEDEQVLGTILTLGNGQLGVRGEFEL--ERSPYGTIVSGVYDYT-PYFYR  55
CsKP    1 ----------MKLSEREWLIEQDKL--EASGKFETCFALTNGYIGIRGINEEVFCEETPGTYIAGVFDKS-TAQVT  63
TbKP    1 MVKHMFLEDVNNLISDDKWLIFQNEYNTEVNPRYETLFTLTNGYMGVRGTFEEGSEGERSGNFIAGIFDKS-DAQVR  76

YcjT   64 ELVNLPDVVGMEIAINGEVFSLSH---EAWQRELDFASGELRRNVVWRTSNGSGYTIASRRFVSADQLPLIALEITI 137
PhKP   56 ELVNGPRTIGMIIIIDGELINPSSQKVKEFQRELDIEKGLLRTHLEIETKNGNKILYKSTRIVHMKRKNLILLDFEL 132
CsKP   64 ELVNLPNPIGLRIYINREFLNPLKCEILEFKRVLDLKQGILYRKLRLKDVKGRITTIEGFRFVSMNNKNLIVQKYDV 140
TbKP   77 EIVNAQNWLRIKLYVEGEELSLDKCQLIEFKRILDMKKGILFRSMLIKDSKDRITRIEGYRFISRSDLHRSAIKLFV 153

YcjT  138 TPLDADASVLISTGIDATQTNHGRQH------LDETQVRVFGQHLMQGSYTTQDGRSDVAISCCCKVSGD-----VQ 203
PhKP  133 KASK-GGIAVVVNPIEFNTANPGFIDEIMIKHYRVDSIKET-EEGVYARVKTLDNKYTLEIASSLVPSEYTSR---- 203
CsKP  141 VCENYSAVLNVESFIDATTVNSKDVPNDRVKHYEIDKKKDF-ADGIYLGITTKDKKYKVGIASSTKVLLN-----NQ 211
TbKP  154 TPVNYSGVVGIESIIDGTVLNSADSPKHRVKHLKVADNSSLNKSGVYLETATIDDDIRIATGSAVRLYHYEDKEKNN 230

YcjT  204 QCYTAKERRLL-----QHTSAQLHAGETMTLQKLVWIDWRDDRQAALDEWGSASLRQLEMCAQQSYDQLLAASTENW 275
PhKP  204 -STFR-TDNEIG----EIYIVKLKPGKTYKFTKYVTV------------SKGAALEELKDVKRLGFEKLYEEHINSW 262
CsKP  212 RCYFNRFTKDLGYIITENFEVEAKQGERYEIEKLTVLVS--SREKNVGDVFETCTNKLKEFETKSAEKLLFEHIEEY 286
TbKP  231 IAKFKRF-LPLGEMSIEYFEFDGTENKTVVIDKFIITYT--SRDVKKGLLKSTVEKELFAFAGEGIDKELQRHIEVY 304

YcjT  276 RQWWQKRRITVNGGEAHDQQALDYALYHLRIMTPAHDERSSIAAKGLTGEGYKGHVFWDTEVFLLPFHLFSDPTVAR 352
PhKP  263 KRIWEKVKVEIEGDKDL-ENALNFNIFHLIQSLPPTD-KVSLPARGIHGFGYRGHIFWDTEIYALPFFIFTMPKEAR 337
CsKP  287 KRLWDVANIDIVGDEVA-NKSVKFNIFHLISMANPEDEHVSLGAKGLHGEGYKGHVFWDTEIFMLPFYIYTNPAAAK 362
TbKP  305 EELWSVADINIEGDEEA-DKALRFNIFHLMSSVNENDPMVSIAAKALHGEGYKGHVFWDTEIFMLPFFIYVHPKAAK 380

YcjT  353 SLLRYRWHNLPGAQEKARRNGWQGALFPWESARSGEEETPEFAAINIRTGLRQKVASAQAEHHLVADIAWAVIQYWQ 429
PhKP  338 RLLLYRCNNLDAAKENAKMNGYQGVQFPWESADDGREATPSEIPLDMLGRKIVRIYTGEEEHHITADIAYIVDFYYQ 414
CsKP  363 AMLMYRYNLLDAARENARKNGYKGAQFPWESADTGEEETPK-WGYDYLGN-PVRIWTGDIEYHISADIAYAVMNYVR 437
TbKP  381 TLLMYRYNMLDAARKNAALNGYKGAQYPWESADTGEEETPK-WGFDYMGN-PVRIWTGDLEHHITADIAFAVWEYFR 455

