Manuscript I

A study of a type II restriction endonuclease from the cold-adapted organism *Psychrobacter arcticus*

Grgic Miriam, Altermark Bjørn, Leiros Ingar

Department of Chemistry, UiT the Arctic University of Norway, Tromsø, Norway

Introduction

DpnI is classified as a type II restriction enzyme and is found in a few prokaryotes. DpnI acts without a methylation counterpart and is a so-called solitary endonuclease. It was first discovered and isolated in the 1980s from the human pathogen *Streptococcus* (*Diplococcus) pneumoniae*, and named DpnI based on its origin (1-3). DpnI is one of the rare endonucleases that cleave methylated DNA. It cleaves DNA at the recognition sequence Gm6A↓TC (where ↓ indicates the cleavage site, m6 indicates methylation on 6^{th} carbon of the adenine ring, in the following termed GATC) when the adenine is methylated on both strands. At high concentrations *in vitro*, DpnI can cleave hemimethylated DNA as well (4). Due to the fact that it cleaves at the recognition site, and not further away, and leaves blunt ends at the cleaving site, it is classified as a type IIM restriction enzyme and is the most studied example from this group (1). Furthermore, DpnI is reported to act as a monomer in solution (2, 4). DpnI is not accompanied by a methyltransferase and it is assumed that its main function is to protect bacteria from phages propagated in Dam^+ (DNA adenine methyltransferases) bacteria hosts (5). This relies on DpnI being specific for the Dam methylated GATC sites. Due to its GATC sequence specificity, DpnI is widely used in several biotechnological applications. Two known examples are application in mutagenesis and in restriction-free cloning (6) utilizing the ability of DpnI to cleave parental DNA whereas newly amplified DNA that is not methylated is not susceptible for restriction (6, 7). Additionally, it is used accompanied with m6A methyltransferases, in the study of of interactions between DNA and DNA binding proteins (8). For example, DpnI can accompany Dam methyltransferase and be used for *in vivo* mapping of binding sites for various transcription factors, such as in DamID, the DNA adenine methyltransferase identification technique (8, 9).

So far, there is only one crystal structure of DpnI (from *S. pneumoniae*) reported while the mechanism of methylation-dependent cleavage is still relatively unknown (10). The DpnI is a two-domain protein, it contains an N-terminal catalytic (PD…(D/E)XK) domain, which is common to many type II restriction enzymes and have a typical fold for this family (11, 12) (Figure 1, left side). The C-terminal (HTH) domain has a winged helix fold that consists of 3 α-helices and a 3-stranded β sheet (Figure 1, right side). Both of these domains can bind DNA independently (5, 10). The DpnI enzyme has also a metal binding motif, where most commonly Mg^{2+} binds, as well as a DNA binding Zn-finger domain (Figure 1) (2, 10).

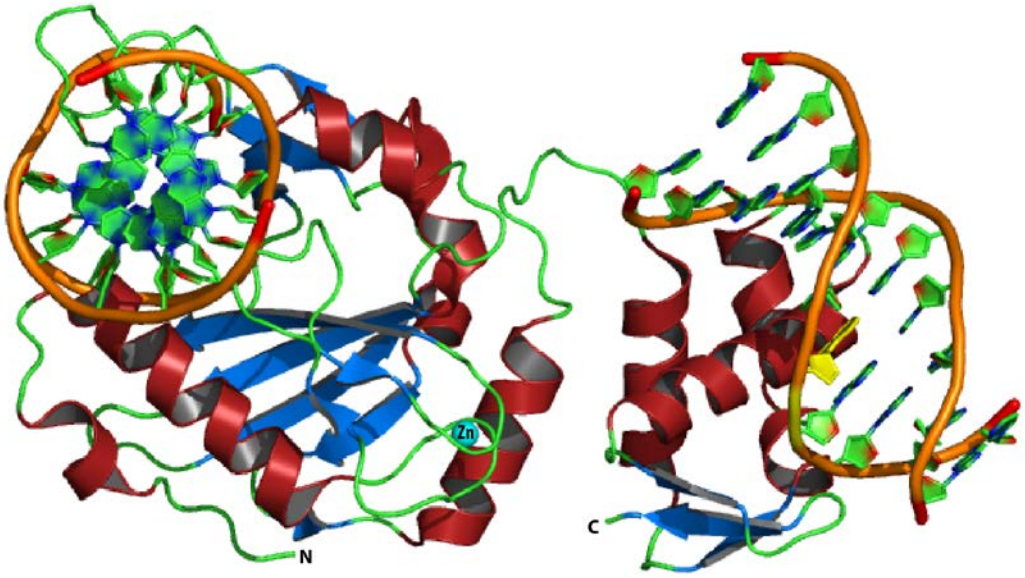

Figure 1. Structure of DpnI from *S. pneumoniae*. (PDB ID: 4KYW). DpnI is a two domain protein with an N-terminal catalytic domain (shown to the left) which contains the (PD-(D/E)XK) motif and a C-terminal (HTH) domain (shown to the right). DNA is presented in orange with bases presented in green and methylated adenine inyellow. One DNA molecule is bound to each domain. Zn^{2+} is presented as a turquoise sphere.

By similarity searches using BLAST it can be noticed that the DpnI enzyme from *S. pneumoniae* shares high sequence similarity with several other enzymes that can be placed in the DRP (dam replacing protein) superfamily consisting of a restriction endonuclease that is flanked by pseudo-transposable small repeat elements (13). These enzymes occur only in Dam-deficient strains (14). Besides being present in both *Neisseria* and *Streptococcus*, there are also DRP proteins found in a few other pathogenic bacteria, *e.g. Treponema* and *Francisella.* Additionally, we found two proteins with high sequence similarity to the DpnI enzyme from *S. pneumoniae* in the bacteria from the genus *Psychrobacter*. Members of this genus include the well-characterized cold-adapted bacterium *P. arcticus* in addition to the mesophilic strain *P.* sp. PRwf-1 (15, 16).

The aim of the project

Originating from the cold adapted organism *P. arcticus*, we hypothesized that this type II DpnI-like enzyme could provide insight into cold adaptation. The three-dimensional structure of the DpnI enzyme from both the psychrophilic *P. arcticus* and the mesophilic *P.* sp. PRwf-1 strains could point to structural differences that might explain differences in the activity and general behavior of the enzyme that can be regarded as adaptation to a cold environment. In addition, we aimed at doing a thorough characterization of the enzymes regarding their activity and specificity profiles and shed some light on the mechanism for modification-dependent restriction. Finally, being putatively salt-tolerant and heat-labile, a DpnI homologue from a psychrophilic, marine organism could have a

potential in applications within biotechnology and molecular biology (17).

Materials and methods

Screening different *Psychrobacter* strains for *dpnC* homologues

In order to identify homologous *dpnC* genes (the gene encoding for DpnI) in six related *Psychrobacter* isolates (Appendix, Table 1), we performed PCR reactions using a set of degenerate primers that were designed based on a specific conserved flanking regions of the DpnI gene (Appendix, primers 1-5, Table 2). Genomic DNA from all strains was isolated using the Wizard Genomic DNA kit (Promega). A PCR was run according to the protocol for Taq polymerase (NEB) using genomic DNA as template and specifically designed degenerate primers in total volumes of 25 µL. The sequences acquired by amplification with degenerate primers were sequenced and used to make a multiple DNA sequence alignment using ESPript 3.x(18).

