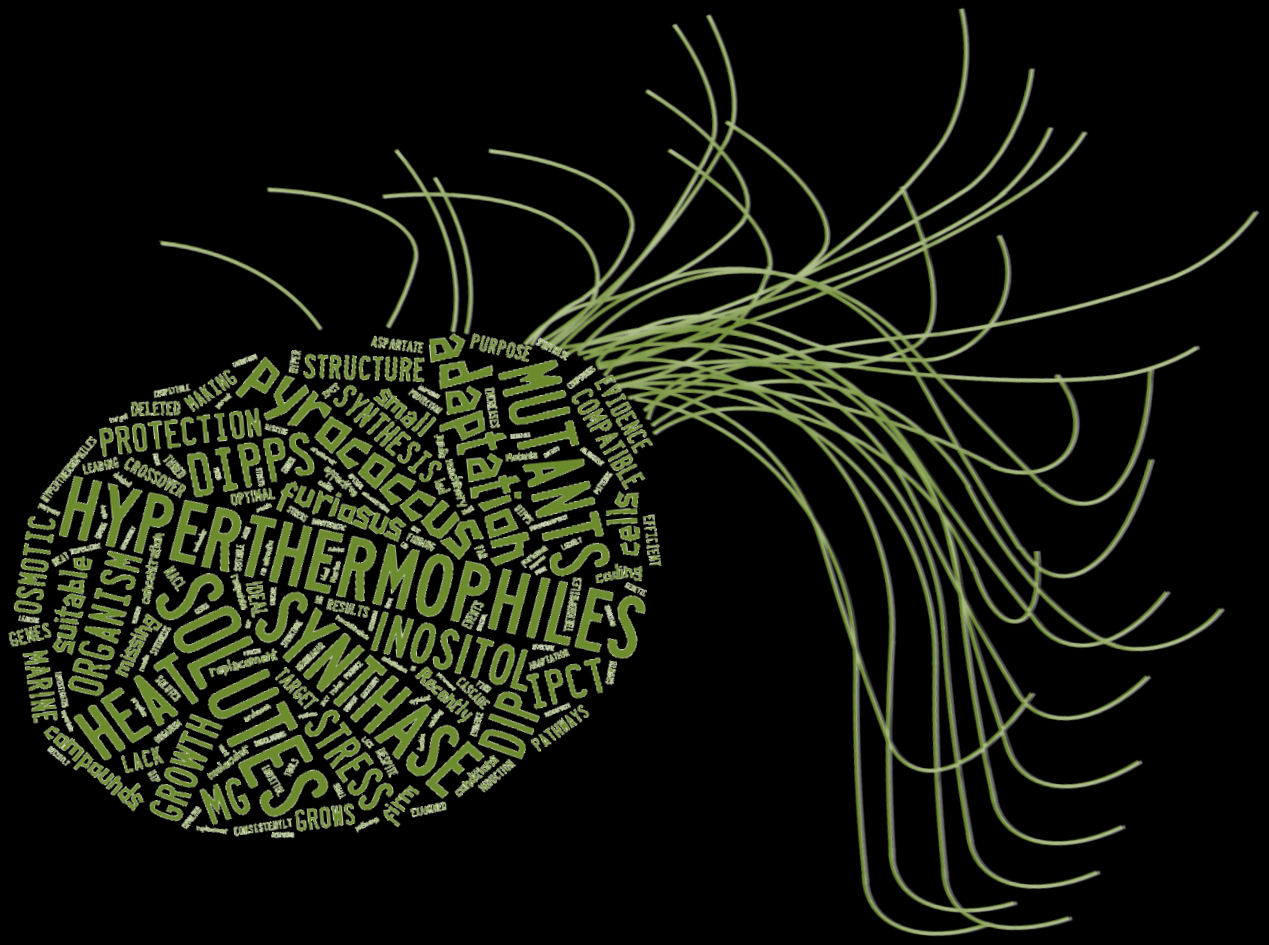
Insights into the biochemical strategies of adaptation to heat stress in the hyperthermophilic archaeon *Pyrococcus furiosus*

Ana Maria da Silva Esteves



Dissertation presented to obtain the Ph.D. degree in Biochemistry

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,
February, 2015

INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
ANTÓNIO XAVIER /UNL

Knowledge Creation

Insights into the biochemical strategies of adaptation to heat stress in the hyperthermophilic archaeon *Pyrococcus furiosus*

Ana Maria da Silva Esteves

Supervisor: Professora Helena Santos
Co-supervisor: Dr. Nuno Borges

Dissertation presented to obtain the Ph.D degree in Biochemistry
Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras, February, 2015

From left to right: Prof. Volker Müller, Dr. Emmanouil Matzapetakis, Dr. Nuno Borges (Co-supervisor), Ana M. Esteves, Prof. Helena Santos (Supervisor), Prof. Pedro Moradas Ferreira, Prof. Beate Averhoff, Prof. Hermínia de Lencastre. Oeiras, 9th of February 2015.

Apoio financeiro da Fundação para a Ciência e Tecnologia e do FSE no âmbito do Quadro Comunitário de apoio, Bolsa de Doutoramento com a referência SFRH/BD/61742/2009.

In loving memory of my grandmother Elisa

Acknowledgements

First, I would like to thank Prof. Helena Santos, my supervisor, for accepting me as a Ph.D. student in her laboratory. Also, I want to express my appreciation to Prof. Helena Santos for being an excellent team leader, and for her constant effort to put together the resources and expertise her students need to carry out their projects with success. Her guidance, rigor, patience and enthusiasm for science definitely contributed to my development as a young scientist.

I want to thank Dr. Nuno Borges, my co-supervisor, for all the support, patience, positive criticism, and also for his friendship. I underline the precious help I received from Dr. Nuno Borges, always encouraging me to work more accurately. I also appreciate his “engineering” qualities, which were essential for setting-up the experiments reported in this dissertation.

I thank Prof. Michael Adams, Dr. Sanjeev Chandrayan, and Patrick McTernan for the great collaboration and for making so pleasant and productive my stay at the Adams Lab.

To Dr. Pedro Lamosa, Dr. Luís G. Gonçalves, Dr. Carla Jorge, Dr. Marta Rodrigues, Dr. Tiago Pais, and Dr. Ana Lúcia Carvalho for valuable advice and helpful discussions. To Dr. Gonçalo Graça for helping me with the NMR experiments and data analysis. I also want to thank all of them for their team spirit, and especially for making me feel welcome in the laboratory.

To Inês Torcato for the great time we spent working together and for the help she gave me in the last year. I must not forget to acknowledge her dedication to work.

To Dr. Paula Fareleira for scientific advice and for her interest in anaerobic organisms.

To Marta Conchinha and Teresa Baptista for technical support.

To Dušica Radoš, Sara Rebelo, Dário Neves, Cristiana Faria, Joana Sousa, and other colleagues of the Cell Physiology & NMR Lab for creating such a great working atmosphere.

To former members of the group, Ana Mingote and Sónia Neto, for their friendship, to Dr. Luís Fonseca for scientific advice, to Pedro Quintas, Andreia Cepeda, Márcia Oliveira, and Patrícia Almeida for their support.

To Fundação para a Ciência e a Tecnologia for the financial support provided by the Ph.D. grant, and to Instituto de Tecnologia Química e Biológica for providing conditions to pursue scientific excellence.

Aos meus amigos Gi e Vasco que partilharam muitos dos bons e menos bons momentos deste percurso.

Por último, mas não menos importante, agradeço a toda a minha (enorme) família pelo apoio incondicional ao longo destes últimos anos. Agradeço em particular aos meus pais por todo o carinho e dedicação, ao Rui por

estar presente na minha vida e à minha avó Elisa, por ter sido um bom exemplo de força e sobretudo por sempre me ter feito sentir única.

Abstract

Organisms that thrive optimally at temperatures above 80°C are called hyperthermophiles. These prokaryotes have been isolated from a variety of hot environments, such as marine geothermal areas, hence they are usually slightly halophilic. Like other halophiles, marine hyperthermophiles have to cope with fluctuations in the salinity of the external medium and generally use low-molecular mass organic compounds to adjust cell turgor pressure. These compounds can accumulate to high levels without interfering with cell metabolism, thereby deserving the designation of compatible solutes. Curiously, the accumulation of compatible solutes also occurs in response to supraoptimal temperatures. Moreover, (hyper)thermophiles display a clear preference for negatively charged solutes, such as diglycerol phosphate, glycerol-phospho-inositol, mannosyl-di-*myo*-inositol-phosphate, di-*myo*-inositol-phosphate (DIP) and mannosylglycerate (MG). The two latter solutes are the most common amongst (hyper)thermophiles. The distribution of DIP is restricted to organisms with optimal growth temperature above 60°C and it accumulates primarily at supraoptimal temperatures, thus a thermoprotective role has often been proposed.

One of the goals of this thesis was to clarify the physiological role of ionic compatible solutes during adaptation of hyperthermophiles to heat stress. In the last decade, many studies demonstrated the efficacy of negatively charged solutes for the stabilization of model enzymes and proteins. However, the role of these solutes *in vivo* has yet to be

demonstrated. An earlier attempt to determine the role of compatible solutes used a mutant strain of *Thermococcus kodakarensis* unable to synthesize DIP, but this study was not fully conclusive since the absence of DIP was compensated by the accumulation of aspartate. Recently, genetic tools became available for the manipulation of the hyperthermophilic archaeon *Pyrococcus furiosus*. The patterns of solute accumulation in *P. furiosus* DSM 3638 under stressful conditions have been investigated, revealing the preferential accumulation of DIP and mannosylglycerate (MG) under heat and osmotic stress, respectively. Therefore, *P. furiosus* seems a promising target organism as it does not accumulate amino acids from the external medium during stress adaptation. Herein, a variant of the type strain *P. furiosus* DSM 3638, which is naturally competent, was used as parent strain and characterized with regard to the growth profile and patterns of solute accumulation. We verified that the profiles of solute accumulation by variant and type strains were similar. The ability of the variant strain to produce either MG or DIP was disrupted by deleting the genes coding for the respective synthases. To assess the effect of the missing solutes, the growth profile and solute pattern of each mutant strain were analyzed under optimal growth conditions, heat stress, and osmotic stress and compared with the parent strain. The results showed that the absence of DIP was offset by MG accumulation at supraoptimal temperatures with similar efficacy. However, the substitution of MG by DIP and aspartate resulted in a less efficient growth at supraoptimal NaCl concentration. This study demonstrates that

MG and DIP have interchangeable roles in thermoprotection, while MG is better suited for osmoprotection. This is one more illustration of solute replacement, a strategy generally used by organisms that rely on solute accumulation for osmoregulation.

The patterns of compatible solute accumulation by *Pyrococcus furiosus* in response to heat stress are known, but the effects at the level of the global metabolome have not been investigated. To gain insight into the metabolite changes caused by heat stress, the type strain of *Pyrococcus furiosus* was grown under optimal growth conditions (90ºC) until mid-exponential phase; subsequently, the cells were subjected to a thermal shock at 97°C for approximately 45 minutes. The internal metabolite pools before and after the heat challenge were analyzed by nuclear magnetic resonance (NMR) and mass spectrometry (MS). As these analytical techniques have very different sensitivity limits, sampling, quenching and extraction procedures were adapted to each technique. ^{1}H-NMR analysis allowed the assignment of 35 metabolites, and quantification of 27 of them. A large proportion (37%) of the metabolites identified were amino acids, but organic acid, nucleotides, cofactors, polyamines, sugar nucleotides, compatible solutes, and sugars were also observed. On the other hand, the targeted analysis of central metabolites by mass spectrometry allowed the identification of 15 compounds. Among these, 9 were quantified, including 3 nucleotides (ATP, ADP, AMP), a combination of 2 trioses (2-phosphoglycerate + 3-phosphoglycerate), GDP-mannose, glycerol-3-phosphate, phosphoenolpyruvate, ribose-1-phosphate, and trehalose-6-

phosphate. Integration of the results obtained using the two techniques supported the conclusion that supraoptimal temperatures favor the synthesis of compatible solutes and N-acetylated NDP-sugars that may be channeled to the synthesis of the cell wall. In contrast, the pool sizes of metabolites associated with the glycolysis or the Krebs cycle were negatively affected. To our knowledge, this is the first study on the metabolome of a hyperthermophilic organism under heat stress conditions. Further work must be done to improve the statistical significance of the results, extend the number of metabolites identified and validate the preliminary conclusions.

An additional challenge of this thesis was the elucidation of the crystal structure and catalytic mechanism of the key enzyme in DIP synthesis, i.e., the bifunctional protein IPCT/DIPPS. The cytosolic domain (IPCT) shows high similarity to the nucleotidyltransferase family, which comprise enzymes that catalyze the transfer of nucleotides onto phospho-sugars. The second domain (DIPPS) possesses six transmembrane segments and catalyses the transfer of the inositol group from CDP-inositol to inositol-1-phosphate. The measurement of the enzyme activity demanded the synthesis and purification of these substrates. The C-terminal domain contains a conserved sequence pattern $DG(x)_2AR(x)_8G(x)_3D(x)_3D$ typical of the CDP-alcohol phosphatidyltransferase family, comprising integral membrane proteins found in all Domains of Life, mostly involved in the synthesis of phospholipids. Despite the large number of proteins belonging to this family, no representative structure was available at the time this

work plan started. Therefore, the resolution of the structure of the membranar DIPPS domain was deemed highly relevant and challenging. The structure of the recombinant bifunctional protein IPCT/DIPPS of *Archaeoglobus fulgidus* with and without bound CDP-inositol was solved by X-ray crystallography to a resolution of 2.5 and 2.65 Å, respectively. The DIPPS domain comprises six transmembrane (TM) α-helices and one α-helix parallel to the membrane plane, and forms a dimer via its transmembrane helices. The structure with bound CDP-inositol revealed the location of the active site, which is a hydrophilic cavity widely open to the cytoplasm with a magnesium ion surrounded by four highly conserved aspartate residues from helices TM2 and TM3. Site-directed mutagenesis of conserved amino acid residues from the family consensus sequence or residues considered relevant for substrate binding validated the results obtained with bound CDP-inositol and allowed us to propose the catalytic mechanism for the CDP-alcohol phosphatidyltransferases.

In summary, the results presented in this thesis represent a significant advance in our knowledge of the strategies of adaptation to heat stress in hyperthermophilic archaea. The role of DIP and MG in the protection against stressful conditions was established. Moreover, the analysis of the heat stress metabolome of a model hyperthermophilic archaeon showed that the metabolism is mainly diverted to the synthesis of compatible solutes (DIP and MG), and probably cell wall components. The elucidation of the three-dimensional structure and the catalytic mechanism of the DIPPS represent breakthroughs for the construction of models of the

homologous phosphatidylinositol-phosphate synthases. This enzyme is essential for the viability of *Mycobacterium tuberculosis*, hence the structure determination of the DIPPS homologue is a crucial step towards the design of novel anti-tuberculosis drugs.

Resumo

Os microrganismos adaptados a proliferar optimamente em ambientes com temperaturas superiores a 80°C são designados hipertermófilos. Estes organismos têm sido isolados a partir de diversos ambientes aquecidos, tais como zonas geotermais marinhas. Por terem sido isolados de habitats marinhos, estes microrganismos costumam ser também ligeiramente halófilos. Assim, tal como os halófilos, os organismos hipertermófilos marinhos sobrevivem a variações externas de sal através da acumulação intracelular de compostos orgânicos de baixa massa molecular que, por não interferirem com o normal funcionamento celular, são denominados solutos compatíveis. Curiosamente, alguns destes solutos são preferencialmente acumulados em resposta a temperaturas supra-óptimas. Além disso, os hipertermófilos têm uma clara preferência pela acumulação de solutos com carga eléctrica negativa tais como diglicerol fosfato, glicero-fosfo-inositol, manosil-di-*mio*-inositol-fosfato, di-*mio*-inositol-fosfato (DIP) e manosilglicerato (MG), sendo os dois últimos solutos os mais comuns. É importante realçar que o DIP nunca foi encontrado em organismos que proliferam optimamente em temperaturas inferiores a 60°C. Para além disso, este soluto responde primariamente em condições de choque térmico e, por isso, tem-lhe sido atribuído um papel em termoproteção celular.

Um dos objectivos desta tese foi elucidar a importância fisiológica dos solutos compatíveis iónicos na adaptação dos hipertermófilos a condições de stresse térmico e salino. Durante a última década, vários

estudos demonstraram a eficácia dos solutos compatíveis com carga eléctrica negativa na estabilização de enzimas e proteínas modelo. Contudo, a escassez de ferramentas genéticas que permitam manipular hipertermófilos tem adiado a validação fisiológica destes resultados. Numa tentativa prévia para determinar o papel dos solutos compatíveis, foi usada uma estirpe mutada de *Thermococcus kodakarensis* incapaz de sintetizar DIP; no entanto, em situação de choque térmico a ausência de DIP foi colmatada pela acumulação de aspartato. O recente desenvolvimento de ferramentas genéticas para manipular o hipertermófilo *Pyrococcus furiosus* possibilitou a realização deste trabalho. A acumulação de solutos pela estirpe *P. furiosus* DSM 3638 em condições de choque térmico e osmótico foi caracterizada, revelando a acumulação preferencial de DIP e MG, respectivamente. Assim, *P. furiosus* parece ser um candidato promissor para este estudo, uma vez que não acumula aminoácidos do meio externo. Neste trabalho, a estirpe usada como estirpe parental derivou da estirpe selvagem *P. furiosus* DSM 3638 e é naturalmente competente. A estirpe foi caracterizada em relação ao perfil crescimento e de acumulação de solutos e, tal como na estirpe selvagem, o DIP respondeu primariamente ao choque térmico, enquanto o soluto MG acumulou em condições de elevada salinidade. As estirpes construídas foram manipuladas na sua capacidade de produzir MG ou DIP, através da deleção dos genes que codificam as sintases em cada uma das vias de síntese. Posteriormente, as estirpes deficientes na síntese de MG ou DIP foram caracterizadas em condições de crescimento óptimo, choque

térmico e choque salino e comparadas com a estirpe parental. Em condições de temperatura supra-óptima, a ausência de DIP foi colmatada pelo MG e o crescimento não foi afectado. No entanto, em condições de salinidade elevada a substituição de MG por DIP e aspartato não foi suficiente para manter o crescimento. Concluindo, este trabalho demonstrou a capacidade superior do MG em osmoproteção enquanto DIP e MG têm papéis semelhantes na termoprotecção.

Presentemente, a resposta ao choque térmico em hipertermófilos está pouco caracterizada. No caso do arqueão *Pyrococcus furiosus*, para além da informação recolhida sobre o perfil de acumulação de solutos compatíveis, existe apenas um estudo sobre o efeito do choque térmico na transcrição de genes. Assim, um dos objectivos desta tese foi elucidar o efeito do choque térmico no metabolismo de *P. furiosus*. A estirpe DSM 3638 foi crescida em condições de temperatura óptima (90°C) e, na fase exponencial, as células foram sujeita a um choque térmico (97°C) durante 45 minutos. O conteúdo intracelular das duas condições de crescimento foi analisado por ressonância magnética nuclear (NMR) e espectrometria de massa (MS). Visto que as duas técnicas diferem em termos de sensibilidade na detecção, os métodos de recolha de células, paragem do metabolismo e extração dos metabolitos foram adaptados para cada análise. A análise por ^{1}H NMR permitiu a atribuição de 35 metabolitos, mas apenas 27 foram quantificados com sucesso, sendo que 37% dos metabolitos identificados foram aminoácidos. Foram também identificados ácidos orgânicos, nucleótidos, cofactores, poliaminas, açúcares, açúcares

nucleotídicos e solutos compatíveis. Em contraste, a análise efectuada por espectrometria de massa permitiu a identificação de 15 moléculas. Destas, 9 foram quantificadas com sucesso, incluindo 3 nucleótidos (ATP, ADP, AMP), uma combinação de duas trioses (2/3-fosfoglicerato), GDP-manose, glicerol-3-fosfato, fosfoenolpiruvato, ribose-1-fosfato e trealose-6-fosfato. A integração dos resultados obtidos através das duas técnicas suporta a observação de que temperaturas supra-óptimas favorecem a síntese de solutos compatíveis e açúcares nucleotídicos que podem estar envolvidos na síntese da parede celular. Em contraste, os metabolitos envolvidos na glicólise e ciclo de Krebs diminuíram com o aumento da temperatura. Do que temos conhecimento, este é o primeiro trabalho em que o metaboloma de um organismo hipertermófilo foi estudado sob condições de choque térmico. No entanto, é necessária a continuação do trabalho para aumentar o valor estatístico dos resultados, ampliar o número de metabolitos identificados e validar as conclusões retiradas.

Um desafio adicional nesta tese foi a elucidação da estrutura tridimensional da proteína bifuncional IPCT/DIPPS, uma das proteínas envolvidas na síntese do DIP. O domínio citosólico apresenta uma elevada semelhança com a família nucleotidiltransferase, uma família que engloba enzimas que catalisam a transferência de nucleótidos para açúcares fosfatados. O segundo domínio DIPPS contém seis segmentos transmembranares e catalisa a transferência do grupo inositol do CDP-inositol para o inositol-1-fosfato. Os substratos inositol-1-fosfato e CDP-inositol foram sintetizados para serem utilizados na caracterização

funcional da proteína e também nos ensaios de co-cristalização. Este domínio C-terminal contém uma sequência conservada $DG(x)_2AR(x)_8G(x)_3D(x)_3D$ típica da família de proteínas CDP-álcool fosfatidiltransferase. As proteínas que pertencem a esta família estão integradas na membrana, podem ser encontradas em todos os Domínios da Vida e normalmente estão envolvidas na síntese de fosfolípidos. Apesar do grande número de proteínas que pertencem a esta família, até à data do início deste trabalho nenhuma estrutura tinha sido elucidada. Assim, a resolução da estrutura do domínio DIPPS é muito relevante. Neste trabalho, foram obtidas as estruturas da proteína IPCT/DIPPS recombinante de *A. fulgidus* sem e com o ligando CDP-inositol por cristalografia de raio-X numa resolução de 2.65 e 2.5 Å, respectivamente. A estrutura do domínio N-terminal contém seis hélices transmembranares (TM) e uma α-hélice paralela ao plano da membrana. A proteína forma um dímero por intermédio das suas hélices transmembranares. A estrutura com o CDP-inositol ligado revelou a localização do centro activo numa cavidade hidrofílica amplamente aberta para o citoplasma com um ião de magnésio rodeado por quatro resíduos de aspartato altamente conservados localizados nas hélices TM2 e TM3. A mutagénese dirigida dos resíduos de aminoácidos da sequência de consenso da família ou de resíduos que foram considerados relevantes para a ligação de substrato validou os resultados obtidos com a ligação do CDP-inositol. Os dados obtidos suportam também a proposta de um mecanismo catalítico universal para as proteínas da família CDP-álcool fosfatidiltransferase.

Resumindo, os resultados apresentados nesta tese representam um avanço significativo no nosso conhecimento das estratégias de adaptação ao stresse térmico em arqueões hipertermófilos. O papel de DIP e MG na proteção contra condições de stresse foi estabelecido. Além disso, a resposta do metaboloma a condições de choque térmico num organismo considerado modelo parecem indicar que o metabolismo é desviado principalmente para a síntese de solutos compatíveis (DIP e MG) e componentes da parede celular. A elucidação da estrutura tridimensional e do mecanismo catalítico das DIPPS representam avanços para a construção de modelos para proteínas homólogas das sintases de fosfatidilinositol-fosfato. Esta enzima é essencial para a viabilidade de *Mycobacterium tuberculosis* e, consequentemente, a determinação da estrutura do homólogo DIPPS é um passo fundamental para a concepção de novos fármacos anti-tuberculose.

Contents

Abbreviations

2-AAA, 2-aminoadipate

2-OHGlu, 2-hydroxyglutarate

2PG, 2-phosphoglycerate

3PG, 3-phosphoglycerate

5-FOA, 5-Fluoroorotic Acid

Aco, aconitate

ADP, adenosine 5’-diphosphate

AMP, adenosine 5’-monophosphate

ATP, adenosine 5’-triphosphate

BPG, 1,3-bisphosphoglycerate

CDP-DAG, cytidine 5’-diphospho-diacylglycerol

CDP-gly, cytidine 5’-diphosphoglycerol

CDP-Ino or CDP-inositol, cytidine 5’-diphospho-*myo*-inositol

CDP-OH_P_trans, CDP-alcohol phosphatidyltransferase

Cit, citrate

COSY, ^{1}H-^{1}H correlation spectroscopy

CTP, cytidine 5’-triphosphoglycerol

DAG, diacylglycerol

DDM, n-dodecyl β-D-maltopyranoside

DFR, down flanking region

DHAP, dihydroxyacetone phosphate

diC8PC, 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine

DIP, di-*myo*-inositol-phosphate or di-*myo*-inositol-1,3'-phosphate

DIPP or DIP-phosphate, di-*myo*-inositol-1,3'-phosphate-1-phosphate or di-*myo*-inositol-phosphate-phosphate

DIPPS or DIPP synthase, CDP-L-*myo*-inositol:L-*myo*-inositol-1-phosphate transferase or di-*myo*-inositol-1,3'-phosphate-1-phosphate synthase or di-*myo*-inositol-phosphate-phosphate synthase

DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine

ED, Entner-Doudoroff Pathway

EM, Embden-Meyerhof pathway

FID, free induction decay

GAP, glyceraldehyde-3-phosphate

GAPDH, glyceraldehyde-3-phosphate dehydrogenase

GAPN, NADP-dependent glyceraldehyde-3-phosphate dehydrogenase

GAPOR, ferredoxin-dependent glyceraldehyde-3-phosphate oxidoreductase

GDH, glutamate dehydrogenase

GDP, guanosine 5'-diphosphoglycerol

GDP-man, guanosine-5’-diphospho-D-mannose

Gly3P, glycerol-3-phosphate

HSP, heat shock protein

HSP20 or sHSP, heat shock protein 20 kDa

HSP60 or thermosome, heat shock protein 60 kDa

HSQC, heteronuclear single quantum coherence spectroscopy

HSR, heat shock response

IDMS, isotope dilution mass spectrometry

Ino1P, inositol-1-phosphate or L-*myo*-inositol-1-phosphate

IPCT, CTP:L-*myo*-inositol-1-phosphate cytidylyltransferase

IPTG, isopropyl β-D-1-thiogalactopyranoside

Isocit, isocitrate

LC/MS, liquid chromatography/mass spectrometry

LDAO, lauridyldimethyl amine oxide

Mal, malate

MG, mannosylglycerate

MGS, mannosylglycerate synthase

MIPS, L-*myo*-inositol-1-phosphate synthase

MPG, mannosyl-3-phosphoglycerate

MPGP, mannosyl-3-phosphoglycerate phosphatase

MPGS, mannosyl-3-phosphoglycerate synthase

NADP, nicotinamide adenine dinucleotide phosphate

NAD, nicotinamide adenine dinucleotide

NMR, nuclear magnetic resonance

NOESY, nuclear overhauser enhancement spectroscopy

OD, optical density

OTG, n-octyl β-D-thioglucopyranoside

PCR, polymerase chain reaction

PEP, phosphoenolpyruvate

P_i, inorganic phosphate

PK, pyruvate kinase

POR, pyruvate ferredoxin oxidoreductase

PP_i, pyrophosphate

PTS, phosphotransferase system

pyrF, orotidine-5'-phosphate decarboxylase

Rib1P, ribose-1-phosphate

RT-PCR, reverse transcriptase polymerase chain reaction

Succ, succinate

TM, transmembrane

TMAO, trimethylamine N-oxide

TOCSY, ^{1}H-^{1}H total correlation spectroscopy

Tre6P, trehalose-6-phosphate

TSP, sodium 3-(trimethylsilyl) proprionate-2,2,3,3-d_4

UDPGalNac, UDP-*N*-acetylgalactosamine

UDPGlcNac, UDP-*N*-acetylglucosamine

UFR, up flanking region

VAT or *aaa*$^{+}$ATPase, VCP-like ATPase of *Thermoplasma acidophilum*

Amino acids

Ala or A, Alanine

Arg or R, Arginine

Asn or N, Asparagine

Asp or D, Aspartic acid

Cys or C, Cysteine

Glu or E, Glutamic acid

Gln or Q, Glutamine

Gly or G, Glycine

His or H, Histidine

Ile or I, Isoleucine

Leu or L, Leucine

Lys or K, Lysine

Met or M, Methionine

Phe or F, Phenylalanine

Pro or P, Proline

Ser or S, Serine

Thr or T, Threonine

Trp or W, Tryptophan

Tyr or Y, Tyrosine

Val or V, Valine

CHAPTER 1

General introduction

Contents

Life in extreme environments

In the last 50 years, scientists discovered life in inhospitable places which were previously considered incapable of supporting life. These places included deep sea black smokers, alkaline lakes, acidic springs, hot sulfurous springs, and salt lakes (Schäfer 1992). Such environments are characterized by extreme temperatures, pressures, salinities, and pH values. Therefore, organisms proliferating under these extreme conditions are called “extremophiles” and can be divided according to their optimal growth conditions (Table 1.1). Although the majority of extremophiles belong to the domain Archaea, several members of Bacteria and Eukarya are able to cope with extreme conditions.

The unique physiological properties of extremophiles have attracted the interest of the scientific community. Deciphering the strategies used by these organisms to cope with such extreme conditions is a challenging task. In this thesis we will focus mainly on hyperthermophiles.

TABLE 1.1 Examples of extreme environmental conditions and classification of extremophiles (adapted from Rothschild and Mancinelli 2001).

Environmental parameter	Classification	Minimum	Maximum	Examples
Temperature	Psychrophile	-	15ºC	*Psychrobacter*, some insects
	Thermophile	60ºC	80ºC	*Rhodothermus marinus*
	Hyperthermophile	80ºC	113 ºC	*Pyrolobus fumarii*
Salinity	Halophile	2 M	5.1 M	Halobacteriaceae, *Dunaliella salina*
Radiation	-	-	10 kGy*	*Deinococcus radiodurans*
Pressure	Piezophiles	-	50 MPa	*Shewanella benthica*
Desiccation	Xerophile	-	0.605 a_w[a]	*Artemia salina*, nematodes, microbes, fungi, lichens
pH	Alkaliphile	9	14	*Natronobacterium*, *Bacillus firmus* OF4, *Spirulina* spp.
	Acidophile	0	4	*Cyanidium caldarium*, *Ferroplasma* sp. (both pH 0)

* Values for Gamma radiation. [a] Water activity defined by the equation a_w= vapour pressure of the solution/vapour pressure of the water (Stevenson et al. 2014)

Hyperthermophilic organisms

By 1980, several thermophilic organisms growing optimally at temperatures up to 75°C had already been isolated (Brock 1978) being *Sulfolobus acidocaldarius*, a thermoacidophilic archaeon living optimally at temperatures ranging from 70-75°C and at pH 2, the most thermophilic organism characterized so far (Brock et al. 1972). Stetter and Zillig in their first trip to Iceland, in 1980, collected samples from boiling springs and mud pools and they were able to isolate the first organism exhibiting

fastest growth above 80°C, which was named *Methanothermus fervidus* (Stetter et al. 1981). Hyperthermophilic organisms turned out to be very common in hot terrestrial and submarine environments and, since the early eighties, more than 137 species have their genomes completely sequenced. The most extreme hyperthermophilic organism characterized so far belongs to the genus *Pyrolobus*. The archaeon *Pyrolobus fumarii* was isolated from deep-sea hydrothermal vents and is able to grow up to 113°C (Blöchl et al. 1997).

The construction of a phylogenetic tree based on the comparison of 16S rRNA sequences revealed the location of hyperthermophilic organisms in all the short deep branches closest to the root (Fig. 1.1). Their position within the phylogenetic tree in combination with their growth requirements led Stetter to hypothesize the existence of similar hyperthermophilic organisms on the early Earth, 3.9 billion years ago (for useful reviews see Stetter 2006, 2013).

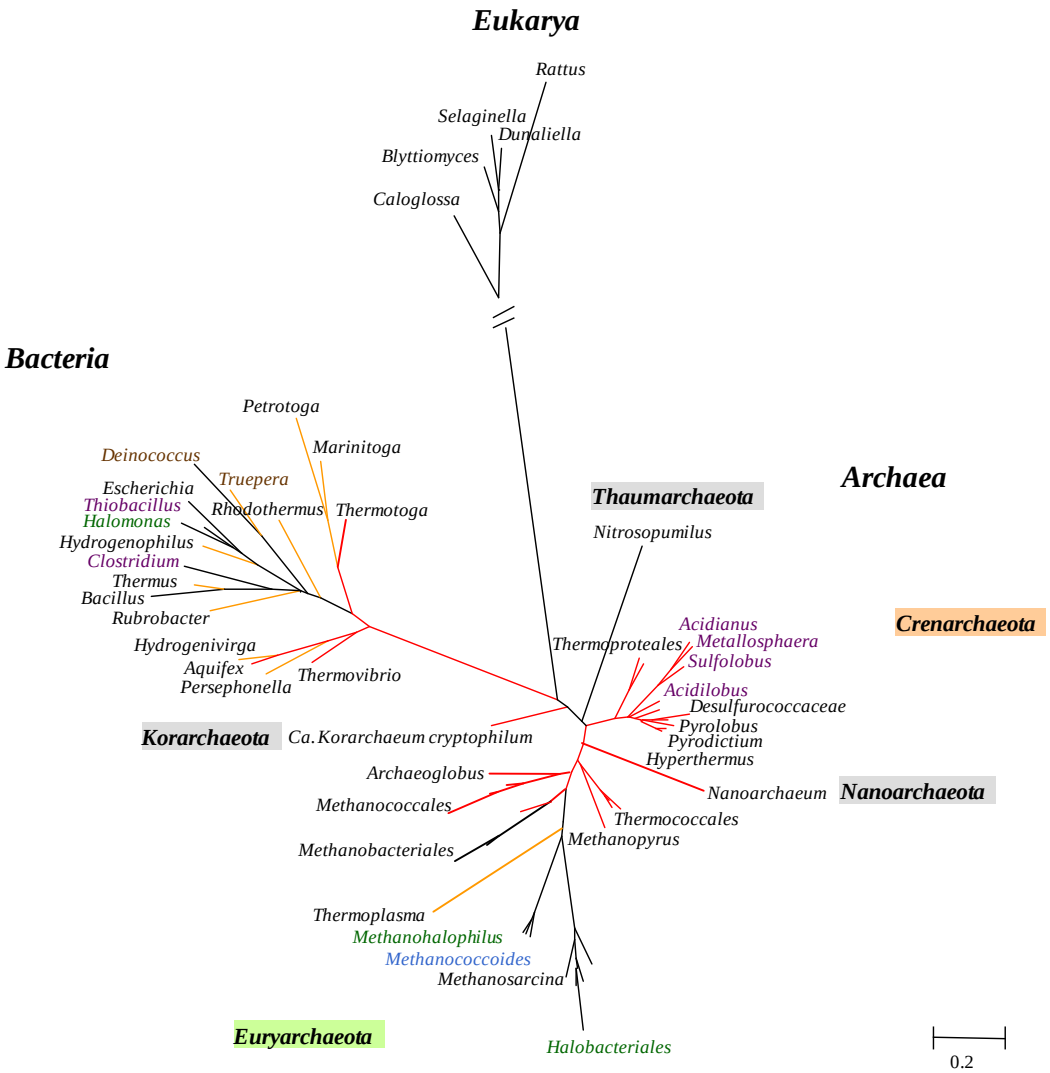


FIGURE 1.1 Small subunit ribosomal RNA-based phylogenetic tree. Hyperthermophilic and thermophilic branches are represented in red and orange, respectively. Desulfurococcaceae include the genera *Ignisphaera*, *Aeropyrum*, *Ignicoccus*, and *Staphylothermus*; Methanococcales include the genera *Methanocaldococcus*, *Methanotorris*, and *Methanococcus*; Methanobacteriales include the genera *Methanosphaera*, *Methanobacterium*, and *Methanothermus*;

Halobacteriales include the genera *Halobacterium*, *Natrobacterium*, *Haloferax*, *Natronococcus*, and *Halococcus*. Reproduced with permission from Rodrigues 2011.

Most hyperthermophilic organisms isolated thus far belong to the domain Archaea but the two bacterial genera, *Thermotoga* and *Aquifex*, include hyperthermophilic species. Within Eukarya, a few thermophilic organisms have been described. One example is the red alga *Cyanidium caldarium*, which grows optimally at 56°C (Doemel and Brock 1970). It is interesting to note that, the only two orders within hyperthermophilic Bacteria (Aquificales and Thermotogales) are close to the root of the phylogenetic tree. Therefore, these hyperthermophilic bacteria are considered good candidates for evolutionary studies (Vieille and Zeikus 2001, Stetter 2006). Archaea are divided into 5 phyla, 34 genera and 10 orders. They are distributed among the phyla Crenarchaeota, Euryarchaeota, Korarchaeota, Nanoarchaeota, and Thaumarchaeota. Most of the well-studied hyperthermophiles belong either to the Crenarchaeota or the Euryarchaeota, while the phylum Thaumarchaeota was proposed only in 2008 and gathers many mesophilic and thermophilic archaea (Brochier-Armanet et al. 2008). Representatives of the order Thermococcales were isolated from marine hot water ecosystems and, so far, form the most numerous group of hyperthermophilic archaea (Bertoldo and Antranikian 2006). They belong to the branch of the Euryarchaea and can be divided into 3 genera: *Thermococcus*, *Pyrococcus* and *Palaeococcus*. Presently, the three genera are represented by 39 species,

with two belonging to the genus *Palaeococcus*, 7 belonging to *Pyrococcus* and 30 to the genus *Thermococcus*.

Adaptation of hyperthermophiles to high temperature

Due to their small size and relatively simple morphology, hyperthermophilic organisms are not able to protect themselves by isolation against environmental high temperatures. Therefore, cell components from hyperthermophilic organisms must be heat resistant (Stetter 1999). The molecular adaptations responsible for the survival and growth of hyperthermophiles at high temperatures are not completely clarified. Sometimes, it is difficult to distinguish features that are associated with (hyper)thermophily from those related with bacterial and archaeal phylogeny (Daniel and Cowan 2000). A good example is the lipid structure of bacterial and archaeal membranes. The ester-linkages typically found in Bacteria and Eukarya are replaced by ether-linkages in Archaea (Koga 2012). Despite the higher stability of ether-linkages, it is impossible to consider this feature as an adaptation to thermophily since ether-linkages are present in all members of the domain Archaea, not only in those thriving at high temperatures. Moreover, even though certain thermophilic bacteria have in their membranes ether-linkages rather than the typical ester-linkages, those with ester lipids can also survive under high temperatures (Daniel and Cowan 2000, Koga 2012). Therefore, the presence of ether lipids in cell membranes is most likely a phylogenetic feature. Most archaea growing under moderate conditions exhibit two C20 hydrocarbon chains attached to glycerol through ether-linkages. In

contrast, thermophilic members of Archaea have in their membranes C40 hydrocarbon chains, forming cyclic tetraether bilayers (Konings et al. 2002, Matsuno et al. 2009).

Nucleic acids can be severely affected by high temperatures. For instance, under high temperatures, DNA undergoes denaturation by strand separation. Therefore, hyperthermophilic organisms had to adapt to maintain genome integrity. Two strategies that can prevent unfolding of DNA and RNA strands are: increase the GC content of molecules; and enhance intracellular salt concentration (Forterre 2009). However, none of these strategies were generally adapted by hyperthermophilic organisms (Daniel and Cowan 2000). Typically, hyperthermophilic organisms have GC content between 30 to 60%, which is similar or even lower than mesophilic organisms (Daniel et al. 2004). No correlation is observed between optimal growth temperature and GC content. For instance, *Pyrococcus furiosus* and *Pyrolobus fumarii*, which grow optimally between 100°C and 106°C, have a GC content of 38 and 53%, respectively (Daniel et al. 2004). Comparative genome-based studies showed that reverse gyrase, a protein responsible for the introduction of positive supercoils into dsDNA, was the only protein specific for hyperthermophilic organisms (Forterre et al. 2000, Forterre 2002). In an attempt to unravel the physiological importance of reverse gyrase in hyperthermophiles, Atomi and co-workers deleted the gene encoding reverse gyrase in *Thermococcus kodakarensis*. They found that the disruption of the target protein did not result in cell death. Actually, its absence only affected the

growth rate, an effect that was more evident at higher growth temperatures (Atomi et al. 2004). Therefore, the authors concluded that reverse gyrase is not a prerequisite for life at high temperatures.

When compared with proteins from mesophilic organisms, proteins from hyperthermophiles are generally more stable. For instance, an amylase from *Pyrococcus woesei* is still active at 130°C (Koch et al. 1991). Therefore, understanding the mechanism underlying protein stability is relevant from an academic point of view but is also important for industrial applications since hyperthermophiles are a good source of thermostable enzymes (for useful review sees Atomi et al. 2011). Thermostability has been correlated with a combination of structural parameters, such as, increased number of ionic interactions, increased extent of hydrophobic-surface, increased number of proline and glutamine residues, improved core packing, greater rigidity, extended secondary structures, shorter surface loops, and higher states of oligomerization (Luke et al. 2007). A statistical analysis of the thermophilic protein sequences revealed an overall enrichment in charged amino acids (Greaves and Warwicker 2007). A study comprising 93 proteins from meso-, thermo-, and hyperthermophilic origin led to the conclusion that those proteins with hyperthermophilic origin have stronger ion pairing, fewer cavities, higher β-sheet contents, and shorter loops (Szilágyi and Závodszky 2000). However, a general concept of how proteins stabilize their fold at high temperature is still missing.

A number of low molecular mass metabolites are unstable at temperatures around 100°C. The most critical metabolites seem to be ATP and NAD, having half-lives below 30 minutes (*in vitro* studies) (Stetter 1999, Daniel et al. 2004). The cleavage of phosphate-phosphate bonds present in ATP, ADP and AMP provides the cell the energy necessary to catalyze many other reactions. Therefore, cells must adapt to high temperatures maintaining the necessary energy sources. From the phosphorylated compounds (ATP, ADP, AMP, PP_i), ATP is the most unstable at temperatures near 100°C. Several thermophilic archaea use more stable compounds, such as PP_i and ADP, rather than ATP. For instance, the hexokinase and phosphofructokinase in the genus *Pyrococcus* are ADP- rather than ATP-dependent (Daniel et al. 2004). NAD(P) serves as an electron carrier that transports electrons between molecules for oxidation-reduction reactions. Many thermophilic organisms, such as *Thermotoga maritima* and *P. furiosus*, have enzymes dependent on the reduction of ferredoxin instead of NAD(P) (Buckel and Thauer 2013).

Stress response

Microorganisms living in soil and water environments are often exposed to changes in environmental conditions. These changes may be physical, like temperature variation, or chemical, such as changes of salinity, pH or oxygen concentration. Therefore, to survive and function under unfavorable conditions, microorganisms had to develop stress responses

to protect cell components but also to repair damages caused by sudden changes in environmental conditions.

Heat stress

The heat shock response (HSR) was first observed in *Drosophila* when chromosomal puffs were formed and specific proteins were induced under supraoptimal temperatures (Ritossa 1962). Ever since, the HSR has been studied in several species of Eukarya, Bacteria, and Archaea (Laksanalamai et al. 2004b). Curiously, comparative genomics between different species revealed a high degree of conservation of the basic features of HSR (Feder and Hofmann 1999). In agreement, transcriptional studies using model organisms across the three domains showed that approximately 50-200 genes are significantly induced (Richmond et al. 1999, Gasch et al. 2000, Tabuchi et al. 2008). For example, an up-shift in the growth temperature of *Escherichia coli* results in the induction of over 120 genes (Meyer and Baker 2011). These heat-induced genes code for proteins that can be divided into seven classes, according to their function (Richter et al. 2010). The class designated as heat shock proteins (HSP) is preferentially synthesized across species. This group is usually composed by molecular chaperones, which are responsible for the refolding of denatured or newly synthesized proteins, usually through cycles of binding and release, without being a part of the final native structure (Young et al. 2004). Most organisms have five different groups of HSPs, usually classified according to their molecular mass: HSP100 (80-110 kDa),

HSP90 (82-96 kDa), HSP70 (67-76 kDa), HSP60 (58-65 kDa), and the small HSP (below 40 kDa) (Trent 1996, Laksanalamai et al. 2004a). With the exception of mesophilic archaea, where homologues of HSP100 and HSP70 can be found, most archaeal species lack the larger HSPs (Table 1.2). In contrast, proteins homologous to the HSP60 and small HSP (sHSP) are usually found in meso-, thermo-, and hyperthermophilic archaea (Laksanalamai et al. 2004a).

TABLE 1.2 Distribution of HSPs in the three Domains (adapted from Laksanalamai et al. 2004a).

HSP	Bacteria	Eukarya	Archaea
HSP100	ClpA, ClpB, HsIU	HSP100	Absent*
HSP90	HtpG	HSP90, HSP83	Absent
HSP70	DnaK	HSP70, HSC70	HSP70[#]
HSP60	GroEL	TCP-1	TF55, thermosome
sHSP	lbpA, lbpB	sHSP, α-crystallin	sHSP

* ClpA/B homologues are found in *Methanothermobacter thermautotrophicus*. [#] HSP70 are absent from most thermophiles and all the hyperthermophiles, with the exception of *Aeropyrum pernix*, in which a putative truncated HSP70 has been detected. TCP-1, T-complex protein 1; TF55, thermophilic factor 55.

The HSP60 proteins, also known as chaperonins, are a subdivision of molecular chaperones characterized by a double-ring structure, which are divided into two groups. The group I is commonly known as GroEL/ES complex and is found in Bacteria, some Eukarya and several mesophilic

Archaea, while Group II is found in most eukaryotes, and archaea (Laksanalamai et al. 2004a). In hyperthermophilic archaea, the group II is usually referred as “thermosome”. In opposition to eukaryotes, where as many as eight homologues can be found, hyperthermophilic archaea have usually between 1 and 3 thermosome-encoding homologous genes. For instance, the genus *Pyrococcus* has only a homologue of HSP60 gene, while *Sulfolobus solfataricus* and *Sulfolobus tokodaii* have on their genome three genes coding for HSP60 protein (Laksanalamai et al. 2004a). Transcriptional studies carried out using *P. furiosus* and *Archaeoglobus fulgidus* showed that genes coding for thermosome and sHSP are induced by heat stress (Shockley et al. 2003, Rohlin et al. 2005). The sHSP in archaea resemble the α-crystallin protein family from Eukarya. Although the amino acid sequences of sHSPs diverge among archaea, all of them share a sequence motif with $\alpha\beta$ crystalline protein of the vertebrate eyes lens (Narberhaus 2002). Their molecular mass varies in the range 15 to 40 kDa and, *in vivo*, they appear to form multimeric complexes (Laksanalamai et al. 2004b). Like HSP60, the number of *shsp* homologues in the genome is variable within hyperthermophilic archaea. Most organisms have one to two *shsp* homologues, being *Pyrobaculum aerophilum* the only exception with three homologues (Laksanalamai et al. 2004b). Several studies using sHSP from hyperthermophilic organisms, such as *Methanocaldococcus jannaschii*, *Pyrococcus furiosus*, and *Thermococcus* sp. KS-1, revealed their ability preventing aggregation of mesophilic proteins at high temperatures (Kim et al. 1998, Laksanalamai et al. 2001, Usui et al. 2001).

Therefore, it has been suggested that sHSP from extremophilic organisms (Bacteria and Archaea) act as molecular chaperones. Actually, this role had been already established for α-crystallins and sHSP from plants, mammals, and mesophilic prokaryotes (Haslbeck 2002, Laksanalamai et al. 2004b). The second class of proteins preferentially synthesized in response to heat stress is composed by components of the proteolytic system, like proteases. They are responsible for the clearance of misfolded and irreversibly aggregated proteins (Richter et al. 2010). The third class encloses DNA- and RNA-repair proteins, while the fourth class is composed by metabolic enzymes. This class presents more variation among species. The remaining classes include regulatory proteins that are involved in the activation and or/repression of target genes involved in stress response, proteins that sustain cellular structures, such as the cytoskeleton, and proteins involved in membrane-related processes, such as transport and structure (Richter et al. 2010). Actually, the membrane stability is an important feature for the cell. An increase in temperature makes the membrane more fluid, and consequently less stable. In yeast, the protein HSP12 binds directly to the membrane, decreasing its fluidity. When soluble, this protein is unstructured, but when it contacts with lipids becomes α-helical (Welker et al. 2010). Independent studies using membranes from *T. kodakarensis* and *A. fulgidus* showed that the growth temperature modulates the lipidic composition of the membrane. This condition affects the length of the hydrocarbon chains. For example, the

ratio of cardarchaeol (C40) to archaeol (C20) increases with increasing growth temperatures (Lai et al. 2008, Matsuno et al. 2009).

When compared with Eukarya and Bacteria, the heat shock response in Archaea, especially in those organisms living at high temperatures, is poorly characterized. So far, studies were carried out on *P. furiosus*, *A. fulgidus*, *S. solfataricus*, and *M. jannaschii* (Shockey et al. 2003, Rohlin et al. 2005, Boonyaratanakornkit et al. 2005, 2007, Tachdjian and Kelly 2006). Among hyperthermophilic bacteria, *T. maritima* is the only species with a genome-wide investigation under heat stress (Pysz et al. 2004). A general trend was observed in *P. furiosus*, *A. fulgidus*, and in *M. janaschii*, where the HSP60, sHSP, and the proteasome were up-regulated under heat stress. In contrast, in *A. fulgidus* the genes coding for a prefoldin were down-regulated, while in *M. janaschii* the homologous genes were up-regulated under high temperatures. The expression pattern of prefoldin subunits was examined in *T. kodakarensis* grown at different temperatures (Danno et al. 2008). The study revealed that two subunits were not affected while the remaining two subunits were heat-induced. Curiously, some elements of the mesophilic heat shock response are conserved in *T. maritima*. However, several differences were also noticed. For instance, most genes encoding ATP-dependent proteases were down-regulated. Moreover, the regulatory strategies adopted by *T. maritima* seem to be different from those usually found in mesophilic bacteria (Pysz et al. 2004).

Regulation of the heat shock response

The regulatory mechanisms underlying the activation of heat stress response are still poorly understood. The transcriptional activation of heat shock genes plays an essential role on the global response to elevated temperatures. Therefore, a great effort was directed to understanding the transcriptional regulation of heat shock genes. In eukaryotes, a heat shock factor called Hsf1 was early discovered and characterized (Wu 1984). This regulon binds to DNA sequences through the recognition of a pentameric heat shock element (HSE), initiating the assembly of the transcription machinery (Sorger and Pelham 1987). Later, a second group of regulatory proteins, known as MSN2 and MSN4, was identified in *S. cerevisiae*. In contrast to the Hsf1, which is specific for heat shock response, the MSN2/4 regulon is activated by other environmental stresses, like oxidative stress (Morano et al. 2012). Bacteria control the transcription of heat shock genes by both activation and repression mechanisms. For example, in *E. coli* the σ^{32} is the main transcriptional regulator involved in the activation of heat shock proteins. At elevated temperatures, σ^{32} binds to RNA polymerase and, the resulting complex is able to recognize and bind to a specific promoter defined by the sigma factor (Guisbert et al. 2008). In contrast, transcription repressors can be found in many bacterial species, such as *Bacillus* and *Streptomyces*. The HrcA, HspR, and RheA are a few examples of repressor proteins (Servant and Mazodier 2001). In hyperthermophilic archaea, Phr is the only transcription regulator known that is involved in heat stress response. It was first identified in *P. furiosus*

and, it has homologues among Euryarchaea (Vierke et al. 2003). The mechanism of action of Phr will be described below.

Interestingly, the genome of *T. maritima* has no homologues of sigma factor regulators usually found in gram-negative mesophilic bacteria (Klinkert and Narberhaus 2009) suggesting that *T. maritima* has an alternative mechanism for the regulation of the heat shock response.

Osmotic stress

Organisms living either in water- or soil-environments are subjected to variations on environmental osmolarity and, since biological membranes are permeable to water, variations on external osmotic pressure may lead to the influx or efflux of water. The influx of water occurs when cells are subjected to hypo-osmotic conditions, while under hyper-osmotic conditions the efflux of water causes a reduction in the cell volume. Therefore, to avoid either cell lysis or dehydration, respectively, cells had to develop strategies to cope with changes in external salt concentrations. The most common strategy adopted by organisms to survive hyper-osmotic shock involves the intracellular accumulation of ions or organic solutes to counteract the external osmotic pressure.

