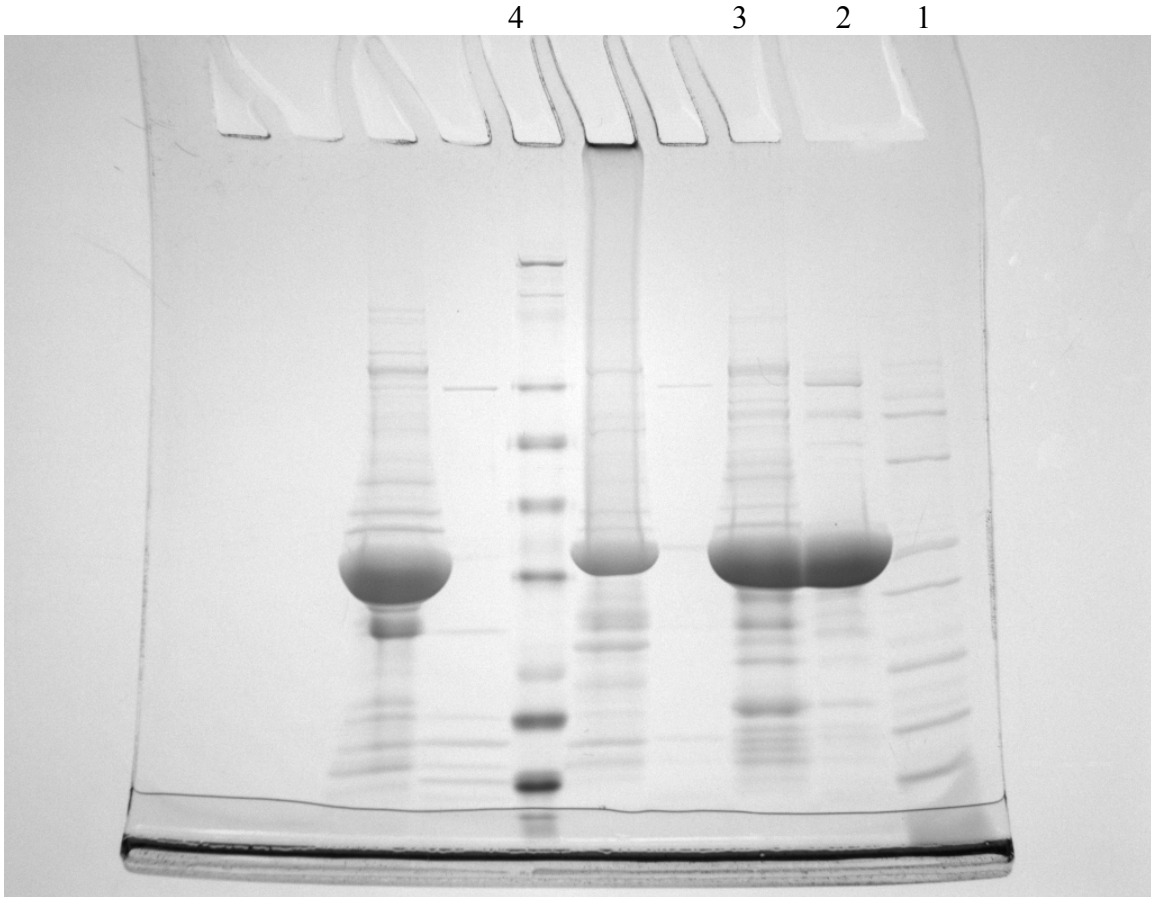


**Supplemental Figure S1. SDS-PAGE analysis of sugar kinases from *T. maritima*.