Codon optimization of the *dpn*C gene from *Psychrobacter* species

Rare codons, ten for the *P. arcticus dpnC* gene and thirteen for the *P.* sp. PRwf-1 *dpnC* gene, were discovered by analysis using the Graphical Codon Usage Analyzer[1]. Rare codons can be overcome by synthetic gene optimization or co-expression of rare codons-tRNA. We decided to purchase the codon-optimized *dpnC* gene from *P. arcticus* from the Mr. Gene Company (now affiliated with Life Technologies Corporation), who delivered the optimized gene in the pDONR221 Gateway vector (Appendix, Table 3).

E. coli strain engineering

In order to be able to produce the DpnI proteins, we had to modify the SCS110 and the BL21 (DE3) *E. coli* strains that were available for recombinant expression (Appendix, Table 1). The λDE3 Lysogenisation kit (Novagen) was used to integrate the bacteriophage T7 polymerase gene (gene1) into the *E. coli* SCS110 strain host chromosome by lysogenisation with the λDE3 prophage. The λDE3 phage was made by inserting the T7 RNA polymerase gene behind the *lacUV5* promoter into the *Bam*HI cloning site of λD69. The lysogenized host, *E. coli* SCS110 (DE3) (Appendix, Table 1) was used to overexpress genes cloned into T7 promoter driven expression vectors. In order to prepare lysogens, *E. coli* was co-infected with three bacteriophages, λDE3, a helper and a selection phage. The helper phage provides the *int* function that helps on integration and excision of the λDE3 phage from the chromosome. The selection phage kills the λDE3 host range mutant that otherwise would be among the surviving cells. The host/phage mixture was incubated at 37 °C for 20 min, which

[1] http://gcua.schoedl.de/

allows the phage to infect the host, after which the mixture was plated on an LB plate and incubated at 37 °C overnight. The successful integration of lysogens to the host chromosome was confirmed by amplification of the inserted gene by using gene1-specific primers (Appendix, Table 2) in a colony PCR modified from Vethanayagam and Flower (19). Individual colonies were used as template in a colony PCR method, which was done following the protocol for Taq polymerase (NEB). The second engineering approach to improve heterologous protein expression was to inactivate the *dam* gene in the *E. coli* BL21 (DE3) strain using the Red/ET Recombination system (Gene Bridges). The Dam protein methylates the DpnI recognition sequence, and as DpnI would digest methylated DNA, the *dam* gene is incompatible with recombinant DpnI expression. The *dam* gene was deleted by directed homologous recombination in strains of *E. coli*, which express phage-derived protein pairs, either RecE/RecT from the Rac prophage or Redα/Redβ from the λ phage. Two 50bp linkers complementary to the insertion site on the chromosome were added to the linear kanamycin resistance gene insert by PCR. *E. coli* BL21 (DE3) was transformed with the expression plasmid pRedET, a plasmid that carries genes for recombination, which was induced by the addition of 10 % (v/v) L-arabinose to the culture. After induction, the PCR product carrying the linker regions was electroporated into the host. Successful deletion of the *dam* gene was verified on LB plates supplemented with kanamycin.

Cloning and recombinant expression of *Psychrobacter* DpnI

To overexpress DpnI protein from *Psychrobacter* species, the *dpnC* gene had to be cloned into a non-methylated plasmid and subsequently propagated and expressed in strains that are lacking the Dam methylase. The non-methylated vector was obtained by propagating the expression vectors pTrc99a and later pDest14 in *E. coli* SCS110 (Agilent Technologies) (Appendix, Table 1), a strain that lacks both Dam and Dcm, the two endogenous DNA methyltransferases. The gene *dpnC* encoding the DpnI was amplified using DpnI gene specific primers (Appendix, Table 2) and genomic DNA from *P. arcticus* and *P.* sp. PRwf-1, following the protocol for Phusion polymerase (Thermo Fisher Scientific). The genes were cloned as both a native construct and as a fusion construct with an N-terminal histidine tag to ease purification. The vector pTrc99a and the insert were both digested with *SalI* and *NcoI* restriction enzymes (NEB) and subsequently ligated with T4 DNA ligase (NEB) following the manufacturer's protocols. The *dpnC* gene was also sub-cloned into the pDest14 vector (primers listed Appendix in Table 2) downstream the T7 promotor using Gateway cloning (Invitrogen); the construct was cloned in fusion with the histidine-tag at the C-terminus of the protein as well as a native construct using Gateway cloning (20). The *E. coli* SCS110 strain was used for propagation of all constructs. Sanger sequencing at UiT sequencing facilities (data not shown) confirmed cloning of all constructs to be successful. For expression of *dpnC* cloned into the pTrc99a vector, we used either the SCS110

strain or *E. coli* DB24, a strain that possess no known *E. coli* MTase genes (Appendix, Table 1) (21). For expression of *dpnC* cloned into pDEST14, we used the engineered strains, SCS110 (DE3) or BL21 (DE3)Δ*dam* (Appendix, Table 1). Vectors harboring the *dpnC* gene were transformed into these cells following a standard transformation protocol (22). All vectors and bacterial strains used are listed in tables 1 and 3 in Appendix, respectively. The recombinant protein was expressed in 1 L cultures of Lysogeny broth (LB) media and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Takara Bio), followed by overnight (16-20h) expression at 20 °C.
Codon optimized genes were sub-cloned into the pDEST14 vector using Gateway cloning and the constructs was transformed into SCS110 (DE3) cells which were used for protein expression using same protocol as for other constructs.

Purification of recombinantly expressed *Psychrobacter* DpnI

After an overnight expression, the cells were harvested by centrifugation at 6000 rpm for 25 min at 4 °C and resuspended in 25 ml of lysis buffer (50 mM Tris HCl, pH 7.5, 250 mM NaCl, 5 mM β-mercaptoethanol (β-me), 10 mM imidazole) containing 1x Complete protease inhibitor cocktail (Roche) and DNaseI (Sigma Aldrich). The cells were disrupted by 30 minutes of sonication (10s on and 10s off, with ~25% output and temperature below 20 °C) by Sonics Vibra-Cell VC 750 sonicator, and centrifuged (Avanti J-26XP centrifuge, Beckman Coulter) at 9000x*g* for 30 min at 4 °C after which the supernatant was used for affinity purification. Constructs with histidine tag were purified using a 5 ml HisTrap crude FF column (GE healthcare) on an ÄKTA Basic purification system. The protein was eluted across a gradient of elution buffer containing 50 mM Tris-HCl pH 7.5, 250 mM NaCl and 500 mM imidazole. For purification of native proteins, a SP sepharose ion exchange column (GE Healthcare) was used on the abovementioned purification system and proteins were eluted across a NaCl gradient in buffer containing 25 mM Tris-HCl pH 8.5, 5 mM β-me, 1 M NaCl. Purified proteins were confirmed by MS/MS and by immunoblot analysis using an antibody against the histidine-tags (Antibodies-online[2])**.**

