

Subsystem: N-Acetyl-D-Glucosamine Utilization

Olga Zagnitko¹, Andrei Osterman^{1,2}

¹FIG, ²The Burnham Institute

I. Introduction

N-Acetyl-D-Glucosamine (GlcNAc) is a major component of complex carbohydrates. GlcNAc salvage and recycling pathways are present in many species. This subsystem is associated with numerous variations and nonorthologous gene displacements. In eukaryotes, GlcNAc salvage is directly linked (via PAGAM, see diagram) to the biosynthesis of UDP-GlcNAc. This route does not exist in bacteria (which do not have PAGAM), and utilization of GlcNAc proceeds via catabolism to Fructose-6P [1,2].

II. Subsystem notes

A preliminary analysis of this subsystem projected over a broad range of bacterial genomes reveals a number of open problems (missing genes) and allows us to make several functional predictions. For example, the analysis of this subsystem and proteins encoded within an extensive gene cluster in *Thermotoga maritima* (this analysis was performed in collaboration with JCSG team, www.jcsg.org) led to the following conjectures*:

1. Glucosamine-6-phosphate deaminase (EC 3.5.99.6) (G6PD, gene *nagB* in *E.coli*) is functionally replaced by a nonorthologous gene (TM0813). Structural analysis of TM0813 (at JCSG) suggests that this protein, a truncated (single-domain) homolog of a biosynthetic enzyme (*glmU* gene in *E.coli*, TM0148 in *T.maritima*), does not contain a Gln utilization domain required for an amidotransferase reaction. A comparison with the structure-function analysis of GlmU domains described in [3,4] suggests that TM0813 can catalyze only a deaminase/isomerase reaction, consistent with its inferred role in a catabolic pathway of *T.maritima*. Anticorrelation in the occurrence profiles of NagB and TM0813 homologs in a number of genomes are in agreement with this conjecture.

2. Acetyl-D-Glucosamine uptake in *T.maritima* and a number of other bacteria (such as rhizobiaceae) is predicted to be driven by a specialized ABC transport system. *T.maritima* does not contain a GlcNAc-specific component of PTS system (*nagE* gene in *E.coli*), which appears to be replaced by an ABC cassette (genes TM0810-0812), clustered with other *nag*-genes in *T.maritima*.

3. Bacterial N-acetylglucosamine kinase (EC 2.7.1.59) gene has not been previously identified, although the corresponding activity was detected. This activity is absolutely required in *T.maritima* and other bacteria, which lack PTS system (as discussed see above). At least two NAGK candidates have been identified: TM0808 (*nagC* homolog) and TM1280 (a distant homolog of recently characterized eukaryotic NAGK [5,6]). An evidence for the first candidate (currently known solely as GlcNAc operon repressor in *E.coli*) is based on the strong chromosomal clustering and long-range similarity analysis (ROK-kinase family).

For additional comment, definition of functional variant, etc see notes in SEED and at the Subsystem Forum (<http://brucella.uchicago.edu/SubsystemForum>).

*in experimental verification by an on-going collaborative effort at the Burnham Institute and Genome Institute of Novartis Foundation (GNF) in San Diego

