Proteomic Analysis of Cellulose Metabolism in *Clostridium thermocellum*

Euan Burton

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Biology) at

Concordia University

Montréal, Québec, Canada

November 2010

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Euan Burton

Entitled: Proteomic Analysis of Cellulose Metabolism in *Clostridium thermocellum*

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Titorenko____________________ Chair

Dr. Powlowski____________________ External Examiner

Dr. Storms_______________________ Examiner

Dr. Tsang________________________ Examiner

Dr. Martin_______________________ Supervisor

Approved by Dr. Gulick_______________________

Graduate Program Director

November, 2010 Dr. Lewis________________________

Dean of Faculty

ABSTRACT

Proteomic Analysis of Cellulose Metabolism in *Clostridium thermocellum*

Euan Burton

The cytoplasmic and membrane proteomes of the cellulolytic bacterium *Clostridium thermocellum* grown on either cellulose or cellobiose were quantitatively compared using a metabolic ^{15}N isotope labelling method in conjunction with nanoLC-ESI-MS/MS. In total 1255 proteins were identified in the study and 95 of those were able to have their relative abundance per cell compared in at least one cellular compartment in response to the substrates provided. Three cellulosome anchoring proteins, OlpB, Orf2P, and SdbA, were all found to be more abundant in cellulose grown cells. Cellodextrin transport complex B, appearing to be the major cellodextrin transporter in the organism, did not demonstrate any difference in abundance on either substrate. Cellodextrin transport complex A, putatively a strict cellotriose transporter, demonstrated increased abundance in cellobiose grown cells. This study reveals that cells grown on cellulose increase their abundance of phospoenolpyruvate carboxykinase while decreasing the abundance of pyruvate kinase, suggesting that the organism diverts carbon flow into a transhydrogenase malate cycle that can increase the production of the biosynthetic intermediates NADPH and GTP. Glutamate dehydrogenase and copper amine oxidase proteins were also found to have increased abundance in cellulose grown cells, suggesting that the assimilation of ammonia is up-regulated during growth on the substrate. The up-regulation of enzymes leading to increased production of biosynthetic intermediates and the assimilation of ammonia reveal how the organism produces the substrates necessary to increase the production of cellulosomes required for the hydrolysis of crystalline cellulose.

ACKNOWLEDGEMENTS

I would like to thank my supervisor and mentor Dr. Martin for all of his support and expertise on this project, his thoughtful advice, and for the exciting, welcoming and helpful research environment he has created.

I am very grateful to Dr. Reginald Storms and Dr. Adrian Tsang for sitting on my committee.

I would like to thank all my fellow lab members for their friendship and support. Especial gratitude goes to Nicholas Gold, who helped me a great deal in learning what I needed to begin this project, and to Dr. Heng Jiang, whose expertise in proteomics helped me exceedingly in completing it.

Most importantly I would like to thank my family for their endless love and support. I am shaped profoundly by my late grandfather who, with great patience and thoughtfulness, would always try to answer the endless curiosities I had about the world as seen through a child’s eyes.

I dedicate this work to him.

Table of Contents

List of figures

List of tables

List of tables (cont.)

List of equations

List of abbreviations

AA	amino acid
ABC	adenosine-binding cassette
ACN	acetonitrile
ADP	adenosine diphosphate
AMP	adenosine monophosphate
ATP	adenosine triphosphate
CBP	consolidated bioprocessing
CBM	cellulose binding module
CEM	cellulose-enzyme-microbe
CID	collision induced dissociation
CV	coefficient of variation
DOE	U.S. Department of Energy
dSDS	double sodium dodecyl sulphate
DTT	dithiolthreitol
ELISA	enzyme-linked immunosorbent assay
ESI	electrospray ionization
FA	formic acid
GDP	guanosine diphosphate
GMP	guanosine monophosphate
GTP	guanosine triphosphate
GHG	greenhouse gas
IQR	interquartile range
JGI	Joint Genome Institute
MS	mass spectrometer or mass spectrometry
MS/MS	tandem mass spectrometry
MSD	membrane spanning domain
m/z	mass to charge ratio
nanoLC	nanovolume liquid chromatography
NCBI	National Center for Biotechnology Information
NBD	nucleotide binding domain
PAGE	polyacrylamide gel electrophoresis
PQD	pulsed-Q dissociation
POR	pyruvate:ferredoxin oxidoreductase
PTM	post-translational modification
PPCK	phosphoenolpyruvate carboxykinase
PK	pyruvate kinase
RT-PCR	reverse transcriptase polymerase chain reaction
SBP	substrate binding protein
SDS	sodium dodecyl sulphate

1. Objectives

The aim of this study was to better understand how *Clostridium thermocellum* alternatively modulates its proteome in response to growth on either Avicel (a microcrystalline cellulose model substrate) or on cellobiose. The proteins of the cell were interrogated in two ways. First, a broad study of the composition of the cytoplasmic and membrane cellular fractions during growth on either substrate was investigated by nanoLC-ESI-MS/MS peptide sequencing. Second, a quantitative method was employed that would allow the measurement of changes in expression of proteins found in both the cytoplasmic and membrane fractions of cells grown on either substrate. This quantitative assessment was accomplished using a metabolic isotope-labelling strategy in conjunction with nanoLC-ESI-MS/MS. By comparing the proteomes of Avicel-grown and cellobiose-grown *C. thermocellum* cells, we expected to identify several hundred proteins utilized by the organism, and to uncover patterns in protein expression that improve our understanding of anaerobic cellulose metabolism. The soluble fraction of the cells were expected to provide clues of the intracellular catabolism taking place during growth on cellulose, while the membrane fraction was expected to possess information about the transporters utilized to internalize oligosaccharides hydrolyzed from cellulose and the protein machinery involved in the secretion of cellulase enzymes and cellulosomes.

2. Introduction

Global biofuel production contributed 1.5% of worldwide transport fuel consumption in 2008, a more than 400% increase compared to the production volumes from 2000 [1]. The dramatic increase in the growth of this industry in the last decade has been a result of ambitious support policies from governments such as Canada, the U.S and the E.U. [2]. Many other emerging economies (e.g. Brazil, China, Thailand, Indonesia, and India) have also implemented policies supportive of biofuel production [2]. The intentions of such policies include; reducing the carbon footprint of global energy consumption, increasing domestic energy production, and invigorating rural economies with new markets for their crops. Virtually all biofuel produced today can be described as "first-generation" in that they are produced from the fermentation of sugars that are not derived from lignocellulose. First generation biofuels are produced from crops such as corn, wheat, sugar beet and sugar cane [3]. While 1st generation biofuels may prove to be economically sustainable (depending upon the relative cost of oil and context of supportive policy), heavy criticism has recently been drawn upon them regarding both their social and environmental impacts, with the potential exception of ethanol from Brazilian sugarcane [4]. Increases in the commodity price of agricultural products has initiated worldwide debate as to whether 1st generation biofuels can be produced without endangering food production. Estimates of the effect of biofuels on the increased cost of food prices range from 0-75%, with most literature agreeing on a 15-25% range [2]. Also debated are the net CO_2 reductions of 1st generation biofuels relative to conventional fossil fuels, as it has been shown that some biofuels have very low net lifecycle benefits [1]. Persuasive arguments have been put forward arguing that the actual carbon footprint of current biofuels is actually much worse than that of fossil fuels after a correct budgeting of direct and indirect land use changes are accounted for, depending upon the biofuel feedstock and the type of land use transition [5-6]. The effects of these land use changes are also thought to negatively affect biodiversity due to accelerating deforestation, in addition to increasing competition for scarce water in some regions of the world. In light of these difficulties, efforts are underway to create a new generation of biofuel production technologies that improve yield, environmental sustainability, and originate from non-edible biomass material.

2.1. Second generation biofuels

In response to the limitations of 1st generation biofuels, 2nd generation biofuels are being developed which allow for the conversion of lignocellulose from plant biomass into fuels such as ethanol, butanol and hydrogen. Minimizing the land use appropriated for biofuel production is of utmost importance to the environmental and social sustainability of its use. Low cost agricultural and forestry wastes, wood processing wastes and the organic component of municipal solid wastes can all be used as lignocellulosic feedstocks without the need of any additional arable land [2]. Furthermore up to 1 billion hectares of marginal and degraded lands unsuitable for food production could be utilized for the production of selected perennial energy crops, such as switchgrass and miscanthus [3]. It has been estimated that by up to 25% of the world's primary energy resources could be supplied by biomass by 2050 using available farm, forest and urban residues and by growing perennial energy crops on marginal land [3]. The primary barrier to the commercialization of 2nd generation biofuels into the global economy, and the social and environmental sustainability such commercialization would bring, is the absence of low cost technology to overcome the recalcitrance of lignocellulose.

2.2. Consolidated bioprocessing of cellulosic biomass

Lignocellulose is the major component of plant biomass and is composed of the carbohydrate polymers cellulose and hemicellulose, which are tightly bound to the chemically complex, non carbohydrate polymer, lignin. Due to the complexity of the chemical structure of the lignocellulosic biomass feedstock, and the crystalline structure of the cellulose within it, multiple substrate processing steps need to be achieved to obtain the final biofuel product (e.g. ethanol). The first step in converting biomass to ethanol is pre-treatment, involving cleaning and breakdown of the structure of the lignocellulosic materials. A combination of physical and chemical (e.g. acid hydrolysis) processes are typically applied, which allow a crude separation of the biomass into its cellulose, hemicellulose and lignin components [7]. Some hemicellulose can be converted into its component sugars in this step, and the lignin is removed and can be burned as fuel onsite in place of fossil fuels, to offset the energy and carbon costs of the process [4]. The pre-treatment of the biomass renders the remaining polysaccharide portions more amenable to subsequent enzymatic hydrolysis. After pre-treatment the remaining cellulose and hemicelluloses from the feedstock is hydrolysed into fermentable sugars in the major saccharification step. In process designs where saccharification and fermentation of the biomass occur separately, it is referred to as separate hydrolysis and fermentation (SHF). An important process modification made for the enzymatic hydrolysis of biomass was the introduction of simultaneous saccharification and fermentation (SSF) which allows saccharification and fermentation of the biomass to occur in the same vessel. The processing configuration can be further improved to include the co-fermentation of multiple sugar substrates including pentoses present in hemicellulose in a process termed simultaneous saccharification and co-fermentation (SSCF). In the SSCF process, cellulose, cellulase enzymes and co-fermenting microbes are combined, reducing the required number of vessels and their associated costs. Cocktails of cellulolytic enzymes are available commercially and their costs have reduced more than 30-fold in recent years [8]. Despite this dramatic reduction in cost, enzymatic hydrolysis of pre-treated cellulose and the cost of making such cellulase enzymes is second only to the pre-treatment itself in terms of cost-barriers to the commercialization of 2nd generation biofuels [9]. In other words, the three most expensive processing operations are for overcoming the natural resistance of plants to the biological breakdown of sugars.

It is because of these costs that researchers are now looking at the possibility of a consolidated bioprocessing (CBP) approach that could produce all required enzymes within the reactor vessel, allowing a single microbe or microbial community to both produce the enzymes necessary to break down cellulose to sugars and to ferment all the available sugars into a biofuel, such as ethanol. Consolidated bioprocessing merges four biologically mediated transformations, allowing them to take place within the same reactor: cellulase and hemicellulase production; hydrolysis of pretreated biomass into sugars; the fermentation of hexose sugars; and the fermentation of pentose sugars [10]. CBP is seen as the next paradigm of biomass conversion technology and has great potential for improving the efficiency of the bioconversion process and reduction of its cost [11]. Research is underway to better understand which microbes are capable of the transformations necessary for CBP and how these organisms and their adaptations can be applied to commercialization of second generation biofuels using biotechnology.

2.3. *Clostridium thermocellum*

One microorganism that has received substantial consideration for implementation in CBP is *Clostridium thermocellum*, a gram-positive, anaerobic, thermophilic bacterium with an optimum growth temperature of 60°C. *C. thermocellum* is widespread in nature as its habitat is organic plant material in decomposition [12]. Investigators have found strains in municipal/agricultural wastes [6], sewage digestion sludge [13], soil [14], cotton bales [15], river mud [16], and hot springs [17]. It is able to use cellulose as a sole carbon source, and is capable of diverse product fermentation, producing various amounts of acetate, ethanol, formate, lactate, H_2 and CO_2, under different growth conditions [21]. The genome of *C. thermocellum* has been sequenced by the Joint Genome Institute (JGI). The high growth temperature of the organism is an attractive feature for a potential CBP industrial strain as it would reduce the cooling costs necessary during the bioconversion process, and would also make the process less susceptible to contamination from other microorganisms. Furthermore, the higher temperature also makes ethanol removal and recovery more economical than it is for other industrial organisms grown at lower temperatures. The thermophilic nature of *C. thermocellum* also results in enzymes from the organism being very robust. Anaerobiosis is an advantage for large-scale culture as one of the most expensive steps in industrial fermentations is that of supplying sufficient oxygen transfer, e.g., for cellulase production [21].

C. thermocellum has attracted research interest due mainly to its cellulolytic and ethanologenic abilities. The organism's capability of directly converting cellulosic substrates, such as dilute acid pre-treated hardwood, into ethanol offers a way forward towards the commercial realization of 2nd generation biofuels [20]. *C. thermocellum* posses the enzymatic activities necessary to degrade hemicellulose, though it does not grow on the pentose sugars released from this substrate. Already work is being done to find a hemicellulolytic thermophilic partner, such as a genetically modified strain of *Thermoanaerobaterium saccharolyticum,* to cooperate within a stable co-culture capable of fermenting all solubilised sugars from cellulose into fuel [22]. The ethanol yield of *C. thermocellum* could theoretically be improved by eliminating competing pathways of acetate and lactate production using recently developed genetic tools [23-25]. Cellulose degradation in *Clostridium thermocellum* is accomplished by a large extracellular cellulase system called the cellulosome [26]. The cellulosome of *Clostridium thermocellum* has been extensively studied and has become a paradigm for our understanding of extracellular, complex-mediated cellulose hydrolysis [21, 27-30].

2.3.1. Cellulosome mediated cellulose hydrolysis

The cellulosome of *Clostridium thermocellum*, first described by Lamed in 1983 [26], is composed of multiple enzymatic subunits bound tightly to a large non-catalytic scaffoldin protein (CipA), via type I dockerin-cohesin protein domain interactions. Type I dockerin domains present on the cellulosomal enzymes allow membership to the complex by docking with extreme affinity to nine modularly organized type I cohesion modules found on the scaffoldin core [30]. The scafoldin protein is cell associated by virtue of a second class of cohesion-type II domains on the scaffoldin protein with strict specificity for a second class of type II dockerin domains found on anchoring proteins attached to the cell [27]. The genome of *C. thermocellum* encodes five proteins with Type II cohesins. Four posses S-layer homology domains (SdbA, OlpB, Orf2p and Cthe0735) for cell-surface anchoring of type II dockerin bearing CipA, and the other (Cthe_0736) is without an SLH domain. Some of these CipA anchoring proteins posses more than one type II cohesin domain, and can thus potentially anchor more than one cellulosome to the cell surface. The cellulosome complex is targeted to the cellulose substrate by means of a cellulose binding domain (CBD) found on the scaffoldin protein in addition to specific cellulosomal enzymes that themselves contain a CBD (Figure 1) [31]. On the cell surface, cellulosomes appear as polycellulosomal aggregates which mediate the attachment of cells to cellulose fibres [32].

Cellulosomes allow a highly efficient method of cellulose hydrolysis compared to the secretion of non-complexed cellulase enzymes as seen in aerobic cellulolytic fungi such as *T. reesei.* Multiple layers of synergistic improvement are afforded to dockerin containing cellulase and hemicellulase enzymes as they cooperate to degrade complex lignocellulose substrates within a broad cellulose-enzyme-microbe complex (CEM) [18]. This is accomplished by the targeting of diverse cellulolytic activities in close proximity to their substrates and each other, maximizing the synergistic effect of those activities at the cellulose-enzyme interface. Also, by keeping cells nearby the source of released sugars, the CEM complex optimizes active transport of extracellular hydrolytic end products into the cell [33]. These soluble oligosaccharide end-products, cellodextrins ranging from two to seven β-1,4 linked glucose monomers, are also known to have inhibitory properties on the cellulolytic enzymes generating them [18, 34-35], thus active internalization of the soluble cellodextrin end-products of cellulose hydrolysis serves the activity of the cellulosome

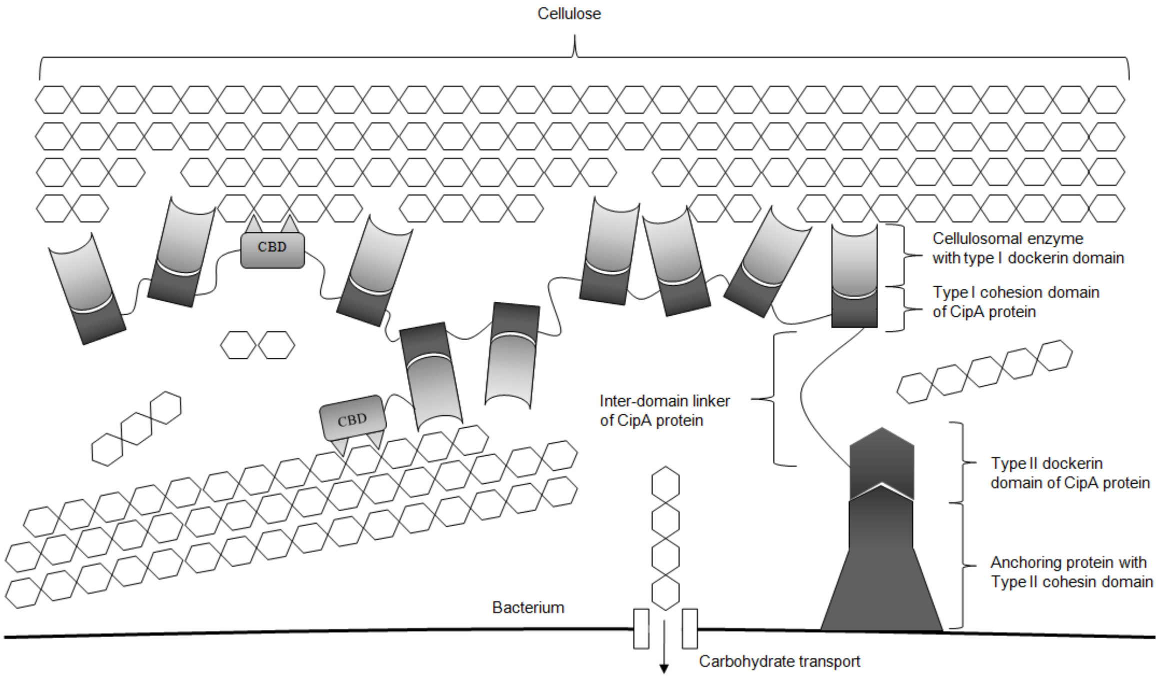


Figure 1. A schematic model of a *C. thermocellum* cellulosome and the cellulose-enzyme-microbe (CEM) interaction. The scaffoldin protein, CipA, is shown, bound to the bacterial surface by an anchoring protein via a single type II cohesin-dockerin interaction. The scaffoldin protein contains a cellulose binding domain (CBD), interlinked with nine cohesion domains by way of inter-domain linkers. Various dockerin-domain containing cellulosomal enzymes bind the scaffoldin protein, mediated by the high affinity type I cohesin-dockerin interaction. Some cellulosomal enzymes posses their own CBD.

2.3.2. Internalization and catabolism of soluble cellodextrins

Transport of cellodextrins by *C. thermocellum* into the cell occurs via ATP binding cassette (ABC) oligosaccharide transporters and is ATP dependent [36]. ABC transporters are multi-component systems composed of two membrane spanning domains (MSD) that forms the channel required for transport, and two nucleotide binding domains (NBD) responsible for hydrolysing ATP and initiating the conformational change allowing active transport. These four domains can each be encoded by a separate polypeptide, or domains may be fused together in any number of ways into multi-domain polypeptides [37-38]. Auxiliary substrate binding proteins (SBP) have been recruited into many ABC transporters for the purpose of presenting the substrate to the channel outside of the cell [39-40]. Five sugar ABC transporters have been identified in *C. thermocellum* based on the published genome sequence, and their substrate affinities have been inferred based on *in vitro* microcalorimetric titration studies of their corresponding heterologously expressed SBP sequences [21]. It was found that four of the SBPs (CbpA, CbpB, CbpC, and CbpD) possess specific substrate affinities for cellodextrins and the other possesses affinity for laminaribose (Lbp) [41]. These five SBPs are encoded by genes found in operons that also encode the remaining domains necessary for the complete functioning of an ABC transporter (Figure 2), with the exception of the operon containing gene cbpB, which is missing a gene that would encode a NBD, and so it must be encoded elsewhere in the genome for this ABC transporter to function.

Once inside the cell, cellodextrins undergo phosphorolytic cleavage [42-46] allowing for the energy in the β 1-4 glycosidic bonds to be conserved in the form of a glucose-1-phosphate product. This product can then be isomerised to glucose-6-

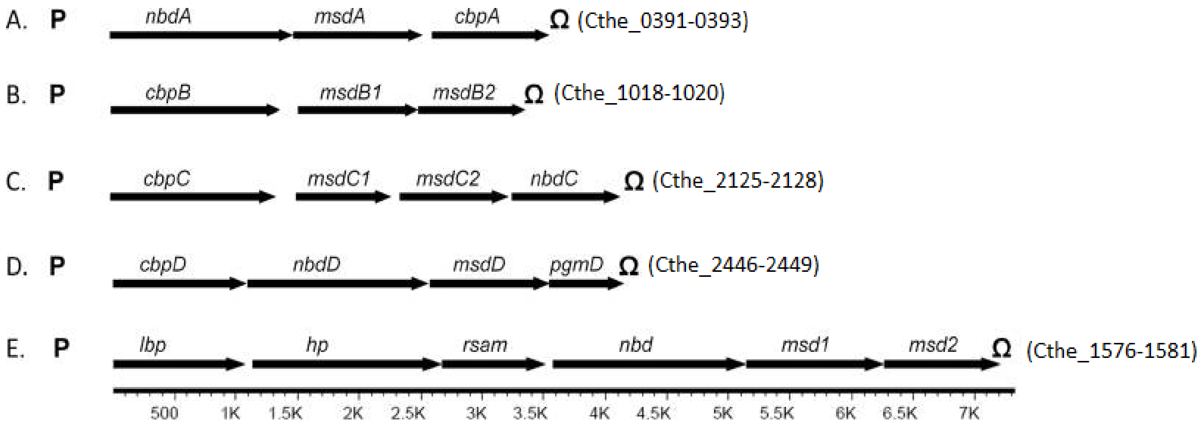


Figure 2. Genetic organization of five putative sugar ABC transporters in *C. thermocellum (*Sugar ABC transport complexes *A-E).* Figure reproduced from [41]. *cbp* encodes cellodextrin-binding protein, *lbp* encodes laminaribiose-binding protein, *msd* encodes membrane-spanning doamin protein, *nbd* encodes nucleotide-binding domain protien, *pgm* encodes phosphoglycerate mutase, *hp* encodes a hypothetical protein, and *rsam* encodes radical *S*-adenosylmethionine. Gene numbers, as annotated in the *C. thermocellum* genome database (version NC_009012.1), of component genes within the operon are provided at the far right of the figure. The letters P and Ω symbolize potential promoters and rho-independent terminators, respectively.

phosphate by the action of phosphoglucomutase (EC 5.4.2.2). Phosphorolytic cleavage is catalyzed by cellobiose phosphorylase (CbP) (EC 2.4.1.20) and cellodextrin phosphorylase (CdP) (EC 2.4.1.49) according to the following reactions:

Eq. 1
$$G_2 + P_i \overset{CbP}{\Leftrightarrow} G + G1P$$

Eq. 2
$$G_n + P_i \overset{CdP}{\Leftrightarrow} G_{n-1} + G1P$$

where G_n denotes a cellodextrin oligomer of length *n*, P_i denotes inorganic phosphate, and G1P denotes glucose-1-phosphate.

Phosphorolytic metabolism is superior to typical hydrolytic metabolism in the context of feeding glucose to the Embden-Meyerhof pathway, since the cellular ATP pool is invested in the phosphorylation of only a single glucose molecule per cellodextrin to glucose-6-phophate. Thus it is expected that *C. thermocellum* could produce greater net ATP per glucose molecule from the consumption of longer chain cellodextrins because of net savings of ATP/glucose in active transport, and the reduced investment of ATP/glucose for the phosphorylation of glucose to glucose-1-phosphate. It has been demonstrated that cellodextrins of an average length of n=4 are internalized in *C. thermocellum*, and that the energetic benefits of efficient transport and phosphorolytic cleavage of cellodextrins outweigh the increased energetic costs of cellulosome synthesis when the organism is grown on cellulose as compared with cellobiose [47]. Rates of phosphorolytic cleavage of cellodextrins in cells of *C. thermocellum* grown on Avicel are 20 fold higher than rates of hydrolysis [43].

Understanding how *C. thermocellum* modifies its metabolism in response to the availability of an insoluble cellulose substrate will offer insights into the specific proteins necessary for cellulose hydrolysis and transport in this organism, and how cellulose derived sugars are converted into energy and fermentative end-products. Such extensive analysis of cellular proteins and pathways is now possible in biological research using mass spectrometry.

2.4. Quantitative proteomics using mass spectrometry

Mass spectrometry (MS) is an important developing method for large-scale determination of gene and cellular function directly at the protein level. Shotgun MS methodologies have been developed for the study of whole proteomes or sub-proteome fractions [48-50]. While many pioneering studies demonstrated the usefulness of MS to create valuable inventories of protein content of subcellular structures and organelles [51], MS approaches are emerging as powerful methods assessing quantitative differences in protein abundance between cells or subcellular fractions by comparing the proteomes in two or more physiological states. For example, comparative proteomic analysis of *E.coli* grown on alternative carbon sources has been investigated revealing several enzymes and binding proteins related to lactose metabolism [52]. MS analysis in conjunction with 2D-PAGE of *E. coli* proteomes differentiated by growth on either glucose or oleic acid demonstrated 52 proteins with altered expression during growth on the long chain fatty acid as sole carbon source [53]. Quantitative MS analysis of *S. cerevisiae* in response to the stressor furfural revealed 107 differentially expressed proteins including those involved in glucose fermentation, the tricarboxylic acid cycle, stress response, and sulphur amino acid biosynthesis [54]. Quantitative proteomics has also led to the discovery of signaling pathways involving insulin and an insulin-like growth factor modulating longevity in *C. elegans* [55], and of protein substrates for protein kinases involved in DNA repair in response to radioactivity in human cell lines [56].

Recent studies have begun to apply MS based proteomics to the identification of proteins involved in *C. thermocellum* physiology including; the identification of ethanol-adapted cellular proteins demonstrating gel-spot density profile changes on denaturing SDS-PAGE gels [57], and the identification of cellulosomal proteins purified by 2D-PAGE [58]. Quantitative comparisons of *C. thermocellum* have been restricted to the profiling of cellulosomal proteins, with the relative changes in composition of cellulosomes being examined when grown on Avicel compared to cellobiose [59], and later examining the effects of growth on Avicel with added hemicellulose or pre-treated switchgrass on cellulosome composition [60].

2.4.1 Peptide sequencing by tandem mass spectrometry

In the shotgun proteomics MS/MS approach, a protein mixture is cleaved into smaller peptide fragments, typically enzymatically using trypsin. The peptide mixture is then resolved using high performance liquid chromatography (HLPC) before elution and ionization from a micro-capillary tip to which a voltage is applied. In this process of electrospray ionization (ESI), gas phase peptide ions are non-destructively generated prior to entering the mass spectrometer through an ion transfer tube. In ion trap MS such as the Thermo LTQ used in this study, ions entering the MS are focused into the linear ion trap, a region where combined electric and magnetic fields capture ions within a confined region under vacuum. Peptide ions within the trap are surveyed in the initial MS event by the electron multiplier, and a selected ion species of interest is held within the trap while the remaining ions are ejected. The isolated parent ion can then be fragmented by collision with helium dampening gas. In the second MS event the fragment ion spectrum is detected, from which the amino acid sequence of the parent ion can be deduced. Thus peptides are sequenced by a tandem of MS events. All the spectra are recorded digitally allowing the sequence information gleaned from MS/MS fragmentation to be matched to theoretical peptides from a protein sequence database cleaved *in silico* with the same cleavage method used experimentally (e.g. trypsin) [61]. The success of an assignment of a peptide MS/MS spectrum to a theoretical peptide in the database can be scored using the SEQUEST algorithm. This cross-correlation score (XC) evaluates the number of fragment ions observed in an MS/MS spectrum, and compares it to what should be observed theoretically. The relative intensities of the fragments also contribute to the score [62]. Such algorithms enable consistent application of stringency conditions to MS/MS spectra, allowing the removal of low confidence peptide and protein identifications.

2.4.2 Relative quantitation using isotope labelled internal standards

The intensity of an observed peptide during an MS experiment is a function not only of its abundance, but also the ionization potential of the peptide, based upon its specific chemical structure, in addition to co-eluting ions causing background and ion suppression effects. For this reason the intensities of observed peptides do not correlate well to their general abundance in the biological sample. An isotopically labelled peptide would, however, co-elute with its non-labelled pair and thus be affected by the same matrix effects, and would itself have the same ionization potential as its chemically identical heavier twin. Any difference in intensity of the pair (having discernible m/z ratios) would thus be a consequence of the relative abundance of the two species. This fact provides the basis for metabolic labelling strategies aimed to uncover fractional differences in the abundance of proteins from two biological samples (Figure 3). The relative intensity of a labelled and unlabelled peptide pair are surveyed multiple times as they co-elute from the column, allowing the construction of ion-current curves for each peptide of the pair. The area under these curves are proportional to the ion's abundance allowing the calculation of the relative abundance of the sample and reference peptides (Figure 4) [63]. Software, such as RelEx [64], is available for the calculation of such peptide ion-current ratios. RelEx is also able to score the overlap of a peptide pair's ion-current curves using a least-squares regression to evaluate the confidence of the quantitative assessment. The use of isotopically labelled internal standards has the additional benefit, if added prior to the extraction procedure, of controlling for losses during sample preparation, handling and injection. The use of internal standards has become a standard method in quantitative proteomics.

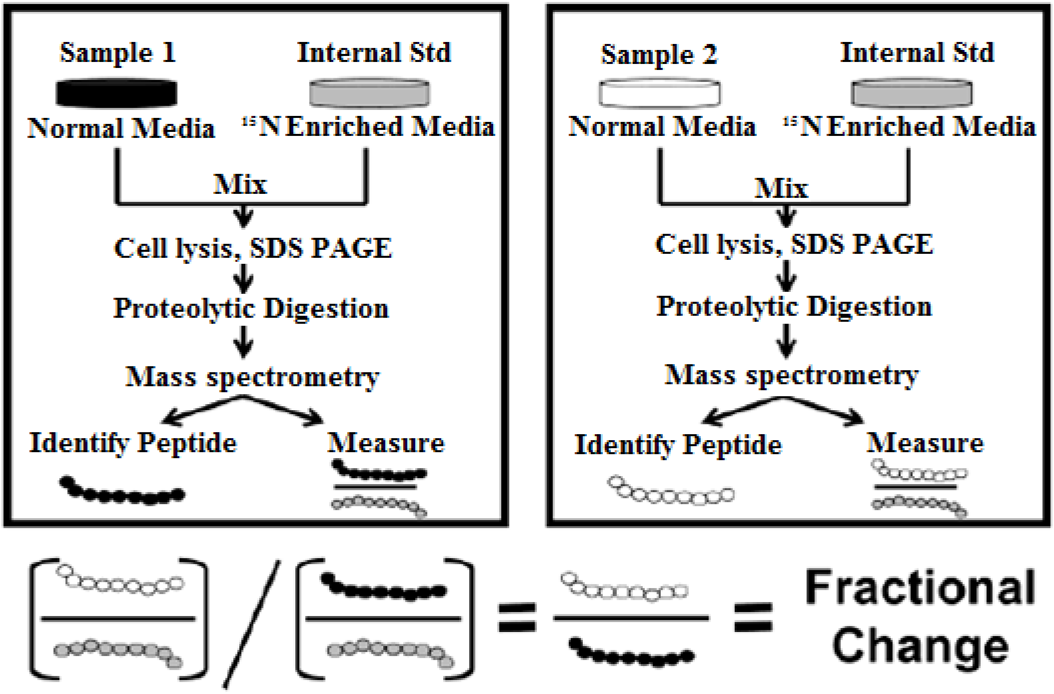
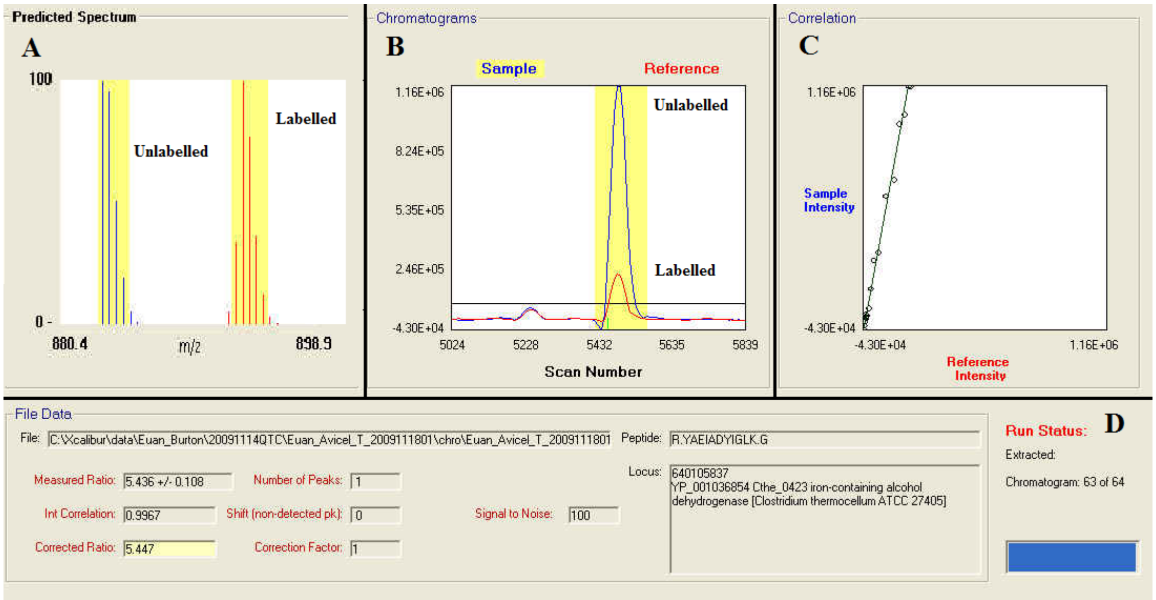


Figure 3. ^{15}N Metabolic labelling strategy used for determining the fractional change in protein abundance between two biological samples, reproduced from [65]. Cells grown under condition 1 and 2 are mixed separately with reference cells enriched in ^{15}N containing media. After protein extraction and enzymatic digestion, measurements of a peptide's intensity relative to a ^{15}N reference peptide can be calculated. The ratio of these intensities relative to the reference provides the actual fractional change of the peptide across both samples.



File: C:\Xcalibur\data\Euan_Burton\20091114QTC\Euan_Avicel_T_2009111801\chro\Euan_Avicel_T_2009111801

Peptide: R.YAEIADYIGLK.G

Measured Ratio: 5.436 +/- 0.108

Number of Peaks: 1

Locus: 640105837
YP_001036854 Cthe_0423 iron-containing alcohol dehydrogenase [Clostridium thermocellum ATCC 27405]

Int Correlation: 0.9967

Shift (non-detected pk): 0

Signal to Noise: 100

Corrected Ratio: 5.447

Correction Factor: 1

Run Status:
Extracted:
Chromatogram: 63 of 64

Figure 4. RelEx software display for calculating the relative abundance of ^{15}N labelled and unlabelled peptides [64]. (A) Display of the isotope distributions of unlabelled and labelled peptides acquired from mass spectra. (B) Overlapping extracted ion chromatogram (XIC) constructed from the signals of a co-eluting peptide isotope pair over time. (C) Graphical display of the correlation factor measuring the overlap of the XICs. (D) Results of calculations on the co-eluting XICs including the relative abundance of the peptide isotope pair calculated as the ratio of the areas under each curve.

3. Materials and Methods

The general strategy for investigating the global cellular response of the proteome of *C. thermocellum* to growth on crystalline cellulose was as follows. *C. thermocellum* was grown in liquid batch culture with either Avicel or cellobiose as the source of carbon. For the purpose of simply identifying proteins present in either of these conditions, these samples were harvested without the inclusion of metabolically labelled reference cells. In separate experiments aimed at discovering the relative abundance of different proteins under the carbon source experimental conditions, each of these sample cultures was mixed with a metabolically labelled ^{15}N reference culture ($^{15}NH_4Cl$) grown on Avicel as sole carbon source. The cells were harvested and their proteins divided into soluble or membrane cellular fractions separated by SDS-PAGE. The gel resolved proteomes were enzymatically digested and the resulting peptides were analysed by nanoLC-ESI-MS/MS. The resulting spectral information was used to identify proteins found in the samples using the SEQUEST algorithm and the most up-to-date *C. thermocellum* protein sequence database. ^{15}N labelled peptides from the reference culture served as internal standards allowing for the quantitative analysis of protein relative abundance between the two samples, using RelEx software.

3.1. Growth media for *C. thermocellum*

Liquid growth media for *C. thermocellum* was prepared as described in [59]. The media is based on ATCC medium 1191 and is prepared as a composite of three distinct solutions: a buffer solution, a vitamin solution, and a mineral solution.

The composition of the buffer solution is provided in Table 1. Sodium hydroxide was added after all other components were completely dissolved. Ninety-five mL of the buffer solution was dispensed into Bellco Glass anaerobic culture bottles. Where Avicel PH101 was to be used as a carbon source, 0.20 g was added to each bottle (0.20% wt/vol final Avicel concentration). The solutions were sparged in the bottles with nitrogen for 5 min, corked with a rubber stopper, and sealed with an aluminum cap using a handheld crimper. The sealed bottles were sterilized by autoclaving on liquid cycle for 15 min. The vitamin solution was made by combining the ingredients listed in Table 2. Ten mL aliquots of the vitamin solution were frozen at -20°C for future use. The ingredients of the mineral solution are provided in Table 3, the nitriloacetic acid was dissolved first by suspending in 500 mL of distilled water and titrating to pH 6.5 with 2 N KOH. Stock mineral solution was filter-sterilized into a sterile bottle and left at room temperature for storage.

Before being added to the buffer solution, 1 mL of the vitamin solution and 10 mL of the mineral solutions was mixed and brought up to a final volume of 100 mL. Vitamin-mineral mixtures to be used for samples using cellobiose (Sigma-Aldrich) as a carbon source had 4.0 g of the disaccharide dissolved in the final volume (4.0% wt/vol final concentration). The mixtures were them sparged with nitrogen for 5 min. At the end of this time, still under continuous nitrogen sparging, the solution was drawn into a 60-mL syringe via a stainless steel cannula. The cannula was then removed and replaced with a 0.2 um filter and syringe needle. The degassed solution within the syringe was then sterilized by passing it through the filter into an anaerobic culture bottle already degassed, stoppered, and autoclaved. Five mL of this final vitamin mineral mixture was added to 95 mL of the sparged and autoclaved buffer solution. For samples containing cellobiose, the final concentration is 0.20% (wt/vol).

3.2. Growth conditions and ^{15}N metabolic labelling.

The effect of carbon source on the proteome of *C. thermocellum* was examined by comparing the relative abundances of proteins discovered in cells grown on either cellulose or cellobiose. Batch cultures of the microbe were grown aerobically at 58°C in 100 mL of full ATCC 1191 medium. Cultures were provided 0.20% (wt/vol) of either Avicel or cellobiose. In addition to these two experimental sample types to be directly compared, an ^{15}N labelled reference sample was grown in minimal ATCC media. This reference sample is different from the experimental Avicel sample in that 99% ^{15}N-enriched NH_4Cl (Cambridge Isotope Laboratories) was used as the nitrogen source and pyridoxal HCL was used in place of pyridoxine HCL. Unlabelled Avicel grown cells were inoculated at 5% (vol/vol) into labelled reference media, and a total of three such passes of ^{15}N enrichment were preformed before inoculation of the reference batch utilized in experiments (estimated enrichment of ^{15}N 98.9%). Cultures were harvested in stationary phase (70h for Avicel, 35h for cellobiose).

Table 1. Buffer solution ingredients, per L, for ATCC 1191 liquid growth medium

	Full medium	Minimal medium
KH_2PO_4	1.58 g	1.58 g
$Na_2HPO_4 \cdot 12H_2O$	4.42 g	4.42 g
$^{14}NH_4Cl$	0.53 g	
$^{15}NH_4Cl$		0.53 g
$MgCl_2 \cdot 6H_2O$	0.19 g	0.19 g
L-cysteine HCL	0.53 g	0.53 g
Yeast extract	2.11 g	
Resazurin (0.1% wt/vol)	1.05 mL	1.05 mL
NaOH (10N)	842 µL	842 µL

Table 2. Vitamin solution ingredients, per L, for ATCC 1191 liquid growth medium

	Full medium	Minimal medium
Biotin	40 mg	40 mg
p-Aminobenzoic acid	100 mg	100 mg
Folic acid	40 mg	40 mg
Pantothenic acid calcium salt	100 mg	100 mg
Nictinic acid	100 mg	100 mg
Vitamin B12	2 mg	2 mg
Thiamine HCl	10 mg	10 mg
Pyridoxine HCL	200 mg	
Pyridoxal HCl		200 mg
Thioctic acid	100 mg	100 mg
Riboflavin	10 mg	10 mg

Table 3. Mineral solution ingredients, per L, for ATCC 1191 liquid growth medium

Nitrilotriacetic acid	1.5 g
$MgSO_4 \cdot 7H_2O$	3 g
$MnSO_4 \cdot H_2O$	500 mg
NaCl	1 g
$FeSO_4 \cdot 7H_2O$	100 mg
$Co(NO_3)_2 \cdot 6H_2O$	100 mg
$CaCl_2$ (anhydrous)	100 mg
$ZnSO_4 \cdot 7H_2O$	100 mg
$CuSO_4 \cdot 5H_2O$	10 mg
$AlK(SO_4)_2$ (anhydrous)	10 mg
Boric Acid	10 mg
$Na_2MoO_4 \cdot 2H_2O$	10 mg
Na_2SeO_3 (anhydrous)	1 mg

3.3. Enrichment of soluble and membrane proteome fractions

Cultures were washed three times in 50 mM Tris buffer (Sigma-Aldrich), pH 8.5, by centrifuging at 12 000 g for 10 minutes at 4°C. The cell pellet was then resuspended in 10 ml lysis buffer (50 mM Tris buffer pH 8.5, 10 mm DTT, 625 units Benzonase (Sigma-Aldrich), 1x EDTA free protease inhibitor cocktail (Roche). For quantitative experiments one half of the suspended cells of each sample were sonicated on ice using a Sonic Dismembrator model 500 (Fisher Scientific) for 2 minutes by twelve 10 second pulses followed by 10 seconds for cooling after each pulse. After lysis the cells were spun down at 16 000 g for 20 minutes at 4°C, and the protein concentration of the supernatant was calculated using the coomassie (Bradford) protein assay. This measurement of protein concentration was used as the basis to equalize the number of cells mixed in the experiment. Equal cellular amounts of unbroken ^{14}N Avicel and ^{14}N cellobiose cell samples were then mixed with unbroken ^{15}N reference cells at a 1:1 protein ratio. The two samples, mixed with the reference, were sonicated as above, spun down at 16 000 G for 20 minutes at 4°C, and the supernatants were centrifuged at 200 000 g for 1 hour at 4°C. The resulting supernatant represents the soluble fraction. The membrane pellet was resuspended in 50 mM Tris pH 8.5 and centrifuged at 200 000 g for 1 hour 3 times to wash the membrane fraction from contaminating soluble fraction protein. The final membrane pellet was stored at -80°C before SDS-PAGE analysis. For non-quantitative experiments aimed at identifying a maximum number of proteins in each sample, cells were not mixed with ^{15}N reference cells, but were fractionated, harvested, and washed in the same way.

3.4 SDS-PAGE separation and in gel digestion of protein fractions.

Quantitative soluble protein fraction data presented in this thesis are compiled from two separate batch culture replicate experiments. Quantitative membrane fraction data was compiled from a single batch culture experiment.

Protein fractions were resolved by 10-20% Tris-HCl SDS-PAGE gradient gel (BIO-RAD). The soluble fraction samples were loaded at a total of 60 µg per lane after boiling in 1x SDS sample buffer (2% SDS, 80 mM Tris pH 6.8, 10% glycerol, 0.002% bromphenol blue, 5% β-mercaptoethanol) for 10 min. In order to prevent membrane protein precipitation and the resulting smearing during SDS-PAGE, the membrane fractions were loaded at 30 µg per lane after incubating with 1x SDS sample buffer at 37°C for 30 min followed by water-bath sonication for 10 min. After staining the gels with Bio-Safe Coomassie (BIO-RAD) the sample lanes were excised and divided into 20 gel bands. The protein in each gel band was reduced and alkylated in order to prevent cysteine thiols from forming oxidized disulfide bonds. Protein from the soluble and membrane fraction samples were digested in-gel with trypsin TPCK (Sigma-Aldrich) as described in appendix A. The resulting peptide mixtures were extracted from the gel pieces using excess buffer, dried, and then reconstituted in 35 µL of 2% (vol/vol) acetonitrile (ACN), 1% (vol/vol) formic acid (FA).

3.5. NanoLC-ESI-MS/MS analysis of gel separated samples

A microcapillary column (New Objective; 75um inner diameter, 30 um tip diameter) was packed 10 cm in length with C18 resin (Magic, 5 um particle size, 300 A pore size) suspended in acetonitrile under pressure using helium in a "packing bomb" as previously described [66]. The microcolumn was positioned in-line with a Finnigan microautosampler and a Finnigan Surveyor MS pump directly in front of an LTQ linear ion-trap mass spectrometer (ThermoElectron). Ten µl of each peptide mixture were loaded individually, on-line, using the autosampler. Each peptide mixture corresponding to a specific gel band was analysed twice by MS (2x10ul) so that identifications from both runs could be pooled together. After being loaded by the autosampler, the samples are cleaned of impurities using a C18 trap (180µm x 2.5 cm) packed with the same material used for the microcolumn by the loading solvent mixture composed of 93% solvent A, and 7% solvent B (solvent A, 1% (vol/vol) formic acid in H_20; solvent B, 1% (vol/vol) formic acid in ACN). The peptides were then eluted to the C18 microcolumn and resolved using a 90-min method in which the ACN gradient rose from 7%-40% of solvent B over 42 min, then from 40%-65% solvent B over 9 min, rising next to 75% solvent B over 4 min, before equilibrating back to 7% solvent B to prepare for the next sample. The LTQ was equipped with an ESI nanosource and operated in positive ion mode with a voltage of 1.1-1.4 kV applied at the liquid junction at the base of the C18 microcolumn. Detection and sequencing of the peptide ions was initiated by full MS survey scans on ions within a 400-1400 m/z window, followed by 8 tandem MS data dependent scans. In this way the 8 most abundant ions from each survey scan were sequenced using pulsed-Q dissociation ion fragmentation. In order to free time for the detector to spend searching for less abundant peptides, ions that triggered a data dependant scan more than once within 30 s were placed on an exclusion list for three minutes.

3.6. Protein database spectral alignment and stringency conditions

The resulting spectra were searched against the 2009/05/07 release of the *C. thermocellum* genome using SEQUEST from BioWorks 3.3 (Thermo Electron). The genome is available at the National center for Biotechnology (NCBI) website courtesy of the U.S. Department of Energy (DOE), JGI (http://www.ncbi.nlm.nih.gov, Version NC_009012.1). The database open reading frames were further annotated by including signal peptide prediction using SignalP 3.0 [67-68], and prediction of the number of transmembrane helix domains using TMHMM 2.0 [69], both available at (http://www.cbs.dtu.dk/). The database was digested *in silico* with trypsin, cleaving at the carboxyl side of lysine and arginine residues. A limit of 3 post translational modifications (PTM) were allowed for peptide matching, and allowable PTMs were the carboxyamidomethylation of cysteine residues (monoisotopic δ mass of 57,0050 Da), and oxidation of methionine residues (to sulfoxide, δ mass of 15,99490). A peptide ion tolerance of ± 2 atomic mass units and a fragment ion tolerance of ± 0.7 atomic mass units were implemented. Charge state analysis was performed during database screening, and stringent filters were applied to define the criteria for successful peptide identification. Singly, doubly, and triply charged peptide ions required SEQUEST cross correlation (XC) scores of at least 1.9, 2.5, and 3.0, respectively. Peptide ions further required a δ CN value of >0.1 for successful assignment. The above stringency conditions satisfied a false positive identification rate of less than 1% based on spectral alignment of *C. thermocellum* samples against an *in silico* reversed *C. thermocellum* database.

3.7. RelEx analysis

For the analysis of quantitative experiments, DTASelect [70] was used to filter the DTA spectrum files, which matched peptides to the protein database using the same stringency conditions as above. After DTASelect screening, the DTA files could then be examined by RelEx [65]. The software produces extracted ion chromatograms of peptide isotope pairs as they are observed during the survey scans over the course of the analytical experiment. The area under the curve of each pair is used to quantify the peptides signal, allowing the calculation of a peptide signal ratio of sample to isotope labelled reference. For a given extracted ion chromatogram pair to be accepted it required a signal to noise ratio above 3 and a regression filter (a measure of the curve overlap) above 0.7 at 1 and 0.4 at 10.

Despite the filters applied to a peptide's sample to reference ratio, the volume of data requires the systematic removal of outlier ratio values that clearly should not be applied to the calculation of a sample to reference ratio of a protein quantified in either the Avicel or cellobiose sample. This is accomplished by considering only peptide ratios within 1.5 times the interquartile range (IQR) below the lower quartile (Q1) and 1.5 times IQR above the upper quartile (Q3) of the population of peptide ratios corresponding to a particular protein in a particular sample.

As multiple peptides can be matched to a protein, and a given peptide’s sample to reference ratio can be observed more than once over the entire experiment, protein ratios of sample protein to reference protein within a sample were calculated as the geometric mean of all satisfactory peptides ratios matched to them.

The geometric mean (GM) of sample to reference ratios (r) of n peptides (a_1, a_2, ... a_n) for a particular protein (A) can be defined by

$$GM = \sqrt[n]{\prod_{i=1}^{n} a_i} = \sqrt{a_1 a_2 \cdots a_n} = 10^{\left[\frac{1}{n}\sum_{i=1}^{n}\log a_i\right]} = r_A$$

Eq. 3

The geometric mean can be thought of as the antilog of the arithmetic mean of the logs of the peptide ratios contributing to the geometric average. It is appropriate to calculate the geometric mean for a given protein's ratio of sample to reference, as opposed to the more traditional arithmetic mean, in order to give each peptide ratio equal weight [71].

The ratio of each unlabelled Avicel-grown protein over ^{15}N-labled Avicel-grown protein was divided by the ratio of the corresponding unlabelled cellobiose-grown protein over ^{15}N Avicel-grown protein. The quotient of these ratios is the ratio of unlabelled Avicel-grown protein over cellobiose-grown protein and thus the relative abundance of that protein across the two biological samples.

$$R_A = \frac{r_{A,Avicel}}{r_{A,cellobiose}}$$

Eq. 4

Where R_A is the overall ratio of the abundance of protein A on Avicel relative to cellobiose.

Variation within the experiment is measured using the % Coefficient of Variation (%CV) of r_A , given by

Eq. 5

$$\%CV_{Sample/reference} = 100 \cdot (10^{SD_{sample/reference}} - 1)$$

Where $SD_{sample/reference}$ represents the standard deviation of the log transformed peptide ratios ($log\ a_1$, $log\ a_2$,...$log\ a_i$).

The cumulative standard coefficient of variation in the measurement of the final ratio R_A is given by

Eq. 6

$$\%SCV_{Avicel\,/\,cellobiose} = 100 \cdot \sqrt{\left(\%CV_{Avicel\,/\,reference}/100\right)^2 + \left(\%CV_{cellobiose/reference}/100\right)^2}$$

The variation in measurement of the relative abundance of Protein A on Avicel relative to cellobiose can then be described as

Eq. 7

$$R_A \times/\div \left(\%SCV_{Avicel\,/\,cellobiose}/100 + 1\right)$$

The two-tailed Student's t-test was used to determine the probability that the difference in the ratios calculated for growth on Avicel and for growth on cellobiose was due to real differences or could be explained instead by variation in the data. The t-distribution value was calculated as

Eq. 8

$$\left| \frac{r_{A,Avicel} - r_{A,cellobiose}}{\sqrt{\dfrac{\left(\dfrac{1}{N_{A,Avicel}} + \dfrac{1}{N_{A,cellobiose}}\right)\left[\left(N_{A,Avicel} - 1\right)\cdot SD^2_{A,Avicel} + \left(N_{A,cellobiose} - 1\right)\cdot SD^2_{A,cellobiose}\right]}{df}}} \right|$$

Where N is the number of peptides used for the calculation of the ratio r, and df is the degrees of freedom, described by

Eq. 9

$$df = N_{A,Avicel} + N_{A,cellobiose} - 2$$

4. Results

4.1. Identification of soluble and membrane fraction proteins of *C. thermocellum*

In order to maximize the number of proteins discovered from cells grown on Avicel and cellobiose, samples were separately analysed without the addition of the ^{15}N reference sample necessary for quantitative comparison of protein abundance. Because proteomic samples containing ^{15}N reference peptides have twice the number of unique mass/charge species to detect, but still the same amount of unique peptide sequences, we would expect to identify a greater fraction of the proteome in samples without labelled reference. As can be seen in Table 4, there are indeed a greater number of unique peptide and protein identifications found in samples without reference cells.

In total, 864 unique proteins were identified in the ^{15}N reference-free soluble fraction; 162 discovered uniquely on Avicel, 206 uniquely on cellobiose, and 496

Table 4. Unique proteomic identifications from samples grown on alternative carbon sources.

Soluble fraction	**Total proteins**	**Unique peptides**	**Membrane fraction**	**Total proteins**	**Unique peptides**
Avicel	658	1654	Avicel	639	1692
Cellobiose	702	1714	Cellobiose	681	1804
Avicel with ^{15}N reference*	563	1383	Avicel with ^{15}N reference*	600	1517
Cellobiose with ^{15}N reference*	533	1350	Cellobiose with ^{15}N reference*	605	1612

* These samples were mixed ^{15}N metabolically labelled Avicel grown cells to enable subsequent quantitative comparison.

common to both samples. Within the ^{15}N reference-free membrane fraction 890 unique proteins were discovered; 209 discovered uniquely on Avicel, 251 uniquely on cellobiose, and 430 common to both samples. Lists of identified proteins from the membrane fraction are included in appendix E and F, representing cells grown on Avicel and cellobiose respectively. The enrichment of membrane proteins in the membrane protein fractions was investigated by determining the amount of integral membrane proteins that were identified. Of the 890 proteins discovered in that fraction, 257 (29%) were identified that possessed at least 1 trans-membrane helix. Of the 864 proteins identified in the soluble protein fraction only 154 (18%) proteins were identified possessing at least one trans-membrane helix.

A total of 1255 unique proteins were identified in this study from the non-redundant sum of all experiments. *Clostridium thermocellum* putatively contains 3238 open reading frames based on the published genome sequence, thus over 38% of the total proteome was observed in the present investigation. It must be stressed that the lack of identification of a particular protein from a sample does not alone justify a conclusion that it is not present. In order to examine the functional distribution of all identified proteins on the two substrates, the 1255 proteins were manually annotated based upon sequence alignment to 1 of 224,847 orthologous groups of proteins present in 630 species using version 2 of the eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups) database [72].

Figure 5 illustrates the distribution of proteins identified in the soluble fraction of both Avicel and cellobiose grown cells. While there were 368 unique protein identifications between the two samples in the soluble fraction, overall the distribution of

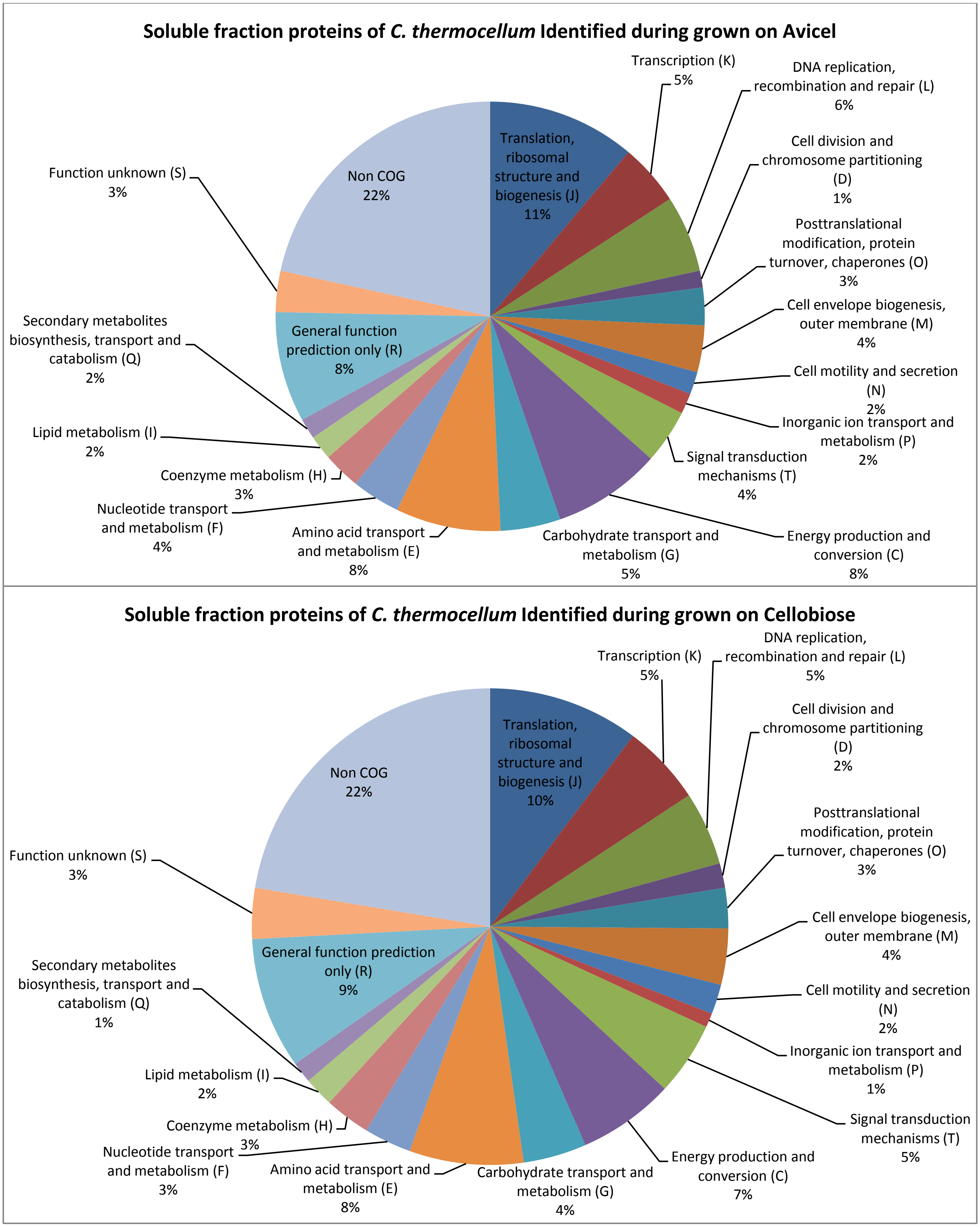


Figure 5. Functional distribution of the soluble fraction proteome of *C. thermocellum* identified during growth on Avicel and cellobiose as annotated using eggNOG v2.0. Functional annotations and their respective one letter codes are labelled on the pie charts. The non-COG group represent proteins to which a cluster of orthologous genes could not be assigned.

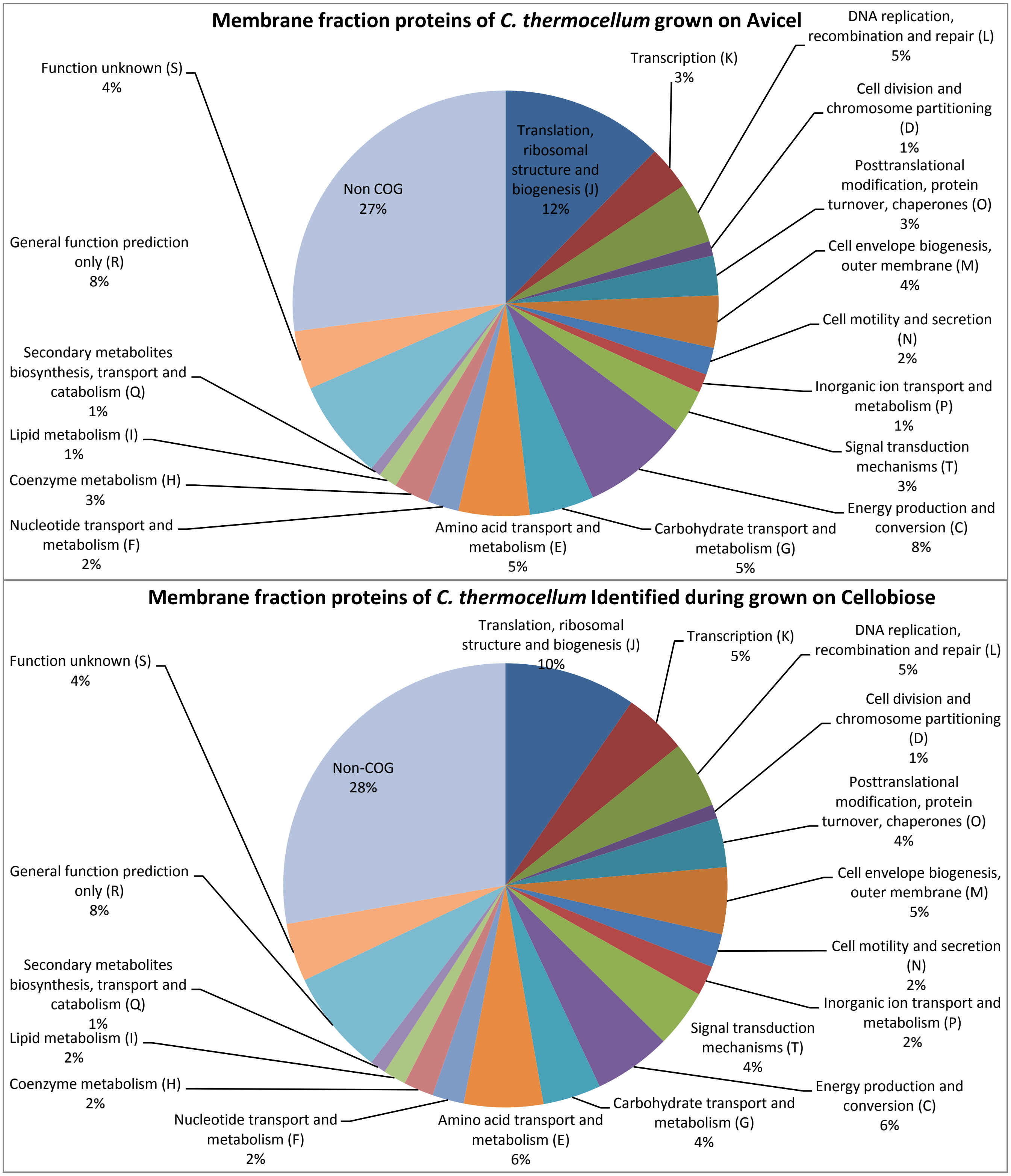


Figure 6. Functional distribution of the membrane fraction proteome of *C. thermocellum* identified during growth on Avicel and cellobiose as annotated using eggNOG v2.0. Functional annotations and their respective one letter codes are labelled on the pie charts. The non-COG group represent proteins to which a cluster of orthologous genes could not be assigned.

protein functions of the Avicel and cellobiose grown samples are very similar. The same trend was true of the Avicel and cellobiose grown cells from the membrane fraction (figure 6). A comparison of figures 5 and 6 show clearly that the membrane fraction had a greater proportion of proteins that were unable to be assigned to a COG (cluster of orthologous genes).