Purification of native DpnI protein from *Psychrobacter* isolates

In order to verify whether *P. arcticus* translates a functional DpnI enzyme we tried to isolate it from the original organism, *P. arcticus* (Appendix, Table 1). In addition, we wanted to test if similar *Psychrobacter* strains produced functional protein. For this purpose, we chose *P.* sp. SR2003-1-2 isolated at Svalbard. Proteins from both strains were purified following a protocol developed by dr. C. Bakermans (23). The *Psychrobacter* isolates were grown in 600 ml LB at 15 ºC. After 3 days, cells

[2] http://www.antibodies-online.com

were collected by centrifugation at 6000×*g* for 10 min at 4 °C, resuspended in 35 ml ice-cold Buffer A (10 mM Tris-HCl pH 7.4, 0.2 mM $MgCl_2$, 0.2 mM EDTA, 2.0 mM β-me) supplemented with 1x Complete protease inhibitor cocktail (Roche). The cell mixture was lysed by sonication for 30 minutes in 10 second on/off pulses, with ~25% output keeping the temperature below 20 °C. Cell debris was removed by centrifugation at 1000 ×*g* for 45 min at 4 °C and after centrifugation the supernatant was transferred to a clean 50 ml centrifuge tube. To precipitate nucleic acids, NaCl to a final concentration of 0.1 mM NaCl and a final concentration of 1% (v/v) polyethyleneimine were added, mixed gently and removed by centrifugation at 15000-×*g* for 10 min at 4 °C. The supernatant was transferred to a clean 250 ml bottle and 100% saturated $(NH_4)_2SO_4$ was added slowly during gentle mixing at 4 °C for 2h to a final concentration of 70% in order to precipitate proteins, which were collected by centrifugation at 25,000×g for 30 min at 4 °C. The precipitated proteins were resuspended in 10 ml Buffer B (20 mM KH_2PO_4, 0.2 mM $MgCl_2$, 0.2 mM EDTA, 2.0 mM ß-me, pH 7.4) while the precipitated nucleic acid pellet was re-extracted with 10 ml of Buffer A plus 0.6 M NaCl. Nucleic acids were removed by centrifugation 25000 ×*g* for 10 min at 4 °C. The resuspended protein pellet and the supernatant (from the nucleic acid precipitation step) were pooled and dialyzed overnight in a Slide-A-Lyzer dialysis cassette (MWCO 10,000 Da) against 1 L Buffer B, which was exchanged twice. Dialyzed samples were loaded on a heparin column (GE Healthcare) equilibrated with Buffer B using the ÄKTA Basic purification system. The column was washed with Buffer B plus 10% glycerol, and proteins were eluted across a NaCl gradient in Buffer C (Buffer B plus 1.2 M NaCl and 10% glycerol).

Activity assay

To determine the activity of DpnI we used a crude restriction assay. pTrc99A, a vector methylated by endogenous MTases during propagation in *E. coli* DH5α , was purified using the QIAprep Spin Miniprep Kit (Qiagen) and subsequently used as substrate in the activity assay. The non-methylated vector, which was propagated in the MTase negative *E. coli* SCS110, was used as negative control. To confirm the methylation of the pTrc99A vector, we used a commercially available DpnI (NEB) that will restrict the vector at GATC sites once they are adenine methylated. The suggested protocol for the commercial enzyme from NEB was followed: 0.5 μL enzyme (10 units) and 50 ng DNA were incubated in reaction buffer (10 mM Tris pH 7.6, 50mM NaCl, 6 mM $MgCl_2$, 5 mM β-me) in a final volume of 25 μL for 1 h at 37°C. Products from DpnI treatment were analyzed on a 1% agarose gel in 1x Tris-acetate-EDTA (TAE) buffer.

Results and Discussion

Screening for genes encoding DpnI-homologues in different *Psychrobacter* strains

In this experiment, we screened for the *dpnC* gene in six different *Psychrobacter* species (Appendix, Table 1). The gene was present in several members of the genus, with diverse amino acid sequences. The sequence similarity was ~60% and an alignment can be seen in Figure 2. Also, in all homologs *dpnC* genes we could detect part of the PD-(D/E)XK motif, a motif that is conserved among type II REases. Additionally, all genes encode the conserved Zn^{2+} finger motif (Figure 2).

