Eur. J. Biochem. **267**, 3502–3512 (2000) © FEBS 2000

A DNA ligase from the psychrophile *Pseudoalteromonas haloplanktis* gives insights into the adaptation of proteins to low temperatures

D. Georlette[1], Z. O. Jónsson[2,3], F. Van Petegem[4], J.-P. Chessa[1], J. Van Beeumen[4], U. Hübscher[2] and C. Gerday[1]

[1]*Laboratory of Biochemistry, Institute of Chemistry, B6a Université de Liège, Sart-Tilman, Belgium;* [2]*Institute of Veterinary Biochemistry, Universität Zürich-Irchel, Zürich, Switzerland;* [3]*Department of Pathology, Brigham & Women's Hospital, Boston, MA, USA;* [4]*Laboratory for Protein Biochemistry and Protein Engineering, Department of Biochemistry Physiology and Microbiology, Gent, Belgium*

The cloning, overexpression and characterization of a cold-adapted DNA ligase from the Antarctic sea water bacterium *Pseudoalteromonas haloplanktis* are described. Protein sequence analysis revealed that the cold-adapted *Ph* DNA ligase shows a significant level of sequence similarity to other NAD^+-dependent DNA ligases and contains several previously described sequence motifs. Also, a decreased level of arginine and proline residues in *Ph* DNA ligase could be involved in the cold-adaptation strategy. Moreover, 3D modelling of the N-terminal domain of *Ph* DNA ligase clearly indicates that this domain is destabilized compared with its thermophilic homologue. The recombinant *Ph* DNA ligase was overexpressed in *Escherichia coli* and purified to homogeneity. Mass spectroscopy experiments indicated that the purified enzyme is mainly in an adenylated form with a molecular mass of 74 593 Da. *Ph* DNA ligase shows similar overall catalytic properties to other NAD^+-dependent DNA ligases but is a cold-adapted enzyme as its catalytic efficiency (k_{cat}/K_m) at low and moderate temperatures is higher than that of its mesophilic counterpart *E. coli* DNA ligase. A kinetic comparison of three enzymes adapted to different temperatures (*P. haloplanktis*, *E. coli* and *Thermus scotoductus* DNA ligases) indicated that an increased k_{cat} is the most important adaptive parameter for enzymatic activity at low temperatures, whereas a decreased K_m for the nicked DNA substrate seems to allow *T. scotoductus* DNA ligase to work efficiently at high temperatures. Besides being useful for investigation of the adaptation of enzymes to extreme temperatures, *P. haloplanktis* DNA ligase, which is very efficient at low temperatures, offers a novel tool for biotechnology.

Keywords: NAD^+-dependent DNA ligase; psychrophile; thermophile; structural comparison; overexpression.

Previous studies of enzymes adapted to cold have revealed that they are characterized by a higher catalytic efficiency (k_{cat}/K_m) than their mesophilic counterparts at low and moderate temperatures, and by a greater susceptibility to thermal denaturation (reviewed in [1–3]). Both these properties appear to be due to a more flexible protein structure [4]. This has been corroborated by recent crystallographic data from the psychrophilic *Alteromonas haloplanktis* α-amylase [5], DS2-3R citrate synthase [6], *Vibrio marinus* triose-phosphate isomerase [7] and *Aquaspirillium arcticum* malate dehydrogenase [8]. Intracellular enzymes are less exposed to external conditions than extracellular enzymes and thus constitute an attractive tool for the investigation of adaptation to cold. The fact that several mesophilic and thermophilic DNA ligases are available, and have been studied biochemically in some detail, prompted us to isolate a psychrophilic DNA ligase and compare these three classes of enzymes side by side. In this paper we report the isolation and characterization of a novel NAD^+-dependent DNA ligase from the psychrophilic bacterium *Pseudoalteromonas haloplanktis* (*Ph*) and compare its characteristics to those of *Escherichia coli* (*Ec*) and *Thermus scotoductus* (*Ts*) DNA ligases.