The influx of inorganic ions, such as K^+ and Cl^-, is preferentially used by several halophilic Archaea from the family Halobacteriaceae, by bacteria from the order Haloanaerobiales, and by the bacterium *Salinibacter ruber* (Oren 1999, da Costa et al. 1998, Oren and Mana 2002). These organisms can accumulate inorganic ions to high

concentrations, and for that reason, most of their proteins evolved on a way to cope with such conditions. For instance, halophilic proteins are usually enriched with acidic amino acids, and for most of them, the presence of K^+ and/or Na^+ is a prerequisite for enzyme stability and activity (Eisenberg and Wachtel 1987, da Costa et al. 1998).

In contrast, the accumulation, either by transport or synthesis, of specific organic solutes is widespread in nature, being found in many organisms including archaea, bacteria, fungi, yeast and algae (Martin et al. 1999). The organic solutes are usually low-molecular mass molecules, highly soluble in water, neutral or zwitterionic at physiological pH, and like inorganic ions, may accumulate to high levels. As their presence in the cytoplasm does not interfere with the normal function of the cell, they were designated “compatible solutes” (Brown 1976, Yancey et al. 1982). According to their chemical nature, compatible solutes can be divided into amino acids and derivatives, sugars and derivatives, polyols, betaines, and ectoines (da Costa et al. 1998). Besides their role in osmoadaptation, compatible solutes can also protect cell components against other stress conditions, like heat stress, oxidative stress, freezing and desiccation (Santos and da Costa 2002).

The topics osmosensing and osmoregulation have been extensively studied in several mesophilic bacteria, such as *E. coli* and *Bacillus subtilis* (for useful reviews see Morbach and Krämer 2004, Krämer 2010). In Archaea, little is known about the molecular mechanisms underlying accumulation of compatible solutes. The haloarchaea *Haloferax vulcanii* is

a model organism in this field, but little information on this topic is available. Transcriptional and proteomic studies were carried out using *H. vulcanii* grown at different salinity conditions (Ferrer et al. 1996, Mojica et al. 1997, Bidle et al. 2008). A homologue of the mesophilic protein PspA, which is involved in osmoregulation, was identified in *H. vulcanii*. Bidle and collaborators demonstrated that the protein levels are salt-mediated, but the mechanisms underlying this response are most-likely different from those found in mesophilic bacteria (Bidle et al. 2008).

Diversity and distribution of compatible solutes

The strategy of organic solute accumulation is preferentially used by the majority of living organisms during osmoadaptation since it does not interfere with cell machinery or function. Energetically, the uptake of compatible solutes from the external medium is more favorable. However, several organisms prefer *de novo* synthesis of organic solutes to cope with osmotic stress. In the last decades, with the discovery of new organisms, mainly those living at high temperature, the list of compatible solutes has increased significantly. Even though, the structural diversity of compatible solutes discovered so far is limited, as all compounds belong to two major classes: i) sugars and polyols and ii) amino acids and their derivatives (Roeßler and Müller 2001). This is likely a consequence of the fundamental features that compatible solutes must own, i.e., compatibility with macromolecules and cellular functions (Le Rudulier et al. 1984, Roeßler and Müller 2001). Some compatible solutes like trehalose, proline, glycine

betaine, and ectoine are widespread in the nature, while others, such as di-*myo*-inositol-phosphate, appear to be confined to organisms living under high temperatures.

The amino acids glutamate, proline, glutamine, and alanine are commonly found in mesophilic bacteria during osmotic stress (Empadinhas and da Costa 2008). The accumulation of α-glutamate was described for many mesophiles and hyperthermophiles during adaptation to low osmotic pressure (da Costa et al. 1998). It was suggested that α-glutamate could function as a counterion for K^+. The accumulation of the β-isomer of glutamate has been also described but only in a few members of mesophilic and hyperthermophilic bacteria, such as *Petrotoga mobilis* and *Aquifex pyrophilus* (Lamosa et al. 2006, Fernandes et al. 2010). In opposition to alanine and glutamine, proline can accumulate to high levels (in the order of Molar), in members of gram-positive bacteria. In *Streptomyces* sp., the three amino acids are increased under osmotic stress (Killham and Firestone 1984). Within prokaryotes, proline was first identified in members of the halophilic *Bacillus* species, as well as in the non-halophilic *Bacillus subtilis*. Later, researchers found that proline is usually accumulated by *Bacillus* spp. but only in combination with the main osmolyte, ectoine (Galinski 1995). Aspartate is rarely used as a compatible solute. To date, it was only identified in a few members of hyperthermophilic archaea, such as members of the order *Thermococcales* and in some *Methanothermococcus* spp. (Lamosa et al. 1998, Martin et al. 1999, Neves et al. 2005). Glycine betaine has been described to act as

osmoprotectant in plants, mammals, algae, bacteria and archaea. However, the strategy for glycine betaine accumulation is not conserved across the different domains of life. While the vast majority of living cells can capture glycine betaine from the external environment, few can synthesize this molecule either from choline or methylation of glycine. *B. subtilis* and *Methanothermobacter marburgensis* are examples of bacterial and archaeal species, respectively, which synthesize glycine betaine from choline (Roeßler and Müller 2001). In addition to glycine betaine, ectoin is one of the most used compatible solutes, despite its absence in archaeal species. Ectoine and hydroxyectoine are cyclic forms of *N*-acetylated amino acids, which can be either taken up from external medium or synthesized (da Costa et al. 1999).

The non-reducing glucose disaccharide trehalose occurs in a wide variety of organisms, such as mesophilic bacteria (*Corynebacterium glutamicum*), thermophilic bacteria (*Rhodothermus marinus*), and hyperthermophilic archaea (*Palaeococcus ferrophilus*) (Nunes et al. 1995, Neves et al. 2005, Avonce et al. 2006). A variety of functions have been proposed for trehalose, from carbon storage to cell protector against a diversity of stresses, like oxidation, dehydration, heat and cold (da Costa et al. 1998, Elbein et al. 2003). Sucrose, sulfotrehalose, glucosylglycerate, mannosylglycerate, and the unusual mannosucrose are examples of others carbohydrates used as compatible solutes by organisms (da Costa et al. 1998). Sucrose is accumulated by non-halophilic cyanobacteria while mannosucrose has been detected in *Agrobacterium tumefaciens*, a

mesophilic bacterium that can tolerate up to 2% (w/v) of NaCl (Smith et al. 1990). The organic solute glucosylglycerate is a structural analogue of mannosylglycerate and its role as a compatible solute has been demonstrated in several halophilic organisms, namely in the archaeon *Methanohalophilus portucalensis* and in the bacterium *Chromohalobacter salexigens* (Cánovas et al. 1999, Empadinhas and da Costa 2011). Currently, it is known that glucosylglycerate accumulates in response to osmotic stress in nitrogen-limited environments (Empadinhas and da Costa 2011). The presence of mannosylglycerate has been described in Archaea, Bacteria, and Eukarya. Among eukaryotes, the organic solute has been detected in 44 genera of the red algae, but its physiological role remains unclear (Borges et al. 2014). *Rhodothermus marinus*, *Thermus thermophilus*, and *Rubrobacter xylanophilus* represent the thermophilic bacteria in which the presence of mannosylglycerate was confirmed (Nunes et al. 1995, Santos and da Costa 2002). Within archaea, mannosylglycerate appears only in marine hyperthermophiles of the genera *Pyrococcus*, *Thermococcus*, *Palaeococcus*, *Archaeoglobus*, *Aeropyrum,* and *Stetteria* (Borges et al. 2014). With the exception of *Rubrobacter xylanophilus*, in which the pool of mannosylglycerate is not affected by the temperature or salinity of growth, all the organisms responded to osmotic stress by increasing the pool of mannosylglycerate. Moreover, in *R. marinus* and *P. ferrophilus* the organic solute responded not only to osmotic stress but also to an increase in the growth temperature (Nunes et al. 1995, Neves et al. 2005). To investigate the role

of mannosylglycerate in osmoadaptation, mutants of *T. thermophilus* unable to synthesize this organic solute were constructed. Results showed that mannosylglycerate is essential for osmoadaption at low salinity (Alarico et al. 2007).

Glycerol, arabitol, *myo*-inositol, sorbitol, and mannitol are examples of compatible solutes belonging to the class of polyols, typically found within the domain Eukarya, namely in yeast, algae, plants, and mammals (da Costa et al. 1998). The distribution of polyols among bacteria is restricted to a few species, such as *Zymomonas mobilis* and *Pseudomonas putida*. The former utilizes sorbitol as osmoprotector, while the latter organism accumulates mannitol during osmotic stress (da Costa et al. 1998).

In contrast, polyol phosphodiesters represent a group of compatible solutes that are restricted to thermophilic and hyperthermophilic species. Di-*myo*-inositol-phosphate (DIP) is one example of a compatible solute widespread in hyperthermophilic organisms and does not appear in mesophilic organisms. It was initially identified in *Pyrococcus woesei* in 1992 (Scholz et al. 1992) and, since then, it has been described in many archaea from the genera *Archaeoglobus*, *Thermococcus*, *Pyrococcus*, *Pyrolobus*, *Hyperthermus*, *Stetteria*, *Aeropyrum,* and *Methanotorris*. Among bacteria it was identified in the genera *Aquifex*, *Thermotoga*, and *Rubrobacter* (Santos et al. 2011). In most organisms studied so far, di-*myo*-inositol-phosphate levels increased in response to heat stress. For instance, in *P. furiosus*, an increase of 6°C in the growth temperature resulted in a 20-fold increase in DIP's level (Martins and Santos 1995).

Two novel compounds, diglycerol phosphate and glycero-phospho-inositol, were first identified in the hyperthermophilic archaeon *A. fulgidus* and in the bacterium *Aquifex pyrophylus* (Martins et al. 1997). To date, diglycerol phosphate was encountered only in members of the genus *Archaeoglobus* under osmotic stress conditions (Gonçalves et al. 2003, Lamosa et al. 2006), while glycero-phospho-inositol is present in species of the genera *Archaeoglobus* and *Aquifex* and increases mainly in response to a combination of heat and osmotic stresses (Lamosa et al. 2006).

Accumulation of compatible solutes by (hyper)thermophilic organisms

The accumulation of compatible solutes seems to be a very successful strategy adopted by a large variety of organisms to cope with osmotic stress. However, the role of these organic compounds goes beyond osmoprotection. During the last decades, the discovery of (hyper)thermophilic organisms led to the identification of novel compatible solutes, highly restricted to microorganisms from hot environments. Interestingly, some of these solutes are accumulated in hyperthermophilic organisms in response to supraoptimal temperatures (Santos and da Costa 2002, Santos et al. 2011), suggesting a thermoprotective role. In fact, this early speculation has been supported by many *in vitro* studies, which demonstrated the ability of compatible solutes to protect proteins against heat denaturation (Shima et al. 1998, Ramos et al. 1997, Borges et al. 2002, Faria et al. 2008). For instance, di-*myo*-inositol-phosphate and

mannosylglycerate, two of the most widespread organic solute in (hyper)thermophiles, increased the thermal stability of glyceraldehydes-3-phosphate dehydrogenase from *P. woesei* and lactate dehydrogenase from rabbit muscle, respectively (Scholz et al. 1992, Borges et al. 2002). However, contradictory results can be found in the literature concerning the thermostabilizing properties of di-*myo*-inositol-phosphate. This solute had no effect on the thermostability of hydrogenase and ferredoxin oxidoreductase from the bacterium *T. maritima*, and a negative effect on lactate dehydrogenase from rabbit muscle (Ramakrishnan et al. 1997, Borges et al. 2002).

It is interesting to note that compatible solutes from hyperthermophilic bacteria and archaea are negatively charged (Fig. 1.2), in opposition to those found in mesophilic organisms, which are generally neutral or zwitterionic at physiological pH (Santos and da Costa 2002). This anionic feature is usually conferred by the presence of a carboxylate, phosphate, or sulfate group. Again, *in vitro* studies using model proteins showed the superior ability of negatively charged solutes to act as thermo-stabilizers (Faria et al. 2003, Faria et al. 2008). Faria and co-workers assessed the effect of compatible solutes either neutral (e.g. trehalose, glycerol, ectoine) or negatively charged (di-*myo*-inositol-phosphate, diglycerol phosphate, mannosylglycerate) on the thermostability of mitochondrial malate dehydrogenase from pig heart, recombinant staphylococcal nuclease A, and hen-egg lysozyme (Faria et al. 2008). The authors concluded that negatively charged compounds are superior thermo-stabilizers, and also

reinforced the concept that the mechanism underlying protein stabilization by negatively charged compounds may be dependent on specific solute/enzyme interaction.

Normally, (hyper)thermophilic organisms show a response pattern for solute accumulation, i.e., a specific solute tends to respond to a specific stress response. For example, diglycerol phosphate, mannosylglycerate, aspartate, and glutamate generally accumulate in response to an increase in the salinity of the medium, while the pools of di-*myo*-inositol-phosphate and derivatives are higher under heat stress conditions. However, some exceptions to this rule have been detected. For example, when *P. ferrophilus* is challenged with supraoptimal temperatures, mannosylglycerate and glutamate replace the missing di-*myo*-inositol-phosphate (Neves et al. 2005). The composition of the growth medium can also influence the pattern of solute accumulation in hyperthermophiles. This situation has been reported for *A. fulgidus* strain 7324 and for *Thermococcus litoralis*. Under osmotic stress conditions, the former is able to accumulate primarily diglycerol phosphate, but only when lactate is used as carbon source. When lactate is replaced by starch, mannosylglycerate accumulates preferentially (Gonçalves 2008). In the case of *T. litoralis*, trehalose accumulates when present in the growth medium. The same is observed for the compatible solutes hydroxyproline and β-galactopyranosyl-5-hydroxylysine (Lamosa et al. 1998).

The import of compatible solutes from the external medium rather than *de novo* synthesis is energetically favorable for organisms (Oren

1999). However, the transport systems in hyperthermophilic organisms are poorly investigated in comparison to mesophilic organisms, where the transport of proline, glycine betaine, or ectoine has been well characterized (Lai et al. 2000, Bremer and Krämer 2000, Roeßler and Müller 2001). In contrast, the biosynthetic pathways of many compatible solutes found in (hyper)thermophiles have been described, namely mannosylglycerate, di-*myo*-inositol-phosphate, diglycerol phosphate, glucosylglycerate, among others (reviewed by Santos et al. 2011).

Despite the availability of many *in vitro* studies reinforcing the role of compatible solutes in thermoprotection, a definite prove is still missing, mainly due to the lack of genetic tools to manipulate hyperthermophiles. Recently, the availability of a genetic system to manipulate *T. kodakarensis* allowed the construction of a mutant unable to accumulate DIP. However, results were inconclusive since the mutant strain was able to substitute DIP by aspartate under heat stress conditions, with similar efficacy (Borges et al. 2010). Therefore, the physiological role of compatible solutes in hyperthermophilic organisms is still unclear.

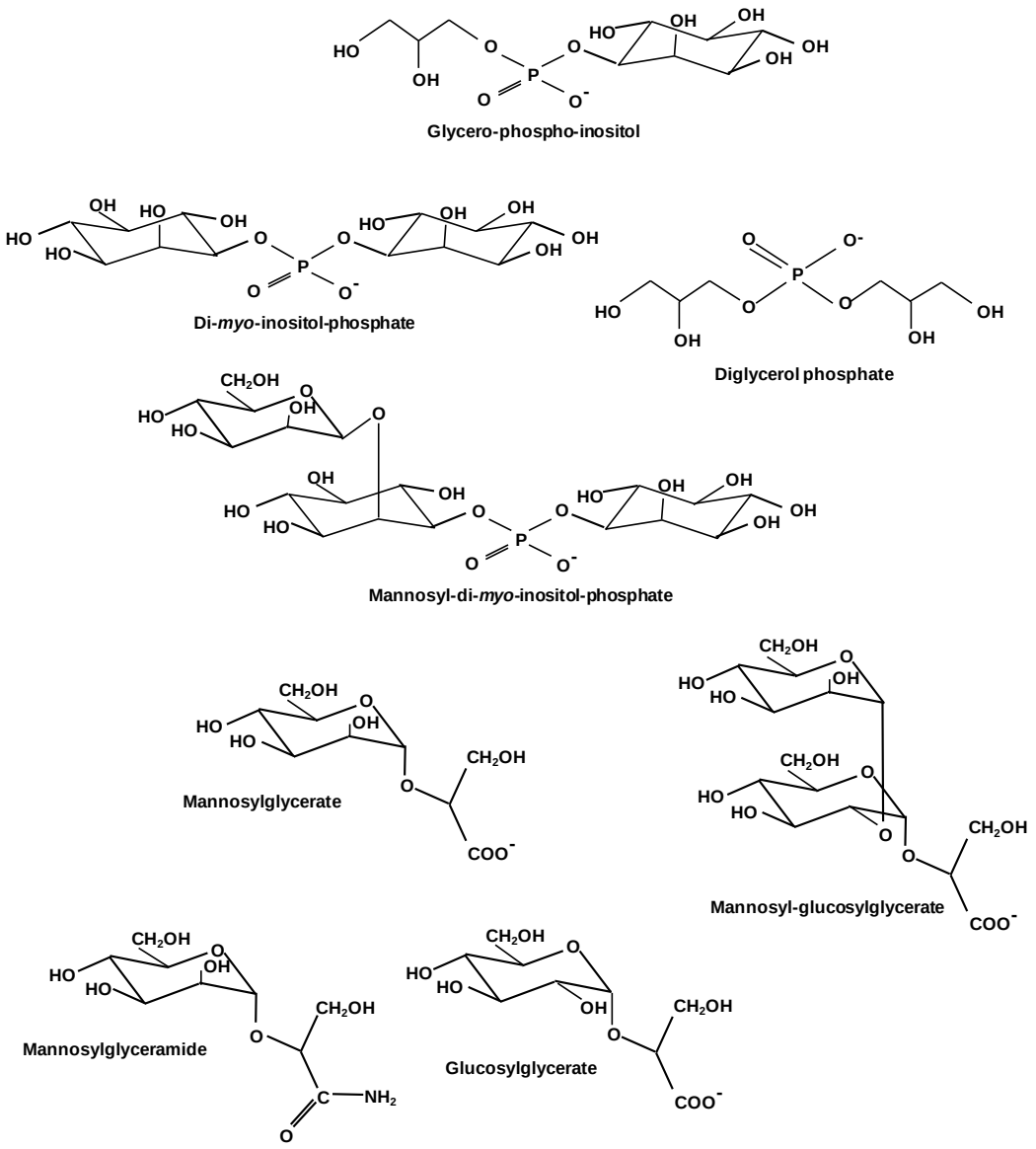


FIGURE 1.2 Structures of compatible solutes typically found in (hyper)thermophilic organisms.

Biosynthesis of di-*myo*-inositol-phosphate

In 1998, two distinct pathways for di-*myo*-inositol-phosphate biosynthesis were published. Chen and co-workers investigated the crude cell extract of *Methanotorris igneus* for enzymatic activities and proposed a two-step biosynthetic pathway, where L-*myo*-inositol-1-phosphate is activated by CTP and the resulting CDP-L-*myo*-inositol reacts with L-*myo*-inositol to yield di-*myo*-inositol-phosphate. However, the authors were not able to prove the existence of CDP-inositol (Chen et al. 1998). The second proposed pathway for di-*myo*-inositol-phosphate synthesis was investigated in *Pyrococcus woesei* by Scholz and colleagues. They proposed a single-step pathway in which two molecules of L-*myo*-inositol-1-phosphate are condensed at the expense of an NTP (Scholz et al. 1998). Both studies confirmed the activity of L-*myo*-inositol-1-phosphate synthase, the enzyme that catalyzes the formation of L-*myo*-inositol-1-phosphate from glucose-6-phosphate. In 2006, our team investigated the biosynthesis of di-*myo*-inositol-phosphate in *Archaeoglobus fulgidus* and firmly established a two-step pathway: glucose-6-phosphate is converted into L-*myo*-inositol-1-phosphate by the action of L-*myo*-inositol-1-phosphate synthase (MIPS). This activity had been reported previously (Chen et al. 1998, Scholz et al. 1998). Then, L-*myo*-inositol-1-phosphate is activated by CTP to produce the intermediate CDP-L-*myo*-inositol, which is then condensed with another molecule of L-*myo*-inositol-1-phosphate to yield the phosphorylated form of di-*myo*-inositol-phosphate. The last step is catalyzed by a phosphatase, which dephosphorylates di-*myo*-inositol-phosphate-phosphate into di-*myo*-inositol-phosphate (Fig. 1.3) (Borges et

al. 2006). This pathway was further validated with the identification of the genes encoding CTP:inositol-1-phosphate cytidylyltransferase (IPCT), the enzyme that catalyzes the condensation of CTP and L-*myo*-inosito-1-phosphate into CDP-L-*myo*-inositol, and CDP-inositol:inositol-1-phosphate transferase (DIPPS, also known as DIPP synthase) that produces the phosphorylated intermediate of DIP from CDP-L-*myo*-inositol and L-*myo*-inositol-1-phosphate (Rodionov et al. 2007, Rodrigues et al. 2007). In most organisms accumulating di-*myo*-inositol-phosphate, the two activities (cytidylyltransferase and DIPP synthase) are fused in a single gene product, but separate genes can also be found in a few organisms, namely in *Aeropyrum pernix*, *Hyperthermus butylicus*, *Pyrolobus fumarii* and several *Thermotoga* spp. (Rodionov et al. 2007, Rodrigues et al. 2007, Gonçalves et al. 2012).

Activity assays showed that the DIPPS domain is specific for L-*myo*-inositol-1-phosphate, but it has a certain flexibility regarding the CDP-alcohol molecule since it recognizes either CDP-inositol or CDP-glycerol, producing the phosphorylated forms of di-*myo*-inositol-phosphate or glycerol-phospho-inositol, respectively (Rodrigues et al. 2007). The sequence analysis of the DIPPS domain showed a conserved motif that is characteristic of the CDP-alcohol phosphatidyltransferase family (Rodrigues et al. 2007). Enzymes of this family are usually located in the membrane and are responsible for the displacement of CMP from a CDP-alcohol to a second alcohol with the formation of a phosphodiester bond and concomitant hydrolysis of pyrophosphate. In agreement, hydropathic

profiles predicted the presence of three transmembrane helices in the DIPPS domain. Recently, the first three dimensional structure of an IPCT was resolved by X-ray (Brito et. al. 2011).

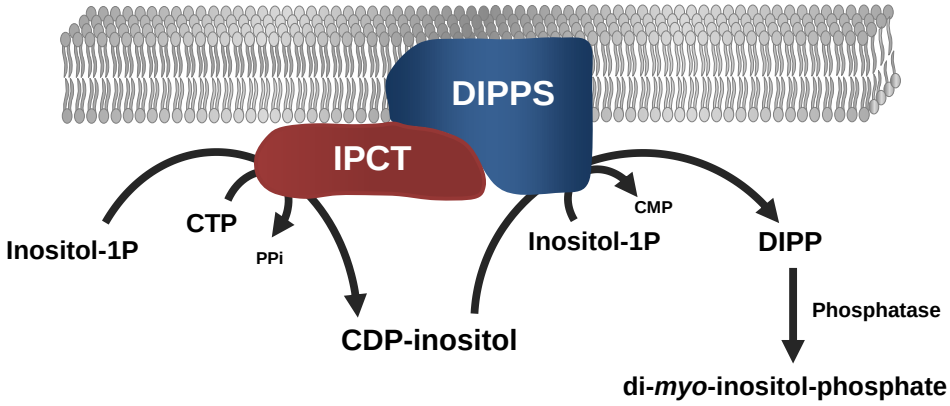


FIGURE 1.3 Schematic representation of di-*myo*-inositol-phosphate biosynthesis. The first step is catalyzed by IPCT (CTP:inositol-1-phosphate cytidylyltransferase), which condensates a molecule of L-*myo*-inositol-1-phosphate with CTP to yield the intermediate CDP-L-*myo*-inositol. In a second step, DIPPS (CDP-inositol:inositol-1-phosphate transferase) is responsible for the condensation of CDP-L-*myo*-inositol with L-*myo*-inositol-1-phosphate, producing a phosphorylated form of di-*myo*-inositol-phosphate (DIPP). In many organisms, both activities are present in a single polypeptide chain, which was predicted to have a cytosolic domain (IPCT) and a transmembrane domain (DIPPS).

The team of M. Roberts proposed that the last step of di-*myo*-inositol-phosphate is catalyzed by an inositol monophosphatase-like phosphatase (IMPase) (Rodionov et al. 2007).

Biosynthesis of mannosylglycerate

The first reports on the occurrence of mannosylglycerate in thermophilic prokaryotes were published in 1995 (Nunes et al. 1995, Martins and Santos 1995). This solute was identified in cell extracts of two thermophilic bacteria, *R. marinus* and *T. thermophilus*, and in one hyperthermophilic archaeon, *P. furiosus*. Later, the genes involved in mannosylglycerate biosynthesis were identified in *R. marinus* and *P. horikoshii*, and two alternative pathways were established (Martins et al. 1999, Empadinhas et al. 2001): a single-step pathway, where the enzyme mannosylglycerate synthase (MGS) catalyzes the transference of the mannosyl moiety from GDP-mannose to D-glycerate, producing mannosylglycerate; and a two-step pathway, in which the mannosyl moiety is transferred from GDP-mannose to 3-phosphoglycerate by the action of mannosyl-3-phosphoglycerate synthase (MPGS) yielding a phosphorylated intermediate, mannosyl-3-phosphoglycerate. The intermediate is further dephosphorylated to mannosylglycerate in a reaction catalyzed by a phosphatase (MPGP) (Fig. 1.4) (Santos et al. 2011, Borges et al. 2014).

The single-step pathway is almost exclusive to algae and early land plants, and thus far, it has been only identified in the thermophilic bacterium *R. marinus* (Martins et al. 1999), in the mesophilic red alga *Caloglossa leprieurii* (Borges et al. 2014), and in the plant *Selaginella moellendorffii* (Nobre et al. 2013). Despite the presence of the biosynthetic genes, the accumulation of mannosylglycerate was never detected in *S. moellendorffii*.

In contrast, the two-step pathway is present in many organisms accumulating mannosylglycerate, namely in *R. marinus*, *T. thermophilus*, *R. xylanophilus* and in archaea from the order Thermococcales (Martins et al. 1999, Empadinhas et al. 2003, Neves et al. 2005, Borges et al. 2014). To date, *Dehalococcoides ethenogenes* is the only organism known to have a single gene coding for MPGS and MPGP. The other organisms have separate genes for the synthase and phosphatase activities (for a review see Borges et al. 2014). Currently, two structures of the MPGS of the archaeon *P. horikoshii* and the bacterium *T. thermophilus* are available (Kawamura et al. unpublished, Gonçalves et al. 2010) and the two structures are highly similar (P*ho*MPGS, PDB 2ZU7; *Tth*MPGS, PDB 2WVM). Representatives of MPGP from the same species were also structurally characterized (Kawamura et al. 2008, Gonçalves et al. 2011).

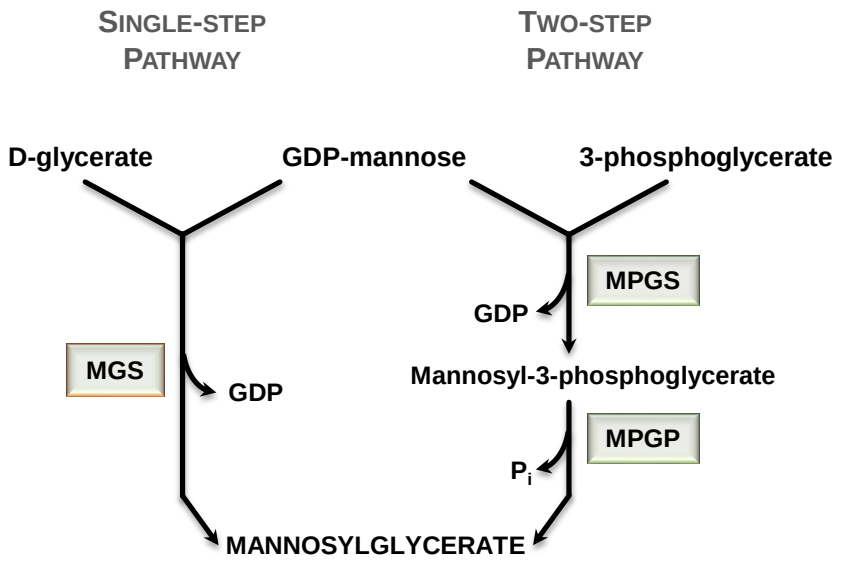


FIGURE 1.4 Pathways for mannosylglycerate biosynthesis. In the single-step pathway, the enzyme mannosylglycerate synthase (MGS) catalyzes the condensation of D-glycerate and GDP-mannose to produce mannosylglycerate, while in the two-step pathway mannosyl-3-phosphoglycerate synthase (MPGS) and mannosyl-3-phosphoglycerate phosphatase (MPGP) catalyze the formation and dephosphorylation of the intermediate mannosyl-3-phosphoglycerate, respectively.

The hyperthermophilic archaeon *Pyrococcus furiosus*

Isolation and morphology

Pyrococcus furiosus is a strictly anaerobic archaeon that grows optimally at temperatures near the boiling point of water. It was first isolated in 1986 by Stetter and co-workers from a shallow marine solfatara at Vulcano Island in

southern Italy (Fiala and Stetter 1986) and is one of the most extensively studied hyperthermophiles since it grows rapidly achieving good growth yields (generation times of 35 minutes and cell density higher than 10^8 cells/ml) (Fiala and Stetter 1986, Raven et al. 1992). Cells are motile flagellated cocci of 0.8 to 2.5 μm in width and occur often in pairs (Fig. 1.5). Each cell contains approximately 50 monopolar polytrichous flagella, each measuring approximately 7 nm in width and 7 μm in length (Fiala and Stetter 1986).

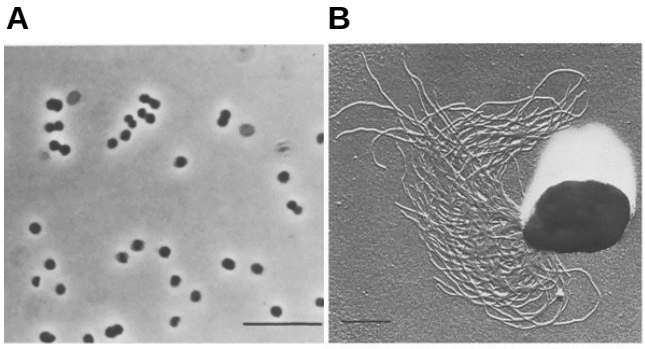


Figure 1.5 A, Phase contrast microscopy of cells of *Pyrococcus furiosus*. Bar corresponds to 10 μm. **B**, Electron microscopy of *P.* furiosus, showing a monopolar polytricous flagellation. Bar corresponds to 0.5 μm. Adapted from Fiala and Stetter 1986.

Growth conditions

The hyperthermophilic archaeon *P. furiosus* grows at temperatures ranging from 70 to 105ºC with an optimum around 95-100ºC (Fiala and Stetter 1986, Martins and Santos 1995). Growth was reported at pH values

between 5 and 9 with optimal growth at pH 7. It possesses a chemoorganoheterotrophic metabolism, conserving energy by fermentation of either peptides or complex sugars such as maltose, cellobiose, starch, and glycogen (Fiala and Stetter 1986, Schönheit and Schäfer 1995). Although ^{14}C and ^{13}C-labeling studies have shown that *P. furiosus* is able to transport and metabolize glucose, no growth was detected when glucose or other monomeric sugars (fructose, galactose, sorbose, ribose, arabinose, and xylose) were used as carbon source (Kengen et al. 1994, Schäfer et al. 1994, Kengen et al. 1996). In opposition to most *Thermococcales*, the presence of elemental sulfur (S^o) during *P. furiosus* growth is facultative and, when provided, it is utilized as the terminal electron acceptor, leading to the formation of H_2S (Fiala and Stetter 1986, Adams et al. 2001). Several studies demonstrated that growth performed with elemental sulfur resulted in higher final cell density (Fiala and Stetter 1986, Adams et al. 2001, Chou et al. 2007). Fermentation occurs through a modified Embden-Meyerhof (EM) pathway and the final products are H_2 or H_2S (absence or presence of S^o, respectively), acetate, alanine and CO_2.

Carbohydrate uptake

Analysis of the completed genome sequence of several members of Archaea, including *P. furiosus*, demonstrated that homologues of the phosphoenolpyruvate-dependent phosphotransferase system (PTS), typically found in Bacteria are absent in these organisms (Koning et al. 2001). Instead, Archaea utilize binding protein-dependent ATP-binding

cassette (ABC) transporters for sugar uptake. In particular, *P. furiosus* possesses ABC transporters specific for trehalose/maltose (TM system, PF1739 to PF1741), maltodextrins (MD system, PF1933 and PF1936 to PF1938), and cellobiose/β-glucosides uptake (Lee et al. 2006, Koning et al. 2001). The TM and MD systems are homologous to the *mal* transporter that transports maltose in *E. coli*. Therefore, the TM system is also known as Mal-I transporter while the MD system can also be called as Mal-II transporter (Lee et al. 2006). Both transporters are regulated by the transcriptional repressor TrmB. Inhibition of Mal-I is released by the presence of maltose and trehalose, while sucrose and maltodextrin relieve the repression of Mal-II transporter (Lee et al. 2003). Transcriptional studies using *P. furiosus* cells grown on maltose or starch showed the upregulation of PF1935, a gene encoding an extracellular amylopullulanase. Therefore, this enzyme was suggested to act as the major starch-hydrolyzing enzyme in *P. furiosus* generating maltodextrin and maltose. Also, two transmembrane permeases (PF1748 and PF1749) adjacent to the Mal-I transporter operon were upregulated, suggesting their involvement in maltose and starch assimilation (Lee et al. 2006).

Carbohydrate metabolism

The Embden-Meyerhof (EM) pathway is the most common route for glucose degradation in all domains of life. However, alternative pathways such as the Entner-Doudoroff (ED) pathway can be found in several prokaryotes. ^{13}C NMR studies combined with enzyme activity

measurements in crude cell extracts revealed that *P. furiosus* degrade glucose exclusively via a modified Embden-Meyerhof (EM) pathway (Kengen et al. 1994). Although catalyzing the same conversions, the glycolytic enzymes in *P. furiosus* differ from those found in classical EM pathway. For instance, the key-enzymes glucokinase and the phosphofructokinase are dependent on ADP rather than ATP and produce AMP instead of ADP (Fig. 1.6) (Kengen et al. 1994). An even more remarkable difference is the glycolytic conversion of gyceraldehyde-3-phosphate (GAP). In Bacteria and Eukarya, the enzymes gyceraldehyde-3-phosphate dehydrogenase (GAPDH) and 3-phosphoglycerate kinase (PGK) catalyze the two-step conversion of GAP to 3-phosphoglycerate (3PG) with the formation of the intermediate 1,3-bisphosphoglycerate (BPG) and ATP. In contrast, *P. furiosus* converts GAP to 3PG through a direct and irreversible oxidation, omitting the formation of the intermediate 1,3-bisphosphoglycerate and ATP. This reaction can be catalyzed either by $NAD(P)^+$-dependent nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) or by ferredoxin (Fd)-dependent glyceraldehyde-3-phosphate oxidoreductase (GAPOR) (Kengen et al. 1994, van der Oost et al. 1998). Although the first enzyme recognizes both NAD^+ and $NADP^+$, it possesses a clear preference for $NADP^+$. GAPOR is a unique tungsten-containing enzyme and it has no analog in mesophilic archaea or bacteria (Mukund and Adams 1995). With the exception of GAPOR and GAPN, the enzymes involved in the lower part of glycolysis are homologous to their classical counterparts in Bacteria and Eukarya (for a review see Bräsen et

al. 2014). The conversion of phosphoenolpyruvate (PEP) to pyruvate is commonly catalyzed by pyruvate kinase (PK). However, transcriptional studies using *P. furiosus* grown on maltose indicated an increase in the concentration of mRNA encoding PEP synthase, an enzyme involved in the conversion of pyruvate to PEP during gluconeogenesis (Robinson et al. 1994, Robinson and Schreier 1994). Therefore, it was suggested that PEP synthase not only played a role in gluconeogenesis but also in glycolysis by converting PEP to pyruvate. The deletion of PEP synthase in *T. kodakarensis* resulted in a strain unable to grow under fermentation conditions, emphasizing the suggestion that PEP synthase is essential for glycolysis (Imanaka et al. 2006). Controversially, enzymatic characterization of PEP synthase from *P. furiosus* showed a clear preference for the gluconeogenic direction leading to PEP formation (Hutchins et al. 2001).

Finally, pyruvate is oxidatively decarboxylated in a ferredoxin-dependent reaction catalyzed by pyruvate ferredoxin-oxidoreductase (POR) to CO_2 and acetyl-CoA. The conversion of acetyl-CoA to acetate occurs via an ADP-dependent acetyl-CoA synthetase with production of ATP.

Nowadays, GAPOR and GAPDH are the only glycolytic enzymes known to be regulated. Previous studies showed an increase in GAPOR and a decrease in GAPDH activities when *P. furiosus* was grown on cellobiose relative to pyruvate. Expression of the gene encoding GAPOR is

regulated at transcript level while GAPDH appears to be regulated post-transcriptionally (van der Oost et al. 1998).

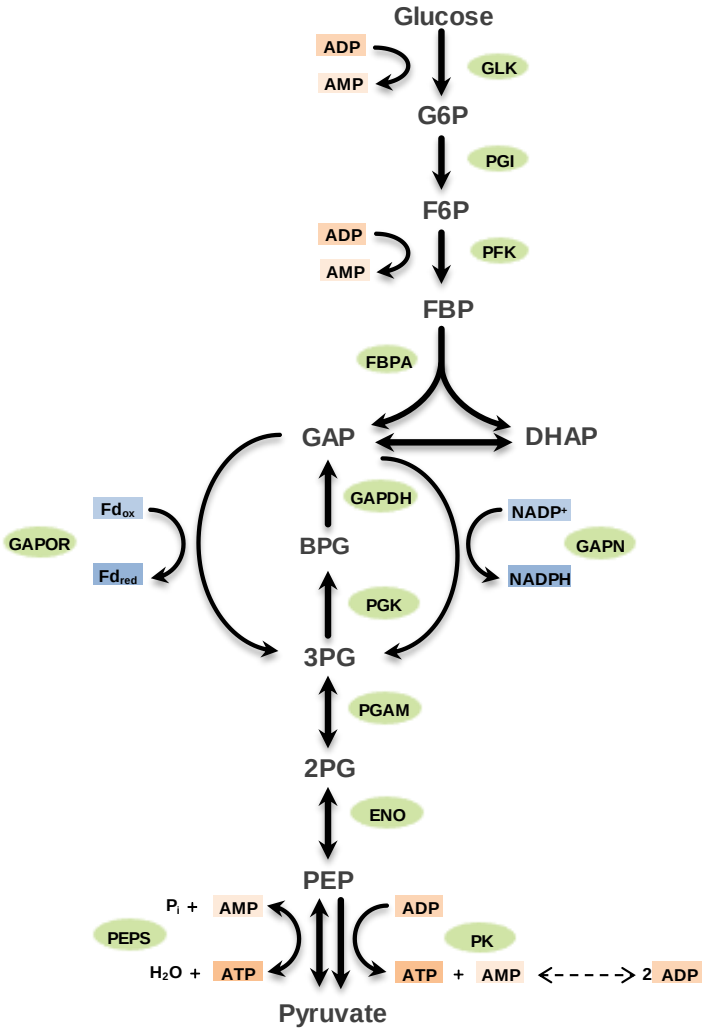


Figure 1.6 Glucose consumption in the anaerobic hyperthermophilic archaeon *Pyrococcus furiosus* through a modified Embden-Meyerhof pathway for glycolysis. Most Archaea employ the enzymes PGK and GAPDH for gluconeogenesis

(Bräsen et al. 2014). ENO, enolase; FBPA, fructose-1,6-bisphosphate aldolase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GAPN, nonphosphorylating GAPDH; GAPOR, GAP:Fd oxidoreductase; GLK, glucose kinase; PEPS, PEP synthetase; PFK, phosphofructokinase; PGI, phosphoglucose isomerase; PGAM, phosphoglycerate mutase PK, pyruvate kinase; G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; FBP, fructose-1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde-3-phosphate; BPG, 1,3-bisphosphoglycerate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; Fd, ferredoxin.

Energy conservation

Pyrococcus furiosus obtains energy by substrate-level phosphorylation and the amount of ATP formed during degradation of glucose to pyruvate through a modified Embden-Meyerhof pathway is still unclear (Schut et al. 2013, Bräsen et al. 2014). The amount of ATP formed per molecule of glucose may be either 0 or 2 molecules depending on the enzyme catalyzing the conversion of PEP to pyruvate (Fig. 1.6). The usage of pyruvate kinase results in no net gain of energy since 4 ADPs are consumed by the two ADP-dependent sugars kinases (upper part of glycolysis) and pyruvate kinase while the same number of ADP is regenerated from 2 AMPs (ADP-dependent sugar kinases) and 2 ATPs (pyruvate kinase) via the adenylate kinase (Bräsen et al. 2014). In the case of PEP synthase, 2 ADPs are consumed by the ADP-dependent sugar kinases (upper part of glycolysis) and the 2 AMPs produced are used by PEP synthase, leading to the formation of 2 ATPs. In addition, 2 ATPs molecules are conserved during the conversion of pyruvate to acetate. The reason why *P. furiosus* relies on ADP rather than ATP as phosphoryl donor

has been discussed. From an energetic point of view, ADP and ATP are similar since the energy released by ADP and ATP hydrolysis is comparable (Lehninger 1975). The hypothesis that ADP could be more stable at higher temperatures was raised (Kengen et al. 1994) and, in fact, the half-life of ADP and ATP at 90°C in the presence of $MgCl_2$ is 750 and 115 minutes, respectively (Kengen et al. 1996). However, *in vivo* this argument loses importance since the turnover rate for these molecules is in the order of seconds.

The degradation of glucose to acetate generates reducing equivalents at two levels: during the oxidation of GAP to 3PG via GAPOR; and during the oxidation of pyruvate to acetyl-CoA via POR. Ferredoxin in place of the expected NAD(P) is used as the electron acceptor. Subsequently, *P. furiosus* is able to transfer reducing equivalents from ferredoxin to either protons or elemental sulfur (if present in the medium) to produce H_2 or H_2S, respectively. Mbh, the enzyme responsible for H_2 production in *P. furiosus*, is a membrane-bound hydrogenase complex encoded by a 14-gene cluster. This protein complex has been extensively studied by Adams and co-workers (Sapra et al. 2000, Sapra et al. 2003, Schut et al, 2013; McTernan et al. 2014). Mbh is a unique hydrogenase since it uses the energy generated by H_2 production to translocate either H^+ or Na^+ across the membrane and ultimately synthesize ATP via oxidative phosphorylation. Therefore, the use of ferredoxin as electron acceptor instead of NAD(P) is advantageous for the organism since reduced ferredoxin can be used directly for ATP synthesis through

production of H_2 while NAD(P)H cannot because it has a higher redox potential than the H_2/H^+ couple (Sapra et al. 2003).

The pentose phosphate pathway (PPP) which is utilized by Bacteria and Eukarya to generate reducing agents (NADPH) and precursors for nucleotide and aromatic amino acids biosynthesis is usually divided into two phases: the oxidative PPP, where glucose-6-phosphate undergoes oxidative decarboxylation to form ribulose-5-phosphate plus reduced NADPH; and the non-oxidative PPP, where an enzyme with isomerase activity converts ribulose-5-phosphate to ribose-5-phosphate. In parallel, the excess of ribulose-5-phosphate is converted into fructose-6-phosphate and glyceraldehyde-3-phosphate, with concomitant production of erythrose-4-phosphate, the precursor for aromatic amino acids (Soderberg 2005). However, in Archaea the oxidative PPP seems to be absent while the non-oxidative PPP is incomplete in most organisms (Yu et al. 1994, Soderberg 2005, Bräsen et al. 2014). Archaea losing parts or even the complete non-oxidative PPP synthesize ribulose-5-phosphate through the reverse ribulose monophosphate pathway (Soderberg 2005, Orita et al. 2006). This pathway was first identified in Bacteria as a formaldehyde fixation pathway, where ribulose-5-phosphate and formaldehyde are reversibly converted to D-arabino-3-hexulose-6-phosphate, which is further isomerized to fructose-6-phosphate. Like in *T. kodakarensis*, genes encoding two of the four enzymes involved in the non-oxidative PPP (transaldolase and ribulose-5-phosphate epimerase) are missing in *P.*

furiosus, while the activities for ribulose monophosphate pathway are present (Orita et al. 2006) (Fig. 1.7).

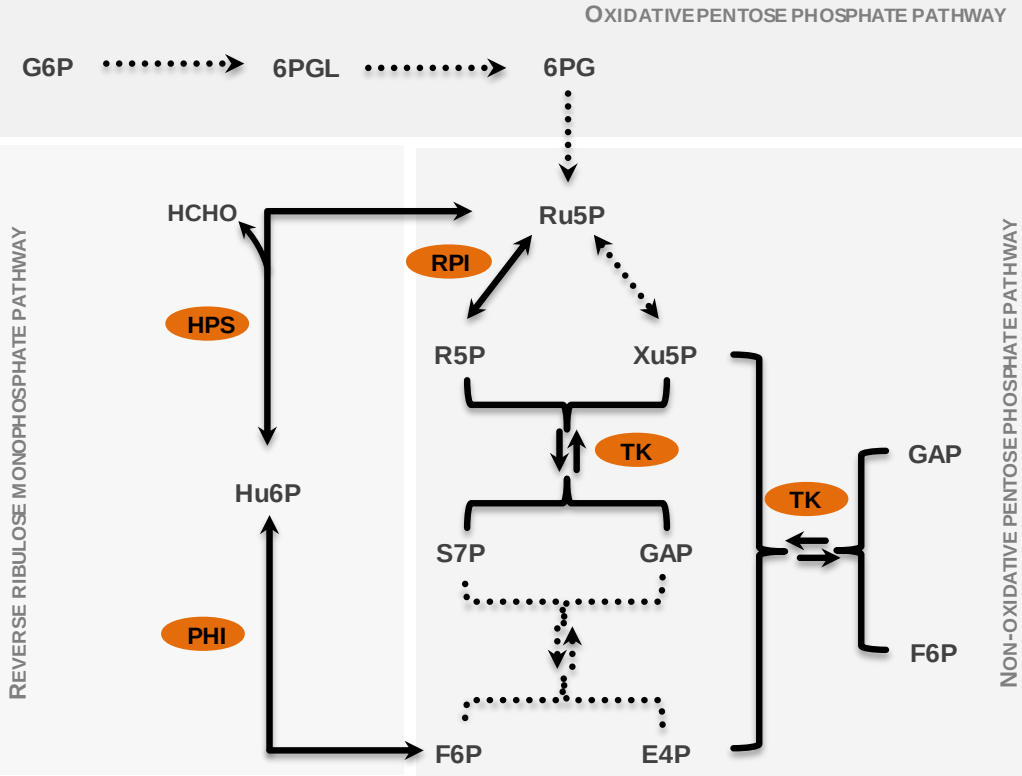


FIGURE 1.7 Proposed pentose phosphate pathway in *P. furiosus*. Depicted are the reverse ribulose monophosphate pathway, the oxidative and the non-oxidative pentose phosphate pathway. Dashed lines display the reactions catalyzed by enzymes which are not present in *P. furiosus* genome. G6P, glucose-6-phosphate; 6PGL, 6-phosphogluconate-D-lactonase; 6PG, 6-phosphogluconate; Ru5P, ribulose-5-phosphate; R5P, ribose-5-phosphate; Xu5P, xylulose-5-phosphate; GAP, glycehaldehyde-3-phosphate; S7P, sedoheptulose-7-phosphate; F6P, fructose-6-phosphate; E4P, erythrose-4-phosphate; RPI, ribose-

5-phosphate isomerase; TK, transketolase; HPS, 3-hexulose-6-phosphate synthase; PHI, 6-phospho-3-hexuloisomerase.

Metabolism of peptides

Growth of *P. furiosus* is usually carried out on carbohydrates; however, peptides can also be used as carbon and energy source. For this reason, the metabolism of carbohydrates has received more attention during the last years. The pathway proposed for peptide catabolism involves hydrolysis into individual amino acids followed by their oxidative deamination to 2-keto acids. These CoA derivatives, together with acetyl-CoA produced via the oxidation of pyruvate during glycolysis are transformed to their corresponding organic acids by the action of two acetyl-CoA synthases, with concomitant formation of ATP (Adams et al. 2001, Scott et al. 2014).

Genome of *P. furiosus*

Currently, 7 species belonging to the genus *Pyrococcus* have their genome completely sequenced (two species of *P. furiosus*, *P. horikoshii*, *P. abyssi*, *Pyrococcus* sp. NA2, *Pyrococcus* sp. STO4, and *P. yayanosii*). The genome of *P. furiosus* was the first to become public, in 2002, and it is composed by approximately 1.9 million of base pairs (Robb et al. 2001). The original genome annotation contained 2,065 open reading frames (ORFs) labeled from PF0001 to PF2065 and, from those, 1,041 were annotated as encoding proteins with known function. In addition, 869 ORFs were designated as encoding conserved proteins, while the remaining 155

were considered unique to *P. furiosus*. With the increasing number of sequenced genomes, 25 of the 155 unique ORFs were reclassified and are no longer unique to the species *P. furiosus* (Poole et al 2005). The availability of a genome sequence triggered the characterization of *P. furiosus* at several “omics” levels, including transcriptomics (Schut et al. 2003, Weinberg et al. 2005, Lee et al. 2006, Strand et al. 2010), comparative genomics (Zivanovic et al. 2002), proteomics (Lancaster et al. 2011, Menon et al. 2009) and structural genomics (Jenney and Adams 2008). At transcript level, a DNA microarray approach was used to study the effect of cold shock (Weinberg et al. 2005), oxidative stress (Strand et al. 2010), and growth on carbohydrates and peptides (Schut et al. 2003, Lee et al. 2006) on gene expression.

The development of genetic tools for hyperthermophilic organisms presents many challenges given the extreme conditions where these organisms usually proliferate. For instance, antibiotic selection usually applied in bacteria is ineffective due to the low thermostability of drugs and the heterologously expressed resistance proteins (Leigh et al. 2011, Lipscomb et al. 2011). To date, simvastatin is one of the few drugs successfully applied in archaea (Lam and Doolittle 1989, Peck et al. 2000, Matsumi et al. 2007). This drug is a competitive inhibitor of the 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase, which is involved in the biosynthesis of isoprenoides, the main component of the archaeal membrane lipids (Lipscomb et al. 2011). Selection based on nutritional markers may substitute the use of antibiotics. Selection based on uracil

auxotrophy is the most useful since they can be counter-selected using 5-fluoroorotic acid (5-FOA) (for useful reviews see Allers and Mevarech 2005, Leigh et al. 2011).

Recently, a variant of the *P. furiosus* NCBI strain (DSM 3638) was isolated. The variant strain had the remarkable ability of being naturally competent and highly efficient for DNA uptake. In comparison to Bacteria, where transformation via natural competence was already reported in over 60 species, only 3 archaeal species exhibit this feature, *T. kodakarensis* being one of them (Johnsborg et al. 2007, Sato et al. 2003). Therefore, Lipscomb and co-workers took advantage from the natural competence of the variant strain and the availability of a shuttle vector system for *P. furiosus* based on simvastatin resistance to generate a genetically manipulated strain by target gene disruption of the *pyrF* gene. The resulting uracil auxotrophic strain was designated as *P. furiosus* COM1 and its genome was recently sequenced (Bridger et al. 2012). The natural competence of the variant strain was probably conferred by the disruption of five genes coding for CRISPR-associated proteins (Cas), a mechanism used by several bacteria and archaea to prevent invasion by foreign nucleic acids (Makarova et al. 2011). Genetic tools to manipulate *P. furiosus* were already applied on the deletion (Bridger et al. 2011, Thorgersen et al. 2014), insertion (Basen et al. 2012) and over-expression (McTernan et al. 2014) of several genes, allowing their physiological characterization.

Heat shock response

Like many hyperthermophilic archaea, the heat shock response of *P. furiosus* has been poorly characterized. The first study about the heat stress response of *P. furiosus* was published in 1995 by Martins and Santos (1995) and focused on the accumulation of intracellular solutes. The authors analyzed cell extracts of *P. furiosus* grown under optimal and stressful conditions by ^{1}H and ^{13}C NMR and reported a positive correlation between di-*myo*-inositol-phosphate accumulation and temperature increase. Under optimal growth conditions (95°C and 2.8% NaCl, according to the growth rates reported on the same study), mannosylglycerate and di-*myo*-inositol-phosphate were the predominant solutes detected in cell extracts of *P. furiosus*, while at higher temperatures (100 and 101°C), di-*myo*-inositol-phosphate was the sole solute reported. Its pool increased approximately 20-fold from the optimum 95°C to 101°C. Conversely, mannosylglycerate was the primary solute accumulated under osmotic stress (Martins and Santos 1995).