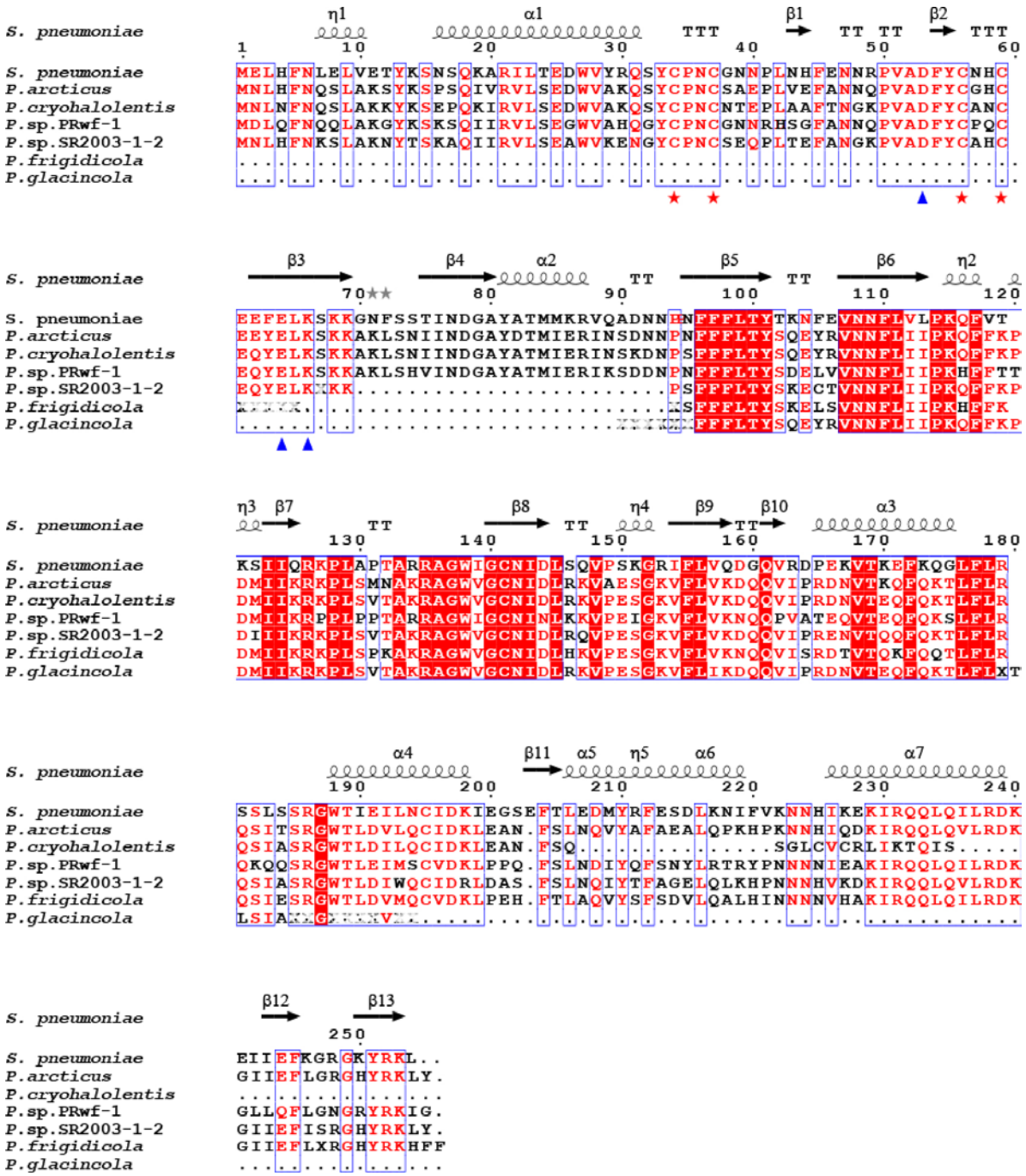

Figure 2. Protein sequence alignment comparing type II DpnI homologs from six *Psychrobacter* species and *S. pneumoniae*. Residues with red background are showing high sequence similarity while residues in blue squares are shoving partial similarity. Secondary structure elements on top are annotated from *S. pneumoniae* DpnI enzyme (PDB ID: 4KYW), where α presents α-helices and β presents β-strands (4). Bottom secondary structure elements are predicted using the protein sequence of *P. arcticus* DpnI and PsiPred (24). Red stars indicate cysteine residues involved in the Zn-finger motif. Blue triangles point to residues that are involved in the PD…(D/E)XK motif. The internal gaps sequenced in the sequence of *P. glacincola* and *P. frigidicola* are due to areas that are not sequenced. The sequences were aligned using T-coffee (25) and the figure was generated using ESPript (18).

We decided to continue working with the dpnC gene from two strains, namely P. arcticus (psychrophilic) and P. sp. PRwf-1 (mesophilic) hoping that a detailed characterization would provide insight into enzymatic mechanisms and structural features of enzymes adapted to different temperature environments (warm and cold).

Recombinant expression of native and his-tagged *Psycrobacter arcticus* DpnI in *E. coli* SCS110

To ease purification, a polyhistidine (his) tagged version of the DpnI protein was designed. In case the polyhistidine tag would disrupt proper protein folding or activity, a native version of the DpnI protein was designed as well. The *dpnC* gene was successfully cloned in frame with an N-terminal his-tag in the pTrc99a vector using restriction-ligation cloning. The N-terminally his-tagged protein was purified by affinity chromatography (Figure 3A) and the recombinant expression of the tagged protein was confirmed by immunoblot analysis using an antibody against the N-terminal histidine-tag (Figure 3B). Native recombinant protein was purified using ion exchange chromatography. Both recombinant proteins were confirmed by mass spectrometry analysis (MS/MS). Despite the non-optimal expression system, we managed to produce the recombinant protein in small amounts (Figure 3 and 4), which was sufficient for preliminary characterization experiments. The recombinant protein was, however, not pure enough for crystallization screens (Figure 3A), which was one of our aims of this project.

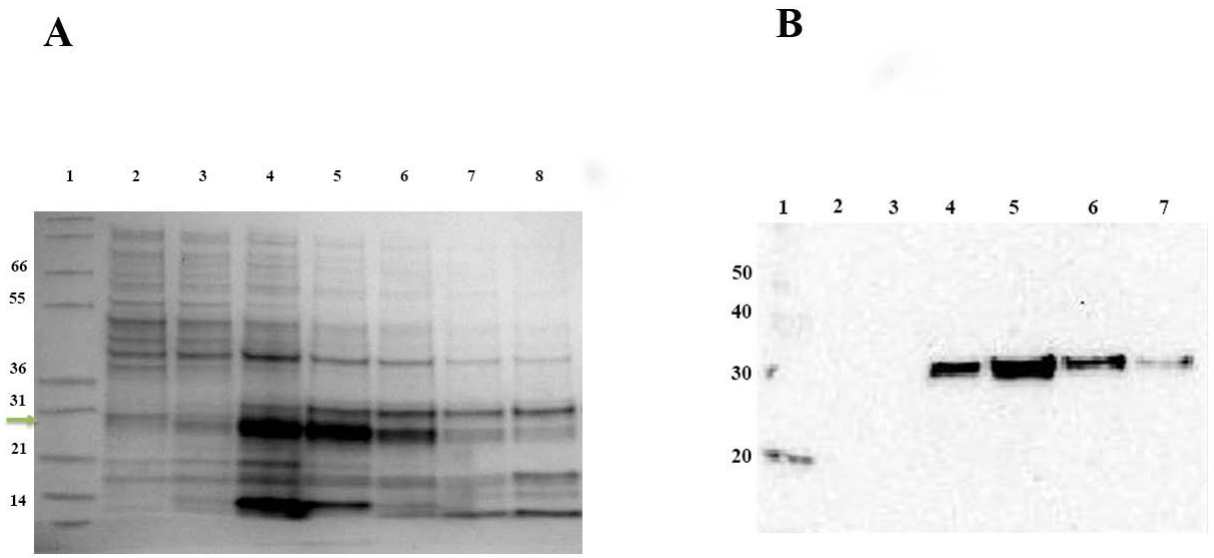

Figure 3. Purification of the recombinant *P. arcticus* DpnI protein with N-terminal histidine tag. A. An SDS-PAGE gel showing the IMAC elution fractions of recombinant DpnI with N-terminal histidine tag. The green arrow points to fractions with DpnI protein (lanes 4-6). Lanes 2 and 3 contain flow through and washing fraction from purification.**B. Immunoblot analysis of IMAC elution fractions using antibodies against the N-terminal his-tag.** Lanes2 and 3 contain flow through and washing fractions from purification, while lanes 4-6 contains selected fractions from elution gradient that contain DpnI.The theoretical size of the protein is 29.8 kDa, calculated using ProtParam tool[3].