YcjT  430 TTGDESFIAHEGMALLLETAKFWISRAVR--VNDRLEIHDVIGPDEYTEHVNNNAYTSYMARYNVQQALNIARQFGC 504
PhKP  415 VSGDLEFMNRCGLEIIFETARFWASRVEFEEGK-GYVIKKVIGPDEYHEHVNNNFFTNLMAKHNLELAIRYFRE--S 488
CsKP  438 ATDDIDFLLNYGSEIIIETARFWASICKYNKEKGRYEINDVIGPDEFHEHCNNNAYTNYLAKWNLLKASELCNLLLE 514
TbKP  456 ATEDIEFMLNYGAEVIFETARFWVSRCEYVKELDRYEINNVIGPDEFHEHVDNNAYTDYLAKWNIKKGLELINMLKE 532

YcjT  505 SDDAFI--------------HRAEMFLKELWMPEIQPDGVLPQDDSFMAKPAINLAKYKAAAGKQTI-LLDYSRAEV 566
PhKP  489 KNREPWKKIVEKLNIREEEVEKWEEIAKNMYIPRKID-GVFEEFDGYFELMDFEVDPFNIG--EKTL-PEEI-RNNI 560
CsKP  515 KYPKYFEKLSKKINLSDEEPFVWQEIASKIYIPYHPDKKLIEQFEGYFNLKDFVIKEYDQN--NMPVWPEGVELDKL 589
TbKP  533 KYPEHYHAISNKKCLTNEEMEKWKEVEEKIYIPYDKDKKLIEQFEGYFDKKDYVIDKFDEN--NMPIWPEGVDITKL 607

YcjT  567 NEMQILKQADVVMLNYMLPEQFSAASCLANLQFYEPRTIHDSSLSKAIHGIVAARCGLLTQSYQFWREGTEIDLGAD 643
PhKP  561 GKTKLVKQADVIMAQYLLKDYFSPEEIKSNFNYYIRRTTHASSLSMPPYAIIATWIGEVKIAYEYFKRCANIDLKNV 637
CsKP  590 NNYQLIKQADVVMLLYLLGEEFDDQTKKINYDYYEKRTMHKSSLSPSIYALMGVRVGETNRAYINFMRTALTDLEDN 666
TbKP  608 GDTQLIKQADVVMLMLLLGEEFDEETKRINYEYYEKRTMHKSSLGPSMYAIMGLKVGDHKNAYQSFMRSANVDLVDN 684

YcjT  644 PHSCDDGIHAAATGAIWLGAIQGFAGVSVRD-GELHLNPALPEQWQQLSFPLFWQGCELQVTLDAQRIAIRT---SA 716
PhKP  638 YGNTAEGFHLATAGGTWQVLVRGFCGLNVKG-NKIELNPNLPEKWKYVKFRIFFKGSWIEFKISRKKVRARMLEGSR 713
CsKP  667 QGNTHLGIHAASLGGTWQALVFGFGGISIEKDDVLSVNPWLPEKWESLKFSIWWKGNLLDFKITKDNVEVKKRVEKG 743
TbKP  685 QGNTKEGLHAASAGGTWQVVVFGFGGMEIDKEGALNINSWLPEKWDKLSYKVFWKGNLIEVIVTKQEVTVKKLKGKG 761

YcjT  717 PVSLRLNGQLITVAEESVF-CLGDFILPFNGTATKHQEDE                                      755
PhKP  714 KVKISSFGKEVDLYPGKEVVIVAN---------------                                       737
CsKP  744 NVKLKIKGQEAII--------------------------                                       756
TbKP  762 NIKVKVKGKELTIE-------------------------                                       775
```

Figure S2: Multiple sequence alignment of YcjT, and kojibiose phosphorylases from *Pyrococcus horikoshii* (PsKP), *Caldicellulosiruptor saccharolyticus* (CsKP), and *Thermoanaerobacter brockii* (TbKP) created using Clustal omega. The residues interacting with kojibiose, as observed in the CsKP structure (2IWW), are conserved in YcjT. These residues are indicated in pink. YcjT has Ser409 in place of a conserved threonine residues as indicated in blue. The catalytic glutamate residue is shown in green.