1. Functional Roles, Abbreviations, Subsets and Alternative Forms of Enzymes

N-Acetyl-D-Glucosamine Utilization

Alternative forms

Subsets of roles

Column	Abbrev	Functional Role	
1	BHA	Beta-hexosaminidase (EC 3.2.1.52)	
*NAGK	2	NAGKe	N-acetylglucosamine kinase euakryotic type (EC 2.7.1.59)
	3	NAGKb	N-acetylglucosamine kinase bacterial type predicted (EC 2.7.1.59)
	4	NAGPD	N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)
*G6PD	5	G6PD	Glucosamine-6-phosphate deaminase (EC 3.5.99.6)
	6	G6PDa	Glucosamine-6-phosphate deaminase [isomerizing], alternative (EC 3.5.99.6)
*PTS_NAG variations	7	PTS_NAG	PTS system, N-acetylglucosamine-specific IIBC component (EC 2.7.1.69)
	8	PTS_NAGt	PTS system, N-acetylglucosamine-specific IIBC component (EC 2.7.1.69)
	9	PTS_NAGa	PTS system, N-acetylglucosamine-specific IIA component (EC 2.7.1.69)
10	PTS_PPT	Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)	
11	PTS_PCP	Phosphocarrier protein of PTS system	
*ABCNAG Transporter cassette (specificity inferred)	12	ABCNAGc	N-Acetyl-D-glucosamine ABC transport system, periplasmic sugar-binding protein
	13	ABCNAGb	N-Acetyl-D-glucosamine ABC transport system, permease protein 1
	14	ABCNAGb.	N-Acetyl-D-glucosamine ABC transport system, permease protein 2
	15	ABCNAGa	N-Acetyl-D-glucosamine ABC transport system ATP-binding protein
16	NAG_R	Predicted transcriptional regulator of N-Acetylglucosamine utilization, GntR family	
17	NagC	Transcriptional regulator of N-acetylglucosamine utilization	
18	NagD	Phosphatase NagD predicted to act in N-acetylglucosamine utilization subsystem	
19	NAMPE	N-acetylmannosamine-6-phosphate 2-epimerase (EC 5.1.3.9)	
20	PAGAM	Phosphoacetylglucosamine mutase (EC 5.4.2.3)	

2. Subsystem spreadsheet

N-Acetyl-D-Glucosamine Utilization

a fragment of the SEED display with selected examples

Genome ID	Organism	Variant Code	BHA	*NAGK	NAGPD	*G6PD	*PTS_NAG	PTS_PPT	PTS_PCP	*ABCNAG	NAG_R	NagC	NagD	NAMPE		
158878.1	Staphylococcus aureus subsp. aureus Mu50 [B]	11111		44-3	701	569-5	265-7	1084	1083				929	318		
158879.1	Staphylococcus aureus subsp. aureus N315 [B]	11111		43-3	672	541-5	262-7	964	963				812	314		
211586.1	Shewanella oneidensis MR-1 [B]	11128	2042 , 3200	3198-2	3196	3197-6	2031-9	1235 , 2032	2033				2513			
83333.1	Escherichia coli K12 [B]	121110	1092	1580-3, 666-3	667	668-5	669-7, 2385-9	2356 , 2384 , 2782 , 3866	2383				1580 , 666	665	3167	
243274.1	Thermotoga maritima MSB8 [B]	23128	802	1269-2, 405-3, 801-3	807	806-6				803-12 , 804-13 , 805-14			1724			
198094.1	Bacillus anthracis str. Ames [B]	12118		2270-2	3936	3935-5	3931-7, 466-8, 5148-9	3929	3930 , 4982			3934		4803		
272558.1	Bacillus halodurans C-125 [B]	12118	675	1094-3, 2758-3, 700-3	421	420-5	844-7, 422-8, 673-8	3073	3074 , 3566			419	1094 , 2758 , 700	3428		
1491.1	Clostridium botulinum ATCC 3502 [B]	12118		3497-2	665	664-5	669-8, 2768-9, 668-9	1466	926			666				
266834.1	Sinorhizobium meliloti 1021 [B]	23128	2819	1494-2, 1240-3, 4359-3	1497	1496-6		3906				1492-12 , 1491-13 , 1490-14 , 1488-15	1495			
272626.1	Listeria innocua Clip11262 [B]	31111		215-3	2199 , 945	2435-5 , 2850-5 , 870-5 , 946-5	1006-7	367 , 992	991			1829-12 , 218-12	947	215	2483	2913
169963.1	Listeria monocytogenes EGD-e [B]	31131		176-3	2099 , 948	2348-5 , 2710-5 , 870-5 , 949-5 , 35-6	1009-7	348 , 995	994			1722-12	950	176	2391	2790
243365.1	Chromobacterium violaceum ATCC 12472 [B]	33128	2073 , 259	2896-2, 2895-3	556	557-6	559-7, 558-9, 980-9	2311 , 3052 , 558 , 816 , 980				262-12 , 261-13 , 260-14	555		3244	
74547.1	Prochlorococcus marinus str. MIT 9313 [B]	81111	1956	1816-2	2031	1418-5								1329		