[3] http://web.expasy.org/protparam/

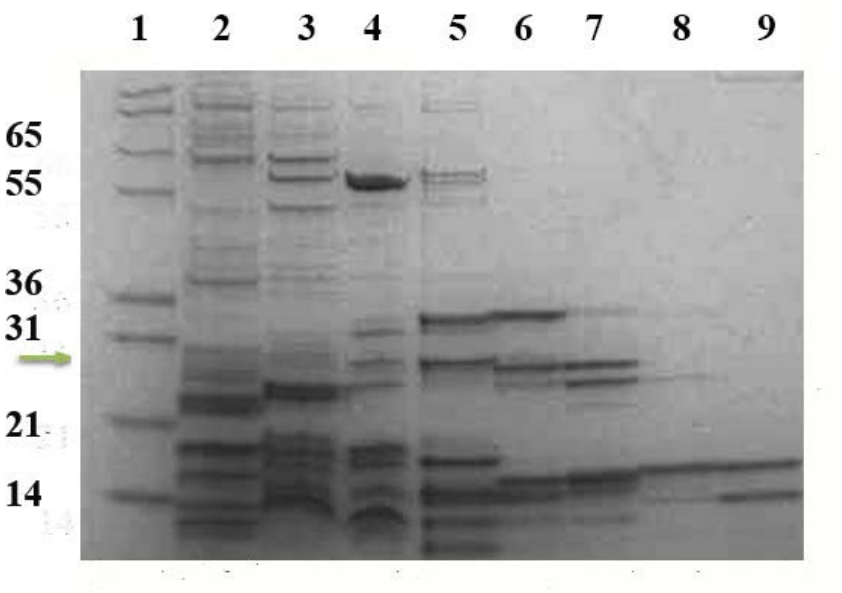

Figure 4. SDS-PAGE gel showing fractions of DpnI eluting from the ion exchange column in purification of the native recombinant DpnI protein. Green arrow points to fractions with DpnI protein (lanes 4, 5 and 6) based on size (DpnI is 29.8 kDa). Mark12 was used as molecular weight marker (lane 1).

The synthesized and *E. coli* optimized *dpnC* genes were sub- cloned into the pDest14 vector using Gateway cloning. The constructs were transformed into *E. coli* SCS110 (DE3) but only preliminary test expression, which was not satisfactory, was done before the project was terminated.

Recombinant expression of *Psychrobacter arcticus* DpnI in engineered strains

The system we set up for recombinant expression, expressing DpnI from the pTrc99a vector in the *E. coli* SCS110 strain, did not overexpress recombinant proteins in levels typically expected for the *E. coli* BL21/pET system (26). The recombinant expression level and yield of soluble DpnI was generally observed as low and significantly contaminated with endogenous proteins from the host (Figure 3A and 4). In order to improve our expression systems, that would enable us to produce higher quantities of pure protein we engineered two *E. coli* strains that were used in protein production, SCS110 and BL21 (DE3), respectively. One of the two approaches we set up was to modify the *E. coli* SCS110 expression host by introducing the lambda bacteriophage gene-1 encoding T7 RNA polymerase behind the *lacUV*5 promoter, and thus generating the engineered *E. coli* strain SCS110 (DE3). This system would allow a T7 driven expression from any vector encoding a T7 promoter upstream of the gene of interest. Our second approach was to inactivate the *dam* gene in the BL21 (DE3) expression strain, giving an engineered *E. coli* BL21 (DE3) Δ*dam.* This approach would allow an already T7 optimized expression host to be unaffected by the methyl-adenine restriction activity upon induction of functional, recombinant DpnI. The strain engineering was carried out by λDE3 lysogenization (Figure 5) and Red/ET recombination (data not shown), respectively. Both strain engineering approached were successful.

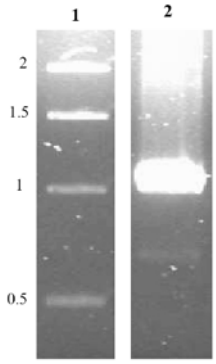

Figure 5. Confirming the presence of gene-1 inserted into the host chromosome of the engineered *E. coli* SCS110 (DE3) strain. The verification of the insertion of the lambda bacteriophage gene-1, encoding the T7 RNA polymerase, to the host chromosome was done by PCR using specific primers complementary to internal regions of gene-1. The theoretical size of the product was 1100 bp. The black arrow is pointing to the product that corresponds to the expected theoretical size. Lane 1 is 1 Kb DNA ladder and in lane 2 is PCR product.

We managed to express the recombinant *P. arcticus* DpnI fused to a C-terminal His-tag in this modified expression systems as well (Figure 6). However, amounts expressed were not sufficient, contamination from *E. coli* proteins was still substantial, and due to that, this approach was aborted. The native construct from *P. arcticus* cloned into pDEST14 was also transformed to *E. coli* SCS110 (DE3). Despite optimization trials, we did not observe improvements when it came to the expression levels or protein stability. The protein was quite unstable and significant precipitation was observed after storing the protein for 48h at 4 °C. Additional optimization of the purification conditions were not carried out.

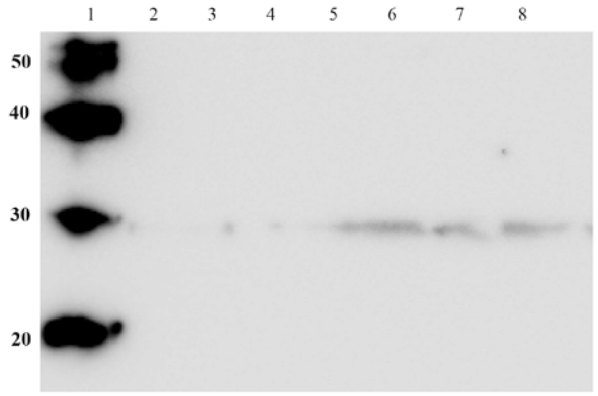

Figure 6. Immunoblot analysis of IMAC elution fractions using antibodies against the C-terminal his-tag. The *dpnC* gene was cloned into pDEST14 and subsequently transformed into BL21(DE3)Δ*dam*. In lane 1 the MagicMark™ XP Western Protein Standard (Invitrogen) was used as molecular weight marker. Lanes 2 and 3 contain fractions from flow through and washing steps of IMAC purification, respectively. Lanes 5-8 contains fractions from the elution gradient. The theoretical size of the DpnI protein is 29, 8 kDa, and is indicated with an arrow.

Isolation of DpnI from *P. arcticus* and *P.* sp. *SR-2003-1-2*

Our final approach for obtaining DpnI protein for characterization studies was to isolate the enzyme from the host of origin. Two strains were available for cultivation, and we proceeded with isolation of the DpnI from both, *P. arcticus* (data not shown) and *Psychrobacter* sp. SR-2003-1-2 from Svalbard (Figure 7) using heparin affinity purification. Isolation of DpnI from both organisms was successful as we managed to detect activity in fractions containing DpnI (Figure 9). However, it was not possible to measure the exact amount of DpnI due to a substantial amount of *Psychrobacter* proteins that co-purified with it. Further purification steps were not attempted due to the low yield of DpnI.

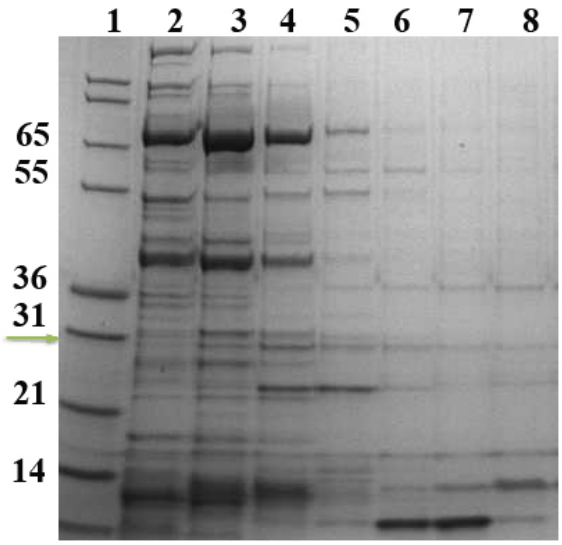

Figure 7. Purification of native DpnI isolated from *Psychrobacter* sp. SR-2003-1-2. SDS-PAGE analysis showing fractions from the NaCl gradient in heparin purification (lanes 1-8). Fractions that were assumed to contain DpnI based on the size of the protein (~30 kDa,lanes 3-5) are marked by green arrow. These fractions were used in activity assay. Mark 12 was used as molecular weight (MW) marker.