**

**A. TM0067 / KdgK**

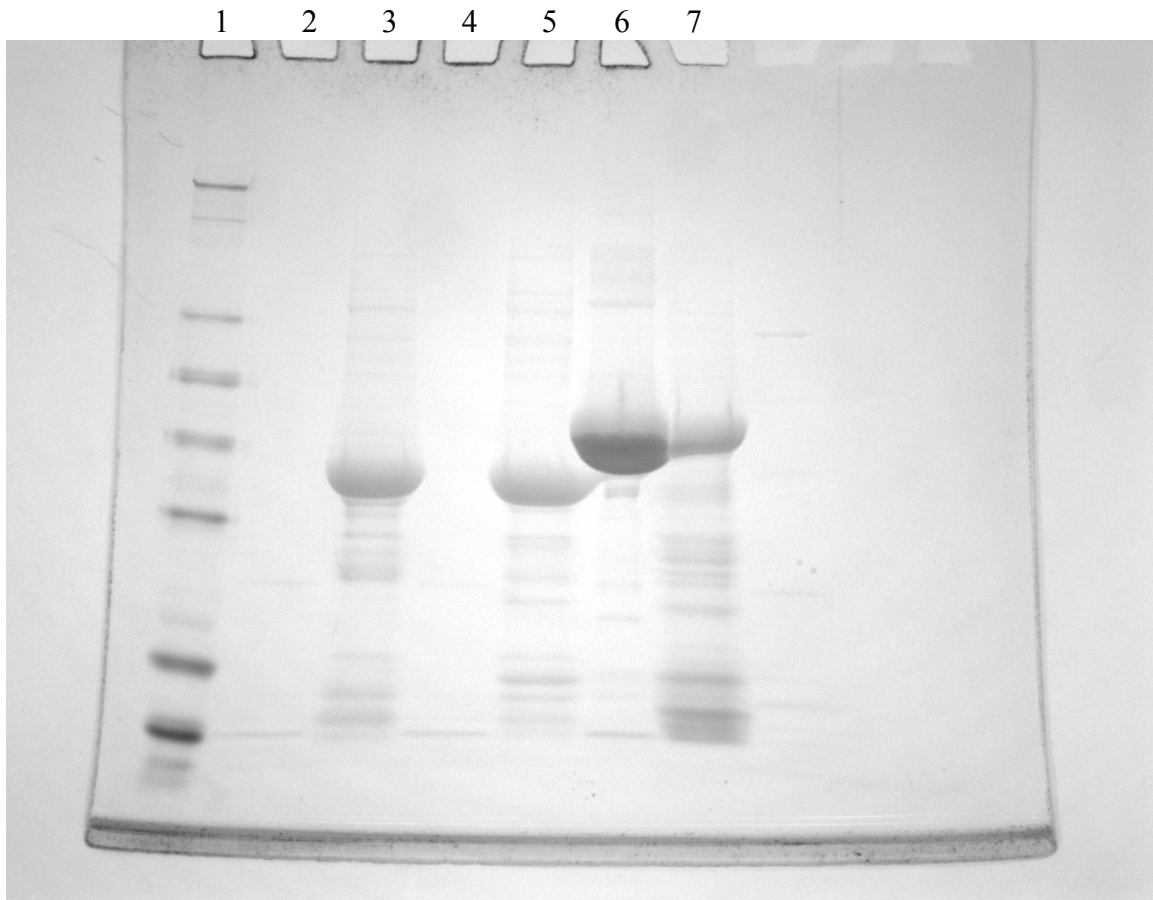


1&4: protein marker

2: TM0067 soluble part (10ul)

3: TM0067 insoluble part (10ul)

