A non-canonical pathway for aromatic amino acid biosynthesis in haloarchaea

Identification and function of essential genes

Miriam (Mirit) Kolog Gulko

Tel-Aviv

2010

Erklärung

Diese Dissertation wurde im Sinne von §13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Dr. Dieter Oesterhelt betreut.

Ehrenwörtliche Versicherung

Diese Dissertationn wurde selbständig, ohne unerlaubte Hilfe erarbeitet.

München, am ....

....
Miriam (Mirit) Kolog Gulko

Dissertation eingereicht am: 25/03/2010

1. Gutacher: Prof. Dr. Dieter Oesterhelt

2. Guttacher: Prof. Dr. Nediljko Budisa

Mündliche Prüfung am: 07/06/2010

Ricardo, as always, for all things-
past, present and yet to come

"The outcome of any serious research can only be to make two
questions grow where only one grew before"

Thorstein Veblen, (1857-1929) U.S. economist and sociologist

List of Figures

List of Tables

1 Abbreviations

ABZ- 4-aminobenzoate (folate precursore)

Amp- ampicillin

AroAAs- Aromatic amino acids

ASA- L-aspartate semialdehyde

Chlorm- chloramphenicol

CHY- conserved hypothetical protein

CSPD- disodium3-(4-methoxyspiro{1,2-dioxetane-2,3'-(5-hloro)tricycle [3.3.1.13, 17] decan}4-yl) phenyl phosphate

DAHPS-3-deoxy-D-arabion-heptulosonate-7-phosphate synthase

DHA- dihydroxyacetone

DHAP - dihydroxyacetone phosphate

DHQ-3- dehydroquinate

DHS-3- dehydroshikimate

DIG- dioxygenin

DKFP- 6-deoxy-5-ketofructose 1-phosphate

DS-downstream region

E-4-P- erythrose-4-phosphate

F-1,6-P- fructose 1,6-bisphosphate

GAP- glyceraldehyde 3-phosphate

GDH- glycerol dehydrogenase

Gn- generation time

GPDH- glycerol-1-phosphate dehydrogenase

H. s - *Halobacteruim salinarum* strain R1

HA- hydroxyacetone (2 keto 1-propanol)

HPAP- hydroxpyruvaldehyde phosphate

HY- hypothetical protein

IPTG- isopropyl-β-D-thiogalactopyranoside

Mev- mevinolin

NBHA - O-(4-nitrobenzyl) hydroxylamine hydrochloride

PEP- phosphoenolpyruvate

PPP - pentose phosphate pathway

PRPP- phosphorribosylpyrophosphate

RT- room temperature

SPH - spheroplasting solution

TBS- N-methyl-N-(tert-butyldiamethylsilyl) trifluroacetamide

TMS- N-methyl-N-(trimethylsilyl) trifluroacetamide

US- upstream region

WT- wild type

X-Gal- 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, $C_{14}H_{15}BrClNO_6$

2 Acknowledgments

I wish to express my gratitude to the many people who made this thesis possible:

I would like to gratefully acknowledge the supervision of Prof. Oesterhelt. I am grateful for numerous stimulating discussions, for the generous support and for his true gift to find mistakes in chemical structures of compounds and pathways.

I would like to thank Prof. Mike Dyall-Smith for more than few extensive conversations over the last two years, which very kindly read this thesis cover to cover, and his comments have led to the introduction of numerous improvements and new ideas.

To the Budisa group: Prof. Nediljko Budisa, Traudl Wenger, Dr. Sandra Lepthien, Dr. Birgit Wiltschi, Dr. Christina Wolschner, Dr. Lars Merkel, Michael Hösl, and Carlos G. Acevedo-Rocha. My "adopted group", but unlike adopted children I was honored to choose my own "family".

I owe my deepest gratitude to Jurgen Musiol, Lissy Weyher-Stingl (both from the core facility) and Sigrid Bauer. They have opened their hearts and their doors, supplying me with resources I never could have found without them.

To my invaluable network of supportive, forgiving, generous and loving friends without whom I could not have survived the process: Dr. Stephanie Bleicken, Dr. Barbara Mulinacci, Dr. Katarina Furtwangler, Dr. Rita Schwaiger, Dr. Wilfried Staudinger, Dr. Valery Tarasov, Dr. Lars Merkel, Ervil Jovkovic, Dr. Makdes Debela, Dr. Delwyn Dyall-Smith, Tanja Oberwinkler, Katrin Krzak, Peter Reichelt, Cacilia Kostler and Kathrin Klee. I wish to thank them all, for their support and especially for showing me that one could finish a thesis with their wits intact. Needless to mention that the order of the names is meaningless.

Special thanks to Dr. Lars Merkel, Dr. Frank Siedler, Jurgen Musiol and Prof. Moroder for numerous useful discussions about structures, chemical reactions and chemistry.

To the bioinformatic group: Dr. Ricardo del Rosario, Dr. Orland Gonzales, Dr. Alberto Marin Sanguino and Locedie Mansueto. I am happy I got the courage to open the door to your office, it was valuable. Although sometimes an unfamiliar language is been spoken (mathematics, computation, simulations…), it is still a pleasure. Special thanks to Dr. Friedhelm Pfeiffer for always been available to answer any question regarding the halolex and haloarchaea in general.

To the former Oe lab members: Bettina Brustmann, Brigitte Kessler and Susanne von Gronau. Thank you for introducing me to the unexpected complicated word of "halos", teaching me all you know and answering all my questions.

To Dr. Angela Witte and her group, I am grateful for the fruitful collaboration and for believing (long after I already gave up) that *N. pharaonis* can be transformed.

And finally, to Ricardo, my dearest friend, cheerleader, voice of reason, counselor, and life raft. Thank you for always coming to my defense.

3 Summary

Eukaryotes and bacteria are known to utilize the shikimate pathway for synthesizing aromatic amino acids (AroAAs). Of the 17 enzymes involved in the well-studied *E. coli* pathway, haloarchaea possess only 13 recognizable homologs, and these cover the steps necessary to convert 3-dehydroquinate (DHQ) to AroAAs. No homologs were found for the initial reactions in the pathway responsible for the biosynthesis of DHQ, suggesting that haloarchaea use a non-canonical pathway for biosynthesis of AroAAs.

To assign function to all individual genes of this de novo biosynthesis pathway in *H. salinarum*, a number of strategies were employed, including targeted mutations of the proposed first three genes, nutrient requirements, phenotypes, AroAAs uptake assays, enzyme activities of the purified gene products, and global examination of AroAAs-related genes using a genome-wide DNA microarray for *H. salinarum*.

In summary, the DNA microarray data provided solid evidence to support the function of all 13 gene homologues that are required downstream to DHQ. These were differentially expressed when AroAAs were omitted from the growth media. In addition, ORFs involved in specific and unspecific AroAAs transport systems were found to be strongly regulated including the *E. coli* aroP homolog, as well as *dpp*$A_3B_3C_3D_3F_3$ operon (ABC transporter) and various Na^+-dependent transporters. Finally, the transcription regulators TrmB (OE3018F) and TBPe (OE4146F) were strongly regulated when AroAAs were added to the growth media, while LrpA2 (OE2776F) was regulated when AroAAs were omitted from the growth media.

It was hypothesized that the first three steps of the de novo AroAAs pathway were catalyzed by ORFs OE1472F, OE1475F and OE1477R, respectively, based on sequence similarity to *M. jannaschii*, and genomic context. These predictions were tested by determining the role of these ORFs both in vivo and in vitro.

Several lines of evidence indicated that ORF OE1472F has a role in the AroAAs biosynthesis pathway: (1) genetically disturbing ORF OE1472F resulted in an auxotrophic mutant that required AroAAs for growth. (2) the purified recombinant protein specified by this ORF exhibits aldolase activity in the catabolic direction, and uses the phosphate sugar F-1,6-P as substrate. This might indicates that the ORF is

involved in the formation of DKFP, which is one of the precursors in the pathway. Nevertheless, the aldol like condensation between a hydroxyacetone fragment of DKFP with ASA needs to be tested. (3) this ORF was part of the *trp* operon identified in *H. salinarum*, which was transcriptionally regulated by AroAAs in the medium. (4) alignment with the transaldolase from *M. jannaschii* (which has known 3D structure) allowed OE1472F to be classified as an archaeal-type aldolase class 1A (transaldolase), supporting the possibility of this ORF catalyzing the first step in the de novo pathway.

The role of the second ORF in the de novo pathway was confirmed by the phenotype of an insertion mutant OE1475F::pMG501, which showed reduced ability to grow without AroAAs. In the presence of AroAAs it could grow as well as WT, confirming that this ORF specifies a protein that is important for AroAAs biosynthesis.

The predicted enzymatic function of the protein encoded by the 3rd ORF, OE1477R, was demonstrated in vitro. The purified gene product of ORF OE1477R could remove water molecule from DHQ to form 3- dehydroshikimate, demonstrating it is a 3-dehydroquinate dehydratase type I enzyme (EC 4.2.1.10). Deletion of the ORF resulted in AroAAs auxotrophy. This mutant utilized only low concentrations of Phe when growing on synthetic medium supplemented with only AroAAs, while the presence of AroAAs plus either DHQ or shikimate induced the uptake of Phe.

Together, the data for these three ORFs showed that, as predicted, they catalyze three consecutive enzymatic reactions responsible for the production of DHQ in the de novo pathway of AroAAs in *H. salinarum*.

4 Introduction

4.1 The three domains of life

The classification of cellular life has changed several times over the last 60 years as more information accumulated about cellular structure biology and genetics. In the late 1970s the classification of life was radically revised by Woese et al., (Woese and Fox, 1977; Woese et al., 1990) who found evidence for a previously unknown group of single-celled organisms, which were so different from the known bacteria that they required classification at the highest level of division. The distinctiveness of archaea was supported by 16S rRNA gene sequence, then later by phylogenies based on a number of proteins (such as elongation factors EF-1/Tu and EF-2/G, RNA polymerase subunits II and III and ATPases type F and V, (Gogarten et al., 1989; Iwabe et al., 1989; Puhler et al., 1989). The later studies also predicted that the root of the tree lay between the Bacteria lineage on the one side, and lineage leading to Archaea and Eukarya on the other (figure #1), indicating that Archaea were more related to Eukarya than to Bacteria. Apart from sequence data, all Archaea show a biochemical characteristic of major significance, the cell composition of the membrane. Unlike eubacterial and eukaryotic lipids, which usually consist of fatty acids that are linked to glycerol by ester bonds, archaeal membrane lipids are chemically very different, being composed of isoprenyl glycerol ethers (Gambacorta et al., 1994). Above the lipid linker, the archaeal cell is surrounded either by a proteinaceous surface layer (S-layer), or a rigid cell wall sacculus formed by polymers like pseudomurein or heteropolysaccharide, or nothing at all (Kandler, 1994). These and other data underpinned Woese's original 1977 proposal to split the prokaryotes into two distinct kingdoms (later Domains), the Bacteria and the Archaea (Woese et al., 1990). In the past few years, with the sequencing of complete genomes from numerous archaea (see Genome OnLine database at http://genomesonline.org/index2.htm), an enormous amount of data has become available that has been used to critically examine the evolutionary relationships among various organisms. These studies have confirmed the three-domain proposal, and it is now generally accepted that life is classified into three domains: Bacteria, Archaea, Eukarya.

The domain Archaea comprises 4 phyla (previously kingdoms) as seen in figure #1: Korarchaeota, Nanoarchaeota, Crenarchaeota and Euryarchaeota. Euryarchaeota unite a diverse group of archaea characterized by their ability to generate methane (methanogens) or by their growth in extreme habitats with high salt concentrations (extreme Halophiles) or at high (hyperthermophilic) and low (psychrophilic) temperatures. In contrast, crenarchaeota comprise almost exclusively hyperthermophilic species.
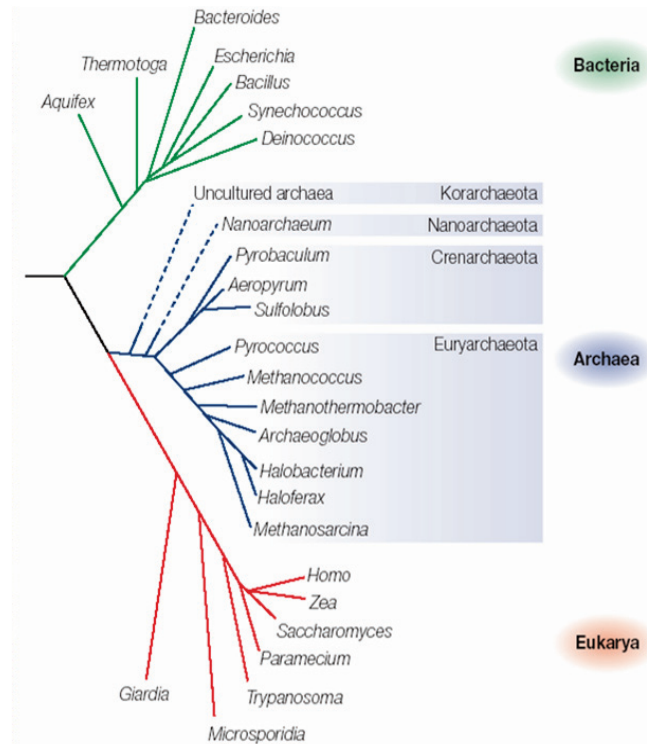
Figure # 1: The rRNA tree of life (Allers, 2005). Molecular phylogeny clearly distinguishes Archaea as a distinct domain of life independent from Bacteria and Eukarya.

4.2   Diversity and unity

As indicated by the tree of life (figure #1) the Archaea show a closer relationship to eukaryotes than to Bacteria. However, Archaea are prokaryotic in cell architecture. They are single-celled, contain no nucleus or other eukaryotic organelles, are of approximately the same size as most Bacteria, have circular DNA genomes, have genes organized into operons, and use pathways such as energy conservation and biosynthesis that are similar to Bacteria (Allers, 2005). In contrast, their information-processing machinery such as RNA polymerase, resemble their eukaryotic counterparts, although they are far simpler in structure (Allers, 2005).

Genome sequence comparisons revealed that Archaea might have acquired many of their genes by horizontal gene transfer (Koonin et al., 2001), seen especially in metabolically diverse methanogenic and halophilic archaea which cohabit with bacteria, and from which they may have acquired many genes for novel metabolic functions (Deppenmeier et al., 2002; Rosenshine et al., 1989). In particular, the haloarchaea appear to have evolved from the anaerobic methanogen lineage (figure #1), and must have recruited many genes for aerobic metabolism (see next section).

A perplexing issue arising from genome sequence data is that as much as 50% of the annotated archaeal genes have no identifiable functions. In many cases, metabolic pathways show gaps where no gene homologues can be found by sequence comparison, but functional homologues are presumably present among the complement of hypothetical proteins. For example, central pathways such as the pentose-phosphate pathway and shikimate pathway (Allers, 2005; Falb et al., 2008; Gonzalez et al., 2008).

4.3 Metabolism of halophilic archaea

Members of the family Halobacteriaceae are commonly referred to as extremely halophilic archaea, or halobacteria. Currently, 28 genera of this family are recognized and include members isolated from salt lakes, soda lakes (high pH), saltern crystallizer ponds, salty soils and salted foods (NCBI taxonomy website: http://www.ncbi.nlm.nih.gov/sites/entrez). Haloarchaea thrive in extremely saline environments (up to 4.5M NaCl) and share physiological adaptations such as acidic proteins that can function at the high internal salt concentrations found in the cytoplasm. Surprisingly, the nutritional demands and metabolic pathways were found to vary considerably among diffrent haloarchaeal species (for review see (Falb et al., 2008)).

H. salinarum, like other haloarchaea requires high salt concentrations for growth. It is an aerobic hetrotroph that degrades carbon sources such as amino acids, glycerol and organic acids via the TCA cycle (Ghosh and Sonawat, 1998) using oxygen as terminal electron acceptor (Schafer et al., 1996). Due to low the solubility of oxygen in salt-saturated brines, it easily becomes a limiting factor for respiration, and alternative terminal acceptors, such as DMSO and to some extent fumarate, have been reported to be utilized (Oren, 1991; Oren and Truper, 1990). *H. salinarum* employs

two further modes of energy conservation under anaerobic condition. First, it is able of photophosphorylate using the light-driven proton pump bacteriorhodopsin, building up a proton motive force for ATP generation (Oesterhelt, 1988). As a second possibility, *H. salinarum* is able to ferment arginine via the arginine deaminase pathway (Dundas and Halvorson, 1966; Ruepp and Soppa, 1996).

The metabolic diversity of halophilic archaea was investigated at genomic level through systematic metabolic reconstruction and comparative analysis of *H. salinarum*, *H. marismortui*, *H. walsbyi* and the haloalkaliphile *N. pharaonis* (for review see (Falb et al., 2008; Gonzalez et al., 2008)). The metabolic reconstruction showed that in spite their common hypersaline environments, halophilic archaea are surprisingly different in their nutritional demands. For example, while *N. pharaonis* can synthesis all amino acids de novo (Falb et al., 2005), *H. salinarum* has reduced capabilities to synthesize amino acid and therefore requires 5 amino acids in order to grow in synthetic medium (Grey and Fitt, 1976; Oesterhelt and Krippahl, 1973). Specifically, *H. salinarum* lacks gene clusters for valine, leucine, isoleucine, lysine, and arginine synthesis (supplementary material S2 in (Falb et al., 2008)). Furthermore, the predicted set of essential amino acids (lysine, methionine, valine, leucine and isoleucine) fits well to the set of amino acids that can be sensed by *H. salinarum*, except for lysine, which is not an attractant signal, and cysteine which is sensed by *Bas*T (Kokoeva et al., 2002).

The metabolic reconstruction by Falb et al., (Falb et al., 2008; Gonzalez et al., 2008)), also revealed some interesting metabolic pathways different to bacteria. For example, the biosynthesis of ribose-5-phosphate (R5P), the precursor of ribonucleotides, is usually synthesized through the pentose phosphate pathway (PPP). However, from the reconstruction of the PPP in *H. salinarum* it was observed that *H. salinarum* lacks the non-oxidative branch, and possesses only two enzymes of the oxidative branch. Those two enzymes constitute the final steps in producing ribulose-5-phosphate. The absence of the non-oxidative branch of the PPP pathway also results in the absence of the pathway intermediate, erythrose-4-phosphate, which is one of the precursors of aromatic amino acids (AroAAs) biosynthesis in bacteria, known as the shikimate pathway.

4.3.1 Shikimate pathway of aromatic amino acids (AroAAs) biosynthesis

In *E. coli*, phenylalanine, tyrosine, and tryptophan are synthesised by a common pathway, the shikimate pathway, as depicted in figure #2. The pathway derives its name from *Illicium religiosum* (in Japanese, shikimi-no-ki) from which shikimic acid was first isolated in 1885. The initial step is the condensation of phosphoenolpyruvate (PEP, a glycolytic intermediate) with erythrose-4-phosphate (E-4-P, a pentose phosphate pathway intermediate). The resulting C_7 open-chain sugar loses its phosphoryl group and cyclises to 3-dehydroquinate (DHQ). Dehydration then yields 3-dehydroshikimate (DHS), which is reduced by NADPH to shikimate. A second molecule of phosphoenolpyruvate then condenses with 5-phosposhikimate to give an intermediate that loses its phosphoryl group, yielding chorismate. Chorismate is the precursor for either prephenate, or anthranilate.
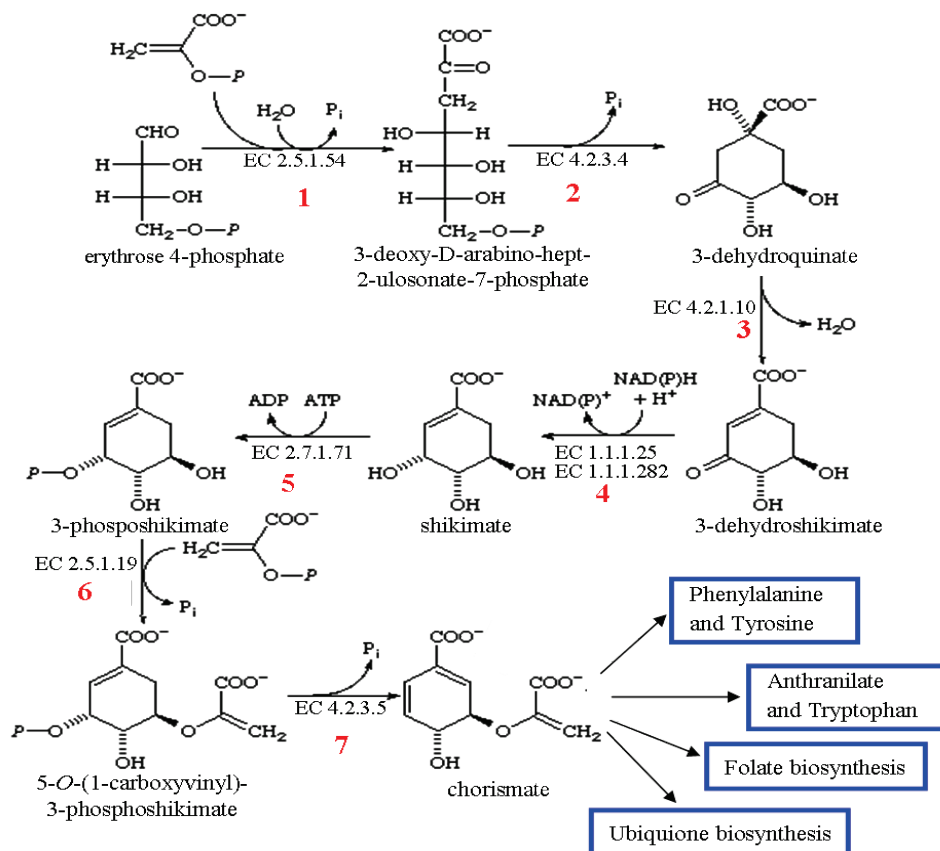
Figure # 2: The synthesis of chorismate in *E. coli*. (Modified from IUBMB Enzyme Nomenclature website, http://www.chem.qmul.ac.uk/iubmb/enzyme). The end products derived from the precursor chorismate are marked in blue boxes. The seven steps necessary to convert the precursors to chorismate are numbered in red. In fungi, such as *Saccharomyces cerevisiae*, steps 2-6 are catalyzed by the pentafunctional enzyme denoted AROM (Banerji et al., 1993; Charles et al., 1986).

In the prephenate branch (figure #3), a mutase converts chorismate into prephenate, the immediate precursor of the aromatic ring of phenylalanine and tyrosine. Dehydration and decarboxylation yields phenylpyruvate. Alternatively, prephenate can be oxidatively decarboxylated to yield *p*-hydroxyphenylpyruvate. These α-keto acids are then transaminated to yield phenylalanine and tyrosine.
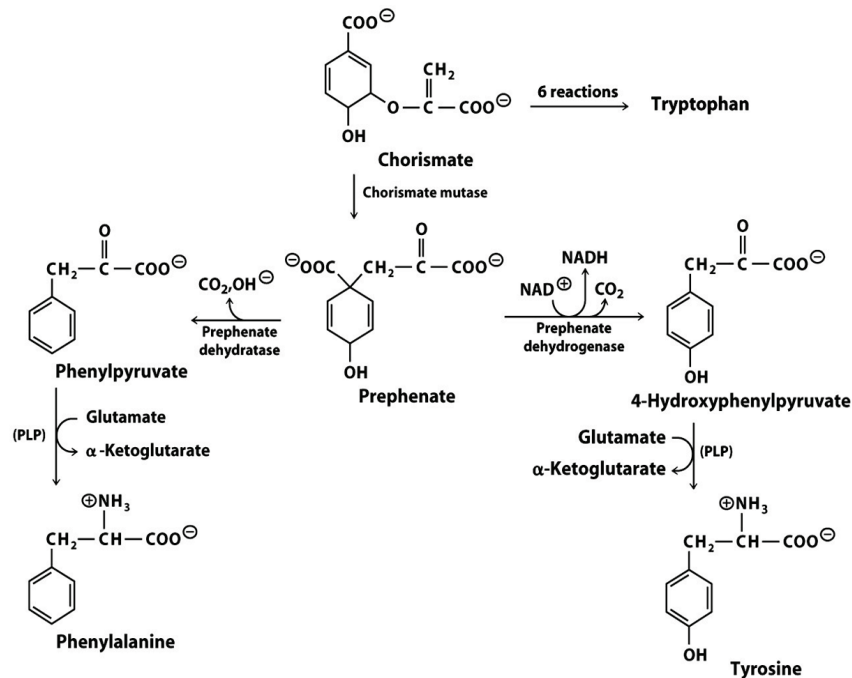
Figure # 3: The synthesis of phenylalanine and tyrosine from chorismate (The prephenate branch). (http://sandwalk.blogspot.com/2007/03/how-cells-make-tryptophan-phenylalan ine .html).

In the anthranilate branch (figure #4), chorismate acquires an amino group from the side chain of glutamine to form anthranilate. Anthranilate then condenses with phosphoribosyl-pyrophosphate (PRPP), an activated form of ribose phosphate. PRPP is also a key intermediate in the synthesis of histidine, purine nucleotides, and pyrimidine nucleotides. The C-1 atom of ribose-5-phosphate becomes bound to the nitrogen atom of the anthranilate in a reaction that is driven by the hydrolysis of pyrophosphate. The ribose moiety of phosphoribosylanthranilate undergoes rearrangement to yield 1-(o-carboxyphenylamino)-1-deoxyribulose 5-phosphate. This intermediate is dehydrated and decarboxylated to form indole-3-glycerol phosphate. Finally, indole-3-glycerol phosphate reacts with serine to form tryptophan. The glycerol phosphate side chain of indole-3-glycerol phosphate is replaced by the

carbon skeleton and amino acid group of serine. This reaction is catalysed by tryptophan synthase.
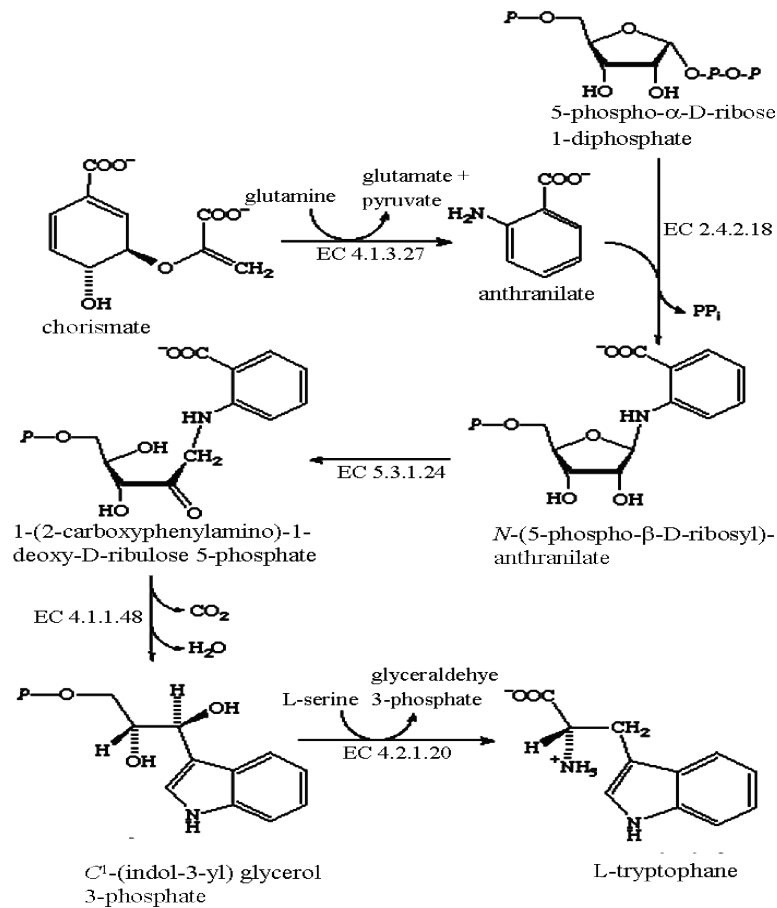
Figure # 4: The synthesis of tryptophan from chorismate (anthranilate branch). From IUBMB Enzyme Nomenclature website, http://www.chem.qmul.ac.uk/iubmb/enzyme.

4.3.2 Pathway reconstruction from genomic data

The seven enzymatic steps in the chorismate pathway have been extensively studied in *E. coli*, and the corresponding genes have been identified (Pittard, 1996). All of the participating enzymes, including shikimate kinase, are conserved in a broad range of organisms, such as bacteria, yeast and plants.

In 1996, the first archaeal genome sequence was published, that of the methanogenic archaeon *Methanocaldococcus jannaschii* (*M.j*) (Bult et al., 1996). In this organism, the genes for the first two steps of the chorismate pathway and for shikimate kinase (the fifth step) appeared to be missing. Selkov et al. (Selkov, 1997) even declared that the pathway starts from DHQ. Attempts to identify the gene encoding shikimate kinase using similarity to known examples of this enzyme yielded

no plausible candidates in any of the archaeal genomes so far sequenced (Daugherty et al., 2001).

In the case of the missing archaeal shikimate kinase gene, the problem was approached on basis of reconstruction of metabolic pathways and analysis of gene clustering on the chromosome. Daugherty *et al.*, (Daugherty et al., 2001) reconstructed the chorismate pathway from the genomic data as known for *E. coli* (figure #5), using more than 50 complete and partial microbial genomes. From the data presented in figure #5 one can see that the chorismate pathway is remarkably conserved over broad range of organisms, and homologs of all seven genes could be easily identified by sequence comparison. As summarized in figure #5, most prokaryotes possess either all of the genes (for example: *B. subtilis* and *A. fulgidus*), or none (for example: *T. pallidum* and *P.horikoshii*) and this is consistent with their growth requirements (Gonzalez et al., 1998). The missing archaeal shikimate kinase was identified by Daugherty *et al.*, (Daugherty et al., 2001), and verified experimentally by purification of the enzyme and determination of its kinetic parameters. Moreover, its orthologs in other archaea have been pointed out (Daugherty et al., 2001), figure#5).

GENOME	Step 1 EC 4.1.2.15 2-Dehydro-3-deoxyphosphoheptonate aldolase *aroH, aroG, aroF*	Step 2 EC 4.6.1.3 3-Dehydroquinate synthase *aroB*	Step 3 EC 4.2.1.10 3-Dehydroquinate dehydratase *aroD*	Step 4 EC 1.1.1.25 Shikimate 5-dehydrogenase *aroE, ydiB*	Step 5 EC 2.7.1.71 Shikimate kinase (Bacterial and Eukaryotic) *aroK, aroL*	Step 6 EC 2.5.1.19 Phosphoshikimate 1-carboxyvinyl transferase *aroA*	Step 7 EC 4.6.1.4 Chorismate synthase *aroC*	(Step 5) (EC 2.7.1.71) Predicted Shikimate kinase (Archaeal) -
E. coli	REC01661 REC05569 REC00721	REC05984	REC01650	REC05912 REC01649	REC05985 REC00372	REC00874	REC05421	-
B. subtilis	RBS02969	RBS02266	RBS02304 RBS02442	RBS02559	RBS00316	RBS02256	RBS02267	-
T. maritima	RTM00236	RTM00229	RTM00228	RTM00231	RTM00229	RTM00232	RTM00230	-
S. pneumoniae	RPN00965-66	RPN00386	RPN00384	RPN00385	RPN00391	RPN00390	RPN00387	-
T. pallidum	-	-	-	-	-	-	-	-
S. cerevisiae	RSC01644 RSC08655	RSC06906 (Pentafunctional Enzyme)					RSC05895	-
M. jannaschii	-	-	RMJ05308	RMJ00483	-	RMJ00806	RMJ07769	RMJ07785
A. fulgidus	-	-	RAG18799	RAG27692	-	RAG27692	RAG50410	RAG45918
A. pernix	RAP00399	RAP00398	RAP00397	RAP00396	-	RAP00394	RAP00393	RAP00395
P. furiosus	RPF01413	RPF01411-12	RPF01410	RPF01409	-	RPF01402	RPF01401	RPF01407-08
P. horikoshii	-	-	-	-	-	-	-	-

(Bacteria: *E. coli*, *B. subtilis*, *T. maritima*, *S. pneumoniae*, *T. pallidum*, *S. cerevisiae*; Archaea: *M. jannaschii*, *A. fulgidus*, *A. pernix*, *P. furiosus*, *P. horikoshii*)

Figure # 5: Reconstruction of the first steps in shikimate pathway in bacteria and archaea. The reconstruction was done using genomic data and chromosomal clustering of chorismate biosynthesis genes. Gene names shown in italic are those of *E. coli*. Modified from (Daugherty et al., 2001). The EC number of the first enzyme in the pathway has been changed to EC 2.5.1.54, and not as indicated in the figure above. Black boxes within a genome row represent proximity on the chromosome.

4.3.2.1 Shikimate pathway in *Methanocaldococcus* and *Methanococcus*

Although the chromosomal clustering approach proved to be useful in identifying the shikimate kinase in *M. jannaschii*, it failed to identify the first two genes in the pathway. This is most likely because a different pathway for the formation of DHQ is used, which is consistent with labelling studies of *M. jannaschii* showing that E-4-P is not the precursor of chorismate. Since the enzymes for the two steps preceding and the two steps following the phosphoylation of shikimate are present in all archaea, a complete pathway for synthesis of aromatic amino acid most likely exists, and so needs to be identified.

In 2004 White (White, 2004) reported that 6-deoxy-5-ketofructose 1-phosphate (DKFP) and L-aspartate semialdehyde (ASA) were precursors to DHQ. He demonstrated that DKFP supplies a hydroxyacetone moiety which, via a transaldolase reaction, undergoes an aldol condensation with ASA (figure #6). White has demonstrated that the recombinant protein product of *M. jannaschii* ORF MJ0400 catalyzes the transaldolase reaction while the recombinant protein product of ORF MJ1249 was responsible for the oxidative deamination and the cyclization reactions. Enzymatic function was demonstrated in vitro by quantifying the DHQ generated when the recombinant protein products of both these ORFs were incubated with DKFP and ASA (table #2 (White, 2004)). Addition of the individual proteins produced no product (0.012 nmole), whereas the mixture of the two proteins produced 7 nmole DHQ. The production of DHQ was greatly stimulated (330 nmole) when NAD was added. From the data presented, White estimated the specific activity of the MJ0400 enzyme was > 860 nmole/min/mg protein (White, 2004).

The in vivo function of the ORFs suggested by White (White, 2004) for the early steps in the novel pathway for the biosynthesis of AroAAs in *M. jannaschii*, were characterized both biochemically and genetically in *M. maripaludis* (Porat et al., 2006). Strong evidence was found to support the in vivo role of DKFP in this pathway. Firstly, phenylalanine was found to be labelled after the incorporation of labelled acetate. Secondly, a mutant with a deletion of the first ORF (aroA′, corresponding to MJ0400) required AroAAs for growth. Complementation with this gene restored the WT phenotype. Unexpectedly, deletion of the second ORF (aroB′ corresponding to MJ1249) did not convert the strain to auxotrophy, suggesting this gene was not required for AroAAs biosynthesis. Although the bioinformatic analysis

supports the role of this gene, it might be that there is another, unidentified enzyme that is able to perform the same function. Since there are no other ORFs homologous to aroB′, the proposed second enzyme is unrelated.
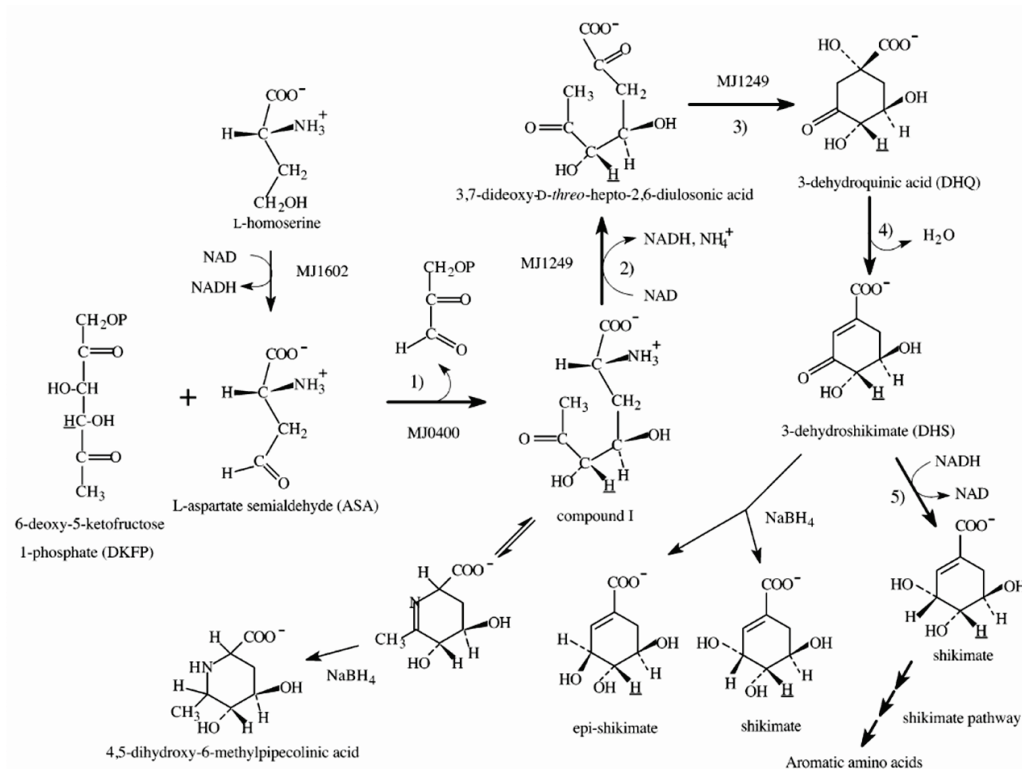
Figure # 6: The alternative pathway for the biosynthesis of AroAAs in *M. jannaschii* as suggested by White (White, 2004).

4.3.2.2 The Shikimate pathway in Haloarchaea

Several lines of circumstantial evidence indicate that haloarchaea possess a novel pathway for the early steps of AroAAs biosynthesis. First, in the archaeal domain, the oxidative pentose phosphate pathway (PPP), as seen in figure #7, does not appear to be present. The genes for most of this pathway cannot be detected in archaeal genomes (indicated in red), including those responsible for the production of E-4-P. Unless there is an alternative pathway for production of E-4-P, it can not be used as the precursor of the AroAAs pathway (step 1 in figure #2).
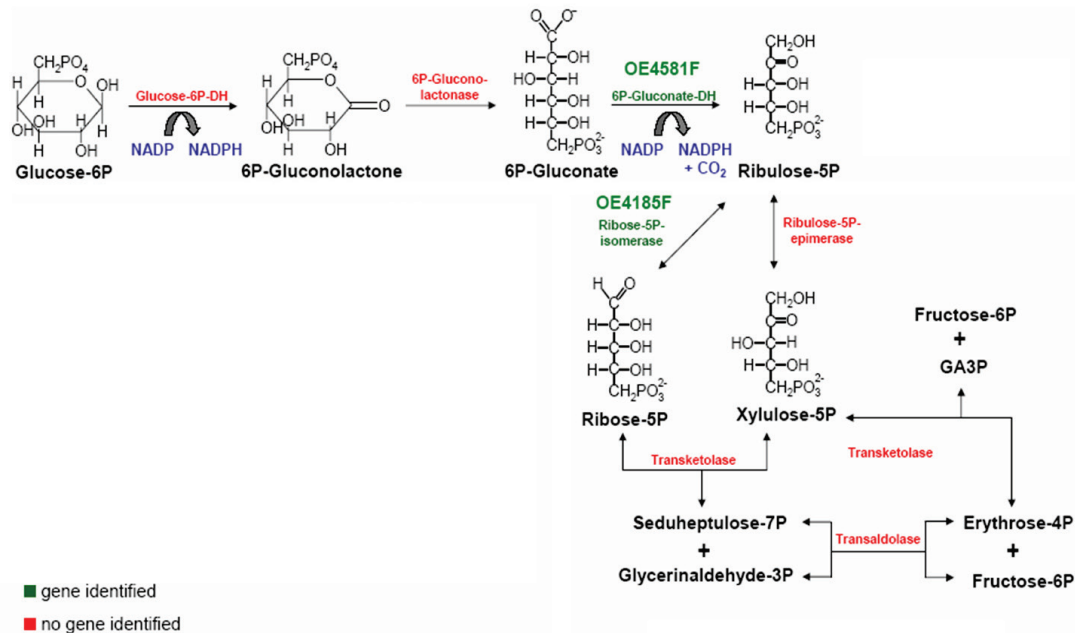
Figure # 7: The pentose phosphate pathway (PPP) in *H. salinarum*, the source of E-4-P. The enzymes for which genes cannot be found are indicated in red. Based on the experimental results of Dr. Daniela Breckau.

Second, homologues for all 13 genes necessary to convert DHQ to AroAAs are present in all sequenced haloarchaeal genomes, indicating that downstream to DHQ, the canonical pathway for synthesis of AroAAs takes place, as shown in figure #8.

Third, attempts to identify the genes encoding the first two steps using sequence similarity to known enzymes yielded no plausible candidates in any of the archaeal genomes (www.halolex.mpg.de), which indicates that as yet unidentified and non-orthologous enzymes are likely to span these early steps.

Fourth, in M. jannaschii, ORFs MJ40400 and MJ1249 have been shown to play a role in the alternative pathway of DHQ biosynthesis (White, 2004). Homologues, displaying high levels of sequence similarity (31-44%), are encoded in haloarchaea genomes, as seen in table #1.
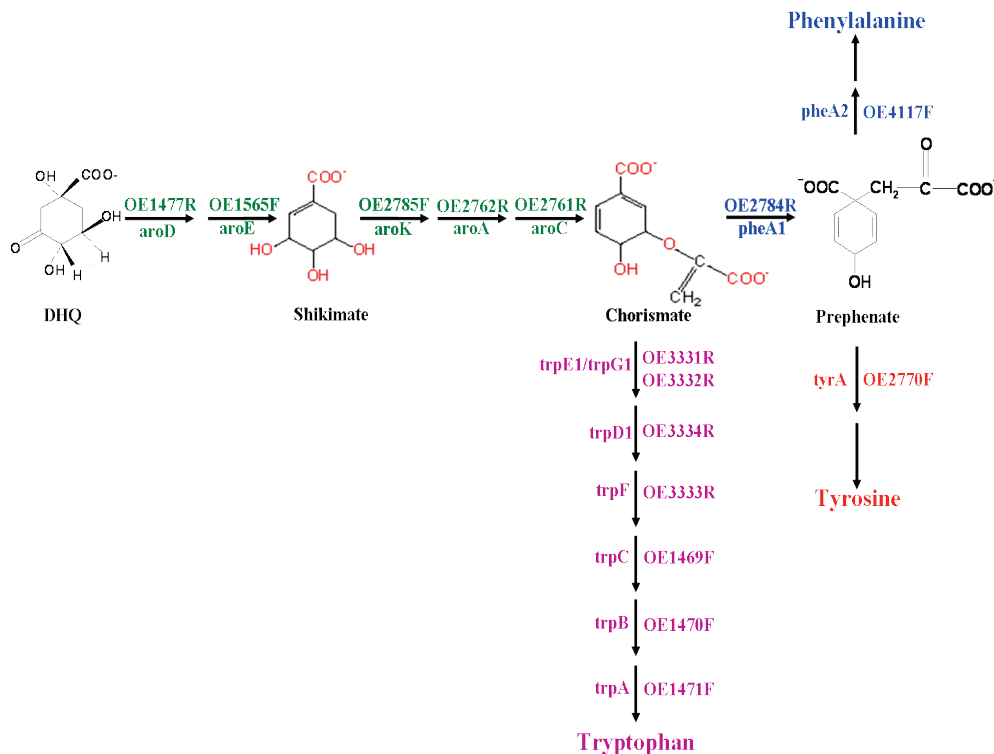
Figure # 8: Homologues ORFs of *H. salinarum* downstream of DHQ. Green represents the ORFs in the chorismate pathway starting from DHQ, pink- the ORFs of the tryptophan branch, blue- the ORFs in the phenylalanine branch, and red- the ORFs in the tyrosine branch.

While some haloarchaea have single homologues of *M. jannaschii* genes, *H. salinarum* and *N. pharaonis* each have two homologs of MJ0400 (aroA′). ORFs OE2019F and OE1472F were annotated as fructose bisphosphate aldolase I and II, respectively. Examination of the gene organization in *H. salinarum* and *N. pharaonis*, show that OE1472F and NP3160A are linked to their respective *trp*CBA operons (figure #9), indicating that they are likely to be involved in the *trp* pathway, while OE2019F and NP1594A, respectively, are involved in gluconeogensis. However, in *M. jannaschii* and *M. maripaludis*, the homologs to aroA′ are not linked to the genes of tryptophan biosynthesis (Porat et al., 2006). ORF OE1475F was annotated as conserved protein (www.halolex.mpg.de).

18

Table # 1: Homologs for the *M. jannaschii* ORFs MJ0400 and MJ1294. 31-44% amino acid identity was found.

	Fructose bisphosphate aldolase EC 4.1.2.13	3- dehydroquinate synthase EC 4.2.3.4
M. jannaschii	**MJ0400**	**MJ1249**
H. salinarum	**OE1472F**, OE2019F	**OE1475F**
N. pharaonis	NP3160A, NP1594A	NP2238A
Haloarcula marismortui	rrnAC1881	rrnAC1879
Haloquadratum walsbyi	HQ1156A	HQ1155A
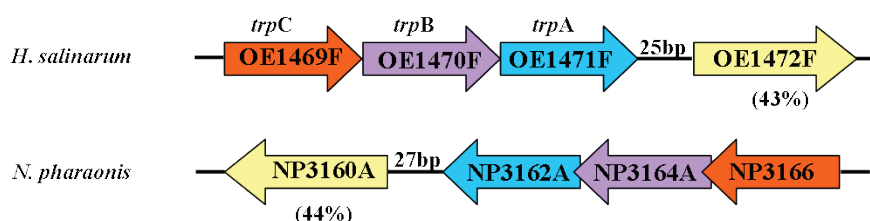
Figure # 9: ORFs OE1472F and NP3160A relative to the *trp*CBA operon in *H. salinarum* and *N. pharaonis*, respectively. The corresponding genes in the *trp* operon have the same color. The % of amino acid identity to *M. jannaschii* ORF MJ0400 is indicated in parenthesis. Note that the linker between *trp*A and OE1472F or NP3160A do not contain a putative TATA box.

The genomic contexts of all ORFs involved in the de novo pathway of AroAAs biosynthesis pathway in *H. salinarum* are represented in the figures 10 and 11 below.

Figure # 10, illustrates the ORFs responsible for the conversion of the precursors to chorismate. As explained above, several lines of circumstantial evidence indicates that haloarchaea possess a novel pathway for the early steps of AroAAs biosynthesis. These early steps are assumed to be responsible for the production of DHQ, using ORFs OE1472F and OE1475F (marked with red circles), homologs of ORFs in *M. jannaschii* (White, 2004). ORF OE1477R represents the first step for which a homolog has been identified (figure #11A-B). The genomic contexts of the ORFs responsible for the conversion of DHQ to chorismate are represented in figure #10C-E.

In figure #11, F-I represents the ORFs assigned for the tryptophan, tyrosine and phenylalanine branches, respectively. Figure #11J, on the other hand illustrates the genomic context of ORF OE2019F, homolog to OE1472F (see table #1).
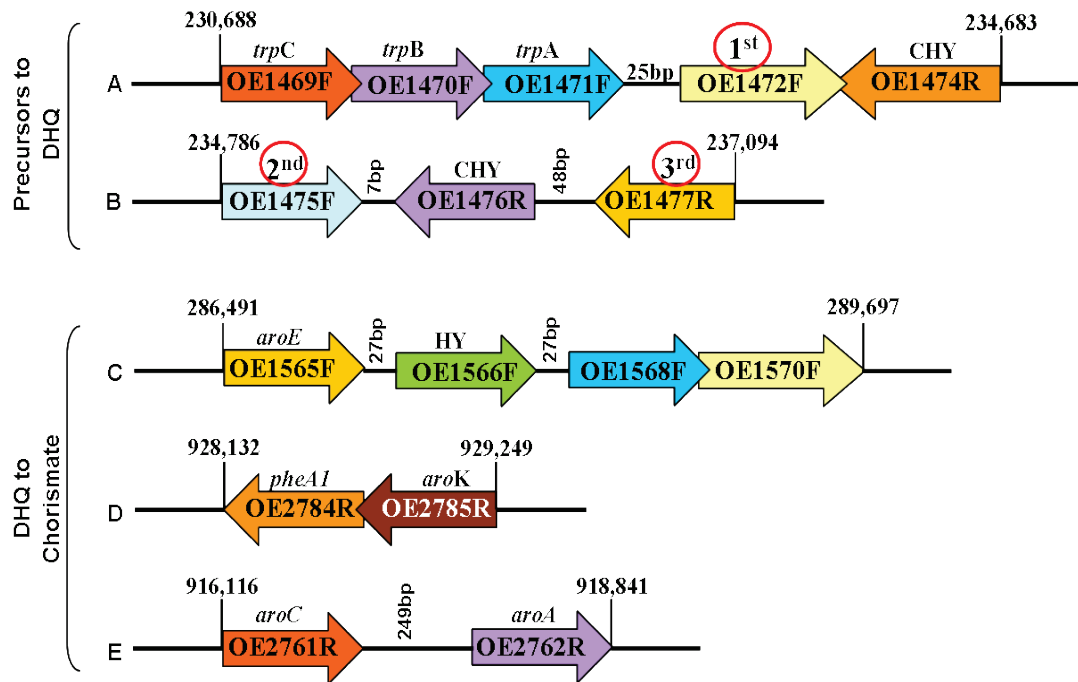
Figure # 10: The genomic context of ORFs involved in the conversion of the precursors to chorismate. **A-B**: ORFs assigned to convert the precursors to DHQ. The first and second ORFs (marked in red circles) are homologs to *M. jannaschii* ORFs (White, 2004) and part of the non canonical part of the pathway. ORF OE1477R is the 3rd ORF discussed in this thesis. **C-D**: ORFs needed to convert DHQ to chorismate. The numbers represent the positions of the ORFs in the chromosomal DNA of *H. salinarum*. **CHY**- conserved hypothetical protein, **HY**- hypothetical protein. The locations and directions of open reading frames are indicated by arrows.
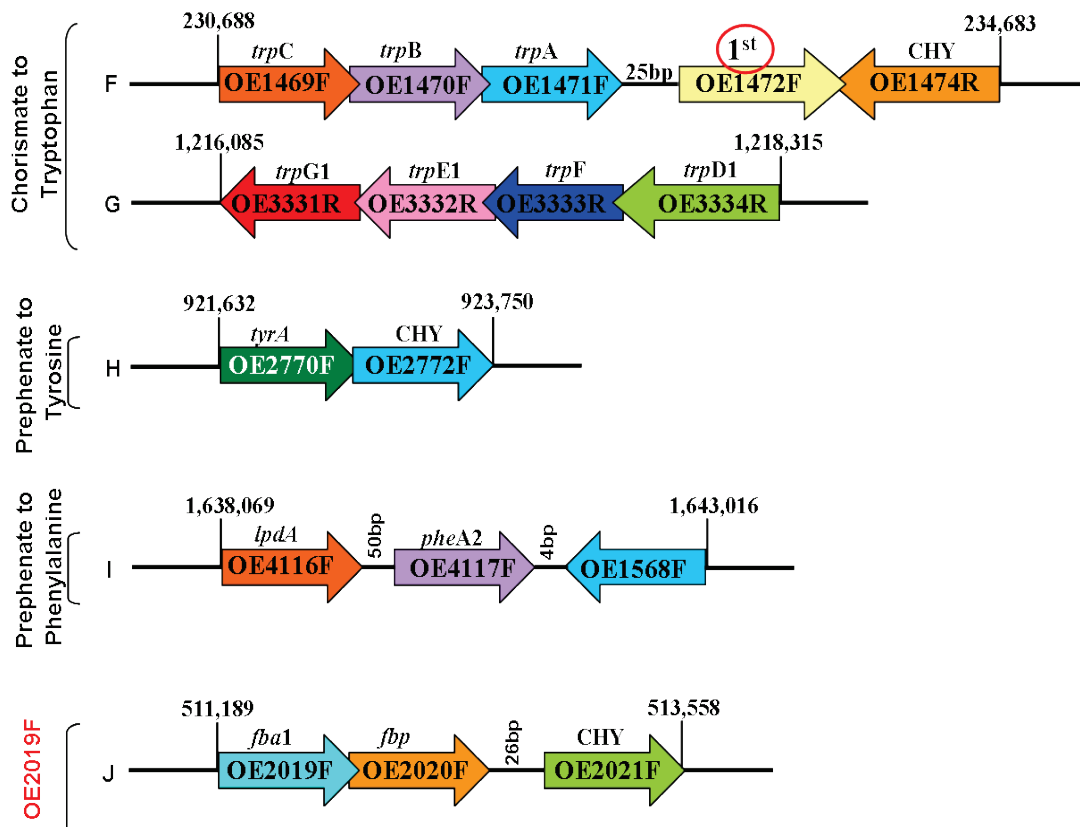
Figure # 11: The genomic context of the ORFs involved in the conversion of chorismate to tryptophan, tyrosine and phenylalanine. **F-G**: the tryptophan branch, **H**- the tyrosine branch, **I**-the phenylalanine branch. **J**- the genomic context of ORF OE2019F, homolog to OE1472F. The numbers represent the positions of the ORFs in the chromosomal DNA of *H. salinarum*. CHY- conserved hypothetical protein. The locations and directions of open reading frames are indicated by arrows.

4.3.2.3 Postulating de novo pathway in Haloarchaea

On the basis of the evidence presented above, ORFs OE1472F and OE1475F, were proposed as candidates for the first two steps in the biosynthesis of AroAAs in haloarchaea. This was consistent with the pathway suggested by White for *M. jannaschii* (White, 2004), and provided a testable hypothesis: that haloarchaea homologous would catalyse the same reactions as those shown experimentally for *M. jannaschii*.

The first step would then be predicted to be an aldol like condensation between HA (hydroxyacetone) moiety of DKFP (6-deoxy-5-ketofructose 1-phosphate) to ASA (aspartate semialdehyde), using the enzyme specified by ORF OE1472F. Hydroxpyruvaldehyde phosphate (HPAP) would be released and compound I formed. In the second stage, oxidative deamination and the cyclization using ORF OE1475F is

predicated to form DHQ. From this stage on, the canonical pathway can be used for the biosynthesis of AroAAs.
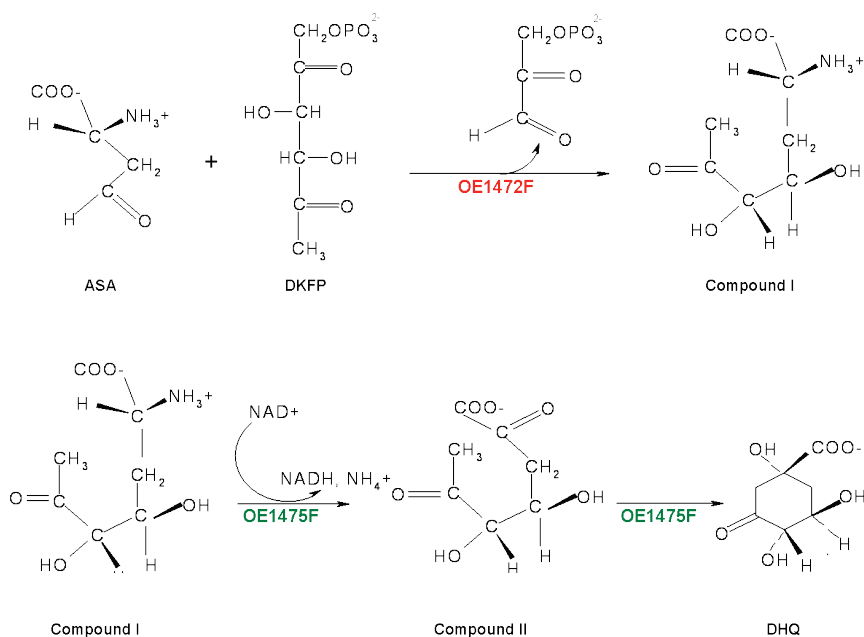
Figure # 12: Initial steps in the de novo pathway of AroAAs biosynthesis in haloarchaea according to (White, 2004). **ASA**- L-aspartate semialdehyde, **DKFP**-6-deoxy-5-ketofructose 1-phosphate.

4.4 Enzyme classification

The enzymes relevant to the first three steps in the predicted pathway are aldolase/transaldolase, dehydroquinate synthase, and dehydroquinate dehydratase, enzymes specified by ORFs OE1472F, OE1475F and OE1477R, respectively.

4.4.1 Aldolases EC 4.1.2.13

Aldolase is an enzyme which is capable of cleaving fructose 1-phosphate to form glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). The reaction is reversible and part of glycolysis.

Two distinct classes of aldolases are found in nature, which differ in their molecular properties and mechanisms (Rutter, 1964a). **Class I aldolases**, as represented by rabbit muscle aldolase, catalyses the cleavage of FBP through a covalently bound intermediate. The intermediate is stabilized by a Schiff base formed

between the substrate and an active site residue of the enzyme, Lys237. Enzymes which are part of this class are inactivated by borohydride ($NaBH_4$) (Rutter, 1964b). **Class II aldolases**, represented by yeast aldolase, depend on divalent metal ions to stabilize the carbanion intermediate, and are inhibited by EDTA (Lebherz and Rutter, 1969).

Early studies of extreme halophiles described the distribution of class I and II aldolases in various haloarchaea. Their activity was measured in cell-free extracts in the presence of EDTA or $NaBH_4$ (Altekar W., 1987). These studies showed that the *H. salinarum*, *H. volcanii* and *H. mediterranei* enzymes exhibit the properties of class II aldolase (i.e. their activities were metal dependent), while the aldolase of *H. vallismortis* was classified as a class I enzyme, due to inactivation by $NaBH_4$. Although class I and II aldolase activities have been demonstrated in archaea (Altekar W., 1987; Dhar and Altekar, 1986; Fuchs et al., 1983; Krishnan and Altekar, 1991; Schafer and Schonheit, 1992), no genes encoding classical class I and II enzymes have been identified in any of the sequenced archaeal genomes, suggesting that archaea possess a novel type of aldolase. In 1998, Thomson et al., (Thomson et al., 1998) described a new type of FBP aldolase in *E. coli*, which belongs to the class I family. Although it uses a Schiff-base mechanism, its shows only low sequence similarity to other members of this class. This aldolase was originally misannotated in the *E. coli* genome as dehydrin (DhnA). Homologs were identified in archaea, as seen in the multiple alignments done by Siebers et al., (Siebers et al., 2001). Remarkably, *M. jannaschii*, *A. fulgidus*, *H. salinarum* NRC-1 and *H. salinarum* each possess two paralogous genes of this family, while *Pyrococcus furiosus* and *Aeropyrum pernix* each possess one gene. The enzymes from *Pyrococcus furiosus* and *Aeropyrum pernix* (Siebers et al., 2001) were expressed in *E. coli*, and showed preferred substrate specificity for FBP in the catabolic direction and exhibited metal-independent class I aldolase activity via a Schiff-base mechanism. Consequently, class I aldolases were grouped into two different sequence families, one with only eukaryotic members and one with both bacterial and archaeal members. The latter were designated **class IA aldolase (DhnA family),** possessing a catalytic lysine residue (Lys237) (Siebers et al., 2001; Thomson et al., 1998).

4.4.2 Dehydroquinate synthase EC 4.6.1.3

This enzyme catalyses the conversion of 3-deoxy-D-arabino-heptulosonate 7-phosphate to 3-dehydroquinate (DHQ) in the shikimate pathway of bacteria and eukaryotes (see figure #2, second step). The overall reaction is complex and involves alcohol oxidation, phosphate elimination, carbonyl reduction, ring opening and intermolecular aldol condensation, in a single active site (Srinivasan et al., 1963; Widlanski et al., 1989). In prokaryotes, the enzyme is encoded by a single function gene, *aro*B (Bentley, 1990). In some eukaryotes such as *Neurospora crassa*, *Aspergillus nidulans* and *Saccharomyces cerevisiae*, dehydroquinate synthase is produced as part of the pentafunctional AROM protein which catalyse steps 2-6 (figure #2) (Banerji et al., 1993; Charles et al., 1986).

ORF OE1475F in *H. salinarum* was annotated as a conserved hypothetical protein but displayed a 41% sequence similarity to MJ1294 from *M. jannaschii*. The latter enzyme has shown to be responsible for the synthesis of DHQ in the de novo pathway of AroAAs in *M. jannaschii* (White, 2004).

4.4.3 dehydroquinate dehydratase EC 4.2.1.10

A 3-dehydroquinate dehydratase is an enzyme that catalyzes the dehydration of 3-dehydroquinate (DHQ) to form 3-dehydroshikimate (DHS) while releasing a water molecule. This enzymatic reaction is the 3rd step in the shikimate pathway (figure #2, 3rd step).

Two distinct types of 3-dehydroquinate dehydratase are found in nature which catalyze the same overall reaction but are unrelated at the sequence level and utilize completely different mechanisms (Gourley et al., 1999). Both enzyme types function as part of the shikimate pathway but while type I is found in archaea, plants and some bacteria, type II is found in bacteria and fungi. 3-dehydroquinate dehydratase from the bacterium *Salmonella typhimurium* was the first type I enzyme-product complex to be crystallized and its structure solved to 2.1Å resolution (Gourley et al., 1999; Lee et al., 2002). In 2004, the structure of another type I enzyme from *Staphylococcus aureus* was solved (Nichols et al., 2004), showing that type I enzymes belong to the $(\beta\alpha)_8$-barrel fold superfamily.

ORF OE1477R (*aro*D) in *H. salinarum* was annotated as a 3-dehydroquinate dehydratase, and predicted to be responsible for the formation of DHS.

4.5 Microarray analysis

The aromatic amino acids phenylalanine, tyrosine and tryptophan are synthesized in *E. coli* via the shikimate pathway, a well-studied and widely conserved biosynthetic pathway that involves 17 enzymes, and multiple levels of regulation (Bentley, 1990; Knaggs, 1999; Pittard, 1996). When analysing similar pathways in novel microorganisms, particularly those with non-canonical enzymatic steps such as haloarchaea, a number of strategies need to be considered. At the transcriptional level one might use traditional methods for detecting gene expression changes such as Northen blotting or differential RNA display, which allow analysis of only one or few genes at a time. DNA microarrays on the other hand are a much more powerful tool for studying genome-wide differential gene expression. Microarrays allow high throughput, parallelism, speed and automation and by encompassing the whole genome it eliminates bias associated with preselecting a subset of genes believed to be involved in certain cellular event. Nonetheless, microarray experiments produce voluminous datasets which are frequently difficult to analyse and can lead to confusing hypotheses and conclusions (Dharmadi and Gonzalez, 2004).

DNA Microarrays were developed in the early nineties of the last century and have been used successfully to study gene expression in many organisms, including plants (Aharoni, 2001), pathogen-infected hosts (Diehn and Relman, 2001), humans (Kurella et al., 2001), and bacteria, such as *Bacillus subtilis* and *E. coli* (Eymann et al., 2002; Laub et al., 2000; Rohlin, 2002; Ye et al., 2000). Until 2001, no genome-wide transcriptome analysis had been performed on any archaeal species, and only one example was published using an array based on 271 ORFs from the hyperthermophilic archaeon *Pyrococcus furiosus* (Schut et al., 2001). In recent years, DNA microarrays of archaeal species have become an important tool for studying complex biological processes, such as understanding the regulatory network underlying the biochemical pathways for phototropy in *H. salinarum* NCR-1 strain (array based on 2413 ORFs (Baliga et al., 2002)). Microarrays were used as well to investigate chromosome replication in the hyperthermophilic acidophiles *S. acidocaldarius* and *S. solfataricus* using arrays based on 1,914 and 2,488 ORFs,

respectively (Lundgren et al., 2004). In 2006, the genome-wide DNA microarray (1722 ORFs) made of *Methanococcus maripaludis* was published, which validated proteomic data by microarray and RT-PCR (Xia et al., 2006). This was followed by a publication from Twellmeyer et al., (Twellmeyer et al., 2007)which established the first genome-wide DNA microarray of *H. salinarum* R1 (2774 ORFs). This array has been used to analyse the transcriptional differences between cells grown under aerobic and phototrophic conditions (Twellmeyer et al., 2007), the phosphate-dependent gene expression in *H. salinarum* (Wende et al., 2009), the global and specific transcriptional regulatory effects of Lrp-like proteins in *H. salinarum* (Schwaiger, 2010), and the AroAAs-related genes of the shikimate pathway (this study).

4.6 Objectives

The pathway leading to the biosynthesis of aromatic amino acids, known as the shikimate pathway, is well understood in both, Eukaryotes and Bacteria, where the genes, enzymes and the reactions, for each step have been intensively studied (Daugherty et al., 2001; Pittard, 1996). This pathway is critical for growth as, in addition to the biosynthesis of AroAAs, it is the source of *p*-aminobenzoic acid (the precursor of folic acid), and *p*-hydroxybenzoic acid, the precursor of the quinones, which are members of the electron transport chain. Analysis of the haloarchaeal genomes has failed to identify the first two genes in this pathway, but from studies on the euryachaeota *M. jannaschii* (White, 2004), a pathway in haloarchaea can be postulated. This proposed non-canonical pathway is the subject of this investigation in *Halobacterium salinarum*, with the aims to provide genetic and biochemical evidence for the full de novo pathway of aromatic amino acids in haloarchaea.

The specific objectives related to verifying the proposed pathway were:

1) The construction of mutants of each of the first three genes: OE1472F, OE1475F and OE1477R (figure #10A-B). Deletion of these genes would be expected to convert *H. salinarum* to aromatic amino acid auxotrophy.

2) Determining the phenotypes of these mutants including nutrient requirements, growth rates and their ability to import AroAAs from the medium.

3) Assessing the differential gene expression of all assigned genes for the AroAAs biosynthesis pathway using a DNA microarray of *H. salinarum*.

4) Expression and purification of the gene products of OE1472F, OE1477R and OE2019F in order to examine their enzymatic activities.

5) Developing the relevant biological assays to determine the activities, kinetics and substrate specificities of the purified OE1472F, OE1477R and OE2019F proteins.