Activity assay

Recombinant and native DpnI isolated from *Psychrobacter* sp. SR2003-1-2 were subjected for activity testing in a crude restriction assay. The N-terminally his-tagged DpnI protein expressed in *E. coli* SCS110 did not show any activity in a restriction assay regardless of experimental conditions, such as variable protein concentration, additives or temperature and incubation time (Figure 8).

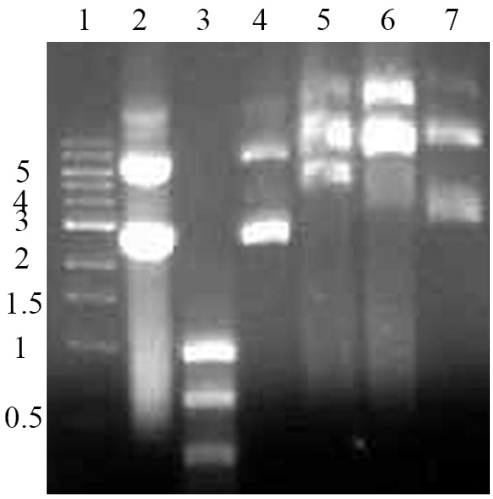

Figure 8. Crude restriction assay with recombinant DpnI from *P. arcticus*. Lane 1, 1Kb DNA ladder; lane 2, non-methylated pTrc99A (negative control), lane 3, methylated pTrc99A plasmid treated with commercial DpnI (positive control); lanes 4-7,methylated pTrc99A plasmid treated with different elution fractions from purification that contained recombinant N-terminally histagged DpnI from *P. arcticus*. Bands above 3Kb represent different states of the plasmid:: relaxed, supercoiled and circular.

We hypothesized that the N-terminal histidine-tag was interfering with the protein folding and/or is blocking the active sites for substrate or cofactor. To try to overcome these problems we investigated the native version of DpnI expressed from pDEST14 in *E. coli* SCS110 (DE3) in the crude restriction assay. When the methylated pTrc99A vector, used as a substrate for restriction, was treated with the purified, recombinant DpnI, catalytic activity was observed (Figure 9A). The resulting cleavage pattern complied with the expected cleavage pattern as compared to the pattern found from treating methylated pTrc99A with the commercial DpnI. Both recombinantly expressed native DpnI from *P. arcticus* and commercial DpnI digested the vector into several fragments, with two main signature fragments -between 500-1000 bp while the rest were smaller than 500 bp (Figure 9A). In addition, some elution fractions from heparin purification containing DpnI isolated from the *Psychrobacter* sp. SR2003-1-2 showed weak activity in the same assay (Figure 9B). These data might indicate that the small amount of DpnI isolated from *Psychrobacter* sp. SR2003-1-2 was active. It cannot be ruled out, however, that other co-purified proteins in *Psychrobacter* sp. SR2003-1-2 contain restriction activity.

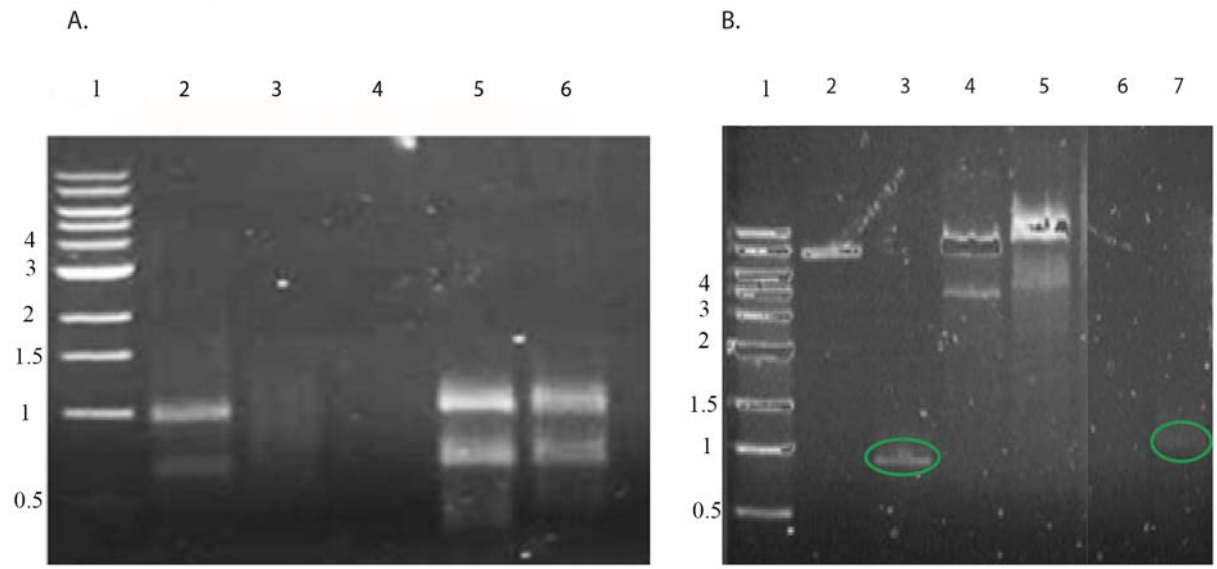

Figure 9. Activity assay for recombinant DpnI from *P. arcticus* and isolated native DpnI from *Psychrobacter* sp. SR2003-1-2. A. Agarose gel from assay with the recombinant native DpnI from ***P. arcticus***. Lane 1, 1Kb DNA ladder; lane 2, methylated pTrc99A plasmid treated with the commercial DpnI (positive control); lanes 3-6, methylated pTrc99A plasmid treated with native *P. arcticus* DpnI from elution fractions of IMAC purification. **B. Agarose gel from assay with native DpnI isolated from *P.* sp. SR 2003-1-2**. Lane 1 1Kb DNA ladder; lane 2, non-methylated vector (negative control); lane 3, methylated vector treated with commercial DpnI (positive control); lanes 4-7, methylated vector treated with DpnI from elution fractions from heparin purification. Blue arrow points to expected band after restriction.

Conclusion and considerations for future work

In the end, despite having purification fractions, both from recombinant overexpression and from native isolation, containing active DpnI protein, we found the yields to be too low and/or too contaminated with host proteins to continue with functional and structural characterization. In addition, we experienced that the purified proteins were unstable and showed a tendency to precipitate shortly after purification. One plausible explanation for the low expression levels and generally low stability of these proteins may be the fact that they are cold-adapted. Generally, cold - adapted proteins have an intrinsic instability, enabling them to cope with the high flexibility required for activity at low temperatures. Due to the increased flexibility compared to the mesophilic and thermophilic homologues, cold-adapted proteins are also exhibiting low thermal stability (27). After a long-term effort in vector and host optimization for expression and purification, we decided to terminate the project.

