

>Cellulose Synthase | Uniprot: Q3J125 | Rhodobacter sphaeroides  
MTVRAKARSPLRVVPLLFLWVALLVFPGLLAAAPVAPSAQGLIALSAVVLVALLKPFA  
DKMVPFRLLL SAASMLVMRYWFWRLFETLPPPALDASFLFALLLFAVETFSISIFFLNGF  
LSADPTDRPFPRPLQPEELPTVDILVPSYNEPADMLSVTLAAAKNMIYPARLRVTVLCCDD  
GGTDQRCMSDPPELAQKAQERRRELQQLCRELGVVYSTRERNEHAKAGNMSAALERLKGE  
LVVVFDADHVPSRDFLARTVGYFVEDPDLFLVQTPHFFINPDPPIQRNLALGDRCPPENEM  
FYGKIHRGLDRWGGAFFCGSAAVLRRRALDEAGGFAGETITEDAETALEIHSRGWKSLEYI  
DRAMIAQLQPETFASFIQQRGRWATGMMQMLLLKNPLFRRGLGIAQRLCYLNSMSFWFFP  
LVRMMFLVAPLIYLLFFGIEIFVATFEEVLAYMPGYLAVSFLVQNALFARQRWPLVSEVYE  
VAQAPYLARAIVTLLRPRSARFAVTAKDETLSENYISPIYRPLLFTFLLCLSGVLATLV  
RWVAFPGDRSVLLVVGWAVLNVLLVGFALRAVAEKQRRRAAPRVQMEVPAEAQIPAFGN  
RSLTATVLDASTSGVRLVRLPGVGDHPALEAGGLIQFQPKFPDAPQLERMVRGRIRSA  
RREGGTVMVGVI FEAGQPIAVRET VAYLIFGESAHWRMREATMRPIGLLHG MARILWMA  
AASLPKTARDFMDEPARRRRRHEEPKEKQAHLLAFGTFDFSTEPDWAGELLDPTAQVSARP  
NTVAWGSN

>Curdlan synthase | uniprot: Q9X2V0 | Agrobacterium sp. ATCC 31749  
MYFSAEGDVQSVLYVNLTIAGAILFALLADPRKMVDRLAFSIIIMLLSLGVYIVWRATDT  
LPPPELSLETLCWYTYFTFELISVLYAMGSILILLRRTDWSAVADQGEAYLAGNPHAPLV  
DVFICTYNEPLNVLEKSIIAAQAMDYPRLRVFCDDTRRGEVPTYCEAAGVNYVTRPDNK  
HAKAGNLNALLHTNALEEVSDFIMVLDADFAFQANFLRRVTGLFSDPKVAVVQTPQFYF  
NSDPIQHNLGIDKSFVDDQRFVDFDDQPAKDAVGCAFRVGT SFVRRRAAVNGIGGFPTDA  
LTEDMLLTYRLMERYVTRWLNEKWSVGLSAEGVPEYITQTRWCLGTIQIGLLRTGPLW  
RGNFTLTQRLHYLHGLFCWLSKPLILCLLLAPSIYWLTVGSALQVDELMMFKLGLSSLAL  
FWTYSTWISGKRTLPLFTEVTHALTAVPITITL FQAIRKPFGRPFKVTEKGGDRSQVRVH  
LPTGIFFAFVTLSSAVSIVLAVYGLDAPSELSSRDCLNIWSAVAMVIAFTSFICCIELP  
RFGKEEMIGVDFRQRLRSASSTRPVRTGLSTENITLAAVPSSSDVKDVFVPEAGWMRIS  
PAEHAQNSGKFDIHPSEQRRSILRLLFRKAPENVAEQDLMKSMRILLARAFG

>Mixed-linkage glucan synthase | Uniprot: Q92WG2 | Sinorhizobium meliloti  
MVQSLLALAPTLIVIAFFLLGPFNWSRHSWSRAVTCAFVAAIALRYMLWRFTETVLPYP  
NDGPNFYWVWFVIVEFLAFSEVVLFLILMSRYVDRSAEGNRLERQFFERDERELPTVDV  
FIPTYNEPLDVLERTIVGALALDHPKDKLNVYVLDGRRDWLRTFCEGRGAIHVTRSDNA  
HAKAGNMNGLRVSSGDFIAVFDADFV PYSFLRRTL PFFMDDTIGIVQTPQHFFNVDPI  
QSNL GLENLWPDEQRLFFDEIAPSRDGDVDFCCGSCSIARRKAVDAIGGFPTESITEDL  
LTTLSMLNRGFKTRYLNERLSMGLAAENLTGYFVQRERWCQGGIQTLYLHNGPLRGPGLS  
LFQRMFLPMSWL VQYLVR FIVLLIPIVYLLWFGALPLYFTDVADYVSNQVPLLAAYFLM  
FWLTPTRYLPLVSTAVGTFFSTFRMLPTVLSLVRPFGKPFKVTPKGSSNEANVFDAYTFT  
WIAGFIVVTALGLLINIVPETARVEGSFSAIAALWSGINIVVLLIIASLICFEKPRRLLQA  
FKLDEPVDV DVGVEGR LVSLSDKAVVAVSTETRFKSTKVRLNIEGFAPLEADLKQVTQRR  
GDITRTGDKQRYYLHLHYDLRGVERDKMIKLYTGRYSRDVPDIDKIAVSVNLLLRAFG

**S1 Fig. Bacterial GT2 query sequences.** Full-length sequences for cellulose synthase (RsBcsA), curdlan synthase (AtumCrdS) and mixed-linkage glucan synthase (SmBgsA) used as inputs into EVFold for retrieval of homologous sequences.