5 Materials and Methods

5.1 Materials

5.1.1 Chemicals

All chemicals used in this study were purchased from Sigma-Aldrich, Merck, or Riedel de Haen at the highest grade available. Exceptions are listed below:

Bacto^{TM} agar	Difco, Detroit, MI, USA
Bacto^{TM} trypton	Difco, Detroit, MI, USA
Bacto^{TM} yeast extract	Difco, Detroit, MI, USA
BMA Seakem^{TM} LE agarose	Biozym, Hess. Oldendorf, Germany
disodium3-(4-methoxyspiro{1,2-dioxetane-2,3'-(5- hloro)tricycle [3.3.1.13,17] decan}4-yl) phenyl phosphate (CSPD)	Roche Diagnostics, Mannheim, Germany
DIG-11-dUTP	Roche Diagnostics, Mannheim, Germany
DNA ladder 1Kb	Invitrogen, Karlsruhe, Germany
Polyethylene glycol (PEG$_{600}$)	Sigma Aldrich, St. Louis, USA
HisTrap HP-1ml	GE Healthcare, Freiburg, Germany
Albumin, bovine fraction V	SERVA Electrophoresis GmbH, Heidelberg, Germany

5.1.2 Kits and Enzymes

All restriction enzymes were purchased from New England Biolabs, Beverly, MI, USA. Other enzymes or kits used in this study are listed below:

ABI PRISM BIG DYE^{TM} Cycle Sequencing Ready Reaction Kit	Applied Biosystems, Foster City, CA, USA
BCA Reagent Assay	Pierce, Rockford, IL, USA
DIG-11-dUTP	Boehringer Mannheim Biochemicals
LA Taq Polymerase	TaKaRa Bio Inc, Japan

QIAprep[TM] Spin Miniprep Kit	Qiagen, Hilden, Germany
QIAquick[TM] Gel Extraction Kit	Qiagen, Hilden, Germany
Wizard® SV Gel and PCR Clean Up System	Promega, Madison Wisconsin, USA
ECL Plus Western blot Detection system	GE Healthcare, Freiburg, Germany

5.1.3 Media, Buffers and Stock solutions

5.1.3.1 Media

Media were prepared with distilled water and sterilized in an autoclave, with the exception of the synthetic medium where sterilization was achieved by filtration. For plates, 15g of Bacto-Agar were added to 1 liter of medium in a 2 liter flask. Antibiotics were added to the medium after it cooled down to ~60°C. Before autoclaving, the pH was adjusted to pH=7.0, when necessary.

LB medium for E. coli

	Per 1 liter
Bacto-Tryptone	10g
Yeast extract	5g
NaCl	10g

„Halo medium" for Halobacterium salinarum (Oesterhelt, 1983)

	Per 1 liter
NaCl	250g
$MgSO_4$ x $7H_2O$	20g
Na_3-citrate x $2H_2O$	3g
KCl	2g
Oxoid-Peptone	10g
pH adjusted to 7-7.2 with NaOH	

__Synthetic medium for Halobacterium salinarum + 1.1mM AroAAs__

(Modified from (Oesterhelt, 1983))

Solution 1: The substances were dissolved in 800ml of distilled water, and the pH was adjusted to pH =7.0 with 1M NaOH.

	MW (g/mole)	Amount (g)	Final concentration *(mM)
NaCl	58.44	235	3904.1
$MgSO_4$ x $7H_2O$	246.48	20	78.78
KCl	74.55	2	26.05
KNO_3	101.11	0.1	0.96
Na_3-citrate x $2H_2O$	294.1	0.5	1.65
Glycerol	92.09	20ml of 50%(W/V)	109
800ml of this solution was used for preparing 1.03 liter medium			

Solution 2: The substances were dissolved in 100ml of distilled water, and autoclaved.

	MW (g/mole)	Amount (g)	Final concentration * (mM)
KH_2PO_4	136.09	2.86	0.408
K_2HPO_4 x $3H_2O$	228.19	6.62	0.563
2ml of this stock solution was used for preparing 1.03 liter medium			

Solution 3: The trace elements were dissolved one after the other in ~80 ml distilled water which contained 4 drops of 25% HCl. After all the substances were dissolved, the final volume was adjusted to 100ml.

	MW (g/mole)	Amount (g)	Final concentration * (μM)
$CuSO_4 \times 5H_2O$	249.68	0.005	0.1944
$FeCl_2 \times 4H_2O$	198.81	0.23	11.23
$MnSO_4 \times H_2O$	169	0.03	1.723
$ZnSO_4 \times 7H_2O$	287.54	0.044	1.485
$Na_2MoO_4 \times 2H_2O$	241.95	0.024	0.963
$CoCl_2 \times 6H_2O$	237.93	0.024	0.979
1ml of this stock solution was used for preparing 1.03 liter medium			

Solution 4: The amino acids were dissolved in 750ml distilled water, and filter sterilized.

	MW (g/mole)	Amount (g)	Final concentration * (mM)
L-Arg	174.2	2	2.23
L-Ile	131.18	2.2	3.26
L-Leu	131.18	4	5.92
L-Lys-HCl	182.65	1.25	1.33
L-Thr	119.12	2.5	4.08
L-Ser	105.9	3.05	5.64
L-Ala	89.09	1.11	2.42
L-Gly	75.07	0.38	0.98
L-Pro	115.13	0.52	0.88
L-Asp	133.1	2	2.92
L-Glu	147.13	6.62	9
150ml of this stock solution was used for preparing 1.03 liter medium			

Solution 5: L-Tyr, 27.59mM stock solution

1.41g Tyr (MW=181.19g/mole) were dissolved in 49ml 1M HCl. The total volume was adjusted to 282ml with distilled water.

41ml of this stock solution was used to prepare 1.03liter solution with 1.1mM Tyr.

Solution 6: Vitamins were dissolved in 10ml distilled water, and sterilized by filtration.

	MW (g/mole)	Amount (g)	Final concentration * (μM)
Thiamine	337.3	0.1	14.39
Folic acid	441.4	0.1	11
Biotin	244.31	0.01	1.99
0.5ml of this stock solution was used for preparing 1.03 liter medium This solution was added last after the pH was adjusted to 7.0			

Solution 7: L-Met, 134.04mM stock solution

2 g Met (MW= 149.21g/mole) were dissolved in 100ml distilled water with ~ 1 drop of 1M NaOH. The final pH=6.85.

10ml of this stock solution was used to prepare 1.03 liter solution with 1.3mM Met.

Solution 8: L-Phe, 150mM stock solution

4.956g Phe (MW=165.19 g/mole) were added to ~150ml distilled water with 4 drops of 1N NaOH. Phe was added gradually. After all the Phe had dissolved, the total volume was adjusted to 200ml.

7.55ml of this stock solution was used to prepare 1.03 liter solution with 1.1mM Phe.

Solution 9: L-Trp, 134mM stock solution

1.37g Trp (MW=204.23 g/mole) were dissolved in ~30ml distilled water with 9ml 1M HCl. When all the Trp had dissolved, the volume was adjusted to 50ml and filter sterilized. **8.45ml** of this stock solution was used to prepare 1.03liter solution with 1.1mM Trp.

Solution 10: Val was dissolved in 100ml distilled water.

	MW (g/mole)	Amount (g)	* (mM)
L-Val	117.15	2.93	2.428
10ml of this stock solution was used for preparing 1.03 liter medium		

* The concentration in each table is the final concentration in 1.03 liter medium

**M205 medium for Natronomonase pharaonis** (according to DSMZ http://www.dsmz.de)

	Per 1 liter
Casamino acids	15g
Na_3-citrate x $2H_2O$	3.0g
Glutamic acid	2.5g
$MgSO_4$ x $7H_2O$	2.5g
KCl	2.0g
NaCl	250.0g
pH adjusted to 8.7-8.8 with 20% sterilized Na_2CO_3	

5.1.3.2 Buffers

**Basal salt solution**

Reagent	Amounts (g)	Final concentration
NaCl	250	4.278M
$MgCl_2$ x $7H_2O$	20	0.081M
Na_3-citrate x $2H_2O$	3	0.01M
KCl	2	0.0268M
H_2O	to 1liter	
pH adjusted to 7-7.2 with NaOH		

__Buffers for NiNTa column__

The amounts of the compounds are calculated for 1 liter buffer. The pH of all buffers was adjusted to pH=8.0.

Buffer/compounds	NaH_2PO_4 x H_2O	NaCl	KCl	Imidazole
NP (Binding buffer)	6.9g (50mM)	17.54g (300mM)	--------	--------
NPI$_{20}$ (Wash buffer)	6.9g (50mM)	17.54g (300mM)	--------	1.36g (20mM)
NPI$_{250}$ (Elution buffer)	6.9g (50mM)	17.54g (300mM)	--------	17.02g (250mM)
NP-KCl (Binding buffer)	6.9g (50mM)	--------	74.55 (1M)	--------
NPI$_{20}$-KCl (Wash buffer)	6.9g (50mM)	--------	74.55 (1M)	1.36g (20mM)
NPI$_{250}$-KCl (Elution buffer)	6.9g (50mM)	--------	74.55 (1M)	17.02g (250mM)

5.1.3.3 Antibiotics

Antibiotics	Stock** (mg/ml)	Solvent	Per liter	Final concentration
Ampicillin (Amp)	100	H_2O	1ml	100µg/ml
Chloramphenicol (Chlor)	34	H_2O	1ml	34µg/ml
Novobiocin (Nov)	20	H_2O	1ml	20µg/ml
Mevinolin (Mev)*	10	EtOH	1ml	10µg/ml

* Mevinolin (MKevinacor^{TM}): 6 tablets were ground to a powder, and mixed with total volume of 12ml EtOH. The mixture was stirred well for 10min at RT and then centrifuged (15min, 4000rpm). The supernatant was collected and filtered (0.45µM). Aliquots were stored at -20°C.

** Aliquots of the stock solutions were stored at (-20°C), until use.

5.1.4 Strains, Vectors and Oligonucleotides

5.1.4.1 Strains

Strain	Genotype or Phenotype	Source or reference
E. coli		
ElectroMax DH10B	F⁻*mcr*A (mrr-hsdRMS-mcrBC), O80*lac*ZΔM15, Δ*lac*X74, *rec*A1, *end*A1,*ara*Δ139, Δ(*ara*, *leu*7697) , *gal*U, *gal*kλ, *rps*L, *mup*G, ☐	Invitrogen (Electrocompetent)
ElectroMax StbI4	mBcrA Δ(*mcr*BC-*hsd*RMS-mrr), *rec*A1, *end*A1,*gyr*A96, *gal*⁻, *thi*⁻, *sup*E44λ⁻, *rel*A1, Δ☐*lac*-*pro*AB)/F' *pro*AB⁺ , *lac*I⁹ZΔM15, *Tn1*0(Tet^R)	Invitrogen (Electrocompetent)
One Shot®TOP10Electrocomp™	F⁻*mcr*AΔ(*mrr-hsd*RMS-*mcr*BC), φ80*lac*ZΔM15 Δ*lac*X74*rec*A1, *ara*D139Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L(Str^R)*end*A1 *nup*G	Invitrogen (Electrocompetent)
H. salinarum		
R1	BR⁺, HR⁺, SRI⁺, SRII⁺, Car+, Rub⁺, Ret⁺, Vac⁻	(Stoeckenius et al., 1979)
R1, ΔOE1477R	R1 derivative with ΔOE1477R, BR⁺, HR⁺, SRI⁺, SRII⁺, Car+, Rub⁺, Ret⁺, Vac⁻	This study
R1,OE1475F::pMG501	R1 derivative with OE1475F::pMG501, BR⁺, HR⁺, SRI⁺, SRII⁺, Car+, Rub⁺, Ret⁺, Vac⁻	This study
R1, OE1471F:: Terminator	R1 derivative with OE1471F:: Terminator (No promoter/terminator for OE1472F), BR⁺, HR⁺, SRI⁺, SRII⁺, Car+, Rub⁺, Ret⁺, Vac⁻	This study
N. pharaonis		
WT	Strain 2160 http://www.dsmz.de	(Falb et al., 2005; Kamekura et al., 1997; Soliman and Truper, 1982; Tindall et al., 1984)

5.1.4.2 Plasmids

Name	Relevant description	Source
Plasmids for _H. salinarum_		
pUS-Mev	6.3Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, bop	(Pfeiffer et al., 1999; Schweiger, 1996)
pWL-102	10.5Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R	(Lam and Doolittle, 1989)
pBPH-M	6.06Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, bop promoter, OE3349F as integration site, 7xHis	(Besir, 2001)
pMG550	6.832Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, bop promoter, OE3349F as integration site, **OE1472F**, 7xHis, parental plasmid - pBPH-M	This study
pMG650	7.198Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, bop promoter, OE3349F as integration site, **OE1475F**, 7xHis, parental plasmid - pBPH-M	This study
pMG750	6.715Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, bop promoter, OE3349F as integration site, **OE1477R**, 7xHis, parental plasmid - pBPH-M	This study
pMKK100	7.2Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, _Bga_H	(Koch and Oesterhelt, 2005)
pMG501	7.777Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, _Bga_H, **509bp of OE1471F (trpA), modified terminator of FlgA**, parental plasmid -pMKK100	This study
pMG601	7.729Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, _Bga_H, **500bp of middle OE1475F**, parental plasmid -pMKK100	This study
pMG500	8.185Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, _Bga_H, **flanking regions of OE1472F**, parental plasmid -pMKK100	This study
pMG600	8.184Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, _Bga_H, **flanking regions of OE1475F**, parental plasmid -pMKK100	This study
pMG700	7.213Kbp; _H. salinarum_ shuttle vector; Amp^R, Mev^R, _Bga_H, **flanking regions of OE1477R**, parental plasmid -pMKK100	This study
pET22b(+)	5.49Kbp; _E. coli_ expression vector; Amp^R, 6xHis	Novogen
pMG560	6.165Kbp; _E. coli_ expression vector; Amp^R, **OE1472F**, 6xHis, parental plasmid - pET22b(+)	This study
pMG760	6.048Kbp; _E. coli_ expression vector; Amp^R, **OE1477R**, 6xHis, parental plasmid - pET22b(+)	This study

pMG860	6.159Kbp; *E. coli* expression vector; Amp^R, **OE2019F**, 6xHis, parental plasmid - pET22b(+)	This study
Plasmids for *N. pharaonis*		
pUS-Mev	6.3Kbp; *H. salinarum* shuttle vector; Amp^R, Mev^R, bop	(Pfeiffer et al., 1999; Schweiger, 1996)
pMKK100	7.2Kbp; *H. salinarum* shuttle vector; Amp^R, Mev^R, *Bga*H	(Koch and Oesterhelt, 2005)
pNB102	9.1Kbp, *N. pharaonis* shuttle vector, Amp^R, Mev^R, colE1.ori, pNB101 (ori of N. pharaonis strain AS7091, 2.5Kbp)	(Zhou et al., 2004a; Zhou et al., 2004b)
pMG100	8.84Kbp, *N. pharaonis* shuttle vector, pNB101 (ori of N. pharaonis strain AS7091, 2.5Kbp), Amp^R, Mev^R, and bop gene from pUS-Mev plasmid.	This study
pMG200	9.751Kbp, *N. pharaonis* shuttle vector, pNB101 (ori of N. pharaonis strain AS7091, 2.5Kbp), Amp^R, Mev^R, and BgaH gene from pMKK100 plasmid.	This study

5.1.4.3 Oligonucleotides

All oligonucleotides used in this study were ordered from Metabion International AG (www.mymetabion.com). For the list of oligonucleotides used in this study, see appendix section 8.3 table #19.

5.2 Microbiological methods

5.2.1 Cultivation of *E. coli*

For the isolation of vector DNA, *E.coli* was cultivated in LB medium in 4.5ml or 50ml cultures inoculated at 37°C, 245rpm. Where appropriate, the LB medium was supplemented with antibiotics. For expression of proteins in *E.coli*, 1 liter volumes of LB medium (in 5 liter flasks) were inoculated with an overnight pre-culture and incubated at 37°C, 245rpm.

5.2.2 Cultivation of *H. salinarum* and *N. pharaonis*

Cultures were grown aerobically in the dark at 40°C and with constant shaking (100rpm). Cultivation of *H. salinarum* was carried out in 10ml medium (25ml flask), 35ml medium (100ml flask) or 1liter medium (2liter flask), and were inoculated with pre-cultures at rates of 1/50, 1/35 or 1/29 of the culture volume, respectively. Cell

growth was monitored by measuring optical density at 600nm (OD_{600}) in a spectrophotometer. In anaerobically grown cultures, the growth rate of the cells was monitored by light scattering of the culture using a Klett photometer.

Cultures of *N. pharaonis* were grown aerobically in the dark at 40°C and with constant shaking (100rpm), in 37ml medium (100ml flask). Cell growth was monitored by measuring optical density at 600nm (OD_{600}) in a spectrophotometer. To check the affect of Bacitracin (Merck) and protease (from *Streptomyces griseus*, Biochemika, cat#81748) on cell morphology, both were added to the growing medium in different concentrations.

5.3 Molecular biological methods

5.3.1 Transformation of *E. coli*

5.3.1.1 Preparation of chemically-competent cells

A single XL-10 colony from an LB plate was used to inoculate 25ml LB medium, and this was incubated at 37°C, 250rpm for 7.5h. This was used was used to inoculate two 1liter flasks, each containing 250ml of SOB medium (Super Optimal Broth used for initial growth). The first flask received 10ml of the starter culture, and the second received 4ml of the starter culture. The two 1liter flasks were incubated at 22°C, 150rpm.

The following morning, growth of the cultures was monitored every 45min until one of the flasks reached OD_{600}=0.5-0.55. Then the culture vessel was transferred to an ice-water bath for 10min. Cells were harvested by centrifugation at 2500g, 10min at 4°C (Sorvall rotor). The medium was poured off, and the inverted centrifuge bottle was left on a stack of paper towels for 2min, to remove any drops of remaining medium.

The pellet was resuspended gently in 80ml of ice cold transformation buffer, and the cells were harvested by centrifugation at 2500g, 10min at 4°C (Sorvall rotor). The medium was decanted, and the centrifuge bottle again allowed to drain on a stack of paper towels for 2min, to remove any drops of remaining medium. The pellet was resuspended in 20ml of ice cold transformation buffer, 1.5ml of DMSO added, and the bacterial suspension mixed by swirling followed by 10min incubation on ice. 50µl

aliquots of the suspension were dispensed into chilled, sterile microcentrifuge tubes. The competent cells were immediately snap-frozen by immersing the tightly closed tubes in a bath of liquid nitrogen. The tubes were stored at -70°C until needed.

The transformation efficiency of chemically competent cells was measured, and found to be 6.5×10^6 CFU/μg plasmid.

SOB medium (Super Optimal Broth)

	Per 1 liter (g)
Bacto-Tryptone	20
Yeast extract	5
NaCl	0.5

Powders were dissolved in 950ml water, and 10ml of 250mM KCl solution [1.86g of KCl in 100ml deionized water] was added and the pH adjusted pH=7.0 with 5N NaOH. The volume was made to 1liter with deionized water, and the medium sterilized by autoclaving (20min at 15psi). Just before use, 5ml of a sterile solution of 2M $MgCl_2$ [19g of $MgCl_2$ in 90ml of deionized water] was added.

Transformation buffer

a) Preparation of 0.5M PIPES [piperazine-1,2-bis(2-ethanesulfonic acid)] (pH 6.7): 15.1g of PIPES was dissolved in 80ml of pure water (Milli-Q), and the pH of the solution adjusted to 6.7 with 5M KOH. The volume was made to 100ml and the solution sterilized by filtration (0.45μm pore size), and divided into 20ml aliquots and stored at -20°C.

b) Transformation buffer was prepared by dissolving all of the solutes listed below in 800ml of MiliQ water, and then the volume adjusted to 1 liter. The solution was sterilized by filtration (0.45μm pore size), divided into 100ml aliquots and stored at -20°C.

Transformation buffer

Reagent	Amounts pre liter (g)	Final concentration
MnCl$_2$ x 4H$_2$O	10.88	55mM
CaCl$_2$ x 2H$_2$O	2.20	15mM
KCl	18.65	250mM
PIPES (0.5M pH 6.7)	20ml	10mM
H$_2$O	to 1 liter	

5.3.1.2 Transformation of chemically-competent cells

A tube of competent cells was removed from the -70°C freezer, and thawed by holding the tube in the palm of the hand, then transferred immediately to ice for 10min. Up to 25ng of DNA per 50µl competent cells was used, then the tubes were swirled gently several times, and stored on ice for additional 20min. The tube (competent cells+DNA) was transferred to a rack placed in a preheated 42°C circulating water bath for exactly 90 seconds, and immediately incubated on ice for 2min. 800µl of LB was added, and the mixture incubated for 1h, 37°C, 250rpm. 0.5-5µl of transformation mixture was plated on LB medium plates with appropriate antibiotics.

5.3.1.3 Transformation of electro-competent cells

20µl of electro competent cells was thawed on ice and 1-2µl of DNA was added to the cells, and mixed by gently tapping the tube. The cells/DNA mixture was pipetted carefully into a chilled 0.1cm cuvette. The cuvette was gently tapped to remove air bubbles, then electroporated in a Gene Pulser^{TM} (Biorad, Munich Germany) using the following settings: 2.0KV, 200Ω, 25µF. A typical time constants were 3.9msecond. Immediately after electroporation, 1ml of SOC medium (glucose-rich SOB medium containing 20 mM glucose, see 5.3.1.1) was added to the cuvette, and the mixture was transferred into a 15ml snap-cap tube, and incubated at 37°C, 250rpm for 1h. 0.5-5µl

(in the case of plasmid propagation) or 50µl-200µl (in case of ligation reactions) was spread on selective plates.

5.3.2 Transformation of *H. salinarum*

Transformation of *H. salinarum* was performed according to the PEG method (Cline, 1987; Cline et al., 1989) with minor modifications. 1.5ml of culture (OD$_{600}$ 0.6-0.8) was centrifuged (5min 16000g, RT) and the supernatant removed. Residual medium was removed after additional centrifugation (1min, 16000g, RT) using a pipette. The cell pellet was then gently resuspended in 150µl spheroplasting solution (SPH). To produce spheroplasts, 30µl of 0.25M EDTA (pH 8.0 in SPH solution) was added and the mixture was incubated for 5min at RT. Then 10µl of DNA (in SPH solution, ~1µg DNA) was added and the mixture incubated at RT for 5min. 190µl of 60%PEG$_{600}$ (in SPH solution), was added to the tube, followed by rapid mixing of the contents to prevent lysis of the cells due to high local concentration of PEG$_{600}$. After 20min incubation at RT, 1ml of Halo medium+15% sucrose was added to the tube in order to dilute the PEG. The mixture was centrifuged for 2min, 16000g at RT, and the supernatant removed. After an additional centrifugation (1min, 16000g, RT), residual medium was removed by a micropipette. Finally, the pellet was resuspended gently with 1ml of Halo medium+15% sucrose and incubated at 37°C, 250rpm for 12-16h to allow the cells to restore their S-layer.

Most of the plasmids in this study contained the reporter gene *Bga*H (coding for the halophilic galactosidase *Bga*H, from *Haloferax lucentense*, (Holmes and Dyall-Smith, 2000; Patenge et al., 2000), therefore, before spreading the transformation mixture, the plates which contained 10µg/ml Mev, were smeared with X-gal (150µl of 20mg/ml, followed by 30min incubation under the hood). After cell spreading, the plates were incubated at 40°C in a closed transparent box with ~10ml water to prevent the plates from drying out. Usually, transformants became visible after 10-14 days.

Similar protocol was used for *N. pharaonis* with the following changes: the SPH buffer contained 2M Betaine, cells were recovered in M205 medium with either 15% sucrose or 2M Betaine, and cells were plated on M205 plates with 10µg/ml Mev.

Unbuffered SPH solution

Reagent	Amounts (per 100ml)	Final concentration
NaCl	11.7g	2M
KCl	2.5ml of 1M	25mM
Sucrose	15g	15% (w/v)
H$_2$O	to 100ml	

The solution was sterilized by filtration and stored at 4°C until use.

5.3.3 Transformation of *N. pharaonis*

5.3.3.1 PEG method

The PEG method was similar to the one described for *H. salinarum* (section 5.3.2) but with the with the following changes for *N. pharaonis*: the SPH buffer contained 2M Betaine; cells were recovered in M205 medium with either 15% sucrose or 2M Betaine; and cells were plated on M205 plates with 10µg/ml Mev.

5.3.3.2 Electroporation

Cells grown in M205 medium (OD$_{600nm}$ 0.5-0.7), were centrifuged (2min, 10000g, RT) and the supernatant was discarded. The cells were centrifuged again to remove residual medium and the pellet was resuspended with 4.3M Betaine (Fluka, cat#14300) in 50mM Tris HCl pH 8.7. Plasmid DNA was added to the cell suspensions and the tubes were incubated in RT for 10min. The mixture was electroporated (25µF, 1.5KV, different resistance, see appendix table #20). Cells were recovered in 800µl M205 medium with 15% sucrose or 2M Betaine (14-16h, 37°C, 100rpm). The cells were centrifuge and the pellet resuspended in 100µl supernatant, then the cells were plated on M205 plates+10µg/ml Mev.

5.3.3.3 Gun particle

1.5-3x10^8 cells/ml of *E. coli*, *N. pharaonis* and *H. salinarum* were plated on LB, halo medium and M205 medium plates, respectively, and then incubated at 37°C (14 days for haloarchaea and o.n for *E. coli*).

The microcarriers (either M5 Tungsten or gold particles) were prepared one day before the bombardment, and coated with DNA on the day of the bombardment. The microcarriers were coated according to Bio-Rad recommended protocol ((Kikkert, 1993). 500µg of microcarriers were used per bombardment, based on the method of Sanford et al., (Sanford et al., 1993).

0.625µg DNA was use for each bombardment event. The Biolistic®PDS-100/He device (Bio-Rad) was used (Kikkert, 1993), using helium pressure was 1100psi, gap distance (i.e. the distance between the helium source and the flying disk) of 1cm, and target distance (i.e. the distance between the partical launch site and target cell) of 6cm. Cells were stamped to selective plates (10µg/ml Mev) and colonies were visible after 8 weeks.

5.3.4 Preparation of genomic DNA from *H. salinarum*

300µl of culture in the stationary phase (~ OD_{600} =1) was centrifuged for 2min 16000g at RT. The supernatant was discarded and the pellet recentrifuged (1min, 16000g, RT) to remove residual medium. Cells were lysed with 200µl sterile autoclaved H_2O (bi- distilled). The crude DNA, without further purification, was stored at 4°C.

5.3.5 Isolation of plasmid DNA from *E. coli*

Single colonies of *E.coli* transformants were picked from LB agar plates containing antibiotic. Cells were grown overnight (37°C, 250rpm) in 4.5ml LB medium with antibiotics and harvested by centrifugation (15min, 3200rpm, 4°C). Plasmid isolation was performed with QIAprep Spin Miniprep Kit (Qiagen), according to the manufacturer's detailed guide. Elution was done with 50µl of EB buffer (10mM Tris HCl, pH 8.5).

5.3.6 Agarose gel electrophoresis

Horizontal agarose gels slabs were used for analysis and separation of PCR products. The gel was prepared by melting the agarose with TAEx1 (running buffer). After the gel had solidified, the DNA samples were mixed with sample buffer and pipeted into the wells. For Southern blots, 1.5µg of genomic DNA per digest was used for each well, while 0.1µg of digested plasmid (used as controls for the specificity of

the labeled probes) were loaded into each well. 1Kb DNA MW ladder (Life Technologies Inc) was used as the size standard. The DNA pattern was visualized by staining the gel for 15min in 0.1mg/ml ethidium bromide followed by washing for 15min in H_2O and illumination by UV light at 302nm.

50 x TAE

Reagent	Amounts (per 1 liter)	Final concentration
Tris base	242g	2M
Acetic acid	57.1ml	0.997M
EDTA (disodium salt)	18.61g	50mM
H_2O	to 1liter	
The pH was adjusted to 8.0 using acetic acid		

Sample buffer

Reagent	Amounts/ volumes	Final concentration
Sucrose	50g	50%
0.5M EDTA (disodium salt)	20µl	0.1mM
Bromophenolblue	10mg	0.1% (w/v)
10% SDS	1ml	0.1% (w/v)
H_2O	to 100ml	
pH was adjusted to 7.0 with NaOH		

5.3.7 Isolation of DNA from preparative agarose gels

Preparative gels (1-2.5%), were performed for purification of PCR products and the products of endonuclease digests. The isolation of DNA fragments was carried out using the Wizard®SV Gel and PCR Clean-Up System kit (Promega), according to the protocol of the manufacture.

5.3.8 Determination of DNA concentration

DNA concentration was determined by analytic agarose gel electrophoresis. A 1µl volume of each DNA sample was loaded onto a gel that included 10µl of an appropriate DNA ladder. The concentration of the individual bands in the ladder allowed estimation of the DNA concentration. Alternatively, the concentration was determined using a NanoDrop®ND-1000 spectrophotometer (NanoDrop) using 1-5µl sample, measuring the absorbance at 260 and 280nm. The concentration of the DNA sample was calculated using an extinction coefficient of 50ng/ml (dsDNA) and 33ng/ml for single-stranded DNA (ssDNA).

5.3.9 Sequencing of DNA

An ABI PRISM BigDye Terminator Cycle Sequencing Kit (Applied Bioscience, Foster City, CA, USA) was used for sequencing. The Terminator Ready Reaction Mix contained labeled ddNTP's (BigDye terminator), dNTP's mixture, AmpliTaq DNA polymerase and $MgCl_2$ in Tris HCl pH 9.0. Usually 2-3µl of mini-prep DNA was added as template before the reaction mixture was applied to the thermo-cycler. The product of the sequencing reaction was diluted 1:2 with 10µl H_20 (bi-distilled) or 0.4% SDS and sent to the in-house core facility to generate the DNA sequences chromatograms. These were analysed using the program BioEdit (Hall, 1999).

Sequencing reaction

1µl reaction mix (enzyme, dNTP's)

1µl buffer

2µl 5M Betaine

2-3µl DNA after mini prep

2µl 5pmole/µl primer

to 10µl H_2O

Thermocycler steps		
96°C	DNA melting	30 cycles
60°C	Annealing and elongation	
4°C	Pause	

5.3.10   Polymerase chain reaction (PCR)

Depending on the aim of the PCR, either Taq polymerase or the proof-reading polymerase LA taq were used. These were added to the reaction mixture immediately before starting the PCR.

Taq polymerase in colony PCR

Transformant colonies were screened for gene deletions by direct colony PCR using primers targeting the flanking regions of the particular genes. The PCR mixture was prepared as a master mix, and then 10μl volumes were transferred into individual PCR tubes. Colonies were picked by sterile toothpick, restreaked on plates and subsequently rolled in the reaction mix (10μl).

Reaction mixture	
Thermopol buffer (x10)	1μl
Forward/reverse primer (5pmole/μl)	1μl
dNTP's (1.25mM each)	0.8μl
DMSO	0.4μl
Taq polymerase (5U/μl)	0.2μl
SBx5 for colony PCR	2μl
DNA from 1 colony	
H$_2$O (bidis, autoclaved)	to 10μl

Stock SBx5 *
5.8ml Glycerol (stock 86% V/V)
3.8ml ddH$_2$O
100μl of Bromophenol blue (saturated solution)
100μl xylen cyanol (saturated solution)
200μl 1M TrisHCl pH 7.2

* The colony PCR procedure was used for initial screening. The presence of the dyes did not interfere with the amplification. Candidates for deletion mutants were further checked using LA Taq polymerase.

PCR Temperature profile		
step	Temperature & Time	cycles
Initial denaturation	94°C, 4min	x1
Denaturation	94°C, 30sec	x 39
Annealing	Tm+4°C, 40sec	
Elongation	72°C, X min**	x1
Final elongation	72°C, 10min	

** The times needed for the elongation step were adjusted according to the product length, using the formula: 1min per 1Kbp.

LA taq polymerase

Reaction mixture	
LA Taq buffer II (x10) (Mg^{2+} plus)	5µl
Forward/reverse primer (5pmole/µl)	5µl each
dNTP's (1.25mM each)	8µl
DMSO	1µl
Taq polymerase (5U/µl)	0.25µl
Genomic DNA (5.3.3)	1-2µl
H$_2$O (bidis, autoclaved)	to 50µl

PCR Temperature profile		
step	Temperature & Time	cycles
Initial denaturation	95°C, 6min	
Denaturation	95°C, 1min	x 24
Annealing	Tm+4°C, 40sec	
Elongation	72°C, X min**	
Final elongation	72°C, 7min	

** The times needed for the elongation step were adjusted according to the product length, using the formula: 1min per 1Kbp.

5.3.10.1 Ligation of DNA fragments

Ligation reactions were performed with ~100ng DNA, 1U of T4 DNA ligase (Invitrogen), 2µl of reaction buffer (x5) in total volume of 10µl. Before the reaction, the concentrations of the insert and vector were determined either by agarose gel electrophoresis or a Nanodrop spectrophotometer. Inserts and vectors with compatible cohesive ends were usually incubated in molar ratios of 4:1 to 2:1 overnight at 10°C.

5.3.10.2 Digestion of DNA by restriction endonucleases

A typical digestion reaction (40µl) contained 4µl (40-80U) of restriction enzyme, and 4µl of the corresponding 10x reaction buffer (as recommended by the manufacture). The concentrations of DNA were determined either by agarose gel electrophoresis or by Nanodrop spectrophotometer, to 10µg of vector and 0.12-0.38µg of insert.

5.3.10.3 ORFs OE1472F, OE1475F, OE1477R, and OE2019F from *H. salinarum*

For homologous overexpression, the genes OE1472F, OE1475F and OE1477R (http://halolex.mpg.de) were amplified by a one step PCR, using the proofreading DNA polymerase LA-Taq, genomic DNA from *H. salinarum strain* R1 as a template, and the primers listed in appendix table #18, (pMG550, pMG650, and pMG750, respectively). The entire ORF's were amplified, but without a stop codon, then digested with *HindIII* and *BamHI* and subcloned into the *E. coli-H. salinarum* shuttle vector pBPH-M. The pBPH-M vector contains an inducible bop promoter at the 5΄ end of the cloned gene, and a 7xHis C-terminal purification tag followed by a stop codon. Selection is via a Mev^R marker, and part of OE3349F is present to allow integration into the chromosome.

For heterologous overexpression, the genes OE1472F, OE1475F, OE1477R and OE2019F were amplified by PCR using the proofreading DNA polymerase LA-Taq, genomic DNA from *H. salinarum strain* R1 as a template, and the primers listed in appendix table#18, (pMG560, pMG660, pMG760, and pMG860, respectively). The genes (without the stop codon) were amplified, digested with *NdeI* and *XhoI* and subcloned into a similarly treated pET22b(+) expression vector.

5.3.10.4 Plasmids for gene deletion

Single, in-frame deletions were achieved by homologous recombination using "suicide plasmids" carrying the flanking regions of each gene. The upstream (US) and downstream (DS) regions of the target genes were amplified separately by PCR (~500bp each) and specific restriction sites were introduced at each end. *BamHI* and *PstI* sites were introduced at the 3' end and 5' end (respectively) of the US region.

PstI and *HindIII* sites were introduced at the 3' end and 5' end (respectively) of the DS region, as seen in the figure #13A- B below.
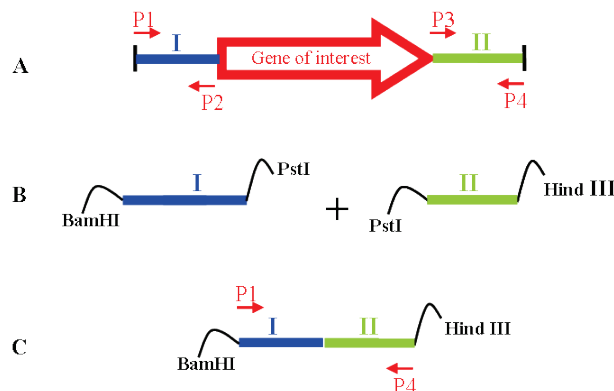
Figure # 13: The construction of gene deletion vector. The US (I) and DS (II) regions of a specific gene (**A**) were amplified with restriction sites using primers P1+P2, and P3+P4, respectively (**B**), then US (I) and DS (II) regions were ligated after cleavage with *Pst*I (**C**).

The US and DS amplified sequences were purified from a 1% preparative agarose gels (figure #14A), then cleaved with *PstI* and ligated to each other using T4 DNA ligase. The overnight ligation was followed by PCR using the ligated US+DS regions as template with primers P1 and P4.
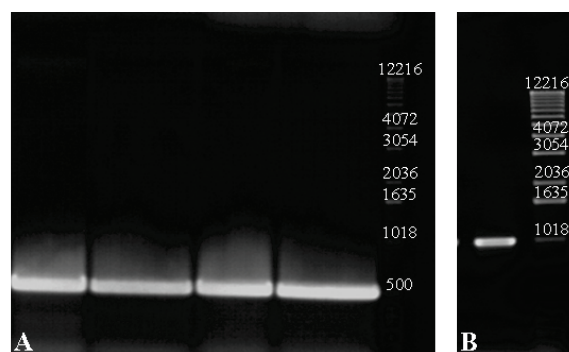
Figure # 14: **A** - A preparative agarose gel with US (lane 1,2) and DS (lane 3,4) regions after amplification by PCR. **B**- Analytical 1% agarose gel of ligated US+DS sequences after PCR amplification using P1 and P4 as primers. The positions of MW size marker bands are indicated at the right of each panel (in bp). 1% agarose gel.

The PCR amplified US+DS regions (figure #14B) of each gene were purified, and then ligated into pCR 2.1-TOPO vector. Positive candidates were chosen according to the size of the inserts obtained after the cleavage with *EcoRI*. A correct construct will give fragments of 3.9Kbp, and ~1Kbp, as seen in figure #15 below.
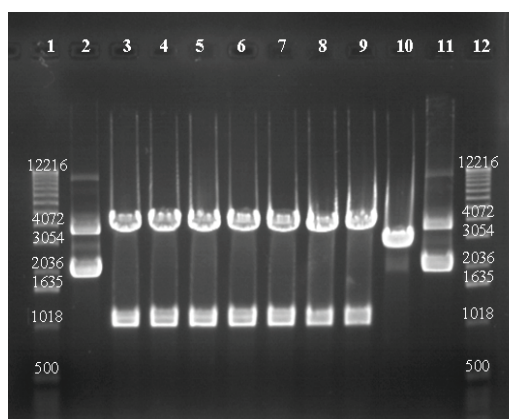
Figure # 15: *EcoRI*-digested plasmids from independently picked transformants after TOPO cloning (lanes 2-11). Correct clones contain 3.9Kbp, and 1Kbp fragments. MW markers are shown in lanes 1 and 12 (in bp), 1% agarose gel.

The constructs shown in lanes 3-9 (figure #15) were sequenced and confirmed to have the correct sequence. One of these was selected, cleaved by *BamHI* and *HindIII*, separated on a 1% preparative agarose gel, and the 1031bp cloned fragment purified.

The plasmid pMKK100, which was used as backbone, was cleaved as well with *BamHI* and *HindIII* and dephosphorylated. The purified, cleaved US+DS regions were ligated into the dephosphorylated and cleaved pMKK100 vector. The ligated DNA was introduced into competent *E. coli* XL-10 cells. Transformed colonies carrying plasmids containing the US+DS regions were identified by electrophoresis of plasmid extracts, similar to the example shown in figure #15. The correct construct was chosen according to the cleavage pattern of the purified plasmid with *BamHI* and *HindIII*. Correct constructs show a ~1Kbp fragment (the insert) and a ~7.1Kbp fragment (the vector pMKK100).

The plasmids used for in-frame ORF deletions, contained the US+DS regions of OE1472F, OE1475F and OE1477R, respectively. In addition, each plasmid contained Amp^R, and Mev^R regions as selection markers, and the *Bga*H gene as the reporter gene for blue/ red screening, as seen in the figure #16.
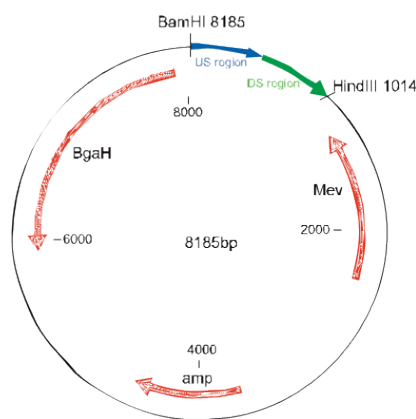
Figure # 16: Diagrammatic map illustrating the plasmids pMG500, pMG600, and pMG700. The plasmids were used to generate single in-frame deletions of OE1472F, OE1475F and OE1477R, respectively.

5.3.10.4.1   *Blue/red screening of transformant colonies*

The gene deletion plasmids directed at ORFs OE1472F, OE1475F and OE1477R, were introduced into *H. salinarum* R1 cells using the PEG method (materials and methods section 5.3.2), and plated on complex medium containing mevinolin (Mev) and X-gal. A successful first cross-over event would produce blue colored transformants colonies. During the first cross-over, the complete plasmid would integrate into the genome via homologous recombination between either the upstream region (US) or the downstream region (DS) on the plasmid and the corresponding region in the genome. Thus, positive transformants would be resistant to mevinolin and develop a blue color on plates containing X-gal.

Blue colored transformants were then transferred into medium without Mev, which allowed a second cross-over event to take place, excising the plasmid together with its Mevinolin-resistance and β-galactosidase genes from the genome. Even after several generations, second cross-over events occur only in small fraction of cells. On plates containing X-gal without Mevinolin, cells which had undergone a second cross-over were identified based on their red color (due to bacteriorhodopsin and bacterioruberin) in contrast to blue colonies that still contained the plasmid. As the second cross-over can occur in two ways (figure #17), deletion of the target gene is expected to occur in only ≤50% of the red colonies.

While two successive cross-over events between the same homologous regions (i.e. two US or two DS cross-over events) would lead to restoration of the parental genotype, an US cross-over followed by a DS cross-over (or vice versa) would lead to

deletion of the target gene. After the second cross-over event, the target gene bounded by the flanking regions will be eliminated and replaced by the direct fusion of the flanking regions as present on the plasmid. Gene deletions were verified by PCR analysis and Southern blot.
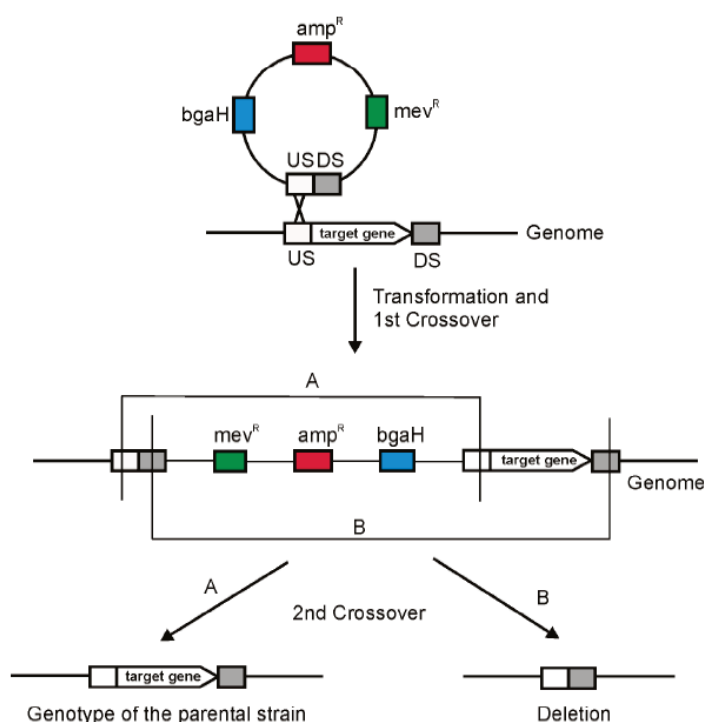
Figure # 17: Schematic illustration of the first and second cross-over events. The plasmid contains the fused upstream (**US**) and downstream (**DS**) regions of the target gene. The first cross-over is exemplified by a cross-over of the US region in the plasmid and US in the genome, but can also occur between the DS regions. If the second cross over occurs between the US region again (**A**), the genotype of the parent strain is restored. If the second cross-over occurs between the DS regions (**B**), the target gene is deleted. Diagram prepared by Dr. Wilfried Staudinger

5.3.10.5 Southern blot analysis

1.5µg of chromosomal DNA from WT and gene deletion candidates, were cleaved with a restriction enzyme, and the DNA fragments separated on a 1% agarose gel. The gel was vacuum-blotted onto a nylon membrane (Hybond-N, Amersham), using essentially the method of Southern (Southern, 1975). For size determination of the DNA fragments, digoxigenin-labeled DNA molecular marker VIII (Roche diagnostic, Mannheim, Germany) was included on the gel. All gel pretreatment steps, such depurination, denaturation and neutralization were done as recommended by Southern

(Southern, 1975). After UV-crosslinking the DNA to the membrane, all further steps, including prehybridzation, hybridization, washing and chemiluminesecent detection with CSPD, were performed according to the manufacturers recommendations (Roche diagnostic, Mannheim, Germany).

Probes for hybridization were generated by incorporation of digoxigenin-labeled dUTP in a PCR, using DIG labeling mix [Plus] (Roche). Two 500bp labeled probes were designed to hybridize to the up-stream region of the gene and to the internal region of the corresponding gene.

5.3.11 Microarrays

5.3.11.1 Microarray design

The microarray was designed as described by Twellmeyer et al., (Twellmeyer et al., 2007 119). Each γ - amino-silane coated CMT-GAPS-II glass slide was spotted with five replicates of 2774 DNA probes, manufactured as described by Twellmeyer et al., (Twellmeyer et al., 2007).

5.3.11.2 Isolation of total RNA

9ml cultures of *H. salinarum* R1, grown in different media and to different OD_{600nm}, were harvested by centrifugation (5min, 8000rpm, 4°C). The pellet was resuspended with 6ml of peqGOLD RNApure (peqLAB Biotechnology, Erlangen, Germany), and the mixture was frozen in liquid nitrogen and stored at -80°C. Total RNA was extracted using chloroform/phenol extractions, washing with 75% EtOH and finally the RNA dissolved in 100μl DEPC-treated water. The RNA was DNase treated with DNAase kit DNA-free (Ambion, Huntington, United Kingdom) and the absence of DNA was confirmed by a PCR using LA Taq, primers P1 and P2 for ORF OE1472F (500bp, appendix table #18), 35 cycles and RNA before and after digestion as templates (figure #18A). The quality of the DNA-free total RNA was assessed by 1% denaturating agarose gel in TBE buffer. Intact total RNA run on a denaturating gel will have sharp, clear 23S and 16S rRNA bands (figure #15B). The concentration of intact total RNA was determined using NanoDrop$^®$ND-1000 spectrophotometer (NanoDrop). The concentration of the RNA sample was calculated using the formula $1OD_{260nm}=40ng/ml$ ssRNA.
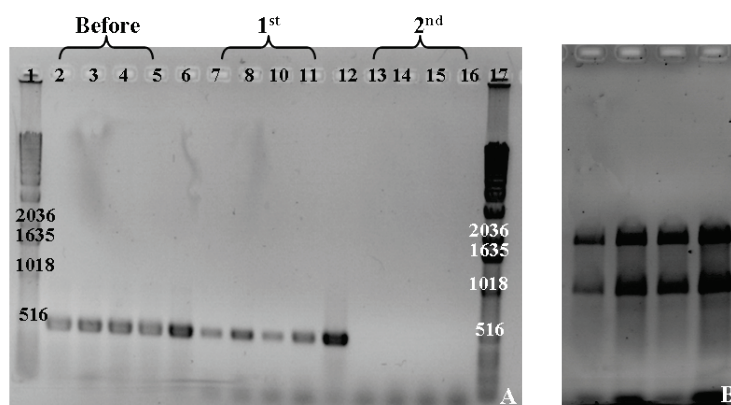
Figure # 18: Examples of RNA preparations. **A-** 1% agarose gel in TAE buffer with 4 isolated RNA samples before and after 1st and 2nd DNA digest. Lanes 6 and 12 are positive controls, Chromosomal DNA of R1. The MW markers are shown in lanes 1 and 17 (in bp). **B-** 1% denaturing agarose gel in TBE buffer with 4 RNA samples after 2nd DNAse digestion.

10xTBE buffer

Reagent	Amounts (per 1liter)	Final concentration
Tris base	108g	0.89mM
Boric acid	55g	0.889mM
0.5M EDTA pH=8.0 in DEPC water	40ml	2mM
DEPC-water	to 1liter	

1% denaturing agarose gel in TBE buffer

To 100ml 1% agarose gel in TBE buffer (after melting and cooling down) add 0.236g guanidine thiocyanate (MW=118g/mole, final concentration, 20mM).

RNA samples

RNA samples were diluted 1/10 with DEPC-treated water, the SB was added, and the mixture was heated to 75°C for 5 min and immediately loaded on the gel and electrophoresed.

5.3.11.3 Microarray analysis

Adapted cultures of *H. salinarum* were grown in synthetic medium with or without AroAAs, and cells were harvested by centrifugation at OD_{600nm}=0.2 and 0.58, respectively. Additionally, the deletion strain ΔOE1477R and WT *H. salinarum* were

grown in synthetic medium supplemented with 1.1mM AroAAs+1.1mM Shikimate. RNA was isolated from cell samples as described above.

3μg of total RNA was reversed transcribed into cy5 or Cy3-labled cDNA using a CyScribe first-strand cDNA synthesis kit with random nonamer primers and Cy5/Cys3-dUTP (Amersham Biosciences, Freiburg, Germany). RNA was removed by alkaline hydrolysis, and the cDNA purified and concentrated as described by Zaigler et al., (Zaigler et al., 2003). After prehybridization, labeled cDNA was pipetted onto the microarray which was then sealed in hybridization chamber for overnight at 64°C. The microarrays were washed and dried by centrifugation (5min, 1500rpm), as described by (Zaigler et al., 2003). Labeled cDNA from cells grown with AroAAs was hybridized with labeled cDNA from cells grown without AroAAs (in two different OD's). Additionally, labeled cDNA from R1 grown with AroAAs and shikimate was hybridized with labeled cDNA from mutant ΔOE1477R grown in the same medium (in two different OD's).

Hybbridized microarray slides were scanned for Cy5 and Cy3 florescence signals using a GenePix 4000B scanner (Biozym Seientific GmbH, Hessisch Oldendrof, Germany). Image processing was done with GenePix Pro 6 (Biozym Scientific GmbH) and the florescence values were processed in the R environment using a program written by G. Welzl (Twellmeyer et al., 2007). Significant gene regulations were identified by a one-sample t-test as implemented in the TIGR MultiExperiment viewer (Saeed et al., 2003). A gene was considered to be significantly regulated if the regulation factor was three fold or higher, and the calculated *P* value of the regulated gene was *P* <0.001.

5.3.12   RT-PCR

1μg total RNA was reversed transcribed with 400ng random hexamer primer and 50U of Reverse-iT^{TM} RTase Blend (ABgene House, UK). Quantitative PCRs were performed in Bio-Rad iCycler MyiQ™ single-color Real-time PCR detection system using a SYBR green PCR master mix kit (from AB Applied biosystems). The final reaction volume was 25μl with 0.5μl of the reverse transcription reaction serving as template. Primers were designed with Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) and were used at a final

concentration of 0.92 pmole/µl (see appendix table #18 for the sequence of the primers).

Transcript level differences were calculated by a relative quantification approach using an internal standard gene, OE1160R, which encodes ribosomal protein L10.eR. For all calculations, the mean-C_t of 2 replicate reactions per primer pair was used.

5.4 Protein Methods

5.4.1 SDS-PAGE of proteins

The SDS-polyacrylamide gel electrophoresis system used in this study essentially followed the method of Laemmli (Laemmli, 1970; Schagger and von Jagow, 1987). Protein samples, at a concentration of 10µg/ml (determined by BCA method), were mixed with sample buffer, boiled for 10min at 95°C, followed by 5min centrifugation (16000g, RT), and the supernatant loaded into the wells. In addition to this, *E.coli* cell lysates were sonicated (5min in sonication bath) before the centrifugation. The protein samples were loaded on ready-to-use 4-12% NuPAGE®Novex®Bis-Tris Mini Gels (Invitrogen), with running buffer (MES SDS running buffer) prepared according to the manufacturer's recommendation. Running conditions: 10min at 100V constant voltage, followed by ~55min at 200V constant voltage. 5µl of prestained protein marker (PageRuler, Fermentas) was also loaded. Electrophoresis was stopped when the lower band in the marker reached the bottom of the gel. The gel was stained for 25min with Coomassie Blue and protein bands were visualized after the gel was subsequently treated with destaining solution.

5xSDS-PAGE sample buffer

0.5M Tris-HCl pH 6.8
10% glycerol
2.3% (w/v) SDS
5% DTT
0.1% bromophenyl blue
Staining solution

Coomassie blue R-250

Destaining solution (1liter)
125ml Isopropanol

100ml Acetic acid

775ml RO water

5.4.2 Western blot analysis

Protein samples were separated on SDS-PAGE as described above but with two differences: 1µg of protein was separated, and only 3µl of the prestained marker was used. The protein bands from the polyacrylamide gel were transferred onto a PVDF membrane (Immobilon^{TM}-P, Millipore, 0.45µm pore size) for immunodetection. The transfer method was based on the semidry electrobloting system using an XCell II Blot Module (Invitrogen). Transfers were performed for 2h, RT, 170mA, 25V. The membrane was incubated with blocking reagent overnight at 4°C.

His-tagged proteins were detected using the ECL Plus Western blot Detection system (GE Healthcare). The membrane was incubated with mouse anti-His Tag monoclonal IgG1 antibody (diluted 1:2500 in 3%BSA in TBST) for 1h at RT with slow shaking. Unbound antibody was removed by washing (3x20ml TBST) followed by a second series of washes (3x25ml 3%BSA in TBST, for 10min each at RT). Then the membrane was incubated with horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (diluted 1:12500 in 3%BSA in TBST) for 1h at RT. Unbound secondary antibody was removed by washing in 3x20ml TBST followed by 3x25ml TBST, for 10min each at RT. The membrane was treated with 1ml of detection reagent mixture for 1min. After draining off the excess reagent, the membrane was covered with plastic wrap and exposed to x-ray film for 30sec up to 2min.

Semi dry blotting buffer:

25mM Tris (3.03g)

192mM Glycine (14.4g)

20% (v/v) methanol

To 1liter with pure water

The pH of the solution was 8.3.

10xTBST buffer, pH 7.4

100mM TrisHCl

1.5M NaCl

2% (w/v) Triton X100

Blocking solution

3% (w/v) BSA in 200ml TBST (for one blot)

5.4.3 Determination of protein concentration

The concentration of the protein was analyzed by the BCA Protein Assay according to the protocol of the manufacture (Pierce), using Bovine serum albumin (BSA) to generate a standard curve.

5.4.4 EasyXpress protein Synthesis Kit (Qiagen)

Following the manufacturer's recommendations, 0.6µg of each plasmid was mixed with *E. coli* extract and the reaction mixture incubated at 37°C for 1h, shaking at 500rpm. The in-vitro expression was analyzed both by SDS-PAGE and Western blot.

5.4.5 Expression and purification of OE1472F, OE1475F and OE1477R from *R1*

E. coli DH3-Rosetta (Invitrogen) was used as host strain for the expression plasmids: OE1472F-His $_{C-6xhis}$/ pET22b(+), OE1477R-His $_{C-6xhis}$/ pET22b(+), and OE2019F-His $_{C-6xhis}$/ pET22b(+), respectively. Cells from single colonies were inoculated into 50ml of LB with antibiotics and incubated overnight, 37°C, 250rpm. These were used to inoculate 1liter LB (with antibiotics) in 5liter flasks, giving an initial cell density of $0.1OD_{600nm}$. The cells were induced with 1mM IPTG for 3h (37°C, 200rpm), after the $OD6_{00nm}$ reached ~ 0.6. The cells were harvested by centrifugation (4000rpm, 40min, and 4°C). The cell pellet from each 2 liter culture was stored at -70°C until further use.

Before purification on NiNTa affinity column, the cell pellet was thawed on ice and resuspended with 50ml NP buffer containing DNAase. PMSF (Phenylmethylsulfonyl fluoride) in 2-propanol was added to the suspension to a final concentration of 0.2mM and the cells were lysed using a French press (2 passages, 1000psi). After centrifugation at 40000rpm, 40min, 4°C (Ti50), the resulting supernatant was diluted to a final volume of 250ml with NP buffer. The NiNTa column (HisTrp^{TM}HP) was equilibrated with NP buffer, and the bound fraction eluted with a step gradient between NPI_{200} and NPI_{250}.

For homologous overexpression of OE1472F, OE1475F and OE1477R, the recombinant vectors OE1472F-His $_{C-7xhis}$/pBPH-M, OE1475F-His $_{C-7xhis}$/pBPH-M and OE1477R-His $_{C-7xhis}$/pBPH-M, were introduced into *H. salinarum* R1 cells using the

PEG method. The cells were plated on Halo medium plates containing 10μg/ml Mev, and transformants were grown up in 35ml Halo medium +10μg/ml Mev. These cultures were used to prepare total DNA, which was used to identify colonies that contained the plasmid. 1liter of halo medium+10μg/ml Mev was inoculated, and the cells were harvested by centrifugation (7200rpm, 20min, 4°C). The cell pellet was resuspended in 3ml 3.4M KCl+10mM NaCl+25mM Tris HCl pH 7.2. The cells were lysed by sonication (till ~90% of the cells were broken), DNAase was added and the solution was mixed 30min at RT, shaking. After centrifugation (45000rpm, 1h, 12°C, TI50), the concentration of total proteins in the cell-free extract was determined by the BCA protein assay (Pierce).

5.4.6 N-terminal sequencing

N-terminal sequencing by the Edman degradation method was performed by the core facility in the Institute. Proteins were either sequenced from solution of from PVDF membranes as solid support.

5.4.7 Amino acid analysis

200μl samples were taken from cultures of R1 and deletion strains cultures grown in various media, at different time points (~5 points per growing curve). The samples were centrifuged (5min, 16000g, RT), and the supernatant from each time point was divided into two tubes. The supernatants were stored at -20°C until all samples were collected. Amino acid analysis was done by Mr. Wolfgang Strasshofer from the core facility, using an Amino Acid Analyzer (Biotronik LC3000). The data in figures #31, 38 and 41 represents aggregate data from 3-4 cultures that were prepared and processed separately.

5.4.8 Activity assays

5.4.8.1 Aldolase activity using a coupled enzyme assay

The activity of commercial Aldolase (Sigma) was measured using the coupled enzyme assay suggested by the manufacturer. The coupled assay involved 3 enzymes: **Aldolase** (from rabbit muscle) which cleaves F-1,6-P into DHAP and GAP; **Trioesphosphate Isomerase** (TIM, EC 5.3.1.1), which is responsible for the

isomerasation of GAP to DHAP; and **Glyerophosphate Dehydrogenase** (GDH, EC1.1.1.8), which reacts with DHAP in the presence of NADH as cofactor, as seen below:

$$F\text{-}1,6\text{-}P \; + \; H_2O \; \rightarrow \; GAP \; + \; DHAP$$
$$GAP \; \rightarrow \; DHAP$$
$$2 \, DHAP + 2NADH \; \rightarrow \; 2 \, Glycerophosphate + \, 2NAD^+$$
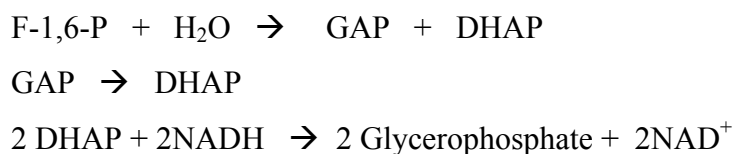
In a 1.0 ml reaction mix, the final concentrations were 84mM Tris, 1.92mM F-1,6-P, 0.15mM NADH, 1.67U GDH, 16.7U TIM, and 14.84mU Aldolase. The activity was measured at room temperature following the decrease in the absorbance (340nm) when Aldolase was added to the reaction mixture. The slope of the graph ($\Delta A_{(340nm)}$/min) was used for calculating the specific activity of the Aldolase. The extinction coefficient of NADH used for the calculation was $6.22mM^{-1} \times cm^{-1}$. One unit will convert 1.0μmole of F-1,6-P to DHAP and GAP per minute at pH 7.2 at RT.

5.4.8.2 Colorimetric assay of Aldolase

Aldolase activity was assayed by the colorimetric procedure of Silbley and Lehninger (Sibley, 1949) with some modifications. Activity was assayed in a 0.5ml reaction mixture in the presence of 2M KCl with 78.7mM Tris HCl (pH=7.2), 1mM Cysteine, 0.5mM $FeCl_2$, 56μM hydrazine sulfate (in 50mM Tris HCl pH=7.2), and 1.2μM F-1,6-P, as substrate. The reaction (37°C, 30min) was terminated with 2ml of 10% TCA, and the chromogens developed were read at 540nm. Blank tubes contained 1.2μM F-1,6-P, which was added after the addition of TCA. Aldolase activity was expressed as μmoles DHAP formed/mg protein/30min at 37°C.

5.4.8.3 3-Dehydroquinate dehydratase

The activity of 3-dehydroquinate dehydratase was assayed by monitoring the formation of 3-dehydroshikimate (DHS) at 234nm ($\varepsilon = 12 \times 10^3 \, M^{-1}cm^{-1}$) at RT. The standard assay mixture (1ml) contained 2.625M KCl+43.75mM Tris-HCl pH=7.2, 10mM 3-dehydroquinate (DHQ, synthesized by Mr. Jurgen Musiol) (Bottomley et al., 1996).

5.4.8.4 LC-MS

The enzyme activity of recombinant proteins was measured after derivatization of the reaction products and subsequent analysis by LC-MS. ~51μg enzyme in 3M KCl +50mM Tris HCl pH=7.2 were mixed with 2.5mM F-1,6-P or 2.5mM DKFP and 2.5mM ASA, in a total volume of 60μl, and incubated at 37°C for 30min or o.n. For the derivatization, 2mM o-(-4-nitrobenzyl)-hydroxylamin-hydrochlorid (NBHA) in MeOH was added, in a final volume of 150μl. The mixture was centrifuged (5min, 16000g) and the supernatant was incubated for 1h at 60°C.

Up to 10μl of the derivatized mixture was loaded on a C18 reverse phase column (25°C, flow-250μl/min). Derivatized reactants and products were eluted with ACN (2min 10% ACN, 28min gradient from 10%-30%, 5min 70% ACN, 2min 90% ACN, 11min 10% ACN), and monitoring was at 210nm and 254nm.

The activity of commercial Aldolase was measured using LC-MS. 16.7U/ml TIM, 1.92mM F-1,6-P, 14.84Mu/ml Aldolase in 84mM Tris-HCl pH=7.2 in total volume of 60μl were incubated at 37°C, for 30min. The derivatization with NBHA, and the elution was done as described above. The samples were measured by Mrs. Lissy Weyher-Stingl

5.4.8.5 GC-MS

A two-step derivatization of the samples was used to aid the identification of reactants and products. The first step was methoximation of the carbonyl group with freshly prepared methoxyamin in pyridin. The second step was derivatization with TMS or TBS (Macherey Nagel) to form the corresponding derivatives.

Briefly, the desired volume of sample was dried down using a Spedi-vac, and then resuspended in 10μl of freshly prepared methoxyamin (20mg in 1ml pyridine, MW=83.52g/mole). The mixture was stirred for 90min at 30°C and 900rpm. 90μl of TMS or TBS were added to each sample, and the tubes were mixed for a further 30min, 900rpm at 37°C (for TMS) or 75°C (for TBS). 30μl of the derivatized sample was transferred to a glass vial. 1μl of the derivatized sample was injected to the GC-MS, eluting with step gradient of temperature. GC-MS was carried out using a TermoQuest gas chromatograph (TraceGC) equipped with a Rtx®-5Sil MS W/Integra-Guard ™ column (Restek, 30mx0.25mmx0.25μm) that was directly

connected to a ion trap mass spectrometer (Polaris Q from Finnigen). The injection volume was 1µl at a carrier gas flow of 1.2ml/min helium with a splitless injector. The initial oven temperature of 60°C was maintained for 2min and then raised to 120°C at 8°C/min. Several temperature gradient were used including raising the temperature to 155°C at 6°C/min and maintaining the temperature for 5min, than raising the temperature to 250°C at 10°C/min for 12min, and finally raising the temperature to 300°C at 15°C/min and maintaining the temperature for 3min. Mass spectra were analyzed in the rang of 60-750 atom mass units. MS data were processed using the program Xcalibure.

The samples were both derivatized and measured by Mrs. Sigrid Bauer.

5.5 Synthesis

Numerous compounds used in this thesis were synthesized by Mr. Jurgen Musiol. Among them were: **ASA** (L-aspartate semialdehyde), **DKFP** (6-deoxy-5-ketofructose 1-phosphate), **DHQ** (3- dehydroquinate) and **HPAP**- hydroxpyruvaldehyde phosphate.

6 Results

6.1 Mutations of key ORFs in AroAAs biosynthesis pathway

In the proposed pathway of aromatic amino acids (AroAAs) synthesis, seen in figure #19, the first three steps are catalyzed by enzymes encoded by ORF's OE1472F, OE1475F, and OE1477R, respectively. Deletion of these genes should convert *H. salinarum* to aromatic amino acid auxotrophy, and may lead to the build up of precursor metabolites. This was tested by the construction and analysis of mutant strains deleted in each of the three genes. In the first step, the phenotype of the WT strain was tested and then plasmids were engineered to enable in-frame deletion or insertion mutants of the target genes.
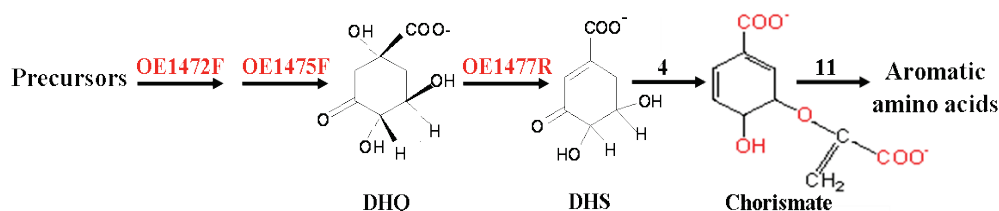
Figure # 19: The proposed pathway of AroAAs biosynthesis in *H. salinarum*. The first three steps of the pathway are shown along with the *H. salinarum* ORFs (in red) believed to specify required enzymes. The number of enzymes needed for synthesis of the later intermediates is indicated above the arrows.