**B. TM0952 / Dr1K**

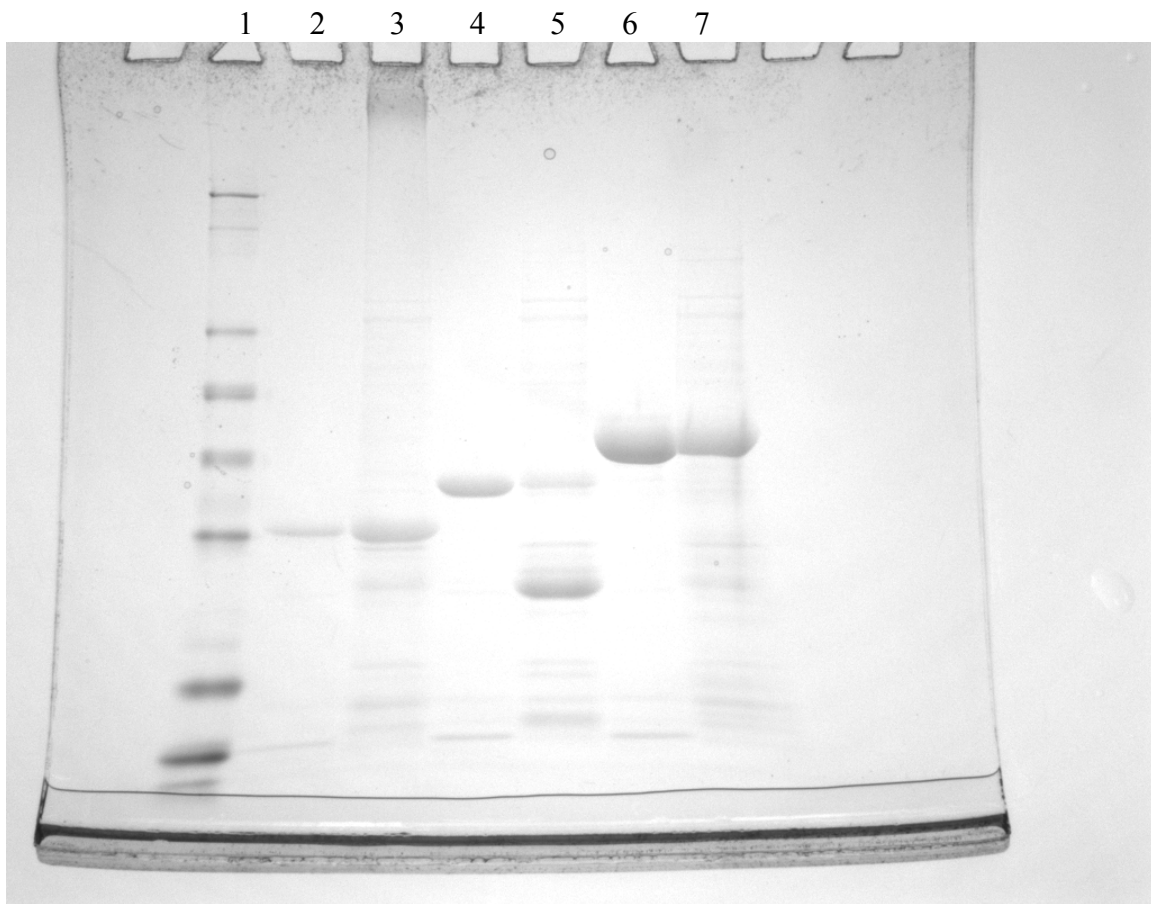


1: protein marker

6: TM0952 soluble part (10ul)

7: TM0952 insoluble part (10ul)

**C. TM0067 / KdgK, TM1190 / GalK, and TM1430 / GlpK**



1: protein marker

2: TM0960 soluble part (10ul)

3: TM 0960 insoluble part (10ul)

4: TM1190 soluble part (10ul)

5: TM1190 insoluble part (10ul)

6: TM1430 soluble part (10ul)

7: TM1430 insoluble part (10ul)