In this regard, it should also be mentioned, that according to a recent publication regarding? *S. pneumoniae* recombinant DpnI was produced in sufficient amounts for crystallization studies (4, 5). However, similar problems in both expression and purification of DpnI were encountered, such as lacking a DpnI-optimized expression system. The authors eventually managed to overcome their problems with expression by adding a cleavage site after the N-terminal his-tag for tag-removal as well as adding more steps in protein purification (4). Based on their success, we believe there are

some additional tests that could be performed in order to increase protein expression such as designing and cloning alternative recombinant constructs. Other alternative approached to improve expression include using removable solubility tags (*e.g.* maltose binding protein (MBP) or the sumo fusion partner or cloning of the *dpnC* gene downstream a cold-shock promoter for low-temperature expression (28). Alternative expression systems such as yeast or cell-free protein expression could also be used in order to further explore recombinant expression of DpnI as a way of avoiding methylation of foreign, recombinant DNA.

Acknowlegments

We are grateful to Gro Elin Kjæreng Bjerga for her generous help with writing the manuscript.

Reference

1. Roberts, R. J., Belfort, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S. K., Dryden, D. T. F., and Dybvig, K. (2004) A Nomenclature for Restriction Enzymes, DNA Methyltransferases, Homing Endonucleases, and Their Genes. In *Restriction Endonucleases* pp. 1-18, Springer
2. De la Campa, A. G., Springhorn, S. S., Kale, P., and Lacks, S. A. (1988) Proteins encoded by the DpnI restriction gene cassette. Hyperproduction and characterization of the DpnI endonuclease. *Journal of Biological Chemistry* **263**, 14696-14702
3. Lacks, S., and Greenberg, B. (1975) A deoxyribonuclease of Diplococcus pneumoniae specific for methylated DNA. *Journal of Biological Chemistry* **250**, 4060-4066
4. Siwek, W., Czapinska, H., Bochtler, M., Bujnicki, J. M., and Skowronek, K. (2012) Crystal structure and mechanism of action of the N6-methyladenine-dependent type IIM restriction endonuclease R. DpnI. *Nucleic acids research* **40**, 7563-7572
5. Mierzejewska, K., Siwek, W., Czapinska, H., Kaus-Drobek, M., Radlinska, M., Skowronek, K., Bujnicki, J. M., Dadlez, M., and Bochtler, M. (2014) Structural basis of the methylation specificity of R.DpnI. *Nucleic Acids Research*
6. Van Den Ent, F., and Löwe, J. (2006) RF cloning: a restriction-free method for inserting target genes into plasmids. *Journal of biochemical and biophysical methods* **67**, 67-74
7. Weiner, M. P., Costa, G. L., Schoettlin, W., Cline, J., Mathur, E., and Bauer, J. C. (1994) Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. *Gene* **151**, 119-123
8. Jeltsch, A., Jurkowska, R. Z., Jurkowski, T. P., Liebert, K., Rathert, P., and Schlickenrieder, M. (2007) Application of DNA methyltransferases in targeted DNA methylation. *Applied microbiology and biotechnology* **75**, 1233-1240
9. Vogel, M. J., Peric-Hupkes, D., and van Steensel, B. (2007) Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nat. Protocols* **2**, 1467-1478
10. Siwek, W., Czapinska, H., Bochtler, M., Bujnicki, J. M., and Skowronek, K. (2012) Crystal structure and mechanism of action of the N6-methyladenine-dependent type IIM restriction endonuclease R. DpnI. *Nucleic acids research* **40**, 7563-7572
11. Pingoud, A., Fuxreiter, M., Pingoud, V., and Wende, W. (2005) Type II restriction endonucleases: structure and mechanism. *Cellular and molecular life sciences* **62**, 685-707
12. Kosinski, J., Feder, M., and Bujnicki, J. M. (2005) The PD-(D/E) XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. *BMC bioinformatics* **6**, 172
13. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410
14. Cantalupo, G., Bucci, C., Salvatore, P., Pagliarulo, C., Roberti, V., Lavitola, A., Bruni, C. B., and Alifano, P. (2001) Evolution and function of the neisserial *dam*-replacing gene. *FEBS letters* **495**, 178-183
15. Bowman, J. P. (2006) The genus Psychrobacter. In *The Prokaryotes* pp. 920-930, Springer
16. Ayala-del-Río, H. c. L., Chain, P. S., Grzymski, J. J., Ponder, M. A., Ivanova, N., Bergholz, P. W., Di Bartolo, G., Hauser, L., Land, M., and Bakermans, C. (2010) The genome sequence of Psychrobacter arcticus 273-4, a psychroactive Siberian permafrost bacterium, reveals mechanisms for adaptation to low-temperature growth. *Applied and environmental microbiology* **76**, 2304-2312
17. Cavicchioli, R., Charlton, T., Ertan, H., Omar, S. M., Siddiqui, K., and Williams, T. (2011) Biotechnological uses of enzymes from psychrophiles. *Microbial biotechnology* **4**, 449-460
18. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**, W320-W324
19. Vethanayagam, J., and Flower, A. (2005) Decreased gene expression from T7 promoters may be due to impaired production of active T7 RNA polymerase. *Microbial Cell Factories* **4**, 3

20. Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (2000) [34] GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes. *Methods in enzymology* **328**, 575-IN577
21. Kong, H., Lin, L.-F., Porter, N., Stickel, S., Byrd, D., Posfai, J., and Roberts, R. J. (2000) Functional analysis of putative restriction modification system genes in the Helicobacter pylori J99 genome. *Nucleic acids research* **28**, 3216-3223
22. Services, S. T. (2004) SCS110 Competent Cells.
23. Bakermans, C., Sloup, R. E., Zarka, D. G., Tiedje, J. M., and Thomashow, M. F. (2009) Development and use of genetic system to identify genes required for efficient low-temperature growth of Psychrobacter arcticus 273-4. *Extremophiles* **13**, 21-30
24. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *Journal of molecular biology* **292**, 195-202
25. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-217
26. Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113-130
27. Smalås, A. O., Leiros, H. K., Os, V., and Willassen, N. P. (2000) Cold adapted enzymes. *Biotechnol Annu Rev* **6**, 1-57
28. Qing, G., Ma, L.-C., Khorchid, A., Swapna, G., Mal, T. K., Takayama, M. M., Xia, B., Phadtare, S., Ke, H., and Acton, T. (2004) Cold-shock induced high-yield protein production in Escherichia coli. *Nature biotechnology* **22**, 877-882

Appendix

Table 1. Bacterial strains used in this study.