6.1.1 *H. salinarum* strain R1 can grow without AroAAs

The ability of R1 to grow in synthetic medium with and without AroAAs was determined by monitoring the growth of R1 after adaptation to synthetic medium with different concentrations of AroAAs. The results are presented in figure #20.

The addition of AroAAs to the synthetic medium improved the growth rate of *H. salinarum* R1. In synthetic medium with 1.1mM AroAAs the generation time (Gn) was two fold higher than without AroAAs (11.8±0.26h and 23.8±0.3h, respectively). No difference in the Gn time was seen when the synthetic medium was supplemented with 1.1mM, 0.11mM or 0.011mM AroAAs (Gn=11.8±0.26h, 11.7±0.02h and 11.4±0.16h, respectively). Together, the data indicate the following: (1) WT can uptake one or more of the supplied AroAAs (figures #31, 38 and 41), (2) growth

stimulation only requires low concentrations of AroAAs, (3) 1.1mM AroAAs was a reasonable level to use in the supplemented media, and (4) while a WT strain can clearly grow without AroAAs, deletion strains are expected to be auxotrophic for AroAAs supplementation
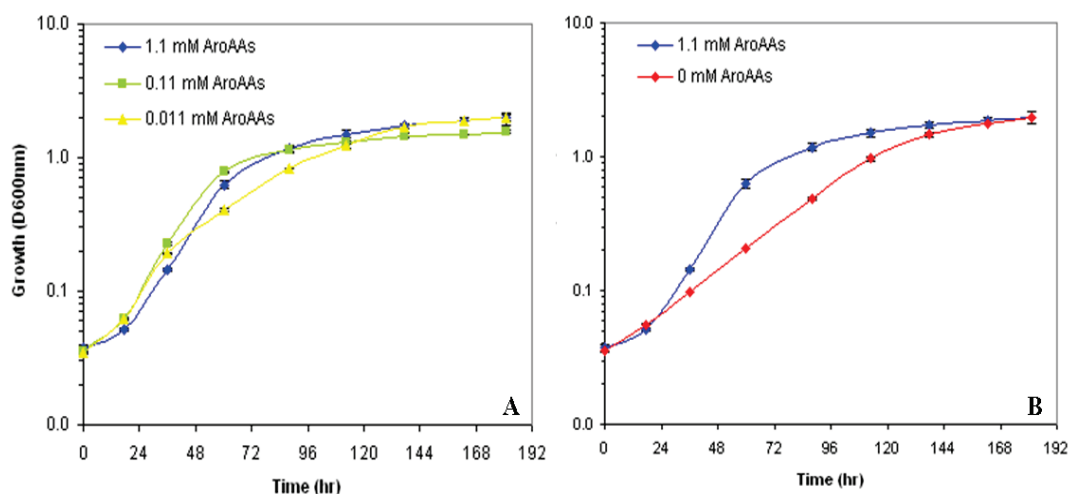
Figure # 20: Growth of R1 in synthetic medium with different concentrations of AroAAs. The cell grow in 35ml, 40°C, 100rpm, in the dark. Points represent the average absorbance at each time point. Vertical error bars represent the standard deviation.

6.1.2 Construction of mutants deleted for genes OE1472F, OE1475F and OE1477R

6.1.2.1 Plasmids for gene deletions

Single, in-frame deletions of ORF's OE1472F, OE1475F and OE1477R (figure #19) were attempted by recombination using integration plasmids carrying flanking regions of each gene. Flanking regions were amplified separately by PCR, ligated to each other, cloned into plasmid pMKK100 (figure #21A, B and C, respectively), and the plasmid introduced into *E. coli*. Transformed *E. coli* colonies were screened for plasmid containing the correct insert by restriction with *BamHI* and *HindIII* (figure # 21D, lane 4). As seen in figure #21C, the final constructs contained Amp^R and Mev^R markers, and the *Bga*H gene as the reporter gene for blue/ red screening, but no replication genes for haloarchaea.
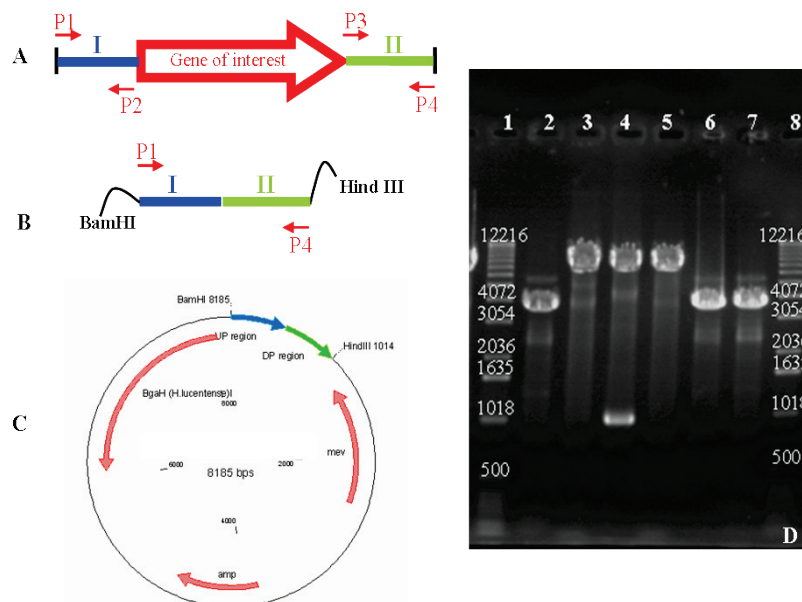
Figure # 21: Constructing plasmids for in-frame deletion. **A-** Illustration of the flanking US (I) and DS (II) regions of a specific gene with introduced restriction sites at each end. **B-** Illustration of the ligated flanking regions which were introduced to a plasmid. **C-** General map to illustrate the plasmids used for in-frame deletions of ORF's OE1472F, OE1475F and OE1477R. **D-** Example of a 1% agarose gel of purified plasmids from *E. coli* transformants after *Bam*HI+*Hind*III digestion. Fragments of ~ 1Kbp and ~7Kbp are seen for the positive candidates. MW markers are shown in lanes 1 and 8 (in bp).

6.1.2.2 Selection and screening of deletion strains

The conditions needed for selecting and screening of deletion strains after transformation of the suicide plasmids into R1 needed to be considered carefully for the following reasons. Firstly, deletion of these genes should convert *H. salinarum* to aromatic amino acid auxotrophy, requiring the synthetic medium to be supplemented with AroAAs. Secondly, chorismate is a precursor of two additional pathways (figure #22): (a) the production of menaquinone (MQ), which is the mobile carrier of electrons in the respiratory chain, and (b) the production of tetrahydrofolate (THF), which is converted to several other folate coenzymes that are involved in biosynthesis of amino acids, nucleotides and coenzyme A. This means that THF (or a precursor for THF) would also be predicted to be required by such mutants. In addition, to overcome the dependence upon the respiration chain, cells could be grown anaerobically. Fortunately, R1 is photosynthetic. It can synthesise bacteriorhodopsin as a light driven proton pump. This generates a proton motive force that is utilized for ATP synthesis (reviewed by (Oesterhelt, 1998)).

Therefore, the synthetic medium used for recovery of deletion mutants was supplemented with 4-aminobenzoic acid (ABZA, a precursor for THF), 1.1mM AroAAs and 1.1mM DHQ. Cells were grown anaerobically under strong illumination.
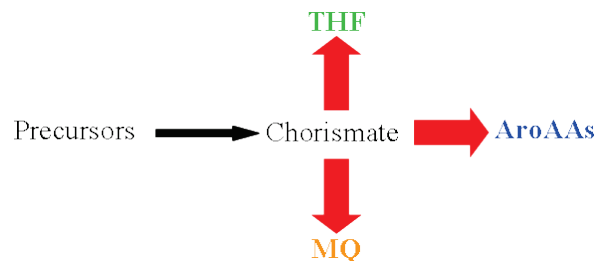
Figure # 22: The precursor of AroAAs, menaquione (MQ) and tetrahydrofolate (THF). Modified diagram from GenomeNet Database resources (KEGG), *H. salinarum* map 00400 (http://www.genome.ad.jp/kegg/pathway/map/map00400.html)

6.1.2.2.1 Mutants could not be recovered by anaerobic growth under white light

Strain R1 was transformed with suicide plasmids designed to create in-frame deletion of OE1472F, OE1475F and OE1477R. Mev^R transformants with integrated plasmids were recovered on selective plates and screened for the presence of the *Bga*H gene. Mev^R and *Bga*H$^+$ transformants were then grown anaerobically in 8ml volume sealed glass tubes (40°C, 100rpm, figure #23B) containing 4ml of synthetic medium with 1.1mM AroAAs + 1.1mM DHQ and 0.5mM ABZA (SM + AroAAs + DHQ+ABZA) without drug selection, to allow further recombination to eliminate the target genes, leaving only the flanking regions. Cultures were inoculated into fresh medium after the turbidity had reached ~50 klett units, for a total of three passages. The growth of the recombinant strains after the 3rd transfer was recorded and plotted (figure #23A). As controls, the anaerobic growth of R1 in a complex medium and in synthetic medium with 1.1mM AroAAs was also recorded. The generation time (Gn) of R1 in synthetic medium was 13.07h. The strains that had been transformed with plasmids targeting ORFs OE1472F, OE1475F and OE1477R, had Gn times of 12.54h, 7.44h and 5.45h, respectively.

Although the tubes were not flushed with N_2, the growth conditions rapidly become anaerobic because the inoculum (2 Klett unit in 4 ml medium) consumes all the O_2 in the medium within 17h (the consumption rate of O_2 is 0.63μmole O_2/day x

Klett, solubility of O_2 in high salt medium is 25µM O_2, Loccedie Mansueto, data not published). Furthermore, the tubes were sealed tightly with a screw capped lid.
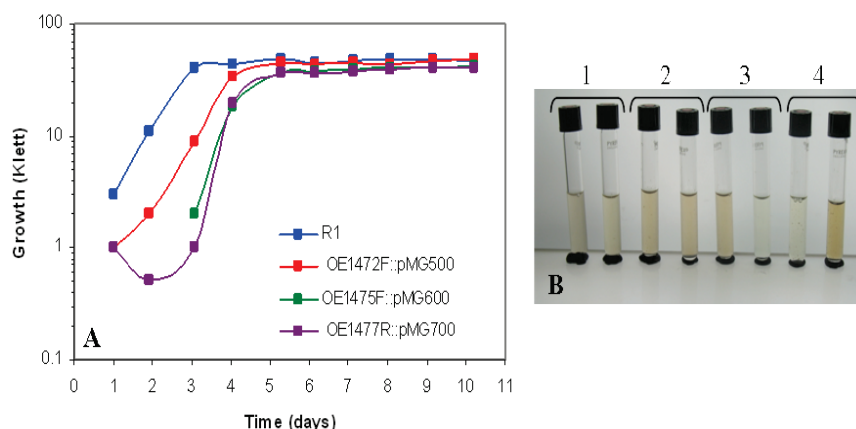
Figure # 23: The anaerobic growth of adapted *H. salinarum* cells in sealed glass tubes. **A-** The growth curves after 3rd transfer into synthetic medium+1.1mM AroAAs+1.1mM DHQ+0.5mM 4 aminobenzoic acid ,at 40°C, 100rpm in sealed glass tubes. Cells grew up to a density of 50 Klett units. **B-** The tubes in duplicates used for the anaerobic growth under white light. **1**-R1, **2**-OE1472F:pMG500, **3**- OE1475F:pMG600, and **4**-OE1477R:pMG700, respectively.

To differentiate between WT and deletion strains, the adapted anaerobic cultures were plated on supplemented synthetic medium plates (SM + AroAAs + DHQ+ABZA) containing X-gal but without Mevinolin. Cells that had undergone a second cross over and eliminated the target gene would also have eliminated the *Bga*H gene and so would have a red color phenotype (due to bacteriorhodopsin and bacterioruberin) on X-gal plates. Blue colonies would be strains that still retained the plasmid. Synthetic medium plates with 1.1mM AroAAs, 1.1mM DHQ and 0.5mM ABZA were inoculated with adapted anaerobic cultures from the glass tubes and incubated in anaerobic chambers, under white light, as seen in figure # 24.

The temperature in the sealed anaerobic chambers was regulated to 37°C using circulating water reservoir at the same temperature. The air in the chambers was removed by flushing the chambers with N_2, and sealing them, so there was no possible exchange of air with the outside. Each chamber contained a colored indicator to detect oxygen. Trace amounts of oxygen were removed by the included Gas Pak, which produced CO_2+ H_2 when the O_2 in the chamber reacted with sodium borohydride, sodium bicarbonate and citric acid tablets in the Gas Pak. The oxygen indicator remained white during the experiment (i.e. no oxygen present).
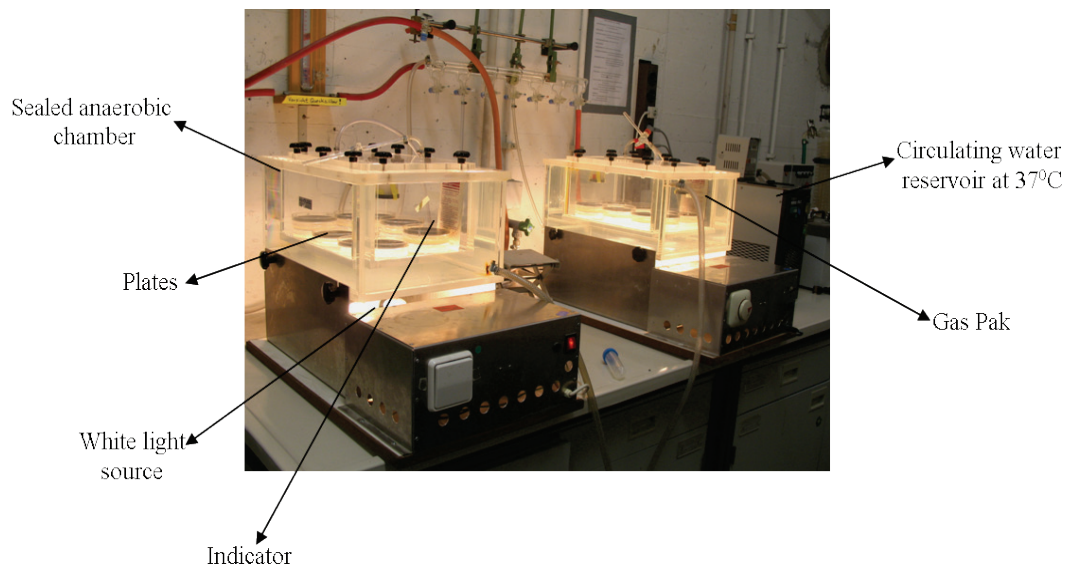
Figure # 24: The anaerobic chambers used to differentiate WT from deletion strains grown on synthetic medium plates with 1.1mM AroAAs+1.1mM DHQ+0.5mM ABZA, at 37°C, under white light. (Oesterhelt, 1982).

As a growth control, WT cells were plated on both complex medium and synthetic medium with 1.1mM AroAAs and incubated in the same chambers. Colonies of WT were visible on plates after 9 weeks in the anaerobic chambers. However, only two colonies of grew after 7 months of incubation, most probably WT. Those colonies did not re-grow when subcultured into liquid medium under anaerobic condition in sealed glass tubes (figure #23B), so there was no possibility to check whether the cells were WT or recombinant strain. Nonetheless the experiment was not successful, so recovery under aerobic conditions was attempted.

6.1.2.2.2 Recovery of deletion mutants by aerobic growth in the dark

Next, attempts to obtain knockout strains were done using aerobic growth in synthetic medium supplemented with 1.1mM AroAAs and 1.1mM DHQ (SM+AroAAs+DHQ). Assuming that DHQ can be taken up by cells from the medium, the production of chorismate was expected to support the production of AroAAs, MQ and THF in mutants lacking ORFs OE1472F, OE1475F and OE1477R (figure #22).

Strain R1, after transformation with plasmids pMG500, pMG600 and pMG700, were plated on complex medium plates containing mevinolin (Mev) and X-gal. Mixture of cells with the *Bga*H[+] phenotype (via homologous recombination of the plasmid) were subcultured into 10ml of fresh medium (in 25ml flask, 40°C, 100rpm,

dark) for a total of three passages. The growth of the recombinant strains after the 3rd transfer was recorded and plotted as seen in figure #25.
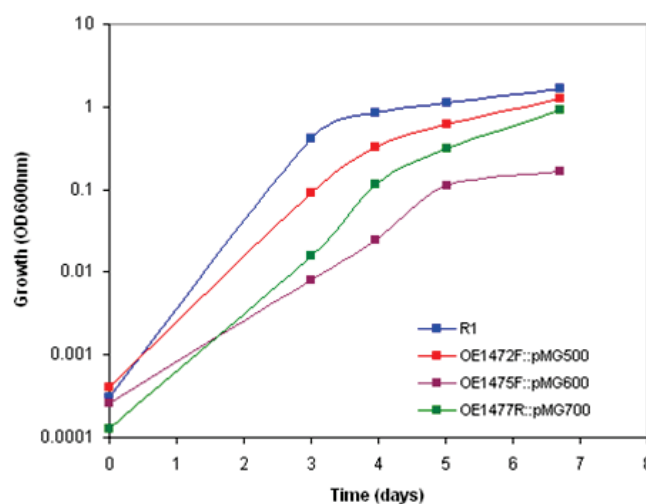
Figure # 25 : The growth of recombinants in synthetic medium+1.1mM AroAAs+1.1mM DHQ. Conditions: 40°C, 100rpm, aerobic. The initial OD presented in the graph was calculated.

The adapted cells from liquid medium were plated on synthetic medium plates which contained 1.1mM AroAAs, 1.1mM DHQ and X-gal but without mevinolin. The plates were incubated at 37°C for 20 days. $BgaH^+$ and $BgaH^-$ colonies were obtained for all three gene deletion constructs. The $BgaH^+$ phenotype indicated strains where the plasmids were still integrated into the chromosome, while the red phenotype indicated either a WT strain or a deletion strain. The strains with the $BgaH^-$ phenotype were tested by colony PCR and Southern blot hybridization, to distinguish WT from the desired deletion strain.

6.1.3 In vivo mutation analysis

6.1.3.1 Deletion mutant ΔOE1477R

6.1.3.1.1 PCR analysis to confirm the deletion in ΔOE1477R

$BgaH^-$ colonies were screened by PCR to confirm the in-frame deletion of OE1477R. Two sets of colony PCR reactions were performed. The first PCR used primers P1+P4 to differentiate WT from a deletion candidate. In a WT strain, a full length PCR fragment of 1.69Kbp was expected, while in a potential deletion

candidate, only a 1Kbp fragment would be amplified (the length of the flanking regions). A second PCR reaction used primers P5+P4, in which WT amplification would generate a 1.2Kbp product, while no product would be expected in a deletion strain.

PCRs using primers P1+P4 are shown in figure #26B lanes 2, 4, and 6, respectively. WT DNA gave the expected product (1.69Kbp), while a 1Kbp fragment was amplified from the potential deletion candidates #3 and #11 (figure #26B, lanes 4 and 6, respectively). Note that in strain #11 (lane 6) two additional bands are seen and one of them has the same size as WT (lane 2). PCRs using primers P5+P4 are shown in lanes 3, 5, 7, respectively, only WT DNA gave a product. According to colony PCR analysis, both colonies are potential deletion candidates, and they were further analysed by Southern blot hybridization.
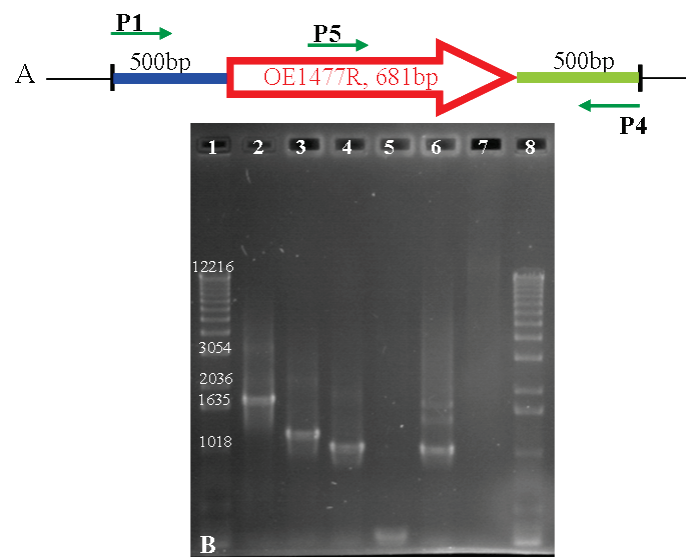
Figure # 26: PCR analysis of potential ΔOE1477R strain. **A-** The relative position of the primers used to confirm the deletion of ORF OE1477R. **B-**1% agarose gel of 1st PCR reaction using primers P1+P4 and total DNA from WT, strain #3, and strain #11 (lanes 2,4,6, respectively). 2nd PCR reaction using primers P5+P4 and DNA of WT, clone #3, and clone #11 (lanes 3,5,7, respectively). MW markers are shown in lanes 1 and 8 (in bp).

6.1.3.1.2 Southern blot analysis

To verify the deletion of OE1477R in the potential deletion candidates (strain # 3 and 11), a Southern blot analysis was performed. Total DNA was isolated from liquid cultures (SM+AroAAs+Shikimate+ABZA) of each strain, digested with *SacI*, separated on a 1% agarose gel, and transferred to a membrane. Two DIG labeled

probes were used: one for the flanking region of OE1477R, and a second for the gene itself. In the WT strain, the probe for the flanking region should detect a 1.7Kbp fragment, while a shorter fragment is expected (0.98Kbp) in a deletion candidate. The probe for the gene should detect a 1.7Kbp fragment in the WT strain and nothing in a deletion strain.

As shown in figure #27A, strains #3 and #11 produced a 0.98Kbp band when a DIG labeled probe of the flanking region of OE1477R was used (figure #27A). Nevertheless, only in strain #3 was the gene deleted entirely, as seen from figure #27B (lane 3), since no complementary DNA segment was detected when using DIG labeled probe for the gene itself. In strain #11, a 1.7Kbp fragment was detected with the gene probe, indicating that it is a mixture of WT cells and a deletion strain.

Controls in the Southern blot included DNA from plasmid pMG700 (which contains the flanking regions of OE1477R) and plasmid pMG750 (which contains the gene OE1477R). As expected, the DNA segment of pMG700 was detected only when using DIG labeled probe for the flanking region (figure #27A, lane 5), while the DNA segment of pMG750 was detected only when using DIG labeled probe for the gene (figure #27B, lane 6).
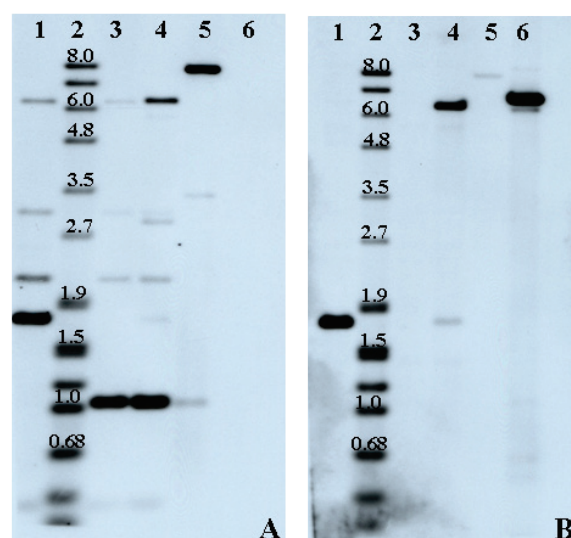
Figure # 27: Southern blot analysis of WT vs. deletion strains. Lane 1: *SacI* digested DNA from WT, lane 2: DIG VII size marker (in bp), lane 3: *SacI* digested DNA from deletion strain #3, lane 4: *SacI* digested DNA from deletion strain #11, lane 5: *HindIII* digested pMG700, lane 6: *HindIII* digested pMG750. **A**- Hybridization with a DIG labeled probe of the US flanking region of OE1477R. **B**- Hybridization with a DIG labeled probe of part of the gene.

6.1.3.1.3 Phenotype of the ΔOE1477R strain

In order to check the phenotype of the ΔOE1477R mutant, strain #3 was grown in synthetic medium with 1.1mM AroAAs and 1.1mM shikimate and plated on synthetic medium with or without AroAAs. If OE1477R is part of the AroAAs biosynthesis pathway, this strain should be auxotrophic for AroAAs, and will not be able to grow in synthetic medium without AroAAs. As seen in figure #28, the ΔOE1477R mutant was unable to grow on plate without AroAAs, even after 28 days incubation, while WT cells did not grow well under the same conditions but growth was clearly visible (lower right panel). As a control for the viability of the cells and to demonstrate that the deletion strain is able to grow on those plates, cells were grown in liquid synthetic medium supplemented with AroAAs and shikimate, and plated as well (left and middle panels of figure #25).
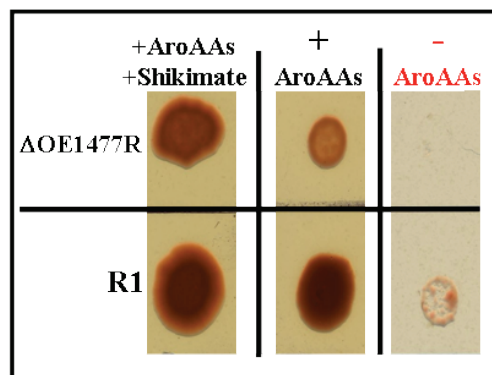
Figure # 28: Phenotype of ΔOE1477R compared to WT, growing on plates. Cells were grown in liquid synthetic medium supplemented with 1.1mM AroAAs and 1.1mM shikimate until OD_{600nm}=1.565, and 1.315 (deletion and WT, respectively) (40°C, 100rpm). Cells were washed with basal salt before diluting $1/10^{-3}$ and spotting 3μl from each dilution on plates containing AroAAs+shikimate, only AroAAs and without AroAAs .

Next, the average Gn of the deletion ΔOE1477R strain #3 was measured after inoculation into liquid synthetic medium supplemented as follows: (1) 1.1mM AroAAs, (2) synthetic medium without AroAAs, (3) AroAAs +DHQ, and AroAAs + Shikimate (4). As shown in figure #29B, the deletion strain ΔOE1477R was not capable of growing without AroAAs in the medium, but could grow at half the rate of the WT when supplemented with AroAAs. In order for the deletion strain to grow as well as the WT the synthetic medium needed to be supplemented with 1.1mM

72

AroAAs and either 1.1mM DHQ or 1.1mM Shikimate (figure # 29C and D, respectively).
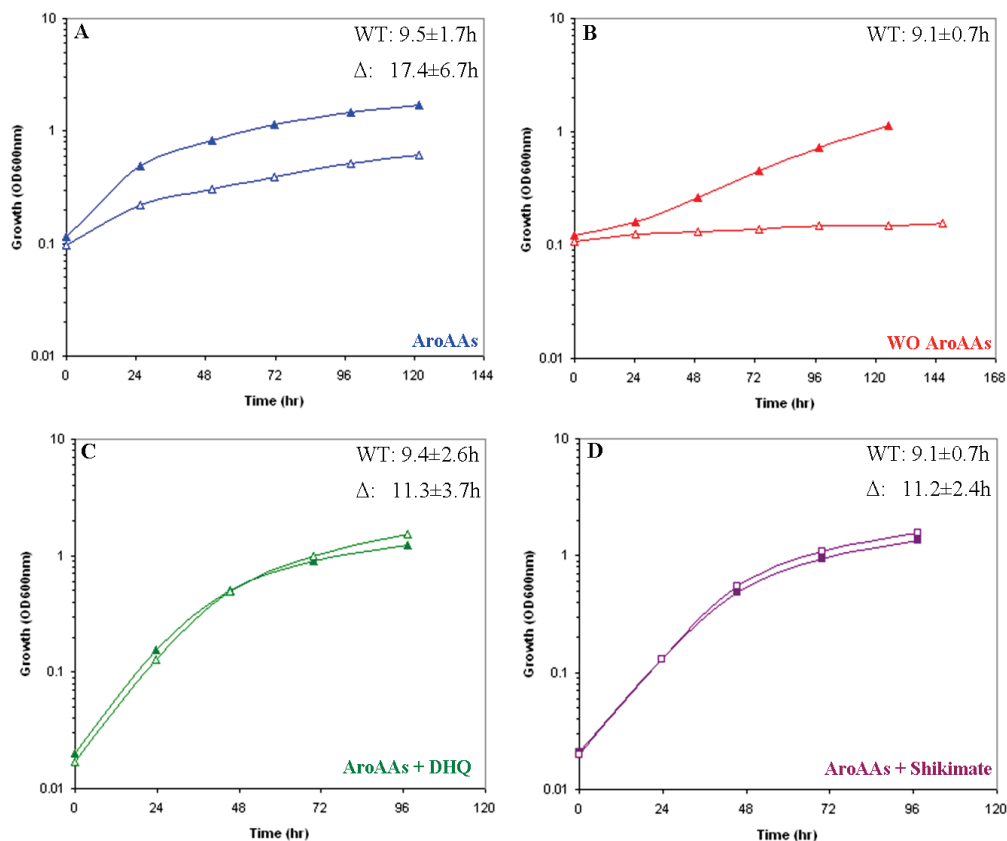
Figure # 29: The growth of ΔOE1477R in supplemented synthetic medium. The growth of WT (▲) and ΔOE1477R(△) in synthetic medium with 1.1mM AroAAs (A), without AroAAs (B), with 1.1mM AroAAs +1.1mM DHQ (C), and with 1.1mM AroAAs + 1.1mM Shikimate (D). The inoculum was washed with basal salt solution, before resuspending in the appropriate medium. The numbers at the top left are average Gn times. The cells were grown in 25ml flasks containing 10ml medium, at 40°C, 100rpm.

Both, deletion and WT strains had similar average Gn times when grown in synthetic medium with AroAAs + Shikimate (figure # 29D), suggesting that the enzyme encoded by OE1477R acts on a substrate upstream to Shikimate, since Shikimate was able to compensate for the loss of the gene.

Additionally, the ability of the deletion strain to grow in synthetic medium with AroAAs + DHQ (figure #29C) as well as in medium containing only DHQ (figure #30B), indicate that the gene product of OE1477R acts on a substrate upstream to DHQ, which is not in agreement to the reconstructed pathway (figure #19). This discrepancy can be explained by the reaction presumably catalyzed by OE1477R, either synthesis of DHQ or dehydration of DHQ (figure #30A). If the former is true than DHQ is expected to compensate for the loss of the gene, as demonstrated by the

growth of the mutant in synthetic medium supplemented with AroAAs + DHQ (figure #29C). If the later is true the cells have been fed with both DHQ and DHS. The later was examined in more details by LC-MS. The two possible sources for DHS are contamination in DHQ or spontaneous release of water molecule when DHQ was incubated at 37°C during the growth of the cells. In either case there was a need to determine the amount of DHS available to the cells.
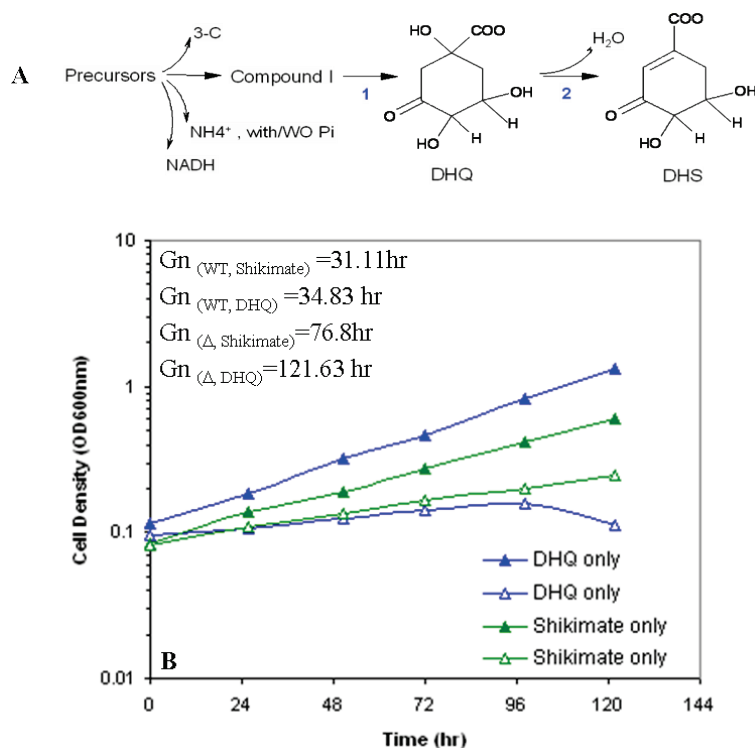
Figure # 30: **A**- Biosynthesis of DHS. The numbers in the figure refer to the enzymes: (1) 3-dehydroquinate synthase (EC 4.6.1.3), and (2) 3-dehydroquinase dehydratase (EC 4.2.1.10). **B**- Growth requirements of ΔOE1477R mutant (△) and WT (▲). The inoculum was washed with basal salt solution before resuspending in synthetic medium supplemented with either DHQ or shikimate. Cells were grown in 25ml flasks containing 10ml medium, at 40°C, 100rpm.

As seen in table # 2, 0.4% DHS was detected after derivatization of DHQ. One can not distinguish between contamination in DHQ with DHS or formation of DHS during the derivatization procedure. In either case the concentration of DHS is low. On the other hand, when incubating DHQ for 1h at 37°C, 3% of DHS were formed. No increase in the formation of DHS was observed when DHQ was incubated for 23h at the same temperature, indicating that the reaction either achieved equilibrium (although, water being bad nucleophile most likely will not attack double boned), or out of the total percentage of DHS formed most of it is been transformed to other

intermediate leaving 3% of DHS. In either of the explanation when feeding the cells with DHQ we were also providing at least 3% of DHS. The 3% of DHS corresponds to 0.033mM DHS, meaning 0.2mM DHS per OD (0.033mM DHS/0.159OD, figure#30B). As shown by Gonzalez et al., (figure #4 in (Gonzalez et al., 2009)), only 0.1mM Phe/OD is needed for biomass (whether integrated into proteins or as free metabolites), demonstrating that by feeding the cells with DHQ containing 3% of DHS, the cells can grow.

Table # 2: Detection DHS formed under different conditions

[1] Pre-treatment	[2] Detection (%)	
	DHQ	DHS
no pre-incubation	99.6	0.4
90°C, 1h in 0.1M HCl	76	24
37°C, 1h	97	3
37°C, 23h	97	3

[1]- 2.5mM DHQ in 1.75M KCl+29mM Tris-HCl was pretreated before derivatization, according the conditions specified in the table.

[2]- Derivatization was carried out at 60°C for 1h, using 2mM NBHA [O-(4-nitrobenzyl) hydroxylamine hydrochloride] in MeOH. The derivatized compounded were separated on C18 column using ACN. DHQ eluted after 15min and DHS after 20.2min. The % represents area detected at 254nm.

6.1.3.1.4 Aromatic amino acid analysis of the ΔOE1477R mutant

The concentrations of aromatic amino acids, phenylalanine (Phe) and tyrosine (Tyr) were measured in the medium of cultures grown in different synthetic media supplemented with AroAAs, Shikimate, DHQ, and combinations of these. Media were inoculated with either R1 or the deletion strain ΔOE1477R, and the concentrations of Phe and Tyr from three biological repeats (represented by the symbols circles, triangles and diamonds) at different time points during the growth,

were plotted as shown in figure #31. The experimentally measured concentrations were modeled (represented by the dash lines, done by Dr. Orland Gonzalez) which allowed the calculation of uptake rates of Phe and Tyr in different supplemented media.

No significant differences were observed in the extracellular concentration of Tyr between cultures of the WT and deletion strain, in either of the synthetic media (figure 31B, D and E, table #2). Moreover, no difference in the deletion strain, in either the extracellular concentration or the uptake rate of Phe (table #3), was observed in synthetic media containing AroAAs+Shikimate or AroAAs+DHQ (figure #31C and E), while a significant difference was seen in the concentration of Phe in synthetic medium supplemented only with AroAAs (figure #31A). In the case of the WT, the Phe levels decreases steadily after 2 days of growth (OD_{600}=0.514), while in the mutant ΔOE1477R the concentration of Phe remained steady during the growth. The uptake rate of the R1 strain was 9.2 nmole Phe $OD^{-1}ml^{-1}h^{-1}$, while no uptake was seen in ΔOE1477R grown in the same medium (table #3), unless either DHQ or shikimate were added to the medium in addition to AroAAs.

Table # 3: Uptake rates of Phe and Tyr by R1 and ΔOE1477R. Cells grow in supplemented synthetic media. The uptake rates were calculated from the corresponding model stimulations.

Synthetic media supplemented with	Phe (nmole OD^{-1} ml^{-1} h^{-1})		Tyr (nmole OD^{-1} ml^{-1} h^{-1})	
	R1	ΔOE1477R	R1	ΔOE1477R
+AroAAs	9.2	0	0.6	0
+AroAAs+Shikimate	8.7	7.5	0	0
+AroAAs+DHQ	12.5	9.4	0.6	0
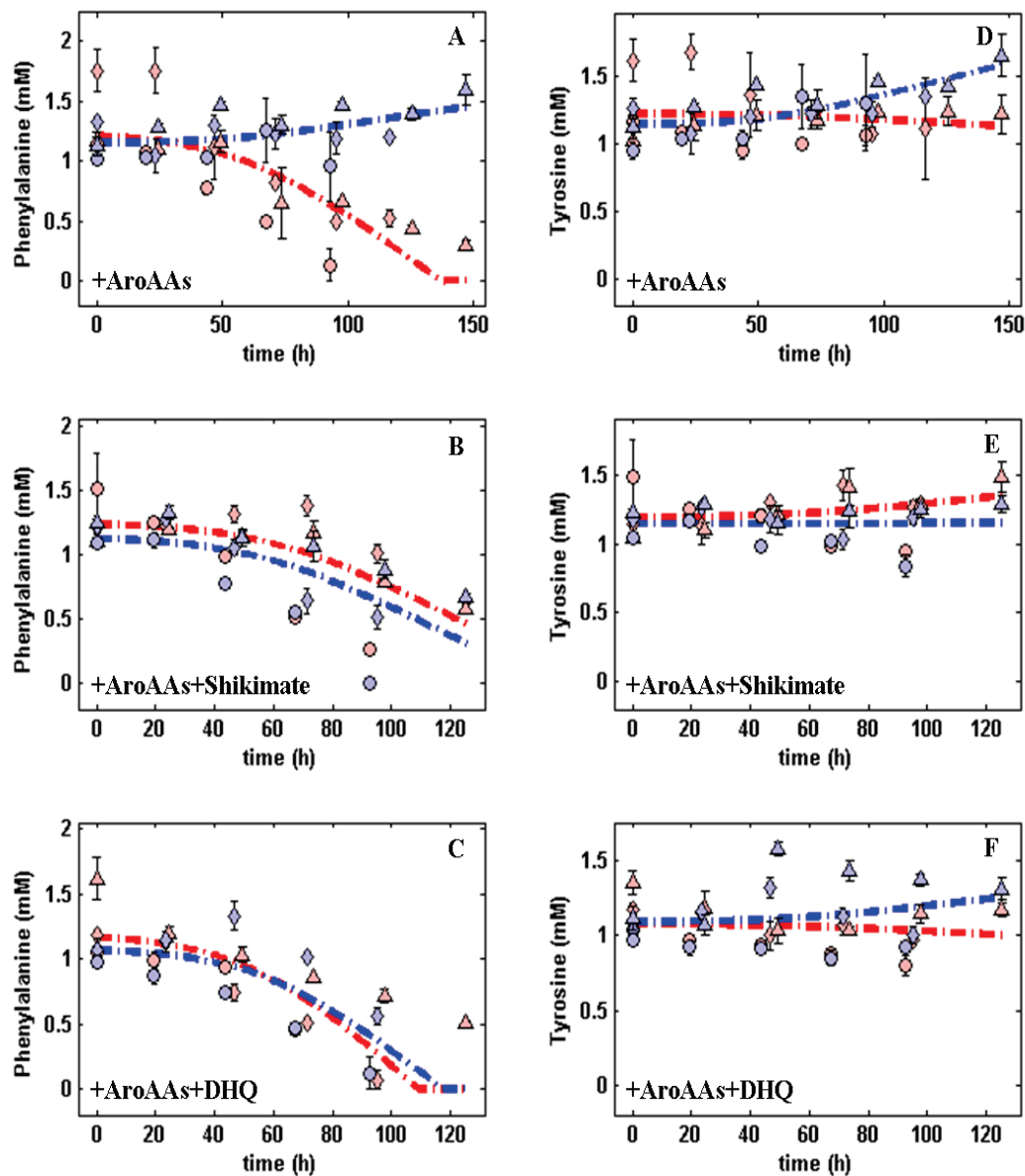
Figure # 31: Phe and Tyr consumption during the growth of WT and mutant ΔOE1477R. The experimental data for Phe (left panel, A, B, and C) and Tyr (right panel, D, E and F) are shown using circles, triangles and diamonds (representing biological repeats) , red for WT and blue for the mutant. Error bars represent the deviation of the average. Simulations curves are shown using dashed lines, red for WT, and blue for ΔOE1477R. The uptake rates were calculated from the corresponding model stimulations. Note that small differences in the concentration of AroAAs are hard to detect, therefore low uptake rats need to be carefully addressed. Nevertheless, the differences in the trends between WT and deletion strain are more trustworthy. Done by Dr. Orland Gonzalez.

6.1.3.1.5 DHQ uptake by H. salinarum and ΔOE1477R

DHQ in the supernatants of late exponential phase cultures of R1 and ΔOE1477R, grown in synthetic medium with 1.1mM AroAAs and 1.1mM DHQ (figure #32A), was derivatized using the method discussed in materials and methods. The concentration of DHQ in the medium was calculated using derivatized DHQ.

The data confirm uptake of DHQ from the medium by both WT and mutant ΔOE1477R. As shown in table #2, DHQ is in equilibrium with DHS. The DHS can be used by the mutant to be converted to AroAAs.
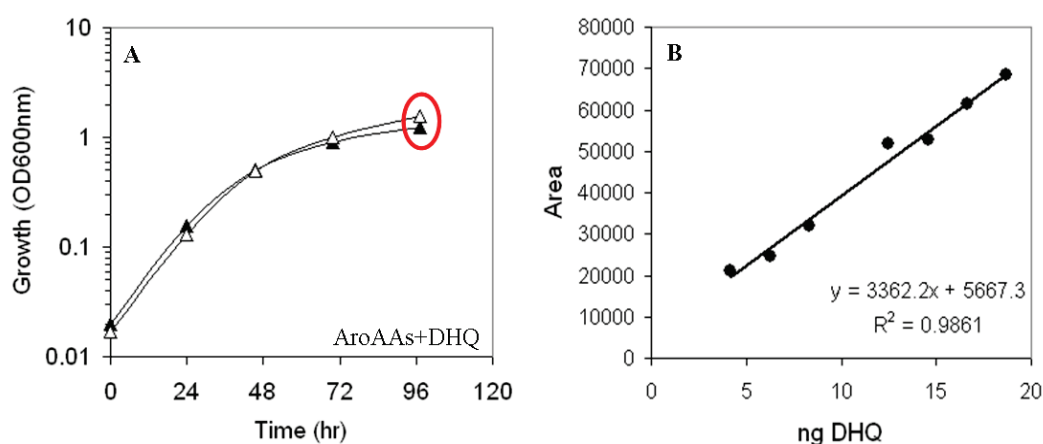
Figure # 32: Growth of R1 and ΔOE1477R in medium supplemented with AroAAs +DHQ. **A**-growth curves of R1 (▲) and ΔOE1477R (△), respectively. Cells were grown in 25ml flasks containing 10ml medium, at 40°C, 100rpm. **B**-calibration curve of derivatized DHQ.

Table # 4 : Concentration of DHQ in the medium after the growth of R1 and ΔOE1477R in synthetic medium with 1.1mM AroAAs and 1.1mM DHQ.

	OD600$_{nm}$	mM DHQ	% DHQ used
R1	1.22	0.728	42
ΔOE1477R	1.54	0.265	79
Blank [1]		1.245	

[1] Synthetic medium with 1.1mM AroAAs and 1.1mM DHQ

6.1.3.2 ORFs OE1472F and OE1475F were disturbed by insertions

Attempts to obtain deletion strains of OE1472F and OE1475F were not successful, even when using the conditions that allowed the isolation of ΔOE1477R i.e., aerobic growth in synthetic medium supplemented with 1.1mM AroAAs and 1.1mM DHQ, or

1.1mM AroAAs+1.1mM Shikimate + 0.5mM 4-aminobenzoate. While colonies displaying a BgaH⁻ phenotypes were obtained after culturing the primary transformants in medium without mevinolin, PCR analyses revealed that these were genotypically WT, indicating the second cross-over had only restored the WT genotype.

Looking at the genomic context of OE1472F and OE1475F (figure #33), it was possible that complete deletion of these genes might interfere with expression of nearby genes. As seen in figure #33, deletion of OE1472F, could disturb the *trp*CBA operon, or OE1474R (CHY), while deletion of OE1475F might interfere with the transcription of OE1476R and OE1477R (which most probably form an operon). Therefore, instead of full deletions, the target genes were disturbed by insertions, with the aim of obtaining mutants showing reduced activity of the corresponding enzymes.
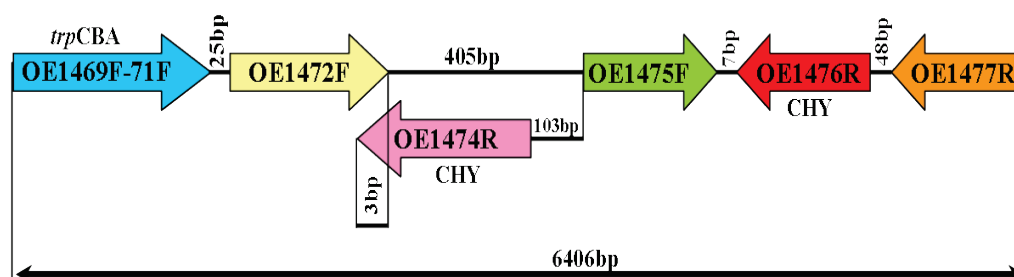
Figure # 33: Schematic representation of annotated ORFs in *H. salinarum* R1 genome (genome coordinates 230688-237094bp)

6.1.3.3 Insertion mutant OE1471F:: pMG501

Four ORF's are included in the *trp* operon (figure #34) OE1469F (*trp*C), OE1470F (*trp*B), OE1471F (*trp*A) and OE1472F. By introducing a transcription terminator sequence at the 3' end of the gene OE1471F (*trp*A), transcription of OE1472F would be reduced but the upstream genes should not be affected.

500bp of the 3' end of OE1471F were amplified by PCR and ligated to a 55bp long sequence containing the terminator of the *flg*A operon. The resulting 583bp fragment was amplified by PCR using LA Taq polymerase, purified from a preparative 1% agarose gel, cloned into TOPO vector and sequenced to confirm that no errors had been introduced. It was then subcloned into pMKK100, generating plasmid pMG501, which contained Amp^R, Mev^R markers and *Bga*H as a reporter gene. pMG501 was

introduced into strain R1, and transformants were recovered on complex medium plates with Mev and Xgal. Colonies were screened by PCR to confirm the in-frame integration of the terminator using the primers indicated in figure #34 (green arrows). In-frame integration should show a 1566bp fragment, whereas the WT sequence should not generate a product. As shown in figure #35A, all transformants had integrated the plasmid pMG501. PCR controls included chromosomal DNA from the WT, and plasmid pMG501. As expected, no product was amplified from WT DNA, while plasmid DNA gave a strong band. To exclude the possibility that pMG501 could survive in R1 without integration into the chromosome (although pMG501 did not contain a replication ori for haloarchaea), a second PCR was performed using the primers indicated in figure #34 (red arrows). In-frame integration should give a 2209bp fragment, while no product is expected for the WT. As shown in figure #35B, plasmid pMG501 was found to be integrated in all transformants tested. Controls in the PCR included chromosomal DNA from the WT strain, and plasmid pMG501. As expected no amplification was visible when using WT or plasmid DNA.
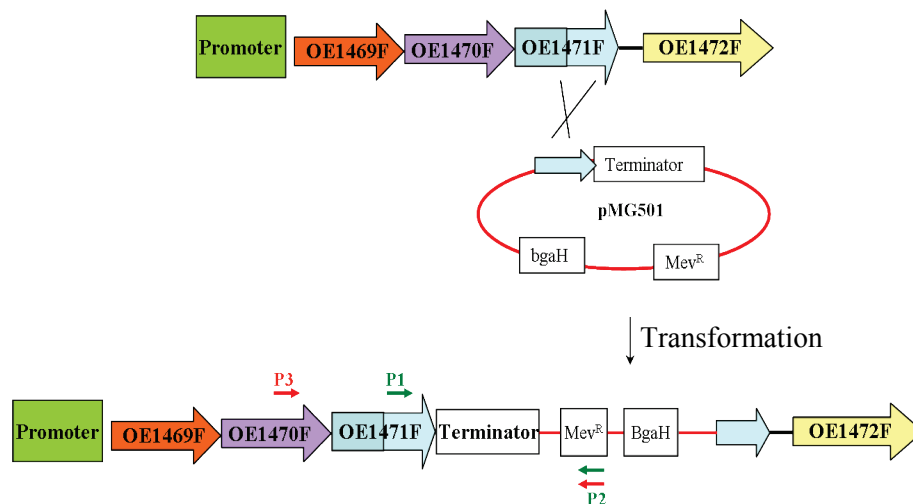
Figure # 34: Schematic representations of the single cross-over event between the chromosomal DNA of R1 and pMG501. Arrows indicate the relative positions of the primers used.
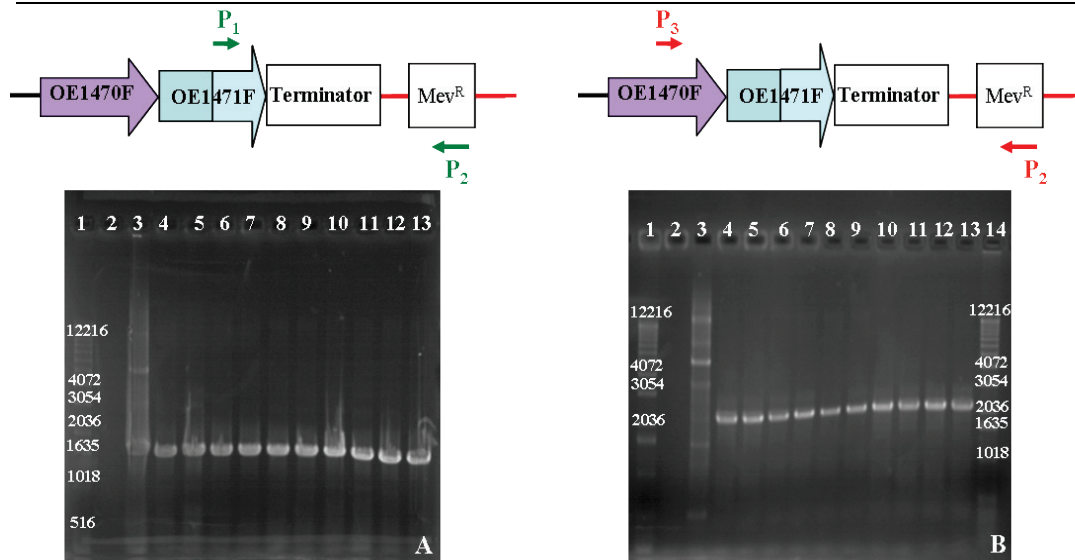
Figure # 35: PCR analysis of cells transformed with pMG501 using **A** - primers P1 and P2 (green arrows), and **B**- primers P3 and P2 (red arrows). Lanes 1 and 14: MW markers (in bp). The templates used in the PCRs were, lane2: Chromosomal DNA of R1, lane 3: plasmid pMG501, lanes 4-13- transformed colonies. !% agarose gel.

6.1.3.3.1 Transcription level of ORF OE1472F was reduced in the mutant

The expression of OE1472F in OE1471F::pMG501 mutant was compared to the WT strain. Both were cultivated in synthetic medium without AroAAs. The transcript level differences were calculated by a relative quantification approach using an internal standard gene, OE1160R (ribosomal protein L10.eR). As seen in figure# 36, the expression of OE1472F in the mutant strain was down-regulated 4.7 fold and 44 fold when the culture was harvested at $OD_{600nm}=0.4$ and 1.0, respectively. The down-regulation indicates that integration of pMG501 into the chromosomal DNA of R1 (figure #35) had indeed reduced the transcription level of OE1472F. This mutant was then analysed for its ability to grow with and without AroAAs.

Figure # 36: Transcription levels of ORF OE1472F in mutant OE1471F::pMG501 at OD600nm=0.4 and 1.0. Results were obtained using RT-qPCR.
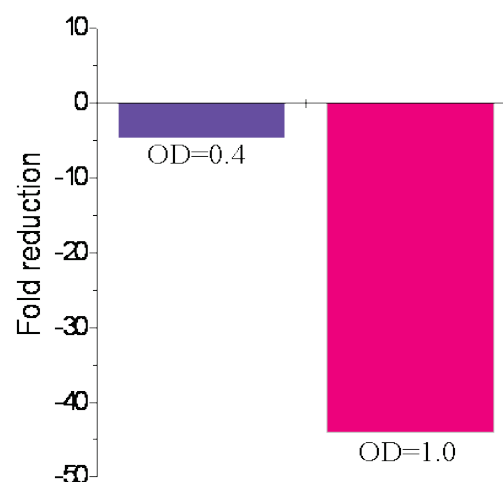
6.1.3.3.2 Phenotype of OE1471F::pMG501 mutant

Introduction of the upstream terminator was shown to reduce or completely block transcription of OE1472F and, if OE1472F is part of the AroAAs biosynthesis pathway, the mutant should grow poorly or not at all in media lacking AroAAs.

The phenotype of OE1471F::pMG501 was tested after inoculation of the mutant strain into synthetic medium with and without AroAAs, or into synthetic medium supplemented with 1.1mM AroAAs and either 1.1mM Shikimate or 1.1mM DHQ. After each growth the genotype was conformed by PCR.
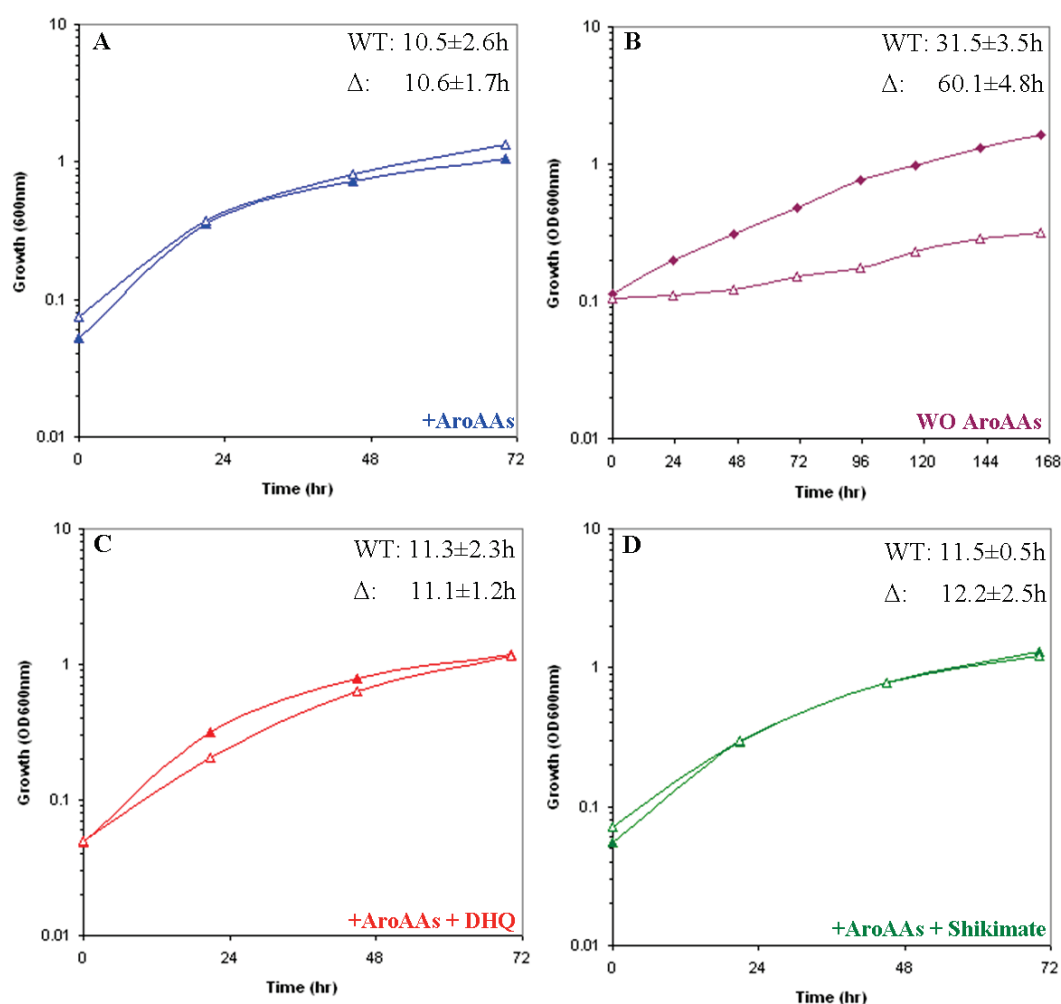
Figure # 37: Growth requirements of WT (▲) and mutant OE1471F::pMG501 (△) in synthetic media supplemented with **A**-AroAAs, **B**-without AroAAs, **C**-AroAAs+DHQ, and **D**-AroAAs+Shikimate. The inoculum was washed with basal salt solution before resuspending in the appropriate medium. The numbers at the top left are average Gn times. Cells were grown in 25ml flasks containing 10ml medium, at 40°C, 100rpm.

As seen in figure#37A, C and D, similar average Gn times were obtained for both WT and the mutant when grown in media with AroAAs, and in media with AroAAs and either DHQ or Shikimate. But in media without AroAAs, the mutant OE1471F::pMG501 grows at about half the rate of WT cells (figure #37B). Together, the data indicate a reduced ability to synthesis AroAAs, and this is alleviated by supplementation of the medium with AroAAs. It also suggests that either transcription of OE1472F was not completely blocked by the integrated terminator, or that another gene in *H. salinarum* (possibly OE2019F, see introduction table #1) can complement, to a limited extent, for the loss of OE1472F.

6.1.3.3.3 Utilization of Phe and Tyr by the OE1471F::pMG501 mutant

The extracellular concentrations of Phe and Tyr were measured in cultures grown in different synthetic media supplemented with AroAAs, Shikimate, DHQ, and combinations of these. Media were inoculated with either R1 or the OE1471F::pMG501 mutant. The extracellular concentrations of Phe and Tyr (figure #38) were measured three times (represented by the symbols circles, triangles and diamonds) from three independent cultures. The experimentally measured changes in Phe and Tyr concentrations were modeled (represented by the dash line) which allowed the calculation of uptake rates in different supplemented media (table #5, done by Dr. Orland Gonzalez). Note that small differences in the concentration of AroAAs are hard to detect, therefore low uptake rats need to be carefully addressed. Nevertheless, the differences in the trends between WT and deletion strain are more trustworthy.

Table # 5: Uptake rates of Phe and Tyr by R1 and OE1471F::pMG501. Cells grow in supplemented synthetic media. The uptake rates were calculated from the corresponding model stimulations.

Synthetic media supplemented with	Phe (nmole OD^{-1} ml^{-1} h$^{-1)}$)		Tyr (nmole OD^{-1} ml^{-1} h$^{-1)}$)	
	R1	OE1471F::pMG501	R1	OE1471F::pMG501
+AroAAs	11.5	7.5	0.9	0.2
+AroAAs+Shikimate	9.6	5.2	1.5	0
+AroAAs+DHQ	13.8	11.8	2.0	2.6
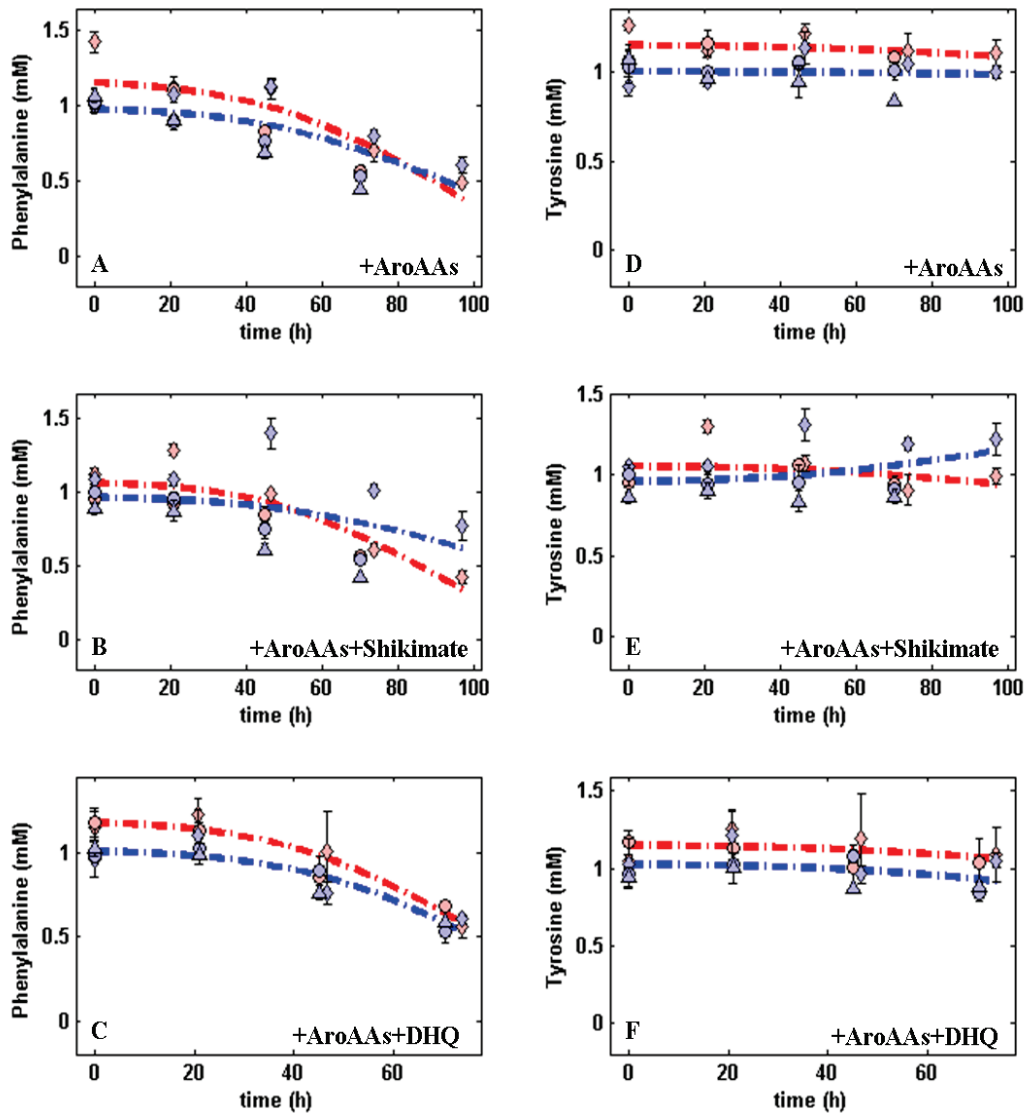
Figure # 38: Phe and Tyr consumption during the growth of WT and mutant OE1471F::pMG501. The experimental data for Phe (left panel, A, B, and C) and Tyr (right panel, D, E and F) are shown using circles, triangles and diamonds, red for WT and blue for the mutant. Error bars represent one standard deviation. Simulations curves are shown using dashed lines, red for WT, and blue for mutant OE1471F::pMG501. The uptake rates were calculated from the corresponding model simulations. Done by Dr. Orland Gonzalez.

As seen in figure #38D-F, the concentration of Tyr remained constant in all media for both the WT and mutant strains. The extracellular concentration of Phe on the other hand, decreased over time, in different media (Figure #38A-C, table#5), indicating uptake of Phe by both WT and the OE1471F::pMG501 mutant. From the data presented in table #5, the uptake rate of Phe by OE1471F::pMG501 was 30-45%

lower than that of the WT in synthetic medium supplemented with either AroAAs, or AroAAs with shikimate. No differences in the uptake rates were seen in synthetic medium supplemented with AroAAs + DHQ.