Matching colors highlight genes that occur close to each other on the chromosome. Genes (proteins) assigned with respective functional roles are shown by unique SEED IDs. Alternative forms are indicated by additional numbers, dash-separated. Two examples (red box) are further illustrated by projection on a subsystem diagram.

3. Subsystem diagram

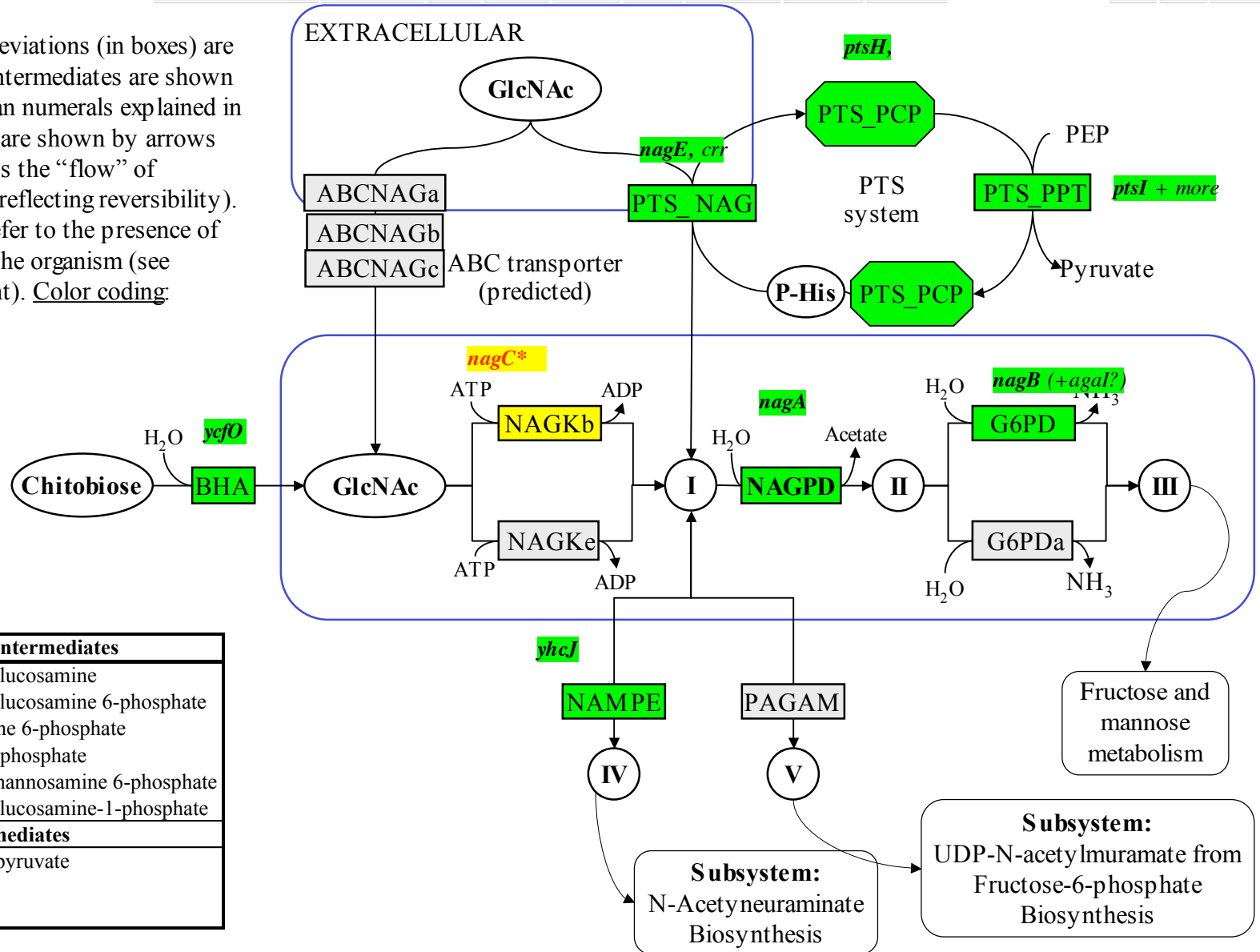
Example: *E. coli* K12

N-Acetyl-D-Glucosamine Utilization

Organism	Variant Code	BHA	*NAGK	NAGPD	*G6PD	*PTS_NAG	PTS_PPT	PTS_PCP	*ABCNAG	NAG_R	NagC	NagD	NAMPE
Escherichia coli K12 [B]	121110	1092	1580-3, 666-3	667	668-5	669-7, 2385-9	2356, 2384, 2782, 3866	2383			1580, 666	665	3167