In 2003, Shockley and co-workers analyzed the effect of heat shock on the transcription of 201 genes using a targeted cDNA microarray in conjunction with northern blot. Results revealed the differential expression of several genes, namely those involved in protein rescue (molecular chaperones), proteolysis, and stabilization (synthesis of compatible solutes) (Shockley et al. 2003). The genes encoding the thermosome (HSP60-like protein), a small heat shock protein, and two molecular chaperones belonging to the CDC48/p97 branch of the AAA^{+} family were

strongly induced. In addition, the gene encoding L-*myo*-inositol-1-phosphate synthase, which is the enzyme catalyzing the synthesis of the precursor for di-*myo*-inositol-phosphate production was also strongly induced by heat stress. Curiously, and despite the over-expression of many genes related to carbohydrate transport, the gene encoding glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) was down-regulated at supraoptimal temperatures.

Currently, most of the information gathered about heat shock response regulation comes from bacteria and eukaryotes. Archaea lack homologues to the regulatory proteins found in those organisms, but a novel regulator of the heat shock response, called Phr, was identified in *P. furiosus* and it is conserved among Euryarchaeota. The protein acts as a transcription repressor that, under optimal growth temperatures, is bound to the promoter region of specific heat shock genes, abolishing gene transcription. With temperature up-shift, Phr releases the promoter region allowing transcription machinery to bind and initiation of transcription (Vierke et al. 2003). The three dimensional structure of Phr from *P. furiosus* was determined and it revealed a winged helix DNA binding domain with an N-terminal helical extension, similar to the bacterial SmtB/ArsR family (Liu et al. 2007). The protein exists as a dimer and it recognizes a palindromic nucleotide sequence with 29 bp in length. Currently, the mechanism underlying Phr dissociation at high temperatures remains unclear. Due to the lack of genetic tools to manipulate *P. furiosus*, Keese and collaborators (Keese et al. 2009) identified the targets of Phr by

cell-free transcription of fragmented chromosomal DNA in the presence or absence of the regulon. The results confirmed that genes encoding Phr, small heat shock protein, and AAA^+ ATPase are targets of the regulon Phr, as previously suggested by Vierke and co-workers (Vierke et al. 2003). Additionally, six open reading frames were newly identified (PF0624, PF1042, PF1291, PF1292, PF1488, and PF1616). Most of the open reading frames are annotated as hypothetical proteins while the gene PF1616 codes for the L-*myo*-inositol-1-phosphate synthase.

CHAPTER 2

Mannosylglycerate and di-*myo*-inositol-phosphate have interchangeable roles during adaptation of *Pyrococcus furiosus* to heat stress

This chapter is published in:

Esteves AM, Chandrayan SK, McTernan PM, Borges N, Adams MW, Santos H. 2014. Mannosylglycerate and di-*myo*-inositol phosphate have interchangeable roles during adaptation of *Pyrococcus furiosus* to heat stress. Applied Environmental Microbiology, 80:4226-4233.

Contents

Summary

Marine hyperthermophiles accumulate small organic compounds, known as compatible solutes, in response to supraoptimal temperature or salinity. *Pyrococcus furiosus* is a hyperthermophilic archaeon that grows optimally at temperatures near 100°C. This organism accumulates mannosylglycerate (MG) and di-*myo*-inositol-phosphate (DIP) in response to osmotic and heat stress, respectively. It has been assumed that MG and DIP are involved in cell protection; however, firm evidence for the roles of these solutes in stress adaptation is still missing, largely due to the lack of genetic tools to produce suitable mutants of hyperthermophiles. Recently, such tools were developed for *P. furiosus*, making this organism an ideal target for that purpose. In this work, genes coding for the synthases in the biosynthetic pathways of MG and DIP were deleted by double-crossover homologous recombination. The growth profiles and solute patterns of the two mutants and the parent strain were investigated under optimal growth conditions and also at a supraoptimal temperature and NaCl concentration. DIP was a suitable replacement for MG during heat stress, but substitution of MG by DIP and aspartate led to a less efficient growth under osmotic stress. The results suggest that the cascade of molecular events leading to MG synthesis is tuned for osmotic adjustment, while the machinery for induction of DIP synthesis responds to either stress agent. MG is as effective as DIP to protect cells against heat, despite the finding that DIP consistently increases in response to heat stress in the nine (hyper)thermophiles examined thus far.

Introduction

Many thermophiles and hyperthermophiles were isolated from marine geothermal areas and, accordingly, grow optimally in media containing 2 to 4% (w/v) NaCl. Like the vast majority of halophiles, these organisms accumulate compatible solutes to balance the external osmotic pressure (Santos and da Costa 2001, Neves et al. 2005, Gonçalves et al. 2008). However, the organisms from hot habitats accumulate unusual negatively charged solutes (sometimes designated thermolytes), which contrast with the neutral or zwitterionic nature of the solutes typical of mesophiles. Moreover, the intracellular content of solutes from (hyper)thermophiles increases not only with the NaCl concentration of the medium, but also with the growth temperature, suggesting that the role of thermolytes goes beyond osmoprotection (Santos et al. 2011).

Screening for new solutes in (hyper)thermophiles showed that di-*myo*-inositol-phosphate (DIP) and mannosylglycerate (MG) are the most widespread components of the solute pools in organisms adapted to grow at temperatures above 60ºC (Santos et al. 2011). Typically, heat stress conditions lead to the accumulation of DIP and DIP-derivatives, while MG increases in response to supraoptimal salinity in the growth medium; curiously, the thermophilic bacterium *Rhodothermus marinus* and the hyperthermophilic archaeon *Palaeococcus ferrophilus*, which lack DIP biosynthesis, accumulate MG to cope with heat stress (Neves et al. 2005, Borges et al. 2004). On the other hand, *Archaeoglobus fulgidus* strains unable to synthesize MG, use diglycerol phosphate as an alternative solute for osmo-adjustment (Gonçalves et al. 2003).

DIP and MG are the two major solutes in *Pyrococcus furiosus* (DSM 3638), an extensively studied hyperthermophilic archaeon, able to grow at temperatures up to 103ºC, and regarded as a model organism for investigating the molecular basis of adaptation to high temperatures (Fiala and Stetter 1986, Martins and Santos 1995, Leigh et al. 2011). Upon osmotic or heat stress, the total solute pool increased in concentration by approximately 10-fold. MG represented 70% of the total pool under salt stress, while DIP was the only solute accumulating at supraoptimal temperatures. Minor amounts of glutamate were used only for adjustment to low salinity (Martins and Santos 1995). Therefore, the profile of the stress response in *P. furiosus* is representative of the general trend, insofar as MG and DIP are preferentially associated with osmoprotection and thermoprotection, respectively (Neves et al. 2005, Rodrigues et al. 2009). However, a definite proof of their physiological roles is still lacking.

The present work was designed to address the following questions: do DIP and MG play a role in cell protection against stress? If so, do they have specialized activities in osmoadaptation and thermoadaptation? To answer these questions, mutants deficient in the synthesis of specific compatible solutes are required. Fortunately, the development of tools for the genetic manipulation of hyperthermophiles has advanced significantly in recent years (Leigh et al. 2011, Sato et al. 2003, Allers and Mevarech 2005). In particular, genetic tools have been developed to manipulate *P. furiosus* (Lipscomb et al. 2011). The genetic system is based on a variant of *P. furiosus* (DSM 3638), designated COM1, which is highly competent

for external DNA uptake and recombination. The genome sequence of strain COM1 showed extensive changes, however, the major metabolic features of the wild-type strain appear to be conserved (Bridger et al. 2012).

P. furiosus is an ideal target organism for investigating the physiological role of MG and DIP because these are the major components of its solute pool, they are produced by *de novo* synthesis, and, importantly, the organism does not accumulate amino acids or other biomolecules from the external medium to cope with osmotic and heat stress (Martins and Santos 1995). The preferential import of alternative solutes could complicate the interpretation of results, as reported for a DIP-deficient mutant of *Thermococcus kodakarensis* (Borges et al. 2010). The biosynthetic pathways of DIP and MG are depicted in Figure 2.1.

In this study, we deleted the genes encoding the key-enzymes in the biosynthesis of MG and DIP in *P. furiosus* COM1 and studied the impact of these disruptions on growth as well as on the pattern of solute accumulation as a means to obtain insight into the physiological role of those ionic solutes.

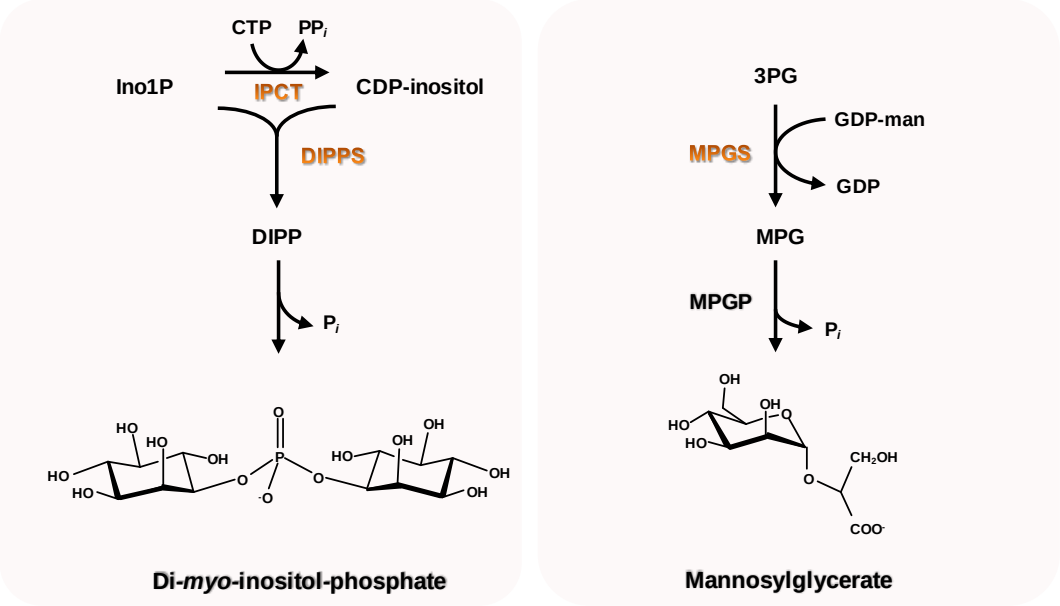


FIGURE 2.1 Biosynthetic pathways for di-*myo*-inositol-phosphate (left panel) and mannosylglycerate (right panel). Enzymes: IPCT, CTP:L-*myo*-inositol-1-phosphate cytidylyltransferase; DIPPS, di-*myo*-inositol-phosphate-phosphate synthase; the gene encoding the phosphatase activity is unknown; MPGS, mannosyl-3-phosphoglycerate synthase; and MPGP, mannosyl-3-phosphoglycerate phosphatase. Other abbreviations: DIPP, di-*myo*-inositol-phosphate-phosphate; 3PG, 3-phosphoglycerate; MPG, mannosyl-3-phosphoglycerate; GDP-man, GDP-mannose.

Materials and Methods

Strains and culture conditions

P. furiosus strain COM1, a uracil auxotrophic strain, was used as the host strain for genetic manipulation (Farkas et al. 2012). This organism was transformed following the method of Lipscomb et al. (2011).

The physiologic studies of *P. furiosus* were performed with cells cultivated in the complex medium described previously (Adams et al. 2001), with the following modifications: 0.5% (w/v) of maltose and 0.25% (w/v) of yeast extract were used as carbon sources and growth was performed without elemental sulfur. Cultures were grown in a 2 liter fermentation vessel with continuous gassing with a mixture of N_2 (80%) and CO_2 (20%), and stirring at 80 rpm. Cultures used as pre-inoculum were successively passed in fresh medium, at least 3 times before inoculation. Prior to inoculation (10% of an overnight culture), the medium was supplemented with cysteine (0.5 g/l) and Na_2S (0.5 g/l) and the pH adjusted to 6.8. During growth, the pH was maintained at 6.8 by the addition of $NaHCO_3$ 10% (w/v). Optical density (OD) at 600 nm was used to assess cell growth. The temperature of the growth medium was monitored with a thermocouple probe Fluke 80PK-22 (Fluke Corporation, WA, US). The accuracy of measurements over the range 85 – 100ºC is better than ±0.5ºC.

To study the effect of NaCl concentration on the growth of *P. furiosus* COM1 (parent strain), cells were grown at 90ºC in medium with different NaCl concentrations (1.0, 2.8, 3.6, 4.5, and 5.5% (w/v), corresponding to 0.171, 0.479, 0.616, 0.770 and 0.941 M, respectively); the effect of temperature was investigated at different temperatures, 88, 90, 93, 95, 98, and 100ºC, in medium with the optimum NaCl concentration (2.8% NaCl). To determine the effect of osmotic and heat stress on the growth parameters and solute accumulation cells were grown at optimal conditions

(2.8% NaCl, 90ºC), under osmotic (4.5% NaCl, 90ºC) and heat stress (2.8% NaCl, 98ºC).

Construction of linear fragments for gene deletions

The linear fragment, ae∆*mpgs*, was constructed to delete the gene encoding the mannosyl-3-phosphoglycerate synthase (MPGS) (pfc_02085) (Fig. 2.2). The upstream and downstream flanking regions (1.5 kbp) of the *mpgs* gene were amplified from the genomic DNA. The P_{gdh}*pyrF* marker cassette was also amplified from plasmid pGLW021. The 3'-primer of the upstream (∆mpgs_prev_UFR) and the 5' primer of the downstream (∆mpgs_pfow_DFR) fragments have 20 to 30 bases homologous to the 5'- and 3'- ends of the marker cassette, respectively. The three PCR fragments were joined by PCR using the 5'-primer of the upstream (∆mpgs_pfow_UFR) and 3'-primers of the downstream (∆mpgs_prev_DFR) fragments. This strategy was also used to construct a second linear fragment, ae∆*ipct/dipps*, to obtain the mutant deficient in DIP synthesis by deleting the gene encoding the bifunctional enzyme CTP:L-*myo*-inositol-1-phosphate cytidylyltransferase/di-*myo*-inositol-phosphate-phosphate synthase, IPCT/DIPPS (pfc_04525). The primer sequences used in this work are presented in Table 2.1.

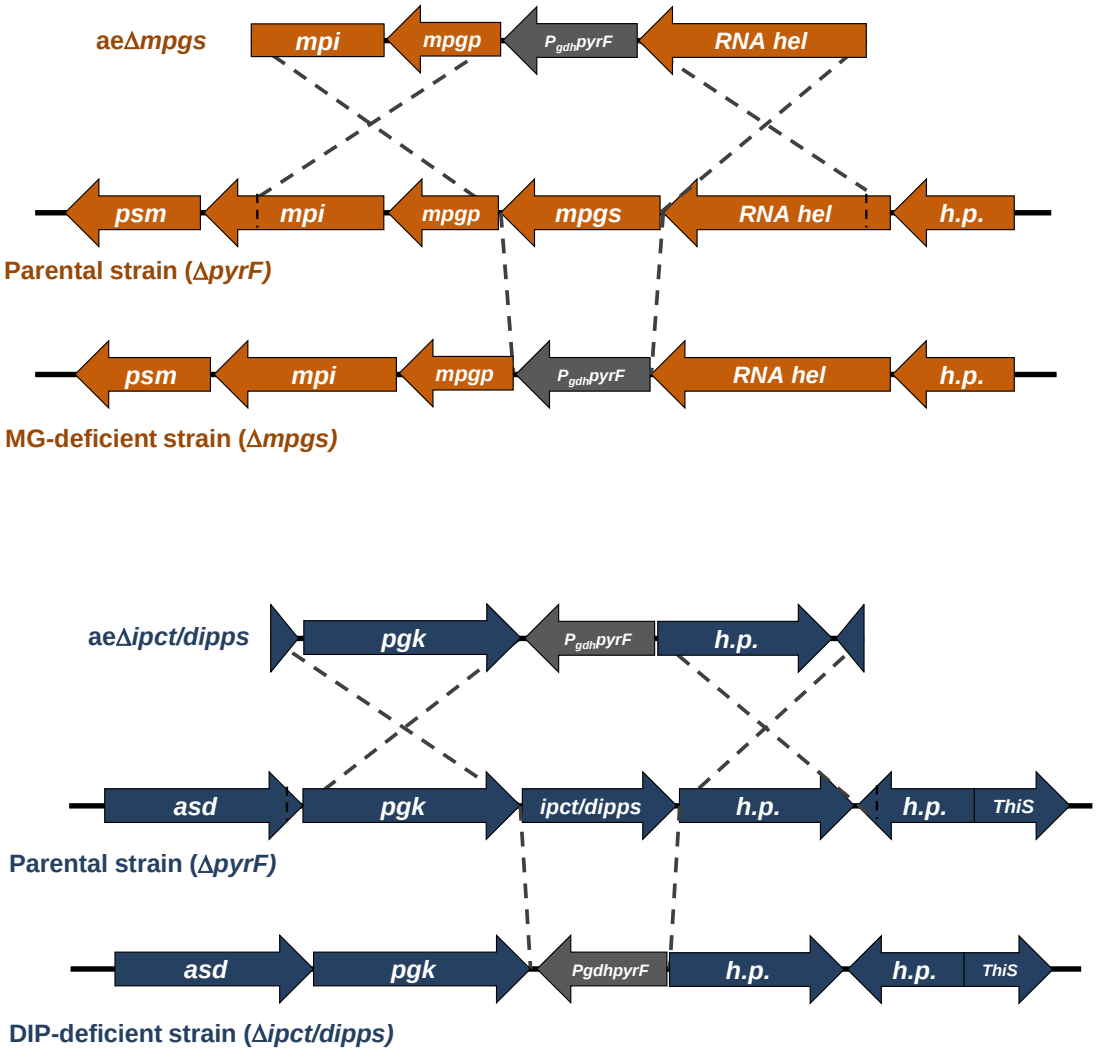


FIGURE 2.2 Genetic organization of *P. furiosus* mutants after double homologous recombination using linear fragments aeΔ*mpgs* (MG-deficient mutant) and aeΔ*ipct/dipps* (DIP-deficient mutant). *psm*, phospho-sugar mutase; *mpi*, mannose-6-phosphate isomerase; *mpgs*, mannosyl-3-phosphoglycerate synthase; *mpgp*, mannosyl-3-phosphoglycerate phosphatase; *RNA hel*, RNA helicase; *h.p.*, hypothetical protein; *asd*, aspartate-semialdehyde dehydrogenase;

pgk, phosphoglycerate kinase; *ThiS*, sulfur carrier protein; P_{gdh}, promoter region of glutamate dehydrogenase; *pyrF*, orotidyne 5'-phosphate decarboxylase.

TABLE 2.1 Primers used in this work.

Designation	Primer sequence (5'-3')
Δmpgs_pfow_UFR	TCAACCCCAACGGTTGCTGTTAACCTTTCAGTTAGAATTAGG
Δmpgs _prev_UFR	ATGGAGCTCAAGATAAAATCTTTTTTATGGATGAAAGTTGTCGAG AACAACTCCGA
Δmpgs _pfow_marker	AAAAAAGATTTTATCTTGAGCTCCATTCTTTCACCTCCTCGAAAAT CTTCTTAGCG
Δmpgs _prev_marker	CTCTTGTATATTCTTTTTGGTGGTTAAGATTGAAAATGGA GTGAGCTGAG
Δmpgs _pfow_DFR	TTAACCACCAAAAAGAATATACAAGAGGGAGTATATA
Δmpgs_prev_DFR	GGGCCGGAAGGGATCCAAATAGAATTGCAA
Δipctdipps_pfow_UFR	TAGCGACAACAAGCCCTTCAGGACAATTAG
Δipctdipps_prev_UFR	GGAGCTCAAGATAAAATCTTTTTTAGGATTCTACTTTTAAATGTCT GGTT
Δipctdipps_pfow_marker	AAAAAAGATTTTATCTTGAGCTCCATTCTTTCAC
Δipctdipps_prev_marker	CGGTGATGAGAGTTAGGTCAACCATGATTGAAAATGGAGTGAGC TGAGTT
Δipctdipps_pfow_DFR	ATGGTTGACCTAACTCTCATCACCGAAGAAGCCGT
Δipctdipps_prev_DFR	ACTATTGAGCTTCCAGAACTTGGAATAACAAT
sHSP_Pfow	ATGGTGAGGAGAATAAGAAGATGGG
sHSP_Prev	CTATTCAACTTTAACTTCGAATCCTTC
HSP60_Pfow	ATGGCCCAGTTAGCAGGCCAACCCATTC
HSP60_Prev	TCAGTCTAGATCACTGCTGAAGTCCTCGC
HSPx_Pfow	ATGTTATTATTATGGGCAATATTCTTGC
HSPx_Prev	CTAAATCTCTAAGCTTCTGTACATCTC
Phr_Pfow	ATGGGAGAGGAGCTAAACAG G
Phr_Prev	TTAAATGTTAATGTTTAGGAATTTTTCC
POR_Pfow	ATGATAGAAGTTCGCTTTCACGGTAGAGG
POR_Prev	TTAAAGCTCAAATACTTCTGTTTTTTCG
pgk1	GAGAGAATAAAGGCTATGAAACCTGGGGAA
pgk2	CTACTTTTAAATGTCTGG
mpgp1	GTCTCCAAAGCTTTCGTCGATAACTTTGAG
mpgp2	GAGTAATATTCCTGGATCTAGACAAGACTCTC

Transformation of strains

P. furiosus COM1 was transformed with ae∆*mpgs* or ae∆*ipct/dipps* by natural competence to obtain the MG-deficient mutant or the DIP-deficient mutant, respectively. Selection was performed on solid defined medium without uracil. Colonies were picked into 4 ml of defined liquid medium without uracil and incubated overnight for DNA isolation and PCR screening. Genomic DNA isolation was performed as previously described (Lipscomb et al. 2011). Briefly, cells from 1 ml of overnight *P. furiosus* culture were harvested and suspended in 100 µl of buffer A (25% sucrose, 50 mM Tris-HCl, 40 mM EDTA, pH 7.4) followed by the addition of 250 µl of 6 M guanidinium HCl–20 mM Tris, pH 8.5. The mixture was incubated at 70ºC for 5 minutes. Genomic DNA was extracted with 350 µl of phenol-chloroform-isoamyl alcohol (25:24:1; buffered at pH 8), followed by ethanol precipitation, and suspended in 50 µl of 10 mM Tris buffer pH 8.0. PCR analyses confirmed that the MG- and DIP-deficient mutants lack the *mpgs* and *ipct/dipps* genes, respectively. The positive mutants of both transformations were further purified by successive cultivation on solid medium without uracil. Additionally, each flank of the recombination region (1.7-kbp) was amplified and sequenced for each mutant. Sequence analysis revealed that the *mpgs* and *ipct/dipps* genes were correctly deleted and replaced by the *pyrF* cassette.

Southern blot analysis

Chromosomal DNA was extracted by the method previously described (Johansen and Kibenich 1992). Chromosomal DNAs of the MG and DIP-deficient mutants were digested with *Xmn*I/*Hind*III/*Bgl*II and *Hind*III/*Xmn*I/*Sac*I, respectively. The chromosomal DNA of the *P. furiosus* COM1 was also digested with both sets of restriction enzymes. The resulting DNA fragments were separated onto 1% agarose gel by electrophoresis, and transferred to a nylon membrane. The probe (500-bp), against a region close to the 3'-end of the upstream region of the target gene, was amplified by PCR and labeled following the protocol supplied by the manufacturer (ECL Direct nucleic labeling and detection system, GE Healthcare). Probe hybridization was carried out as previously described (Sambrook et al. 1989). The signals were detected with the enhanced chemiluminescence (ECL) system (GE Healthcare).

Extraction, identification, and determination of intracellular solutes

Cells were harvested at the end of exponential phase by centrifugation (6,370×*g*, 15 min) and washed twice with a solution containing all medium components except carbon sources. Cell pellets were re-suspended in water and disrupted by sonication. Total protein concentration was estimated using the Pierce BCA protein assay kit (Thermo). Intracellular solutes were extracted with boiling ethanol 80%, as described previously (Santos et al. 2006). Solvent was removed by rotary evaporation, and the residue freeze-dried. The dry residue was extracted twice with a mixture of

water-chloroform (2:1). After centrifugation, the aqueous phase was lyophilized and the residue was dissolved in $^{2}H_2O$ for ^{1}H-NMR analysis. Formate was used as concentration standard (Borges et al. 2010).

Reverse transcription-PCR (RT-PCR) experiments

The parent and DIP-deficient strains were grown under heat stress conditions (98ºC, 2.8% NaCl) as described above. Cells were harvested at the end of exponential phase and total RNA was extracted using RNeasy Mini Kit (Qiagen, Hilden, Germany). RNA samples were treated with Turbo DNase (Life Technologies™, Grand Island NY, US) to remove residual chromosomal DNA. To confirm the absence of chromosomal DNA in the RNA preparations, RNA samples were used as DNA templates for PCR using specific primers for gene pfc_04085, which encodes pyruvate ferredoxin oxidoreductase (POR), (see Table 2.1). Genomic DNA from *P. furiosus* COM1 was used as positive control. RT-PCR experiments were carried out using the OneStep RT-PCR Kit (Qiagen, Hilden, Germany). Briefly, aliquots of total RNA (60 ng) were mixed with 400 µM of each dNTP, 0.6 µM of specific primers, 1× buffer containing 2.5 mM of Mg^{2+}, and 1 µl of Qiagen OneStep RT-PCR enzyme mix for a reaction volume of 25 µl. Reverse transcription occurred at 50ºC for 30 minutes followed by a PCR activation step at 95ºC for 15 minutes. PCR conditions consisted of 30 repetitive cycles of denaturation at 94ºC for 30 seconds, annealing at 55ºC for 30 seconds, extension at 72ºC for 1 minute, and a final extension at 72ºC for 10 minutes. The primers used to amplify internal fragments of

genes pfc_09170 (coding for HSP20), pfc_08710 (for HSP60), pfc_07215 (for HtpX), pfc_04085 (for POR), and pfc_09175 (for *aaa*$^{+}$ATPase), are shown in Table 2.1. The gene encoding POR was used as control. RT-PCR products were separated by electrophoresis and visualized with ethidium bromide. Semiquantitave analysis was performed by using the Quantity One software (Bio-Rad, Hercules, CA). Results are expressed as a ratio of the target PCR product to POR. These experiments were performed in duplicate.

Results

Effect of temperature and NaCl concentration on growth and solute accumulation by the parent strain

The growth profiles of *P. furiosus* COM1 between 88 and 100ºC were examined in medium containing 2.8% (w/v) NaCl. In accordance with the properties of the original strain of *P. furiosus* (Fiala and Stetter 1986), the growth rate did not change appreciably in the temperature range 88 to 98ºC (Table 2.2 and results not shown). However, maximum cell density was observed at around 90ºC and growth within 8 hours was not observed at 100ºC (Table 2.2, Fig. 2.3A).

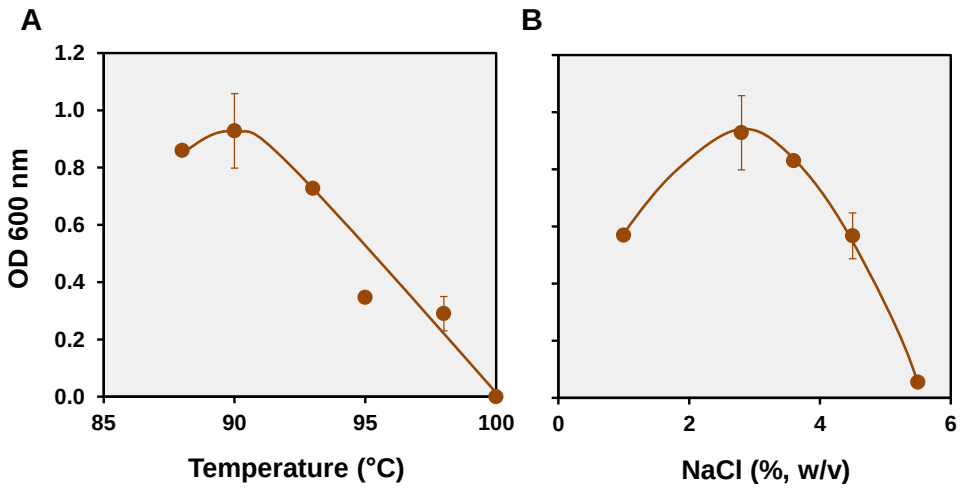


FIGURE 2.3 Maximal cell density (OD_{600}) for *P. furiosus* COM1 as a function of the growth temperature (A) and the NaCl concentration in the growth medium (B). The following NaCl concentrations were examined: 1.0, 2.8, 3.6, 4.5 and 5.5% (w/v), corresponding to 0.171, 0.479, 0.616, 0.770 and 0.941 M, respectively. The bars represent standard deviations obtained from three to six independent experiments. In the other cases a single experiment was performed.

At the temperature for maximum growth, 90ºC, strain COM1 grew in media containing up to about 5.5% NaCl, with optimal growth at approximately 2.8% NaCl (Table 2.2, Fig. 2.3B), a behavior typical of slightly halophilic organisms. Unlike for the temperature dependence, the cell density and the growth rate displayed parallel trends when the NaCl concentration was varied (Fig. 2.3B and results not shown).

TABLE 2.2 Growth rate and maximal cell density (OD measured at 600 nm) of *P. furiosus* COM1 (parent strain) and mutants under optimal growth conditions, heat stress, and osmotic stress.

***P. furiosus* strain**	**Growth rate (h^{-1}) (maximal cell density)[a]**		
	90ºC, 2.8% NaCl	**98ºC, 2.8% NaCl**	**90ºC, 4.5% NaCl**
Parent strain	0.65 ± 0.05 (0.88 ± 0.13)	0.70 ± 0.08 (0.32 ± 0.06)	0.39 ± 0.03 (0.49 ± 0.08)
MG-deficient	0.56 ± 0.06 (0.71 ± 0.05)	0.62 ± 0.08 (0.29 ± 0.03)	0.25 ± 0.04 (0.39 ± 0.04)
DIP-deficient	0.73 ± 0.06 (1.42 ± 0.06)	0.77 ± 0.03 (0.45 ± 0.07)	0.43 ± 0.03 (0.63 ± 0.15)

[a] The values are averages from three to six independent experiments. Optimal temperature for growth: 90ºC; optimal NaCl concentration: 2.8% (w/v).

In view of the observed growth profiles, 90ºC and 2.8% NaCl were selected as the optimal temperature and salt conditions for growth of the COM1 strain. Heat stress was imposed by increasing the growth temperature to 98ºC in medium containing 2.8% NaCl, while the effect of salt stress was studied by increasing the NaCl concentration to 4.5%, at the optimal growth temperature. Typical growth curves obtained at optimal and stress conditions are shown in Figure 2.4A. An increase of 8ºC above the optimal temperature resulted in a clear impairment on the final cell density (OD600 varied from 0.88 to 0.32), while not affecting significantly the growth rate (Table 2.2). The lack of effect of temperature on the growth rate over the range 75 - 90ºC was also reported for *T. kodakarensis* by Kanai et al. (2010). On the other hand, in medium containing 4.5% NaCl

both the growth rate and the final OD were reduced by about 40% when compared with optimal conditions.

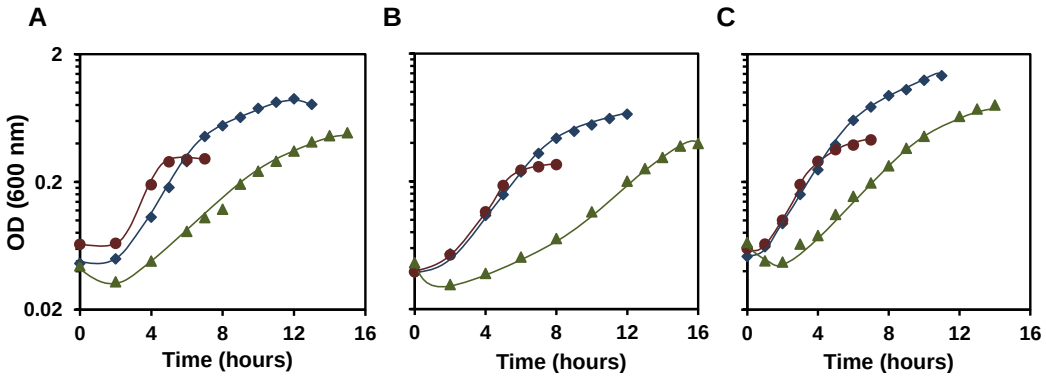


FIGURE 2.4 Growth curves of *P. furiousus* COM1 (**A**), the MG-deficient strain (**B**), and the DIP-deficient strain (**C**) at optimal growth conditions (90°C; 2.8% NaCl) (blue), under heat stress (98°C; 2.8% NaCl) (red), and under osmotic stress (90°C; 4.5% NaCl) (green). The growth rates were determined and are shown in Table 5.2. OD means the optical density. For each strain and condition, one representative curve of a total of three to six is shown.

In parallel, the accumulation of compatible solutes was investigated in cells collected at the end of the exponential phase (Table 2.3). MG was the predominant solute accumulated by strain COM1 under all growth conditions examined, corresponding to about 65%, 47% and 78% of the total organic solute pool under optimal, heat stress and salt stress conditions, respectively. Under optimal growth conditions DIP was detected in vestigial amounts, but its concentration increased 7-fold between the optimal temperature and heat stress conditions, constituting 20% of the total solute pool at 98ºC; in contrast, the concentration of MG

decreased by 24% in response to the heat stress condition. Alanine, lactate, and aspartate were also detected in the ethanol extract, but in lesser amounts, and their concentrations did not change significantly with either of the stresses imposed (Table 2.3, Fig. 2.5). Accumulation of lactate is intriguing, since the genomes of *P. furiosus* lack genes encoding a lactate dehydrogenase homolog (Basen et al. 2012). We postulate that lactate was taken up from the medium. NMR analysis confirmed the presence of lactate in yeast extract (17.9 mg of lactate per gram), leading to a final concentration of 0.5 mM lactate in the culture medium. The concentrations of intracellular lactate were highly variable among experimental replicates and hence the respective standard deviations are much greater than those estimated for the other solutes.

The total solute pools at optimal growth temperature and under heat stress conditions were identical (0.65 and 0.68 µmol/mg of protein, respectively), but in response to elevated salinity the total pool of solutes increased approximately 2.2-fold. This increment is attributed largely to higher levels of MG, which increased 2.7-fold. DIP also increased (3.5-fold), but the final concentration was much lower than that of MG, contributing only 5% to the total pool of solutes.

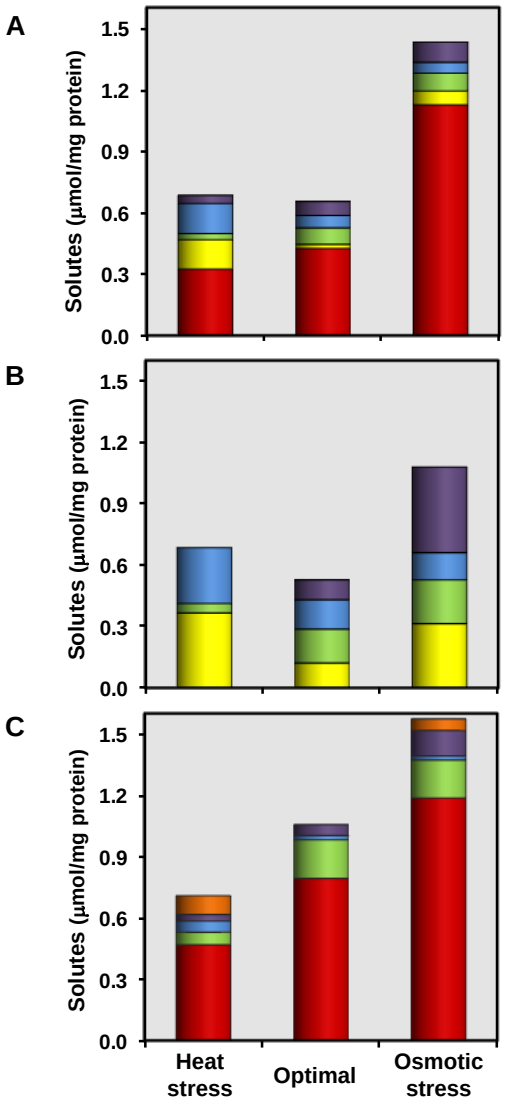


FIGURE 2.5 Composition of organic solutes in *P. furiosus* COM1 (parent strain) (**A**), the MG-deficient mutant (**B**), and the DIP-deficient mutant (**C**), under optimal growth conditions (90°C; 2.8% NaCl), heat stress (98°C; 2.8% NaCl), and osmotic stress (90°C; 4.5% NaCl). Mannosylglycerate (red); di-*myo*-inositol-phosphate (yellow); alanine (green); lactate (blue); aspartate (purple); and L-*myo*-inositol-1-

phosphate (orange). The values are the mean of two to three independent experiments. Standard deviations are presented in Table 2.3.

Construction of mutants deficient in MG or DIP synthesis

The contribution of compatible solutes during osmo- and thermo-adaptation was assessed by constructing mutants of *P. furiosus* COM1 unable to synthesize MG or DIP. To this end, the genes encoding the synthases involved in DIP and MG synthesis, IPCT/DIPPS and MPGS, respectively, (Fig. 2.1), were deleted by double-crossover homologous recombination. The construction of the two mutants was confirmed by different techniques. PCR analysis corroborated the absence of the deleted genes. In addition, each flank of the recombination region was sequenced to show the correct replacement of the target gene by the *pyrF* cassette. Moreover, the genotype of each mutant was verified by Southern blot analysis using a 500-bp fragment upstream of the target gene and three restriction enzymes (Fig. 2.6). The predicted hybridization pattern for single insertion of the marker cassette into the genome was obtained for both deletion mutants. Finally, analysis of the solute pool of the two mutant strains confirmed the absence of the targeted solute (MG or DIP) in cell-extracts, as described below.

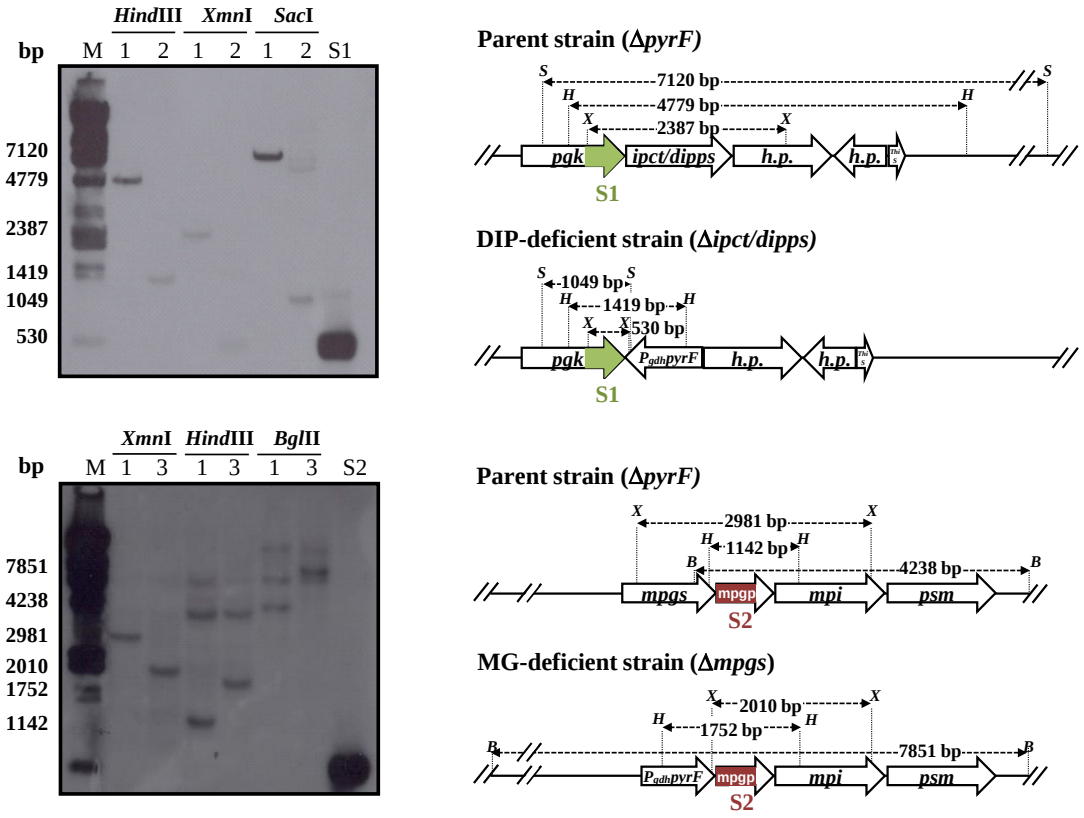


FIGURE 2.6 Genetic confirmation of the *P. furiosus* mutant strains by Southern-blot analysis (left panel). The total DNA was digested with the enzymes indicated, separated on 1% agarose gel, blotted and hybridized with labeled probes S1 (*pgk*, pair primer pgk1/pgk2) or S2 (*mpgp*, pair primer mpgp1/mpgp2). The restriction enzymes were selected to distinguish between the parental strain (1) and the mutant strains, DIP-deficient strain (2) and MG-deficient strain (3). The molecular size of the hybridized fragments are indicated. M, molecular size marker. Predicted restriction map of the recombination regions in the parent and mutant strains (right panel). Abbreviations: *pgk*, phosphoglycerate kinase; *ipct/dipps*, CTP:L-*myo*-inositol-1-phosphate cytidylyltransferase/di-*myo*-inositol-phosphate-phosphate synthase; *ThiS*, sulfur carrier protein; P_{gdh}, promoter region of glutamate dehydrogenase; *pyrF*, orotidyne 5'-phosphate decarboxylase; *h.p.*, hypothetical protein; *psm*, phospho-sugar mutase; *mpi*, mannose-6-phosphate

isomerase; *mpgs*, mannosyl-3-phosphoglycerate synthase; *mpgp*, mannosyl-3-phosphoglycerate phosphatase; *B*, *Bgl*II; *H*, *Hind*III, *X*, *Xmn*I; *S*, *Sac*I.

Effect of heat and salinity stress on growth and solute accumulation by mutants

DIP-deficient and MG-deficient mutants were grown under optimal and under stress conditions as described above for the parent strain. At optimal conditions the MG-deficient mutant exhibited growth rate and final cell density similar to those of the parent strain. Curiously, the DIP-deficient mutant reached a higher final OD (1.42 to compare with 0.88), and the growth rate was also slightly higher (Table 2.2 and Fig. 2.4). Under heat stress conditions the three strains exhibited similar growth parameters. Upon salinity stress, the growth of the MG-deficient strain was clearly impaired, but again the DIP-deficient mutant grew slightly better than the parent strain. It seems as though MG is more important than DIP for the performance of this archaeon.

To interpret the results of the growth performance of the three strains we compared the composition of the solute pools under optimal and stress conditions. At optimal growth conditions, the absence of MG in the MG-deficient mutant was offset by an increase in the pools of DIP, alanine and lactate (6-, 2.1- and 2.3-fold, respectively) in comparison with the parent strain (Table 2.3 and Fig. 2.5). While MG was by far the predominant solute in the parent strain, the solute pool in the MG deficient strain comprised four solutes, DIP, alanine, aspartate, and lactate, in similar proportions. Like the parent strain, increasing the growth temperature led

to a strong increase in the level of DIP (3-fold), which became the major solute corresponding to about 55% of the total solute pool. Besides DIP, lactate and alanine were detected, but aspartate decreased to an undetectable concentration.

TABLE 2.3 Quantification of organic solutes in *P. furiosus* COM1 (parent strain) and the MG deficient and DIP-deficient mutants, under optimal growth conditions, heat stress and osmotic stress.

Strain	Growth temp (°C)	NaCl concn (% w/v)	Solutes (μmol/mg of protein)[a]						
			MG	DIP	Ala	Lac	Asp	Ino1P	Total
Parent	90	2.8	0.42 ± 0.03	0.02 ± 0.01	0.08 ± 0.01	0.06 ± 0.01	0.07 ± 0.01	n.d.	0.65 ± 0.04
	98	2.8	0.32 ± 0.06	0.14 ± 0.01	0.03 ± 0.02	0.15 ± 0.13	0.04 ± 0.01	n.d.	0.68 ± 0.14
	90	4.5	1.12 ± 0.00	0.07 ± 0.00	0.09 ± 0.04	0.05 ± 0.00	0.10 ± 0.04	n.d.	1.43 ± 0.06
MG-deficient	90	2.8	n.d.	0.12 ± 0.03	0.17 ± 0.02	0.14 ± 0.02	0.10 ± 0.01	n.d.	0.53 ± 0.04
	98	2.8	n.d.	0.37 ± 0.05	0.05 ± 0.03	0.27 ± 0.10	n.d.	n.d.	0.68 ± 0.12
	90	4.5	n.d.	0.31 ± 0.05	0.21 ± 0.02	0.13 ± 0.04	0.42 ± 0.07	n.d.	1.08 ± 0.10
DIP-deficient	90	2.8	0.79 ± 0.19	n.d.	0.19 ± 0.02	0.02 ± 0.01	0.05 ± 0.01	n.d.	1.06 ± 0.19
	98	2.8	0.47 ± 0.06	n.d.	0.06 ± 0.02	0.05 ± 0.03	0.03 ± 0.01	0.09 ± 0.03	0.70 ± 0.07
	90	4.5	1.19 ± 0.11	n.d.	0.19 ± 0.00	0.02 ± 0.00	0.12 ± 0.00	0.06 ± 0.01	1.58 ± 0.11

[a] Values are the mean of two to three independent experiments. MG, mannosylglycerate; DIP, di-*myo*-inositol-phosphate; Ala, alanine; Lac, lactate; Asp, aspartate; and Ino1P, L-*myo*-inositol-1-phosphate. n.d., not detected.

The total concentrations of solutes were similar in the MG-deficient strain when heat stress and optimal conditions are compared, but the total concentrations clearly increased upon salt stress, primarily due to the increment in the pool of aspartate (4-fold) and DIP (3-fold). Even so, the total solute pool in the MG-deficient strain was significantly lower than that of the parent strain when subjected to the same osmotic stress (1.08 to compare with 1.43 µmol/mg of protein).

Except for the absence of DIP, the DIP-deficient mutant showed patterns of solute accumulation generally similar to those of the parent strain (Fig. 2.5). MG was the major solute under all growth conditions examined, corresponding to about 75, 67 and 75% of the total organic solute pool under optimal, heat stress and salt stress conditions, respectively. Notably, a significant increase of 1.6-fold (1.06 ± 0.19 *vs* 0.65 ± 0.04 µmol/mg of protein; Table 2.3) in total solute concentration was observed for the DIP-deficient mutant in comparison with COM1 under optimal growth conditions. In response to salinity stress, the level of MG increased by 50% and the total pool of solutes was about 50% and 10% higher than that of the MG-deficient and parent strains, respectively. L-*myo*-inositol-1-phosphate, the substrate of the IPCT/DIPPS enzyme encoded by the deleted gene, was detected under heat and salt stress but not at optimal conditions. Rising the temperature from 90 to 98ºC caused a

decrease in the MG pool (1.7-fold). The total solute pool was similar to those observed in the wild-type and MG-deficient strains (0.70, 0.68 and 0.68 µmol/mg of protein), when challenged also with heat stress.

The rationale behind this study relies on the assumption that the effects observed on the growth profiles of the mutants under stress conditions are caused by the deletion of the target solutes and not by the differential induction of other stress factors, such as heat shock proteins and chaperonins, which would undermine the interpretation of results. The response of *P. furiosus* to heat stress was studied at the transcription level by Shockley et al. (2003). Based on the results of these authors we selected four genes that were clearly induced (4-7 fold) under heat stress, i.e., genes coding for a Hsp60-like chaperonin (PF1974), a Hsp20-like small heat shock protein (PF1883), one molecular chaperone belonging to the AAA^+ family known as VAT (PF1882) and an ATP-independent protease (PF1597). RT-PCR experiments confirmed that the transcription levels of the corresponding genes in strain COM1 were identical in the mutant and control strains (Fig. 2.7).

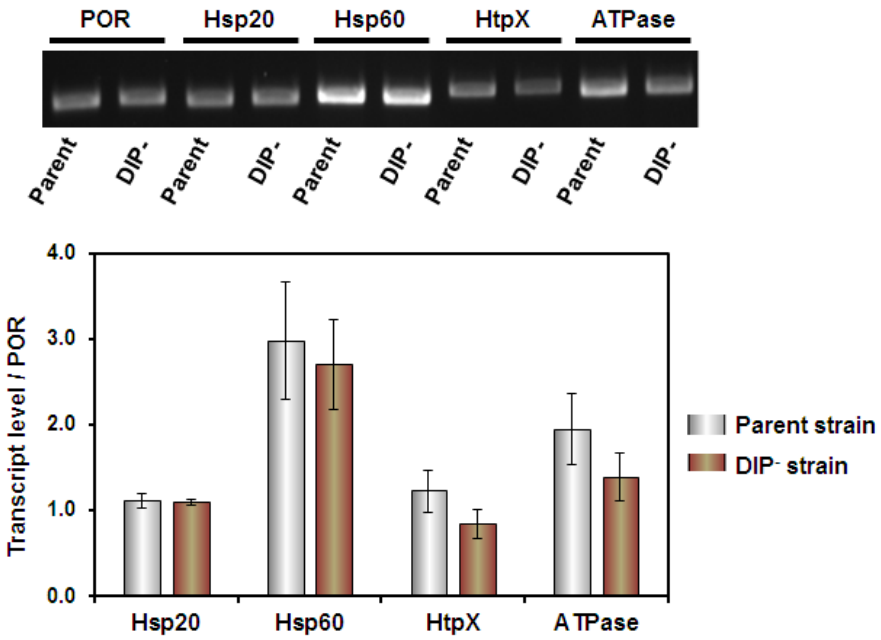


FIGURE 2.7 Comparison of the transcriptional levels of four genes involved in heat stress response between the Parent and DIP-deficient strains grown under heat stress (98°C, 2.8% NaCl). Upper panel, RT-PCR experiments were performed using primers specific for: Hsp20, Hsp20-like small heat shock protein (pfc_09170); Hsp60, thermosome (pfc_08710); HtpX, ATP-independent protease (pfc_07215); ATPase, molecular chaperone known as VAT (pfc_09175). Pyruvate ferredoxin oxidoreductase (POR) was used as a control. Lower panel, semiquantitave analysis of transcript levels relative to POR. The experiments were done in duplicate.

Discussion

The variant strain *P. furiosus* COM1 showed a salinity profile similar to that of the original wild-type strain, *P. furiosus* DSM 3638^T, but the temperature profile was shifted downwards. Although both strains showed similar growth rates over the 88 – 98°C range, the temperature for maximum cell

density of COM1 was 90ºC and growth was not observed at 100°C (after 8 hr), while the maximum cell density of the wild-type strain is at 95ºC and some growth is observed at 100°C (Fiala and Stetter 1986, Martins and Santos 1995). The explanation for the observed phenotypic differences could be related to the large number of changes encountered at the genome level (Bridger et al. 2012). Exploring these differences by complementing the COM1 strain with genes that are functional in the DSM 3638^T strain provides an avenue to approach this problem.

As in *P. furiosus* DSM 3638, the COM1 variant accumulates MG in response to salt stress and DIP in response to supraoptimal temperature. However, while DIP is by far the predominant solute in the wild-type strain at optimal and supraoptimal temperatures, MG is the major solute (over 65% of the total pool) in the variant strain under all conditions examined. It appears that the variant strain relies heavily on MG both for osmotic and thermal protection and DIP has a much smaller contribution. A comparison of the regions (360-bp) immediately upstream of the genes encoding IPCT/DIPPS and MPGS in the genomes of the wild-type DSM 3638 and the variant COM1 showed a perfect match. In a further attempt to find clues for the differentiated behaviour of the strains, we looked for changes in putative transcription regulators: 43 out of the 44 candidates annotated in the NCBI database were identical, and differences (58% identity) were found only in PF0054. For the moment, the origin of the phenotypic differences between the two strains remains obscure.