6.1.3.4 Insertion mutant OE1475F:: pMG601

The central part OE1475F was amplified by PCR, sequenced and cloned in pMKK100, to obtain the integration vector pMG601. It contained the Amp^R, Mev^R markers and the reporter gene BgaH for red/blue screening. Successful integration of this plasmid into the R1 chromosome would slice OE1475F into two pieces, separated by plasmid sequences (figure #39A). As shown in the figure #39A, the insertion is likely to block expression of a functional enzyme, although about 2/3 of the ORF could still be transcribed via the natural promoter.

pMG601 was introduced into R1 cells and 9 transformants were analysed by PCR to verify the single cross-over integration of pMG601. The PCR was designed using one primer which anneals to the beginning of the gene (which is in the chromosomal DNA) and a second primer that targets the Mev^R marker in the plasmid (figure#39A, green arrows). A 1.8Kbp fragment is expected after the integration of pMG601. PCR controls included chromosomal DNA from the WT, and DNA from plasmid pMG601. As shown in figure #39B, a 1.8Kbp fragment was amplified in all 9 transformants, indicating they contained pMG601 integrated into the correct site of the R1 chromosome.
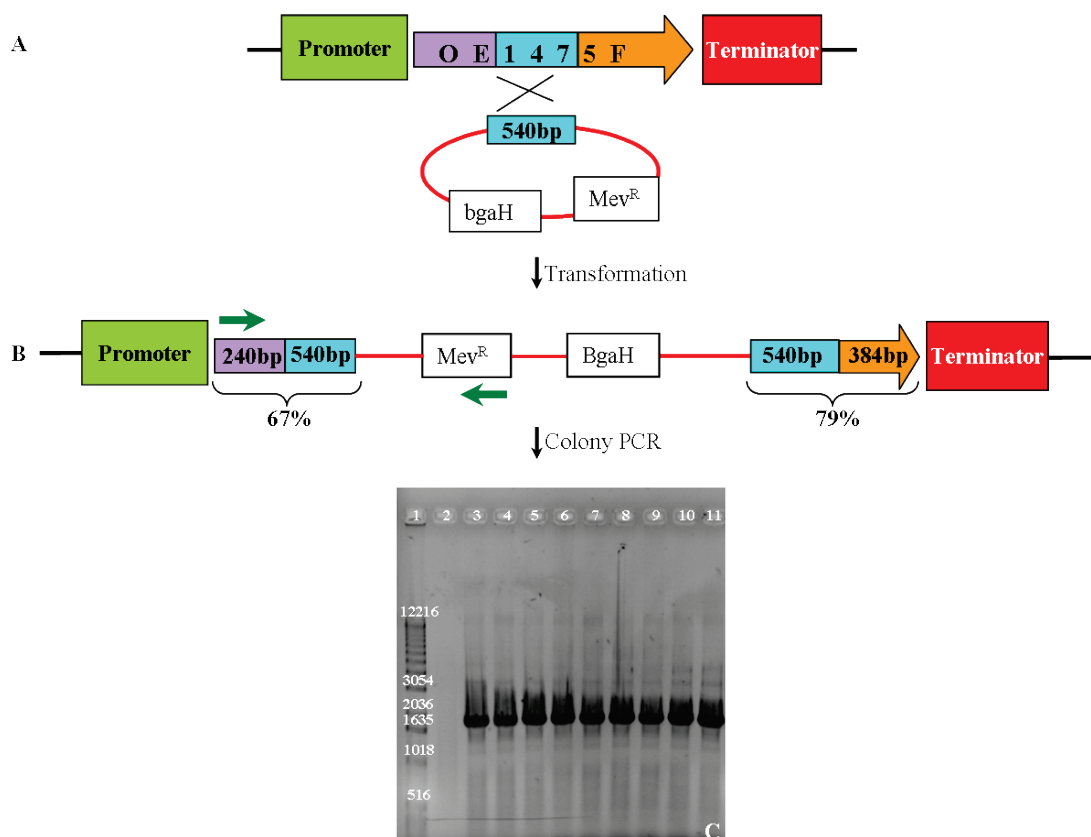
Figure # 39: **A**- Schematic representations of the single crossover event with pMG601. **B**-The expected genotype of the mutant after transformation. The length (in bp) of the sliced gene is indicated. **C**-1% agarose gel of PCR products from chromosomal DNA of R1 (lane 2), and chromosomal DNA from 9 BgH+ colonies after integration of pMG601 into R1 (lane 3-11), MW marker is shown in lane 1 (in bp). The percentages r See the text for details of the primers used.

6.1.3.4.1 Phenotype of OE1475F::pMG601 mutant

To test whether ORF OE1475F plays a role in the AroAAs biosynthesis pathway, the growth of the mutant OE1475F::pMG601 was investigated in synthetic medium alone, or supplemented with 1.1mM AroAAs, and either 1.1mM Shikimate or 1.1mM DHQ. The mutant OE1475F::pMG601 could grow on synthetic medium with only 1.1mM AroAAs, as indicated form the average Gn times recorded for both mutant and WT (figure #40A). However, the growth rate of the mutant was about half that of the WT in synthetic medium without AroAAs (figure #40B), indicating that ORF OE1475F play a role in AroAAs biosynthesis pathway. Both the mutant and the WT grow equally well in synthetic medium supplemented with AroAAs and either shikimate or DHQ (figure #40C-D). The genotype of the WT and the mutant strains was confirmed by PCRs.
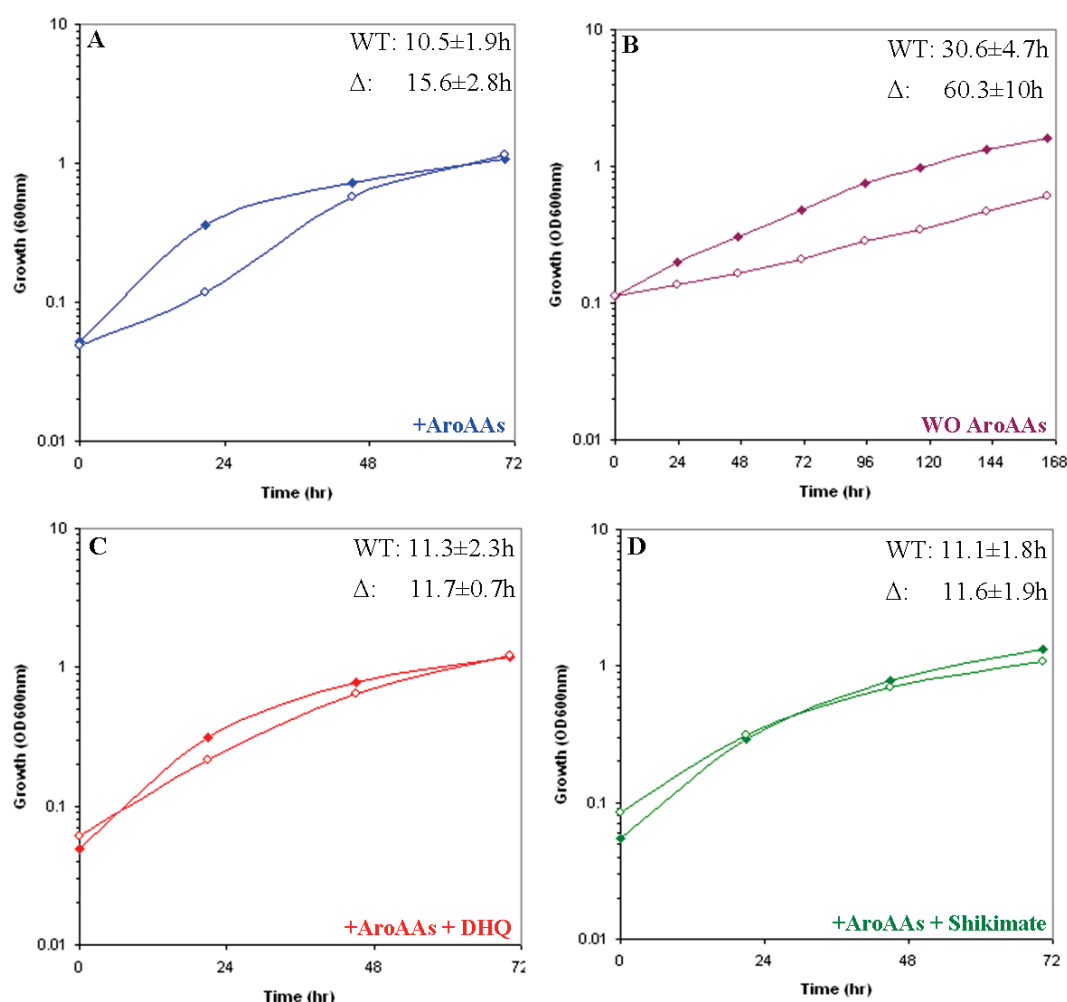
Figure # 40: Growth requirements of mutant OE1475F::pMG601 (△) and WT (▲). The cells were grown in 25ml flasks containing 10ml medium, at 40°C, 100rpm. **A**- synthetic medium with 1.1mM AroAAs, **B**- synthetic medium without AroAAs, **C**-synthetic medium with 1.1mM AroAAs +1.1mM DHQ, and **D**- synthetic medium with 1.1mM AroAAs +1.1mM Shikimate. The cells used for inoculation of C were from synthetic medium with 1.1mM AroAAs. The numbers at the top left are average Gn times. The inoculum was washed with basal salt, before resuspending in the appropriate medium.

6.1.3.4.2 Utilization of Phe and Tyr by OE1475F::pMG601 mutant

As described in the previous section, the effect of medium supplemented with AroAAs, or AroAAs with shikimate on the growth of the OE1475F::pMG601 mutant, indicated it was able to take up AroAAs and use them for growth. To establish that uptake was occurring, the extracellular concentrations of Phe and Tyr were measured during the growth. Amino acid measurements were preformed three times (represented by the symbols circles, triangles and diamonds) from three independent cultures. The experimentally measured concentrations were modeled (represented by

the dash line) which allowed the calculation of uptake rates of Phe and Tyr in different supplemented media (Done by Dr. Orland Gonzalez).

As seen in figure #41D-F, no significant difference in the concentration of Tyr was seen during the growth of either the WT or the mutant. A decrease in the extracellular concentration of Phe was seen in synthetic medium supplemented with either AroAAs+shikimate or AroAAs+DHQ, for both WT and mutant OE1475F::pMG601 (figure #41B,C). No change in the concentration of Phe was observed in mutant OE1475F::pMG601 grown in synthetic medium supplemented with only AroAAs. From the data presented in table #6, the uptake of the mutant grown in synthetic medium supplemented only with AroAAs, was 12% (1.3nmole Phe OD^{-1} ml^{-1} h^{-1}) of the uptake rate of the WT grown in the same media. Note that small differences in the concentration of AroAAs are hard to detect, therefore low uptake rats need to be carefully addressed. Nevertheless, the differences in the trends between WT and deletion strain are more trustworthy.

Table # 6: Uptake rates of Phe and Tyr by R1 and OE1475F::pMG601. Cells grow in supplemented synthetic media. The uptake rates were calculated from the corresponding model stimulations.

Synthetic media supplemented with	Phe (nmole OD^{-1} ml^{-1} h^{-1})		Tyr (nmole OD^{-1} ml^{-1} h^{-1})	
	R1	OE1475F::pMG601	R1	OE1471F::pMG501
+AroAAs	10.8	1.3	0	0
+AroAAs+Shikimate	9.1	12.2	0	0
+AroAAs+DHQ	11.9	16.5	2.9	4.2
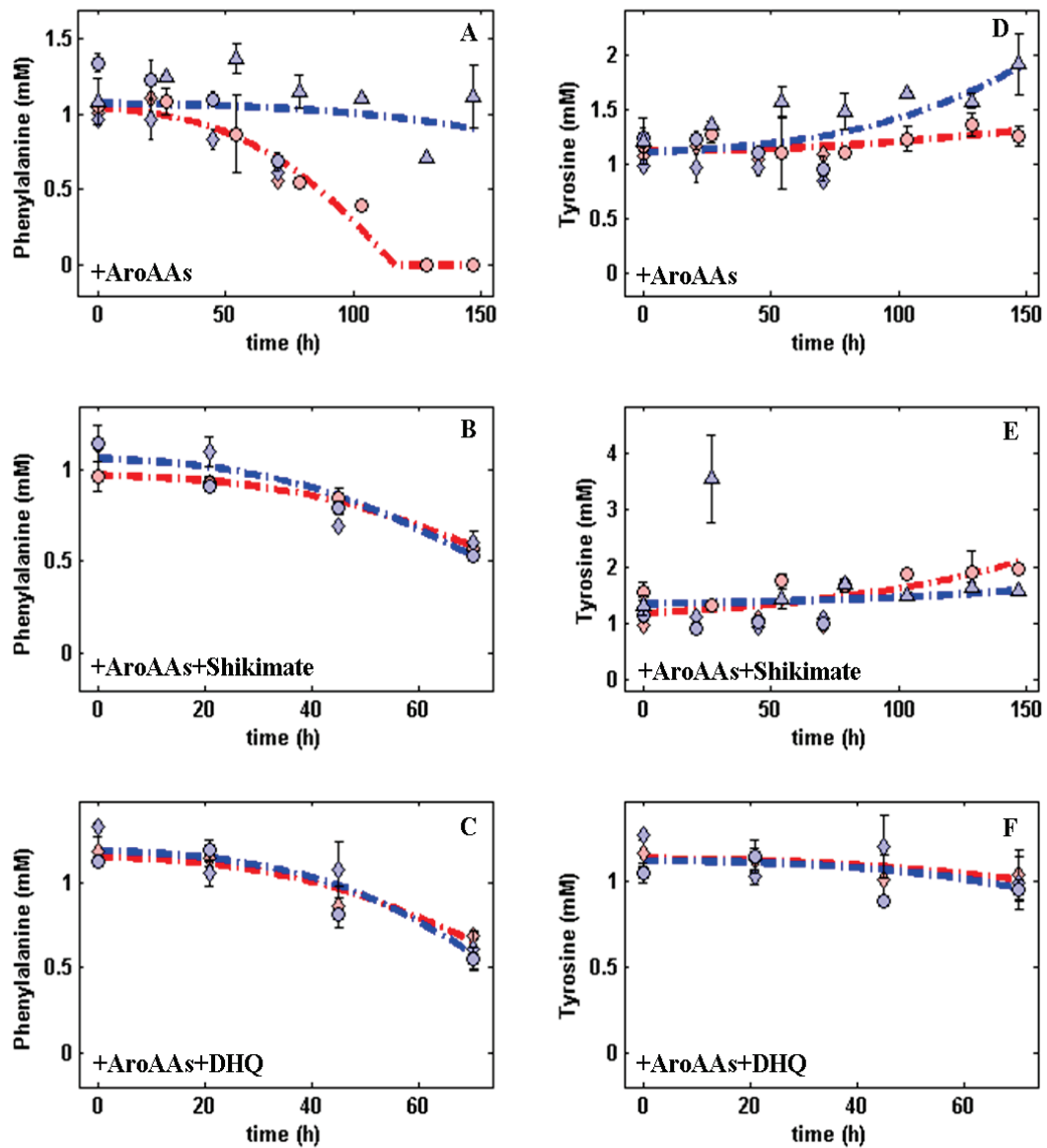
Figure # 41: Phe and Tyr consumption data during the growth of WT and mutant OE1475F::pMG601. The experimental data for Phe (left panel, A, B, and C) and Tyr (right panel, D, E and F) are shown using circles, triangles and diamonds, red for WT and blue for the mutant. Error bars represent one standard deviation. Simulations curves are shown using dashed lines, red for WT, and blue for mutant OE1475F::pMG601. Note the different time scales in the left panel. The uptake rates were calculated from the corresponding model simulations. Done by Dr. Orland Gonzalez.

6.2 Genome-wide transcriptional studies (microarray)

Regulation of gene expression is essential for normal growth, replication of cells, and for adaptation to changes in the environment. By expressing certain genes only when necessary, cells conserve both energy and resources. To study the differential gene expression of all assigned genes for the biosynthesis of AroAAs, a DNA microarray covering 3072 ORFs of *H. salinarum* R1 was used (Twellmeyer et al., 2007). This was expected to assist in identifying the pathways of AroAAs production, as well as shed a light on associated regulatory or metabolic genes.

In the first series of experiments (see below), cell samples of *H. salinarum* grown in synthetic medium with and without AroAAs were taken at two different growth phases (OD_{600nm}=0.2, and 0.58). RNA was isolated and used to prepare cDNA. The cDNA from cells grown in medium without AroAAs (at specific OD) was compared to the cDNA prepared from cells grown in medium with AroAAs on a DNA microarray. In the second series of experiments (section 6.2.1.6), the cDNA prepared from WT cells was compared to cDNA of mutant ΔOE1477R (*aro*D, 3rd gene in the pathway, see figure #19) both grown in synthetic medium with AroAAs and shikimate. In all experiments, the criterion, which was used to identify whether a gene was regulated or not, was a combination of statistical significance and fold induction (see materials and methods section 5.3.10.3).

In the first comparison, using cDNA's from early exponential phase cells grown with and without AroAAs, 898 ORFs gave signals that were either below background or rejected for technical reasons, 1645 ORFs gave variable signals with P-values (calculated from the replicate spots) above 10^{-3}, and considered not significant, while 529 ORFs gave statistically significant signals ($P \leq 10^{-3}$). As seen in table #7, within the latter group, 36% were conserved hypothetical proteins (CHY) or hypothetical proteins (HY). 360 ORFs showed no change in the transcription level (fold change $\leq$1.5), while a fold-change of 4.0 or higher was observed for 40 ORFs in synthetic medium without AroAAs relative to synthetic medium with AroAAs.

Table # 7: Distribution of significantly regulated ORFs ($P \leq 10^{-3}$) in *H. salinarum* grown in synthetic medium without AroAAs, observed at OD_{600nm}=0.2 and 0.58.

Fold induction	OD_{600nm}=0.2*			OD_{600nm}=0.58*		
	Assigned ORFs	CHY	HY	Assigned ORFs	CHY	HY
≥ 4.0	26	11	3	11	4	1
1.5-4.0	64	49	16	70	33	12
≤ 1.5	248	92	20	215	120	37
Total	**338**	**152**	**39**	**296**	**157**	**50**
Total	**529**			**503**		

*-The AroAAs-related genes regulated at OD_{600nm}=0.2 and 0.58, are shown in table #7

6.2.1 Transcription of AroAAs-related genes

The transcription of AroAAs-related genes of *H. salinarum* grown in synthetic medium without AroAAs relative to with AroAAs (OD_{600nm}=0.2 and 0.58) is represented in table #8.

In the **chorismate pathway**, high values of transcriptional regulation were observed only for *aro*K, *aro*A, and *aro*C (3.6-6 fold induction). *aro*E and *aro*D gave a low value (1.3, 1.6 fold induction, respectively), while no changes were detected in ORFs OE1472F and OE1475F. In the case of ORF OE1472F, the intensities obtained were below background, while for ORF OE1475F no significant regulation was found (fold induction=0.96, p=0.318). It was shown that OE1472F, OE1475F and OE1477R (*aro*D) are essential for AroAAs biosynthesis, so the transcription of these ORFs, was assessed as well by an independent method, RT-qPCR. As seen in table #8, the expression OE1472F was found to be strongly regulated (12.5±0.7 fold induction), as calculated from two biological repeats. On the other hand, fold induction of 1.5±1.2, and 2.3±1.7 was observed for OE1475F and OE1477R, respectively, indicating constitutive transcriptions of these genes.

In the branch leading from **chorismate to tryptophan,** all the 7 ORFs were significantly induced at early exponential phase growth, while a decrease in the transcription was seen in the mid-exponential phase ($OD_{600nm}=0.58$). The parallel changes in gene expression and their genomic position indicate that the 7 ORFs are organized into two separated operons, consisting of: trpCBA-OE1472F, and trpD$_1$FE$_1$G$_1$ (see figure #81A and B).

In the proposed de novo pathway for AroAAs biosynthesis, the first step is catalyzed by the OE1472F protein, which cleaves a six carbon substrate, DKFP, followed by condensation with ASA. According to the genomic data of *H. salinarum* (www.halolex.mpg.de), ASA can be synthesized either from homoserine (involving ORF OE4722R) or L-aspartate (involving ORFs OE4333R and OE3063F), but the microarray data in the current study showed no changes in the transcription of those ORFs. Moreover, no regulation was observed for ORF OE2500R, which was predicted to synthesize methyllglyoxal from GAP (the first is the precursor for DKFP), indicating constitutive induction of theses genes.

6.2.1.1 Transport systems for AroAAs

Five transport systems of AroAAs are known in *E. coli* (Ikeda, 2006): (1) *mtr* and *tnaB* which are responsible for the transport of tryptophan, (2) *tyrP* for the transport of tyrosine, (3) *pheP* for the transport of phenylalanine and (4) *aroP* for the transport of all aromatic amino acids. No strong homologues were found for *mtr*, *tnaB*, *tyrP* or *pheP* in *H. salinarum*.

ORF OE2779F (homolog to *aro*P and *phe*P) was predicted to be an amino acid transport protein. This ORF was 3.5 fold up regulated in R1 cells grown without AroAAs and harvested at $OD_{600nm}=0.2$. Similar up regulation (fold=3.3) was found for this ORF in cells harvested at $OD_{600nm}=0.58$, indicating that the OE2779F protein could be responsible for the transport of AroAAs, and is regulated when no AroAAs are present in the medium. Unlike the other genes shown in the table #7, there was no change in the degree of regulation of this ORF at different growth phases, suggesting that AroAAs are still been taken up at mid-exponential phase growth. Moreover, ORF OE4384R (a weak homolog of *tna*B) was 4 fold up regulated ($p=2.4\times10^{-4}$) in the deletion strain ΔOE1477R (synthetic medium with AroAAs+shikimate at

$OD_{600nm}=0.2$). This ORF was also 2.6 fold up regulated in WT cells grown in synthetic medium supplemented with AroAAs.

6.2.1.1.1 ABC transporters and Na$^+$-dependent transporters

Two categories of ABC transporters exist: ABC-transporters involved in the uptake of solutes from the medium, and transporters which export metabolites to the medium. Both types of transporters couple solute movement to ATP hydrolysis. In the annotation of microbial genome sequences, ORFs representing ABC transporters are usually reliably predicted, but the specific substrate of individual ABC transporters is much more difficult to predict by sequence homology, and additional evidence is needed to permit more accurate assignments. Keeping this in mind, it was found that seven ORFs predicted to be ABC transporters were strongly up regulated when AroAAs were omitted from the synthetic medium. These were annotated as oligopeptides/dipeptide transporters which would be consistent with the cells trying to scavenge AroAAs from the medium (table #8). Note that some of the ABC-transport genes are arranged in an operon, the dppA$_3$B$_3$C$_3$D$_3$F$_3$ operon, seen in figure #42E. All the genes in this operon, except for ORF OE4317F (dppD3), were significantly regulated. This operon was not found to be regulated either in ΔOE1477R mutant or WT cells grown in synthetic medium supplemented with AroAAs and shikimate.

In addition to the ABC-transporters, three ORFs annotated as Na$^+$-dependent transporters were strongly regulated, when grown R1 in synthetic medium without AroAAs, as seen in table #8. Na$^+$-dependent transporters may also be able to transport amino acids, using the Na$^+$ gradient produced during phototrophic growth as a motive force ((Lanyi, 1977) and references within).

Table # 8: Expression of AroAAs-related genes and transport-related genes in *H. salinarum* R1 cells grown in synthetic medium without AroAAs relative to synthetic medium with AroAAs.

ORF	Gene	[1] **Predicted function**	[2] **Fold induction at OD$_{600nm}$**	
			0.2	0.58
Precursors ⟶ DHQ				
OE1472F	*fba*2	2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase	[3] **12.5±0.7**	----
OE1475F	------	3-dehydroquinate synthase (EC 4.2.3.4) type II	0.96, NS [3] **1.5±1.2**	----
DHQ ⟶ Chorismate				
OE1477R	*aro*D	3-dehydroquinate dehydratase (EC 4.2.1.10)	1.6 [3] **2.3±1.7**	1.3
OE1565F	*aro*E	shikimate 5-dehydrogenase (EC 1.1.1.25)	1.35 NS	2.56 NS
OE2785R	*aro*K	shikimate kinase (EC 2.7.1.71), archaeal-type	6.0	4.4
OE2762R	*aro*A	3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19)	3.6 (P=0.003)	2.1
OE2761R	*aro*C	chorismate synthase (EC 4.2.3.5)	6.3	4.4
Chorismate ⟶ Prephenate				
OE2784R	*phe*A1	chorismate mutase (EC 5.4.99.5)	2.0	3.8
Prephenate ⟶ Phenylpyruvate				
OE4117F	*phe*A2	prephenate dehydratase (EC 4.2.1.51)	0.7	1 NS
Prephenate ⟶ 4-hydroxyl Phenylpyruvate				
OE2770F	*tyr*A	prephenate dehydrogenase (EC 1.3.1.12)	6.5	2.7

Chorismate ⟶ Tryptophan				
OE3331R	*Trp*G1	anthranilate synthase (EC 4.1.3.27) component II	2.1	1.4
OE3332R	*Trp*E1	anthranilate synthase (EC 4.1.3.27) component I	10	4.1
OE3334R	*Trp*D1	anthranilate phosphoribosyltransferase (EC 2.4.2.18)	6.2	4.5
OE3333R	*Trp*F	phosphoribosylanthranilate Isomerase (EC 5.3.1.24)	8.5	ND
OE1469F	*Trp*C	indole-3-glycerol-phosphate synthase (EC 4.1.1.48)	7.7	
OE1470F	*Trp*B	tryptophan synthase (EC 4.2.1.20) beta subunit	8.9	ND
OE1471F	*Trp*A	tryptophan synthase (EC 4.2.1.20) α subunit	6.7	3.6
Transport system of AroAAs				
OE2779F	*aro*P	amino acid transport protein (probable phenylalanine transport protein)	3.5	3.3
Sodium dependent transporters				
OE3173F		sodium-dependent transporter	3.2	13.2
OE5143F		sodium-dependent transporter	12.5	ND
OE3511F	*pan*F	probable sodium/pantothenate symporter	4.5	ND
ABC-transporters				
OE4311F	*dpp*A3	ABC-type transport system periplasmic substrate-binding protein (probable substrate dipeptide/oligopeptide/nickel)	5.6	ND
OE4313F	*dpp*B3	ABC-type transport system permease protein (probable substrate : dipeptide/ oligopeptide/nickel)	4.2	ND
OE4316F	*dpp*C3	ABC-type transport system permease protein	4.9	ND
OE4317F	*dpp*D3	ABC-type transport system ATP-binding	ND	ND

		protein (probable substrate : dipeptide/ oligopeptide/nickel)		
OE4318F	*dpp*F3	ABC-type transport system ATP-binding protein (probable substrate :dipeptide/ oligopeptide/nickel)	4.2	1.3
OE4555F	*dpp*C2	ABC-type transport system permease protein (probable substrate : dipeptide/ oligopeptide/nickel)	3.2	1.6
OE3631F	*pot*A	ABC-type transport system ATP-binding protein (probable substrate: spermidine/ putrescine)	2.9	ND

[1] ORF predictions are derived from the genome annotation at www.halolex.mpg.de.

[2] The numerical data represent mean expression ratios in WT cells grown in medium without AroAAs vs cells grown in medium with AroAAs. These were determined at two different cell densities. All fold changes had p values of $\leq 10^{-3}$, except were indicated.

[3] Values in bold were determined by RT-qPCR. For all calculation the average-C_t (cycle threshold) of 2 replicate reactions per primer pair was used. The standard deviation was calculated from two biological repeats.

ND-not detected

NS- not statistically significant (p>0.01)

6.2.1.2 Conserved hypothetical proteins and hypothetical proteins

Although the genome sequence of *H. salinarum* R1 has been determined and carefully annotated, there remains a considerable fraction of the ORFs that cannot be assigned to function due to the lack of sequence homologs of known function. These are annotated as hypothetical (HY) or conserved hypothetical (CHY) proteins. If these are real genes, then they should be transcribed (at least under some environmental conditions), and this can be detected using DNA microarrays which allow parallel investigation of all genes involved in controlling the physiological state of a culture.

As shown in table #9, 7 ORFs annotated as CHY or HY were found to be significantly up regulated in the absence of AroAAs in the growth medium. These data not only show transcription of these hypothetical ORFs but also that they are likely to be important in the cell response to AroAAs limitation. It was also found that some members of this group are found in close proximity to each other, and are likely to be co-transcribed, i.e. form operons (figure #82A-D).

Table # 9: Transcription of hypothetical (HY) and conserved hypothetical (CHY) proteins in *H. salinarum* cells grown in synthetic medium without AroAAs relative to synthetic medium with AroAAs.

ORF	[1]Predicted function	[2]Fold induction	[3]Operon
OE2776F	probable transcription regulator (*Lrp*A2)	14.9	
OE1775R	**CHY**	11.1	
OE4414R	**CHY**	3.8	
OE4412R	threonine synthase (EC 4.2.3.1)	3.2	
OE2770F	prephenate dehydrogenase (EC 1.3.1.12) (*tyr*A)	6.5	A
OE2772F	**CHY**	17.9	A
OE4418R	**HY**	9.5	B
OE4419R	argininosuccinate lyase (EC 4.3.2.1) (*arg*H)	3.4	B
OE4420R	argininosuccinate synthase (EC 6.3.4.5) (*arg*G)	12	B
OE1772A1R	**CHY**	4.2	C
OE1774R	**CHY**	6.0	C
OE1448R	**CHY**	8.0	D
OE1447R	probable metallo-beta-lactamase family hydrolase	6.1	D

[1] ORF predictions are derived from the genome annotation at www.halolex.mpg.de.

[2] The numerical data represent induction ratios in WT cells grown in medium without AroAAs *vs.* cells grown in medium with AroAAs. These were determined at cell density of $0.2\text{OD}_{600\text{nm}}$. All fold changes had p values of $\leq 10^{-3}$.

[3] ORFs organized in operons, see figure #42.
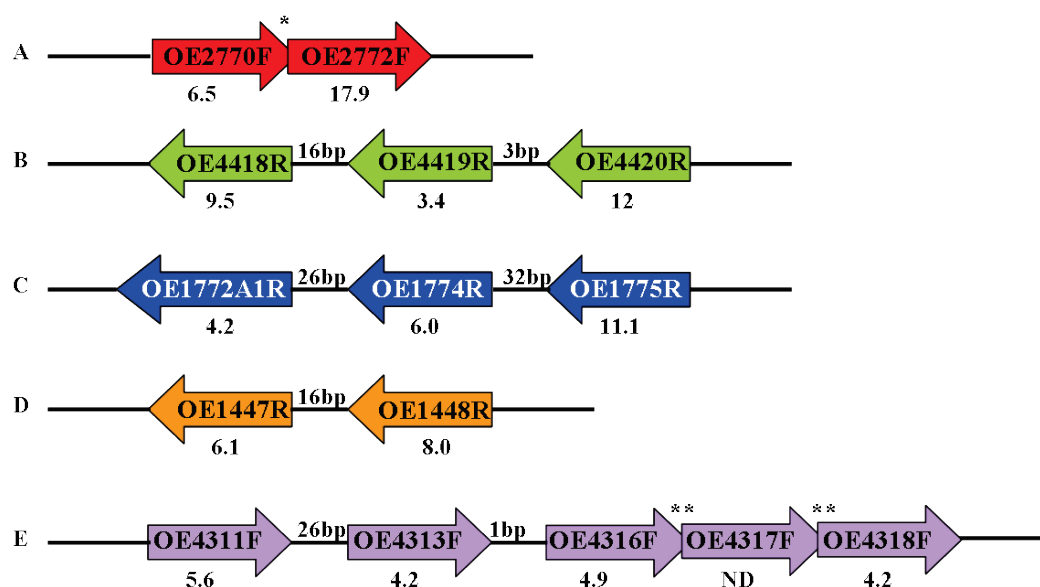
Figure # 42: AroAAs regulated CHY and HY ORFs that are closely adjacent, and possibly organized in operons. The fold regulations are indicated below. * , **-represents a 10bp, and 3bp overlapping regions, respectively. A-D see table #8.

6.2.1.3 Growth of ΔOE1477R in synthetic medium with AroAAs plus shikimate

When cDNA of WT *H. salinarum* was hybridized against cDNA of the deletion strain ΔOE1477R (page 76, figure #28), both grown in synthetic medium with AroAAs and shikimate, none of the AroAAs-related genes shown in table #8 were regulated in the mutant compared to the WT. This is in agreement with the uptake assay (table #3, and figure #31) showing similar uptake rates of Phe by both the WT and the mutant, which can support the growth without activating the AroAAs biosynthesis pathway, or by converting Shikimate to AroAAs. On the other hand, several CHY, HY, and transposase were significantly induced in the mutant (table #10).

Table # 10: Expression of genes in ΔOE1477R mutant cells relative *H. salinarum* cells grown in synthetic medium with AroAAs and shikimate.

ORF	Gene	(1) Predicted function	(2) Fold induction at OD$_{600nm}$	
			0.2	0.69
OE2469F	flgA1	flagellin A1	2.2	3.1
OE4073R	hcpB	halocyanin hcpB	3.5	1.6
OE4390F	polA1	DNA-directed DNA polymerase (EC 2.7.7.7) type III subunit 1	6	ND
OE7033R	gvpD1	regulatory protein gvpD	2.9	ND
Transposase (ISH3)				
OE5027R		transposase (ISH3)	1.9	3.9
OE5225R		transposase (ISH3)	2.2	3.2
OE5383F		transposase (ISH3)	ND	3.2
OE7048F		transposase (ISH3)	1.8	4.3
CHY				
OE1388R		conserved hypothetical protein	ND	3.8
OE1416F		conserved hypothetical protein	0.7	15.8
OE4384R		conserved hypothetical protein	4.0	1.5
OE5134F		conserved hypothetical protein	9.6	ND
OE6130F		conserved hypothetical protein	2.9	4.2
OE7003R		conserved hypothetical protein	3.2	ND
HY				
OE4077F		hypothetical protein	3	ND
OE6020R		hypothetical protein	3.8	ND

[1] ORF predictions are derived from the genome annotation at www.halolex.mpg.de.

[2] The numerical data represent mean expression ratios in ΔOE1477R cells grown in synthetic medium with AroAAs and shikimate vs WT cells grown in the same medium. These were determined at two different cell densities. All fold changes had p values of $\leq 10^{-3}$, except were indicated.

ND-not detected

6.3 Expression and purification of proteins encoded by OE1472F, OE2019F & OE1477R

Both OE1472F and OE2019F were annotated as fructose bisphosphate aldolase (EC 4.1.2.13, Figure#18), but only ORF OE1472F is believed to be involved in the AroAAs biosynthesis pathway (see introduction). ORF OE1477R, annotated as 3-dehydroquinate dehydratase (EC 4.2.1.10, figure #19), is the 3rd ORF in the pathway.

To study in detail the structural and functional properties of the proteins encoded by OE14772F, OE2019F and OE1477R, they needed to be purified. An overexpression system was essential because the natural level in the cells was probably low, and no assays were available that could be used to rapidly identify active fractions. It is known that heterologous overexpression may fail, or be inefficient for a number of reasons, such as the absence of cofactors, use of rare codons and consequent poor translation, differences in co- or post translation modification and processing, and improper protein folding due to different physiological conditions (such as presence of high salt). In consideration of those issues, an expression system in *H. salinarum* was tested first, using the vector pBPH-M (Besir, 2001). Gene expression was controlled by the inducible bacteriorhodopsin (BR) promoter that is highly active when cells are grown under illumination (white light) and/or oxygen depletion.

6.3.1 Expression of OE1472F in *H. salinarum*

A C-terminal His-tagged OE1472F was constructed by a one step PCR using the proofreading DNA polymerase LA-Taq. The 795bp gene, without the stop codon, was amplified by PCR, sequenced and subcloned into the *E. coli-H. Salinarum* shuttle vector, pBPH-M (figure #43A).

The pBPH-M vector contains an inducible bop promoter, a 7xHis C-terminal purification tag, a Mev^R marker and part of OE3349F, to allow integration into the chromosome. The OE1472F-His $_{C-7xhis}$/pBPH-M recombinant vector (pMG550) was transformed into *H. salinarum* and plated on media containing Mev. pMG550 can integrate into the chromosome using either OE1472F or OE3349F as a recombination site, but only recombination with OE3349F will allow OE1472F to be overexpressed and carry a His Tag. To differentiate between the possible cross-over events, a PCR

was designed using primers which target both the bop promoter and OE3349F (red arrows, figure #43A). No amplification is expected for either a WT or an OE1472F recombinant, while a 972bp fragment is expected for an OE3349F recombinant.

Total DNA was isolated from liquid cultures (Halo medium+10µg/ml Mev) of 5 transformants and used for PCR with the primers shown in figure#43A. The gel analysis (figure #43B) shows that a 972bp fragment was amplified in four out of five transformants. PCR controls included chromosomal DNA from the WT, and DNA from plasmid pMG550.
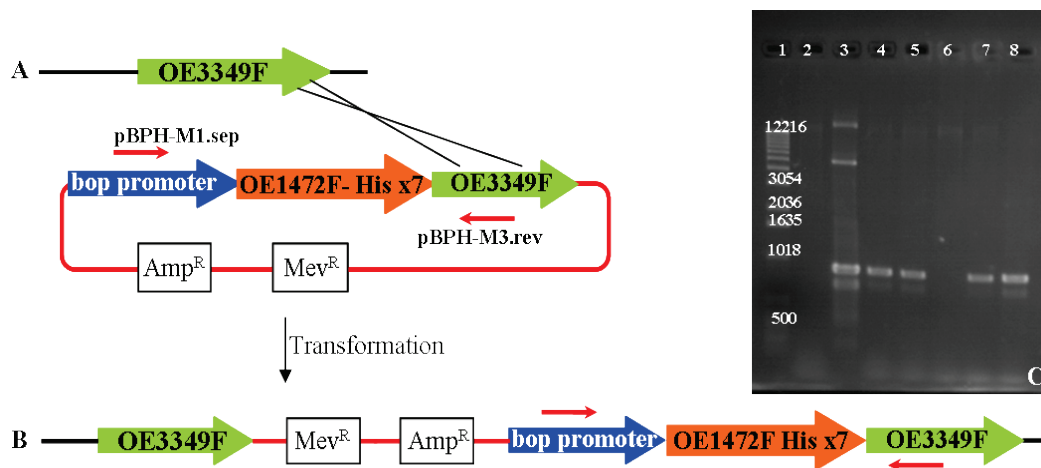
Figure # 43: The desired integration of pMG550 into the chromosome of *H. salinarum.* **A-** Schematic illustration of the integration. **B-** The expected genotype in the case of successful transformation. **C-** 1% agarose gel of PCR products using template DNA from WT (lane 2), pMG550 (lane 3) and transformants E1-E5 (lanes 4-8). MW markers are shown in lanes 1 (in bp).

All four transformants with the correctly integrated plasmid (E1-2, E4-5) were tested for the overexpression of His-tagged OE1472F. They were grown to OD_{600nm}=1.0 under oxygen depleted conditions to activate the bop promoter, cell lysates prepared, and proteins separated by SDS-PAGE (figure #44A). A similar gel was used for Western blotting and the His-tagged proteins were detected using an anti His-tag antibody (figure #44B). The calculated molecular weight of 7xHis tagged OE1472F is 29,055 Da, and as seen in figure #44B, E4 cells showed overexpression of proteins with molecular weight of ~30-40KDa. To identify this tagged protein, 8.5ml of the cell-free extract in NPI_{10}-KCl was loaded twice on a 1ml NiNTa affinity column equilibrated with the same binding buffer. The unbound protein was washed

away, while the bound protein eluted over 30min linear gradient between NPI_{10}-KCl and NPI_{500}-KCl.
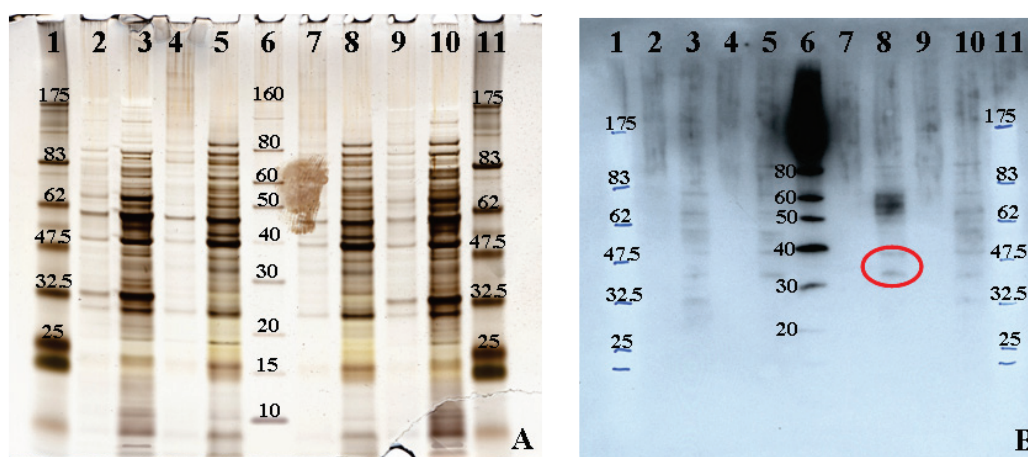
Figure # 44: Overexpression of 7xHis tagged OE1472F in *H. salinarum*. **A-** 4-12% polyacrylamide SDS-PAGE and **B-** Western blot of 1μg cell-free extract of E1, E2, E4 and E5 (lanes 2,4,7 and 9, respectively), and 10μg cell-free extract of E1, E2, E4 and E5 (lanes 3,5,8 and 10, respectively). Lanes 1 and 11- prestained protein marker (in KDa), lane 6-His tagged marker.

The tagged protein eluted in a single peak over an Imidazole concentration range of 60-143mM. A total of 0.132mg/ml protein was eluted, which was 6% of the total protein loaded on the column. 10μg of total protein from each elution fraction was separated on a 4-12% SDS-PAGE, and the gel stained by coomassie blue (figure #44A). The proteins were transferred to a PVDF membrane and probed with anti His antibody (figure #45A). As shown in figure #45A, the eluted protein was relatively pure.

However, according to N terminal sequencing, the bands circled in red had an amino acid sequence (TDAASL), which corresponded to hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34, OE3637R). This has a MW of 41067Da, and contains three histidine residues. Elution fractions 15 and 16 were subjected to mass spectrometry, revealing a main mass of 40286.9 Da, and 37559.7 Da, respectively. Neither of these values were the expected mass of the tagged OE1472F, i.e. 29055Da.

102
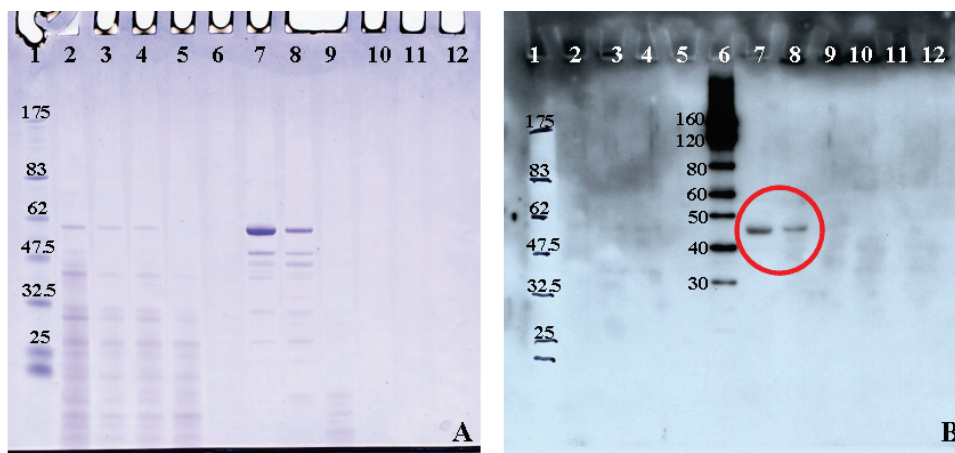
Figure # 45: Purification of OE3637R from cell-free extracts. **A**-4-12% polyacrylamide SDS-PAGE, and **B**-Western blot of: **lane 1**: prestained protein marker (in KDa), **lane2**:1st load of NiNTa column, **lane3**: 2nd load of NiNTa column, **lane4**: wash fraction washed with NPI$_{10}$-KCL buffer, **lane5**: His marker (in KDa), **lanes 6-12**: elution fractions 15-20, eluted in 30min linear gradient of NPI$_{10}$-KCL-NPI$_{500}$-KCL buffer.

To exclude the possibility that the His-tag in OE1472F was buried in the 3D structure of the expressed protein, and therefore was unable to bind to the NiNTA column, 1ml of the unbound cell-free extract from culture E4 after the second load (figure #45A, lane 3) was denatured with 8M urea and mixed with 1ml of NiNTa beads. The mixture was incubated overnight at 4°C, followed by batch separation. The beads were washed with NPI$_{20}$+8M urea, and elution was done in a step gradient with NPI$_{500}$+8M urea. Fractions collected from the purification were then analysed by SDS-PAGE, which was then stained with coomassie blue (figure #46A), and with a silver stain (figure #44C). To detect His-tagged proteins, the gel was also processed for Western blot analysis (figure #46B). Although the protein eluted in fractions 2-4 appears to be His-tagged, the molecular weight of the main band according to the ESI-MS was found to be 27927.5Da, which was 1127Da shorter than expected for the tagged protein.

Further attempts to over-express tagged OE1472F in cells grown under white light illumination (in order to induce the bop promoter), were unsuccessful in detecting His-tagged proteins on the western blot (results not shown, F189, F190). Given the difficulties found with this expression system, it was decided to try a more conventional expression system (see next section).
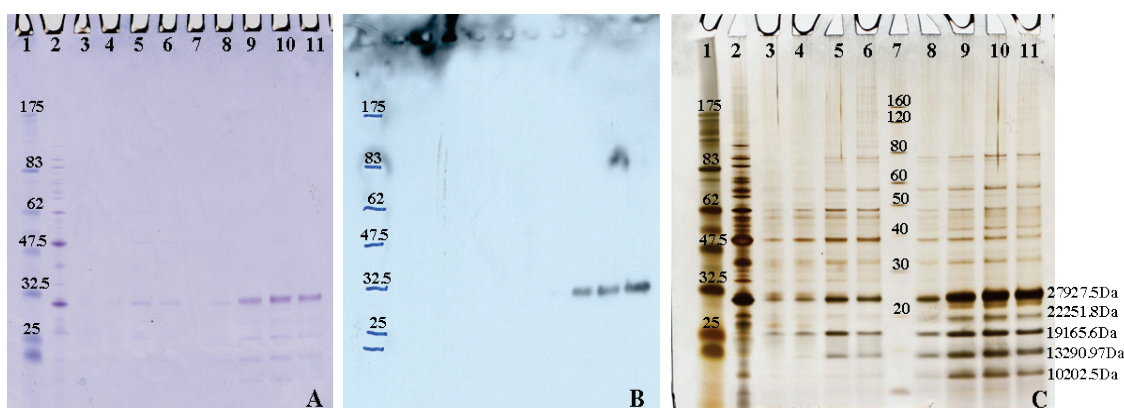
Figure # 46: Purification of over-expressed OE1472F after denaturation. A 4-12% polyacrylamide SDS-PAGE stained with **A**-Coomassie blue and **C**- with silver stain. **B**-Western blot. **Lane1**: pre-stained protein marker (in KDa), **lane 2**: crude cell extract, **lane3**: 1st flow through fraction, **lanes 4-6**: wash fractions 1, 2 and 7, in NPI_{20}-8M urea, **lanes 8-11**: elution fractions 1-4, eluted in step gradient with NPI_{500}-8M urea. 10μg (according to OD_{280nm}) were loaded on the SDS-PAGE.

6.3.2 Cell-free expression of OE1472F, OE2019F and OE1477R

Due to the difficulties in expressing OE1472F in *H. salinarum*, we tried next the *E. coli* as expression system, and as a first step, an in-vitro transcription-translation system was tested for its ability to express full-length haloarchaeal proteins. The EasyXpress System Kit (Qiagen) was selected for this purpose.

The genes of interest were amplified by PCR from R1 chromosomal DNA, and sequenced to check for errors. They were then cloned into TOPO vector, and subcloned into pET22b(+) (Novogen), using ElectroMax DHB10 cells. The transformants were checked for plasmids containing the correct insert, 6xHis-tag and the stop codon. These were then used for propagation and large scale isolation of plasmids OE1472F-His $_{C-6xhis}$/ pET22b(+), OE1477R-His $_{C-6xhis}$/ pET22b(+), and OE2019F-His $_{C-6xhis}$/ pET22b(+).

Expression of the three genes was tested by adding the purified plasmids to the transcription-translation reactions (EasyXpress protein Synthesis Kit, Qiagen), following the manufacturer's recommendations. In-vitro expression was analysed by SDS-PAGE and Western blot (figure #47A and B, respectively). As shown in figure# 47B, plasmids OE1472F-His $_{C-6xhis}$/ pET22b(+), OE1477R-His $_{C-6xhis}$/ pET22b(+), and OE2019F-His $_{C-6xhis}$/ pET22b(+) directed the synthesis of correctly sized products, at readily detectable levels, and the expressed proteins were indeed His-tagged. A ladder-like pattern of faint bands was also seen in the western blot (figure #47B, lanes

8-12), including in the negative control (which contains no DNA template), and presumably represent *E. coli* proteins that react weakly with the anti His-tag antibody.

The detected molecular weights of all the ORF's were in the range of the calculated molecular weights of 28,298Da, 24,244Da, and 29,279Da (OE1472F, OE1477R and OE2019F, respectively). It is known that haloarchaea proteins migrate slower on SDS-PAGE due to their high content of acidic amino acids. While the three proteins being expressed were very acidic, the percentage of acidic residue varied (23.7%, 17.2%, and 16.7%, respectively). Together, the data indicate that OE1472F, OE1477R and OE2019F can be expressed in *E. coli*, as full-length His-tagged proteins.
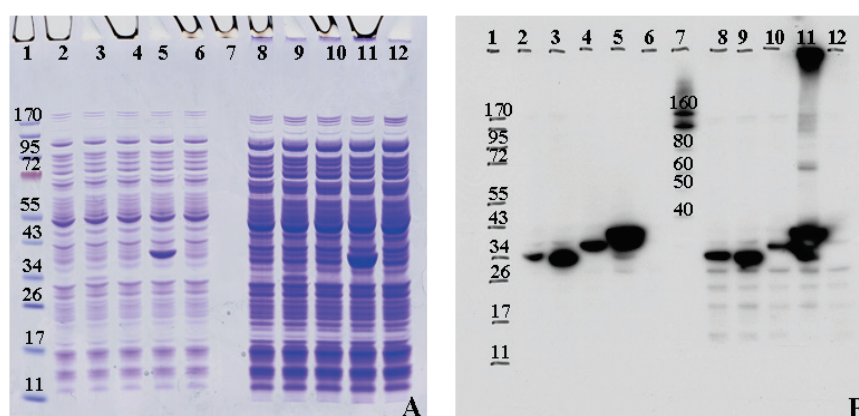
Figure # 47: Analysis of in-vitro synthesized products directed by OE1472F, OE1477R and OE2019F. **A**-SDS-PAGE (4-12%), **B**-western blot of **lane 1**: prestained protein marker (KDa), **lanes 2-4**: 1µl expression mixture of OE1472F, OE1477R, OE2019F, **lanes 5-6**: positive control, and negative control (no template), **lane 7**: His-tagged marker, **lanes 8-10**: 5µl expression mixture of OE1472F, OE1477R, OE2019F, **lanes 11-12**: positive control, and negative control (no template).

6.3.2.1 Batch purification from cell-free extracts

The expressed proteins generated in the cell-free reaction mixtures were purified via their His-tags using a batch protocol. The cell-free reaction mixtures (32µl) were diluted 1:3 with NP buffer and mixed with NiNTa beads (2hr RT), which were equilibrated with the same buffer. The unbound fraction was removed by centrifugation (1min, 5000g, RT), and the beads washed five times with NPI_{20}. The bound protein was eluted using a two step gradient: first between 20-200mM Imidazole, and second between 200-500mM Imidazole. Proteins were analysed by SDS-PAGE and western blot. As shown in figure #48, His-tagged proteins were

detected in Western blots (figure #48D-F), and found to elute between Imidazole concentrations of 20-200mM. Some differences in the relative yields of the purified proteins were observed. A moderate yield was seen for OE1472F, while a high yield was obtained for OE1477R and a very low yield for OE2019F. None of the proteins were visible by coomassie blue staining, so the expression was poor in absolute terms, but expression was possible in this in vitro system, and it was expected that in vivo expression in *E. coli* would be more efficient.
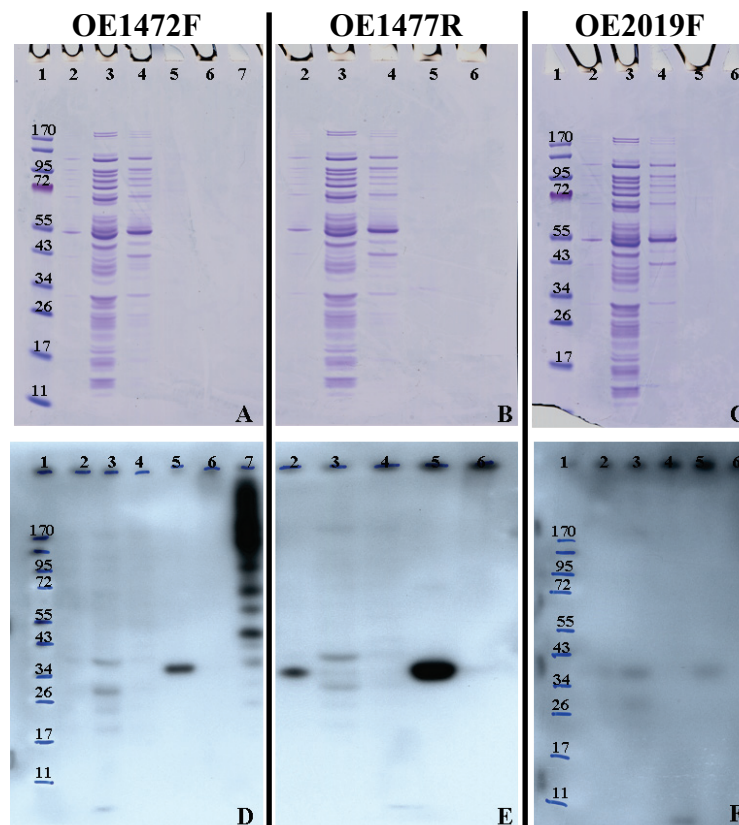
Figure # 48: Purification of in-vitro His-tag synthesized proteins OE1472F, OE1477R and OE2019F. **A-C**: SDS-PAGE after batch purification, **D-F**: Western blots. **Lane 1**: pre-stained protein marker in KDa, **lane 2**: sample before purification, **lane 3**: unbound fraction, **lane 4**: 1st wash fraction using 20mM Imidazole, **lane 5**: 1st elution fraction using a step gradient of 20-200mM Imidazole, **lane 6**: 4th elution fraction using a step gradient of 200-500mM Imidazole, **lane 7**: His tagged marker.

6.3.3 In vivo expression of OE1472F, OE2019F and OE1477R in *E. coli*

The previous section demonstrated that OE1472F, OE1477R and OE2019F, were all able to be expressed as a full-length His-tagged proteins in a cell-free system derived from *E. coli*. In vivo expression in *E. coli* was then tested using the same plasmids, and the DH3-Rosetta strain (Invitrogen) as the host. Unlike, the cell-free system, which was not optimized for expression of foreign genes, the Rosetta strain contains various combinations of rare tRNA to enhance the expression of genes isolated from organisms with AT or GC rich genomes.

6.3.3.1 Transformation and expression in the DH3-Rosetta strain

Plasmids OE1472F-His $_{C-6xhis}$/ pET22b(+), OE1477R-His $_{C-6xhis}$/ pET22b(+), and OE2019F-His $_{C-6xhis}$/ pET22b(+), were introduced into *E. coli* DH3-Rosetta cells, and transformants were grown in 2ml of LB with antibiotics (overnight, 37°C, 250rpm). These pre-cultures were used to inoculate 10ml LB (+ antibiotics) in 100ml flasks, starting at a cell density of $0.1OD_{600nm}$. When the OD_{600nm} reached ~ 0.6, gene expression was induced by the addition of 1mM IPTG and cultures incubated for 3h (37°C, 200rpm).

Culture samples were collected before and after induction, and were analysed on 4-12% SDS-PAGE gels (figure #49). Comparison of the coomassie stained proteins of non-induced cells and induced cells, shows that bands corresponding to the expected size of proteins OE1472F, OE1477R and OE2019F were expressed at high levels, as evidenced by the thick bands (indicated by the red boxes in figure #49). The calculated molecular weights of the 6xHis tagged OE1472F, OE1477R and OE2019F were 28,298.3Da, 24,244.7Da, and 29,279.2Da, respectively. All the above tagged proteins migrated slower on the SDS-PAGE compared to the calculated molecular weight.
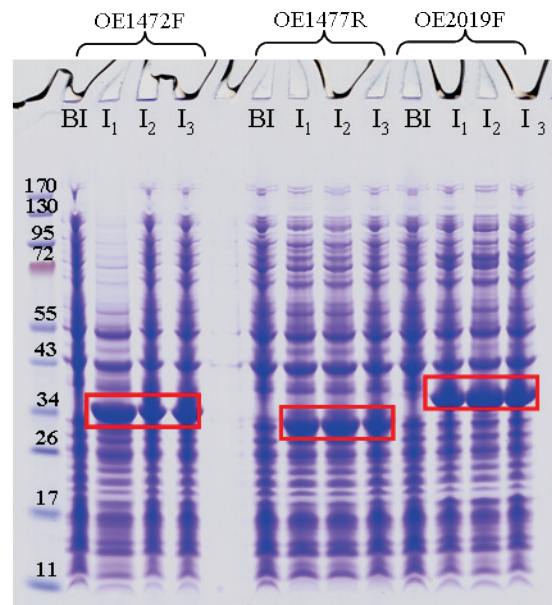
Figure # 49: Purification of in-vivo His-tag proteins OE1472F, OE1477R and OE2019F. The proteins are *E. coli* expressed OE1472F, OE1477R, and OE2019F-His $_{C-6xhis}$/ pET22b(+) plasmids before and after induction. **BI**-before induction, I_1-I_3- cultures 1-3 after induction. MW markers were run in the left most lane. The expressed His tagged OE1472F, OE1477R and OE2019F are indicated by the red boxes. SDS-PAGE, 4-12%.

6.3.3.2 Solubility of the *E. coli* expressed proteins

To check whether the over-expressed proteins were soluble or expressed as inclusion bodies, transformants were grown in 50ml of LB medium (with antibiotics) until OD_{600nm}=0.6, and then induced for 3h (37°C, 200rpm) with 1mM IPTG. The cells were broken by sonication, and centrifuged (4°C, 40000rpm, 40min) to separate the soluble cytoplasmic fraction from the cell precipitated cell debris. The pellet was resuspended in 8M urea to the same final volume as the supernatant. Figure #50 shows the protein profiles of samples taken before induction, after induction, the supernatants after sonication, and the pellet after sonication. Comparison of the soluble and particulate fraction showed that all three proteins (OE1472F, OE1477R, and OE2019F) were over-expressed predominantly as soluble proteins.
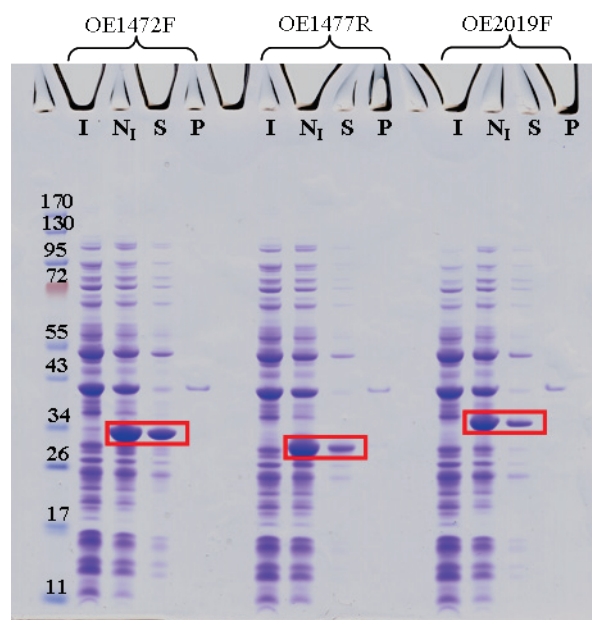
Figure # 50: Solubility of recombinant His tagged OE1472F, OE1477R and OE2019F proteins analyzed by SDS-PAGE on a 4-12% gel. **I**- induced samples, **N_I**- Non-induced samples, **S**-supernatants after sonication, and **P**- pellets after sonication. Coomassie blue stain. The expressed His tagged OE1472F, OE1477R and OE2019F are indicated by the red boxes.

6.3.4 Purification of His-tagged haloarchaeal proteins by affinity chromatography

From the results of the preceding section, it was now possible to scale up the production of the three haloarchaeal proteins using in vivo expression in *E. coli*. Once purified, they could be used to examine enzyme activity, structure, and protein-protein interactions.

6.3.4.1 Small scale purification of OE1472F, OE2019F and OE1477R proteins

Initial purification attempts were performed using 50ml cultures over-expressing OE1472F, OE1477R and OE2019F. Cells were collected by centrifugation, sonicated to release soluble proteins, and the supernatants loaded on 1ml NiNTa columns, equilibrated with PBS. The flow through was washed with NPI_{20}, and the bound protein eluted with NPI_{250}. As shown in figure #51, proteins OE1472F, OE1477R, and OE2019F elute with high purity in the third elution fraction of Imidazole a step gradient (20-250mM). Some contaminating *E. coli* proteins were seen in SDS-PAGE gels, running above and below the proteins of interest.
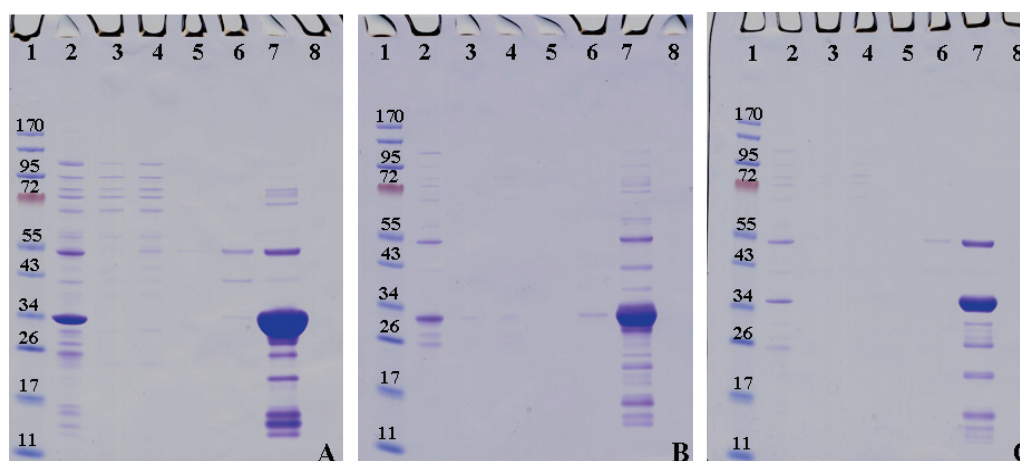
Figure # 51: Purification of OE1472F (**A**), OE1477R (**B**) and OE2019F (**C**) on NiNTa columns. **Lane 1**: Pre-stained protein marker in KDa, **lane 2**: cell supernatant, **lane 3**: unbound proteins, **lane 4**: 1st wash fraction in NPI$_{20}$, **lanes 5-8**: elution fractions 1-4 eluted in step gradient using NPI$_{20}$-NPI$_{250}$. 4-12% acrylamide SDS-PAGE gels, coomassie blue stain.

Table # 11: Molecular masses and N-terminal sequences of His-tagged OE1472F, OE1477R and OE2019F, as measured by ESI-MS, and the sequence of the first N terminal amino acid determined by N-terminal sequence [1].

Expressed protein	Calculated mass	Observed mass (ESI-MS)	# of Methionine	sequence of first 6 amino acids [1]
OE1472F-6XHis	28,298.31 Da	28,298.80 Da	+2	**MM**TDAG
OE1477R-6XHis	24,244.72 Da	24,241.47 Da	+2	**MM**EFQD
		24,110.93 Da	+1	**M**EFQD
OE2019F-6XHis	29,279.25 Da	29,279.30 Da	+2	**MM**RPFE
				MRPFE

The N-terminal sequences of each of the purified proteins, and their molecular masses were determined by ESI-MS. These are presented in table #11. There was a close correspondence between the calculated and experimentally determined masses.

All proteins had two methionines, one intrinsic methionine and a second methionine, which was introduced during the in-frame cloning with the restriction site *NdeI*. Post translation modification was not expected, due to the cloning strategy

applied. While cloning the genes of interest into pET22b(+), the signal peptide sequence for periplasmic localization, was removed, so expressed proteins would not be exported and post translation modification will not occur. But, proteins OE1472F, and OE2019F, showed some processing of the first methionine to give N-termini with just one methionine.

The N terminal sequences results were in agreement with the ESI-MS findings (table#11), except for OE2019F, where an additional species was found containing only one methionine. Both, ESI-MS, and N-terminal sequencing, confirmed that the proteins were His-tagged and full length.

6.3.4.2 Separation under native conditions using 1M KCl

To investigate whether His-tagged OE1472F could be recovered in high purity using native, high salt conditions, it was purified from *E. coli* cells in the presence of 1M KCl and separated on NiNTa using Imidazole which contained 1M KCl (see materials and methods section 5.1.3.2).

Before sonication, the cell pellet was resuspended in NPI_{20}-KCl. After sonication, the cell lysate loaded on a 1ml NiNTa column equilibrated with NPI_{20}-KCl. Unbound protein was washed through, and two step elution gradients were performed; the first from NPI_{20-103}-KCl, and the second from $NPI_{103-250}$-KCl. As shown in figure#52, the protein eluted in both of the step gradients. No significant improvement in the purity was seen when comparing the protein eluted in high salt, to the protein eluted in low salt (figure #52, lane 7, 11, and 12).

Figure # 52: Purification of His-tagged OE1472F under native conditions with high salt buffers. **Lane 1**: prestained protein marker in KDa, **lane 2**: cell lysate from the induced culture, **lane 3**: sample before separation in NPI_{20}-KCl, **lane 4**: unbound fraction in NPI_{20}-KCl, **lane 5**:1st wash fraction in NPI_{20}-KCl, **lane 6-8**: 14th, 15th and 16th wash fractions in NPI_{20-103}-KCl, **lanes 9-11**: Elution fractions 1-3 in $NPI_{103-250}$-KCl, **lane 12**: control, 3rd elution fraction in NPI_{20-250} purified in low salt.
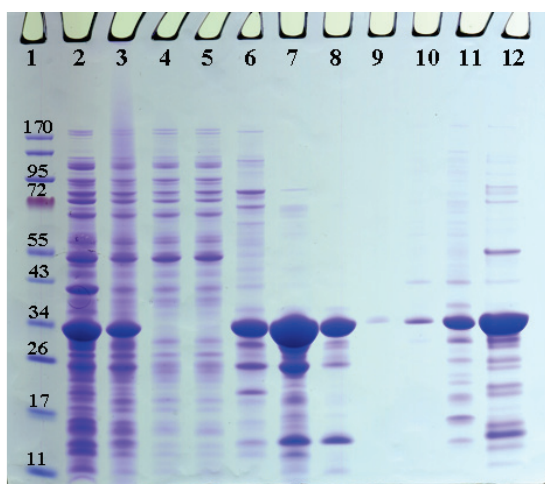
6.3.4.3 Large scale purification of OE1472F and OE1477R

Given the previous results showing good expression and purification from small volume cultures, the next step was to scale up in order to produce sufficient quantities of protein for detailed study.