4.2. Substrate-induced changes to the relative abundance of proteins in *C. thermocellum*

Quantitative differences in protein expression of *C. thermocellum* grown on cellobiose and Avicel were measured based upon the equal incorporation of ^{15}N peptides of a *C. thermocellum* reference sample grown on Avicel. The calculation of the relevant ^{14}N to ^{15}N ratios of peptide XIC's by Relex provides a reliable method to compare the abundance of particular peptides, and their proteins, in the two biological samples. Because of the additional stringency imposed on the data for the positive assignment of a peptide with its co-eluting ^{15}N partner, and the requirement that at least two peptide-ratio measurements per sample are made to allow a quantitative evaluation of a particular protein’s relative abundance, far fewer proteins are quantified than are identified. Proteins were considered to have increased expression if they were at least 1.4 fold more abundant in one sample compared to another. This threshold was arbitrarily chosen so as to not include proteins with small changes in expression that may be statistically significant but not biologically relevant.

4.2.1. Substrate-induced changes to the relative abundance of soluble-fraction proteins

Soluble-fraction proteins were denatured and separated by SDS-PAGE (Figure 7). Proteins from gel bands excised from the gel were trypsin digested and extracted for analysis. A total of 72 non-ribosomal proteins found in cells grown on both substrates were quantified. Table 5 shows the 22 proteins found to possess at least 1.4 fold greater expression on Avicel. In Table 6 are found the 8 proteins found to posses 1.4 fold or greater expression on cellobiose. Of these 31 proteins identified to have an increase in expression on a particular substrate, all but 2 had p-values ≥ 0.5, as determined by the two-tailed Student's t-test, thus rejecting the null hypothesis. The remaining 39 proteins quantified as being within these alternative ranges are regarded as not having important differences in expression and are listed in Table 7. Ribosomal proteins quantified in the study did not show any meaningful change in expression under the conditions tested.

The cellulosomal scaffold protein CipA (gene Cthe_3077) demonstrated the greatest relative change in expression of all the proteins quantified under the conditions tested. The more than 8-fold greater abundance of CipA when grown on Avicel compared to cellobiose is consistent with previous measurements of batch cultures using an enzyme-linked immunosorbent assay (ELISA) with an antibody raised against a segment of CipA [73-74]. An examination of purified non-anchored cellulosome complexes from Avicel and cellobiose grown *C. thermocellum* cells showed that CipA was 5 fold more abundant in samples grown on Avicel [59].

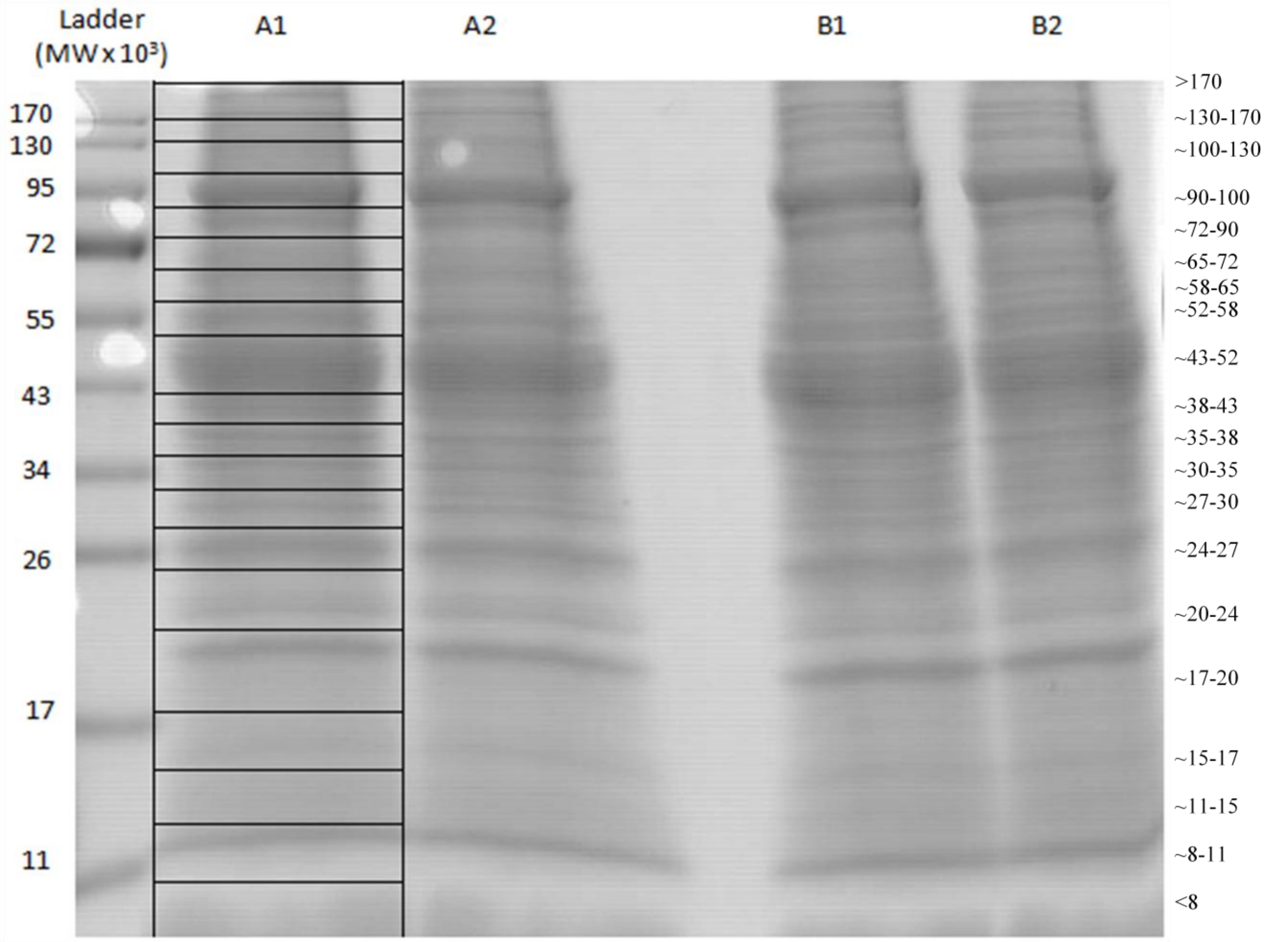


Figure 7. Soluble protein fractions from *C. thermocellum* batch cultures grown to stationary phase separated by SDS-PAGE (10-20% Tris-HCl), stained with Bio-Safe Coomassie. Lane A1-2, 1:1 mixture of unlabelled Avicel-grown cells and ^{15}N labelled Avicel grown cells, 60 µg total protein per lane. Lane B1-2, 1:1 mixture of unlabelled cellobiose-grown cells with ^{15}N labelled Avicel grown cells 60 µg total protein per lane. Molecular weight markers are labelled on the left of the figure in kD. The pattern of gel excision used for all lanes is marked for lane A1. Approximate MW sizes of excised gel bands is reported on the right of the figure.

As might be expected, cellodextrin phosphorylase (Cthe_2989) was induced in cells grown on Avicel. Cellodextrin phosphorylase has been cloned, sequenced and purified from *C. thermocellum* [44, 75], confirming the gene Cthe_2989 as the sequence responsible for its activity. Growth of *C. thermocellum* on cellulose would result in the internalization of cellodextrins larger than cellobiose into the cell, and thus an increased requirement for phosphorylase activity to feed glucose-1-phosphate and glucose to the Embden-Meyerhof pathway. Enzymes of the Embden-Meyerhof pathway did not demonstrate differences in expression during growth on Avicel and cellobiose, with the exception of pyruvate kinase. The identification and quantification of proteins in the Embden-Meyerhof pathway is summarized in Table 8. Phosphofructokinase (Cthe_0347), Glyceraldehyde-3-phosphate dehydrogenase (Cthe_0137), Phosphoglycerate kinase (Cthe_0138), and Enolase (Cthe_0143), were all successfully quantified as not having significant change in expression. The remaining enzymes of the pathway were all identified in this study; Glucokinase (Cthe_2938), Phosphoglucomutase (Cthe_1265), Glucose-6-phosphate isomerase (Cthe_0217), Fructose-1,6-bisphosphate aldolase (Cthe_0349), Triose phosphate isomerise (Cthe_0139). Six candidate sequences in the *C. thermocellum* genome may encode Phosphoglycerate mutase activity (Cthe_0140, Cthe_0707, Cthe_0946, Cthe_1292, Cthe_1435, Cthe_2449); Cthe_0140 and Cthe_0946 were identified in this study.

Table 5. Proteins from the soluble fraction demonstrating increased expression during growth of *C. thermocellum* on Avicel relative to cellobiose as determined by RelEx.

		^{14}N Avicel/^{15}N Avicel			^{14}N cellobiose/^{15}N Avicel			^{14}N Avicel/^{14}N cellobiose		
Gene Number	Protein name or type	Ratio 1	%CV	no. pep.	Ratio 2	%CV	no. pep.	Ratio1/Ratio2	%SCV	*p*-value
Cthe_1368	S-layer-like domain containing protein	1.359	16.220	24	0.564	40.037	4	**2.410**	43.197	0.0121
Cthe_1965	Alkyl hydroperoxide reductase	0.953	34.354	41	0.407	23.901	16	**2.343**	41.851	<0.0001
Cthe_0904	Protein-export membrane protein SecD	1.621	12.665	3	0.749	16.939	5	**2.163**	21.151	<0.0001
Cthe_0681	Inosine 5-monophosphate dehydrogenase	1.027	21.317	12	0.505	34.943	6	**2.034**	40.932	0.0001
Cthe_1100	Prepilin-type cleavage/methylation	1.584	4.425	2	0.844	15.849	2	**1.876**	16.455	0.0435
Cthe_3202	CRISPR-associated protein, Csh2 family	1.501	23.590	28	0.940	29.868	15	**1.597**	38.060	<0.0001
Cthe_0388	Alcohol dehydrogenase GroES-like protein	2.395	47.950	48	1.525	16.161	5	**1.571**	50.600	<0.0001
	Energy production and conversion, carbon metabolism									
Cthe_3077	Cellulosome scaffold protein CipA	1.629	19.025	18	0.193	36.818	3	**8.437**	41.443	0.0005
Cthe_3120	Pyruvate:ferredoxin oxidoreductase	1.403	50.717	7	0.665	15.166	3	**2.110**	52.936	0.0003
Cthe_2989	Cellodextrin Phosphorylase	0.650	22.792	25	0.316	8.588	8	**2.053**	24.357	<0.0001
Cthe_2874	Phosphoenolpyruvate carboxykinase (GTP)	1.718	24.157	40	1.064	28.291	20	**1.614**	37.202	<0.0001
Cthe_1425	Membrane-bound proton-translocating pyrophosphatase	2.252	42.544	2	1.480	20.090	3	**1.522**	47.049	0.0448
Cthe_2392	Pyruvate:ferredoxin oxidoreductase, alpha subunit	1.276	16.485	26	0.893	17.199	31	**1.428**	23.823	<0.0001
Cthe_2393	Pyruvate:ferredoxin oxidoreductase, beta subunit	1.457	54.562	12	0.956	22.832	15	**1.524**	49.147	0.0075
Cthe_2390	Pyruvate:ferredoxin oxidoreductase, gamma subunit	1.290	9.968	14	0.900	17.175	14	**1.433**	19.858	0.0381
	Nitrogen metabolism									
Cthe_0374	Glutamate dehydrogenase (NADP+)	1.094	21.545	57	0.357	12.804	20	**3.066**	25.062	<0.0001
Cthe_3226	Copper amine oxidase-like protein	1.630	15.751	14	1.079	9.411	4	**1.511**	18.348	0.0001
	Unknown function									
Cthe_3230	Hypothetical protein	2.107	22.139	11	0.891	32.925	3	**2.364**	39.676	0.0230
Cthe_0076	Hypothetical protein	1.455	16.052	24	0.786	46.773	13	**1.852**	48.999	0.0003
Cthe_1970	Hypothetical protein	1.006	15.178	4	0.670	24.391	3	**1.502**	28.728	0.0612
Cthe_2709	Hypothetical protein	1.663	34.575	5	1.133	17.575	7	**1.468**	38.786	0.0413
Cthe_2383	Hypothetical protein	1.900	19.726	14	1.351	31.150	9	**1.406**	36.871	0.0057

%CV, Coefficient of variation; %SCV, standard coefficient of variation; *p*-value, probability that the null hypothesis is true, based on 2-tailed Student t-test of ratio 1 versus ratio 2.

Table 6. Proteins from the soluble fraction demonstrating decreased expression during growth of *C. thermocellum* on Avicel relative to cellobiose as determined by RelEx.

		^{14}N Avicel/^{15}N Avicel			^{14}N cellobiose/^{15}N Avicel			^{14}N Avicel/^{14}N cellobiose		
Gene Number	Protein name or type	Ratio 1	%CV	no. pep.	Ratio 2	%CV	no. pep.	Ratio1/Ratio2	%SCV	*p*-value
Cthe_1104	Prepilin-type cleavage/methylation	1.246	16.510	12	1.940	9.859	7	**0.642**	19.230	<0.0001
	Energy production and conversion, carbon metabolism									
Cthe_0423	Bifunctional acetaldehyde-CoA/alcohol dehydrogenase	0.992	31.372	250	1.652	42.994	258	**0.600**	53.223	<0.0001
Cthe_1308	Pyruvate kinase	1.449	61.873	30	2.283	46.058	30	**0.635**	67.133	0.0002
Cthe_0701	Oxaloacetate decarboxylase (ATP)	0.708	23.004	4	1.473	51.885	7	**0.512**	56.756	0.0038
	Cellodextrin transport									
Cthe_0393	CbpA (Cellodextrin transport complex A)	1.958	36.190	41	3.614	23.831	41	**0.542**	43.331	<0.0001
Cthe_0392	MsdA (Cellodextrin transport complex A)	2.447	132.779	2	5.712	50.827	11	**0.218**	142.174	0.0262
	Unknown function									
Cthe_0820	Hypothetical protein	0.847	16.681	10	1.287	31.543	17	**0.658**	35.682	<0.0001
Cthe_2090	Hypothetical protein	0.816	61.397	12	1.554	32.853	9	**0.525**	69.634	0.0012

%CV, Coefficient of variation; %SCV, standard coefficient of variation; *p*-value, probability that the null hypothesis is true, based on 2-tailed Student t-test of ratio1 versus ratio2

Table 7. List of quantified proteins from the soluble fraction demonstrating no change in expression on either Avicel or cellobiose (^{14}N Avicel/^{14}N cellobiose ratio between 1.4 to 0.71).

Gene Number	Protein name or type	^{14}N Avicel/^{14}N cellobiose	SCV%
Cthe_0004	YD repeat containing protein	1.349	37.592
Cthe_1504	Linocin_M18 bacteriocin protein	1.348	27.244
Cthe_0445	Cell division protein FtsZ	1.310	158.322
Cthe_0856	Branched-chain amino acid aminotransferase	1.291	18.472
Cthe_1912	Copper amine oxidase-like protein	1.185	43.637
Cthe_2518	Ketol-acid reductoisomerase	1.148	41.678
Cthe_0700	Biotin/lipoyl attachment protein	1.147	15.304
Cthe_0699	Carboxyl transferase	1.120	38.745
Cthe_0424	Aminoglycoside phosphotransferase	1.037	48.325
Cthe_2657	Histone-like DNA-binding protein	1.059	19.751
Cthe_0395	RbsD or FucU transport	0.972	15.592
Cthe_1005	Translation elongation factor Ts	0.964	31.233
Cthe_2892	Chaperonin GroEL	0.908	32.394
Cthe_1105	Type II secretion system protein	0.888	92.997
Cthe_2724	DNA-directed RNA polymerase, beta subunit	0.883	38.055
Cthe_0819	ABC transporter related protein	0.803	30.635
Cthe_0418	Polynucleotide phosphorylase/polyadenylase	0.786	50.197
Cthe_0714	S1 RNA-binding domain protein	0.783	41.545
	Energy production and conversion, carbohydrate metabolism		
Cthe_0143	Enolase	1.229	28.027
Cthe_0342	Hydrogenase, Fe-only	1.176	50.112
Cthe_0347	Phosphofructokinase	1.149	43.056
Cthe_0137	Glyceraldehyde-3-phosphate dehydrogenase	1.081	39.440
Cthe_0341	NADH dehydrogenase quinone	1.029	93.760
Cthe_0344	Malic Enzyme	0.989	47.158
Cthe_0285	Isocitrate dehydrogenase, NADP-dependent	0.950	18.648
Cthe_2267	V-type ATP synthase subunit A	0.869	54.713
Cthe_0138	Phosphoglycerate kinase	0.831	50.413
Cthe_1862	ABC transporter related protein	0.798	35.138
Cthe_1862	ABC transporter related protein	0.798	35.138
Cthe_0357	Alpha-glucan phosphorylases	0.794	32.370
Cthe_2608	F0F1 ATP synthase subunit beta	0.772	25.724
	Cellodextrin transport		
Cthe_1020	CbpB (Cellodextrin transport complex B)	1.005	29.621
Cthe_1019	MsdB1 (Cellodextrin transport complex B)	0.864	69.781
	Unknown function		
Cthe_0858	Protein of unknown function DUF1432	1.371	32.202
Cthe_3096	Hypothetical protein	1.347	32.502
Cthe_2422	Hypothetical protein	1.312	12.353
Cthe_1192	Hypothetical protein	1.157	29.909
Cthe_0836	Hypothetical protein	1.165	12.291
Cthe_2130	Hypothetical protein	0.960	30.399
Cthe_1098	Hypothetical protein	0.908	28.645

%SCV, standard coefficient of variation.

Interestingly, pyruvate kinase (PK, gene Cthe_1308) was shown to have reduced expression when grown on Avicel, while a GTP producing phosphoenolpyruvate carboxykinase (PPCK, gene Cthe_2874) has increased expression under the same conditions. As illustrated in Figure 8, this would redirect carbon flow from the Embden-Meyerhof pathway into a transhydrogenase malate cycle through the subsequent reactions of malate dehydrogenase (Cthe_0345) and malic enzyme (Cthe_0344), directing carbon back to pyruvate. Carbon flow through this cycle has been previously hypothesized [76] and would increase the ratio of NADPH/NADH production in the cell. The relative abundance of malic enzyme was successfully quantified in this study, and did not change under the conditions tested. Malate dehydrogenase was not quantified, but was identified in both Avicel and cellobiose grown cells. Malate dehydrogenase and malic enzyme activities have been observed previously in *C. thermocellum* [76], and the malic enzyme has been purified and characterized as ammonium activated [77]. Gene numbers Cthe_0344 and Cthe_0345 are both annotated by conceptual translation as "malate dehydrogenase" in the current version of the JGI *C. thermocellum* genome (NC_009012.1). However, the product of Cthe_0344 shares 78% amino acid identity with the malic enzyme of *Clostridium cellulovorans* (E-value = 6e-164) and possess a conserved malic enzyme domain (cd05311, E-value = 7e-95), suggesting that gene Cthe_0344 encodes the observed malic enzyme activity. The putative role of genes Cthe_0345 and Cthe_0344 as malate dehydrogenase and malic enzyme, respectively, is consistent with their proximity in an apparent malate cycle operon.

Table 8. Summary of the enzymes of the Embden-Meyerhof pathway detected in the soluble protein fraction of *C. thermocellum*.

Enzyme	Gene number	Identification	Quantification
Phosphoglucomutase	Cthe_1265	A, C	NQ
Glucokinase	Cthe_2938	A, C	NQ
Phosphoglucose isomerase	Cthe_0217	A, C	NQ
Phosphofructokinase	Cthe_0347	A, C	-
Frucotose bisphosphate aldolase	Cthe_0349	A, C	NQ
Triose phosphate isomerase	Cthe_0139	A, C	NQ
Glyceraldehyde phospate dehydrogenase	Cthe_0137	A, C	-
Phospoglycerate kinase	Cthe_0138	A, C	-
Phospoglycerate mutase	Cthe_0140, Cthe_0946	A, C A, C	NQ NQ
Enoalse	Cthe_0143	A, C	-
Pyruvate kinase	Cthe_1308	A, C	↑ C (1.563)

A, Avicel grown cells; C, Cellobiose grown cells; NQ, not quantified; -, quantified as having no significant change in expression.

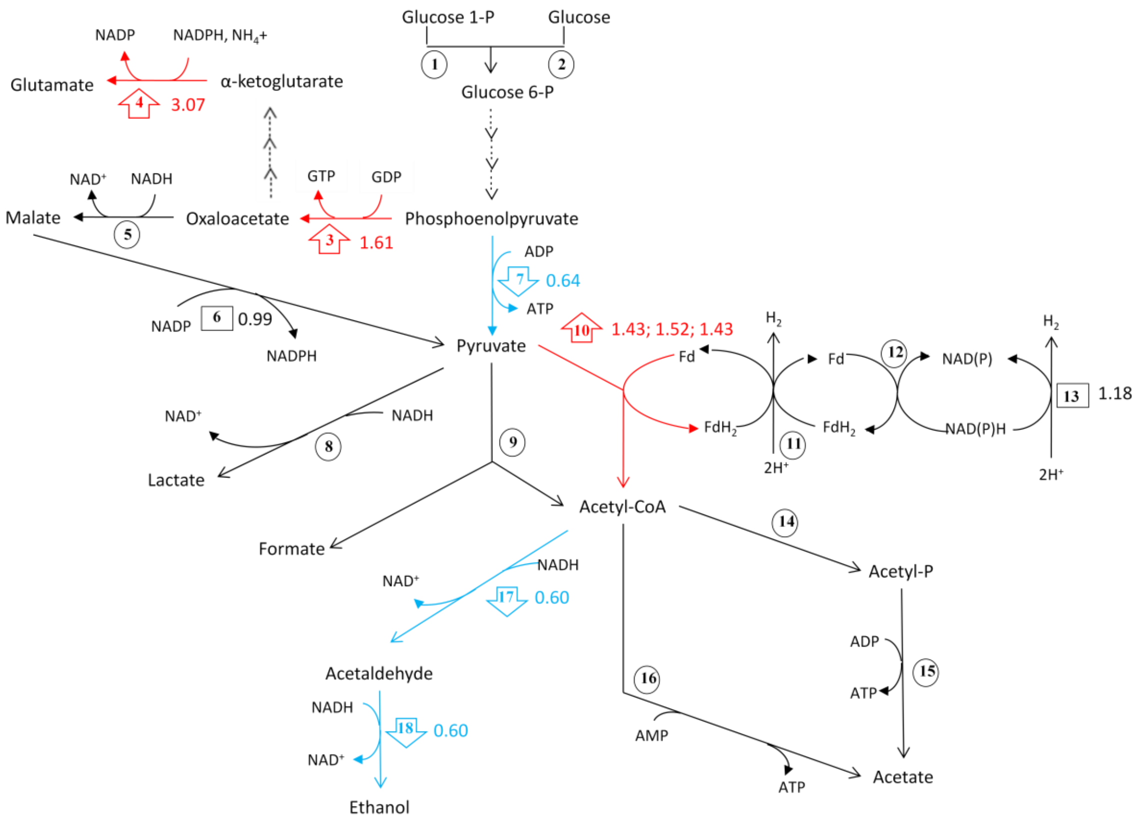


Figure 8. Conversion of glucose and glucose-1-P into fermentation products by *C. thermocellum* during growth on Avicel. Enzymes are labelled by numbers. Arrow shaped borders surrounding the number indicates if the expression of the enzyme increased (highlighted in red) or decreased (highlighted in blue) during growth on Avicel in the soluble protein fraction, with the ratio of protein in Avicel grown cells relative to cellobiose grown cells at right. Square borders indicate the enzyme was quantified as having no change in expression, circles borders indicate the enzyme was not quantified, but proteins with the putative activity were identified by MS.

1, phosphoglucomutase, (EC 5.4.2.2); **2**, glucokinase, (EC 2.7.1.2); **3**, phosphoenolpyruvate carboxykinase, (EC 4.1.1.32); **4**, Glutamate dehydrogenase, (E.C 1.4.1.3); **5**, Malate dehydrogenase (EC 1.1.1.37); **6**, Malic enzyme, (EC 1.1.1.40); **7**, pyruvate kinase, (EC 2.7.1.40); **8**, Lactate dehydrogenase, (EC 1.1.1.27); **9**, pyruvate formate lyase, (EC 2.3.1.54); **10**, pyruvate:ferredoxin oxidoreducatase (subunits alpha, beta, and gamma respectively), (EC 1.2.7.1); **11**, ferredoxin hydrogenase (EC1.12.7.2); **12**, ferredoxin:NAD(P)$^+$ oxidoreductase (EC 1.18.1.2, EC 1.18.1.3); **13**, hydrogen:NAD(P)$^+$ oxidoreductase (EC 1.12.1.2, EC 1.12.1.3); **14**, phosphotransacetylase (EC 2.3.1.8); **15**, acetate kinase, (EC 2.7.2.1); **16**, acetate thiokinase (EC 6.2.1.1); **17**, acetaldehyde dehydrogenase (EC 1.2.1.10); **18**, alcohol dehydrogenase (EC 1.1.1.1); Fd, oxidized ferredoxin; FDH$_2$, reduced ferredoxin.

The increased carbon flux through the described malate cycle shunt during growth on Avicel is further supported by the increased expression of inosine 5-monophosphate dehydrogenase (gene Cthe_0681) during growth on this substrate. Inosine 5-monophosphate dehydrogenase activity is the rate-limiting step in the *de novo* synthesis of the guanine nucleotides [78], and the described malate shunt would be increasing the relative contribution of these nucleotides in conserving energy from substrate level phosphorylation in the cell by the action of PPCK (Cthe_2874).

This malate cycle is likely affected by another 465 amino acid product of gene Cthe_0701, oxaloacetate decarboxylase, which has a decreased expression on Avicel as reported in Table 6. The putative activity of this protein, however, is elusive. The protein encoded by gene number Cthe_0701 is a member of the phylogenetic protein cluster PRK12331 which includes 33 homologous sequences across 8 genres. The putative activity of enzymes in this cluster is that of another PPCK similar to the already mentioned PPCK enzyme Cthe_2874, but instead produces ATP as opposed to GTP (E.C number 4.1.1.49). The existence of two PPCK activities that are oppositely regulated and produce alternative nucleotide triphosphates would suggest an important node in the carbon flux of the cell and alternative demands on the cellular ATP and GTP metabolite pools during growth on the two carbon sources tested. Alternatively, another likely putative function of the gene product of Cthe_0701 is instead the decarboxylation of oxaloacetate by conversion to pyruvate (E.C. number 4.1.1.32) also based upon domain sequence homology. This activity would allow the malate cycle to redirect to pyruvate without the transhydrogenation of electrons from NADH to NADPH, restricting the production of biosynthetic cofactors during growth on cellobiose. The two putative functions of Cthe_0701 are illustrated in Figure 9.

Glutamate dehydrogenase (Cthe_0374) was found to have more than 3-fold increased expression during growth on Avicel suggesting that nitrogen assimilation is induced during cellulose catabolism as seen in Figure 8. The metabolite oxaloacetate present in the described malate cycle represents an important biosynthetic precursor for several molecules in the cell including the production of α-ketoglutarate, through which the nitrogen source ammonia is assimilated into the cell's biomass by the action of glutamate dehydrogenase,

Eq. 10 $\alpha\text{-ketoglutarate} + \mathrm{NH_4} + \mathrm{H^+} + \mathrm{NADPH} \Leftrightarrow \mathrm{Glutamate} + \mathrm{NADP^+} + \mathrm{H_2O}$

α-ketoglutarate is formed from oxaloacetate by the subsequent reactions preformed by citrate synthase (Cthe_3027), aconitase (Cthe_3158), and isocitrate dehydrogenase (Cthe_0285). The latter three enzymes were successfully identified under both growth conditions and isocitrate dehydrogenase was quantified in this study and its relative abundance was unchanged across the conditions tested.

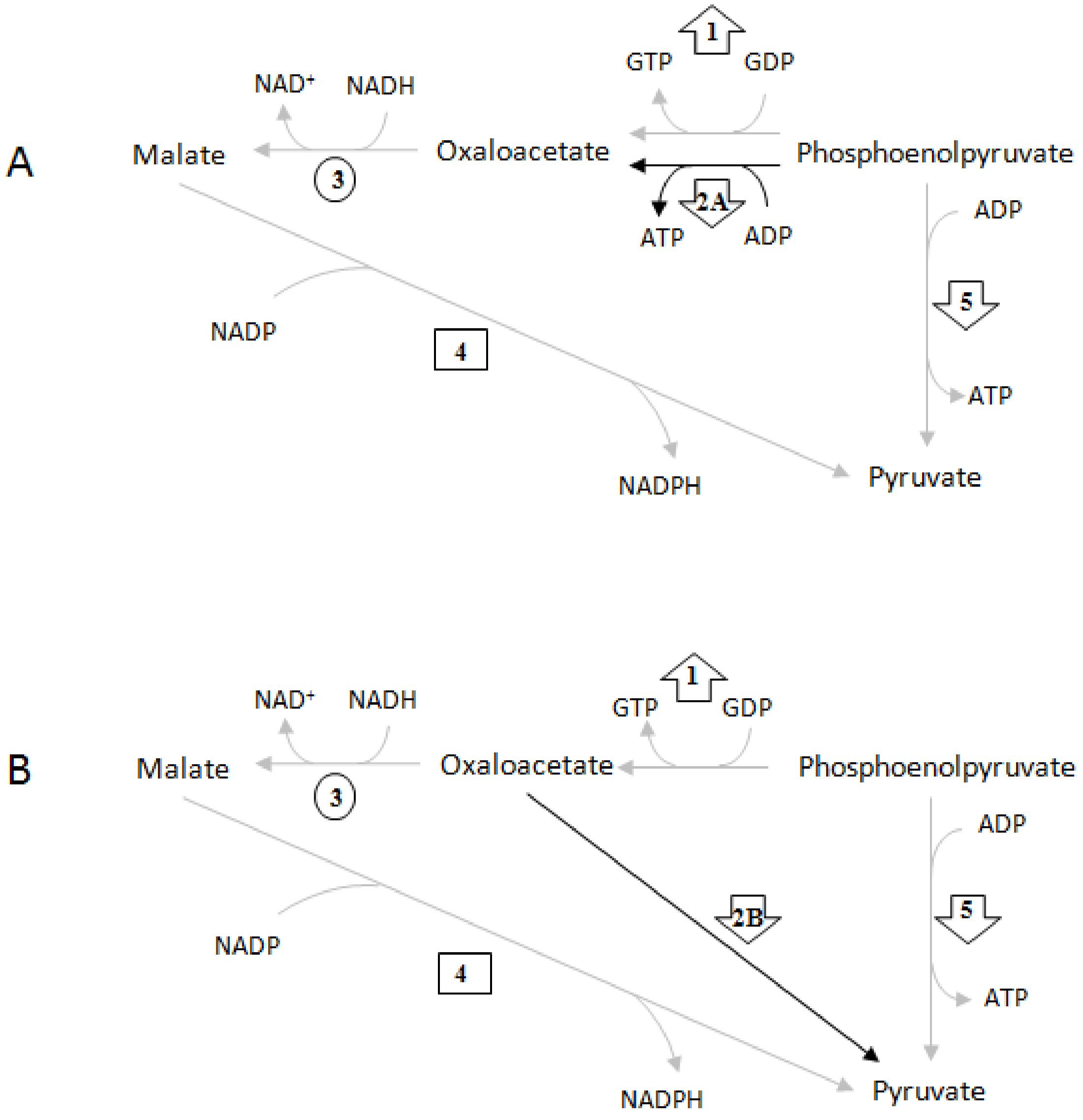


Figure 9. Two putative enzymatic functions of the gene product of Cthe_0701 which demonstrates decreased expression on Avicel. **A**, Cthe_0701 with putative ATP producing PPCK activity; B, Cthe_0701 with putative oxaloacetate decarboxylase activity. Enzymes are labelled by numbers. Arrow shaped borders surrounding the number indicates if the expression of the enzyme increased or decreased during growth on Avicel in the soluble protein fraction, square borders indicate the enzyme was quantified as having no change in expression, circles borders indicate the enzyme was not quantified, but proteins with the putative activity were identified by MS.

1, phosphoenolpyruvate carboxykinase (GTP), (EC 4.1.1.32); **2A**, phosphoenolpyruvate carboxykinase (GTP), (E.C number 4.1.1.49); **2B** Oxaloacetate decarboxylase, (E.C. number 4.1.1.32); **3**, Malate dehydrogenase (EC 1.1.1.37); **4**, Malic enzyme, (EC 1.1.1.40); **5**, pyruvate kinase, (EC 2.7.1.40).

As shown in Table 5, three of the proteins successfully quantified are the protein products of genes Cthe_2390, Cthe_2392 and Cthe_2393 which all demonstrated increased expression on Avicel. These proteins represent members of a putative pyruvate:ferredoxin oxidoreductase (POR) complex; the γ, α, and β subunits respectively. POR is an important activity in pyruvate metabolism and catalyzes the following reaction;

Eq. 11

$$\text{pyruvate} + \text{CoA} + 2\,\text{oxidized ferredoxin} \Leftrightarrow \text{acetyl-CoA} + \text{CO}_2 + 2\,\text{reduced ferredoxin} + 2\,\text{H}^+$$

The genomic context of Cthe_2390, Cthe_2392 and Cthe_2393 suggests a bacterial POR operon which includes Cthe_2391, the δ subunit of the complex. This δ subunit Cthe_2391 was not quantified, nor was it even identified in any of the fractions. This can, however, simply be due to a lack of "discoverability" of the protein, in part because of its small size (101 amino acids, aa) compared to the rest of the proteins present in the operon; γ (189 aa), α (394 aa), and β (311 aa). In addition to this multi-subunit putative POR, a single large putative POR protein (1178 aa, gene Cthe_3120) was found to be more abundant in the Avicel grown culture. This single subunit POR contains the domains found on the multi-subunit POR including a 4Fe-4S center, a thiamine pyrophosphate binding domain, and catalytic or regulatory domain and is similar to a single subunit POR purified from *C. acetobutilicum* [79-80]. A summary of the proteins identified and quantified from the soluble fraction involved in pyruvate metabolism is presented in Table 9.

An 873 amino acid product of gene number Cthe_0423, a putative bifunctional acetaldehyde-CoA/alcohol dehydrogenase, was more abundant in cellobiose grown cells than Avicel grown cells. This gene product has been proposed to be the key enzyme involved in ethanol synthesis in *C. thermocellum* [79], allowing the NADH-mediated reduction of acetyl-CoA, the product of POR, to acetaldehyde, followed by subsequent NADH-mediated reduction to ethanol.

Two proteins possessing prepilin-type cleavage/methylation motifs were quantified in this study. Cthe_1100 was found to have increased expression on Avicel grown cells, while Cthe_1104 had increased expression on cellobiose grown cells as seen in Table 5 and 6 respectively. Prepilin-type cleavage/methylation motifs direct N-terminal cleavage and methylation at a conserved phenylalanine residue. This motif is most often found at the N-terminus of pilins and other proteins involved in secretion [81]. SecD (Cthe_0904), a component of the Sec protein-translocation pathway, was also found to have increased expression on Avicel.

Four hypothetical proteins were quantified as having increased expression on Avicel and *p*-values below 0.05; Cthe_0076, Cthe_2383, Cthe_2709, Cthe_3230. Two hypothetical proteins demonstrated increased expression on cellobiose; Cthe_0820, and Cthe_2090.

Table 9. Summary of identified and quantified enzymes involved in pyruvate metabolism found in the soluble protein fraction of *C. thermocellum*

Enzyme	Gene number	Identification	Quantification
Lactate dehydrogenase	Cthe_1053	A, C	NQ
Pyruvate formate lyase	Cthe_0505	A, C	NQ
Pyruvate formate lyase – Activating enzyme	Cthe_0506	ND	NQ
Pyruvate:ferredoxin oxidoreductase	Cthe_2390-2393	A, C *	↑A *(1.433; ND;1.428; 1.524)
	Cthe_2794-2797	ND	NQ
	Cthe_3120	A, C	↑A (2.110)
Fe-only hydrogenase (NADH)	Cthe_0342	A, C	-
	Cthe_0430	A, C	NQ
Fe-only hydrogenase (NADPH)	Cthe_3003	A, C	NQ
Bifunctional acetaldehyde-CoA/alcohol dehydrogenase	Cthe_0423	A, C	↑ C (1.667)
Iron containing alcohol dehydrogenase	Cthe_0101	ND	NQ
	Cthe_0394	ND	NQ
	Cthe_2579	ND	NQ
Acetate kinase	Cthe_1028	A, C	NQ
Phosphotransacetylase	Cthe_1029	A, C	NQ
Acetate Thiokinase	Cthe_0551	A, C	NQ

* The δ subunit, Cthe_2391, of the Pyruvate:ferredoxin oxidoreductase complex was not identified or quantified. A, Avicel grown cells; C, Cellobiose grown cells. ND, not detected; NQ, not quantified.; -, quantified as having no significant change in expression.

4.2.2. Substrate-induced changes to the relative abundance of membrane-fraction proteins

Membrane-fraction proteins were denatured and separated by SDS-PAGE (figure 10). Data reported here is the result of proteins excised from the gel, digested with trypsin, and extracted for MS analysis. A total of 47 non-ribosomal proteins found in cells grown on both substrates were quantified, representing the sum of proteins shown in Tables 10, 11, and 12. Table 10 shows the 13 proteins found to possess at least 1.4 fold greater expression on Avicel. In Table 11 are found the 10 proteins found to posses 1.4 fold greater expression on cellobiose. Of the 31 proteins identified to have an increase in expression on a particular substrate, all but 2 had p-values ≥ 0.5, as determined by the two-tailed Student’s t-test, thus rejecting the null hypothesis. The remaining 23 proteins quantified as being within these alternative ranges are regarded as not having important differences in expression and are listed in Table 12.

Three cellulosomal anchoring proteins were found to be more abundant in the membranes of Avicel grown cells (Table 10); OlpB (Cthe_3078), Orf2p (Cthe_3079), Sdba (Cthe_1307). Sdba is a 65 kD MW protein with 2 SLH (S-layer homology domains) and a single type II cohesion domain available to bind the type II dockerin domain of the cellulosome. The abundance of this protein in the membrane fraction of Avicel cells was 2 fold greater than in cellobiose grown cells. OlpB and Orf2p demonstrated more than 5-fold increase in expression on Avicel grown cells. Orf2p is a 72 kD MW protein possessing 3 SLH domains and two type II cohesion domains potentially for binding

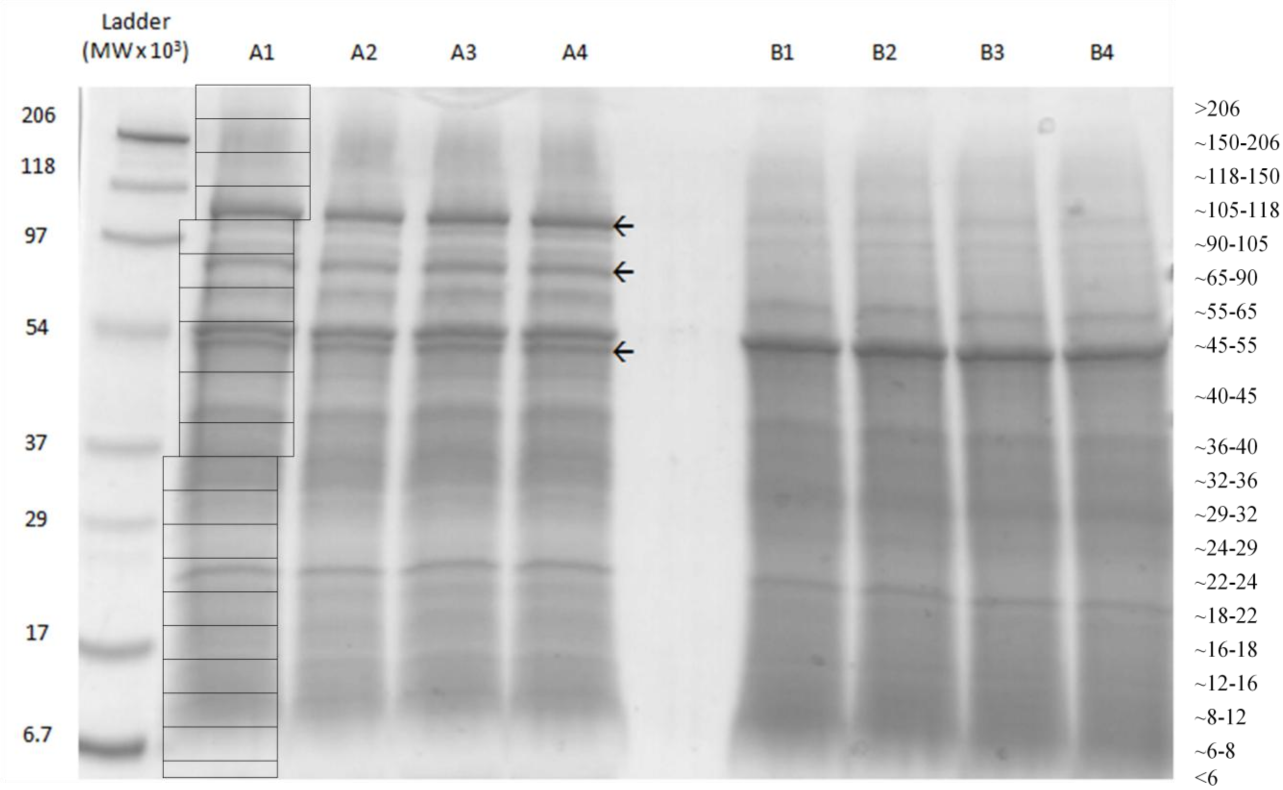


Figure 10. Membrane protein fractions from *C. thermocellum* batch cultures grown to stationary phase separated by SDS-PAGE (10-20% Tris-HCl), stained with Bio-Safe Coomassie. Lane A1-4, 1:1 mixture of unlabelled Avicel-grown cells and ^{15}N labelled Avicel grown cells, 30 µg total protein per lane. Lane B1-4, 1:1 mixture of unlabelled cellobiose-grown cells with ^{15}N labelled Avicel grown cells 30 µg total protein per lane. Molecular weight markers are ladled on the left of the figure. Three black arrows indicate protein bands with visible changes in expression.

Table 10. Proteins from the Membrane fraction demonstrating increased expression during growth of *C. thermocellum* on Avicel relative to cellobiose as determined by RelEx.

		^{14}N Avicel/^{15}N Avicel			^{14}N cellobiose/^{15}N Avicel			^{14}N Avicel/^{14}N Cellobiose		
Gene Number	Protein name or type	Ratio 1	%CV	no. pep.	Ratio 2	%CV	no. pep.	Ratio1/Ratio2	%SCV	*p*-value
Cthe_1102	Fimbrial assembly protein	3.761	33.08	23	1.668	33.35	15	**2.255**	46.97	<0.0001
Cthe_2892	Chaperonin GroEL	2.602	31.00	8	1.112	85.79	3	**2.336**	91.22	0.1327
Cthe_2236	Flagellin-like protein	7.470	61.11	15	3.595	13.50	3	**2.076**	62.58	0.0002
Cthe_2408	Phage shock protein A, PspA	2.861	24.01	44	1.410	38.86	15	**2.032**	45.68	<0.0001
Cthe_0904	Protein-export membrane protein, SecD	3.724	28.21	24	1.903	24.41	22	**1.956**	37.30	<0.0001
Cthe_1368	S-layer-like domain containing protein	2.413	33.83	55	1.257	55.73	12	**1.920**	65.19	0.0304
	Cellulosome anchoring proteins									
Cthe_3078	Cellulosome anchoring protein, OlpB	2.902	19.61	19	0.513	15.47	10	**5.655**	24.98	<0.0001
Cthe_3079	Cellulosome anchoring protein, Orf2P	2.304	8.87	4	0.554	23.14	4	**4.160**	24.78	0.0002
Cthe_1307	Cellulosome anchoring protein, SdbA	2.851	31.79	14	1.430	27.58	3	**1.990**	42.08	0.0197
	Nitrogen Metabolism									
Cthe_1827	Copper amine oxidase-like protein	3.050	2.47	5	0.580	12.60	4	**5.260**	12.84	<0.0001
Cthe_3227	Copper amine oxidase-like protein	4.003	1.70	3	1.849	12.32	10	**2.161**	12.44	<0.0001
Cthe_3226	Copper amine oxidase-like protein	2.911	22.22	19	1.768	33.72	17	**1.645**	40.38	<0.0001
	Unknown function									
Cthe_2383	Hypothetical protein	4.472	29.20	15	3.003	34.21	16	**1.489**	44.97	0.0004

%CV, Coefficient of variation; %SCV, standard coefficient of variation; *p*-value, probability that the null hypothesis is true, based on 2-tailed Student t-test of ratio 1 versus ratio 2.

Table 11. Proteins from the Membrane fraction demonstrating decreased expression during growth of *C. thermocellum* on Avicel relative to cellobiose as determined by RelEx.

		^{14}N Avicel/^{15}N Avicel			^{14}N cellobiose/^{15}N Avicel			^{14}N Avicel/^{14}N Cellobiose		
Gene Number	Protein name or type	Ratio 1	%CV	no. pep.	Ratio 2	%CV	no. pep.	Ratio1/Ratio2	%SCV	*p*-value
Cthe_2932	DNA-directed RNA polymerase	1.391	23.45	3	4.419	49.58	12	**0.315**	54.85	0.0004
Cthe_2998	ABC transporter related protein	1.615	21.17	2	3.763	15.91	4	**0.429**	26.48	0.0488
Cthe_0016	Ferritin and Dps	3.167	26.30	5	7.344	17.90	2	**0.431**	31.81	0.0150
Cthe_1504	Linocin_M18 bacteriocin protein	1.016	157.13	2	2.111	51.54	21	**0.481**	165.37	0.4688
Cthe_0701	Oxaloacetate decarboxylase (ATP)	1.797	28.41	2	2.666	61.65	3	**0.674**	67.88	0.0389
Cthe_1363	lipopolysaccharide biosynthesis	3.038	12.49	6	4.495	59.99	9	**0.676**	61.27	0.0392
	Cellodextrin transport									
Cthe_0393	CbpA (Cellodextrin transport complex A)	2.169	35.87	26	3.714	43.98	32	**0.584**	56.75	<0.0001
Cthe_0392	MsdA (Cellodextrin transport complex A)	1.971	60.24	5	2.985	48.62	13	**0.660**	77.41	0.0062
	Unknown function									
Cthe_3096	Hypothetical protein	1.990	39.55	3	6.021	30.25	20	**0.330**	49.79	0.0209
Cthe_1431	Hypothetical protein	1.838	2.89	2	4.369	15.76	2	**0.421**	16.03	0.0668

%CV, Coefficient of variation; %SCV, standard coefficient of variation; *p*-value, probability that the null hypothesis is true, based on 2-tailed Student t-test of ratio 1 versus ratio 2.

Table 12. List of quantified proteins from the Membrane fraction demonstrating no change in expression on either Avicel or cellobiose (^{14}N Avicel/^{14}N cellobiose ratio between 1.4 to 0.71).

Gene Number	Protein name or type	^{14}N Avicel/ ^{14}N cellobiose	%SCV
Cthe_0424	Aminoglycoside phosphotransferase	1.303	56.38
Cthe_2518	Ketol-acid reductoisomerase	1.272	60.07
Cthe_2730	Translation elongation factor Tu	0.785	86.16
Cthe_1251	Xanthine/uracil/vitamin C permease	0.741	35.12
	Lipid transport and metabolism		
Cthe_0937	Fatty acid/phospholipid synthesis protein PlsX	1.329	42.98
Cthe_0699	Carboxyl transferase	0.751	40.99
	Energy production and conversion, carbohydrate metabolism		
Cthe_2604	ATP synthase F0, B subunit	1.366	38.86
Cthe_0341	NADH dehydrogenase quinone	1.127	67.72
Cthe_0423	Bifunctional acetaldehyde-CoA/alcohol dehydrogenase	1.028	68.89
Cthe_0137	Glyceraldehyde-3-phosphate dehydrogenase, type I	0.876	61.56
Cthe_0347	Phosphofructokinase	0.869	32.62
Cthe_1862	ABC transporter related protein	0.868	32.52
Cthe_0143	Phosphopyruvate hydratase	0.758	59.98
	Cellodextrin transport		
Cthe_1020	CbpB (Cellodextrin transport complex B)	0.922	50.35
Cthe_1018	MsdB2 (Cellodextrin transport complex B)	0.846	51.70
Cthe_1019	MsdB1 (Cellodextrin transport complex B)	0.777	36.69
	Unknown function		
Cthe_1098	Hypothetical protein	1.503	52.68
Cthe_2130	Hypothetical protein	1.334	40.38
Cthe_1101	Hypothetical protein	1.356	40.41
Cthe_0637	Protein of unknown function DUF909	0.767	53.53
Cthe_2357	Hypothetical protein	1.161	70.50
Cthe_2422	Hypothetical protein	0.790	52.85

%SCV, standard coefficient of variation.

more than 1 cellulosome. OlpB, the largest of the anchoring proteins with an estimated MW of 245 kD, contains 3 SLH domains in addition to 7 type II cohesion domains, potentially allowing for a large network of cellulosomes to bind at the cell surface via this anchoring unit. A visible increase in the intensity of three protein bands of Avicel grown cells, compared to cellobiose grown cells, can be seen in the SDS-PAGE gel of the membrane fraction (Figure 10). The molecular weights of these protein bands are estimated to be ~110 kD, ~85 kD, and 54kD based upon the relative migration of the protein ladder. These three bands likely correspond to the three anchoring proteins OlpB, Orf2p, and Sdba respectively, based upon the quantified relative expression observed for these proteins using the isotope labelled ESI-MS/MS method and the fact that peptides for each of these proteins were observed in the appropriate gel bands. Linking the theoretical MW of a protein based upon its aa sequence with the observed electrophoretic migration in SDS-PAGE is complicated by several factors, including post-translational modifications of the protein such as glycosylation and proteolysis. The migration of OlpB corresponding to 110 kD, instead of the estimated 245 kD size of the protein, is likely the result of a proteolytic cleavage. This cleavage may be involved in the observed release of cellulosomes from the cell surface during stationary phase. Also found to have increased expression on Avicel was the S-layer protein Cthe_1368. An increase in S-layer synthesis would support the increased number of SLH domain-containing anchoring proteins present during growth on Avicel.

Three putative copper-amine oxidase proteins were demonstrated to be more abundant on Avicel than cellobiose grown cells; Cthe_3227, Cthe_3226, Cthe_1827.

Copper-amine oxidases are a family of oxidoreductases that liberate ammonia from primary amines in the following reaction (E.C. 1.4.3.21),

Eq.12

$$RCH_2NH_2 + H_2O + O_2 \Leftrightarrow RCHO + NH_3 + H_2O_2$$

Table 13 lists all the putative Copper-amine oxidases from the *C. thermocellum* genome, including the 15 that were identified in this study. The prediction of signal peptides and a transmembrane helix in the majority of these sequences suggest a role in acquiring ammonia from plant cell wall proteins and other sources of primary amines in the extracellular environment [82-83]. This process would also scavenge oxygen for the strict anaerobe and produce hydrogen peroxide outside of the cell, a reactive chemical species that has been suggested to aid in the depolymerisation of cellulose by cellulolytic microbes including brown-rot fungi [84-85]. The induction of mechanisms to acquire nitrogen from the extracellular environment during growth on Avicel would support the active and costly biosynthetic demand for nitrogen from the increased synthesis of cellulosomes on this substrate.

Proteins from two of the four putative ABC cellodextrin transporters of *C. thermocellum* were observed and quantified in the membrane fraction. All three of the gene products of the operon encoding cellodextrin transporter B (MsdB1, MsdB2, and CbpB; Figure 2), were quantified in this study and none demonstrated any significant change in expression under either growth condition (Table 11). CbpB and MsdB2 were also quantified in the soluble fraction, also with no change in relative expression (Table 7). Two of the proteins in the complex, MsdB1 and MsdB2, each possess a membrane spanning domain with six transmembrane helices and together form the transmembrane channel for the transporter. CbpB is the cellodextin binding protein for the complex. The

Table 13. Summary MS identification and quantification of putative copper-amine oxidases from the *C. thermocellum* genome.

Gene number	#TMH	SP prediction	Identification	Quantification
Cthe_1827	1	Y	A, C	↑**A** (5.250)
Cthe_3226	1	Y	A, C	↑**A** (2.161)
Cthe_3227	1	Y	A, C	↑**A** (1.645)
Cthe_0045	1	Y	A, C	NQ
Cthe_0047	1	Y	A, C	NQ
Cthe_0048	1	Y	A, C	NQ
Cthe_0243	1	Y	A, C	NQ
Cthe_0402	1	Y	A, C	NQ
Cthe_0824	1	Y	A, C	NQ
Cthe_1374	0	N	A, C	NQ
Cthe_1833	1	Y	A, C	NQ
Cthe_1912	0	Y	A, C	-
Cthe_3228	1	Y	A, C	NQ
Cthe_1909	1	Y	C	NQ
Cthe_2109	1	Y	C	NQ
Cthe_0593	0	N	ND	NQ
Cthe_0744	1	Y	ND	NQ
Cthe_0799	1	Y	ND	NQ
Cthe_1258	1	Y	ND	NQ
Cthe_1778	1	N	ND	NQ
Cthe_1908	1	Y	ND	NQ
Cthe_2146	0	N	ND	NQ
Cthe_2424	1	Y	ND	NQ
Cthe_3188	0	N	ND	NQ

A, Avicel grown cells; C, Cellobiose grown cells. ND, not detected; NQ, not quantified.; -, quantified as having no significant change in expression.

substrate specificity for this protein has been previously determined *in vitro* to be for cellodextrins of lengths n=2 (cellobiose) to n=5 (cellopentose), where n represents the number of glucose units in the oligosaccharide. The three proteins CbpB, MsdB1, and MsdB2 do not together posses all the domains necessary for the proper functioning of an ABC transporter. The cellodextrin transporter complex B also requires a nucleotide binding domain containing polypeptide that must be encoded elsewhere in the genome. Cthe_1862, an estimated 42 kD MW protein possessing an ABC transport related nucleotide binding domain (cd03301, E value 8e-106), is not encoded in an operon alongside other ABC transporter genes and was quantified in this study as having a similar expression pattern to the proteins CbpB, MsdB1 and MsdB2 (Table 7). Cthe_1862 could thus be a strong candidate for the protein carrying the missing nucleotide binding domain of the cellodextrin transporter complex B.

Two protein members of another putative transporter, the cellodextrin transport complex A, were also identified and quantified in the membrane fraction. Both CbpA and MsdA demonstrated increased expression on cellobiose in the membrane fraction (Table 11). Increased expression of these proteins was observed on cellobiose in the soluble fraction as well (Table 6), though the quantification of one of these components, MsdA, in the soluble fraction is invalidated by the high p-value of 0.226. The specificity of the cellodextrin binding protein of this complex, CbpA, has been previously determined, *in vitro,* to be only for cellotroise [41].

Other clues to metabolic changes occurring during growth on Avicel found in the membrane fraction include the observation of individual proteins putatively involved in fimbrial adhesion (fimbrial assembly protein, Cthe_1102), motility (flagellin-like protein, Cthe_2236), protein folding (chaperonin GroEL, Cthe_2892) and protein secretion (protein-export membrane protein SecD, Cthe_0904). The phage shock protein, PspA, (Cthe_2408) was also found to have increased abundance on Avicel. The *Escherichia coli* PspA homolog is synthesized in response to several stresses including heat, ethanol, and osmotic shock [86]. PspA of *E. coli* and has been shown to bind membrane phospholipids and repair proton leakage of membranes damaged by ethanol [87] and be involved in protein export [88]. Protein sequence alignment of *C. thermocellum* and *E. coli* PspA homologs by BLASTp analysis however revealed only a 35% amino acid sequence identity covering only 20 residues with an E score of 6.2, making this functional assignment of the PspA protein in *C. thermocellum* based upon its annotation in the NCBI database dubious at best. A protein with unknown function, Cthe_2383, was found to have increased abundance on Avicel.

5. Discussion

The identification of 1255 proteins from the non-redundant sum of all investigations presented in this thesis represent the broadest reported enquiry into the proteome of *C. thermocellum*. Many proteins were identified uniquely on one of the two substrates, however, this does not provide sufficient evidence that these same proteins are uniquely expressed on the substrate in which they were identified. The functional distribution of the proteins identified in the soluble and membrane fractions grown on both Avicel and cellobiose was investigated by comparing their sequences to the eggNOG v2.0 database [72], which possesses clusters of orthologous genes from a total of 630 organisms (529 bacteria, 46 archaea and 55 eukaryotes). The assignment of functional categories by the identification of orthologous relationships form the basis of most comparative genomic studies [89]. It was hoped this examination might reveal distinct global patterns in the functional distribution of proteins grown on either substrate, however the distribution appears to be highly similar despite the presence of several hundred protein identifications uniquely identified on a particular substrate. It was noted however that a higher proportion of proteins from the membrane fraction, compared to the soluble fraction, were not assigned to a cluster of orthologous genes from the eggNOG database. This would be expected as membrane proteins, as a group, are more poorly characterized than soluble proteins [90]. While the tabulation of the 1255 proteins identified in the described experiments represents a rich supply of information of the biology of *C. thermocellum*, the most valuable clues regarding changes in metabolism of C. *thermocellum* grown on the two substrates is obtained from evidence discovered from quantitative experiments. The identification of proteins under the two growth conditions is used to support this data.

Quantitative comparison of the *C. thermocellum* proteome during growth on Avicel and cellobiose revealed exciting insights into mechanisms controlling the production of biosynthetic intermediates and directing the flow of carbon within the cell. The proposed model of the most important observed results of this study is summarized in figure 11. It appears during growth on Avicel that carbon flux is diverted away from pyruvate kinase and directed towards a transhydrogenase malate cycle via PPCK. The result of this rerouting would shift more electrons from the NADH to the NADPH pools, in addition to a greater ratio of GTP/ATP production from substrate level phosproylation. An increase in NADPH production during growth on Avicel would feed the assimilation of ammonia into cell material via the action of glutamate dehydrogenase (Eq.10) which also had increased expression in Avicel grown cells. The use of the transhydrogenase malate cycle for the purpose of ammonia assimilation is further supported by the fact that the malic enzyme activity in *C. thermocellum* is ammonium activated [77], a property not characteristic of previously described malic enzymes. Increased operation of this cycle in response to growth on Avicel serves the increased nitrogen demand from the biosynthesis of amino acids for the induced production of high MW cellulosome anchors (OlpB, Orf2P, SdbA), scaffolds (CipA), and cellulase proteins on cellulosic substrates [59, 74, 91-92]. The apparent increase in the metabolic mechanisms for assimilating nitrogen into cell material during growth on Avicel is paralleled by an increase in copper-amine oxidases putatively responsible for increasing the availability of ammonia for the cell. Investigation of inorganic nitrogen metabolism in *C. thermocellum* is limited to a study concluding that the organism possesses glutamate dehydrogenase activity [93]. The ultimate fate of the increased production of GTP during growth on Avicel is not entirely clear. An increase in protein translation would present a greater demand for substrates of that process, which include GTP. Ribosomal proteins quantified in this study did not see any changes in expression under the two conditions. GTP may prove to be a substrate for glycosyl transferase enzymes responsible for protein glycosylation in this organism. Glycosylation of *C. thermocellum* cellulase and cellulosomal proteins has been reported and is thought to aid to protect the cellulosome from proteolysis [94-95]. Protein glycosylation in bacteria is poorly understood compared to eukaryotes, and is recently

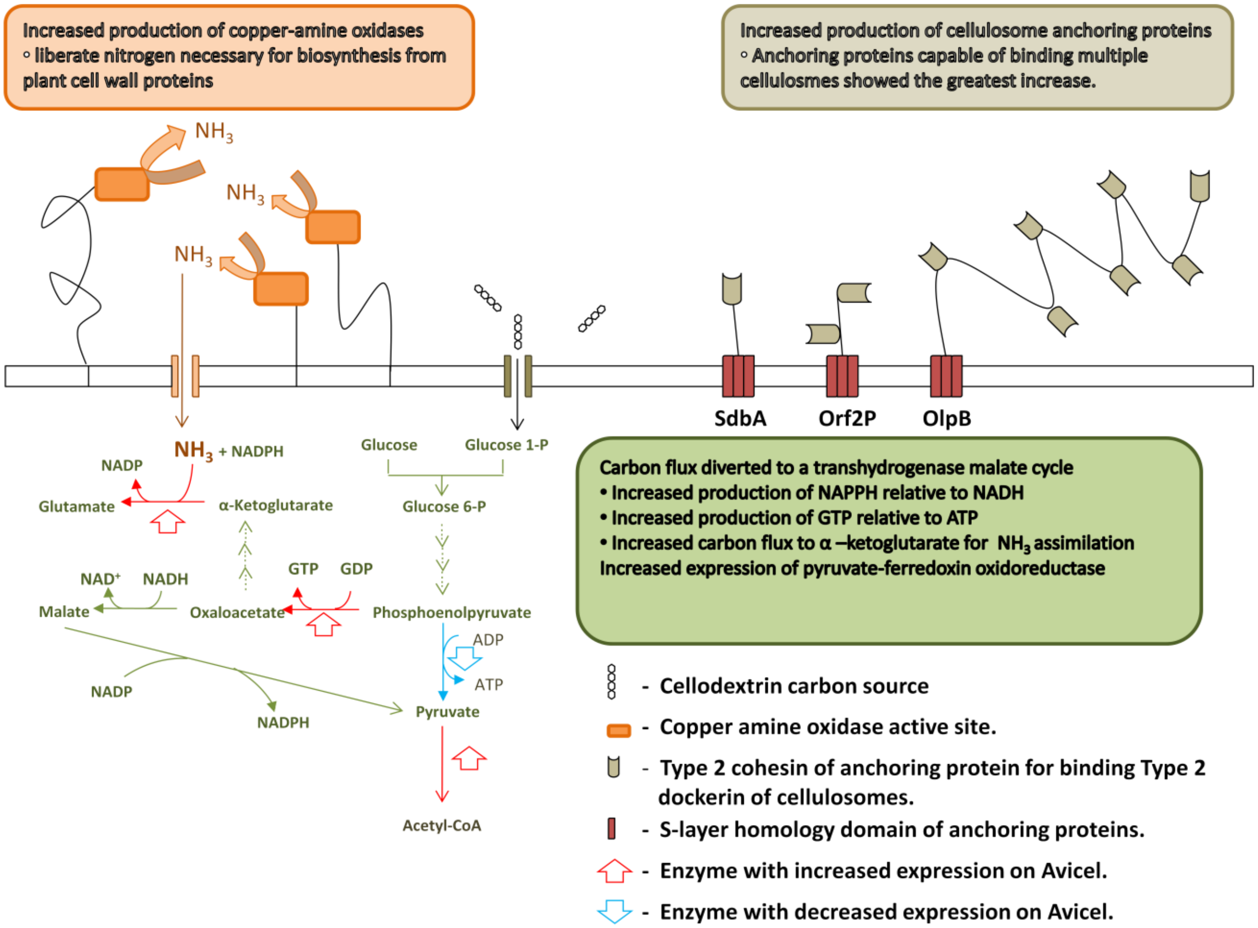


Figure 11. Proposed model of metabolic changes observed in *C. thermocellum* grown on Avicel.

being modeled based on studies in *Campylobacter jejuni,* the first bacterium in which a complete glycosylation pathway has been described [96-97]. While the known glycosylation pathway from this organism does not involve nucleotide sugar biosynthetic pathways that include GTP as a substrate, the more closely related organism *Aneurinibacillus thermoaerophilus* does produce GDP-α-D-rhamnose and GDP-D-mannose as substrates for the glycosylation of S-layer proteins [98-99]**.** Identification of the protein glycosylation pathways involved in *C. thermocellum* including characterization of the nucleotide-activated linkage sugars involved may shed light on the fate of the increased GTP production expected during growth on cellulose. Future investigations of the transhydrogenase malate cycle in *C. thermocellum* should seek to validate its increased operation during growth on cellulose by examining the enzyme activities of PPCK and PK during growth on both Avicel and cellobiose. Furthermore, characterizing the activity of Cthe_0701 by cloning, heterologous expression and His-tag purification would resolve how this enzyme contributes to the metabolism of oxaloacetate and the malate cycle in *C. thermocellum,* after which reduced activity of this enzyme on cellobiose could also be validated by an enzyme assay of *C. thermocellum* lysates grown on both substrates.