Functional role abbreviations (in boxes) are as in Panel 1. Key intermediates are shown in circles with Roman numerals explained in the inset. Reactions are shown by arrows (directionality shows the “flow” of subsystem without reflecting reversibility). Highlighted boxes refer to the presence of respective genes in the organism (see spreadsheet fragment). Color coding:

known
predicted
not present



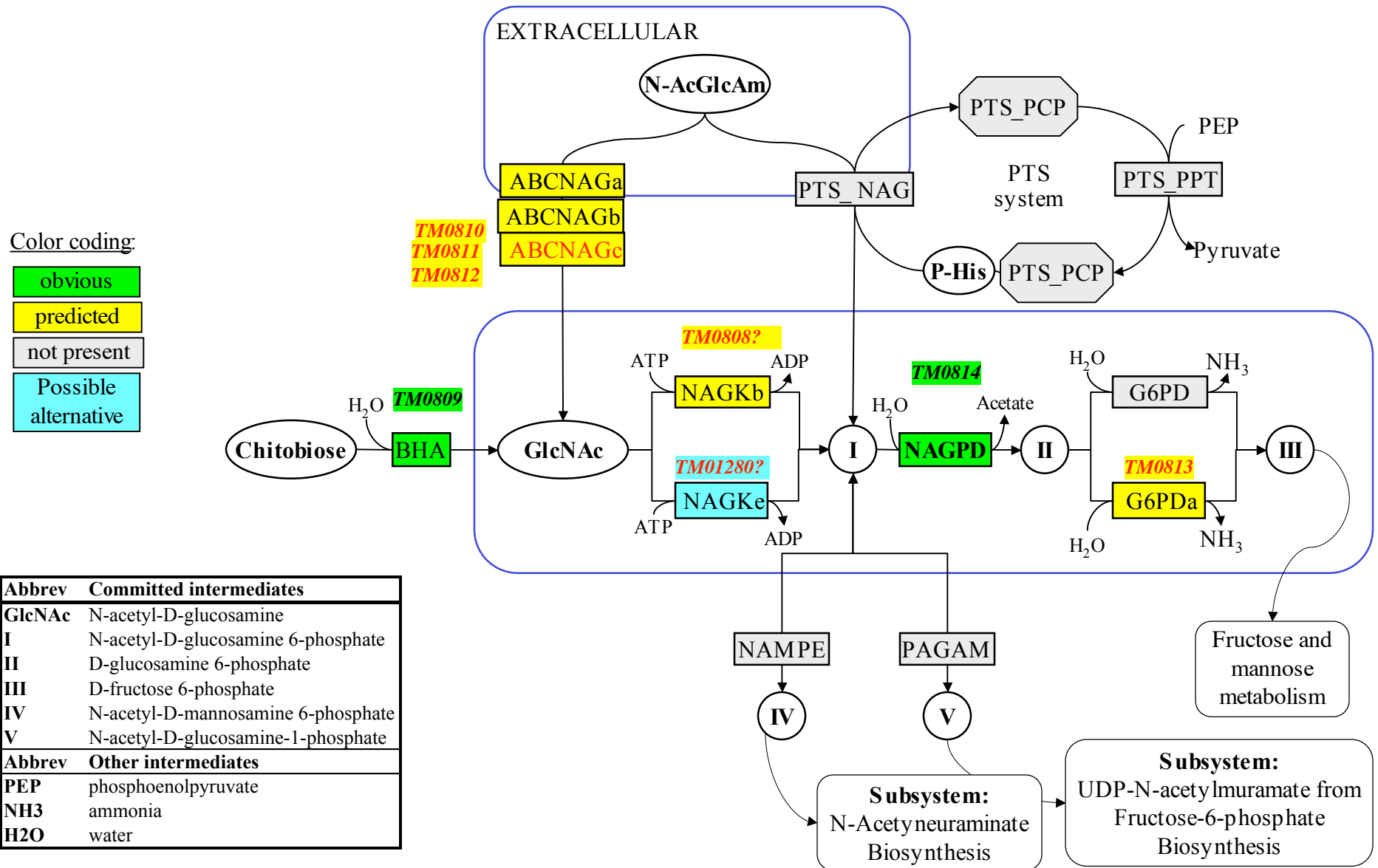
Abbrev	Committed intermediates
GlcNAc	N-acetyl-D-glucosamine
I	N-acetyl-D-glucosamine 6-phosphate
II	D-glucosamine 6-phosphate
III	D-fructose 6-phosphate
IV	N-acetyl-D-mannosamine 6-phosphate
V	N-acetyl-D-glucosamine-1-phosphate
Abbrev	Other intermediates
PEP	phosphoenolpyruvate
NH3	ammonia
H2O	water