OE1472F-His $_{C-6xhis}$ and OE1477R-His $_{C-6xhis}$ were produced from 1 liter and 2 liter induced culture, respectively, as follows. The cells were collected by centrifugation, resuspended in NP buffer with DNAase and 20mM PMSF, and broken open using a French press. After centrifugation, the supernatant was loaded on a 1ml column of NiNTa. Unbound protein was washed through with NPI$_{20}$, and bound protein was eluted using a step gradient of NPI$_{20-250}$. Fractions from the column were separated on a 4-12% SDS-PAGE gel, and stained with coommassie blue. As shown in figure #53A-B, while some of the expressed protein eluted in low Imidazole concentrations (20mM Imidazole), the majority eluted in NPI$_{20-250}$ (figure #53A, lanes8-12 and figure #53B, lanes 7-9).
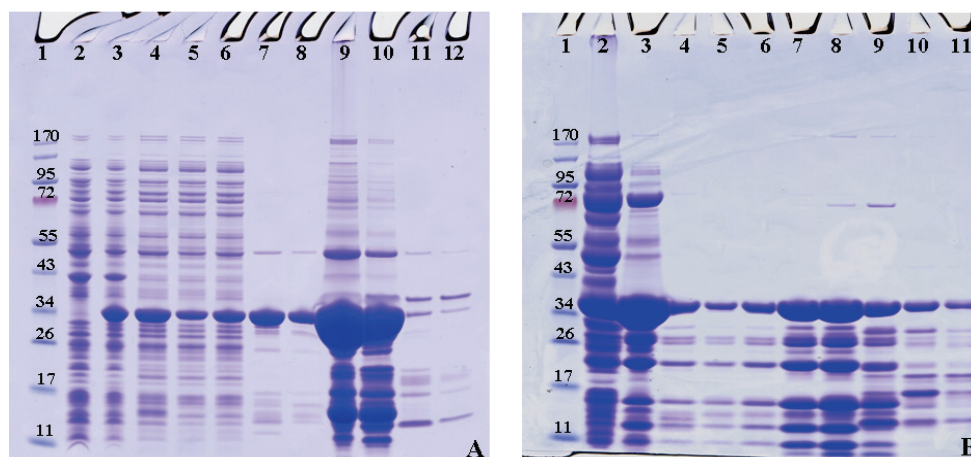
Figure # 53: Purification of His-tagged OE1472F (**A**) and OE1477R (**B**). Bout proteins were purified under native conditions. **Gel A**- **lane 1**: prestained protein marker in KDa, **lane 2**: cell lysate from the non- induced culture, **lane 3**: cell lysate from the induced culture, **lane 4**: sample before separation on NiNTa in NP buffer, **lane 5**: unbound fraction in NP buffer, **lane 6**:1st wash fraction in NPI$_{20}$, **lane 7**: 10th was fraction in NPI$_{20}$, **lanes 8-12**: eluted fractions 1-5 in NPI$_{20-250}$. **Gel B**- **lane 1**: prestained protein marker in KDa, **lane 2**: sample before separation on NiNTa in NP buffer, **lanes 3-5**: wash fractions #2, 10 and 18 in NPI$_{20}$ buffer, **lanes 6-11**: elution fractions #20, 21, 22, 24, 26 and 28 in NPI$_{20-250}$

Now that purification protocols were established for OE1472F, OE2019F and OE1477R the next step was to measured the activities of the purified proteins. Note that the activity of OE1472F and OE2019F was measured after purification from

50ml and 2liter culture (figures #51,53), while the activity of protein OE1477R was measured after purification from 2liter culture, pooling elution fractions 20-24 (Figure #53, lanes 6-9), and dialysis against 3M KCl+50mM Tris HCl pH 7.2.

According to BCA assay 13mg/ml tagged OE1472F and 9.9mg/ml tagged OE1477R were produced from 2liter culture, while ~2.1mg/ml tagged OE2019F was purified from 50ml culture.

6.4 Enzyme activity and specificity

Having purified the proteins of ORFs OE1477R, OE1472F and OE2019F, this section summarizes the development and validation of assays to measure their expected enzyme activities.

ORF OE1477R was annotated as 3-dehydroquinate dehydratase (EC 4.2.1.10) and expected to catalyze the formation of DHS (3-dehydroshikimate). ORF OE1472F was suggested to catalyze the first step in the biosynthesis of AroAAs in *H. salinarum*, using ASA and DKFP as precursors. This step can be carried out by a transaldolase. ORF OE2019F was annotated as F-1,6-P aldolase (EC 4.1.2.13).

6.4.1 Aldolase Assay

ORF OE2019F, was expected to cleave F-1,6-P. As an aldolase it would split F-1,6-P into two triosephosphates: glyceraldehyde phosphate (GAP) and dihydroxyacetone phosphate (DHAP). Aldolase activity can be measured by the colorimetric method of Sibley and Lehninger (Pinto, 1969; Sibley, 1949), which has been shown to efficiently measure aldolase activity in cell-free extracts of haloarchaea (Altekar W., 1987; D'souza S.E., 1998; D'Souza, 1982; Dhar and Altekar, 1986). In this assay, the products (DHAP and GAP), are trapped by reaction with hydrazine, and then converted to their colored dinitrophenylhydrazones. Using 2,4 dinitropheneylhydrazine (DNPH), the hydrazone generated can be detected at 540nm. As shown in figure #54, aldehydes and ketones react with hydrazine to form hydrazone.
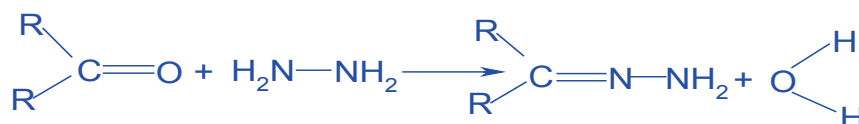
Figure # 54: Derivatization reaction of aldehydes and ketones with hydrazine.

6.4.1.1 Testing the assay using aldolase from rabbit muscle (Sigma)

The ability to detect the formation of DHAP and GAP by the colorimetric assay was checked first with a commercial aldolase (from rabbit muscle, Sigma), using the

114

coupled assay suggested by the manufacturer. The activity of the commercial alodolase was measured by continuous monitoring of NADH at 340nm. Since NADH is consumed in the reaction, there should be a continuous decrease in the absorbance at 340nm. One unit of enzyme reduces one μmole NADH per minute at 25°C in 86.7mM Tris HCl pH=7.4. A decrease in the absorbance of NADH occurred only when the substrate, F-1,6-P, was added to the reaction mixture (figure #55, black arrow). From the corresponding slopes, the activity of the commercial aldolase was calculated to be 0.42±0.03U/ml or 12.1± 0.9U/mg protein (n=5, figure #55, blue curve). As expected, the reaction rate was halved when TIM (Trioesphosphate Isomerase, EC 5.3.1.1) was not added to the reaction mixture (figure #55, red curve).
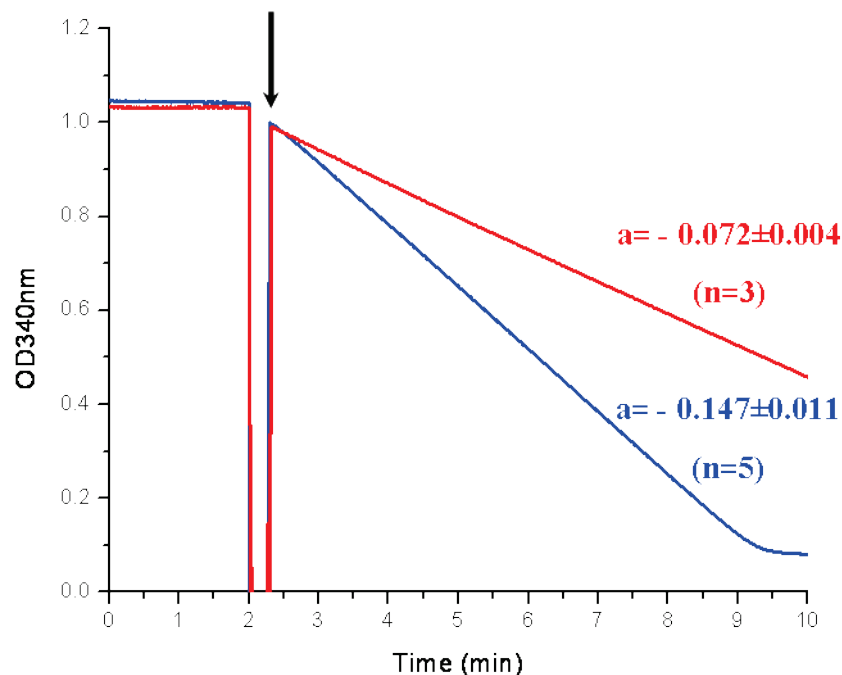
Figure # 55: Activity of commercial aldolase measured with the coupled enzyme assay using 1.92mM F-1,6-P, 16.7U/ml TIM, 1.67U/ml GDH, and 0.15mM NADH. The arrow indicates addition of aldolase to the reaction mixture. **Blue curve**: aldolase activity was measured in the presence of TIM and GDH. **Red curve**: aldolase activity was measured in the presence of GDH only.

For comparison, the same aldolase used in the coupled assay (12.6±1.2 U/mg) was used in the colorimetric assay. As shown in table#12 below, both assays gave very similar values for specific enzyme activity.

Table # 12: Comparison of the colorimetric assay and coupled assay using a rabbit muscle aldolase and TIM.

	U/mg
Colorimetric assay	14.6±2.5
Coupled assay	12.6±1.2

6.4.1.2 *E. coli*-expressed OE1472F and OE2019F show aldolase activity

The ORF OE2019F (*fab1*) was annotated as fructose-bisphosphate aldolase (EC 4.1.2.13) and, by definition, the gene product should be able to cleave F-1,6-P into GAP and DHAP. ORF OE1472F, which is part of the *trp* operon, was reannotated as 2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase, in accordance with the functional evidence of MJ0400 in *M. jannaschii* (White, 2004). As such, the gene product is expected to cleave DKFP, which is assumed to be one of the precursors in the proposed AroAAs biosynthesis pathway, and therefore is not expected to display aldolase activity.

OE2019F and OE1472F were both expressed in *E.coli* and purified on a NiNTa column under low salt conditions using buffers containing 300mM NaCl. The activity of the proteins was measured using the colorimetric assay (low salt, Figure #56A and C). The activity of the freshly eluted protein (**fraction E-3 eluted after step gradient NPI$_{20-250}$, low salt**) was measured directly in elution buffer ("before") and again after exchanging the buffer to one with high salt ("after"). As shown in figure #56A, both proteins possess readily detectable aldolase activity, and are able to cleave F-1,6-P. Their specific activities were very similar, and were increased when the buffer of the enzyme was exchanged to 1M KCl, prior to the assay.

Since the gene product of OE1472F derives from an extreme halophile, it would be expected to function optimally at high salt conditions, similar to the conditions present in the cytoplasm. It is also likely to retain higher activity if purified under high salt condition. For this reason, the specific enzyme activity was measured after purification in 1M KCl. The specific activity of protein eluted in 15th fraction (**W15-wash with NPI$_{20-103}$-KCl buffer**, separation shown in figure #53, lane 7) was compared to the activity of the 3rd elution fraction (**E-3 eluted with NPI$_{103-250}$-KCl buffer**, separation shown in figure #53, lane 7).

116

As shown in figure #56B, no significant difference in the specific activity was seen before and after buffer exchange when the enzyme was purified in 1M KCl, as opposed to low salt purification, where the specific activity was 7 fold higher after buffer exchange (figure # 56A). Note that a significant difference was seen in the specific activity of fraction W-15, compared to fraction E-3. Fraction E-3 was 3-fold more active compared to fraction W-15 (32mU/mg protein and 11mU/mg protein, respectively).

Figure # 56: Specific aldolase activity of *E.coli*-expressed OE1472F and OE2019F purified **A**-at low salt buffers (300mM NaCl) **B**- *E.coli*-expressed OE1472F purified at high salt buffers (1M KCl).

The proteins were purified from 50ml of culture, separated on NiNTa column, and eluted in NPI_{20-250} buffer with 300mM NaCl (**low salt**) or in NiNTa buffer with 1M KCl buffer (**high salt**). Fraction **W-15** eluted in NPI_{20-103}-KCl buffer, while fraction **E-3** eluted in $NPI_{103-250}$-KCl buffer. Aldolase activity is expressed as μmoles DHAP formed/mg protein/30min at 37°C. The colorimetric assay was performed in 2M KCl in 78.7mM Tris HCl (pH=7.2), and 5.9mM NaCl.

Before- activity measured in elution buffer,

After- activity measured after buffer exchange.
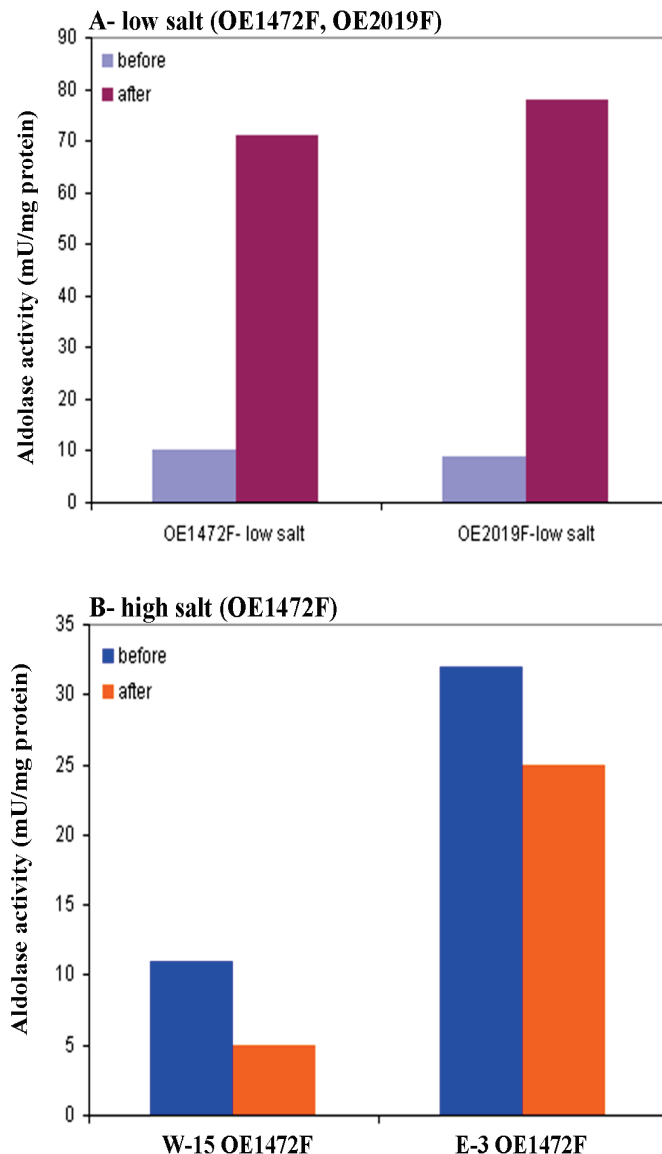
As seen in figure #57, there is no need to purify the enzyme under high salt concentrations (1M KCl). Similar specific activities were obtained when the protein was purified under standard conditions (300mM NaCl) and the buffer was exchanged, compared to purification in high salt (1M KCl), (70mU/mg protein and 43mU/mg protein, respectively. Note that the specific activity of the commercial rabbit muscle aldolase (table #12, 14.6±2.5 U/mg protein) was significantly higher that of OE1472F (figure #56A, 70mU/mg protein). The aldolase from sigma was 206-fold more active compared to OE1472F.
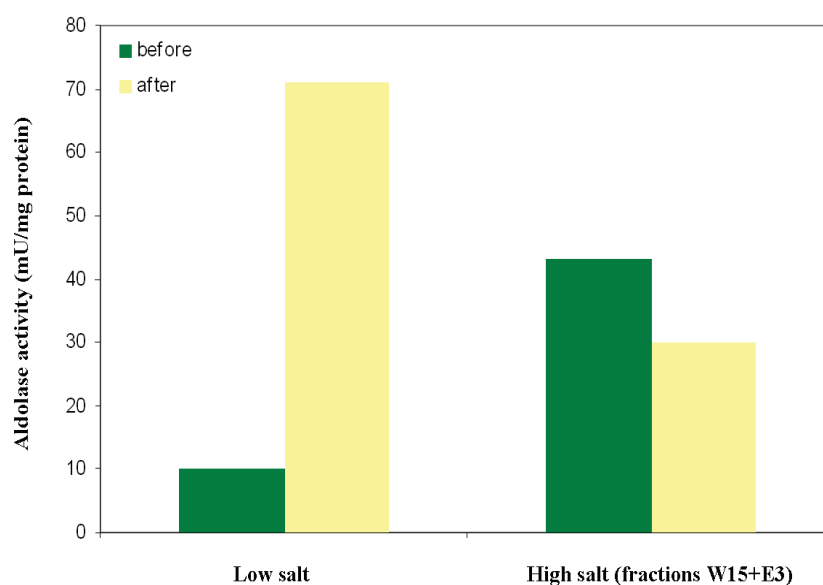
Figure # 57: Specific activity of recombinant OE1472F purified on a NiNTa column using low salt buffers (300mM NaCl) or high salt buffers (1M KCl). The activity was measured before and after buffer exchange using the colorimetric assay. The concentration of the protein was measured by BCA assay. Aldolase activity is expressed as μmoles DHAP formed/mg protein/30min at 37°C. **before-** activity measured in elution buffer, **after-** activity measured after buffer exchange. The proteins used for the measurement are seen in figure #51 for low salt purification, and figure #52 lane 12, for high salt purification lanes 7 and 11.

6.4.1.2.1 Confirming the basis of the colorimetric assay

Sibley and Lehninger (Sibley, 1949) showed that hydrazine traps the triosephosphates GAP and DHAP as soon as they are formed after cleavage of F-1,6-P by aldolase. Hydrazine reacts with the carbonyl group in the triosephosphates to form the corresponding hydrazones. The products of splitting DKFP were expected as well to be trapped by hydrazine and if so be detectable by the colorimetric assay (figure #58).

This assumption was checked by measuring the absorbance of the corresponding hydrazones when 1.2mM DHAP, GAP and HA were mixed with 3mM hydrazine in 2.475M KCl. Since hydrazine does not absorb at 240nm, changes in the absorbance at this wavelength can only occur if the corresponding hydrazones are formed. A plot of the increase in A_{240nm} would then allow the rates of the reaction to be determined from the slope $(a=OD_{240nm}/min)$.
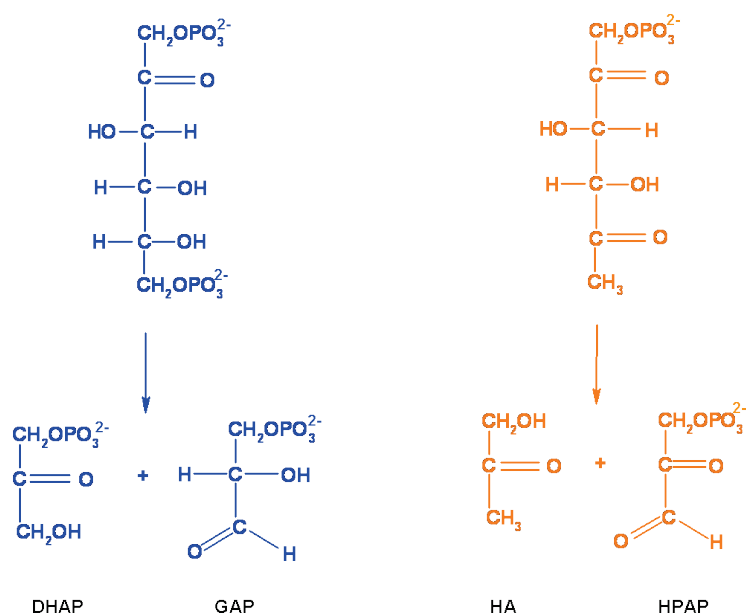
Figure # 58: The structures of F-1,6-P (blue) and DKFP (orange). **DHAP-** dihydroxyacetone phosphate, **GAP-** glyceraldehyde 3-phosphate, **HA-** hydroxyacetone (2 keto 1-propanol), **HPAP-** hydroxpyruvaldehyde phosphate.

As seen in figure #59, the rate of formation of the corresponding hydrazones of DHAP and GAP was not the same as seen from the initial slopes, but the similar final OD240nm was achieved after 10min of incubation, indicating that similar concentrations of the corresponding hydrazones were formed. No change was observed when HA was mixed with hydrazine, meaning that HA-hydrazone was not formed, and thus the colorimetric assay cannot be used to detect cleavage of DKFP by ORF OE1472F.

The formation of the corresponding hydrazones of DKFP and F-1,6-P were measured as well. No hydrazone was detected when F-1,6-P was added to hydrazine, while DKFP formed the corresponding hydrazone (1.47 ± 0.03 OD_{240nm}/min, respectively).
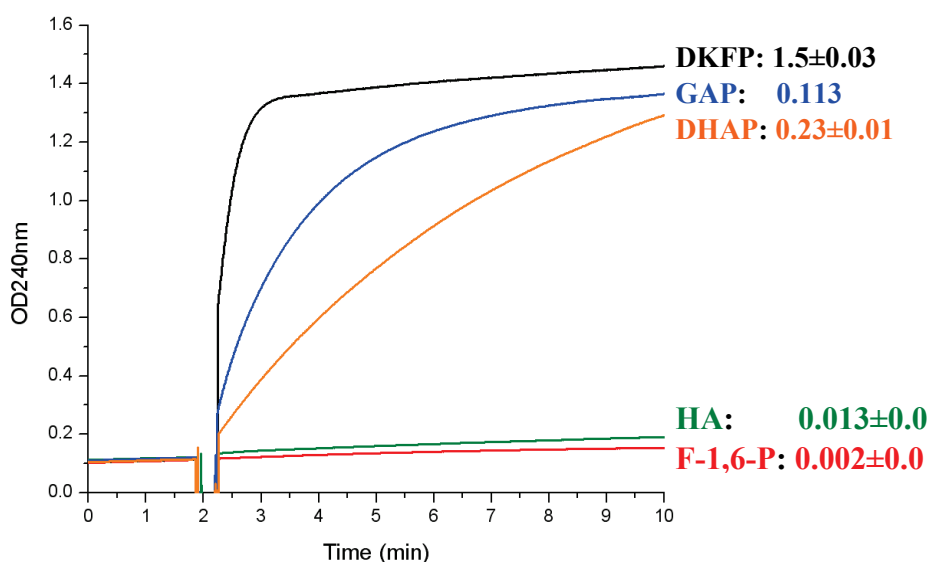
Figure # 59: Detection of the corresponding hydrazones of 1.2mM F-1,6-P, 1.2mM HA, 1.2mM DHAP, 1.2mM GAP, and 1.2mM DKFP. All reactions were carried out at RT, in total volume of 1ml. The reactions contained as well 3mM hydrazine and 2.475M KCl in Tris HCl pH =7.2. The initial slope after adding the substrate was subtracted from the slope obtained in the first 2min (a=0.0054±0.0002OD/min, reaction mixture contained buffer and hydrazine only). A constant increase in the absorbance over time was seen only when substrates were added to the reaction mixtures (after 2 min).

6.4.2 Transaldolase activity

The first step in the proposed pathway is a transaldolase reaction, which catalyzes the transfer a HA moiety from DKFP while releasing hydroxypyruvaldehyde phosphate (figure #60).

Transaldolase activity of the purified recombinant proteins was tested, in the presence of the precursors, ASA and DKFP, using two different assays, LC-MS and GC-MS. LC-MS and GC-MS allow the detection and identification of the derivatized intermediates. Compounds are first separated by HPLC or by GC. In the second dimension (MS), the masses of derivatized compounds can be accurately measured. The advantage of the second dimension is the possibility of detecting a wide range of intermediates by their mass.
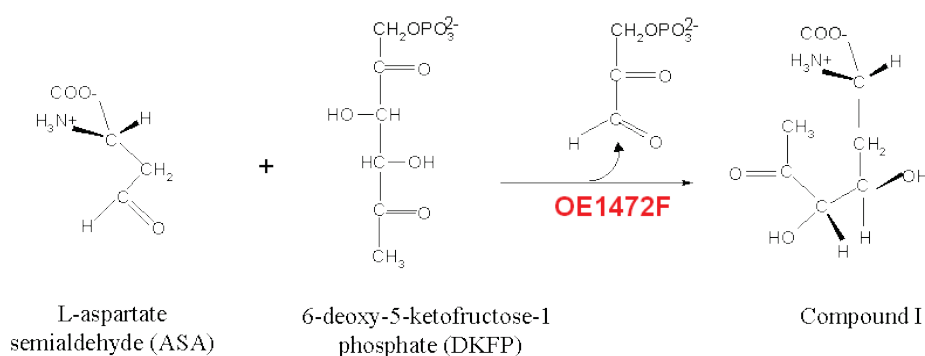
Figure # 60: The transaldolase reaction catalyzed by ORF OE1472F using ASA and DKFP as precursors. Aldole-like condensation between HA moiety and ASA releases HPAP (hydroxpyruvaldehyde phosphate).

6.4.2.1 Validating the LC-MS assay for determining aldolase activity

As a first step the LC-MS method was validated. Pure preparations of ASA, DKFP, F-1,6-P, DHAP, GAP, and DHQ were derivatized according to the protocol described in materials and methods, and the derivatized masses were identified by LC-MS. Their retention times are given in appendix table #15.

Next, the purified tagged OE1472F was incubated with different substrates for 30min at 37°C, followed by derivatization and analysis by LC-MS. None of the expected products or intermediates were detected when the recombinant protein OE1472F was incubated with ASA and DKFP. Moreover, OE1472F showed no aldolase activity with the substrates F-1,6-P. The latter result was unexpected, as the colorimetric assay indicated that OE1472F could cleave F-1,6-P (figure #56A). This discrepancy can be explained either by a difference in sensitivity of the two assays, or tne efficiency of derivatization. Both of those aspects were then examined in more detail.

As seen in figure # 61A and B, ≤12.5 nmole DHAP (in 500µl) can be detected by the colorimetric assay while ≤1 nmole DHAP can be detected by the LC-MS, meaning that the LC-MS method is 12 fold more sensitive compared to the colorimetric assay. The derivatization efficiency can not be measured by the colorimetric assay, but in the LC-MS method it is clear that the efficiency of the derivatization was high, as no excess of NBHA (O-(4-nitrobenzyl) hydroxylamine hydrochloride) was seen after the derivatization when using 2-3mM substrate and 2mM NBHA. Thus, neither the sensitivity nor the derivatization efficiency was the reason for the inconsistency between the assays.
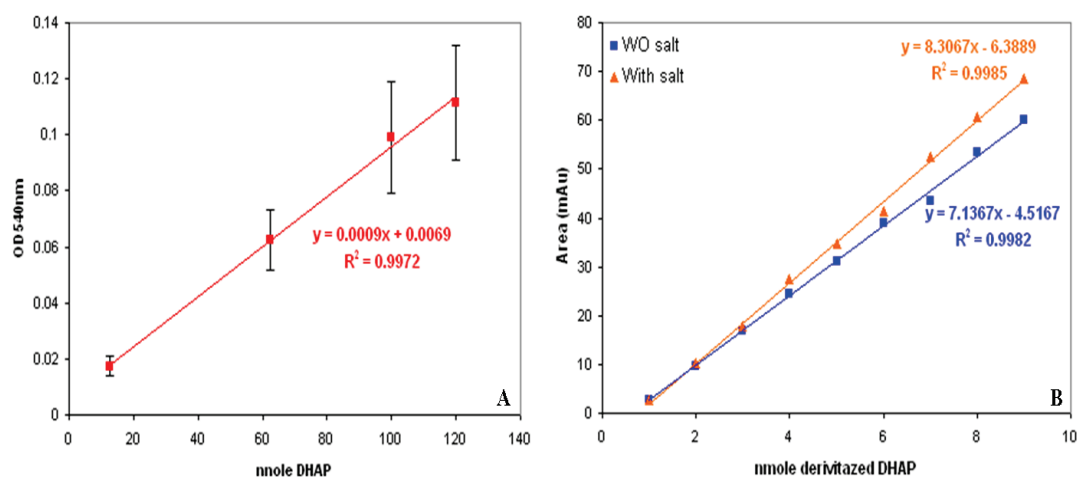
Figure # 61: Calibration curves of derivatized DHAP in **A**-colorimetric assay and **B**-LC-MS.

To further examine the source of the discrepancy, the activity of commercial rabbit muscle aldolase was measured both by the coupled enzyme assay (stock 6.3±0.4U/mg protein, n=5) and LC-MS. The enzyme solution was mixed with 1.92mM F-1,6-P and 16.7U/ml TIM, incubated for 30min at 37°C, then the products derivatized and separated on a C-18 column. The area of the peaks detected at 254nm was calculated and found to be 62.1±2.1 mAu. Using the calibration curve shown in figure #61B (WO salt), the area was found to correspond to 4.0±0.1U/mg protein, which is 63±2% of the total activity determined by coupled enzyme assay.

Nevertheless, the aldolase activity of the commercial enzyme was readily detectable. The fact that protein OE1472F showed no detectable aldolase or transaldolase activity is, most likely due to the low specific activity of this enzyme, combined with the small volumes that can be loaded on the column (up to 9μl). The specific activity of OE1472F determined by the colorimetric assay was 70mU/mg protein, meaning 35.6nmole of DHAP would be formed after 30min at 37°C. Following the derivatization protocol for LC-MS, this amount of DHAP will be further diluted (1/2.5), and since the maximum load volume is 9μl, this would introduce only 0.85nmole DHAP for MS analysis. This is below the limit of detection of the LC-MS. Overnight incubation of reaction mixtures at 37°C did not improve the sensitivity, as no products or intermediates were detected (data not shown). Although it is known that halophilic proteins are often thermostable, it might be that this enzyme is not.

From the above results, it was not possible to determine the aldolase/transaldolase activity of the purified OE1472F using the LC-MS method described here. Improvements that could be tested include: higher protein concentration, and removal of salt from the reaction products before derivatization. Optimizing the enzyme reaction conditions would also be useful to explore.
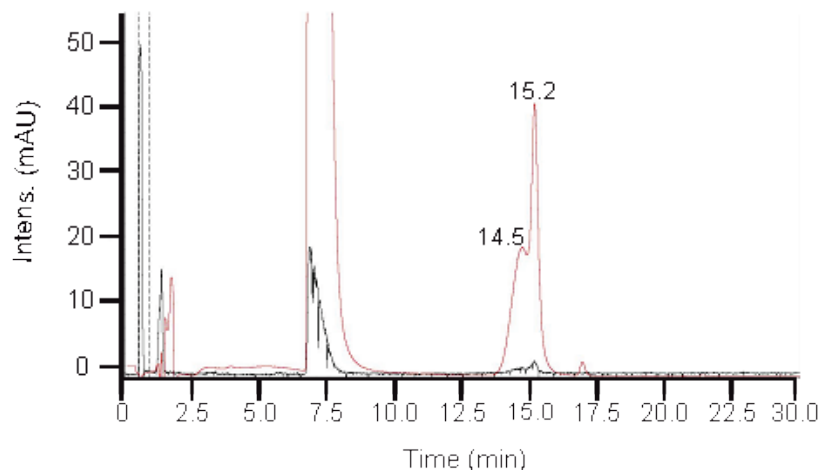
Figure # 62: Derivatized DHAP obtained after cleavage of F-1,6-P with commercial aldolase and TIM. Black curve, the TIC; red curve, absorbance detected at 254nm.

6.4.2.2 Validating the GC-MS assay for determining transaldolase activity

GC-MS analysis requires one or more appropriate derivatization procedures to block active protons and labile keto groups. The combination of methoximation with subsequent silylation is conveniently used. The first step is methoximation of the carbonyl group to form the corresponding oximes (figure #63A). The second step is using silylating agent (TMS or TBS) which replaces the labile hydrogens with a -Si(CH$_3$)$_3$ or -Si(CH$_3$)$_2$-C(CH$_3$)$_3$ group, respectively (figure #63B). Theses derivatization steps are used to prepare volatile and thermally stable derivatives for GC-MS.

In order check whether methoximation was necessary in addition to TMS or TBS derivatization, glucose was used as model molecule in this assay. As seen in figure #64, without methoximation the derivatization of glucose with TMS yielded several products (peak#3 was the desired product), while when using methoximation along with TMS derivatization, only the two stereiosomers of the desired product were detected.

123
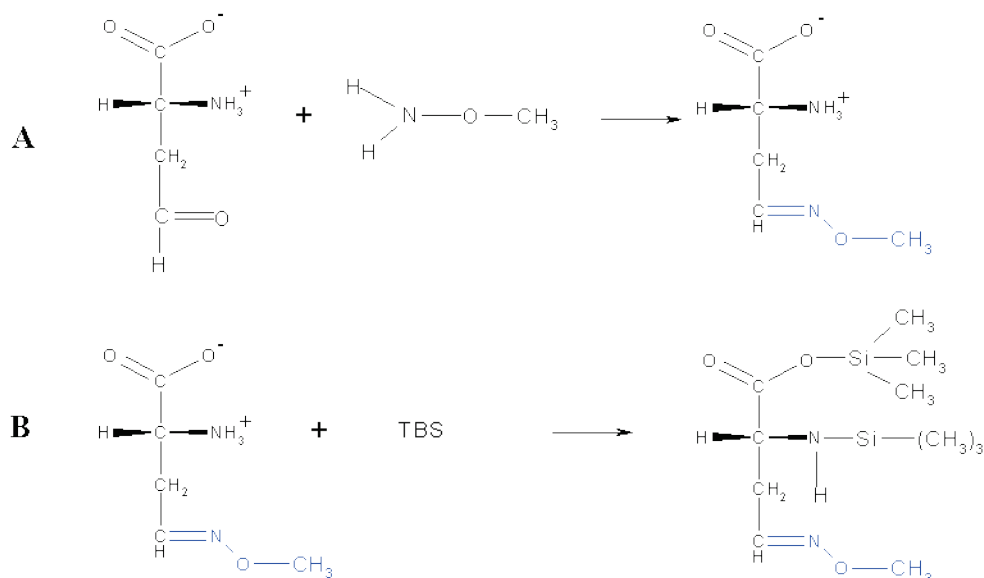
Figure # 63: Two steps derivatization for GC-MS. A- Methoximation with O-methylhydroxyl-hydrochloride. B- silylation with TBS [N-methyl-N-(tert-butyldiamethylsilyl) trifluroacetamide].

Next, a two-step derivatization with either TMS or TBS was tested on the following compounds: ASA, DKFP, F-1,6-P, GAP, DHAP, DHA, DHQ and shikimate. The correct masses of the derivatized substrates were identified by GC-MS, along with the indicated retention times, as shown in the appendix tables #17-18. Methoximation plus TBS derivatization was the preferred method due to the ability to detect all the substrates listed above. This was not possible when the second derivatization was TMS. Moreover the derivatization method used in this method allowed detection of shikimate, which was not possible in the derivatization used in the LC-MS method.
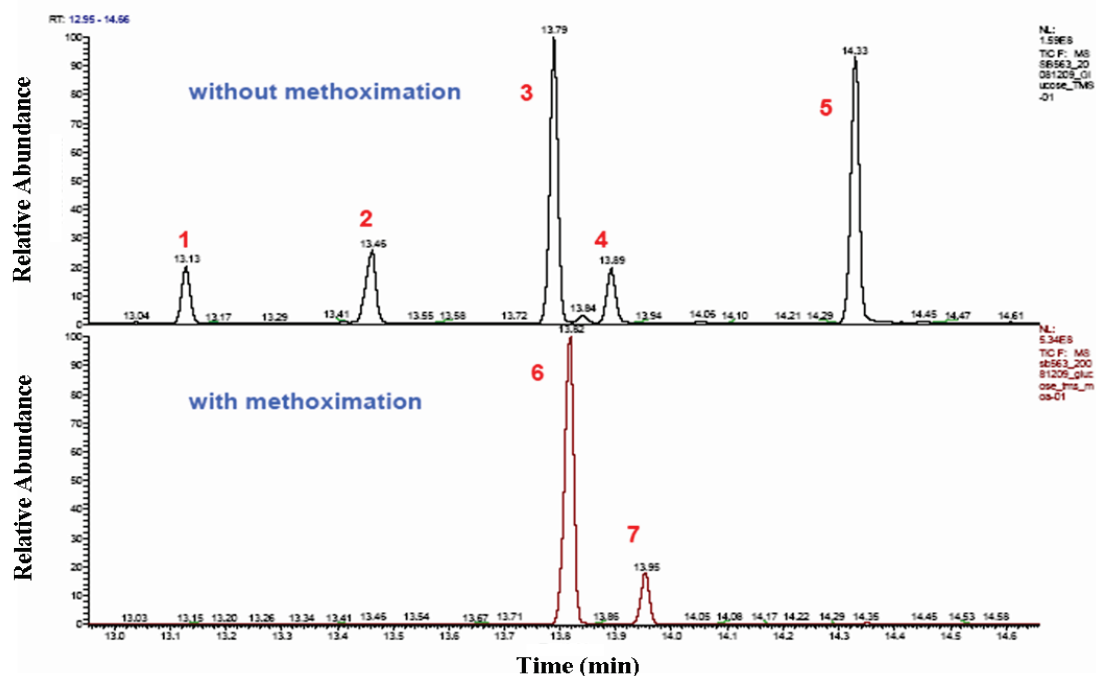
Figure # 64: Derivatization of glucose with and without methoximation followed by a 2nd derivatization with TMS. Peaks 1-2: β-DL-arabinopyranose, 3-D-glucose, 4-1,6-anhydro-β-D-glucose, 5-Glucopyranose, 6-7: D-glucose.

6.4.3 The OE1477R gene product shows dehydroquinase activity

ORF OE1477R was annotated as 3-dehydroquinate dehydratase (EC 4.2.1.10), an enzyme that removes a water molecule from DHQ to form DHS (figure#65). The formation of DHS from DHQ was measured using LC-MS and spectrophotometer.
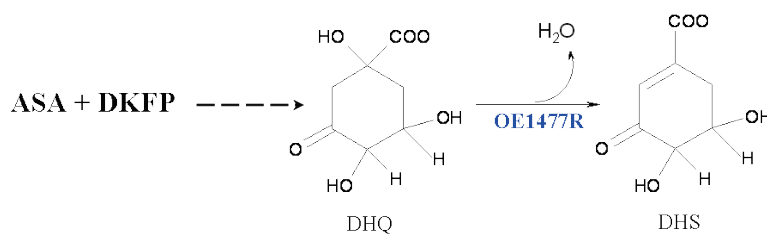
Figure # 65: Schematic representation of the reaction that OE1477R protein might catalyze.

3-dehydroqinase dehydratase activity of the purified protein was assayed first by monitoring the formation of 3-dehydroshikimate (DHS) at 234nm at RT. The standard assay mixture contained 2.57M KCl, 43.75mM Tris HCl (pH=7.2) and 10mM 3-dehydroquinate (DHQ). The specific activity was found to be 55 ± 6mU/mg protein (n=2). 3-dehydroqinase dehydratase activity was also detected using the LC-MS

method (appendix, table #16), where a product corresponding to the mass of the derivatized DHS was detected after incubating DHQ with purified OE1477R.

6.5 Transformation of *N. pharaonis* strain 2160

6.5.1 Background and objectives

The well-studied archaeon *H. salinarum* is an important model organism for the study of haloarchaea. Its genome has been sequenced and annotated ((Falb et al., 2005),www.halolex.mpg.de), and extensive investigations have been preformed on its flagellar biosynthesis, motility, signal transduction, energy metabolism, retinal proteins and so on. Many of those studies were possible because of the availability of the PEG-mediated transformation method introduced by Cline and Doolittle (Cline et al., 1989). In 1982, a novel member of the Halobacteriaceae was isolated from the highly saline soda lakes of Egypt, and named *N. pharaonis* (Soliman, 1982). This aerobic haloalkaliphilic archaeon (phylum euryarchaeota), grows under extreme alkaline and hypersaline conditions (pH 9-11, 4M NaCl). The genome sequence has been determined (www.halolex.mpg.de, (Falb et al., 2005)) but, to date, the study of *N. pharaonis* has been restricted to comparative genomics and structural studies of purified proteins. The lack of genetic tools for *N. pharaonis* prevents essential genetic studies and molecular engineering of this organism.

Developing genetic tools in *N. pharaonis* is of substantial importance for the following reasons. Firstly, the biological basis of its adaptation to high pH and high salinity is of fundamental biological interest and would provide a model system for haloarchaea that live in this doubly extreme environment. Secondly, genetic manipulations would allow the extension of previous studies on the chloride pump halorhodopsin (Lanyi et al., 1990), sensory rhodopsin II ((Hein et al., 2003); (Schmies et al., 2000)) and its transducer HtrII (Klare et al., 2004). Thirdly, there are many intriguing questions arising from close examination of the *N. pharaonis* genome (www.halolex.mpg.de). On the one hand there are many genes with unknown functions, whereas other genes, that would be expected to be present as part of a known biological pathway, are absent. Surprisingly, 42% of theses ORFs are of unknown function or not similar to anything previously known. 24% of the ORFs are conserved hypothetical proteins (CHY) and, 18% are hypothetical proteins (HY). Without methods for analysing the functions of these unknown genes, a full understanding of this organism will not be possible. One example of particular interest revealed by the genome sequence of *N. pharaonis* (Falb et al., 2005), is that it

contains an unknown analogue of complex III, in a functional respiratory chain. The organism also appears to have a unique nitrogen metabolism.

Only one study has described the construction of a potential shuttle vector for haloalkaliphilic archaea (Zhou et al., 2004b). These authors based their vector on a plasmid from *Natronobacterium* sp. Strain AS7091, and while this vector could be introduced into non-haloalkaliphilic hosts, such as *H. salinarum*, the authors were unable to show that it would function in *Natronobacterium* or any other alkaliphilic haloarchaea. No conclusions were given regarding the failure to transform haloalkaliphilic strains with this plasmid but one might speculate that the two most likely reasons are the inability to form competent cells (by EDTA/PEG treatment for uptake of DNA), or that gene for drug resistance did not provide adequate selection.

6.5.2 Mevinolin is a suitable antibiotic for drug selection *N. pharaonis*

The two commonly used antibiotics for drug selection in haloarchaea are Mevinolin (Mev) and Novobiocin (Nov) (Allers, 2005). To test whether these antibiotics can be used for selection of transformants, and to determine the optimum concentration, the growth of *N. pharaonis* was monitored in the presence of different concentrations of these antibiotics (figure #66).

As seen in figure #66, increasing concentrations of both drugs caused reduction in the growth rates. Mevinolin was profoundly effective at 1μg/ml (=2.47μM), extending the Gn time from 12.48h to 28.5h, while 10μg/ml completely blocked the growth. These concentrations (2-4μg/ml) are similar to those used for selection of transformants in other haloarchaea (Dyall-Smith). Novobiocin was less effective, as even at relatively high concentrations (10μg/ml, 15.7μM) cell growth was diminished but not prevented.
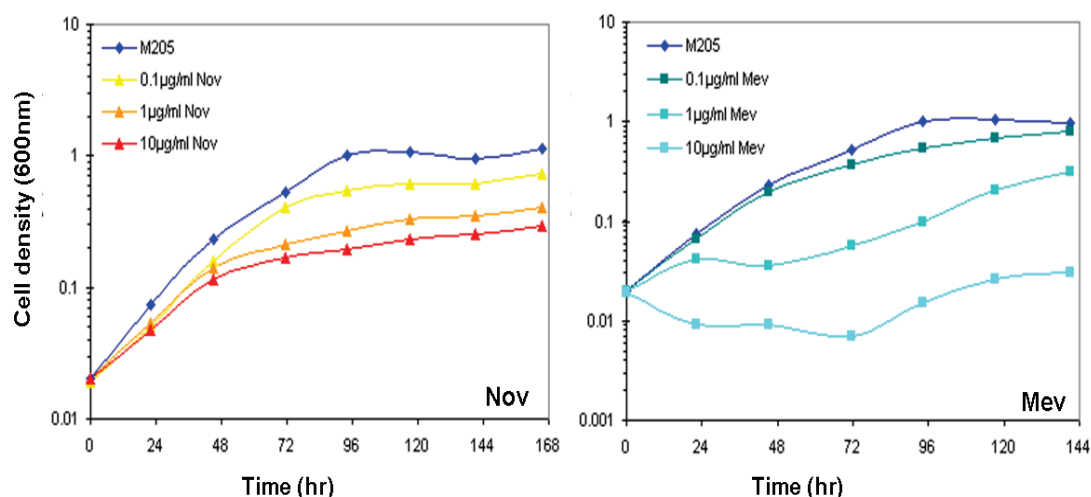
Figure # 66: Growth of *N. pharaonis* in medium M205 (pH=8.9) supplemented with different concentrations of Novobiocin (**Nov**, left panel) and Mevinolin (**Mev**, right panel).

6.5.3 Construction of shuttle vectors for *N. pharaonis*

pNB101 was the first plasmid isolated from the haloalkaliphilic archaeon *Natronobacterium* sp. Strain AS7091(Zhou et al., 2004b). This 2.5Kbp rolling-circle replicating plasmid contains three major open reading frames (ORF1 ,ORF2, and ORF3), as well as three sequence motifs (Motifs I-III) which are common to Rep proteins of RC plasmid family (Zhou et al., 2004b). pNB101 was used to construct the shuttle plasmid pNB102 (Zhou et al., 2004b) which could be used to transform both *E. coli* and *H. salinarum* (strain SNOB).

pNB102, a gift from Dr. Hua Xiang (Institute of Microbiology, Beijing) was used in the current study for constructing plasmids pMG100 and pMG200 (figure#53B, and 53C, respectively). The cleaved Ori of *N. pharaonis* (pNB101, (Zhou et al., 2004a)) was ligated to either pUS-Mev ((Pfeiffer et al., 1999; Schweiger, 1996) or pMKK100 (Koch and Oesterhelt, 2005), as outlined in figure #67 below, and then introduced into competent *E. coli* cells. Transformed colonies, which contained the ligated plasmids, were chosen after mini prep analysis, as seen in figure #68. Transformed *E. coli* colonies were screened for plasmid DNA containing the correct insert by restriction with *BamHI* and *HindIII* for pMG100, or *HindIII* for pMG200 (figure #67A, and B, respectively).
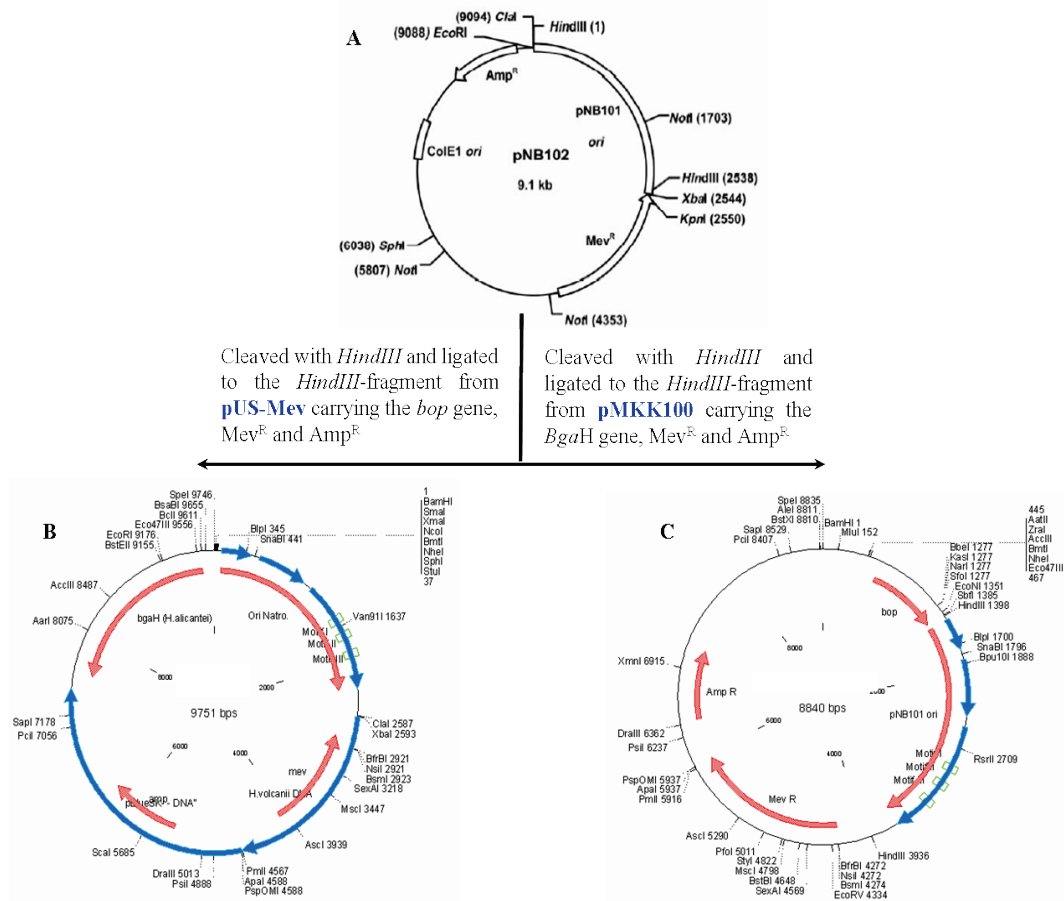
Figure # 67: Construction of shuttle vectors pMG100 (**B**) and pMG200 (**C**) using the *Hind*III-fragment from pNB102 (**A**). pUS-Mev was constructed in-house by Uli Schweiger (Pfeiffer et al., 1999; Schweiger, 1996), and pMKK100 was constructed by Koch and Oesterhelt (Koch and Oesterhelt, 2005).
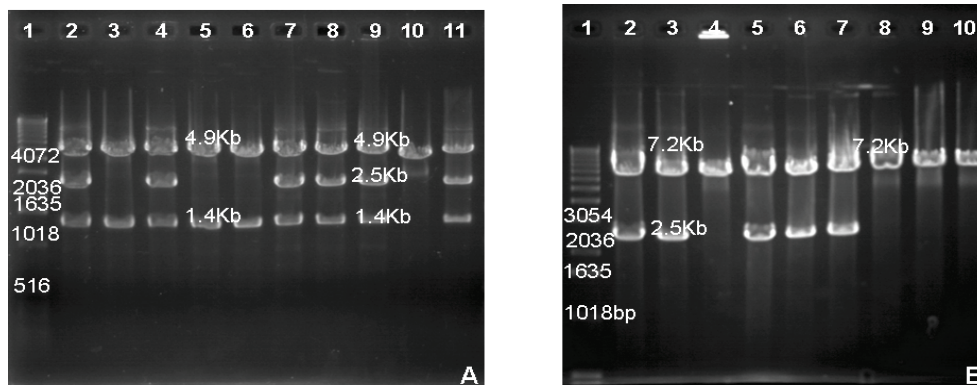
Figure # 68: Mini prep plasmid DNA from separate transformants. The 1% agarose gel showing **A**-*Bam*HI-*Hind*III-digested plasmid DNA from different transformed colonies after introduction of pMG100. **B**- *Hind*III-digested plasmids from 9 colonies after ligation and transformation of pMG200. The loading scheme was the same for both gels. Lane #1: MW size ladder (in bp), lanes 2-11: mini prep plasmid DNA from separate transformants.

6.5.4 Evaluation of transformation methods for *N. pharaonis*

6.5.4.1 EDTA treatment to remove the S-layer

The transformation procedure using the PEG method for halophilic Archaea can be briefly outlined as followed: washing and resuspending cells in spheroplasting solution, adding EDTA to chelate the Mg^{2+} and thereby remove the S-layer (forming spheroplasts), adding DNA, adding PEG_{600} solution (PEG makes the membrane permeable for DNA), washing the spheroplasts (to remove the PEG), and plating.

Although this method has been used successfully with several euryarchaeota that form spheroplasts readily by treating the cells with EDTA (for review see (Allers, 2005)), the conventional method failed to generate spheroplasts of *N. pharaonis* ($\Delta O.D_{600nm}$=0.614, EDTA for 10min at RT, 30 min at RT, and 30min at 37^0C). This could indicate that the S-layer of *N. pharaonis* is different to other haloarchaea in its composition (Falb et al., 2005) or in its subunit interactions. Spheroplast formation and the direct access of DNA to the cell membrane is believed to be an important step of the PEG transformation method, so other techniques to remove the S-layer were investigated.

6.5.4.1.1 Bacitracin forms spheroplast in N. pharaonis

Bacitracin is one of the antibiotics known to interfere with the cell wall formation in Bacteria and Archaea (Pollock et al., 1994; Siewert and Strominger, 1967; Stone and Strominger, 1971; Storm and Strominger, 1973). It forms a complex with a functional phospholipid, which is a carrier that transports sugars from the cytoplasm to the cell membrane (figure #69). By forming this complex, Bacitracin prevents the enzymatic dephosphorylation of the carrier lipid, which is an important step in the biosynthesis of the cell wall. By inhibiting cell wall synthesis, the cells are unable to maintain their normal shape (Upreti et al., 2003). Mescher and Strominger (Mescher and Strominger, 1976) studied the effect of Bacitracin on the morphology of *H. salinarum* and concluded that the cell surface glycoproteins are responsible for maintenance of normal cell morphology.
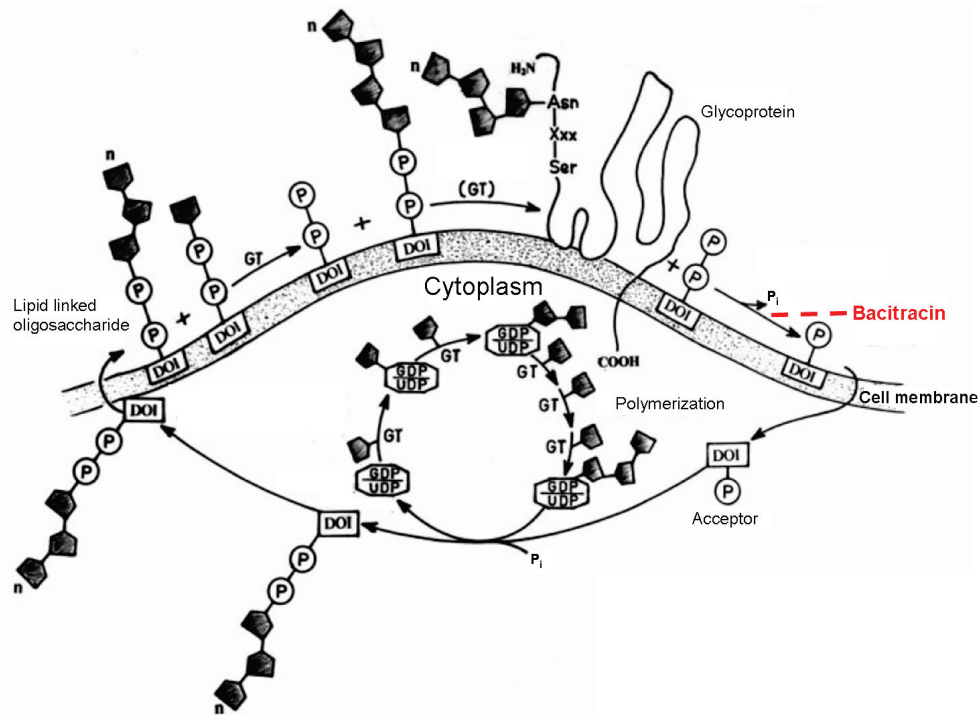
Figure # 69: Peptidoglycan synthesis and the mode of action of Bacitracin. Bacitracin is forming a complex with the lipid carrier and inhibit its dephosphorylation. A lipid (**DOI-P**) or nucleotide (**GDP**, **UDP**) molecule serves as sugar carrier. **GT**-glycosyl transferase, **UDP**-uridine diphosphate, **GDP**-guanidine diphosphate, **Xxx**- any amino acid. From (Upreti et al., 2003).

The effect of Bacitracin on the growth rate and morphology of the cells was tested by growing *N. pharaonis* in the presence of different concentrations of Bacitracin ranging from 5-100µg/ml. As shown in figure #70A, up to 50µg/ml Bacitracin had little effect on the growth of *N. pharaonis*. However, 100µg/ml Bacitracin completely inhibits growth (figure#57A, red curve).
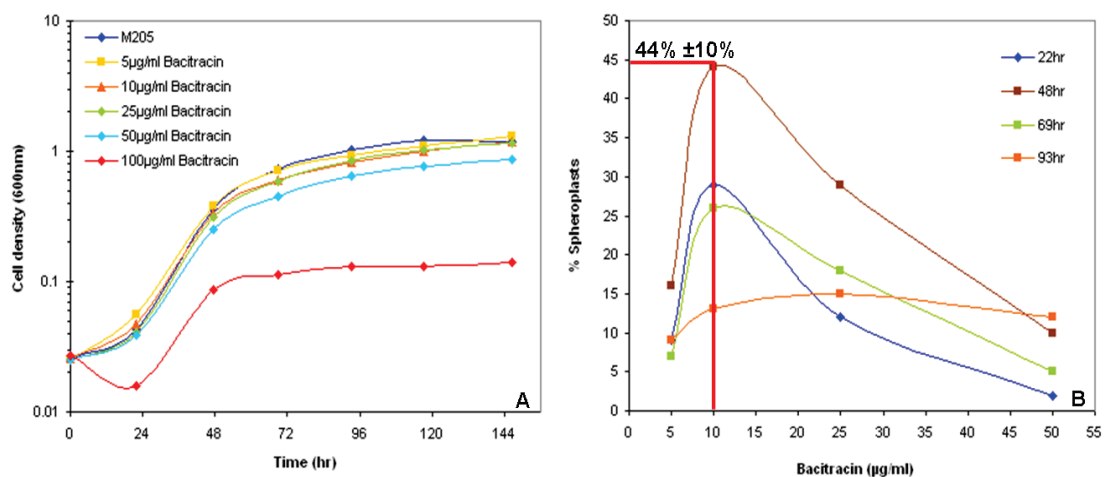
Figure # 70: **A**-Growth of *N. pharaonis* in different concentrations of Bacitracin. **B**-% of spheroplasts *vs*. different concentrations of Bacitracin at different time points.

The effect of Bacitracin on the morphology of *N. pharaonis* was examined in cultures supplemented with different concentrations of this antibiotic. Cell morphology was followed using phase-contrast microscopy (x100 with oil), and the percentage of spheroplasts seen at different time points is summarized in the appendix table #24. As seen in figure#70B, a peak of 44% spheroplasts was observed after 48hr when the medium was supplemented with 10μg/ml Bacitracin. The morphology of the cells at this concentration is shown in figure #71. Bacitracin greatly increased the population of cells that had lost their normal rod shape and become rounded up. Given the known mode of action of Bacitracin (Upreti et al., 2003) the results are consistent with the disturbance of S-layer formation leading to a weakened or absent cell wall, so forming spheroplasts.
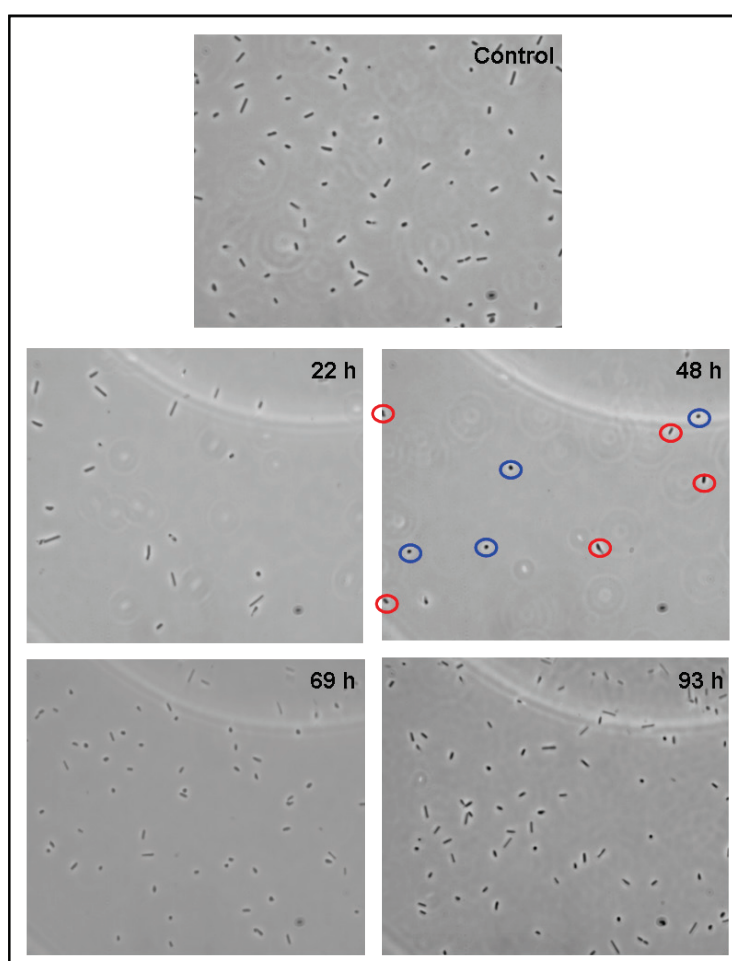
Figure # 71: Cell morphology of *N. pharaonis* at different times during cultivation in medium containing 10μg/ml Bacitracin. Samples taken at different times after inoculation of *N. pharaonis* into M205 medium with 10μg/ml Bacitracin. Time points are shown in the top right hand of each photograph. left Spheroplast (blue circle), rod shape (red circle). Control- *N. pharaonis* after 48h growth in M205 medium. Phase-contrast microscopy photos taken at x100.

A closer look at the relationship between the growing curve and the % of spheroplasts (figure #72B), reveals that the highest % of spheroplasts occurred at the end of the exponential phase and just before the beginning of the stationary phase. The decrease in the % of spheroplasts in the stationary phase may indicate either lysis of spheroplasts or their conversation back to rod shape cells.

To summarize these experiments, Bacitracin appeared to be effective in converting a significant proportion of *N. pharaonis* cells to spheroplasts (20-40%), and is potentially a useful treatment for introducing DNA into the cells.
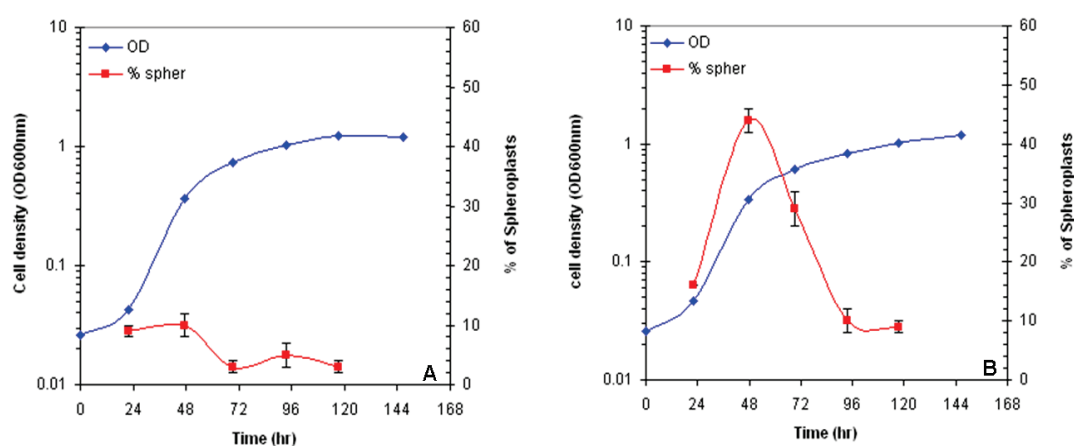
Figure # 72: Correlation between cell density and % of spheroplasts when growing *N. pharaonis* in **A**-medium without Bacitracin, **B**-medium containing 10μg/ml Bacitracin.

6.5.4.1.2 Protease treatment produces spheroplasts in N. pharaonis

Another way of removing the S-layer proteins from haloarchaea is treatment with proteolytic enzymes, such as the protease from *Streptomyces griseus* (Mescher and Strominger, 1976). This protease retains activity at high salt concentrations and has been shown to form spheroplasts of *H. salinarum*. In that study, the effect of protease on cell morphology and glycoprotein content of *H. salinarum* was observed both by electron microscopy and gel electrophoresis. In this way, Mescher and Strominger (Mescher and Strominger, 1976) demonstrated the shape-maintaining role of surface glycoproteins.

In the current study, a soluble protease from *Streptomyces griseus* was selected because it was active at high salt concentration and over a pH range between pH 5-9.

As seen in figure #73A, increasing concentrations of protease slowed the apparent growth rate, causing a 2 fold increase in the Gn times (Gn $_{(WO\ protease)}$=12.9h, Gn $_{(with\ 0.9mg/ml)}$=30.9h). However, this treatment was very effective at generating spheroplasts (appendix table #25). Spheroplast proportions of 69% and 66% were seen after 97h and 72.9h, respectively (figure #73B), when *N. pharaonis* was growing in medium with 0.3mg/ml Protease. Although at first glance, 0.6mg/ml and 0.9mg/ml protease appeared to be useful concentrations due to the high % of spheroplasts (90-100% at 3-5 days), the lower growth at those concentrations made 0.3mg/ml protease the more suitable choice, as it is produced ~70% spheroplasts at 3-4 days.
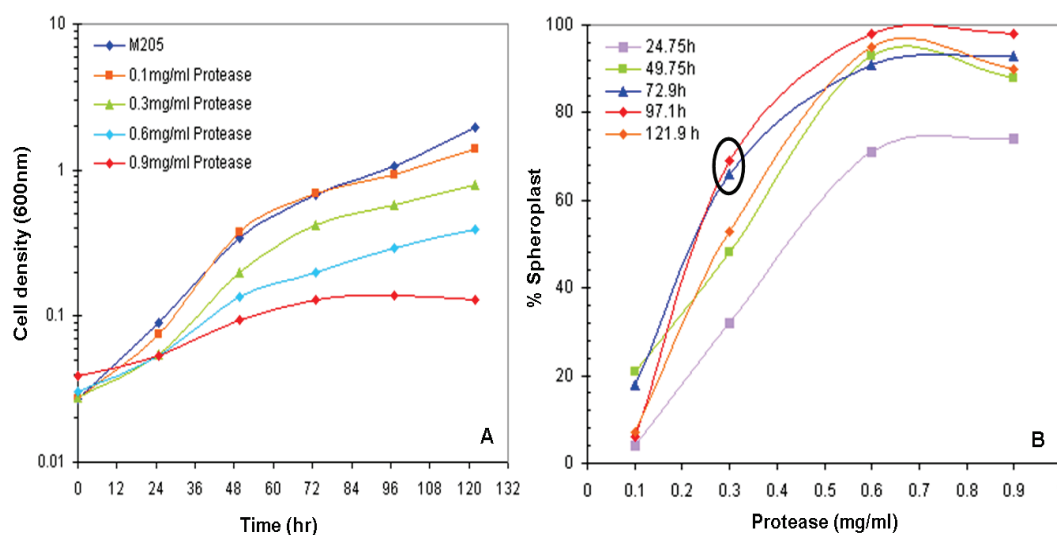
Figure # 73: **A**-Growth of *N. pharaonis* in different concentrations of Protease. **B**-% of spheroplasts *vs*. different concentrations of Protease at different time points.