Carbon flow that is not directed towards α-ketoglutarate for biosynthesis through the malate cycle can instead be directed to pyruvate by the malic enzyme. Pyruvate is a key intermediate during fermentation as it is a metabolic branch point for controlling cellular requirements of cofactors and energy in the cell. Growth on Avicel did demonstrate a slight increase in abundance of two distinct pyruvate:ferrodoxin oxidoreductase systems; a single subunit POR (Cthe_3120) and a multi subunit POR complex (Cthe_2390-2393). Both of these POR enzymes have been shown to be transcribed by *C. thermocellum* in previous literature by reverse transcriptase polymerase chain reaction (RT-PCR) experiments of the organism grown on cellulose [79]. A third POR complex (Cthe_2794-2797) was also shown to be transcribed in the same investigation, but none of its components were identified by MS in this study. The authors performing the RT-PCR experiments demonstrated transcription of this missing POR enzyme up to 49 hours of growth on 1.1g/L cellulose. Perhaps this POR enzyme complex was not detected in this study because the cells were examined at a later stage of growth (75h). While the relative increase in expression of the two POR complexes quantified in this study were small, ranging from 1.4 to 2 fold increased expression, these results were validated by the Students t-test, and all quantified proteins contributing to POR activity demonstrated the same pattern of increased expression. Fermentation enzymes that compete with POR activity for pyruvate in *C. thermocellum* are lactate dehydrogenase and pyruvate formate lyase. The observed decrease in POR expression on cellobiose does not however result in an increase in formate or lactate synthesis, as similar amounts of lactate and formate are produced in batch cultures of *C. thermocellum* grown on cellobiose and cellulose [100-101]. *C. thermocellum* grows rapidly on cellobiose, demonstrating that POR activity under this condition is capable of handling high carbon flux, in contrast to closely related *C. cellulolyticum* [92]. However, *C. thermocellum,* would benefit from increasing cellular POR activity when cellulose, and not cellobiose, is discovered in the natural habitat as cellobiose, in the absence of cellulose, would be a signal of dwindling carbon supply by which the organism could expect reduced carbon flux through the fermentative pathway. Thus it is conceivable that

POR activity, like total cellulase activity [102], is subjected to some form of carbon catabolite repression mechanism by cellobiose that fine tunes the expression of these enzymes.

A bifunctional acetaldehyde-CoA/alcohol dehydrogenase (Cthe_0423) was also observed that demonstrated a slight increase in expression in the soluble fraction of cellobiose grown cells. While *C. thermocellum* in batch culture produces similar end-product profiles when grown on Avicel or cellobiose, growth on cellobiose does produce a higher ethanol to acetate ratio [101], which may be a result of increased alcohol dehydrogenase activity on cellobiose compared to Avicel. This change in expression of Cthe_0423 in the soluble fraction is however in disagreement with the results for the same protein quantified in the membrane fraction which did not result in a change in expression on either substrate (Table 12). The bifunctional acetaldehyde-CoA/alcohol dehydrogenase does not possess any transmembrane helices, signal peptides for secretion, nor is it expected to be membrane associated, but was none the less one of the most discoverable proteins in all experiments and was found to "contaminate" the membrane fraction. Despite the fact that that this enzyme was not enriched for in the membrane fraction, I can see no reason that there would be any bias of ^{15}N labelled vs. ^{14}N proteins entering the fraction, and thus even quantification of soluble proteins in the membrane fraction should remain valid. Therefore the disagreement of the quantification of the bifunctional acetaldehyde-CoA/alcohol dehydrogenase (Cthe0423) in the membrane and soluble fraction reveals that quantification of this protein in this study was inconclusive. What was determined conclusively, however, is that this protein was the only alcohol dehydrogenase detected of the putative alcohol dehydrogenase enzymes involved in ethanol synthesis as suggested from an earlier bioinformatic analysis of the *C. thermocellum* genome [79], suggesting that it is the major enzyme involved in ethanol synthesis in this organism. Future experiments elucidating the regulation of end-product metabolism in *C. thermocellum* could further test and validate the regulation of the enzymes quantified in this study by using their sequences for the construction of primers used in RT-PCR experiments of RNA taken from cells grown on either cellulose or cellobiose at multiple stages of growth. Further examination of enzymatic activities involved in end-product metabolism in *C. thermocellum,* including acetaldehyde-CoA and alcohol dehydrogenase activities can be performed to determine if any increase in these activities are observed during growth on Avicel and cellobiose.

Two proteins possessing signals for prepilin-type cleavage/methylation motifs, Cthe_1100 and Cthe_1104 were quantified in this study with the former having increased abundance on Avicel, and the latter instead having increased abundance on cellobiose. Further characterization of these proteins may illuminate their function and explain their alternative regulation during the growth conditions tested. SecD (Cthe_0904), a component of the Sec protein-translocation pathway, was also found to have increased expression on Avicel. While it is compelling to speculate that this Sec pathway is induced during growth on Avicel for the purpose of increased translocation of cellulosome scaffolds, anchors, and cellulases, the quantification of only a single component of this translocation apparatus emphasizes that further investigation is required to support such a hypothesis.

Proteomics analysis also revealed that of the putative cellodextrin transporters identified by earlier bioinformatic analysis [41], cellodextrin transport complex A and B appear to be the most important under the conditions tested as they were detected in all proteomic experiments while components of the putative cellodextrin transport complexes C and D were never detected. Quantitative proteomic analysis of these two complexes suggests that the expression of complex B capable of transporting n=2 to n=5 cellodextrins is constitutive, while the expression of Complex A, putatively a strict cellotriose transporter, appears to have increased expression on cellobiose, suggesting that its expression is perhaps controlled by an inducer such as cellobiose or cellotriose, or is growth rate dependent with the faster growing cells on cellobiose building up a greater amount of the transport complex per cell. An increase in the abundance of a complex responsible for cellotriose transport during growth on cellobiose is not immediately clear. The accumulation of longer chain cellodextrins in the media during growth on cellobiose has been demonstrated in *C. thermocellum* [42] and other cellulolytic anaerobes such as *C. cellulolyticum* [92], and *F. succinogenes* [103]. Cellodextrin polymerization in these circumstances is the result of reversible cellodextrin phosphorylase activity. Such accumulation has not been observed in organisms growing on solid substrates [92, 104]. Thus it is possible that the observed increase in abundance of a putative cellotriose transporter is the result of induction due to accumulation of cellotriose in cellobiose grown cultures of *C. thermocellum* at stationary phase. Alternatively, the actual *in vivo* substrate affinity of CbpA, when complexed with other members of the ABC transport channel may include cellobiose in addition to cellotriose, in contrast to its *in vitro* characterization of substrate affinity [41]. If so, cellobiose would be an appropriate inducer of this transport complex, which may explain why an increase in abundance of this transporter is observed when grown on cellobiose.

This study, in identifying that cellodextrin transport complex B (Cthe1018-1020) as the only cellodextrin transporter with broad cellodextrin specificity observed during growth on cellulose, is of great value for future experiments that may seek to introduce cellodextrin transport into non-cellulose fermenting microorganisms to obtain novel second generation biofuel producing industrial strains. Experiments seeking to introduce the cellodextrin transport complex B operon genes into a heterologous host would need to include a compatible NBD protein as it is missing from the operon. Cthe_1862 seems the most likely candidate for the NBD protein participating in the complex based upon its domain composition and similarity in relative expression on Avicel and cellobiose to the quantified components of the cellodextrin transport complex B.

Evidence for induction of gene expression by specific inducers in *C. thermcoellum* include the positive effect of laminaribiose on genes encoding cellulases which are also active on β-1,3-glucan [105-106] and more recently the discovery of pairs of alternative σ and anti-σ factors directing cellulosomal gene expression in response to cellulose, xylan, and pectin via extracellular polysaccharide-sensing components [107]. The present proteomic investigation did not result in any quantitative data on any of the later σ and anti-σ factors, however three such σ factors were identified in cellulose grown cells that have been previously demonstrated to have increased transcription in response to cellulose (σ^{I4}, Cthe_0403; σ^{24C}, Cthe_1470; σ^{I6}, Cthe_2120) [107]. The cognate transmembrane anti-σ factor, RsgI6 (Cthe_2119), responsible for the extracellular sensing of cellulose and xylan [108-109] was also identified in the membrane fraction of cellulose grown cells. This work provides the first experimental evidence for the sigma factors σ^{I4}, σ^{24C}, and σ^{I6} at the protein level in *Clostridium thermocellum*, and provides insight into which of the identified sigma factors putatively involved in sensing extracellular polysaccharide are important in the regulation of cellulosmal genes. As σ factors often positively regulate their own expression, an investigation into the promoters of σ^{I4} and σ^{24C} by mapping of their 5′ untranslated regions using a 5′ RACE (Rapid Amplification of cDNA Ends) technique would reveal transcriptional start sites for these genes from which putative -35 and -10 σ factor binding sites could be established in a similar manner to the analysis preformed already for σ^{I6} and σ^{I1} factors in *C. thermocellum* [107]. Comparison of these sequences with the promoters of proteins found to have altered expression during growth on cellulose in the present study (such as the POR enzymes [Cthe_3120; Cthe_2390-2393], or glutamate dehydrogenase [Cthe_0374]) could reveal the mechanism of transcriptional regulation of these proteins.

Inactivation of the anti-σ factor domains of the extracellular polysaccharide-sensing proteins associated with the sigma factors σ^{I4}, σ^{24C}, and σ^{I6} identified in *C. thermocellum* in this study, would allow permanent activation of these sigma factors and thus continuous expression of the cellulosomal genes they regulate according to the regulatory model proposed by Shoham et. al. [107] Such manipulations could prove to be of great value to the use of *C. thermocellum* in an industrial consolidated bioprocess.

The objective of the quantitative proteomic analysis of *Clostridium thermocellum* presented here was to identify cellulose-specific changes in the expression and ultimate relative abundance of the proteins coordinating the metabolism responsible for cellulose hydrolysis. Identifying whether such observed changes are the result of positive or negative induction by the substrate directly, or instead by the different growth rate *C. thermocellum* exhibits when growing on these substrates is difficult to tease apart in batch culture experiments. The same experiment presented here could be improved if cultures of *C. thermocellum* were grown in chemostats where the dilution rate of the limiting carbon substrate can be controlled such that the growth rate of the organism grown on two substrates can be held constant. This would leave the substrates present in the media as the cause of any change in the relative expression of proteins in the cells which could be examined by a similar proteomic analysis. Such continuous cultures of *C. thermocellum* would help in untangling the effects of substrate and growth rate governing a particular protein's regulation and thus could help in elucidating the ultimate molecular mechanisms responsible for such regulation.

Cellulosic ethanol will become a key alternative energy source as the price of oil based energy products increases. The amount of market share in the energy industry that cellulosic ethanol will eventually control is directly tied to the efficiency of the biomass to ethanol process. The research presented in this thesis increases existing understanding of how the cellulolytic ethanogenic bacterium, *C. thermocellum*, reorganizes its metabolism in response to growth on cellulose. By understanding and eventually controlling the expression of gene products necessary for the catabolism of cellulose by *C. thermocellum*, it will be possible to improve upon cell based transformations of biomass to ethanol.

6. References

1. Eisentraut, A., *Sustainable Production of Second Generation Biofuels*. 2010, International Energy Agency. p. 12.
2. Ralph Sims, M.T., Jack Saddler, Warren Mabee, *From 1st to 2nd Generation Biofuel Technologies*. 2008, International Energy Agency. p. 120.
3. Rockefeller Foundation Bellagio Center, I., *A Sustainable Biofuels Consensus.* 2008.
4. Goldemberg, J., *Ethanol for a sustainable energy future.* Science, 2007. **315**(5813): p. 808-10.
5. Fargione, J., et al., *Land clearing and the biofuel carbon debt.* Science, 2008. **319**(5867): p. 1235-8.
6. Searchinger, T., et al., *Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change.* Science, 2008. **319**(5867): p. 1238-40.
7. Kumar, P., et al., *Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production.* Industrial & Engineering Chemistry Research, 2009. **48**(8): p. 3713-3729.
8. Eggleston, G., *Advances in the Industrial Application of Enzymes on Carbohydrate-Based Materials*, in *Industrial Application of Enzymes on Carbohydrate-Based Material*. 2007, American Chemical Society: Washington, DC. p. 1-16.
9. Wyman, C.E., *What is (and is not) vital to advancing cellulosic ethanol.* Trends Biotechnol, 2007. **25**(4): p. 153-7.
10. Lynd, L.R., et al., *Consolidated bioprocessing of cellulosic biomass: an update.* Curr Opin Biotechnol, 2005. **16**(5): p. 577-83.
11. Hahn-Hägerdal, B., et al., *Bio-ethanol - the fuel of tomorrow from the residues of today.* Trends in Biotechnology, 2006. **24**(12): p. 549-556.

12. J. Wiegel, L.G.L., *The Importance of Thermophilic Bacteria in Biotechnology*. CRC Critical Reviews in Biotechnology. Vol. 3. 1986: CRC Press, Inc.

13. Ng, T.K., T.K. Weimer, and J.G. Zeikus, *Cellulolytic and physiological properties of Clostridium thermocellum.* Archives of Microbiology, 1977. **114**(1): p. 1-7.

14. Bender, J., Y. Vatcharapijarn, and T.W. Jeffries, *Characteristics and Adaptability of Some New Isolates of Clostridium-Thermocellum.* Applied and Environmental Microbiology, 1985. **49**(3): p. 475-477.

15. Ljungdahl, L.G., et al., *A Yellow Affinity Substance Involved in the Cellulolytic System of Clostridium-Thermocellum.* Current Microbiology, 1983. **9**(4): p. 195-199.

16. Wiegel, J. and M. Dykstra, *Clostridium-Thermocellum - Adhesion and Sporulation While Adhered to Cellulose and Hemicellulose.* Applied Microbiology and Biotechnology, 1984. **20**(1): p. 59-65.

17. Stainthorpe, A.C. and R.A.D. Williams, *Isolation and Properties of Clostridium-Thermocellum from Icelandic Hot Springs.* International Journal of Systematic Bacteriology, 1988. **38**(1): p. 119-121.

18. Lynd, L.R., et al., *Microbial cellulose utilization: Fundamentals and biotechnology.* Microbiology and Molecular Biology Reviews, 2002. **66**(3): p. 506-+.

19. Lynd, L.R., H.E. Grethlein, and R.H. Wolkin, *FERMENTATION OF CELLULOSIC SUBSTRATES IN BATCH AND CONTINUOUS CULTURE BY CLOSTRIDIUM-THERMOCELLUM.* Applied and Environmental Microbiology, 1989. **55**(12): p. 3131-3139.

20. Lynd, L.R. and H.E. Grethlein, *Hydrolysis of dilute acid pretreated mixed hardwood and purified microcrystalline cellulose by cell-free broth from Clostridium thermocellum.* Biotechnology and Bioengineering, 1987. **29**(1): p. 92-100.

21. Demain, A.L., M. Newcomb, and J.H.D. Wu, *Cellulase, clostridia, and ethanol.* Microbiology and Molecular Biology Reviews, 2005. **69**(1): p. 124-+.

22. Shaw, A.J., et al., *Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield.* Proc Natl Acad Sci U S A, 2008. **105**(37): p. 13769-74.

23. Heap, J.T., et al., *The ClosTron: Mutagenesis in Clostridium refined and streamlined.* J Microbiol Methods, 2010. **80**(1): p. 49-55.

24. Heap, J.T., et al., *The ClosTron: a universal gene knock-out system for the genus Clostridium.* J Microbiol Methods, 2007. **70**(3): p. 452-64.

25. Tyurin, M.V., S.G. Desai, and L.R. Lynd, *Electrotransformation of Clostridium thermocellum.* Applied and Environmental Microbiology, 2004. **70**(2): p. 883-890.

26. Lamed, R., E. Setter, and E.A. Bayer, *Characterization of a cellulose-binding, cellulase-containing complex in Clostridium thermocellum.* Journal of Bacteriology, 1983. **156**(2): p. 828-36.

27. Bayer, E.A., et al., *The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides.* Annu Rev Microbiol, 2004. **58**: p. 521-54.

28. Beguin, P. and M. Lemaire, *The cellulosome: An exocellular, multiprotein complex specialized in cellulose degradation.* Critical Reviews in Biochemistry and Molecular Biology, 1996. **31**(3): p. 201-236.

29. Shoham, Y., R. Lamed, and E.A. Bayer, *The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides.* Trends in Microbiology, 1999. **7**(7): p. 275-281.

30. Doi, R.H. and A. Kosugi, *Cellulosomes: Plant-cell-wall-degrading enzyme complexes.* Nature Reviews Microbiology, 2004. **2**(7): p. 541-551.

31. Desvaux, M., *The cellulosome of Clostridium cellulolyticum.* Enzyme and Microbial Technology, 2005. **37**(4): p. 373-385.

32. Felix, C.R. and L.G. Ljungdahl, *The Cellulosome - the Exocellular Organelle of Clostridium.* Annual Review of Microbiology, 1993. **47**: p. 791-819.

33. Lu, Y.P., Y.H.P. Zhang, and L.R. Lynd, *Enzyme-microbe synergy during cellulose hydrolysis by Clostridium thermocellum.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(44): p. 16165-16169.

34. Mosier, N.S., et al., *Reaction kinetics, molecular action, and mechanisms of cellulolytic proteins.* Adv Biochem Eng Biotechnol, 1999. **65**: p. 23-40.

35. Philippidis, G.P., T.K. Smith, and C.E. Wyman, *Study of the enzymatic hydrolysis of cellulose for production of fuel ethanol by the simultaneous saccharification and fermentation process.* Biotechnol Bioeng, 1993. **41**(9): p. 846-53.

36. Strobel, H.J., F.C. Caldwell, and K.A. Dawson, *Carbohydrate Transport by the Anaerobic Thermophile Clostridium-Thermocellum Lqri.* Applied and Environmental Microbiology, 1995. **61**(11): p. 4012-4015.

37. Higgins, C.F., *ABC transporters: physiology, structure and mechanism - an overview.* Research in Microbiology, 2001. **152**(3-4): p. 205-210.

38. Nikaido, H. and J.A. Hall, *Overview of bacterial ABC transporters.* Abc Transporters: Biochemical, Cellular, and Molecular Aspects, 1998. **292**: p. 3-20.

39. Davidson, A.L. and P.C. Maloney, *ABC transporters: how small machines do a big job.* Trends in Microbiology, 2007. **15**(10): p. 448-455.

40. Hollenstein, K., R.J.P. Dawson, and K.P. Locher, *Structure and mechanism of ABC transporter proteins.* Current Opinion in Structural Biology, 2007. **17**(4): p. 412-418.

41. Nataf, Y., et al., *Cellodextrin and Laminaribiose ABC Transporters in Clostridium thermocellum.* Journal of Bacteriology, 2009. **191**(1): p. 203-209.

42. Zhang, Y.H.P. and L.R. Lynd, *Cellulose utilization by Clostridium thermocellum: Bioenergetics and hydrolysis product assimilation.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(20): p. 7321-7325.

43. Zhang, Y.H.P. and L.R. Lynd, *Kinetics and relative importance of phosphorolytic and hydrolytic cleavage of cellodextrins and cellobiose in cell extracts of Clostridium thermocellum.* Applied and Environmental Microbiology, 2004. **70**(3): p. 1563-1569.

44. Arai, M., K. Tanaka, and T. Kawaguchi, *Purification and Properties of Cellodextrin Phosphorylase from Clostridium-Thermocellum.* Journal of Fermentation and Bioengineering, 1994. **77**(3): p. 239-242.

45. Alexander, J.K., *Purification and specificity of cellobiose phosphorylase from Clostridium thermocellum.* J Biol Chem, 1968. **243**(11): p. 2899-904.

46. Alexander, J.K., *Characteristics of cellobiose phosphorylase.* J Bacteriol, 1961. **81**: p. 903-10.

47. Lu, Y.P., Y.H.P. Zhang, and L.R. Lynd, *Enzyme-microbe synergy during cellulose hydrolysis by Clostridium thermocellum (vol 103, pg 16165, 2006).* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(51): p. 19605-19605.

48. Washburn, M.P., D. Wolters, and J.R. Yates, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nature Biotechnology, 2001. **19**(3): p. 242-247.

49. Wolters, D.A., M.P. Washburn, and J.R. Yates, *An automated multidimensional protein identification technology for shotgun proteomics.* Analytical Chemistry, 2001. **73**(23): p. 5683-5690.

50. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.

51. Andersen, J.S. and M. Mann, *Organellar proteomics: turning inventories into insights.* EMBO Rep, 2006. **7**(9): p. 874-9.

52. Vollmer, M., E. Nagele, and P. Horth, *Differential proteome analysis: two-dimensional nano-LC/MS of E. coli proteome grown on different carbon sources.* J Biomol Tech, 2003. **14**(2): p. 128-35.

53. Han, M.J., et al., *Proteome-level responses of Escherichia coli to long-chain fatty acids and use of fatty acid inducible promoter in protein production.* J Biomed Biotechnol, 2008. **2008**: p. 735101.

54. Lin, F.M., Y. Tan, and Y.J. Yuan, *Temporal quantitative proteomics of Saccharomyces cerevisiae in response to a nonlethal concentration of furfural.* Proteomics, 2009. **9**(24): p. 5471-83.

55. Dong, M.-Q., et al., *Quantitative Mass Spectrometry Identifies Insulin Signaling Targets in C. elegans.* Science, 2007. **317**(5838): p. 660-663.

56. Matsuoka, S., et al., *ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage.* Science, 2007. **316**(5828): p. 1160-6.

57. Williams, T.I., et al., *Proteomic profile changes in membranes of ethanol-tolerant Clostridium thermocellum.* Applied Microbiology and Biotechnology, 2007. **74**(2): p. 422-432.

58. Zverlov, V.V., J. Kellermann, and W.H. Schwarz, *Functional subgenomics of Clostridium thermocellum cellulosomal genes: Identification of the major catalytic components in the extracellular complex and detection of three new enzymes.* Proteomics, 2005. **5**(14): p. 3646-3653.

59. Gold, N.D. and V.J.J. Martin, *Global view of the Clostridium thermocellum cellulosome revealed by quantitative proteomic analysis.* Journal of Bacteriology, 2007. **189**(19): p. 6787-6795.

60. Raman, B., et al., *Impact of Pretreated Switchgrass and Biomass Carbohydrates on Clostridium thermocellum ATCC 27405 Cellulosome Composition: A Quantitative Proteomic Analysis.* Plos One, 2009. **4**(4): p. -.

61. Yates, J.R., 3rd, et al., *Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database.* Anal Chem, 1995. **67**(8): p. 1426-36.

62. Moore, R.E., M.K. Young, and T.D. Lee, *Protein identification using a quadrupole ion trap mass spectrometer and SEQUEST database matching.* Curr Protoc Protein Sci, 2001. **Chapter 16**: p. Unit 16 10.

63. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-62.

64. MacCoss, M.J., et al., *A Correlation Algorithm for the Automated Quantitative Analysis of Shotgun Proteomics Data.* Analytical Chemistry, 2003. **75**(24): p. 6912-6921.

65. MacCoss, M.J., et al., *A correlation algorithm for the automated quantitative analysis of shotgun proteomics data.* Analytical Chemistry, 2003. **75**(24): p. 6912-6921.

66. Delahunty, C. and J.R. Yates Iii, *Protein identification using 2D-LC-MS/MS.* Methods, 2005. **35**(3): p. 248-255.

67. Bendtsen, J.D., et al., *Improved prediction of signal peptides: SignalP 3.0.* Journal of Molecular Biology, 2004. **340**(4): p. 783-795.

68. Nielsen, H., et al., *Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.* Protein Engineering, 1997. **10**(1): p. 1-6.

69. Krogh, A., et al., *Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes.* Journal of Molecular Biology, 2001. **305**(3): p. 567-580.

70. Tabb, D.L., W.H. McDonald, and J.R. Yates, *DTASelect and contrast: Tools for assembling and comparing protein identifications from shotgun proteomics.* Journal of Proteome Research, 2002. **1**(1): p. 21-26.

71. Zar, J.H., *Biostatistical Analysis*. 4th ed. 1999, Upper Saddle River, New Jersey.: Prentice Hall.

72. Muller, J., et al., *eggNOG v2.0: extending the evolutionary genealogy of genes with enhanced non-supervised orthologous groups, species and functional annotations.* Nucleic Acids Res, 2010. **38**(Database issue): p. D190-5.

73. Zhang, Y. and L.R. Lynd, *Quantification of cell and cellulase mass concentrations during anaerobic cellulose fermentation: development of an enzyme-linked immunosorbent assay-based method with application to Clostridium thermocellum batch cultures.* Analytical Chemistry, 2003. **75**(2): p. 219-27.

74. Zhang, Y.H.P. and L.R. Lynd, *Regulation of cellulase synthesis in batch and continuous cultures of Clostridium thermocellum.* Journal of Bacteriology, 2005. **187**(1): p. 99-106.

75. Kawaguchi, T., et al., *Cloning, nucleotide sequence, and expression of the Clostridium thermocellum cellodextrin phosphorylase gene and its application to synthesis of cellulase inhibitors.* Journal of Fermentation and Bioengineering, 1998. **85**(2): p. 144-149.

76. Lamed, R. and J.G. Zeikus, *Ethanol production by thermophilic bacteria: relationship between fermentation product yields of and catabolic enzyme activities in Clostridium*

thermocellum and Thermoanaerobium brockii. Journal of Bacteriology, 1980. **144**(2): p. 569-78.

77. Lamed, R. and J.G. Zeikus, *Thermostable, ammonium-activated malic enzyme of Clostridium thermocellum.* Biochimica Et Biophysica Acta, 1981. **660**(2): p. 251-5.
78. Zhang, R., et al., *Characteristics and crystal structure of bacterial inosine-5'-monophosphate dehydrogenase.* Biochemistry, 1999. **38**(15): p. 4691-700.
79. Carere, C.R., et al., *Pyruvate catabolism and hydrogen synthesis pathway genes of Clostridium thermocellum ATCC 27405.* Indian Journal of Microbiology, 2008. **48**(2): p. 252-266.
80. Meinecke, B., J. Bertram, and G. Gottschalk, *Purification and Characterization of the Pyruvate-Ferredoxin Oxidoreductase from Clostridium-Acetobutylicum.* Archives of Microbiology, 1989. **152**(3): p. 244-250.
81. Hobbs, M. and J.S. Mattick, *Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes.* Molecular Microbiology, 1993. **10**(2): p. 233-243.
82. Samuels, N.M. and J.P. Klinman, *2,4,5-Trihydroxyphenylalanine quinone biogenesis in the copper amine oxidase from Hansenula polymorpha with the alternate metal nickel.* Biochemistry, 2005. **44**(43): p. 14308-17.
83. Hacisalihoglu, A., J.A. Jongejan, and J.A. Duine, *Distribution of amine oxidases and amine dehydrogenases in bacteria grown on primary amines and characterization of the amine oxidase from Klebsiella oxytoca.* Microbiology, 1997. **143 (Pt 2)**: p. 505-12.

84. Veness, R.G. and C.S. Evans, *The Role of Hydrogen Peroxide in the Degradation of Crystalline Cellulose by Basidiomycete Fungi.* Journal of General Microbiology, 1989. **135**(11): p. 2799-2806.

85. Evans, C.S., et al., *Enzymes and small molecular mass agents involved with lignocellulose degradation.* Fems Microbiology Reviews, 1994. **13**(2-3): p. 235-239.

86. Brissette, J.L., et al., *Phage shock protein, a stress protein of Escherichia coli.* Proc Natl Acad Sci U S A, 1990. **87**(3): p. 862-6.

87. Kobayashi, R., T. Suzuki, and M. Yoshida, *Escherichia coli phage-shock protein A (PspA) binds to membrane phospholipids and repairs proton leakage of the damaged membranes.* Molecular Microbiology, 2007. **66**(1): p. 100-9.

88. DeLisa, M.P., et al., *Phage shock protein PspA of Escherichia coli relieves saturation of protein export via the Tat pathway.* Journal of Bacteriology, 2004. **186**(2): p. 366-73.

89. Jensen, L.J., et al., *eggNOG: automated construction and annotation of orthologous groups of genes.* Nucleic Acids Res, 2008. **36**(Database issue): p. D250-4.

90. Santoni, V., M. Molloy, and T. Rabilloud, *Membrane proteins and proteomics: un amour impossible?* Electrophoresis, 2000. **21**(6): p. 1054-70.

91. Dror, T.W., et al., *Regulation of expression of scaffoldin-related genes in Clostridium thermocellum.* Journal of Bacteriology, 2003. **185**(17): p. 5109-5116.

92. Desvaux, M., *Unravelling carbon metabolism in anaerobic cellulolytic bacteria.* Biotechnology Progress, 2006. **22**(5): p. 1229-1238.

93. Bogdahn, M. and D. Kleiner, *Inorganic nitrogen metabolism in two cellulose-degrading clostridia.* Archives of Microbiology, 1986. **145**(2): p. 159-161.

94. Gerwig, G.J., et al., *Primary structure of O-linked carbohydrate chains in the cellulosome of different Clostridium thermocellum strains.* Eur J Biochem, 1991. **196**(1): p. 115-22.

95. Leibovitz, E., et al., *Characterization and subcellular localization of the Clostridium thermocellum scaffoldin dockerin binding protein SdbA.* Journal of Bacteriology, 1997. **179**(8): p. 2519-2523.

96. Weerapana, E. and B. Imperiali, *Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems.* Glycobiology, 2006. **16**(6): p. 91R-101R.

97. Szymanski, C.M. and B.W. Wren, *Protein glycosylation in bacterial mucosal pathogens.* Nat Rev Microbiol, 2005. **3**(3): p. 225-37.

98. Messner, P., et al., *S-layer nanoglycobiology of bacteria.* Carbohydr Res, 2008. **343**(12): p. 1934-51.

99. Kneidinger, B., et al., *Identification of two GDP-6-deoxy-D-lyxo-4-hexulose reductases synthesizing GDP-D-rhamnose in Aneurinibacillus thermoaerophilus L420-91T.* J Biol Chem, 2001. **276**(8): p. 5577-83.

100. Sparling, R., et al., *Formate synthesis by Clostridium thermocellum during anaerobic fermentation.* Canadian Journal of Microbiology, 2006. **52**(7): p. 681-688.

101. Weimer, P.J. and J.G. Zeikus, *Fermentation of cellulose and cellobiose by Clostridium thermocellum in the absence of Methanobacterium thermoautotrophicum.* Appl Environ Microbiol, 1977. **33**(2): p. 289-97.

102. JOHNSON, E.A., F. BOUCHOT, and A.L. DEMAIN, *Regulation of Cellulase Formation in Clostridium thermocellum.* J Gen Microbiol, 1985. **131**(9): p. 2303-2308.

103. Wells, J.E., et al., *Cellodextrin efflux by the cellulolytic ruminal bacterium Fibrobacter succinogenes and its potential role in the growth of nonadherent bacteria.* Appl Environ Microbiol, 1995. **61**(5): p. 1757-62.

104. Matulova, M., et al., *Degradation of wheat straw by Fibrobacter succinogenes S85: a liquid- and solid-state nuclear magnetic resonance study.* Appl Environ Microbiol, 2005. **71**(3): p. 1247-53.

105. Newcomb, M., C.Y. Chen, and J.H.D. Wu, *Induction of the celC operon of Clostridium thermocellum by laminaribiose.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(10): p. 3747-3752.

106. Newcomb, M., and J. H. D. Wu. , *Transcription regulators of the Clostridium thermocellum cellulase system*. Biotechnology of lignocellulose degradation and biomass utilization., ed. K.S. K. Ohmiya, S. Karita, T. Kimura, M. Sakka, and Y. Onishi. 2004, Tokyo, Japan.: Uni Publishers.

107. Nataf, Y., et al., *Clostridium thermocellum cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors.* Proc Natl Acad Sci U S A, 2010. **107**(43): p. 18646-51.

108. Kahel-Raifer, H., et al., *The unique set of putative membrane-associated anti-sigma factors in Clostridium thermocellum suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation.* FEMS Microbiol Lett, 2010. **308**(1): p. 84-93.

109. Bahari, L., et al., *Glycoside hydrolases as components of putative carbohydrate biosensor proteins in Clostridium thermocellum.* J Ind Microbiol Biotechnol, 2010.

Appendices

Appendix A; In-gel trypsin digestion protocol, provided by Dr. Heng Jiang

All reagents should be prepared fresh. The trypsin used should be proteomics grade and modified to resist autolysis and be pure of chymotrypsin activity. All solvents including water should be HPLC-grade or higher. Protocol begins from a stained and washed SDS-PAGE gel containing the proteins for analysis.

1. Cut the gel band (1 mm × 1 mm) and place them into an 1.5-mL tube.
2. Add 300 µL of 10 mM DTT/50 mM ammonium bicarbonate and incubate for 30 min at room temperature.
3. Remove the above solution. Add 300 µL of 50 mM iodoacectimide /50 mM ammonium bicarbonate and incubate protected from light for 30 min at room temperature.
4. Remove the above solution. Add 300 µL of 50 mM ammonium bicarbonate, and incubate for 15 min at room temperature.
5. Add 300 µL of 25mM ammonium bicarbonate/5% acetonitrile and incubate for 15 min at room temperature.
6. Remove the above solution. Add 300 µL of 25mM ammonium bicarbonate/50% acetonitrile and incubate for 30 min at room temperature.
7. Repeat Step 6.
8. Remove the above solution. Add 300 µL of 100% acetonitrile to each tube and incubate for 10 min.
9. Remove acetonitrile and dry gel pieces using a SpeedVac (~10 min).

10. Add ~50 µL (just enough to rehydrate the gel pieces) of trypsin solution (20 µg of trypsin in 2.0 mL of 25 mM ammonium bicarbonate).
11. Incubate overnight at 30°C.
12. Add four-fold volume of trypsin solution of 0.5 % formic acid/60% acetonitrile and incubate at room temperature for 30 min at room temperature.
13. Collect the solution and repeat Step 12.
14. Repeat Step 13.
15. Dry tryptic peptides using a SpeedVac at *medium* temperature. Samples can be stored at -20°C utill MS analysis.

Appendix B; Putative transport proteins of *C. thermocellum*, identified by MS

Putative transport proteins from the *C. thermocellum* were compiled based upon current gene annotation. The presence of a signal peptide and the number of trans-membrane helices were determined for each ORF using by SignalP 3.0 [67-68], TMHMM 2.0 [69] respectively, both available at (http://www.cbs.dtu.dk/). The transport protein components that were identified by MS are listed in Table B1.

Table B1. Putative transport proteins of *C. thermocellum* identified by MS. Proteins shaded in grey are potentially encoded by a transport operon.

Gene Number and Annotation	#TMH	SP	Identification
Cthe_0031 putative membrane transporter	6		
Cthe_0064 Na/Pi-cotransporter II-related protein	9		
Cthe_0384 ABC transporter related protein	0		A
Cthe_0391 ABC transporter related protein	0		A,C
Cthe_0392 inner-membrane translocator	7	Y	
Cthe_0393 sugar ABC transporter sugar-binding protein , CbpA	0	Y	A,C
Cthe_0395 RbsD or FucU transport	0		
Cthe_0396 ABC transporter related protein	3		A,C
Cthe_0397 ABC transporter related protein	5	Y	A,C
Cthe_0534 ABC-bacteriocin transporter	5		A,C
Cthe_0539 ABC transporter related protein	0		
Cthe_0547 periplasmic solute binding protein	1	Y	
Cthe_0548 ABC transporter related protein	0		A,C
Cthe_0549 ABC-3	7		
Cthe_0612 sodium solute transporter superfamily	13	Y	A,C
Cthe_0618 ferrous iron transport protein B	10		A,C
Cthe_0747 extracellular solute-binding protein,	0	Y	A,C
Cthe_0748 transport inner membrane component	6	Y	
Cthe_0749 transport inner membrane component	6		
Cthe_0750 spermidine/putrescine ABC transporter ATPase subunit	0		A,C
Cthe_0819 ABC transporter related protein	0		A,C
Cthe_0910 extracellular solute-binding protein, family 5	1	Y	A,C
Cthe_1018 transport inner membrane component	6	Y	A,C
Cthe_1019 transport inner membrane component	6		A,C
Cthe_1020 extracellular solute-binding protein, family 1 , CbpB	0	Y	A,C
Cthe_1189 ABC transporter related protein	0		A,C
Cthe_1410 cation diffusion facilitator familtransporter	5		A
Cthe_1427 branched-chain amino acid transport	4		
Cthe_1415 ABC transporter related protein	0		
Cthe_1046 extracellular solute-binding protein,	1	Y	
Cthe_1456 ABC transporter related protein	0		A,C
Cthe_1457 polar amino acid ABC transporter, inner membrane subunit	4		A,C
Cthe_1458 extracellular solute-binding protein, family 3	0	Y	
Cthe_1499 ammonium transporter	11	Y	
Cthe_1500 ABC transporter related protein	5		

#TMH, number of trans-membrane helices; SP, presence of a signal peptide for secretion; A, identified by MS in Avicel grown cells; C, Identified by MS in cellobiose grown cells.

Table B1 (cont.). Putative transport proteins of *C. thermocellum* identified by MS. Proteins shaded in grey are potentially encoded by a transport operon.

Gene Number and Annotation	TMH	SP	Identification
Cthe_1501 ABC transporter, transmembrane region	6		
Cthe_1526 ABC transporter related protein	0		
Cthe_1536 ABC transporter related protein	0		
Cthe_1555 NLPA lipoprotein	1	Y	
Cthe_1556 binding-protein-dependent transport inner membrane component	5		
Cthe_1557 ABC transporter related protein	0		
Cthe_1563 ABC transporter related protein	0		C
Cthe_1570 extracellular solute-binding protein, family 3	1	Y	A
Cthe_1571 binding-protein-dependent transport inner membrane component	9		
Cthe_1572 ABC transporter related protein	0		
Cthe_1576 basic membrane lipoprotein, Lbp	0	Y	A
Cthe_1577 hypothetical protein	0		
Cthe_1578 Radical SAM	0		
Cthe_1579 ABC transporter related protein	0		
Cthe_1580 inner-membrane translocator	8	Y	
Cthe_1581 inner-membrane translocator	8		
Cthe_1586 binding-protein-dependent transport inner membrane component	12		
Cthe_1587 ABC transporter related protein	0		
Cthe_1588 extracellular solute-binding protein, famil1	0	Y	
Cthe_1602 phosphate ABC transporter, ATPase subunit	0		
Cthe_1603 phosphate ABC transporter, inner membrane subunit PstA	6		
Cthe_1604 phosphate ABC transporter, inner membrane subunit PstC	6	Y	
Cthe_1605 phosphate ABC transporter binding protein-like protein	0	Y	
Cthe_1667 ABC-2 type transporter	6		
Cthe_1668 ABC transporter related protein	0		
Cthe_1685 ABC transporter related protein	6		A,C
Cthe_1752 ABC transporter related protein	0		C
Cthe_1753 transport spermease protein	8		
Cthe_1754 periplasmic binding protein	1		
Cthe_1762 RND family efflux transporter MFP subunit	1		
Cthe_1763 ABC transporter related protein	0		A,C
Cthe_1799 ABC transporter related protein	0		
Cthe_1801 ABC transporter related protein	0		A,C
Cthe_1802 cobalt ABC transporter, inner membrane subunit CbiQ	4		A
Cthe_1803 cobalamin vitamin B12 biosynthesis CbiM protein	8	Y	A,C
Cthe_1819 ABC transporter related protein	0		

#TMH, number of trans-membrane helices; SP, presence of a signal peptide for secretion; A, identified by MS in Avicel grown cells; C, Identified by MS in cellobiose grown cells.

Table B1 (cont.). Putative transport proteins of *C. thermocellum* identified by MS. Proteins shaded in grey are potentially encoded by a transport operon.

Gene Number and Annotation	TMH	SP	Identification
Cthe_1820 ABC transporter related protein	0		
Cthe_1821 inner-membrane translocator	9		
Cthe_1822 inner-membrane translocator	7		
Cthe_1848 copper-translocating P-type ATPase	7		A,C
Cthe_1849 Heavy metal transport/detoxification protein	7		A,C
Cthe_1862 ABC transporter related protein	0		A,C
Cthe_1917 ATPase, P-type (transporting), HAD superfamily, subfamily IC	8		A,C
Cthe_1919 MgtC/SapB transporter	4		A,C
Cthe_1939 magnesium transporter	3		
Cthe_1957 extracellular solute-binding protein, famil1	1	Y	A,C
Cthe_2110 ABC transporter related protein	0		A
Cthe_2116 binding-protein-dependent transport inner membrane component	6		
Cthe_2117 putative sulfonate transport substrate-binding protein	1	Y	A,C
Cthe_2118 ABC transporter related protein	0		A,C
Cthe_2125 binding-protein-dependent transport inner membrane component	6	Y	
Cthe_2126 binding-protein-dependent transport inner membrane component	6		
Cthe_2128 extracellular solute-binding protein, family 1, CbpC	0	Y	
Cthe_2215 Mg2+ transporter protein, CorA-like protein	2		A,C
Cthe_2262 V-type ATPase, 116 kDa subunit	7		C
Cthe_2263 H+-transporting two-sector ATPase, C subunit	4	Y	A,C
Cthe_2264 H+-transporting two-sector ATPase, E subunit	0		A,C
Cthe_2265 V-type ATP synthase subunit C	0		A,C
Cthe_2266 V-type ATP synthase subunit F	0		
Cthe_2267 V-type ATP synthase subunit A	0		A,C
Cthe_2268 V-type ATP synthase subunit B	0		A,C
Cthe_2269 V-type ATP synthase subunit D	0		
Cthe_2270 ABC transporter related protein	0		C
Cthe_2278 extracellular solute-binding protein, family 3	1	Y	A,C
Cthe_2279 polar amino acid ABC transporter, inner membrane subunit	5		
Cthe_2280 ABC transporter related protein	0		
Cthe_2290 ABC transporter related protein	0		
Cthe_2367 60 kDa inner membrane insertion protein	4		
Cthe_2405 Heavy metal transport/detoxification protein	0		
Cthe_2446 ABC-type sugar transport system, CbpD	1	Y	
Cthe_2447 ABC transporter related protein	0		
Cthe_2448 inner-membrane translocator	9	Y	
Cthe_2449 Phosphoglycerate mutase	0		

#TMH, number of trans-membrane helices; SP, presence of a signal peptide for secretion; A, identified by MS in Avicel grown cells; C, Identified by MS in cellobiose grown cells.

Table B1 (cont.). Putative transport proteins of *C. thermocellum* identified by MS. Proteins shaded in grey are potentially encoded by a transport operon.

Gene Number and Annotation	TMH	SP	Identification
Cthe_2531 sulfate ABC transporter, periplasmic sulfate-binding protein	1	Y	
Cthe_2532 sulfate ABC transporter, inner membrane subunit C	6	Y	
Cthe_2533 sulfate ABC transporter, inner membrane subunit C	6	Y	
Cthe_2534 sulfate ABC transporter, ATPase subunit	0		C
Cthe_2573 ABC transporter related protein	0		
Cthe_2574 binding-protein-dependent transport inner membrane component	6		
Cthe_2664 ABC-1	2		
Cthe_2666 ATPase, P-type (transporting), HAD superfamily, subfamily IC	8		
Cthe_2706 ABC transporter related protein	0		A,C
Cthe_2707 ABC-type transport system multi-copper enzyme maturation,	6		
Cthe_2777 nicotinamide mononucleotide transporter PnuC	8		
Cthe_2790 ABC transporter related protein	0		A,C
Cthe_2791 binding-protein-dependent transport inner membrane component	8		
Cthe_2803 binding-protein-dependent transport inner membrane component	6		
Cthe_2804 ABC transporter related protein	0		
Cthe_2810 Na/Pi-cotransporter II-related protein	9		
Cthe_2934 ABC transporter related protein	0		C
Cthe_2936 ABC transporter related protein	0		
Cthe_2937 cobalt transport protein	4		
Cthe_2942 ABC transporter related protein	0		
Cthe_2943 ABC-2 type transporter	7		
Cthe_2961 extracellular solute-binding protein, famil5	0	Y	A,C
Cthe_2962 oligopeptide/dipeptide ABC transporter, ATPase subunit	0		C
Cthe_2963 oligopeptide/dipeptide ABC transporter, ATPase subunit	0		C
Cthe_2964 binding-protein-dependent transport inner membrane component	6		
Cthe_2965 binding-protein-dependent transport inner membrane component	6	Y	C
Cthe_2970 ABC transporter related protein	0		A
Cthe_2998 ABC transporter related protein	5		A,C
Cthe_3000 phosphate transporter	7		
Cthe_3066 ABC transporter related protein	0		A,C
Cthe_3129 Citrate transporter	9	Y	A,C
Cthe_3147 ABC transporter related protein	6		C
Cthe_3148 ABC transporter related protein	5	Y	
Cthe_3170 ABC transporter related protein	0		
Cthe_3184 cation transporter	10		

#TMH, number of trans-membrane helices; SP, presence of a signal peptide for secretion; A, identified by MS in Avicel grown cells; C, Identified by MS in cellobiose grown cells.

Appendix C; Soluble fraction protien identifications of Avicel grown cells.

Table C1. Complete list of proteins identified in the soluble fraction of *C. thermocellum* grown on Avicel, ranked by XC score.

Protein	Score (XC)[a]	Coverage (%)	MW (daltons)	Peptide (Hits)[b]
YP_001036854 Cthe_0423 iron-containing alcohol dehydrogenase	560.29	49.00	95948.20	926 (926 0 0 0 0)
YP_001039508 Cthe_3120 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	440.28	31.10	130938.90	302 (301 1 0 0 0)
YP_001039121 Cthe_2729 translation elongation factor G	320.25	48.60	77586.17	259 (258 1 0 0 0)
YP_001037445 Cthe_1020 extracellular solute-binding protein, family 1	300.26	49.20	49953.82	235 (232 3 0 0 0)
YP_001039122 Cthe_2730 translation elongation factor Tu	288.27	63.50	44160.61	1085 (1084 1 0 0 0)
YP_001036572 Cthe_0138 Phosphoglycerate kinase	270.26	56.20	42735.32	191 (191 0 0 0 0)
YP_001038203 Cthe_1789 ATPase AAA-2	270.23	34.10	92041.58	121 (121 0 0 0 0)
YP_001036935 Cthe_0505 formate acetyltransferase	260.25	34.50	84351.19	136 (135 0 1 0 0)
YP_00103Cthe_0053 ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent	260.25	29.20	88248.48	173 (173 0 0 0 0)
YP_001037733 Cthe_1308 pyruvate, phosphate dikinase	230.21	27.10	98703.04	129 (129 0 0 0 0)
YP_001036571 Cthe_0137 glyceraldehyde-3-phosphate dehydrogenase, type I	210.25	59.20	36116.36	689 (689 0 0 0 0)
YP_001039116 Cthe_2724 DNA-directed RNA polymerase, beta subunit	200.29	16.10	139869.00	73 (73 0 0 0 0)
YP_001036780 Cthe_0349 fructose-1,6-bisphosphate aldolase, class II	200.24	45.30	33611.15	176 (176 0 0 0 0)
YP_001036772 Cthe_0341 NADH dehydrogenase (quinone)	200.22	35.50	64854.81	145 (142 0 0 2 1)
YP_001039283 Cthe_2892 chaperonin GroEL	198.27	35.30	57438.54	42 (41 1 0 0 0)
YP_001039160 Cthe_2768 metallophosphoesterase	198.21	23.10	98713.68	133 (131 1 1 0 0)
YP_001036778 Cthe_0347 phosphofructokinase	190.22	52.50	45595.20	167 (167 0 0 0 0)
YP_001036717 Cthe_0285 isocitrate dehydrogenase, NADP-dependent	178.23	45.00	45753.04	91 (89 2 0 0 0)
YP_001036775 Cthe_0344 Malate dehydrogenase	170.25	41.50	42073.30	379 (379 0 0 0 0)

YP_001038787 Cthe_2392 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	170.23	34.80	43405.89	79 (79 0 0 0 0)
YP_001037810 Cthe_1385 preprotein translocase, SecA subunit	170.21	21.00	104546.40	86 (86 0 0 0 0)
YP_001038788 Cthe_2393 thiamine pyrophosphate enzyme-like TPP-binding	160.25	39.20	34626.82	105 (105 0 0 0 0)
YP_001037453 Cthe_1028 acetate kinase	160.25	49.10	43937.85	63 (63 0 0 0 0)
YP_001039265 Cthe_2874 Phosphoenolpyruvate carboxykinase (GTP)	160.23	28.90	67583.80	135 (135 0 0 0 0)
YP_001036849 Cthe_0418 Polyribonucleotide nucleotidyltransferase	160.21	22.90	77269.86	51 (51 0 0 0 0)
YP_001037653 Cthe_1228 threonyl-tRNA synthetase	154.21	18.60	73518.59	75 (73 1 1 0 0)
YP_001037636 Cthe_1211 pyridoxal-phosphate dependent TrpB-like enzyme	150.24	34.60	50100.10	215 (215 0 0 0 0)
YP_001036577 Cthe_0143 Phosphopyruvate hydratase	150.23	44.80	47073.19	96 (96 0 0 0 0)
YP_001039001 Cthe_2608 ATP synthase F1, beta subunit	150.23	45.00	50980.69	44 (44 0 0 0 0)
YP_001036574 Cthe_0140 phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	150.21	41.50	57329.97	52 (52 0 0 0 0)
YP_001039117 Cthe_2725 DNA-directed RNA polymerase, beta' subunit	150.20	10.30	130074.00	86 (86 0 0 0 0)
YP_001037430 Cthe_1005 translation elongation factor Ts	148.23	40.50	24321.70	74 (73 1 0 0 0)
YP_001039416 Cthe_3027 Citrate (Si)-synthase	140.23	31.10	52000.72	59 (58 1 0 0 0)
YP_001036861 Cthe_0430 hydrogenase, Fe-only	140.19	27.00	62491.40	47 (47 0 0 0 0)
YP_001036773 Cthe_0342 hydrogenase, Fe-only	138.22	22.30	63972.41	65 (63 1 1 0 0)
YP_001037431 Cthe_1006 ribosomal protein S2	134.18	29.40	28589.95	54 (52 0 1 1 0)
YP_001038276 Cthe_1862 ABC transporter related protein	130.24	37.80	42036.93	171 (171 0 0 0 0)
YP_001036715 Cthe_0283 aldo/keto reductase	130.23	42.20	35432.49	38 (38 0 0 0 0)
YP_001039113 Cthe_2721 ribosomal protein L1	130.22	33.30	25057.48	79 (75 3 0 1 0)
YP_001039323 Cthe_2932 DNA-directed RNA polymerase, alpha subunit	126.19	28.30	35017.31	70 (68 1 1 0 0)
YP_001037523 Cthe_1098 hypothetical protein	120.23	30.10	49674.38	41 (41 0 0 0 0)
YP_001036860 Cthe_0429 NADH dehydrogenase (quinone)	120.23	17.80	68080.16	50 (49 1 0 0 0)
YP_001036855 Cthe_0424 aminoglycoside phosphotransferase	120.23	32.50	28568.92	129 (129 0 0 0 0)

YP_001038999 Cthe_2606 ATP synthase F1, alpha subunit	120.21	20.40	55784.93	31 (31 0 0 0 0)
YP_001037475 Cthe_1050 recA protein	110.23	40.70	37781.88	29 (29 0 0 0 0)
YP_001037388 Cthe_0961 aspartate-semialdehyde dehydrogenase	110.22	40.70	36705.75	80 (80 0 0 0 0)
YP_001039022 Cthe_2630 ribose-phosphate pyrophosphokinase	110.22	23.40	34730.42	31 (31 0 0 0 0)
YP_001037630 Cthe_1205 putative serine protein kinase, PrkA	110.19	12.90	74227.23	33 (33 0 0 0 0)
YP_001037160 Cthe_0732 Chorismate synthase	102.23	26.10	43128.44	20 (19 0 0 0 1)
YP_001037129 Cthe_0701 conserved carboxylase region containing protein	100.26	25.40	52338.23	26 (26 0 0 0 0)
YP_001039156 Cthe_27TROVE domain containing protein	100.22	26.90	55376.65	28 (28 0 0 0 0)
YP_001039306 Cthe_2915 ribosomal protein L5	100.22	48.40	20391.86	41 (37 0 2 0 2)
YP_001038489 Cthe_2089 glycoside hydrolase, family 48	100.22	14.70	83505.41	23 (23 0 0 0 0)
YP_001039295 Cthe_2904 ribosomal protein L4/L1e	100.21	34.60	23295.76	56 (56 0 0 0 0)
YP_001039313 Cthe_2922 ribosomal protein L15	100.20	48.60	15576.57	49 (49 0 0 0 0)
YP_001039466 Cthe_3077 cellulosome anchoring protein, cohesin region	100.19	11.20	196709.70	69 (69 0 0 0 0)
YP_001039590 Cthe_3202 CRISPR-associated protein, Csh2 family	100.19	36.70	34861.18	28 (28 0 0 0 0)
YP_001038496 Cthe_2096 methionyl-tRNA synthetase	98.22	21.30	75095.11	21 (20 1 0 0 0)
YP_001039120 Cthe_2728 ribosomal protein S7	90.26	39.10	17896.51	64 (64 0 0 0 0)
YP_001038208 Cthe_1794 translation elongation factor G	90.25	13.70	76682.54	37 (37 0 0 0 0)
YP_001037148 Cthe_0720 aminotransferase, class V	90.21	20.30	43376.50	46 (46 0 0 0 0)
YP_001037331 Cthe_0904 protein-export membrane protein SecD	90.21	20.40	48261.15	20 (20 0 0 0 0)
YP_001036776 Cthe_0345 L-lactate dehydrogenase	90.21	14.80	34811.13	63 (63 0 0 0 0)
YP_001038280 Cthe_1866 acetylornithine and succinylornithine aminotransferases	90.21	20.50	43892.60	14 (14 0 0 0 0)
YP_001037293 Cthe_0866 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	90.20	30.60	38061.73	23 (23 0 0 0 0)
YP_001039309 Cthe_2918 ribosomal protein L6	80.23	33.30	20135.12	27 (27 0 0 0 0)
YP_001039322 Cthe_2931 ribosomal protein S4	80.22	23.10	23926.69	32 (32 0 0 0 0)

YP_001038913 Cthe_2518 ketol-acid reductoisomerase	80.22	32.00	36199.42	20 (20 0 0 0 0)
YP_001037671 Cthe_1246 phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	80.22	18.50	56306.42	28 (28 0 0 0 0)
YP_001037857 Cthe_1433 short-chain dehydrogenase/reductase SDR	80.22	24.90	27668.10	35 (35 0 0 0 0)
YP_001039320 Cthe_2929 ribosomal protein S13	80.21	29.30	14158.91	30 (30 0 0 0 0)
YP_001038375 Cthe_19FAD-dependent pyridine nucleotide-disulphide oxidoreductase	80.21	14.70	55740.11	22 (22 0 0 0 0)
YP_001038578 Cthe_2183 UTP-glucose-1-phosphate uridylyltransferase	80.20	21.50	32481.11	17 (17 0 0 0 0)
YP_001037193 Cthe_0765 ribosomal protein L19	80.20	49.10	12924.24	40 (40 0 0 0 0)
YP_001037662 Cthe_1237 leucyl-tRNA synthetase	80.19	10.70	94825.20	38 (38 0 0 0 0)
YP_001039324 Cthe_2933 ribosomal protein L17	80.18	34.50	19695.61	37 (37 0 0 0 0)
YP_001037291 Cthe_08pyruvate ferredoxin/flavodoxin oxidoreductase	80.18	32.40	19329.96	44 (44 0 0 0 0)
YP_001037793 Cthe_1368 S-layer-like domain containing protein	78.22	8.30	78075.02	23 (22 1 0 0 0)
YP_001037149 Cthe_0721 nitrogen-fixing NifU-like protein	72.23	56.40	16208.82	56 (54 0 1 0 1)
YP_001039159 Cthe_2767 Methyltransferase type 12	70.22	17.00	53542.80	13 (13 0 0 0 0)
YP_001039304 Cthe_2913 ribosomal protein L14	70.22	32.80	13446.39	53 (53 0 0 0 0)
YP_001039266 Cthe_2875 sigma 54 modulation protein/ribosomal protein S30EA	70.22	49.70	20531.70	24 (24 0 0 0 0)
YP_001039308 Cthe_2917 ribosomal protein S8	70.20	50.80	14209.80	61 (61 0 0 0 0)
YP_001037142 Cthe_0714 hydroxymethylbutenyl pyrophosphate reductase	70.19	7.80	77377.34	45 (45 0 0 0 0)
YP_001038209 Cthe_1795 phospho-2-dehydro-3-deoxyheptonate aldolase	70.18	28.20	36794.52	28 (28 0 0 0 0)
YP_001037151 Cthe_0723 tyrosyl-tRNA synthetase	70.18	14.00	46607.92	15 (15 0 0 0 0)
YP_001036569 Cthe_0135 beta-ketoacyl synthase	70.18	3.10	305695.50	17 (17 0 0 0 0)
YP_001037349 Cthe_0922 diaminopimelate dehydrogenase	70.17	24.00	37201.20	12 (12 0 0 0 0)
YP_001039297 Cthe_2906 ribosomal protein L2	70.17	24.00	30176.28	44 (44 0 0 0 0)
YP_001036833 Cthe_0402 copper amine oxidase-like protein	70.17	10.70	83083.48	11 (11 0 0 0 0)

YP_001037285 Cthe_0858 protein of unknown function DUF1432	68.21	16.40	35281.07	41 (40 1 0 0 0)
YP_001038997 Cthe_2604 ATP synthase F0, B subunit	68.19	25.40	21057.12	23 (22 1 0 0 0)
YP_001037376 Cthe_0949 carbamoyl-phosphate synthase, large subunit	64.16	3.50	118645.00	8 (7 0 0 1 0)
YP_001037389 Cthe_0962 dihydrodipicolinate synthase	60.26	21.20	31954.94	15 (15 0 0 0 0)
YP_001038181 Cthe_1767 hypothetical protein	60.23	14.70	35361.88	20 (20 0 0 0 0)
YP_001039546 Cthe_3158 putative aconitate hydratase	60.22	9.20	68720.59	7 (7 0 0 0 0)
YP_001038241 Cthe_1827 copper amine oxidase-like protein	60.22	12.70	36388.99	44 (44 0 0 0 0)
YP_001037916 Cthe_1495 pyridoxamine 5'-phosphate oxidase-related, FMN-binding	60.22	37.70	14943.53	20 (20 0 0 0 0)
YP_001039300 Cthe_2909 ribosomal protein S3	60.21	34.50	25312.76	69 (69 0 0 0 0)
YP_001039329 Cthe_2938 putative glucokinase, ROK family	60.21	20.20	33872.38	20 (20 0 0 0 0)
YP_001037275 Cthe_0847 translation elongation factor P	60.21	28.10	20640.65	10 (10 0 0 0 0)
YP_001039276 Cthe_2885 phosphoribosylaminoimidazole-succinocarboxamide synthase	60.20	24.50	33151.18	31 (31 0 0 0 0)
YP_001039299 Cthe_2908 ribosomal protein L22	60.20	43.20	14066.90	25 (25 0 0 0 0)
YP_001038256 Cthe_1842 O-acetylhomoserine/O-acetylserine sulfhydrylase	60.20	12.00	46498.20	13 (13 0 0 0 0)
YP_001039467 Cthe_3078 cellulosome anchoring protein, cohesin region	60.20	3.00	248014.30	31 (31 0 0 0 0)
YP_001039588 Cthe_3200 alanyl-tRNA synthetase	60.19	9.30	97985.39	18 (18 0 0 0 0)
YP_001038991 Cthe_2598 uracil phosphoribosyltransferase	60.19	23.00	23078.13	22 (22 0 0 0 0)
YP_001037531 Cthe_1106 twitching motility protein	60.19	28.50	38896.07	11 (11 0 0 0 0)
YP_001037516 Cthe_1091 metal dependent phosphohydrolase	60.19	6.30	58726.74	12 (12 0 0 0 0)
YP_001036876 Cthe_0445 cell division protein FtsZ	60.18	22.60	39703.82	25 (25 0 0 0 0)
YP_001038817 Cthe_2422 hypothetical protein	60.18	24.10	40412.15	15 (15 0 0 0 0)
YP_00103Cthe_0042 small GTP-binding protein	60.17	24.00	44243.27	10 (9 1 0 0 0)
YP_001036594 Cthe_0160 ribosomal protein L21	60.17	58.30	11531.22	20 (20 0 0 0 0)

YP_001037127 Cthe_0699 carboxyl transferase	60.16	15.70	56012.75	9 (9 0 0 0 0)
YP_001039614 Cthe_3226 copper amine oxidase-like protein	58.21	21.40	29604.39	43 (38 5 0 0 0)
YP_001039543 Cthe_3155 beta-lactamase-like protein	52.19	29.10	26983.65	10 (9 0 0 0 1)
YP_001039321 Cthe_2930 ribosomal protein S11	50.26	11.90	14342.67	40 (40 0 0 0 0)
YP_001037283 Cthe_0856 branched-chain amino acid aminotransferase	50.24	21.10	39711.57	41 (41 0 0 0 0)
YP_001037732 Cthe_1307 cellulosome anchoring protein, cohesin region	50.22	7.10	68577.19	26 (26 0 0 0 0)
YP_001037686 Cthe_1261 6-phosphofructokinase	50.22	9.90	34773.23	26 (26 0 0 0 0)
YP_001039293 Cthe_2902 ribosomal protein S10	50.22	31.10	11770.55	16 (16 0 0 0 0)
YP_001037690 Cthe_1265 phosphoglucomutase/phosphomannomutase alpha/beta/alpha domain I	50.21	13.80	64914.06	17 (17 0 0 0 0)
YP_001038324 Cthe_1912 copper amine oxidase-like protein	50.21	10.10	59214.88	13 (13 0 0 0 0)
YP_001038752 Cthe_2357 hypothetical protein	50.20	15.10	44618.16	13 (13 0 0 0 0)
YP_001036804 Cthe_0373 oxidoreductase FAD/NAD(P)-binding	50.19	21.40	30685.13	7 (7 0 0 0 0)
YP_001038939 Cthe_2544 RNA polymerase, sigma-24 subunit, ECF subfamily	50.19	33.70	22184.37	16 (16 0 0 0 0)
YP_001038743 Cthe_2348 Ig-related protein	50.19	6.50	113260.90	11 (11 0 0 0 0)
YP_001037756 Cthe_1331 aspartyl-tRNA synthetase	50.19	4.90	67727.98	20 (20 0 0 0 0)
YP_001038803 Cthe_2408 phage shock protein A, PspA	50.18	24.80	27739.44	36 (36 0 0 0 0)
YP_001039012 Cthe_2619 cell shape determining protein, MreB/Mrl family	50.18	8.10	36744.56	8 (8 0 0 0 0)
YP_001038785 Cthe_2390 pyruvate/ketoisovalerate oxidoreductase, gamma subunit	50.18	24.50	21190.07	38 (38 0 0 0 0)
YP_001037777 Cthe_1352 UDP-glucose 6-dehydrogenase	50.17	17.80	48985.39	6 (6 0 0 0 0)
YP_001036Cthe_0215 phenylalanyl-tRNA synthetase, beta subunit	50.17	8.30	88432.06	9 (9 0 0 0 0)
YP_001036573 Cthe_0139 Triose-phosphate isomerase	50.17	25.50	27116.99	12 (12 0 0 0 0)
YP_001039338 Cthe_2947 prolyl-tRNA synthetase	50.16	5.20	64228.50	11 (11 0 0 0 0)
YP_001037416 Cthe_0991 translation initiation factor IF-2	48.17	4.10	114885.30	5 (4 1 0 0 0)
YP_001037524 Cthe_1099 hypothetical protein	46.19	6.80	94171.84	7 (6 0 1 0 0)

YP_001037527 Cthe_1102 fimbrial assembly protein	40.26	23.50	19870.68	27 (27 0 0 0 0)
YP_001039512 Cthe_3124 AMP-dependent synthetase and ligase	40.24	6.80	61611.04	5 (5 0 0 0 0)
YP_001039311 Cthe_2920 ribosomal protein S5-like protein	40.24	37.30	17531.73	31 (31 0 0 0 0)
YP_001038603 Cthe_2208 LexA DNA-binding region containing protein	40.23	17.40	22036.65	7 (7 0 0 0 0)
YP_001037295 Cthe_0868 type IV pilus assembly PilZ	40.23	18.60	27940.03	12 (12 0 0 0 0)
YP_001039484 Cthe_3096 hypothetical protein	40.22	15.40	33801.52	12 (12 0 0 0 0)
YP_001039115 Cthe_2723 ribosomal protein L7/L12	40.22	21.70	13287.30	10 (10 0 0 0 0)
YP_001037925 Cthe_1504 Linocin_M18 bacteriocin protein	40.21	13.30	30295.85	6 (6 0 0 0 0)
YP_001036788 Cthe_0357 alpha-glucan phosphorylases	40.21	6.90	97651.16	12 (12 0 0 0 0)
YP_001039301 Cthe_2910 ribosomal protein L16	40.21	29.70	16213.71	26 (26 0 0 0 0)
YP_001039114 Cthe_2722 ribosomal protein L10	40.21	16.90	19501.65	31 (31 0 0 0 0)
YP_001038778 Cthe_2383 hypothetical protein	40.21	13.20	33181.84	35 (35 0 0 0 0)
YP_001037104 Cthe_0676 tyrosine recombinase XerD	40.20	12.50	34269.43	19 (19 0 0 0 0)
YP_001038Cthe_2251 Methionine adenosyltransferase	40.20	12.60	43359.20	13 (13 0 0 0 0)
YP_001039461 Cthe_3072 acyl-ACP thioesterase	40.20	17.80	29446.92	11 (11 0 0 0 0)
YP_001037518 Cthe_1093 Methenyltetrahydrofolate cyclohydrolase	40.20	18.60	30767.42	7 (7 0 0 0 0)
YP_001036527 Cthe_0093 septum site-determining protein MinD	40.20	15.80	28892.48	12 (12 0 0 0 0)
YP_001039096 Cthe_2704 Transketolase-like protein	40.19	18.50	31356.90	8 (8 0 0 0 0)
YP_001037043 Cthe_0615 Phenylacetate--CoA ligase	40.19	11.80	48686.20	9 (9 0 0 0 0)
YP_001036771 Cthe_0340 ferredoxin	40.19	32.00	13423.96	12 (12 0 0 0 0)
YP_001038196 Cthe_1782 ribosomal protein S9	40.18	36.90	14656.17	16 (16 0 0 0 0)
YP_001036793 Cthe_0362 transcriptional regulator, AsnC family	40.18	22.00	18091.50	8 (8 0 0 0 0)
YP_001037359 Cthe_0932 beta-ketoacyl synthase	40.18	10.50	43935.96	7 (7 0 0 0 0)
YP_001039424 Cthe_3035 D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	40.18	10.50	42835.39	20 (20 0 0 0 0)
YP_001039097 Cthe_2705 Transketolase, central region	40.17	18.00	33660.42	9 (9 0 0 0 0)