In this work, the genes involved in the synthesis of MG or DIP were disrupted in *P. furiosus* COM1 as a means to obtain insight into the role of these compatible solutes. If a specific solute were an important part of the machinery mounted by the cell to cope with stress, we would expect impaired growth under stressful conditions for a mutant unable to synthesize such a solute. Significantly, growth of the MG-deficient mutant as compared with the parent strain was negatively affected under salt stress, but not under heat stress, an observation that supports a major role of MG in osmoadaptation of this archaeon. In fact, the combined accumulation of DIP and aspartate in response to salt was not sufficient to offset the lack of MG. This observation together with the consistent increase of MG upon osmotic stress in many (hyper)thermophiles (Santos et al. 2011) suggests that the regulatory network implied in MG accumulation is designed to respond effectively to osmotic imbalance. Surprisingly, growth of the DIP deficient mutant was as good, or even slightly better than that of the parent strain under either of the challenges imposed. Comparison of the growth and solute profiles of the parent and the DIP-deficient strains under heat stress conditions suggests that MG can replace DIP for thermoprotection with similar performance. Indeed, the growth profiles were nearly superimposable despite the substantial differences in the composition of the solute pools: while MG and DIP were the predominant solutes in the DIP deficient and MG-deficient strains, respectively, the solute pool of the parent strain comprised a mix of MG and DIP in a proportion of approximately 2:1. Hence, it appears that DIP

and MG can substitute for each other to protect against heat damage, but MG is more closely associated with osmoprotection.

The strategy of solute replacement for osmotic adjustment has been illustrated in a few mesophilic halophiles (Saum et al. 2009, Köcher et al. 2011). For example, in the bacterium *Halobacillus halophilus*, proline can be substituted by glutamate, glutamine and ectoine to cope with osmotic stress (Köcher et al. 2011). In *Methanosarcina mazei* the lack of N^{ε}-acetyl-β-lysine is compensated for by accumulation of glutamate and alanine, but at high salinity the deficient mutant exhibited poorer growth than the parent strain. Thus, in this case the substitution of N^{ε}-acetyl-β-lysine by alternative solutes was not fully adequate for osmotic adjustment (Saum et al. 2009). Replacement of compatible solutes for thermoprotection has also been reported in a DIP-deficient *T. kodakarensis* mutant, where the missing solute was replaced by aspartate, and this substitution had no effect on the ability of the strain to grow at supraoptimal temperatures (Borges et al. 2010).

The construction of *P. furiosus* mutants unable to synthesize either MG or DIP have provided more precise information on the triggers associated with the specific solute accumulation. In the mutant lacking MG synthesis, it is apparent that DIP accumulation is enhanced by heat and also by salt stress, revealing a certain flexibility of the sensing/regulatory mechanisms leading to DIP synthesis. In contrast, MG accumulation was clearly induced at elevated salinity, but reduced under heat stress.

The molecular mechanisms involved in the regulation of DIP synthesis have been investigated in *Thermococcus* spp. as part of the global heat stress response in hyperthermophiles (Kanai et al. 2010). The central role of the transcriptional regulator Phr is well established in members of the order Thermococcales (Kanai et al. 2010, Keese et al. 2010). At optimal temperature, Phr binds to the promoter region of several genes such as, *hsp20* and *aaa*$^{+}$*atpase,* and the gene encoding IPCT/DIPPS, the enzyme implied in DIP synthesis, thereby blocking their transcription. Conversely, at higher temperatures the affinity of Phr to the promoter region is decreased and transcription activated. The activity of Phr in the regulation of DIP synthesis was validated in *T. kodakarensis* by disruption of the respective gene (Kanai et al. 2010). Therefore, the molecular link between heat stress and DIP accumulation is fairly well established at the transcriptional level. On the other hand, the molecular mechanisms leading to enhancement of the DIP pool in response to elevated salinity remain obscure.

Herein we show that in *P. furiosus* the accumulation of MG increased in response to osmotic stress and was reduced at the supraoptimal temperature. This is the general trend for MG accumulation in (hyper)thermophiles; actually, up-regulation of MG synthesis by heat was observed only in *Rhodothermus marinus*, which is unique insofar as it possesses two pathways for the synthesis of this solute (Martins et al. 1999). In this thermophilic bacterium, MG synthesis is regulated at the translational level as the amount of MPGS, the synthase involved in the

2-step pathway, increases in response to elevated salinity and decreases upon heat stress (Borges et al. 2004); on the other hand, the level of the synthase involved in the single-step pathway is selectively increased by heat. As demonstrated in this work, *P. furiosus* synthesises MG exclusively via the 2-step pathway. It is therefore interesting to find that the outcome of MG regulation by stress in *P. furiosus* is identical to that observed for the 2-step pathway in *R. marinus*, although the specific regulatory mechanisms operating in bacteria and archaea are expected to be different.

In conclusion, this work shows that the lack of MG in *P. furiosus* was compensated by DIP accumulation with comparable efficacy in thermoprotection. MG and DIP played interchangeable roles in thermoprotection, while MG was primarily directed at osmoprotection. Future studies must determine the molecular mechanisms in the sequence of events from thermo- and osmo-sensing to the final accumulation of specific solutes.

Acknowledgements and work contributions

I thank M. Adams (University of Georgia, USA) for the fruitful collaboration that resulted in the construction of *Pyrococcus furiosus* mutants. S. K. Chandrayan and P. M. McTernan contributed to mutant constructions.

CHAPTER 3

Metabolomic response of *Pyrococcus furiosus* to heat shock

Manuscript in preparation:

Esteves AM, Peyriga L, Gonçalo G, Torcato I, Borges N, Portais JC, Santos H. 2015. Insights into the metabolomic response of *Pyrococcus furiosus* to heat shock.

Contents

Summary

Metabolome analysis is essential for obtaining a comprehensive picture of cell physiology. In the present study, the metabolomic response of *Pyrococcus furiosus* to heat shock was characterized using nuclear magnetic resonance spectroscopy (NMR) and ion exchange liquid chromatography/mass spectrometry (LC/MS). Data collected using ^{1}H NMR enabled the assignment of 35 metabolites, including 13 amino acids, 5 organic acids, 5 nucleotides, 3 co-factors, 3 polyamines, 2 sugar nucleotides, 2 thermolytes, and 2 sugars, under optimal and heat shock conditions. Aspartate and alanine were the major intracellular metabolites detected in cells grown under optimal conditions. In contrast, at supraoptimal temperatures, the predominant metabolites were di-*myo*-inositol phosphate, mannosylglycerate, maltose, UDP-*N*-acetylglucosamine, and UDP-*N*-acetylgalactosamine. Aspartate was the only metabolite whose concentration decreased under heat shock. LC/MS analysis was used to quantify the metabolites involved in central metabolism. Among the 15 compounds identified, it was possible to quantify 3 nucleotides (ATP, ADP, AMP), a combination of trioses (2-phosphoglycerate/3-phosphoglycerate), GDP-mannose, glycerol-3-phosphate, phosphoenolpyruvate, ribose-1-phosphate, and trehalose-6-phosphate. With the exception of GDP-mannose and ATP, the level of all the metabolites quantified by LC/MS decreased at supraoptimal temperatures. The integration of the NMR and LC/MS data suggests a decrease in the synthesis of metabolites involved in central pathways, such

as glycolysis and Krebs cycle, while the production of compounds related to cell structure and/or protection seems to be favored.

Introduction

The heat shock response has been studied in several hyperthermophiles (Macario and Conway de Macario 1999, Shockley et al. 2003, Laksanalamai et al. 2004a, Rohlin et al. 2005). With the increasing availability of genome sequences, most studies focused on transcriptome analysis using DNA microarray technology (Keese et al. 2010, Kanai et al. 2010). When *Archaeoglobus fulgidus* or *Pyrococcus furiosus* were subjected to supraoptimal temperatures the transcription levels of the genes coding for heat shock proteins, such as Hsp20, thermosome, and proteases, increased (Shockley et al. 2003, Rohlin et al. 2005). Heat stress also affects the expression of genes other than those encoding HSPs. In the case of *A. fulgidus*, 10% of all genes were differentially expressed when the temperature was increased from 78°C to 89°C, denoting a complex process that goes beyond the activation of heat shock proteins (Rohlin et al. 2005). In fact, a remarkable feature of hyperthermophilic organisms is the heat-induced accumulation of unique compatible solutes (Santos et al. 2011), which have been proposed to act as chemical chaperones that protect protein structures against thermal denaturation (Faria et al. 2008). Despite the extensive information on compatible solute accumulation in hyperthermophiles subjected to heat stress conditions

(Martins and Santos 1995, Neves et al. 2005, Borges et al. 2010), the global metabolic response is poorly characterized.

Metabolites are the very end products of enzymatic pathways (Nicholson and Lindon 2008, Tang 2011, Patti et al. 2012), hence, metabolomics is expected to provide a more reliable picture of the actual physiological status of the organism under a specific condition. The number of studies conducted in this field has increased considerably in recent years (Jozefczuk et al. 2010, Brito-Echeverría et al. 2011, Sévin and Sauer 2014). However, at least for microorganisms, most of the initial research effort was directed to the optimization of methods for quenching of metabolism, sampling, and metabolite extraction (Winder et al. 2008, Taymaz-Nikerel et al. 2009, van Gulik 2010). To obtain reliable and significant data, fast quenching of enzymatic activities, separation of exometabolome and endometabolome, and complete extraction of intracellular metabolites are required. The direct sampling of culture into cold methanol or other organic solutions is commonly used for arresting metabolism (Taymaz-Nikerel et al. 2009, van Gulik 2010); however, it has been shown that this method may cause cell lysis and consequent loss of intracellular metabolites (Wittmann et al. 2004, Bolten et al. 2007). The separation of cells from the growth medium can be carried out either by centrifugation or fast-filtration. The latter method has received considerable attention in recent publications (Bolten et al. 2007, Kim et al. 2013) since it avoids the quenching step with cold organic solutions. The subsequent process of metabolite extraction and identification is still challenging due to

the enormous chemical variety of intracellular metabolites and the large range of concentrations at which metabolites can be present. Currently, nuclear magnetic resonance (NMR) and mass spectrometry (MS) combined with a separation method (gas or liquid chromatography), are the most powerful and common analytical techniques applied to the characterization of metabolome (Tang 2011). Both techniques have the advantage of handling a large number of metabolites in a single experiment (Tang 2011), but regrettably neither of them is devoid of disadvantages. While MS requires a more time-consuming sample preparation, NMR has low sensitivity which limits the detection of metabolite concentration to the micromolar range (Wishard 2008, Bingol and Brüschweiler 2014).

The goal of the present work is to obtain insights into the global metabolic response to heat shock in *Pyrococcus furiosus*, a model hyperthermophilic archaeon. Changes in the level of metabolites were investigated by using NMR and ion exchange liquid chromatography/mass spectrometry (LC/MS). It is expected that the combination of these techniques will allow for an improved coverage of the entire metabolome.

Materials and Methods

Strain and culture conditions

P. furiosus DSM 3638 (Fiala and Stetter 1986) was cultivated in complex medium as described previously (Adams et al. 2001), with the following modifications: 0.5% (w/v) of maltose and 0.25% (w/v) of yeast extract were used as carbon sources, the NaCl concentration of the medium was lowered to 1%, and elemental sulfur was not added. The salt concentration was decreased due to incompatibility of salts with LC/MS analysis. Culture growths were performed in a 2-liter fermentation vessel at 90°C with continuous gassing with a mixture of N_2 (80%) and CO_2 (20%), and stirring at 80 rpm. Cultures used as pre-inoculum were successively passed in fresh medium, at least three times before inoculation. Prior to inoculation (10% of an overnight culture), the medium was supplemented with cysteine (0.5 g/l) and Na_2S (0.5 g/l) and the pH adjusted to 6.8. During growth, the pH was maintained at 6.8 by the addition of $NaHCO_3$ 10% (w/v). Optical density (OD) at 600 nm was used to assess cell growth.

To study the effect of heat shock on the metabolism of *P. furiosus*, an aliquot (see below the exact volume for each process) of culture grown at 90ºC in the late-exponential phase (OD ~ 0.220) was collected (control condition), and approximately 250 ml of culture transferred into a 0.5-liter fermentation vessel pre-heated at 97°C. The temperature of the culture reached 97ºC within 10 minutes; the culture was maintained at this temperature for a further 35 minutes. After 45 minutes of incubation, an

aliquot of culture was removed (heat shock condition). The samples were immediately processed to analyze the intracellular metabolites by NMR and LC/MS.

Analysis of the intracellular metabolites by NMR

Cells collected at 90°C (control) or 97°C (heat shock) were processed using the same protocol. Briefly, approximately 250 ml of cell culture was passed through a steel coil immersed in an ice bath, using a vacuum pump. The duration of the transfer step was less than 1 minute and the final temperature of the cell culture was in the range 11 to 16°C. After collection, total broth was transferred into a centrifuge tube and measured gravimetrically. Cells were harvested by centrifugation at 6,370×*g* for 7 minutes at 4°C, washed once with an isotonic solution (1.0% NaCl, 0.35% $MgSO_4 \cdot 7H_2O$, 0.27% $MgCl_2 \cdot 6H_2O$, 0.033% KCl, 0.025% NH_4Cl, and 0.014% $CaCl_2 \cdot 2H_2O$) and centrifuged again at 5,000×*g* for 7 minutes at 4°C. The supernatant was discarded and the intracellular metabolites were extracted immediately from the pellet. The duration of the entire process was approximately 20 minutes. The scheme of the experimental procedure is depicted in Figure 3.1. The data was obtained from three independent biological replicates for each condition (90 and 97°C).

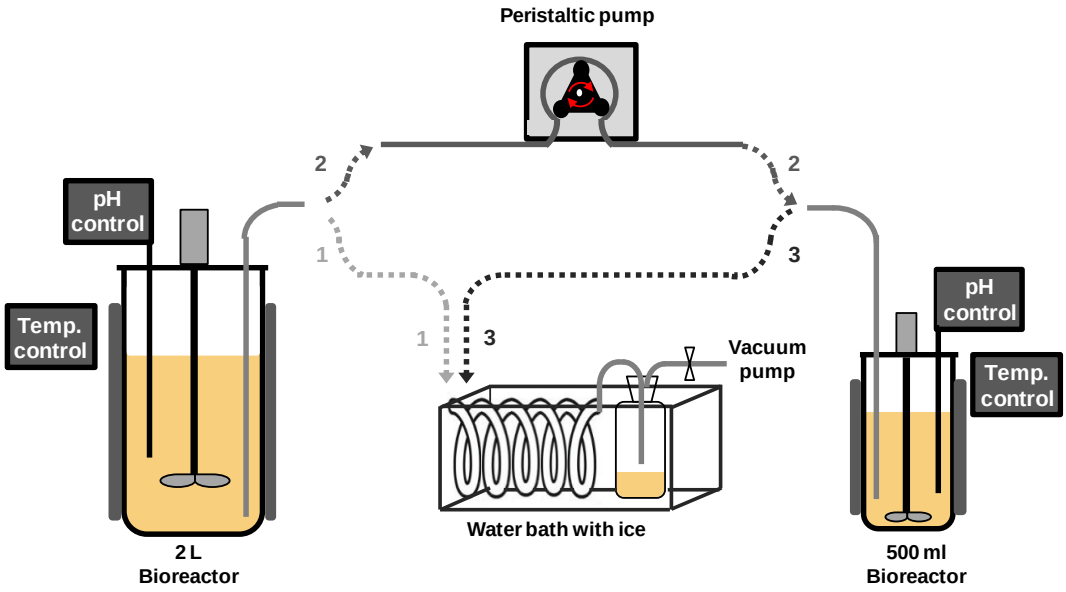


FIGURE 3.1 Schematic representation of the experimental set-up used to impose heat shock and to collect samples for NMR analysis. Dashed arrows represent the several steps of the experiment: **Step 1:** Approximately 250 ml of culture grown at 90°C passed through a cooled steel coil into a collection flask using a vacuum pump and the cells were harvested by centrifugation (control condition); **Step 2:** Immediately after, approximately 250 ml of culture was transferred into a 500 ml bioreactor pre-heated at 97°C using a peristaltic pump; and **Step 3:** After 45 minutes the culture was passed through a cooled steel coil into a collection flask using a vacuum pump and the cells were recovered by centrifugation (heat shock condition).

Extraction of intracellular metabolites

Metabolite extraction was performed using 80% of cold methanol. Cells were re-suspended in 20 ml of methanol and the suspension stirred for 15

minutes in the cold. The mixture was centrifuged at 10,000×*g* for 15 minutes (4°C) and the supernatant was collected. Methanol was removed by low-pressured rotary evaporation and samples were concentrated by lyophilization.

NMR acquisition, data processing and analysis

Residues were dissolved in 600 μl of $^{2}H_2O$ for NMR analysis. An aliquot of 440 μl was introduced into a 5 mm NMR tube and mixed with 50 μl of phosphate buffer, pH 7.0 and 10 μl of sodium 3-(trimethylsilyl) proprionate-2,2,3,3-d_4 (TSP), resulting in a final concentration of 33 mM and 60 μM of phosphate buffer and TSP, respectively. Experiments were conducted on a Bruker AvanceIII 800 NMR spectrometer (Bruker Biospin, Rheinstetten, Germany) operating at 800.33 MHz for ^{1}H, equipped with a 5 mm three-channel probe (TXI-Z H/C/N/-D) and a Bruker temperature control unit set at 298 K. ^{1}H-NMR spectra were acquired using the 1D nuclear Overhauser enhancement spectroscopy (NOESY)-presat sequence with gradients to improve water suppression (*noesygppr1d* pulse sequence from the Bruker library). Free induction decays (FIDs) were collected with 128 transients into 128 k data points using a spectral width of 16 kHz (20 ppm) with a relaxation delay of 4.0 s, an acquisition time of 4.09 s and a mixing time of 10 ms. Two-dimensional and 1D-selective experiments were acquired as described below to aid spectral assignments. The 1D-selective-TOCSY experiments were performed using the Bruker library pulse sequence *selmlgp.2*, 1D homonuclear Hartman-Hahn transfer using a selective

refocusing shaped Gaussian pulse, MLEV17 sequence for mixing and with suppression of zero quantum coherence which minimized peak distortions (Kessler et al. 1986). FIDs were collected with 1 024 transients into 64 k data points using a spectral width of 16 kHz (20 ppm) with a relaxation delay of 2.0 s and a TOCSY mixing time of 60 ms. Depending on the structure of each multiplet, a width of 4-15 Hz was excited with the selective pulse. Two-dimensional NMR experiments, including ^{1}H-^{1}H *J*-resolved, ^{1}H-^{1}H correlation spectroscopy (COSY), ^{1}H-^{1}H total correlation spectroscopy (TOCSY), and ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC), were also acquired. J-resolved spectra were obtained using the pulse sequence *jresgpprqf* from the Bruker library with presaturation during the relaxation delay. FIDs were collected with 2 transients into 8 192 and 40 data points using a spectral width of 13 kHz (16.7 ppm) and 78 Hz for F2 and F1 dimensions, respectively, with a relaxation delay of 2.0 s. COSY experiments were acquired using the *cosygpprqf* pulse sequence from the Bruker library. The 32 FIDs with 4 096 and 400 data points for F2 and F1 dimensions were acquired with a width of 8 kHz (10 ppm), and a relaxation delay of 2 s. TOCSY spectra were performed based on the pulse sequence from the Bruker library *dipsi2esgpph* (homonuclear Hartman-Hahn transfer using DIPSI2 sequence for mixing phase sensitive water suppression using excitation sculpting with gradients) (Shaka et al. 1988). The 64 FIDs were acquired into 2 048 and 512 data points for the dimensions F2 and F1, respectively, with a spectral width of 8.2 kHz (10.2 ppm), 1.5 s of relaxation delay, and

60 ms TOCSY mixing time. The HSQC spectra were acquired with the *hsqcetgpsi* pulse sequence (Bruker library), 2D ^{1}H-^{13}C correlation via double INEPT transfer using sensitivity improvement, phase sensitive using Echo/Antiecho-TPPI gradient selection, with decoupling during acquisition using trim pulses in INEPT transfer (Palmer et al. 1991). The FIDs were collected with 128 transients into 2 048 and 256 data points using a spectral width of 8 kHz (10 ppm) and 33 kHz (165 ppm) for F2 and F1 dimensions respectively, with a relaxation delay of 2.0 s. After collection each spectra was transformed with a specific window function prior to Fourier transformation and line-broadening: exponential functions for 1D spectra and sine functions for 2D spectra. NMR spectra were processed with an exponential line broadening of 0.3 or 1 Hz before Fourier transformation for 1D-NOESY or 1D-TOCSY, respectively. Each spectrum was manually phased, baseline corrected (except for 2D spectra, which were automatically corrected), and the chemical shifts were calibrated by setting the signal of the internal reference (TSP) to 0.0 ppm using TOPSPIN software (Version 3.2, Bruker). The average line width of the TSP peak in all spectra was 1.13±0.18 Hz (average ± SD). Peak assignments were achieved by using information from the two-dimensional spectra, by comparison of chemical shift values, multiplicities and *J*-coupling with online databases such as the Human Metabolome Database (HMDB), and also by using Chenomx NMR suite professional software (Version 8.0, Chenomx Inc.). For metabolite quantification, a ^{1}H-NMR spectrum was acquired with the same parameters described above except

for the relaxation delay, which was increased to 55 seconds to ensure a fully relaxed spectrum. Quantification of the metabolites were obtained from the partially relaxed spectra and then multiplied by the correction coefficient obtained from the comparison of fully relaxed and partially relaxed spectra.

Analysis of the intracellular metabolites by LC/MS

Samples were collected at the same growth stage as for NMR analysis (Fig. 3.2). Briefly, 600 µl of total broth were fast-filtered (Sartolon polyamide 0.45 µm, Sartorius, Goettingen, Germany) and the filter was washed once with 1 ml of a NaCl solution (1%). Subsequently, the washed filter was transferred into a petri dish (50 mm diameter) containing 3 ml of cold methanol solution (80%) and 200 µl of a fully ^{13}C-labeled *E. coli* cell extract (used as internal standard). The extraction of intracellular metabolites was allowed to proceed for 15 minutes at low temperature (below 0°C). The resulting extracts were stored at -80°C. For each condition (90 and 97°C), three technical replicates were collected and analyzed.

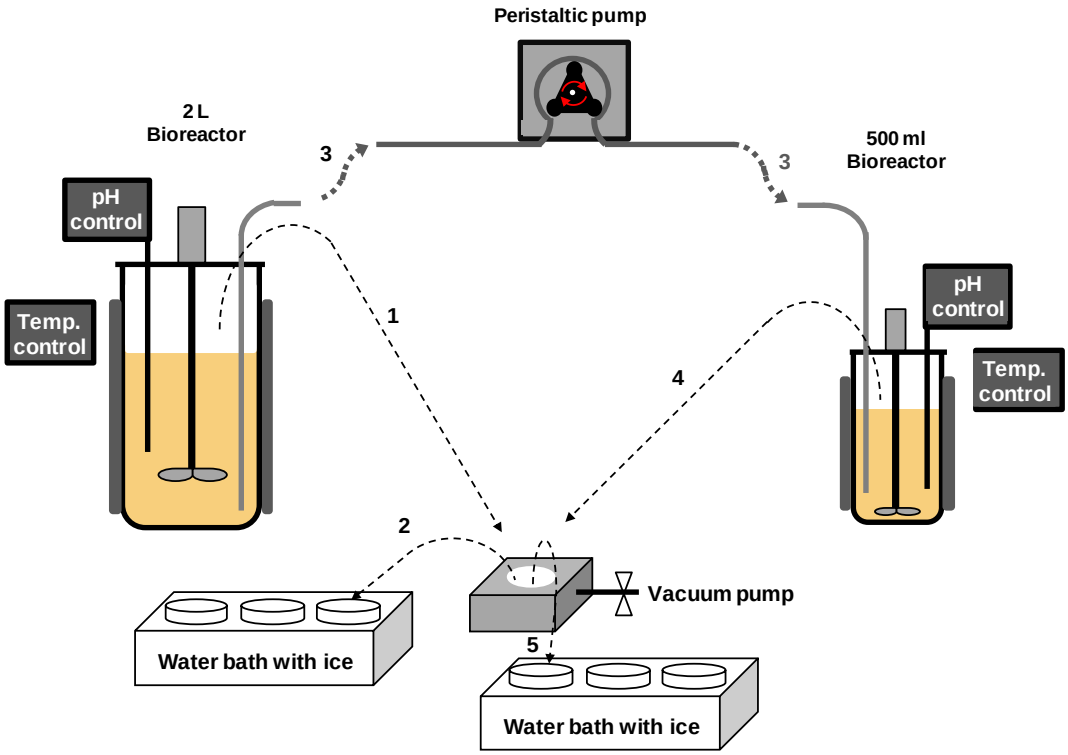


FIGURE 3.2 Schematic representation of the experimental set-up used to impose heat shock and to perform sampling, quenching, and intracellular metabolite extraction for MS analysis. Dashed arrows represent the different steps in the experimental procedure: **Step 1:** An exact culture volume of 0.6 ml was removed from the fermentation vessel at 90°C and fast-filtered; **Step 2:** Immediately, the filter was immersed into cold methanol; **Step 3:** Simultaneously, approximately 250 ml of culture were transferred using a peristaltic pump into a 500 ml bioreactor pre-heated at 97°C; and Step 4: After 45 minutes, 0.6 ml of culture was removed and treated as described above.

Preparation of cell extracts for LC/MS analysis

The methanol present in each of the extracts was evaporated for 4 h (SC110A SpeedVac Plus, ThermoFisher, MA, USA). The remaining aqueous extracts were freeze-dried, re-suspended in 200 µl of milliQ water and stored at −80°C. Intracellular metabolites were analyzed as previously described (Bolten et al. 2007; Kiefer et al. 2007). Briefly, analysis was performed by high performance anion exchange chromatography (Dionex ICS 2000 system, CA, USA) coupled to a triple quadrupole QTrap 4000 (AB Sciex, CA, USA) mass spectrometer. All samples, derived from approximately 40 µg of biomass, were analyzed in the negative mode by multiple reaction monitoring (MRM). For metabolomic experiments, the amounts of metabolites associated with glycolysis (phosphoenolpyruvate (PEP), 1,3-bisphosphoglycerate (BPG), and combined pools of 2- and 3-phosphoglycerate (2PG and 3PG), Pentose Phosphate Pathway (ribose-1-phosphate (Rib1P)), Krebs cycle (malate (Mal), succinate (Succ), aconitate (Aco), and a combination of citrate and isocitrate (Cit and IsoCit)), as well as, nucleotides (ADP, AMP, ATP, GDP-mannose (GDP-man) were determined. Intracellular concentrations of Trehalose-6-phosphate (Tre6P), Glycerol-3-phosphate (Gly3P) and 2-hydroxyglutarate (2-OHGlu) were also measured. To ensure highly accurate quantification, the isotope dilution mass spectrometry (IDMS) method was used (Wu et al. 2005). In this method, an isotopomer of the analyte of interest is used as an internal standard. The required isotopomers were obtained through the cultivation of *E. coli* in 100% ^{13}C-labeled substrates. Therefore, by the addition of

labeled cell extract to both the unlabeled calibration standards and the samples prior to extraction, absolute metabolite concentrations can be obtained.

Protein quantification

Immediately before culture sampling for metabolome analysis, a volume of culture was removed and measured gravimetrically. Then, cells were harvested by centrifugation for 7 minutes at 5,000×*g* and 4°C. The cell pellet was re-suspended in 500 µl of water and disrupted by sonication. The total protein was quantified using the method of Pierce BCA protein assay kit (Thermo).

Results

Methods for culture sampling and metabolic quenching for NMR analysis

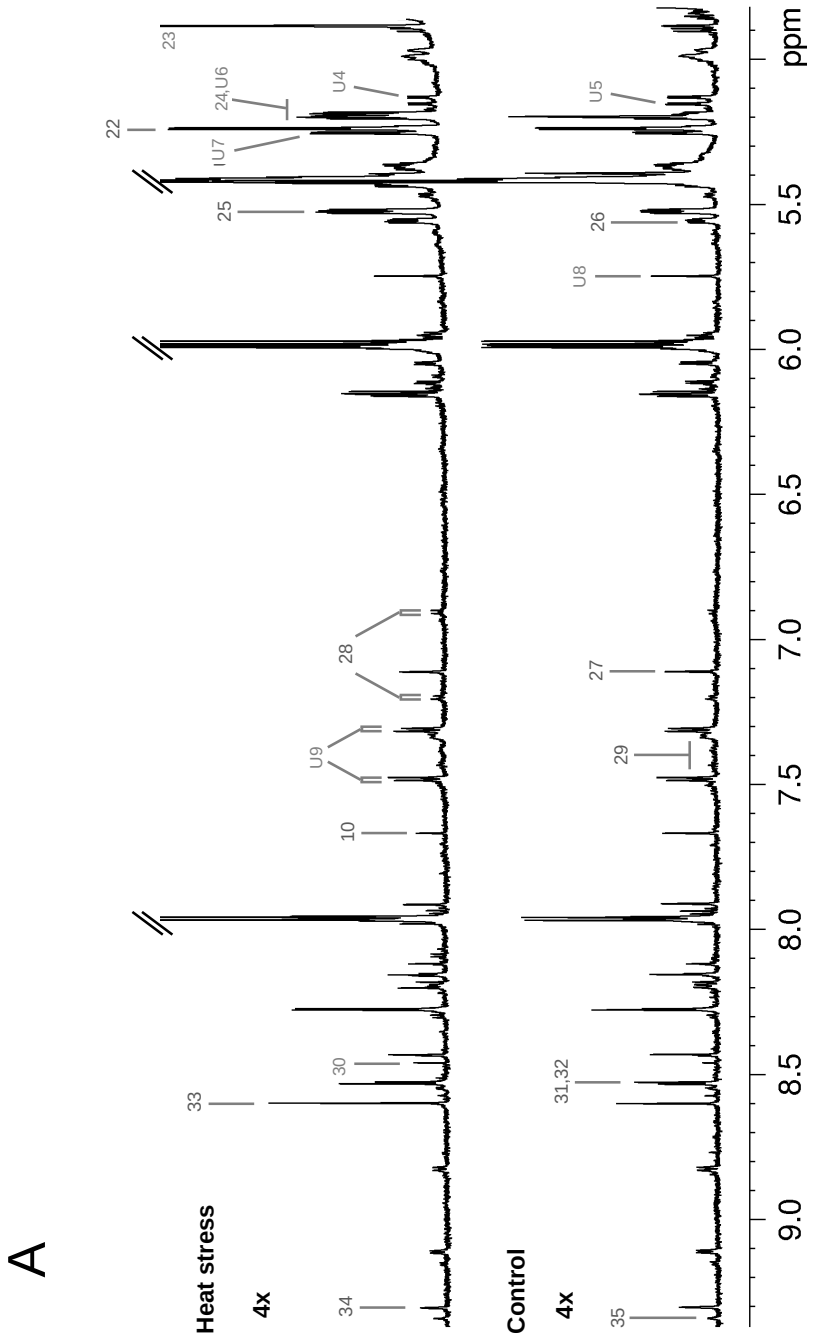
When studying the metabolome of microorganisms, culture sampling and quenching of metabolism are considered as critical steps. The metabolic fluxes occur in the range of seconds, so the sampling/quenching step must happen very fast to immediately arrest metabolism and to ensure a reproducible and reliable snapshot of the physiological condition examined (Taymaz-Nikerel et al. 2009, Mashego et al. 2007). Due to the low cell density reached by *P. furiosus* and also to the low sensitivity of NMR, we

determined that the minimal volume of culture needed to obtain a high quality NMR spectrum is around 250 ml. In addition, two quenching methods were tested: 1) the 250 ml culture at 97ºC was mixed with 250 ml of cold isotonic solution (4ºC). The resulting suspension had a final temperature of around 40ºC; and 2) the 250 ml culture was passed through a steel coil immersed in an ice bath (1ºC); the final temperature of the culture dropped from 97ºC to 11-16ºC in less than 1 minute. The latter method was selected to quench the metabolism of this hyperthermophile since it produced a lower final temperature. Another important issue is the method for cell harvesting. Filtration is the fastest method for sampling/quenching, but it is suitable only for small volumes (from µl to few ml). Nevertheless, we tested this method, but the filter was blocked after passing approximately 100 ml. Therefore, we selected the separation of the cells from the culture medium by centrifugation at 4ºC. Subsequently, the extracellular metabolites were removed by washing the cell pellet with an isotonic solution.

NMR identification of metabolites present in cell extracts of *P. furiosus*

^{1}H-NMR spectra of methanolic extracts of *P. furiosus* grown at 90ºC and subjected to heat shock at 97ºC are shown in Figure 3.3. After an extensive analysis of the data collected from the 1D and 2D NMR spectra, including ^{1}H and ^{13}C chemical shifts, ^{1}H-^{1}H correlations, and *J*-couplings, a total of 35 metabolites were assigned and confirmed with the help of

Chenomx NMR suite software and freely available online databases (Human Metabolome Database, HMDB; and Biological Magnetic Resonance Bank, BMRB). Identified metabolites included amino acids, organic acids, sugars, polyamines and other organic molecules. The resonances of six metabolites (ATP, ADP, AMP, succinate, riboflavin, and pyridoxine) were confirmed in conditions that mimic the cell extracts (pH and ionic strength). This strategy was also used to assign the resonance at 3.27 ppm; in the mimetic solution glycine betaine produces a signal close to this value, but it is not possible to put forward a definite assignment. According to the Human Metabolome Data Base the resonance of trimethylamine N-oxide (TMAO) has a chemical sift of 3.25 ppm. However, this hypothesis requires validation since TMAO has never been reported in prokaryotes. Interestingly, we detected maltose, which is the carbon source for growth. As the cell pellet was washed with an isotonic solution, we assume that the maltose observed in cell extracts is due to intracellular accumulation. Chemical shifts (^{1}H and ^{13}C) and *J*-couplings for metabolites are presented in Table 3.1.


A
Heat stress
4x
Control
4x
22
23
24,U6
U4
U7
25
U9
28
10
33
30
34
27
29
31,32
35
U8
26
U5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
ppm

B

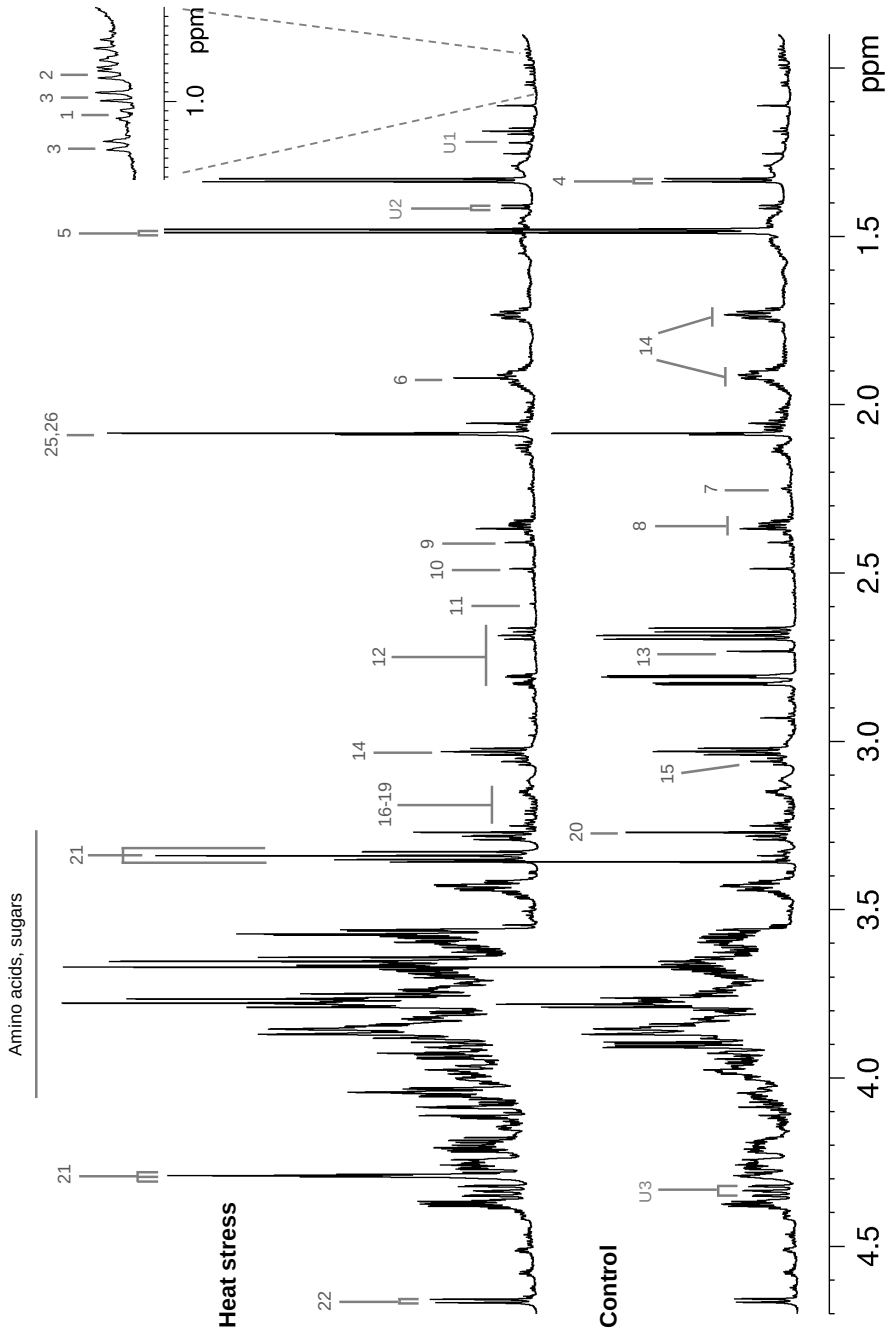

Heat stress
Control
Amino acids, sugars
21
22
25,26
5
6
14
16-19
12
11
10
9
20
15
13
8
7
U1
U2
U3
4
3
2
1
1.0
ppm
4.5
4.0
3.5
3.0
2.5
2.0
1.5
ppm

Figure 3.3 Down- (**A**) and up-fields (**B**) of the water peak (4.8 ppm) of ^{1}H-NMR spectra of the cell extracts of *P. furiosus* grown under 90ºC (optimal) and 97ºC (heat stress) conditions. Extraction of intracellular metabolites was performed with 80% of cold methanol. The resonance assignments are given in Table 3.1. 1, isoleucine; 2, leucine; 3, valine; 4, lactate; 5, alanine; 6, acetate; 7, 2-aminoadipate; 8, glutamate; 9, succinate; 10, pyridoxine; 11, riboflavin; 12, aspartate; 13, dimethylamine; 14, lysine; 15, ornithine; 16, arginine; 17, cystine; 18, choline; 19, putrescine; 20, putative glycine betaine; 21, di-*myo*-inositol-phosphate (DIP); 22, maltose; 23, mannosylglycerate (MG); 24, trehalose; 25, UDP-*N*-acetylglucosamine (UDPGlcNac); 26, UDP-*N*-acetylgalactosamine (UDPGalNac); 27, histidine; 28, tyrosine; 29, phenylalanine; 30, formate; 31, ATP; 32, ADP; 33, AMP; 34, $NADP^{+}$; 35, NAD^{+}. U, unassigned peaks. Spectra region between 4.8 and 9.4 ppm was amplified 4 fold.

Despite our efforts, 9 signals remain unassigned (Table 3.2). Signal labeled as U6 appears at 5.19 ppm and according to the *J*-resolved spectrum is part of a multiplet. The HSQC spectra show a ^{1}H-^{13}C correlation at 5.19 - 103.07 ppm, while TOCSY revealed a correlation between this signal and the region between 3.58 and 3.61 ppm. According to *J*-resolved spectra, the second unknown pattern, U3, is a double doublet (4.32 and 4.35 ppm). TOCSY spectra showed correlations between signals at 4.32-4.35 and 4.83 ppm; in addition, HSQC unveiled correlations between a carbon signal at 72.06 ppm and a broad proton signal at 4.32-4.35 ppm, and between 80.6 ppm and 4.83 ppm (Table 3.1). Unfortunately data collected from NMR spectra was not sufficient to determine the structure of unknown U3 and U6. To assist the identification of unknown signals, 1D selective TOCSY was performed for all the unassigned resonances (Sandusky et al. 2011). This method is based on the selective irradiation of a single resonance and it unravels ^{1}H-^{1}H correlations that may be obscured by overlapping resonances in 2D TOCSY spectra. In most cases, this methodology was unsuccessful. The selective irradiation of the resonance corresponding to U3 showed a proton-proton coupling with a signal at 4.26 ppm (data not shown).

TABLE 3.1 Proton and carbon chemical shifts, multiplicity, and *J*-coupling of metabolites identified.

Nº	Metabolite	Assigment	δ ^{1}H (multiplicity, *J*)	δ ^{13}C
1	Isoleucine	γCH_3	1.01 (d, 7.15 Hz)	n.d.
2	Leucine	αCH_3	0.98 (d, 6.50 Hz)	n.d.
3	Valine	$\gamma' CH_3$	1.05 (d, 7.04 Hz)	n.d.
4	Lactate	CH_3	1.33 (d, 6.94 Hz)	22.89
5	Alanine	CH_3	1.48 (d, 7.28 Hz)	18.8
6	Acetate	CH_3	1.92 (s)	23.4
7	2-aminoadipate	$\gamma' CH_2$	2.24 (t, 7.43 Hz)	n.d.
8	Glutamate	γCH_2	2.35 (m)	36
9	Succinate	CH_2	2.41 (s)	n.d.
10	Pyridoxine	CH_3	2.49 (s)	n.d.
11	Riboflavin	CH_3	2.49 (s)	n.d.
12	Aspartate	CH_2	2.68 (dd, 8.92, 17.44 Hz)	39.5
13	Dimethylamine	CH_3	2.73 (s)	41.5
14	Lysine	εCH_2	3.03 (t, 7.65 Hz)	41.65
15	Ornithine	εCH_2	3.06 (t, 7.65 Hz)	41.65
16	Arginine	δCH_2	3.25 (t, 6.79 Hz)	43.5
17	Cystine	CH_2	3.20 (dd, 8.12,14.80 Hz)	n.d.
18	Choline	CH_3	3.20 (s)	n.d.
19	Putrescine	CH_2	3.11 (t)	n.d.
20	Putative Glycine betaine	CH_3	3.27 (s)	56.2
21	DIP	C_5H	3.34 (t, 9.32 Hz)	77
22	Maltose	C_1H	4.66 (d, 8.02 Hz)	98.3
23	MG	C_1H, Man	4.88 (d, 1.67 Hz)	101.3
24	Trehalose	C_1H	5.20 (d, 3.92 Hz)	96.1
25	UDPGlcNac	C_1H, Glc	5.52 (q, 3.35 Hz)	97.1
26	UDPGalNac	C_1H, Gal	5.55 (q, 3.35 Hz)	97.2
27	Histidine	C_5H, ring	7.11 (s)	n.d.
28	Tyrosine	$C_{2,6}H$, ring	7.20 (d, 8.55 Hz)	n.d.
29	Phenylalanine	$C_{3,5}H$, ring	7.42 (d, 7.43 Hz)	n.d.
30	Formate	CH	8.46 (s)	n.d.
31	ATP	C_8H, Ade	8.52 (s)	n.d

TABLE 3.1 Continuation.

Nº	Metabolite	Assigment	δ ^{1}H (multiplicity, *J*)	δ ^{13}C
32	ADP	C_8H, Ade	8.53 (s)	n.d
33	AMP	C_8H, Ade	8.59 (s)	n.d
34	$NADP^+$	CH	9.30 (s)	n.d
35	NAD^+	CH	9.34 (s)	n.d

s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; n.d., not determined. The meaning of abbreviations is indicated in the legend of Figure 3.2.

TABLE 3.2 Proton and carbon chemical shifts, multiplicity, and *J*-coupling of unidentified signals.

Peak	δ ^{1}H (multiplicity, *J*)	δ ^{13}C
U1	1.22 (s)	n.d
U2	1.41 (d, 6.92 Hz)	24.5
U3	4.34 (ddd, 2.22 Hz)	n.d
U4	5.13 (d, 4.10 Hz)	n.d
U5	5.15 (d, 4.10 Hz)	n.d
U6	5.19 (m)	n.d
U7	5.25 (d, 3.98 Hz)	n.d
U8	5.75 (s)	104.1
U9	7.48 (d, 8.03 Hz)	n.d

s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; n.d., not determined.

NMR analysis of the effect of heat shock in the metabolism of *P. furiosus*

Among the 35 metabolites identified by NMR, 27 were successfully quantified. The intracellular level of 18 metabolites changed with the heat shock imposed. The intracellular concentration of the two well-known compatible solutes of *P. furiosus* increased in response to heat shock

(from 90 to 97ºC): di-*myo*-inositol-phosphate (DIP) and mannosylglycerate (MG) increased approximately 5- and 7-fold, respectively (Fig. 3.4). The intracellular concentration of maltose increased slightly (2.4-fold) upon heat shock (from 0.0067 to 0.0163 µmol/mg of protein). The amino acids (aspartate and alanine) were identified as major metabolites under optimal growth conditions. However, a sudden increment in temperature had a strong negative effect in the intracellular concentration of both amino acids, the level of aspartate decreasing nearly 3-fold (from 0.0247 to 0.0087 µmol/mg of protein). In the case of alanine, the results were highly variable among replicates (Fig. 3.4), which could indicate that factors other than heat shock affect the level of this amino acid. New experiments are planned to improve the precision of the alanine values under heat shock.

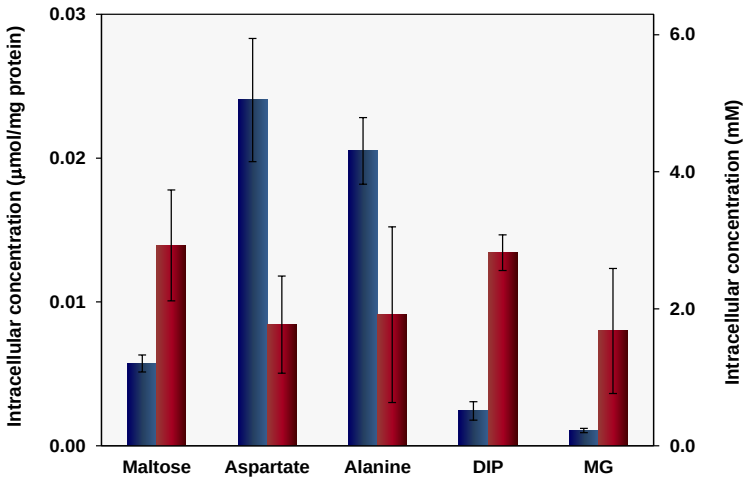


FIGURE 3.4 Intracellular concentration of major metabolites identified in cell extracts of *P. furiosus*. Optimal growth condition is represented in blue, while heat shock is depicted in red. Data correspond to the average of three independent

replicates for each condition examined. Sodium 3-(trimethylsilyl) proprionate-2,2,3,3-d_4 (TSP) was used as external reference for metabolite quantification. Intracellular concentration (mM) was determined based on a cell volume of 4.5 µl/mg protein (Martins and Santos 1995).

The amino acids lysine, ornithine, and glutamate showed an increment in response to heat shock, their concentration increasing approximately by 2 fold under heat shock (Fig. 3.5A). This result (around 2-fold increase) was also observed for UDP-*N*-acetylglucosamine (UDPGlcNac) and UDP-*N*-acetylgalactosamine (UDPGalNac).

Sixteen metabolites were detected in vestigial amounts (Fig. 3.5B); for example, formate has a signal to noise (S/N) of 4.91 as compared to S/N of 351.65 observed for UDPGlcNac, a major metabolite. Due to the low signal to noise, the pool sizes of these metabolites showed large standard deviations (> 30%). Only 8 metabolites (AMP, ADP, leucine, tyrosine, NAD^+, 2-aminoadipate, formate, and riboflavin) responded positively to an increase of temperature. Signals due to most of the nine unassigned compounds did not respond to the imposed heat shock; however, resonances labeled as U3 and U6 had a positive correlation with heat shock, the areas increasing from 0.166 and 0.331 to 0.268 and 0.661, respectively.

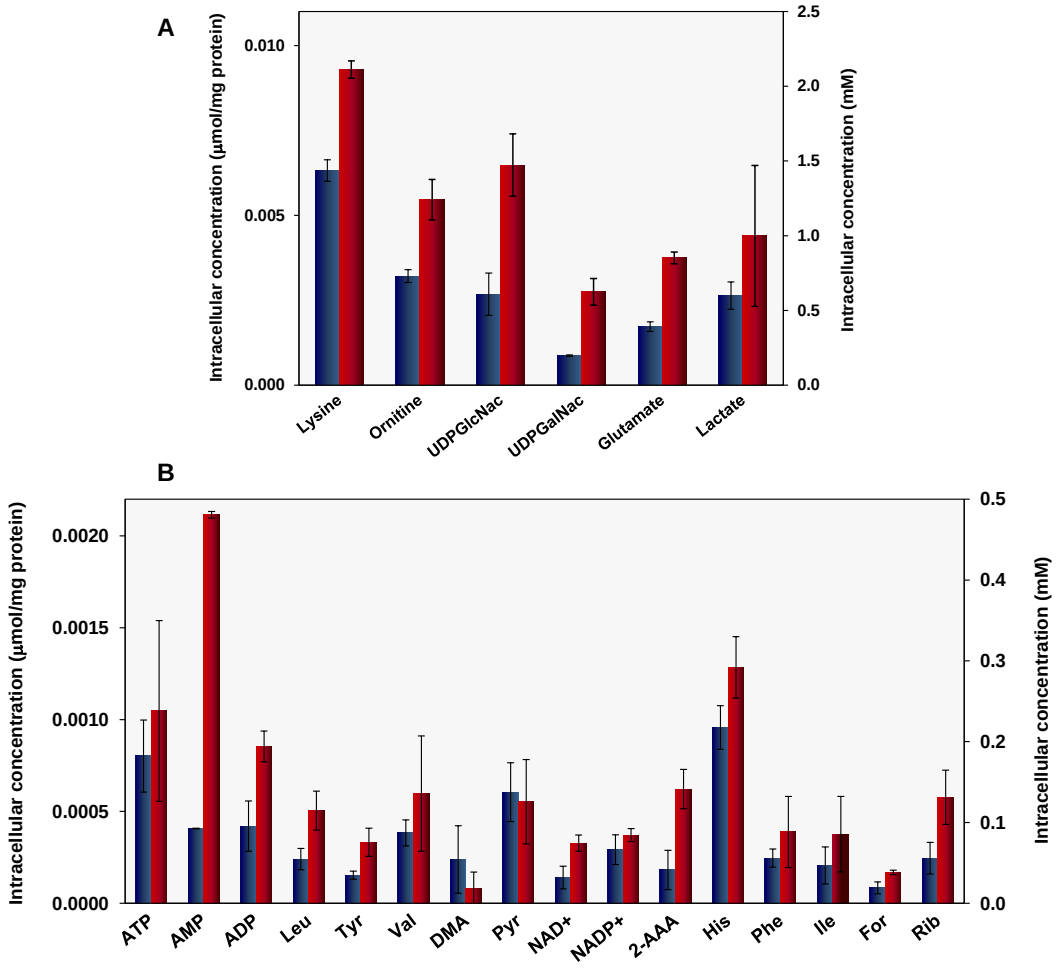


FIGURE 3.5 Intracellular concentration of minor metabolites identified in cell extracts of *P. furiosus*. Optimal growth condition is represented in blue, while heat shock is depicted in red. Data correspond to the average of three independent replicates for each condition examined. Sodium 3-(trimethylsilyl) proprionate-2,2,3,3-d_4 (TSP) was used as external reference for metabolite quantification. DMA, dimethylamine; Pyr, pyridoxine; 2-AAA, 2-aminoadipate; Phe, phenylalanine; For, formate; and Rib, riboflavin. Intracellular concentration (mM)

was determined based on a cell volume of 4.5 µl/mg protein (Martins and Santos 1995).

The effect of heat shock in the metabolism of *P. furiosus* as investigated by LC/MS

In total, the absolute concentrations of 9 metabolites were determined by mass spectrometry, including, ATP, ADP, AMP, GDP-mannose, 2-phosphoglycerate plus 3-phosphoglycerate, phosphoenolpyruvate, glycerol-3-phosphate, ribose-1-phosphate, and trehalose-6-phosphate (Table 3.3). With the exception of GDP-man and ATP, all these metabolites decreased at supraoptimal temperature.

TABLE 3.3 Absolute quantification of metabolites carried out by LC/MS in cell extracts of *P. furiosus* under optimal and heat shock conditions.