As seen in figure#74B, in contrast to Bacitracin, the % of spheroplasts increased with the O.D of the culture till the end of the stationary phase (96h, ΔO.D600nm=0.58, 69% spheroplast). The morphology of the cells at different time points during growth in 0.3mg/ml Protease is shown in figure #75.
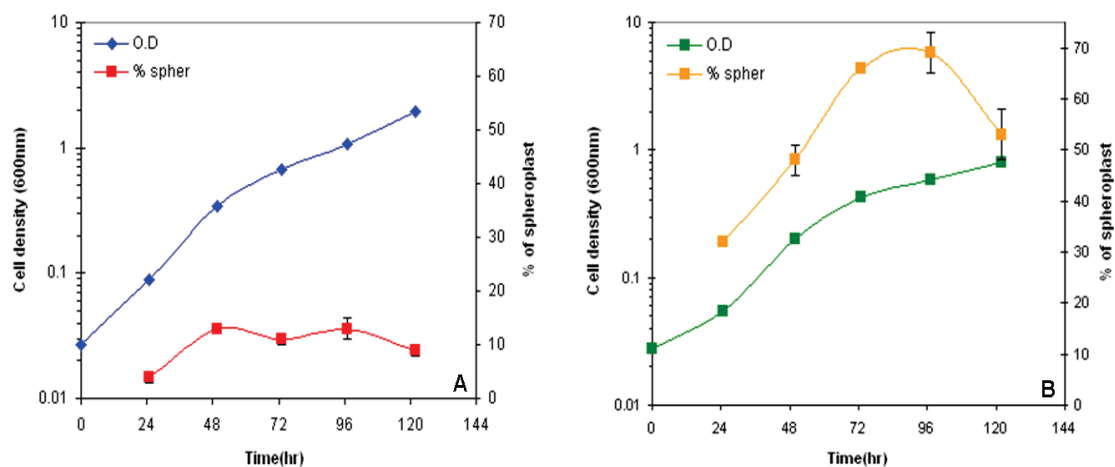
Figure # 74: Correlation between growth density and % of spheroplasts when growing *N. pharaonis* in **A**-standard M205 (40°C, 100rpm), **B**-medium with 0.3mg/ml Protease.
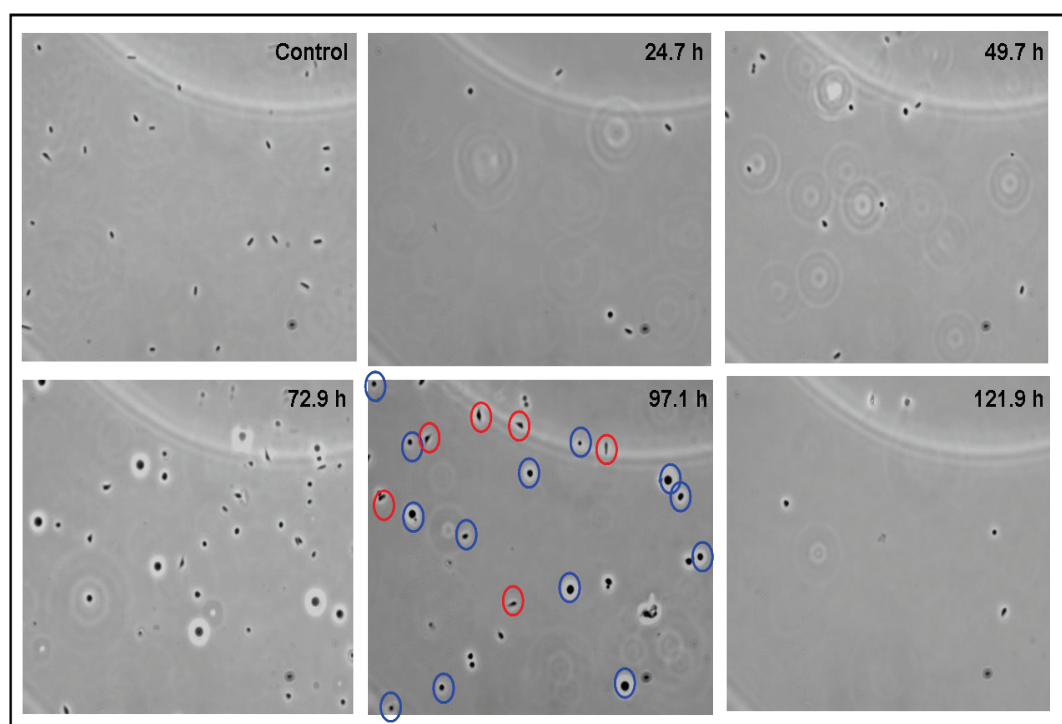
Figure # 75: Cell morphology of *N. pharaonis* at different times during cultivation in medium containing 0.3µg/ml Protease. Spheroplasts (blue circle), rod shape cells (red circle). Control-*N. pharaonis* after 97.1h in M205 medium. Phase-contrast microscopy, photos taken at x100.

136

The spheroplasts formed by *N. pharaonis* in M205 medium supplemented with 0.3mg/ml protease (62%) remained viable and grew again when inoculated into M205 medium without protease (figure# 76, left panel). Most cells regained their rod shape, with only ~20% spheroplasts, which is in the range of the background level seen in untreated cells (figure #76 left panel). When protease-treated cells were subcultured into M205 medium supplemented with 0.3mg/ml protease (figure #76, right panel), there was 48h lag phase in which the % of spheroplasts did not change. In the stationary phase the % of spheroplast was increased again to 62%.
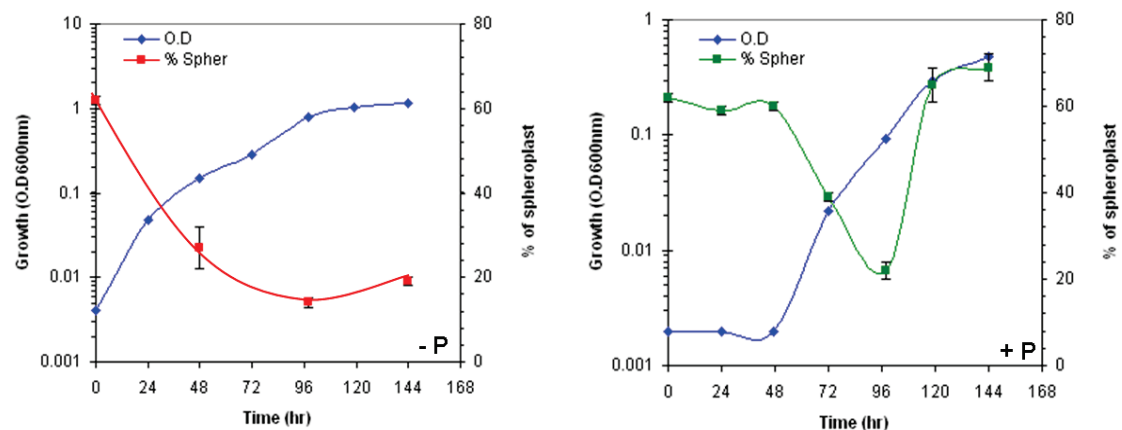
Figure # 76: Protease-treated *N. pharaonis* cells inoculated into M205 medium without Protease (**-P**, left panel), and into M205 medium with 0.3mg/ml Protease (**+P**, right panel).

6.5.4.1.3 Optimizing the composition of spheroplasting buffer (SPH)

Now that spheroplasts of *N. pharaonis* could be formed, other conditions used in the standard PEG method to introduce foreign DNA into *N. pharaonis*, were examined. Sucrose is a component of SPH buffer (2MNaCl, 25mM KCl and 15% (w/v) Sucrose) used to suspended cells before the addition of DNA and PEG_{600}.

While sucrose is known to provide an improvement in cell recovery it was thought that betaine, a naturally occurring compatible solute, may be superior. This was tested as follows. Viable *N. pharaonis* cells ($8.3x10^8$ cells/ml) were incubated in the presence of different concentrations of Basal salt pH-8.8 with or without osmolytes (Sucrose or Betaine), for 28h at RT. The cells were diluted the next day and plated on M205 plates (three different dilutions for each condition), and the viable count used to determine the % of cells which survived the treatment.

As seen in figure #77 (left most bar), the highest % of survival was obtained when *N. pharaonis* was incubated in 2M Basal salts supplemented with 2.3M Betaine (blue bar). No significant difference was seen when the medium included either basal salts or basal salts + sucrose (orange and green bars, respectively). Nevertheless, Betaine was about 3-fold better than 15% sucrose, giving ~28% recovery of viable cells. A high % of survival (24%) was obtained as well when the basal salts concentration was reduced to 0.5M in the presence of 3.8M Betaine. Furthermore, 0.74% of cells survived (6.1×10^6 cell/ml) when the basal salts concentration was reduced down to 0.1M supplemented with 4.2M Betaine. While the cell survival was low, the latter combination of buffers could be suitable for DNA uptake by electroporation.

The effect of lowering the basal salts concentration was also examined. *N. pharaonis* was able to survive at salt concentrations of 1M basal salt, but unable to survive at a salt concentration below 1M NaCl without Betaine (marked with black arrow in figure #77).
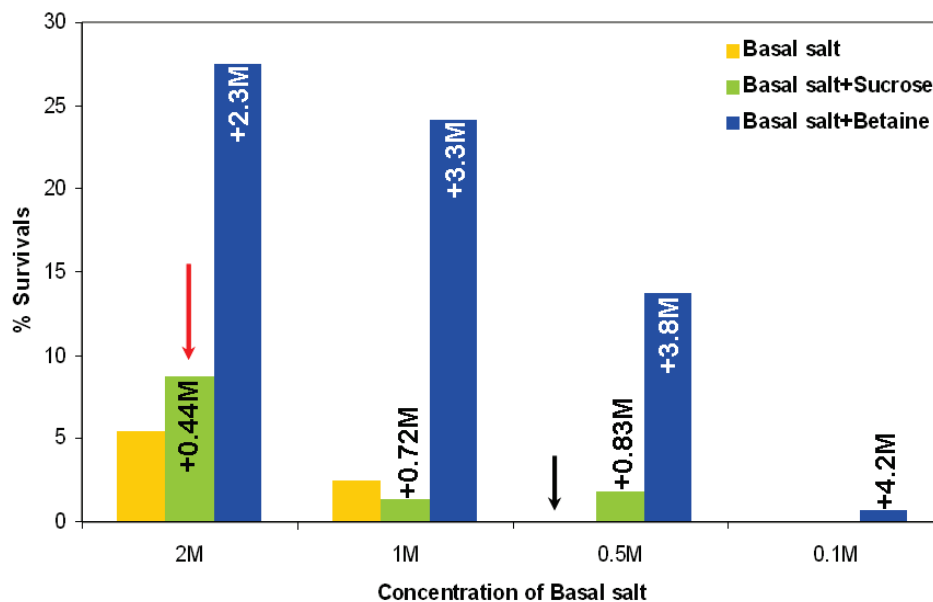
Figure # 77: % of *N. pharaonis* which survived 28h incubation in medium containing different concentrations of basal salt (orange bars) with and without Sucrose (green bars) or Betaine (blue bars). 100%=8.275×10^8 cells/ml. The conventional SPH buffer is indicated with red arrow

6.5.4.1.4 Combination of Bacitracin, protease and the PEG method

At this stage, collaboration with the group of Dr Angela Witte from the university of Vienna (Dep. Microbiology and Immunology), was established. Her laboratory had independently developed a method to transform haloalkaliphilic Archaea and had also used Bacitracin/protease treatment to form spheroplasts. The plasmids pMG100 and pMG200 were sent to her laboratory and were used to transform *N. pharaonis*. Potential transformants were screened for plasmid pMG100 as follows. Plasmid extracts from 3 different transformants colonies were used as template in PCRs with plasmid-specific primers. A 2393bp fragment was expected if the plasmid was present, while no amplification was expected if the cells did not contain plasmid. As seen in figure # 78, the expected fragment was visible only in pMG100 transformants (lanes 4-6). The PCR also generated some smaller products. No amplification was seen in the negative control (lane 3), while the plasmid control (positive control, lane 2) gave the expected band plus many shorter products. Regardless of the smaller PCR products, the PCR results gave convincing evidence that the transformants colonies harbored the pMG100 plasmid, showing that not only did the transformation method work for *N. pharaonis*, but also that the plasmid replication and Mev^R marker functioned as anticipated.
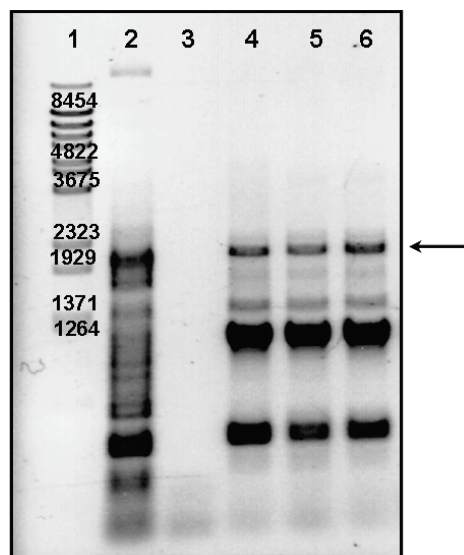
Figure # 78: PCR to detect pMG100 in *N. pharaonis* transformants. **Lane 1**: Lamda BstEII (in bp), **Lane 2**: positive control, isolated pMG100 plasmid, **Lane 3**-negative control, DNA from *N. pharaonis*, **Lanes 4-6**: plasmids from 3 transformed colonies after introduction of pMG100 (method of A. Witte).The expected fragment is indicated by an arrow.

6.5.4.2 Is electroporation possible in haloarchaea?

Electroporation is a common method for the introduction of foreign DNA into host cells using high-voltage electric pulses. The electric field induces transient pores in the cell wall and thereby increases the permeability of the host cells to macromolecules, such as DNA (Gehl, 2003; Mir, 2008; Neumann et al., 1982).

For known reasons, electroporation is not applicable in haloarchaea, as the method requires a high voltage (so a very low salt concentrations) and haloarchaea can not tolerate salt concentrations less than 1M NaCl ((Allers, 2005) and figure #77 indicated by black arrow). It is clear that in order to perform electroporation the salt concentration needs to be decreased to ≤100mM NaCl, while at the same time cell lysis needs to be prevented. The experiments summarized in figure #77 indicated that this could be achieved, as the external salt can be replaced by Betaine.

Betaine is highly soluble in water ($\leq$ 5M) and is a zwitterionic organic osmolyte, so by definition at natural pH it contains both negatively and positively charged moieties. This natural behavior of zwitterionic osmolytes makes them suitable candidates for use in electroporation solutions, as they have no net charge but can maintain the osmotic balance necessary for cell survival.

To test the possibility of replacing salts in the medium with Betaine, viable cells corresponding to 5.34×10^{8} cell/ml, were centrifuged, resuspended in M205 medium without or with different concentrations of Betaine, incubated 1h at RT, and than plated on M205 plates. As seen in figure #79A, 50.1%±1.7% (n=2) survived in M205 medium with 4.3M NaCl, whereas, 2.2%±1.7% (n=2) survived 1h incubation in 4.3M Betaine alone. Even though the % of survival is low, 2.2% cells is equal to 1.8×10^{7} cells/ml. Reducing betaine concentrations to 3M and 2M lowered the % of survival by 2 orders of magnitude, and 3 orders of magnitude, respectively (figure #79B).
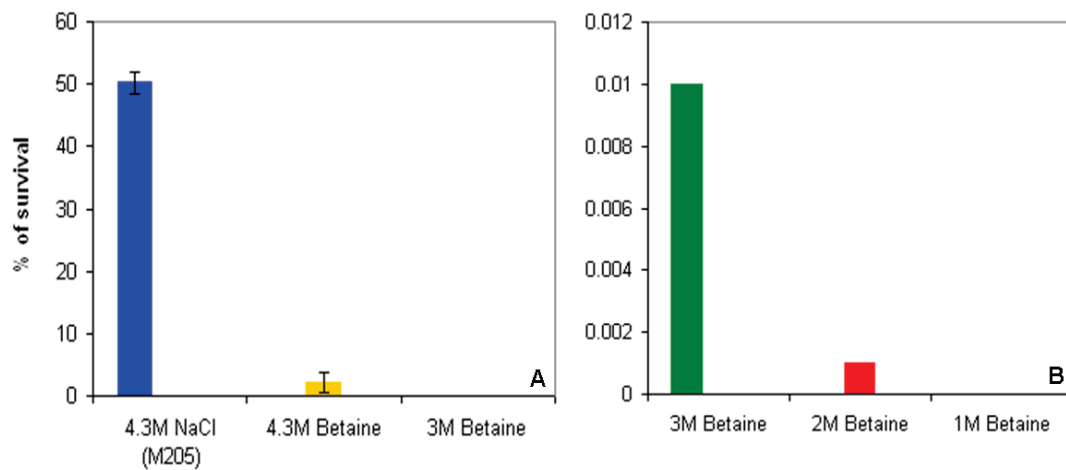
Figure # 79: % of *N. pharaonis* cells surviving after 1h incubation in M205 (Blue), and at different concentrations of Betaine. 100%=5.34x10^8 cells/ml.

6.5.4.2.1 Developing a protocol for electroporation of N. pharaonis

The results of the preceding experiments were used to develop a method that would allow electroporation of *N. pharaonis*. Cells grown in M205 medium (OD_{600nm} 0.5-0.7), were centrifuged (2min, 10000g, RT) and the supernatant was discarded. The cells were centrifuged again to remove residual medium and the pellet was resuspended with 4.3M Betaine in 50mM Tris HCl pH 8.7. Plasmid DNA was added to the cell suspensions and the tubes were incubated in RT for 10min. The mixture was electroporated and cells recovered then plated as described in section below.

6.5.4.2.1.1   Settings and conditions

Optimum electroporation conditions can differ not only between type and species but also for strains within species. Therefore, it is important to determine the best parameters of electroporation, such as the time constant and the field strength applied to the sample (Lurquin, 1997). The time constant is dependent on the total resistance of the sample and the capacitance selected on the electroporation device. The field strength is dependent on the initial voltage delivered by the electroporator and the distance between the electrodes of the cuvette (Lurquin, 1997).

The time constant was varied by using different resistence values while the capacitance and the voltage been constant (25μF, 1.5kV, see appendix table #20). Additionally, from 1-3 pulses were tried in order to compensate for the low observed time constants, and the different incubation buffers which contained either sucrose or Betaine. Controls included both *N. pharaonis* and *H. salinarum* but without the

addition of plasmid DNA. These were plated on non-selective media, and colonies counted after 14 days (C-71). Although, cells survived the electroporation, no successful transformation of the plasmids were obtained either with *N. pharaonis* or with *H. salinarum*. A small number of colonies of *N. pharaonis* and *H. salinarum* were obtained after electroporation (3 pulses of ∞ resistance with 1sec break in between) on selective plates (10μg/ml Mev) after 6 weeks incubation at 37°C (C-169), but there was no time available to determine if they contained plasmids.

6.5.4.3 Biolistic-particle delivery system (Gun particle)

Biolistic transformation is considered to be a useful tool for transformation of plants and animal cells (Daniell et al., 1990; Sanford, 1988, 1990; Williams et al., 1991; Yang et al., 1990). In 1990, this method was first used to transform the prokaryote *Bacillus megaterium*, which was chosen due to its large cell size (1.5x5.0μm) and the difficulty to transform this gram-positive by other methods (Shark K.B., 1990). Two years later, the same lab succeeded in transforming five bacterial species including *E. coli*, *Agrobacterium tumefaciens*, *Erwinia amylovora*, *Erwinia stewartii* and *Pseudomonas syringae* pv. *syringae*. They reported transformation rates of 10^4 transformants per 0.8μg DNA for *E. coli*, and lower transformation rates for the other species ($>10^2$ transformants per 0.8μg DNA).

The Biolistic®PDS-100/He device seen in figure #80 (Kikkert, 1993), is marketed by Bio-Rad laboratories, and is based on the PDS-1000 model, which was designed by Sanford at el. (Sanford J. C., 1987). Briefly, helium pressure and vacuum circuits are used to accelerate microcarriers, coated with DNA, into the target cells. Activating the Fire switch allows helium to flow into the gas acceleration tube. The gas is held until a rupture disk is burst. This burst generates a helium shock wave, which hits the microcarriers and launches them toward the target cells.
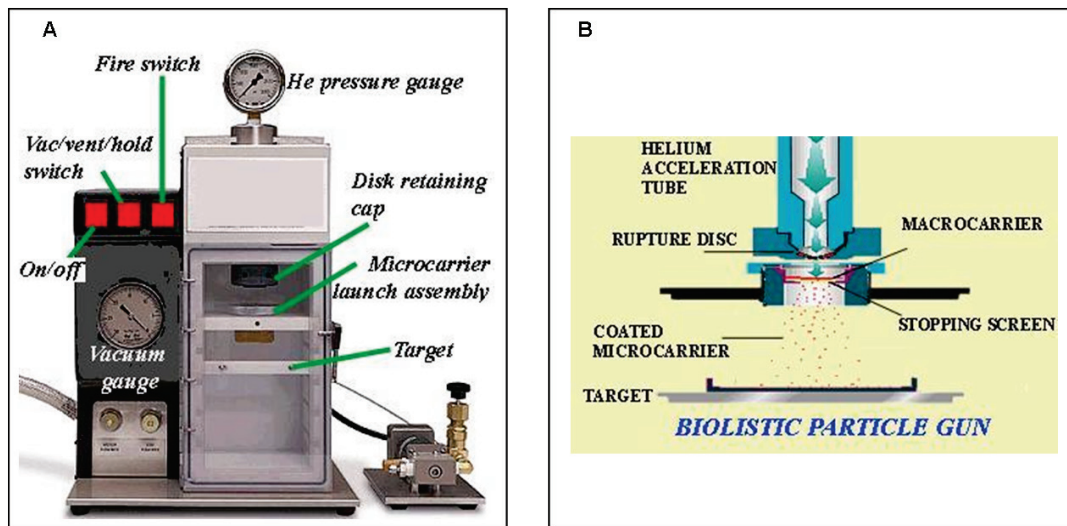
Figure # 80: **A**-The Biolistic®PDS-1000/He device. **B**-Components of the Biolistic®PDS-1000/He device: acceleration mechanism and main chamber.

6.5.4.3.1.1 Settings and conditions

After the bombardment, $1.5-3 \times 10^8$ cells/ml of *E. coli*, *N. pharaonis* and *H. salinarum* were plated and then incubated at 37°C (14 days for haloarchaea and o.n for *E. coli*).

500µg of microcarriers and 0.625µg DNA were used per bombardment, based on the method of Sanford et al., (Sanford et al., 1993). In this study, either M5 Tungsten particles or gold particles were used. M5 tungsten particles are characterized by a mean diameter of 0.71µm, a median of 0.362µm, and a mode in the range of 0.1-0.2µm (figure#2 in (Smith et al., 1992). The gold particles are characterized by a tight bimodoal distribution of particles, with peaks of 1.0 and 0.2µm (figure#2 in (Smith et al., 1992). Despite the disadvantages of M5 tungsten particles (more heterogeneous in size and shape compared to gold, and its ability to catalytically degrade DNA over time), it has been reported that M5 tungsten particles give significantly more transformants than cells bombarded with gold (transformants per plate- M5 tungsten particles: 3488, gold: 142, (Smith et al., 1992)).

Colonies were visible on selective plates (10µg/ml Mev) after 8 weeks, but time constraint prevented further analysis.

143

6.5.4.4 Lipofection in *N. pharaonis*

Metcalf et al., (Metcalf et al., 1997) successfully adapted the liposome-mediated transformation method for use with species of *Methanosarcina*. This method allowed transformation of 9 of the 11 strains tested, as well as *E. coli*. In 2005, Ehlers et al., was successful in using this method to transform *M. mazei* by improving the plating efficiency and optimizing the liposome-mediated transformation protocol (Ehlers et al., 2005). Even electrocompetent *E. coli* cells were successfully transformed using the liposome-mediated lipofection method, without the use of electroporation (transformation efficiency $2x10^4$-$1x10^5$ transformants/µg DNA) (Kawata et al., 2003).

In the current study, *N. pharaonis* cells were sonicated to disturb their cell walls followed by lipofection. Early exponential phase cells (OD_{600nm}=0.362), were sonicated (duty cycle-50%, output control-5) for 30sec, 70sec and 190sec, corresponding to 16.5%, 20.35% and 52% broken cells, respectively. As seen in table #13, sonication lowered the viable count from 184 to $24.5x10^6$ cell/ml, after 190sec of sonication. No significant difference in viable count was seen between 30 and 70sec of sonication. Colonies were visible after 12 days.

Next, the influence of sonication and lipofection with DOTPA (Roche) was tested. After transformation with the plasmids, treated cells were incubated in M205 medium with 15% sucrose o.n at 37°C, 250rpm. The cells were plated the next day and colonies were visible after 3-4 weeks. As seen in table #14, no significant difference was seen in the viable cell count whether DOTPA was added or not. There was no possibility to pursue these experiments further due to time constraint (G149, H-105).

Table # 13: The effect of sonication on viable cells of *N. pharaonis*. Cells were plated after dilution of $1/10^6$, in duplicates on M205 plates.

Tube no	Sonication (sec)	Colonies
9 (control, no sonication)	-	$184.5x10^6$ cells/ml
18	30	$126.5x10^6$ cells/ml
19	70	$116.5x10^8$ cells/ml
20	190	$24.5x10^7$ cells/ml

Table # 14: The effect of sonication and lipofaction with DOTPA on viable cells of *N. pharaonis*.

Sonication (sec)	DOTPA	# of colonies on selective plates [1]	
		pMG200[2]	pMG300[2]
0 (control, WO sonication)	-	full [c]	80
30	-	32	187
70	-	>200	5
190	-	32	ND
0 (control, WO sonication)	+	full [c]	full [c]
30	+	12	165
70	+	0	0
190	+	0	0

[1] treated cells were plated on M205+10μg/ml Mev+X gal, except for the controls, which were plated on unselective M205 plate.

[2] 3.7μg of pMG200 and pMG300 were used for transformation

[3] the plates contained too many colonies to be possible to count.

6.5.5 Future plans

To allow fundamental genetic studies and molecular engineering of *N. pharaonis*, genetic tools are essential i.e. shuttle vectors and a transformation method. The data presented in this chapter demonstrated that the plasmids constructed for *N. pharaonis* are indeed able to function as shuttle vectors (figure #67 and 78). These plasmids were used in different transformation methods, including the PEG method combined with bacitracin and protease, electroporation, gun particle delivery and lipofection. To date, it seems that the most promising transformation method was the PEG method combined with bacitracin and protease (figure #78). The other methods tested may still be useful especially because they are less time consuming procedures. All the methods presented need to be examined father to determine the optimum efficiency of transformation, but the basic obstacles to genetic manipulation of this host are now overcame.

7 Discussion

7.1 Microarray analysis of AroAAs-related genes

Of the 17 enzymes involved in shikimate pathway in *E. coli*, 13 homologues were found in *H. salinarum* (figure #8). These encompass the steps necessary to convert DHQ to AroAAs, indicating that down stream of DHQ, the canonical pathway for biosynthesis of AroAAs is used by *H. salinarum*. In the current study, a number of strategies were employed to assign function to individual genes of this pathway, ranging from targeted mutation of specific genes to a global examination of AroAAs regulated genes using a genome-wide microarray for *H. salinarum* (Twellmeyer et al., 2007). As seen in figure #81, the DNA microarray data provided solid evidence to support the function of the 13 genes downstream of DHQ, as these were differentially expressed ($p \leq 0.001$, fold induction ≥ 3.0) when AroAAs were omitted from the growing media of R1. Additional insights retrieved from the microarray experiments concerning regulation, transport systems and transcription factors are described in more detail in the following sections.

7.1.1 Regulation of gene expression, uptake and metabolic flow

Microorganisms require mechanisms that allow fast and efficient adjustments to rapid changes in external conditions. These are necessary for survival, and to conserve both energy and resources. Most microorganisms dedicate an appreciable fraction of their genetic material to regulatory functions, and by doing so, they ensure that they will have the flexibility to respond appropriately to changes in pH, temperature, available substrates, etc. Common regulatory checkpoints occur at various stages of gene expression (transcription, translation), and at the level of enzyme activity (feedback inhibition, enzyme activation). The intracellular levels of metabolites can also be utilized to activate or inhibit metabolic pathways. The interplay of multiple regulatory systems acting on specific enzymes and pathways can be extremely complex (Adhya, 1999). At an organizational level, the clustering of genes with related functions into single transcription units (operons) allows simultaneous regulation, and is an efficient means of producing the correct proteins ratios, such as the enzymes required for a specific biosynthetic pathway.
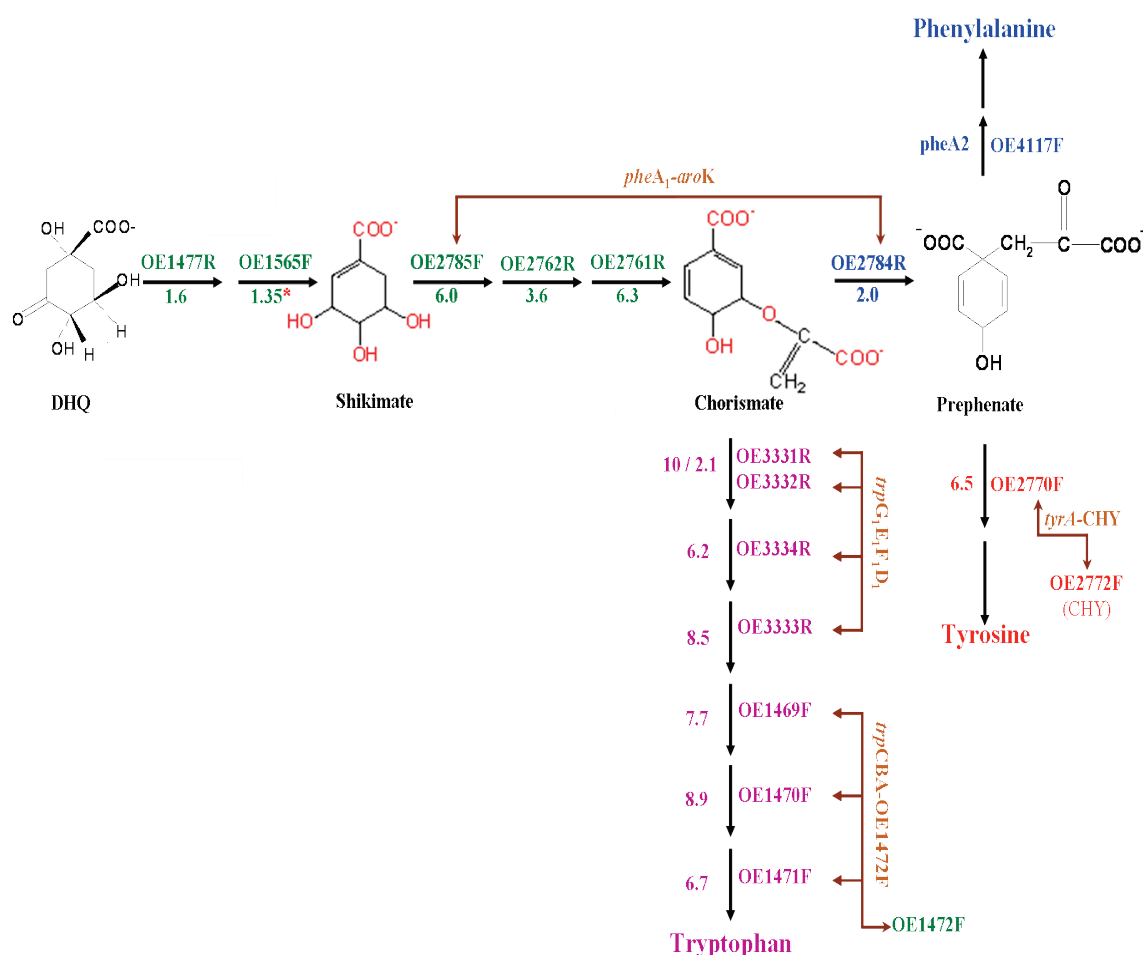
Figure # 81: Transcription of the differentially expressed AroAAs-related genes. All of the ORFs showed significant changes ($p \leq 0.001$, fold induction ≥ 3.0), except for OE1565F (aroE, $p=0.017$), and OE2762R (aroA, $p=0.003$). **Green** represents the ORFs in the chorismate pathway starting from DHQ, **pink**- the ORFs of the tryptophan branch, **blue**- the ORFs in the phenylalanine branch, and **red**- the ORFs in the tyrosine branch. ORFs organized in operons are marked with brown arrows. The operons are discussed with more details in the following sections.

7.1.1.1 Trp, tyrA and aroK-pheA1 operons

Based on the genetic context, two *trp* gene clusters were identified which contain the enzymes on the branch leading from chorismate to tryptophan. As seen in figure #82A, the first *trp* cluster consists of the overlapping ORFs OE1469F, OE1470F and OE1471F, followed by a 25bp sequence which connects ORF OE1472F to the cluster. The second *trp* cluster seen in figure #82B, $trpD_1FE_1G_1$, consists of four overlapping ORFs. Transcription of both clusters was strongly induced by the absence of AroAAs in the growth medium (figure #82).

Each of the *trp* clusters can be described as operons, as each possesses a common promoter region. In the first cluster (figure #82A), a putative TATA box promoter element was identified at position -33 from the starting codon of ORF OE1469F, while the 25bp sequence between ORF OE1471F and OE1472F, does not contain any TATA box sequence. The second *trp* cluster (figure #82B), contained a putative TATA box promoter element sequence at position -49. Both TATA box motifs were consistent with the conserved 8 bp promoter sequence element (TTTAWATR, with W=A or T, R=A or G) found in most Archaea (Gregor and Pfeifer, 2005). In a study by Koide et al., (Koide et al., 2009; Price et al., 2005), both of the *trp* clusters were identified as operons in *H. salinarum* sp. NCR-1, based on genomic-specific distance models and transcription start and end points of annotated operons by Koide et al,.
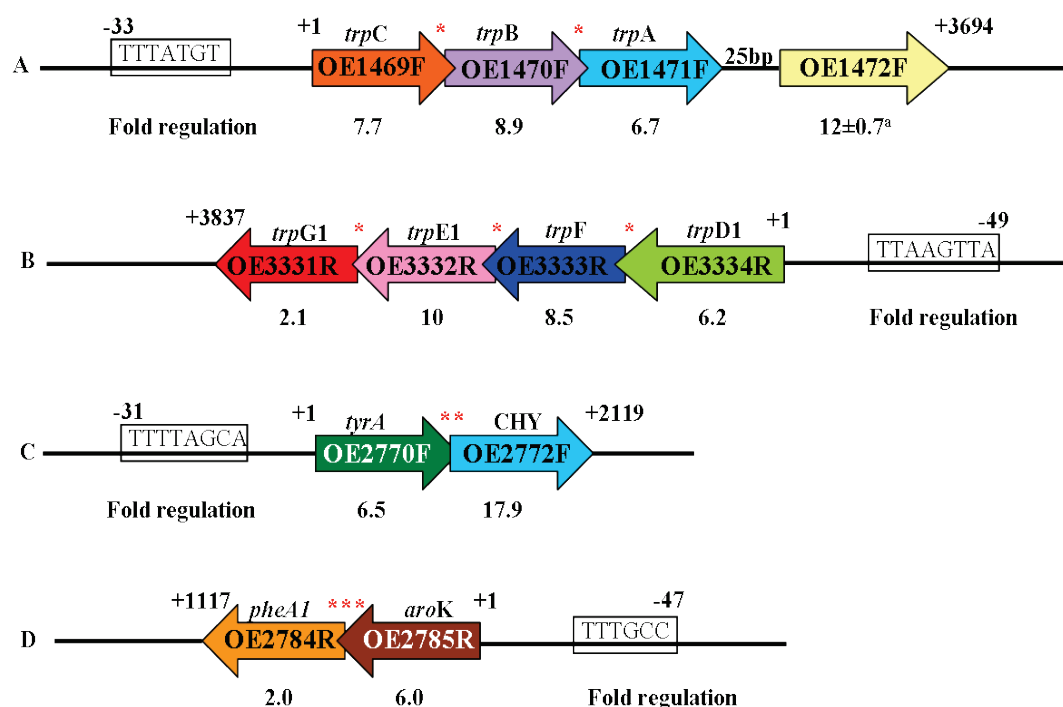
Figure # 82: AroAAs biosynthesis genes that display operon-like organization in *H. salinarum*. **A** and **B**: the two *trp* operons, **C**: the *tyr*A operon, **D**-*aro*K-*phe*A1 operon. The fold regulations as found in the DNA microarray experiment are indicated below the ORFs. [a]-fold regulation was calculated from RT-qPCR experiment using two biological repeats. *-represents a 4bp overlapping region, **-represents a 10bp overlapping region, ***-represents a 66bp overlapping region. The putative TATA box promoter motifs are indicted. CHY-conserved hypothetical protein.

In *H. salinarum*, *H. salinarum* sp NCR-1, *H. marismortui*, *H. walsbyi* and *N. pharaonis*, the genes specifying the seven enzymatic reactions leading to the

formation of tryptophan from chorismate are organized in a similar arrangement, meaning split into two separated operons (figure #83). This is different to the arrangement of *E. coli* and *methanococcus*. For example, *Methanobacterium thermoautotrophicum*, possesses all seven *trp* genes adjacent to each other in the order *trp*EGCFBAD (Gast et al., 1994; Meile et al., 1991), as do some other Archaea (figure #5, (Xie et al., 2003)). Surprisingly, the *trp* operon in *M. maripaludis* is organized in the same way as in *E. coli*, *trp*EDCBA but, in contrast to *E. coli*, it does not contain a 14 amino acid leader peptide immediately preceding *trp*E (figure #85). Although it is believed that essential metabolic pathways have been likely transferred horizontally between Archaea and Bacteria, the histidine operon in *E. coli*, *his*GDCNBHAFI, is another example were genes clustered in operon, were either scattered thorough the halobacterials genome (i.e. *H. salinarum*, *N. pharaonis*, and *H. walsbyi*), or partially clustered (i.e. *P. furiosus*) (Fondi et al., 2009).
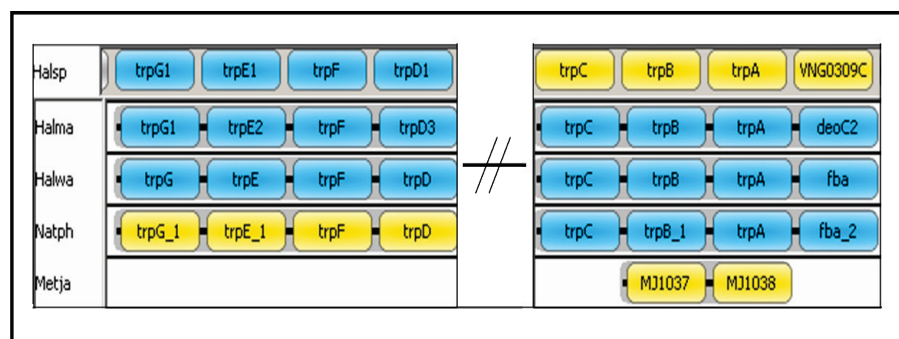
Figure # 83: *trp* operons in different archaea. $trpG_1E_1FD_1$ (left panel) and *trp*CBA-X (right panel) in different archaea. **Halsp-** *Halobacterium* sp. NRC-1, **Halma-***Haloarcula marismortui* ATCC 43049, **Halwa-** *Haloquadratum walsbyi* DSM 16790, **Natph-***Natronomonas pharaonis* DSM2160, **Metja-** *Methanocaldococcus jannascii* DSM 2661. Modified from WWW.Synteview.u-psud.fr (Lemoine et al., 2008).

Based on the microarray results, the overlapping gene pairs, *tyrA*-CHY, and *aro*K-*phe*A1 are also likely to form operons, as depicted in figure #82C, D, respectively. Both clusters were identified as operons in *H. salinarum* sp NCR-1 based on transcription start and end points (Koide et al., 2009; Price et al., 2005). The *tyrA*-CHY operon is composed of ORF OE2770F (*tyr*A) and a conserved hypothetical protein (CHY, ORF OE2772F). Both ORFs are regulated according to the DNA microarray results obtained from synthetic medium lacking AroAAs, indicating that the CHY is a utilized ORF and potentially significant for biosynthesis or regulation of AroAAs. One the basis of its position, it is possible that this CHY could catalyse the

formation of tyrosine from 4-hydroxyphenylpyruvate, which is catalyzed by *tyr*B in *E. coli* (EC 2.6.1.57 or EC 2.6.1.5, aromatic-amino-acid transaminase, and tyrosine transaminase, respectively). No candidates for *tyr*B have been identified in haloarchaea by sequence similarity to the *E. coli* gene, suggesting that ORF OE2772F might be either a new member of transaminase family (although no conserved motifs were found), or has some other, as yet unknown, role in the AroAAs pathway. In *E. coli*, *tyr*A clusters with ar*o*F which catalyses the first step in aromatic amino acid biosynthesis from the precursors PEP and E-4-P.

As seen in figure #82D, transcription of the gene cluster *aro*K-*phe*A1, was regulated by AroAAs and was found to be conserved in haloarchaea but not in *M. maripaludis*, *M. jannaschii* or *E. coli* (figure #84).
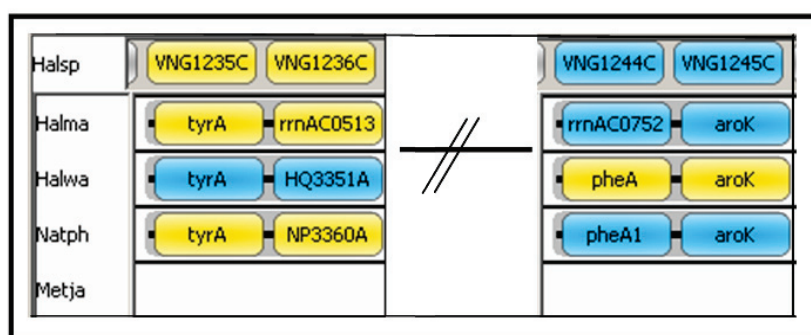
Figure # 84: *tyr*A and *aro*K-*phe*A1 operons in different archaea. **Halsp**- *Halobacterium* sp. NRC-1, **Halma**-*Haloarcula marismortui* ATCC 43049, **Halwa**- *Haloquadratum walsbyi* DSM 16790, **Natph**- *Natronomonas pharaonis* DSM2160, **Metja**- *Methanocaldococcus jannaschii* DSM 2661, **Esc05**-*E.coli* K12. Modified from WWW.Synteview.u-psud.fr (Lemoine et al., 2008).

As seen in figure #83 and 84, the *trp*$G_1E_1FD_1$, *trp*CBA-OE1472F, *tyr*A-CHY and *aro*K-*phe*A1 operons are conserved in *Halobacterium* sp. *NRC-1*, *H. salinarum*, *Haloarcula marismortui*, *Haloquadratum walsbyi*, and *Natronomonas pharaonis*. Differences are seen in *M. jannaschii*, in which the operons are missing or partly missing.

An arrangement of genes in operons confers both advantages and disadvantages. The most obvious advantage is that genes with similar function are transcribed together. The greatest disadvantage is that, unless some further level of regulation exist (differences in the amount of mRNA or its stability, the strength of ribosomal binding sites, and so on), the amount of polypeptides from those genes will be the same even though resultant enzymes may have different catalytic rates (Glansdorff,

150

1999). The enzymes with slower rates will be the limiting factor in the pathway. As a result, when genes are transcribed together, an excess of some enzymes is likely to occur. However, the amount of the mRNA and polypeptide synthesis is only one aspect of the control of the tryptophan pathway. In *E. coli*, besides these, there are two other levels of control that affect the amount of tryptophan synthesis within the cell. The first is feed-back inhibition which influence the activity of the first reaction (figure #86), and thereby the amount of metabolites flowing through the pathway. The second is the formation of multi-enzyme complexes that greatly increases the catalytic efficiency of the various reactions. In complexes, the product of one reaction can be used directly by the next enzyme because the concentration of the substrate in the vicinity of the second enzyme is much higher than would occur were the two enzymes separate. Examples of such complexes are *trp*E-*trp*G(D) and *trp*A-*trp*B (Ikeda, 2006).

The transcription of *trp* operon and *phe*A in *E.coli* are regulated by attenuation, in response to changing intracellular concentrations of trp and phe, in the cell (Yanofsky, 1981). A leader peptide contains "control codons" that specify the amino acid end product of the enzymes encoded in the operon. When the supply of amino acid and thus of cognate aminoacyl-tRNA is deficient, the ribosome stalls at the control codons, the leader RNA forms a preemptor 2:3 hairpin, and transcription starts (figure #85C). If the aminoacyl-tRNA is abundant, the ribosome stops at the end of the leader peptide, thereby forming the prohibiting 3:4 hairpin, and transcription of the remainder of the operon does not occur (attenuation, figure #85B). In *E. coli*, the leader peptide of the *trp* operon contains 2 trp codons, while 7 phe codons precede the Phe ORF (Gollnick and Babitzke, 2002). It seems that transcriptions of the *trp* operons in *H. salinarum* are not regulated by attenuation, as no leader peptide was identified, and moreover, no trp codons were identified upstream to the first ORF in the operons.
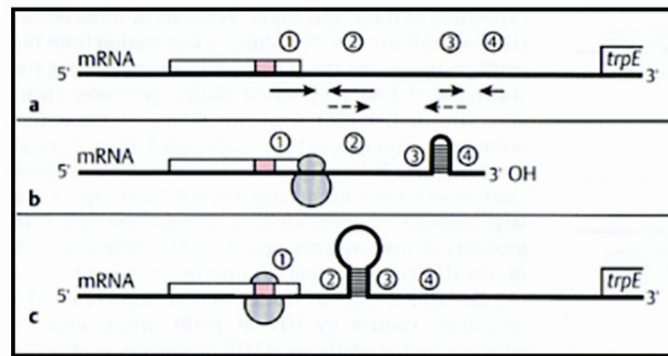
Figure # 85: Regulation by transcription attenuation in *E. coli*. **a**-5′-end of *trp* mRNA containing the leader region is shown. Open bar, leader peptide; red bar-control codons. The arrows indicate the regions (1-4) that form various potential RNA hairpin structures. **b**- if ribosomes (shaded) complete the translation of leader peptide (i.e., there is sufficient charged tRNA^{trp}), they allow 3:4 terminator hairpin. **c**-if the ribosome stalls at the control codons because of the lack of charged tRNA^{trp}, they permit 2:3 hairpin formation, which precludes attenuation. (Adhya, 1999).

7.1.1.2 AroAAs Transport systems

In *E. coli*, at least five different systems are known for the aromatic amino acid transport (figure #86). These include the specific transport systems for L-tyrosine (the *tyr*P gene product), L-phenylalanine (the *phe*P gene product), and L-tryptophan (the *mtr* and *tna*B gene products), as well as a general transport system (the *aro*P gene product) (Ikeda, 2006; Whipp and Pittard, 1977).

A search of the *H. salinarum* genome sequence using the *E. coli aro*P and *phe*P protein sequences (by BLASTP) revealed three potential candidate for AroAAs transporters: **OE2779F**, annotated as an amino acid transport protein (probable phenylalanine transport protein); **OE5101R** and **OE1288F**, both annotated as probable cationic amino acid transport protein (with overall 22-24% amino acid identity). BLASTP searches for homologs of *E. coli tyr*P, *mtr* and *tna*B proteins, retrieved additional ORFs: OE1476R, OE2891F, and OE4384R, respectively, all annotated as conserved hypothetical proteins.

If these genes are AroAAs transporters then it might be expected that their expression would be regulated by the external levels of one or more of the amino acids they transport. As shown by the microarray results, this was true for at least some of the predicted ORFs. ORF OE2779F (homolog to *aro*P and *phe*P) was 3.5 fold up regulated in R1 cells grown without AroAAs. Moreover, ORF OE4384R (conserved hypothetical) was 4 fold up regulated in the deletion strain ΔOE1477R

grown in synthetic medium supplemented with AroAAs and shikimate. It might be that in spite of the low similarity, this conserved hypothetical ORF has an unknown function associated with AroAAs biosynthesis, and one can speculate that ORF OE4384R could be a candidate for a second low-affinity transport system of tryptophan, while *aro*P (OE2779F) is the high affinity transport system.
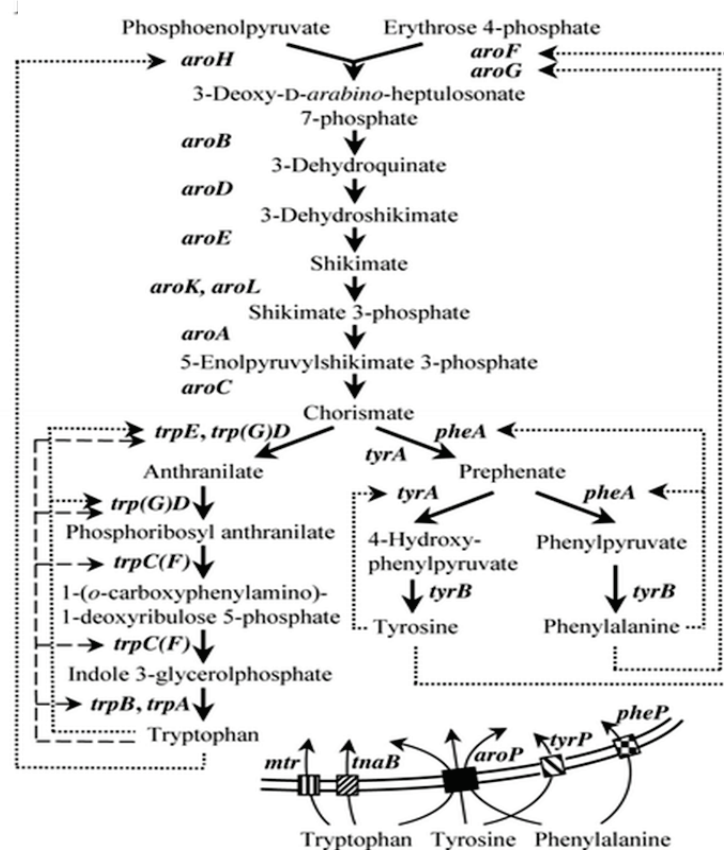
Figure # 86: Biosynthetic pathways, regulation and transport systems of aromatic amino acids in *E. coli*, modified from (Ikeda, 2006)). The dotted lines and the dashed lines indicate feedback inhibition and repression, respectively. Symbols for genes follow mostly the *E. coli* K-12 linkage map.

7.1.1.3 Cellular control mechanisms of AroAAs biosynthesis

The regulation mechanisms of AroAAs biosynthesis of *E. coli* and *M. maripaludis* would be discussed in some details while comparing to *H. salinarum* with the aim to recognize similar regulation patterns i.e. regolons.

7.1.1.3.1 Regulation of AroAAs biosynthesis in E. coli

In *E. coli*, the control of metabolic flow in AroAAs biosynthesis occurs mainly by end-product inhibition of bifunctional enzymes, such as chorismate mutase-prephenate dehydratase (specified by *phe*A and *tyr*A), and anthranilate synthase (specified by *trp*E, *trp*G(D)). This is illustrated in figure #86.

The first step in the pathway is carried out by three DAHPS isoenzymes (DAHPS-3-deoxy-D-arabion-heptulosonate-7-phosphate synthase, *aroF*, *aroG* and *aroH*, EC 2.4.1.54 (Davies and Davidson, 1982; Shultz et al., 1984; Zurawski et al., 1981*)),* each feedback-regulated by one of the aromatic amino acids Tyr (*aro*F), Phe (*aro*G), and *Trp* (*aro*H), respectively (Byng et al., 1983). The Phe-inhibited *aro*G (79%) and the Tyr-inhibited *aro*F (20%) deliver nearly all of the DAHPS activity for the cell. The Trp-inhibited *aro*H contributes only 1% of the cellular enzyme activity (Tribe et al., 1976). Inhibition of *aro*G (Phe-inhibited) and *aro*F (Tyr-inhibited) is highly effective and abolishes >95% of their specific enzyme activity (McCandliss et al., 1978; Schoner and Herrmann, 1976), whereas the Trp inhibition of *aro*H is only ≈60%, thereby resulting in a 40% constitutive *aro*H activity (Pittard, 1996).

In addition to feedback control, the expression of genes relevant to aromatic amino acid biosynthesis and transport are regulated at a global level, by regulons. By definition, regulons control multiple operons by one common, but specific regulator. The purpose of this type of regulation is to extend the same coordinated control to more than one operon that encodes a similar function. In *E.coli*, two AroAAs regulons have been identified: *Tyr*R and *Trp*R. The genes regulated by *Tyr*R comprise at least eight transcription units, each of which is regulated in a distinctive manner by the *Tyr*R proteins. The nature of regulation of each promoter (repression or activation) is determined by the binding site of the RNA polymerase and the different forms of the *Tyr*R protein (Pittard et al., 2005). Genes regulated by *Tyr*R include *aro*F, *aro*L, *tyr*P, *mtr*, *aro*P, *tyr*B and *aro*G. While the *Tyr*R gene product can act both as a repressor and as an activator of transcription, *Trp*R mediates only repression. The genes regulated by *Trp*R include *aro*H, the *trp* operon, *aro*L, and *mtr* (Ikeda, 2006). Once the transcription of *E. coli trp* operon initiated (regulated by *Trp*R), the elongating transcription complex is subjected to control by transcription attenuation (see section 7.1.1.1). The combined actions of repression (80-fold) and transcription attenuation (8-fold) allow approximately 600-fold regulation in response to changing

154

concentrations of intracellular tryptophan (Gollnick and Babitzke, 2002). No orthologs of *Tyr*R and *Trp*R have been identified in haloarchaea.

7.1.1.3.2 Regulation of AroAAs biosynthesis in M. maripaludis

Porat at el., (Porat et al., 2004) tested the specific activities of three enzymes of AroAAs biosynthesis in *M. maripaludis*; CM (chorismate mutase), PDT (prephenate dehydratase), and PDH (prephenate dehydrogenase). Measurements were made in cell extracts of *M. maripaludis* after growth in synthetic medium supplemented with AroAAs (table #3 in (Porat et al., 2004)). These data were used to elucidate the regulatory features of AroAAs biosynthesis, and are been summarized in figure #87. The presence of tyrosine was not found to affect the activity of CM, but inhibited the activity of PHD (4 fold), while activating PDT (2.2 fold). Supplementing the medium with phenylalanine had no influence on the specific activity of CM or PDH, but inhibited strongly the activity of PDT (3.4 fold).

Additionally, it was found that aryl acids are substrates of a second pathway for the biosynthesis of AroAAs in *M. maripaludis* (figure #87). This was proven by growth experiments where three aryl acids fulfilled the requirement for AroAAs by the auxotrophic mutant. The three aryl acids, indoleacetate, p-hydroxyphenylacetate, and phenylacetate are activated to their CoA derivatives, and then reductively carboxylated by the enzyme indolepyruvate oxidreductase (IOR, EC 1.2.7.8) to indolepyruvate, p-hydroxyphenylpyruvate, and phenylpyruvate, respectively. These substrates are then converted to tryptophan, tyrosine and phenylalanine, respectively, by aromatic aminotransferases (AroAT). It was shown experimentally by Porat at el., (Porat et al., 2004), that addition of p-hydroxyphenylacetate to the synthetic medium did not effect the activity of either CM or PDT but inhibited the activity of PDH (2.3 fold), while addition p-hydroxyphenylpyruvate inhibited the activity of PDH as well (2.3 fold). Furthermore, addition of aryl acids to the growth medium reduced the activity of DHQ dehydratase by one-third. Porat et al., (Porat et al., 2006), suggested that the expression of fructose-1,6-phosphate (ORF MJ1585) was regulated by aryl acids, due to the nearly complete absence of labeled shikimate and dehydroshikimate in cells following growth with aryl acids (Porat et al., 2006).

In halophilic archaea on the other hand, the occurrence of AroAAs synthesis via indolepyruvate oxidoreductase is unlikely, since only two ferredoxin- dependent

oxidoreductase complexes, pyruvate- and 2-oxoglutarate-ferredoxin oxidoreductase, have been identified within their genomes (OE2623R, OE2622R, and OE1711R, OE1710R, respectively). No homologs of indolepyruvate- ferredoxin oxidoreductase (*iro*A2, EC 1.2.7.8, MMP0713, and MMP0714, the α and the β subunits, respectively) have been found in *H. salinarum*. If aryl acids can be utilized by *H. salinarum* for AroAAs biosynthesis then the enzymes are unrelated to those of *M. maripaludis*. Alternatively, another still unknown additional pathway for synthesis of AroAAs may be present.

From the regulation of AroAAs biosynthesis in both *E. coli* and *M. maripaludis*, one can conclude that neither of them can provide a model for the regulation of AroAAs in *H. salinarum*.
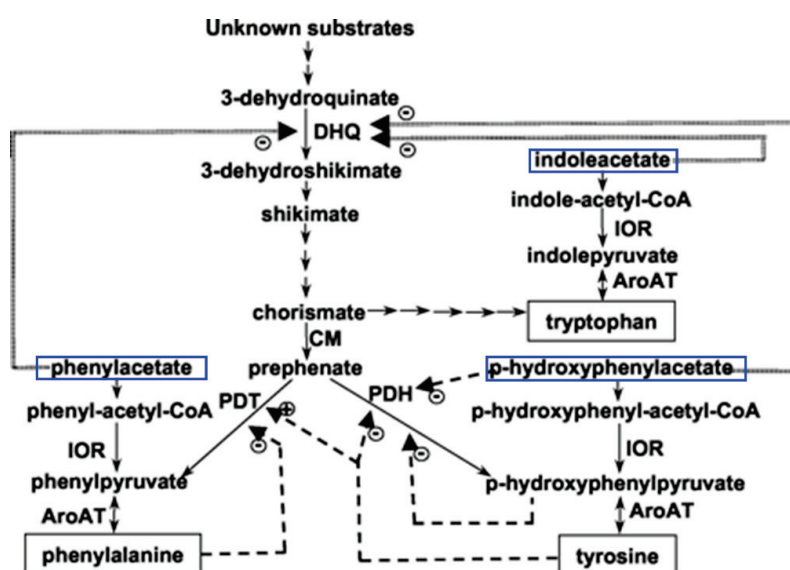
Figure # 87: Pathways for the biosynthesis of AroAAs in *M. maripaludis* as presented in (Porat et al., 2004). The aryl acid pathway starts from phenylacetate, p-hydroxyphenylacetate, and indoleacetate (marked with blue boxes), precursors for phenylalanine, tyrosine, and tryptophan, respectively. The enzymes shown are **DHQ** (3-dehydroquinate dehydratase), **CM** (chorismate mutase), **PDT** (prephenate dehydratase), **PDH** (prephenate dehydrogenase), **AroAT** (aromatic aminotransferase), and **IOR** (indolepyruvate oxidreductase). Possible transcriptional regulation of DHQ is indicated by double dashed lines. Inhibition (–) or activation (+) of the PDT and PDH enzyme activities are indicated by single dashed lines. (Porat et al., 2004).

7.1.1.4 Transcription factors: TBP, TFB and TrmB

Transcription in archaea is catalyzed by a eukaryotic type RNA polymerase, that requires only two transcription factors, TATA-binding protein (TBP) and transcription factor B (TFB). However, multiple transcription factor homologues have

been identified in *H. salinarum* R1 (6 TBPs, and 7 TFBs) (Baliga et al., 2000; Facciotti et al., 2007). A systematic knock out of each of the TBPs and TFBs has revealed that two TFBs (-f and -g) and three TBPs (-b, -e, and -f) appear to be essential for growth under standard laboratory conditions (Facciotti et al., 2007). It was also observed that TBPe interacts with numerous TFBs (TFBb, -c, -d, -e, and -g), suggesting a dominant role among TBPs for TBPe. This hypothesis was supported as well in the current study, where DNA microarray results of R1 revealed that only TBPe (ORF OE4146F) was regulated (3.2 fold) by addition of AroAAs to the growth medium, while transcription of the TFBs remained unchanged (1.0 fold regulation).

A TrmB-like global transcriptional regulator (VNG1451C) was identified in *H. salinarum* NRC-1, and was predicted to act on 113 different promoters involved in diverse metabolic pathways, such as glycolysis, TCA cycle, amino acid biosynthesis pathways and the biosynthesis of their cognate cofactors {Schmid, 2009 #236} Interestingly, some of the ORF involved in AroAAs biosynthesis were noted by Facciotti et al., to be directly controlled by trmB. These include: OE1475F (3-dehydroquinate synthase in *H. salinarum*, EC 4.6.6.3), OE2761R (*aro*C), OE1565F (*aro*E, trmB acts as repressor), and OE2563R (*men*D, one of the ORFs involved in menaquinone biosynthesis), while ORFs OE1472F (1st gene involved in the de novo pathway), and OE2019F (an F-1,6-P aldolase, trmB acts as activator), were regulated by TrmB but not directly controlled by it (supplementary material table #6 {Schmid, 2009 #236}. In the current study, the archaeal TrmB ortholog OE3018F in *H. salinarum* R1 was found to be weakly induced (1.9 fold regulation) when AroAAs were added to the growing medium of *H. salinarum*.

7.1.1.5 Lrp-like proteins

The *trh* (or LrpA/AsnC) family of transcription regulators was first discovered in *E. coli* (Willins et al., 1991) and controls the expression of large number of operons, directly or indirectly. They can act as both repressors and activators. A recently published DNA microarray analysis revealed that Lrp affects transcription of at least 10% of all *E. coli* genes, including most of the genes that are expressed upon entrance into stationary phase (Tani et al., 2002). The *E. coli* Lrp regulon conntrols genes that are involved in amino acid metabolism. *H. salinarum* possesses 7 Lrp-like transcriptional regulators (*trh*1 to *trh*7), which have homologues that are widely

distributed across bacterial and archaeal species (Bateman et al., 2004). Their specific role in the regulation of *H. salinarum* strain NRC-1 was unknown until a study by Bonneau et al.,(Bonneau et al., 2006) which predicted that 4 *trh* proteins (*trh*3, *trh*4, *trh*5 and *trh*7) plays a significant role in coordinating the expression of diverse cellular processes with competing transport processes. By showing the subnetwork surrounding *trh*3, *trh*4, *trh*5 and *trh*7, they found that each *trh* protein regulates a unique set of bioclusters. For example, they predicted that trh3 is a repressor of phosphate and amino acid uptake systems and that it is co-regulated with diverse metabolic processes involving phosphate consumption. Moreover, Schwaiger et al., (Schwaiger, 2010) showed that Lrp (*trh*4, OE3923F) is responsible for activating the expression of glutamine synthetase gene *gln*A, for regulating the peptide-and phosphate transport, and for activating the expression of the transcription regulator *sir*R. While LrpA1 (*trh*7, OE2621R) repressed the expression of the aspartate transaminase gene *asp*B3.

The DNA microarray results obtained in the current study showed that ORF OE2776F (*trh1*=LrpA2) was strongly regulated (14.9 fold) when R1 was grown in synthetic medium without AroAAs, indicating that LrpA2 is a regulator of AroAAs biosynthesis. Furthermore, in cells grown in synthetic medium supplemented with AroAAs and shikimate, ORF OE2621R (*trh*7=LrpA1) was also strongly regulated (15 fold) in the WT strain compared to the deletion mutant ΔOE1477R. ORF MJ0723 (*Prt*2) of *M. jannaschii* (homolog of OE2621R with 32% amino acid identity), was shown in vitro to positively regulate the archaeal transcription apparatus by recruitment of the TATA-binding protein (Ouhammouch et al., 2003).

7.2 The first three genes in the pathway

The genome of *H. salinarum* possesses homologs for five out of seven genes of the de novo pathway of AroAAs biosynthesis up to chorismate. ORFs OE1472F and OE1475F were assigned according to *M. jannaschii*'s ORFs MJ40400 and MJ1249, respectively, which have been shown to play a role in the alternative pathway of DHQ biosynthesis (Porat et al., 2006; White, 2004), while ORF OE1477R is homologous to *aro*D, which encodes the third enzyme in the pathway.

While the presence, sequence similarities and gene organization are reasonable evidence for assigning putative function for these genes, solid proof must come from experimental data. In this study, a number of strategies were employed to determine the role of these genes, ranging from targeted mutation of specific genes, nutritional requirements, phenotype, and uptake assays to specific activity of the purified gene product. In the following paragraphs the results of these studies will be discussed separately for each ORF.
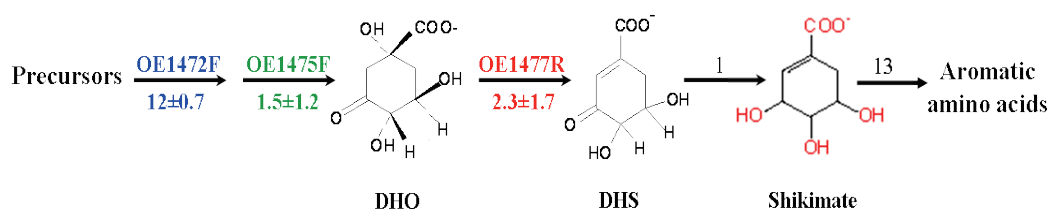
Figure # 88: ORFs catalyzing the first three steps of the proposed de novo pathway for the biosynthesis of AroAs in *H. salinarum*. The fold induction of these ORFs, as calculated from RT-qPCR experiments, is indicated below the arrows. The numbers above the arrows downstream to DHS indicates the numbers of enzymatic steps required to synthesis the intermediates.