YP_001039021 Cthe_2629 UDP-N-acetylglucosamine pyrophosphorylase	40.17	7.50	51493.20	8 (8 0 0 0 0)
YP_001038776 Cthe_2381 seryl-tRNA synthetase	40.17	8.50	48456.13	6 (6 0 0 0 0)
YP_0010372Cthe_0836 hypothetical protein	40.17	19.80	20373.11	18 (18 0 0 0 0)
YP_001039315 Cthe_2924 adenylate kinases	40.16	9.70	24396.61	5 (5 0 0 0 0)
YP_001038376 Cthe_1965 alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	40.16	26.20	20791.48	7 (7 0 0 0 0)
YP_001037747 Cthe_1322 chaperone protein DnaK	40.15	4.80	65641.27	7 (7 0 0 0 0)
YP_001038605 Cthe_2210 3-isopropylmalate dehydratase, small subunit	38.19	29.60	20861.62	11 (10 1 0 0 0)
YP_001038372 Cthe_1961 Nucleotidyl transferase	38.17	4.90	90642.27	25 (9 5 5 2 4)
YP_001039296 Cthe_2905 Ribosomal protein L25/L23	34.19	13.70	13281.25	16 (13 2 1 0 0)
Cthe_3088	30.24	23.00	7948.42	19 (19 0 0 0 0)
YP_001039393 Cthe_3004 ferredoxin	30.22	8.70	54617.13	6 (6 0 0 0 0)
YP_001038631 Cthe_2236 flagellin-like protein	30.22	10.30	29549.11	21 (21 0 0 0 0)
YP_001038197 Cthe_1783 ribosomal protein L13	30.21	9.70	16763.18	9 (9 0 0 0 0)
YP_001038251 Cthe_1837 hypothetical protein	30.21	21.30	20171.65	9 (9 0 0 0 0)
YP_001037304 Cthe_0877 GTP-binding protein YchF	30.21	8.80	40506.09	12 (12 0 0 0 0)
YP_001039468 Cthe_3079 cellulosome anchoring protein, cohesin region	30.21	5.40	74924.76	7 (7 0 0 0 0)
YP_001036620 Cthe_0186 UDP-glucose 4-epimerase	30.20	10.70	39089.43	6 (6 0 0 0 0)
YP_001039469 Cthe_3080 cellulosome anchoring protein, cohesin region	30.20	7.40	48500.00	17 (17 0 0 0 0)
YP_001037122 Cthe_0694 spermidine synthase	30.20	11.60	31334.13	13 (13 0 0 0 0)
YP_001036538 Cthe_0104 riboflavin biosynthesis protein RibD	30.20	11.00	39237.94	3 (3 0 0 0 0)
YP_001036781 Cthe_0350 signal peptidase I	30.19	6.20	22212.42	3 (3 0 0 0 0)
YP_001037788 Cthe_1363 lipopolysaccharide biosynthesis	30.19	3.00	52051.22	9 (9 0 0 0 0)
YP_001037806 Cthe_1381 threonine synthase	30.19	4.80	55667.31	3 (3 0 0 0 0)
YP_001036742 Cthe_0311 excinuclease ABC, A subunit	30.18	4.90	104767.60	17 (17 0 0 0 0)
YP_001039049 Cthe_2657 histone-like DNA-binding protein	30.18	31.90	10081.48	12 (12 0 0 0 0)

YP_001037483 Cthe_1058 Glycine hydroxymethyltransferase	30.18	10.90	45379.45	5 (5 0 0 0 0)
YP_001037362 Cthe_0935 malonyl CoA-acyl carrier protein transacylase	30.18	13.50	33369.34	7 (7 0 0 0 0)
YP_001039615 Cthe_3227 copper amine oxidase-like protein	30.17	11.60	30240.66	13 (13 0 0 0 0)
YP_001037887 Cthe_1465 Redoxin	30.17	27.30	16775.50	6 (6 0 0 0 0)
YP_001039294 Cthe_2903 ribosomal protein L3	30.17	11.30	23551.69	14 (14 0 0 0 0)
YP_001037076 Cthe_0glutamyl-tRNA synthetase	30.17	4.20	63361.20	4 (4 0 0 0 0)
YP_001039551 Cthe_3163 Carbohydrate binding family 25	30.17	25.20	17688.08	13 (9 4 0 0 0)
YP_001037526 Cthe_1101 hypothetical protein	30.17	8.60	48018.17	12 (12 0 0 0 0)
YP_001037737 Cthe_1312 glycyl-tRNA synthetase	30.16	8.40	54044.84	6 (6 0 0 0 0)
YP_001038696 Cthe_2301 CRISPR-associated autoregulator, DevR family	30.16	4.40	32508.38	4 (4 0 0 0 0)
YP_001038Cthe_2252 thioesterase family protein	30.15	16.40	14553.56	4 (4 0 0 0 0)
YP_001038212 Cthe_1798 CoA-binding protein	30.15	10.10	24616.96	6 (6 0 0 0 0)
YP_001039111 Cthe_2719 NusG antitermination factor	30.15	13.00	20117.46	8 (8 0 0 0 0)
YP_001039392 Cthe_3003 hydrogenase, Fe-only	30.15	3.40	71733.04	7 (7 0 0 0 0)
YP_001039078 Cthe_2686 type IV pilus assembly protein PilM	30.14	6.50	41419.14	5 (5 0 0 0 0)
YP_001036887 Cthe_0456 protein of unknown function UPF0047	28.20	16.70	15799.08	3 (2 1 0 0 0)
YP_001037799 Cthe_1374 copper amine oxidase-like protein	28.16	10.30	30003.46	7 (3 4 0 0 0)
YP_001039545 Cthe_3157 pyruvate carboxyltransferase	28.15	7.40	52253.63	6 (5 1 0 0 0)
YP_001037675 Cthe_1250 phosphoribosylaminoimidazole carboxylase, catalytic subunit	26.22	15.00	18407.83	9 (8 0 1 0 0)
YP_001038177 Cthe_1763 ABC transporter related protein	26.16	4.90	27658.55	12 (2 2 5 3 0)
YP_001038441 Cthe_2039 UvrD/REP helicase	26.15	1.80	143255.10	15 (3 1 9 2 0)
YP_001038528 Cthe_2130 hypothetical protein	24.20	6.00	48488.74	5 (3 0 0 1 1)
YP_001036522 Cthe_0088 cell shape determining protein, MreB/Mrl family	24.16	8.80	36283.29	5 (2 2 1 0 0)
YP_001037460 Cthe_1035 DNA ligase, NAD-dependent	24.14	2.10	74795.39	4 (3 0 0 1 0)

YP_001036666 Cthe_0234 AMP-dependent synthetase and ligase	22.15	1.40	95168.88	4 (1 2 0 1 0)
YP_001037361 Cthe_0934 3-oxoacyl-(acyl-carrier-protein) reductase	20.24	9.30	26061.62	2 (2 0 0 0 0)
YP_001038215 Cthe_1801 ABC transporter related protein	20.22	5.90	28160.66	2 (2 0 0 0 0)
YP_001037296 Cthe_0869 hypothetical protein	20.22	9.50	25289.15	3 (3 0 0 0 0)
YP_001038540 Cthe_2143 hypothetical protein	20.22	23.90	11988.26	6 (6 0 0 0 0)
YP_001037770 Cthe_1345 adenine phosphoribosyltransferase	20.22	8.80	19068.20	11 (11 0 0 0 0)
YP_001039131 Cthe_2739 trigger factor	20.22	2.80	48813.22	4 (4 0 0 0 0)
YP_001039305 Cthe_2914 ribosomal protein L24	20.21	22.40	12807.98	11 (11 0 0 0 0)
YP_001038774 Cthe_2379 hypothetical protein	20.21	20.10	18996.12	9 (9 0 0 0 0)
YP_001036501 Cthe_0067 Silent information regulator protein Sir2	20.20	5.40	26808.54	5 (5 0 0 0 0)
YP_001037390 Cthe_0963 Dihydrodipicolinate reductase	20.20	6.70	27775.55	14 (14 0 0 0 0)
YP_001037532 Cthe_1107 type II secretion system protein E	20.20	5.00	88291.56	3 (3 0 0 0 0)
YP_001036510 Cthe_0076 hypothetical protein	20.20	23.30	13699.52	11 (11 0 0 0 0)
YP_001037849 Cthe_1425 Inorganic diphosphatase	20.20	3.90	71337.66	5 (5 0 0 0 0)
YP_001039279 Cthe_2888 imidazoleglycerol phosphate synthase, cyclase subunit	20.20	10.70	27690.37	8 (8 0 0 0 0)
YP_001038728 Cthe_2333 two component transcriptional regulator, winged helix family	20.20	10.30	26690.06	2 (2 0 0 0 0)
YP_001037454 Cthe_1029 phosphate acetyltransferase	20.19	8.90	38641.03	19 (19 0 0 0 0)
YP_001037500 Cthe_1075 protein of unknown function DUF881	20.19	11.20	26506.01	6 (6 0 0 0 0)
YP_001037530 Cthe_1105 type II secretion system protein	20.19	6.40	45131.49	6 (6 0 0 0 0)
YP_001039000 Cthe_2607 ATP synthase F1, gamma subunit	20.19	5.40	34138.72	2 (2 0 0 0 0)
YP_001037652 Cthe_1227 Thymidylate synthase	20.19	8.00	31913.28	2 (2 0 0 0 0)
YP_001037775 Cthe_1350 single-strand binding protein	20.18	20.10	15435.49	5 (5 0 0 0 0)
YP_001037140 Cthe_0712 cytidylate kinase	20.18	11.50	25239.33	4 (4 0 0 0 0)
YP_001038581 Cthe_2186 single-strand binding protein	20.18	14.50	15880.94	9 (9 0 0 0 0)
YP_001039593 Cthe_3205 hypothetical protein	20.18	6.80	25549.67	3 (3 0 0 0 0)

YP_001036842 Cthe_0411 hypothetical protein	20.17	6.40	39706.39	6 (6 0 0 0 0)
YP_001038254 Cthe_1840 cysteine synthase A	20.17	9.00	33377.69	4 (4 0 0 0 0)
YP_001038356 Cthe_1945 Thioredoxin-disulfide reductase	20.17	8.80	32476.80	3 (3 0 0 0 0)
YP_001039157 Cthe_2765 hypothetical protein	20.17	6.80	60495.86	2 (2 0 0 0 0)
YP_001038198 Cthe_1784 hypothetical protein	20.17	10.70	15136.84	4 (4 0 0 0 0)
YP_001038466 Cthe_2066 serine O-acetyltransferase	20.17	12.10	27382.19	5 (5 0 0 0 0)
YP_001039098 Cthe_2706 ABC transporter related protein	20.17	3.30	41203.11	8 (8 0 0 0 0)
YP_001038912 Cthe_2517 acetolactate synthase, small subunit	20.17	13.50	18609.04	6 (6 0 0 0 0)
YP_001037608 Cthe_1183 ATPase associated with various cellular activities, AAA_3	20.17	8.70	35005.28	4 (4 0 0 0 0)
YP_001037478 Cthe_1053 L-lactate dehydrogenase	20.16	6.30	35053.30	6 (6 0 0 0 0)
YP_001039493 Cthe_3105 exsB protein	20.16	12.20	24239.59	2 (2 0 0 0 0)
YP_001039133 Cthe_2741 ATP-dependent Clp protease, ATP-binding subunit ClpX	20.16	2.10	47861.47	3 (3 0 0 0 0)
YP_001039105 Cthe_2713 dihydroxy-acid dehydratase	20.16	2.90	58842.55	4 (4 0 0 0 0)
YP_001036722 Cthe_0290 Homoserine dehydrogenase	20.16	4.40	47059.77	2 (2 0 0 0 0)
YP_001036857 Cthe_0426 putative PAS/PAC sensor protein	20.16	5.20	62518.02	3 (3 0 0 0 0)
YP_001037084 Cthe_0656 type IV pilus assembly protein PilM	20.16	6.00	41929.30	3 (3 0 0 0 0)
YP_001036997 Cthe_0568 methionyl-tRNA formyltransferase	20.16	6.10	34385.84	8 (8 0 0 0 0)
YP_001037057 Cthe_0629 type II secretion system protein E	20.16	5.40	55014.36	2 (2 0 0 0 0)
YP_001036758 Cthe_0327 S-layer-like domain containing protein	20.16	1.60	94141.88	2 (2 0 0 0 0)
YP_001039278 Cthe_2887 phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase	20.16	11.40	25584.55	4 (4 0 0 0 0)
YP_001037042 Cthe_0614 pyruvate ferredoxin/flavodoxin oxidoreductase	20.16	11.50	21269.39	8 (8 0 0 0 0)
YP_001038475 Cthe_2075 protein of unknown function DUF28	20.15	7.40	26973.46	6 (6 0 0 0 0)
YP_001036740 Cthe_0309 excinuclease ABC, B subunit	20.15	2.30	75466.09	6 (6 0 0 0 0)
YP_001037330 Cthe_0903 protein-export membrane protein SecF	20.15	4.10	34990.80	13 (13 0 0 0 0)

YP_001036803 Cthe_0372 glutamate synthase (NADPH), homotetrameric	20.15	6.00	50484.87	2 (2 0 0 0 0)
YP_001036984 Cthe_0555 PpiC-type peptidyl-prolyl cis-trans isomerase	20.15	3.10	47318.00	2 (2 0 0 0 0)
YP_001039530 Cthe_3142 hypothetical protein	20.15	2.00	18432.98	7 (7 0 0 0 0)
YP_001036859 Cthe_0428 NADH dehydrogenase (ubiquinone), 24 kDa subunit	20.15	2.00	18338.32	5 (5 0 0 0 0)
YP_001037381 Cthe_0954 Uracil phosphoribosyltransferase	20.15	6.50	20402.04	2 (2 0 0 0 0)
YP_001036848 Cthe_0417 ribosomal protein S15	20.15	13.80	10378.75	4 (4 0 0 0 0)
YP_001037373 Cthe_0946 Phosphoglycerate mutase	20.15	5.20	27615.73	3 (3 0 0 0 0)
YP_001036579 Cthe_0145 metal dependent phosphohydrolase	20.14	2.00	20528.80	9 (5 0 3 1 0)
YP_001036806 Cthe_0375 GMP synthase, large subunit	20.14	2.20	57519.82	2 (2 0 0 0 0)
YP_001037041 Cthe_0613 thiamine pyrophosphate enzyme-like TPP-binding	20.14	2.00	65986.80	7 (7 0 0 0 0)
YP_001036736 Cthe_0305 DNA topoisomerase (ATP-hydrolyzing)	20.14	3.20	76560.08	4 (4 0 0 0 0)
YP_001038756 Cthe_2361 DNA gyrase, A subunit	20.14	1.50	93408.07	3 (3 0 0 0 0)
YP_001038563 Cthe_2168 Propeptide, PepSY amd peptidase M4	20.14	1.40	83073.66	2 (2 0 0 0 0)
YP_001037797 Cthe_1372 hypothetical protein	20.14	14.50	17235.16	3 (3 0 0 0 0)
YP_001039119 Cthe_2727 ribosomal protein S12	20.14	15.50	15415.67	9 (9 0 0 0 0)
YP_001036999 Cthe_0570 peptidase, membrane zinc metallopeptidase, putative	20.14	2.00	24747.35	3 (3 0 0 0 0)
YP_001037150 Cthe_0722 tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	20.14	5.80	40110.61	2 (2 0 0 0 0)
YP_001037587 Cthe_1162 glucosamine--fructose-6-phosphate aminotransferase, isomerizing	20.14	2.30	66987.34	2 (2 0 0 0 0)
YP_001037244 Cthe_0816 NAD(+) kinase	20.13	3.80	31649.52	2 (2 0 0 0 0)
YP_001037412 Cthe_0987 riboflavin biosynthesis protein RibF	20.13	5.80	35892.63	2 (2 0 0 0 0)
YP_001037392 Cthe_0965 ATP:corrinoid adenosyltransferase BtuR/CobO/CobP	20.13	13.10	19668.35	2 (2 0 0 0 0)
YP_001039513 Cthe_3125 heat shock protein Hsp20	20.13	8.10	17300.84	2 (2 0 0 0 0)
YP_001036792 Cthe_0361 hypothetical protein	20.13	5.30	21979.22	2 (2 0 0 0 0)
YP_001037220 Cthe_0792 Protein of unknown function UPF0001	20.13	3.40	26403.80	3 (3 0 0 0 0)

YP_001037672 Cthe_1247 phosphoribosylglycinamide formyltransferase	20.13	9.60	23334.32	2 (2 0 0 0 0)
YP_001038773 Cthe_2378 parB-like partition protein	18.17	7.30	32855.74	5 (4 1 0 0 0)
YP_001039566 Cthe_3178 protein of unknown function DUF342	18.16	2.60	59841.41	7 (1 6 0 0 0)
YP_001036743 Cthe_0312 ATPase AAA-2	18.16	2.00	90556.23	4 (1 3 0 0 0)
YP_001037102 Cthe_0674 NUDIX hydrolase	18.14	12.60	20289.62	2 (1 1 0 0 0)
YP_001038873 Cthe_2478 phage minor structural GP20	18.14	16.80	21347.13	4 (1 3 0 0 0)
YP_001038171 Cthe_1757 peptidase M23B	18.13	3.20	34708.10	4 (1 1 1 1 0)
YP_001036726 Cthe_0295 phosphoserine aminotransferase	16.16	4.30	41362.31	2 (1 0 1 0 0)
YP_001036827 Cthe_0396 ABC transporter related protein	16.14	2.00	63672.71	4 (3 0 1 0 0)
YP_001038417 Cthe_2010 hypothetical protein	14.17	3.90	48522.36	3 (0 1 2 0 0)
YP_001039589 Cthe_3201 CRISPR-associated protein, Csh1 family	14.14	2.00	71508.29	3 (0 0 2 0 1)
YP_001037319 Cthe_0892 protein of unknown function DUF34	14.13	2.00	40998.22	2 (0 1 1 0 0)
YP_001036980 Cthe_0551 AMP-dependent synthetase and ligase	12.16	2.50	63015.91	2 (1 0 0 0 1)
YP_001038341 Cthe_1930 carboxyl-terminal protease	12.13	1.60	55784.72	2 (0 1 0 1 0)
YP_001038957 Cthe_2563 hypothetical protein	12.13	1.90	68880.94	2 (0 1 0 1 0)
YP_001037196 Cthe_0768 nucleic acid binding protein, containing KH domain	10.23	23.70	8420.60	6 (6 0 0 0 0)
YP_001037328 Cthe_0901 pantoate--beta-alanine ligase	10.21	5.30	31643.65	5 (5 0 0 0 0)
YP_00103Cthe_0045 copper amine oxidase-like protein	10.20	8.20	30544.85	2 (2 0 0 0 0)
YP_001036578 Cthe_0144 preprotein translocase, SecG subunit	10.20	20.50	8797.02	7 (7 0 0 0 0)
YP_001039421 Cthe_3032 hypothetical protein	10.20	10.40	15321.67	4 (4 0 0 0 0)
YP_001039132 Cthe_2740 ATP-dependent Clp protease, proteolytic subunit ClpP	10.20	7.20	21587.15	1 (1 0 0 0 0)
YP_001038182 Cthe_1768 NifU-related domain containing protein	10.20	6.50	25046.62	4 (4 0 0 0 0)
YP_001037700 Cthe_1275 H+-ATPase subunit H	10.20	9.70	16877.05	3 (3 0 0 0 0)
YP_001037835 Cthe_1411 tryptophan synthase, alpha subunit	10.20	6.60	28167.31	1 (1 0 0 0 0)
YP_001039557 Cthe_3169 short-chain dehydrogenase/reductase SDR	10.19	5.20	27026.97	5 (5 0 0 0 0)

YP_001039481 Cthe_3093 Adenylosuccinate synthase	10.19	4.00	47011.32	3 (3 0 0 0 0)
YP_001036613 Cthe_0179 Argininosuccinate synthase	10.19	3.20	45593.73	2 (2 0 0 0 0)
YP_001037307 Cthe_0880 phospho-2-dehydro-3-deoxyheptonate aldolase	10.18	4.40	36848.65	7 (7 0 0 0 0)
YP_001037403 Cthe_0978 UDP-N-acetylmuramyl-tripeptide synthetases	10.18	3.50	53768.75	1 (1 0 0 0 0)
YP_001037650 Cthe_1225 translation initiation factor IF-3	10.18	7.90	18832.10	1 (1 0 0 0 0)
YP_001039434 Cthe_3045 FHA domain containing protein	10.18	9.20	19762.67	4 (4 0 0 0 0)
YP_0010374Cthe_1039 ribosomal protein S20	10.18	13.00	12988.50	5 (5 0 0 0 0)
YP_001037525 Cthe_1100 prepilin-type cleavage/methylation	10.18	2.00	19349.18	2 (2 0 0 0 0)
YP_001038087 Cthe_1671 Recombinase	10.17	3.10	59958.98	5 (4 0 1 0 0)
YP_001038119 Cthe_1705 hypothetical protein	10.17	2.00	65110.92	1 (1 0 0 0 0)
YP_001038632 Cthe_2237 flagellin-like protein	10.17	4.40	29413.22	5 (5 0 0 0 0)
YP_001039390 Cthe_3001 protein of unknown function DUF47	10.17	6.80	24092.60	1 (1 0 0 0 0)
YP_001037384 Cthe_0957 preprotein translocase, YajC subunit	10.17	14.10	10950.06	6 (6 0 0 0 0)
YP_001037619 Cthe_1194 hypothetical protein	10.17	4.20	45013.45	3 (3 0 0 0 0)
YP_001036511 Cthe_0077 hypothetical protein	10.17	8.70	13916.57	7 (7 0 0 0 0)
YP_001038259 Cthe_1845 Homoserine O-succinyltransferase	10.17	5.60	35848.21	1 (1 0 0 0 0)
YP_001039002 Cthe_2609 ATP synthase F1, epsilon subunit	10.16	7.40	15369.17	4 (4 0 0 0 0)
YP_00103Cthe_0019 regulatory protein, P-II family	10.16	13.00	11659.06	2 (2 0 0 0 0)
YP_001036891 Cthe_0460 DNA topoisomerase I	10.16	2.00	79261.14	4 (4 0 0 0 0)
YP_001039106 Cthe_2714 acetolactate synthase, large subunit, biosynthetic type	10.16	3.10	60696.15	1 (1 0 0 0 0)
YP_00103Cthe_0036 hybrid cluster protein	10.16	2.20	59404.46	2 (1 0 1 0 0)
YP_001038358 Cthe_1947 Redoxin	10.16	7.60	22258.21	3 (3 0 0 0 0)
YP_001037292 Cthe_0865 3-methyl-2-oxobutanoate dehydrogenase (ferredo	10.16	5.60	26948.71	1 (1 0 0 0 0)
YP_001036694 Cthe_0262 gamma-glutamyl phosphate reductase	10.16	2.80	47972.01	5 (5 0 0 0 0)
Cthe_2016	10.16	2.00	24309.88	1 (1 0 0 0 0)
YP_001038263 Cthe_1849 Heavy metal transport/detoxification protein	10.16	5.80	54009.14	1 (1 0 0 0 0)

YP_001039452 Cthe_3063 Acetyl xylan esterase	10.16	4.40	36006.41	1 (1 0 0 0 0)
YP_001039201 Cthe_2809 glycoside hydrolase, family 16	10.16	2.00	147694.60	2 (2 0 0 0 0)
YP_001037297 Cthe_0870 NADPH-dependent FMN reductase	10.16	2.00	21013.74	1 (1 0 0 0 0)
YP_001039482 Cthe_3094 glycosyl transferase, family 2	10.16	4.00	37622.82	1 (1 0 0 0 0)
YP_001038258 Cthe_1844 transcriptional regulator, BadM/Rrf2 family	10.16	5.90	16944.71	3 (3 0 0 0 0)
YP_00103Cthe_0048 copper amine oxidase-like protein	10.16	3.70	30342.87	4 (4 0 0 0 0)
YP_001037178 Cthe_0750 spermidine/putrescine ABC transporter ATPase subunit	10.16	2.80	39800.65	3 (3 0 0 0 0)
YP_001037447 Cthe_1022 Glycerol-3-phosphate dehydrogenase (NAD(P)(+))	10.16	2.00	36516.76	1 (1 0 0 0 0)
YP_001036769 Cthe_0338 NADH-quinone oxidoreductase, E subunit	10.16	7.30	18057.11	4 (4 0 0 0 0)
YP_001038799 Cthe_2404 transcriptional regulator, GntR family	10.15	5.70	26246.02	1 (1 0 0 0 0)
YP_001037003 Cthe_0574 serine/threonine protein kinase	10.15	2.00	78419.30	1 (1 0 0 0 0)
YP_001038384 Cthe_1973 Tetratricopeptide TPR_2	10.15	2.20	58122.05	1 (1 0 0 0 0)
YP_001038972 Cthe_2579 iron-containing alcohol dehydrogenase	10.15	4.90	40879.99	1 (1 0 0 0 0)
YP_001038423 Cthe_2018 hypothetical protein	10.15	3.20	48846.50	2 (2 0 0 0 0)
YP_001039517 Cthe_3129 Citrate transporter	10.15	2.00	41296.19	1 (1 0 0 0 0)
YP_001037822 Cthe_1398 cellulosome enzyme, dockerin type I	10.15	2.00	92306.83	1 (1 0 0 0 0)
YP_001038723 Cthe_2328 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	10.15	2.00	44355.70	2 (2 0 0 0 0)
YP_001038185 Cthe_1771 Rubrerythrin	10.15	12.00	20380.99	1 (1 0 0 0 0)
YP_001039476 Cthe_3087 response regulator receiver protein	10.15	2.00	30441.89	1 (1 0 0 0 0)
YP_001039207 Cthe_2815 lysyl-tRNA synthetase	10.15	2.00	61656.02	1 (1 0 0 0 0)
YP_001037656 Cthe_1231 Serine-type D-Ala-D-Ala carboxypeptidase	10.15	2.00	47954.54	1 (1 0 0 0 0)
YP_001038211 Cthe_1797 3-phosphoshikimate 1-carboxyvinyltransferase	10.15	4.00	45504.41	2 (2 0 0 0 0)
YP_001039289 Cthe_2898 anti-sigma-factor antagonist	10.15	10.60	11802.36	2 (2 0 0 0 0)
YP_001036889 Cthe_0458 exodeoxyribonuclease III Xth	10.15	4.40	29295.52	2 (2 0 0 0 0)

YP_001037292 Cthe_0865 3-methyl-2-oxobutanoate dehydrogenase (ferredo	10.15	5.60	26948.71	1 (1 0 0 0 0)
YP_001037792 Cthe_1367 PHP-like protein	10.15	4.30	29478.46	5 (0 1 0 2 2)
YP_001038281 Cthe_1867 carbamoyl-phosphate synthase, small subunit	10.15	4.80	39325.05	1 (1 0 0 0 0)
YP_00103Cthe_0055 hypothetical protein	10.15	3.90	20568.21	1 (1 0 0 0 0)
YP_001039381 Cthe_2992 RNA polymerase, sigma-24 subunit, ECF subfamily	10.15	4.60	20428.45	6 (4 2 0 0 0)
YP_001039004 Cthe_2611 Fibronectin, type III	10.15	0.70	183582.80	1 (1 0 0 0 0)
YP_001039455 Cthe_3066 ABC transporter related protein	10.15	5.00	27035.17	1 (1 0 0 0 0)
YP_001039085 Cthe_2693 hypothetical protein	10.15	5.50	17809.54	2 (2 0 0 0 0)
YP_001036831 Cthe_0400 hypothetical protein	10.15	3.20	39513.86	1 (1 0 0 0 0)
YP_001037105 Cthe_0677 phosphopentomutase	10.15	2.30	42561.62	3 (3 0 0 0 0)
YP_00103Cthe_0056 Ig-like, group 2	10.15	2.00	493538.80	3 (1 2 0 0 0)
YP_001036503 Cthe_0069 Aspartate--ammonia ligase	10.15	2.00	39008.13	1 (1 0 0 0 0)
YP_001037444 Cthe_1019 binding-protein-dependent transport systems inner membrane component	10.14	2.00	37355.85	3 (3 0 0 0 0)
YP_001037222 Cthe_0794 aluminium resistance protein	10.14	2.10	47156.19	2 (2 0 0 0 0)
YP_001036988 Cthe_0559 single-strand binding protein/Primosomal replication protein n	10.14	4.20	27479.06	2 (2 0 0 0 0)
YP_001037259 Cthe_0831 Polyprenyl synthetase	10.14	3.40	32474.66	1 (1 0 0 0 0)
YP_001037515 Cthe_1090 hypothetical protein	10.14	5.00	28844.71	1 (1 0 0 0 0)
YP_001039303 Cthe_2912 ribosomal protein S17	10.14	10.60	10034.51	2 (2 0 0 0 0)
YP_001038815 Cthe_2420 HD superfamily phosphohydrolases-like protein	10.14	2.00	70494.38	4 (4 0 0 0 0)
YP_001037498 Cthe_1073 protein of unknown function DUF1429	10.14	14.30	12049.66	5 (1 3 1 0 0)
YP_001037049 Cthe_0621 putative translation initiation factor, aIF-2BI family	10.14	2.60	37833.65	1 (1 0 0 0 0)
YP_001037369 Cthe_0942 MiaB-like tRNA modifying enzyme YliG	10.14	2.00	51752.77	1 (1 0 0 0 0)
YP_001038580 Cthe_2185 ribosomal protein S18	10.14	11.70	10856.08	2 (2 0 0 0 0)
YP_001038282 Cthe_1868 carbamoyl-phosphate synthase, large subunit	10.14	2.00	117506.10	1 (1 0 0 0 0)
YP_001037247 Cthe_0819 ABC transporter related protein	10.14	5.20	32076.80	1 (1 0 0 0 0)

YP_001039298 Cthe_2907 ribosomal protein S19	10.14	8.50	10589.65	3 (3 0 0 0 0)
YP_001038460 Cthe_2060 RNA methyltransferase, TrmH family, group 3	10.14	4.70	30197.23	1 (1 0 0 0 0)
YP_001036977 Cthe_0548 ABC transporter related protein	10.14	4.60	28795.18	1 (1 0 0 0 0)
YP_001039314 Cthe_2923 preprotein translocase, SecY subunit	10.14	2.00	47339.89	4 (4 0 0 0 0)
YP_001036585 Cthe_0151 protein of unknown function DUF965	10.14	2.00	10363.24	6 (6 0 0 0 0)
YP_001038468 Cthe_2068 8-oxoguanine DNA glycosylase-like protein	10.14	4.40	34247.96	2 (2 0 0 0 0)
YP_001038971 Cthe_2578 Ppx/GppA phosphatase	10.14	3.80	57132.26	1 (1 0 0 0 0)
YP_001036751 Cthe_0320 hypothetical protein	10.14	2.00	55540.01	1 (1 0 0 0 0)
YP_001038906 Cthe_2511 MscS Mechanosensitive ion channel	10.14	2.00	31334.22	1 (1 0 0 0 0)
YP_001036534 Cthe_0100 hypothetical protein	10.14	3.50	30310.38	1 (1 0 0 0 0)
YP_001037666 Cthe_1241 nicotinate (nicotinamide) nucleotide adenylyltransferase	10.14	5.30	23673.40	2 (2 0 0 0 0)
YP_001038266 Cthe_1852 Hsp33 protein	10.14	5.10	32212.42	1 (1 0 0 0 0)
YP_001037378 Cthe_0951 orotidine 5'-phosphate decarboxylase	10.14	4.50	35151.93	2 (2 0 0 0 0)
YP_001038634 Cthe_2239 carbon storage regulator, CsrA	10.14	2.00	8649.78	1 (1 0 0 0 0)
YP_001036994 Cthe_0565 hypothetical protein	10.14	8.60	15597.69	1 (1 0 0 0 0)
YP_001037614 Cthe_1189 ABC transporter related protein	10.14	3.10	36518.56	4 (0 1 1 0 2)
YP_001036983 Cthe_0554 phosphoribosylformylglycinamidine synthase	10.14	0.80	139261.80	1 (1 0 0 0 0)
YP_001037685 Cthe_1260 4-hydroxybenzoyl-CoA thioesterase	10.14	8.00	16215.26	1 (1 0 0 0 0)
YP_001039491 Cthe_3103 2-hydroxyglutaryl-CoA dehydratase, D-component	10.14	2.00	37446.18	2 (2 0 0 0 0)
YP_001039261 Cthe_2870 protein of unknown function DUF21	10.14	2.00	48670.55	1 (1 0 0 0 0)
YP_001039020 Cthe_2628 SpoVG	10.14	11.70	10462.28	3 (3 0 0 0 0)
YP_001037212 Cthe_0784 hypothetical protein	10.14	3.40	33757.39	2 (2 0 0 0 0)
YP_001037610 Cthe_1185 hypothetical protein	10.14	3.10	50416.20	1 (1 0 0 0 0)
YP_001037553 Cthe_1128 hypothetical protein	10.14	2.00	14799.74	2 (2 0 0 0 0)
YP_001037757 Cthe_1332 histidyl-tRNA synthetase	10.14	2.00	46805.88	1 (1 0 0 0 0)

YP_001038779 Cthe_2384 S-layer-like domain containing protein	10.14	6.30	21039.21	1 (1 0 0 0 0)
YP_001036910 Cthe_0479 response regulator receiver protein	10.14	2.00	13205.91	1 (1 0 0 0 0)
YP_001038148 Cthe_1734 phage / plasmid primase, P4 family	10.14	2.40	72976.33	1 (1 0 0 0 0)
YP_001037265 Cthe_0837 protein of unknown function DUF322	10.14	8.20	14220.37	2 (2 0 0 0 0)
YP_001039100 Cthe_2708 hypothetical protein	10.14	2.00	53616.59	3 (3 0 0 0 0)
YP_001036853 Cthe_0422 CoA-binding protein	10.13	5.70	24791.90	6 (6 0 0 0 0)
YP_001038771 Cthe_2376 DNA gyrase, B subunit	10.13	2.00	72446.46	1 (1 0 0 0 0)
YP_00103Cthe_0054 protein of unknown function DUF187	10.13	1.40	82321.66	1 (1 0 0 0 0)
YP_001037668 Cthe_1243 GCN5-related N-acetyltransferase	10.13	2.00	17625.19	1 (1 0 0 0 0)
YP_001037281 Cthe_0854 shikimate 5-dehydrogenase	10.13	4.00	33301.66	2 (2 0 0 0 0)
YP_001037443 Cthe_1018 binding-protein-dependent transport systems inner membrane component	10.13	2.00	31904.56	3 (3 0 0 0 0)
YP_001037301 Cthe_0874 glutamine amidotransferase of anthranilate synthase	10.13	2.00	22686.02	2 (2 0 0 0 0)
YP_001036630 Cthe_0196 glutamine synthetase, catalytic region	10.13	2.30	78506.49	1 (1 0 0 0 0)
YP_001039088 Cthe_2696 carbohydrate kinase, YjeF related protein	10.13	1.70	54543.56	2 (2 0 0 0 0)
YP_001036540 Cthe_0106 GTP cyclohydrolase II	10.13	3.10	45817.61	2 (1 1 0 0 0)
YP_001037944 Cthe_1523 glycosyl transferase, group 1	10.13	2.00	71313.41	1 (1 0 0 0 0)
YP_001037993 Cthe_1574 hypothetical protein	10.13	5.70	27775.38	2 (2 0 0 0 0)
YP_001038Cthe_2246 anti-sigma-28 factor, FlgM	10.13	2.00	10775.75	2 (2 0 0 0 0)
YP_001037803 Cthe_1378 response regulator receiver modulated metal dependent phosphohydrolase	10.13	2.00	58100.63	1 (1 0 0 0 0)
YP_001037183 Cthe_0755 aminotransferase, class I and II	10.13	3.00	42895.99	1 (1 0 0 0 0)
YP_001037440 Cthe_1015 hypothetical protein	10.13	9.30	14662.84	1 (1 0 0 0 0)
YP_001039017 Cthe_2625 beta-hydroxyacyl-(acyl-carrier-protein) dehydratase FabZ	10.13	6.20	15911.54	1 (1 0 0 0 0)
YP_001037472 Cthe_1047 hypothetical protein	10.13	8.30	20674.68	1 (1 0 0 0 0)

YP_001037322 Cthe_0895 RNA polymerase, sigma 38 subunit, RpoS	10.13	3.60	41084.67	1 (1 0 0 0 0)
YP_001037240 Cthe_0812 response regulator receiver protein	10.13	2.00	30139.98	2 (2 0 0 0 0)
YP_001039277 Cthe_2886 imidazole glycerol phosphate synthase, glutamine amidotransferase subunit	10.13	4.40	22409.65	1 (1 0 0 0 0)
YP_001037215 Cthe_0787 isoleucyl-tRNA synthetase	10.13	1.30	106730.50	1 (1 0 0 0 0)
YP_001038962 Cthe_2568 hypothetical protein	10.13	2.00	25385.09	1 (1 0 0 0 0)
YP_001037009 Cthe_0580 aminotransferase, class I and II	10.13	2.80	44743.46	1 (1 0 0 0 0)
YP_001038495 Cthe_2095 hydrolase, TatD family	10.13	3.10	29094.93	1 (1 0 0 0 0)
YP_001037794 Cthe_1369 histidine triad (HIT) protein	10.13	10.50	12631.70	2 (2 0 0 0 0)
YP_001037480 Cthe_1055 protein of unknown function DUF58	10.13	2.00	47172.58	1 (1 0 0 0 0)
YP_001038243 Cthe_1829 Chromosome segregation ATPases-like protein	10.13	2.00	124587.60	1 (1 0 0 0 0)
YP_001039048 Cthe_2656 MazG family protein	10.13	4.60	30447.53	1 (1 0 0 0 0)
YP_001036960 Cthe_0531 hypothetical protein	10.13	2.00	87439.81	1 (1 0 0 0 0)
YP_001036770 Cthe_0339 histidine kinase	10.13	2.00	20975.22	1 (1 0 0 0 0)
YP_001036Cthe_0208 single-stranded-DNA-specific exonuclease RecJ	10.13	1.20	86561.65	2 (2 0 0 0 0)
YP_001038272 Cthe_1858 peptidase M23B	10.13	2.00	41616.22	1 (1 0 0 0 0)
YP_001038604 Cthe_2209 3-isopropylmalate dehydrogenase	10.13	2.80	39066.03	1 (1 0 0 0 0)
YP_001037073 Cthe_0homocysteine S-methyltransferase	10.13	1.10	87269.55	1 (1 0 0 0 0)
YP_001039218 Cthe_2826 hypothetical protein	10.13	2.70	48947.83	1 (1 0 0 0 0)
YP_001038328 Cthe_1916 two component transcriptional regulator, winged helix family	10.13	7.00	26245.90	1 (1 0 0 0 0)
YP_001037989 Cthe_1570 extracellular solute-binding protein, family 3	10.13	3.10	39028.01	1 (1 0 0 0 0)
YP_001038955 Cthe_2561 CDP-glucose 4,6-dehydratase	10.13	4.70	43318.39	1 (1 0 0 0 0)
YP_001036737 Cthe_0306 DNA gyrase/topoisomerase IV, subunit A	10.13	1.10	83209.58	1 (1 0 0 0 0)
YP_001038121 Cthe_1707 phage minor structural protein	10.13	2.00	71351.81	1 (1 0 0 0 0)
YP_001037907 Cthe_1485 beta-lactamase-like protein	10.13	3.40	47381.80	1 (1 0 0 0 0)
YP_001037740 Cthe_1315 Guanylate kinase	10.13	4.90	23239.91	1 (1 0 0 0 0)

YP_001038725 Cthe_2330 hypothetical protein	10.13	2.00	57264.16	1 (1 0 0 0 0)
YP_001038261 Cthe_1847 two component transcriptional regulator, winged helix family	10.13	3.40	26565.03	1 (1 0 0 0 0)
YP_001036791 Cthe_0360 thioredoxin	10.13	2.00	12173.26	1 (1 0 0 0 0)
YP_001036533 Cthe_0099 5,10-methylenetetrahydrofolate reductase	10.13	3.80	32764.93	1 (1 0 0 0 0)
YP_001037411 Cthe_0986 peptidase M16-like protein	10.13	2.00	49371.58	1 (1 0 0 0 0)
YP_001037184 Cthe_0756 Membrane dipeptidase	10.13	2.00	34538.65	1 (1 0 0 0 0)
YP_001037424 Cthe_0999 1-deoxy-D-xylulose 5-phosphate reductoisomerase	10.13	2.40	41864.34	1 (1 0 0 0 0)
YP_001038996 Cthe_2603 ATP synthase F0, C subunit	10.13	2.00	7218.07	1 (1 0 0 0 0)
YP_001038472 Cthe_2072 hypothetical protein	10.13	4.40	27549.79	1 (1 0 0 0 0)
YP_001038128 Cthe_1714 hypothetical protein	10.13	2.00	12935.55	1 (1 0 0 0 0)
YP_001039475 Cthe_3086 3-phosphoshikimate 1-carboxyvinyltransferase	10.13	2.00	45886.62	1 (1 0 0 0 0)
YP_001037808 Cthe_1383 Tetratricopeptide TPR_2	10.13	3.10	45162.93	1 (1 0 0 0 0)
YP_001036580 Cthe_0146 ribonuclease R	10.13	2.00	86939.25	1 (1 0 0 0 0)
YP_001039141 Cthe_2749 hypothetical protein	10.13	2.80	59077.16	1 (1 0 0 0 0)
YP_001038635 Cthe_2240 protein of unknown function DUF180	10.13	6.90	18155.34	1 (1 0 0 0 0)
YP_001037134 Cthe_0706 transcriptional regulator, RpiR family	10.13	4.00	33631.35	1 (1 0 0 0 0)
YP_001036596 Cthe_0162 ribosomal protein L27	10.13	14.10	10032.62	1 (1 0 0 0 0)
YP_001037100 Cthe_0672 pyrroline-5-carboxylate reductase	10.13	4.10	29459.34	1 (1 0 0 0 0)
YP_001036707 Cthe_0275 glycosyltransferase 36	10.12	1.20	92790.86	2 (2 0 0 0 0)
YP_001037689 Cthe_12DNA polymerase III, alpha subunit	10.12	2.00	134313.30	1 (1 0 0 0 0)
YP_001037996 Cthe_1577 hypothetical protein	10.12	2.00	58166.91	1 (1 0 0 0 0)
YP_001039310 Cthe_2919 ribosomal protein L18	10.12	8.20	13690.56	1 (1 0 0 0 0)
YP_001037874 Cthe_1451 GCN5-related N-acetyltransferase	10.12	5.10	18051.00	4 (2 2 0 0 0)
YP_001038663 Cthe_2268 Sodium-transporting two-sector ATPase	10.11	2.80	50691.77	1 (1 0 0 0 0)
YP_001039425 Cthe_3036 methyl-accepting chemotaxis sensory transducer	8.18	2.00	62462.38	5 (0 5 0 0 0)
YP_00103 Cthe_0003 ankyrin repeat protein	8.17	3.60	35561.51	32 (0 31 0 1 0)

YP_001037749 Cthe_1324 heat-inducible transcription repressor HrcA	8.17	4.60	39225.57	4 (0 1 1 0 2)
YP_001039432 Cthe_3043 hypothetical protein	8.16	2.00	24448.82	1 (0 1 0 0 0)
YP_001039035 Cthe_2Nucleotidyl transferase	8.16	2.00	40256.01	1 (0 1 0 0 0)
YP_001037053 Cthe_0625 glycoside hydrolase, family 9	8.16	2.80	79760.16	1 (0 1 0 0 0)
YP_001039373 Cthe_2983 methyl-accepting chemotaxis sensory transducer	8.16	2.00	61326.02	6 (0 5 1 0 0)
YP_001036683 Cthe_0251 transglutaminase-like protein	8.16	2.00	96666.69	1 (0 1 0 0 0)
YP_001038226 Cthe_1812 Urease accessory protein UreD	8.16	10.70	29686.90	1 (0 1 0 0 0)
YP_00103Cthe_0020 biotin synthase	8.15	5.00	35720.15	1 (0 1 0 0 0)
YP_001038Cthe_2254 hypoxanthine phosphoribosyltransferase	8.15	2.00	20691.94	1 (0 1 0 0 0)
YP_001038671 Cthe_2276 AAA ATPase, central region	8.15	2.00	43847.11	2 (0 2 0 0 0)
YP_001038917 Cthe_2522 membrane associated protein	8.14	2.00	61939.47	2 (0 1 1 0 0)
YP_001039525 Cthe_3137 hypothetical protein	8.14	5.70	21252.79	5 (0 4 1 0 0)
Cthe_2984	8.14	2.00	22648.09	1 (0 1 0 0 0)
YP_001037913 Cthe_1492 NAD(P)H dehydrogenase (quinone)	8.14	2.00	61292.02	4 (0 4 0 0 0)
YP_001036667 Cthe_0235 Glutaredoxin-like protein, YruB-family	8.14	2.00	8440.23	1 (0 1 0 0 0)
YP_001039474 Cthe_3085 response regulator receiver protein	8.14	2.00	35092.44	1 (0 1 0 0 0)
YP_001038470 Cthe_2070 hypothetical protein	8.14	2.00	32039.12	3 (0 3 0 0 0)
YP_001037566 Cthe_1141 hypothetical protein	8.14	0.70	157008.00	3 (0 3 0 0 0)
YP_001037691 Cthe_1266 methyl-accepting chemotaxis sensory transducer	8.14	2.00	91886.56	1 (0 1 0 0 0)
YP_001037223 Cthe_0795 alpha amylase, catalytic region	8.14	1.90	67491.74	5 (0 5 0 0 0)
YP_001036962 Cthe_0533 Radical SAM	8.14	2.00	57598.87	1 (0 1 0 0 0)
YP_001037479 Cthe_1054 ATPase associated with various cellular activities, AAA_3	8.14	2.00	35777.95	3 (0 3 0 0 0)
YP_001037456 Cthe_1031 glutamyl-tRNA(Gln) amidotransferase, B subunit	8.14	2.00	54020.93	2 (0 1 0 1 0)
YP_001037114 Cthe_0686 tryptophanyl-tRNA synthetase	8.14	4.60	37373.24	1 (0 1 0 0 0)
YP_001036763 Cthe_0332 phosphoribulokinase/uridine kinase	8.14	2.00	64507.49	2 (0 2 0 0 0)
YP_0010393Cthe_2974 hypothetical protein	8.14	3.50	45212.02	2 (0 2 0 0 0)

YP_001037490 Cthe_1065 type IV pilus assembly PilZ	8.14	2.00	25076.02	2 (0 2 0 0 0)
YP_001038806 Cthe_2411 metallophosphoesterase	8.14	2.00	43678.01	3 (0 3 0 0 0)
YP_001039042 Cthe_2650 polysaccharide biosynthesis protein CapD	8.14	2.50	68185.34	2 (0 2 0 0 0)
YP_001038729 Cthe_2334 polysaccharide biosynthesis protein CapD	8.14	2.60	68628.90	3 (0 2 1 0 0)
YP_001038769 Cthe_2374 DNA replication and repair protein RecF	8.13	4.30	43059.54	1 (0 1 0 0 0)
YP_001038278 Cthe_18acetylglutamate kinase	8.13	2.00	33029.62	1 (0 1 0 0 0)
YP_001037133 Cthe_0705 hypothetical protein	8.13	15.40	9275.10	1 (0 1 0 0 0)
YP_001038019 Cthe_1600 two component transcriptional regulator, winged helix family	8.13	2.00	25125.33	1 (0 1 0 0 0)
YP_001037098 Cthe_0670 spore germination protein	8.13	3.00	40973.39	1 (0 1 0 0 0)
YP_001037748 Cthe_1323 GrpE protein	8.13	5.30	25432.41	1 (0 1 0 0 0)
YP_001038153 Cthe_1739 SNF2-related protein	8.13	2.00	44230.29	2 (0 2 0 0 0)
YP_001036597 Cthe_0163 GTP1/OBG subdomain containing protein	8.13	2.10	46746.25	2 (1 1 0 0 0)
YP_001037337 Cthe_0910 extracellular solute-binding protein, family 5	8.13	2.20	63116.98	1 (0 1 0 0 0)
YP_001038239 Cthe_1825 multi-sensor hybrid histidine kinase	8.13	2.00	106314.10	1 (0 1 0 0 0)
YP_001038120 Cthe_1706 hypothetical protein	8.13	2.90	42932.78	1 (0 1 0 0 0)
YP_001038684 Cthe_2289 hypothetical protein	8.13	5.00	27862.32	3 (0 3 0 0 0)
YP_001037002 Cthe_0573 protein serine/threonine phosphatases	8.13	2.00	27359.65	1 (0 1 0 0 0)
YP_001038480 Cthe_2080 NUDIX hydrolase	8.13	2.00	22258.02	1 (0 1 0 0 0)
YP_001038Cthe_2247 regulatory protein, MerR	8.13	2.00	16866.37	2 (0 2 0 0 0)
YP_001038518 Cthe_2118 ABC transporter related protein	8.13	3.60	28509.21	1 (0 1 0 0 0)
YP_001037720 Cthe_1295 5-formyltetrahydrofolate cyclo-ligase	8.13	2.00	21804.43	1 (0 1 0 0 0)
YP_001039540 Cthe_3152 cobalamin 5'-phosphate synthase	8.13	4.40	26592.71	1 (0 1 0 0 0)
YP_0010387Cthe_2369 ribonuclease P protein component	8.13	2.00	15109.09	1 (0 1 0 0 0)
YP_001037714 Cthe_1289 hypothetical protein	8.13	17.20	7786.11	1 (0 1 0 0 0)
YP_001037771 Cthe_1346 single-stranded-DNA-specific exonuclease RecJ	8.13	2.00	93542.28	1 (0 1 0 0 0)

YP_001038507 Cthe_2107 thymidylate kinase	8.13	2.00	26517.36	1 (0 1 0 0 0)
YP_001038493 Cthe_2093 hypothetical protein	8.13	4.30	38749.77	1 (0 1 0 0 0)
YP_001037097 Cthe_0669 spore germination B3 GerAC like	8.13	3.40	48730.63	1 (0 1 0 0 0)
YP_001039351 Cthe_2961 extracellular solute-binding protein, family 5	8.13	2.00	60372.57	1 (0 1 0 0 0)
YP_001038026 Cthe_1609 Recombinase	8.13	2.00	60124.05	1 (0 1 0 0 0)
YP_001036941 Cthe_0511 histidine kinase	8.13	2.00	50721.29	3 (0 3 0 0 0)
Cthe_1395	8.13	23.80	7466.82	1 (0 1 0 0 0)
YP_001036921 Cthe_0490 CheA signal transduction histidine kinases	8.13	2.00	77010.76	1 (0 1 0 0 0)
YP_001036963 Cthe_0534 ABC-type bacteriocin transporter	8.13	2.00	80764.15	1 (0 1 0 0 0)
YP_001038678 Cthe_2283 methyl-accepting chemotaxis sensory transducer	6.17	2.00	154468.80	14 (1 5 4 3 1)
YP_001039371 Cthe_2981 hypothetical protein	6.16	5.50	19707.33	12 (0 6 6 0 0)
YP_001037948 Cthe_1527 periplasmic sensor signal transduction histidine kinase	6.16	2.90	40255.13	12 (0 2 10 0 0)
YP_001038323 Cthe_1911 Carbohydrate binding family 6	6.16	2.00	146588.20	2 (0 0 2 0 0)
Cthe_2134	6.15	6.30	20822.68	6 (0 1 2 2 1)
YP_001037234 Cthe_0806 PAS/PAC sensor hybrid histidine kinase	6.15	1.30	102250.20	3 (0 0 1 1 1)
YP_001036951 Cthe_0521 helicase-like protein	6.15	2.00	126703.70	1 (0 0 1 0 0)
YP_001038436 Cthe_2032 hypothetical protein	6.14	0.60	223041.40	4 (0 2 2 0 0)
YP_001037765 Cthe_1340 hypothetical protein	6.14	4.30	41295.08	2 (0 0 1 0 1)
YP_001039552 Cthe_31two component transcriptional regulator, AraC family	6.14	2.00	62914.19	2 (0 0 2 0 0)
YP_001037963 Cthe_1543 aspartyl-tRNA synthetase	6.14	2.00	66454.89	3 (0 1 2 0 0)
YP_001038781 Cthe_2386 VanW	6.14	2.00	62732.77	2 (0 1 1 0 0)
YP_001036874 Cthe_0443 protein of unknown function DUF1290	6.14	2.00	13788.51	1 (0 0 1 0 0)
YP_001038846 Cthe_2451 SNF2-related protein	6.14	2.00	126155.20	1 (0 0 1 0 0)
YP_001037914 Cthe_1493 transcriptional regulator, PadR-like family	6.14	6.90	24163.14	2 (0 0 0 1 1)
YP_001036668 Cthe_0236 aldo/keto reductase	6.13	2.00	42386.54	1 (0 0 1 0 0)
YP_001039240 Cthe_2848 phage major capsid protein, HK97	6.13	3.20	48301.61	2 (0 0 1 1 0)
YP_001036582 Cthe_0148 transposase, mutator type	6.13	2.00	47187.59	1 (0 0 1 0 0)

YP_001039087 Cthe_2695 hypothetical protein	6.13	4.80	25498.29	2 (0 0 2 0 0)
YP_001038367 Cthe_1956 Tagatose-6-phosphate kinase	6.13	2.00	33741.50	1 (0 0 1 0 0)
YP_001038154 Cthe_1740 BRO-like protein	6.13	2.00	29829.33	1 (0 0 1 0 0)
YP_001037931 Cthe_1510 short-chain dehydrogenase/reductase SDR	6.13	2.00	27380.86	2 (0 1 1 0 0)
YP_001036759 Cthe_0328 peptide chain release factor 3	6.13	2.00	60603.03	1 (0 0 1 0 0)
YP_001038442 Cthe_2040 ATP-dependent exonuclease synthesis protein A	6.13	2.00	130363.10	1 (0 0 1 0 0)
YP_001039548 Cthe_3160 putative RNA methylase, NOL1/NOP2/sun family	6.13	2.00	51578.11	1 (0 0 1 0 0)
YP_001038290 Cthe_1876 Tn7-like transposition protein A	6.13	3.90	32890.93	1 (0 0 1 0 0)
YP_001038114 Cthe_1700 Recombinase	6.13	2.00	51453.39	1 (0 0 1 0 0)
YP_001038103 Cthe_1688 Radical SAM	6.13	2.00	36818.96	1 (0 0 1 0 0)
YP_001038914 Cthe_2519 2-isopropylmalate synthase/homocitrate synthase family protein	6.11	2.00	59595.77	2 (0 1 1 0 0)
YP_001037586 Cthe_1161 hypothetical protein	4.15	2.00	93283.76	2 (0 0 1 1 0)
YP_001037497 Cthe_1072 putative stage IV sporulation YqfD	4.14	4.50	45042.68	1 (0 0 0 1 0)
YP_001037883 Cthe_1461 FAD dependent oxidoreductase	4.14	2.10	58448.48	8 (0 0 2 5 1)
YP_001038149 Cthe_1735 phage DNA polymerase	4.14	2.00	66028.44	3 (0 0 1 2 0)
YP_001036656 Cthe_0222 hypothetical protein	4.14	2.00	9503.24	2 (0 1 0 1 0)
YP_001036798 Cthe_0367 NADP oxidoreductase, coenzyme F420-dependent	4.14	5.90	31848.73	1 (0 0 0 1 0)
YP_001037276 Cthe_0848 peptidase M24	4.14	2.00	40148.85	2 (0 0 0 2 0)
YP_001038731 Cthe_2336 glycosyl transferase, group 1	4.14	2.00	43607.84	4 (0 0 1 3 0)
YP_001037Cthe_1221 glycosyltransferase 36	4.14	0.60	333880.10	1 (0 0 0 1 0)
YP_001038969 Cthe_2576 hypothetical protein	4.14	2.00	50069.20	1 (0 0 0 1 0)
YP_001036896 Cthe_0465 flagellar M-ring protein FliF	4.13	1.70	57786.86	1 (0 0 0 1 0)
YP_001037024 Cthe_0596 GTP-binding protein	4.13	2.00	24130.91	1 (0 0 0 1 0)
YP_001036979 Cthe_0550 heat shock protein Hsp90	4.13	2.00	73792.48	3 (0 0 0 1 2)
YP_001039158 Cthe_2766 hypothetical protein	4.13	14.80	8968.65	2 (0 0 0 2 0)
YP_001036834 Cthe_0403 DNA-directed RNA polymerase sigma factor	4.13	4.00	32431.13	2 (0 0 0 1 1)

YP_001036563 Cthe_0129 metal dependent phosphohydrolase	4.13	2.00	58142.20	5 (0 0 0 3 2)
YP_001036948 Cthe_0518 type III restriction enzyme, res subunit	4.13	2.00	114120.70	3 (0 1 1 1 0)
YP_001039128 Cthe_2736 phosphoenolpyruvate-protein phosphotransferase	4.13	2.00	63495.36	1 (0 0 0 1 0)
YP_001036745 Cthe_0314 glycosyltransferase 28-like protein	4.13	2.00	44624.88	1 (0 0 0 1 0)
YP_001037850 Cthe_1426 AzlC-like protein	4.13	2.00	26220.13	2 (0 0 0 2 0)
YP_001037466 Cthe_1041 UDP-N-acetylmuramoylalanine--D-glutamate ligase	4.13	2.40	51752.06	1 (0 0 0 1 0)
YP_001038951 Cthe_2557 glycosyl transferase, family 2	4.13	2.00	57454.98	1 (0 0 0 1 0)
YP_001036895 Cthe_04flagellar hook-basal body complex subunit FliE	4.13	2.00	11494.90	1 (0 0 0 1 0)
YP_001038624 Cthe_2229 N-acetylneuraminate synthase	2.16	2.00	39121.83	3 (0 0 0 2 1)
YP_001036993 Cthe_05Trans-hexaprenyltranstransferase	2.14	3.50	35809.44	2 (0 0 0 0 2)
YP_001039574 Cthe_3186 Radical SAM	2.14	1.40	67315.30	1 (0 0 0 0 1)
YP_001038240 Cthe_1826 response regulator receiver sensor signal transduction histidine kinase	2.13	2.00	41724.54	4 (0 0 2 0 2)
YP_001037421 Cthe_0996 DNA polymerase III, alpha subunit	2.13	0.70	163280.50	1 (0 0 0 0 1)
YP_001038178 Cthe_17Outer membrane protein-like protein	2.13	2.00	41742.34	1 (0 0 0 0 1)
YP_001038758 Cthe_2363 methyltransferase GidB	2.13	2.00	27335.71	1 (0 0 0 0 1)
YP_001036635 Cthe_0201 glutamate synthase, alpha subunit-like protein	2.13	4.20	26583.20	2 (0 1 0 0 1)
YP_001039010 Cthe_2617 peptidase M23B	2.13	2.00	33303.01	1 (0 0 0 0 1)
YP_001037008 Cthe_0579 nitroreductase	2.13	4.60	19786.45	1 (0 0 0 0 1)
YP_001036595 Cthe_0161 protein of unknown function DUF4	2.13	2.00	12500.59	1 (0 0 0 0 1)

[A] XC, cross correlation score. Xc score as determined by Bioworks 3.3 program. The Xc score represents the cross correlation value between the observed peptide fragment mass spectrum and the one theoretically predicted. [b] Total peptide hits observed for the proteins (columns within parethesis represent the number of primary, secondary, tertiary,etc. peptide hits based spectral quality of the peptides spectra).

Appendix D; Soluble fraction protien identifications of cellobiose grown cells.

Table D1. Complete list of proteins identified in the soluble fraction of *C. thermocellum* grown on cellobiose, ranked by XC score.