DNA ligases are enzymes required for important cellular processes such as DNA replication, DNA recombination and DNA repair. They catalyse the formation of a phosphodiester bond between adjacent 5′-phosphoryl and 3′-hydroxyl groups in double stranded DNA. The reaction mechanism can be divided into three distinct catalytic steps. First, the ligase is activated through the formation of a covalent adenylated intermediate by transfer of the adenyl group of NAD^+ or ATP to the ε-NH_2 of a lysine residue in the DNA ligase. Second, the AMP moiety is transferred from the DNA ligase to the 5′-phosphate group at the single-strand break site, creating a new pyrophosphate bond. Third, the phosphodiester bond formation is achieved through an attack of the 3′-OH group of the DNA on the activated 5′-group, with the concomitant release of AMP [9].

The importance of DNA ligases is illustrated by the fact that mammalian cells contain five DNA ligases (I, II, III_α, III_β and IV) (reviewed in [10,11]). All the eukaryotic DNA ligases, as well as those from bacteriophages, archaea and viruses, utilize ATP as a cofactor [12–15]. Most eubacteria, however, contain only a single DNA ligase that is NAD^+-dependent [16–21], although genomic sequencing projects have revealed genes encoding ATP-dependent DNA ligases in *Haemophilus influenzae* [22] and *Aquiefex aeolicus* [23,24]. These

Correspondence to C. Gerday, Laboratory of Biochemistry, Institute of Chemistry, B6a Université de Liège, B-4000 Sart-Tilman, Belgium. Fax: + 32 43663364; Tel.: + 32 43663340; E-mail: ch.gerday@ulg.ac.be

Abbreviations: *Ph*, *Pseudoalteromonas haloplanktis*; *Ec*, *Escherichia coli*; *Bs*, *Bacillus stearothermophilus*; *Ts*, *Thermus scotoductus*; RF-C, replication factor C.

Note: the *Pseudoalteromonas haloplanktis* DNA ligase nucleotide sequence reported in this paper has been deposited in the GenBank database under accession number AF126866.

(Received 2 February 2000, revised 5 April 2000, accepted 6 April 2000)

eubacterial ATP-dependent DNA ligases may be derived from integration of viral genes into the bacterial genomes.

NAD^+-dependent DNA ligases consist of at least two biochemically independent domains [18]. The N-terminal domain contains a KXDG box in which the lysine residue becomes adenylated, and a universally conserved NPRNAAAGS motif [20]. The KXDG box is the only conserved motif among NAD^+-dependent and ATP-dependent DNA ligases. The C-terminal domain contains four conserved cysteine residues that form a zinc-finger-like motif which could be implicated in the recognition of nicked DNA [18] and a DNA binding domain showing sequence similarity to the DNA binding domain of the eukaryotic replication factor C (RF-C) large subunit [18–20,25]. To date, several NAD^+-dependent DNA ligases from mesophilic (*Ec* [26–29]) and thermophilic *Bacillus stearothermophilus* (*Bs*) [18–30], *Rhodothermus marinus* [20–31], *Thermus filiformis* [21], *Ts* [20] and *Thermus thermophilus* [19–37]) bacteria have been characterized to various extents. In addition several NAD^+-dependent *lig* (gene encoding DNA ligase) genes have been identified in bacterial genome sequencing projects. The sizes of the deduced proteins from these *lig* genes are similar and they all show strong sequence similarity.

MATERIALS AND METHODS

Sources of strains and enzymes

The strain *Ph* TAE 72 was isolated from Antarctic sea water at the Dumont d'Urville Antarctic Station; *Ec* DNA ligase was from New England Biolabs (205 S); *Ts* DNA ligase was a kind gift from G. Eggertsson (Reykjavik, Iceland) [20].

Gene cloning and sequencing of *Ph* DNA ligase

The psychrophilic DNA ligase gene was isolated by PCR screening. Two degenerate oligonucleotides (5′-GGIGCIACIC-GIGGIGACGG-3′ and 5′-CCIGCIGCIGCGTTIGCIGGGTT-3′) were designed on the basis of an alignment of several amino-acid sequences from bacterial DNA ligases [20]. A PCR reaction, performed on genomic DNA using these two oligonucleotides, amplified, as expected, a 220-bp fragment. This fragment was cloned into a PCRScript vector (Stratagene) and sequencing revealed 65% identity with the *Ec* DNA ligase gene (BESTFIT comparison). Next, a genomic library was constructed by cloning *Hin*dIII-digested TAE 72 genomic DNA into a pUC19 vector. XL1-blue cells (Stratagene) were transformed with the library. All the clones were spread on 20 agar plates, each containing 100 clones. Colonies scraped off replica agar plates were pooled, and the plasmid DNA was isolated and screened by PCR: out of the 20 original plates, four plates were positive for the presence of the DNA ligase gene. Each positive plate was then divided onto five plates, each of them containing 20 clones. The pooled colonies of each replica plate were still submitted to a PCR reaction which identified four positive plates (80 clones). All replica clones from these plates were then tested by PCR, resulting in the identification of four positive clones. Restriction analysis showed that they were identical. Complete sequencing of both strands (ALF DNA sequencer; Pharmacia) revealed a full length ligase gene (2019 bp) with putative promoter and terminator sequences.