7.2.1 ORF OE1472F - the first ORF in the AroAAs pathway

ORF OE1472F was annotated in the *H. salinarum* genome as fructose bis-phosphate aldolase II (EC 4.1.2.13), an enzyme that cleaves F-1,6-P to the corresponding triose phosphates. The sequence of OE1472F is 43% identical with that of MJ0400 from *M. jannaschii* (Porat et al., 2006; White, 2004), which has been demonstrated to be member of the superfamily an aldolase/transaldolase {Morar, 2007 #224}, catalyzing ASA and DKFP via an elimination of hydroxypyruvaldehyde phosphate (HPAP) to yield 2-amino-3,7-dideoxy-D-threo-hept-6-ulosonic acid

(ADH). In the current study, the role of the gene product OE1472F in *H. salinarum* was assessed experimentally and the results discussed below.

7.2.1.1 Classification of OE1472F as an archaeal type aldolase class IA

The presence of two class IA aldolase paralogues in *M. jannaschii*, *A. fulgidus*, *H. salinarum* NRC-1 and *H. salinarum* R1 (Siebers et al., 2001) raises the question why there is a need for two aldolases? Galperin et al., (Galperin et al., 2000), predicted that although both of the orthologs are aldolases, one member will be an aldolase whereas the other will be a deoxyribose phosphate aldolase or a transaldolase. Like transaldolase, aldolase would cleave F-1,6-P into GAP and DHAP. Whereas an aldolase releases both of these molecules as products, a transaldolase releases only GAP and transfers the DHAP group to an aldose acceptor. Thus, aldolases do not require an aldose acceptor substrate for turnover to occur, whereas a transaldolase does.

The two paralogues genes in *H. salinarum*, OE2019F and OE1472F (table #1), were annotated as fructose bis-phosphate aldolase I, and II, respectively. Although both ORFs showed similar aldolase activity (figure #56A), ORF OE2019F is expected to be an aldolase, while ORF OE1472F is likely to be a transaldolase that initiates the AroAAs biosynthesis pathway. The latter assignment is based on the gemonic position of this ORF is part of the multicistronic *trp*CBA-OE1472F operon.

Recent work by White (White, 2004) on *M. jannaschii* showed that the gene product MJ0400, annotated as 2-amino-3,7-dideoxy-D-threo-hept-6-ulosonic acid (ADH) synthase, catalyses a transaldolase reaction between DKFP and ASA to yield ADH. A crystal structure of ADH synthase (2.6Å) from *M. jannaschii* in a complex with F-1,6-P and with DHAP was solved by Morar et al (Morar et al., 2007), who also compared the crystal structure of the *M. jannaschii* ADH synthase to the enzyme from *Thermoproteus tenax* (*T. tenax*). Both of the enzymes were homodecamers containing the active site at the top of the barrel (on strand β6), with approximately 60% conserved residues in the active site (Morar et al., 2007). The *M. jannaschii* enzyme was confirmed to be a member of the class the 1A aldolase/ transaldolase family based on the fold of the enzyme, the utilization of a Schiff base intermediate, and the presence and localization of the catalytic Lys-Tyr residues.

The sequence alignment of OE1472F to MJ0400 from *M. jannaschii* is presented in figure #89, and shows the following: (1) 43% of the amino acids are completely conserved while 61% of the residues are similar. (2) all the active site residues described by Morar et al for *M. jannaschii* (Morar et al., 2007) can be identified in OE1472F from *H. salinarum* and are all completely conserved. (3) the conserved sequence motifs in archaea described by Siebers et al (Siebers et al., 2001) were all identified. (4) when superposition of the MJ0400-F1,6-P complex (figure #89, lower panel, in blue) with the OE1472F sequence (figure #89, lower panel, in red), one can see almost no differences except in the unstructured loop regions. On the basis of sequence alignment, presence of active site residues, conserved motifs and that the purified protein can form octamers (appendix section 8.8), *H. salinarum* ORF OE1472F is likely a member of the class 1A aldolase/ transaldolase family (figure #89).
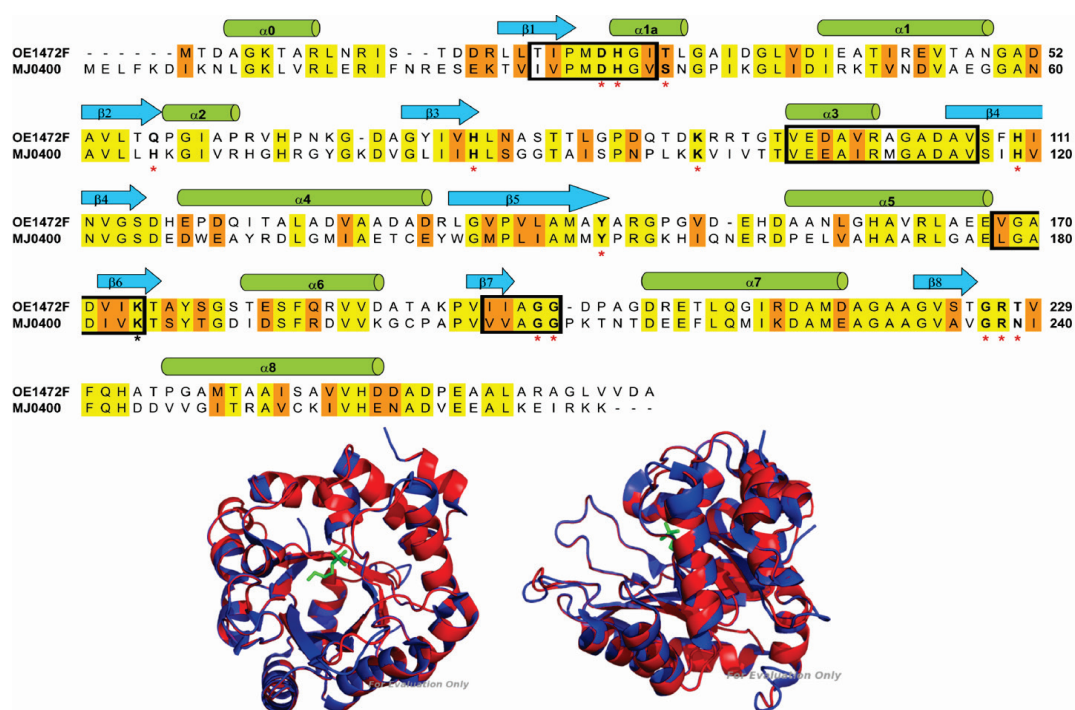
Figure # 89: Sequence alignment of OE1472F from *H. salinarum* and MJ0400 from *M. jannaschii*. Conserved sequence motifs in archaea described by Siebers et al (Siebers et al., 2001) are marked with black boxes. Active site residues described by Morar et al for *M. jannaschii* (Morar et al., 2007) are marked in boldface letters and *. The catalytic lysine residue (Lys237) determined for the *E. coli* class IA aldolase (DhnA type aldolase), is marked with a black asterisk (Thomson et al., 1998). Completely conserved residues are colored in yellow (43%); similar residues are colored in orange (61%). The alignment was done using ClustalW multiple sequence alignment program (http://us.expasy.org). The secondary structure elements according to MJ0400 (Morar et al., 2007) are shown at the top of the alignment. **Lower panel**-Superposition of MJ0400-F1,6-P complex with OE1472F. Blue-MJ0400, red-OE1472F, green-F-1,6-P. The PDB file of OE1472F was generated by Phyre (http://us.expasy.org) and 3D structure was viewed by Pymol.

7.2.1.2 Enzyme activity of the OE1472F and OE2019F proteins

To determine the biological function of ORFs OE1472F and OE2019F, the gene products were expressed in *E. coli* and purified. Assuming that the proposed de novo pathway of *M. jannaschii* (White, 2004) takes place also in *H. salinarum*, OE1472F (homolog to MJ0400) was predicted to catalyze a transaldolase reaction between DKFP and ASA, while OE2019F was predicted to be an aldolase. These predictions were tested using *E. coli*-expressed and purified proteins.

Unexpectedly, both OE1472F and OE2019F purified proteins displayed similar F-1,6-P aldolase activity as measured by the colorimetric assay (table #15, *H. salinarum*). From the comparison to *M. jannaschii*, OE1472F (homolog of MJ0400) showed 70-fold higher specific activity compared to MJ0400, while the specific activity of OE2019F (homolog of MJ1585) was 6.9-fold, 2.9-fold and 7.4-fold lower compared to aldolases from *M. jannaschii, T. tenax* and *P. furiosus*, respectively. Although one can clearly identify MJ1585 from *M. jannaschii* as an aldolase, this conclusion can not be made for OE1472F of *H. salinarum*, because of its similar specific activity towards F-1,6-P.

Table # 15: Aldolase and transaldolase activity detected in different archaea.

Organism	Assay	Specific activity (mU/mg Protein)		Reference
		ORF	**activity**	
M. jannaschii	coupled assay, recombinant proteins	[1] MJ0400 [2] MJ1585	<0.1 540	(Porat et al., 2006)
M. maripaludis	coupled assay, cell extract	WT, [1] ΔMMP0686	6.6-7.2 6.6-7.2	(Porat et al., 2006)
M. jannaschii [4]	GC-MS, purified enzymes	[1] MJ0400+ [2] MJ1249 MJ0400+MJ1249+NADP MJ0400+MJ1249+NAD	4.8 103.1 226.8	(White, 2004)
H. salinarum	colorimetric assay, purified proteins	[1] OE1472F [2] OE2019F	71 78	This study

| T. tenax[5] | coupled assay, recombinant proteins | AJ310483[7] | 230 | (Siebers et al., 2001) |
| P. furiosus[6] | coupled assay, recombinant proteins | AF368259[7] | 580 | (Siebers et al., 2001) |

[1]**MJ0400** from *M. jannaschii* is homologous to OE1472F from *H. salinarum* and MMP0686 from *M. maripaludis*. It is believed to be an transaldolase that catalyzes the first reaction in the AroAAs biosynthesis in these organisms.

[2]**MJ1585** from *M. jannaschii* is homologous to OE2019F from *H. salinarum* and MMP0293 from *M. maripaludis*. It is an aldolase.

[3]**MJ1249** is homologous to OE1475F from *H. salinarum* and MMP0006 from *M. maripaludis*. It is believed to catalyze the second reaction in the AroAAs biosynthesis, synthesizing DHQ.

[4] The formation of DHQ was measured using 1.6mM ASA and 2.2mM DKFP as substrate (White, 2004), while in all others the substrate was F-1,6-P.

[5] *Thermoproteus tenax*, crenarchaeon

[6] *Pyrococcus furiosus*, euryarchaeon
[7] GenBank accession number

The fact that OE1472F displayed aldolase activity towards F-1,6-P may indicate that F-1,6-P is playing a role in the biosynthesis of AroAAs. Moreover, as demonstrated by Morar et al ((Morar et al., 2007) and figure #89), both transaldolases (OE1472F and MJ0400) can bind F-1,6-P in the active site. While MJ0400 showed no detectable aldolase activity, OE1472F displayed aldolase activity towards F-1,6-P (table #15).

Together, the data may indicate either (1) ASA and F-1,6-P are the precursors in the AroAAs biosynthesis pathway in *H. salinarum*, or (2) ORF OE1472F has an additional role in the AroAAs biosynthesis pathway, such as the formation of DKFP. If the former is true, an aldole-like condensation between DHAP moiety of F-1,6-P and ASA is expected. GAP will be released, and compound I will be formed. However, oxidative deamination and phosphorylation by ORF OE1475F will not lead to the formation of DHQ (figure #90A). This will only be possible from ASA and F-1,6-P if HPAP (hydroxpyruvaldehyde phosphate) is released after the aldole-like condensation (figure #90B). But this is not likely as OE1472F is classified as an aldolase/transaldolase, and is expected to cleave F-1,6-P to GAP and DHAP.
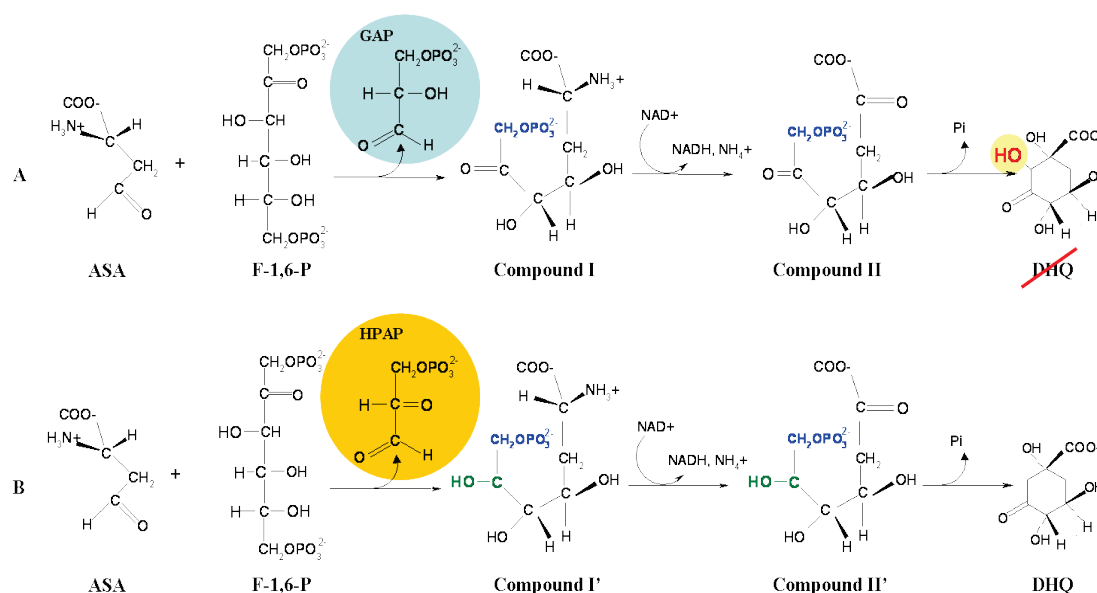
Figure # 90: The expected products and intermediates after cleavage of F-1,6-P and condensation with ASA. **A-** Aldoe-like condensation of DHAP moiety from F-1,6-P to ASA dose not result in DHQ. **B-** Formation of DHQ is possible only when compound II' is formed after releasing HPAP and oxidative deamination of compound I'. Compound II' is identical to the first intermediate formed in the canonical pathway when PEP and E-4-P are condensed (IUBMB Enzyme Nomenclature website, http://www.chem.qmul.ac.uk/iubmb/enzyme, EC 2.5.1.54). **ASA**-L-aspartate semialdehyde, **F-1,6-P-** fructose 1,6-bisphosphate, **GAP-** glyceraldehyde 3-phosphate, **HPAP-** hydroxpyruvaldehyde phosphate, **DHQ**-3 dehydroquinate, **PEP-** phosphoenolpyruvate, **E-4-P-** erythrose-4-phosphate.

The second explanation for an aldolase activity of ORF OE1472F is the possibility that this ORF is involved as well in the formation of DKFP. White and Xu (White and Xu, 2006) demonstrated that the recombinant MJ1585 gene product from *M. jannaschii* catalyzes the condensation of methylglyoxal with a DHAP fragment (derived from F-1-P and/or F-1,6-P), to generating DKFP (table #2, experiment 5 and 6 in (White and Xu, 2006)). This observation was consist with the MJ1585 gene product also functioning as a F-1,6-P aldolase ((Porat et al., 2006) and table #15). Thus, White and Xu (White and Xu, 2006) concluded that the glycolytic enzyme MJ1585 is a multifunctional enzyme, responsible for both cleavage of F-1,6-P (figure #91 reaction #1) and the formation of DKFP from methylglyoxal and the DHAP moiety derived from F-16-P and/or F-1-P (figure #91, reaction #3). On the other hand, a transaldolase reaction between ASA and DKFP (precursors of the AroAAs biosynthesis) is mediated by the product of the MJ0400 gene (White, 2004), which showed no detectable aldolase activity ((Porat et al., 2006). In contrast to *M. jannaschii*, both of the homologs in *H. salinarum*: OE2019F (homolog to MJ1585)

164

and OE1472F (homolog to MJ0400), demonstrated similar aldolase activities (table #15). Thus one can speculate that the two aldolase reactions depicted in figure #91 might be catalyzed either by OE2019F (reaction #1) and OE1472F (reaction #3) or only by OE1472F (reaction #1 and #3). Some indirect evidence to support the possibility that ORF OE1472F can catalyze reaction #3 was given in the paper published by Morar et al (Morar et al., 2007), where MJ0400 (homolog to OE1472F) was crystallized with DHAP, resulting from soaking the crystal with F-1-P. The structure contained a clear DHAP electron density linked to Lys184 (figure #89, located on strand β6, shown to be one of the active site residues).

The methylglyoxal used for the production of DKFP can derived from specific enzymatic reactions. One source might be methylglyoxal synthase (EC 4.2.3.3), which converts DHAP to pyruvate via methylglyoxal and D-lactate during glycolysis in *E. coli* (Cooper and Anderson, 1970; Kalapos, 1999). The probability of this to be the source of methylglyoxal in haloarchaea is low due to the absence of methylglyoxal synthase in the genomes of haloarchaea. On the other hand, methylglyoxal may arise enzymatically from GAP via elimination of phosphate catalyzed by triose phosphate isomerase, using OE2500R (EC 5.3.1.1) (figure #91 reaction #2 and (Richard, 1991).
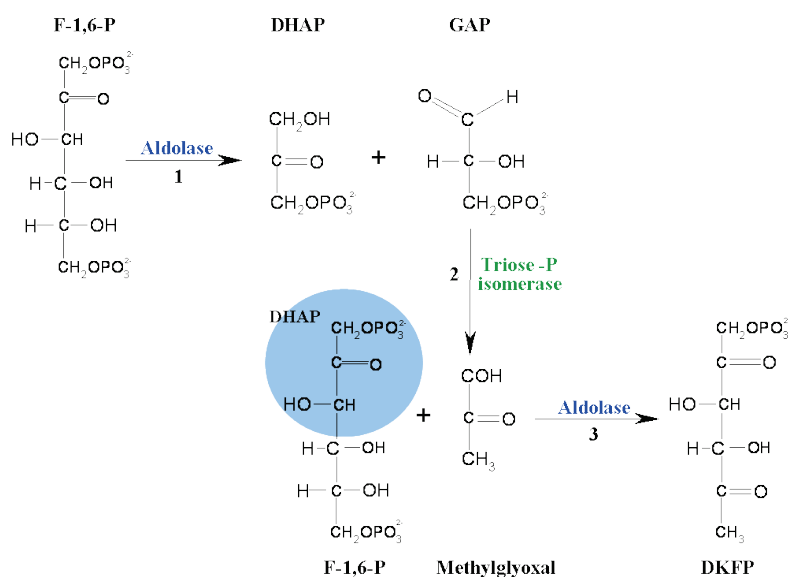
Figure # 91: Proposed pathways for the formation of Methylglyoxal and DKFP. **Reactions #1 & 2**: were suggested to be catalyzed by the multifunctional glycolytic enzyme, MJ1585 in *M. jannaschii*. **Reaction #2**: was suggested to be catalyzed by triose phosphate Isomerase EC 5.3.1.1 in *M. jannaschii* and *H. salinarum* by ORFs MJ1528 and OE2500R, respectively. Modified from White and Xu (White and Xu, 2006).

Although OE1472F showed aldolase activity, in order to initiate the AroAAs biosynthesis pathway in *H. salinarum*, this enzyme should demonstrate a transaldolase activity with ASA and DKFP as substrates. Since the standard transaldolase coupled assay could not be used, an LC-MS assay was established in which substrates and or intermediate could be derivatized, detected and identified in high salt. Although the method was validated, OE1472F showed no detectable aldolase/ transaldolase activity. This was most likely due to its low specific activity, which was shown to be below the limit of detection of this assay. Although no transaldolase activity was demonstrated for ORF OE1472F, it is most likely transaldolase, as predicted from the sequence similarity of ORF OE1472F to MJ0400 (figure #89). Nevertheless, improvements to the LC-MS methods need to be tested in order to determine experimentally whether or not the gene product of OE1472F is a transaldolase. From the data presented in table #15 for *M. jannaschii*, White (White, 2004) was able to demonstrate that the combination of MJ0400 and MJ1249 (homologs of OE1472F and OE1475F, respectively), could stimulate the production of DHQ, thus the gene products of MJ0400 and MJ1249 are the first steps needed for formation of DHQ in *M. jannaschii*. This approach need to be tested as well in H. salinarum, where OE1472F and OE1475F, will be expected to generate DHQ, using ASA and DKFP as precursors (figure #88).

7.2.1.3 Phenotype of the OE1471F::pMG501 mutant

Even though transaldolase activity was not demonstrated for the OE1472F protein, its role in the AroAAs biosynthesis pathway was determined in mutants where the corresponding gene was disturbed by insertion (figure #35). Assuming that no additional pathways exist for synthesizing AroAAs in *H. salinarum*, a mutant defective in one of the genes in the AroAAs biosynthesis pathway would be expected to be auxotrophic for one or more AroAAs.

The mutant, OE1471F::pMG501, showed reduced ability to grow without AroAAs, while addition of AroAAs allowed growth at the same level as WT, thereby confirming the role of this ORF in the AroAAs pathway. Furthermore, RT-qPCR results showed that ORF OE1472F was 12-fold induced when AroAAs were omitted from the growing medium. Mutant OE1471F::pMG501 grows to some extent without AroAAs due to either incomplete blocking of OE1472F transcription or by limited

compensation by some other gene product. In the later case, this could be by OE2019F, as it is one of the paralogous genes to class the 1A aldolase family, which was shown in this study to cleave F-1,6-P (figure #56A).

Supplementing the synthetic medium with AroAAs plus either shikimate or DHQ, was sufficient for the mutant to grow as well as WT, indicating that the OE1472F enzyme acts on a substrate upstream of DHQ and shikimate.

The ability of the mutant to take up Phe from the medium supplemented only by AroAAs, suggests that while ORF OE1475F and OE1477R can regulate the specific transporters of AroAAs, the gene product of OE1472F has little or no regulatory effect on the Phe transport, allowing cells to take up Phe from the medium.

7.2.2 ORF OE1475F- the second ORF in the AroAAs pathway

ORF OE1475F was annotated in *H. salinarum* genome as a conserved hypothetical protein. It was found to be homologous to MJ1294 from *M. jannaschii*, with which it shares 41% amino acid identity. MJ1294 was annotated as 3-dehydroquinate synthase (EC 4.6.1.3) and shown to catalyze the NAD dependent oxidative deamination and the cyclization reaction of 2-amino-3,7-dideoxy-D-threo-hept-6-ulosonic acid (NDH) to form 3-dehydroquinate (figure #3 in (White, 2004)). The role of the gene product of OE1475F in *H. salinarum* was determined experimentally.

7.2.2.1 Phenotype of the OE1475F::pMG601 mutant

The role of ORF OE1475F in the AroAAs biosynthesis pathway was determined in mutants where the target gene was disturbed by insertion (figure #39). The mutant OE1475F::pMG501 showed reduced ability to grow without AroAAs, but when AroAAs were added to the synthetic medium, cells grew as well as the WT. These results confirmed the role of this ORF in the AroAAs pathway.

The mutant was capable of growing on synthetic medium supplemented with all three AroAAs, and although the uptake of Phe was poor, as observed from uptake assays (table #6), it was sufficient to support the growth of the mutant. As shown by Gonzales et al., (Gonzalez et al., 2009) for WT *H. salinarum*, for most of the supplied amino acids, the uptake rates exceeds the rate at which they are directly incorporated. More specifically, it was proven experimentally that only 0.2mM Phe were

incorporated into biomass, allowing cell growth to an optical density of $1.8OD_{600nm}$. This represents a consumption of 0.11mM Phe OD^{-1} cell biomass. In the case of mutant OE1475F::pMG601, 0.094mM Phe OD^{-1} was consumed by the cells when grown in synthetic medium supplemented only by AroAAs (table#6). This level of Phe consumption is sufficient to support growth, as shown by Gonzalez et al., (figure #4 in (Gonzalez et al., 2009)). Interestingly, the consumption rate of Phe by the mutant was elevated in the presence of AroAAs plus either DHQ or shikimate (table #6), suggesting that the intracellular concentration of DHQ and shikimate induces the uptake of Phe by the mutant.

The ability of the mutant to grow in the presence of AroAAs plus either shikimate or DHQ, indicated that the gene product of OE1475F acts on substrate upstream (i.e. before) of DHQ or shikimate.

In contrast to OE1475F and MJ1249 (*H. salinarum* and *M. jannaschii*, respectively), the homolog from *M. maripaludis* (ORF MMP0006) was demonstrated not to be involved in AroAAs biosynthesis. Deletion of this gene did not result in a growth requirement for AroAAs, the mutant grew as well as the WT strain (figure #5B in (Porat et al., 2006)). In addition, labeling experiments showed that shikimate and dihydroshikimate (DHS) pools were labeled to nearly the same extent as in the WT. Therefore, Porat et al (Porat et al., 2006) concluded that the gene MMP0006 was not required for AroAAs biosynthesis in *M. maripaludis*.

7.2.2.2 Enzyme activity and structure

The enzyme responsible for the formation of DHQ, 3-dehydroquinate synthase (EC 4.6.1.3), is of a great interest because it catalyses five individual reactions in a single active site involving alcohol oxidation, phosphate elimination, carbonyl reduction, ring opening and intermolecular aldol condensation (Srinivasan et al., 1963; Widlanski et al., 1989). To date, 3D structures for this enzyme are available from *Aspergillus nidulans* (PDB accession code 1SG6), *Emericella nidulans* (PDB codes: 1NVB, 1NVD, 1NVE), *Thermus thermophilus* (PDB code 1UJN), and *Staphylococcus aureus* (PDB codes: 1XAG, 1XAH, 1XAI, 1XAJ). No crystal structure is available for the archaeal-type 3-dehydroquinate synthase. Submitting the sequence of OE1475F to 3D-JIGSAW (server which builds three-dimensional models for proteins based on homologues of known structure,

http://bmm.cancerresearchuk.org/~3djigsaw/) retrieved no homologous sequences in PFAM, PDB or SCOP.
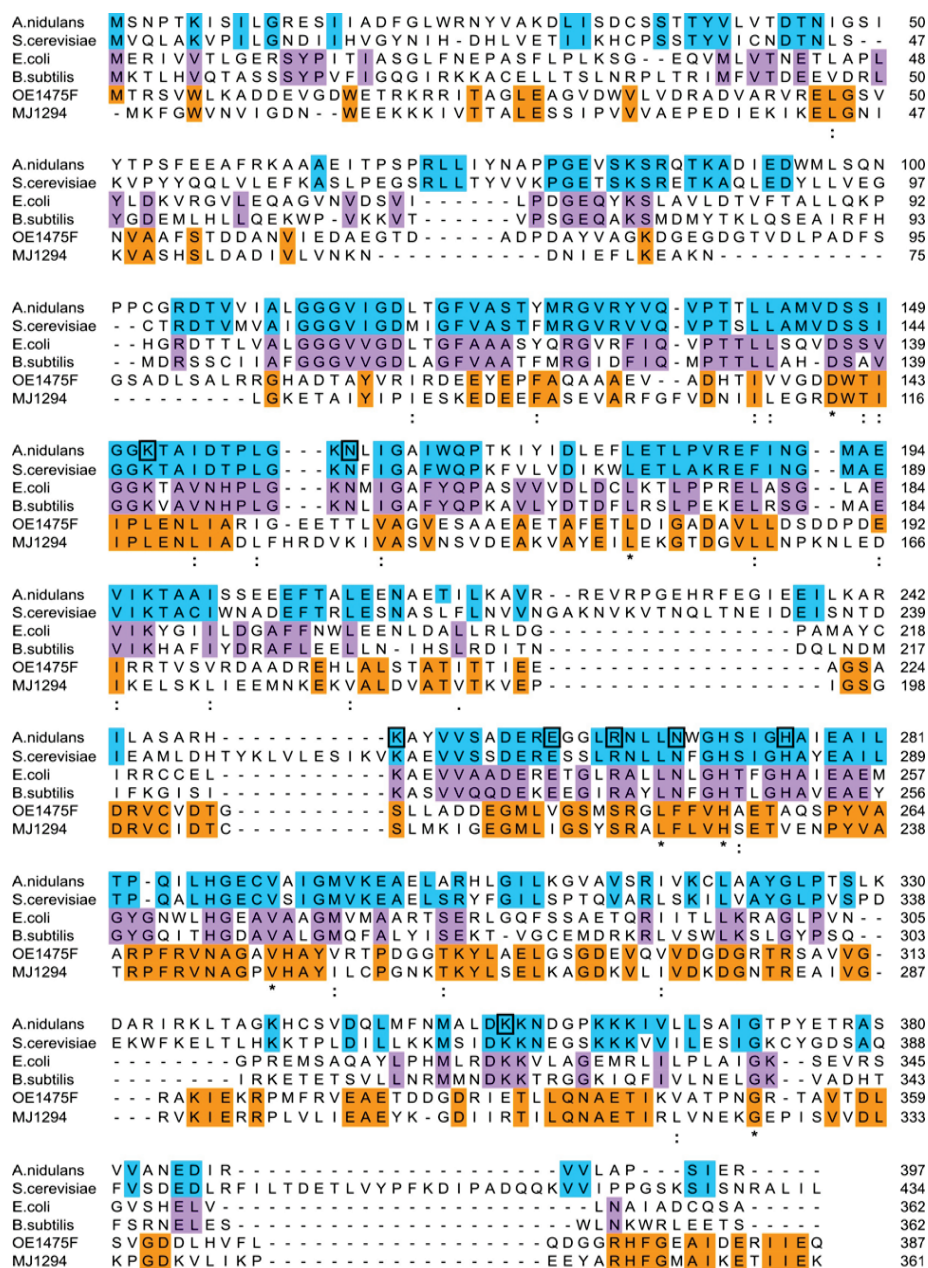
Figure # 92: Sequence comparison of 3-dehydroquinate synthases from eukaryotic, prokaryotic and archaeal origins. *Aspergillus nidulans* (UniProt accession P07547), *Saccharomyces cerevisiae* (UniProt accession PO85566), *E. coli* (GenBank accession P07639), *Bacillus subtilis* (P31102), OE1475F (*H. salinarum*, www.halolex.mpg.de), and MJ1249 (from *M. jannaschii*). *****- identical amino acid in all six sequences, **:**-similar amino acids in all six sequence. Identical amino acid in fungal, prokaryotes and archaea are marked in blue, violet and orange, respectively. The catalytic residues identified in *A. nidulans* are marked with black boxes (Nichols et al., 2003). The fungi sequences are the N-terminal subunit of the representative multi-enzyme complexes. Catalytic residues: Asn162, Asn268, His275, Arg264, Arg130, Lys152, Lys356, Lys250 and Glu260. The alignment was done using ClustalW multiple sequence alignment program (http://us.expasy.org).

As seen in figure #92, OE1472F showed low overall identity with homologs of *A. nidulans* and *S. ceervisiae* (11%), and with *E. coli* and *B. subtilis* (12%), while 41% sequence identity was found with MJ0400 from *M. jannaschii*. But ORF OE1475F shows 75-65% sequence identity to homologs found in other haloarchaea such as *N. pharaonis*, *Haloarcula marismortui* and *Haloquadratum walsbyi* (www.halolex.mpg.de). Nevertheless, some regions are highly conserved among all sequences. Downstream to the first conserved region there is an insertion in the sequences of *E. coli* , *B. subtilis* , *H. salinarum* and *M. jannaschii* of different length. This insertion is followed by another conserved region.

9 catalytic residues are responsible for the reactions leading to the formation of DHQ, these were identified in the crystal structure of 3-dehydroquinate synthase from *A. nidulans* (Carpenter et al., 1998; Nichols et al., 2003). As seen in figure #92 (in black boxes), all the catalytic residues are completely conserved among the fungi and prokaryotic sequence, but not in the archaeal sequences (figure #92), which is not surprising due to the high number of Lys residues involved. Out of 9 residues 3 are Lys (Lys152, Lys356 and Lys250), which are involved in more then one reaction: Lys152 is involved in reactions 2, 4 and 5, Lys250 is involved in reactions 4 and 5, while Lys356 is involved in the second reaction. As shown earlier (Madern et al., 2000; Madern et al., 1995), the primary sequence of halophilic proteins is characterized by an increased level of acidic residues along with a drastic reduction of lysine residues, compared to their mesophilic contraparts.

7.2.3 ORF OE1477R- the third ORF in the AroAAs pathway

ORF OE1477R (*aro*D) was annotated as 3-dehydroquinate dehydratase (EC 4.2.1.10), and by definition was expected to remove a water molecule from 3-dehydroquinate (DHQ) to form 3-dehydroshikimate (DHS), as shown in figure #88. It was predicted to catalyse the 3rd step in the de novo biosynthesis pathway of AroAAs in *H. salinarum,* so a deletion of this ORF should result in AroAAs auxotrophy. It is also possible that expression of this gene might be regulated by the exogenous levels of AroAAs.

From the RT-qPCR results shown in figure #88, no significant regulation was observed for ORF OE1477R, whether AroAAs were present or not in the synthetic medium, suggesting that this gene might be constitutively expressed, as opposed to a

genetic switch, which assures that the expression of a specific gene is turned on when a signal, (e.g., a nutrient), is present and turned off when the signal is absent.

7.2.3.1 Phenotype of the deletion mutant ΔOE1477R

In-frame deletion of ORF OE1477R was achieved by recombination, and the mutant was recovered by aerobic growth in the dark in synthetic medium supplemented with 1.1mM AroAAs and 1.1mM DHQ. The genotype of ΔOE1477R was confirmed by PCR and Southern blot analysis.

The phenotype of ΔOE1477R was studied by cultivation in synthetic medium with and without AroAAs. The inability of the mutant to grow without AroAAs confirmed the role of OE1477R in the biosynthesis of AroAAs.

The deletion strain was capable of growing slowly on synthetic medium supplemented with AroAAs, by utilizing only low concentrations of Phe sufficient for supporting this low growth (figure #31 and table #3). Nevertheless, when the mutant grew on synthetic medium supplemented with AroAAs plus either DHQ or shikimate the mutant was capable of utilizing all the Phe supplied in the medium. This observation can be interpreted as evidence that the internal levels of DHQ and shikimate regulates the uptake of Phe from the medium. This allows the incorporation of AroAAs into proteins or as free metabolites, as well as serving as carbon skeleton donors.

7.2.3.2 Enzyme activity of OE1477R

As expected, the purified gene product of ORF OE1477R was shown to possess dehydroquinate dehydratase activity, i.e. formation of DHS from DHQ. The specific activity of the purified gene product of OE1477R from *H. salinarum* was 55±6mU/mg protein. In comparison, the specific activity of dehydroquinate dehydratase found in cell extracts of *M. maripaludis* was 6.0±0.2mU/mg protein (Porat et al., 2004). No detectable activity was found when *aro*D was deleted (<0.5mU/mg protein) (Porat et al., 2004). A purified enzyme from *E. coli* displayed a similar specific activity (89mu/mg protein) compared to OE1477R from *H. salinarum* (Mitsuhashi and Davis, 1954).

```
                                 10.           20.  +  +  + 31.
Fulg ---------- -----MKLVA TLS------S PEELELAE-- --KADVVELR IDLFDFSG--
Halo ---------- -------MDL AAD------- PLAALDDYDG --VLPVLAIN RADWEGGA--
Pyro ---------- -------MPK LAGVIIAKTI EEAVKKIKLH --EADLYELR VDYLESMEN-
Meth ---------- -------MIC LPVVEDSVEK AIKTAEKYLE --IADIVEHR IDMLKEVS--
Solf ---------- ------MRPL IVASLPIKKI EDLKLIENFL --DADLIELR LDYLREREVS
Aqui ---------- --------ML IAVPLDDTNF SENLKKAKEK --GADIVELR VDQFSDTSLN
Coli MKTVTVKDLV IGTGAPKIIV SLMAKDIASV KSEALAYREA --DFDILEWR VDHYADLSNV
Salm MKTVTVKNLI IGEGMPKIIV SLMGRDINSV KAEALAYREA --TFDILEWR VDHFMDIAST
Staf ---------- --MTHVEVVA TITPQLSI-E ETLIQKINHR IDAIDVLEIR IDQFENV-TV
                         s1====s   h1=========h      s2====s      h2

                         40.  +  +  **       60.        70.        80.
Fulg ---------- --ARVDKEKI LTRRVSDGG KFEGDERERI EKMKRAFDSL NPDYVDLESD
Halo ---------- ----------AADGG DRIDALAEAA RTDCVAAVDI ERSALVDDDT
Pyro --------IE LLLPHSNKLI VTMRKVEEGG AKRIEEQERL KLFEKFAKIK P-AYFDVELY
Meth -------EED IEKFAKYPCI ITVRADWEGG YWKGNNEERL NLIKKAIECN -AKFVDIELR
Solf ---LISDYYE FLDKYKKKLI VTRDKGEGG INQLADELKI KILNELYERQ YLYDIEVSFL
Aqui ---YVKEKLE EVHSQGLKTI LTIRSPEEGG REVKNREELF EELSPLSDYT ---DIELSSR
Coli ESVMAAAKIL RETMPEKPLL FTHRSAKEGG EQAISTEAYI ALNRAAIDSG LVDMIDLELF
Salm QSVLTAARVI RDAMPDIPLL FTHRSAKEGG EQTITTQHYL TLNRAAIDSG LVDMIDLELF
Staf DQVAEMITKL KVMQDSFKLL VTMRTKLQGG YGQFTNDSYL NLISDLANIN GIDMIDIEWQ
     ==============h   s3===s          h3==============h   s4====s

                         90.         *  *  *     **  110.          120. *  +
Fulg --LPDSAFDF --------NC RIIESMHNFI RTPDYSELKG IVEGR---RG DLVKIATMGK
Halo --ADGAEALA AAR---STDT TTVVSAHIFD GTPSLSAMAD LLGEACSL-G DVGKIAVTPQ
Pyro --SNIVEDVI N--LSSKIGS KVILSMHNFS KTPSYQTLQS ILDSMIELEP DIIKIVTYAR
Meth --EEKNKELV KFRDEIGSKT KIIISMHFE KTPSKEKLVE IVEKALSI-G DIAKFATMAN
Solf --QKYDIPYD N--------- -RIVSMHVFN YLPTLEKIKE IVSKFSEK-A FSVKIAVPSL
Aqui --GLLVKLY- --NITKEAGK KLIISMHNFE IIPPNWIIRE VLREGYRY-G GIPKIAVKAN
Coli --TGDDQVKE TVAYAHAHDV KVVMSNHDFH KTPSAEEIIA RLRKMQSFDA DIPKIALMPQ
Salm --TGDADVKA TVDYAHAHNV YVVMSNHDFH QTPSAEEMVS RLRKMQALGA DIPKIAVMPQ
Staf ADIDIEKHQR IITHLQQYNK EVIISHHNFE STPPLDELQF IFFKMQKFNP EYVKIAVMPH
     h4=============h    s5=====s       h5==============h    s6====s

            130. +     140.            150. *        * 164. ++        177. +
Fulg -SKRDVETIVR ILTNY----- -DDVVAFIMG ERFSFTRVLA AYLGSPFIYC YVG--SPKA
Halo -DRGDALDVIR VTHEYSAAG- -MTVATMGMG DLGRHTRAVT PLYGSKLGYA PVADGETTA
Pyro -DFRDNIEVIK LYENSENL-- ----IAFCMG EKGKISRIFS ALYSPFTYVS IDN---SVA
Meth -SKEDVLNILE VINKYPGK-- ---IIGIGMG EKGKLTRILG VYFGSILTFA SYKG-KSSA
Solf KGYKEVLLPLL EYENVTVIP- --------MS NNS-LERIAV GLLGSKLVYS YAI--EPIA
Aqui -SYEDVARLLC ISRQVEGE-- ---KILISMG DYGKISRLAG YVFGSVITYC SLE--KAFA
Coli -STSDVLTLLA ATLEMQEQYA DRPIITMSMA KTGVISRLAG EVFGSAATFG AVK--KASA
Salm -SKHDVLTLLT ATLEMQQHYA DRPVITMSMA KEGVISRLAG EVFGSAATFG AVK--QASA
Staf -NKNDVLNLLQ AMSTFSDT-M DCKVVGISMS KLGLISRTAQ GVFGALTYG CIG--EPQA
     h6================h   s7==s    h7==========h s8==s

     +**  187.
Fulg PGQISLDDARE IISRLG- -------- 196
Halo PGQYAPAALQA LIADLQ- -------- 195
Pyro PGQFTVGEMRK ILELLGE TK------ 217
Meth PGQVDIDTLKE IWRLMDL K------- 220
Solf QGQLYYKKVIQ IFNYIND ITTSSLVT 220
Aqui PGQIPLEEMVE LRKKFYR L------- 219
Coli PGQISVNDLRT VLTILHQ A------- 252
Salm PGQIAVNDLRS VLMILHN A------- 252
Staf PGQIDVTDLKA QVTLY-- -------- 238
     h8============h
```

Figure # 93: Multiple alignment of archaeal and bacterial dehydroquinase sequences, modified from (Smith and Gallagher, 2008). Five archaeal sequences are at the top, followed by four bacterial sequences. The organism names are abbreviated as follows: **Fulg**-*Archaeoglobus fulgidus*, **Halo**- *Halobacterium* sp. NRC-1, **Pyro**- *Pyrococcus furiosus*, **Meth**-*Methanocaldococcus jannaschii*, **Solf**- *Sulfolobus solfataricus*, **Aqui**- *Aquifex aeolicus*, **Coli**-*Escherichia coli*, **Salm**- *Salmonella typhimurium*, and **Staf**- *Staphylococcus aureus*. The three known structures are Fulg (at the top), Salm and Staf (the last two). Residue numbers along the top are for the *A. fulgidus* sequence. Residues that are completely conserved among the nine sequences are shown in bold and marked both with an asterisk and with a blue box, while those that are conserved with one or two exceptions are marked both with a plus sign and red box. The secondary structure below the alignment, giving numbers for the eight strands and helices of the barrel, corresponds to the Salmonella enzyme and corresponds approximately with the others. For the three known structures (Fulg, Salm and Staf), the alignment was structure-based, while for the other sequences it was calculated by ClustalW based on sequence homology. The active site residues are marked in green boxes, corresponding to Glu23, Arg45, His98, Lys122, and Arg159 in *A. fulgidus*. Non-conserved active site residues are marked in black box (Ser22, and Ser232 in *S. typhimurium*).

7.2.3.3 OE1477R- a dehydroquinate dehydratase type I enzyme

Additional confrmation that ORF OE1477R is indeed dehydroquinate dehydratase is given in figure #93. The alignment of type I 3-dehydroquinate dehydratase from the thermophilic archaeon *Archaeoglobus fulgidus* with multiple bacterial homologs, demonstrates that the active site is generally similar in its key residues and their geometry. Out of the seven residues within the product-complex structure of the *S. typhimurium* enzyme, five are well conserved in the alignment, and are expected to play a similar role in the *H. salinarum* enzyme, (figure #93, marked with green boxes). The greatest difference is seen in two of the active site residues in *S. typhimurium* (Ser21, and Ser232), which are nonconserved, (figure #93 marked with black boxes) (Smith and Gallagher, 2008). Additionally, one can see that many of the helical regions, especially helices 2 and 4, appear to be absent or significantly shorter in the archaeal protein (figure #93 and (Smith and Gallagher, 2008)).

7.2.4 A non- canonical pathway for AroAAs biosynthesis in *H. salinarum*

The putative functions of ORFs OE1472F, OE1475F and OE1477R were confirmed experimentally, both in vivo and in vitro. All three ORFs were found to be part of a non-canonical pathway for AroAAs biosynthesis in *H. salinarum*, and to act on substrates upstream to DHS, as demonstrated from the phenotype of mutant strains defective in these ORFs. The set of reactions leading to the formation of DHQ is represented in figure #94.

In summary, ORF OE1472F was classified as a member of the archaeal aldolase/transaldolase family based on the known 3D structure of transaldolase from *M. jannaschii*, and that it displays an aldolase activity towards F-1,6-P as substrate. Although, transaldolase activity was not shown experimentally, it is predicted from the data presented that ORF OE1472F is the first gene in the de novo pathway in *H. salinarum*. Moreover, the unexpected aldolase activity of this enzyme may indicate involvement in the formation of DKFP.

In addition, ORF OE1475F was found to play a role in the biosynthesis of AroAAs in *H. salinarum* as the mutant required both AroAAs plus either shikimate or DHQ for growth. This conserved protein is believed to be a 3-dehydroquinate synthase, and

as such, responsible for the oxidative deamination and the cyclization reactions, to form DHQ. Up to date, no crystal structure of the archaeal-type enzyme is available.

Finally, evidence was provided to show that the gene product of OE1477R catalyses the third step in the de novo pathway of AroAAs biosynthesis in *H. salinarum*. The enzyme was classified as 3-dehydroquinate dehydratase, and the purified gene product was capable of forming DHS from DHQ.

Based on the microarray analysis and RT-qPCR results, it was shown that several regulatory mechanisms are used when *H. salinarum* utilize AroAAs via a non-canonical biosynthesis pathway. In summary; (1) all 13 homologous genes downstream to DHQ were strongly regulated when AroAAs were omitted from growth medium of R1; (2) upstream to DHQ, while ORF OE1472F was strongly regulated in the absence of AroAAs, ORFs OE1475F and OE1477R were expressed constitutively; (3) the involvement of specific and unspecific transporters were shown, including *aro*P, various Na-dependent transporters and ABC transporters, which would be consistent with the cells trying to scavenge AroAAs from the medium; (4) global regulators such as TrmB and Lrp-like proteins were found to be regulate. While TrmB was weakly induced when AroAAs were added to the growing medium of R1, LrpA2 was strongly induced when AroAAs were absent in the growing medium, and LrpA1 was strongly up regulated when R1 was grown in synthetic medium supplemented with AroAAs and shikimate.

174
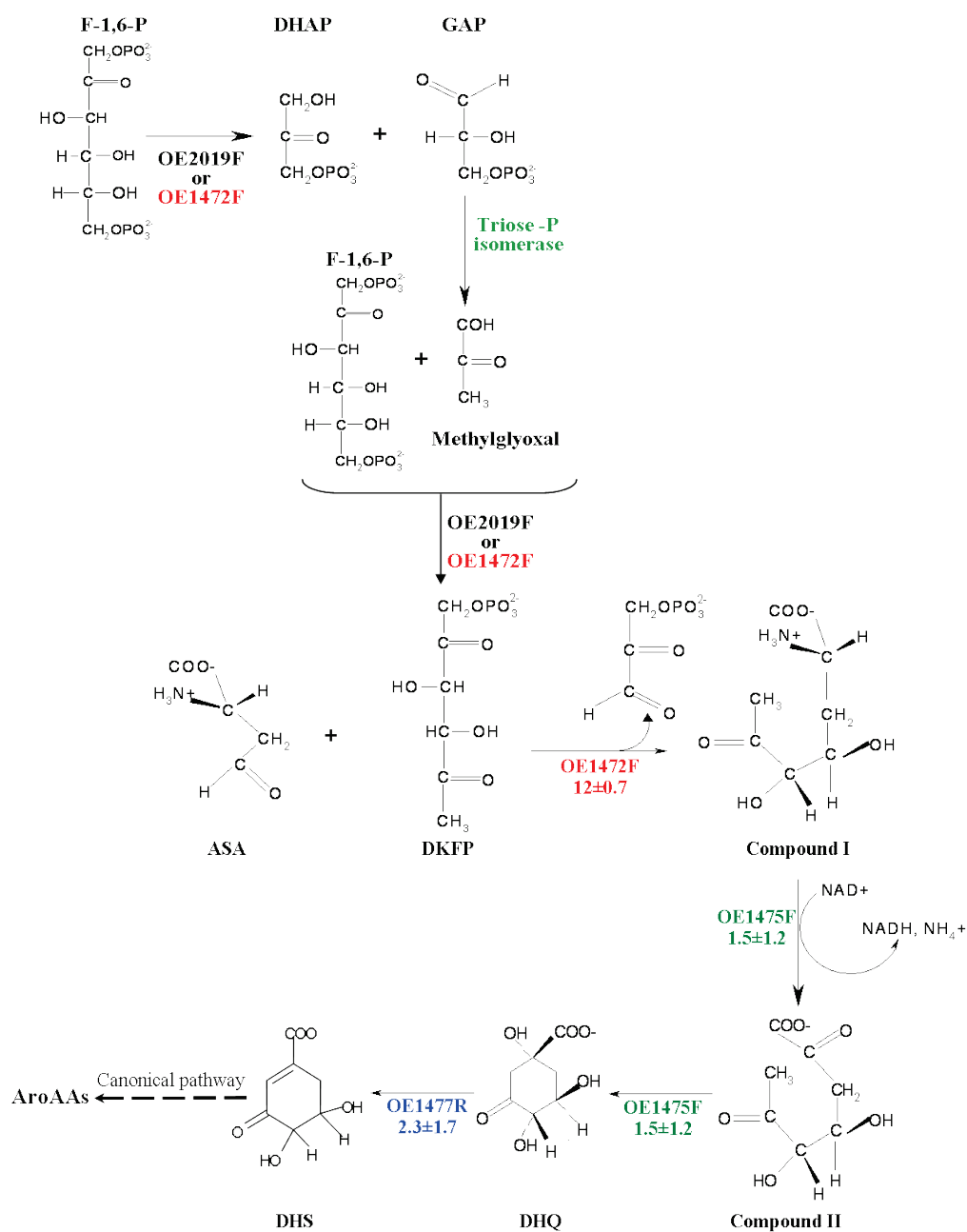
Figure # 94: A non-canonical pathway for AroAAs biosynthesis in *H. salinarum*. The ORF catalyzing the specific steps are indicated below the arrows. The canonical pathway takes place downstream to DHQ.

8 Appendix

8.1 Derivatized masses identified by LC-MS

The derivatized masses of the pure substrates identified by GC-MS with the indicated retention times. DHS was the only compound in the table which was not injected as a pure substance. DHS was detected (with the indicated mass and retention time), in reaction with protein OE1477R and DHQ.

Table # 16 : The mass detected by LC-MS after derivatization with NBHA

	Formula	Exact Mass (g/mole)	Exact Mass after derivatization [1] (g/mole)	Mass detected by LC-MS [2] (g/mole)	RT (min)
ASA	$C_4H_7O_3N$	117.10	267.24	268.09	12.5
DKFP	$C_6H_{11}O_8P$	242.12	542.39	543.11 [3]	28.2
F-1,6-P	$C_6H_{14}O_{12}P_2$	340.11	490.25	491.04	3.7
DHAP, GAP	$C_3H_7O_6P$	170.05	320.19	321.05	14.5, 15.2
DHQ	$C_7H_{10}O_6$	190.15	340.29	341.09	14.8
DHS [5]	$C_7H_8O_5$	172.13	322.27	323.08 [4]	20.1
NBHA	$C_7H_8N_2O_3$	168.15	--------	169.06	5.6

[1]-Calculated exact mass.

[2]-Calculated exact mass after derivatization and release of water molecules.

[3]-The exact mass as detected by LC-MS.

[4]- DKFP was derivatized twice and released of two molecules of water.

[5]- DHS was not used as a standard, but was identified by mass after the cleavage of DHQ by OE1477R.

8.2 Derivatized masses identified by GC-MS

The pure substrates ASA, DKFP, F-1,6-P, GAP, DHAP, DHA, DHQ and Shikimate were derivatized by methoximation followed by TMS (table#6) or TBS (table #7). The derivatized masses, and the retention times are listed in the tables below.

Table # 17: The mass detected by GC-MS after derivatization with methoximation and TMS.

	Formula	Exact Mass (g/mole)	Exact Mass after derivatization [1] (g/mole)	Mass detected by GC-MS [2] (g/mole)	RT (min)
ASA	$C_4H_7O_3N$	117.042	290.148	nd	
DKFP	$C_6H_9O_8P$	240.003	588.230	nd	
F-1,6-P	$C_6H_{10}O_{12}P_2$	335.964	873.299	nd	
GAP	$C_3H_5O_6P$	167.982	415.143	415.1431	12.46, 12.54
DHAP	$C_3H_5O_6P$	167.982	415.143	nd	
DHA	$C_3H_6O_3$	90.031	263.137		8.67
DHQ	$C_7H_9O_6$	189.039	507.232	507	13.71
DHS	$C_7H_7O_5$	171.1306	417.1823	-------	------
Shikimate	$C_7H_9O_5$	173.045	462.210	462	13.17

[1]-Calculated exact mass after derivatization by methoximation and TMS.

[2]-The exact mass as detected by GC-MS.

Table # 18 : The mass detected by GC-MS after derivatization by methoximation and TBS.

	Formula	Exact Mass (g/mole)	Exact Mass after derivatization [1] (g/mole)	Mass detected by GC-MS [2] (g/mole)	RT (min)
ASA	$C_4H_7O_3N$	117.042	374.242	317.242, 289.242[3]	24.86
DKFP	$C_6H_9O_8P$	240.003	756.418	699.418	43.38
F-1,6-P	$C_6H_{10}O_{12}P_2$	335.964	1167.627	nd mass to high	
GAP	$C_3H_5O_6P$	167.982	541.284	456.284	30.48
DHAP	$C_3H_5O_6P$	167.982	541.284	484.284	30.72, 30.95
DHA	$C_3H_6O_3$	90.031	374.231	290.231	20.87
DHQ	$C_7H_9O_6$	189.039	675.420	675.420, 618.420[3]	36.90-36.94
DHS	$C_7H_7O_5$	171.1306	543.323	486.323, 458.323	36.91
Shikimate	$C_7H_9O_5$	173.045	630.398	630.399, 573.399[3]	34.52

[1]-Calculated exact mass after derivatization by methoximation and TBS

[2]-The exact mass as detected by GC-MS.

[3] Variation from the calculated mass are due to cleavage of 57Da [$C(CH_3)_3$], or 85Da [$C(CH_3)_3CO$].

8.3 Oligonucleotides

Table # 19 : List of oligonucleotides used in this study

Primer name	Sequence	Tm
P1a of pMG550.for	GGCCGG_ **AGCTT**ATGACTGACGCCGGGAAAACC GCACGACTC (HindIII)	85.0
P2a of pMG550.rev	CCGGCC_**GGATCC**_C_GGCGTCGACGACGAGACC GGCCCGG (BamHI)	91.0
P1a of pMG650.for	GGCCGG_ **AAGCTT** ATGACACGGTCCGTTTGGCTG AAAGCTGACG (HindIII)	84.0
P2a of pMG650.rev	CCGGCC_**GGATCC**_C_CTGCTCGATGATGCGCTCG TCGATGGCCTCC (BamHI)	89.0
P1a of pMG750.for	GGCCGG_ **AAGCTT**ATGGAGTTCCAGGATTTTCTGC TCGCCGCG (HindIII)	84.0
P2a of pMG750.rev	CCGGCC_**GGATCC** _C _CTGTAGGTCGGCGATCAGC GCCTGTAGC (BamHI)	88.0
pMG560		
pET-OE1472F.For	GGCCGG_**CATATG**ATGACTGA CGCCGGGAAAACC GCACGAC TC (NdeI)	85.0
pET-OE1472F.rev	CCGGCC_**CTCGAG**GGCGTCGACGACGAGACCGGC CCGG (XhoI)	90.0
pMG660		
pET-OE1475F.For	GGCCGG_**CATATG**ATGACACG GTCCGTTTGG CTG AAAGCTG ACG (NdeI)	84.0
pET-OE1475F.rev	CCGGCC_**CTCGAG**CTGCTCGATGATGCGCTCGTCG ATGGCCTCC (XhoI)	88.0
pMG760		
pET-OE1477R.For	GGCCGG_**CATATG**ATGGAGTTCCAGGATTTTCTGC TCGCCG CG (NdeI)	84.0
pET-OE1477R.rev	CCGGCC_**CTCGAG**CTGTAGGTCGGCGATCAGCGC CTGTAGC (XhoI)	87.0
pMG860		
pET-OE2019F.For	GGCCGG_**CATATG**ATGCGACCCTTCGAGGACTCAC CGATCT CCCGC (NdeI)	87.0
pET-OE2019F.rev	CCGGCC_**CTCGAG**GATGCGTCCGAGCGCCTCGTCG ACGCTGG (XhoI)	90.0
pMG500		
P1 of OE1472F.For	GGCCGG _ **GGATCC**CGGGCGTCTCCGGGTTCGTGG	81.9

	(BamHI)	
P2 of OE1472F.rev	CCGGCC _ **CTGCAG**TGGTACTGGCTAGCAAACGTTGGG (PstI)	77.4
P3 of OE1472F.For	GGCCGG _ **TGCAG**CGCTCGGCGTCAGTCACCTGCC (PstI)	81.5
P4 of OE1472F.rev	CCGGCC _ **AAGCTT**CAGTCGACGCCGGCCTCCAGG (HindIII)	79.4
pMG600		
P1 of OE1475F.For	GGCCGG _ **GGATCC**CCCGGCGCGATGACCGCCGCC (BamHI)	84.4
P2 of OE1475F.rev	CCGGCC _ **TGCAG**TGCCGGATGGAGGCGGCACGCG (PstI)	82.7
P3 of OE1475F.For	GGCCGG _ **TGCAG**CCGCGG TCAGTCCAGCTCCCCG (PstI)	82.7
P4 of OE1475F.rev	CCGGCC _ **AGCTT**GGCGGCGCTACAGGCGCTGATC (HindIII)	79.1
pMG700		
P1 of OE1477R.For	GGCCGG_**GGATCC**TCGAGGGCGT CGCCACGTCC T G (BamHI)	81.5
P2 of OE1477R.rev	CCGGCC_**CTGCAG**ACCCGGCTCACACGCGGCCAG (PstI)	81.9
P3 of OE1477R.For	GGCCGG _**CTGCAG**GTGGCCGCG TTCGGTGGGT TTAC (PstI)	80.0
P4 of OE1477R.rev	CCGGCC_**AAGCTT**CGCCACTTCGGGGAGGCCATCG (HindIII)	79.1
pMG501		
Terminator (XbaI). for	ctagaCCGCGCGGTATTTTCTCCGAATCGGGGTTTTTT TCGTGTGTAGCTGGCActgca	
Terminator (PstI). rev	gTGCCAGCTACACACGAAAAAAACCCCGATTCGGA GAAAATACCGCGCggt	
P1of trpA.For	GGCCGG_**GGATCC**CGGCGTTCGTCTCCGCGGCGGC CGAGGCGGGC (BamHI)	95.0
P2of trpA.rev	CCGGCC_**TCTAGA**CTATTTGTGTTCCGGTTCCGGC ACGC (XbaI)	82.0
pMG601		
P7 of mid OE1475F.For	GGCCGG_**GGATCC**GACGGCGAAGGCGACGGCACC G(BamHI)	87.0

P8 of mid OE1475F.rev	CCGGCC_**GCATGC**CGACTGGGCGGTTTCGGCGTGC (SphI)	86.0
Primers for RT-PCR		
RT- L10ER.for	GTACCGGAAGATCGACAAGC	60.0
RT- L10ER.rev	GGAACAAGGCTGATCTGTACG	61.0
RT- OE1472F.for	ATCAGCACCGACGACAGAC	59.0
RT- OE1472F.rev	CGTTGAGGTGGACGATGTAG	60.0
RT- OE1475F.for	GAGGAAACCACGCTCGTC	58.0
RT- OE1475F.rev	CTCCTCGATGGTGGTGATG	59.0
RT- OE1477R.for	ACGACTACGACGGAGTGTTG	60.0
RT- OE1477R.rev	GTTGTGGTGTCTGTCGAACG	60.0

8.4 Volcano plot
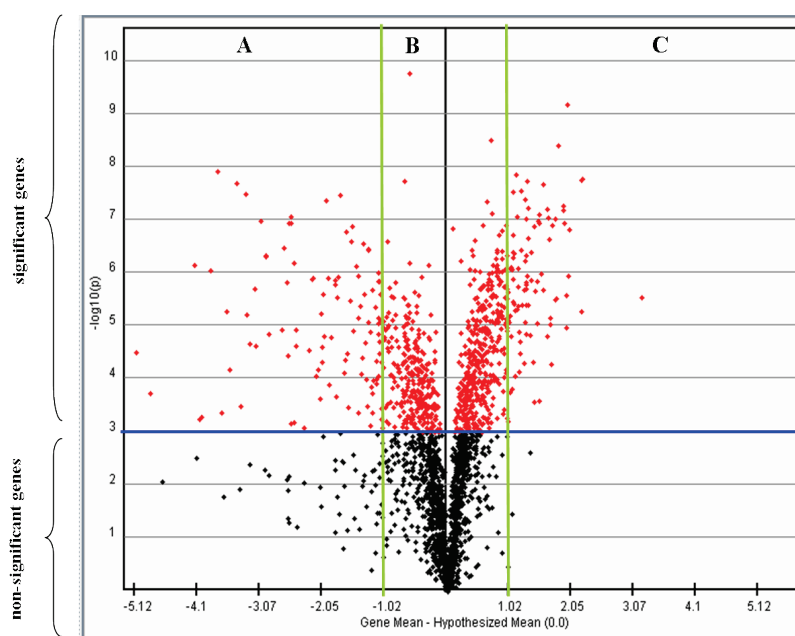
<u>Figure # 95</u>: Volcano plot generated by TIGR MEV. The data used for the plot was cDNA from cells grown in medium with AroAAs (at OD600nm=0.2) compared to the cDNA prepared from cells grown in medium without AroAAs. The plot can be divided into non-statistically significant genes (p≥0.001), and statistically significant genes (p≤0.001), below and above the blue line, respectively. The statistically significant genes can be divided into three groups: **A**-genes with gene mean ≤-1.02 (equals ≤0.49 fold regulation), **B**- genes with gene mean between -1.02 to 1.02 (fold regulation of 0.49-2.03), and **C**-genes with gene mean ≥1.02 (fold regulation ≥2.03).

8.5 Time constants

Table # 20: Time constants when electroporating *N. pharaonis* and *H. salinarum* (R1) using: 25µF, 1.5Kv, 0.2cm cuvette and different resistance values.

Resistance (Ω)	Calculated time constant (ms)	Displayed time constant (ms)	
		N. pharaonis	*H. salinarum* (TOM)
100	2.5	1.7	1.9
200	5	2.5	3.2
400	10	3	4.3
600	15	3.4	4
800	20	3.5	4.3
1000	25	3.6	5.2

8.6 Growth of OE1471F::pMG501 & OE1475F::pMG601 mutants on plates

The phenotypes of OE1471F::pMG501 and OE1475F::pMG601 mutants were compared to a WT and R1 transformed with pST1-pReport plasmid (constructed by Katarina Furtwangler) as controls for WT cells transformed with plasmids containing the Mev^R cassette.

The mutants were grown in synthetic medium with 1.1mM AroAAs, 1.1mM shikimate and 10µg/ml Mev. R1 was grown on the same medium without Mev. All cells were then plated on synthetic medium with or without AroAAs. As seen in figure #91, R1 did not grow on either of the plates due to the presence of Mev. Both of the mutants were able to grow on all media, although growth was slower on synthetic medium without AroAAs. The control, R1 transformed with pST1-pReport plasmid grow better on synthetic medium without AroAAs compared to the mutants.

As a control for the viability of the cells and to demonstrate the cells can grow on those plates, cells grown in liquid synthetic medium supplemented with AroAAs and shikimate were plated as well.
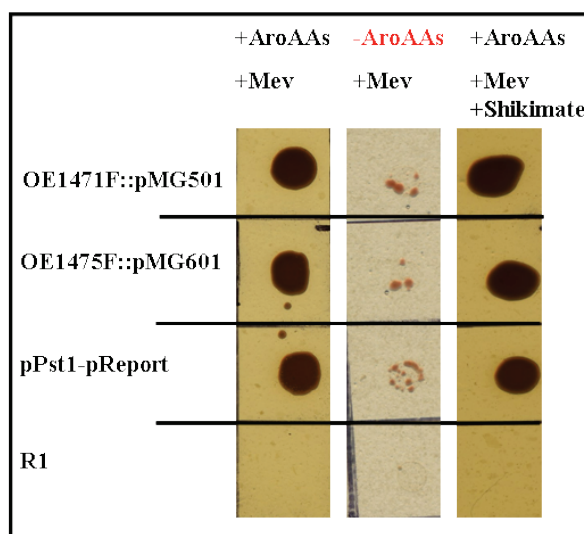
Figure # 96: Phenotype of OE1471F::pMG501 and OE1475F::pMG601 compared to R1 transformed with plasmid pPST1-pReport (plasmid containing Mev^R) and WT strain R1. Cells were grown in liquid synthetic medium supplemented with 1.1mM AroAAs and 1.1mM shikimate (40°C, 100rpm). Cells were washed with basal salt before diluting $1/10^{-2}$ with basal salt and spotting 3µl from each dilution.

8.7 Generation times and the uptake rates of Phe in WT and different mutants.

8.7.1 Growth of the WT and ΔOE1477R

Table # 21: Generation times (Gn) and the uptake rates (Y) of Phe in cultures of *H. salinarum* and ΔOE1477R mutant.

Synthetic medium supplemented with	Gn (h)		Phe uptake rates ($\mu MxOD^{-1}xh^{-1}$)	
	R1	ΔOE1477R	R1	ΔOE1477R
AroAAs	9.53±1.7	17.38±6.7	9.2	0
Without AroAAs	23.39±7.3	------		
AroAAs+Shikimate	11.17±2.4	9.09±0.7	8.7	8.47
AroAAs+DHQ	11.35±3.7	9.42±2.6	12.5	8.19

8.7.2 Growth of the WT and OE1471F::pMG501 mutant

Table # 22: Generation times (Gn) and the uptake rates (Y) of Phe in cultures of *H. salinarum* and OE1471F::pMG501 mutant.