Protein	Score(XC)[a]	Coverage (%)	MW (daltons)	Peptide (Hits)[b]
YP_001036854 Cthe_0423 iron-containing alcohol dehydrogenase	520.27	54.00	95948.20	574 (572 2 0 0 0)
YP_001039122 Cthe_2730 translation elongation factor Tu	370.26	65.00	44160.61	1224 (1223 1 0 0 0)
YP_001036772 Cthe_0341 NADH dehydrogenase (quinone)	340.25	47.40	64854.81	331 (331 0 0 0 0)
YP_001037445 Cthe_1020 extracellular solute-binding protein, family 1	300.24	59.50	49953.82	265 (265 0 0 0 0)
YP_001039116 Cthe_2724 DNA-directed RNA polymerase, beta subunit	290.34	32.70	139869.00	113 (113 0 0 0 0)
YP_001039508 Cthe_3120 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	288.24	23.10	130938.90	132 (129 2 0 1 0)
YP_001038203 Cthe_1789 ATPase AAA-2	270.23	34.50	92041.58	100 (100 0 0 0 0)
YP_001036780 Cthe_0349 fructose-1,6-bisphosphate aldolase, class II	260.25	60.80	33611.15	238 (236 2 0 0 0)
YP_001036571 Cthe_0137 glyceraldehyde-3-phosphate dehydrogenase, type I	248.25	63.70	36116.36	740 (737 3 0 0 0)
YP_001039117 Cthe_2725 DNA-directed RNA polymerase, beta' subunit	240.24	25.80	130074.00	104 (104 0 0 0 0)
YP_001039160 Cthe_2768 metallophosphoesterase	240.23	24.00	98713.68	152 (151 1 0 0 0)
YP_001036778 Cthe_0347 phosphofructokinase	230.24	44.30	45595.20	116 (114 2 0 0 0)
YP_001037127 Cthe_0699 carboxyl transferase	228.20	41.30	56012.75	57 (55 2 0 0 0)
YP_001037129 Cthe_0701 conserved carboxylase region containing protein	220.28	41.70	52338.23	83 (83 0 0 0 0)
YP_001039121 Cthe_2729 translation elongation factor G	220.25	30.00	77586.17	176 (175 1 0 0 0)
YP_001037653 Cthe_1228 threonyl-tRNA synthetase	218.21	26.60	73518.59	127 (124 3 0 0 0)
YP_001036572 Cthe_0138 Phosphoglycerate kinase	210.25	39.50	42735.32	156 (156 0 0 0 0)

YP_001036775 Cthe_0344 Malate dehydrogenase	200.25	54.40	42073.30	367 (365 2 0 0 0)
YP_001036717 Cthe_0285 isocitrate dehydrogenase, NADP-dependent	200.22	48.30	45753.04	95 (95 0 0 0 0)
YP_001036773 Cthe_0342 hydrogenase, Fe-only	200.20	38.80	63972.41	147 (147 0 0 0 0)
YP_001037376 Cthe_0949 carbamoyl-phosphate synthase, large subunit	200.19	19.30	118645.00	70 (70 0 0 0 0)
YP_001037532 Cthe_1107 type II secretion system protein E	190.23	25.20	88291.56	51 (50 1 0 0 0)
YP_001038788 Cthe_2393 thiamine pyrophosphate enzyme-like TPP-binding	180.27	54.30	34626.82	127 (126 1 0 0 0)
YP_001038787 Cthe_2392 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	170.23	34.80	43405.89	82 (81 1 0 0 0)
YP_001037733 Cthe_1308 pyruvate, phosphate dikinase	170.20	19.80	98703.04	51 (51 0 0 0 0)
YP_001037430 Cthe_1005 translation elongation factor Ts	166.23	41.90	24321.70	80 (79 0 1 0 0)
YP_001037431 Cthe_1006 ribosomal protein S2	162.19	45.20	28589.95	60 (55 4 0 1 0)
YP_001036860 Cthe_0429 NADH dehydrogenase (quinone)	160.23	26.40	68080.16	81 (81 0 0 0 0)
YP_00103Cthe_0053 ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent	160.22	21.30	88248.48	37 (37 0 0 0 0)
YP_001039265 Cthe_2874 Phosphoenolpyruvate carboxykinase (GTP)	160.22	25.50	67583.80	68 (67 1 0 0 0)
YP_001036849 Cthe_0418 Polyribonucleotide nucleotidyltransferase	150.21	22.60	77269.86	71 (71 0 0 0 0)
YP_001037636 Cthe_1211 pyridoxal-phosphate dependent TrpB-like enzyme	140.23	37.70	50100.10	194 (193 1 0 0 0)
YP_001039588 Cthe_3200 alanyl-tRNA synthetase	140.23	18.00	97985.39	51 (50 1 0 0 0)
YP_001037453 Cthe_1028 acetate kinase	140.22	32.60	43937.85	48 (48 0 0 0 0)
YP_001036861 Cthe_0430 hydrogenase, Fe-only	136.23	24.40	62491.40	52 (51 0 1 0 0)
YP_001039416 Cthe_3027 Citrate (Si)-synthase	134.24	29.40	52000.72	54 (52 1 1 0 0)
YP_00103Cthe_0042 small GTP-binding protein	130.25	29.50	44243.27	37 (37 0 0 0 0)
YP_001036574 Cthe_0140 phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	130.22	23.50	57329.97	33 (33 0 0 0 0)
YP_001036573 Cthe_0139 Triose-phosphate isomerase	130.21	43.80	27116.99	59 (58 1 0 0 0)
YP_001039338 Cthe_2947 prolyl-tRNA synthetase	128.20	28.10	64228.50	55 (53 2 0 0 0)
YP_001039156 Cthe_27TROVE domain containing protein	120.24	27.90	55376.65	30 (30 0 0 0 0)

YP_001039276 Cthe_2885 phosphoribosylaminoimidazole-succinocarboxamide synthase	120.23	35.70	33151.18	33 (33 0 0 0 0)
YP_001038276 Cthe_1862 ABC transporter related protein	120.22	30.80	42036.93	141 (141 0 0 0 0)
YP_001039323 Cthe_2932 DNA-directed RNA polymerase, alpha subunit	120.22	42.20	35017.31	48 (48 0 0 0 0)
YP_001039113 Cthe_2721 ribosomal protein L1	120.21	37.70	25057.48	63 (59 3 1 0 0)
YP_001036569 Cthe_0135 beta-ketoacyl synthase	120.17	3.60	305695.50	28 (28 0 0 0 0)
YP_001037388 Cthe_0961 aspartate-semialdehyde dehydrogenase	110.27	40.70	36705.75	64 (64 0 0 0 0)
YP_001037475 Cthe_1050 recA protein	110.26	33.50	37781.88	32 (32 0 0 0 0)
YP_001037671 Cthe_1246 IMP cyclohydrolase	110.26	24.10	56306.42	44 (44 0 0 0 0)
YP_001036776 Cthe_0345 L-lactate dehydrogenase	110.26	22.30	34811.13	48 (48 0 0 0 0)
YP_001039022 Cthe_2630 ribose-phosphate pyrophosphokinase	110.24	33.10	34730.42	56 (56 0 0 0 0)
YP_001039392 Cthe_3003 hydrogenase, Fe-only	110.23	14.80	71733.04	39 (39 0 0 0 0)
YP_001037686 Cthe_1261 6-phosphofructokinase	110.21	28.70	34773.23	49 (49 0 0 0 0)
YP_001036855 Cthe_0424 aminoglycoside phosphotransferase	100.25	32.90	28568.92	113 (113 0 0 0 0)
YP_001038578 Cthe_2183 UTP-glucose-1-phosphate uridylyltransferase	100.24	43.60	32481.11	30 (30 0 0 0 0)
YP_001038Cthe_2251 Methionine adenosyltransferase	100.24	22.90	43359.20	38 (38 0 0 0 0)
YP_001039266 Cthe_2875 sigma 54 modulation protein/ribosomal protein S30EA	100.24	43.50	20531.70	45 (45 0 0 0 0)
YP_001039120 Cthe_2728 ribosomal protein S7	100.23	36.50	17896.51	27 (27 0 0 0 0)
YP_001037416 Cthe_0991 translation initiation factor IF-2	100.20	6.90	114885.30	44 (44 0 0 0 0)
YP_001037788 Cthe_1363 lipopolysaccharide biosynthesis	100.19	16.40	52051.22	42 (37 4 1 0 0)
YP_001036935 Cthe_0505 formate acetyltransferase	96.26	17.80	84351.19	29 (28 0 1 0 0)
YP_001039424 Cthe_3035 D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	90.26	27.40	42835.39	42 (42 0 0 0 0)
YP_001037361 Cthe_0934 3-oxoacyl-(acyl-carrier-protein) reductase	90.25	39.30	26061.62	30 (30 0 0 0 0)
YP_001038991 Cthe_2598 uracil phosphoribosyltransferase	90.24	33.50	23078.13	37 (37 0 0 0 0)
YP_001038496 Cthe_2096 methionyl-tRNA synthetase	90.22	20.30	75095.11	21 (21 0 0 0 0)

YP_001036876 Cthe_0445 cell division protein FtsZ	90.22	27.90	39703.82	66 (66 0 0 0 0)
YP_001037523 Cthe_1098 hypothetical protein	90.21	15.20	49674.38	21 (21 0 0 0 0)
YP_001039324 Cthe_2933 ribosomal protein L17	90.21	34.50	19695.61	37 (37 0 0 0 0)
YP_001038241 Cthe_1827 copper amine oxidase-like protein	90.20	24.80	36388.99	24 (24 0 0 0 0)
YP_001038785 Cthe_2390 pyruvate/ketoisovalerate oxidoreductase, gamma subunit	90.20	35.90	21190.07	41 (41 0 0 0 0)
YP_001039546 Cthe_3158 putative aconitate hydratase	90.18	17.00	68720.59	17 (17 0 0 0 0)
YP_001039159 Cthe_2767 Methyltransferase type 12	90.18	28.60	53542.80	22 (22 0 0 0 0)
YP_001039329 Cthe_2938 putative glucokinase, ROK family	88.22	18.90	33872.38	37 (35 2 0 0 0)
YP_001039131 Cthe_2739 trigger factor	88.22	21.50	48813.22	28 (27 1 0 0 0)
YP_001039308 Cthe_2917 ribosomal protein S8	88.20	57.60	14209.80	29 (28 1 0 0 0)
YP_001037285 Cthe_0858 protein of unknown function DUF1432	80.23	33.40	35281.07	47 (47 0 0 0 0)
YP_001037160 Cthe_0732 Chorismate synthase	80.22	23.10	43128.44	21 (21 0 0 0 0)
YP_001037531 Cthe_1106 twitching motility protein	80.21	30.20	38896.07	26 (26 0 0 0 0)
YP_001037732 Cthe_1307 cellulosome anchoring protein, cohesin region	80.20	10.50	68577.19	25 (25 0 0 0 0)
YP_001039295 Cthe_2904 ribosomal protein L4/L1e	80.20	21.20	23295.76	31 (31 0 0 0 0)
YP_001037810 Cthe_1385 preprotein translocase, SecA subunit	80.20	7.00	104546.40	29 (29 0 0 0 0)
YP_001036771 Cthe_0340 ferredoxin	80.20	59.80	13423.96	39 (38 1 0 0 0)
YP_001037359 Cthe_0932 beta-ketoacyl synthase	80.20	15.80	43935.96	34 (34 0 0 0 0)
YP_001037756 Cthe_1331 aspartyl-tRNA synthetase	80.19	9.40	67727.98	24 (24 0 0 0 0)
YP_001038208 Cthe_1794 translation elongation factor G	80.18	14.50	76682.54	25 (25 0 0 0 0)
YP_001037378 Cthe_0951 orotidine 5'-phosphate decarboxylase	80.17	29.00	35151.93	22 (22 0 0 0 0)
YP_001038372 Cthe_1961 Nucleotidyl transferase	78.17	11.00	90642.27	23 (17 2 1 1 2)
YP_001039311 Cthe_2920 ribosomal protein S5-like protein	70.25	49.40	17531.73	14 (14 0 0 0 0)
YP_001039299 Cthe_2908 ribosomal protein L22	70.24	50.40	14066.90	37 (37 0 0 0 0)
YP_001037283 Cthe_0856 branched-chain amino acid aminotransferase	70.24	23.30	39711.57	34 (34 0 0 0 0)

YP_001037142 Cthe_0714 hydroxymethylbutenyl pyrophosphate reductase	70.23	10.70	77377.34	37 (37 0 0 0 0)
YP_001039114 Cthe_2722 ribosomal protein L10	70.23	31.50	19501.65	43 (43 0 0 0 0)
YP_001039309 Cthe_2918 ribosomal protein L6	70.23	27.30	20135.12	28 (28 0 0 0 0)
YP_001037104 Cthe_0676 tyrosine recombinase XerD	70.23	21.30	34269.43	29 (29 0 0 0 0)
YP_001036527 Cthe_0093 septum site-determining protein MinD	70.23	32.30	28892.48	25 (24 1 0 0 0)
YP_001039300 Cthe_2909 ribosomal protein S3	70.23	42.90	25312.76	58 (58 0 0 0 0)
YP_001037857 Cthe_1433 short-chain dehydrogenase/reductase SDR	70.22	26.50	27668.10	25 (25 0 0 0 0)
YP_001039322 Cthe_2931 ribosomal protein S4	70.22	21.60	23926.69	39 (39 0 0 0 0)
YP_001037193 Cthe_0765 ribosomal protein L19	70.20	38.60	12924.24	30 (30 0 0 0 0)
YP_001037737 Cthe_1312 glycyl-tRNA synthetase	70.20	13.90	54044.84	24 (24 0 0 0 0)
YP_001037041 Cthe_0613 thiamine pyrophosphate enzyme-like TPP-binding	70.19	11.30	65986.80	29 (29 0 0 0 0)
YP_001038752 Cthe_2357 hypothetical protein	70.19	20.50	44618.16	16 (16 0 0 0 0)
YP_001037516 Cthe_1091 metal dependent phosphohydrolase	70.19	16.20	58726.74	28 (28 0 0 0 0)
YP_001037630 Cthe_1205 putative serine protein kinase, PrkA	70.18	9.00	74227.23	15 (15 0 0 0 0)
YP_001037293 Cthe_0866 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	70.18	23.10	38061.73	15 (15 0 0 0 0)
YP_001039313 Cthe_2922 ribosomal protein L15	70.18	32.90	15576.57	25 (25 0 0 0 0)
YP_001037291 Cthe_08pyruvate ferredoxin/flavodoxin oxidoreductase	70.17	20.10	19329.96	20 (20 0 0 0 0)
YP_001037331 Cthe_0904 protein-export membrane protein SecD	70.16	10.60	48261.15	15 (15 0 0 0 0)
YP_001036842 Cthe_0411 hypothetical protein	68.18	23.30	39706.39	13 (12 1 0 0 0)
YP_001039306 Cthe_2915 ribosomal protein L5	64.21	34.60	20391.86	22 (21 0 0 1 0)
Cthe_3088	60.26	45.90	7948.42	19 (19 0 0 0 0)
YP_001037389 Cthe_0962 dihydrodipicolinate synthase	60.25	18.80	31954.94	17 (17 0 0 0 0)
YP_001037362 Cthe_0935 malonyl CoA-acyl carrier protein transacylase	60.24	29.90	33369.34	25 (25 0 0 0 0)
YP_001037122 Cthe_0694 spermidine synthase	60.24	17.50	31334.13	28 (28 0 0 0 0)

YP_001038913 Cthe_2518 ketol-acid reductoisomerase	60.24	26.30	36199.42	17 (17 0 0 0 0)
YP_001036697 Cthe_0265 aminotransferase, class V	60.23	21.70	41967.30	12 (12 0 0 0 0)
YP_001039157 Cthe_2765 hypothetical protein	60.23	12.20	60495.86	10 (10 0 0 0 0)
YP_001037304 Cthe_0877 GTP-binding protein YchF	60.22	10.20	40506.09	10 (10 0 0 0 0)
YP_001037148 Cthe_0720 aminotransferase, class V	60.21	13.50	43376.50	18 (18 0 0 0 0)
YP_001036726 Cthe_0295 phosphoserine aminotransferase	60.20	14.20	41362.31	12 (12 0 0 0 0)
YP_001036594 Cthe_0160 ribosomal protein L21	60.20	46.60	11531.22	26 (26 0 0 0 0)
YP_001037381 Cthe_0954 Uracil phosphoribosyltransferase	60.20	40.20	20402.04	19 (19 0 0 0 0)
YP_001036577 Cthe_0143 Phosphopyruvate hydratase	60.20	15.50	47073.19	34 (34 0 0 0 0)
YP_001037151 Cthe_0723 tyrosyl-tRNA synthetase	60.19	14.20	46607.92	18 (18 0 0 0 0)
YP_001036804 Cthe_0373 oxidoreductase FAD/NAD(P)-binding	60.19	18.10	30685.13	19 (19 0 0 0 0)
YP_001039078 Cthe_2686 type IV pilus assembly protein PilM	60.19	18.80	41419.14	15 (15 0 0 0 0)
YP_001037483 Cthe_1058 Glycine hydroxymethyltransferase	60.19	12.60	45379.45	12 (12 0 0 0 0)
YP_001038209 Cthe_1795 phospho-2-dehydro-3-deoxyheptonate aldolase	60.18	17.00	36794.52	23 (23 0 0 0 0)
YP_001037747 Cthe_1322 chaperone protein DnaK	60.18	10.00	65641.27	16 (16 0 0 0 0)
YP_001039320 Cthe_2929 ribosomal protein S13	60.18	15.40	14158.91	22 (22 0 0 0 0)
YP_001037084 Cthe_0656 type IV pilus assembly protein PilM	60.18	17.10	41929.30	13 (13 0 0 0 0)
YP_001037806 Cthe_1381 threonine synthase	60.17	9.80	55667.31	14 (14 0 0 0 0)
YP_001039297 Cthe_2906 ribosomal protein L2	60.17	20.00	30176.28	14 (14 0 0 0 0)
YP_00103Cthe_0004 YD repeat containing protein	60.17	2.80	223448.00	6 (6 0 0 0 0)
YP_001039021 Cthe_2629 UDP-N-acetylglucosamine pyrophosphorylase	60.17	16.30	51493.20	11 (11 0 0 0 0)
YP_001038999 Cthe_2606 ATP synthase F1, alpha subunit	60.16	12.50	55784.93	12 (12 0 0 0 0)
YP_001039012 Cthe_2619 cell shape determining protein, MreB/Mrl family	58.20	12.80	36744.56	15 (14 1 0 0 0)
YP_001039393 Cthe_3004 ferredoxin	50.26	11.40	54617.13	42 (42 0 0 0 0)
YP_001036788 Cthe_0357 alpha-glucan phosphorylases	50.25	8.50	97651.16	7 (7 0 0 0 0)

YP_001038776 Cthe_2381 seryl-tRNA synthetase	50.23	7.80	48456.13	10 (10 0 0 0 0)
YP_001037275 Cthe_0847 translation elongation factor P	50.23	33.00	20640.65	14 (14 0 0 0 0)
YP_001036579 Cthe_0145 metal dependent phosphohydrolase	50.21	33.90	20528.80	18 (18 0 0 0 0)
YP_001039049 Cthe_2657 histone-like DNA-binding protein	50.21	46.20	10081.48	18 (18 0 0 0 0)
YP_001039315 Cthe_2924 adenylate kinases	50.19	26.30	24396.61	10 (10 0 0 0 0)
YP_001038803 Cthe_2408 phage shock protein A, PspA	50.19	20.70	27739.44	19 (19 0 0 0 0)
YP_001037375 Cthe_0948 oxidoreductase FAD/NAD(P)-binding	50.19	20.30	28830.65	11 (11 0 0 0 0)
YP_001037042 Cthe_0614 pyruvate ferredoxin/flavodoxin oxidoreductase	50.18	24.10	21269.39	11 (11 0 0 0 0)
YP_001036833 Cthe_0402 copper amine oxidase-like protein	50.18	8.90	83083.48	6 (6 0 0 0 0)
YP_001039378 Cthe_2989 glycosyltransferase 36	50.18	5.20	111670.40	10 (10 0 0 0 0)
YP_001037662 Cthe_1237 leucyl-tRNA synthetase	50.18	5.80	94825.20	10 (10 0 0 0 0)
YP_001037349 Cthe_0922 diaminopimelate dehydrogenase	50.16	19.80	37201.20	12 (12 0 0 0 0)
YP_001037392 Cthe_0965 ATP:corrinoid adenosyltransferase BtuR/CobO/CobP	50.16	30.70	19668.35	12 (12 0 0 0 0)
YP_001036769 Cthe_0338 NADH-quinone oxidoreductase, E subunit	50.15	15.20	18057.11	11 (11 0 0 0 0)
YP_001037109 Cthe_0681 IMP dehydrogenase/GMP reductase	50.15	11.50	55083.21	6 (6 0 0 0 0)
YP_001039304 Cthe_2913 ribosomal protein L14	50.15	25.40	13446.39	31 (31 0 0 0 0)
YP_001039614 Cthe_3226 copper amine oxidase-like protein	48.22	24.40	29604.39	19 (16 3 0 0 0)
YP_001038215 Cthe_1801 ABC transporter related protein	40.28	10.60	28160.66	20 (20 0 0 0 0)
YP_001037527 Cthe_1102 fimbrial assembly protein	40.26	23.50	19870.68	34 (34 0 0 0 0)
YP_001039321 Cthe_2930 ribosomal protein S11	40.24	11.90	14342.67	24 (24 0 0 0 0)
YP_001039557 Cthe_3169 short-chain dehydrogenase/reductase SDR	40.22	15.90	27026.97	9 (9 0 0 0 0)
YP_001037793 Cthe_1368 S-layer-like domain containing protein	40.22	6.50	78075.02	17 (17 0 0 0 0)
YP_001038540 Cthe_2143 hypothetical protein	40.21	39.80	11988.26	10 (10 0 0 0 0)
YP_001038817 Cthe_2422 hypothetical protein	40.20	17.90	40412.15	15 (15 0 0 0 0)
YP_001039543 Cthe_3155 beta-lactamase-like protein	40.20	24.60	26983.65	16 (16 0 0 0 0)

YP_001036857 Cthe_0426 putative PAS/PAC sensor protein	40.19	7.60	62518.02	9 (9 0 0 0 0)
YP_001037134 Cthe_0706 transcriptional regulator, RpiR family	40.19	14.00	33631.35	8 (8 0 0 0 0)
YP_001039551 Cthe_3163 Carbohydrate binding family 25	40.19	31.30	17688.08	44 (39 5 0 0 0)
YP_001037777 Cthe_1352 UDP-glucose 6-dehydrogenase	40.19	13.10	48985.39	6 (6 0 0 0 0)
YP_001039303 Cthe_2912 ribosomal protein S17	40.18	29.40	10034.51	8 (8 0 0 0 0)
YP_001038211 Cthe_1797 3-phosphoshikimate 1-carboxyvinyltransferase	40.18	11.30	45504.41	14 (14 0 0 0 0)
YP_001037709 Cthe_1284 glycogen/starch synthases, ADP-glucose type	40.18	9.20	55860.04	5 (5 0 0 0 0)
YP_001038256 Cthe_1842 O-acetylhomoserine/O-acetylserine sulfhydrylase	40.18	11.20	46498.20	7 (7 0 0 0 0)
YP_001036707 Cthe_0275 glycosyltransferase 36	40.17	7.30	92790.86	7 (7 0 0 0 0)
YP_001038773 Cthe_2378 parB-like partition protein	40.17	17.60	32855.74	15 (15 0 0 0 0)
YP_001039390 Cthe_3001 protein of unknown function DUF47	40.17	16.40	24092.60	10 (10 0 0 0 0)
YP_001039111 Cthe_2719 NusG antitermination factor	40.17	14.70	20117.46	9 (9 0 0 0 0)
YP_001037373 Cthe_0946 Phosphoglycerate mutase	40.16	19.70	27615.73	9 (9 0 0 0 0)
YP_001039466 Cthe_3077 cellulosome anchoring protein, cohesin region	40.16	5.80	196709.70	9 (9 0 0 0 0)
YP_001037518 Cthe_1093 Methenyltetrahydrofolate cyclohydrolase	40.15	7.70	30767.42	8 (8 0 0 0 0)
YP_001036910 Cthe_0479 response regulator receiver protein	40.15	30.80	13205.91	4 (4 0 0 0 0)
YP_001036997 Cthe_0568 methionyl-tRNA formyltransferase	40.14	6.10	34385.84	6 (6 0 0 0 0)
YP_001036921 Cthe_0490 CheA signal transduction histidine kinases	38.18	3.00	77010.76	10 (9 1 0 0 0)
YP_001036542 Cthe_0108 type IV pilus assembly PilZ	38.16	15.00	24709.01	6 (3 3 0 0 0)
YP_001036848 Cthe_0417 ribosomal protein S15	38.14	29.90	10378.75	9 (8 1 0 0 0)
YP_001037344 Cthe_0917 glutaminyl-tRNA synthetase	34.16	9.10	66415.21	5 (4 0 0 1 0)
YP_001038528 Cthe_2130 hypothetical protein	32.23	6.00	48488.74	10 (8 0 0 1 1)
YP_001037652 Cthe_1227 Thymidylate synthase	30.24	14.50	31913.28	10 (10 0 0 0 0)
YP_001037149 Cthe_0721 nitrogen-fixing NifU-like protein	30.22	20.80	16208.82	24 (24 0 0 0 0)

YP_001037675 Cthe_1250 phosphoribosylaminoimidazole carboxylase, catalytic subunit	30.22	26.00	18407.83	6 (6 0 0 0 0)
YP_001038631 Cthe_2236 flagellin-like protein	30.22	10.30	29549.11	25 (25 0 0 0 0)
YP_001039483 Cthe_3095 glycosyl transferase, family 39	30.22	2.60	110274.30	6 (6 0 0 0 0)
YP_001038728 Cthe_2333 two component transcriptional regulator, winged helix family	30.21	6.50	26690.06	5 (5 0 0 0 0)
YP_001039279 Cthe_2888 imidazoleglycerol phosphate synthase, cyclase subunit	30.21	10.70	27690.37	13 (13 0 0 0 0)
YP_001037700 Cthe_1275 H+-ATPase subunit H	30.21	23.40	16877.05	7 (7 0 0 0 0)
YP_001039293 Cthe_2902 ribosomal protein S10	30.21	33.00	11770.55	18 (18 0 0 0 0)
YP_001037916 Cthe_1495 pyridoxamine 5'-phosphate oxidase-related, FMN-binding	30.21	21.50	14943.53	11 (11 0 0 0 0)
YP_001038778 Cthe_2383 hypothetical protein	30.20	13.20	33181.84	8 (8 0 0 0 0)
YP_001038603 Cthe_2208 LexA DNA-binding region containing protein	30.20	22.60	22036.65	8 (8 0 0 0 0)
YP_001037374 Cthe_0947 dihydroorotate dehydrogenase family protein	30.20	11.40	32496.19	10 (10 0 0 0 0)
YP_001038251 Cthe_1837 hypothetical protein	30.20	14.80	20171.65	12 (12 0 0 0 0)
YP_001038605 Cthe_2210 3-isopropylmalate dehydratase, small subunit	30.20	14.80	20861.62	11 (11 0 0 0 0)
YP_001037526 Cthe_1101 hypothetical protein	30.20	9.30	48018.17	15 (15 0 0 0 0)
YP_001037082 Cthe_0654 biotin and thiamin synthesis associated protein	30.19	4.60	52781.02	4 (4 0 0 0 0)
YP_001037786 Cthe_1361 NAD-dependent epimerase/dehydratase	30.19	13.50	39572.03	5 (5 0 0 0 0)
YP_001036516 Cthe_0082 ATP-dependent protease La	30.19	3.80	91737.00	5 (5 0 0 0 0)
YP_001039283 Cthe_2892 chaperonin GroEL	30.18	7.90	57438.54	5 (5 0 0 0 0)
YP_001039544 Cthe_3156 methyl-accepting chemotaxis sensory transducer	30.18	6.70	46200.70	5 (5 0 0 0 0)
YP_00103Cthe_0036 hybrid cluster protein	30.18	6.80	59404.46	9 (6 0 2 0 1)
YP_001037770 Cthe_1345 adenine phosphoribosyltransferase	30.18	8.80	19068.20	12 (12 0 0 0 0)
YP_0010372Cthe_0836 hypothetical protein	30.17	19.20	20373.11	14 (14 0 0 0 0)
YP_001036792 Cthe_0361 hypothetical protein	30.17	11.70	21979.22	7 (7 0 0 0 0)
YP_001039484 Cthe_3096 hypothetical protein	30.17	12.50	33801.52	12 (12 0 0 0 0)
YP_001037797 Cthe_1372 hypothetical protein	30.16	7.20	17235.16	4 (4 0 0 0 0)

YP_001039512 Cthe_3124 AMP-dependent synthetase and ligase	30.16	0.00	61611.04	4 (4 0 0 0 0)
YP_001038580 Cthe_2185 ribosomal protein S18	30.16	12.80	10856.08	7 (7 0 0 0 0)
YP_001037741 Cthe_1316 protein of unknown function DUF370	30.16	30.10	10347.41	8 (8 0 0 0 0)
YP_001038756 Cthe_2361 DNA gyrase, A subunit	30.16	3.30	93408.07	8 (8 0 0 0 0)
YP_001039294 Cthe_2903 ribosomal protein L3	30.16	11.30	23551.69	8 (8 0 0 0 0)
YP_001037058 Cthe_0630 hypothetical protein	30.16	10.50	37146.05	4 (4 0 0 0 0)
YP_001038212 Cthe_1798 CoA-binding protein	30.15	17.50	24616.96	9 (9 0 0 0 0)
YP_001036694 Cthe_0262 gamma-glutamyl phosphate reductase	30.15	5.60	47972.01	8 (8 0 0 0 0)
YP_001037046 Cthe_0618 ferrous iron transport protein B	30.15	1.70	78985.87	6 (6 0 0 0 0)
YP_001037202 Cthe_0774 RNA chaperone Hfq	30.15	12.20	9361.92	8 (8 0 0 0 0)
YP_001037066 Cthe_0638 FHA domain containing protein	30.14	0.70	175123.40	5 (5 0 0 0 0)
YP_001039115 Cthe_2723 ribosomal protein L7/L12	28.22	19.40	13287.30	13 (9 4 0 0 0)
YP_001039097 Cthe_2705 Transketolase, central region	28.16	8.00	33660.42	3 (2 1 0 0 0)
YP_001037259 Cthe_0831 Polyprenyl synthetase	28.15	6.40	32474.66	4 (3 1 0 0 0)
YP_001037799 Cthe_1374 copper amine oxidase-like protein	28.14	10.30	30003.46	5 (4 1 0 0 0)
YP_001036743 Cthe_0312 ATPase AAA-2	26.17	3.50	90556.23	6 (1 5 0 0 0)
YP_001037205 Cthe_0777 DNA mismatch repair protein MutS	26.16	4.40	99327.57	4 (2 2 0 0 0)
YP_001039589 Cthe_3201 CRISPR-associated protein, Csh1 family	24.14	4.40	71508.29	6 (2 1 3 0 0)
YP_001036666 Cthe_0234 AMP-dependent synthetase and ligase	22.18	4.70	95168.88	7 (6 0 0 0 1)
YP_001037276 Cthe_0848 peptidase M24	22.14	3.10	40148.85	5 (1 3 0 1 0)
YP_001039310 Cthe_2919 ribosomal protein L18	20.22	13.10	13690.56	6 (6 0 0 0 0)
YP_001037196 Cthe_0768 nucleic acid binding protein, containing KH domain	20.22	23.70	8420.60	3 (3 0 0 0 0)
YP_001037365 Cthe_0938 regulatory protein, DeoR	20.21	12.20	21123.69	8 (8 0 0 0 0)
YP_001039096 Cthe_2704 Transketolase-like protein	20.21	8.20	31356.90	6 (6 0 0 0 0)
YP_001037657 Cthe_1232 AMP-dependent synthetase and ligase	20.21	5.20	64466.99	6 (6 0 0 0 0)
YP_001039305 Cthe_2914 ribosomal protein L24	20.20	20.70	12807.98	9 (9 0 0 0 0)

YP_001037356 Cthe_0929 Stage V sporulation protein S	20.20	20.90	8862.88	7 (7 0 0 0 0)
YP_001038774 Cthe_2379 hypothetical protein	20.20	20.10	18996.12	8 (8 0 0 0 0)
YP_001037214 Cthe_0786 3-dehydroquinate synthase	20.20	6.90	40065.85	6 (6 0 0 0 0)
YP_001037524 Cthe_1099 hypothetical protein	20.20	3.30	94171.84	4 (4 0 0 0 0)
YP_001037140 Cthe_0712 cytidylate kinase	20.20	13.70	25239.33	5 (5 0 0 0 0)
YP_001037295 Cthe_0868 type IV pilus assembly PilZ	20.20	9.30	27940.03	3 (3 0 0 0 0)
YP_001037452 Cthe_1027 protein of unknown function DUF177	20.19	14.40	19025.28	6 (6 0 0 0 0)
YP_001037454 Cthe_1029 phosphate acetyltransferase	20.19	4.70	38641.03	11 (11 0 0 0 0)
YP_001037076 Cthe_0glutamyl-tRNA synthetase	20.19	5.70	63361.20	4 (4 0 0 0 0)
YP_001037296 Cthe_0869 hypothetical protein	20.19	14.90	25289.15	6 (6 0 0 0 0)
YP_001038356 Cthe_1945 Thioredoxin-disulfide reductase	20.19	4.70	32476.80	5 (5 0 0 0 0)
YP_001037608 Cthe_1183 ATPase associated with various cellular activities, AAA_3	20.18	5.20	35005.28	3 (3 0 0 0 0)
YP_001036Cthe_0215 phenylalanyl-tRNA synthetase, beta subunit	20.18	2.50	88432.06	4 (4 0 0 0 0)
YP_001037690 Cthe_1265 phosphoglucomutase/phosphomannomutase alpha/beta/alpha domain I	20.18	2.80	64914.06	2 (2 0 0 0 0)
YP_001039106 Cthe_2714 acetolactate synthase, large subunit, biosynthetic type	20.18	5.00	60696.15	4 (4 0 0 0 0)
YP_001037835 Cthe_1411 tryptophan synthase, alpha subunit	20.18	13.70	28167.31	6 (6 0 0 0 0)
YP_001038196 Cthe_1782 ribosomal protein S9	20.18	10.80	14656.17	11 (11 0 0 0 0)
YP_001039579 Cthe_3191 ATPase involved in chromosome partitioning-like protein	20.18	4.70	45193.98	3 (3 0 0 0 0)
YP_001037390 Cthe_0963 Dihydrodipicolinate reductase	20.18	6.70	27775.55	10 (10 0 0 0 0)
YP_001038198 Cthe_1784 hypothetical protein	20.18	17.60	15136.84	6 (6 0 0 0 0)
YP_001038761 Cthe_2366 single-stranded nucleic acid binding R3H	20.18	9.70	23558.82	4 (4 0 0 0 0)
YP_001037500 Cthe_1075 protein of unknown function DUF881	20.18	5.60	26506.01	5 (5 0 0 0 0)
YP_001037907 Cthe_1485 beta-lactamase-like protein	20.18	5.50	47381.80	8 (8 0 0 0 0)
YP_001039104 Cthe_2712 hypothetical protein	20.17	11.70	26108.48	6 (6 0 0 0 0)
YP_001037849 Cthe_1425 Inorganic diphosphatase	20.17	3.90	71337.66	5 (5 0 0 0 0)

YP_001039098 Cthe_2706 ABC transporter related protein	20.17	8.90	41203.11	6 (6 0 0 0 0)
YP_001037065 Cthe_0637 protein of unknown function DUF909	20.17	13.00	11050.80	8 (8 0 0 0 0)
YP_00103Cthe_0045 copper amine oxidase-like protein	20.17	12.60	30544.85	2 (2 0 0 0 0)
YP_001037478 Cthe_1053 L-lactate dehydrogenase	20.17	6.30	35053.30	4 (4 0 0 0 0)
YP_00103Cthe_0019 regulatory protein, P-II family	20.17	25.00	11659.06	5 (5 0 0 0 0)
YP_001039467 Cthe_3078 cellulosome anchoring protein, cohesin region	20.17	0.80	248014.30	6 (6 0 0 0 0)
YP_001038Cthe_2254 hypoxanthine phosphoribosyltransferase	20.16	10.30	20691.94	6 (6 0 0 0 0)
YP_001037183 Cthe_0755 aminotransferase, class I and II	20.16	5.60	42895.99	5 (5 0 0 0 0)
YP_001038567 Cthe_2172 ribosomal protein L31	20.16	34.30	7926.99	5 (5 0 0 0 0)
YP_001036781 Cthe_0350 signal peptidase I	20.16	9.80	22212.42	3 (3 0 0 0 0)
YP_001036806 Cthe_0375 GMP synthase, large subunit	20.16	2.50	57519.82	2 (2 0 0 0 0)
YP_001039469 Cthe_3080 cellulosome anchoring protein, cohesin region	20.16	4.70	48500.00	6 (6 0 0 0 0)
YP_001037443 Cthe_1018 binding-protein-dependent transport systems inner membrane component	20.16	8.70	31904.56	15 (15 0 0 0 0)
YP_001037666 Cthe_1241 nicotinate (nicotinamide) nucleotide adenylyltransferase	20.16	10.20	23673.40	3 (3 0 0 0 0)
YP_001039493 Cthe_3105 exsB protein	20.16	4.50	24239.59	6 (6 0 0 0 0)
YP_001037353 Cthe_0926 signal recognition particle-docking protein FtsY	20.16	7.60	32882.91	4 (4 0 0 0 0)
YP_001037472 Cthe_1047 hypothetical protein	20.16	6.60	20674.68	5 (5 0 0 0 0)
YP_001038799 Cthe_2404 transcriptional regulator, GntR family	20.16	5.70	26246.02	7 (7 0 0 0 0)
YP_001037836 Cthe_1412 tryptophan synthase, beta subunit	20.16	6.10	43091.86	2 (2 0 0 0 0)
YP_001038771 Cthe_2376 DNA gyrase, B subunit	20.16	3.60	72446.46	3 (3 0 0 0 0)
YP_001038415 Cthe_2008 conserved hypothetical protein, CF-22 family	20.16	11.00	20068.01	3 (3 0 0 0 0)
YP_001039615 Cthe_3227 copper amine oxidase-like protein	20.16	4.90	30240.66	6 (6 0 0 0 0)
YP_001037439 Cthe_1014 MutS2 family protein	20.16	2.50	88635.20	2 (2 0 0 0 0)
YP_001036980 Cthe_0551 AMP-dependent synthetase and ligase	20.15	2.50	63015.91	2 (2 0 0 0 0)
YP_001037102 Cthe_0674 NUDIX	20.15	12.10	20289.62	3 (3 0 0 0 0)

hydrolase				
YP_001038584 Cthe_2189 diguanylate cyclase with GAF sensor	20.15	5.80	47299.96	3 (3 0 0 0 0)
YP_001038475 Cthe_2075 protein of unknown function DUF28	20.15	4.50	26973.46	7 (7 0 0 0 0)
YP_001038569 Cthe_2174 transcription termination factor Rho	20.15	3.70	72542.66	4 (4 0 0 0 0)
YP_001039312 Cthe_2921 ribosomal protein L30	20.15	22.00	6614.67	10 (10 0 0 0 0)
YP_001037145 Cthe_0717 MCP methyltransferase, CheR-type	20.15	8.10	30329.86	4 (4 0 0 0 0)
YP_001037009 Cthe_0580 aminotransferase, class I and II	20.15	5.50	44743.46	4 (4 0 0 0 0)
YP_001039133 Cthe_2741 ATP-dependent Clp protease, ATP-binding subunit ClpX	20.15	2.10	47861.47	5 (5 0 0 0 0)
YP_001037330 Cthe_0903 protein-export membrane protein SecF	20.15	4.10	34990.80	9 (9 0 0 0 0)
YP_001036742 Cthe_0311 excinuclease ABC, A subunit	20.14	3.40	104767.60	4 (4 0 0 0 0)
YP_001036856 Cthe_0425 hypothetical protein	20.14	16.00	8920.56	7 (7 0 0 0 0)
YP_001037444 Cthe_1019 binding-protein-dependent transport systems inner membrane component	20.14	3.10	37355.85	6 (6 0 0 0 0)
YP_001039119 Cthe_2727 ribosomal protein S12	20.14	7.70	15415.67	5 (5 0 0 0 0)
YP_001039314 Cthe_2923 preprotein translocase, SecY subunit	20.14	0.00	47339.89	8 (8 0 0 0 0)
YP_001037569 Cthe_1144 restriction modification system DNA specificity domain	20.14	0.00	54427.77	2 (2 0 0 0 0)
YP_001038Cthe_2252 thioesterase family protein	20.14	16.40	14553.56	4 (4 0 0 0 0)
YP_001038743 Cthe_2348 Ig-related protein	20.14	2.60	113260.90	2 (2 0 0 0 0)
YP_001036832 Cthe_0401 methyl-accepting chemotaxis sensory transducer	20.14	1.90	85087.83	3 (3 0 0 0 0)
YP_001036999 Cthe_0570 peptidase, membrane zinc metallopeptidase, putative	20.14	5.30	24747.35	3 (3 0 0 0 0)
YP_001037220 Cthe_0792 Protein of unknown function UPF0001	20.13	5.60	26403.80	3 (3 0 0 0 0)
YP_001037212 Cthe_0784 hypothetical protein	20.13	6.50	33757.39	2 (2 0 0 0 0)
YP_001037415 Cthe_0990 ribosome-binding factor A	20.13	17.10	14132.42	3 (3 0 0 0 0)
YP_001037215 Cthe_0787 isoleucyl-tRNA synthetase	20.13	1.30	106730.50	2 (2 0 0 0 0)
YP_001037114 Cthe_0686 tryptophanyl-tRNA synthetase	18.18	9.70	37373.24	3 (1 1 1 0 0)
YP_001038466 Cthe_2066 serine O-acetyltransferase	18.18	6.00	27382.19	5 (3 1 1 0 0)
YP_00103Cthe_0056 Ig-like, group 2	18.16	0.50	493538.80	11 (4 7 0 0 0)

YP_001037963 Cthe_1543 aspartyl-tRNA synthetase	18.15	2.10	66454.89	4 (2 1 0 1 0)
YP_001039281 Cthe_2890 putative transcriptional regulator	18.15	0.00	43839.23	2 (1 1 0 0 0)
YP_001038696 Cthe_2301 CRISPR-associated autoregulator, DevR family	18.15	3.70	32508.38	3 (2 1 0 0 0)
YP_001037369 Cthe_0942 MiaB-like tRNA modifying enzyme YliG	18.14	2.00	51752.77	3 (2 1 0 0 0)
YP_001038980 Cthe_2587 hypothetical protein	18.14	17.00	17817.86	2 (1 1 0 0 0)
YP_001037479 Cthe_1054 ATPase associated with various cellular activities, AAA_3	18.14	3.80	35777.95	3 (1 2 0 0 0)
YP_001037490 Cthe_1065 type IV pilus assembly PilZ	18.14	4.50	25076.02	6 (3 3 0 0 0)
YP_001038177 Cthe_1763 ABC transporter related protein	16.16	7.30	27658.55	8 (2 0 2 3 1)
YP_001038574 Cthe_2179 Pectate lyase/Amb allergen	16.16	2.90	99828.96	3 (1 1 1 0 0)
YP_001038441 Cthe_2039 UvrD/REP helicase	16.15	0.90	143255.10	10 (1 1 6 1 1)
YP_001036887 Cthe_0456 protein of unknown function UPF0047	16.15	0.00	15799.08	3 (2 0 1 0 0)
YP_001036941 Cthe_0511 histidine kinase	16.13	0.00	50721.29	2 (0 2 0 0 0)
YP_001037307 Cthe_0880 phospho-2-dehydro-3-deoxyheptonate aldolase	14.18	8.00	36848.65	5 (4 0 0 1 0)
YP_001036948 Cthe_0518 type III restriction enzyme, res subunit	14.16	3.60	114120.70	2 (0 1 1 0 0)
YP_001038624 Cthe_2229 N-acetylneuraminate synthase	14.16	3.10	39121.83	4 (3 0 0 1 0)
YP_001037792 Cthe_1367 PHP-like protein	14.15	5.80	29478.46	2 (1 0 0 1 0)
YP_001036682 Cthe_0250 protein of unknown function UPF0027	14.13	2.30	44820.60	2 (1 0 0 1 0)
YP_001036827 Cthe_0396 ABC transporter related protein	14.13	0.00	63672.71	2 (1 0 0 1 0)
YP_001036979 Cthe_0550 heat shock protein Hsp90	12.15	1.90	73792.48	4 (2 0 0 1 1)
YP_001038500 Cthe_2100 transcriptional regulator, AbrB family	10.23	16.50	8910.75	10 (10 0 0 0 0)
YP_001037650 Cthe_1225 translation initiation factor IF-3	10.21	7.90	18832.10	4 (4 0 0 0 0)
YP_001038182 Cthe_1768 NifU-related domain containing protein	10.21	6.50	25046.62	3 (3 0 0 0 0)
YP_001038351 Cthe_1940 RDD domain containing protein	10.20	6.10	28171.33	6 (6 0 0 0 0)
YP_001039421 Cthe_3032 hypothetical protein	10.20	10.40	15321.67	3 (3 0 0 0 0)
YP_001038807 Cthe_2412 SMC protein-like protein	10.20	2.20	101479.00	1 (1 0 0 0 0)

YP_001036578 Cthe_0144 preprotein translocase, SecG subunit	10.20	20.50	8797.02	9 (9 0 0 0 0)
YP_001037328 Cthe_0901 pantoate--beta-alanine ligase	10.20	5.30	31643.65	4 (4 0 0 0 0)
YP_001037801 Cthe_1376 Homoserine dehydrogenase	10.20	3.90	45085.58	1 (1 0 0 0 0)
YP_001039481 Cthe_3093 Adenylosuccinate synthase	10.19	4.00	47011.32	6 (6 0 0 0 0)
YP_001036522 Cthe_0088 cell shape determining protein, MreB/Mrl family	10.19	4.40	36283.29	4 (3 1 0 0 0)
YP_001038694 Cthe_2299 CRISPR-associated helicase Cas3	10.19	2.30	87169.28	5 (4 1 0 0 0)
YP_001039079 Cthe_2687 thioesterase superfamily	10.19	7.80	18915.82	4 (4 0 0 0 0)
YP_001038376 Cthe_1965 alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	10.19	7.50	20791.48	4 (4 0 0 0 0)
YP_001038558 Cthe_2163 anti-sigma-factor antagonist	10.19	14.40	12690.81	3 (3 0 0 0 0)
YP_001036858 Cthe_0427 serine phosphatase	10.19	3.60	43202.31	3 (3 0 0 0 0)
YP_001036501 Cthe_0067 Silent information regulator protein Sir2	10.19	5.40	26808.54	4 (4 0 0 0 0)
YP_001036981 Cthe_0552 transcriptional regulator, XRE family with cupin sensor	10.19	8.20	20560.73	2 (2 0 0 0 0)
YP_001036685 Cthe_0253 ATPase associated with various cellular activities, AAA_3	10.18	4.60	36907.50	1 (1 0 0 0 0)
YP_001037530 Cthe_1105 type II secretion system protein	10.18	3.50	45131.49	3 (3 0 0 0 0)
YP_001039593 Cthe_3205 hypothetical protein	10.18	6.80	25549.67	2 (2 0 0 0 0)
YP_001038726 Cthe_2331 hypothetical protein	10.18	8.00	19873.26	1 (1 0 0 0 0)
YP_001036581 Cthe_0147 protein of unknown function DUF523	10.18	0.00	15404.86	1 (1 0 0 0 0)
YP_001037727 Cthe_1302 hypothetical protein	10.18	2.00	61496.71	1 (1 0 0 0 0)
YP_0010374Cthe_1039 ribosomal protein S20	10.18	13.00	12988.50	8 (8 0 0 0 0)
YP_001038632 Cthe_2237 flagellin-like protein	10.18	4.40	29413.22	6 (6 0 0 0 0)
YP_001037277 Cthe_0849 3-dehydroquinate dehydratase, type II	10.18	10.60	15646.33	1 (1 0 0 0 0)
YP_001039296 Cthe_2905 Ribosomal protein L25/L23	10.18	13.70	13281.25	5 (5 0 0 0 0)
YP_001039461 Cthe_3072 acyl-ACP thioesterase	10.18	5.50	29446.92	2 (2 0 0 0 0)
YP_001038181 Cthe_1767 hypothetical protein	10.18	4.80	35361.88	1 (1 0 0 0 0)

YP_001037027 Cthe_0599 thiazole biosynthesis protein ThiH	10.17	3.80	42819.91	1 (1 0 0 0 0)
YP_001036511 Cthe_0077 hypothetical protein	10.17	8.70	13916.57	6 (6 0 0 0 0)
YP_001038423 Cthe_2018 hypothetical protein	10.17	3.20	48846.50	4 (4 0 0 0 0)
YP_001037429 Cthe_1004 uridylate kinase	10.17	8.10	25621.45	1 (1 0 0 0 0)
YP_001036613 Cthe_0179 Argininosuccinate synthase	10.17	3.20	45593.73	8 (8 0 0 0 0)
YP_001039482 Cthe_3094 glycosyl transferase, family 2	10.17	4.00	37622.82	4 (4 0 0 0 0)
YP_0010365Cthe_0130 3-oxoacyl-(acyl-carrier-protein) synthase	10.17	4.40	38867.86	2 (2 0 0 0 0)
YP_001036538 Cthe_0104 riboflavin biosynthesis protein RibD	10.17	4.90	39237.94	2 (2 0 0 0 0)
YP_001038272 Cthe_1858 peptidase M23B	10.16	3.50	41616.22	2 (2 0 0 0 0)
YP_001038912 Cthe_2517 acetolactate synthase, small subunit	10.16	7.10	18609.04	4 (4 0 0 0 0)
YP_001038204 Cthe_1790 ATP:guanido phosphotransferase	10.16	3.20	38662.09	4 (4 0 0 0 0)
YP_001037845 Cthe_1421 signal peptide peptidase SppA, 36K type	10.16	4.30	35780.34	2 (2 0 0 0 0)
YP_001037043 Cthe_0615 Phenylacetate--CoA ligase	10.16	0.00	48686.20	2 (2 0 0 0 0)
YP_001038259 Cthe_1845 Homoserine O-succinyltransferase	10.16	5.60	35848.21	2 (2 0 0 0 0)
YP_001038529 Cthe_2131 hypothetical protein	10.16	9.80	13757.98	5 (5 0 0 0 0)
YP_001038832 Cthe_2437 shikimate kinase	10.16	6.50	18864.94	4 (4 0 0 0 0)
YP_001039002 Cthe_2609 ATP synthase F1, epsilon subunit	10.16	7.40	15369.17	2 (2 0 0 0 0)
YP_001036547 Cthe_0113 Uncharacterized P-loop ATPase protein UPF0042	10.16	5.50	33114.67	1 (1 0 0 0 0)
YP_001036988 Cthe_0559 single-strand binding protein/Primosomal replication protein n	10.16	4.20	27479.06	2 (2 0 0 0 0)
YP_001039067 Cthe_2675 hypothetical protein	10.16	10.40	25930.29	1 (1 0 0 0 0)
YP_001038626 Cthe_2231 Methyltransferase type 11	10.16	6.10	25433.77	1 (1 0 0 0 0)
YP_001038822 Cthe_2427 protein of unknown function DUF1385	10.16	3.10	35839.37	4 (4 0 0 0 0)
YP_001037514 Cthe_1089 Stage V sporulation protein S	10.16	17.00	9140.00	1 (1 0 0 0 0)
YP_001037588 Cthe_1163 phosphoglucosamine mutase	10.16	3.10	48744.14	2 (2 0 0 0 0)
YP_001038504 Cthe_2104 PSP1	10.16	3.80	32969.09	2 (2 0 0 0 0)
YP_001037746 Cthe_1321 chaperone protein DnaJ	10.16	3.60	42060.23	1 (1 0 0 0 0)

YP_001037784 Cthe_1359 glycosyltransferase 28-like protein	10.16	0.00	18935.60	1 (1 0 0 0 0)
YP_001039301 Cthe_2910 ribosomal protein L16	10.16	10.30	16213.71	3 (3 0 0 0 0)
YP_001037895 Cthe_1473 NADPH-dependent FMN reductase	10.16	8.00	20921.93	1 (1 0 0 0 0)
YP_001038119 Cthe_1705 hypothetical protein	10.16	0.00	65110.92	1 (1 0 0 0 0)
YP_001036652 Cthe_0218 metallophosphoesterase	10.16	0.00	30473.25	1 (1 0 0 0 0)
YP_001037105 Cthe_0677 phosphopentomutase	10.16	2.30	42561.62	4 (4 0 0 0 0)
YP_001039587 Cthe_3199 protein of unknown function UPF0118	10.16	6.60	38545.74	1 (1 0 0 0 0)
YP_001037525 Cthe_1100 prepilin-type cleavage/methylation	10.16	7.50	19349.18	4 (4 0 0 0 0)
YP_001037800 Cthe_1375 aspartate kinase	10.16	2.00	49629.48	4 (4 0 0 0 0)
YP_001037440 Cthe_1015 hypothetical protein	10.16	9.30	14662.84	1 (1 0 0 0 0)
YP_001038205 Cthe_1791 UvrB/UvrC protein	10.16	8.30	19332.59	2 (2 0 0 0 0)
YP_001039201 Cthe_2809 glycoside hydrolase, family 16	10.15	1.00	147694.60	2 (2 0 0 0 0)
YP_001038772 Cthe_2377 Cobyrinic acid a,c-diamide synthase	10.15	6.60	28474.09	4 (4 0 0 0 0)
YP_001038197 Cthe_1783 ribosomal protein L13	10.15	9.00	16763.18	3 (3 0 0 0 0)
YP_001037426 Cthe_1001 undecaprenyl diphosphate synthase	10.15	0.00	29048.47	1 (1 0 0 0 0)
YP_001038179 Cthe_1765 hypothetical protein	10.15	0.00	44742.39	2 (1 1 0 0 0)
YP_001038324 Cthe_1912 copper amine oxidase-like protein	10.15	0.00	59214.88	1 (1 0 0 0 0)
YP_001036722 Cthe_0290 Homoserine dehydrogenase	10.15	2.30	47059.77	2 (2 0 0 0 0)
YP_001038582 Cthe_2187 ribosomal protein S6	10.15	12.60	11131.85	1 (1 0 0 0 0)
YP_001037494 Cthe_1069 protein of unknown function UPF0054	10.15	7.40	19165.63	2 (2 0 0 0 0)
YP_001037048 Cthe_0620 iron (metal) dependent repressor, DtxR family	10.15	0.00	15040.75	1 (1 0 0 0 0)
YP_001037403 Cthe_0978 UDP-N-acetylmuramyl-tripeptide synthetases	10.15	3.50	53768.75	1 (1 0 0 0 0)
YP_001039298 Cthe_2907 ribosomal protein S19	10.15	14.90	10589.65	1 (1 0 0 0 0)
YP_001036546 Cthe_0112 UDP-N-acetylenolpyruvoylglucosamine reductase	10.15	3.00	33466.49	4 (4 0 0 0 0)
YP_001037808 Cthe_1383 Tetratricopeptide TPR_2	10.15	0.00	45162.93	1 (1 0 0 0 0)

YP_001037762 Cthe_1337 type II secretion system protein	10.15	4.80	32758.53	1 (1 0 0 0 0)
YP_001036889 Cthe_0458 exodeoxyribonuclease III Xth	10.15	4.40	29295.52	2 (2 0 0 0 0)
YP_001039477 Cthe_3089 UspA	10.15	10.00	14712.77	1 (1 0 0 0 0)
YP_001038520 Cthe_2120 RNA polymerase sigma factor	10.15	5.10	29950.84	1 (1 0 0 0 0)
YP_001038344 Cthe_1933 Zn-dependent hydrolase of the beta-lactamase fold	10.15	0.00	23397.92	1 (1 0 0 0 0)
YP_001038509 Cthe_2109 copper amine oxidase-like protein	10.15	0.00	94460.30	1 (1 0 0 0 0)
YP_001039554 Cthe_3166 glucose-1-phosphate adenylyltransferase	10.15	2.80	47612.06	1 (1 0 0 0 0)
YP_001038030 Cthe_1613 glycosyl hydrolase-like protein	10.15	0.00	92386.96	1 (1 0 0 0 0)
YP_001039001 Cthe_2608 ATP synthase F1, beta subunit	10.15	0.00	50980.69	3 (3 0 0 0 0)
YP_001036859 Cthe_0428 NADH dehydrogenase (ubiquinone), 24 kDa subunit	10.15	0.00	18338.32	2 (2 0 0 0 0)
YP_001039100 Cthe_2708 hypothetical protein	10.15	3.00	53616.59	2 (2 0 0 0 0)
YP_001037265 Cthe_0837 protein of unknown function DUF322	10.15	8.20	14220.37	2 (2 0 0 0 0)
YP_001037201 Cthe_0773 SOS-response transcriptional repressor, LexA	10.15	5.20	24015.77	3 (3 0 0 0 0)
YP_001038Cthe_2253 ATP-dependent metalloprotease FtsH	10.15	1.70	66641.17	1 (1 0 0 0 0)
YP_001036533 Cthe_0099 5,10-methylenetetrahydrofolate reductase	10.15	4.10	32764.93	2 (2 0 0 0 0)
YP_001039280 Cthe_2889 phosphoribosyl-AMP cyclohydrolase	10.15	5.90	25576.21	1 (1 0 0 0 0)
YP_001036506 Cthe_0072 phage shock protein C, PspC	10.15	6.80	18036.66	1 (1 0 0 0 0)
YP_001038651 Cthe_2256 replicative DNA helicase	10.15	3.80	49924.21	1 (1 0 0 0 0)
YP_001038873 Cthe_2478 phage minor structural GP20	10.15	8.60	21347.13	1 (1 0 0 0 0)
YP_001037049 Cthe_0621 putative translation initiation factor, aIF-2BI family	10.15	2.60	37833.65	3 (3 0 0 0 0)
YP_001039105 Cthe_2713 dihydroxy-acid dehydratase	10.15	1.40	58842.55	1 (1 0 0 0 0)
YP_001037993 Cthe_1574 hypothetical protein	10.15	5.70	27775.38	2 (2 0 0 0 0)
YP_001037003 Cthe_0574 serine/threonine protein kinase	10.15	0.00	78419.30	1 (1 0 0 0 0)
YP_001037676 Cthe_1251 Xanthine/uracil/vitamin C permease	10.15	2.60	49393.75	1 (1 0 0 0 0)
YP_001036960 Cthe_0531 hypothetical protein	10.14	1.80	87439.81	1 (1 0 0 0 0)

YP_001039417 Cthe_3028 Pyridoxal-dependent decarboxylase	10.14	3.30	45200.77	2 (2 0 0 0 0)
YP_001036853 Cthe_0422 CoA-binding protein	10.14	5.70	24791.90	3 (3 0 0 0 0)
YP_001036698 Cthe_0266 methyl-accepting chemotaxis sensory transducer	10.14	3.00	55424.34	3 (3 0 0 0 0)
YP_001039468 Cthe_3079 cellulosome anchoring protein, cohesin region	10.14	1.50	74924.76	2 (2 0 0 0 0)
YP_001037614 Cthe_1189 ABC transporter related protein	10.14	5.60	36518.56	4 (0 1 0 0 3)
YP_001036994 Cthe_0565 hypothetical protein	10.14	8.60	15597.69	2 (2 0 0 0 0)
YP_001038723 Cthe_2328 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	10.14	0.00	44355.70	1 (1 0 0 0 0)
YP_001036875 Cthe_0444 cell division protein FtsA	10.14	0.00	44730.68	1 (1 0 0 0 0)
YP_001036708 Cthe_0276 D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	10.14	3.40	35661.52	1 (1 0 0 0 0)
YP_001039085 Cthe_2693 hypothetical protein	10.14	5.50	17809.54	4 (4 0 0 0 0)
YP_001039381 Cthe_2992 RNA polymerase, sigma-24 subunit, ECF subfamily	10.14	4.60	20428.45	1 (1 0 0 0 0)
YP_001036567 Cthe_0133 AMP-dependent synthetase and ligase	10.14	2.20	55084.16	1 (1 0 0 0 0)
YP_001038378 Cthe_1967 DNA segregation ATPase FtsK/SpoIIIE and related proteins	10.14	3.50	66691.32	1 (1 0 0 0 0)
YP_001039405 Cthe_3016 (NiFe) hydrogenase maturation protein HypF	10.14	0.00	84348.26	1 (1 0 0 0 0)
YP_001038322 Cthe_1910 hypothetical protein	10.14	0.00	46643.71	2 (2 0 0 0 0)
YP_001037713 Cthe_1288 two component transcriptional regulator, winged helix family	10.14	4.80	26420.85	1 (1 0 0 0 0)
YP_001037715 Cthe_1290 hypothetical protein	10.14	0.00	12047.38	1 (1 0 0 0 0)
YP_001036770 Cthe_0339 histidine kinase	10.14	5.90	20975.22	1 (1 0 0 0 0)
YP_0010377Cthe_1339 type II secretion system protein E	10.14	2.70	46617.81	1 (1 0 0 0 0)
YP_001036598 Cthe_01protein of unknown function UPF0044	10.14	9.40	10681.73	1 (1 0 0 0 0)
YP_001039618 Cthe_3230 hypothetical protein	10.14	1.10	109655.20	2 (1 0 1 0 0)
YP_001039530 Cthe_3142 hypothetical protein	10.14	7.40	18432.98	2 (2 0 0 0 0)
YP_001036503 Cthe_0069 Aspartate--ammonia ligase	10.14	3.20	39008.13	2 (2 0 0 0 0)
YP_001039020 Cthe_2628 SpoVG	10.14	0.00	10462.28	4 (4 0 0 0 0)
YP_001039088 Cthe_2696 carbohydrate kinase, YjeF related protein	10.14	0.00	54543.56	1 (1 0 0 0 0)

YP_00103Cthe_0048 copper amine oxidase-like protein	10.14	3.70	30342.87	2 (2 0 0 0 0)
YP_001037044 Cthe_0616 amino acid-binding ACT	10.14	7.70	15451.36	3 (3 0 0 0 0)
YP_001037619 Cthe_1194 hypothetical protein	10.14	2.60	45013.45	1 (1 0 0 0 0)
YP_001037116 Cthe_0688 putative transcriptional regulator	10.14	5.10	22425.68	1 (1 0 0 0 0)
YP_001039491 Cthe_3103 2-hydroxyglutaryl-CoA dehydratase, D-component	10.14	0.00	37446.18	2 (2 0 0 0 0)
YP_001037611 Cthe_1186 Alpha/beta hydrolase fold-3	10.14	5.10	35160.46	1 (1 0 0 0 0)
YP_001036620 Cthe_0186 UDP-glucose 4-epimerase	10.14	3.10	39089.43	4 (4 0 0 0 0)
YP_001038965 Cthe_2572 hypothetical protein	10.14	5.80	24096.17	1 (1 0 0 0 0)
YP_001037765 Cthe_1340 hypothetical protein	10.14	0.00	41295.08	1 (1 0 0 0 0)
YP_001038328 Cthe_1916 two component transcriptional regulator, winged helix family	10.14	4.80	26245.90	1 (1 0 0 0 0)
YP_001037292 Cthe_0865 3-methyl-2-oxobutanoate dehydrogenase (ferredo	10.14	5.60	26948.71	1 (1 0 0 0 0)
YP_001036829 Cthe_0398 protein of unknown function DUF74	10.14	0.00	11285.74	2 (2 0 0 0 0)
YP_001038472 Cthe_2072 hypothetical protein	10.14	4.40	27549.79	1 (1 0 0 0 0)
YP_001037315 Cthe_0888 type IV pilus assembly PilZ	10.14	3.90	26967.06	1 (1 0 0 0 0)
YP_001037240 Cthe_0812 response regulator receiver protein	10.14	0.00	30139.98	1 (1 0 0 0 0)
YP_001037755 Cthe_1330 signal peptidase I	10.14	6.90	19480.25	1 (1 0 0 0 0)
YP_001038741 Cthe_2346 O-antigen polymerase	10.14	0.00	114378.30	1 (1 0 0 0 0)
YP_001038283 Cthe_1869 ornithine carbamoyltransferase	10.14	3.30	34220.27	5 (5 0 0 0 0)
YP_001037501 Cthe_1076 protein of unknown function DUF881	10.14	0.00	27007.32	1 (1 0 0 0 0)
YP_001037769 Cthe_1344 (p)ppGpp synthetase I, SpoT/RelA	10.14	1.50	82982.93	1 (1 0 0 0 0)
YP_001036777 Cthe_0346 hypothetical protein	10.14	0.00	13843.44	1 (1 0 0 0 0)
YP_001036616 Cthe_0182 Holliday junction DNA helicase RuvB	10.14	2.40	36494.54	1 (1 0 0 0 0)
YP_001038280 Cthe_1866 acetylornithine and succinylornithine aminotransferases	10.14	0.00	43892.60	2 (2 0 0 0 0)
YP_001038206 Cthe_1792 Firmicute transcriptional repressor of class III stress genes	10.14	6.50	17037.77	1 (1 0 0 0 0)

YP_001037238 Cthe_0810 CheA signal transduction histidine kinases	10.14	1.40	79494.73	1 (1 0 0 0 0)
YP_001037446 Cthe_1021 stage IV sporulation protein A	10.13	0.00	56148.24	1 (1 0 0 0 0)
YP_001039227 Cthe_2835 hypothetical protein	10.13	7.80	22630.74	1 (1 0 0 0 0)
YP_001037247 Cthe_0819 ABC transporter related protein	10.13	5.20	32076.80	1 (1 0 0 0 0)
YP_001038759 Cthe_23glucose inhibited division protein A	10.13	3.00	70739.05	6 (5 1 0 0 0)
YP_001036814 Cthe_0383 protein of unknown function DUF214	10.13	4.50	22495.28	1 (1 0 0 0 0)
YP_001039458 Cthe_3069 two component transcriptional regulator, winged helix family	10.13	0.00	26281.78	2 (2 0 0 0 0)
YP_001036714 Cthe_0282 glycogen/starch synthases, ADP-glucose type	10.13	2.30	55693.16	1 (1 0 0 0 0)
YP_001037740 Cthe_1315 Guanylate kinase	10.13	0.00	23239.91	1 (1 0 0 0 0)
YP_001037958 Cthe_1538 transcriptional regulator, XRE family	10.13	0.00	42503.86	1 (1 0 0 0 0)
YP_001039475 Cthe_3086 3-phosphoshikimate 1-carboxyvinyltransferase	10.13	0.00	45886.62	1 (1 0 0 0 0)
YP_001037139 Cthe_0711 chorismate mutase	10.13	0.00	13087.62	1 (1 0 0 0 0)
YP_001038906 Cthe_2511 MscS Mechanosensitive ion channel	10.13	0.00	31334.22	2 (2 0 0 0 0)
YP_001039277 Cthe_2886 imidazole glycerol phosphate synthase, glutamine amidotransferase subunit	10.13	4.90	22409.65	1 (1 0 0 0 0)
YP_001037436 Cthe_1011 Peptidoglycan glycosyltransferase	10.13	1.50	60568.02	1 (1 0 0 0 0)
YP_001039429 Cthe_3040 hypothetical protein	10.13	0.00	16338.30	1 (1 0 0 0 0)
YP_001036891 Cthe_0460 DNA topoisomerase I	10.13	0.00	79261.14	2 (2 0 0 0 0)
YP_001037318 Cthe_0891 response regulator receiver protein	10.13	6.10	16925.05	1 (1 0 0 0 0)
YP_001037785 Cthe_1360 polysaccharide biosynthesis protein CpsF	10.13	5.50	19315.26	1 (1 0 0 0 0)
YP_001037637 Cthe_1212 hypothetical protein	10.13	3.50	36834.32	1 (1 0 0 0 0)
YP_001039140 Cthe_2748 SsrA-binding protein	10.13	7.10	17851.74	1 (1 0 0 0 0)
YP_001037670 Cthe_1245 phosphoribosylamine--glycine ligase	10.13	0.00	46106.55	1 (1 0 0 0 0)
YP_001039506 Cthe_3118 hemerythrin-like metal-binding protein	10.13	10.10	15662.79	1 (1 0 0 0 0)
YP_001037056 Cthe_0628 hypothetical protein	10.13	0.00	31420.28	2 (2 0 0 0 0)
YP_001037794 Cthe_1369 histidine triad (HIT) protein	10.13	10.50	12631.70	3 (3 0 0 0 0)