Metabolites	Intracellular concentration (µmol/mg protein)[a]	
	Optimal growth (90°C)	Heat shock (97°C)
2/3PG	0.0076 ± 0.0023 (1.68)	0.0029 ± 0.0003 (0.64)
ADP	0.0174 ± 0.0005 (3.88)	0.0129 ± 0.0015 (2.88)
AMP	0.0053 ± 0.0013 (1.18)	0.0027 ± 0.0010 (0.60)
ATP	0.0075 ± 0.0010 (1.67)	0.0111 ± 0.0008 (2.48)
GDP-Man	0.0004 ± 4.9E-05 (0.09)	0.0007 ± 0.0001 (0.15)
Gly3P	0.0103 ± 0.0016 (2.30)	0.0051 ± 0.0008 (1.13)
PEP	0.0065 ± 0.0011 (1.46)	0.0024 ± 0.0003 (0.52)
Rib1P	0.0003 ± 0.0003 (0.11)	n.d.
Tre6P	0.0016 ± 0.0017 (0.53)	n.d.

[a] Values are averages of three technical replicates.
The intracellular concentrations (mM), based on a cell volume of 4.5 µl/mg protein (Martins and Santos 1995), are given in parentheses.

Due to technical difficulties, absolute quantification was not possible for 5 metabolites (malate, citrate/isocitrate, succinate, cis-aconitate, and 1,3-bisphosphoglycerate); relative quantification data are presented (Table 3.4). The LC/MS analysis was performed by Dr. Lindsay Peyriga at MetaToul (Metabolomic and Fluxomic platform, Toulouse).

TABLE 3.4 Relative quantification of metabolites by LC/MS in cell extracts of *P. furiosus*.

Metabolites	Relative quantification	
	Optimal growth (90°C)	Heat shock (97°C)
Malate*	0.1122 ± 0.0198	0.0232 ± 0.0118
Succinate*	0.2185 ± 0.0994	0.0208 ± 0.0035
Citrate+Isocitrate*	0.5386 ± 0.3308	0.0379 ± 0.0108
BPG*	0.1297 ± 0.0119	0.0523 ± 0.0084
Cis-aconitate#	679.8 ± 449.8	76.9 ± 24.6
2-OHGlu#	2996.4 ± 411.4	638.3 ± 78.1

* Relative quantification: ratio ^{12}C area /^{13}C area per mg of protein

Relative quantification: ^{12}C area per mg of protein

Discussion

Metabolomic studies have usually two main objectives: to find new metabolites and identify meaningful changes in the concentration of all the metabolites in response to an external or internal stimulus. Here, we studied the metabolome of a model hyperthermophilic archaeon, *Pyrococcus furiosus*, under optimal and heat shock conditions. To our knowledge, this is the first metabolome analysis of the heat shock response in a hyperthermophilic organism. So far, only two studies have been reported on the metabolome of two hyperthermophiles, *P. furiosus*

and *Sulfolobus solfataricus*, but they had the goal of investigating the effect of cold shock (Kalisiak et al. 2008, Zaparty et al. 2010).

In this work, the metabolome of *P. furiosus* was studied using two complementary techniques (NMR and MS). The severity of the heat shock challenge was the same (45 minutes at 97ºC) regardless the analytical method used, however the disparity of sampled volumes imposed very different methods for cell separation, i.e., centrifugation for NMR and fast-filtration for LC/MS. These differences are due to the low sensitivity of NMR and the low cell yield of *P. furiosus*. For NMR analysis, the culture sample (250 ml) was pushed through a steel coil immersed into an ice bath, then cells were harvested and the pellet was treated to extract intracellular metabolites. For MS analysis, the culture sample (0.6 ml) was fast-filtered and the intracellular metabolites extracted immediately. The direct culture sampling into a cold organic solution is the most common process used for quenching metabolism, but has the disadvantage of provoking metabolite leakage (Taymaz-Nikerel et al. 2009). Therefore, the use of cold organic solutions for quenching was not considered in this work. MS analysis was used to identify and quantify the metabolites with high turnover rates and low intracellular concentrations, while the NMR analysis is more suitable for the determination of the pool of final metabolites. Using these two techniques, a total of 35 metabolites were quantified (Table 3.5).

TABLE 3.5 Metabolite quantification by NMR and LC/MS (*) under optimal and heat shock conditions. Color represents positive (red) and negative (blue) variation of metabolites under heat shock.

Category	Metabolite	Optimal condition (μmol/mg protein)	Heat shock condition (μmol/mg protein)	Log_2 (Heat/ Optimal)	Fold
Hexose derivative	MG	0.0011 ± 0.0001	0.0080 ± 0.0043	**2.91**	**7.50**
Polyol	DIP	0.0024 ± 0.0006	0.0134 ± 0.0012	**2.47**	**5.54**
Nucleotide	AMP	0.0004 ± 0.0000	0.0017 ± 0.0007	**2.05**	**4.15**
Sugar nucleotide	UDPNacGal	0.0009 ± 0.0000	0.0027 ± 0.0004	**1.66**	**3.16**
Disaccharide	Maltose	0.0057 ± 0.0006	0.0139 ± 0.0039	**1.28**	**2.43**
Sugar nucleotide	UDPNacGlc	0.0027 ± 0.0006	0.0065 ± 0.0009	**1.27**	**2.42**
Cofactor	Riboflavin	0.0002 ± 0.0001	0.0006 ± 0.0001	**1.23**	**2.35**
Cofactor	NAD^+	0.0001 ± 0.0001	0.0003 ± 0.0000	**1.22**	**2.34**
Amino acid	Glutamate	0.0017 ± 0.0001	0.0037 ± 0.0002	**1.12**	**2.18**
Amino acid	Tyrosine	0.0002 ± 0.0000	0.0003 ± 0.0001	**1.12**	**2.17**
Amino acid	Leucine	0.0002 ± 0.0001	0.0005 ± 0.0001	**1.07**	**2.10**
Nucleotide	ADP	0.0004 ± 0.0001	0.0009 ± 0.0001	**1.02**	**2.03**
Organic acid	Formate	0.0001 ± 0.0000	0.0002 ± 0.0000	**1.00**	**1.99**
Amino acid	Isoleucine	0.0002 ± 0.0001	0.0004 ± 0.0002	**0.87**	**1.83**
Polyamine	Betaine[?]	0.0006 ± 0.0002	0.0011 ± 0.0000	**0.83**	**1.78**
Sugar nucleotide	GDP-Man*	0.0004 ± 0.0000	0.0007 ± 0.0001	**0.82**	**1.76**
Amino acid	Ornitine	0.0032 ± 0.0002	0.0055 ± 0.0006	**0.77**	**1.70**
Organic acid	Lactate	0.0026 ± 0.0004	0.0044 ± 0.0021	**0.74**	**1.67**
Amino acid	Phenylalanine	0.0002 ± 0.0000	0.0004 ± 0.0002	**0.65**	**1.57**
Amino acid	Valine	0.0004 ± 0.0001	0.0006 ± 0.0003	**0.64**	**1.56**
Organic acid	2 aminoadipate	0.0004 ± 0.0004	0.0006 ± 0.0001	**0.63**	**1.54**
Nucleotide	ATP*	0.0075 ± 0.0010	0.0111 ± 0.0008	**0.56**	**1.48**
Amino acid	Lysine	0.0063 ± 0.0003	0.0093 ± 0.0003	**0.56**	**1.47**
Amino acid	Histidine	0.0010 ± 0.0001	0.0013 ± 0.0002	**0.42**	**1.34**
Nucleotide	ATP	0.0008 ± 0.0002	0.0010 ± 0.0005	**0.38**	**1.31**
Cofactor	$NADP^+$	0.0003 ± 0.0001	0.0004 ± 0.0000	**0.35**	**1.27**
Cofactor	Pyridoxin	0.0006 ± 0.0002	0.0006 ± 0.0002	**-0.13**	**-1.09**
Nucleotide	ADP*	0.0174 ± 0.0005	0.0130 ± 0.0015	**-0.43**	**-1.35**
Nucleotide	AMP*	0.0053 ± 0.0013	0.0027 ± 0.0010	**-0.98**	**-1.97**
Polyol	Gly-3P*	0.0104 ± 0.0016	0.0051 ± 0.0008	**-1.03**	**-2.04**
Amino acid	Alanine	0.0205 ± 0.0023	0.0091 ± 0.0061	**-1.17**	**-2.25**
Triose	2/3-PG*	0.0076 ± 0.0024	0.0029 ± 0.0003	**-1.39**	**-2.62**
	PEP*	0.0066 ± 0.0011	0.0024 ± 0.0004	**-1.47**	**-2.77**
Amino acid	Aspartate	0.0240 ± 0.0043	0.0084 ± 0.0034	**-1.51**	**-2.85**
polyamine	Dimethylamine	0.0002 ± 0.0002	0.0001 ± 0.0001	**-1.55**	**-2.93**

* Metabolites detected and quantified using LC/MS.

[?] Putative glycine betaine.

Typical metabolomic studies of mesophilic bacteria and plants allow the identification/quantification of a higher number of metabolites (Sévin and Sauer 2014, Wu et al. 2014). For instance, Sévin and Sauer studied the response of *E. coli* to hyperosmotic shock using non-targeted mass spectrometry analysis and claimed to detect approximately 1000 metabolites (Sévin and Sauer 2014).

An overview of the metabolic changes caused by heat shock in *P. furiosus* is shown in Figure 3.6. The results indicate that heat shock led to a reduction in the pool of metabolites below 3PG and to an increase in the metabolites derived from the upper part of glycolysis. This metabolic shift led to an increase in the accumulation of two compatible solutes, DIP and MG. The preferential accumulation of DIP and DIP-derivatives during adaptation to supraoptimal temperatures has been amply demonstrated, while MG responds preferentially to supraoptimal salinity conditions (Santos et al. 2011). Therefore, it is curious to observe an increment of MG upon heat shock in *P. furiosus* strain DSM 3638. Actually, this result agrees with an earlier study that compared the pools of organic solutes in *P. furiosus* cells grown at 90º and 98ºC (Martins and Santos 1995). However, it should be noted that at temperatures higher than 98ºC the MG pool vanishes (Martins and Santos 1995).

We were surprised by the low levels of DIP and MG accumulating in cells subjected to heat shock, i.e., around 3.0 mM and 1.8 mM, respectively. These pools are about 30 to 40 times lower when compared with *P. furiosus* cells grown at optimal (90ºC) and heat stress conditions

(98ºC). We think that these differences are primarily related with the higher salinity in the growth medium used by Martins and Santos (1995), i.e., 2.8% NaCl versus 1% NaCl in the present study. Moreover, the long term adaptation allowed for in the earlier studies is not comparable with the short term response (45 min at 97ºC), imposed in the heat shock experiment.

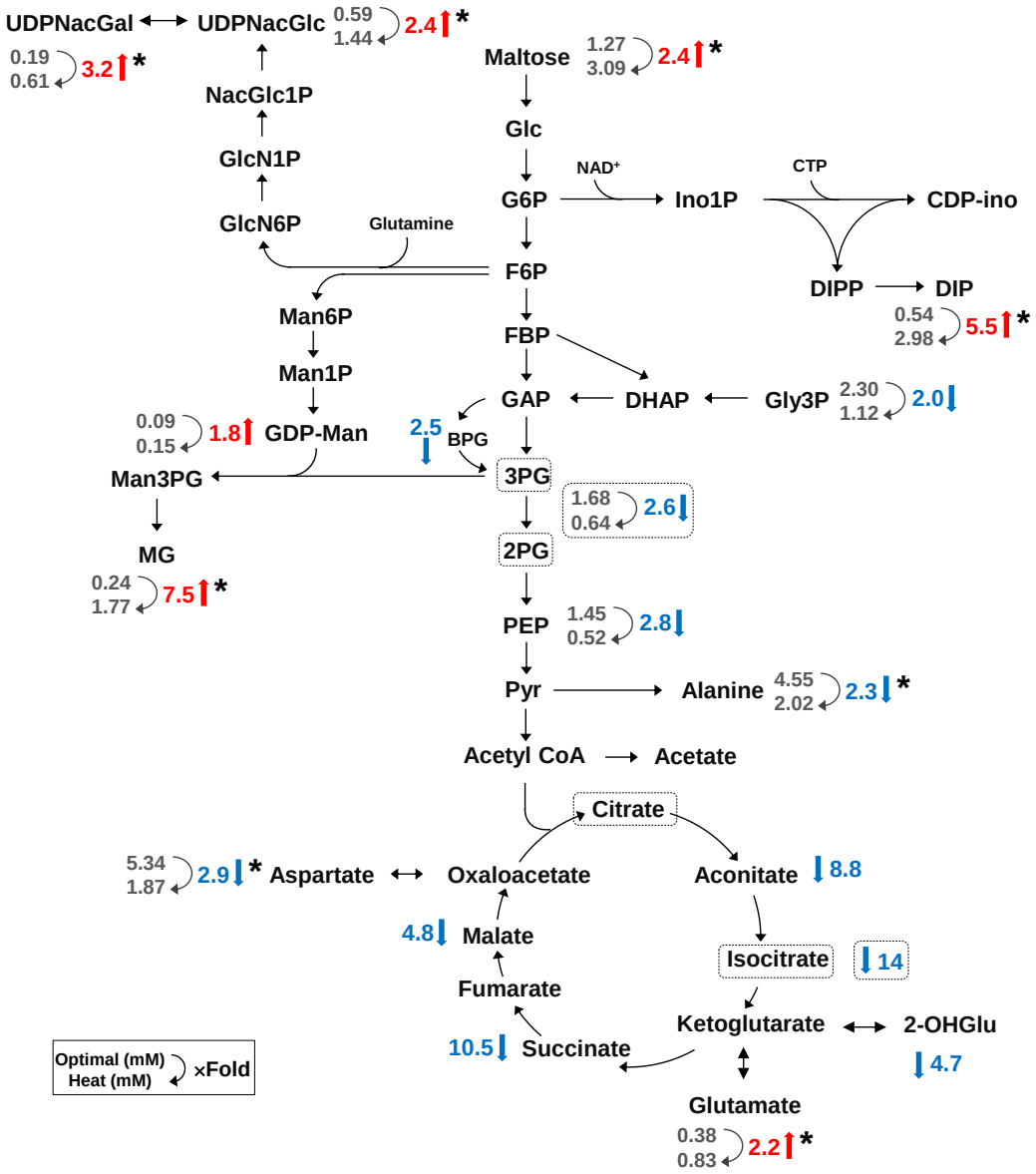


FIGURE 3.6 Overview of the metabolic changes induced by heat shock. Red arrows represent metabolites which responded positively to heat shock, while blue arrows represent the opposite variation. Metabolites detected by NMR are

labeled with an asterisk. Variation of 1,3-bisphosphoglycerate, malate, succinate, combination of citrate and isocitrate, 2-OHGlu, and cis-aconitate were obtained from relative quantification by MS (Table 3.5).

Curiously, GDP-mannose, one of the substrates for MG synthesis, increased slightly with heat shock while the pool of the second substrate, 2/3PG, decreased. Moreover, the pool sizes of glycolytic metabolites below 2/3PG or involved in the Krebs cycle were negatively affected. In accordance, the transcript levels of the gene encoding GAPOR, the enzyme that leads to the formation of 3PG, decreases at supraoptimal temperatures (Shockley et al. 2003). Since the final pool of MG increased it is clear that the net flux towards MG was enhanced under heat shock. In *P. horikoshii* the activity of MPGS did not change significantly in cells grown at 90ºC or at 97ºC (Empadinhas et al. 2001), but the effect in *P. furiosus* has not been investigated. We performed RT-PCR of the *mpgs* gene, but unfortunately the results were highly irreproducible (results not shown). It would be interesting to complement our results with transcriptomic and proteomic data of the enzymes involved in the synthesis of MG and its precursors.

In this study, we describe a positive correlation between maltose concentration and an increase in temperature. This disaccharide was used as the main carbon and energy sources for cell growth. The uptake and metabolism of maltose was characterized in *P. furiosus* at the transcriptional level (Lee et al. 2006): the import of maltose is mediated by Mal-I transporter (PF1739 to PF1744), which has homology to bacterial

transporters Mal (Noll et al. 2008). Intracellular maltose is cleaved by the action of α-glucosidase (PF0132) into glucose which is subsequently phosphorylated and enters glycolysis (Lee et al. 2006). As mentioned above, the heat shock response was studied at the transcriptional level in *P. furiosus* with cells grown on tryptone and yeast extract (Shockley et al. 2003). As expected, there was an increase in the expression level of genes coding for chaperone-related proteins and proteases; importantly, a drastic increase (around 5-fold), in the expression levels of maltose transporters and glycoside hydrolase genes was observed. The increment in the expression of genes related to carbohydrate transport and degradation was interpreted as a need to counterbalance the ATP depletion caused by heat stress (Shockley et al. 2003). The assumption of a higher energy demand by cells under heat stress is supported by several transcriptomic and proteomic studies carried out with *E. coli* (Nonaka et at. 2006, Hasan and Shimizu 2008). In particular, Ye and co-authors (2012) reported the drop of the intracellular glucose pool and proposed an increase in the glycolytic flux on the basis of augmented expression levels of glycolytic genes. In this context, our results on the accumulation of maltose in *P. furiosus* at supraoptimal temperature were unexpected. It seems as though heat shock slows down maltose metabolism. We speculate that maltose contributes to the mechanism of cell protection at elevated temperature. The osmoprotective role of maltose was confirmed in the bacterium *Sinorhizobium meliloti* (Gouffi et al. 1999).

The compounds UDP-*N*-acetylglucosamine and UDP-*N*-acetylgalactosamine were identified in cell extracts of *P. furiosus* grown under optimal temperatures at the late logarithmic phase (Ramakrishnan et al. 1997). In our work, the pool of these metabolites falls in the same concentration range of DIP. Interestingly, when *P. furiosus* cells were shocked at supraoptimal temperatures the presence of these two metabolites increased by 3-fold. It is known that UDPGlcNac and UDPGalNac are involved in the biosynthesis of morphological structures, such as peptidoglycan and S-layers proteins (Hartamann and König 1990, König et al. 1994, Ramakrishnan et al. 1997). Interestingly, the relationship between thermal death and degradation of the cell structure was investigated in *Thermococcus* sp. strain Tc-1-95 (Mitsuzawa et al. 2006), and a considerable deformation of the S-layer was caused by lethal thermal stress (i.e. from 92ºC to 119ºC for 1.5 seconds). Therefore, the increased pools of UDPGlcNac and UDPGalNac could be related with a defense mechanism aimed at maintaining the integrity of the cell structure.

The pool of aspartate decreased in response to heat shock. Aspartate can be taken up from the medium or synthesized *de novo* by the aspartate aminotransferase (PF0522) through transamination of oxaloacetate, an intermediate of the Krebs cycle. The osmoprotecting role of aspartate has been reported in halophilic bacteria (Joghee and Jayaramann 2014) and also in a few members of the order Thermococcales (*Palaeococcus ferrophilus* and *Thermococcus kodakarensis*) (Borges et al. 2010). Therefore, it seems that aspartate is more directed to osmoprotection. So

far, the only exception to this trend was observed in a DIP-deficient *T. kodakarensis* mutant where the lack of DIP was replaced by aspartate under heat stress conditions (Borges et al. 2010).

The identification of 2-aminoadipate (2-AAA) in cell extracts of *P. furiosus* is interesting. It suggests that lysine synthesis may occur through the 2-AAA pathway, usually found in fungi and yeast, while most prokaryotes synthesize this amino acid via the diaminopimelic acid (DAP) pathway, using aspartic acid as precursor. In fact, the synthesis of lysine from 2-oxoglutaric acid and acetyl-CoA via the 2-AAA pathway was earlier reported for the thermophilic bacterium *Thermus thermophilus* (Kosuge and Hoshino 1998). Moreover, the genes encoding the enzymes implicated in the 2-AAA pathway were identified in the genomes of *T. thermophilus* and *P. horikoshii,* a hyperthermophilic archaeon closely related to *P. furiosus* (Kosuge and Hoshino 1998, Nishida et al. 1999). A BLAST search using as reference the diaminopimelic acid decarboxylase from *Archaeoglobus fulgidus*, an enzyme from the DAP pathway, revealed no homologues in the *P. furiosus* genome. This evidence reinforces the view that the synthesis of lysine in this archaeon proceeds via the 2-AAA pathway.

In *P. furiosus*, glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) catalyzes the oxidation of glyceraldehyde-3-phosphate to 3-phosphoglycerate, skipping the formation of the intermediate 1,3-bisphosphoglycerate (Mukund and Adams 1995). Several studies justify the differences in glycolysis as an adaptation of the

organisms to high temperatures since 1,3-bisphosphoglycerate, glyceraldehyde-3-phosphate, and dihydroxyacetone phosphate (DHAP) are known to be unstable at high temperatures (Kouril et al. 2013). However, we observed the presence of 1,3-bisphosphoglycerate in cell extracts of *P. furiosus*, suggesting that glyceraldehyde-3-phosphate dehydrogenase (GAPDH), the enzyme catalyzing the conversion of glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate, is active. This observation is in agreement with literature data that reported this activity in *P. furiosus* (Schäfer and Schönheit 1992).

To conclude, a method for sampling and quenching of large culture volumes was optimized. An integration of the results obtained using NMR and MS supported the conclusion that heat shock favors the synthesis of compatible solutes and N-acetylated NDP-sugars that may be involved in the synthesis of the cell wall. In contrast, the pool sizes of metabolites associated with the glycolysis or the Krebs cycle were negatively affected. Whether these decreased pools are associated with lower fluxes through central metabolism lacks demonstration by complementation with non trivial studies on carbon flux analysis. To our knowledge, this is the first study on the metabolome of a hyperthermophilic organism under heat shock conditions. Further work must be done to improve the statistical significance of the results, extend the number of metabolites identified and validate the preliminary conclusions.

Work contributions

This work results from a collaboration with the group of J. C. Portais from the MetaToul facility (Toulouse). L. Peyriga performed the mass spectrometry analysis. G. Graça contributed to the acquisition and analysis of NMR spectra. I. Torcato helped during the process of culture sampling for NMR and MS analyses.

CHAPTER 4

X-ray structure of the IPCT/DIPPS bifunctional enzyme of *Archaeoglobus fulgidus* and insights into its catalytic mechanism

Part of this chapter is published in:

Nogly P, Gushchin I, Remeeva A, Esteves AM, Borges N, Ma P, Ishchenko A, Grudinin S, Round E, Moraes I, Borshchevskiy V, Santos H, Gordeliy V, Archer M. 2014. X-ray structure of a CDP-alcohol phosphatidyltransferase membrane enzyme and insights into its catalytic mechanism. Nature Communications. 5:4169.

Contents

Summary

Many marine hyperthermophiles accumulate di-*myo*-inositol-phosphate in response to heat stress. The biosynthetic pathway for di-*myo*-inositol-phosphate involves the formation of the intermediate metabolites CDP-inositol and di-*myo*-inositol-phosphate-phosphate via CTP:L-*myo*-inositol-1-phosphate cytidylyltransferase (IPCT) and CDP-L-*myo*-inositol:L-*myo*-inositol-1-phosphate transferase (DIPPS), respectively. Curiously, the C-terminal domain contains a conserved sequence pattern, $DG(x)_2AR(x)_8G(x)_3D(x)_3D$, typical of the CDP-alcohol phosphatidyltransferase family, a group of integral membrane proteins mostly involved in phospholipid biosynthesis. Representatives of this family are found in all domains of life and catalyze the displacement of CMP from a CDP-alcohol by a second alcohol with formation of a phosphodiester bond and concomitant breaking of a phosphoanhydride bond. Despite the crucial role played by these membrane enzymes, no structure was available when this workplan was designed. Here, we report the structures of the IPCT/DIPPS from *Archaeoglobus fulgidus* with and without bound CDP-inositol to 2.5 Å and 2.65 resolution, respectively. The protein was produced in *E. coli*, solubilized with triton X-100 and purified by affinity chromatography. Crystals were grown using an *in meso* crystallization method. The bifunctional protein dimerizes through the DIPPS domains, each comprising six transmembrane (TM) α-helices. The active site cavity is hydrophilic and widely open to the cytoplasm with a magnesium ion surrounded by four highly conserved aspartate residues from helices TM2

and TM3. The structure of IPCT/DIPPS with bound CDP-inositol revealed the location of this substrate binding site in the DIPPS domain. Results were validated by mutagenesis studies. A structure-based catalytic mechanism is proposed.

Introduction

Di-*myo*-inositol-phosphate (DIP) is an organic solute widespread in marine hyperthermophilic bacteria and archaea, and has not been found in microorganisms living optimally at temperatures below 60°C. The strong association between DIP and thermophily is further emphasized by the finding that its level generally increases in response to supraoptimal temperatures (Santos et al. 2011). Therefore, DIP and derivatives are regarded as hallmarks of microbial adaptation to very hot environments.

The synthesis of DIP starts from inositol-1-phosphate and proceeds in three steps (Borges et al. 2006, Rodionov et al. 2007). First, inositol-1-phosphate is activated to CDP-inositol by CTP:inositol-1-phosphate cytidylyltransferase (IPCT), then CDP-inositol reacts with inositol-1-phosphate to yield DIP-phosphate (DIPP) by the action of CDP-inositol:inositolphosphotransferase, also named DIPP synthase (DIPPS), and finally DIPP is dephosphorylated to DIP. In most DIP-producers, such as *Pyrococcus furiosus* and *Archaeoglobus fulgidus*, both activities (IPCT and DIPPS) are present in a single gene product, but

separate genes may also be found (Gonçalves et al. 2012, Rodrigues et al. 2007).

The N-terminus of the bifunctional enzyme IPCT/DIPPS has homology with proteins belonging to the nucleotidyltransferase family and the crystal structure of *A. fulgidus* IPCT has been solved (Brito et al. 2011). Interestingly, five transmembrane helices are predicted in the C-terminal domain of IPCT/DIPPS, which led to the hypothesis that DIPPS is a membrane protein (Rodrigues et al. 2007, Brito et al. 2011). Moreover, DIPPS proteins possess a motif characteristic of the CDP-alcohol phosphatidyltransferase (CDP-OH_P_trans) family, which typically comprises integral membrane proteins mostly involved in the synthesis of phospholipids, such as phosphatidylinositol, phosphatidylethanolamine or phosphatidylcholine. CDP-alcohol phosphatidyltransferases catalyze the displacement of CMP from a CDP-alcohol by a second alcohol with formation of a phosphodiester bond and concomitant breaking of a phosphoanhydride bond (Fig. 4.1).

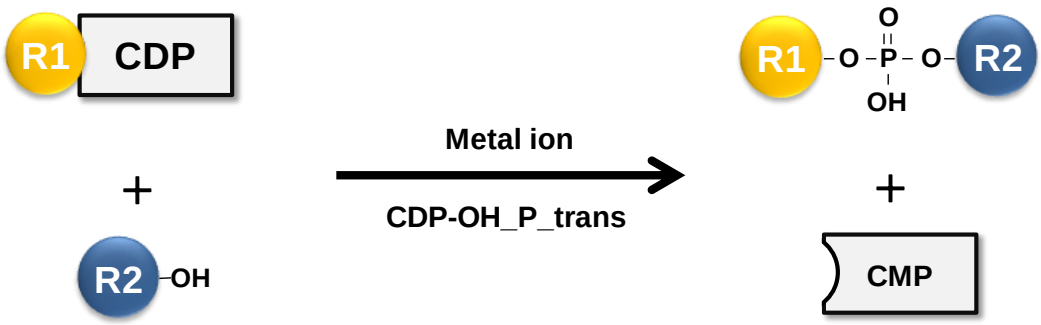


FIGURE 4.1 Representation of the reaction catalyzed by CDP-alcohol phosphatidyltransferases: CMP is displaced from CDP-alcohol (**R1**) by a second alcohol (**R2**) with formation of a phosphodiester bond and concomitant breaking of the phosphoanhydride bond. **R** may represent inositol, inositol-1-phosphate, choline, diacylglycerol, amongst others (adapted from Sciara et al. 2014).

To date, more than 12750 sequences are classified as CDP-OH_P_trans in UniProtKB, but no structure has been yet characterized for any member of this family. Therefore, although our main interest is to understand the structural basis of DIP synthesis, the actual relevance of this study surpasses the field of adaptation to high temperature since it will provide ground-breaking insights into the biosynthesis of phospholipids, major membrane components in all domains of life.

Here we present the three-dimensional structure of an archaeal CDP-alcohol phosphatidyltransferase (DIPPS) coupled with a cytidylyltransferase (IPCT) in its apo form and in complex with CDP-inositol at 2.65 and 2.5 Å resolution, respectively. We analyze the functional role of highly conserved amino acid residues by site-directed mutagenesis studies

and propose a catalytic mechanism. This structural information opens new avenues for homology modeling of other CDP-OH phosphatidyltransferases, namely those with potential medical interest.

Materials and Methods

Production and purification of the DIPPS substrates, L-*myo*-inositol-1-phosphate and CDP-L-*myo*-inositol

L-*myo*-inositol-1-phospate and CDP-L-*myo*-inositol (hereafter abbr. as inositol-1-phosphate and CDP-inositol) were synthesized enzymatically via L-*myo*-inositol-1-phosphate synthase (MIPS) and the truncated IPCT from *A. fulgidus*, respectively.

The production of the recombinant L-*myo*-inositol-1-phosphate synthase was performed as previously described (Rodrigues et al. 2007). Briefly, *E. coli* BL21(DE3) cells, bearing the plasmid with the *mips* gene, were grown at 37ºC in LB medium supplemented with ampicillin (100 μg/ml) to an OD of 0.6 at 600 nm, then treated with 1 mM IPTG and allowed to grow further for 4 hours before harvesting. The cells were harvested by centrifugation, re-suspended in 20 mM Tris-HCl (pH 7.6) containing 10 mM $MgCl_2$ and disrupted in a French press. Cells debris was removed by centrifugation (10,000×*g*, 10 min) and the supernatant solution was heated at 75ºC for 10 minutes to eliminate thermolabile proteins of *E. coli*. After centrifugation (10,000×*g*, 10 min), the supernatant was finally used to obtain inositol-1-phosphate. The reaction mixture had a final

concentration of 20 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 10 mM glucose-6-phosphate, and 5 mM NAD^+. The reaction was allowed to proceed for 1 hour at 85ºC. The production of inositol-1-phosphate was confirmed by ^{1}H-NMR. The mixture obtained was applied onto a QAE-Sephadex A-25 column, equilibrated with buffer A (5 mM sodium bicarbonate pH 9.8). Elution was carried out with a linear gradient from buffer A to buffer B (1 M sodium bicarbonate pH 9.8). Collected fractions were analyzed by ^{1}H-NMR; the samples containing inositol-1-phosphate were applied onto a Dowex 50W-X8 column previously treated with 1 M HCl and washed with distilled water until the pH of the eluate was approximately 6. The desired compound was eluted with distilled water; the sample pH was corrected to 5.3 prior to lyophilization. Inositol-1-phoshate was quantified by ^{1}H-NMR in the final preparation, using formate as a concentration standard.

The recombinant IPCT from *A. fulgidus* was expressed and produced as described previously (Brito et al. 2011). Briefly, *E.* coli cells harboring the plasmid with the *ipct* gene were grown in LB medium (100 μg/ml ampicillin) at 37ºC to an OD of 0.6 at 600 nm. Protein production was induced with 1 mM of IPTG and cell growth was allowed to proceed for further 4 hours. Cells were harvested by centrifugation (10,000×*g*, 10 min), re-suspended in 20 mM Tris-HCl (pH 7.6) containing 10 mM $MgCl_2$, and disrupted in a French press. Cells debris was removed by centrifugation (10,000×*g*, 10 min) and the supernatant was heated at 85ºC for 30 minutes. After centrifugation (10,000×*g*, 10 min), the cell extract was

applied onto a His-Trap column (GE Healthcare BioScience AB, Uppsala Sweden) and the pure IPCT was used for CDP-inositol production. A 2 ml reaction mixture containing 5 mM of inositol-1-phosphate, 5 mM of CTP, 25 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, and 160 μg of IPCT was incubated at 85ºC for 30 minutes. After confirmation of CDP-inositol production by ^{31}P-NMR, the reaction mixture was filtered through a 10 kDa Omega Nanosep filter (Pall Life Sciences, Hampshire, UK) to remove proteins. The filtrate was incubated at 37ºC for 30 minutes with 12 μL of alkaline phosphatase (Sigma) to dephosphorylate the residual amounts of inositol-1-phosphate and CTP. The final sample of CDP-inositol was filtered as described above, lyophilized and quantified by ^{31}P-NMR, using diglycerol phosphate as a concentration standard.

Production of IPCT/DIPPS from *Archaeoglobus fulgidus*

The *A. fulgidus* AF0263 gene encoding IPCT/DIPPS was amplified by PCR (forward primer: 5'-TCCCCCCGGGATGATAAATGTTGACGGAGAATAC-3', *Xma*I site underlined; reverse primer: 5'-GCGGAGCTCTTTAGAAACCAAAACAGCAAGTAAAG-3', *Sac*I site underlined). The PCR product was cloned into the expression vector pET52b(+) (Novagen) using *Sac*I and *Xma*I restriction sites, with N-terminal Strep tag II and C-terminal His_{10} tag. The correct sequence of the resulting plasmid was verified by DNA sequencing. The expression vector was transformed into *E. coli* C43 DE3. Cells were cultivated at 37ºC

in terrific broth (TB) medium supplied with 100 μg/ml of ampicillin. When an OD of 1.0 at 600 nm was reached, cells were treated with 0.5 mM of IPTG and allowed to grow further for 6 hours at 30ºC before harvesting by centrifugation (1,000×*g*, 10 min).

Solubilization of the IPCT/DIPPS from *E. coli* membrane fractions

The detergents, n-Dodecyl β-D-Maltopyranoside (DDM), triton X-100, cymal-5, n-octyl β-D-thioglucopyranoside (OTG), and lauryldimethyl amine oxide (LDAO) were examined to look for the most effective solubilizing agent. Cell pellets were re-suspended in 50 mM Bis/Tris/propane (pH 8.0) containing 5% glycerol and 100 mM NaCl, and disrupted in a French press. Cell debris was removed by centrifugation (10,000×*g*, 15 min, 4ºC); the supernatant was collected, divided into five equal aliquots, and ultracentrifuged (150,000×*g*, 1.5 h, 4ºC). The supernatant was discarded. To test the efficacy of each detergent, each pellet was solubilized for 2 h at 4ºC in 50 mM Bis/Tris/propane (pH 8.0) containing 5% glycerol, 300 mM NaCl, 1 mM $MgCl_2$, and 2% DDM, or 2% triton X-100, or 2% cymal-5 or 2% LDAO or 0.8% of OTG. Solubilized suspensions were ultracentrifuged (100,000×*g*, 1 h, 4ºC) and the supernatants were used for activity measurements to assess the level of IPCT/DIPPS solubilization. The best solubilization detergents was used for the purification of the bifunctional enzyme.

Purification of IPCT/DIPPS

The recombinant protein was produced as described above. Cell pellets were re-suspended in 50 mM Bis/Tris/propane (pH 8.0) containing 5% glycerol and 100 mM NaCl, and disrupted in a French press. Cell debris was removed by centrifugation (10,000×*g*, 15 min, 4ºC); the supernatant was collected and ultracentrifuged (150,000×*g*, 1.5 h, 4ºC). The supernatant was discarded and the sediment solubilized for 2 h at 4ºC in 50 mM Bis/Tris/propane (pH 8.0) containing 5% glycerol, 100 mM NaCl, 50 mM imidazole, 1 mM $MgCl_2$, and 2% of the best solubilization detergents. A second ultracentrifugation step was performed (100,000×*g*, 1 h, 4ºC) to remove insoluble materials. The supernatant was applied onto a HisTrap HP column (GE Healthcare). Non-specifically bound proteins were removed by washing the column with 50 mM Bis/Tris/propane (pH 8.0) containing 5% glycerol, 50 mM imidazole, 100 mM NaCl, and 0.05% of the best solubilization detergents (Buffer A). Elution was carried out with 50 mM Bis/Tris/propane (pH 8.0) containing 5% glycerol, 100 mM NaCl, 1 mM $MgCl_2$, 500 mM imidazole, and 0.05% of the best solubilization detergent (DDM, triton X-100 or cymal-5, see below). Fractions, containing the pure protein, were collected and dialyzed against buffer A. The purity of the protein preparations was judged by SDS-PAGE. The resulting IPCT/DIPPS preparations were used for activity assays to assessing possible disrupting effects of the purification procedure on enzyme functionality.

Reconstitution of the purified IPCT/DIPPS into lipid vesicles

The pure protein solubilized in triton X-100 was obtained as described above. A fraction (160 µg of protein) was dialyzed to remove imidazole in 50 mM of Bis/Tris/propane (pH 8.0) containing 5% of glycerol, 100 mM of NaCl, and 0.1% of triton X-100. Lipid vesicles were prepared with either 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) or 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine (diC8PC) (Fig. 4.2) as follows: 2.0 mg of DMPC were added to 1 ml of 50 mM Bis/Tris/propane (pH 8.0) containing 5% of glycerol, 100 mM NaCl, and 1 mM $MgCl_2$. Vesicles of diC8PC were obtained by dissolving 4.6 mg of the lipid in 1 ml of the same buffer solution. In either case triton X-100 was added to the mixture to obtain a final lipid:detergent ratio of 1:1.5 (w/w). The samples were vortex-mixed vigorously for approximately 60 seconds to disperse the lipids and incubated at room temperature for 90 minutes. The resulting mixtures were incubated on ice for 90 minutes and then centrifuged (10,000×*g*, 2 min) to remove insoluble aggregates. Finally, the protein solution was added to the lipid vesicle preparation to a ratio of 2000:1 (lipid:protein), vortex-mixed for 30 seconds, and incubated at room temperature for 15 minutes, followed by further incubation at 4°C for 90 minutes. After this period, triton X-100 was removed by adsorption on Bio-beads SM (Bio-Rad) at 4°C for 16 hours (Ghosh et al. 2009). The reconstituted vesicles were separated from Bio-beads by low-speed centrifugation and used for IPCT and DIPPS activity assays.

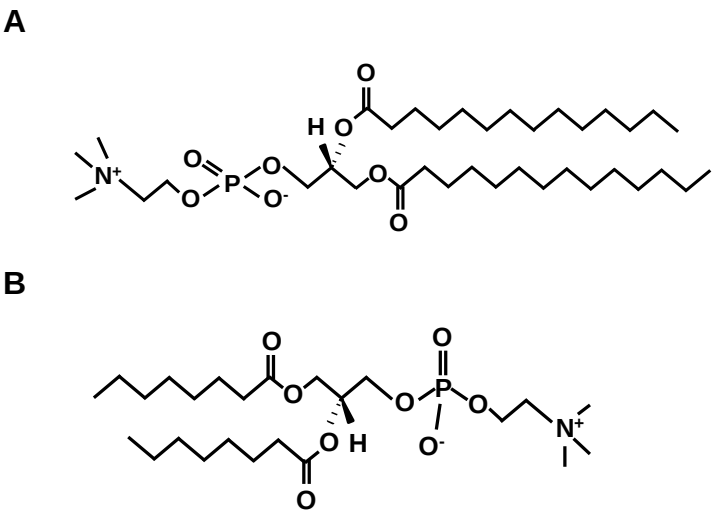


FIGURE 4.2 Structures of: **A)** 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC); and **B**) 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine (diC8PC).

Activity assays for the recombinant IPCT/DIPPS

To assess stability of the recombinant enzyme throughout the purification process, the activity of IPCT/DIPPS was measured in cell extracts of *E. coli*, and also after solubilization and purification. Reaction mixtures (0.3 ml) containing 25 mM Bis/Tris/Propane-HCl (pH 8.0), 10 mM $MgCl_2$, 2.7 mM CTP, 7 mM inositol-1-phosphate, and either 5-10 µg of purified IPCT/DIPPS or 250 µg of solubilized proteins, or 900 µg of total protein in cell extracts were incubated at 85ºC for different periods of time (between 30 seconds and 1 hour). The reactions were stopped on ice prior to the addition of deuterated water (200 µl) and 25 mM EDTA (pH 8.0). CDP-inositol and DIPP, the end products of IPCT and DIPPS, respectively, were quantified by ^{31}P-NMR using diglycerol phosphate as an internal

concentration standard. Protein content was estimated by the Bradford method (Bradford, 1976).

Purification of IPCT/DIPPS used for the crystallization trials

Cells were cultivated in TB medium at 37°C in a 30 liter fermenter with pH controlled at 7.4, harvested at 6 h after IPTG induction and re-suspended in lysis buffer (50 mM sodium citrate, 50 mM sodium phosphate, 200 mM NaCl, 5% glycerol, 100 mM sucrose, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.5 mM phenylmethanesulfonylfluoride (PMFS), 2 mM β-mercaptoethanol and protease inhibitors cocktail tablet SIGMAFAST; pH 7.5). Cells were disrupted using a constant flow cell disruptor and cell debris and inclusion bodies were removed by centrifugation at 10,000×*g* for 30 min (4°C). *E. coli* membranes were isolated by ultracentrifugation at 200,000×*g* for 2 h (4°C). The protein was extracted from the membranes by solubilization in 2% triton X-100, 20% glycerol, 1 M NaCl, 50 mM phosphate buffer, 50 mM imidazole, 100 mM sucrose, 2 mM β-mercaptoethanol and 1 mM $MgCl_2$ (final pH 7.4 at 4°C) for 2 h. Insoluble fraction was separated by 1 hour centrifugation at 100,000×*g* (4°C). The supernatant was loaded onto HisTrap HP column (GE Healthcare). The loaded sample was washed with buffer A (0.1% triton X-100, 50 mM phosphate buffer, 50 mM imidazole, 10% glycerol and 1 M NaCl; pH 7.4), followed by an additional wash with buffer A supplemented with 110 mM imidazole. The protein was eluted with buffer B containing 0.1% triton X-100, 50 mM phosphate buffer, 500 mM imidazole, 10% glycerol, 300 mM

NaCl and 1 mM $MgCl_2$; pH 7.4. Immediately after elution, 1 mM β-mercaptoethanol was added and the protein sample concentrated to ~20 mg/ml by using a centrifugal filter unit with a membrane of 100 kDa molecular weight cutoff. The major differences relative to the purification protocol described above are the following ones: (i) growth was carried out in 30 liter fermentation vessels with pH controlled at 7.4 instead of 2 liter Erlenmeyers without pH control; (ii) a continuous flow cell disrupter was used instead of a French press; (iii) the membrane fraction was isolated by centrifugation at 200,000×*g* for 2 hours, while in the protocol described above separation was carried out at 150,000×*g* for 1.5 hours; (iv) sodium citrate or phosphate replaced Bis/Tris/propane in the buffers; (v) the concentrations of NaCl and glycerol were changed in the solubilization buffer and in buffer A.

Crystallization and structure determination

The crystals of IPCT/DIPPS were grown using an *in meso* crystallization method (Landau and Rosenbusch 1996), similarly to Gordeliy et al. (2002). The protein-mesophase mixture contained concentrated enzyme solution and monoolein. High-throughput crystallization was performed with an NT8 robot (Formulatrix, USA). Plate-like crystals reaching the size of 5×20×200 µm were grown in 50 nl protein-containing mesophase drops overlaid by precipitant solution (4-12% PEG 550-6,000, 0.2-2 M sodium malonate, pH 5.8-8.2) at 295 K for 30-300 days. The crystals were harvested directly from the mesophase matrix and flashfrozen in liquid nitrogen. The best

X-ray diffraction data were collected from two crystals at 100 K, 0.8726 Å wavelength, on ID23-1 and ID23-2 beamlines of the European Synchrotron Radiation Facility (Grenoble, France). Data of these two crystals were merged and integrated to 2.65 Å (R_{pim} in the highest resolution shell is 46.9% with CC of 51.3% and I/σI of 1.7; see below Table 4.3), according to Karplus and Diederichs (2012), with MOSFLM (Leslie et al. 2002) and scaled using Scala (Evans 2006) from CCP4 software suite (Winn et al. 2011). Initial phases were obtained via Molecular Replacement with MOLREP (Vagin and Teplyakov 1997) using the crystallographic structure of the soluble domain (PDB ID 2XME) as a starting model. The electron density maps revealed that the soluble domains form layers with a separation of ~ 70 Å that is enough for the membrane-like layer containing TM domains. Additional search for α-helical fragments resulted in establishment of position of five TM α-helices. Refinement of the fragments' structure with Refmac5 (Murshudov et al. 1997) and PHENIX (Adams et al. 2010) resulted in improvement of the electron density maps, and following identification of the sixth TM helix and two α-helices parallel to the membrane plane. Eventually, identities of the side chains were determined. The final R-factors are 24.3 and 30.0% (R and R_{free}, respectively). Ramachandran plot shows that 96.3% of the amino-acid residues are in favoured regions, 3.2% are in allowed regions and 0.5% in not allowed regions. All structural figures were prepared with CCP4mg (McNicholas et al. 2011) and PyMOL (DeLano 2002).

Construction and characterization of IPCT/DIPPS mutants

Point mutations were introduced using QuickChange I kit (Agilent Technologies) and confirmed by DNA sequencing (STAB VIDA). Primers designed for the mutants are listed in Table 4.1. *E. coli* C43 harboring wild type or mutant variants of the *ipct/dipps* gene from *A. fulgidus* were cultivated at 37ºC in TB medium supplied with 100 μg/ml of ampicillin. When an OD of 1.0 at 600 nm was reached, cells were treated with 0.5 mM of IPTG and allowed to grow further for 6 hours at 30ºC before harvesting by centrifugation (1,000×*g*, 10 min). SDS-PAGE and His-tag western blotting were performed on the membrane preparations derived from cell extracts, confirming that all mutants were expressed in the *E. coli* membranes (data not shown). Protein content was estimated by the Bradford method (Bradford 1976). The IPCT/DIPPS activities were assessed in cell extracts of *E. coli* C43. The cell pellets were suspended in 50 mM Bis/Tris/propane-HCl (pH 8.0) containing 5% (v/v) glycerol and 100 mM NaCl; cells were disrupted in a French press and debris were removed by centrifugation at 10,000×*g* for 15 min (4°C). Supernatants were examined for IPCT/DIPPS activity. Reaction mixtures containing 0.9 mg of total protein were prepared as described above and incubated at 85°C for 1 and 30 min. The production of CDP-inositol and DIPP was assessed by ^{31}P-NMR using diglycerol phosphate as an internal concentration standard. All the assays were performed in duplicate with cells derived from independent growths.

TABLE 4.1 Forward and reverse primers used for site-directed mutagenesis. Mutated nucleotides are underlined.

Substitution	Oligonucleotide primer sequences (5'-3')
S308A	CAAGGAAAATCAACAGGAAAATCGCGACGAGAATTTCGGC GCCGAAATTCTCGTCGCGATTTTCCTGTTGATTTTCCTTG
S308M	GGAAAATCAACAGGAAAATCATGACGAGAATTTCGGCAGCCA TGGCTGCCGAAATTCTCGTCATGATTTTCCTGTTGATTTTCC
D357A	CTCCTCAATTCTCGCCGGGTGTGATGGGG CCCCATCACACCCGGCGAGAATTGAGGAG
D357N	GTTCTCCTCAATTCTCAACGGGTGTGATGGGG CCCCATCACACCCGTTGAGAATTGAGGAGAAC
D360A	TTCTCGACGGGTGTGCTGGGGAGATAGCAAG CTTGCTATCTCCCCAGCACACCCGTCGAGAA
D360N	CAATTCTCGACGGGTGTAATGGGGAGATAGCAAGA TCTTGCTATCTCCCCATTACACCCGTCGAGAATTG
A364M	TCTCGACGGGTGTGATGGGGAGATAATGAGAGCCTCACTGA TCAGTGAGGCTCTCATTATCTCCCCATCACACCCGTCGAGA
R365M	CGGGTGTGATGGGGAGATAGCAATGGCCTCACTGAAA TTTCAGTGAGGCCATTGCTATCTCCCCATCACACCCG
S371P	GAGCCTCACTGAAAATGCCCAAAAAAGGCGGGTA TACCCGCCTTTTTTGGGCATTTTCAGTGAGGCTC
D378A	AAAAAAGGCGGGTATGTTGCTTCAATTCTGGACAGGTTC GAACCTGTCCAGAATTGAAGCAACATACCCGCCTTTTTT
D378N	CCAAAAAAGGCGGGTATGTTAAATTCAATTCTGGACAGGTTC GAACCTGTCCAGAATTGAATTAACATACCCGCCTTTTTTGG
D382A	GTATGTTGATTCAATTCTGGCCAGGTTCGTCGATTTCCTTT AAAGGAAATCGACGAACCTGGCCAGAATTGAATCAACATAC
D382N	GGTATGTTGATTCAATTCTGAACAGGTTCGTCGATTTCCTT AAGGAAATCGACGAACCTGTTCAGAATTGAATCAACATACC
R443A	GCTGAACTACATTCCCGGAAAAGCGGATGAGAGGATTTTTCTG CAGAAAAATCCTCTCATCCGCTTTTCCGGGAATGTAGTTCAGC

Co-crystallization of IPCT/DIPPS with substrates

For crystallization of IPCT/DIPPS with inositol-1-phosphate and CDP-inositol (separately and in combination), the protein was incubated with 8.6 mM inositol-1-phosphate or 8 mM of CDP-inositol for 1 hour (while on ice). In case of co-crystallization trials with both substrates, the protein solution was first incubated with 8 mM of CDP-inositol for 1 hour on ice. After this period, 8.6 mM of inositol-1-phosphate was added and the solution was kept on ice for a further 30 minutes. Crystallization was carried out as described above.

Results

Assessment of IPCT/DIPPS functionality throughout the purification procedure

Purification of integral membrane proteins requires the screening of appropriate solubilizing agents, such as detergents and liposomes, which facilitate the extraction of proteins from the cellular membranes, ideally without affecting its native structure and biological function (Ghosh et al. 2009). Naturally, protein extraction from membranes depends on the ability of detergents to penetrate and cross membranes (le Maire et al. 2000); moreover, solubilization of the membrane protein depends on the micellar size and other physicochemical properties of the detergent. Indeed, several studies describe the inactivating properties of zwitterionic

compounds in opposition to non-ionic detergents, which are generally considered as non-denaturating agents (le Maire et al. 2000, Alexandrov et al. 2008). Therefore, screening for a suitable detergent is mandatory for a successful outcome in the characterization of membrane proteins. In this study, five detergents (DDM, triton X-100, Cymal-5, OTG, and LDAO) were tested for solubilization of IPCT/DIPPS from *E. coli* membranes. Relevant chemical and physical properties of these detergents are shown in Table 4.2.

TABLE 4.2 Physicochemical properties of the detergents examined in this work. Data from Sigma-Aldrich Corporation.

	Triton X-100	DDM	Cymal-5	OTG	LDAO
Nature	Non-ionic	Non-ionic	Non-ionic	Non-ionic	zwitterionic
CMC*	0.2-0.9 (0.0125-0.058)	0.15 (0.0095)	2.4-5.0 (0.12-0.23)	23-25 (0.67-0.73)	n.a.
Aggregation number	625	511	66	n.a.	n.a.
AMW	80	50	32.6	n.a.	n.a.

n.a. data not available.
***** The critical micelle concentration (mM), i.e., the concentration at which micelles are spontaneously formed in water at 20-25ºC. The values between parentheses correspond to % w/v.
Aggregation number is the average number of monomers in a micelle.
AMW, average micellar weight in kDa.

We compared the solubilizing ability of DDM, triton X-100, Cymal-5, OTG, and LDAO in terms of IPCT/DIPPS activity and protein recovery from

membrane fractions. These two properties were evaluated by quantification of the end products of IPCT and DIPPS after protein solubilization. For solubilization, *E. coli* membranes were incubated with buffer containing 2% of detergent except for OTG which was used at a final concentration of 0.8% (w/v). For all detergents examined, the final concentration was above the critical micelle concentration (Table 4.2). Non-solubilized proteins were removed by ultracentrifugation and product formation (CDP-inositol and DIPP) was determined by ^{31}P-NMR. The total amount of CDP-inositol plus DIPP reflects IPCT activity, while the level of DIPP correlates with DIPPS activity.