Synthetic medium supplemented with	Gn (h)		Phe uptake rates ($\mu M \times OD^{-1} \times h^{-1}$)	
	R1	OE1471F::pMG501	R1	OE1471F::pMG501
AroAAs	10.4±2.6	10.6±1.6	10.8	7.5
Without AroAAs	**31.5±3.5**	**60.13±4.8**		
AroAAs+Shikimate	11.5±0.5	12.18±2.5	9.6	5.2
AroAAs+DHQ	11.3±2.3	11.14±1.2	13.8	11.8

8.7.3 Growth of the WT and OE1475F::pMG601 mutant

Table # 23: Generation times (Gn) and the uptake rates (Y) of Phe in cultures of *H. salinarum* and OE1475F::pMG601 mutant.

Synthetic medium supplemented with	Gn (h)		Phe uptake rates ($\mu M \times OD^{-1} \times h^{-1}$)	
	R1	OE1475F::pMG601	R1	OE1475F::pMG601
AroAAs	10.5±1.9	15.6±2.8	10.8	1.3
Without AroAAs	**30.6±4.7**	**63.3±10**		
AroAAs+Shikimate	11.1±1.8	11.6±1.9	9.1	12.2
AroAAs+DHQ	11.3±2.3	11.7±0.7	11.9	16.5

8.8 Size exclusion chromatography of purified OE1472F

The enzyme was purified from 1 liter culture and loaded on NiNTa column. The bound protein eluted after a step gradient between 20-250mM Imidazole. The protein elution buffer (50mM NaH_2PO_4 + 300mM NaCl+250mM Imidazole pH=8.0) was

exchange to 1M KCl+100mM Tris HCl (pH 7.2)+5mM MgCl$_2$ by dialysis o.n at 4°C. 100μl was loaded on S-200. The protein was eluted using the same buffer.

Elution Volume (ml)	%	# of monomers
1.43ml	43%	x 4
1.64ml	30%	x 2
1.80ml	27%	x 0.5
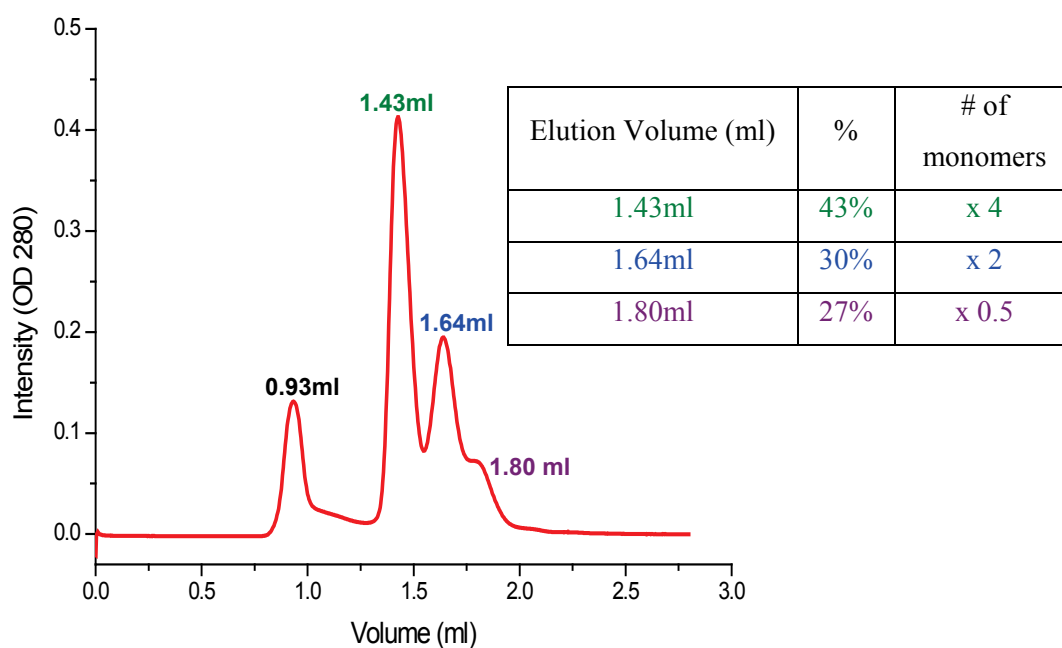
~4mg/ml protein was unfolded with 8M Urea using dialysis and folding in 2M KCl+20mM Tris-HCl (pH 7.2)+5mM MgCl2. 100μl was loaded on S-200. The protein was eluted using the same buffer, forming 14% of octamers.

Elution Volume (ml)	%	# of monomers
1.26ml	14%	x 8
1.55ml	18%	x 2
1.71ml	55%	x 1
2.05 ml	13%	x 0.5 or smaller
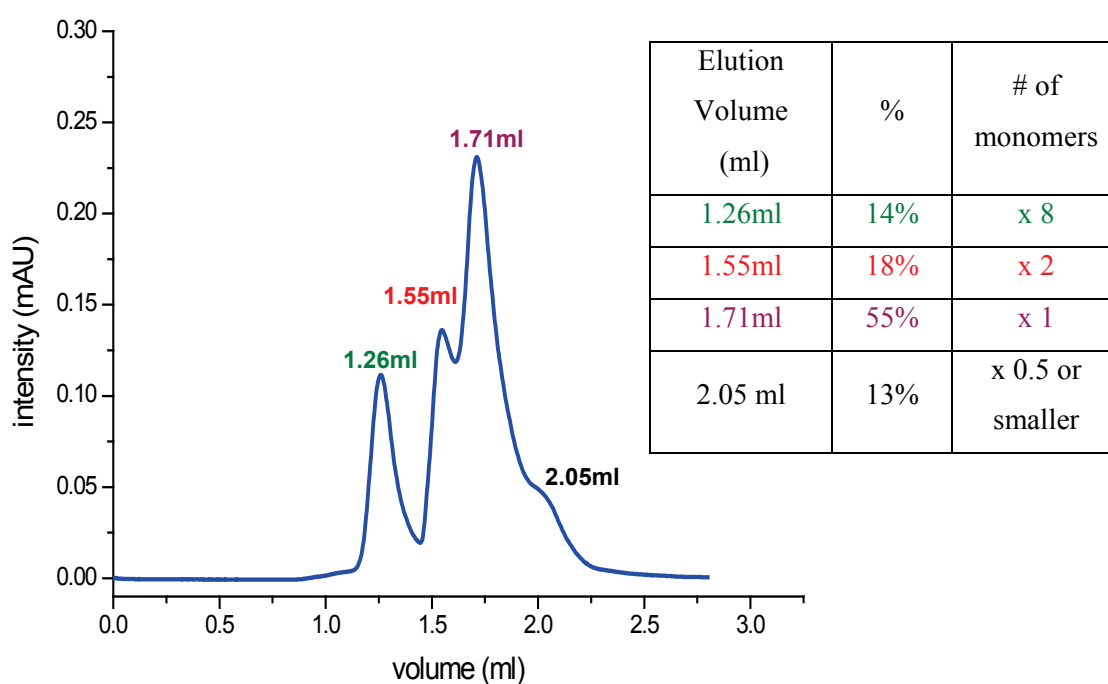
8.9 Percentage of spheroplasts after treatments to remove the S-layer

8.9.1 Bacitracin treatment

Table # 24: % of spheroplasts in different concentration of Bacitracin, after 48h growth. Sum: total of cell count, Sphero:spheroplast, SD: standard deviation.

OD_{600nm}=0.37 0µg/ml Bacitracin		
Flask #1:	rod	Sphero
Sum (n=4)	241	26
total	267	
% Sphero of total		**10**
SD	11	2

OD_{600nm}=0.25 50µg/ml Bacitracin		
Flask #5:	rod	Sphero
Sum (n=13)	239	37
total	276	
% Sphero of total		**13**
SD	7	2

OD_{600nm}=0.38 5µg/ml Bacitracin		
Flask #2:	rod	Sphero
Sum (n=12)	213	86
total	299	
% Sphero of total		**29**
SD	5	2

OD_{600nm}=0.09 100µg/ml Bacitracin		
Flask #6:	rod	Sphero
Sum (n=7)	23	3
total	26	
% Sphero of total		**12**
SD	2	1

OD_{600nm}=0.34 10µg/ml Bacitracin		
Flask #3:	rod	Sphero
Sum (n=11)	71	56
total	127	
% Sphero of total		**44**
SD	3	2

OD_{600nm}=0.31 25µg/ml Bacitracin		
Flask #4:	rod	Sphero
Sum (n=13)	243	86
total	329	
% Sphero of total		**26**
SD	5	2

8.9.2 Protease treatment

Table # 25: % of spheroplast in different concentration of Protease, after 97.1h growth. Samples #1 and #2 were diluted ¼ before counting the cells. Sum: total cell count, Sphero: spheroflast, SD: standard deviation.

OD$_{600nm}$=1.064		
0mg/ml Protease		
Flask #1:	rod	Sphero
Sum (n=5)	149	22
Sumx4	596	88
total	684	
% Sphero of total		**13**
SD	8	2

OD$_{600nm}$=0.92		
0.1mg/ml Protease		
Flask #2:	rod	Sphero
Sum (n=6)	125	8
Sumx4	500	32
total	532	
% Sphero of total		**6**
SD	11	2

OD$_{600nm}$=0.58		
0.3mg/ml Protease		
Flask #3:	rod	Sphero
Sum (n=11)	552	117
total	169	
% Sphero of total		**69**
SD	3	4

OD$_{600nm}$=0.29		
0.6µg/ml Protease		
Flask #4:	rod	Sphero
Sum (n=14)	1	57
total	58	
% Sphero of total		**98**
SD	0	2

OD600nm=0.138		
0.9µg/ml Protease		
Flask #5:	rod	Sphero
Sum (n=10)	1	42
total	43	
% Sphero of total		**98**
SD	0	2

8.10 Activity of GPDH in cell-free extracts

8.10.1.1 Activity of GPDH measured in basal salt

The activity of glycerol-1-phosphate dehydrogenase (GPDH) in cell-free extracts was recorded by measuring the decrease in the absorbance of NADPH at 340nm, using basal salt as buffer, as seen in figure #97. 0.15mM NADPH was added to 800µl Basal salt solution and 25mM Tris-HCl (pH 7.2), at RT. The absorbance at 340mM was measured continuously for 3min (a_1). Then 20µl of cell-free extract (=1 mg) was added to the reaction mixture, and the absorbance was measured again continuously for ~2.5min (a_2). The reaction was initiated by adding 1.2mM DHAP (a_3).

The activity after each addition was calculated (a_1, a_2 and a_3, respectively), as seen in table #26, and the specific activity of GPDH in the cell-free extract was calculated by subtracting the unspecific activities (after adding NADPH, after adding cell-free extract, respectively) from the activity calculated after adding the substrate. The activity of GPDH in cell-free extract was found to be 88 mU/ml. Approximately 50% of the detected activity (177 mU/ml) is unspecific, and caused by oxidization of NADPH in the cell-free extract.
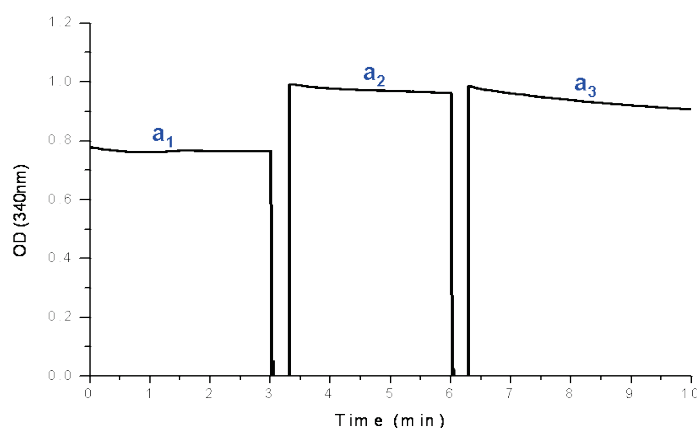
Figure # 97: The rates of NADH reduction after each addition (a_1, a_2 and a_3) are given in table #19

Table # 26: Activity of cell-free extract in basal salt

	Unspecific activity (mU/ml) when adding			**Specific activity (mU/ml)**
	NADPH (0.15mM) a_1	Cell extract (20μl) a_2	Precursor (1.2mM DHAP) a_3	$a_3\text{-}a_2\text{-}a_1$
In NaCl (#90)	16 mU/ml	73 mu/ml	177 mU/ml	88 mU/ml

8.10.1.2 What is the cofactor for GPDH?

To check the preferred cofactor for GPDH, the activity of GPDH was monitored continuously at 340nm using either NADH or NADPH, as cofactors. The activity (mU/ml) after each addition was calculated, and the specific activity was calculated by subtracting the unspecific activities from the activity measured after the precursor was added. As seen in figure #98 addition of cofactor to the cell-free extract resulted in activity, most probably because of metabolites in the cell-free extract, which can use NADH/NADPH as reducing agents. But only when NADPH was used as cofactor did the addition of DHAP result in an additional decrease in absorbance.

Table # 27: The activity of GPDH in cell-free extract using NADH or NADPH as cofactors. Activity was measure in basal salt, at RT.

	Unspecific activity (mU/ml) when adding			Specific activity (mU/ml)
	Cofactor (0.15mM)	Cell extract (20μl)	Precursor (1.2mM DHAP)	
With NADH (# 88)	39	129	153	0
With NADPH (#85)	16	73	177	88
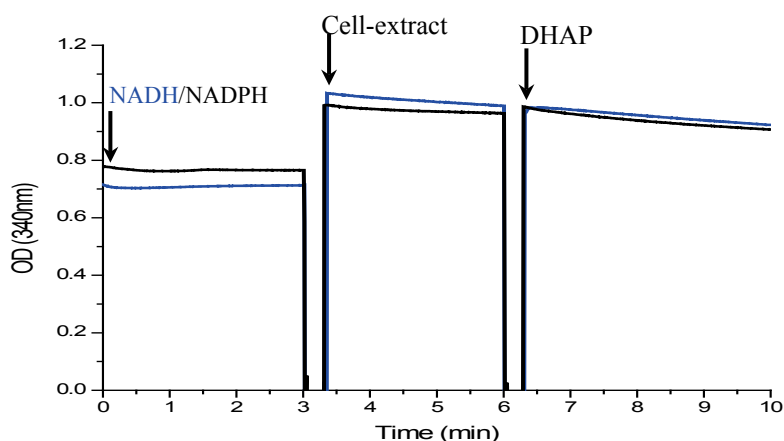
Figure # 98: Activity of GPDH in cell-free extract using NADH or NADPH as cofactors. The decrease in the absorbance at 340nm was measured at RT in Basal salt. Arrows indicates were additions were made.

8.10.1.3 Activity of GPDH measured in 2.8M KCl

As GPDH is an intracellular enzyme, it might be possible to increase the detected specific activity of GPDH in the cell-free extracts by changing the buffer to one containing KCl. This hypothesis was confirmed when the activity of GPDH was measured in the presence of 2.8M KCl.

0.15mM NADPH in 2.8M KCl and 25mM Tris-HCl (pH 7.2) were added to 20μl of cell-free extract (=1mg). The reaction was initiated by adding 1.2mM DHAP to the reaction mixture. The absorbance was measured continuously over 10min, at 340nm RT. A specific activity of 247±18 mU/ml (n=3) was detected (table #21, figure #84), 26% of the total activity detected, was unspecific, due to the cell-free extract, and 5% of the total activity detected was due to oxidation of the NADPH.
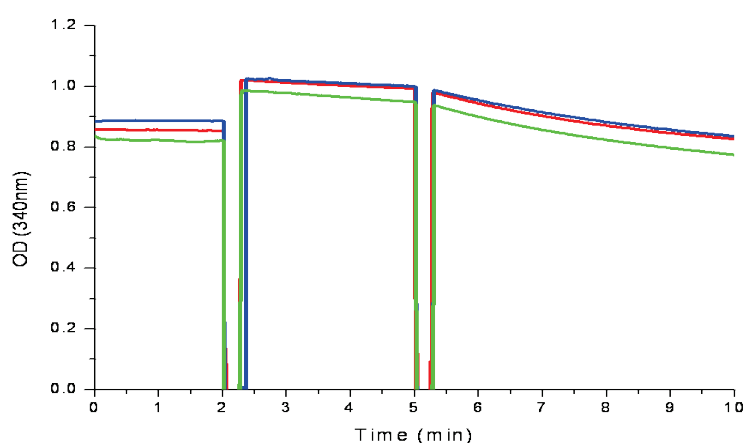
Figure # 99: Activity of GPDH in cell-free extracts using NADPH as cofactors. The decrease in the absorbance at 340nm was measured at RT in 2.8M KCl.

190

Table # 28: The specific activity of GPDH in 2.8M KCl buffer.

	Unspecific activity (mU/ml) when adding			Specific activity (mU/ml)
	NADPH (0.15mM)	Cell extract (20µl)	Precursor (1.2mM DHAP)	
# 140	8	81	347	258
#145	0.8	81	339	257
#169	48	121	395	226
Average	**19±25**	**94±23**	**360±30**	**247±18**

The specific activity of GPDH detected in cell-free extracts was compared to the unspecific activity detected in mixtures without added precursor, or without cell-free extract, or without added cofactor. The absorbance at 340nm of those samples was recorded, as seen in figure #99.
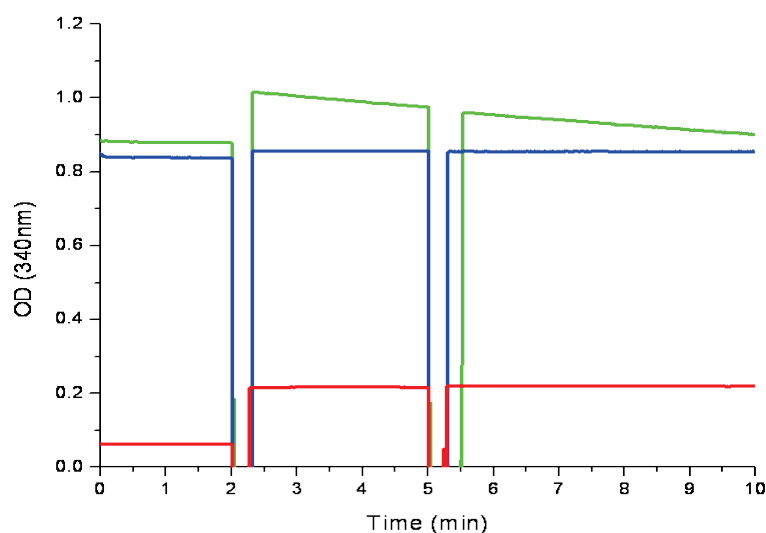
Figure # 100: Activity of GPDH in cell-free extract using NADPH as cofactors. The decrease in the absorbance at 340nm was measured at RT in 2.8M KCl. Activity of GPDH without DHAP, without cell-free extract, and without NADPH.

The maximum unspecific activity of the cell-free extract without DHAP was 1mU/ml (figure #100, green curve, table #29), whereas, unspecific activity of 8mU/ml was detected when NADPH was added to DHAP. By far, the higher contribution to the unspecific activity was seen when cell-free extract was added to NADPH. This activity was 36% of the specific activity detected when DAHP was added as well to the reaction mixture. A slight decrease in the absorbance was seen when there was

191

only NADPH or DHAP in the reaction mixture, suggesting the NADPH was not oxidized during the 10min reaction, and only the combination of NADPH + DHAP + cell free extract resulted in specific activity.

Table # 29: The unspecific activity of GPDH due to addition cell-extract, NADPH and DHAP.

	Unspecific activity (mU/ml) when adding			Specific activity (mU/ml)
	NADPH (0.15mM)	Cell extract (20µl)	Precursor (1.2mM DHAP)	
# 170 (WO DHAP)	16	121	105	
#171 (WO cell extract)	8	0	0	8
#172 (WO NADPH)	1	2	0	1

The activity of GPDH was measured both in basal salts and in 2.8M KCl. As seen in table # 30, slight differences are seen in the unspecific activities when the activity of GPDH was measured either in basal salts, or in KCl. But, the steeper slope seen when DHAP was added to the reaction mixture containing KCl, resulted in ~3-fold higher specific activity (258mU/ml vs 88mU/ml, respectively).

Table # 30: Effect of NaCl and KCl on the activity of GPDH

	[a]Unspecific activity (mU/ml) when adding			[b]Specific activity (mU/ml)
	NADPH (0.15mM)	Cell extract (20µl)	Precursor (1.2mM DHAP)	
In NaCl (#90)	16	73	177	88
In KCl (# 140) (n=3)	19±25	94±23	360±30	247±18

[a]- The activity after each addition was calculated: after adding NADPH, cell-free extract, and precursor, in the concentrations indicated in the table.

[b]-The specific activity of GPDH in the cell-free extracts was calculated by subtracting the unspecific activities (after adding NADPH, after adding cell-free extract, respectively) from the activity calculated after adding the precursor.

8.10.1.4 Optimizing the conditions for GPDH

8.10.1.4.1 Effect of increasing the volume of cell-extract

The influence of different volumes of cell-free extract (stock 50mg/ml) was measured, in order to find the optimal volume of the cell-free extract for assessing GPDH activity. 0.15mM NADPH was added to the reaction mixture which contained 2.8M KCl+ 25mM TrisHCl (pH 7.2), then, different volumes of cell-free extract were added (between 10µl-150µl), followed by addition of 1.2mM DHAP. The absorbance at 340nm was recorded, as seen in figure #101.
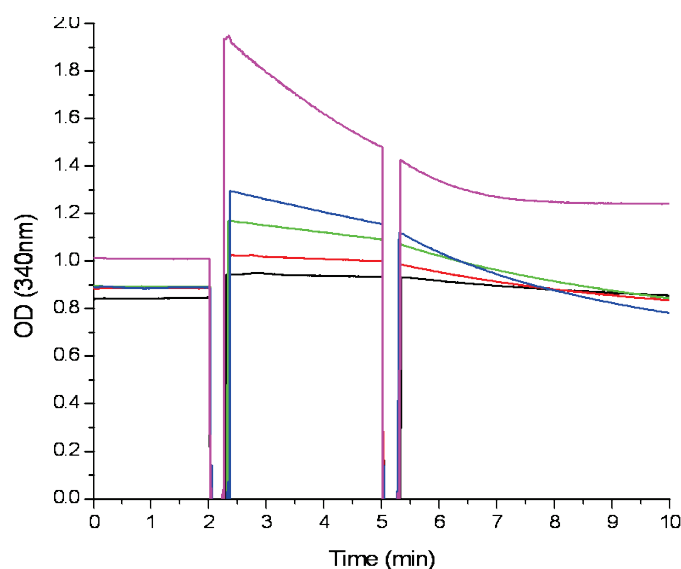
Figure # 101: The effect of different volumes of cell extract on the activity of Glycerol-1-phosphate dehydrogenase (GPDH). Extract additions were color coded: 10µl, 20µl, 40µl, 60µl, 150µl, respectively.

As seen from table #31, 10µl or 20µl cell-free extract (equals to 0.5, and 1mg/ml total protein according to BCA assay) were the optimal choices for the concentration of protein used in the assay, as seen from the linear increase of the ΔA/min, which corresponded to the increase in the volume. Higher cell-free extract volumes did not follow the expected correlation. The lowest unspecific activity (~120mU/ml) was obtained when using 10-40µl cell extract (=0.5-2mg/ml total protein).

Table # 31: The correlation between the volume of the cell extracts and the slopes obtained after adding DHAP.

Sample #	Vol. of cell extract (µl)	Slope when adding DHAP (ΔA/min)	Unspecific Activity (mu/ml)	Activity* (mU/ml)
146	10µl	-0.021	113	340
145	20µl	-0.042	81	340
147	40µl	-0.062	117	250
148	60µl	-0.1	145	270
144	150µl	-0.089	186	96

* The unspecific activity was not subtracted

8.10.1.4.2   *Effect of increasing the concentration of substrate*

The influence of different substrate concentration on the specific activity of glycerol 1-P dehydrogenase, was tested 20µl (=1mg/ml total protein) of cell-free extract, as seen in figure #102. As seen in table #32, the specific activity correlated to the increase in the concentration of DHAP, when using concentrations between 6-12mM DHAP.
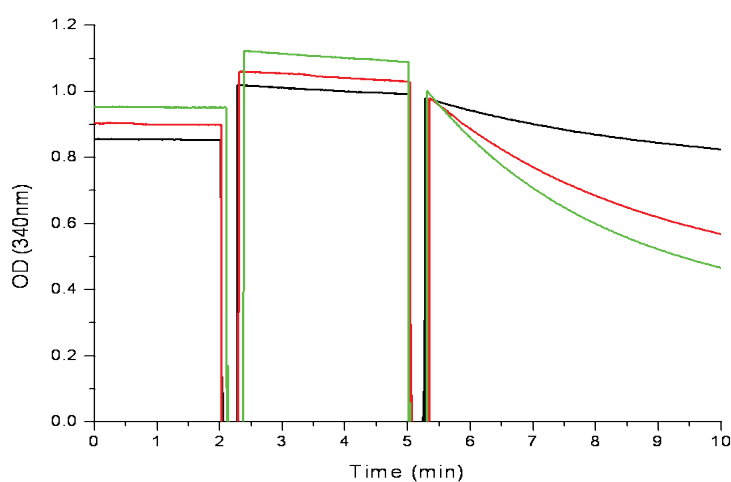
Figure # 102: The effect of different concentration of DHAP (1.2mM, 6mM, 12mM, respectively) on the activity of Glycerol-1-phosphate dehydrogenase (GPDH).

Table # 32: The specific activity of GPDH calculated after addition of different concentration of DHAP.

	Unspecific activity (mU/ml) when adding			Specific activity (mU/ml)
	NADPH (0.15mM)	Cell extract (20μl)	Precursor (DHAP)	
0.12mM (#149)	0	97	113	16
1.2mM (#140)	8	81	347	258
6mM (# 143)	24	105	1016	887
12mM (#150)	8	97	1427	1322

8.10.1.5 GPDH from cells grown in complex medium vs. synthetic medium WO AroAAs

According to MS results of A. Tebbe, the regulation factor of the GPDH gene OE1620F (EC 1.1.1.261) is +2.17 (synthetic medium/complex medium). This fact might indicate that cell-free extracts prepared from synthetic medium will contain more enzyme compared to cells from complex medium. As seen in figure #103 and table #33, the enzyme activity was ~1.6 higher in cell extracts derived from cells growing in complex medium.
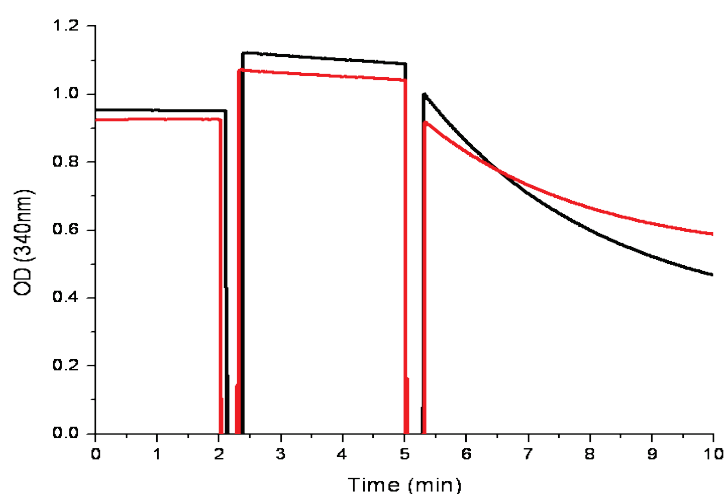
Figure # 103: The activity of Glycerol-1-phosphate dehydrogenase (OE1620F) in cell-free extract from complex medium *vs*. synthetic medium WO AroAAs.

Table # 33: Activity of glycerol-1-phosphate dehydrogenase (GPDH) in cell- free extract from Halo medium vs synthetic medium without AroAAs.

	Unspecific activity (mU/ml) when adding			Specific activity (mU/ml)
	NADH (0.15mM)	Cell extract (20μl)	Precursor (12mM DHAP)	
Halo medium (# 150)	8	96	1427	1322
Synthetic WO AroAAs (# 174)	10	80	940	850

8.10.1.6 Activity assay for GPDH from cell-free extract in 2.8M KCl

825μl 3.4M KCl + 10mM NaCl + 25mM Tris pH=7.2

15μl 10mM NADPH (0.15mM)

20μl cell extract (1mg total protein according to BCA assay)

100μl 120mM DHAP (12mM)

Total volume: 1000μl (RT)

8.10.1.7 Aldolase activity in cell free extract of *H. salinarum*

In the current study, aldolase activity was detected in pure preparations of proteins OE1472F and OE2019F, so it was of interest to see if aldolase activity could be detected directly in cell extracts of *H. salinarum*, using a spectrometric assay. To check this, a coupled enzyme assay was performed, where the formation of DHAP was detected by the disappearance of NADPH. DHAP, one of the products obtained after the cleavage of F-1,6-P is converted to glycerol-1-P by glycerol-1-phosphate dehydrogenase (GPDH, OE1602F, EC 1.1.1.261), as shown in the reaction below.

$$F\text{-}1,6\text{-}P \rightarrow GAP + DHAP$$

$$DHAP + NADPH \rightarrow Glycerol\text{-}1\text{-}P + NADP^+$$

The aldolase activity in cell-free extracts of R1 grown in complex medium, was measured using the coupled enzyme assay with GPDH in basal salt and 2.8M KCl and 1.2mM F-1,6-P. Very low activity (8mU/ml) was detected when the buffer was basal

salt (table #34, sample #94 pink), but no activity was detected when the assay was preformed in 2.8M KCl (table #34). That may be due to low turnover number of the Aldolase, thus lower amounts of DHAP were released. As indicated from sample 169 (bottom row of table #33), GPDH was still active in cell-free extracts, when aldolase activity was measured.

Table # 34: Aldolase activity in cell free-extracts of H. salinarum

Substrates	[a]Unspecific activity (mU/ml) when adding			[b]Specific activity (mU/ml)
	NADPH (0.15mM)	Cell extract (20μl)	Precursors (1.2mM each)	
1.2mM F-1,6-P (# 94)	0 mU/ml	72 mU/ml	80 mU/ml	8
1.2mM F-1,6-P (# 166)	8 mU/ml	113mU/ml	97mU/ml	0
WO F-1,6-P (# 168)	2.5 mU/ml	137 mU/ml	97 mU/ml	0
1.2mM DHAP (# 169) 7 days after cell extract preparation	48 mU/ml	121 mU/ml	395 mU/ml	226

[a]- The activity after each addition was calculated after adding NADPH, cell-free extract, and precursor, in the concentrations indicated in the table.

[b]-The specific activity of GPDH in the cell-free extracts was calculated by subtracting the unspecific activities (after adding NADPH, after adding cell-free extract, respectively) from the activity calculated after adding the precursor.

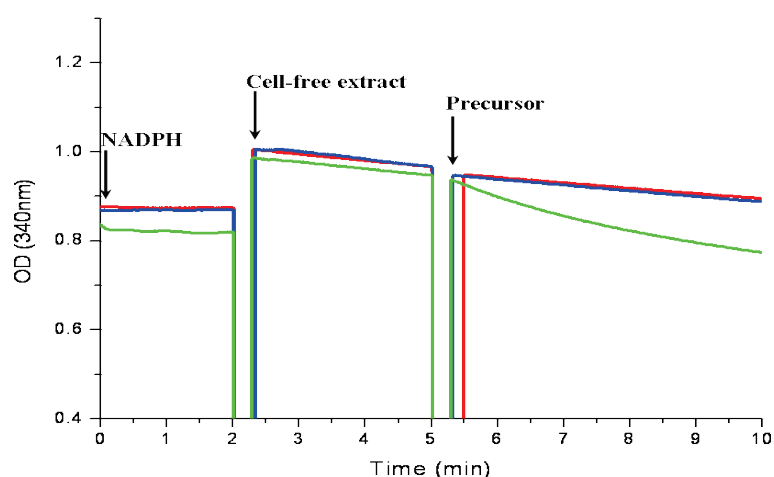**Figure # 104:** Aldolase activity in *H. salinarum* cell free-extracts as detected by the coupled enzyme assay coupled with GDPH. 2.8M KCl, NADPH, cell free-extracts, and precursors were added at the indicated times and the decrease in absorbance at 340nm followed at RT. Red- 1.2mM F-1,6-P, Blue-without precursor, Green-positive control for the activity of GDPH using 1.2mM DHAP. Arrows indicates the addition time point.

9 Bibliography

Adhya, S. (1999). Regulation of gene expression:operons and regulons. In Biology of the prokaryotes, J.W. Lengeler, Drews, G. and Schlegel, H.G., ed. (Stuttgart, Germany: Georg Thieme Verlg), pp. 437-467.

Aharoni, A., and Vorst, O. (2001). DNA microarrays for functional plant genomics. Plant Molecular Bioloby *48*, 99-118.

Allers, T., and Mevarech, M. (2005). Archaeal genetics - the third way. Nat Rev Genet *6*, 58-73.

Altekar W., and Dhar, N.M. (1987). Archaebacterial class I and class II aldolase from extreme halophiles. Origins of Life and Evolution of the Biosphere *18*, 59-64.

Baliga, N.S., Goo, Y.A., Ng, W.V., Hood, L., Daniels, C.J., and DasSarma, S. (2000). Is gene expression in Halobacterium NRC-1 regulated by multiple TBP and TFB transcription factors? Mol Microbiol *36*, 1184-1185.

Baliga, N.S., Pan, M., Goo, Y.A., Yi, E.C., Goodlett, D.R., Dimitrov, K., Shannon, P., Aebersold, R., Ng, W.V., and Hood, L. (2002). Coordinate regulation of energy transduction modules in Halobacterium sp. analyzed by a global systems approach. Proc Natl Acad Sci U S A *99*, 14913-14918.

Banerji, S., Wakefield, A.E., Allen, A.G., Maskell, D.J., Peters, S.E., and Hopkin, J.M. (1993). The cloning and characterization of the arom gene of Pneumocystis carinii. J Gen Microbiol *139*, 2901-2914.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L.*, et al.* (2004). The Pfam protein families database. Nucleic Acids Res *32*, D138-141.

Bentley, R. (1990). The shikimate pathway-a metabolic tree with many branches. Crit Rev Biochem Mol Biol *25*, 307-384.

Besir, H. (2001). Investigation of lipid-mediated crystallization of the ion pumps bacteriorhodopsin and halorhodopsin from *Halobacterium salinarum.* (Munich, LMU).

Bonneau, R., Reiss, D.J., Shannon, P., Facciotti, M., Hood, L., Baliga, N.S., and Thorsson, V. (2006). The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. Genome Biol *7*, R36.

Bottomley, J.R., Clayton, C.L., Chalk, P.A., and Kleanthous, C. (1996). Cloning, sequencing, expression, purification and preliminary characterization of a type II dehydroquinase from Helicobacter pylori. Biochem J *319 (Pt 2)*, 559-565.

Byng, G.S., Johnson, J.L., Whitaker, R.J., Gherna, R.L., and Jensen, R.A. (1983). The evolutionary pattern of aromatic amino acid biosynthesis and the emerging phylogeny of pseudomonad bacteria. J Mol Evol *19*, 272-282.

Carpenter, E.P., Hawkins, A.R., Frost, J.W., and Brown, K.A. (1998). Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis. Nature *394*, 299-302.

Charles, I.G., Keyte, J.W., Brammar, W.J., Smith, M., and Hawkins, A.R. (1986). The isolation and nucleotide sequence of the complex AROM locus of Aspergillus nidulans. Nucleic Acids Res *14*, 2201-2213.

Cline, S.W., and Doolittle, W.F (1987). Efficient transfection of the archaebacterium *Halobacterium halobium*. J. Bacteriol *169*, 1341-1344.

Cline, S.W., Lam, W.L., Charlebois, R.L., Schalkwyk, L.C., and Doolittle, W.F. (1989). Transformation methods for halophilic archaebacteria. Can J Microbiol *35*, 148-152.

Cooper, R.A., and Anderson, A. (1970). The formation and catabolism of methylglyoxal during glycolysis in Escherichia coli. FEBS Lett *11*, 273-276.

D'souza S.E., and Altekar W. (1998). A class II fructose -1,6-bisphosphate aldolase from a halophilic archaebacterium *Haloferax mediterranei*. J. Gen. Appl. Microbiol *44*, 235-241.

D'Souza, S.E., and Altekar, W. (1982). A halophilic fructose 1,6-biphosphate aldolase from Halobacterium halobium. Indian J Biochem Biophys *19*, 135-138.

Daniell, H., Vivekananda, J., Nielsen, B.L., Ye, G.N., Tewari, K.K., and Sanford, J.C. (1990). Transient foreign gene expression in chloroplasts of cultured tobacco cells after biolistic delivery of chloroplast vectors. Proc Natl Acad Sci U S A *87*, 88-92.

Daugherty, M., Vonstein, V., Overbeek, R., and Osterman, A. (2001). Archaeal shikimate kinase, a new member of the GHMP-kinase family. J Bacteriol *183*, 292-300.

Davies, W.D., and Davidson, B.E. (1982). The nucleotide sequence of aroG, the gene for 3-deoxy-D-arabinoheptulosonate-7-phosphate synthetase (phe) in Escherichia coli K12. Nucleic Acids Res *10*, 4045-4058.

Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R.A., Martinez-Arias, R., Henne, A., Wiezer, A., Baumer, S., Jacobi, C.*, et al.* (2002). The genome of Methanosarcina mazei: evidence for lateral gene transfer between bacteria and archaea. J Mol Microbiol Biotechnol *4*, 453-461.

Dhar, N.M., and Altekar, W. (1986). Distribution of Class-I and Class-Ii Fructose Biphosphate Aldolases in Halophilic Archaebacteria. Fems Microbiology Letters *35*, 177-181.

Dharmadi, Y., and Gonzalez, R. (2004). DNA microarrays: experimental issues, data analysis, and application to bacterial systems. Biotechnol Prog *20*, 1309-1324.

Diehn, M., and Relman, D.A. (2001). Comparing functional genomic datasets: lessons from DNA microarray analyses of host-pathogen interactions. Curr Opin Microbiol *4*, 95-101.

Dyall-Smith, D.M. Protocols for haloarchaeal genetics.

Ehlers, C., Weidenbach, K., Veit, K., Deppenmeier, U., Metcalf, W.W., and Schmitz, R.A. (2005). Development of genetic methods and construction of a chromosomal glnK1 mutant in Methanosarcina mazei strain Go1. Mol Genet Genomics *273*, 290-298.

Eymann, C., Homuth, G., Scharf, C., and Hecker, M. (2002). Bacillus subtilis functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. J Bacteriol *184*, 2500-2520.

Facciotti, M.T., Reiss, D.J., Pan, M., Kaur, A., Vuthoori, M., Bonneau, R., Shannon, P., Srivastava, A., Donohoe, S.M., Hood, L.E., and Baliga, N.S. (2007). General transcription factor specified global gene regulation in archaea. Proc Natl Acad Sci U S A *104*, 4630-4635.

Falb, M., Muller, K., Konigsmaier, L., Oberwinkler, T., Horn, P., von Gronau, S., Gonzalez, O., Pfeiffer, F., Bornberg-Bauer, E., and Oesterhelt, D. (2008). Metabolism of halophilic archaea. Extremophiles *12*, 177-196.

Falb, M., Pfeiffer, F., Palm, P., Rodewald, K., Hickmann, V., Tittor, J., and Oesterhelt, D. (2005). Living with two extremes: conclusions from the genome sequence of Natronomonas pharaonis. Genome Res *15*, 1336-1343.

Fondi, M., Emiliani, G., Lio, P., Gribaldo, S., and Fani, R. (2009). The Evolution of Histidine Biosynthesis in Archaea: Insights into the his Genes Structure and Organization in LUCA. J Mol Evol.

Fuchs, G., Winter, H., Steiner, I., and Stupperich, E. (1983). Enzymes of Gluconeogenesis in the Autotroph Methanobacterium-Thermoautotrophicum. Archives of Microbiology *136*, 160-162.

Galperin, M.Y., Aravind, L., and Koonin, E.V. (2000). Aldolases of the DhnA family: a possible solution to the problem of pentose and hexose biosynthesis in archaea. FEMS Microbiol Lett *183*, 259-264.

Gambacorta, A., Trincone, A., Nicolaus, B., Lama, L., and Derosa, M. (1994). Unique Features of Lipids of Archaea. Systematic and Applied Microbiology *16*, 518-527.

Gast, D.A., Jenal, U., Wasserfallen, A., and Leisinger, T. (1994). Regulation of tryptophan biosynthesis in Methanobacterium thermoautotrophicum Marburg. J Bacteriol *176*, 4590-4596.

Gehl, J. (2003). Electroporation: theory and methods, perspectives for drug delivery, gene therapy and research. Acta Physiol Scand *177*, 437-447.

Ghosh, M., and Sonawat, H.M. (1998). Kreb's TCA cycle in Halobacterium salinarum investigated by 13C nuclear magnetic resonance spectroscopy. Extremophiles *2*, 427-433.

Glansdorff, N. (1999). On the origin of operons and their possible role in evolution toward thermophily. J Mol Evol *49*, 432-438.

Gogarten, J.P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E.J., Bowman, B.J., Manolson, M.F., Poole, R.J., Date, T., Oshima, T.*, et al.* (1989). Evolution of the

Vacuolar H+-Atpase - Implications for the Origin of Eukaryotes. Proceedings of the National Academy of Sciences of the United States of America *86*, 6661-6665.

Gollnick, P., and Babitzke, P. (2002). Transcription attenuation. Biochim Biophys Acta *1577*, 240-250.

Gonzalez, J.M., Masuchi, Y., Robb, F.T., Ammerman, J.W., Maeder, D.L., Yanagibayashi, M., Tamaoka, J., and Kato, C. (1998). Pyrococcus horikoshii sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. Extremophiles *2*, 123-130.

Gonzalez, O., Gronau, S., Falb, M., Pfeiffer, F., Mendoza, E., Zimmer, R., and Oesterhelt, D. (2008). Reconstruction, modeling & analysis of Halobacterium salinarum R-1 metabolism. Mol Biosyst *4*, 148-159.

Gonzalez, O., Gronau, S., Pfeiffer, F., Mendoza, E., Zimmer, R., and Oesterhelt, D. (2009). Systems analysis of bioenergetics and growth of the extreme halophile Halobacterium salinarum. PLoS Comput Biol *5*, e1000332.

Gregor, D., and Pfeifer, F. (2005). In vivo analyses of constitutive and regulated promoters in halophilic archaea. Microbiology-Sgm *151*, 25-33.

Grey, V.L., and Fitt, P.S. (1976). Improved Synthetic Growth Medium for Halobacterium-Cutirubrum. Canadian Journal of Microbiology *22*, 440-442.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser *41*, 95-98.

Hein, M., Wegener, A.A., Engelhard, M., and Siebert, F. (2003). Time-resolved FTIR studies of sensory rhodopsin II (NpSRII) from Natronobacterium pharaonis: implications for proton transport and receptor activation. Biophys J *84*, 1208-1217.

Ikeda, M. (2006). Towards bacterial strains overproducing L-tryptophan and other aromatics by metabolic engineering. Appl Microbiol Biotechnol *69*, 615-626.

Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., and Miyata, T. (1989). Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc Natl Acad Sci U S A *86*, 9355-9359.

Kalapos, M.P. (1999). Methylglyoxal in living organisms: chemistry, biochemistry, toxicology and biological implications. Toxicol Lett *110*, 145-175.

Kamekura, M., Dyall-Smith, M.L., Upasani, V., Ventosa, A., and Kates, M. (1997). Diversity of alkaliphilic halobacteria: proposals for transfer of Natronobacterium vacuolatum, Natronobacterium magadii, and Natronobacterium pharaonis to Halorubrum, Natrialba, and Natronomonas gen. nov., respectively, as Halorubrum vacuolatum comb. nov., Natrialba magadii comb. nov., and Natronomonas pharaonis comb. nov., respectively. Int J Syst Bacteriol *47*, 853-857.

Kandler, O. (1994). Cell-Wall Biochemistry and 3-Domain Concept of Life. Systematic and Applied Microbiology *16*, 501-509.

Kawata, Y., Yano, S., and Kojima, H. (2003). Escherichia coli can be transformed by a liposome-mediated lipofection method. Bioscience Biotechnology and Biochemistry *67*, 1179-1181.

Kikkert, J.R. (1993). The Biolistic(R) Pds-1000 He Device. Plant Cell Tissue and Organ Culture *33*, 221-226.

Klare, J.P., Gordeliy, V.I., Labahn, J., Buldt, G., Steinhoff, H.J., and Engelhard, M. (2004). The archaeal sensory rhodopsin II/transducer complex: a model for transmembrane signal transfer. FEBS Lett *564*, 219-224.

Knaggs, A.R. (1999). The biosynthesis of shikimate metabolites. Nat Prod Rep *16*, 525-560.

Koch, M.K., and Oesterhelt, D. (2005). MpcT is the transducer for membrane potential changes in Halobacterium salinarum. Mol Microbiol *55*, 1681-1694.

Koide, T., Reiss, D.J., Bare, J.C., Pang, W.L., Facciotti, M.T., Schmid, A.K., Pan, M., Marzolf, B., Van, P.T., Lo, F.Y.*, et al.* (2009). Prevalence of transcription promoters within archaeal operons and coding sequences. Mol Syst Biol *5*, 285.

Kokoeva, M.V., Storch, K.F., Klein, C., and Oesterhelt, D. (2002). A novel mode of sensory transduction in archaea: binding protein-mediated chemotaxis towards osmoprotectants and amino acids. EMBO J *21*, 2312-2322.

Koonin, E.V., Makarova, K.S., and Aravind, L. (2001). Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol *55*, 709-742.

Krishnan, G., and Altekar, W. (1991). An Unusual Class-I (Schiff-Base) Fructose-1,6-Bisphosphate Aldolase from the Halophilic Archaebacterium Haloarcula-Vallismortis. European Journal of Biochemistry *195*, 343-350.

Kurella, M., Hsiao, L.L., Yoshida, T., Randall, J.D., Chow, G., Sarang, S.S., Jensen, R.V., and Gullans, S.R. (2001). DNA microarray analysis of complex biologic processes. J Am Soc Nephrol *12*, 1072-1078.

Lam, W.L., and Doolittle, W.F. (1989). Shuttle vectors for the archaebacterium Halobacterium volcanii. Proc Natl Acad Sci U S A *86*, 5478-5482.

Lanyi, J.K. (1977). Transport in halobacterium halobium: light-induced cation-gradients, amino acid transport kinetics, and properties of transport carriers. J Supramol Struct *6*, 169-177.

Lanyi, J.K., Duschl, A., Hatfield, G.W., May, K., and Oesterhelt, D. (1990). The primary structure of a halorhodopsin from Natronobacterium pharaonis. Structural, functional and evolutionary implications for bacterial rhodopsins and halorhodopsins. J Biol Chem *265*, 1253-1260.

Laub, M.T., McAdams, H.H., Feldblyum, T., Fraser, C.M., and Shapiro, L. (2000). Global analysis of the genetic network controlling a bacterial cell cycle. Science *290*, 2144-2148.

Lebherz, H.G., and Rutter, W.J. (1969). Distribution of fructose diphosphate aldolase variants in biological systems. Biochemistry *8*, 109-121.

Lemoine, F., Labedan, B., and Lespinet, O. (2008). SynteBase/SynteView: a tool to visualize gene order conservation in prokaryotic genomes. BMC Bioinformatics *9*, 536.

Lundgren, M., Andersson, A., Chen, L., Nilsson, P., and Bernander, R. (2004). Three replication origins in Sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination. Proc Natl Acad Sci U S A *101*, 7046-7051.

Lurquin, P.F. (1997). Gene transfer by electroporation. Mol Biotechnol *7*, 5-35.

Madern, D., Ebel, C., and Zaccai, G. (2000). Halophilic adaptation of enzymes. Extremophiles *4*, 91-98.

Madern, D., Pfister, C., and Zaccai, G. (1995). Mutation at a single acidic amino acid enhances the halophilic behaviour of malate dehydrogenase from Haloarcula marismortui in physiological salts. Eur J Biochem *230*, 1088-1095.

McCandliss, R.J., Poling, M.D., and Herrmann, K.M. (1978). 3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase. Purification and molecular characterization of the phenylalanine-sensitive isoenzyme from Escherichia coli. J Biol Chem *253*, 4259-4265.

Meile, L., Stettler, R., Banholzer, R., Kotik, M., and Leisinger, T. (1991). Tryptophan gene cluster of Methanobacterium thermoautotrophicum Marburg: molecular cloning and nucleotide sequence of a putative trpEGCFBAD operon. J Bacteriol *173*, 5017-5023.

Mescher, M.F., and Strominger, J.L. (1976). Structural (shape-maintaining) role of the cell surface glycoprotein of Halobacterium salinarium. Proc Natl Acad Sci U S A *73*, 2687-2691.

Metcalf, W.W., Zhang, J.K., Apolinario, E., Sowers, K.R., and Wolfe, R.S. (1997). A genetic system for Archaea of the genus Methanosarcina: liposome-mediated transformation and construction of shuttle vectors. Proc Natl Acad Sci U S A *94*, 2626-2631.

Mir, L.M. (2008). Application of electroporation gene therapy: past, current, and future. Methods Mol Biol *423*, 3-17.

Mitsuhashi, S., and Davis, B.D. (1954). Aromatic biosynthesis. XIII. Conversion of quinic acid to 5-dehydroquinic acid by quinic dehydrogenase. Biochim Biophys Acta *15*, 268-280.

Morar, M., White, R.H., and Ealick, S.E. (2007). Structure of 2-amino-3,7-dideoxy-D-threo-hept-6-ulosonic acid synthase, a catalyst in the archaeal pathway for the biosynthesis of aromatic amino acids. Biochemistry *46*, 10562-10571.

Neumann, E., Schaefer-Ridder, M., Wang, Y., and Hofschneider, P.H. (1982). Gene transfer into mouse lyoma cells by electroporation in high electric fields. EMBO J *1*, 841-845.

Nichols, C.E., Lockyer, M., Hawkins, A.R., and Stammers, D.K. (2004). Crystal structures of Staphylococcus aureus type I dehydroquinase from enzyme turnover experiments. Proteins *56*, 625-628.

Nichols, C.E., Ren, J., Lamb, H.K., Hawkins, A.R., and Stammers, D.K. (2003). Ligand-induced conformational changes and a mechanism for domain closure in Aspergillus nidulans dehydroquinate synthase. J Mol Biol *327*, 129-144.

Oesterhelt, D. (1982). Anaerobic Growth of Halobacteria. Methods in Enzymology *88*, 417-420.

Oesterhelt, D. (1988). The structure and mechanisim of the family of retinal proteins from halophilic archaea. Cuur. Op. Struct. Biol. *8*, 489-500.

Oesterhelt, D. (1998). The structure and mechanism of the family of retinal proteins from halophilic archaea. Curr Opin Struct Biol *8*, 489-500.

Oesterhelt, D., and Krippahl, G. (1983). Phototrophic growth of halobacteria and its use for isolation of photosynthetically-deficient mutants. Ann Microbiol (Paris) *134B*, 137-150.

Oesterhelt, D., and Krippahl, G. (1973). Light Inhibition of Respiration in Halobacterium-Halobium. Febs Letters *36*, 72-76.

Ouhammouch, M., Dewhurst, R.E., Hausner, W., Thomm, M., and Geiduschek, E.P. (2003). Activation of archaeal transcription by recruitment of the TATA-binding protein. Proc Natl Acad Sci U S A *100*, 5097-5102.

Pfeiffer, M., Rink, T., Gerwert, K., Oesterhelt, D., and Steinhoff, H.J. (1999). Site-directed spin-labeling reveals the orientation of the amino acid side-chains in the E-F loop of bacteriorhodopsin. Journal of Molecular Biology *287*, 163-171.

Pinto, P.V.C., . Van Dereal P.A., Kaplan, A. (1969). Aldolase I. Colorematric Determination. Clinical Chemistry *15*, 339-348.

Pittard, A.J. (1996). Biosynthesis of the aromatic amino acids, Vol 1, 2 edn (Washington, D.C. : ASM Press, Washington, D.C.).

Pittard, J., Camakaris, H., and Yang, J. (2005). The TyrR regulon. Mol Microbiol *55*, 16-26.

Pollock, T.J., Thorne, L., Yamazaki, M., Mikolajczak, M.J., and Armentrout, R.W. (1994). Mechanism of bacitracin resistance in gram-negative bacteria that synthesize exopolysaccharides. J Bacteriol *176*, 6229-6237.

Porat, I., Sieprawska-Lupa, M., Teng, Q., Bohanon, F.J., White, R.H., and Whitman, W.B. (2006). Biochemical and genetic characterization of an early step in a novel pathway for the biosynthesis of aromatic amino acids and p-aminobenzoic acid in the archaeon Methanococcus maripaludis. Mol Microbiol *62*, 1117-1131.

Porat, I., Waters, B.W., Teng, Q., and Whitman, W.B. (2004). Two biosynthetic pathways for aromatic amino acids in the archaeon Methanococcus maripaludis. J Bacteriol *186*, 4940-4950.

Price, M.N., Huang, K.H., Alm, E.J., and Arkin, A.P. (2005). A novel method for accurate operon predictions in all sequenced prokaryotes. Nucleic Acids Res *33*, 880-892.

Puhler, G., Leffers, H., Gropp, F., Palm, P., Klenk, H.P., Lottspeich, F., Garrett, R.A., and Zillig, W. (1989). Archaebacterial DNA-Dependent Rna-Polymerases Testify to the Evolution of the Eukaryotic Nuclear Genome. Proceedings of the National Academy of Sciences of the United States of America *86*, 4569-4573.

Richard, J.P. (1991). Kinetic parameters for the elimination reaction catalyzed by triosephosphate isomerase and an estimation of the reaction's physiological significance. Biochemistry *30*, 4581-4585.

Rohlin, O.M., kAO, l., AND lIAO, j. (2002). Global expression profiling of acetate-grown *Escherichia coli*. J. Biol Chem *277*, 13175-13183.

Rosenshine, I., Tchelet, R., and Mevarech, M. (1989). The mechanism of DNA transfer in the mating system of an archaebacterium. Science *245*, 1387-1389.

Rutter, W.J. (1964a). Evolution of Aldolase. Fed Proc *23*, 1248-1257.

Rutter, W.J.a.G., W.E. (1964b). Taxonomic biochemistry and serology, Vol 417 (New York: N.Y. Ronald Press).

Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M.*, et al.* (2003). TM4: a free, open-source system for microarray data management and analysis. Biotechniques *34*, 374-378.

Sanford J. C., K.T.M., Wolf D.E., and Allen N. (1987). Delivery of substances into cells and tissues using a particle bombardment process. Particulate Science and Technology *5*, 27-37.

Sanford, J.C. (1988). The Biolistic Process. Trends in Biotechnology *6*, 299-302.

Sanford, J.C. (1990). Biolistic Plant Transformation. Physiologia Plantarum *79*, 206-209.

Sanford, J.C., Smith, F.D., and Russell, J.A. (1993). Optimizing the biolistic process for different biological applications. Methods Enzymol *217*, 483-509.

Schafer, G., Purschke, W.G., Gleissner, M., and Schmidt, C.L. (1996). Respiratory chains of archaea and extremophiles. Biochim Biophys Acta *1275*, 16-20.

Schafer, T., and Schonheit, P. (1992). Maltose Fermentation to Acetate, Co2 and H2 in the Anaerobic Hyperthermophilic Archaeon Pyrococcus-Furiosus - Evidence for the Operation of a Novel Sugar Fermentation Pathway. Archives of Microbiology *158*, 188-202.

Schmies, G., Luttenberg, B., Chizhov, I., Engelhard, M., Becker, A., and Bamberg, E. (2000). Sensory rhodopsin II from the haloalkaliphilic natronobacterium pharaonis: light-activated proton transfer reactions. Biophys J *78*, 967-976.

Schoner, R., and Herrmann, K.M. (1976). 3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from Escherichia coli. J Biol Chem *251*, 5440-5447.

Schut, G.J., Zhou, J., and Adams, M.W. (2001). DNA microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus: evidence for anNew type of sulfur-reducing enzyme complex. J Bacteriol *183*, 7027-7036.

Schwaiger, R., Schwarz, C., Furtwanngler, K., Tarasov, V., Wende, A., and Oesterhelt, D. (2010). Transcriptional control by two leucine-responsive regulatory proteins in *Halobacterium salinarum* R1. Submitted.

Schweiger, U. (1996). Functional characterization of the retinal binding pocket of bacteriorhodopsin via specific mutagenesis. (Innsbruck, LFU).

Selkov, E., N. Maltsev, G. J. Olsen, R. Overbeek and W.B. Whitman (1997). A reconstruction of the methbolism of *Methanococcus jannaschii* from sequence data. Genes Cells *197*, GC11-GC26.

Shark K.B., S., F. D., Harpending, P.R., Rasmussen, J.L. and Sanford, J.C. (1990). Biolistic transformation of a procaryote: *Bacillus megaterium*. Applied and Environmental Microbiology *53*, 480-485.

Shultz, J., Hermodson, M.A., Garner, C.C., and Herrmann, K.M. (1984). The nucleotide sequence of the aroF gene of Escherichia coli and the amino acid sequence of the encoded protein, the tyrosine-sensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. J Biol Chem *259*, 9655-9661.

Sibley, J.A., and Lehninger, A.L., (1949). Determination of aldolase in animal tissues. J. Biol. Chem *177*, 859.

Siebers, B., Brinkmann, H., Dorr, C., Tjaden, B., Lilie, H., van der Oost, J., and Verhees, C.H. (2001). Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. J Biol Chem *276*, 28710-28718.

Siewert, G., and Strominger, J.L. (1967). Bacitracin: An Inhibitor of the Dephosphorylation of Lipid Pyrophosphate, an Intermediate in the Biosynthesis of the Peptidoglycan of Bacterial Cell Walls. Proc Natl Acad Sci U S A *57*, 767-773.

Smith, F.D., Harpending, P.R., and Sanford, J.C. (1992). Biolistic transformation of prokaryotes: factors that affect biolistic transformation of very small cells. J Gen Microbiol *138*, 239-248.

Smith, N.N., and Gallagher, D.T. (2008). Structure and lability of archaeal dehydroquinase. Acta Crystallogr Sect F Struct Biol Cryst Commun *64*, 886-892.

Soliman, G.S.H., and Truper, H.G. (1982). Halobacterium pharaonis sp. nov., a new, extremely haloalkaliphilic archaebacterium with low magnesium requirement. Zentralblatt fur Bakterologie Mikrobiologie und Hygiene I Abteilung Originale C *3*, 318-329.

Soliman, G.S.H., and Truper, H.G. (1982). Halobacterium-Pharaonis Sp-Nov, a New, Extremely Haloalkaliphilic Archaebacterium with Low Magnesium Requirement. Zentralblatt Fur Bakteriologie Mikrobiologie Und Hygiene I Abteilung Originale C-Allgemeine Angewandte Und Okologische Mikrobiologie *3*, 318-329.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol *98*, 503-517.

Srinivasan, P.R., Rothchild, J., and Sprinson, D.B. (1963). The enzymatic conversion of 3-deoxy-D-arabino-heptulosinic acid 7-phosphate to 5-dhydroquinate. J. Biol Chem *238*, 3176-3182.

Stoeckenius, W., Lozier, R.H., and Bogomolni, R.A. (1979). Bacteriorhodopsin and the purple membrane of halobacteria. Biochim Biophys Acta *505*, 215-278.

Stone, K.J., and Strominger, J.L. (1971). Mechanism of action of bacitracin: complexation with metal ion and C 55 -isoprenyl pyrophosphate. Proc Natl Acad Sci U S A *68*, 3223-3227.

Storm, D.R., and Strominger, J.L. (1973). Complex formation between bacitracin peptides and isoprenyl pyrophosphates. The specificity of lipid-peptide interactions. J Biol Chem *248*, 3940-3945.

Tani, T.H., Khodursky, A., Blumenthal, R.M., Brown, P.O., and Matthews, R.G. (2002). Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. Proc Natl Acad Sci U S A *99*, 13471-13476.

Thomson, G.J., Howlett, G.J., Ashcroft, A.E., and Berry, A. (1998). The dhnA gene of Escherichia coli encodes a class I fructose bisphosphate aldolase. Biochem J *331 (Pt 2)*, 437-445.

Tindall, B.J., Ross, H.N.M., and Grant, W.D. (1984). Natronobacterium Gen-Nov and Natronococcus Gen-Nov, 2 New Genera of Haloalkaliphilic Archaebacteria. Systematic and Applied Microbiology *5*, 41-57.

Tribe, D.E., Camakaris, H., and Pittard, J. (1976). Constitutive and repressivle enzymes of the common pathway of aromatic biosynthesis in Escherichia coli K-12: regulation of enzyme synthesis at different growth rates. J Bacteriol *127*, 1085-1097.

Twellmeyer, J., Wende, A., Wolfertz, J., Pfeiffer, F., Panhuysen, M., Zaigler, A., Soppa, J., Welzl, G., and Oesterhelt, D. (2007). Microarray analysis in the archaeon Halobacterium salinarum strain R1. PLoS One *2*, e1064.

Upreti, R.K., Kumar, M., and Shankar, V. (2003). Bacterial glycoproteins: functions, biosynthesis and applications. Proteomics *3*, 363-379.

Wende, A., Furtwangler, K., and Oesterhelt, D. (2009). Phosphate-Dependent Behavior of the Archaeon Halobacterium salinarum Strain R1. Journal of Bacteriology *191*, 3852-3860.

Whipp, M.J., and Pittard, A.J. (1977). Regulation of aromatic amino acid transport systems in Escherichia coli K-12. J Bacteriol *132*, 453-461.

White, R.H. (2004). L-Aspartate semialdehyde and a 6-deoxy-5-ketohexose 1-phosphate are the precursors to the aromatic amino acids in Methanocaldococcus jannaschii. Biochemistry *43*, 7618-7627.

White, R.H., and Xu, H. (2006). Methylglyoxal is an intermediate in the biosynthesis of 6-deoxy-5-ketofructose-1-phosphate: a precursor for aromatic amino acid biosynthesis in Methanocaldococcus jannaschii. Biochemistry *45*, 12366-12379.

Widlanski, T., Bender, S.L., and Knowles, J.R. (1989). Dehydroquinate synthase: the use of substrate analogues to probe the late steps of the catalyzed reaction. Biochemistry *28*, 7572-7582.

Williams, R.S., Johnston, S.A., Riedy, M., DeVit, M.J., McElligott, S.G., and Sanford, J.C. (1991). Introduction of foreign genes into tissues of living mice by DNA-coated microprojectiles. Proc Natl Acad Sci U S A *88*, 2726-2730.

Willins, D.A., Ryan, C.W., Platko, J.V., and Calvo, J.M. (1991). Characterization of Lrp, and Escherichia coli regulatory protein that mediates a global response to leucine. J Biol Chem *266*, 10768-10774.

Woese, C.R., Kandler, O., and Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A *87*, 4576-4579.

Xie, G., Keyhani, N.O., Bonner, C.A., and Jensen, R.A. (2003). Ancient origin of the tryptophan operon and the dynamics of evolutionary change. Microbiol Mol Biol Rev *67*, 303-342, table of contents.

Yang, N.S., Burkholder, J., Roberts, B., Martinell, B., and McCabe, D. (1990). In vivo and in vitro gene transfer to mammalian somatic cells by particle bombardment. Proc Natl Acad Sci U S A *87*, 9568-9572.

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. Nature *289*, 751-758.

Ye, R.W., Tao, W., Bedzyk, L., Young, T., Chen, M., and Li, L. (2000). Global gene expression profiles of Bacillus subtilis grown under anaerobic conditions. J Bacteriol *182*, 4458-4465.

Zaigler, A., Schuster, S.C., and Soppa, J. (2003). Construction and usage of a onefold-coverage shotgun DNA microarray to characterize the metabolism of the archaeon Haloferax volcanii. Mol Microbiol *48*, 1089-1105.

Zhou, M., Xiang, H., Sun, C., Li, Y., Liu, J., and Tan, H. (2004a). Complete sequence and molecular characterization of pNB101, a rolling-circle replicating plasmid from the haloalkaliphilic archaeon Natronobacterium sp. strain AS7091. Extremophiles *8*, 91-98.

Zhou, M., Xiang, H., Sun, C., and Tan, H. (2004b). Construction of a novel shuttle vector based on an RCR-plasmid from a haloalkaliphilic archaeon and transformation into other haloarchaea. Biotechnol Lett *26*, 1107-1113.

Zurawski, G., Gunsalus, R.P., Brown, K.D., and Yanofsky, C. (1981). Structure and regulation of aroH, the structural gene for the tryptophan-repressible 3-deoxy-D-arabino-heptulosonic acid-7-phosphate synthetase of Escherichia coli. J Mol Biol *145*, 47-73.

Curriculum vitae

Name:	Miriam (Mirit) Kolog Gulko
Address:	Käthe Kollwitz 6, 82152, Martinstried, Germany
Place of birth:	Tel-Aviv, Israel
Nationality:	Israeli

Education

From 10/2005-   **PhD**, **Max Planck Institute of Biochemistry**
Department of Membrane Biochemistry, Germany
<u>Supervisor</u>: Prof. Oesterhelt
<u>Title</u>: A non-canonical pathway for aromatic amino acid biosynthesis in Haloarchaea: identification and function of essential genes

1997-2000   **M.Sc**., **Technion- Israel Institute of Technology**
Department of Food Engineering and Biotechnology, Israel
<u>Supervisor</u>: Prof. Yoval Shoham
<u>Title</u>: Biochemical characterization of an intracellular Xylanase from *Bacillus Steraothermophilus* T-6

1993-1996   **B.Tech**, **Judea and Samaria College**
Department of Chemical Engineering, Israel
<u>Supervisor</u>: Prof. Zeev Wolf
<u>Title</u>: Designing a sulfuric acid manufacture factory, with capability to produce 1000 tons per day

1992-1993   **Practical engineering**, **Judea and Samaria College**
Department of Biotechnology, Israel
<u>Supervisor</u>: Dr. Michael Firer
<u>Title</u>: Cultivation of hybridoma cells in a bioreactor and concentration of the released antibodies

Work experience

2002-2005   **Team Leader-Feasibility and assay development**
Procognia Ltd, Israel
2001-2002   **R&D researcher**
Procognia Ltd, Israel
2002-2001   **Lecturer**
Judea and Samaria College, Israel
1998-2000   **Teaching assistant**
Technion Isreal Institute of Technology, Israel

Martinstried, 08/03/2010