3D modelling

Sequences of the N-terminal domains of *Bs* and *Ph* DNA ligases, as well as all other known sequences of NAD^+-dependent DNA ligases were aligned using CLUSTAL W (1.74) [38]. This alignment, together with the coordinates from the crystal structure of the *Bs* enzyme were used to build a 3D model using MODELLER 4. This program models protein 3D structure by satisfaction of spatial restraints [39–43]. Energy minimization and simulated annealing were carried out using XPLOR version 3.1 [44]. The refinement involved 300 iterations of conjugate gradient minimization followed by molecular dynamics and simulated annealing with an initial heating of the molecule to 700 K and slow cooling to a bath temperature of 300 K. The time step for the simulation was 0.5 fs. Finally, 300 more iterations of conjugate gradient minimization were carried out to produce the refined model. Stereochemistry and geometry of the model at all stages of refinement were checked using the program PROCHECK [45]. Interpretation of the model, comparison with the crystal structure of *Bs* DNA ligase, and calculation of accessibilities were performed using the program TURBO FRODO [46]. Hydrogen bonds and salt-bridges were calculated with the program HBPLUS [47].

Overexpression and purification of *Ph* DNA ligase

The coding region of the psychrophilic DNA ligase gene was PCR amplified with the 5′-primer EXP1 containing an *Nde*I site (5′-ATTGAGGTTCATATGTCTAGCAGCATTAGTG-3′) and the 3′-primer EXP2 including the native Amber stop codon (underlined) (5′-GCTCATTATCGTTTTCCTAACCGT-TATGC-3′). The PCR product was cloned into the PCRScript vector and excised by *Nde*I and *Hin*dIII digestion. The resulting fragment was ligated into the pET28a(+) expression vector (Novagen) previously digested with *Nde*I and *Hin*dIII. Integrity of the *Ph lig* sequence was checked by sequencing of both strands. The recombinant DNA ligase was expressed in *Ec* BL21 (DE3) cells (Stratagene), as follows: an overnight culture (2.5 mL) was inoculated into 150 mL Luria–Bertani medium containing 50 $\mu g \cdot mL^{-1}$ kanamycin and grown at 18 °C until the D_{600} reached 0.6. Expression of the recombinant protein was induced by isopropyl thio-β-D-galactoside (final concentration 0.5 mM) and by further growth at 18 °C for 20 h. The bacterial pellet was resuspended in 20 mL of buffer A (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 20 mM imidazole, 1 mM $PhCH_2SO_2F$) and the cells were disrupted in a prechilled (4 °C) French pressure cell. The insoluble material was removed by centrifugation (30 min, 40 000 ***g***) and the supernatant loaded onto a 1-mL HiTrap chelating column (Pharmacia) charged with Ni^{2+} and equilibrated in buffer A. After loading, the column was washed with 5 mL of buffer A, the His-tagged protein was eluted with a 10-mL linear gradient from 20 to 500 mM imidazole in buffer A, and fractions of 0.5 mL collected. Most of the *Ph* DNA ligase was retained by the column in the presence of 20 mM imidazole and eluted at 250 mM imidazole concentration. The purity of the recombinant *Ph* DNA ligase was tested by SDS/PAGE (8%). Peak fractions containing the His-tagged protein (2 mL) were pooled and diluted to 10 mL in thrombin digestion buffer containing 20 mM Tris/HCl, pH 8.2, 150 mM NaCl and 2.5 mM $CaCl_2$. Digestion was started by adding 300 U of thrombin from bovine plasma (T4648, Sigma) and the sample incubated for 16 h at 16 °C. A precipitate that formed during incubation was removed by centrifugation (30 min, 40 000 ***g***). The supernatant containing the recombinant DNA ligase was passed again through a 1-mL HiTrap chelating column charged with Ni^{2+} to separate the thrombin-digested DNA ligase from the undigested fraction. The flow-through (10 mL) was loaded onto a 1-mL FPLC MonoQ HR 5/5 column (Pharmacia)