YP_001037731 Cthe_1306 hypothetical protein	10.13	0.00	14990.78	3 (3 0 0 0 0)
YP_001037538 Cthe_1113 SEC-C motif containing protein	10.13	0.00	74335.48	1 (1 0 0 0 0)
YP_001037363 Cthe_0936 3-oxoacyl-(acyl-carrier-protein) synthase III	10.13	2.70	35975.23	1 (1 0 0 0 0)
YP_001038488 Cthe_2088 protein of unknown function DUF710	10.13	8.10	15702.12	3 (3 0 0 0 0)
YP_001038697 Cthe_2302 hypothetical protein	10.13	0.00	64955.30	1 (1 0 0 0 0)
YP_001037780 Cthe_1355 glycosyl transferase, family 2	10.13	2.90	36200.88	1 (1 0 0 0 0)
YP_001038972 Cthe_2579 iron-containing alcohol dehydrogenase	10.13	3.00	40879.99	2 (2 0 0 0 0)
YP_001036587 Cthe_0153 Holliday junction resolvase YqgF	10.13	7.80	15848.54	1 (1 0 0 0 0)
YP_001039278 Cthe_2887 phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase	10.13	0.00	25584.55	2 (2 0 0 0 0)
YP_001037012 Cthe_0583 response regulator receiver protein	10.13	2.00	43858.34	1 (1 0 0 0 0)
YP_001037984 Cthe_1565 Nitrogenase	10.13	0.00	54581.06	1 (1 0 0 0 0)
YP_001038120 Cthe_1706 hypothetical protein	10.13	4.40	42932.78	1 (1 0 0 0 0)
YP_001037887 Cthe_1465 Redoxin	10.13	0.00	16775.50	1 (1 0 0 0 0)
YP_001039507 Cthe_3119 flavin reductase-like, FMN-binding	10.13	4.10	24647.10	1 (1 0 0 0 0)
YP_001036Cthe_0214 phenylalanyl-tRNA synthetase, alpha subunit	10.13	2.70	38290.86	2 (2 0 0 0 0)
YP_001037412 Cthe_0987 riboflavin biosynthesis protein RibF	10.13	2.60	35892.63	1 (1 0 0 0 0)
YP_001036665 Cthe_0233 GCN5-related N-acetyltransferase	10.13	5.10	22946.86	1 (1 0 0 0 0)
YP_001038128 Cthe_1714 hypothetical protein	10.13	11.00	12935.55	1 (1 0 0 0 0)
YP_001038606 Cthe_2211 3-isopropylmalate dehydratase, large subunit	10.13	3.10	45122.21	1 (1 0 0 0 0)
YP_001039050 Cthe_2658 RNA-binding S4	10.13	13.90	8950.93	2 (2 0 0 0 0)
YP_001039210 Cthe_2818 CheA signal transduction histidine kinases	10.13	0.00	67789.99	1 (1 0 0 0 0)
YP_001037242 Cthe_0814 DNA repair protein RecN	10.13	1.40	64789.03	1 (1 0 0 0 0)
YP_001039545 Cthe_3157 pyruvate carboxyltransferase	10.13	2.20	52253.63	1 (1 0 0 0 0)
YP_001039476 Cthe_3087 response regulator receiver protein	10.13	4.50	30441.89	1 (1 0 0 0 0)
YP_001037775 Cthe_1350 single-strand binding protein	10.13	0.00	15435.49	1 (1 0 0 0 0)
YP_001039118 Cthe_2726 ribosomal protein L7Ae/L30e/S12e/Gadd45	10.13	11.40	8566.64	2 (2 0 0 0 0)

YP_001037188 Cthe_0760 hypothetical protein	10.13	0.00	61915.95	1 (1 0 0 0 0)
YP_001038136 Cthe_1722 phage Terminase	10.13	0.00	97269.15	1 (1 0 0 0 0)
YP_001037168 Cthe_0740 transcriptional regulator, GntR family	10.13	0.00	28146.60	1 (1 0 0 0 0)
YP_001038746 Cthe_2351 hypothetical protein	10.13	0.00	16448.10	1 (1 0 0 0 0)
YP_001037878 Cthe_1456 ABC transporter related protein	10.13	3.80	29346.55	1 (1 0 0 0 0)
YP_001037629 Cthe_1204 protein of unknown function DUF444	10.13	3.10	48132.70	1 (1 0 0 0 0)
Cthe_0225	10.13	10.00	13876.33	1 (1 0 0 0 0)
YP_001037368 Cthe_0941 CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	10.13	0.00	22551.58	1 (1 0 0 0 0)
YP_001038992 Cthe_2599 CMP/dCMP deaminase, zinc-binding	10.13	9.20	16954.66	1 (1 0 0 0 0)
YP_001036793 Cthe_0362 transcriptional regulator, AsnC family	10.13	6.30	18091.50	1 (1 0 0 0 0)
YP_00103Cthe_0046 hypothetical protein	10.13	1.30	102790.20	1 (1 0 0 0 0)
YP_001036580 Cthe_0146 ribonuclease R	10.13	0.00	86939.25	1 (1 0 0 0 0)
YP_001037757 Cthe_1332 histidyl-tRNA synthetase	10.13	2.60	46805.88	1 (1 0 0 0 0)
YP_001039430 Cthe_3041 UbiA prenyltransferase	10.13	0.00	36326.88	1 (1 0 0 0 0)
YP_001037263 Cthe_0835 hypothetical protein	10.13	10.70	9453.02	1 (1 0 0 0 0)
YP_001036751 Cthe_0320 hypothetical protein	10.13	2.50	55540.01	1 (1 0 0 0 0)
YP_001038652 Cthe_2257 ribosomal protein L9	10.13	0.00	16353.93	1 (1 0 0 0 0)
YP_001038452 Cthe_2052 hypothetical protein	10.13	0.00	15188.59	1 (1 0 0 0 0)
YP_001038252 Cthe_1838 glycoside hydrolase, family 10	10.13	1.80	69474.37	1 (1 0 0 0 0)
YP_001038997 Cthe_2604 ATP synthase F0, B subunit	10.13	5.50	21057.12	1 (1 0 0 0 0)
YP_001039003 Cthe_2610 hypothetical protein	10.12	0.00	66553.36	2 (1 0 0 1 0)
YP_001038115 Cthe_1701 Recombinase	10.11	2.60	49310.72	1 (1 0 0 0 0)
YP_001036515 Cthe_0081 arginine biosynthesis bifunctional protein ArgJ	10.11	3.20	42652.67	1 (1 0 0 0 0)
YP_001038413 Cthe_2004 transposase	10.10	2.40	57187.78	1 (1 0 0 0 0)
YP_001038226 Cthe_1812 Urease accessory protein UreD	8.20	10.70	29686.90	2 (1 1 0 0 0)
YP_001036738 Cthe_0307 hypothetical protein	8.17	24.00	14014.25	1 (0 1 0 0 0)
YP_001039425 Cthe_3036 methyl-accepting chemotaxis sensory transducer	8.16	0.00	62462.38	3 (0 3 0 0 0)
YP_00103Cthe_0003 ankyrin repeat protein	8.16	3.60	35561.51	20 (1 19 0 0 0)

YP_001039566 Cthe_3178 protein of unknown function DUF342	8.16	2.60	59841.41	6 (1 5 0 0 0)
YP_001038671 Cthe_2276 AAA ATPase, central region	8.16	0.00	43847.11	2 (0 2 0 0 0)
YP_001037465 Cthe_1040 DNA polymerase III, delta subunit	8.16	0.00	38875.81	1 (0 1 0 0 0)
YP_001038323 Cthe_1911 Carbohydrate binding family 6	8.16	1.20	146588.20	3 (0 1 2 0 0)
YP_001039568 Cthe_3180 hypothetical protein	8.16	0.00	35437.56	1 (0 1 0 0 0)
YP_001037223 Cthe_0795 alpha amylase, catalytic region	8.15	0.00	67491.74	9 (0 9 0 0 0)
YP_00103Cthe_0016 Ferritin and Dps	8.15	11.90	18594.23	1 (0 1 0 0 0)
Cthe_2134	8.15	6.30	20822.68	4 (0 1 2 1 0)
YP_001037504 Cthe_1079 Nucleotidyl transferase	8.15	0.00	91250.67	1 (0 1 0 0 0)
YP_001038368 Cthe_1957 extracellular solute-binding protein, family 1	8.15	3.60	51207.38	1 (0 1 0 0 0)
YP_001038740 Cthe_2345 DegT/DnrJ/EryC1/StrS aminotransferase	8.15	0.00	43377.93	1 (0 1 0 0 0)
YP_001038535 Cthe_2138 glycoside hydrolase, family 43	8.15	2.20	65074.28	1 (0 1 0 0 0)
YP_001036993 Cthe_05Trans-hexaprenyltranstransferase	8.15	3.50	35809.44	2 (0 1 0 1 0)
YP_001036663 Cthe_0231 Radical SAM	8.15	0.00	49723.59	1 (0 1 0 0 0)
YP_001037239 Cthe_0811 response regulator receiver protein	8.15	7.70	13041.80	2 (0 2 0 0 0)
YP_001038019 Cthe_1600 two component transcriptional regulator, winged helix family	8.15	0.00	25125.33	1 (0 1 0 0 0)
YP_001039154 Cthe_2762 transcriptional regulator-like protein	8.15	0.00	52842.75	1 (0 1 0 0 0)
YP_001039042 Cthe_2650 polysaccharide biosynthesis protein CapD	8.14	0.00	68185.34	1 (0 1 0 0 0)
YP_001038357 Cthe_1946 FAD-dependent pyridine nucleotide-disulphide oxidoreductase	8.14	2.60	47018.04	1 (0 1 0 0 0)
YP_001036763 Cthe_0332 phosphoribulokinase/uridine kinase	8.14	2.20	64507.49	3 (0 3 0 0 0)
YP_001038684 Cthe_2289 hypothetical protein	8.14	0.00	27862.32	1 (0 1 0 0 0)
YP_001038171 Cthe_1757 peptidase M23B	8.14	0.00	34708.10	3 (0 3 0 0 0)
YP_001039094 Cthe_2702 polysaccharide pyruvyl transferase	8.14	0.00	84022.88	1 (0 1 0 0 0)
YP_001037913 Cthe_1492 NAD(P)H dehydrogenase (quinone)	8.14	0.00	61292.02	4 (0 4 0 0 0)
YP_001038436 Cthe_2032 hypothetical protein	8.14	0.00	223041.40	2 (0 2 0 0 0)
YP_001036668 Cthe_0236 aldo/keto reductase	8.14	0.00	42386.54	2 (0 1 1 0 0)
YP_001038433 Cthe_2029 hypothetical protein	8.14	5.20	31588.09	2 (1 1 0 0 0)

YP_001037495 Cthe_1070 metal dependent phosphohydrolase	8.14	0.00	82339.86	1 (0 1 0 0 0)
YP_001036874 Cthe_0443 protein of unknown function DUF1290	8.14	12.00	13788.51	1 (0 1 0 0 0)
YP_001037473 Cthe_1048 Rhomboid-like protein	8.14	0.00	57844.07	1 (0 1 0 0 0)
YP_001039525 Cthe_3137 hypothetical protein	8.14	0.00	21252.79	1 (0 1 0 0 0)
YP_001038729 Cthe_2334 polysaccharide biosynthesis protein CapD	8.14	0.00	68628.90	5 (0 5 0 0 0)
YP_001037566 Cthe_1141 hypothetical protein	8.14	0.70	157008.00	3 (0 3 0 0 0)
YP_001038153 Cthe_1739 SNF2-related protein	8.14	0.00	44230.29	1 (0 1 0 0 0)
YP_001038806 Cthe_2411 metallophosphoesterase	8.14	0.00	43678.01	2 (0 2 0 0 0)
Cthe_2984	8.14	6.30	22648.09	1 (0 1 0 0 0)
YP_001036923 Cthe_0492 CheC, inhibitor of MCP methylation	8.14	0.00	21787.68	2 (0 2 0 0 0)
YP_001037749 Cthe_1324 heat-inducible transcription repressor HrcA	8.14	0.00	39225.57	3 (0 3 0 0 0)
YP_001038243 Cthe_1829 Chromosome segregation ATPases-like protein	8.14	0.00	124587.60	1 (0 1 0 0 0)
YP_001037594 Cthe_1169 Beta-lactamase class A-like protein	8.14	0.00	44134.89	3 (0 3 0 0 0)
YP_001036630 Cthe_0196 glutamine synthetase, catalytic region	8.14	0.00	78506.49	1 (0 1 0 0 0)
YP_001037449 Cthe_1024 small GTP-binding protein	8.13	3.40	49796.14	1 (0 1 0 0 0)
YP_001038453 Cthe_2053 hypothetical protein	8.13	0.00	47377.91	1 (0 1 0 0 0)
YP_001036963 Cthe_0534 ABC-type bacteriocin transporter	8.13	0.00	80764.15	1 (0 1 0 0 0)
YP_001038758 Cthe_2363 methyltransferase GidB	8.13	0.00	27335.71	2 (0 2 0 0 0)
YP_001038383 Cthe_1972 hypothetical protein	8.13	0.00	12143.36	2 (0 2 0 0 0)
YP_001036865 Cthe_0434 ATPase	8.13	3.00	41745.81	2 (0 2 0 0 0)
YP_001036984 Cthe_0555 PpiC-type peptidyl-prolyl cis-trans isomerase	8.13	0.00	47318.00	1 (0 1 0 0 0)
YP_001036838 Cthe_0407 Radical SAM	8.13	0.00	49617.13	1 (0 1 0 0 0)
YP_001039590 Cthe_3202 CRISPR-associated protein, Csh2 family	8.13	0.00	34861.18	1 (0 1 0 0 0)
YP_001037456 Cthe_1031 glutamyl-tRNA(Gln) amidotransferase, B subunit	8.13	0.00	54020.93	1 (0 1 0 0 0)
YP_001038725 Cthe_2330 hypothetical protein	8.13	0.00	57264.16	1 (0 1 0 0 0)
YP_001036704 Cthe_0272 Serine-type D-Ala-D-Ala carboxypeptidase	8.13	0.00	42145.81	2 (0 2 0 0 0)
YP_001037Cthe_1224 ribosomal protein	8.13	21.50	7351.37	1 (0 1 0 0 0)

L35				
YP_001038367 Cthe_1956 Tagatose-6-phosphate kinase	8.13	4.80	33741.50	1 (0 1 0 0 0)
YP_001036550 Cthe_0116 protein of unknown function DUF199	8.13	0.00	35701.01	1 (0 1 0 0 0)
YP_001038634 Cthe_2239 carbon storage regulator, CsrA	8.13	0.00	8649.78	1 (0 1 0 0 0)
YP_001039365 Cthe_2975 DNA-directed RNA polymerase sigma factor	8.13	0.00	30081.62	3 (0 3 0 0 0)
YP_001037178 Cthe_0750 spermidine/putrescine ABC transporter ATPase subunit	8.13	2.80	39800.65	2 (1 1 0 0 0)
YP_001038660 Cthe_2265 H+-transporting two-sector ATPase, C (AC39) subunit	8.13	0.00	38965.83	3 (0 3 0 0 0)
YP_001038239 Cthe_1825 multi-sensor hybrid histidine kinase	8.13	1.80	106314.10	1 (0 1 0 0 0)
YP_001037488 Cthe_1063 thiamine biosynthesis/tRNA modification protein ThiI	8.13	0.00	44373.72	1 (0 1 0 0 0)
YP_001036570 Cthe_0136 4'-phosphopantetheinyl transferase	8.13	5.80	28027.55	1 (0 1 0 0 0)
YP_001038Cthe_2249 helicase, RecD/TraA family	8.13	2.00	84297.03	1 (0 1 0 0 0)
YP_001037231 Cthe_0803 pseudouridine synthase, RluA family	8.13	0.00	35832.23	1 (0 1 0 0 0)
YP_001038Cthe_2247 regulatory protein, MerR	8.13	0.00	16866.37	3 (0 3 0 0 0)
YP_001039284 Cthe_2893 response regulator receiver protein	8.13	3.70	40751.88	1 (0 1 0 0 0)
YP_001037586 Cthe_1161 hypothetical protein	8.13	0.00	93283.76	2 (0 1 0 1 0)
YP_001036901 Cthe_0470 hypothetical protein	8.13	0.00	31496.69	1 (0 1 0 0 0)
YP_001036895 Cthe_04flagellar hook-basal body complex subunit FliE	8.13	0.00	11494.90	1 (0 1 0 0 0)
YP_001037582 Cthe_1157 hypothetical protein	8.13	12.60	11318.65	3 (0 3 0 0 0)
YP_001038712 Cthe_2317 response regulator receiver protein	8.13	0.00	54371.08	2 (0 2 0 0 0)
YP_001037002 Cthe_0573 protein serine/threonine phosphatases	8.13	0.00	27359.65	1 (0 1 0 0 0)
YP_001036794 Cthe_0363 aminotransferase, class I and II	8.13	4.60	43954.24	1 (0 1 0 0 0)
YP_001038727 Cthe_2332 multi-sensor signal transduction histidine kinase	8.13	2.50	68353.57	1 (0 1 0 0 0)
YP_001038917 Cthe_2522 membrane associated protein	8.13	2.60	61939.47	2 (0 1 0 1 0)
YP_001037579 Cthe_1154 transcriptional regulator, XRE family	8.13	0.00	9614.31	1 (0 1 0 0 0)
YP_001038974 Cthe_2581 dihydropteroate synthase	8.13	3.60	43330.46	1 (0 1 0 0 0)

YP_001038602 Cthe_2207 isoaspartyl dipeptidase	8.13	3.60	41918.98	1 (0 1 0 0 0)
YP_001037474 Cthe_1049 regulatory protein RecX	8.13	0.00	24532.88	1 (0 1 0 0 0)
YP_001037243 Cthe_0815 arginine repressor, ArgR	8.13	7.80	17053.89	1 (0 1 0 0 0)
YP_001037250 Cthe_0822 IstB-like ATP-binding protein	8.13	5.40	27700.50	1 (0 1 0 0 0)
YP_0010382Cthe_1850 hypothetical protein	8.13	0.00	17668.95	1 (0 1 0 0 0)
YP_001038341 Cthe_1930 carboxyl-terminal protease	8.13	0.00	55784.72	1 (0 1 0 0 0)
YP_001036945 Cthe_0515 transposase IS66	8.13	0.00	59400.95	1 (0 1 0 0 0)
YP_001039533 Cthe_3145 transcriptional repressor, CopY family	8.13	0.00	14033.30	1 (0 1 0 0 0)
YP_001038187 Cthe_1773 peptidase S16, lon-like protein	8.13	0.00	91962.27	1 (0 1 0 0 0)
YP_001039580 Cthe_3192 hypothetical protein	8.13	0.00	24663.38	1 (0 1 0 0 0)
YP_001038561 Cthe_2166 putative PAS/PAC sensor protein	8.13	0.00	43255.23	1 (0 1 0 0 0)
YP_001036635 Cthe_0201 glutamate synthase, alpha subunit-like protein	8.13	4.20	26583.20	1 (0 1 0 0 0)
YP_001038389 Cthe_1978 hypothetical protein	8.13	0.00	9514.85	1 (0 1 0 0 0)
YP_001038820 Cthe_2425 MotA/TolQ/ExbB proton channel	8.13	0.00	30827.41	1 (0 1 0 0 0)
YP_001037668 Cthe_1243 GCN5-related N-acetyltransferase	8.13	0.00	17625.19	1 (0 1 0 0 0)
YP_001039625 Cthe_3237 Relaxase/mobilization nuclease domain containing protein	8.13	2.20	53453.32	1 (0 1 0 0 0)
YP_001037208 Cthe_0780 HAD-superfamily hydrolase, subfamily IA, variant 3	8.13	0.00	26253.77	1 (0 1 0 0 0)
YP_001038348 Cthe_1937 glutamate racemase	8.13	6.70	29790.56	1 (0 1 0 0 0)
YP_001037689 Cthe_1264 DNA polymerase III, alpha subunit	8.13	0.00	134313.30	3 (1 2 0 0 0)
YP_001038026 Cthe_1609 Recombinase	8.13	0.00	60124.05	1 (0 1 0 0 0)
YP_001039351 Cthe_2961 extracellular solute-binding protein, family 5	8.13	0.00	60372.57	1 (0 1 0 0 0)
YP_001038914 Cthe_2519 2-isopropylmalate synthase/homocitrate synthase family protein	8.11	1.10	59595.77	2 (1 1 0 0 0)
YP_001037278 Cthe_0851 hypothetical protein	8.11	0.00	9941.12	1 (0 1 0 0 0)
YP_001037874 Cthe_1451 GCN5-related N-acetyltransferase	8.11	0.00	18051.00	1 (0 1 0 0 0)
YP_001038813 Cthe_2418 ATPase	8.10	0.00	40165.99	1 (0 1 0 0 0)
YP_001038678 Cthe_2283 methyl-accepting chemotaxis sensory transducer	6.17	0.00	154468.80	5 (0 2 3 0 0)

YP_001039371 Cthe_2981 hypothetical protein	6.16	5.50	19707.33	4 (0 0 4 0 0)
YP_001037948 Cthe_1527 periplasmic sensor signal transduction histidine kinase	6.16	2.90	40255.13	4 (0 3 1 0 0)
YP_001037463 Cthe_1038 spore protease	6.16	0.00	37867.05	1 (0 0 1 0 0)
YP_001039373 Cthe_2983 methyl-accepting chemotaxis sensory transducer	6.15	0.00	61326.02	4 (0 2 2 0 0)
YP_001036Cthe_0208 single-stranded-DNA-specific exonuclease RecJ	6.15	0.00	86561.65	1 (0 0 1 0 0)
YP_0010387Cthe_2369 ribonuclease P protein component	6.14	0.00	15109.09	2 (0 0 2 0 0)
YP_001038224 Cthe_1810 peptidase M56, BlaR1	6.14	0.00	85821.74	2 (0 0 2 0 0)
YP_001039087 Cthe_2695 hypothetical protein	6.14	0.00	25498.29	3 (0 0 3 0 0)
YP_001038957 Cthe_2563 hypothetical protein	6.14	0.00	68880.94	1 (0 0 1 0 0)
YP_001036719 Cthe_0287 multi-sensor hybrid histidine kinase	6.14	1.40	132798.20	1 (0 0 1 0 0)
YP_001038290 Cthe_1876 Tn7-like transposition protein A	6.14	0.00	32890.93	2 (0 0 2 0 0)
YP_001037498 Cthe_1073 protein of unknown function DUF1429	6.14	0.00	12049.66	6 (0 4 2 0 0)
YP_001036834 Cthe_0403 DNA-directed RNA polymerase sigma factor	6.14	0.00	32431.13	2 (0 1 1 0 0)
YP_001037337 Cthe_0910 extracellular solute-binding protein, family 5	6.14	0.00	63116.98	2 (0 0 1 0 1)
YP_001036918 Cthe_0487 Cobyrinic acid a,c-diamide synthase	6.14	0.00	33111.78	1 (0 0 1 0 0)
YP_001038495 Cthe_2095 hydrolase, TatD family	6.14	5.90	29094.93	1 (0 0 1 0 0)
YP_001039355 Cthe_2965 binding-protein-dependent transport systems inner membrane component	6.14	5.20	33754.50	1 (0 0 1 0 0)
YP_001038781 Cthe_2386 VanW	6.14	0.00	62732.77	1 (0 0 1 0 0)
YP_001036928 Cthe_0498 protein of unknown function DUF342	6.14	0.00	59572.50	1 (0 0 1 0 0)
YP_001038951 Cthe_2557 glycosyl transferase, family 2	6.13	0.00	57454.98	3 (0 0 1 1 1)
YP_001038877 Cthe_2482 hypothetical protein	6.13	11.80	15546.20	2 (0 1 1 0 0)
YP_001037053 Cthe_0625 glycoside hydrolase, family 9	6.13	2.80	79760.16	1 (0 0 1 0 0)
YP_001037054 Cthe_0626 hypothetical protein	6.13	0.00	52981.45	2 (0 0 2 0 0)
YP_001039240 Cthe_2848 phage major capsid protein, HK97	6.13	0.00	48301.61	2 (0 0 2 0 0)
YP_001039597 Cthe_3209 hypothetical protein	6.13	0.00	46140.90	2 (0 0 2 0 0)

YP_001036896 Cthe_0465 flagellar M-ring protein FliF	6.13	0.00	57786.86	1 (0 0 1 0 0)
YP_001038731 Cthe_2336 glycosyl transferase, group 1	6.13	0.00	43607.84	1 (0 0 1 0 0)
YP_001037Cthe_1217 ATP-dependent Clp protease, ATP-binding subunit clpA	6.13	0.00	88140.23	1 (0 0 1 0 0)
YP_00103Cthe_0028 uncharacterized protein, YcgL-like protein	6.13	4.20	39198.25	1 (0 0 1 0 0)
YP_001038627 Cthe_2232 polysaccharide biosynthesis protein CapD	6.13	0.00	37398.55	1 (0 0 1 0 0)
YP_001038278 Cthe_18acetylglutamate kinase	6.12	3.00	33029.62	1 (0 0 1 0 0)
YP_001037883 Cthe_1461 FAD dependent oxidoreductase	4.14	2.10	58448.48	4 (0 0 1 3 0)
YP_001038491 Cthe_2091 hypothetical protein	4.14	2.50	49445.48	1 (0 0 0 1 0)
YP_001039253 Cthe_2862 hypothetical protein	4.14	0.00	23932.85	2 (0 0 1 1 0)
YP_001036563 Cthe_0129 metal dependent phosphohydrolase	4.14	0.00	58142.20	2 (0 0 0 2 0)
YP_001039004 Cthe_2611 Fibronectin, type III	4.14	0.00	183582.80	2 (0 0 0 1 1)
YP_001037748 Cthe_1323 GrpE protein	4.14	0.00	25432.41	2 (0 0 1 1 0)
YP_001037133 Cthe_0705 hypothetical protein	4.13	0.00	9275.10	2 (0 0 0 2 0)
YP_001038979 Cthe_2586 amidohydrolase 2	4.13	0.00	29355.27	1 (0 0 0 1 0)
YP_001038506 Cthe_2106 protein of unknown function DUF327	4.13	0.00	17618.43	1 (0 0 0 1 0)
YP_001038732 Cthe_2337 hypothetical protein	4.13	0.00	49997.88	1 (0 0 0 1 0)
YP_001039158 Cthe_2766 hypothetical protein	2.15	0.00	8968.65	1 (0 0 0 0 1)
YP_001037268 Cthe_0840 hypothetical protein	2.14	0.00	24288.16	2 (0 1 0 0 1)
YP_001038149 Cthe_1735 phage DNA polymerase	2.14	0.00	66028.44	4 (0 0 0 2 2)
YP_001039548 Cthe_3160 putative RNA methylase, NOL1/NOP2/sun family	2.14	0.00	51578.11	2 (0 1 0 0 1)
YP_001039167 Cthe_2775 transcriptional regulator, Crp/Fnr family	2.14	6.90	26629.95	1 (0 0 0 0 1)
YP_001037234 Cthe_0806 PAS/PAC sensor hybrid histidine kinase	2.14	0.00	102250.20	2 (0 0 0 0 2)
YP_001039518 Cthe_3130 hypothetical protein	2.14	0.00	28162.02	1 (0 0 0 0 1)
YP_001036951 Cthe_0521 helicase-like protein	2.13	0.50	126703.70	1 (0 0 0 0 1)
YP_001037421 Cthe_0996 DNA polymerase III, alpha subunit	2.13	0.70	163280.50	1 (0 0 0 0 1)
YP_001037660 Cthe_1235 Cellulose 1,4-beta-cellobiosidase	2.13	0.00	758397.60	1 (0 0 0 0 1)

YP_001038240 Cthe_1826 response regulator receiver sensor signal transduction histidine kinase	2.13	0.00	41724.54	1 (0 0 0 0 1)
YP_001038920 Cthe_2525 glutamyl-tRNA reductase	2.13	3.30	38000.81	1 (0 0 0 0 1)

[A] XC, cross correlation score. Xc score as determined by Bioworks 3.3 program. The Xc score represents the cross correlation value between the observed peptide fragment mass spectrum and the one theoretically predicted. [b] Total peptide hits observed for the proteins (columns within parethesis represent the number of primary, secondary, tertiary,etc. peptide hits based spectral quality of the peptides spectra).

Appendix E; Membrane fraction protien identifications of Avicel grown cells.

Table E1. Complete list of proteins identified in the membrane fraction of *C. thermocellum* grown on Avicel, ranked by XC score.

Protein	Score (XC)[a]	Coverage (%)	MW (daltons)	Peptide (Hits)[b]
YP_001036854 Cthe_0423 iron-containing alcohol dehydrogenase	698.31	73.10	95948.2	91 (89 1 1 0 0)
YP_001037445 Cthe_1020 extracellular solute-binding protein, family 1	380.29	65.6	49954	49 (49 0 0 0 0)
YP_001038276 Cthe_1862 ABC transporter related protein	280.25	55.10	42036.9	31 (31 0 0 0 0)
YP_001039508 Cthe_3120 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	266.26	24.90	130938.9	28 (27 0 1 0 0)
YP_001036772 Cthe_0341 NADH dehydrogenase quinone	260.25	42.90	64854.81	28 (28 0 0 0 0)
YP_001036833 Cthe_0402 copper amine oxidase-like protein	238.22	40.3	83083	25 (24 1 0 0 0)
YP_001037127 Cthe_0699 carboxyl transferase	230.21	47.50	56012.8	25 (25 0 0 0 0)
YP_001039122 Cthe_2730 translation elongation factor Tu	208.23	59.30	44160.6	23 (22 1 0 0 0)
YP_001036780 Cthe_0349 fructose-1,6-bisphosphate aldolase, class II	200.24	65.00	33611.2	23 (23 0 0 0 0)
YP_001039121 Cthe_2729 translation elongation factor G	190.20	36.30	77586.2	19 (19 0 0 0 0)
YP_001036571 Cthe_0137 glyceraldehyde-3-phosphate dehydrogenase, type I	180.27	64.00	36116.4	24 (24 0 0 0 0)
YP_001038743 Cthe_2348 Ig-related protein	180.22	21.0	113261	19 (19 0 0 0 0)
YP_001038324 Cthe_1912 copper amine oxidase-like protein	180.21	40.2	59215	19 (19 0 0 0 0)
YP_001036773 Cthe_0342 hydrogenase, Fe-only	176.20	37.80	63972.4	18 (16 2 0 0 0)
YP_001036855 Cthe_0424 aminoglycoside phosphotransferase	170.24	64.70	28568.9	20 (20 0 0 0 0)
YP_001037810 Cthe_1385 preprotein translocase, SecA subunit	170.21	21.90	104546.4	18 (18 0 0 0 0)
YP_001037733 Cthe_1308 pyruvate, phosphate dikinase	164.18	23.20	98703.0	17 (15 1 1 0 0)

YP_001038787 Cthe_2392 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	160.21	39.30	43405.9	18 (18 0 0 0 0)
YP_001037285 Cthe_0858 protein of unknown function DUF1432	150.27	46.8	35281	18 (18 0 0 0 0)
YP_001038999 Cthe_2606 ATP synthase F1, alpha subunit	150.26	35.20	55784.9	16 (16 0 0 0 0)
YP_001039469 Cthe_3080 cellulosome anchoring protein, cohesin region	150.23	34.5	48500	18 (18 0 0 0 0)
YP_001038997 Cthe_2604 ATP synthase F0, B subunit	150.23	49.70	21057.1	16 (16 0 0 0 0)
YP_001036510 Cthe_0076 hypothetical protein	140.33	64.3	13700	19 (19 0 0 0 0)
YP_001038788 Cthe_2393 thiamine pyrophosphate enzyme-like TPP-binding	140.22	50.80	34626.8	16 (16 0 0 0 0)
YP_001036572 Cthe_0138 Phosphoglycerate kinase	140.21	44.60	42735.3	16 (16 0 0 0 0)
YP_001039283 Cthe_2892 chaperonin GroEL	140.19	35.30	57438.5	14 (14 0 0 0 0)
YP_001037523 Cthe_1098 hypothetical protein	130.21	33.8	49674	14 (14 0 0 0 0)
YP_00103 Cthe_0053 ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent	130.19	16.90	88248.5	15 (15 0 0 0 0)
YP_001039294 Cthe_2903 ribosomal protein L3	128.20	43.90	23551.7	14 (13 1 0 0 0)
YP_001038631 Cthe_2236 flagellin-like protein	124.27	39.70	29549.1	18 (16 1 1 0 0)
YP_001037331 Cthe_0904 protein-export membrane protein SecD	120.24	27.9	48261	14 (14 0 0 0 0)
YP_001039113 Cthe_2721 ribosomal protein L1	120.21	37.70	25057.5	14 (14 0 0 0 0)
YP_001039265 Cthe_2874 Phosphoenolpyruvate carboxykinase GTP	120.19	25.10	67583.8	13 (13 0 0 0 0)
YP_001039614 Cthe_3226 copper amine oxidase-like protein	116.22	38.2	29604	14 (11 3 0 0 0)
YP_001038778 Cthe_2383 hypothetical protein	110.23	41.8	33182	12 (12 0 0 0 0)
YP_001037732 Cthe_1307 cellulosome anchoring protein, cohesin region	110.23	19.7	68577	12 (12 0 0 0 0)
YP_001036778 Cthe_0347 phosphofructokinase	110.22	40.70	45595.2	12 (12 0 0 0 0)
YP_001039117 Cthe_2725 DNA-directed RNA polymerase, beta subunit	110.18	9.40	130074.0	11 (11 0 0 0 0)
YP_001039323 Cthe_2932 DNA-directed RNA polymerase, alpha subunit	110.16	32.40	35017.3	11 (11 0 0 0 0)
YP_001036775 Cthe_0344 Malate dehydrogenase	100.23	28.70	42073.3	11 (11 0 0 0 0)
YP_001036989 Cthe_0560 FAD-dependent pyridine nucleotide-disulphide oxidoreductase	100.22	21.00	66711.5	11 (11 0 0 0 0)

YP_001039295 Cthe_2904 ribosomal protein L4/L1e	100.20	47.10	23295.8	12 (12 0 0 0 0)
YP_001036577 Cthe_0143 Phosphopyruvate hydratase	100.18	33.00	47073.2	10 (10 0 0 0 0)
YP_001037430 Cthe_1005 translation elongation factor Ts	98.21	37.70	24321.7	11 (10 1 0 0 0)
YP_001037793 Cthe_1368 S-layer-like domain containing protein	96.22	11.4	78075	14 (12 2 0 0 0)
YP_001039306 Cthe_2915 ribosomal protein L5	90.24	51.60	20391.9	12 (12 0 0 0 0)
YP_001038197 Cthe_1783 ribosomal protein L13	90.22	54.50	16763.2	12 (12 0 0 0 0)
YP_0010372 Cthe_0836 hypothetical protein	90.22	33.5	20373	10 (10 0 0 0 0)
YP_001038803 Cthe_2408 phage shock protein A, PspA	90.19	37.60	27739.4	10 (10 0 0 0 0)
YP_001037142 Cthe_0714 hydroxymethylbutenyl pyrophosphate reductase	90.16	14.60	77377.3	9 (9 0 0 0 0)
YP_001039466 Cthe_3077 cellulosome anchoring protein, cohesin region	80.21	17.7	196710	9 (9 0 0 0 0)
YP_001037530 Cthe_1105 type II secretion system protein	80.21	30.00	45131.5	12 (12 0 0 0 0)
YP_001039467 Cthe_3078 cellulosome anchoring protein, cohesin region	80.21	5.3	248014	13 (13 0 0 0 0)
YP_001038938 Cthe_2543 Spore germination protein-like protein	80.20	19.40	52690.6	9 (9 0 0 0 0)
YP_001039322 Cthe_2931 ribosomal protein S4	80.20	28.40	23926.7	8 (8 0 0 0 0)
YP_001039300 Cthe_2909 ribosomal protein S3	80.19	41.20	25312.8	9 (9 0 0 0 0)
YP_001039308 Cthe_2917 ribosomal protein S8	80.18	50.80	14209.8	10 (10 0 0 0 0)
YP_001036935 Cthe_0505 formate acetyltransferase	80.18	11.30	84351.2	9 (9 0 0 0 0)
YP_001037193 Cthe_0765 ribosomal protein L19	80.17	55.30	12924.2	8 (8 0 0 0 0)
YP_001037527 Cthe_1102 fimbrial assembly protein	70.28	37.20	19870.7	8 (8 0 0 0 0)
YP_001038241 Cthe_1827 copper amine oxidase-like protein	70.23	22.6	36389	7 (7 0 0 0 0)
YP_001037849 Cthe_1425 Inorganic diphosphatase	70.22	8.50	71337.7	7 (7 0 0 0 0)
YP_001036574 Cthe_0140 phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	70.22	15.30	57330.0	9 (9 0 0 0 0)
YP_001038785 Cthe_2390 pyruvate/ketoisovalerate oxidoreductase, gamma subunit	70.21	41.10	21190.1	9 (9 0 0 0 0)
YP_001039297 Cthe_2906 ribosomal protein L2	70.20	25.80	30176.3	9 (9 0 0 0 0)

YP_001037330 Cthe_0903 protein-export membrane protein SecF	70.20	19.50	34990.8	8 (8 0 0 0 0)
YP_001039484 Cthe_3096 hypothetical protein	70.20	24.4	33802	9 (9 0 0 0 0)
YP_001037924 Cthe_1503 hypothetical protein	70.19	66.40	13222.8	8 (8 0 0 0 0)
YP_001036776 Cthe_0345 L-lactate dehydrogenase	70.16	18.9	34811	7 (7 0 0 0 0)
YP_001038490 Cthe_2090 hypothetical protein	62.16	25.00	21466.3	8 (7 0 0 0 1)
YP_001039120 Cthe_2728 ribosomal protein S7	60.22	39.10	17896.5	6 (6 0 0 0 0)
YP_001039324 Cthe_2933 ribosomal protein L17	60.20	38.50	19695.6	7 (7 0 0 0 0)
YP_001039201 Cthe_2809 glycoside hydrolase, family 16	60.20	7.6	147695	6 (6 0 0 0 0)
YP_001039299 Cthe_2908 ribosomal protein L22	60.20	51.20	14066.9	7 (7 0 0 0 0)
YP_001039468 Cthe_3079 cellulosome anchoring protein, cohesin region	60.20	12.8	74925	7 (7 0 0 0 0)
YP_001038817 Cthe_2422 hypothetical protein	60.20	25.2	40412	6 (6 0 0 0 0)
YP_001038375 Cthe_19FAD-dependent pyridine nucleotide-disulphide oxidoreductase	60.19	13.90	55740.1	6 (6 0 0 0 0)
YP_001038913 Cthe_2518 ketol-acid reductoisomerase	60.19	20.50	36199.4	6 (6 0 0 0 0)
YP_001039320 Cthe_2929 ribosomal protein S13	60.19	35.00	14158.9	7 (7 0 0 0 0)
YP_001037526 Cthe_1101 hypothetical protein	60.19	10.00	48018.2	7 (7 0 0 0 0)
YP_001039615 Cthe_3227 copper amine oxidase-like protein	60.17	15.0	30241	7 (7 0 0 0 0)
YP_001038376 Cthe_1965 alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	60.17	40.60	20791.5	6 (6 0 0 0 0)
YP_001036771 Cthe_0340 ferredoxin	56.18	45.10	13424.0	8 (7 0 1 0 0)
YP_001038752 Cthe_2357 hypothetical protein	54.16	20.3	44618	6 (5 0 0 1 0)
YP_001039310 Cthe_2919 ribosomal protein L18	50.23	36.90	13690.6	5 (5 0 0 0 0)
YP_001036849 Cthe_0418 Polyribonucleotide nucleotidyltransferase	50.22	9.10	77269.9	5 (5 0 0 0 0)
YP_001038912 Cthe_2517 acetolactate synthase, small subunit	50.20	30.00	18609.0	5 (5 0 0 0 0)
YP_001037662 Cthe_1237 leucyl-tRNA synthetase	50.20	8.00	94825.2	5 (5 0 0 0 0)
YP_001036594 Cthe_0160 ribosomal protein L21	50.19	51.50	11531.2	5 (5 0 0 0 0)
YP_001039311 Cthe_2920 ribosomal protein S5-like protein	50.18	49.40	17531.7	7 (7 0 0 0 0)

YP_001039116 Cthe_2724 DNA-directed RNA polymerase, beta subunit	50.18	4.70	139869.0	5 (5 0 0 0 0)
YP_00103Cthe_0045 copper amine oxidase-like protein	50.18	27.5	30545	6 (6 0 0 0 0)
YP_001038939 Cthe_2544 RNA polymerase, sigma-24 subunit, ECF subfamily	50.18	26.70	22184.4	5 (5 0 0 0 0)
YP_001039482 Cthe_3094 glycosyl transferase, family 2	50.17	13.90	37622.8	6 (6 0 0 0 0)
YP_001039100 Cthe_2708 hypothetical protein	50.17	12.5	53617	5 (5 0 0 0 0)
YP_001039001 Cthe_2608 ATP synthase F1, beta subunit	50.17	14.00	50980.7	5 (5 0 0 0 0)
YP_001037630 Cthe_1205 putative serine protein kinase, PrkA	50.17	9.30	74227.2	5 (5 0 0 0 0)
YP_001039483 Cthe_3095 glycosyl transferase, family 39	50.17	6.5	110274	5 (5 0 0 0 0)
YP_001039304 Cthe_2913 ribosomal protein L14	50.17	33.60	13446.4	6 (6 0 0 0 0)
YP_001036549 Cthe_0115 glycogen debranching enzyme, putative	50.16	6.40	76556.8	5 (5 0 0 0 0)
YP_001037129 Cthe_0701 conserved carboxylase region containing protein	50.15	14.60	52338.2	5 (5 0 0 0 0)
YP_001038580 Cthe_2185 ribosomal protein S18	50.15	43.60	10856.1	5 (5 0 0 0 0)
YP_001036860 Cthe_0429 NADH dehydrogenase quinone	48.20	10.60	68080.16	5 (4 1 0 0 0)
YP_001036717 Cthe_0285 isocitrate dehydrogenase, NADP-dependent	48.17	13.40	45753.0	6 (5 1 0 0 0)
YP_001037444 Cthe_1019 binding-protein-dependent transport systems inner membrane component	46.15	12.20	37355.9	5 (4 0 1 0 0)
YP_001038196 Cthe_1782 ribosomal protein S9	44.17	33.10	14656.2	5 (4 0 0 1 0)
YP_001037676 Cthe_1251 Xanthine/uracil/vitamin C permease	42.18	6.9	49394	6 (5 0 0 0 1)
YP_001039305 Cthe_2914 ribosomal protein L24	40.23	39.70	12808.0	4 (4 0 0 0 0)
YP_001036620 Cthe_0186 UDP-glucose 4-epimerase	40.22	11.60	39089.4	4 (4 0 0 0 0)
YP_001039321 Cthe_2930 ribosomal protein S11	40.21	21.60	14342.7	5 (5 0 0 0 0)
YP_001039115 Cthe_2723 ribosomal protein L7/L12	40.21	27.10	13287.3	6 (5 1 0 0 0)
YP_001039114 Cthe_2722 ribosomal protein L10	40.21	24.20	19501.7	6 (6 0 0 0 0)
YP_001038247 Cthe_1833 copper amine oxidase-like protein	40.20	20.4	30568	4 (4 0 0 0 0)
YP_001036715 Cthe_0283 aldo/keto reductase	40.20	22.90	35432.5	4 (4 0 0 0 0)

YP_001039313 Cthe_2922 ribosomal protein L15	40.20	34.20	15576.6	5 (5 0 0 0 0)
YP_001037525 Cthe_1100 prepilin-type cleavage/methylation	40.19	22.50	19349.2	4 (4 0 0 0 0)
YP_001039301 Cthe_2910 ribosomal protein L16	40.18	46.90	16213.7	4 (4 0 0 0 0)
YP_001039266 Cthe_2875 sigma 54 modulation protein/ribosomal protein S30EA	40.18	36.70	20531.7	4 (4 0 0 0 0)
YP_001037887 Cthe_1465 Redoxin	40.17	32.00	16775.5	4 (4 0 0 0 0)
YP_001039510 Cthe_3122 S-layer-like domain containing protein	40.17	8.2	77594	4 (4 0 0 0 0)
YP_001039416 Cthe_3027 Citrate Si -synthase	40.17	11.40	52000.7	4 (4 0 0 0 0)
YP_001036861 Cthe_0430 hydrogenase, Fe-only	40.17	6.40	62491.4	4 (4 0 0 0 0)
YP_001038996 Cthe_2603 ATP synthase F0, C subunit	40.17	21.9	7218	4 (4 0 0 0 0)
YP_001039101 Cthe_2709 hypothetical protein	40.17	11.0	55116	4 (4 0 0 0 0)
YP_001037653 Cthe_1228 threonyl-tRNA synthetase	40.17	7.10	73518.6	4 (4 0 0 0 0)
YP_001036782 Cthe_0351 PpiC-type peptidyl-prolyl cis-trans isomerase	40.16	8.9	53834	4 (4 0 0 0 0)
YP_001036792 Cthe_0361 hypothetical protein	40.16	16.50	21979.2	4 (4 0 0 0 0)
YP_001037431 Cthe_1006 ribosomal protein S2	40.15	13.10	28590.0	4 (4 0 0 0 0)
YP_001039314 Cthe_2923 preprotein translocase, SecY subunit	38.18	7.00	47339.9	4 (3 1 0 0 0)
YP_001038466 Cthe_2066 serine O-acetyltransferase	34.19	20.20	27382.2	4 (2 1 1 0 0)
YP_001037128 Cthe_0700 biotin/lipoyl attachment protein	32.18	32.60	13653.3	4 (3 0 0 0 1)
YP_001038262 Cthe_1848 copper-translocating P-type ATPase	32.16	5.10	79962.7	4 (2 1 0 1 0)
YP_001037671 Cthe_1246 phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	32.14	8.20	56306.4	4 (3 0 0 0 1)
YP_001037149 Cthe_0721 nitrogen-fixing NifU-like protein	30.21	28.90	16208.8	3 (3 0 0 0 0)
YP_001039293 Cthe_2902 ribosomal protein S10	30.20	33.00	11770.6	3 (3 0 0 0 0)
YP_001039309 Cthe_2918 ribosomal protein L6	30.20	21.30	20135.1	3 (3 0 0 0 0)
Cthe_3088	30.20	32.4	7948	4 (4 0 0 0 0)
YP_001036939 Cthe_0509 sodium ion-translocating decarboxylase, beta subunit	30.20	15.90	38047.1	3 (3 0 0 0 0)
YP_001039131 Cthe_2739 trigger factor	30.20	7.50	48813.2	3 (3 0 0 0 0)
YP_001037524 Cthe_1099 hypothetical protein	30.20	5.5	94172	3 (3 0 0 0 0)

YP_001038180 Cthe_1766 glutamate 5-kinase	30.20	17.00	29832.5	3 (3 0 0 0 0)
YP_001037283 Cthe_0856 branched-chain amino acid aminotransferase	30.19	14.90	39711.6	3 (3 0 0 0 0)
YP_001036666 Cthe_0234 AMP-dependent synthetase and ligase	30.19	5.80	95168.9	3 (3 0 0 0 0)
YP_001039296 Cthe_2905 Ribosomal protein L25/L23	30.19	35.00	13281.3	4 (4 0 0 0 0)
YP_001037500 Cthe_1075 protein of unknown function DUF881	30.19	16.3	26506	3 (3 0 0 0 0)
YP_001039085 Cthe_2693 hypothetical protein	30.18	13.90	17809.5	3 (3 0 0 0 0)
YP_0010374 Cthe_1039 ribosomal protein S20	30.18	20.90	12988.5	3 (3 0 0 0 0)
YP_001037925 Cthe_1504 Linocin_M18 bacteriocin protein	30.18	13.70	30295.9	3 (3 0 0 0 0)
YP_0010373Cthe_0937 fatty acid/phospholipid synthesis protein PlsX	30.17	7.70	36458.3	3 (3 0 0 0 0)
YP_001039298 Cthe_2907 ribosomal protein S19	30.16	31.90	10589.7	4 (4 0 0 0 0)
YP_001037443 Cthe_1018 binding-protein-dependent transport systems inner membrane component	30.16	11.8	31905	3 (3 0 0 0 0)
YP_001039557 Cthe_3169 short-chain dehydrogenase/reductase SDR	30.16	15.10	27027.0	3 (3 0 0 0 0)
YP_001037788 Cthe_1363 lipopolysaccharide biosynthesis	30.15	8.40	52051.2	3 (3 0 0 0 0)
YP_001039110 Cthe_2718 preprotein translocase, SecE subunit	30.15	23.80	9175.1	3 (3 0 0 0 0)
YP_001038203 Cthe_1789 ATPase AAA-2	30.15	3.80	92041.6	3 (3 0 0 0 0)
YP_001039119 Cthe_2727 ribosomal protein S12	30.14	20.40	15415.7	3 (3 0 0 0 0)
YP_001039303 Cthe_2912 ribosomal protein S17	30.14	18.80	10034.5	3 (3 0 0 0 0)
YP_001039025 Cthe_2633 hypothetical protein	30.13	7.7	41666	3 (3 0 0 0 0)
YP_001039588 Cthe_3200 alanyl-tRNA synthetase	28.15	3.00	97985.4	3 (2 1 0 0 0)
YP_001039285 Cthe_2894 hypothetical protein	26.16	17.60	20119.8	3 (2 0 1 0 0)
YP_001038818 Cthe_2423 hypothetical protein	26.13	3.0	111319	3 (2 0 1 0 0)
YP_001038181 Cthe_1767 hypothetical protein	26.13	9.00	35361.9	3 (2 0 1 0 0)
YP_001038528 Cthe_2130 hypothetical protein	24.18	8.40	48488.7	3 (2 0 0 1 0)
YP_001038678 Cthe_2283 methyl-accepting chemotaxis sensory transducer	24.15	1.90	154468.8	3 (2 0 0 1 0)
YP_001038243 Cthe_1829 Chromosome segregation ATPases-like protein	22.14	1.6	124588	3 (1 1 0 1 0)

YP_001037284 Cthe_0857 protein of unknown function DUF107	20.21	16.00	17807.0	2 (2 0 0 0 0)
YP_001038998 Cthe_2605 ATP synthase F1, delta subunit	20.21	15.40	21896.0	3 (3 0 0 0 0)
YP_001038774 Cthe_2379 hypothetical	20.20	15.40	18996.1	2 (2 0 0 0 0)
YP_001037 Cthe_1224 ribosomal protein L35	20.20	23.10	7351.4	2 (2 0 0 0 0)
YP_001039022 Cthe_2630 ribose-phosphate pyrophosphokinase	20.20	9.10	34730.4	2 (2 0 0 0 0)
YP_001037151 Cthe_0723 tyrosyl-tRNA synthetase	20.20	6.90	46607.9	2 (2 0 0 0 0)
YP_001039000 Cthe_2607 ATP synthase F1, gamma subunit	20.20	8.50	34138.7	2 (2 0 0 0 0)
YP_001037295 Cthe_0868 type IV pilus assembly PilZ	20.19	9.30	27940.0	2 (2 0 0 0 0)
YP_001038582 Cthe_2187 ribosomal protein S6	20.17	13.70	11131.9	2 (2 0 0 0 0)
YP_001037341 Cthe_0914 hypothetical protein	20.17	2.50	96251.1	2 (2 0 0 0 0)
YP_001037390 Cthe_0963 Dihydrodipicolinate reductase	20.17	10.30	27775.6	3 (3 0 0 0 0)
YP_001037293 Cthe_0866 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	20.17	6.60	38061.7	2 (2 0 0 0 0)
YP_001037566 Cthe_1141 hypothetical protein	20.16	1.00	157008.0	2 (2 0 0 0 0)
YP_001038759 Cthe_23glucose inhibited division protein A	20.16	4.00	70739.1	2 (2 0 0 0 0)
YP_001039390 Cthe_3001 protein of unknown function DUF47	20.16	12.60	24092.6	2 (2 0 0 0 0)
YP_001037344 Cthe_0917 glutaminyl-tRNA synthetase	20.16	5.50	66415.2	2 (2 0 0 0 0)
YP_001036541 Cthe_0107 Riboflavin synthase	20.16	16.80	16428.7	2 (2 0 0 0 0)
YP_001036781 Cthe_0350 signal peptidase I	20.16	10.40	22212.4	2 (2 0 0 0 0)
YP_001037291 Cthe_08pyruvate ferredoxin/flavodoxin oxidoreductase	20.16	12.80	19330.0	2 (2 0 0 0 0)
YP_001038761 Cthe_2366 single-stranded nucleic acid binding R3H	20.16	11.70	23558.8	2 (2 0 0 0 0)
YP_001037349 Cthe_0922 diaminopimelate dehydrogenase	20.16	6.60	37201.2	2 (2 0 0 0 0)
YP_001038577 Cthe_2182 Ig-like, group 2	20.16	2.3	101483	2 (2 0 0 0 0)
YP_001038652 Cthe_2257 ribosomal protein L9	20.15	10.80	16353.9	2 (2 0 0 0 0)
YP_001037141 Cthe_0713 1-acyl-sn-glycerol-3-phosphate acyltransferase	20.15	13.30	22765.5	2 (2 0 0 0 0)
YP_001036876 Cthe_0445 cell division protein FtsZ	20.15	9.00	39703.8	2 (2 0 0 0 0)

YP_001039543 Cthe_3155 beta-lactamase-like protein	20.14	11.50	26983.7	2 (2 0 0 0 0)
YP_001039544 Cthe_3156 methyl-accepting chemotaxis sensory transducer	20.14	6.7	46201	2 (2 0 0 0 0)
YP_001037109 Cthe_0681 IMP dehydrogenase/GMP reductase	20.14	3.40	55083.2	2 (2 0 0 0 0)
YP_001037376 Cthe_0949 carbamoyl-phosphate synthase, large subunit	20.14	2.00	118645.0	2 (2 0 0 0 0)
YP_0010368Cthe_0433 glycoside hydrolase, family 9	20.14	2.9	89432	2 (2 0 0 0 0)
YP_001037148 Cthe_0720 aminotransferase, class V	20.13	4.80	43376.5	2 (2 0 0 0 0)
YP_001036698 Cthe_0266 methyl-accepting chemotaxis sensory transducer	20.13	1.6	55424	2 (2 0 0 0 0)
YP_001036848 Cthe_0417 ribosomal protein S15	18.16	20.70	10378.8	2 (1 1 0 0 0)
YP_001037160 Cthe_0732 Chorismate synthase	18.15	2.80	43128.4	2 (1 1 0 0 0)
YP_001037650 Cthe_1225 translation initiation factor IF-3	18.15	20.10	18832.1	2 (1 1 0 0 0)
YP_001039589 Cthe_3201 CRISPR-associated protein, Csh1 family	18.15	4.40	71508.3	2 (1 1 0 0 0)
YP_001038153 Cthe_1739 SNF2-related protein	18.14	5.90	44230.3	2 (1 1 0 0 0)
YP_0010393 Cthe_2974 hypothetical protein	18.14	6.50	45212.0	2 (1 1 0 0 0)
YP_001037693 Cthe_1268 histidine kinase	18.14	5.90	45215.0	2 (1 1 0 0 0)
YP_001038282 Cthe_1868 carbamoyl-phosphate synthase, large subunit	18.13	2.30	117506.1	2 (1 1 0 0 0)
YP_001039129 Cthe_2737 excinuclease ABC, C subunit	18.13	1.80	71725.7	2 (1 1 0 0 0)
YP_001036546 Cthe_0112 UDP-N-acetylenolpyruvoylglucosamine reductase	18.13	3.30	33466.5	2 (1 1 0 0 0)
YP_001038544 Cthe_2148 Carbohydrate-binding, CenC-like protein	18.13	8.3	28392	2 (1 1 0 0 0)
YP_001038239 Cthe_1825 multi-sensor hybrid histidine kinase	18.13	1.80	106314.1	2 (1 1 0 0 0)
YP_001037332 Cthe_0905 metal dependent phosphohydrolase	16.15	12.20	24726.8	3 (2 0 1 0 0)
YP_001038491 Cthe_2091 hypothetical protein	16.14	3.00	49445.5	2 (0 2 0 0 0)
YP_001037518 Cthe_1093 Methenyltetrahydrofolate cyclohydrolase	14.18	9.80	30767.4	2 (1 0 0 1 0)
YP_001038917 Cthe_2522 membrane associated protein	14.13	5.20	61939.5	2 (0 1 1 0 0)
YP_001038706 Cthe_2311 hypothetical protein	14.13	7.5	27741	2 (0 1 1 0 0)
YP_001038171 Cthe_1757 peptidase M23B	14.13	5.50	34708.1	2 (0 1 1 0 0)
YP_00103 Cthe_0056 Ig-like, group 2	12.16	0.5	493539	2 (1 0 0 0 1)
YP_001036844 Cthe_0413 glycoside hydrolase, family 9	12.13	2.0	137030	2 (1 0 0 0 1)

YP_001039616 Cthe_3228 copper amine oxidase-like protein	10.26	8.6	30072	2 (2 0 0 0 0)
YP_001036578 Cthe_0144 preprotein translocase, SecG subunit	10.23	20.5	8797	1 (1 0 0 0 0)
YP_001038658 Cthe_2263 H+-transporting two-sector ATPase, C subunit	10.21	12.3	15967	1 (1 0 0 0 0)
YP_001038500 Cthe_2100 transcriptional regulator, AbrB family	10.21	16.50	8910.8	2 (2 0 0 0 0)
YP_001038182 Cthe_1768 NifU-related domain containing protein	10.20	6.50	25046.6	1 (1 0 0 0 0)
YP_001037389 Cthe_0962 dihydrodipicolinate synthase	10.20	7.20	31954.9	1 (1 0 0 0 0)
YP_001037690 Cthe_1265 phosphoglucomutase/phosphomannomutase alpha/beta/alpha domain I	10.19	2.90	64914.1	1 (1 0 0 0 0)
YP_001036793 Cthe_0362 transcriptional regulator, AsnC family	10.19	8.20	18091.5	1 (1 0 0 0 0)
YP_001039049 Cthe_2657 histone-like DNA-binding protein	10.19	16.50	10081.5	1 (1 0 0 0 0)
YP_001038489 Cthe_2089 glycoside hydrolase, family 48	10.18	1.8	83505	1 (1 0 0 0 0)
YP_001038802 Cthe_2407 AAA ATPase, central region	10.18	2.00	61796.0	1 (1 0 0 0 0)
YP_001037043 Cthe_0615 Phenylacetate--CoA ligase	10.18	4.20	48686.2	2 (2 0 0 0 0)
YP_001037158 Cthe_0730 hypothetical protein	10.18	8.40	24840.0	2 (2 0 0 0 0)
YP_001038632 Cthe_2237 flagellin-like protein	10.18	4.40	29413.2	1 (1 0 0 0 0)
YP_001037938 Cthe_1517 type I phosphodiesterase/nucleotide pyrophosphatase	10.18	6.50	42290.1	1 (1 0 0 0 0)
YP_001037292 Cthe_0865 3-methyl-2-oxobutanoate dehydrogenase ferredoxin	10.18	5.60	26948.7	1 (1 0 0 0 0)
YP_001037384 Cthe_0957 preprotein translocase, YajC subunit	10.18	14.10	10950.1	1 (1 0 0 0 0)
YP_001037677 Cthe_1252 Auxin Efflux Carrier	10.18	3.90	39935.9	1 (1 0 0 0 0)
YP_001039530 Cthe_3142 hypothetical protein	10.18	15.40	18433.0	1 (1 0 0 0 0)
YP_001039424 Cthe_3035 D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	10.17	3.10	42835.4	1 (1 0 0 0 0)
YP_001039513 Cthe_3125 heat shock protein Hsp20	10.17	8.10	17300.8	1 (1 0 0 0 0)
YP_001036788 Cthe_0357 alpha-glucan phosphorylases	10.17	1.40	97651.2	1 (1 0 0 0 0)
YP_001039104 Cthe_2712 hypothetical protein	10.17	5.40	26108.5	1 (1 0 0 0 0)
YP_00103 Cthe_0036 hybrid cluster protein	10.17	2.20	59404.5	1 (1 0 0 0 0)

YP_001037388 Cthe_0961 aspartate-semialdehyde dehydrogenase	10.17	4.50	36705.8	1 (1 0 0 0 0)
YP_001037303 Cthe_0876 glycosyl transferase, family 39	10.17	2.10	49970.4	1 (1 0 0 0 0)
YP_001037453 Cthe_1028 acetate kinase	10.17	4.80	43937.9	1 (1 0 0 0 0)
YP_001037754 Cthe_1329 putative CoA-substrate-specific enzyme activase	10.17	1.80	160075.3	1 (1 0 0 0 0)
YP_001038441 Cthe_2039 UvrD/REP helicase	10.17	1.50	143255.1	1 (1 0 0 0 0)
YP_001037777 Cthe_1352 UDP-glucose 6-dehydrogenase	10.17	3.20	48985.4	1 (1 0 0 0 0)
YP_001038905 Cthe_2510 protein of unknown function DUF368	10.17	4.20	30692.6	1 (1 0 0 0 0)
YP_001037907 Cthe_1485 beta-lactamase-like protein	10.16	3.40	47381.8	1 (1 0 0 0 0)
YP_001038119 Cthe_1705 hypothetical protein	10.16	3.60	65110.9	1 (1 0 0 0 0)
YP_001036769 Cthe_0338 NADH-quinone oxidoreductase, E subunit	10.16	7.30	18057.1	1 (1 0 0 0 0)
YP_001037755 Cthe_1330 signal peptidase I	10.16	6.90	19480.3	2 (2 0 0 0 0)
YP_00103 Cthe_0003 ankyrin repeat protein	10.16	7.4	35562	2 (0 1 0 0 1)
YP_001036511 Cthe_0077 hypothetical protein	10.16	8.70	13916.6	1 (1 0 0 0 0)
YP_001037455 Cthe_1030 TPR repeat domain containing protein	10.16	8.00	31289.3	1 (1 0 0 0 0)
YP_001038799 Cthe_2404 transcriptional regulator, GntR family	10.16	4.80	26246.0	1 (1 0 0 0 0)
YP_001038346 Cthe_1935 arginyl-tRNA synthetase	10.16	2.50	63997.0	1 (1 0 0 0 0)
YP_001038371 Cthe_1960 Peptidoglycan-binding domain 1	10.16	3.3	38782	1 (1 0 0 0 0)
YP_001036751 Cthe_0320 hypothetical protein	10.16	2.30	55540.0	1 (1 0 0 0 0)
YP_001038423 Cthe_2018 hypothetical protein	10.16	3.20	48846.5	1 (1 0 0 0 0)
YP_001036596 Cthe_0162 ribosomal protein L27	10.16	14.10	10032.6	1 (1 0 0 0 0)
YP_001037362 Cthe_0935 malonyl CoA-acyl carrier protein transacylase	10.16	4.20	33369.3	1 (1 0 0 0 0)
YP_00103 Cthe_0048 copper amine oxidase-like protein	10.16	3.7	30343	1 (1 0 0 0 0)
YP_001037380 Cthe_0953 aspartate carbamoyltransferase	10.16	7.10	34054.3	1 (1 0 0 0 0)
YP_001038991 Cthe_2598 uracil phosphoribosyltransferase	10.16	5.30	23078.1	1 (1 0 0 0 0)
YP_001038757 Cthe_2362 parB-like partition protein	10.16	4.30	31921.3	1 (1 0 0 0 0)
YP_001039135 Cthe_2743 putative metalloendopeptidase, glycoprotease family	10.16	6.80	36092.9	1 (1 0 0 0 0)