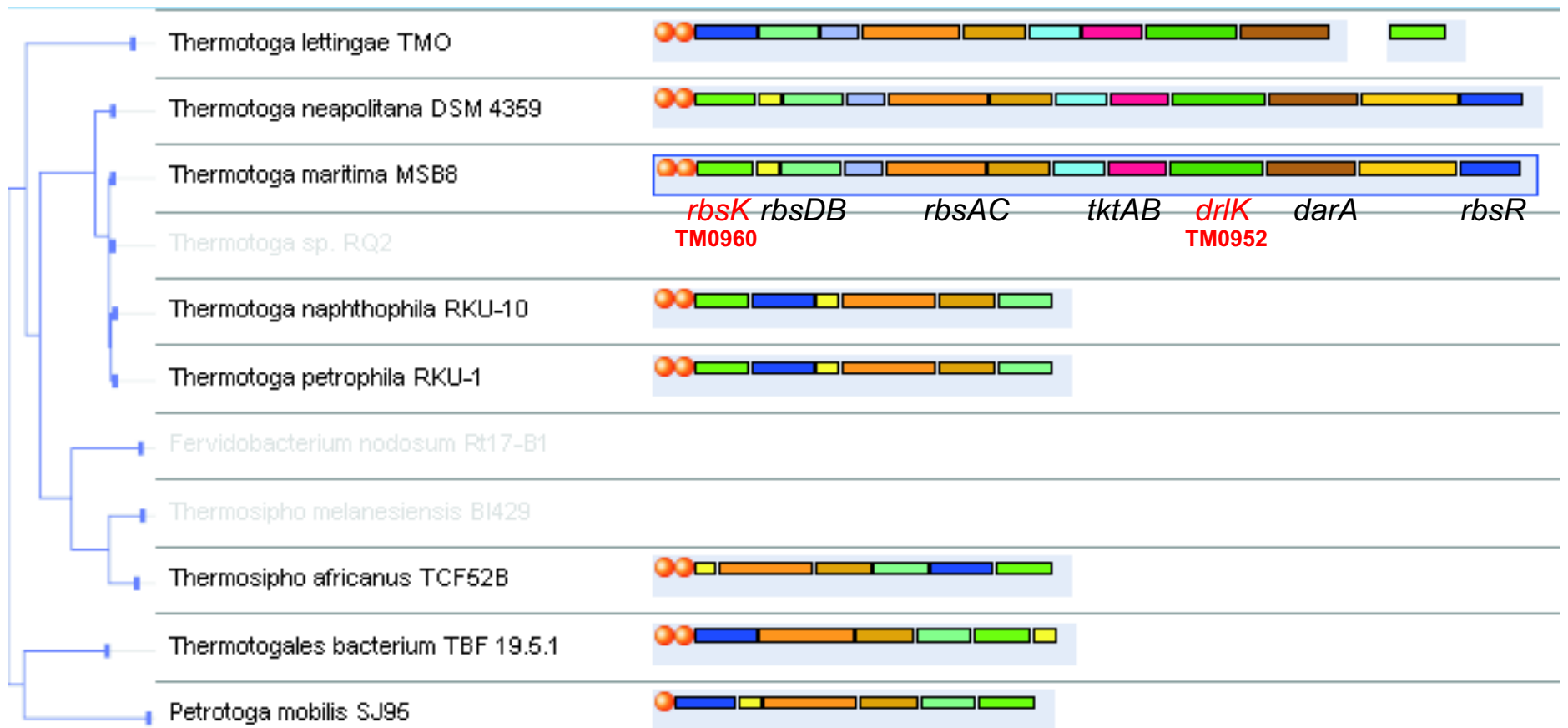
```

      R L Y S K W K E A V S R S L G W E K R Stop
TM1430  AGGCTGTATTCAAAAATTGAAAAGAGGCAGTGTCTGAGATCCCTCGGATGGGAAAAACGATTGA
TRQ2_1322 AGGCTGTATTCAAAAATGGAAAAGAGGCAGTGTCTGAGATCCCTTGGATGGGAAAAACGATTGA
Tpet_1364 AGGCTGTATTCAAAAATGGAAAAGAGGCAGTGTCTGAGATCCCTTGGATGGGAAAAACAGTGA
CTN_1064  AGGTTGTATTCAAAAATGGAAAAGAGGCTGTGTCTGAGGTCTCTCGGATGGGAAAAGCAGTGA
*****

```

**Supplemental Figure S2. Multiple alignment of 3'-coding region of TM1430 and its orthologs in *Thermotoga* spp.**

The respective amino acids of TM1430 protein are shown on the top. The tryptophan-coding codon which has a nonsense mutation in the *T. maritima* genome is highlighted in gray. The stop codons are in bold font. Abbreviations for the Thermotogales genomes are given in Table 2.