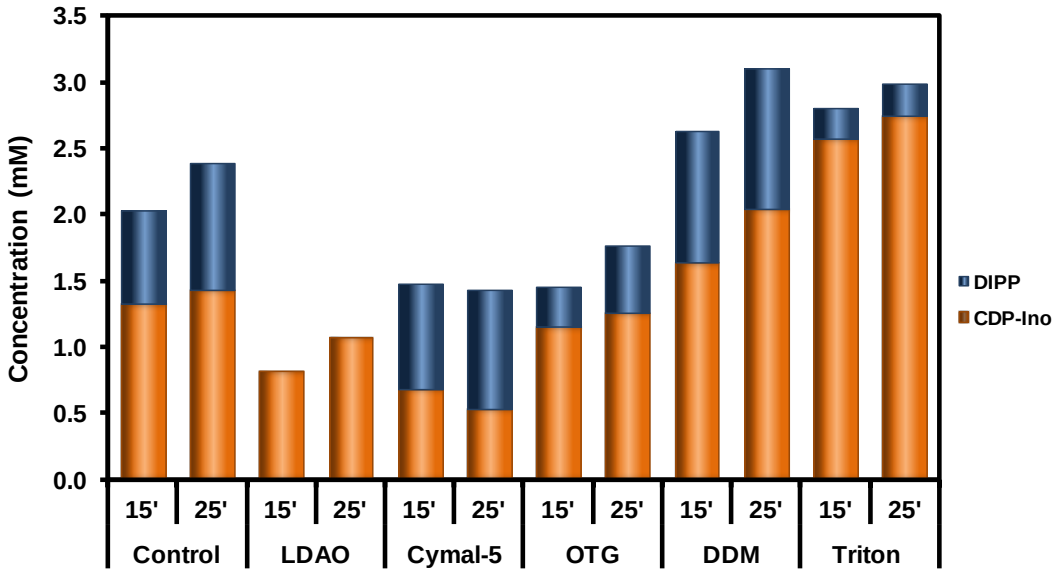


FIGURE 4.3 Production of CDP-inositol (orange) and DIPP (blue) by detergent-solubilized membrane proteins of *E. coli* expressing *ipct/dipps* after 15 and 25 minutes of incubation at 85ºC. A control assay was prepared without addition of

detergent during the solubilization step. Tested detergents: LDAO (lauryldimethyl amine oxide); OTG (n-octyl beta-D-thioglucopyranoside); Cymal-5; Triton X-100 (triton X-100); and DDM (n-dodecyl-ß-maltoside). Products were quantified by ^{31}P-NMR.

The non-ionic detergents DDM and Cymal-5 resulted in the two highest yields of DIPP, while the use of the zwitterionic detergent LDAO led to full inactivation of DIPPS (Fig. 4.3). This result is not surprising since the inactivating properties of zwitterionic detergents are well known (le Maire et al. 2000, Alexandrov et al. 2008). Triton X-100 and DDM resulted in high and comparable activity of the cytoplasmic domain (IPCT), but the functionality of DIPPS was significantly lower when triton X-100 was used. In spite of this drawback, this latter detergent was selected for large-scale protein production due to its lower cost. Surprisingly, substantial amounts of DIPP were detected in the control assay even though solubilizing agents were absent. Probably, membrane proteins were solubilized to some extent with the *E. coli* lipids. Since the same amount of total protein was used in all the assays, these results show the superior ability of DDM and triton X-100 to solubilize IPCT/DIPPS. Cymal-5 seemed less efficient for protein extraction but was especially advantageous to preserve the activity of the membrane domain. In view of these findings, batches of pure IPCT/DIPPS solubilized in Cymal-5, DDM or triton X-100 were obtained and their IPCT and DIPPS activities evaluated. Formation of CDP-inositol was observed in all cases, but the final product, DIPP, was consistently absent. Thus, we conclude that the second activity was lost during the

purification process. In a further attempt to recover DIPPS activity, the purified IPCT/DIPPS was reconstituted into lipid vesicles (DMPC and diC8PC). Supposedly, these vesicles mimic a more natural membrane environment, and consequently they may preserve better the native protein structure (Wang et al. 2013). Triton X-100, which was used throughout the purification process, was removed by adsorption on neutral polymeric beads. Unfortunately, DIPP production was not recovered.

Structure determination

The bifunctional enzyme IPCT/DIPPS from *A. fulgidus* was produced in *E. coli* and purified by immobilized metal affinity chromatography. The recombinant protein was crystallized from sample preparations using triton X-100 for membrane extraction and protein solubilization. Crystals were grown using the *in meso* crystallization method (Landau and Rosenbusch 1996). They belong to the orthorhombic space group $P2_12_12$ with unit cell parameters of $a = 41.37$, $b = 107.58$ and $c = 23.95$ Å (Table 4.3). One IPCT/DIPPS molecule is present in the crystal asymmetric unit corresponding to a Matthews coefficient (Matthews 1968) of 2.87 $Å^2\ Da^{-1}$ and a solvent content of ~ 57%. This molecule interacts extensively with another molecule related by crystallographic symmetry. The protein molecules are organized in the crystal as a stack of membrane layers, typical for *in meso* crystallization (type I membrane protein crystals (Ostermeier and Michel 1997), with interlayer contacts relying entirely on the IPCT cytoplasmic domain (data not shown). The structure was solved

by molecular replacement using the coordinates of the IPCT domain that we reported earlier (Brito et al. 2011) and refined to a resolution of 2.65 Å with R_{work}/R_{free} values of 24.5% and 30%, respectively (Table 4.3). The final model comprises 408 amino acid residues, 1 magnesium ion, 9 water molecules and 4 lipid fragments. The electron density maps are generally of good quality (data not shown). The soluble IPCT domain is somewhat less ordered than the membrane DIPPS one (average B-factors of IPCT and DIPPS backbone atoms are 70.9 and 50.5 $Å^2$, respectively).

TABLE 4.3 Data collection and refinement statistics.

Data collection	
Space group	$P2_12_12$
Cell dimensions	
a, *b*, *c* (Å)	41.37, 107.58, 123.95
α, β, γ (°)	90, 90, 90
Resolution (Å)	61.82–2.65 (2.79–2.65)*
No. unique reflections	15,952
R_{merge} [⊥]	0.203 (1.063)
R_{pim} [∴]	0.081 (0.469)
CC^a	0.991 (0.513)
$I/\sigma I$	7.7 (1.7)
Completeness (%)	99.9 (99.9)
Redundancy	7.2 (5.8)
Refinement	
Resolution (Å)	53.85–2.65
Rwork/Rfree (%)	24.3/30.0
Number of atoms	
Protein	3,112
Mg^{2+}	1
Lipid fragments	4
Water	9
Average B-factor (Å²)	
Protein	61.4
Mg^{2+}	58.7
Lipid fragments	58.5
Water	48.3
Ramachandran	
Residues in favored regions (%)	96.3
Residues in disallowed regions (%)	0.5
R.m.s. deviations	
Bond lengths (Å)	0.014
Bond angles (°)	0.791

*Highest resolution shell is shown in parenthesis.

[⊥] $R_{merge} = \sum_{hkl} \sum_i \left| li(hkl) - \overline{l(hkl)} \right| / \sum_{hkl} \sum_i li(hkl)$

[∴] $R_{pim} = \sum_{hkl} \left[1/(N-1)\right]^{1/2} \sum_i \left| li(hkl) - \overline{l(hkl)} \right| / \sum_{hkl} \sum_i li(hkl)$

Calculated with the program SCALA, R_{merge} and R_{pim} are indicators of the precision of the final merged and average data set, where $li(hkl)$ is the observed intensity of the i^{th} measurement, $\overline{l(hkl)}$ is the average intensity of multiple observations of symmetry-related reflections and *N* is redundancy.
[a] *CC* Coefficient as described in Karplus and Diederichs (2012).

Overall fold

The IPCT/DIPPS bifunctional protein forms a dimer via its transmembrane (TM) domains (Fig. 4.4A,B). The dimeric interface is formed by helix TM4 of one protomer protruding between helices TM3' and TM4' of the other protomer and flanked by TM6. The interface is stabilized mostly by hydrophobic contacts within the lipidic bilayer and several H-bonds between polar/charged residues located at the cytoplasmic extramembrane side with a total contact area of 2,050 Å^2. The DIPPS domain contains six TM α-helices TM1-TM6 and one shorter bent juxtamembrane helix (connector helix 2; Fig. 4.4A). TM2 shows a kink around Asp357, nearby the membrane interface. The intracellular side of the TM domain is large and contains many polar and charged residues, whereas the opposite side is mostly embedded in the membrane. All TMs cross the membrane, except TM5, which is shorter. Connector helices 1 and 2 lie nearly parallel to the membrane plane, and link the soluble and membrane domains. We assigned connector helix 1 to IPCT and connector helix 2 to DIPPS, based on their distribution in the organisms where IPCT and DIPPS do not belong to the same polypeptide chain (data not shown).

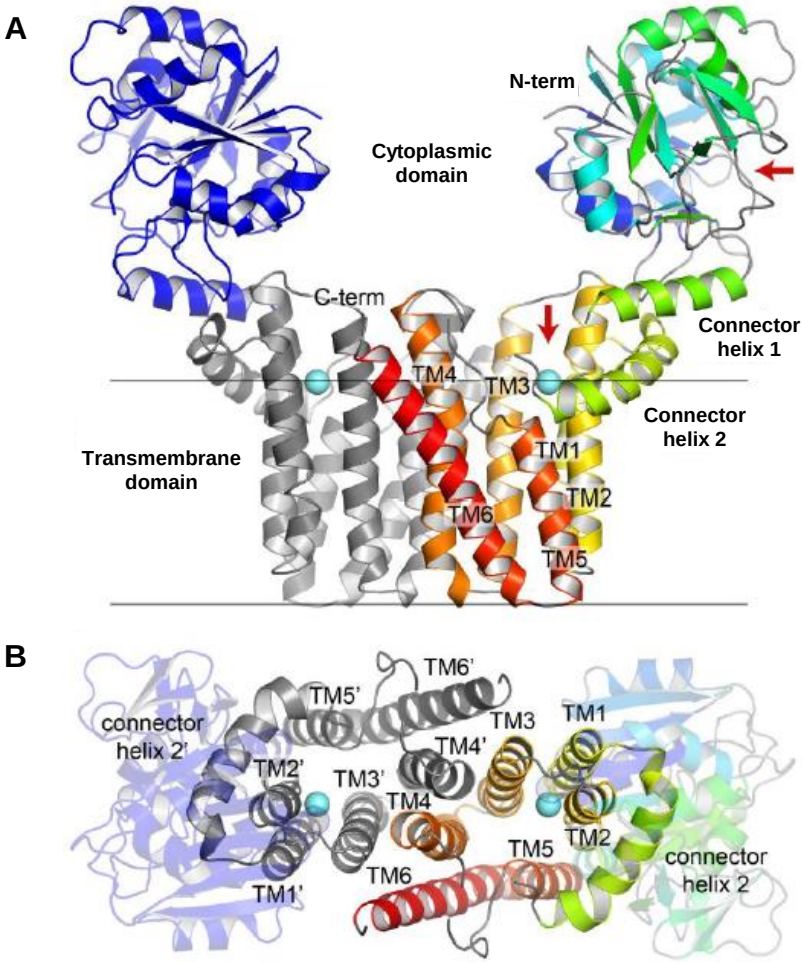


FIGURE 4.4 X-ray structure of IPCT/DIPPS. The protein is dimerized via the TM domain. **A**, One monomer is rainbow colored and the other one (related by crystallographic symmetry) is in blue for the cytoplasmic IPCT domain and grey for the membrane DIPPS domain. Red arrows indicate the access to their active sites with a putative Mg^{2+} drawn as a cyan sphere. Membrane surface position

(grey) was calculated with PPM server (Lomize et al. 2012). **B**, Top view of the DIPPS domain (perpendicular to the membrane plane).

Although connector helix 2 is conserved among the DIPP synthases, it is absent in many other CDP-alcohol phosphatidyltransferases. It wraps around TM2 and the N-terminus of TM5. Both connector helices are amphipathic, the hydrophobic side of connector helix 2 faces the membrane and is proximal to two lipid fragments, whereas its hydrophilic side points towards the cytoplasm and connector helix 1, whose apolar residues face the proximal β-sheet of IPCT.

The cytoplasmic IPCT domain binds at the edge of the DIPPS domain (Fig. 4.4A). It comprises a core region formed by a central seven-stranded mixed β-sheet flanked by six α-helices. This arrangement is reminiscent of a dinucleotide-binding Rossmann fold, typical of nucleotidyltransferases, with an additional two-stranded β-sheet and a 30-residue long stretch. The sugar-binding region comprises a short antiparallel three-stranded β-sheet lying against the face of the central β-sheet where the nucleotide binds. The overall fold of the IPCT domain is very similar to the structure of that domain alone, which we previously reported (Brito et al. 2011) (root mean square deviation, r.m.s.d. of 0.6 Å for Cα atoms). Interestingly, the IPCT domains dimerized in the crystals of the truncated one-domain construct (Brito et al. 2011) via the same interface that connects IPCT and DIPPS domains in the full-length protein.

Active site architecture and consensus sequence motif

The six TM helices of DIPPS form a central hydrophilic cavity ~ 20 Å deep. This cavity is wide open (~10-20 Å) to the cytoplasm (Fig. 4.4A). Its entrance is flanked by three loops between TM2-TM3 (residues 369-373), connectors helix 1-helix 2 (residues 291-298), and TM4-TM5 (residues 430-440). The last two loops are less ordered suggesting flexibility (data not shown), and residues 292-294 and 430-433 were not included in the model as no electron density was visible in the maps.

Remarkably, at one side of the cavity there are four aspartate residues (Asp357, Asp360, Asp378 and Asp382) pointing to the same patch of the electron density (Fig. 4.5A). The last three aspartate residues are part of the conserved sequence pattern of the CDP-alcohol phosphatidyltransferase family, $DG(x)_2AR(x)_{7,8}G(x)_3D(x)_3D$, as defined in Prosite (Sigrist et al. 2013), but all four aspartate residues are highlighted in the HMM logo of this family (Fig. 4.5B). Since Mg^{2+} had been used throughout the purification of IPCT/DIPPS, we assigned a magnesium ion to this blob of density lying at the membrane interface. Mg^{2+} is coordinated to the carbonyl of Asp357 and carboxylate groups of Asp357, Asp360 and Asp378 (Fig. 4.5A), where it is expected to decrease the pKa of the surrounding aspartate residues. Asp382 is not directly coordinated to the metal ion, but is within H-bonding distances to Asp357 and Asp378, and is ~ 4.5 Å away from Arg443. This cluster of negatively charged residues will tend to increase the pKa of neighboring amino acid residues. The

coordination sphere of the magnesium ion includes the pyrophosphate moiety of CDP-inositol (see below Fig. 4.6), as observed in some glycosyltransferases (Lairson et al. 2008) and kinases (Auerbach et al. 1997).

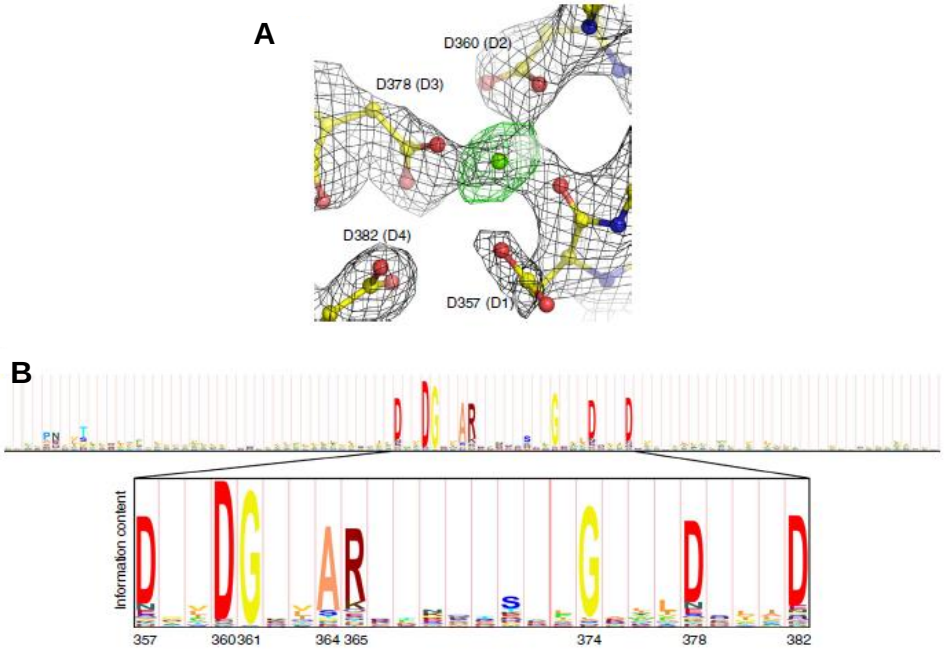


FIGURE 4.5 DIPPS active site along with HMM logo. **A**, Zoomed view of the electron density map around the conserved aspartate residues. The maps were calculated before insertion of the magnesium ion in the model. The $2F_o–F_c$ density is shown in black and contoured at 1.5σ, the $F_o–F_c$ density is green and contoured at 3σ; electron density around Asp357 is not very well defined. **B**, HMM logo50 around the consensus sequence of CDP-alcohol phosphatidyltransferase family based on the alignment of 124 representative Pfam protein families as seeds (Punta et al. 2012). The most conserved residues are located at the intracellular ends of TM2 and TM3.

Interestingly, the entries to the active sites of both enzymes that catalyze consecutive reactions in DIP synthesis are oriented in different directions (Fig. 4.4A, red arrows). It is currently not known whether cross-talk exists between the two domains or if *ipct/dipps* gene fusion/separation produces any differences in the overall catalytic efficiency. Ligand binding could trigger rearrangements in the relative orientation of the two domains, as the interactions between IPCT and DIPPS domains are relatively weak (interface area of 479 $Å^2$, solvation free-energy gain is -7.3 kcal mol^{-1} calculated with PISA (Krissinel and Henrick 2007). However, the structural information available for the protein bound to CDP-inositol does not support this speculation. Final conclusions can be drawn only when co-crystallization with the two substrates is accomplished.

Substrate binding site

While searching for potential substrate binding sites, we identified three pockets on the surface of DIPPS using HOLLOW (Ho and Gruswitz 2008) (Fig. 4.6A). The co-crystallization of IPCT/DIPPS with the substrate CDP-inositol showed that pocket 1 is occupied by the cytidine moiety (Fig. 4.6B,C,D).This pocket is flanked by the loop between TM2 and TM3, and is lined by the family consensus residues Gly361, Ala364, and Gly374. The residue Gly374 interacts with one face of the planar cytosine ring while Ala364 and Gly361 are located in the opposite side (Figs. 4.6C,D and 4.7).

Pocket 2 is at the bottom of the hydrophilic cavity formed by strictly conserved residues among DIPPS proteins, namely Asn304, Arg305, Ser308, Ser354, Asp357 and Arg443 (Figs. 4.6 and 4.7), sited on connector helix 2, TM2 and TM5. This pocket is occupied by the inositol moiety of CDP-inositol (Fig. 4.6C,D). The position of the CDP moiety is in agreement with results reported by Sciara and co-workers (2014) for a different member of the CDP-alcohol phosphatidyltransferase family. Pocket 3 is surrounded by helices TM4, TM5 and TM6, and loop TM5–TM6. Amino acid residues from this loop are not conserved, and the pocket itself is far from the catalytic magnesium ion. Several attempts to co-crystallize DIPPS with inositol-1-phosphate were unsuccessful. Further efforts are underway to identify the binding site of this substrate.

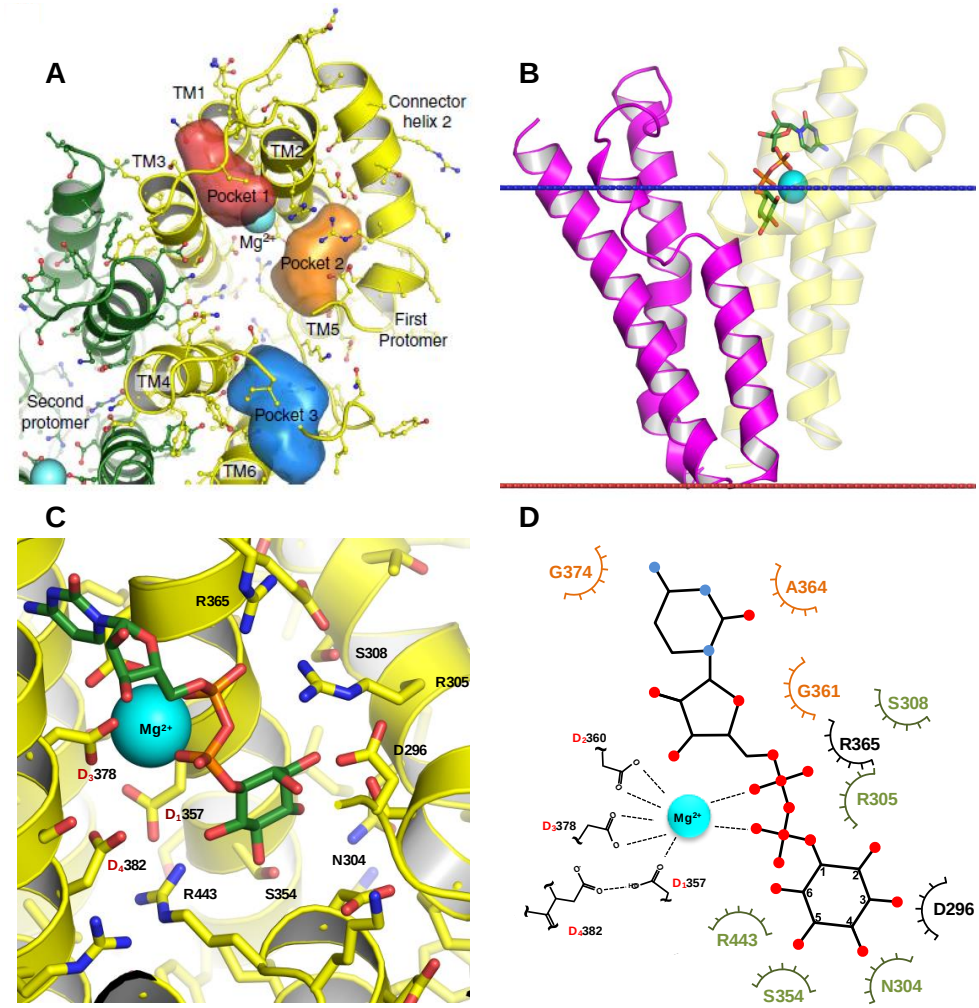


FIGURE 4.6 Pockets visualization and binding of CDP-inositol in the active site. **A**, Putative ligand-binding pockets of DIPPS assigned with HOLLOW (Ho and Gruswitz 2008) (top view from the cytoplasm-facing side). The pockets are shown in pink (1), orange (2) and blue (3). **B**, Position of CDP-inositol relative to the bilayer surface. The CDP moiety is above the bilayer surface, whereas the inositol group of CDP-inositol is slightly underneath it. The membrane position was estimated using the PPM server (Lomize et at. 2012). **C**, Detailed view of the binding of CDP-inositol into pockets 1 and 2. The strictly conserved residues

Gly361, Ala364 and Gly374, belonging to the family consensus sequence, form the surface of the pocket 1, where cytidine moiety binds. The inositol moiety occupies pocket 2. **D**, Schematic representation of CDP-inositol location in pocket 1 (orange) and 2 (green).

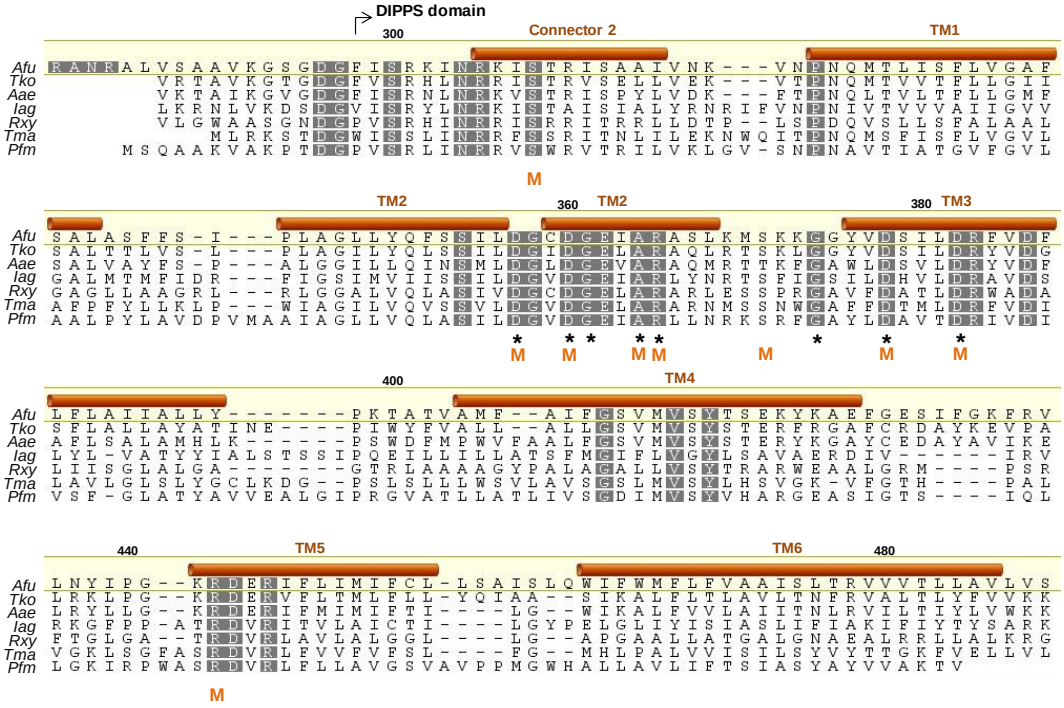


FIGURE 4.7 Alignment was done with Geneious 5.6 (Kearse et al. 2012) based on DIPPS sequences from 6 organisms with confirmed or putative DIPP activity (Gonçalves et al. 2012). The selected sequences represent the major branches of phylogenetic tree (Gonçalves et al. 2012). *Afu, Archaeoglobus fulgidus*; *Tko, Thermococcus kodakarensis*; *Aae., Aquifex aeolicus*; *Iag, Ignisphaera aggregans*; *Rxy, Rubrobacter xylanophilus*; *Tma, Thermotoga maritima*; and *Pfm, Pyrolobus fumarii*. Strictly conserved residues among all IPCT/DIPPS sequences are shown with a grey background, and those from the consensus pattern of the CDP-

alcohol phosphatidyltransferase family are marked with an asterisk while mutated residues are marked with an orange M. Secondary structure elements represented with red tubes (α-helices) were assigned with DSSP (Kabsch and Sander 1983).

Observation of IPCT/DIPPS activity in cell extracts of *E. coli*

As mentioned above, the activity of the membrane domain of recombinant IPCT/DIPPS was lost during purification. Fortunately, when a crude cell-free extract was incubated with CTP and inositol-1-phosphate, both IPCT and DIPPS activities were detected. Therefore, we decided to proceed to monitoring the time course of the enzymatic reactions directly in cell extracts of the heterologous host. Reaction mixtures containing cell extract of *E. coli* (900 µg of total protein), 2.7 mM CTP and 7 mM inositol-1-phosphate were incubated at 85°C for time periods in the range 30 seconds to 30 minutes and product formation was followed by ^{31}P-NMR (Fig 4.8).

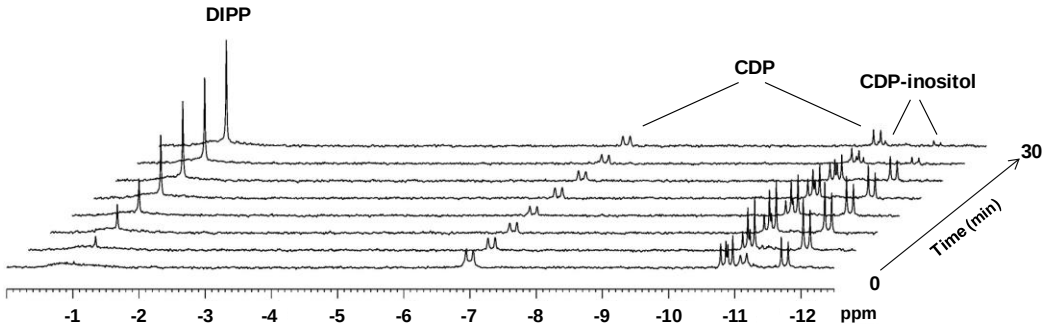


FIGURE 4.8 Time course of product formation from 2.7 mM CTP and 7 mM inositol-1-phosphate by an extract of *E. coli* cells producing the *A. fulgidus* IPCT/DIPPS, as determined by ^{31}P-NMR spectroscopy. The reaction mixture was incubated at 85°C for different time periods.

The concentrations of the final product (DIPP) and the intermediate product (CDP-inositol) of IPCT/DIPPS are plotted as a function of the reaction time (Fig. 4.9). The level of CDP-inositol was maximal at about 1.5 min of incubation; within 30 minutes this metabolite was virtually converted into the phosphorylated form of DIP. The detection of considerable amounts of CDP already in the initial spectra is due to the presence of this contaminant in the commercial CTP preparation.

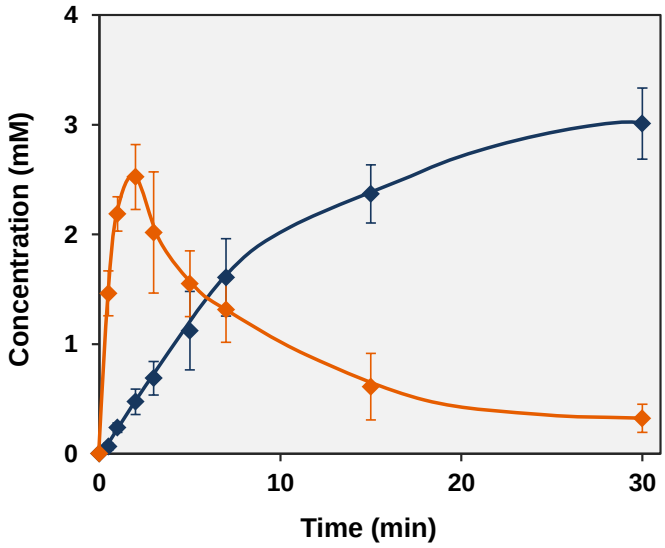


FIGURE 4.9 Time courses of CDP-inositol (orange) and the phosphorylated form of DIP (DIPP, blue) formed at 85°C by a cell extract of *E. coli* harboring *ipct/dipps* from *A. fulgidus*. The rate of CDP-inositol production is considerably higher than that of DIPP production for short incubation times, leading to a transient peak of CDP-inositol.

Site-directed mutagenesis and functional assays

Site-directed mutagenesis on conserved amino acid residues of the family sequence motif or considered functionally relevant to DIPPS (for example, for substrate binding) was carried out. The expression of IPCT/DIPPS mutants was validated by western blot using antiboby against His-tag (data not shown). The production of CDP-inositol is assumed to be a reliable indicator of the IPCT/DIPPS yield for the different mutants since none of the point mutations targeted the IPCT domain. We found that the formation

of CDP-inositol, the product of the reaction catalyzed by IPCT, was fairly constant for the 14 mutants examined (Fig. 4.10).

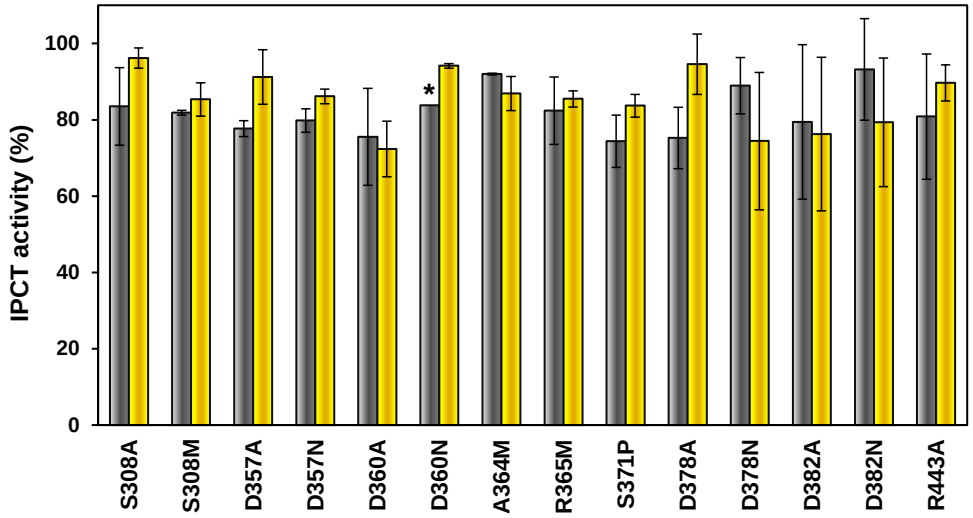


FIGURE 4.10 Production of CDP-inositol by mutants of IPCT/DIPPS after 1 (grey) and 30 (yellow) minutes of incubation at 85°C. Values are relative to the wild type protein, which was considered 100%. Values are averages of two independent replicates. The amount of total protein used (0.9 mg) was the same for all the assays. The bar labeled with an asterisk (*) corresponds to a single experiment.

HMM logo defines the conserved sequence of CDP-alcohol phosphatidyltransferases as $D1^{357}(x)_2D2^{360}G^{361}(x)_2A^{364}R^{365}(x)_8G^{374}(x)_3D3^{378}(x)_3D4^{382}$ (superscript numbering corresponding to IPCT/DIPPS sequence; Fig. 4.5B). Mutations of these aforementioned four aspartate residues to alanine or asparagine inactivated the enzyme, except for D357N, which resulted in a very low

activity (Fig. 4.11). These aspartate residues are part of the hydrophilic active site cavity and are close to the magnesium ion. Moreover, mutations of consensus residues Ala364 and Arg365 to methionine caused a considerable decrease of DIPPS activity (Fig. 4.11). These residues are located at TM2 and lining pocket 1 (Fig. 4.6C,D). We speculate that A364M may obstruct the accessibility of the CDP moiety from CDP-inositol to the active site, whereas R365M may affect the coordination of the negatively charged diphosphate moiety. Mutation of equivalent residues (Ala to Gly and Arg to Ala) in *Saccharomyces cerevisiae* choline phosphatidyltransferase (CPT) also led to reduced catalytic activity, suggesting that both residues have a role in substrate binding or positioning (Williams and McMaster 1998). Mutation of S371P within TM2-TM3 loop resulted in a significant reduction of the enzyme activity (Fig. 4.11). This is probably owing to restraints on the loop flexibility that could affect the substrates entry. A similar result was also observed for the corresponding allele mutation in *E. coli* phosphatidylglycerophosphate synthase (Usui et al. 1994).

Related to pocket 2, we mutated Ser308, a strictly conserved amino acid residue among enzymes with putative DIPPS activity, located on connector helix 2 that is adjacent to binding site of the diphosphate moiety (Figs 4.6C and 4.7). Decrease of DIPPS activity in S308M could result from introducing a bulkier side chain in the binding site, thus preventing substrate binding (Fig. 4.11). Moreover, Arg443 sits at the bottom of the hydrophilic cavity and is within hydrogen bonding distances to the

carboxylate groups of Asp357 and Asp386 and to one hydroxyl of the inositol moiety in CDP-inositol. Importantly, the substitution of Arg443 by alanine (R443A) abolishes DIPPS activity.

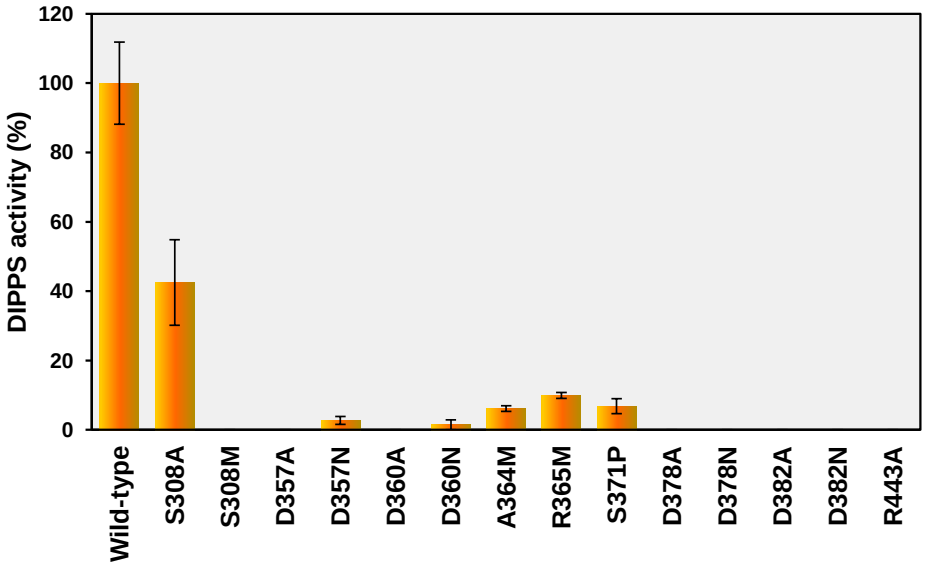


FIGURE 4.11 Diagram of DIPPS activity as assessed in 30 min reactions to study the effect of point mutations; values are expressed as % relative to the wild type enzyme in the presence of 10 mM Mg^{2+}. Values are averages of experiments derived from two independent growths.

Discussion

In this work, we report the 3-dimensional structure of the bifunctional enzyme IPCT/DIPPS, which comprises a membrane domain in addition to a cytoplasmic domain. This is an important achievement, especially if the very low success rate in the crystallization of membrane proteins is considered, i.e., 3 ‰.

Despite our efforts, the membrane domain was inactive in the purified enzyme (Table 4.4). This inactivation can be due to structural disruption of DIPPS because of the deficient ability of detergent micelles to reproduce the cell membrane environment. On the other hand, it is highly unlikely that detergent micelles are stable at the high temperature required for activity of hyperthermophilic enzymes. For example, DM micelles have a melting temperature of 50ºC. (F. Mancia, personal communication). Thus, the functionality of this type of proteins can only be tested efficiently when highly stable artificial micelles are available. Limited effort was directed to reconstitute the pure IPCT/DIPPS into lipids extracted from *A. fulgidus*, but the DIPPS activity was not recovered (data not shown).

TABLE 4.4 Qualitative assessment of **IPCT/DIPPS** functionality throughout the process of protein production for crystallization trials.

	***E. coli* cell extracts**	**Solubilized in detergents**	**Purified**	**Reconstituted in lipid vesicles**
IPCT	+	+	+	+
DIPPS	+	+	-	-

Concurrently but totally independently of our team, the group of F. Mancia at the Columbia University elucidated the structure of AF2299 from *A. fulgidus*, another representative of the CDP-alcohol phosphatidyltransferase family. The two structures were published in the same issue of Nature Communications (Nogly et al. 2014, Sciara et al. 2014). Like IPCT/DIPPS, the AF2299 protein comprises a cytoplasmic and a membrane domain, but the soluble domain is small and probably has no independent physiological role. The physiological function of AF2299 remains obscure and therefore the functionality of the pure protein was not verified. The AF2299 shares 36% of amino acid identity with IPCT/DIPPS and, the two proteins show super imposable structural architectures. This striking coincidence reinforces our view that the structure determined for IPCT/DIPPS is physiologically meaningful.

The membrane character of DIPPS is puzzling, since all its substrates and products are water soluble. A recent evolution analysis for DIPPS domain has shown that these enzymes evolved from phosphatidylinositol-phosphate synthases that are membrane proteins, so DIPPS kept the membrane-spanning character of its ancestor (Gonçalves et al. 2012). In contrast, CDP-OH_P_trans enzymes involved in phospholipid biosynthesis usually have a lipophilic substrate that can either be a donor substituent or alcohol acceptor (for example, diacylglycerol (DAG)) that reacts with a water-soluble one (typically a phospholipid head group, for example, glycerol or serine). However, some enzymes, such as cardiolipin synthases, have two lipidic substrates, CDP-DAG and

phosphatidylglycerol. The X-ray structure of IPCT/DIPPS revealed a large polar active site cavity that lies near the membrane interface and should be a common feature to other CDP-alcohol phosphatidyltransferases. This positioning is strategic to allow accommodation and binding of both hydrophilic and lipophilic substrates. The catalytic reaction occurs at the hydrophilic-hydrophobic interface of the membrane with Mg^{2+} lying on the border (Fig. 4.6B), so the lipophilic chains of diacylglycerol or other phospholipids could easily be embedded in the lipid bilayer.

The HMM logo sequence of CDP-alcohol phosphatidyltransferases (Fig. 4.5B) compared with Prosite (Sigrist et al. 2013) reveals the conservation of an extra aspartate residue (D1, Asp357), which is not yet included in the consensus motif of this family. This Asp was shown to be essential for the catalytic activity of other phosphatidyltransferases (Solis-Oviedo et al. 2012), likely having a role in substrate binding and positioning. We therefore propose to add this first aspartate residue in the consensus sequence pattern of CDP-alcohol phosphatidyltransferases, as previously suggested (Solis-Oviedo et al. 2012). Furthermore, some phosphatidylcholine synthases were shown to have 12 amino acid residues between Arg and Gly (Solis-Oviedo et al. 2012), so a more general conserved sequence pattern for CDP-alcohol phosphatidyltransferase family should be **D**$(x)_2$DG$(x)_2$AR$(x)_{7,8}$G$(x)_3$D$(x)_3$D, to include all members of this family identified thus far (proposed changes highlighted in bold).

Our site-directed mutagenesis studies on the consensus amino acid residues (Fig. 4.11) showed that all four aspartate, alanine and arginine residues are functionally relevant, as their loss abolishes or considerably reduces the catalytic activity (Fig. 4.11). These results generally agree with data from the *Sinorhizobium meliloti* bacterial choline phosphotransferase (CPT) (Solis-Oviedo et al. 2012), *S. cerevisiae* CPT1 (Williams and McMaster 1998) and *Brassica napus* plant aminoalcoholphosphatidyltransferase (AAPT1) (Qi et al. 2003), but the degree of reduction in their enzyme activities caused by the mutations may depend on the source and type of assays (*in vivo* or *in vitro*). A striking difference occurs when the second Asp (D2) is substituted by Ala; the small effect observed in yeast CPT1 and plant AAPT1 contrasts with the complete enzyme ablation of archaeal IPCT/DIPPS and bacterial CPT (also observed with D2N mutation). Indeed, this is the most conserved residue within CDP-alcohol phosphatidyltransferases (Fig. 4.5B). These results imply that the eukaryotic CPT1 and AAPT1 should still be able to retain the divalent ion in the catalytic site of D2A mutants. Furthermore, no mutations were performed on the first Asp (D1) of yeast CPT1 and plant AAPT1, as it was not yet considered to be part of the family consensus motif.

Notably, other family of peripheral membrane enzymes named phospholipase D, phosphatidylserine synthase type, which includes *E. coli* phosphatidylserine synthase, prokaryotic cardiolipin synthase, phospholipase D and some nucleases (Williams and McMaster 1998)

shares a different motif ($HxK(U)_4D(x)_4UUGO$). These enzymes catalyze the transfer of a substituted phosphate group, other than diphosphate or nucleotidyl residues, from one compound (donor) to another (acceptor). Although they appear capable of forming an identical phosphoester bond (Koonin 1996, Ponting and Kerr 1996) they do not require a divalent ion for catalysis and follow a ping-pong reaction type with an enzyme-bound intermediate (Raetz et al. 1987, Dechavigny et al. 1991, Larson and Dowhan 1976).

In contrast, kinetics studies on members of the CDP-alcohol phosphatidyltransferase family, such as synthases of phosphatidylserine from yeast (Raetz et al. 1987, Bae-Lee and Carman 1984) and *Bacillus licheniformis* (Dutt and Dowhan 1985), phosphatidylglycerophosphate from *E. coli* (Hirabayashi et al, 1976) and phosphatidylcholine from rat (Pontoni et al. 1985), suggest a sequential Bi–Bi mechanism for these enzymes. Assuming also a single displacement S_N2-like mechanism for DIPPS, we propose that Asp382 functions as base catalyst and promotes the direct nucleophilic attack of the hydroxyl group (on C3) of inositol-1-phosphate on the β-phosphoryl of CDP-inositol with formation of DIPP and CMP.

We expect a high degree of structural conservation among CDP-alcohol phosphatidyltransferases around the consensus region, located at the N-terminus of their sequences (TM2-TM3 in DIPPS, pocket 1) with subtle or more significant differences on the ligand-binding sites depending on their substrates. Members of this family are involved in the synthesis of phospholipids, which have important roles in cell viability and

infectivity of pathogenic bacteria (Gibellini and Smith 2010, Heath et al. 2001, Morii et al. 2013, Conover et al. 2008). Therefore, DIPPS structure may be of utmost value to structural homology modeling and in providing molecular targets for the rational design of novel drugs.

Work contributions

The X-ray structure of IPCT/DIPPS resulted from a collaborative effort between the team of H. Santos and the teams of the crystallographers M. Archer and V. Gordeliy. N. Borges and P. Nogly cloned the *ipct/dipps* gene from *A. fulgidus*. P. Nogly expressed and purified the protein for crystallization trials and constructed the mutants of IPCT/DIPPS for activity assays with the assistance of P. Ma. A. Remeeva and E. Round performed *in meso* crystallization. I. Gushchin, A. Ishchenko and V. Borshchevskiy collected crystallization data and solved the structure. A. M. Esteves produced and purified inositol-1-phosphate and CDP-inositol used for activity assays and co-crystallization studies, purified the protein, tested several detergents for solubilization, reconstituted the protein into lipid vesicles and carried out all the activity assays.

CHAPTER 5

General discussion

Contents

Already three decades ago the scientific community was amazed with the demonstration of microbial life and cell proliferation at temperatures near 100ºC by the team of Karl Stetter (Stetter et al. 1981). This discovery fueled a great deal of interest to understanding the biochemical basis of Life at such elevated temperatures, which were expected to disrupt irreversibly any biological complex organization. Today, we know that these extremely robust cells are formed by the same building blocks, i.e., amino acids, nucleotides, etc, as those of mesophiles, but the specific mechanisms underlying cell thermoresistance remain largely obscure yet. Probably, the main obstacle to progress in this field stems from the multitude of strategies and small interactions that contribute to the final thermostability of the cell and its structural components.

Our team invested many years of work in the characterization of compatible solutes of organisms adapted to hot environments, especially those isolated from shallow or deep hot springs in the sea. The view that compatible solutes of hyperthermophiles, particularly di-*myo*-inositol-phosphate (DIP) and mannosylglycerate (MG), play a role in thermoadaptation developed from the following basic information: i) their ionic nature contrasts with the neutral character of the vast majority of osmoregulators used by mesophiles; ii) the intracellular pools of specific solutes increase notably in response to heat stress; iii) ionic solutes are much better than uncharged compounds for thermostabilization of model proteins and other biomaterials.

One of the main goals of this thesis was to clarify the postulated involvement of compatible solutes, especially DIP, in thermoadaptation of hyperthermophiles. To this end, suitable mutants of *Pyrococcus furiosus* were constructed and characterized. Additionally, the catalytic mechanism of DIP synthesis was studied by setting efforts to determine the three-dimensional structure of the membrane enzyme, DIPPS, which catalyses the key biosynthetic step. Initially, we were also interested in understanding the regulatory mechanisms of DIP synthesis and searched for regulators of the heat shock response in hyperthermophiles, but this line of research did not advance at the pace required for a PhD work and was abandoned. Finally, we set out to investigate the heat shock response of a DIP-producer at the metabolomic level, with the purpose to obtain clues on metabolic strategies used by hyperthermophiles to cope with high temperature.

The membrane protein involved in DIP synthesis is the first representative of the CDP-alcohol phosphatidyltransferase family for which a 3D-structure has been solved

The biosynthesis of di-*myo*-inositol-phosphate (DIP) has been established in several hyperthermophilic organisms (Borges et al. 2006, Rodrigues et al. 2007, Rodionov et al. 2007). This compatible solute is synthesized from CTP and L-*myo*-inositol-1-phosphate via a three-step pathway, comprising CDP-inositol and DIP-phosphate (DIPP) as intermediate metabolites (see Chapter 1, Fig. 1.3). The two first reactions are catalyzed by

L-*myo*-inositol-1-phosphate cytidylyltransferase (IPCT) and CDP-inositol:L-*myo*-inositol-1-phosphate transferase (DIPPS). The genes encoding the IPCT and DIPPS enzymes have been identified (Rodrigues et al. 2007, Rodionov et al. 2007). As of January 7, 2015, there are 33 hyperthermophiles in which the two activities are encoded by a single gene, while separate genes are found in 8 hyperthermophilic organisms.

IPCT catalyses the synthesis of CDP-inositol, a metabolite assumed exclusive to DIP synthesis until a recent report demonstrated its involvement in the synthesis of glycerophosphoinositides of the thermophilic bacterium *Rhodothermus marinus* (Jorge et al. 2014). Interestingly, homologues of this novel synthase are predicted in the genomes of two bacterial classes, Sphingobacteria and Alphaproteobacteria, suggesting that this rare metabolite, CDP-inositol, is not restricted to organisms adapted to high temperature.

The three-dimensional structure of the soluble domain of IPCT/DIPPS from *A. fulgidus* has been solved at a resolution of 1.9 Å, however, the whole protein that comprises the membranar DIPPS domain resisted numerous attempts to obtaining good quality crystals (Brito et al. 2011). Therefore, the elucidation of the three-dimensional structure of IPCT/DIPPS is one of the main achievements of the present thesis. The strong efforts directed to the solubilization and purification of the bifunctional enzyme from *A. fulgidus* resulted in the determination of the structure of the apo- as well as the complex with CDP-inositol at 2.65 and 2.5 Å resolution, respectively (Chapter 4). This is the first structure of a

protein belonging to the CDP-alcohol phosphatidyltransferase family. Coincidently, a team of the Columbia University, USA, reported the structure of a second representative of the same family i.e., AF2299 from *A. fulgidus* (Sciara et al. 2014). Co-crystallization with CDP-glycerol was successful, but the authors failed to identify the function of AF2299 (Sciara et al. 2014). We speculate that AF2299 could be involved in the synthesis of diglycerol phosphate, the major solute used by *A. fulgidus* for osmoadaptation (Martins et al. 1997, Gonçalves et al. 2003). Unfortunately, all attempts to validate this hypothesis through heterologous expression in several hosts were unsuccessful. Similar to IPCT/DIPPS, the AF2299 contains an N-terminal cytosolic domain and a C-terminal membrane domain. Therefore, the presence of a cytosolic domain facilitates crystal development, probably by providing essential contact between the layers of molecules within crystals (Nogly et al. 2014, Sciara et al. 2014). It seems as though the cytosolic domain works as a crystallization chaperone (Filippo Mancia, personal communication).

The DIPPS domain is composed by six transmembrane α-helices and one shorter juxtamembrane helix. The transmembranar helices form a hydrophilic cavity in the center, which is open to the cytoplasm. Interestingly, the three aspartate residues belonging to the consensus sequence are located in one side of this cavity. An additional aspartate is located in this cavity and, despite its absence in the consensus sequence of the CDP-alcohol phosphatidyltransferases, is conserved among proteins of this family. Site-directed mutagenesis of the four conserved aspartate

residues to either alanine or asparagine residues showed the importance of those amino acids for DIPPS activity. Based on these findings we propose a new conserved sequence domain for the CDP-alcohol phosphatidyltransferases family: **D**$(x)_2$DG$(x)_2$AR$(x)_{7\text{-}12}$G$(x)_3$D$(x)_3$D, where the alteration is highlighted in bold. Co-crystallization of IPCT/DIPPS bound with CDP-inositol provided insight into the catalytic mechanism. The structure of the complex validated the location of the active site on the above-mentioned hydrophilic cavity lined by the four conserved aspartate residues. The amino acids D_1, D_2 and D_3 (where 1, 2, and 3 correspond to the position of the aspartate residues in the consensus sequence), together with the pyrophosphate group of CDP, form the coordination sphere of the magnesium ion. In contrast, D_4 does not appear to have a role on substrate binding due to its distant position to the ligand (see chapter 4, Fig. 4.6C). Interestingly, the position of the CDP moiety from CDP-inositol is in agreement with the results published by Sciara and co-workers (2014) for AF2299. These authors reported the three-dimensional structure of this protein in the apo-, CMP-bound, CDP-bound, and CDP-glycerol-bound states. Moreover, they reported that D_4 has a direct role in the mechanism of phosphotransfer in proteins of the CDP-alcohol phosphatidyltransferase family. This interpretation is supported by earlier kinetic studies on members of this family that suggest a sequential Bi-Bi mechanism involving a nucleophilic attack of the acceptor on the β-phosphorous of the CDP-alcohol donor (Pontoni et al. 1985, Schlame et al. 1995). Based on our co-crystallization studies (not yet published), and the

results published by Sciara and co-workers (2014), we propose the following catalytic mechanism for DIPPS: the residue D_4 (Asp382, according to DIPPS sequence) acts as a catalytic base, removing a proton from the hydroxyl group (C3) of the L-*myo*-inositol-1-phosphate (acceptor alcohol), facilitating the nucleophilic attack by the activated L-*myo*-inositol-1-phosphate on the β-phosphorus of CDP-inositol. This process results in the formation of CMP and the phosphorylated form of DIP via an associative (S_2N-like) transition state (Fig. 5.1). Thus far, our attempts to obtain crystals of IPCT/DIPPS with both ligands failed, but we think that this information is essential to substantiate the proposed transition state. Curiously, the proposed catalytic mechanism as well as the active site configuration for CDP-alcohol phosphatidyltransferases is similar to those described for phosphotransferases (e.g. kinases), despite the unrelated structures of these two classes of enzymes (Matte et al. 1998).