Strain	Application	Genotype	References
Psychrobacter arcticus	Amplification of *dpnC* gene and expression of native DpnI protein		Dr. C. Bakermans and DSMZ, Germany
P.cryohalolentis	Screening for homolog of *dpnC* gene		Dr. C. Bakermans
P. sp. PRwf-1	Screening for homolog of *dpnC* gene and amplification of *dpnC* gene		Dr. C. Bakermans
P. sp. SR2003-1-2	Screening for homolog of *dpnC* gene and expression of native DpnI protein		Dr. B. Landfald
P. glacincola	Screening for homolog of *dpnC* gene		DSMZ, Germany
P. frigidicola	Screening for homolog of *dpnC* gene		DSMZ, Germany
E. coli SCS110	Propagation of nonmethylated vectors and recombinant expression of *dpnC*	*rpsL* (Strr) *thr leu endA thi-1 lacY galK galT ara tonA tsx dam dcm supE44* Δ (*lac-proAB*) [F´ *traD36 proABlacI*qZ Δ *M15*].	Agilent Technologies
E. coli SCS110 (DE3)	Recombinant expression of *dpnC*	*rpsL* (Strr) *thr leu endA thi-1 lacY galK 18al Tara tonA tsx dam dcm supE44* Δ (*lac-proAB*) [F´ *traD36 proAB lacI*qZ Δ *M15*] λ (DE3[lacI lacUV5-T7 gene 1 ind1 sam7 nin5]).	This study
E. coli BL21(DE3) Δ*dam*	Recombinant expression of *dpnC*	F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ (DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) Δdam	This study
E .coli DB24 (ER 2796)	Recombinant expression of *dpnC*	*dam-16::Kan trp-31 his-1 fhuA2 rpsL104 D(lacZ)r1 glnV44 xyl-7 mtl-2 metR1 mcr-62 argG6 D(mcrB-hsd-mrr)114 dcm-6 zed-501::Tn10*]. *Escherichia coli* ER2502 [*F– l– fhuA2 ara-14 leu* Δ *(gpt-proA)62 lacY1 glnV44 galK2 rpsL20endA1R(zgb210::Tn10)Tet S xyl-5 mtl-1* Δ *(mcrC-mrr)HB101*] and ER2566 {*F– l– fhuA2 [lon] ompT lacZ::T7 gene1 gal sulA11* Δ *(mcrC-mrr)114::IS10 R(mcr-73::miniTn10–*Tet^S*)2 R(zgb-210::Tn10–*Tet^S*) endA1 [dcm]*}	Dr. R. Roberts

Table 2. Primers used in this study. Gene-specific sequences are in bold.

	Primer name	Sequence (5'-3')	Description
1	BA_DpnI_Deg1F	AAATGTGAGTTYTATAGCTGTATG[a]	Degenerate forward primer
2	BA_DpnI_Deg2F	ATCATRGTGKCATAAGCRCCATC	Degenerate forward primer, binds the anti-sense strand of the *dpnC*-gene
3	BA_DpnI_Deg1R	CATACCATARATAATGGTAGGATC[a]	Degenerate reverse primer
4	BA_DpnI_Deg2R	ATAATGGTAGGATCCGTCTG	Degenerate reverse primer
5	BA_DpnI_Deg3R	GATGGYGCTTATGMCACYATGAT	Degenerate reverse primer, binds sense-strand of the *dpnC*-gene
6	BA_PsArc_F3	CATGCCATGGCT*CACCACCACCACCACC*AT AATCTACATTTTAATC[b]	Forward primer for cloning of the *dpnC* genes from *P. arcticus* to pTrc99A
7	BA_PsArc_R1	CATGGTCGACTTAATATAATTTTCGATAAT GTCCACG[b]	Primer for cloning of the *dpnC* genes from *P. arcticus* to pTrc99A
8	BA_PsSp_F3	CATGCCATGGCT*CACCACCACCACCACC*AT GATTTACAGTTTAATC[b]	Primer for cloning of the *dpnC* genes from *P.* sp. PRwf-1 to pTrc99A
9	BA_PsSp_R1	CATCGTCGACTCAGCCTATTTTGCGATACC GCCCATT[b]	Primer for cloning of the *dpnC* genes from *P.* sp. PRwf-1 to pTrc99A
10	P.arct-Forw1	GGAGATAGAACC**ATGAATCTACATTTTAA TCAAAGT**	Primer for Gateway cloning of the *P. arcticus* to pDest14
11	BA-Adapt-CHIS-Forw2	GGGGACAAGTTTGTACAAAAAAGCAGGCT TCGAAGGAGATAGAACC[c]	Adapter primers are primers with *att* sites using in Gateway cloning to pDest14
12	P.arct-Rev1	GAAAGCTGGGTG**TTAATATAATTTTCGAT AATGTCC**	Primers for Gateway cloning of the *P. arcticus* to pDest14
13	BA-Adapt-NHIS-Rew2	GGGGACCACTTTGTACAAGAAAGCTGGGT GTTA[c]	Adapter primers are primers with *att* sites using in Gateway cloning to pDest14
14	P.arct-CHIS-R1	TTAGTGGTGGTGGTGGTGGTG**ATATAATTT TCGATAATGTCCACG**	Primers for Gateway cloning of the *P. arcticus* to pDest14
15	BA-Adapt-CHIS-Rew2	GGGGACCACTTTGTACAAGAAAGCTGGGT CTTAGTGGTGGTGGTGGTGGTG[c]	Adapter primers are primers with *att* sites using in Gateway cloning to pDest14
16	P.spPrwf1-Forw1	GGAGATAGAACCTCAGCCTATTTTGCGATA CCGCCC	Primers for Gateway cloning of the *P.* sp. PRwf-1 to pDest14
17	P. spPrwf1-Rev1	GAAAGCTGGGTG **ATG GAT TTA CAG TTT AAT CAA GAC**	Primers for Gateway cloning of the *P.* sp. PRwf-1 to pDest14
18	P. spPrwf1-CHIS-Rev1	ATGGTGGTGGTGGTGGTGG **GAT TTA CAG TTT AAT CAA CAG**	Primers for Gateway cloning of the *P.* sp. PRwf-1 to pDest14
19	t7pol1	GATTAACATCGCTAAGAACG	Primers for amplification of integrated gene1 into SCS110 *E. coli* genome
20	t7pol2	GATTCATGTCGATGTCTTCC	Primers for amplification of integrated gene1 into SCS110 *E. coli* genome

a) Y means pyrimidines (C or T) and R means purines (A or G).
b) *NcoI* and *SalI* restriction sites are underlined and polyhistidine-encoded regions are italicized.

Table 3. Vectors used in this study.

Vector	Application
pDONR221	Gateway® donor vector with attP1 and attP2 sites and a kanamycin resistance marker
pTrc99A	Bacterial expression vector with inducible lacI promoter, ampicillin resistance marker; restriction enzyme cloning
pDEST14	Gateway® destination vector for inducible high-level expression of proteins in bacterial cells with ampicillin resistance marker
pRedET (kan)	Helper vector; Red/ET expression plasmid. Used for deletion of *dam* gene
FRT-PGK-gb2-neo-FRT template DNA	PCR-template (plasmid DNA) for generating a FRT-flanked PGK-gb2-neo cassette. Used for deletion of *dam* gene