YP_001037722 Cthe_1297 Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	10.16	6.30	37328.4	1 (1 0 0 0 0)
YP_001039578 Cthe_3190 helicase, RecD/TraA family	10.16	2.50	79629.2	1 (1 0 0 0 0)
YP_001037799 Cthe_1374 copper amine oxidase-like protein	10.15	3.80	30003.5	1 (1 0 0 0 0)
YP_001037121 Cthe_0693 Pseudouridine synthase, Rsu	10.15	6.30	27352.5	1 (1 0 0 0 0)
YP_001037122 Cthe_0694 spermidine synthase	10.15	3.60	31334.1	1 (1 0 0 0 0)
YP_0010387 Cthe_2369 ribonuclease P protein component	10.15	6.30	15109.1	1 (1 0 0 0 0)
YP_001037074 Cthe_0646 anaerobic ribonucleoside-triphosphate reductase	10.15	1.10	79998.1	2 (2 0 0 0 0)
YP_001037531 Cthe_1106 twitching motility protein	10.15	4.00	38896.1	1 (1 0 0 0 0)
YP_001038578 Cthe_2183 UTP-glucose-1-phosphate uridylyltransferase	10.15	5.2	32481	1 (1 0 0 0 0)
YP_001038456 Cthe_2056 hypothetical protein	10.15	6.10	36306.5	1 (1 0 0 0 0)
YP_001039035 Cthe_2643 Nucleotidyl transferase	10.15	5.20	40256.0	1 (1 0 0 0 0)
YP_001039387 Cthe_2998 ABC transporter related protein	10.15	2.40	63369.4	1 (1 0 0 0 0)
YP_001036983 Cthe_0554 phosphoribosylformylglycinamidine synthase	10.15	1.50	139261.8	1 (1 0 0 0 0)
YP_001037916 Cthe_1495 pyridoxamine 5'-phosphate oxidase-related, FMN-binding	10.15	10.00	14943.5	1 (1 0 0 0 0)
YP_001038451 Cthe_2051 CRISPR-associated RAMP protein, SSO1426 family	10.15	10.80	29381.0	1 (1 0 0 0 0)
YP_001039381 Cthe_2992 RNA polymerase, sigma-24 subunit, ECF subfamily	10.15	4.60	20428.5	1 (1 0 0 0 0)
YP_001037275 Cthe_0847 translation elongation factor P	10.15	6.50	20640.7	2 (2 0 0 0 0)
YP_001037710 Cthe_1285 metal dependent phosphohydrolase	10.15	2.6	53393	1 (1 0 0 0 0)
YP_001038574 Cthe_2179 Pectate lyase/Amb allergen	10.15	1.1	99829	1 (1 0 0 0 0)
YP_001038253 Cthe_1839 Radical SAM	10.15	5.70	40201.2	1 (1 0 0 0 0)
YP_001039002 Cthe_2609 ATP synthase F1, epsilon subunit	10.15	7.40	15369.2	1 (1 0 0 0 0)
YP_001036683 Cthe_0251 transglutaminase-like protein	10.15	1.3	96667	1 (1 0 0 0 0)
YP_001038366 Cthe_1955 RNA binding S1	10.15	1.80	80671.0	1 (1 0 0 0 0)
YP_001037008 Cthe_0579 nitroreductase	10.15	6.40	19786.5	1 (1 0 0 0 0)
YP_001037454 Cthe_1029 phosphate acetyltransferase	10.15	4.70	38641.0	1 (1 0 0 0 0)
YP_001038918 Cthe_2523 hypothetical	10.15	7.90	14542.0	1 (1 0 0 0 0)

protein				
YP_001036675 Cthe_0243 copper amine oxidase-like protein	10.15	4.6	31681	1 (1 0 0 0 0)
YP_001038380 Cthe_1969 hypothetical protein	10.15	2.20	52057.6	1 (1 0 0 0 0)
YP_001037610 Cthe_1185 hypothetical protein	10.15	2.10	50416.2	1 (1 0 0 0 0)
YP_001038507 Cthe_2107 thymidylate kinase	10.15	3.10	26517.4	1 (1 0 0 0 0)
YP_001038972 Cthe_2579 iron-containing alcohol dehydrogenase	10.14	3.50	40880.0	1 (1 0 0 0 0)
YP_001037202 Cthe_0774 RNA chaperone Hfq	10.14	12.20	9361.9	1 (1 0 0 0 0)
YP_001038258 Cthe_1844 transcriptional regulator, BadM/Rrf2 family	10.14	5.90	16944.7	1 (1 0 0 0 0)
YP_001037337 Cthe_0910 extracellular solute-binding protein, family 5	10.14	1.6	63117	1 (1 0 0 0 0)
YP_001038569 Cthe_2174 transcription termination factor Rho	10.14	1.40	72542.7	1 (1 0 0 0 0)
YP_001038330 Cthe_1918 Orn/Lys/Arg decarboxylase, major region	10.14	3.50	53290.5	1 (1 0 0 0 0)
YP_001036701 Cthe_0269 glycoside hydrolase, family 8	10.14	2.3	52561	1 (1 0 0 0 0)
YP_001037636 Cthe_1211 pyridoxal-phosphate dependent TrpB-like enzyme	10.14	2.90	50100.1	1 (1 0 0 0 0)
YP_001036540 Cthe_0106 GTP cyclohydrolase II	10.14	2.90	45817.6	1 (1 0 0 0 0)
YP_001038384 Cthe_1973 Tetratricopeptide TPR_2	10.14	2.20	58122.1	1 (1 0 0 0 0)
Cthe_2124	10.14	2.2	50925	1 (1 0 0 0 0)
YP_001038220 Cthe_1806 cellulosome enzyme, dockerin type I	10.14	0.9	235980	1 (1 0 0 0 0)
YP_001038807 Cthe_2412 SMC protein-like protein	10.14	1.70	101479.0	1 (1 0 0 0 0)
YP_001037042 Cthe_0614 pyruvate ferredoxin/flavodoxin oxidoreductase	10.14	5.20	21269.4	1 (1 0 0 0 0)
YP_001037448 Cthe_1023 protein of unknown function DUF205	10.14	5.90	23592.1	1 (1 0 0 0 0)
YP_001039181 Cthe_2789 NLPA lipoprotein	10.14	2.5	39797	1 (1 0 0 0 0)
YP_001037737 Cthe_1312 glycyl-tRNA synthetase	10.14	1.90	54044.8	1 (1 0 0 0 0)
YP_001037845 Cthe_1421 signal peptide peptidase SppA, 36K type	10.14	2.8	35780	1 (1 0 0 0 0)
YP_001037361 Cthe_0934 3-oxoacyl- acyl-carrier-protein reductase	10.14	4.90	26061.6	1 (1 0 0 0 0)
YP_001037184 Cthe_0756 Membrane dipeptidase	10.14	5.50	34538.7	1 (1 0 0 0 0)
YP_001038779 Cthe_2384 S-layer-like domain containing protein	10.14	6.3	21039	1 (1 0 0 0 0)

YP_001037030 Cthe_0602 thiamine biosynthesis protein ThiC	10.14	2.30	48266.6	1 (1 0 0 0 0)
YP_001037102 Cthe_0674 NUDIX hydrolase	10.14	6.60	20289.6	1 (1 0 0 0 0)
YP_001036573 Cthe_0139 Triose-phosphate isomerase	10.14	6.40	27117.0	1 (1 0 0 0 0)
YP_001036562 Cthe_0128 Integrase, catalytic region	10.14	4.60	44028.6	1 (1 0 0 0 0)
YP_001039445 Cthe_3056 transposase, IS204/IS1001/IS1096/IS1165	10.14	2.10	55648.1	1 (1 0 0 0 0)
YP_001039268 Cthe_2877 S-layer-like domain containing protein	10.14	2.1	67015	1 (1 0 0 0 0)
YP_001036501 Cthe_0067 Silent information regulator protein Sir2	10.14	2.90	26808.5	1 (1 0 0 0 0)
YP_001037756 Cthe_1331 aspartyl-tRNA synthetase	10.14	1.30	67728.0	1 (1 0 0 0 0)
YP_001039325 Cthe_2934 ABC transporter related protein	10.14	2.30	33338.6	1 (1 0 0 0 0)
YP_001039021 Cthe_2629 UDP-N-acetylglucosamine pyrophosphorylase	10.14	1.30	51493.2	1 (1 0 0 0 0)
YP_001037529 Cthe_1104 prepilin-type cleavage/methylation	10.14	6.3	19225	1 (1 0 0 0 0)
YP_001037685 Cthe_1260 4-hydroxybenzoyl-CoA thioesterase	10.14	8.00	16215.3	1 (1 0 0 0 0)
YP_001036534 Cthe_0100 hypothetical protein	10.14	4.2	30310	1 (1 0 0 0 0)
YP_001037044 Cthe_0616 amino acid-binding ACT	10.14	7.70	15451.4	1 (1 0 0 0 0)
YP_001037816 Cthe_1391 2-isopropylmalate synthase	10.14	1.80	56422.7	1 (1 0 0 0 0)
YP_001036883 Cthe_0452 cellulosome anchoring protein, cohesin region	10.14	6.6	28451	1 (1 0 0 0 0)
YP_001036516 Cthe_0082 ATP-dependent protease La	10.14	1.70	91737.0	1 (1 0 0 0 0)
YP_001036843 Cthe_0412 glycoside hydrolase, family 9	10.13	1.2	100558	1 (1 0 0 0 0)
YP_001037516 Cthe_1091 metal dependent phosphohydrolase	10.13	1.30	58726.7	1 (1 0 0 0 0)
YP_001038446 Cthe_2046 hypothetical protein	10.13	5.90	19878.0	1 (1 0 0 0 0)
YP_001038605 Cthe_2210 3-isopropylmalate dehydratase, small subunit	10.13	7.40	20861.6	1 (1 0 0 0 0)
YP_001036500 Cthe_0066 hypothetical protein	10.13	7.6	18835	1 (1 0 0 0 0)
YP_001036999 Cthe_0570 peptidase, membrane zinc metallopeptidase, putative	10.13	5.30	24747.4	1 (1 0 0 0 0)
YP_001036960 Cthe_0531 hypothetical protein	10.13	1.8	87440	1 (1 0 0 0 0)
Cthe_0850	10.13	4.2	32313	1 (1 0 0 0 0)

YP_001038870 Cthe_2475 phage portal protein, SPP1 family	10.13	2.30	56081.0	1 (1 0 0 0 0)
YP_001038995 Cthe_2602 ATP synthase F0, A subunit	10.13	5.20	28445.0	1 (1 0 0 0 0)
YP_001036996 Cthe_0567 peptide deformylase	10.13	5.90	19073.1	1 (1 0 0 0 0)
YP_001039095 Cthe_2703 protein of unknown function DUF161	10.13	6.00	33046.1	1 (1 0 0 0 0)
YP_001038128 Cthe_1714 hypothetical	10.13	11.00	12935.6	1 (1 0 0 0 0)
YP_001036613 Cthe_0179 Argininosuccinate synthase	10.13	3.50	45593.7	1 (1 0 0 0 0)
YP_001037Cthe_1223 ribosomal protein L20	10.13	6.80	13356.4	1 (1 0 0 0 0)
YP_001038805 Cthe_2410 protein of unknown function DUF342	10.13	2.10	58771.0	1 (1 0 0 0 0)
YP_001039068 Cthe_2676 GumN	10.13	1.9	53716	1 (1 0 0 0 0)
YP_001038278 Cthe_18acetylglutamate kinase	10.13	3.00	33029.6	1 (1 0 0 0 0)
YP_001037514 Cthe_1089 Stage V sporulation protein S	10.13	17.00	9140.0	1 (1 0 0 0 0)
YP_001036827 Cthe_0396 ABC transporter related protein	10.13	1.20	63672.7	1 (1 0 0 0 0)
YP_001036984 Cthe_0555 PpiC-type peptidyl-prolyl cis-trans isomerase	10.13	2.2	47318	1 (1 0 0 0 0)
YP_001039312 Cthe_2921 ribosomal protein L30	10.13	16.90	6614.7	1 (1 0 0 0 0)
YP_001039020 Cthe_2628 SpoVG	10.13	11.70	10462.3	1 (1 0 0 0 0)
YP_001038087 Cthe_1671 Recombinase	10.13	3.10	59959.0	1 (1 0 0 0 0)
YP_001038598 Cthe_2203 GTP cyclohydrolase I	10.13	6.40	20797.1	1 (1 0 0 0 0)
YP_001037553 Cthe_1128 hypothetical protein	10.13	7.80	14799.7	1 (1 0 0 0 0)
YP_001037363 Cthe_0936 3-oxoacyl- acyl-carrier-protein synthase III	10.13	2.70	35975.2	1 (1 0 0 0 0)
YP_001037809 Cthe_1384 FolC bifunctional protein	10.13	2.80	47865.1	1 (1 0 0 0 0)
YP_001038958 Cthe_25hypothetical protein	10.13	5.70	26823.2	1 (1 0 0 0 0)
YP_001038161 Cthe_1747 DNA mismatch endonuclease vsr	10.13	8.00	16582.7	1 (1 0 0 0 0)
YP_001037252 Cthe_0824 copper amine oxidase-like protein	10.13	1.5	67339	1 (1 0 0 0 0)
YP_001039338 Cthe_2947 prolyl-tRNA synthetase	10.13	1.00	64228.5	1 (1 0 0 0 0)
YP_001037112 Cthe_0684 Polynucleotide adenylyltransferase region	10.13	1.60	99087.3	1 (1 0 0 0 0)
YP_001037Cthe_1221 glycosyltransferase 36	10.13	0.8	333880	2 (0 1 0 0 1)
YP_001036503 Cthe_0069 Aspartate--ammonia ligase	10.13	2.40	39008.1	1 (1 0 0 0 0)

YP_001039566 Cthe_3178 protein of unknown function DUF342	10.13	2.10	59841.4	1 (1 0 0 0 0)
YP_001038285 Cthe_1871 Tn7-like transposition protein D	10.13	2.70	73259.0	1 (1 0 0 0 0)
YP_001038314 Cthe_1902 Pyridoxal-5'-phosphate-dependent enzyme, beta subunit	10.13	5.50	32755.1	1 (1 0 0 0 0)
YP_001037670 Cthe_1245 phosphoribosylamine--glycine ligase	10.13	2.40	46106.6	1 (1 0 0 0 0)
YP_001036765 Cthe_0334 putative anti-sigma regulatory factor, serine/threonine protein kinase	10.13	8.70	16257.4	1 (1 0 0 0 0)
YP_001037946 Cthe_1525 protein of unknown function DUF214	10.13	1.80	70514.4	1 (1 0 0 0 0)
YP_001037258 Cthe_0830 acid phosphatase/vanadium-dependent haloperoxidase related	10.13	5.40	16108.6	1 (1 0 0 0 0)
YP_001038885 Cthe_2490 hypothetical protein	10.13	4.50	36296.3	1 (1 0 0 0 0)
YP_001038272 Cthe_1858 peptidase M23B	10.13	3.5	41616	1 (1 0 0 0 0)
YP_001036800 Cthe_0369 protein of unknown function DUF111	10.13	3.20	45603.7	1 (1 0 0 0 0)
YP_00103Cthe_0009 YD repeat containing protein	10.13	0.30	209626.4	1 (1 0 0 0 0)
YP_001038547 Cthe_2151 hypothetical protein	10.13	1.60	83178.1	1 (1 0 0 0 0)
YP_001036615 Cthe_0181 Holliday junction DNA helicase RuvA	10.13	6.40	22143.8	1 (1 0 0 0 0)
YP_001036875 Cthe_0444 cell division protein FtsA	10.13	1.50	44730.7	1 (1 0 0 0 0)
YP_001038260 Cthe_1846 periplasmic sensor signal transduction histidine kinase	10.13	2.6	51783	1 (1 0 0 0 0)
YP_001036506 Cthe_0072 phage shock protein C, PspC	10.13	6.80	18036.7	1 (1 0 0 0 0)
YP_001037462 Cthe_1037 cell wall hydrolase/autolysin	10.13	3.9	30734	1 (1 0 0 0 0)
YP_001038896 Cthe_2501 hypothetical protein	10.13	6.30	16731.8	1 (1 0 0 0 0)
YP_001037281 Cthe_0854 shikimate 5-dehydrogenase	10.13	5.70	33301.7	1 (1 0 0 0 0)
YP_001037977 Cthe_1558 transcriptional regulator, BadM/Rrf2 family	10.13	7.90	16772.9	1 (1 0 0 0 0)
YP_001039539 Cthe_3151 Adenosylcobinamide-phosphate guanylyltransferase	10.13	6.40	20876.8	1 (1 0 0 0 0)
YP_001037800 Cthe_1375 aspartate kinase	10.13	2.90	49629.5	1 (1 0 0 0 0)
YP_001038493 Cthe_2093 hypothetical protein	10.13	2.6	38750	1 (1 0 0 0 0)
YP_001039065 Cthe_2673 hypothetical protein	10.13	12.80	11320.7	1 (1 0 0 0 0)

YP_00103Cthe_0054 protein of unknown function DUF187	10.13	2.3	82322	1 (1 0 0 0 0)
YP_001037598 Cthe_1173 type IV pilus assembly PilZ	10.12	2.80	24807.1	1 (1 0 0 0 0)
YP_001038Cthe_2249 helicase, RecD/TraA family	10.12	1.20	84297.0	1 (1 0 0 0 0)
YP_001038741 Cthe_2346 O-antigen polymerase	10.12	0.80	114378.3	1 (1 0 0 0 0)
YP_001039519 Cthe_3131 von Willebrand factor, type A	10.12	1.1	62240	1 (1 0 0 0 0)
YP_001036842 Cthe_0411 hypothetical protein	10.12	2.00	39706.4	1 (1 0 0 0 0)
YP_001037372 Cthe_0945 metallophosphoesterase	10.12	2.10	37881.6	1 (1 0 0 0 0)
YP_001038442 Cthe_2040 ATP-dependent exonuclease synthesis protein AddB superfamily I helicase	10.12	0.40	130363.1	1 (1 0 0 0 0)
YP_001039383 Cthe_2994 hypothetical protein	10.12	1.2	75927	1 (1 0 0 0 0)
YP_001036719 Cthe_0287 multi-sensor hybrid histidine kinase	10.12	0.6	132798	1 (1 0 0 0 0)
YP_001038914 Cthe_2519 2-isopropylmalate synthase/homocitrate synthase family protein	10.12	1.10	59595.8	1 (1 0 0 0 0)
YP_001038755 Cthe_2360 glycoside hydrolase, family 9	10.12	0.80	104527.3	1 (1 0 0 0 0)
YP_001036867 Cthe_0436 Tetratricopeptide TPR_2	10.12	0.70	122890.1	1 (1 0 0 0 0)
YP_001036595 Cthe_0161 protein of unknown function DUF4	10.12	4.40	12500.6	1 (1 0 0 0 0)
YP_001037206 Cthe_0778 hypothetical protein	10.11	4.60	15107.5	1 (1 0 0 0 0)
YP_001036708 Cthe_0276 D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	10.11	1.90	35661.5	1 (1 0 0 0 0)
YP_001037278 Cthe_0851 hypothetical protein	10.11	8.00	9941.1	1 (1 0 0 0 0)
YP_00103Cthe_0047 copper amine oxidase-like protein	10.11	2.2	30435	1 (1 0 0 0 0)
YP_001038216 Cthe_1802 cobalt ABC transporter, inner membrane subunit CbiQ	10.11	2.20	30388.6	1 (1 0 0 0 0)
YP_001037806 Cthe_1381 threonine synthase	10.11	2.20	55667.3	1 (1 0 0 0 0)
YP_001039186 Cthe_2794 pyruvate/ketoisovalerate oxidoreductase, gamma subunit	10.11	4.40	19950.5	1 (1 0 0 0 0)
YP_001039396 Cthe_3007 ErfK/YbiS/YcfS/YnhG	10.11	5.20	25407.9	1 (1 0 0 0 0)
YP_001038479 Cthe_2079 signal peptidase I	10.10	4.3	21673	1 (1 0 0 0 0)
YP_001037118 Cthe_0690 hypothetical	10.10	2.50	23518.3	1 (1 0 0 0 0)

protein				
YP_001038695 Cthe_2300 CRISPR-associated protein Cas5	10.10	2.10	26796.9	1 (1 0 0 0 0)
YP_001039474 Cthe_3085 response regulator receiver protein	8.15	3.20	35092.4	1 (0 1 0 0 0)
YP_001036898 Cthe_0467 Flagellar biosynthesis/type III secretory pathway protein-like protein	8.15	4.20	29946.8	2 (0 0 0 2 0)
YP_001038372 Cthe_1961 Nucleotidyl transferase	8.15	1.30	90642.3	1 (0 1 0 0 0)
YP_001037948 Cthe_1527 periplasmic sensor signal transduction histidine kinase	8.15	2.90	40255.1	1 (0 1 0 0 0)
YP_001038177 Cthe_1763 ABC transporter related protein	8.15	4.10	27658.6	1 (0 1 0 0 0)
YP_001038556 Cthe_2161 hypothetical protein	8.15	3.40	33780.1	1 (0 1 0 0 0)
YP_001037968 Cthe_1548 Enoyl-CoA hydratase/isomerase	8.15	6.50	27621.2	1 (0 1 0 0 0)
YP_001037749 Cthe_1324 heat-inducible transcription repressor HrcA	8.15	1.70	39225.6	1 (0 1 0 0 0)
YP_001038237 Cthe_1823 Extracellular ligand-binding receptor	8.15	2.6	46053	1 (0 1 0 0 0)
YP_001038436 Cthe_2032 hypothetical protein	8.14	0.6	223041	1 (0 1 0 0 0)
YP_001038201 Cthe_1787 glycoside hydrolase 15-related	8.14	2.20	73633.7	1 (0 1 0 0 0)
YP_001039240 Cthe_2848 phage major capsid protein, HK97	8.14	3.20	48301.6	1 (0 1 0 0 0)
YP_001038806 Cthe_2411 metallophosphoesterase	8.14	3.40	43678.0	1 (0 1 0 0 0)
YP_001037885 Cthe_1463 protein of unknown function DUF891	8.14	13.00	13440.2	1 (0 1 0 0 0)
YP_001037133 Cthe_0705 hypothetical protein	8.14	15.40	9275.1	1 (0 1 0 0 0)
YP_001036834 Cthe_0403 DNA-directed RNA polymerase sigma factor	8.14	4.00	32431.1	1 (0 1 0 0 0)
YP_001037354 Cthe_0927 chromosome segregation protein SMC	8.14	1.50	136868.0	2 (0 0 1 0 1)
YP_001038394 Cthe_1983 transposase IS116/IS110/IS902	8.14	3.00	45903.8	1 (0 1 0 0 0)
YP_001037792 Cthe_1367 PHP-like protein	8.14	3.10	29478.5	1 (0 1 0 0 0)
YP_001036629 Cthe_0195 RNA polymerase, sigma-24 subunit, ECF subfamily	8.14	10.30	11336.3	1 (0 1 0 0 0)
YP_001037245 Cthe_0817 diguanylate cyclase	8.14	0.80	209216.5	1 (0 1 0 0 0)
YP_001038378 Cthe_1967 DNA segregation ATPase FtsK/SpoIIIE and related proteins	8.14	2.60	66691.3	1 (0 1 0 0 0)
YP_001038626 Cthe_2231 Methyltransferase type 11	8.14	3.70	25433.8	1 (0 1 0 0 0)

YP_001039590 Cthe_3202 CRISPR-associated protein, Csh2 family	8.14	2.60	34861.2	1 (0 1 0 0 0)
YP_001036569 Cthe_0135 beta-ketoacyl synthase	8.14	0.40	305695.5	1 (0 1 0 0 0)
YP_001037276 Cthe_0848 peptidase M24	8.14	3.10	40148.9	2 (0 0 1 0 1)
YP_001039042 Cthe_2650 polysaccharide biosynthesis protein CapD	8.14	2.50	68185.3	1 (0 1 0 0 0)
YP_001036688 Cthe_0256 histidine kinase	8.14	1.80	80768.8	1 (0 1 0 0 0)
YP_001039050 Cthe_2658 RNA-binding S4	8.14	10.10	8950.9	1 (0 1 0 0 0)
YP_001039450 Cthe_3061 transcriptional repressor, CopY family	8.14	6.10	17249.2	1 (0 1 0 0 0)
YP_001039004 Cthe_2611 Fibronectin, type III	8.14	0.6	183583	1 (0 1 0 0 0)
YP_001038100 Cthe_1685 ABC transporter related protein	8.14	1.70	64734.1	1 (0 1 0 0 0)
YP_001037996 Cthe_1577 hypothetical protein	8.14	1.40	58166.9	1 (0 1 0 0 0)
YP_001039029 Cthe_2637 glycosyl transferase, group 1	8.14	2.00	38858.4	1 (0 1 0 0 0)
YP_001038694 Cthe_2299 CRISPR-associated helicase Cas3	8.13	0.80	87169.3	1 (0 1 0 0 0)
YP_001038986 Cthe_2593 peptide chain release factor 1	8.13	3.30	40998.9	1 (0 1 0 0 0)
YP_001037773 Cthe_1348 ribosomal protein S21	8.13	22.40	6901.8	1 (0 1 0 0 0)
YP_001036832 Cthe_0401 methyl-accepting chemotaxis sensory transducer	8.13	2.40	85087.8	1 (0 1 0 0 0)
YP_001037714 Cthe_1289 hypothetical protein	8.13	17.20	7786.1	1 (0 1 0 0 0)
YP_001039218 Cthe_2826 hypothetical protein	8.13	2.90	48947.8	1 (0 1 0 0 0)
YP_001039160 Cthe_2768 metallophosphoesterase	8.13	1.80	98713.7	1 (0 1 0 0 0)
YP_001038723 Cthe_2328 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	8.13	3.60	44355.7	1 (0 1 0 0 0)
YP_001038123 Cthe_1709 Phage-related protein-like protein	8.13	2.40	81749.4	1 (0 1 0 0 0)
YP_001038240 Cthe_1826 response regulator receiver sensor signal transduction histidine kinase	8.13	3.00	41724.5	1 (0 1 0 0 0)
YP_001038298 Cthe_1884 transposase IS116/IS110/IS902	8.13	3.50	48534.1	1 (0 1 0 0 0)
YP_001039211 Cthe_2819 methyl-accepting chemotaxis sensory transducer	8.13	0.50	154334.8	1 (0 1 0 0 0)
YP_001037960 Cthe_1540 glutamyl-tRNA Gln amidotransferase, B subunit	8.13	2.10	52645.8	1 (0 1 0 0 0)
YP_001036887 Cthe_0456 protein of unknown function UPF0047	8.13	7.20	15799.1	1 (0 1 0 0 0)
YP_001039069 Cthe_2677 histidine kinase	8.13	4.60	26818.0	1 (0 1 0 0 0)

YP_001036616 Cthe_0182 Holliday junction DNA helicase RuvB	8.13	4.20	36494.5	1 (0 1 0 0 0)
YP_001037930 Cthe_1509 protein of unknown function DUF438	8.13	1.90	47519.6	1 (0 1 0 0 0)
YP_00103Cthe_0040 Cellulase., Cellulose 1,4-beta-cellobiosidase	8.13	1.4	98470	1 (0 1 0 0 0)
YP_001038Cthe_2247 regulatory protein, MerR	8.13	6.80	16866.4	1 (0 1 0 0 0)
YP_001037057 Cthe_0629 type II secretion system protein E	8.13	1.50	55014.4	1 (0 1 0 0 0)
YP_001039253 Cthe_2862 hypothetical protein	8.13	5.50	23932.9	1 (0 1 0 0 0)
YP_001037892 Cthe_1470 RNA polymerase, sigma-24 subunit, ECF subfamily	8.13	5.90	23943.7	1 (0 1 0 0 0)
YP_001037998 Cthe_1579 ABC transporter related protein	8.13	2.30	57089.3	1 (0 1 0 0 0)
YP_001038144 Cthe_1730 hypothetical protein	8.13	9.20	16451.4	1 (0 1 0 0 0)
YP_001036941 Cthe_0511 histidine kinase	8.13	2.30	50721.3	1 (0 1 0 0 0)
YP_00103Cthe_0024 biotin biosynthesis protein BioC	8.13	3.90	32622.6	1 (0 1 0 0 0)
YP_001037675 Cthe_1250 phosphoribosylaminoimidazole carboxylase, catalytic subunit	8.13	6.40	18407.8	1 (0 1 0 0 0)
YP_001037569 Cthe_1144 restriction modification system DNA specificity domain	8.13	2.10	54427.8	1 (0 1 0 0 0)
YP_001038217 Cthe_1803 cobalamin vitamin B12 biosynthesis CbiM protein	8.13	2.0	37379	1 (0 1 0 0 0)
YP_001037251 Cthe_0823 hypothetical protein	8.13	3.00	58854.5	1 (0 1 0 0 0)
YP_001038974 Cthe_2581 dihydropteroate synthase	8.13	3.60	43330.5	1 (0 1 0 0 0)
YP_001037771 Cthe_1346 single-stranded-DNA-specific exonuclease RecJ	8.13	1.00	93542.3	1 (0 1 0 0 0)
YP_001036711 Cthe_0279 hypothetical protein	8.13	23.80	6918.7	1 (0 1 0 0 0)
YP_001038866 Cthe_2471 hypothetical protein	8.13	2.30	50488.0	1 (0 1 0 0 0)
YP_001037913 Cthe_1492 NAD P H dehydrogenase quinone	8.13	2.00	61292.0	1 (0 1 0 0 0)
YP_001039373 Cthe_2983 methyl-accepting chemotaxis sensory transducer	8.13	1.60	61326.0	1 (0 1 0 0 0)
YP_001038904 Cthe_2509 DNA topoisomerase	8.13	3.90	37812.1	1 (0 1 0 0 0)
YP_001037457 Cthe_1032 glutamyl-tRNA Gln amidotransferase, A subunit	8.13	2.30	52926.1	1 (0 1 0 0 0)
YP_001039465 Cthe_3076 Radical SAM	8.13	3.30	49202.7	1 (0 1 0 0 0)
YP_001038476 Cthe_2076 periplasmic sensor signal transduction histidine kinase	8.13	2.10	43752.4	1 (0 1 0 0 0)

YP_001039281 Cthe_2890 putative transcriptional regulator	8.13	3.90	43839.2	1 (0 1 0 0 0)
YP_001039496 Cthe_3108 phage SPO1 DNA polymerase-related protein	8.13	8.50	21456.2	1 (0 1 0 0 0)
YP_001037076 Cthe_0glutamyl-tRNA synthetase	8.13	3.10	63361.2	1 (0 1 0 0 0)
YP_001038769 Cthe_2374 DNA replication and repair protein RecF	8.13	3.00	43059.5	1 (0 1 0 0 0)
YP_001037660 Cthe_1235 Cellulose 1,4-beta-cellobiosidase	8.12	0.10	758397.6	1 (0 1 0 0 0)
YP_001038501 Cthe_2101 Protein of unknown function UPF0011	8.12	2.10	31968.8	1 (0 1 0 0 0)
YP_001036921 Cthe_0490 CheA signal transduction histidine kinases	8.12	1.30	77010.8	1 (0 1 0 0 0)
YP_001036743 Cthe_0312 ATPase AAA-2	8.12	0.60	90556.2	1 (0 1 0 0 0)
YP_001036655 Cthe_0221 hypothetical protein	8.12	7.00	12916.6	1 (0 1 0 0 0)
YP_001037491 Cthe_1066 DNA repair protein RecO	8.11	2.40	28524.7	1 (0 1 0 0 0)
YP_001039457 Cthe_3068 periplasmic sensor signal transduction histidine kinase	8.11	1.5	54782	1 (0 1 0 0 0)
YP_001037847 Cthe_1423 protein of unknown function UPF0118	8.11	2.40	41042.1	1 (0 1 0 0 0)
YP_001037154 Cthe_0726 peptidase M18, aminopeptidase I	8.11	1.40	55802.7	1 (0 1 0 0 0)
YP_001039463 Cthe_3074 Cof-like hydrolase	8.11	1.60	35220.0	1 (0 1 0 0 0)
YP_001038139 Cthe_1725 hypothetical protein	8.10	4.90	11961.2	1 (0 1 0 0 0)
YP_001036669 Cthe_0237 response regulator receiver protein	8.10	4.60	15180.1	1 (0 1 0 0 0)
YP_001037883 Cthe_1461 FAD dependent oxidoreductase	6.16	2.10	58448.5	1 (0 0 1 0 0)
YP_001036993 Cthe_05Trans-hexaprenyltranstransferase	6.15	3.10	35809.4	1 (0 0 1 0 0)
YP_001038624 Cthe_2229 N-acetylneuraminate synthase	6.14	2.90	39121.8	1 (0 0 1 0 0)
YP_001038290 Cthe_1876 Tn7-like transposition protein A	6.14	3.90	32890.9	1 (0 0 1 0 0)
YP_001038756 Cthe_2361 DNA gyrase, A subunit	6.14	1.60	93408.1	1 (0 0 1 0 0)
YP_001037748 Cthe_1323 GrpE protein	6.14	5.30	25432.4	1 (0 0 1 0 0)
YP_001036895 Cthe_04flagellar hook-basal body complex subunit FliE	6.14	9.60	11494.9	1 (0 0 1 0 0)
YP_001038475 Cthe_2075 protein of unknown function DUF28	6.13	2.90	26973.5	1 (0 0 1 0 0)
YP_001037581 Cthe_1156 zinc finger, CHC2-type	6.13	1.30	79193.2	1 (0 0 1 0 0)
YP_001038886 Cthe_2491 hypothetical protein	6.13	6.10	15120.7	1 (0 0 1 0 0)

YP_001038846 Cthe_2451 SNF2-related protein	6.13	1.20	126155.2	1 (0 0 1 0 0)
YP_001036980 Cthe_0551 AMP-dependent synthetase and ligase	6.13	1.80	63015.9	1 (0 0 1 0 0)
YP_001036853 Cthe_0422 CoA-binding protein	6.13	4.80	24791.9	1 (0 0 1 0 0)
YP_001038154 Cthe_1740 BRO-like protein	6.13	3.50	29829.3	1 (0 0 1 0 0)
YP_001039242 Cthe_2850 transposase IS116/IS110/IS902	6.13	1.90	48736.2	1 (0 0 1 0 0)
YP_001037369 Cthe_0942 MiaB-like tRNA modifying enzyme YliG	6.13	1.80	51752.8	1 (0 0 1 0 0)
YP_001038035 Cthe_1618 hypothetical protein	6.13	1.40	80501.7	1 (0 0 1 0 0)
YP_001038773 Cthe_2378 parB-like partition protein	6.13	2.80	32855.7	1 (0 0 1 0 0)
YP_001036618 Cthe_0184 type IV pilus assembly protein PilM	6.13	1.90	41304.3	1 (0 0 1 0 0)
YP_001036756 Cthe_0325 NAD+ synthetase	6.12	0.80	72447.9	1 (0 0 1 0 0)
YP_001038183 Cthe_1769 hypothetical protein	6.12	6.20	15065.2	1 (0 0 1 0 0)
YP_001039441 Cthe_3052 YD repeat protein	6.12	0.30	334010.6	1 (0 0 1 0 0)
YP_001038545 Cthe_2149 amine oxidase	6.12	1.60	49732.1	1 (0 0 1 0 0)
YP_001037097 Cthe_0669 spore germination B3 GerAC like	6.11	1.4	48731	1 (0 0 1 0 0)
YP_001039054 Cthe_2662 RNA binding S1	6.11	4.30	15590.3	1 (0 0 1 0 0)
YP_001037187 Cthe_0759 flagellar biosynthesis	6.11	6.40	10391.6	1 (0 0 1 0 0)
YP_001037983 Cthe_15GCN5-related N-acetyltransferase	6.10	3.40	20189.4	1 (0 0 1 0 0)
YP_001037223 Cthe_0795 alpha amylase, catalytic region	4.15	1.40	67491.7	1 (0 0 0 1 0)
YP_001037614 Cthe_1189 ABC transporter related protein	4.14	3.10	36518.6	1 (0 0 0 1 0)
YP_001038178 Cthe_17Outer membrane protein-like protein	4.14	2.20	41742.3	1 (0 0 0 1 0)
YP_001036948 Cthe_0518 type III restriction enzyme, res subunit	4.13	1.00	114120.7	1 (0 0 0 1 0)
YP_001037405 Cthe_0980 hypothetical protein	4.13	6.80	18874.1	1 (0 0 0 1 0)
YP_001037487 Cthe_1062 VanW	4.13	1.70	54159.7	1 (0 0 0 1 0)
YP_0010382Cthe_1850 hypothetical protein	4.13	6.7	17669	1 (0 0 0 1 0)
YP_001038286 Cthe_1872 Tn7-like transposition protein C	4.13	1.10	62732.8	1 (0 0 0 1 0)
YP_001039010 Cthe_2617 peptidase M23B	4.13	3.3	33303	1 (0 0 0 1 0)
YP_001039210 Cthe_2818 CheA signal transduction histidine kinases	4.13	1.70	67790.0	1 (0 0 0 1 0)
YP_001037730 Cthe_1305 Cupin 2, conserved barrel	4.13	6.10	12422.5	1 (0 0 0 1 0)

Cthe_2984	4.13	6.3	22648	1 (0 0 0 1 0)
YP_001039366 Cthe_2976 hypothetical protein	4.11	7.60	9460.2	1 (0 0 0 1 0)
YP_001036505 Cthe_0071 Cellulose 1,4-beta-cellobiosidase	4.11	0.6	105671	1 (0 0 0 1 0)
YP_001036691 Cthe_0259 Rhomboid-like protein	4.10	3.1	21771	1 (0 0 0 1 0)
YP_001037496 Cthe_1071 PhoH-like protein	2.15	3.60	37547.6	1 (0 0 0 0 1)
YP_001038929 Cthe_2534 sulfate ABC transporter, ATPase subunit	2.14	1.70	39932.3	1 (0 0 0 0 1)
YP_001037903 Cthe_1481 membrane protein-like protein	2.14	0.8	83438	1 (0 0 0 0 1)
YP_001038659 Cthe_22H+-transporting two-sector ATPase, E subunit	2.14	3.50	22306.0	1 (0 0 0 0 1)
YP_001038957 Cthe_2563 hypothetical protein	2.14	1.9	68881	1 (0 0 0 0 1)
YP_001038149 Cthe_1735 phage DNA polymerase	2.13	2.00	66028.4	1 (0 0 0 0 1)
YP_001037931 Cthe_1510 short-chain dehydrogenase/reductase SDR	2.13	4.00	27380.9	1 (0 0 0 0 1)
YP_001039553 Cthe_3165 PpiC-type peptidyl-prolyl cis-trans isomerase	2.13	2.60	44023.3	1 (0 0 0 0 1)
YP_001039271 Cthe_2880 histidyl-tRNA synthetase	2.13	1.90	47276.4	1 (0 0 0 0 1)
YP_001038396 Cthe_1985 phage major capsid protein, HK97	2.13	1.90	46529.7	1 (0 0 0 0 1)
YP_001037024 Cthe_0596 GTP-binding protein	2.13	5.30	24130.9	1 (0 0 0 0 1)
YP_001039284 Cthe_2893 response regulator receiver protein	2.13	2.60	40751.9	1 (0 0 0 0 1)
YP_001038280 Cthe_1866 acetylornithine and succinylornithine aminotransferases	2.13	3.00	43892.6	1 (0 0 0 0 1)
YP_001037995 Cthe_1576 basic membrane lipoprotein	2.13	2.8	37576	1 (0 0 0 0 1)
YP_001038Cthe_2250 hypothetical protein	2.11	3.70	15634.2	1 (0 0 0 0 1)
YP_001038837 Cthe_2442 carbohydrate kinase, FGGY	2.11	1.40	54300.4	1 (0 0 0 0 1)

[A] XC, cross correlation score. Xc score as determined by Bioworks 3.3 program. The Xc score represents the cross correlation value between the observed peptide fragment mass spectrum and the one theoretically predicted. [b] Total peptide hits observed for the proteins (columns within parethesis represent the number of primary, secondary, tertiary,etc. peptide hits based spectral quality of the peptides spectra).

Appendix F; Membrane fraction protien identifications of cellobiose grown cells.

Table F1. Complete list of proteins identified in the membrane fraction of *C. thermocellum* grown on cellobiose, ranked by XC score.

Protein	Score (XC)	Coverage (%)	MW (daltons)	Peptide (Hits)
YP_001036854 Cthe_0423 iron-containing alcohol dehydrogenase	658.30	71.40	95948.2	84 (83 1 0 0 0)
YP_001038743 Cthe_2348 Ig-related protein	470.25	52.3	113261	57 (57 0 0 0 0)
YP_001039508 Cthe_3120 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	458.27	44.70	130938.9	52 (51 1 0 0 0)
YP_001037445 Cthe_1020 extracellular solute-binding protein, family 1	408.28	63.6	49954	55 (54 1 0 0 0)
YP_001036833 Cthe_0402 copper amine oxidase-like protein	350.25	47.3	83083	41 (41 0 0 0 0)
YP_001036772 Cthe_0341 NADH dehydrogenase quinone	340.23	51.80	64854.81	39 (39 0 0 0 0)
YP_001036860 Cthe_0429 NADH dehydrogenase quinone	328.24	51.60	68080.16	38 (37 1 0 0 0)
YP_001036773 Cthe_0342 hydrogenase, Fe-only	328.21	60.70	63972.4	38 (37 1 0 0 0)
YP_001036861 Cthe_0430 hydrogenase, Fe-only	308.23	45.90	62491.4	34 (33 1 0 0 0)
YP_001038276 Cthe_1862 ABC transporter related protein	270.27	56.50	42036.9	33 (32 1 0 0 0)
YP_001039122 Cthe_2730 translation elongation factor Tu	258.23	66.00	44160.6	36 (35 1 0 0 0)
YP_001039117 Cthe_2725 DNA-directed RNA polymerase, beta' subunit	254.19	24.60	130074.0	29 (28 0 0 1 0)
YP_001037127 Cthe_0699 carboxyl transferase	240.21	48.60	56012.8	26 (26 0 0 0 0)
YP_001039116 Cthe_2724 DNA-directed RNA polymerase, beta subunit	230.22	25.40	139869.0	23 (23 0 0 0 0)
YP_001038324 Cthe_1912 copper amine oxidase-like protein	230.21	46.5	59215	24 (24 0 0 0 0)
YP_001037810 Cthe_1385 preprotein translocase, SecA subunit	200.21	24.50	104546.4	21 (21 0 0 0 0)
YP_001036855 Cthe_0424 aminoglycoside phosphotransferase	192.25	69.90	28568.9	24 (22 1 0 1 0)
YP_001036571 Cthe_0137 glyceraldehyde-3-phosphate dehydrogenase, type I	190.23	64.00	36116.4	25 (25 0 0 0 0)
YP_001038631 Cthe_2236 flagellin-like protein	184.28	43.80	29549.1	25 (22 3 0 0 0)
YP_001037285 Cthe_0858 protein of unknown function DUF1432	170.29	52.6	35281	20 (20 0 0 0 0)
YP_001037530 Cthe_1105 type II secretion system protein	170.24	52.00	45131.5	22 (22 0 0 0 0)
YP_001037446 Cthe_1021 stage IV sporulation protein A	170.21	44.50	56148.2	19 (19 0 0 0 0)
YP_001039614 Cthe_3226 copper amine oxidase-	164.21	38.2	29604	19 (16 3 0 0 0)

like protein				
YP_001036857 Cthe_0426 putative PAS/PAC sensor protein	160.25	37.90	62518.0	17 (17 0 0 0 0)
YP_001036989 Cthe_0560 FAD-dependent pyridine nucleotide-disulphide oxidoreductase	160.23	34.90	66711.5	19 (19 0 0 0 0)
YP_001039323 Cthe_2932 DNA-directed RNAPα	160.22	49.50	35017.3	16 (16 0 0 0 0)
YP_001038241 Cthe_1827 copper amine oxidase-like protein	150.22	48.3	36389	16 (16 0 0 0 0)
YP_001039283 Cthe_2892 chaperonin GroEL	150.19	43.30	57438.5	15 (15 0 0 0 0)
YP_001038997 Cthe_2604 ATP synthase F0, B subunit	140.24	44.80	21057.1	19 (19 0 0 0 0)
YP_001037732 Cthe_1307 cellulosome anchoring protein, cohesin region	140.23	26.8	68577	15 (15 0 0 0 0)
YP_001036792 Cthe_0361 hypothetical protein	140.23	54.30	21979.2	15 (15 0 0 0 0)
YP_001037849 Cthe_1425 Inorganic diphosphatase	140.22	18.90	71337.7	14 (14 0 0 0 0)
YP_001036776 Cthe_0345 L-lactate dehydrogenase	140.21	44.7	34811	15 (15 0 0 0 0)
YP_001038379 Cthe_1968 cell divisionFtsK/SpoIIIE	136.20	15.80	121871.1	15 (13 2 0 0 0)
YP_001037129 Cthe_0701 conserved carboxylase region containing protein	130.25	30.30	52338.2	17 (17 0 0 0 0)
YP_001036510 Cthe_0076 hypothetical protein	128.34	67.4	13700	17 (16 1 0 0 0)
YP_001039113 Cthe_2721 ribosomal protein L1	128.20	43.30	25057.5	14 (13 1 0 0 0)
YP_001039306 Cthe_2915 ribosomal protein L5	120.23	52.20	20391.9	17 (17 0 0 0 0)
YP_001038778 Cthe_2383 hypothetical protein	120.22	42.4	33182	15 (15 0 0 0 0)
YP_001039100 Cthe_2708 hypothetical protein	120.21	32.6	53617	12 (12 0 0 0 0)
YP_0010372Cthe_0836 hypothetical protein	110.22	48.4	20373	13 (13 0 0 0 0)
YP_001039469 Cthe_3080 cellulosome anchoring protein, cohesin region	110.22	27.1	48500	12 (12 0 0 0 0)
YP_001036778 Cthe_0347 phosphofructokinase	110.21	38.10	45595.2	12 (12 0 0 0 0)
YP_001038803 Cthe_2408 phage shock protein A, PspA	110.19	43.40	27739.4	13 (13 0 0 0 0)
YP_001037636 Cthe_1211 pyridoxal-phosphate dependent TrpB-like enzyme	108.22	34.60	50100.1	11 (10 1 0 0 0)
YP_001038203 Cthe_1789 ATPase AAA-2	108.19	15.30	92041.6	11 (10 1 0 0 0)
YP_001037341 Cthe_0914 hypothetical protein	104.23	13.20	96251.1	13 (12 0 0 1 0)
YP_001037331 Cthe_0904 protein-export membrane protein SecD	100.25	28.3	48261	12 (12 0 0 0 0)
YP_001036775 Cthe_0344 Malate dehydrogenase	100.24	28.70	42073.3	10 (10 0 0 0 0)
YP_001039483 Cthe_3095 glycosyl transferase, family 39	100.23	15.7	110274	10 (10 0 0 0 0)
YP_001039294 Cthe_2903 ribosomal protein L3	100.21	31.60	23551.7	11 (11 0 0 0 0)
YP_001037523 Cthe_1098 hypothetical protein	100.20	27.7	49674	10 (10 0 0 0 0)
YP_001039322 Cthe_2931 ribosomal protein S4	100.20	37.00	23926.7	10 (10 0 0 0 0)
YP_001037788 Cthe_1363 lipopolysaccharide biosynthesis	100.19	24.60	52051.2	11 (11 0 0 0 0)
YP_001039482 Cthe_3094 glycosyl transferase, family 2	100.18	35.20	37622.8	11 (11 0 0 0 0)
YP_0010373Cthe_0937 fatty acid/phospholipid synthesis protein PlsX	98.24	33.00	36458.3	11 (10 1 0 0 0)
YP_001037793 Cthe_1368 S-layer-like domain containing protein	98.22	13.4	78075	13 (12 1 0 0 0)
YP_001036849 Cthe_0418 Polyribonucleotide nucleotidyltransferase	92.20	14.70	77269.9	10 (8 0 2 0 0)

YP_001038788 Cthe_2393 thiamine pyrophosphate enzyme-like TPP-binding	90.21	41.50	34626.8	10 (10 0 0 0 0)
YP_001038999 Cthe_2606 ATP synthase F1, alpha subunit	80.25	20.80	55784.9	8 (8 0 0 0 0)
YP_00103Cthe_0045 copper amine oxidase-like protein	80.23	35.7	30545	9 (9 0 0 0 0)
YP_001039324 Cthe_2933 ribosomal protein L17	80.21	43.70	19695.6	9 (9 0 0 0 0)
YP_001038566 Cthe_2171 type III restriction enzyme, res subunit	80.20	10.80	111910.8	8 (8 0 0 0 0)
YP_001039313 Cthe_2922 ribosomal protein L15	80.20	47.30	15576.6	9 (9 0 0 0 0)
YP_001039466 Cthe_3077 cellulosome anchoring protein, cohesin region	80.20	11.7	196710	9 (9 0 0 0 0)
YP_001039320 Cthe_2929 ribosomal protein S13	80.19	30.90	14158.9	9 (9 0 0 0 0)
YP_001037193 Cthe_0765 ribosomal protein L19	80.18	55.30	12924.2	8 (8 0 0 0 0)
YP_001037444 Cthe_1019 binding-protein-dependent transport systems inner membrane component	80.17	13.80	37355.9	9 (9 0 0 0 0)
YP_001036549 Cthe_0115 glycogen debranching enzyme, putative	80.16	10.80	76556.8	8 (8 0 0 0 0)
YP_001038652 Cthe_2257 ribosomal protein L9	76.16	45.30	16353.9	8 (7 0 1 0 0)
YP_001037527 Cthe_1102 fimbrial assembly protein	70.27	33.90	19870.7	9 (9 0 0 0 0)
YP_001039114 Cthe_2722 ribosomal protein L10	70.23	34.30	19501.7	9 (9 0 0 0 0)
YP_001038376 Cthe_1965 alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	70.23	52.90	20791.5	9 (9 0 0 0 0)
YP_001038787 Cthe_2392 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	70.22	23.10	43405.9	8 (8 0 0 0 0)
YP_001038197 Cthe_1783 ribosomal protein L13	70.21	33.80	16763.2	9 (9 0 0 0 0)
YP_001039297 Cthe_2906 ribosomal protein L2	70.21	32.40	30176.3	8 (8 0 0 0 0)
YP_001039308 Cthe_2917 ribosomal protein S8	70.20	50.80	14209.8	8 (8 0 0 0 0)
YP_001039311 Cthe_2920 ribosomal protein S5-like protein	70.20	61.40	17531.7	10 (10 0 0 0 0)
YP_001037416 Cthe_0991 translation initiation factor IF-2	70.19	7.80	114885.3	7 (7 0 0 0 0)
YP_001038817 Cthe_2422 hypothetical protein	70.19	27.5	40412	8 (8 0 0 0 0)
YP_001037248 Cthe_0820 hypothetical protein	70.19	16.40	77770.0	8 (8 0 0 0 0)
YP_001039615 Cthe_3227 copper amine oxidase-like protein	70.18	19.1	30241	7 (7 0 0 0 0)
YP_001038196 Cthe_1782 ribosomal protein S9	70.18	44.60	14656.2	7 (7 0 0 0 0)
YP_001039484 Cthe_3096 hypothetical protein	68.18	24.4	33802	7 (6 1 0 0 0)
YP_001037Cthe_1221 glycosyltransferase 36	68.16	3.0	333880	7 (6 1 0 0 0)
YP_001039530 Cthe_3142 hypothetical protein	60.24	39.50	18433.0	6 (6 0 0 0 0)
YP_001039467 Cthe_3078 cellulosome anchoring protein, cohesin region	60.22	3.4	248014	7 (7 0 0 0 0)
YP_001037388 Cthe_0961 aspartate-semialdehyde dehydrogenase	60.21	20.20	36705.8	6 (6 0 0 0 0)
YP_001037330 Cthe_0903 protein-export membrane protein SecF	60.21	16.40	34990.8	7 (7 0 0 0 0)
YP_001038752 Cthe_2357 hypothetical protein	60.19	19.5	44618	6 (6 0 0 0 0)
YP_001037531 Cthe_1106 twitching motility protein	60.19	27.40	38896.1	7 (7 0 0 0 0)
YP_001039295 Cthe_2904 ribosomal protein L4/L1e	60.18	33.20	23295.8	7 (7 0 0 0 0)
YP_001037107 Cthe_0679 Serine-type D-Ala-D-Ala carboxypeptidase	60.17	18.1	44618	6 (6 0 0 0 0)

YP_001036780 Cthe_0349 fructose-1,6-bisphosphate aldolase, class II	60.17	27.50	33611.2	6 (6 0 0 0 0)
YP_001036782 Cthe_0351 PpiC-type peptidyl-prolyl cis-trans isomerase	60.16	12.7	53834	6 (6 0 0 0 0)
YP_001037653 Cthe_1228 threonyl-tRNA synthetase	54.17	10.40	73518.6	7 (6 0 0 1 0)
YP_001039301 Cthe_2910 ribosomal protein L16	50.24	46.90	16213.7	5 (5 0 0 0 0)
YP_001039305 Cthe_2914 ribosomal protein L24	50.24	50.90	12808.0	5 (5 0 0 0 0)
YP_001037524 Cthe_1099 hypothetical protein	50.21	6.9	94172	5 (5 0 0 0 0)
YP_001037924 Cthe_1503 hypothetical protein	50.21	37.80	13222.8	6 (6 0 0 0 0)
YP_001038247 Cthe_1833 copper amine oxidase-like protein	50.21	25.7	30568	6 (6 0 0 0 0)
YP_001039309 Cthe_2918 ribosomal protein L6	50.21	29.50	20135.1	5 (5 0 0 0 0)
YP_001039299 Cthe_2908 ribosomal protein L22	50.20	50.40	14066.9	6 (6 0 0 0 0)
YP_001039300 Cthe_2909 ribosomal protein S3	50.20	28.80	25312.8	6 (6 0 0 0 0)
YP_001037526 Cthe_1101 hypothetical protein	50.20	9.50	48018.2	5 (5 0 0 0 0)
YP_001037516 Cthe_1091 metal dependent phosphohydrolase	50.19	9.90	58726.7	6 (6 0 0 0 0)
YP_001036876 Cthe_0445 cell division protein FtsZ	50.19	18.10	39703.8	5 (5 0 0 0 0)
YP_001037202 Cthe_0774 RNA chaperone Hfq	50.19	81.70	9361.9	5 (5 0 0 0 0)
YP_001036891 Cthe_0460 DNA topoisomerase I	50.18	8.50	79261.1	5 (5 0 0 0 0)
YP_001036742 Cthe_0311 excinuclease ABC, A subunit	50.18	6.40	104767.6	5 (5 0 0 0 0)
YP_001039510 Cthe_3122 S-layer-like domain containing protein	50.18	11.0	77594	5 (5 0 0 0 0)
YP_001038580 Cthe_2185 ribosomal protein S18	50.18	38.30	10856.1	5 (5 0 0 0 0)
YP_001038466 Cthe_2066 serine O-acetyltransferase	50.17	35.10	27382.2	5 (5 0 0 0 0)
YP_001037033 Cthe_0605 NLP/P60	50.17	17.0	39569	5 (5 0 0 0 0)
YP_001039314 Cthe_2923 preprotein translocase, SecY subunit	50.17	9.40	47339.9	5 (5 0 0 0 0)
YP_001039101 Cthe_2709 hypothetical protein	50.16	12.6	55116	5 (5 0 0 0 0)
YP_001037532 Cthe_1107 type II secretion system protein E	50.16	6.60	88291.6	5 (5 0 0 0 0)
YP_001039304 Cthe_2913 ribosomal protein L14	50.16	33.60	13446.4	6 (6 0 0 0 0)
YP_001037709 Cthe_1284 glycogen/starch synthases, ADP-glucose type	50.16	12.50	55860.0	5 (5 0 0 0 0)
YP_001036729 Cthe_0298 methyl-accepting chemotaxis sensory transducer	50.15	8.80	65032.8	5 (5 0 0 0 0)
YP_001039519 Cthe_3131 von Willebrand factor, type A	50.15	11.5	62240	5 (5 0 0 0 0)
YP_001036569 Cthe_0135 beta-ketoacyl synthase	46.16	2.10	305695.5	5 (3 2 0 0 0)
YP_001039310 Cthe_2919 ribosomal protein L18	40.24	28.70	13690.6	4 (4 0 0 0 0)
YP_001039321 Cthe_2930 ribosomal protein S11	40.23	21.60	14342.7	5 (5 0 0 0 0)
YP_001039387 Cthe_2998 ABC transporter related protein	40.20	10.10	63369.4	4 (4 0 0 0 0)
YP_001038938 Cthe_2543 Spore germination protein-like protein	40.20	8.80	52690.6	4 (4 0 0 0 0)
YP_001038996 Cthe_2603 ATP synthase F0, C subunit	40.19	21.9	7218	5 (5 0 0 0 0)
YP_001039551 Cthe_3163 Carbohydrate binding family 25	40.19	22.70	17688.1	4 (4 0 0 0 0)
YP_001039120 Cthe_2728 ribosomal protein S7	40.19	38.50	17896.5	4 (4 0 0 0 0)

YP_001038785 Cthe_2390 pyruvate/ketoisovalerate oxidoreductase, gamma subunit	40.19	28.10	21190.1	4 (4 0 0 0 0)
YP_001038574 Cthe_2179 Pectate lyase/Amb allergen	40.19	4.9	99829	4 (4 0 0 0 0)
YP_001037525 Cthe_1100 prepilin-type cleavage/methylation	40.19	22.50	19349.2	4 (4 0 0 0 0)
YP_001038939 Cthe_2544 RNA polymerase, sigma-24 subunit, ECF subfamily	40.18	26.70	22184.4	4 (4 0 0 0 0)
YP_001039468 Cthe_3079 cellulosome anchoring protein, cohesin region	40.18	8.4	74925	4 (4 0 0 0 0)
YP_001037128 Cthe_0700 biotin/lipoyl attachment protein	40.18	39.40	13653.3	4 (4 0 0 0 0)
YP_001037676 Cthe_1251 Xanthine/uracil/vitamin C permease	40.18	5.6	49394	5 (5 0 0 0 0)
YP_001038771 Cthe_2376 DNA gyrase, B subunit	40.18	7.50	72446.5	4 (4 0 0 0 0)
YP_001039001 Cthe_2608 ATP synthase F1, beta subunit	40.17	10.60	50980.7	4 (4 0 0 0 0)
YP_001038243 Cthe_1829 Chromosome segregation ATPases-like protein	40.17	4.6	124588	5 (2 2 0 1 0)
YP_001037122 Cthe_0694 spermidine synthase	40.17	20.40	31334.1	4 (4 0 0 0 0)
YP_001036594 Cthe_0160 ribosomal protein L21	40.17	48.50	11531.2	4 (4 0 0 0 0)
YP_001039110 Cthe_2718 preprotein translocase, SecE subunit	40.16	23.80	9175.1	4 (4 0 0 0 0)
YP_001039121 Cthe_2729 translation elongation factor G	40.16	7.50	77586.2	4 (4 0 0 0 0)
YP_001036781 Cthe_0350 signal peptidase I	40.16	22.30	22212.4	4 (4 0 0 0 0)
YP_001039303 Cthe_2912 ribosomal protein S17	40.16	37.60	10034.5	5 (5 0 0 0 0)
YP_001038797 Cthe_2402 Peptidoglycan-binding LysM	40.16	8.80	58641.8	4 (4 0 0 0 0)
YP_001037431 Cthe_1006 ribosomal protein S2	40.15	17.50	28590.0	4 (4 0 0 0 0)
YP_001037104 Cthe_0676 tyrosine recombinase XerD	40.15	15.20	34269.4	4 (4 0 0 0 0)
YP_001039296 Cthe_2905 Ribosomal protein L25/L23	38.19	42.70	13281.3	5 (4 1 0 0 0)
YP_001037Cthe_1223 ribosomal protein L20	38.19	21.40	13356.4	5 (4 1 0 0 0)
YP_001039419 Cthe_3030 methyl-accepting chemotaxis sensory transducer	38.15	5.4	91253	4 (3 1 0 0 0)
YP_001038262 Cthe_1848 copper-translocating P-type ATPase	36.17	6.30	79962.7	4 (3 0 1 0 0)
YP_00103Cthe_0056 Ig-like, group 2	36.16	0.8	493539	4 (2 2 0 0 0)
YP_001038706 Cthe_2311 hypothetical protein	36.16	22.6	27741	4 (3 0 1 0 0)
YP_001038577 Cthe_2182 Ig-like, group 2	36.15	4.6	101483	4 (2 2 0 0 0)
YP_001037293 Cthe_0866 pyruvate flavodoxin/ferredoxin oxidoreductase-like protein	32.17	14.20	38061.7	4 (2 1 0 1 0)
YP_001037359 Cthe_0932 beta-ketoacyl synthase	30.26	9.50	43936.0	3 (3 0 0 0 0)
YP_001037240 Cthe_0812 response regulator receiver protein	30.25	10.00	30140.0	3 (3 0 0 0 0)
YP_001038586 Cthe_2191 1,4-alpha-glucan branching enzyme	30.22	5.80	86603.8	3 (3 0 0 0 0)
YP_001036939 Cthe_0509 sodium ion-translocating decarboxylase, beta subunit	30.22	7.80	38047.1	3 (3 0 0 0 0)
YP_001037475 Cthe_1050 recA protein	30.21	15.20	37781.9	3 (3 0 0 0 0)
YP_001038998 Cthe_2605 ATP synthase F1, delta	30.21	19.70	21896.0	3 (3 0 0 0 0)

subunit				
YP_001038329 Cthe_1917 ATPase, P-type transporting , HAD superfamily, subfamily IC	30.20	4.50	98745.0	3 (3 0 0 0 0)
YP_001038912 Cthe_2517 acetolactate synthase, small subunit	30.20	25.30	18609.0	4 (4 0 0 0 0)
YP_001037283 Cthe_0856 branched-chain amino acid aminotransferase	30.20	14.90	39711.6	3 (3 0 0 0 0)
YP_001037845 Cthe_1421 signal peptide peptidase SppA, 36K type	30.19	12.6	35780	3 (3 0 0 0 0)
YP_001038528 Cthe_2130 hypothetical protein	30.19	8.10	48488.7	3 (3 0 0 0 0)
YP_001039085 Cthe_2693 hypothetical protein	30.19	25.50	17809.5	3 (3 0 0 0 0)
YP_001039392 Cthe_3003 hydrogenase, Fe-only	30.19	7.10	71733.0	3 (3 0 0 0 0)
YP_001036572 Cthe_0138 Phosphoglycerate kinase	30.18	10.60	42735.3	4 (4 0 0 0 0)
YP_001038779 Cthe_2384 S-layer-like domain containing protein	30.18	19.0	21039	3 (3 0 0 0 0)
YP_001038Cthe_2253 ATP-dependent metalloprotease FtsH	30.18	8.0	66641	3 (3 0 0 0 0)
YP_001036858 Cthe_0427 serine phosphatase	30.18	10.00	43202.3	3 (3 0 0 0 0)
YP_001039293 Cthe_2902 ribosomal protein S10	30.18	35.00	11770.6	3 (3 0 0 0 0)
YP_001037141 Cthe_0713 1-acyl-sn-glycerol-3-phosphate acyltransferase	30.18	17.70	22765.5	3 (3 0 0 0 0)
YP_001039298 Cthe_2907 ribosomal protein S19	30.17	31.90	10589.7	5 (5 0 0 0 0)
YP_001037266 Cthe_0838 hypothetical protein	30.17	17.2	24076	3 (3 0 0 0 0)
YP_001036999 Cthe_0570 peptidase, membrane zinc metallopeptidase, putative	30.17	18.50	24747.4	3 (3 0 0 0 0)
YP_001038556 Cthe_2161 hypothetical protein	30.17	13.00	33780.1	3 (3 0 0 0 0)
YP_001038215 Cthe_1801 ABC transporter related protein	30.17	10.60	28160.7	3 (3 0 0 0 0)
YP_001037500 Cthe_1075 protein of unknown function DUF881	30.16	14.6	26506	3 (3 0 0 0 0)
YP_001037384 Cthe_0957 preprotein translocase, YajC subunit	30.16	37.40	10950.1	4 (4 0 0 0 0)
YP_001037799 Cthe_1374 copper amine oxidase-like protein	30.16	9.20	30003.5	3 (3 0 0 0 0)
YP_001038209 Cthe_1795 phospho-2-dehydro-3-deoxyheptonate aldolase	30.16	13.50	36794.5	3 (3 0 0 0 0)
YP_001037430 Cthe_1005 translation elongation factor Ts	30.15	15.80	24321.7	3 (3 0 0 0 0)
YP_001039025 Cthe_2633 hypothetical protein	30.15	8.5	41666	4 (4 0 0 0 0)
YP_001037066 Cthe_0638 FHA domain containing protein	30.15	1.90	175123.4	3 (3 0 0 0 0)
YP_001039119 Cthe_2727 ribosomal protein S12	30.15	15.50	15415.7	3 (3 0 0 0 0)
YP_001038323 Cthe_1911 Carbohydrate binding family 6	30.15	2.5	146588	3 (3 0 0 0 0)
YP_001037381 Cthe_0954 Uracil phosphoribosyltransferase	30.13	11.40	20402.0	3 (3 0 0 0 0)
YP_001038509 Cthe_2109 copper amine oxidase-like protein	28.19	5.7	94460	3 (2 1 0 0 0)
YP_001037365 Cthe_0938 regulatory protein, DeoR	28.19	18.60	21123.7	3 (2 1 0 0 0)
YP_001037055 Cthe_0627 hypothetical protein	28.16	15.80	33161.2	3 (2 1 0 0 0)
YP_001036960 Cthe_0531 hypothetical protein	28.13	2.0	87440	3 (2 1 0 0 0)
YP_001037252 Cthe_0824 copper amine oxidase-like protein	26.14	7.6	67339	3 (2 0 1 0 0)