DNA binding motif of RbsR regulator:



Supplemental Figure S3. Example of comparative genomics reconstruction of RbsR regulon in the Thermotogales.

**Supplemental Table S1. Primers and restriction sites used for cloning of *T. maritima* genes.**

Gene	Plasmid	Sites	Primers <sup>1</sup>
TM1190	pODC29	BspHI/ PstI	F-5'-cgtggatc <u>ATGAAAGTGAAGGCACCAGGAAGAATAAAC</u> ; R-5'-cctcagct <u>gcag</u> TTAGATTTTTTGAACACCGTCTGAAC
TM0952	pET28A	NdeI/ Sal1	F-5'- ggactctgcat <u>ATGTACGTTCTCGCAATAGATC</u> ; R-5'- cgatg <u>tcgac</u> TCACTCGGTGAGTAGTTTTCTG
TM0960	pODC29	PciI/ Sal1	F-5'-gcactctgcaac <u>ATGTTTCTAGTGATTTCCGTCGTG</u> ; R-5'-cacgagg <u>tcgac</u> TTATAAATTCTTCAAAAACGCCTC
TM1430	pODC29	Nco1/ Sal1	F-5'- cgatgtcc <u>ATGgAATATGTTCTCTCTCTCGATCAAG</u> ; R-5'- cgatg <u>tcgac</u> TCATTTTGAATACAGCCTCTCTC
TM1430	pODC29 C-His6 tag	Nco1/ Sal1	F-5'- cgatgtcc <u>ATGgAATATGTTCTCTCTCTCGATCAAG</u> ; R-5 'GATCC <u>GTCGACCGTTTTTCCCATCCGAGGGATC</u>
TM1430 mutagenesi s	pODC29 C-His6 tag	Nco1/ Sal1	F-5'- cgatgtcc <u>ATGgAATATGTTCTCTCTCTCGATCAAG</u> ; R-5' GATCC <u>GTCGACCGTTTTTCCCATCCGAGGGATCTCG</u> ACACTGCCTCTTTCCATTTTGAATACA

<sup>1</sup>Introduced restriction sites are underlined; nucleotides not present in the original *T. maritima* gene sequences are shown in lowercase.

Supplemental Table S2. Genes encoding sugar kinases in 12 Thermotogales genomes.