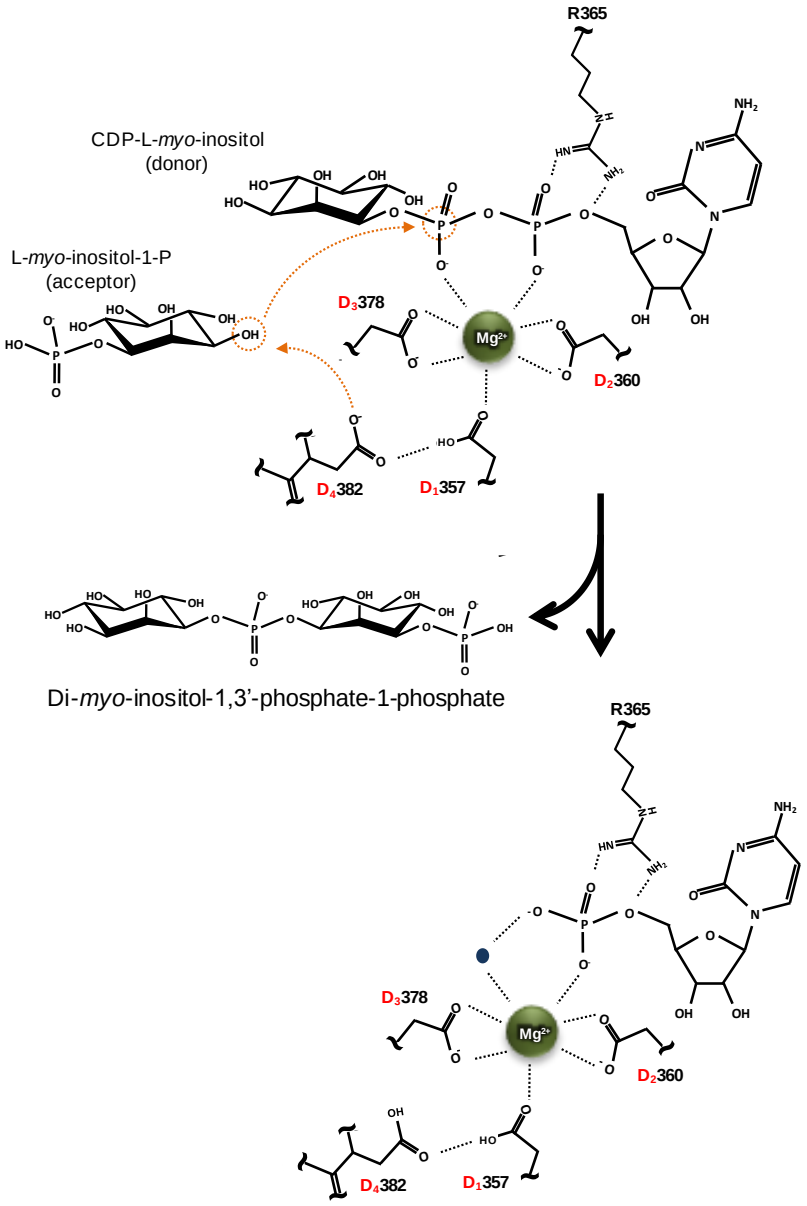


FIGURE 5.1 The catalytic mechanism proposed for DIPPS. The aspartate D_4 deprotonates the C3 position of the acceptor alcohol, L-*myo*-inositol-1-phosphate.

The activated molecule attacks the β-phosphorus of the alcohol donor (CDP-inositol), producing the phosphorylated intermediate, di-*myo*-inositol-1,3'-phosphate-1-phosphate (DIPP). Once the ester bond is formed and DIPP released, a water molecule (blue circle) replaces the β-phosphate in the coordination of the metal ion.

Maintaining the functional structure of membrane proteins throughout the production, extraction and purification procedures is a demanding goal and, unfortunately, we failed to detect DIPPS activity in the pure preparations of IPCT/DIPPS. Nevertheless, we are confident on the physiological significance of the structure determined, since a different team (Sciara et al.) working independently in a different representative of the CDP-alcohol phosphatidyltransferase family, obtained a structure with an identical architecture.

Available data on the biochemical properties of DIPPS are summarized in Table 5.1. The kinetic characterization of DIPPS is not straightforward since it is part of a bifunctional enzyme and the two activities share a common substrate (inositol-1-phosphate). Moreover, the rarity of the substrates (CDP-inositol and inositol-1-phosphate) demanded laborious production and purification methods. Also, ^{31}P-NMR was very useful to monitor time courses of exquisite product formation (DIPP, or CDP-inositol) due to the great analytical power of this technique, but has the drawback of low sensitivity which limits the K_m range that can be investigated. Perhaps, the use of radiolabelled assays should be implemented to take advantage of the high sensitivity of this methodology. However, the total procedure is expected to be extremely time-consuming.

TABLE 5.1 Summary of the biochemical properties of the recombinant IPCT/DIPPS from *A. fulgidus.*

	Substrates	Functionality
IPCT		
L-*myo*-Inositol1P +	CTP [a]	+
L-*myo*-Inositol1P +	ATP or GTP or UTP [a]	-
CTP +	*myo*-inositol or D-*myo*-inositol1P or glycerol or DL-glycerol3P [a]	-
DIPPS		
L-*myo*-Inositol1P +	CDP-L-*myo*-inositol or CDP-D-*myo*-inositol or CDP-glycerol [a]	+
L-*myo*-Inositol1P +	CDP-diacylglycerol [b]	-

Kinetic parameters [c]		K_m (mM)	V_{max} (μmol/min/mg protein)
IPCT	CTP	0.58 ± 0.08	62.9 ± 1.4
	Inositol-1-phosphate	0.87 ± 0.09	
DIPPS	Inositol-1-phosphate	ND	ND
	CDP-L-*myo*-inositol	ND	

[a] Rodrigues et al. 2007.
[b] From radio-labeled assays as described by Jorge et al. 2014 (not shown).
[c] Brito et al. 2011.
ND not determined.

At this stage we deemed it interesting to examine the phylogenetic relationships of DIPPS with the other members of the CDP-alcohol phosphatidyltransferase family. Unlike DIPPS that catalyses the formation of a highly hydrophilic solute, all these proteins are involved in the synthesis of phospholipids. The relevant phylogenetic tree (Fig. 5.2, Sciara et al. 2014) reveals a close link between DIPPS and phosphatidylinositolphosphate synthase (PIPS). Actually an earlier study of our team on the evolutionary history of IPCT/DIPPS proposes the occurrence of a fusion event between an *ipct* gene that evolved from a sugar nucleotidyltransferase and a *dipps* gene that evolved from PIPS (Gonçalves et al. 2012).

Similar to DIP biosynthesis, in bacteria the phosphatidylinositol is synthesized via a two-step pathway from inositol-1-phosphate and CDP-diacylglycerol, involving PIPS and a phosphatase activity (Morii et al. 2010) In contrast, in eukaryotes the synthesis of phosphatidylinositol occurs via a single-step reaction catalyzed by the phosphatidylinositol synthase (PIS), which uses *myo*-inositol and CDP-diacylglycerol. Phylogenetically, the PIS cluster is more closely related with the phosphatidylcholine/ethanolamine synthase from Eukarya.

In the pathogenic bacterium *Mycobacterium tuberculosis*, the presence of phosphatidylinositol is essential for cell viability (Jackson et al. 2000). The distinct pathways used for the synthesis of phosphatidylinositol in Bacteria and Eukarya makes *M. tuberculosis* PIPS a good target for the development of new drugs against tuberculosis. Morii and co-workers

(2013) showed that four chemically synthesized analogues of the substrate inositol-1-phosphate inhibited to some extent the activity of *M. tuberculosis* PIPS. Unfortunately, effective arrest of bacterial growth demanded high concentration of inhibitors, and consequently none of the proposed antimicrobial compounds was useful for clinical applications (Moori et al. 2013). In view of these results, the elucidation of the three-dimensional structure of CDP-alcohol phosphatidyltransferase proteins gains further importance for the rational design of anti-mycobacterium drugs since they may provide good models for the yet unavailable structure of PIPS from *M. tuberculosis*.

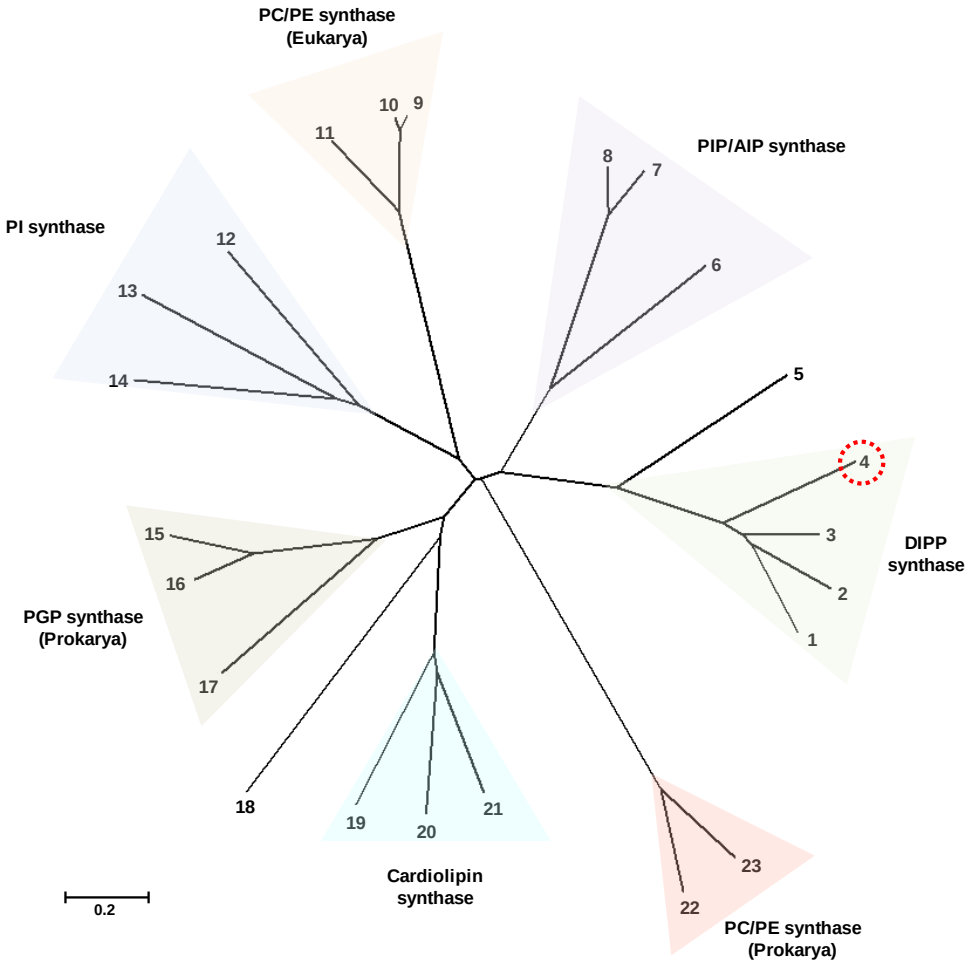


FIGURE 5.2 Maximum likelihood phylogenetic tree of 23 proteins from the CDP-alcohol phosphatidyltransferase family. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987) and distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965). All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6 (Tamura et al. 2013). The species and Uniprot IDs of proteins used to generate the tree are: 1) Q5JDA9, *Thermococcus*

kodakarensis; 2) O67379, *Aquifex aeolicus*; 3) Q8U1Z6, *Pyrococcus furiosus*; 4) O29976, *Archaeoglobus fulgidus*; 5) O27985, *Archaeoglobus fulgidus*; 6) O27726, *Methanothermobacter thermautotrophicus*; 7) P9WPG6, *Mycobacterium tuberculosis*; 8) Q9F7Y9, *Mycobacterium smegmatis*; 9) Q9Y6K0, *Homo sapiens*; 10) Q5ZKD1; *Gallus gallus*; 11) Q8WUD6, *Homo sapiens*; 12) Q8LBA6, *Arabidopsis thaliana*; 13) P06197; *Saccharomyces cerevisiae*; 14) O14735, *Homo sapiens*; 15) P44528, *Haemophilus influenza*; 16) P0ABF8, *Escherichia coli*; 17) Q68XS5, *Ricksettia typhi*; 18) P9WPG4, *Mycobacterium tuberculosis*; 19) O01916, *Caenorhabditis elegans*; 20) Q9UJA2, *Homo sapiens*; 21) Q8MZC4, *Drosophila melanogaster*; 22) Q98MN3, *Rhizobium loti*; 23) Q9KJY8, *Rhizobium meliloti*. DIPP, di-*myo*-inositol-1,3'-phosphate-1-phosphate; PIP, phosphatidylinositol phosphate; AIP, archaeatidylinositol phosphate; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; PGP, phosphatidylglycerophosphate. The IPCT/DIPPS from *Archaeoglobus fulgidus* is highlighted by a red circle. Adapted from Sciara et al. 2014.

Compatible solutes of *Pyrococcus furiosus* play a role in cell adaptation to heat stress

Studies on the accumulation of compatible solutes during growth of marine (hyper)thermophiles under stressful conditions revealed that DIP typically accumulates in response to heat stress, while the level of MG increases primarily in response to osmotic stress (Table 5.2). Notable exceptions to this trend are *Rhodothermus marinus* and *Palaeococcus ferrophilus* that

use MG both for osmo- and thermoadaptation (Silva et al. 1999, Neves et al. 2005).

TABLE 5.2 Accumulation of DIP and MG in (hyper)thermophilic organisms under heat and osmotic stress conditions (adapted from Santos et al. 2011).

	$T_{opt.}$ (ºC)	DIP	MG
Archaea			
Pyrococcus furiosus DSM 3638	98	↑ (Heat)	↑ (Osmotic)
Pyrococcus furiosus COM1	90	↑ (Heat)	↑ (Osmotic)
Methanotorris igneus	88	↑ (Heat)	n.d
Thermococcus stetteri	87	↑ (Heat; Osmotic)	↑ (Osmotic)
Thermococcus celer	87	↑ (Heat)	↑ (Osmotic)
Thermococcus litoralis	85	↑ (Heat)	↑ (Osmotic)
Thermococcus kodakarensis	85	↑ (Heat)	n.d.
Archaeoglobus fulgidus	83	↑ (Heat)	n.d.
Palaeococcus ferrophilus	83	n.d.	↑ (Heat; Osmotic)
Bacteria			
Aquifex pyrophilus	80	↑ (Heat)	n.d.
Rhodothermus marinus	65	n.d.	↑ (Heat; Osmotic)
Rubrobacter xylanophilus	60	↑ (Heat)	n.d.

$T_{opt.}$, Optimal temperature; DIP, di-*myo*-inositol-phosphate; MG, mannosylglycerate. ↑(Osmotic) and ↑(Heat) indicate that the intracellular level of the solute increases in response to osmotic or heat stress, respectively. n.d., not detected.

Under supraoptimal salinity conditions (low external water activity), many halophilic organisms accumulate intracellularly compatible solutes as an osmoregulatory strategy to balance the osmotic pressure between the external environment and the cytosol. The observation that organisms also accumulate compatible solutes under heat stress is puzzling, especially if

we consider that external water activity remains practically constant at high temperatures. Therefore, a correlation between the accumulation of compatible solutes and the thermoprotection of cell components is plausible. In fact, many *in vitro* studies demonstrated that solutes derived from (hyper)thermophiles exhibited superior ability to protect a variety of model proteins against thermal denaturation and aggregation, acting as chemical chaperones (Table 5.3). Recently, the effect of MG on the stabilization of α-synuclein was examined *in vivo* by using a yeast model of Parkinson's disease (Faria et al. 2013). The α-synuclein is a soluble unfolded protein found in the human brain; however, in several pathological conditions, like Parkinson's disease, α-synuclein forms insoluble fibrils that constitute a hallmark of neurodegenerative dysfunctions (Mezey et al. 1998). The yeast model was genetically manipulated to produce MG and, as a consequence, a reduction of 3.3-fold in the number of cells with α-synuclein inclusions was observed, reflecting the chemical chaperone activity of this solute. This was the first demonstration of the anti-aggregating ability of MG in the intracellular milieu (Faria et al. 2013). Despite this elegant *in vivo* study, a definitive proof of the thermoprotection role of thermolytes in the natural producers is still missing.

TABLE 5.3 Effect of MG and DIP on kinetic, thermodynamic and protein conformational stabilization of several proteins.

Protein	Organism/Type of assay	Conc (mM)	MG	DIP	Ref.
	Kinetic stabilization				
Alcohol dehydrogenase	*Pyrococcus furiosus*	500	Stabiliz.		1
	Bacillus stearothermophilus	500	Stabiliz.		1
	Baker yeast	500	Stabiliz.		1
Glutamate dehydrogenase	*Thermotoga maritima*	500	Stabiliz.		1
	Clostridium difficile	500	Stabiliz.		1
	Pyrococcus furiosus	500	No effect		1
Lactate dehydrogenase	Bovine liver	500	Stabiliz.		1
	Rabbit muscle	500	Stabiliz.	Destabiliz.	1,2
GAP dehydrogenase	*Pyrococcus woesei*	115		Stabiliz.	4
Hydrogenase	*Thermotoga maritima*	50		No effect	5
Pyruvate ferredoxin oxidoreductase	*Thermotoga maritima*	50		No effect	5
Malate dehydrogenase	Pig heart	500		Destabiliz.	3
	Thermodynamic stabilization				
Malate dehydrogenase	Pig heart	500	Stabiliz.	Stabiliz.	3
Nuclease A	*Staphylococcus aureus*	500	Stabiliz.	Stabiliz.	3
Lysozyme	Hen egg white	500	Stabiliz.	Stabiliz.	3
	Protein aggregation				
Malate dehydrogenase	Pig heart	500	Inhibition	Inhibition	3
Lactate dehydrogenase	Rabbit muscle	500	Inhibition	Inhibition	3
α-Synuclein	Human	100	Inhibition		6

References: 1, Ramos et al. 1997; 2, Borges et al. 2002; 3, Faria et al. 2008; 4, Scholz et al. 1992; 5, Ramakrishnan et al. 1997; 6, Faria et al., 2013. **Abbreviations:** Stabiliz.,

Stabilization; Destabiliz., Destabilization; Conc, Concentration; MG, mannosylglycerate; DIP, di-*myo*-inositol-phosphate.

So far, *Thermococcus kodakarensis* and *Pyrococcus furiosus* are the only DIP-accumulating organisms for which genetic tools are available. Thus, to examine the physiological role of DIP at supraoptimal temperatures, the gene encoding IPCT/DIPPS was deleted in both hyperthermophilic organisms. Comparison of the growth and solute profiles under heat stress of the parent strain and the DIP-deficient mutant of *T. kodakarensis* revealed that the missing solute was replaced by aspartate. Moreover, the mutant strain had a growth profile similar to the parental strain (Borges et al. 2010). This attempt failed to prove the role of DIP in the natural producer. Most likely, the alternative solute was taken up from the external medium. *P. furiosus* seemed a promising target organism because its pattern of solute accumulation is composed primarily by MG and DIP. In chapter 2 we show that the DIP-deficient strain of *P. furiosus* replaced DIP by MG at supraoptimal temperatures, without growth impair. This result reveals the interchangeable role of DIP and MG under heat stress, but also indicates that DIP is important for heat adaptation in *P. furiosus*, otherwise its replacement by a second solute would be superfluous (Fig. 5.3).

The strategy of solute replacement has been described for osmotic adjustment in a few halophiles (see details in discussion of chapter 2). In *Methanosarcina mazei* the accumulation of glutamate and alanine was not sufficient to compensate the lack of N^{ε}-acetyl-β-lysine under high salinity

conditions, suggesting that the alternative solutes are not fully suitable to cope with the osmotic stress imposed (Saum et al. 2009). A similar behavior was observed with the MG-deficient *P. furiosus* mutant grown under osmotic stress; this mutant showed a defective growth under osmotic stress, despite the substitution of the missing MG by DIP and aspartate. It seems that the accumulation of DIP and aspartate was not sufficient to compensate satisfactorily for the absence of MG in the MG-deficient strain (Chapter 2).

The characterization of solute pattern accumulation by the *P. furiosus* mutants revealed an interesting feature: the DIP pool was enhanced in both stress conditions (heat and osmotic), while the accumulation of MG was only triggered by osmotic stress. Therefore, the sensing/regulatory mechanisms involved in DIP synthesis show certain flexibility. In *P. furiosus*, the transcriptional regulation of the gene encoding IPCT/DIPPS has not been studied. However, a transcriptional repressor of the heat shock response was identified in *P. furiosus* and *T. kodakarensis* (Vierke et al. 2003, Keese et al. 2010, Kanai et al. 2010). Under optimal growth conditions, the transcriptional repressor Phr blocks the transcription of several heat shock related genes, such as *hsp20* and *aaa*$^{+}$*atpase*. Under heat stress, the Phr releases the promoter region and the transcription of heat shock genes is activated (Vierke et al. 2003, Keese et al. 2010). More recently, disruption of the gene encoding the Phr in *T. kodakarensis* resulted in high transcription levels of the *ipct/dipps* gene

(Kanai et al. 2010). Therefore, we assume that Phr also regulates the transcript levels of *ipct/dipps* in *P. furiosus*.

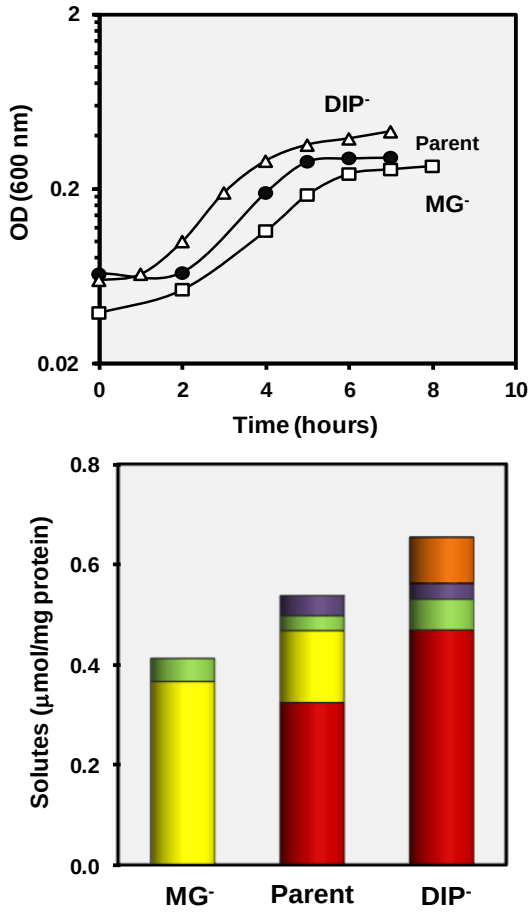


FIGURE 5.3 Growth curves and composition of organic solutes in MG-deficient (MG^-), DIP-deficient (DIP^-), and parent strain of *Pyrococcus furiosus* under heat stress (98ºC, 2.8% NaCl). Mannosylglycerate (red); di-*myo*-inositol-phosphate (yellow); alanine (green): aspartate (purple); L-*myo*-inositol-1-phosphate (orange).

It would be interesting to investigate transcript levels of *ipct/dipps* and mannosyl-3-phosphoglycerate synthase genes in a Phr-deficient strain of

P. furiosus under heat and osmotic stress. This experiment will reveal if Phr responds only to heat stress or also to osmotic stress. It is worth pointing out that the regulatory mechanism(s) leading to the accumulation of compatible solutes in hyperthermophiles under osmotic stress are unknown. Therefore, the clarification of the regulatory mechanisms underlying DIP or MG synthesis is a very interesting goal for further work.

Heat shock response in *Pyrococcus furiosus*: Integration of metabolomic and transcriptomic data

The accumulation of compatible solutes by marine (hyper)thermophilic organisms under heat stress has been extensively studied during the last decades (Santos et al. 2011). However, the metabolome response to heat shock in these (hyper)thermophiles has never been characterized. Thus, we decided to study the changes in the metabolome of *Pyrococcus furiosus* (DSM 3638) in response to heat shock.

One of the main challenges was to optimize the sampling and quenching conditions for a hyperthermophilic organism that reaches low cell densities. *Pyrococcus furiosus* grows to an optical density (600 nm) of 0.3-0.4 under optimal growth conditions. Therefore, the required culture volume for the identification and quantification of metabolites is 3 to 4-fold higher when compared with model mesophilic organisms, like *E. coli*. Large working volumes will affect the sampling time and consequently the promptness of metabolism quenching. The metabolome of *P. furiosus* in response to heat shock (45 minutes at 97ºC) was studied using two

complementary techniques (NMR and MS). The much lower sensitivity of the NMR methodology demanded the sampling of larger culture volumes (250 ml for NMR versus 0.6 ml for MS). This disparity of sample volumes imposed different methods for cell separation, i.e., centrifugation for NMR and fast-filtration for LC/MS. In view of these experimental differences, the integration of NMR and MS results should be regarded with caution.

To our knowledge, the metabolome analysis of *P. furiosus* under heat shock reported here is the first performed with any hyperthermophilic organism. Metabolomic data is typically complemented with other “omic” studies, namely transcriptomic and proteomic, allowing the construction of an overall picture of the microbial physiology. Currently, there is only a study on the effect of heat shock in the transcription of approximately 200 genes of *P. furiosus* (Shockley et al. 2003). Despite the scarcity of information, an integrated scheme of the heat shock response in *P. furiosus* is presented (Fig. 5.4).

Besides the accumulation of DIP and MG, the pools of maltose and two N-acetylated NDP-sugars (UDP-*N*-acetylglucosamine and UDP-*N*-acetylgalactosamine) increased under heat shock. We speculate that N-acetylated NDP-sugars could be channeled to the synthesis of the cell wall possibly to maintain the integrity of the cell structure. The transcript levels of maltose transporters and glycoside hydrolase genes increased in response to heat shock (Shockley et al. 2003). These results were interpreted by the authors as a need to counterbalance the ATP depletion caused by heat stress (Shockley et al. 2003). Therefore, the accumulation

of maltose reported in chapter 3 is puzzling. Despite the increased transcript levels of the glycoside hydrolase genes, it seems that the metabolism of maltose is slower at supraoptimal temperatures.

In contrast, the pools of aspartate and metabolites associated with the glycolysis and Krebs cycle were negatively affected by heat shock. In accordance, the transcript levels of the gene encoding glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR), the enzyme that leads to the formation of 3-phosphoglycerate, decreases at supraoptimal temperatures. Transcriptomic data (Shockley et al. 2003) also showed an increase in the transcript level of the L-*myo*-inositol-phosphate synthase, the enzyme involved in the formation of the precursor for DIP synthesis. Unfortunately, our metabolomic analysis did not provide information on the level of L-*myo*-inositol-1-phosphate. However, it is clear that the flux towards the synthesis of DIP was enhanced under heat stress since the final pool of DIP is higher.

As expected, the transcript levels of the genes encoding molecular chaperones (HSP20, HSP60, AAA^+ ATPase) were induced by heat shock (Shockley et al. 2003). These chaperones are required for the correct folding, assembly, transport, and degradation of proteins within the cell. In the future, it would be interesting to investigate whether the decreased pools of the intermediates of glycolysis and Krebs cycle are associated with lower fluxes through central metabolism. Further work must be done to improve the statistical significance of the results and to extend the number of metabolites identified and validate the preliminary conclusions.

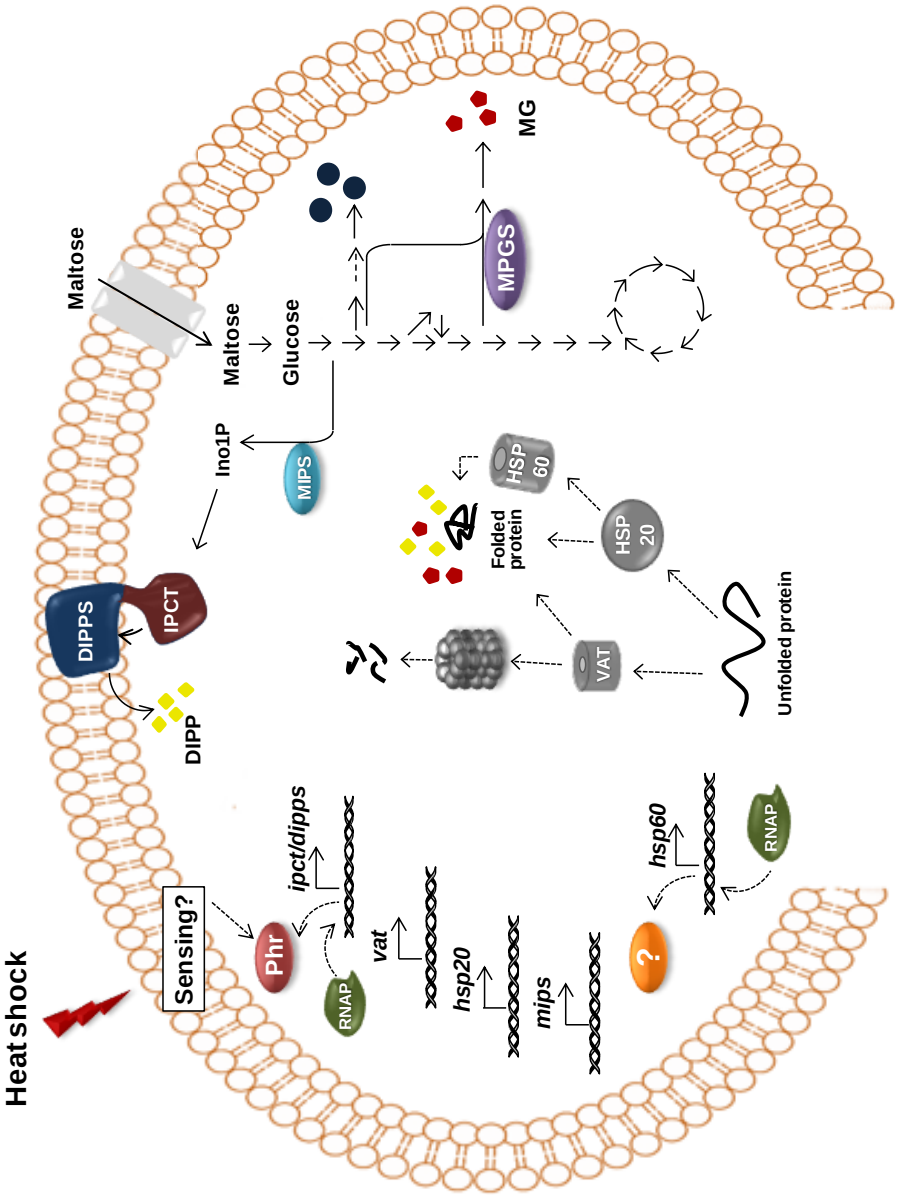


FIGURE 5.4 Schematic overview of the heat shock response in *Pyrococcus furiosus*. DIP (yellow diamonds), MG (red diamonds) levels increased with

temperature, and may serve as chemical chaperones. N-acetylated sugars (blue circles) are also increased. In this hyperthermophile, the transcription repressor (Phr) regulates the expression of the genes encoding AAA^{+}ATPase (VAT homologues), small heat shock proteins (HSP20 or sHSP) and L-*myo*-inositol-1-phosphate synthase (MIPS). In the closely related organism, *T. kodakarensis*, the Phr regulates also the expression of the *ipct/dipps* gene, so it is likely to assume the same function in *P. furiosus*. Although responsive to heat stress, the *hsp60* is not regulated by Phr in *T. kodakarensis* implying the existence of at least another transcriptional regulator. Given the phylogenetic proximity we can hypothesize that a similar regulatory architecture occurs in *P. furiosus* as well. RNAP, RNA polymerase; MPGS, mannosyl-3-phosphoglycerate synthase.

References

Adams MW, Holden JF, Menon AL, Schut GJ, Grunden AM, Hou C, Hutchins MA, Jenney Jr FE, Kim C, Ma K, Pan G, Roy R, Sapra R, Story SV, Verhagen MF. 2001. Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 183:716-724.

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213-221.

Alarico S, Empadinhas N, Mingote A, Simões C, Santos MS, da Costa MS. 2007. Mannosylglycerate is essential for osmotic adjustment in *Thermus thermophilus* strains HB27 and RQ-1. Extremophiles 11:833-840.

Alexandrov AI, Mileni M, Chien EY, Hanson MA, Stevens RC. 2008. Microscale fluorescent thermal stability assay for membrane proteins. Structure 16:351-359.

Allers T, Mevarech M. 2005. Archaeal genetics - the third way. Nat Rev Genet 6:58-73.

Atomi H, Matsumi R, Imanaka T. 2004. Reverse gyrase is not a prerequisite for hyperthermophilic life. J Bacteriol 186:4829-4833.

Atomi H, Sato T, Kanai T. 2011. Application of hyperthermophiles and their enzymes. Curr Opin Biotechnol 22:618-626.

Auerbach G, Huber R, Grättinger M, Zaiss K, Schurig H, Jaenicke R, Jacob U. 1997. Closed structure of phosphoglycerate kinase from *Thermotoga maritima* reveals the catalytic mechanism and determinants of thermal stability. Structure 5:1475-1483.

Avonce N, Mendoza-Vargas A, Morett E, Iturriaga G. 2006. Insights on the evolution of trehalose biosynthesis. BMC Evol Biol 19:109.

Bae-Lee MS, Carman GM. 1984. Phosphatidylserine synthesis in *Saccharomyces cerevisiae*. Purification and characterization of membrane-associated phosphatidylserine synthase. J Biol Chem 259:10857-10862.

Basen M, Sun J, Adams MWW. 2012. Engineering a hyperthermophilic archaeon for temperature-dependent product formation. MBio 3:1-7.

Bertoldo C, Antranikian G. 2006. The order *Thermococcales*. The Prokaryotes 69-81.

Bidle KA, Kirkland PA, Nannen JL, Maupin-Furlow JA. 2008. Proteomic analysis of *Haloferax volcanii* reveals salinity-mediated regulation of the stress response protein PspA. Microbiology 154:1436-1443.

Bingol K, Brüschweiler R. 2014. Multidimensional approaches to NMR-based metabolomics. Anal Chem 86:47-57.

Blöchl E, Rachel R, Burggraf S, Hafenbradl D, Jannasch HW, Stetter KO. 1997. *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 degrees C. Extremophiles 1:14-21.

Bolten CJ, Kiefer P, Letisse F, Portais JC, Wittmann C. 2007. Sampling for metabolome analysis of microorganisms. Anal Chem 79:3843-3849.

Boonyaratanakornkit BB, Simpson AJ, Whitehead TA, Fraser CM, El-Sayed NM, Clark DS. 2005. Transcriptional profiling of the hyperthermophilic methanarchaeon *Methanococcus jannaschii* in response to lethal heat and non-lethal cold shock. Environ Microbiol 7:789-797.

Boonyaratanakornkit BB, Miao LY, Clark DS. 2007. Transcriptional responses of the deep-sea hyperthermophile *Methanocaldococcus jannaschii* under shifting extremes of temperature and pressure. Extremophiles 11:495-503.

Borges N, Ramos A, Raven NDH, Sharp RJ, Santos H. 2002. Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. Extremophiles 6:209-216.

Borges N, Marugg JD, Empadinhas N, da Costa MS, Santos H. 2004. Specialized roles of the two pathways for the synthesis of mannosylglycerate in osmoadaptation and thermoadaptation of *Rhodothermus marinus*. J Biol Chem 279:9892-9898.

Borges N, Gonçalves LG, Rodrigues MV, Siopa F, Ventura R, Maycock C, Lamosa P, Santos H. 2006. Biosynthetic pathways of inositol and glycerol phosphodiesters used by the hyperthermophile *Archaeoglobus fulgidus* in stress adaptation. J Bacteriol 188:8128-8135.

Borges N, Matsumi R, Imanaka T, Atomi H, Santos H. 2010. *Thermococcus kodakarensis* mutants deficient in di-*myo*-inositol phosphate use aspartate to cope with heat stress. J Bacteriol 192:191-197.

Borges N, Jorge CD, Gonçalves LG, Gonçalves S, Matias PM, Santos H. 2014. Mannosylglycerate: structural analysis of biosynthesis and evolutionary history. Extremophiles 18:835-852.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254.

Bräsen C, Esser D, Rauch B, Siebers B. 2014. Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. Microbiol Mol Biol Rev 78:89-175.

Bremer E, Kramer R. 2000. Coping with osmotic challanges: osmoregulation through accumulation and release of compatible solutes in Bacteria. p. 79-96. *In* G. Storz and R. Hengge-Aronis (eds.), Bacterial stress responses. ASM Press, Washington, DC.

Bridger SL, Clarkson SM, Stirrett K, DeBarry MB, Lipscomb GL, Schut GJ, Westpheling J, Scott RA, Adams MW. 2011. Deletion strains reveal metabolic roles for key elemental sulfur-responsive proteins in *Pyrococcus furiosus*. J Bacteriol 193:6498-504.

Bridger SL, Lancaster WA, Poole FL 2nd, Schut GJ, Adams MW. 2012. Genome sequencing of a genetically tractable *Pyrococcus furiosus* strain reveals a highly dynamic genome. J Bacteriol 194:4097-4106.

Brito JA, Borges N, Vonrhein C, Santos H, Archer M. 2011. Crystal structure of *Archaeoglobus fulgidus* CTP:inositol-1-phosphate cytidylyltransferase, a key enzyme for di-*myo*-inositol-phosphate synthesis in (hyper)thermophiles. J Bacteriol 193:2177-2185.

Brito-Echeverría J, Lucio M, López-López A, Antón J, Schmitt-Kopplin P, Rosselló-Móra R. 2011. Response to adverse conditions in two strains of the extremely halophilic species *Salinibacter ruber*. Extremophiles 15:379-389.

Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P. 2008. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. Nat Rev Microbiol 6:245-252.

Brock TD, Brock KM, Belly RT, Weiss RL. 1972. *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol 84:54-68.

Brock TD. 1978. Thermophilic microorganisms and Life at high temperatures. New York: Springer-Verlag.

Brown AD. 1976. Microbial water stress. Bacteriol Rev 40:803-846.

Buckel W, Thauer RK. 2013. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation. Biochim Biophys Acta 1827:94-113.

Cánovas D, Borges N, Vargas C, Ventosa A, Nieto JJ, Santos H. 1999. Role of Nγ-acetyldiaminobutyrate as an enzyme stabilizer and an intermediate in the biosynthesis of hydroxyectoine. Appl Environ Microbiol 65: 3774-3779.

Chen L, Spiliotis ET, Roberts MF. 1998. Biosynthesis of di-*myo*-inositol-1,1'-phosphate, a novel osmolyte in hyperthermophilic archaea. J Bacteriol 180:3785-3792.

Chou CJ, Shockley KR, Conners SB, Lewis DL, Comfort DA, Adams MW, Kelly RM. 2007. Impact of substrate glycoside linkage and elemental sulfur on bioenergetics of and hydrogen production by the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl Environ Microbiol 73:6842-6853.

Conover GM, Martinez-Morales F, Heidtman MI, Luo ZQ, Tang M, Chen C, Geiger O, Isberg RR. 2008. Phosphatidylcholine synthesis is required for optimal function of *Legionella pneumophila* virulence determinants. Cell Microbiol 10:514-528.

da Costa MS, Santos H, Galinski EA. 1998. An overview of the role and diversity of compatible solutes in Bacteria and Archaea. Adv Biochem Eng Biotechnol 61:117-153.

Daniel RM, Cowan DA. 2000. Biomolecular stability and life at high temperatures. Cell Mol Life Sci 57:250–264.

Daniel RM, Eckert RV, Holden JF, Truter J, Crowan DA. 2004. The stability of biomolecules and the implications for life at high temperatures. *In* William S.D. Wilcock, Edward F Delong, Deborah S. Kelley, John A. Baross and S. Craig Cary (eds), The subsea floor biosphere at mid-ocean ridges. American geophysical union, USA.

Danno A, Fukuda W, Yoshida M, Aki R, Tanaka T, Kanai T, Imanaka T, Fujiwara S. 2008. Expression profiles and physiological roles of two types of prefoldins from the hyperthermophilic archaeon *Thermococcus kodakaraensis*. J Mol Biol 382:298-311.

Dechavigny A, Heacock PN, Dowhan W. 1991. Sequence and inactivation of the PSS gene of *Escherichia coli* - phosphatidylethanolamine may not be essential for cell viability. J Biol Chem 266:5323-5332.

DeLano WLT. 2002. The PyMOL Molecular Graphics System (0.98 ed). DeLano Scientific, San Carlos, CA, USA.

Doemel W N, Brock TD. 1970. The upper temperature limit of *Cyanidium caldarium*. Arch Mikrobiol 72:326-332.

Dutt A, Dowhan W. 1985. Purification and characterization of a membrane-associated phosphatidylserine synthase from *Bacillus licheniformis*. Biochemistry 24:1073-1079.

Eisenberg H, Wachtel EJ. 1987. Structural studies of halophilic proteins, ribosomes, and organelles of bacteria adapted to extreme salt concentrations. Annu Rev Biophys Biophys Chem 16:69-92.

Elbein AD, Pan YT, Pastuszak I, Carroll D. 2003. New insights on trehalose: a multifunctional molecule. Glycobiology 13:17R-27R.

Empadinhas N, Marugg JD, Borges N, Santos H, da Costa MS. 2001. Pathway for the synthesis of mannosylglycerate in the hyperthermophilic archaeon *Pyrococcus horikoshii*. Biochemical and genetic characterization of key enzymes. J Biol Chem 276:43580-43588.

Empadinhas N, Albuquerque L, Henne A, Santos H, da Costa MS. 2003. The bacterium *Thermus thermophilus*, like hyperthermophilic archaea, uses a two-step pathway for the synthesis of mannosylglycerate. Appl Environ Microbiol 69:3272-3279.

Empadinhas N, da Costa MS. 2008. Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. Int Microbiol 11:151-61.

Empadinhas N, da Costa MS. 2011. Diversity, biological roles and biosynthetic pathways for sugar-glycerate containing compatible solutes in bacteria and archaea. Environ Microbiol 13:2056-2077.

Evans P. 2006. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr 62:72-82.

Faria C, Jorge CD, Borges N, Tenreiro S, Outeiro TF, Santos H. 2013. Inhibition of formation of α-synuclein inclusions by mannosylglycerate in a yeast model of Parkinson's disease. Biochim Biophys Acta 1830:4065-4072.

Faria TQ, Knapp S, Ladenstein R, Maçanita AL, Santos H. 2003. Protein stabilization by compatible solutes: Effect of mannosylglycerate on unfolding thermodynamics and activity of Ribonuclease A. Chembiochem 4:734-741.

Faria TQ, Mingote A, Siopa F, Ventura R, Maycock C, Santos H. 2008. Design of new enzyme stabilizers inspired by glycosides of hyperthermophilic microorganisms. Carbohydr Res 343:3025-3033.

Farkas J, Stirrett K, Lipscomb GL, Nixon W, Scott RA, Adams MWW, Westpheling J. 2012. Recombinogenic properties of *Pyrococcus furiosus* strain COM1 enable rapid selection of targeted mutants. Appl Environ Microbiol 78:4669-4676.

Feder ME, Hofmann GE. 1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annu Rev Physiol 61:243-282.

Fernandes C, Mendes V, Costa J, Empadinhas N, Jorge C, Lamosa P, Santos H, da Costa MS. 2010. Two alternative pathways for the synthesis of the rare compatible solute mannosylglucosylglycerate in *Petrotoga mobilis*. J Bacteriol 192:1624-1633.

Ferrer C, Mojica FJ, Juez G, Rodríguez-Valera F. 1996. Differentially transcribed regions of *Haloferax volcanii* genome depending on the medium salinity. J Bacteriol 178:309-313.

Fiala G, Stetter KO. 1986. *Pyrococcus furiosus* sp. Nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch Microbiol 145:56-61.

Forterre P, Bouthier De La Tour C, Philippe H, Duguet M. 2000. Reverse gyrase from hyperthermophiles: probable transfer of a thermoadaptation trait from archaea to bacteria. Trends Genet 16:152-154.

Forterre P. 2002. The origin of DNA genomes and DNA replication proteins. Curr Opin Microbiol 5:525-532.

Forterre P. 2009. Strategies of hyperthermophiles in nucleic acids adaptation to high temperature. Enciclopedia of Life Support Systems, UNESCO

Galinski, E. A. 1995. Osmoadaptation in bacteria, p. 273-328. *In* R. K. Poole (ed.), Advances in Microbial Physiology. Academic Press, London.

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. 2000. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11:4241-4257.

Ghosh B, Chakraborti T, Kar P, Dey K, Chakraborti S. 2009. Solubilization, purification, and reconstitution of alpha 2 beta 1 isozyme of Na+/K+ -ATPase from caveolae of pulmonary smooth muscle plasma membrane: comparative studies with DHPC, C12E8, and Triton X-100. Mol Cell Biochem 323:169-184.

Gibellini F, Smith TK. 2010. The Kennedy pathway-*de novo* synthesis of phosphatidylethanolamine and phosphatidylcholine. IUBMB Life 62:414-428.

Gonçalves LG, Borges N, Serra F, Fernandes PL, Dopazo H, Santos H. 2012. Evolution of the biosynthesis of di-*myo*-inositol phosphate, a marker of adaptation to hot marine environments. Environ Microbiol 14:691-701.

Gonçalves LG, Huber R, da Costa MS, Santos H. 2003. A variant of the hyperthermophile *Archaeoglobus fulgidus* adapted to grow at high salinity. FEMS Microbiol Lett 218:239-244.

Gonçalves LG, Lamosa P, Huber R, Santos H. 2008. Di-*myo*-inositol phosphate and novel UDP-sugars accumulate in the extreme hyperthermophile *Pyrolobus fumarii*. Extremophiles 12:383-389.

Gonçalves S, Borges N, Esteves AM, Victor BL, Soares CM, Santos H, Matias PM. 2010. Structural analysis of *Thermus thermophilus* HB27 mannosyl-3-phosphoglycerate synthase provides evidence for a second catalytic metal ion and new insight into the retaining mechanism of glycosyltransferases. J Biol Chem 285:17857-17868.

Gonçalves S, Esteves AM, Santos H, Borges N, Matias PM. 2011. Three-dimensional structure of mannosyl-3-phosphoglycerate phosphatase from *Thermus thermophilus* HB27: a new member of the haloalcanoic acid dehalogenase superfamily. Biochemistry 50:9551-9567.

Gordeliy VI, Labahn J, Moukhametzianov R, Efremov R, Granzin J, Schlesinger R, Büldt G, Savopol T, Scheidig AJ, Klare JP, Engelhard M. 2002. Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex. Nature 419:484-487.

Gouffi K, Pica N, Pichereau V, Blanco C. 1999. Disaccharides as a new class of nonaccumulated osmoprotectants for *Sinorhizobium meliloti*. Appl Environ Microbiol 65:1491-1500.

Greaves RB, Warwicker J. 2007. Mechanisms for stabilisation and the maintenance of solubility in proteins from thermophiles. BMC Struct Biol 29:7-18.

Guisbert E, Yura T, Rhodius VA, Gross CA. 2008. Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. Microbiol Mol Biol Rev 72:545-554.

Hartmann E, König H. 1990. Comparison of the biosynthesis of the methanobacterial pseudomurein and the eubacterial murein. Naturwissenschaften 77:472-475.

Hasan CM, Shimizu K. 2008. Effect of temperature up-shift on fermentation and metabolic characteristics in view of gene expressions in *Escherichia coli*. Microb Cell Fact 2:7:35.

Haslbeck M. 2002. sHsps and their role in the chaperone network. Cell Mol Life Sci 59:1649-1657.

Heath RJ, White SW, Rock CO. 2001. Lipid biosynthesis as a target for antibacterial agents. Prog Lipid Res 40:467-497.

Hirabayashi T, Larson TJ, Dowhan W. 1976. Membrane-associated phosphatidylglycerophosphate synthetase from *Escherichia coli* - purification by substrate affinity chromatography on cytidine 5'-diphospho-1,2-diacyl-*sn*-glycerol sepharose. Biochemistry 15:5205-5211.

Ho BK, Gruswitz F. 2008. HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. BMC Struct Biol 8:49.

Hutchins AM, Mai X, Adams MW. 2001. Acetyl-CoA synthetases I and II from *Pyrococcus furiosus*. Methods Enzymol 331:158-167.

Imanaka H, Yamatsu A, Fukui T, Atomi H, Imanaka T. 2006. Phosphoenolpyruvate synthase plays an essential role for glycolysis in the modified Embden-Meyerhof pathway in *Thermococcus kodakarensis*. Mol Microbiol 61:898-909.

Jackson M, Crick DC, Brennan PJ. 2000. Phosphatidylinositol is an essential phospholipid of mycobacteria. J Biol Chem 275:30092-30099.

Jenney FE Jr, Adams MW. 2008. Hydrogenases of the model hyperthermophiles. Ann N Y Acad Sci 1125:252-266.

Joghee NN, Jayaraman G. 2014. Metabolomic characterization of halophilic bacterial isolates reveals strains synthesizing rare diaminoacids under salt stress. Biochimie 102:102-111.

Johansen E, Kibenich A. 1992. Isolation and characterization of IS1165, an insertion sequence of *Leuconostoc mesenteroides* subsp. *cremoris* and other lactic acid bacteria. Plasmid 27:200-206.

Johnsborg O, Eldholm V, Håvarstein LS. 2007. Natural genetic transformation: prevalence, mechanisms and function. Res Microbiol 158:767-778.

Jorge CD, Borges N, Santos H. 2014. A novel pathway for the synthesis of inositol phospholipids uses CDP-inositol as donor of the polar head group. Environ Microbiol doi: 10.1111/1462-2920.12734.

Jozefczuk S, Klie S, Catchpole G, Szymanski J, Cuadros-Inostroza A, Steinhauser D, Selbig J, Willmitzer L. 2010. Metabolomic and transcriptomic stress response of *Escherichia coli*. Mol Syst Biol 6:1-16.

Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577-2637.

Kalisiak J, Trauger SA, Kalisiak E, Morita H, Fokin VV, Adams MW, Sharpless KB, Siuzdak G. 2008. Identification of a new endogenous metabolite and the characterization of its protein interactions through an immobilization approach. J Am Chem Soc 131:378-386.

Kanai T, Takedomi S, Fujiwara S, Atomi H, Imanaka T. 2010. Identification of the Phr-dependent heat shock regulon in the hyperthermophilic archaeon, *Thermococcus kodakaraensis*. J Biochem 147:361-370.

Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. Science 336:1030-1033.

Kawamura T, Watanabe N, Tanaka I. 2008. Structure of mannosyl-3-phosphoglycerate phosphatase from *Pyrococcus horikoshii*. Acta Crystallogr D Biol Crystallogr 64:1267-1276.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649.

Keese AM, Schut GJ, Ouhammouch M, Adams MW, Thomm M. 2010. Genome-wide identification of targets for the archaeal heat shock regulator Phr by cell-free transcription of genomic DNA. J Bacteriol 192:1292-1298.

Kengen SW, de Bok FA, van Loo ND, Dijkema C, Stams AJ, de Vos WM. 1994. Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus*. J Biol Chem 269:17537-17541.

Kengen SW, Stams AJM, de Vos WM. 1996. Sugar metabolism of hyperthermophiles, FEMS Microbiol Rev 18:119-137.

Kessler H, Oschkinat H, Griesinger C, Bermel W. 1986. Transformation of homonuclear two-dimensional NMR techniques into one-dimensional techniques using Gaussian pulses. J Magn Reson 70:106-133.

Kiefer P, Nicolas C, Letisse F, Portais JC. 2007. Determination of carbon labeling distribution of intracellular metabolites from single fragment ions by ion chromatography tandem mass spectrometry. Anal Biochem 360:182-188.

Killhamt K, Firestone MK. 1984. Salt stress control of intracellular solutes in *Streptomycetes Indigenous* to saline soils. Appl Environ Microbiol 47:301-306.

Kim R, Kim KK, Yokota H, Kim SH. 1998. Small heat shock protein of *Methanococcus jannaschii*, a hyperthermophile. Proc Natl Acad Sci USA 95:9129-9133.

Kim S, Lee do Y, Wohlgemuth G, Park HS, Fiehn O, Kim KH. 2013. Evaluation and optimization of metabolome sample preparation methods for *Saccharomyces cerevisiae*. Anal Chem 85:2169-2176.