YP_001039588 Cthe_3200 alanyl-tRNA synthetase	24.15	4.50	97985.4	3 (2 0 0 1 0)
YP_001037733 Cthe_1308 pyruvate, phosphate dikinase	24.14	4.20	98703.0	3 (1 1 1 0 0)
YP_001037792 Cthe_1367 PHP-like protein	24.14	12.00	29478.5	3 (1 1 1 0 0)
YP_001036771 Cthe_0340 ferredoxin	24.14	27.90	13424.0	3 (1 1 1 0 0)
YP_001037389 Cthe_0962 dihydrodipicolinate synthase	20.23	7.20	31954.9	2 (2 0 0 0 0)
YP_001036578 Cthe_0144 preprotein translocase, SecG subunit	20.23	21.7	8797	2 (2 0 0 0 0)
YP_001036527 Cthe_0093 septum site-determining protein MinD	20.22	12.00	28892.5	2 (2 0 0 0 0)
YP_001037247 Cthe_0819 ABC transporter related protein	20.20	9.80	32076.8	2 (2 0 0 0 0)
YP_001038351 Cthe_1940 RDD domain containing protein	20.20	9.30	28171.3	2 (2 0 0 0 0)
YP_001038281 Cthe_1867 carbamoyl-phosphate synthase, small subunit	20.20	8.50	39325.1	2 (2 0 0 0 0)
YP_001039022 Cthe_2630 ribose-phosphate pyrophosphokinase	20.19	9.40	34730.4	2 (2 0 0 0 0)
YP_001037Cthe_1224 ribosomal protein L35	20.19	23.10	7351.4	2 (2 0 0 0 0)
YP_001037403 Cthe_0978 UDP-N-acetylmuramyl-tripeptide synthetases	20.19	7.20	53768.8	2 (2 0 0 0 0)
Cthe_3088	20.19	21.6	7948	2 (2 0 0 0 0)
YP_001037361 Cthe_0934 3-oxoacyl- acyl-carrier-protein reductase	20.19	8.10	26061.6	2 (2 0 0 0 0)
YP_001037046 Cthe_0618 ferrous iron transport protein B	20.19	4.40	78985.9	2 (2 0 0 0 0)
YP_001037630 Cthe_1205 putative serine protein kinase, PrkA	20.19	3.60	74227.2	3 (3 0 0 0 0)
YP_001036842 Cthe_0411 hypothetical protein	20.18	7.80	39706.4	2 (2 0 0 0 0)
YP_001036859 Cthe_0428 NADH dehydrogenase ubiquinone , 24 kDa subunit	20.18	14.50	18338.3	2 (2 0 0 0 0)
YP_001037727 Cthe_1302 hypothetical protein	20.18	4.90	61496.7	2 (2 0 0 0 0)
YP_001039538 Cthe_3150 cobalamin biosynthesis protein CobD	20.18	6.60	39372.0	3 (3 0 0 0 0)
YP_001037754 Cthe_1329 putative CoA-substrate-specific enzyme activase	20.18	2.40	160075.3	2 (2 0 0 0 0)
YP_001038371 Cthe_1960 Peptidoglycan-binding domain 1	20.18	6.9	38782	2 (2 0 0 0 0)
YP_001037808 Cthe_1383 Tetratricopeptide TPR_2	20.18	4.70	45162.9	2 (2 0 0 0 0)
YP_001036862 Cthe_0431 hypothetical protein	20.18	30.90	9129.7	2 (2 0 0 0 0)
YP_001039451 Cthe_3062 signal transduction histidine kinase regulating citrate/malate metabolism	20.18	7.4	50666	2 (2 0 0 0 0)
YP_0010374Cthe_1039 ribosomal protein S20	20.18	20.90	12988.5	2 (2 0 0 0 0)
YP_001036511 Cthe_0077 hypothetical protein	20.18	16.50	13916.6	2 (2 0 0 0 0)
YP_001039544 Cthe_3156 methyl-accepting chemotaxis sensory transducer	20.17	6.7	46201	2 (2 0 0 0 0)
YP_001039268 Cthe_2877 S-layer-like domain containing protein	20.17	3.6	67015	2 (2 0 0 0 0)
YP_001036596 Cthe_0162 ribosomal protein L27	20.17	22.80	10032.6	2 (2 0 0 0 0)
YP_001037065 Cthe_0637 protein of unknown function DUF909	20.17	32.00	11050.8	3 (3 0 0 0 0)
YP_001037178 Cthe_0750 spermidine/putrescine	20.17	11.10	39800.7	2 (2 0 0 0 0)

ABC transporter ATPase subunit				
YP_001036534 Cthe_0100 hypothetical protein	20.17	7.7	30310	2 (2 0 0 0 0)
YP_001037390 Cthe_0963 Dihydrodipicolinate reductase	20.17	10.30	27775.6	2 (2 0 0 0 0)
YP_00103Cthe_0052 hypothetical protein	20.17	4.5	46894	2 (2 0 0 0 0)
YP_001036769 Cthe_0338 NADH-quinone oxidoreductase, E subunit	20.17	21.20	18057.1	2 (2 0 0 0 0)
YP_001037042 Cthe_0614 pyruvate ferredoxin/flavodoxin oxidoreductase	20.17	19.90	21269.4	2 (2 0 0 0 0)
YP_001037925 Cthe_1504 Linocin_M18 bacteriocin protein	20.17	10.00	30295.9	2 (2 0 0 0 0)
YP_001036506 Cthe_0072 phage shock protein C, PspC	20.17	15.50	18036.7	2 (2 0 0 0 0)
YP_001037223 Cthe_0795 alpha amylase, catalytic region	20.17	1.90	67491.7	2 (2 0 0 0 0)
YP_001037291 Cthe_08pyruvate ferredoxin/flavodoxin oxidoreductase	20.16	12.80	19330.0	2 (2 0 0 0 0)
YP_001038762 Cthe_2367 60 kDa inner membrane insertion protein	20.16	6.50	33187.5	2 (2 0 0 0 0)
YP_001036577 Cthe_0143 Phosphopyruvate hydratase	20.16	4.60	47073.2	2 (2 0 0 0 0)
YP_001037650 Cthe_1225 translation initiation factor IF-3	20.16	7.90	18832.1	2 (2 0 0 0 0)
YP_001039012 Cthe_2619 cell shape determining protein, MreB/Mrl family	20.16	5.80	36744.6	2 (2 0 0 0 0)
YP_001038913 Cthe_2518 ketol-acid reductoisomerase	20.16	7.60	36199.4	2 (2 0 0 0 0)
YP_00103Cthe_0053 ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent	20.16	2.80	88248.5	2 (2 0 0 0 0)
YP_001036675 Cthe_0243 copper amine oxidase-like protein	20.16	9.9	31681	2 (2 0 0 0 0)
YP_001037292 Cthe_0865 3-methyl-2-oxobutanoate dehydrogenase ferredoxin	20.16	9.30	26948.7	2 (2 0 0 0 0)
YP_001038774 Cthe_2379 hypothetical protein	20.16	14.20	18996.1	2 (2 0 0 0 0)
YP_001037443 Cthe_1018 binding-protein-dependent transport systems inner membrane component	20.15	8.7	31905	2 (2 0 0 0 0)
YP_001039160 Cthe_2768 metallophosphoesterase	20.15	2.30	98713.7	2 (2 0 0 0 0)
YP_001039002 Cthe_2609 ATP synthase F1, epsilon subunit	20.15	17.00	15369.2	2 (2 0 0 0 0)
YP_001038547 Cthe_2151 hypothetical protein	20.15	3.40	83178.1	2 (2 0 0 0 0)
YP_001038541 Cthe_2144 DNA polymerase III, subunits gamma and tau	20.15	4.20	61074.7	2 (2 0 0 0 0)
YP_001036751 Cthe_0320 hypothetical protein	20.15	4.90	55540.0	2 (2 0 0 0 0)
YP_001037487 Cthe_1062 VanW	20.15	1.70	54159.7	3 (3 0 0 0 0)
YP_001037433 Cthe_1008 aminodeoxychorismate lyase	20.15	5.9	42861	2 (2 0 0 0 0)
YP_001039312 Cthe_2921 ribosomal protein L30	20.15	35.60	6614.7	2 (2 0 0 0 0)
YP_001038992 Cthe_2599 CMP/dCMP deaminase, zinc-binding	20.15	15.00	16954.7	2 (2 0 0 0 0)
YP_001038544 Cthe_2148 Carbohydrate-binding, CenC-like protein	20.15	9.4	28392	2 (2 0 0 0 0)
YP_001037770 Cthe_1345 adenine phosphoribosyltransferase	20.15	17.00	19068.2	2 (2 0 0 0 0)

YP_001038446 Cthe_2046 hypothetical protein	20.14	17.10	19878.0	2 (2 0 0 0 0)
YP_001037806 Cthe_1381 threonine synthase	20.14	5.00	55667.3	2 (2 0 0 0 0)
YP_001036832 Cthe_0401 methyl-accepting chemotaxis sensory transducer	20.14	3.20	85087.8	2 (2 0 0 0 0)
YP_001036685 Cthe_0253 ATPase associated with various cellular activities, AAA_3	20.14	7.00	36907.5	2 (2 0 0 0 0)
YP_001038804 Cthe_2409 hypothetical protein	20.14	6.0	31530	2 (2 0 0 0 0)
YP_001037529 Cthe_1104 prepilin-type cleavage/methylation	20.13	6.3	19225	2 (2 0 0 0 0)
YP_001039616 Cthe_3228 copper amine oxidase-like protein	20.13	11.3	30072	2 (2 0 0 0 0)
YP_001038773 Cthe_2378 parB-like partition protein	20.13	6.60	32855.7	2 (2 0 0 0 0)
YP_001036579 Cthe_0145 metal dependent phosphohydrolase	20.13	9.80	20528.8	2 (2 0 0 0 0)
YP_001037041 Cthe_0613 thiamine pyrophosphate enzyme-like TPP-binding	20.13	3.70	65986.8	2 (2 0 0 0 0)
YP_001039115 Cthe_2723 ribosomal protein L7/L12	18.18	13.20	13287.3	2 (1 1 0 0 0)
YP_001037660 Cthe_1235 Cellulose 1,4-beta-cellobiosidase	18.17	0.50	758397.6	2 (1 1 0 0 0)
YP_001037625 Cthe_1200 adenosylhomocysteinase	18.16	7.70	45721.3	3 (0 1 1 1 0)
YP_001039403 Cthe_3014 hydrogenase formation HypD protein	18.15	11.20	38157.5	2 (1 1 0 0 0)
YP_001037458 Cthe_1033 glutamyl-tRNA Gln amidotransferase, C subunit	18.15	33.70	11015.8	2 (1 1 0 0 0)
YP_001038917 Cthe_2522 membrane associated protein	18.14	5.00	61939.5	3 (1 0 1 0 1)
YP_001038678 Cthe_2283 methyl-accepting chemotaxis sensory transducer	18.14	1.70	154468.8	2 (1 1 0 0 0)
YP_001038110 Cthe_1695 Radical SAM	18.13	3.10	56024.2	2 (1 1 0 0 0)
YP_001039004 Cthe_2611 Fibronectin, type III	18.13	1.0	183583	2 (1 1 0 0 0)
YP_001039087 Cthe_2695 hypothetical protein	16.15	6.20	25498.3	2 (1 0 1 0 0)
YP_001037319 Cthe_0892 protein of unknown function DUF34	16.15	7.90	40998.2	2 (0 2 0 0 0)
YP_001038217 Cthe_1803 cobalamin vitamin B12 biosynthesis CbiM protein	16.15	6.0	37379	2 (1 0 1 0 0)
YP_001037039 Cthe_0611 RNA-metabolising metallo-beta-lactamase	16.14	4.60	62041.1	2 (1 0 1 0 0)
YP_001036898 Cthe_0467 Flagellar biosynthesis/type III secretory pathway protein-like protein	16.14	4.20	29946.8	2 (1 0 1 0 0)
YP_001037610 Cthe_1185 hypothetical protein	16.14	3.30	50416.2	2 (1 0 1 0 0)
YP_001039279 Cthe_2888 imidazoleglycerol phosphate synthase, cyclase subunit	14.16	7.50	27690.4	2 (1 0 0 1 0)
YP_001039424 Cthe_3035 D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	14.15	6.40	42835.4	2 (1 0 0 1 0)
YP_001038807 Cthe_2412 SMC protein-like protein	14.14	1.90	101479.0	2 (0 1 1 0 0)
YP_001038436 Cthe_2032 hypothetical protein	14.14	0.6	223041	2 (0 1 1 0 0)
YP_001039625 Cthe_3237 Relaxase/mobilization nuclease domain containing protein	14.13	5.40	53453.3	2 (1 0 0 1 0)
YP_001037819 Cthe_1394 two component transcriptional regulator, winged helix family	12.16	4.20	28066.6	2 (0 0 2 0 0)
YP_001037614 Cthe_1189 ABC transporter related protein	12.15	5.60	36518.6	2 (0 1 0 1 0)

YP_001037167 Cthe_0739 hypothetical protein	12.14	1.70	63759.5	2 (0 0 2 0 0)
YP_001038220 Cthe_1806 cellulosome enzyme, dockerin type I	12.14	1.8	235980	2 (0 1 0 1 0)
YP_001037307 Cthe_0880 phospho-2-dehydro-3-deoxyheptonate aldolase	12.13	8.00	36848.7	2 (1 0 0 0 1)
YP_001038658 Cthe_2263 H+-transporting two-sector ATPase, C subunit	10.23	12.3	15967	1 (1 0 0 0 0)
YP_00103Cthe_0047 copper amine oxidase-like protein	10.21	8.2	30435	1 (1 0 0 0 0)
YP_001039409 Cthe_3020 NADH dehydrogenase ubiquinone	10.20	4.20	40421.0	1 (1 0 0 0 0)
YP_001036553 Cthe_0119 putative anti-sigma regulatory factor, serine/threonine protein kinase	10.20	15.40	16526.2	1 (1 0 0 0 0)
YP_001038500 Cthe_2100 transcriptional regulator, AbrB family	10.19	16.50	8910.8	1 (1 0 0 0 0)
YP_001038757 Cthe_2362 parB-like partition protein	10.19	4.30	31921.3	1 (1 0 0 0 0)
YP_001039106 Cthe_2714 acetolactate synthase, large subunit, biosynthetic type	10.18	3.40	60696.2	1 (1 0 0 0 0)
YP_001039421 Cthe_3032 hypothetical protein	10.18	10.40	15321.7	1 (1 0 0 0 0)
YP_001039455 Cthe_3066 ABC transporter related protein	10.18	6.60	27035.2	1 (1 0 0 0 0)
YP_001037423 Cthe_0998 putative membrane-associated zinc metalloprotease	10.18	3.50	46814.3	1 (1 0 0 0 0)
YP_001036726 Cthe_0295 phosphoserine aminotransferase	10.18	4.30	41362.3	1 (1 0 0 0 0)
YP_001039067 Cthe_2675 hypothetical protein	10.18	7.2	25930	1 (1 0 0 0 0)
YP_001036983 Cthe_0554 phosphoribosylformylglycinamidine synthase	10.18	1.50	139261.8	1 (1 0 0 0 0)
YP_001039266 Cthe_2875 sigma 54 modulation protein/ribosomal protein S30EA	10.17	8.50	20531.7	1 (1 0 0 0 0)
YP_001039416 Cthe_3027 Citrate Si -synthase	10.17	3.70	52000.7	1 (1 0 0 0 0)
YP_001038632 Cthe_2237 flagellin-like protein	10.17	4.40	29413.2	1 (1 0 0 0 0)
YP_001039434 Cthe_3045 FHA domain containing protein	10.17	9.20	19762.7	1 (1 0 0 0 0)
YP_001039104 Cthe_2712 hypothetical protein	10.17	5.40	26108.5	1 (1 0 0 0 0)
YP_001039049 Cthe_2657 histone-like DNA-binding protein	10.17	15.40	10081.5	1 (1 0 0 0 0)
YP_001039393 Cthe_3004 ferredoxin	10.17	4.10	54617.1	1 (1 0 0 0 0)
YP_001038119 Cthe_1705 hypothetical protein	10.17	3.60	65110.9	1 (1 0 0 0 0)
YP_00103Cthe_0036 hybrid cluster protein	10.17	2.20	59404.5	1 (1 0 0 0 0)
YP_001038451 Cthe_2051 CRISPR-associated RAMP protein, SSO1426 family	10.17	10.80	29381.0	1 (1 0 0 0 0)
YP_001037717 Cthe_1292 proposed homoserine kinase	10.17	4.50	45158.0	2 (2 0 0 0 0)
YP_001037295 Cthe_0868 type IV pilus assembly PilZ	10.17	5.10	27940.0	1 (1 0 0 0 0)
YP_001037675 Cthe_1250 phosphoribosylaminoimidazole carboxylase, catalytic subunit	10.17	12.70	18407.8	1 (1 0 0 0 0)
YP_001036714 Cthe_0282 glycogen/starch synthases, ADP-glucose type	10.17	2.30	55693.2	1 (1 0 0 0 0)
YP_001038372 Cthe_1961 Nucleotidyl transferase	10.17	2.10	90642.3	1 (1 0 0 0 0)
YP_001037356 Cthe_0929 Stage V sporulation	10.16	12.80	8862.9	1 (1 0 0 0 0)

protein S				
YP_001038319 Cthe_1907 amino acid adenylation domain	10.16	4.30	69674.4	1 (1 0 0 0 0)
YP_001037058 Cthe_0630 hypothetical protein	10.16	4.90	37146.1	1 (1 0 0 0 0)
YP_001037342 Cthe_0915 hypothetical protein	10.16	2.8	46648	1 (1 0 0 0 0)
YP_001039142 Cthe_2750 IstB-like ATP-binding protein	10.16	5.40	28072.8	1 (1 0 0 0 0)
YP_001036788 Cthe_0357 alpha-glucan phosphorylases	10.16	1.40	97651.2	1 (1 0 0 0 0)
YP_001037332 Cthe_0905 metal dependent phosphohydrolase	10.16	9.50	24726.8	1 (1 0 0 0 0)
YP_001037241 Cthe_0813 peptidase S55, SpoIVB	10.16	7.2	48826	1 (1 0 0 0 0)
YP_001037708 Cthe_1283 endonuclease III	10.16	5.60	24307.7	1 (1 0 0 0 0)
YP_001036514 Cthe_0080 CheW protein	10.16	11.20	16400.7	1 (1 0 0 0 0)
YP_001038905 Cthe_2510 protein of unknown function DUF368	10.16	4.20	30692.6	1 (1 0 0 0 0)
YP_001037074 Cthe_0anaerobic ribonucleoside-triphosphate reductase	10.16	1.10	79998.1	2 (2 0 0 0 0)
YP_001037084 Cthe_0656 type IV pilus assembly protein PilM	10.16	2.40	41929.3	1 (1 0 0 0 0)
YP_001036666 Cthe_0234 AMP-dependent synthetase and ligase	10.16	2.80	95168.9	1 (1 0 0 0 0)
YP_001037401 Cthe_0976 phospho-N-acetylmuramoyl-pentapeptide- transferase	10.16	3.30	36463.9	1 (1 0 0 0 0)
YP_001039567 Cthe_3179 Serine-type D-Ala-D-Ala carboxypeptidase	10.16	5.3	46297	1 (1 0 0 0 0)
YP_001038906 Cthe_2511 MscS Mechanosensitive ion channel	10.16	5.00	31334.2	1 (1 0 0 0 0)
YP_001038204 Cthe_1790 ATP:guanido phosphotransferase	10.16	3.20	38662.1	1 (1 0 0 0 0)
YP_001037284 Cthe_0857 protein of unknown function DUF107	10.16	9.80	17807.0	1 (1 0 0 0 0)
YP_001036995 Cthe_0566 primosomal protein N'	10.16	2.90	92407.8	1 (1 0 0 0 0)
YP_001037405 Cthe_0980 hypothetical protein	10.15	6.80	18874.1	1 (1 0 0 0 0)
YP_00103Cthe_0048 copper amine oxidase-like protein	10.15	3.7	30343	1 (1 0 0 0 0)
YP_001036525 Cthe_0091 Peptidoglycan glycosyltransferase	10.15	1.30	79061.7	1 (1 0 0 0 0)
YP_001037762 Cthe_1337 type II secretion system protein	10.15	4.5	32759	1 (1 0 0 0 0)
YP_001038671 Cthe_2276 AAA ATPase, central region	10.15	3.30	43847.1	1 (1 0 0 0 0)
YP_001038320 Cthe_1908 copper amine oxidase-like protein	10.15	2.5	110367	1 (1 0 0 0 0)
YP_001036617 Cthe_0183 protein of unknown function DUF208	10.15	11.60	21341.8	1 (1 0 0 0 0)
YP_001039390 Cthe_3001 protein of unknown function DUF47	10.15	5.30	24092.6	1 (1 0 0 0 0)
YP_001038272 Cthe_1858 peptidase M23B	10.15	3.5	41616	1 (1 0 0 0 0)
YP_001036908 Cthe_0477 flagellar motor switch protein FliM	10.15	5.80	36851.0	1 (1 0 0 0 0)
YP_001037520 Cthe_1095 cell divisionFtsK/SpoIIIE	10.15	0.9	88323	1 (1 0 0 0 0)
YP_001038707 Cthe_2312 hypothetical protein	10.15	6.9	17536	1 (1 0 0 0 0)

YP_001038285 Cthe_1871 Tn7-like transposition protein D	10.15	2.70	73259.0	1 (1 0 0 0 0)
YP_001038227 Cthe_1813 urease accessory protein UreG	10.15	6.90	21968.6	1 (1 0 0 0 0)
YP_001037114 Cthe_0686 tryptophanyl-tRNA synthetase	10.15	4.30	37373.2	1 (1 0 0 0 0)
YP_001039073 Cthe_2681 serine phosphatase	10.15	2.40	87361.5	1 (1 0 0 0 0)
YP_001036870 Cthe_0439 CobB/CobQ-like glutamine amidotransferase	10.15	5.60	27924.9	1 (1 0 0 0 0)
YP_001037158 Cthe_0730 hypothetical protein	10.15	5.30	24840.0	1 (1 0 0 0 0)
YP_001038918 Cthe_2523 hypothetical protein	10.15	7.90	14542.0	1 (1 0 0 0 0)
YP_001037756 Cthe_1331 aspartyl-tRNA synthetase	10.15	1.50	67728.0	1 (1 0 0 0 0)
YP_001037741 Cthe_1316 protein of unknown function DUF370	10.14	15.10	10347.4	1 (1 0 0 0 0)
YP_001039201 Cthe_2809 glycoside hydrolase, family 16	10.14	0.6	147695	1 (1 0 0 0 0)
YP_001038381 Cthe_1970 hypothetical protein	10.14	14.90	11058.4	1 (1 0 0 0 0)
YP_001039325 Cthe_2934 ABC transporter related protein	10.14	2.30	33338.6	1 (1 0 0 0 0)
YP_001037273 Cthe_0845 stage III sporulation protein spoIIIAA	10.14	4.20	37512.9	1 (1 0 0 0 0)
YP_001037846 Cthe_1422 RDD domain containing protein	10.14	3.50	29581.2	1 (1 0 0 0 0)
YP_001036921 Cthe_0490 CheA signal transduction histidine kinases	10.14	1.40	77010.8	1 (1 0 0 0 0)
YP_001038799 Cthe_2404 transcriptional regulator, GntR family	10.14	4.80	26246.0	1 (1 0 0 0 0)
YP_001038599 Cthe_2204 cyanophycin synthetase	10.14	0.80	96427.1	1 (1 0 0 0 0)
YP_001038995 Cthe_2602 ATP synthase F0, A subunit	10.14	5.20	28445.0	1 (1 0 0 0 0)
YP_001039491 Cthe_3103 2-hydroxyglutaryl-CoA dehydratase, D-component	10.14	3.00	37446.2	1 (1 0 0 0 0)
YP_001038603 Cthe_2208 LexA DNA-binding region containing protein	10.14	5.60	22036.7	1 (1 0 0 0 0)
YP_001039159 Cthe_2767 Methyltransferase type 12	10.14	3.00	53542.8	1 (1 0 0 0 0)
YP_001039111 Cthe_2719 NusG antitermination factor	10.14	5.60	20117.5	1 (1 0 0 0 0)
YP_001037448 Cthe_1023 protein of unknown function DUF205	10.14	5.90	23592.1	1 (1 0 0 0 0)
YP_001039024 Cthe_2632 transcription-repair coupling factor	10.14	0.80	134866.6	1 (1 0 0 0 0)
YP_001037162 Cthe_0734 peptidase M23B	10.14	2.4	42399	1 (1 0 0 0 0)
YP_001036984 Cthe_0555 PpiC-type peptidyl-prolyl cis-trans isomerase	10.14	2.2	47318	1 (1 0 0 0 0)
YP_001036843 Cthe_0412 glycoside hydrolase, family 9	10.14	1.2	100558	1 (1 0 0 0 0)
YP_001038128 Cthe_1714 hypothetical protein	10.14	11.00	12935.6	1 (1 0 0 0 0)
YP_001039218 Cthe_2826 hypothetical protein	10.14	2.90	48947.8	1 (1 0 0 0 0)
YP_001039475 Cthe_3086 3-phosphoshikimate 1-carboxyvinyltransferase	10.14	1.90	45886.6	1 (1 0 0 0 0)
YP_001038353 Cthe_1942 hypothetical protein	10.14	4.3	33313	1 (1 0 0 0 0)
YP_001039352 Cthe_2962 oligopeptide/dipeptide	10.14	2.60	38411.3	1 (1 0 0 0 0)

ABC transporter, ATPase subunit				
YP_001036585 Cthe_0151 protein of unknown function DUF965	10.14	14.40	10363.2	1 (1 0 0 0 0)
YP_001039458 Cthe_3069 two component transcriptional regulator, winged helix family	10.14	6.50	26281.8	1 (1 0 0 0 0)
YP_001037657 Cthe_1232 AMP-dependent synthetase and ligase	10.14	2.10	64467.0	1 (1 0 0 0 0)
YP_001037501 Cthe_1076 protein of unknown function DUF881	10.14	5.1	27007	1 (1 0 0 0 0)
YP_001036541 Cthe_0107 Riboflavin synthase	10.14	7.10	16428.7	1 (1 0 0 0 0)
YP_001037232 Cthe_0804 transcriptional regulator, MerR family	10.14	7.90	22515.9	1 (1 0 0 0 0)
YP_001038588 Cthe_2193 Carbohydrate binding family 6	10.14	0.6	103062	1 (1 0 0 0 0)
YP_001039493 Cthe_3105 exsB protein	10.14	4.5	24240	1 (1 0 0 0 0)
YP_001039133 Cthe_2741 ATP-dependent Clp protease, ATP-binding subunit ClpX	10.14	2.60	47861.5	1 (1 0 0 0 0)
YP_001038179 Cthe_1765 hypothetical protein	10.14	2.10	44742.4	1 (1 0 0 0 0)
YP_001038356 Cthe_1945 Thioredoxin-disulfide reductase	10.14	5.70	32476.8	1 (1 0 0 0 0)
YP_001038258 Cthe_1844 transcriptional regulator, BadM/Rrf2 family	10.14	5.90	16944.7	1 (1 0 0 0 0)
Cthe_0850	10.14	4.2	32313	1 (1 0 0 0 0)
YP_001036935 Cthe_0505 formate acetyltransferase	10.14	1.20	84351.2	1 (1 0 0 0 0)
YP_001037566 Cthe_1141 hypothetical protein	10.14	0.60	157008.0	1 (1 0 0 0 0)
YP_001038720 Cthe_2325 Recombinase	10.14	1.70	82190.3	2 (0 1 0 0 1)
YP_001039351 Cthe_2961 extracellular solute-binding protein, family 5	10.14	2.6	60373	1 (1 0 0 0 0)
YP_001036736 Cthe_0305 DNA topoisomerase ATP-hydrolyzing	10.14	1.80	76560.1	1 (1 0 0 0 0)
YP_001038331 Cthe_1919 MgtC/SapB transporter	10.14	8.20	25478.7	1 (1 0 0 0 0)
YP_001038700 Cthe_2305 putative undecaprenol kinase	10.14	4.80	30508.6	1 (1 0 0 0 0)
YP_001039430 Cthe_3041 UbiA prenyltransferase	10.14	4.10	36326.9	1 (1 0 0 0 0)
YP_001037836 Cthe_1412 tryptophan synthase, beta subunit	10.14	2.50	43091.9	1 (1 0 0 0 0)
YP_001039214 Cthe_2822 oxidoreductase-like protein	10.13	4.30	37048.1	1 (1 0 0 0 0)
YP_001038548 Cthe_2152 hypothetical protein	10.13	7.00	20357.9	1 (1 0 0 0 0)
YP_001038380 Cthe_1969 hypothetical protein	10.13	2.80	52057.6	1 (1 0 0 0 0)
YP_001039425 Cthe_3036 methyl-accepting chemotaxis sensory transducer	10.13	2.5	62462	1 (1 0 0 0 0)
YP_001039602 Cthe_3214 hypothetical protein	10.13	1.10	77632.4	1 (1 0 0 0 0)
YP_001037142 Cthe_0714 hydroxymethylbutenyl pyrophosphate reductase	10.13	1.40	77377.3	1 (1 0 0 0 0)
YP_001038815 Cthe_2420 HD superfamily phosphohydrolases-like protein	10.13	1.70	70494.4	1 (1 0 0 0 0)
YP_001036915 Cthe_0484 flagellar biosynthetic protein FlhB	10.13	4.10	44266.0	1 (1 0 0 0 0)
YP_001038Cthe_2254 hypoxanthine phosphoribosyltransferase	10.13	4.90	20691.9	1 (1 0 0 0 0)
YP_001038297 Cthe_1883 SMC protein-like protein	10.13	1.50	79826.3	1 (1 0 0 0 0)
YP_001037842 Cthe_1418 hypothetical protein	10.13	9.60	26459.0	1 (1 0 0 0 0)

YP_001037012 Cthe_0583 response regulator receiver protein	10.13	3.30	43858.3	1 (1 0 0 0 0)
YP_001038268 Cthe_1854 methylated-DNA--protein-cysteine methyltransferase	10.13	9.40	12179.3	1 (1 0 0 0 0)
YP_001039285 Cthe_2894 hypothetical protein	10.13	5.70	20119.8	1 (1 0 0 0 0)
YP_001037435 Cthe_1010 peptidase U32	10.13	5.60	47280.8	1 (1 0 0 0 0)
YP_001039597 Cthe_3209 hypothetical protein	10.13	2.80	46140.9	1 (1 0 0 0 0)
YP_001038818 Cthe_2423 hypothetical protein	10.13	1.0	111319	1 (1 0 0 0 0)
YP_001036550 Cthe_0116 protein of unknown function DUF199	10.13	4.10	35701.0	2 (0 1 0 0 1)
YP_001037704 Cthe_1279 ribosomal protein L28	10.13	16.10	7071.8	1 (1 0 0 0 0)
YP_001037685 Cthe_1260 4-hydroxybenzoyl-CoA thioesterase	10.13	8.00	16215.3	1 (1 0 0 0 0)
YP_001037263 Cthe_0835 hypothetical protein	10.13	10.70	9453.0	1 (1 0 0 0 0)
YP_001036546 Cthe_0112 UDP-N-acetylenolpyruvoylglucosamine reductase	10.13	3.00	33466.5	1 (1 0 0 0 0)
YP_001037587 Cthe_1162 glucosamine--fructose-6-phosphate aminotransferase, isomerizing	10.13	3.80	66987.3	1 (1 0 0 0 0)
YP_001039575 Cthe_3187 putative anti-sigma regulatory factor, serine/threonine protein kinase	10.13	10.00	15737.1	1 (1 0 0 0 0)
YP_001039276 Cthe_2885 phosphoribosylaminoimidazole-succinocarboxamide synthase	10.13	4.80	33151.2	1 (1 0 0 0 0)
YP_001037626 Cthe_1201 purine nucleoside phosphorylase I, inosine and guanosine-specific	10.13	4.00	30421.9	1 (1 0 0 0 0)
YP_001037690 Cthe_1265 phosphoglucomutase/phosphomannomutase alpha/beta/alpha domain I	10.13	2.60	64914.1	1 (1 0 0 0 0)
YP_001039105 Cthe_2713 dihydroxy-acid dehydratase	10.13	2.20	58842.6	1 (1 0 0 0 0)
YP_001037800 Cthe_1375 aspartate kinase	10.13	1.60	49629.5	1 (1 0 0 0 0)
YP_001039084 Cthe_2692 glycosyl transferase, group 1	10.13	4.10	42295.2	1 (1 0 0 0 0)
YP_001039020 Cthe_2628 SpoVG	10.13	7.40	10462.3	1 (1 0 0 0 0)
YP_001038662 Cthe_2267 Sodium-transporting two-sector ATPase	10.13	1.50	65322.7	1 (1 0 0 0 0)
YP_001037267 Cthe_0839 stage III sporulation protein AG	10.13	4.80	22711.1	1 (1 0 0 0 0)
YP_001038530 Cthe_2132 abortive infection protein, putative	10.13	4.20	49977.1	2 (0 1 0 0 1)
YP_001039094 Cthe_2702 polysaccharide pyruvyl transferase	10.13	4.00	84022.9	2 (0 1 0 0 1)
YP_001038208 Cthe_1794 translation elongation factor G	10.13	1.00	76682.5	1 (1 0 0 0 0)
YP_001036798 Cthe_0367 NADP oxidoreductase, coenzyme F420-dependent	10.13	2.80	31848.7	1 (1 0 0 0 0)
YP_001039075 Cthe_2683 hypothetical protein	10.13	5.40	23483.8	1 (1 0 0 0 0)
YP_001037711 Cthe_1286 peptidase S1 and S6, chymotrypsin/Hap	10.13	3.30	54561.9	1 (1 0 0 0 0)
YP_001036563 Cthe_0129 metal dependent phosphohydrolase	10.13	1.40	58142.2	1 (1 0 0 0 0)
YP_001039371 Cthe_2981 hypothetical protein	10.13	5.5	19707	2 (0 0 1 1 0)
YP_001036831 Cthe_0400 hypothetical protein	10.13	2.30	39513.9	1 (1 0 0 0 0)

YP_001037418 Cthe_0993 protein of unknown function DUF448	10.13	14.10	10660.6	1 (1 0 0 0 0)
YP_001036500 Cthe_0066 hypothetical protein	10.13	7.6	18835	1 (1 0 0 0 0)
YP_001036967 Cthe_0538 periplasmic sensor signal transduction histidine kinase	10.13	5.30	38235.3	1 (1 0 0 0 0)
YP_001038869 Cthe_2474 phage terminase, large subunit, PBSX family	10.13	2.20	48589.4	1 (1 0 0 0 0)
YP_001037726 Cthe_1301 protein of unknown function DUF795	10.13	1.90	46426.2	1 (1 0 0 0 0)
YP_001037507 Cthe_1082 hypothetical protein	10.13	8.20	9719.8	1 (1 0 0 0 0)
YP_001037372 Cthe_0945 metallophosphoesterase	10.13	3.00	37881.6	1 (1 0 0 0 0)
YP_001037424 Cthe_0999 1-deoxy-D-xylulose 5-phosphate reductoisomerase	10.13	2.10	41864.3	1 (1 0 0 0 0)
YP_001037184 Cthe_0756 Membrane dipeptidase	10.13	5.50	34538.7	1 (1 0 0 0 0)
YP_001038870 Cthe_2475 phage portal protein, SPP1 family	10.13	2.30	56081.0	1 (1 0 0 0 0)
YP_001037402 Cthe_0977 UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate--D-alanyl-D-alanyl ligase	10.13	2.40	49859.3	1 (1 0 0 0 0)
YP_001038489 Cthe_2089 glycoside hydrolase, family 48	10.13	1.9	83505	1 (1 0 0 0 0)
YP_00103Cthe_0016 Ferritin and Dps	10.13	6.30	18594.2	1 (1 0 0 0 0)
YP_001038455 Cthe_2055 hypothetical protein	10.13	1.70	73451.8	1 (1 0 0 0 0)
YP_001037982 Cthe_1563 ABC transporter related protein	10.13	4.90	24680.9	1 (1 0 0 0 0)
YP_001039433 Cthe_3044 protein of unknown function DUF975	10.13	2.80	28506.0	1 (1 0 0 0 0)
YP_001038366 Cthe_1955 RNA binding S1	10.13	1.80	80671.0	1 (1 0 0 0 0)
YP_001036889 Cthe_0458 exodeoxyribonuclease III Xth	10.13	4.40	29295.5	1 (1 0 0 0 0)
YP_001038653 Cthe_2258 phosphoesterase, RecJ-like protein	10.13	1.3	75381	1 (1 0 0 0 0)
YP_001037323 Cthe_0896 DNA primase	10.13	1.00	68749.2	1 (1 0 0 0 0)
YP_001039021 Cthe_2629 UDP-N-acetylglucosamine pyrophosphorylase	10.13	1.30	51493.2	1 (1 0 0 0 0)
YP_001037548 Cthe_1123 hypothetical protein	10.12	2.90	31835.9	1 (1 0 0 0 0)
YP_001038120 Cthe_1706 hypothetical protein	10.12	1.30	42932.8	1 (1 0 0 0 0)
YP_001037606 Cthe_1181 transglutaminase-like protein	10.12	0.80	87791.0	1 (1 0 0 0 0)
YP_001038349 Cthe_1938 D-alanine--D-alanine ligase	10.12	1.30	41896.3	1 (1 0 0 0 0)
YP_001037148 Cthe_0720 aminotransferase, class V	10.12	1.80	43376.5	1 (1 0 0 0 0)
YP_001038136 Cthe_1722 phage Terminase	10.12	0.60	97269.2	1 (1 0 0 0 0)
YP_001039096 Cthe_2704 Transketolase-like protein	10.12	2.50	31356.9	1 (1 0 0 0 0)
YP_001036520 Cthe_0086 hypothetical protein	10.12	2.40	34040.5	1 (1 0 0 0 0)
YP_001037818 Cthe_1393 multi-sensor signal transduction histidine kinase	10.12	1.70	66259.0	1 (1 0 0 0 0)
YP_001037003 Cthe_0574 serine/threonine protein kinase	10.11	1.00	78419.3	1 (1 0 0 0 0)
YP_001037979 Cthe_1560 Pyridoxal-5'-phosphate-dependent enzyme, beta subunit	10.11	1.70	32798.9	1 (1 0 0 0 0)
YP_001037434 Cthe_1009 O-methyltransferase,	10.11	4.70	24018.9	1 (1 0 0 0 0)

family 3				
YP_001038840 Cthe_2445 Alcohol dehydrogenase GroES-like protein	10.11	3.20	37723.3	1 (1 0 0 0 0)
YP_001039396 Cthe_3007 ErfK/YbiS/YcfS/YnhG	10.11	5.20	25407.9	1 (1 0 0 0 0)
YP_001037938 Cthe_1517 type I phosphodiesterase/nucleotide pyrophosphatase	10.11	1.90	42290.1	1 (1 0 0 0 0)
YP_001036990 Cthe_0561 ApbE-like lipoprotein	10.11	3.4	38212	1 (1 0 0 0 0)
YP_001037278 Cthe_0851 hypothetical protein	10.11	8.00	9941.1	1 (1 0 0 0 0)
YP_001037776 Cthe_1351 acetobutylicum phosphotransbutyrylase	10.11	3.1	17865	1 (1 0 0 0 0)
YP_001037907 Cthe_1485 beta-lactamase-like protein	10.11	1.70	47381.8	1 (1 0 0 0 0)
YP_001036570 Cthe_0136 4'-phosphopantetheinyl transferase	10.11	3.30	28027.6	1 (1 0 0 0 0)
YP_001038956 Cthe_2562 glucose-1-phosphate cytidylyltransferase	10.10	4.20	29905.8	1 (1 0 0 0 0)
YP_001036991 Cthe_0562 protein of unknown function DUF1312	10.10	5.60	14045.4	1 (1 0 0 0 0)
YP_001038831 Cthe_2436 hypothetical protein	10.10	7.70	9058.6	1 (1 0 0 0 0)
YP_001038177 Cthe_1763 ABC transporter related protein	8.16	4.10	27658.6	1 (0 1 0 0 0)
YP_00103Cthe_0003 ankyrin repeat protein	8.16	3.6	35562	1 (0 1 0 0 0)
YP_001038519 Cthe_2119 glycoside hydrolase, family 10	8.16	1.60	85070.1	1 (0 1 0 0 0)
YP_001037981 Cthe_1562 protein of unknown function DUF214	8.16	3.2	44166	1 (0 1 0 0 0)
YP_001038397 Cthe_1986 phage / plasmid primase, P4 family	8.16	2.80	80366.4	1 (0 1 0 0 0)
YP_001037302 Cthe_0875 anthranilate synthase component I	8.16	1.80	55984.8	1 (0 1 0 0 0)
YP_001037133 Cthe_0705 hypothetical protein	8.16	15.40	9275.1	1 (0 1 0 0 0)
YP_001039474 Cthe_3085 response regulator receiver protein	8.15	3.20	35092.4	1 (0 1 0 0 0)
YP_001037714 Cthe_1289 hypothetical protein	8.15	17.20	7786.1	1 (0 1 0 0 0)
YP_001038103 Cthe_1688 Radical SAM	8.15	3.70	36819.0	1 (0 1 0 0 0)
YP_001038904 Cthe_2509 DNA topoisomerase	8.15	3.90	37812.1	1 (0 1 0 0 0)
YP_001038100 Cthe_1685 ABC transporter related protein	8.15	3.30	64734.1	1 (0 1 0 0 0)
YP_001037404 Cthe_0979 Peptidoglycan glycosyltransferase	8.15	1.9	80186	1 (0 1 0 0 0)
YP_001038490 Cthe_2090 hypothetical protein	8.15	5.70	21466.3	1 (0 1 0 0 0)
YP_001038153 Cthe_1739 SNF2-related protein	8.15	4.00	44230.3	1 (0 1 0 0 0)
YP_001037504 Cthe_1079 Nucleotidyl transferase	8.15	1.60	91250.7	1 (0 1 0 0 0)
YP_001038806 Cthe_2411 metallophosphoesterase	8.15	3.40	43678.0	1 (0 1 0 0 0)
YP_001037079 Cthe_0651 hypothetical protein	8.15	9.80	13444.1	1 (0 1 0 0 0)
YP_001036887 Cthe_0456 protein of unknown function UPF0047	8.15	7.20	15799.1	1 (0 1 0 0 0)
YP_001037586 Cthe_1161 hypothetical protein	8.15	1.30	93283.8	1 (0 1 0 0 0)
YP_001039182 Cthe_2790 ABC transporter related protein	8.15	4.50	30112.8	1 (0 1 0 0 0)
YP_001036608 Cthe_0174 sulfatase	8.15	1.60	72949.3	1 (0 1 0 0 0)
YP_001038441 Cthe_2039 UvrD/REP helicase	8.15	0.90	143255.1	1 (0 1 0 0 0)

YP_001038394 Cthe_1983 transposase IS116/IS110/IS902	8.14	3.00	45903.8	1 (0 1 0 0 0)
YP_0010393Cthe_2974 hypothetical protein	8.14	3.50	45212.0	1 (0 1 0 0 0)
YP_001039288 Cthe_2897 transcription elongation factor GreA	8.14	5.70	17965.3	1 (0 1 0 0 0)
YP_001036853 Cthe_0422 CoA-binding protein	8.14	4.80	24791.9	1 (0 1 0 0 0)
YP_001037496 Cthe_1071 PhoH-like protein	8.14	3.60	37547.6	1 (0 1 0 0 0)
YP_001037036 Cthe_0608 peptidase M42	8.14	4.90	38638.7	1 (0 1 0 0 0)
YP_0010387Cthe_2369 ribonuclease P protein component	8.14	6.30	15109.1	1 (0 1 0 0 0)
YP_001037968 Cthe_1548 Enoyl-CoA hydratase/isomerase	8.14	6.50	27621.2	1 (0 1 0 0 0)
YP_001037056 Cthe_0628 hypothetical protein	8.14	3.60	31420.3	1 (0 1 0 0 0)
YP_001038178 Cthe_17Outer membrane protein-like protein	8.14	2.80	41742.3	1 (0 1 0 0 0)
YP_001037913 Cthe_1492 NAD P H dehydrogenase quinone	8.14	2.00	61292.0	1 (0 1 0 0 0)
YP_001036629 Cthe_0195 RNA polymerase, sigma-24 subunit, ECF subfamily	8.14	10.30	11336.3	1 (0 1 0 0 0)
YP_001038367 Cthe_1956 Tagatose-6-phosphate kinase	8.14	4.80	33741.5	1 (0 1 0 0 0)
YP_001039128 Cthe_2736 phosphoenolpyruvate-protein phosphotransferase	8.14	1.80	63495.4	1 (0 1 0 0 0)
YP_001037494 Cthe_1069 protein of unknown function UPF0054	8.14	8.00	19165.6	1 (0 1 0 0 0)
YP_001038979 Cthe_2586 amidohydrolase 2	8.14	2.70	29355.3	1 (0 1 0 0 0)
YP_001038545 Cthe_2149 amine oxidase	8.14	4.00	49732.1	1 (0 1 0 0 0)
YP_001039240 Cthe_2848 phage major capsid protein, HK97	8.14	3.20	48301.6	1 (0 1 0 0 0)
YP_001037374 Cthe_0947 dihydroorotate dehydrogenase family protein	8.14	6.50	32496.2	1 (0 1 0 0 0)
YP_001038616 Cthe_2221 DegT/DnrJ/EryC1/StrS aminotransferase	8.13	2.10	44358.8	1 (0 1 0 0 0)
YP_001036686 Cthe_0254 protein of unknown function DUF58	8.13	2.9	50537	1 (0 1 0 0 0)
YP_001039535 Cthe_3147 ABC transporter related protein	8.13	1.70	64908.9	1 (0 1 0 0 0)
YP_001037053 Cthe_0625 glycoside hydrolase, family 9	8.13	1.4	79760	1 (0 1 0 0 0)
YP_001038183 Cthe_1769 hypothetical protein	8.13	6.20	15065.2	1 (0 1 0 0 0)
YP_001036980 Cthe_0551 AMP-dependent synthetase and ligase	8.13	1.80	63015.9	1 (0 1 0 0 0)
YP_001037985 Cthe_1566 Nitrogenase	8.13	3.80	49392.4	1 (0 1 0 0 0)
YP_0010382Cthe_1850 hypothetical protein	8.13	6.7	17669	1 (0 1 0 0 0)
YP_001037668 Cthe_1243 GCN5-related N-acetyltransferase	8.13	7.20	17625.2	1 (0 1 0 0 0)
YP_001038171 Cthe_1757 peptidase M23B	8.13	3.20	34708.1	1 (0 1 0 0 0)
YP_001039513 Cthe_3125 heat shock protein Hsp20	8.13	10.70	17300.8	1 (0 1 0 0 0)
YP_00103Cthe_0024 biotin biosynthesis protein BioC	8.13	3.90	32622.6	1 (0 1 0 0 0)
YP_00103Cthe_0040 Cellulase., Cellulose 1,4-beta-cellobiosidase	8.13	1.4	98470	1 (0 1 0 0 0)
YP_001039205 Cthe_2813 two component transcriptional regulator, winged helix family	8.13	5.60	26201.9	1 (0 1 0 0 0)

YP_001037421 Cthe_0996 DNA polymerase III, alpha subunit	8.13	0.70	163280.5	2 (0 0 0 2 0)
YP_001036895 Cthe_04flagellar hook-basal body complex subunit FliE	8.13	9.60	11494.9	1 (0 1 0 0 0)
YP_001039050 Cthe_2658 RNA-binding S4	8.13	10.10	8950.9	1 (0 1 0 0 0)
YP_001036503 Cthe_0069 Aspartate--ammonia ligase	8.13	3.80	39008.1	1 (0 1 0 0 0)
YP_001039144 Cthe_2752 beta-lactamase-like protein	8.13	2.30	48437.3	1 (0 1 0 0 0)
YP_001036834 Cthe_0403 DNA-directed RNA polymerase sigma factor	8.13	4.00	32431.1	1 (0 1 0 0 0)
YP_001039553 Cthe_3165 PpiC-type peptidyl-prolyl cis-trans isomerase	8.13	2.60	44023.3	1 (0 1 0 0 0)
YP_001036838 Cthe_0407 Radical SAM	8.13	2.50	49617.1	1 (0 1 0 0 0)
YP_001037369 Cthe_0942 MiaB-like tRNA modifying enzyme YliG	8.13	1.80	51752.8	1 (0 1 0 0 0)
YP_001038462 Cthe_2062 GerA spore germination protein	8.13	2.70	56554.5	1 (0 1 0 0 0)
YP_001038932 Cthe_2537 sulfate adenylyltransferase, large subunit	8.13	4.00	67476.3	1 (0 1 0 0 0)
YP_001038944 Cthe_2550 glycosyltransferase sugar-binding region containing DXD motif	8.13	3.70	38319.5	1 (0 1 0 0 0)
YP_001038731 Cthe_2336 glycosyl transferase, group 1	8.13	2.10	43607.8	1 (0 1 0 0 0)
YP_001037710 Cthe_1285 metal dependent phosphohydrolase	8.13	2.6	53393	1 (0 1 0 0 0)
YP_001038619 Cthe_2224 putative methionyl-tRNA formyltransferase	8.13	6.10	26759.9	1 (0 1 0 0 0)
YP_001036711 Cthe_0279 hypothetical protein	8.13	23.80	6918.7	1 (0 1 0 0 0)
YP_001038885 Cthe_2490 hypothetical protein	8.13	4.50	36296.3	1 (0 1 0 0 0)
YP_001038081 Cthe_1665 Hsp33 protein	8.13	3.20	31930.1	1 (0 1 0 0 0)
YP_001037257 Cthe_0829 hypothetical protein	8.13	4.50	26629.5	1 (0 1 0 0 0)
YP_001038334 Cthe_1922 hypothetical protein	8.13	3.40	36105.4	1 (0 1 0 0 0)
YP_001039481 Cthe_3093 Adenylosuccinate synthase	8.13	1.90	47011.3	1 (0 1 0 0 0)
YP_001039131 Cthe_2739 trigger factor	8.13	2.30	48813.2	1 (0 1 0 0 0)
YP_001038584 Cthe_2189 diguanylate cyclase with GAF sensor	8.13	3.4	47300	1 (0 1 0 0 0)
YP_001039069 Cthe_2677 histidine kinase	8.13	4.60	26818.0	1 (0 1 0 0 0)
YP_001037239 Cthe_0811 response regulator receiver protein	8.13	7.70	13041.8	1 (0 1 0 0 0)
YP_001038974 Cthe_2581 dihydropteroate synthase	8.13	3.60	43330.5	1 (0 1 0 0 0)
YP_001038769 Cthe_2374 DNA replication and repair protein RecF	8.13	3.00	43059.5	1 (0 1 0 0 0)
YP_001036794 Cthe_0363 aminotransferase, class I and II	8.13	2.60	43954.2	1 (0 1 0 0 0)
YP_001038576 Cthe_2181 predicted ATPase of the PP-loop superfamily implicated in cell cycle control	8.13	6.30	27011.6	1 (0 1 0 0 0)
YP_001039401 Cthe_3012 Carbohydrate binding family 6	8.13	1.9	70341	1 (0 1 0 0 0)
YP_001037809 Cthe_1384 FolC bifunctional protein	8.13	2.80	47865.1	1 (0 1 0 0 0)
YP_001037617 Cthe_1192 hypothetical protein	8.13	2.6	38976	1 (0 1 0 0 0)
YP_001036986 Cthe_0557 polysaccharide	8.13	3.9	29640	1 (0 1 0 0 0)

deacetylase				
YP_001037259 Cthe_0831 Polyprenyl synthetase	8.13	3.40	32474.7	1 (0 1 0 0 0)
YP_001038569 Cthe_2174 transcription termination factor Rho	8.13	1.40	72542.7	1 (0 1 0 0 0)
YP_001037076 Cthe_0glutamyl-tRNA synthetase	8.13	1.50	63361.2	1 (0 1 0 0 0)
YP_001037411 Cthe_0986 peptidase M16-like protein	8.12	1.40	49371.6	1 (0 1 0 0 0)
YP_00103Cthe_0051 hypothetical protein	8.12	1.4	47511	1 (0 1 0 0 0)
YP_001037253 Cthe_0825 glycoside hydrolase, family 9	8.12	0.9	72370	1 (0 1 0 0 0)
YP_001038712 Cthe_2317 response regulator receiver protein	8.11	1.30	54371.1	1 (0 1 0 0 0)
YP_001038981 Cthe_2588 putative transcriptional acitvator, Baf family	8.11	2.40	27979.9	1 (0 1 0 0 0)
YP_001036723 Cthe_0291 protein of unknown function DUF255	8.11	0.70	78484.9	1 (0 1 0 0 0)
YP_001037996 Cthe_1577 hypothetical protein	8.11	1.40	58166.9	1 (0 1 0 0 0)
YP_001039339 Cthe_2948 Allergen V5/Tpx-1 related	8.11	2.5	30063	1 (0 1 0 0 0)
YP_001036628 Cthe_0194 hypothetical protein	8.11	6.3	14790	1 (0 1 0 0 0)
YP_001036932 Cthe_0502 hypothetical protein	8.11	2.80	31889.6	1 (0 1 0 0 0)
YP_001038470 Cthe_2070 hypothetical protein	8.11	2.10	32039.1	1 (0 1 0 0 0)
YP_001037123 Cthe_0695 putative agmatinase	8.11	1.70	32119.2	1 (0 1 0 0 0)
YP_001039495 Cthe_3107 Radical SAM	8.11	2.90	24049.2	1 (0 1 0 0 0)
YP_001038471 Cthe_2071 hypothetical protein	8.10	2.60	35949.4	1 (0 1 0 0 0)
YP_001038726 Cthe_2331 hypothetical protein	8.10	3.4	19873	1 (0 1 0 0 0)
YP_001036Cthe_0215 phenylalanyl-tRNA synthetase, beta subunit	6.16	1.50	88432.1	1 (0 0 1 0 0)
YP_001038951 Cthe_2557 glycosyl transferase, family 2	6.15	2.40	57455.0	1 (0 0 1 0 0)
YP_001036827 Cthe_0396 ABC transporter related protein	6.15	2.30	63672.7	1 (0 0 1 0 0)
YP_001037542 Cthe_1117 Tn7-like transposition protein A	6.15	3.20	33337.3	1 (0 0 1 0 0)
YP_001039560 Cthe_3172 protein of unknown function DUF214	6.15	2.3	51635	1 (0 0 1 0 0)
YP_001036979 Cthe_0550 heat shock protein Hsp90	6.14	1.70	73792.5	1 (0 0 1 0 0)
YP_001036993 Cthe_05Trans-hexaprenyltranstransferase	6.14	3.10	35809.4	1 (0 0 1 0 0)
YP_001038290 Cthe_1876 Tn7-like transposition protein A	6.14	3.90	32890.9	1 (0 0 1 0 0)
YP_001038822 Cthe_2427 protein of unknown function DUF1385	6.14	3.10	35839.4	1 (0 0 1 0 0)
YP_001036951 Cthe_0521 helicase-like protein	6.14	1.20	126703.7	2 (0 0 0 1 1)
YP_001037722 Cthe_1297 Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	6.14	1.70	37328.4	1 (0 0 1 0 0)
YP_001038865 Cthe_2470 DNA adenine methylase	6.14	2.90	32557.9	1 (0 0 1 0 0)
YP_001037326 Cthe_0899 hypothetical protein	6.14	8.90	10396.7	1 (0 0 1 0 0)
YP_001036517 Cthe_0083 serine phosphatase	6.14	1.6	64513	1 (0 0 1 0 0)
YP_001037748 Cthe_1323 GrpE protein	6.14	5.30	25432.4	1 (0 0 1 0 0)
YP_001036917 Cthe_0486 GTP-binding signal recognition particle SRP54, G-domain	6.14	3.20	45392.9	1 (0 0 1 0 0)
YP_001039078 Cthe_2686 type IV pilus assembly	6.14	3.00	41419.1	1 (0 0 1 0 0)

protein PilM				
YP_001037595 Cthe_1170 Dihydrodipicolinate reductase	6.13	2.40	32268.4	1 (0 0 1 0 0)
YP_001037771 Cthe_1346 single-stranded-DNA-specific exonuclease RecJ	6.13	1.00	93542.3	1 (0 0 1 0 0)
YP_001036822 Cthe_0391 ABC transporter related protein	6.13	2.20	54772.9	1 (0 0 1 0 0)
YP_001039019 Cthe_2627 purine operon repressor, PurR	6.13	4.40	30003.6	1 (0 0 1 0 0)
YP_001037217 Cthe_0789 RNA-binding S4	6.13	3.80	29684.9	1 (0 0 1 0 0)
YP_001038240 Cthe_1826 response regulator receiver sensor signal transduction histidine kinase	6.13	3.00	41724.5	1 (0 0 1 0 0)
YP_001038986 Cthe_2593 peptide chain release factor 1	6.13	3.30	40998.9	1 (0 0 1 0 0)
YP_001037282 Cthe_0855 HAD superfamily subfamily IIIA phosphatase, TIGR01668	6.13	7.80	19255.3	1 (0 0 1 0 0)
YP_001039147 Cthe_2755 ADP-ribosylation/Crystallin J1	6.13	2.90	30361.2	1 (0 0 1 0 0)
YP_001037106 Cthe_0678 pyrimidine-nucleoside phosphorylase	6.13	2.50	46613.0	1 (0 0 1 0 0)
YP_001036615 Cthe_0181 Holliday junction DNA helicase RuvA	6.13	5.00	22143.8	1 (0 0 1 0 0)
YP_001037735 Cthe_1310 Accessory gene regulator B	6.13	5.50	22258.2	1 (0 0 1 0 0)
YP_001037773 Cthe_1348 ribosomal protein S21	6.13	22.40	6901.8	1 (0 0 1 0 0)
YP_001039550 Cthe_3162 SAM dependent methyltransferase	6.13	1.30	44407.1	1 (0 0 1 0 0)
YP_001038735 Cthe_2340 UDP-glucose/GDP-mannose dehydrogenase	6.13	3.20	49160.6	1 (0 0 1 0 0)
Cthe_2016	6.13	4.6	24310	1 (0 0 1 0 0)
YP_001038144 Cthe_1730 hypothetical protein	6.13	8.50	16451.4	1 (0 0 1 0 0)
YP_001036635 Cthe_0201 glutamate synthase, alpha subunit-like protein	6.11	2.50	26583.2	1 (0 0 1 0 0)
YP_001037884 Cthe_1462 transcriptional regulator, XRE family	6.10	7.70	12043.3	1 (0 0 1 0 0)
Cthe_0224	6.10	9.0	7807	1 (0 0 1 0 0)
YP_001038149 Cthe_1735 phage DNA polymerase	4.15	2.00	66028.4	1 (0 0 0 1 0)
YP_001039552 Cthe_31two component transcriptional regulator, AraC family	4.15	2.00	62914.2	1 (0 0 0 1 0)
YP_001037160 Cthe_0732 Chorismate synthase	4.14	2.50	43128.4	1 (0 0 0 1 0)
YP_001037276 Cthe_0848 peptidase M24	4.14	3.10	40148.9	2 (0 0 0 0 2)
YP_001037308 Cthe_0881 diguanylate cyclase	4.14	3.20	43520.9	1 (0 0 0 1 0)
YP_001038486 Cthe_2086 peptidase U32	4.14	1.50	94067.9	1 (0 0 0 1 0)
YP_001037689 Cthe_12DNA polymerase III, alpha subunit	4.14	1.00	134313.3	1 (0 0 0 1 0)
YP_001038250 Cthe_1836 hypothetical protein	4.13	4.0	37045	1 (0 0 0 1 0)
YP_001036737 Cthe_0306 DNA gyrase/topoisomerase IV, subunit A	4.13	1.20	83209.6	1 (0 0 0 1 0)
YP_00103Cthe_0049 transcriptional regulator, XRE family	4.13	3.60	25229.8	1 (0 0 0 1 0)
YP_001038668 Cthe_2273 hypothetical protein	4.13	1.90	42576.3	1 (0 0 0 1 0)
YP_001036566 Cthe_0132 3-oxoacyl- acyl-carrier-protein synthase	4.13	3.20	38518.9	1 (0 0 0 1 0)

YP_001038929 Cthe_2534 sulfate ABC transporter, ATPase subunit	4.13	1.70	39932.3	1 (0 0 0 1 0)
YP_001037703 Cthe_1278 ATP-dependent DNA helicase RecG	4.13	1.70	78599.6	1 (0 0 0 1 0)
YP_001039531 Cthe_3143 hypothetical protein	4.13	8.60	13669.8	1 (0 0 0 1 0)
YP_001036590 Cthe_0156 Radical SAM	4.13	0.80	70482.8	1 (0 0 0 1 0)
YP_001038181 Cthe_1767 hypothetical protein	4.13	3.60	35361.9	1 (0 0 0 1 0)
YP_001039624 Cthe_3236 hypothetical protein	4.13	9.90	11286.6	1 (0 0 0 1 0)
YP_001037737 Cthe_1312 glycyl-tRNA synthetase	4.12	1.10	54044.8	1 (0 0 0 1 0)
YP_001036902 Cthe_0471 flagellar hook-length control protein	4.12	1.10	58298.2	1 (0 0 0 1 0)
YP_001038465 Cthe_2065 cysteinyl-tRNA synthetase	4.12	1.10	53974.6	1 (0 0 0 1 0)
YP_001039212 Cthe_2820 MCP methyltransferase, CheR-type	4.11	2.50	32085.4	1 (0 0 0 1 0)
YP_001037816 Cthe_1391 2-isopropylmalate synthase	4.11	0.80	56422.7	1 (0 0 0 1 0)
YP_001038624 Cthe_2229 N-acetylneuraminate synthase	2.15	2.90	39121.8	1 (0 0 0 0 1)
YP_001038980 Cthe_2587 hypothetical protein	2.15	6.10	17817.9	1 (0 0 0 0 1)
YP_001037204 Cthe_0776 DNA mismatch repair protein MutL	2.14	1.60	85577.2	1 (0 0 0 0 1)
YP_001037931 Cthe_1510 short-chain dehydrogenase/reductase SDR	2.14	4.00	27380.9	1 (0 0 0 0 1)
YP_001037940 Cthe_1519 putative RNA methylase	2.13	2.40	46143.5	1 (0 0 0 0 1)
YP_001037313 Cthe_0886 DNA polymerase I	2.13	1.30	102096.7	1 (0 0 0 0 1)
YP_001039480 Cthe_3092 conserved hypothetical protein, CF-27 family	2.13	11.80	7510.6	1 (0 0 0 0 1)
YP_001039260 Cthe_2869 hypothetical protein	2.13	10.30	11539.0	1 (0 0 0 0 1)
YP_001036533 Cthe_0099 5,10-methylenetetrahydrofolate reductase	2.13	4.10	32764.9	1 (0 0 0 0 1)
YP_001038249 Cthe_1835 Viral A-type inclusion protein repeat containing protein	2.13	1.0	125254	1 (0 0 0 0 1)
YP_001039470 Cthe_3081 transglutaminase-like protein	2.13	4.3	31840	1 (0 0 0 0 1)
YP_001036624 Cthe_0190 proteinase inhibitor I4, serpin	2.13	1.5	67717	1 (0 0 0 0 1)
YP_001037592 Cthe_1167 Radical SAM	2.13	3.70	33187.9	1 (0 0 0 0 1)
YP_001039589 Cthe_3201 CRISPR-associated protein, Csh1 family	2.12	1.00	71508.3	1 (0 0 0 0 1)
YP_001039445 Cthe_3056 transposase, IS204/IS1001/IS1096/IS1165	2.12	1.70	55648.1	1 (0 0 0 0 1)
YP_001037186 Cthe_0758 protein of unknown function UPF0102	2.12	3.80	15146.0	1 (0 0 0 0 1)
YP_001036639 Cthe_0205 hypothetical protein	2.12	1.10	75783.3	1 (0 0 0 0 1)
YP_001037256 Cthe_0828 deoxyxylulose-5-phosphate synthase	2.12	1.10	68873.5	1 (0 0 0 0 1)
YP_001036Cthe_0210 transcriptional regulator, LacI family	2.11	1.40	39414.3	1 (0 0 0 0 1)

[A] XC, cross correlation score. Xc score as determined by Bioworks 3.3 program. The Xc score represents the cross correlation value between the observed peptide fragment mass spectrum and the one theoretically predicted. [b] Total peptide hits observed for the proteins (columns within parethesis represent the number of primary, secondary, tertiary,etc. peptide hits based spectral quality of the peptides spectra).