Family	Name	<i>Thermotoga maritima</i> MSB8	<i>Thermotoga</i> sp. RQ-2	<i>Thermotoga petrophila</i> RKU-1	<i>Thermotoga naphthophila</i> RKU-10	<i>Thermotoga neapolitana</i> DSM 4359	<i>Thermotoga lettingae</i> TMO	<i>Thermotoga thermarum</i> DSM 5069	<i>Thermosipho africanus</i> TCF52B	<i>Thermosipho melanesiensis</i> BI429	<i>Fervidobacterium nodosum</i> Rt17-B	<i>Petrotoga mobilis</i> SJ95	<i>Kosmotoga olearia</i> TBF 19.5.1	Functional role	Biological process
FGGY	GntK	TM0443	TRQ2_0492	Tpet_0477	Tnap_0225	CTN_0229	-	-	-	-	-	-	-	Gluconokinase	5-keto-D-gluconate utilization
FGGY	AraB	TM0284	TRQ2_0653	Tpet_0628	Tnap_0916	CTN_0400	Tlet_1150	Theth_1115	-	-	-	Pmob_1609	-	L-ribulokinase	L-arabinose utilization
FGGY	XylB	TM0116	TRQ2_0831	Tpet_0808	Tnap_0746	CTN_0574	Tlet_0361	Theth_0157	-	-	-	Pmob_0224	-	Xylulokinase	D-xylose utilization
FGGY	GlpK	TM1430	TRQ2_1322	Tpet_1364	Tnap_1382	CTN_1064	Tlet_1098	Theth_0993	THA_662	Tmel_0617	Fnod_0409	Pmob_0284	Kole_0534	Glycerol kinase	D-glycerol utilization
FGGY	DrkK	TM0952	-	-	-	CTN_1624	Tlet_1331	-	-	-	-	-	-	putative D-ribulokinase	putative D-arabinose utilization
FGGY	RhaB	TM1073	TRQ2_1744	Tpet_1671	Tnap_1694	CTN_1496	-	-	-	-	-	-	-	Rhamnulokinase	L-rhamnose utilization
PfkB	RbsK	TM0960	-	Tpet_1789	Tnap_1790	CTN_1616	Tlet_1905	Theth_0467	THA_674	-	-	Pmob_0924	Kole_0601	Ribulokinase	D-ribose utilization
PfkB	ScrK	TM0296	TRQ2_0635	Tpet_0616	Tnap_0934	CTN_0387	-	-	-	-	-	Pmob_1154	Kole_0610	Fructokinase	Mannitol utilization
PfkB	KdgK	TM0067	TRQ2_0880	Tpet_0857	Tnap_0697	CTN_0626	Tlet_1863	-	-	-	-	-	-	KDG kinase	Glucuronate utilization
PfkB	IoIK	TM0415	-	Tpet_0505	Tnap_0207	CTN_0254	-	Theth_1091	-	-	-	-	Kole_0981	not determined	Inositol utilization
PfkB	LacC	TM0828	TRQ2_0099	Tpet_0099	Tnap_0099	CTN_1748	Tlet_1061	Theth_1239	THA_916	Tmel_0676	Fnod_0971	Pmob_0182	Kole_0346	not determined	-
PfkB	-	TM0795	-	Tpet_0133	Tnap_0134	CTN_1783	Tlet_0330	Theth_1751	-	-	-	Pmob_0041	Kole_0581	not tested	-
COG2971	BglK	TM1280	TRQ2_1539	Tpet_1491	Tnap_1514	CTN_1292	Tlet_0044	Theth_0445	-	-	-	-	-	Glucosamine kinase	Glucoside utilization?
GHMP	Galk	TM1190	TRQ2_1628	Tpet_1562	Tnap_1581	CTN_1385	Tlet_1104	Theth_1358	THA_1581	Tmel_0425	Fnod_0415	Pmob_0830	Kole_0950	Galactokinase	Galactose utilization
ROK	Glk	TM1469	TRQ2_1363	Tpet_1323	Tnap_1340	CTN_1024	Tlet_0610	Theth_0713	THA_1446	Tmel_1141	Fnod_0088	Pmob_1217	Kole_1870	Glucokinase, hexokinase	Glycolysis
COG0469	Pyk	TM0208	TRQ2_0740	Tpet_0716	Tnap_0838	CTN_0477	Tlet_1962	Theth_1880	THA_1204	Tmel_0922	disrupted by f	Pmob_1917	Kole_1625	Pyruvate kinase	Glycolysis
COG0126	Pgk	TM0689	TRQ2_0240	Tpet_0242	Tnap_0485	CTN_1897	Tlet_0997	Theth_0582	THA_1606	Tmel_1265	Fnod_0739	Pmob_1444	Kole_2021	Phosphoglycerate kinase	Glycolysis
COG2379	GckA	TM1585	TRQ2_1248	Tpet_1207	Tnap_1223	CTN_1169	Tlet_1724	Theth_1383	THA_302	Tmel_0085	Fnod_1253	Pmob_0534	Kole_1768	Glycerate-2-kinase	Serine and glycerate utilization
PfkA	PfkA1	TM0209	TRQ2_0739	Tpet_0715	Tnap_0839	CTN_0476	Tlet_1961	Theth_1881	THA_1802	Tmel_1595	Fnod_1661	Pmob_1916	Kole_1624	6-phosphofructokinase	Glycolysis
PfkA	PfkA2	TM0289	TRQ2_0648	Tpet_0623	Tnap_0921	CTN_0395	Tlet_1950	Theth_1889	-	-	-	-	-	6-phosphofructokinase	Glycolysis