Klinkert B, Narberhaus F. 2009. Microbial thermosensors. Cell Mol Life Sci 66:2661-2676.

Koch R, Spreinat K, Lemke K, Antranikian G. 1991. Purification and properties of a hyperthermoactive α-amylase from the archaeobaterium *Pyrococcus woesei*. Arch Microbiol 155:572-578.

Köcher S, Averhoff B, Müller V. 2011. Development of a genetic system for the moderately halophilic bacterium *Halobacillus halophilus*: generation and characterization of mutants defect in the production of the compatible solute proline. Environ Microbiol 13:2122-2131.

Koga Y. 2012. Early evolution of membrane lipids: how did the lipid divide occur? J Mol Evol 72:274-282.

König K, Hartmann E, Kärcher U. 1994. Pathways and principles of the biosynthesis of methanobacterial cell wall polymers. Syst Appl Microbiol 16:510-517.

Koning SM, Elferink MG, Konings WN, Driessen AJ. 2001. Cellobiose uptake in the hyperthermophilic archaeon *Pyrococcus furiosus* is mediated by an inducible, high-affinity ABC transporter. J Bacteriol 183:4979-4984.

Konings WN, Albers SV, Koning S, Driessen AJ. 2002. The cell membrane plays a crucial role in survival of bacteria and archaea in extreme environments. Antonie Van Leeuwenhoek 81:61-72.

Koonin EV. 1996. A duplicated catalytic motif in a new superfamily of phosphohydrolases and phospholipid synthases that includes poxvirus envelope proteins. Trends Biochem Sci 21:242-243.

Kosuge T, Hoshino T. 1998. Lysine is synthesized through the α-aminoadipate pathway in *Thermus thermophilus*. FEMS Microbiol Lett 169:361-367.

Kouril T, Esser D, Kort J, Westerhoff HV, Siebers B, Snoep JL. 2013. Intermediate instability at high temperature leads to low pathway efficiency for an *in vitro* reconstituted system of gluconeogenesis in *Sulfolobus solfataricus*. FEBS J 280:4666-4680.

Krämer R. 2010. Bacterial stimulus perception and signal transduction: response to osmotic stress. Chem Rec 10:217-229.

Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774-797.

Lai D, Springstead JR, Monbouquette HG. 2008. Effect of growth temperature on ether lipid biochemistry in *Archaeoglobus fulgidus*. Extremophiles 12:271-278.

Lai MC, Hong TY, Gunsalus RP. 2000. Glycine betaine transport in the obligate halophilic archaeon *Methanohalophilus portucalensis*. J Bacteriol 182:5020-5024.

Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem 77:521–555.

Laksanalamai P, Maeder DL, Robb FT. 2001. Regulation and mechanism of action of the small heat shock protein from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 183:5198-5202.

Laksanalamai P, Whitehead TA, Robb FT. 2004a. Minimal protein-folding systems in hyperthermophilic archaea. Nat Rev Microbiol 2:315-324.

Laksanalamai P, Robb FT. 2004b. Small heat shock proteins from extremophiles: a review. Extremophiles 8:1-11.

Lam WL, Doolittle WF. 1989. Shuttle vectors for the archaebacterium *Halobacterium volcanii*. Proc Natl Acad Sci USA 86:5478-5482.

Lamosa P, Martins LO, da Costa MS, Santos H. 1998. Effects of temperature, salinity, and medium composition on compatible solute accumulation by *Thermococcus* spp. Appl Environ Microbiol 64:3591-3598.

Lamosa P, Gonçalves LG, Rodrigues MV, Martins LO, Raven ND, Santos H. 2006. Occurrence of 1-glyceryl-1-*myo*-inosityl phosphate in hyperthermophiles. Appl Environ Microbiol 72:6169-6173.

Lancaster WA, Praissman JL, Poole FL 2nd, Cvetkovic A, Menon AL, Scott JW, Jenney FE Jr, Thorgersen MP, Kalisiak E, Apon JV, Trauger SA, Siuzdak G, Tainer JA, Adams MW. 2011. A computational framework for proteome-wide pursuit and prediction of metalloproteins using ICP-MS and MS/MS data. BMC Bioinformatics 12:64.

Landau EM, Rosenbusch JP. 1996. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. Proc Natl Acad Sci USA 93:14532-14535.

Larson TJ, Dowhan W. 1976. Ribosomal-associated phosphatidylserine synthetase from *Escherichia coli*: purification by substrate-specific elution from phosphocellulose using cytidine 5'-diphospho-1,2-diacyl-*sn*-glycerol. Biochemistry 15:5212-5218.

le Maire M, Champeil P, Moller JV. 2000. Interaction of membrane proteins and lipids with solubilizing detergents. Biochim Biophys Acta 1508:86-111.

Le Rudulier D, Strom AR, Dandekar AM, Smith LT, Valentine RC. 1984. Molecular biology of osmoregulation. Science 224:1064-1068.

Lee HS, Shockley KR, Schut GJ, Conners SB, Montero CI, Johnson MR, Chou CJ, Bridger SL, Wigner N, Brehm SD, Jenney FE Jr, Comfort DA, Kelly RM, Adams MW. 2006. Transcriptional and biochemical analysis of starch metabolism in the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 188:2115-2125.

Lee S-J, Engelmann A, Horlacher R, Qu Q, Vierke G, Hebbeln C, Thomm M, Boos W. 2003. TrmB, a sugar-specific transcriptional regulator of the trehalose/maltose ABC transporter from the hyperthermophilic archaeon *Thermococcus litoralis*. J Biol Chem 278: 983-990.

Lehninger AE. 1975. Biochemistry, the molecular basis of cell structure and function. 2nd ed, Worth Publishers, New York.

Leigh JA, Albers SV, Atomi H, Allers T. 2011. Model organisms for genetics in the domain Archaea: methanogens, halophiles, *Thermococcales* and *Sulfolobales*. FEMS Microbiol Rev 35:577-608.

Leslie AG, Powell HR, Winter G, Svensson O, Spruce D, McSweeney S, Love D, Kinder S, Duke E, Nave C. 2002. Automation of the collection and processing of X-ray diffraction data – a generic approach. Acta Crystallogr D Biol Crystallogr 58:1924-1928.

Lipscomb GL, Stirrett K, Schut GJ, Yang F, Jenney FE, Scott R, Adams MW, Westpheling J. 2011. Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. Appl Environ Microbiol 77:2232-2238.

Liu W, Vierke G, Wenke AK, Thomm M, Ladenstein R. 2007. Crystal structure of the archaeal heat shock regulator from *Pyrococcus furiosus*: a molecular chimera representing eukaryal and bacterial features. J Mol Biol 369:474-488.

Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. 2012. OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res 40:D370–D376.

Luke KA, Higgins CL, Wittung-Stafshede P. 2007. Thermodynamic stability and folding of proteins from hyperthermophilic organisms. FEBS Journal 274:4023-4033.

Macario AJ, Conway de Macario E. 1999. The archaeal molecular chaperone machine: peculiarities and paradoxes. Genetics 152:1277-1283.

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV. 2011. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 9:467-477.

Martin DD, Ciulla RA, Roberts MF. 1999. Osmoadaptation in archaea. Appl Environ Microbiol 65:1815-1825.

Martins LO, Santos H. 1995. Accumulation of mannosylglycerate and di-*myo*-inositol-phosphate by *Pyrococcus furiosus* in response to salinity and temperature. Appl Environ Microbiol 61:3299-3303.

Martins LO, Huber R, Huber H, Stetter KO, da Costa MS, Santos H. 1997. Organic solutes in hyperthermophilic archaea. Appl Environ Microbiol 63:896-902.

Martins LO, Empadinhas N, Marugg JD, Miguel C, Ferreira C, da Costa MS, Santos H. 1999. Biosynthesis of mannosylglycerate in the thermophilic bacterium *Rhodothermus marinus*. Biochemical and genetic characterization of a mannosylglycerate synthase. J Biol Chem 274:35407-35414.

Mashego MR, Rumbold K, De Mey M, Vandamme E, Soetaert W, Heijnen JJ. 2007. Microbial metabolomics: past, present and future methodologies. Biotechnol Lett 29:1-16.

Matsumi R, Manabe K, Fukui T, Atomi H, Imanaka T. 2007. Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance. J Bacteriol 189:2683-2691.

Matsuno Y, Sugai A, Higashibata H, Fukuda W, Ueda K, Uda I, Sato I, Itoh T, Imanaka T, Fujiwara S. 2009. Effect of growth temperature and growth phase on the lipid composition of the archaeal membrane from *Thermococcus kodakaraensis*. Biosci Biotechnol Biochem 73:104-108.

Matte A, Tari LW, Delbaere LT. 1998. How do kinases transfer phosphoryl groups? Struct Folding Design 6:413-419.

Matthews BW. 1968. Solvent content of protein crystals. J Mol Biol 33:491-497.

McNicholas S, Potterton E, Wilson KS, Noble ME. 2011. Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr D Biol Crystallogr 67:386-394.

McTernan PM, Chandrayan SK, Wu CH, Vaccaro BJ, Lancaster WA, Yang Q, Fu D, Hura GL, Tainer JA, Adams MW. 2014. Intact functional fourteen-subunit respiratory membrane-bound [NiFe]-hydrogenase complex of the hyperthermophilic archaeon *Pyrococcus furiosus*. J Biol Chem 289:19364-19372.

Menon AL, Poole FL 2nd, Cvetkovic A, Trauger SA, Kalisiak E, Scott JW, Shanmukh S, Praissman J, Jenney FE Jr, Wikoff WR, Apon JV, Siuzdak G, Adams MW. 2009. Novel multiprotein complexes identified in the hyperthermophilic archaeon *Pyrococcus furiosus* by non-denaturing fractionation of the native proteome. Mol Cell Proteomics 8:735-751.

Meyer AS, Baker TA. 2011. Proteolysis in the *Escherichia coli* heat shock response: a player at many levels. Curr Opin Microbiol 14:194–199.

Mezey E, Dehejia A, Harta G, Papp MI, Polymeropoulos MH, Brownstein MJ. 1998. Alpha synuclein in neurodegenerative disorders: murderer or accomplice?. Nat Med 4: 755–757.

Mitsuzawa S, Deguchi S, Horikoshi K. 2006. Cell structure degradation in *Escherichia coli* and *Thermococcus* sp. strain Tc-1-95 associated with thermal death resulting from brief heat treatment. FEMS Microbiol Lett 260:100-105.

Mojica FJ, Cisneros E, Ferrer C, Rodríguez-Valera F, Juez G. 1997. Osmotically induced response in representatives of halophilic prokaryotes: the bacterium *Halomonas elongata* and the archaeon *Haloferax volcanii*. J Bacteriol 179:5471-5481.

Morano KA, Grant CM, Moye-Rowley WS. 2012. The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. Genetics 190:1157-1195.

Morbach S, Krämer R. 2004. Osmoregulation and osmosensing by uptake carriers for compatible solutes in bacteria. pp. 121-153. *In* E. Boles, R. Krämer (eds), Molecular Mechanisms Controlling Transmembrane Transport, Topics in Current Genetics, Berlin.

Morii H, Ogawa M, Fukuda K, Taniguchi H, Koga Y. 2010. A revised biosynthetic pathway for phosphatidylinositol in Mycobacteria. J Biochem 148:593-602.

Morii H, Okauchi T, Nomiya H, Ogawa M, Fukuda K, Taniguchi H. 2013. Studies of inositol 1-phosphate analogues as inhibitors of the phosphatidylinositol phosphate synthase in mycobacteria. J Biochem 153:257-266.

Mukund S, Adams MW. 1995. Glyceraldehyde-3-phosphate ferredoxin oxidoreductase, a novel tungsten-containing enzyme with a potential glycolytic role in the hyperthermophilic archaeon *Pyrococcus furiosus*. J Biol Chem 270:8389-8392.

Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53:240-255.

Narberhaus F. 2002. Alpha-crystallin-type heat shock proteins: socializing minichaperones in the context of a multichaperone network. Microbiol Mol Biol Rev 66:64-93.

Neves C, da Costa MS, Santos H. 2005. Compatible solutes of the hyperthermophile *Palaeococcus ferrophilus*: osmoadaptation and thermoadaptation in the order *Thermococcales*. Appl Environ Microbiol 71:8091-8098.

Nicholson JK, Lindon JC. 2008. Systems biology: Metabolomics. Nature 455:1054-1056.

Nishida H, Nishiyama M, Kobashi N, Kosuge T, Hoshino T, Yamane H. 1999. A prokaryotic gene cluster involved in synthesis of lysine through the amino adipate pathway: a key to the evolution of amino acid biosynthesis. Genome Res 9:1175-1183.

Nobre A, Empadinhas N, Nobre MF, Lourenço EC, Maycock C, Ventura MR, Mingote A, da Costa MS. 2013. The plant *Selaginella moellendorffii* possesses enzymes for

synthesis and hydrolysis of the compatible solutes mannosylglycerate and glucosylglycerate. Planta 237:891-901.

Noll KM, Lapierre P, Gogarten JP, Nanavati DM. 2008. Evolution of *mal* ABC transporter operons in the *Thermococcales* and *Thermotogales*. BMC Evol Biol 8:7.

Nonaka G, Blankschien M, Herman C, Gross CA, Rhodius VA. 2006. Regulon and promoter analysis of the *E. coli* heat-shock factor, σ^{32}, reveals a multifaceted cellular response to heat stress. Genes Dev 20:1776-1789.

Nunes OC, Manaia CM, da Costa MS, Santos H. 1995. Compatible solutes in the thermophilic bacteria *Rhodothermus marinus* and "*Thermus thermophilus*". Appl Environ Microbiol 61:2351-2357.

Oren A. 1999. Bioenergetic aspects of halophilism. Microbiol Mol Biol Rev 63:334-348.

Oren A, Mana L. 2002. Amino acid composition of bulk protein and salt relationships of selected enzymes of Salinibacter ruber, an extremely halophilic bacterium. Extremophiles 6:217-223.

Orita I, Sato T, Yurimoto H, Kato N, Atomi H, Imanaka T, Sakai Y. 2006. The ribulose monophosphate pathway substitutes for the missing pentose phosphate pathway in the archaeon *Thermococcus kodakaraensis*. J Bacteriol 188:4698-4704.

Ostermeier C, Michel H. 1997. Crystallization of membrane proteins. Curr Opin Struct Biol 7:697-701.

Palmer AG, Cavanagh J, Wright PE, Rance M. 1991. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. J Magn Reson 93:151-170.

Patti GJ, Yanes O, Siuzdak G. 2012. Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol 13:263-269.

Peck RF, Dassarma S, Krebs MP. 2000. Homologous gene knockout in the archaeon *Halobacterium salinarum* with ura3 as a counterselectable marker. Mol Microbiol 35:667-676.

Ponting CP, Kerr ID. 1996. A novel family of phospholipase D homologues that includes phospholipid synthases and putative endonucleases: Identification of duplicated repeats and potential active site residues. Protein Sci 5:914–922.

Pontoni G, Manna C, Salluzzo A, del Piano L, Galletti P, De Rosa M, Zappia V. 1985. Studies on enzyme-substrate interactions of cholinephosphotransferase from Rat-liver. Biochim Biophys Acta 836:222-232.

Poole FL 2nd, Gerwe BA, Hopkins RC, Schut GJ, Weinberg MV, Jenney FE Jr, Adams MW. 2005. Defining genes in the genome of the hyperthermophilic archaeon *Pyrococcus furiosus*: implications for all microbial genomes. J Bacteriol 187:7325-7332.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD. 2012. The Pfam protein families database. Nucleic Acids Res 40:D290–D301.

Pysz MA, Ward DE, Shockley KR, Montero CI, Conners SB, Johnson MR, Kelly RM. 2004. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium *Thermotoga maritima*. Extremophiles 8:209-217.

Qi QG, Huang YF, Cutler AJ, Abrams SR, Taylor DC. 2003. Molecular and biochemical characterization of an aminoalcoholphosphotransferase (AAPT1) from *Brassica napus*: effects of low temperature and abscisic acid treatments on AAPT expression in Arabidopsis plants and effects of overexpression of BnAAPT1 in transgenic Arabidopsis. Planta 217:547-558.

Raetz CR, Carman GM, Dowhan W, Jiang RT, Waszkuc W, Loffredo W, Tsai MD. 1987. Phospholipids chiral at phosphorus. Steric course of the reactions catalyzed by phosphatidylserine synthase from *Escherichia coli* and yeast. Biochemistry 26:4022-4027.

Ramakrishnan V, Verhagen MFJM, Adams MWW. 1997. Characterization of di-*myo*-inositol-1,1´-phosphate in the hyperthermophilic bacterium *Thermotoga maritima*. Appl Environ Microbiol. 63:347-350.

Ramos A, Raven NDH, Sharp RJ, Bartolucci S, Rossi M, Cannio R, Lebbink J, Van der Oost J, de Vos WM, Santos H. 1997. Stabilization of enzymes against thermal stress and freeze-drying by mannosylglycerate. Appl Environ Microbiol 63:4020-4025.

Raven N, Ladwa N, Cossar D, Sharp R. 1992. Continuous culture of the hyperthermophilic archaeum *Pyrococcus furiosus*. Appl Microbiol Biotechnol 38:263-267.

Richmond CS, Glasner JD, Mau R, Jin H, Blattner FR. 1999. Genome-wide expression profiling in *Escherichia coli* K-12. Nucleic Acids Res 27:3821-3835.

Richter K, Haslbeck M, Buchner J. 2010. The heat shock response: life on the verge of death. Mol Cell 40:253-266.

Ritossa F. 1962. A new puffing pattern induced by temperature shock and DNP in *Drosophila*. Experientia 18:571-573.

Robb FT, Maeder DL, Brown JR, DiRuggiero J, Stump MD, Yeh RK, Weiss RB, Dunn DM. 2001. Genomic sequence of hyperthermophile, *Pyrococcus furiosus*: implications for physiology and enzymology. Methods Enzymol 330:134-157.

Robinson KA, Robb FT, Schreier HJ. 1994. Isolation of maltose-regulated genes from the hyperthermophilic archaeum, *Pyrococcus furiosus*, by subtractive hybridization. Gene 148:137-141.

Robinson KA, Schreier HJ. 1994. Isolation, sequence and characterization of the maltose-regulated mlrA gene from the hyperthermophilic archaeum *Pyrococcus furiosus*. Gene 151:173-176.

Rodionov DA, Kurnasov OV, Stec B, Wang Y, Roberts MF, Osterman AL. 2007. Genomic identification and *in vitro* reconstitution of a complete biosynthetic pathway for the osmolyte di-*myo*-inositol-phosphate. Proc Natl Acad Sci USA 104:4279-4284.

Rodrigues MV, Borges N, Henriques M, Lamosa P, Ventura R, Fernandes C, Empadinhas N, Maycock C, da Costa MS, Santos H. 2007. Bifunctional CTP: Inositol-1-phosphate cytidylyltransferase/CDP-inositol: Inositol-1-phosphate transferase, the key enzyme for di-*myo*-inositol-phosphate synthesis in several (hyper)thermophiles. J Bacteriol 189:5405-5412.

Rodrigues MV, Borges N, Almeida CP, Lamosa P, Santos H. 2009. A unique β-1,2-mannosyltransferase of *Thermotoga maritima* that uses di-*myo*-inositol phosphate as the mannosyl acceptor. J Bacteriol 191:6105-6115.

Rodrigues MV. 2011. Heat stress adaptation in hyperthermophiles: Biosynthesis of inositol-containing compatible solutes. Ph.D. Thesis, Instituto de Tecnologia Química e Biológica, Portugal.

Roeßler M, Müller V. 2001. Osmoadaptation in bacteria and archaea: common principles and differences. Environ Microbiol 3:743-754.

Rohlin L, Trent JD, Salmon K, Kim U, Gunsalus RP, Liao JC. 2005. Heat shock response of *Archaeoglobus fulgidus*. J Bacteriol 187:6046-6057.

Rothschild LJ, Mancinelli RL. 2001. Life in extreme environments. Nature 409:1092-1101.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual Cold Spring Harbor, N.Y.: Cold Spring Harbor laboratory Press.

Sandusky P, Appiah-Amponsah E, Raftery D. 2011. Use of optimized 1D TOCSY NMR for improved quantitation and metabolomic analysis of biofluids. J Biomol NMR 49:281–290.

Santos H, da Costa MS. 2001. Organic solutes from thermophiles and hyperthermophiles. Methods Enzymol 334:302-315.

Santos H, da Costa MS. 2002. Compatible solutes of organisms that live in hot saline environments. Environ Microbiol 4:501-509.

Santos H, Lamosa P, Borges N. 2006. Characterization and quantification of compatible solutes in (hyper)thermophilic microorganisms, pp. 171-197. In Oren A, Rainey F (eds), Methods in Microbiology: Extremophiles. Elsevier, Amsterdam.

Santos H, Lamosa P, Borges N, Gonçalves LG, Pais T, Rodrigues MV. 2011. Organic compatible solutes of prokaryotes that thrive in hot environments: The importance of ionic compounds for thermostabilization, pp. 497-520. In Horikoshi K, Antranikian G, Bull AT, Robb FT, Stetter KO (eds), Extremophiles Handbook. Springer, Tokyo.

Sapra R, Bagramyan K, Adams MW. 2003. A simple energy-conserving system: proton reduction coupled to proton translocation. Proc Natl Acad Sci USA 100:7545-7550.

Sapra R, Verhagen MF, Adams MW. 2000. Purification and characterization of a membrane-bound hydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 182:3423-8.

Sato T, Fukui T, Atomi H, Imanaka T. 2003. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J Bacteriol 185:210-220.

Saum R, Mingote A, Santos H, Müller V. 2009. A novel limb in the osmoregulatory network of *Methanosarcina mazei* Gö1: N^{ε}-acetyl-β-lysine can be substituted by glutamate an alanine. Environ Microbiol 11:1056-1065.

Schäfer G. 1992. Extremophiles: Fascinating organisms with surprising capabilities. J Bioenerg Biomembr 24:525-527.

Schäfer T, Schönheit P. 1992. Maltose fermentation to acetate, CO_2 and H_2 in the anaerobic hyperthermophilic archaeon *Pyrococcus furiosus*: evidence for the operation of a novel sugar fermentation pathway. Arch Microbiol 158:188-202.

Schäfer T, Xavier KB, Santos H, Schönheit P. 1994. Glucose fermentation to acetate and alanine in resting cell suspensions of *Pyrococcus furiosus*: Proposal of a novel glycoltic pathway based on ^{13}C labelling data and enzyme activates. FEMS Microbiol Lett 121:107–114.

Schlame M, Zhao M, Rua D, Haldar D, Greenberg ML. 1995. Kinetic analysis of cardiolipin synthase: a membrane enzyme with two glycerophospholipid substrates. Lipids 30:633-640.

Scholz S, Sonnenbichler J, Schäfer W, Hensel R. 1992. Di-*myo*-inositol-1,1'-phosphate: a new inositol phosphate isolated from *Pyrococcus woesei*. FEBS Lett 306:239-242.

Scholz S, Wolff S, Hensel R. 1998. The biosynthetic pathway of di-*myo*-inositol-1,1'-phosphate in *Pyrococcus woesei*. FEMS Microbiology Letters 168:37-42.

Schönheit P, Schäfer T. 1995. Metabolism of hyperthermophiles. World J Microbiol Biotechnol 11:26-57.

Schut GJ, Brehm SD, Datta S, Adams MW. 2003. Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: *Pyrococcus furiosus* grown on carbohydrates or peptides. J Bacteriol 185:3935-3947.

Schut GJ, Boyd ES, Peters JW, Adams MW. 2013. The modular respiratory complexes involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their evolutionary implications. FEMS Microbiol Rev 37:182-203.

Sciara G, Clarke OB, Tomasek D, Kloss B, Tabuso S, Byfield R, Cohn R, Banerjee S, Rajashankar KR, Slavkovic V, Graziano JH, Shapiro L, Mancia F. 2014. Structural basis for catalysis in a CDP-alcohol phosphotransferase. Nat Commun 5:4068.

Scott JW, Poole FL, Adams MW. 2014. Characterization of ten heterotetrameric NDP-dependent acyl-CoA synthetases of the hyperthermophilic archaeon *Pyrococcus furiosus*. Archaea. doi: 10.1155/2014/176863.

Servant P, Mazodier P. 2001. Negative regulation of the heat shock response in *Streptomyces*. Arch. Microbiol. 176:237-242.

Sévin DC, Sauer W. 2014. Ubiquinone accumulation improves osmotic-stress tolerance in *Escherichia coli*. Nat Chem Biol 10:266-272.

Shaka AJ, Lee CJ, Pines A. 1988. Iterative schemes for bilinear operators; application to spin decoupling. J Magn Reson 77:274-293.

Shima S, Hérault DA, Berkessel A, Thauer RK. 1998. Activation and thermostabilization effects of cyclic 2,3-diphosphoglycerate on enzymes from the hyperthermophilic *Methanopyrus kandleri*. Arch Microbiol 170:469-472.

Shockley KR, Ward DE, Chhabra SR, Conners SB, Montero CI, Kelly RM. 2003. Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus.* Appl Environ Microbiol 69:2365-2371.

Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. 2013. New and continuing developments at PROSITE. Nucleic Acids Res 41:D344–D347.

Silva Z, Borges N, Martins LO, Wait R, da Costa MS, Santos H. 1999. Combined effect of the growth temperature and salinity of the medium on the accumulation of compatible solutes by *Rhodothermus marinus* and *Rhodothermus obamensis*. Extremophiles 3:163-172.

Smith LT, Smith GM, Madkour MA. 1990. Osmoregulation in *Agrobacterium tumefaciens*: accumulation of a novel disaccharide is controlled by osmotic strength and glycine betaine. J Bacteriol 172:6849-6855.

Soderberg T. 2005. Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes. Archaea 1:347-352.

Solis-Oviedo RL, Martinez-Morales F, Geiger O, Sohlenkamp C. 2012. Functional and topological analysis of phosphatidylcholine synthase from *Sinorhizobium meliloti*. Biochim Biophys Acta 1821:573-581.

Sorger PK, Pelham HR. 1987. Purification and characterization of a heat-shock element binding protein from yeast. EMBO J 6:3035-3041.

Stetter KO, Thomm M, Winter J, Wildgruber G, Huber H, Zillig W, Jané-Covic D, König H, Palm P, Wunderl S. 1981. *Methanothermus fervidus*, sp. nov., a novel extremely thermophilic methanogen isolated from an Icelandic hot springs. Syst Appl Microbiol 2:166-178.

Stetter KO. 1999. Extremophiles and their adaptation to hot environments. FEBS Letters 452:22-25.

Stetter KO. 2006. History of the discovery of the first hyperthermophiles. Extremophiles 10:357-362.

Stetter KO. 2013. A brief history of the discovery of hyperthermophilic life. Biochem Soc Trans 41:416-420.

Stevenson A, Burkhardt J, Cockell CS, Cray JA, Dijksterhuis J, Fox-Powell M, Kee TP, Kminek G, McGenity TJ, Timmis KN, Timson DJ, Voytek MA, Westall F, Yakimov MM, Hallsworth JE. 2014. Multiplication of microbes below 0.690 water activity: implications for terrestrial and extraterrestrial life. Environ Microbiol doi: 10.1111/1462-2920.12598.

Strand KR, Sun C, Li T, Jenney FE Jr, Schut GJ, Adams MW. 2010. Oxidative stress protection and the repair response to hydrogen peroxide in the hyperthermophilic archaeon *Pyrococcus furiosus* and in related species. Arch Microbiol 192:447-459.

Szilágyi A, Závodszky P. 2000. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. Structure 8:493-504.

Tabuchi Y, Takasaki I, Wada S, Zhao QL, Hori T, Nomura T, Ohtsuka K, Kondo T. 2008. Genes and genetic networks responsive to mild hyperthermia in human lymphoma U937 cells. Int J Hyperthermia 24:613-622.

Tachdjian S, Kelly RM. 2006. Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus*. J Bacteriol 188:4553-4559.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725-2729.

Tang J. 2011. Microbial metabolomics. Curr Genomics 12:391-403.

Taymaz-Nikerel H, de Mey M, Ras C, ten Pierick A, Seifar RM, van Dam JC, Heijnen JJ, van Gulik WM. 2009. Development and application of a differential method for reliable metabolome analysis in *Escherichia coli*. Anal Biochem 386:9-19.

Thorgersen MP, Lipscomb GL, Schut GJ, Kelly RM, Adams MW. 2014. Deletion of acetyl-CoA synthetases I and II increases production of 3-hydroxypropionate by the metabolically-engineered hyperthermophile *Pyrococcus furiosus*. Metab Eng 22:83-88.

Trent JD. 1996. A review of acquired thermotolerance, heat-shock proteins, and molecular chaperones in archaea. FEMS Microbiol Rev 18:249-258.

Usui K, Yoshida T, Maruyama T, Yohda M. 2001. Small heat shock protein of a hyperthermophilic archaeon, *Thermococcus* sp. strain KS-1, exists as a spherical 24 mer and its expression is highly induced under heat-stress conditions. J Biosci Bioeng 92:161-166.

Usui M, Sembongi H, Matsuzaki H, Matsumoto K, Shibuya I. 1994. Primary structures of the wild-type and mutant alleles encoding the phosphatidylglycerophosphate synthase of *Escherichia coli*. J Bacteriol 176:3389-3392.

Vagin A, Teplyakov A. 1997. MOLREP: an automated program for molecular replacement. J Appl Crystallogr 30:1022-1025.

van der Oost J, Schut G, Kengen SW, Hagen WR, Thomm M, de Vos WM. 1998. The ferredoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus* represents a novel site of glycolytic regulation. J Biol Chem 273:28149-28154.

van Gulik WM. 2010. Fast sampling for quantitative microbial metabolomics. Curr Opin Biotechnol. 21:27-34.

Vieille C, Zeikus JG. 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1-43.

Vierke G, Engelmann A, Hebbeln C, Thomm M. 2003. A novel archaeal transcriptional regulator of heat shock response. J Biol Chem 278:18-26.

Wang L, Quan C, Liu B, Wang J, Xiong W, Zhao P, Fan S. 2013. Functional reconstitution of *Staphylococcus aureus* truncated AgrC histidine kinase in a model membrane system. PLoS ONE 8:e80400.

Weinberg MV, Schut GJ, Brehm S, Datta S, Adams MW. 2005. Cold shock of a hyperthermophilic archaeon: *Pyrococcus furiosus* exhibits multiple responses to a suboptimal growth temperature with a key role for membrane-bound glycoproteins. J Bacteriol 187:336-348.

Welker S, Rudolph B, Frenzel E, Hagn F, Liebisch G, Schmitz G, Scheuring J, Kerth A, Blume A, Weinkauf S, Haslbeck M, Kessler H, Buchner J. 2010. Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. Mol Cell 39:507-520.

Williams JG, McMaster CR. 1998. Scanning alanine mutagenesis of the CDP-alcohol phosphotransferase motif of *Saccharomyces cerevisiae* cholinephosphotransferase. J Biol Chem 273:13482-13487.

Winder CL, Dunn WB, Schuler S, Broadhurst D, Jarvis R, Stephens GM, Goodacre R. 2008. Global metabolic profiling of *Escherichia coli* cultures: an evaluation of methods for quenching and extraction of intracellular metabolites. Anal Chem 80:2939-2948.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67:235-242.

Wishart DS. 2008. Quantitative metabolomics using NMR. Trends Anal Chem 27:228-237.

Wittmann C, Krömer JO, Kiefer P, Binz T, Heinzle E. 2004. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. Anal Biochem 327:135-139.

Wu C. 1984. Activating protein factor binds *in vitro* to upstream control sequences in heat shock gene chromatin. Nature 311:81-84.

Wu L, Mashego MR, van Dam JC, Proll AM, Winke JL. 2005. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ^{13}C-labeled cell extracts as internal standards. Anal Biochem 336:164–171.

Wu X, Li N, Li H, Tang H. 2014. An optimized method for NMR-based plant seed metabolomic analysis with maximized polar metabolite extraction efficiency, signal-to-noise ratio, and chemical shift consistency. Analyst 139:1769-1778.

Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. 1982. Living with water stress: evolution of osmolyte systems. Science 217:1214-1222.

Ye Y, Zhang L, Hao F, Zhang J, Wang Y, Tang H. 2012. Global metabolomic responses of *Escherichia coli* to heat stress. J Proteome Res 11:2559-2566.

Young JC, Agashe VR, Siegers K, Hartl FU. 2004. Pathways of chaperone-mediated protein folding in the cytosol. Nat Rev Mol Cell Biol 5:781–791.

Yu JP, Ladapo J, Whitman WB. 1994. Pathway of glycogen metabolism in *Methanococcus maripaludis* . J Bacteriol 176:325-332.

Zaparty M, Esser D, Gertig S, Haferkamp P, Kouril T, Manica A, Pham TK, Reimann J, Schreiber K, Sierocinski P, Teichmann D, van Wolferen M, von Jan M, Wieloch P, Albers SV, Driessen AJ, Klenk HP, Schleper C, Schomburg D, van der Oost J, Wright PC, Siebers B. 2010. "Hot standards" for the thermoacidophilic archaeon *Sulfolobus solfataricus.* Extremophiles 14:119-142.

Zivanovic Y, Lopez P, Philippe H, Forterre P. 2002. *Pyrococcus* genome comparison evidences chromosome shuffling-driven evolution. Nucleic Acids Res 30:1902-1910.

Zuckerkandl E, Pauling L. 1965. Molecules as documents of evolutionary history. J Theor Biol 8:357-366.

Appendix

Unraveling the transcription factors involved in the synthesis of di-*myo*-inositol-phosphate in Crenarchaeota and Bacteria

Contents

Summary

In contrast to the extensive knowledge on the biosynthesis of di-*myo*-inositol-phosphate (DIP), little is known about its regulation. To date, only the transcription factor (designated as Phr) that is involved in the regulation of DIP synthesis has been identified in the euryarchaeon *Thermococcus kodakarensis*. However, no homologues were found in members of the Crenarchaeota kingdom of Archaea and Bacteria. To hunt for protein(s) involved in the regulation of DIP synthesis, two different methodologies were followed: 1) the bioinformatics strategy, and 2) the biotin-streptavidin strategy. In the first strategy, BLAST searches in NCBI database were carried out using known heat shock transcription factors as query. In the second strategy, the promoter region of the gene encoding the key-enzyme involved in DIP synthesis was labeled with biotin and mixed with cell extracts of the hyperthermophilic crenarchaeon *Hyperthermus butylicus*. Subsequently, the promoter-transcription factor complexes were purified using streptavidin-linked magnetic beads. Proteins were separated by SDS-PAGE and the promising candidates were extracted and identified by MALDI-TOF analysis. The same approach was carried out using cell extracts of the hyperthermophilic bacterium *Thermotoga maritima.* Two candidates were identified in *H. butylicus*, but none was found in *T. maritima*.

Introduction

Hyperthermophiles possess the remarkable ability to thrive at temperatures near 100ºC. Therefore, it is appealing to investigate the molecular mechanisms underlying this extraordinary trait. The accumulation of compatible solutes has been assumed as a thermoadaptation strategy, and di-*myo*-inositol-phosphate (DIP) is a typical solute accumulated in response to heat stress in these organisms (Santos et al. 2011). The biosynthetic pathway of this compatible solute has been established (Rodrigues et al. 2007, Rodionov et al. 2007). The transcript level of the gene encoding the key-enzyme involved in DIP synthesis (IPCT/DIPPS) increased upon heat stress in *Thermococcus kodakarensis* (Borges et al. 2010). Moreover, the transcription regulator (designated as Phr) implicated in heat shock regulation has been identified in this hyperthermophile (Kanai et al. 2010). Under optimal growth conditions, the Phr binds to the promoter region of several heat shock genes (including the *ipct/dipps* gene), blocking their transcriptions. Under heat stress, the regulator is released and the transcriptions of heat shock genes are activated (Kanai et al. 2010). Therefore, this is a negative regulatory mechanism. This transcription regulator was initially found in the euryarchaeon *Pyrococcus furiosus*, but the induction of the *ipct/dipps* gene was not investigated (Vierke et al. 2003). The three-dimensional structure of the Phr from *P. furiosus* has been elucidated and it is a stable homodimer. Each subunit contains an N-terminal winged helix DNA-binding domain (Helix-Turn-Helix motif, HTH) and a C-terminal antiparallel coiled coil helical domain (Liu et

al. 2007). More recently, the sequence TTT(A/T)ntn(A/T)(A/C)nnnnnGTnAnn(A/T)AAa was identified as the consensus binding sequence for Phr (Keese et al. 2010). In the hyperthermophilic euryarchaeon *Archaeoglobus fulgidus*, a different heat shock transcription factor (designated as HSR1) was identified and it binds to the promoter region of an operon composed by three genes encoding HSR1 itself, AAA^+ ATPase, and small heat shock protein (small HSP), and also to a second small HSP (Rohlin et al. 2005). These authors did not investigate the transcription regulation of the *ipct/dipps* gene by HSR1. Thus far, the transcriptional regulatory mechanism of the *ipct/dipps* gene remains obscure in members of Crenarchaeota and Bacteria. Other heat shock transcription factors have been described in mesophilic Bacteria and Eukarya. Positive regulation has been described in *E. coli* and in most other Gram-negative bacteria (Gross 1996). In these bacteria, the sigma factor (σ^{32}) regulates the transcription of heat shock proteins (Gross 1996). Under optimal growth conditions, this factor binds to DnaK (HSP70) and DnaJ (HSP40), inactivating the activity of these chaperones. After a heat shock, the level of denatured proteins increases remarkably, and as a consequence, the denatured proteins bind to the DnaK and DnaJ, releasing the σ^{32}. Subsequently, the free sigma factor acts as a transcription activator of heat shock genes (Gross 1996). Negative regulation of the heat shock response has been reported in *Bacillus subtilis* and several other Gram-positive and Gram-negative bacteria. Under optimal growth conditions, the transcription repressor HrcA interacts with

the chaperone GroEL and this complex binds to the CIRCE (Controlling Inverted Repeat of Chaperone Expression) element in the promoter region of GroEL, repressing its expression. During the heat shock, the chaperones become engaged in the refolding of unfolded proteins and the HrcA dissociates from the CIRCE elements, allowing the expression of the heat shock genes (Narberhaus, 1999). A somewhat different mechanism was found in *Streptomyces* spp.; in this case, the transcriptional repressor HspR modulates the transcription of the dnaK operon by binding to three partially related inverted repeats (HspR-associated inverted repeats) in the promoter region (Bucca et al. 1995).

The aim of this work was to identify transcription factor(s) involved in the induction of the *ipct*/*dipps* gene in Crenarchaeota and Bacteria. The crenarchaeon *Hyperthermus butylicus* and the bacterium *Thermotoga maritima* were selected as target organisms. To identify the putative transcription regulator(s) in the target organisms, bioinformatics and biotin-streptavidin strategies were carried-out.

Materials and methods

Organisms and growth conditions

The type strain of *H. butylicus* (strain DSM 5456) was obtained from the German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). The organism was grown using a previously described medium supplemented with 1.7% NaCl (Zillig et al. 1990). Cultures were grown in 500 ml static vessels containing 100 ml of medium, at 95ºC (optimal growth

conditions) and 103ºC (heat stress). The cells were harvested by centrifugation and washed twice with an isotonic solution. The cell pellet was re-suspended in 2 ml lysis buffer (50 mM HEPES pH 7.5 containing 100 mM KCl, 1 mM EDTA, 5% glycerol, 0.5 mM DTT, and 0.1% triton X-100) and disrupted by sonication. Cell debris was removed by centrifugation (10,000 rpm, 10 min, 4°C) and the protein content was estimated by the Bradford method (Bradford 1976). *Thermococcus kodakarensis* strain KUW1 was grown at 83°C (optimal growth conditions) and 93°C (heat stress) in 500 ml static vessels, containing 100 ml of a nutrient-rich medium supplemented with sodium pyruvate (0.5%, w/v) (Borges et al. 2010). *Thermotoga maritima* strain DSM 3109 was cultivated at 80°C (optimal growth conditions) and 90°C (heat stress) as described previously (Rodrigues et al. 2009). The cells were harvested by centrifugation and washed twice with an isotonic solution. Crude cell extracts were prepared as described above.

Amplification of the promoter regions of target genes

Promoter regions (200-bp) of the *ipct/dipps* genes were amplified from the *H. butylicus* and *T. kodakarensis* genomes by PCR using the Pfu polymerase (Thermo Scientific). In *T. maritima*, the *ipct* and *dipps* genes are organized in an operon-like structure with the *mips* gene (L-*myo*-inositol-1-phosphate synthase) (Rodionov et al. 2007). The promoter region (200-bp) of this operon was also amplified by PCR from the *T. maritima* genome. The primers used in this study are listed in Table 1. All

the forward primers were 5'-end labeled with biotin (Stab Vida, Costa da Caparica, Portugal).

TABLE 1 List of primers used for the amplification of the promoter regions of the *ipct/dipps* genes from *H. butylicus*, *T. maritima*, and *T. kodakarensis*.

Designation	Primer (5'-3')
P_F_RPromIPCT_H_But	CCGACACTGAGATAGGCCTGGCC
P_R_RPromIPCT_H_But	TGGCCATCACCGCCAGCGGCCTACCAGCAACCCG
P_F_RPromMIPS_Tmar	TTCAAATCGTGAACCTTTTCACG
P_R_RPromMIPS_ Tmar	AAGTCCGGCAACGAATGTACTGGC
P_F_RPromIPCT_ Tkod	CTCAAACCCCTGCTTCAGCG
P_R_RPromIPCT_ Tkod	GGCCATCACCTTTTTAACGG

BLAST searches

Several transcription regulators (σ32, HrcA, and HspR found in mesophilic bacteria, and Phr and HSR1 encountered in hyperthermophilic euryarchaea) were used as query for BLAST searches against the genomes of organisms accumulating DIP, belonging to the Bacteria domain (*Thermotoga maritima*) and to the Crenarchaeota kingdom (*Hyperthermus butylicus* and *Pyrolobus fumarii*).

Biotin-Streptavidin strategy to pull-down putative transcription regulator(s)

For the pull-down assay, the biotinylated *ipct/dipps* promoter (3 μg) was mixed with cell extracts of *H. butylicus* (350 μg) grown at 95ºC (optimal growth condition) and 103ºC (heat stress). These mixtures were performed

in 1.5 ml Eppendorf tubes and each mixture was incubated at 85ºC during 30 minutes. Then, magnetic Dynabeads® M-280 Streptavidin (Invitrogen) were added and incubated at 55ºC for 30 minutes. After incubation, the tubes containing the reaction mixtures were placed into the magnetic separation rack (Invitrogen) for approximately 1 minute and the supernatant was removed. Unbound proteins were removed by three successive washes with 50 mM EPPS (pH 7.5) containing 100 mM KCl, 1 mM EDTA, 5% glycerol, 0.1% triton X-100, and 1 mM of DTT. Elution of DNA-bound proteins was performed at 95ºC for 5 minutes with 1×Laemmli buffer without β-mercaptoethanol. Eluted proteins were separated by Tricine-SDS-PAGE (13%) and silver stained. Only the bands showing molecular weights (MW) below 50 kDa were removed from the gel, since the proteins containing helix-turn-helix motifs (i.e. putative transcription regulators) in *H. butylicus* are within this range of MW. Protein identification was performed by MALDI-TOF (Mass spectrometry facility at ITQB). The same approach was carried out using cell extracts of the bacterium *T. maritima* and the *T. kodakarensis*, and their corresponding promoter regions. The pull-down assay with cell extracts of *T. kodakarensis* was performed as a positive control, since it was demonstrated that the Phr binds to the promoter region of the *ipct*/*dipps* gene (Kanai et al. 2010).

Results

Bioinformatics approach

The known heat shock transcription regulators (Phr, HSR1, HrcA, σ32, and HspR) were used as query for a systematic BLAST search within the genomes of *T. maritima*, *H. butylicus* and *Pyrolobus fumarii*. From this analysis, no clear candidate gene was identified, using HrcA, σ32, and HspR as query. These results are in agreement with the literature data (Liu et al. 2007); in fact, it was suggested that members of the Archaea domain may have developed different regulatory machinery from that found in the Bacteria and Eukarya domains. Moreover, no homologues of Phr were found in the genomes of the target hyperthermophiles. When the HSR1 from *A. fulgidus* was used as template, the Hbut_0073 gene was retrieved. This gene encodes a protein with a helix-turn-helix motif and has homologues in the genomes of two crenarchaeon (*P. fumarii* and *Aeropyrum pernix*) that possess the genes encoding IPCT and DIPPS. The Hbut_0073 showed 62% (e-value of 1e-32) and 46% (e-value of 6e-17) identity with the proteins of *P. fumarii* (Pyrfu_0798) and *A. pernix* (APE_1602a), respectively. Therefore, this protein is a promising candidate for the regulation of DIP synthesis.

The biotin-streptavidin strategy

The biotin-streptavidin strategy was used to hunt for the transcription regulator(s) involved in regulation of DIP synthesis. Briefly, the promoter

region of the gene encoding IPCT/DIPPS was labeled with biotin, and used as bait for pull-down the transcription regulator present in the cell extracts. Subsequently, the promoter-transcription regulator complex(es) were purified with streptavidin-linked magnetic beads and fractionated in a SDS-PAGE. This strategy was carried out using cell extracts derived from *H. butylicus* and *T. maritima* cells cultivated under optimal growth and heat stress conditions to cover the two possible types of regulation: negative and positive.

To validate the method, the pull-down assay was first performed with cell extracts of *T. kodakarensis* and a band corresponding to a protein with a molecular weight of approximately 24 kDa (the MW of Phr is around 23.4 kDa) was obtained (data not shown). Unfortunately, attempts to identify the protein by MALDI-TOF were unsuccessful. The pull-down assays were then carried out with cell extracts of *H. butylicus* and *T. maritima*. No results were obtained with *T. maritima*, but several proteins were pulled-down in the case of *H. butylicus* (Fig. 1). Based on the prediction that transcription regulators in this organism have MWs below 50 kDa, only the bands within this range were selected for MALDI-TOF analysis.

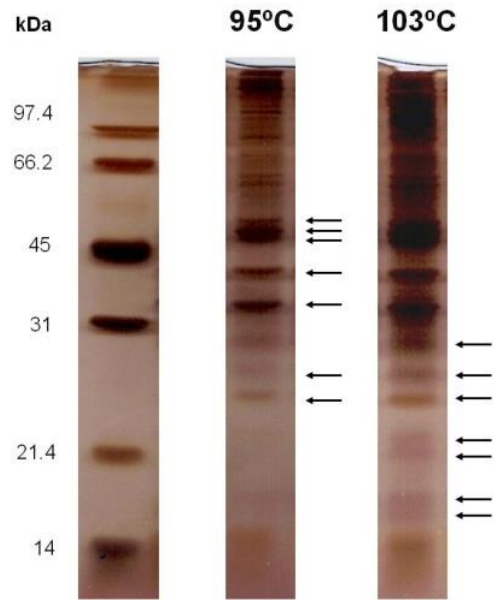


Fig. 1 Silver staining denaturating gel of eluted proteins from the biotin-streptavidin strategy. *Hyperthermus butylicus* was grown at 95ºC (optimal growth temperature) and 103ºC (heat stress). Arrows indicate the bands that were selected for MALDI-TOF analysis.

Of the 17 proteins analyzed by MALDI-TOF, only Hbut_0415 possesses the helix-turn-helix motif characteristic of the transcription factors. Therefore, the gene encoding this protein was selected as a good candidate to encode the transcription regulator involved in DIP regulation.

Table 2 List of enzymes retrieved from the pull-down assay using cell extracts of *H. butylicus*.

Gene	MW (kDa)	Function	Domains	Observations
Hbut_0644	38	Hypothetical Protein	DevR family	CRISPR system
Hbut_1126	31.3	Hypothetical Protein	-	Homology with *Ig. Hospitalis*
Hbut_0470	27.9	Hypothetical Protein	PAC2 superfamily	proteasome assembly chaperone
Hbut_0228	26.4	Probable peroxiredoxin		
Hbut_0415	45.2	Transcriptional regulator	GntR family	
Hbut_0495	37.6	NADPH glutamate synthase		
Hbut_0643	29.8	Hypothetical Protein	Cas5_I superfamily	CRISPR system
Hbut_1356	23	Hypothetical Protein		
Hbut_0049	15	Hypothetical Protein		
Hbut_1161	14.6	Superoxide reductase		
Hbut_1156	17.2	Hypothetical Protein	Byotinilated domain	Byotinilated domain
Hbut_1128	7	Hypothetical Protein	Cren7 protein	DNA packaging

Conclusions

In this work, two strategies were used to identify the transcription regulator(s) implicated in the synthesis of DIP in members of the Crenarchaeota kingdom and Bacteria domain. From these approaches, two promising candidates were obtained. The HBut_0073 gene was selected based on the bioinformatics analysis, while the HBut_0415 gene was identified by the pull-down strategy. The involvement of these putative transcription regulators in the synthesis of DIP must be validated in future work. The putative transcriptional regulators will be produced in *E. coli* and the binding of these proteins to the promoter region of *ipct/dipps* confirmed using the Biacore technique.

References

Borges N, Matsumi R, Imanaka T, Atomi H, Santos H. 2010. *Thermococcus kodakarensis* mutants deficient in di-*myo*-inositol phosphate use aspartate to cope with heat stress. J Bacteriol 192:191-197.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254.

Bucca G, Farina G, Puglia AM, Smith CP. 1995. The dnaK operon of *Streptomyces coelicolor* encodes a novel heat shock protein, which binds to the promoter region of the operon. Mol Microbiol. 17:663–674.

Gross CA. 1996. Function and regulation of the heat shock proteins. pp. 1382-1399. *In* FC Neidhardt, R Curtiss III, JL Ingraham, ECC Lin, KB Low, B Magasanik, WS Reznikoff, M Riley, M Schaechter and HE Umbarger (eds), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, DC,

Kanai T, Takedomi S, Fujiwara S, Atomi H, Imanaka T. 2010. Identification of the Phr-dependent heat shock regulon in the hyperthermophilic archaeon, *Thermococcus kodakaraensis*. J Biochem 147:361-370.

Keese AM, Schut GJ, Ouhammouch M, Adams MW, Thomm M. 2010. Genome-wide identification of targets for the archaeal heat shock regulator Phr by cell-free transcription of genomic DNA. J Bacteriol 192:1292-1298.

Liu W, Vierke G, Wenke AK, Thomm M, Ladenstein R. 2007. Crystal structure of the archaeal heat shock regulator from *Pyrococcus furiosus*: a molecular chimera representing eukaryal and bacterial features. J Mol Biol 369:474-488.

Narberhaus, F. 1999. Negative regulation of bacterial heat shock genes. Mol Microbiol. 31:1–8.

Rodionov DA, Kurnasov OV, Stec B, Wang Y, Roberts MF, Osterman AL. 2007. Genomic identification and *in vitro* reconstitution of a complete biosynthetic pathway for the osmolyte di-*myo*-inositol-phosphate. Proc Natl Acad Sci USA 104:4279-4284.

Rodrigues MV, Borges N, Henriques M, Lamosa P, Ventura R, Fernandes C, Empadinhas N, Maycock C, da Costa MS, Santos H. 2007. Bifunctional CTP: Inositol-1-phosphate cytidylyltransferase/CDP-inositol: Inositol-1-phosphate transferase, the key enzyme for di-*myo*-inositol-phosphate synthesis in several (hyper)thermophiles. J Bacteriol 189:5405-5412.

Rodrigues MV, Borges N, Almeida CP, Lamosa P, Santos H. 2009. A unique β-1,2-mannosyltransferase of *Thermotoga maritima* that uses di-*myo*-inositol phosphate as the mannosyl acceptor. J Bacteriol 191:6105-6115.

Rohlin L, Trent JD, Salmon K, Kim U, Gunsalus RP, Liao JC. 2005. Heat shock response of *Archaeoglobus fulgidus*. J Bacteriol 187:6046-6057.

Santos H, Lamosa P, Borges N, Gonçalves LG, Pais T, Rodrigues MV. 2011. Organic compatible solutes of prokaryotes that thrive in hot environments: The importance of ionic compounds for thermostabilization, pp. 497-520. In Horikoshi K, Antranikian G, Bull AT, Robb FT, Stetter KO (eds), Extremophiles Handbook. Springer, Tokyo.

Vierke G, Engelmann A, Hebbeln C, Thomm M. 2003. A novel archaeal transcriptional regulator of heat shock response. J Biol Chem 278:18-26.

Zillig W, Holz I, Janekovic D, Klenk HP, Imsel E, Trent J, Wunderl S, Forjaz VH, Coutinho R, Ferreira T. 1990. *Hyperthermus butylicus*, a hyperthermophilic sulfur-reducing archaebacterium that ferments peptides. J Bacteriol 172:3959-3965.