4. Subsystem diagram

Example: *T.maritima*

N-Acetyl-D-Glucosamine Utilization

Organism	Variant Code	BHA	*NAGK	NAGPD	*G6PD	*PTS_NAG	PTS_PPT	PTS_PCP	*ABCNAG	NAG_R	NagC	NagD	NAMPE
Thermotoga maritima MSB8 [B]	23128	802	1269-2, 405-3, 801-3	807	806-6				803-12, 804-13, 805-14			1724	



IV. References

- 1: Plumbridge J, Vimr E. Convergent pathways for utilization of the amino sugars N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by *Escherichia coli*. *J Bacteriol.* 1999 Jan;181(1):47-54.
- 2: Plumbridge J, Pellegrini O. Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP. *Mol Microbiol.* 2004 Apr;52(2):437-49.
- 3: Obmolova G, Badet-Denisot MA, Badet B, Teplyakov A. Crystallization and preliminary X-ray analysis of the two domains of glucosamine-6-phosphate synthase from *Escherichia coli*. *J Mol Biol.* 1994 Oct 7;242(5):703-5.
- 4: Teplyakov A, Obmolova G, Badet B, Badet-Denisot MA. Channeling of ammonia in glucosamine-6-phosphate synthase. *J Mol Biol.* 2001 Nov 9;313(5):1093-102.
- 5: Yamada-Okabe T, Sakamori Y, Mio T, Yamada-Okabe H. Identification and characterization of the genes for N-acetylglucosamine kinase and N-acetylglucosamine-phosphate deacetylase in the pathogenic fungus *Candida albicans*. *Eur J Biochem.* 2001 Apr;268(8):2498-505.
- 6: Hinderlich S, Berger M, Schwarzkopf M, Effertz K, Reutter W. Molecular cloning and characterization of murine and human N-acetylglucosamine kinase. *Eur J Biochem.* 2000 Jun;267(11):3301-8.