equilibrated in 20 mM Tris/HCl, pH 7.1. After washing with 5 mL of this buffer, the protein was eluted with a continuous 10 mL gradient to 650 mM NaCl in 20 mM Tris/HCl, pH 7.1, and fractions of 0.3 mL were collected. Pure DNA ligase was eluted at 250 mM NaCl concentration. The DNA ligase-containing pool (0.6 mL) was finally dialysed at 4 °C for 2 h against storage buffer (20 mM Tris/HCl, pH 7.1, 50 mM NaCl, 0.1 mM EDTA and 50% glycerol). Protein concentration was determined with the Bio-Rad protein assay, using BSA as standard. The recombinant DNA ligase was then stored at −20 °C.

Mass spectrometry

The molecular mass of the enzyme was determined via nano electrospray ionization mass spectrometry on a Q-TOF mass spectrometer (Micromass, Manchester, UK). A 10-pmol·μL^{-1} solution (5 μL) was loaded into a Au/Pd coated borosilicate capillary (Protana, Odense, Denmark) which was then placed in the nano electrospray source that came with the instrument. The spray was initiated by breaking slightly the capillary tip after which 1300 V was applied to it. The mass spectrometer was calibrated by preliminary analysis of horse myoglobin.

C-terminal sequencing

Chemical C-terminal sequencing was performed on a Procise 494C Sequencer (PE Biosystems, Foster City, CA, USA) following a protocol slightly modified from that given in [48]. Prior to analysis, the electroblotted sample was treated with phenyl isothiocyanate under basic conditions in order to modify the lysine residues [49].

Atomic absorption spectroscopy

Zinc ion binding by *Ph* DNA ligase was determined by atomic absorption spectroscopy. In order to remove nonspecifically bound Zn^{2+}, *Ph* DNA ligase was subjected to extensive dialysis against EDTA, as described previously by Barnes *et al.* [50]: the protein was dialysed five times for 12 h against 100 mM NaCl, 10 mM 2-mercaptoethanol, 10 mM EDTA and 10 mM Hepes, pH 7.5, followed by dialysis for 12 h against 100 mM NaCl, 0.01 mM 2-mercaptoethanol and 10 mM Hepes, pH 7.5. All the dialysis steps were performed at 4 °C. Untreated *Ph* DNA ligase (10.7 μM) and EDTA-dialysed *Ph* DNA ligase (11.8 μM) were assayed for Zn^{2+} content by atomic absorption spectroscopy (Video 12 aa/ae Spectrophotometer, Instrumentation Laboratory).

DNA ligase assay

For kinetic studies of DNA ligase activity, we modified a fluorimetric assay [51] based on the ligation of a 5′-fluorescein-labelled 17mer (CATCATGGCCCTTACGC) to a 5′-phosphorylated 17mer oligonucleotide (5′-CCAGGGCTC-CACACGTG), annealed to the complementary 34mer oligonucleotide (5′-CACGTGTGGAGCCCTGGGCGTAAGGGCC-ATGATG). All three oligonucleotides were purified by PAGE. Equimolar amounts of these three oligonucleotides were incubated in 20 mM Tris/HCl, pH 7.5 (E199, Amresco), 2 mM $MgCl_2$ (E525, Amresco) and 50 mM NaCl for 5 min at 75 °C, cooled slowly to room temperature and the sample (called 'substrate-template') desalted on a MicroSpin G-25 column (Pharmacia). Each assay was performed in a final volume of 12 μL containing: 100 pmol of substrate-template, 50 mM Tris/HCl, pH 7.8, 26 μM NAD^+, 10 mM $MgCl_2$, 10 mM dithiothreitol, 25 μg·mL^{-1} BSA and DNA ligase to be tested. Samples were incubated for 30 min at 4 °C, and the reactions stopped with 8 μL of 98% (v/v) formamide, 10 mM EDTA and 0.25 mg·mL^{-1} bromophenol blue. After heating for 3 min at 95 °C, the ligated products (34 nucleotides) were separated from unligated substrates (17 nucleotides) by electrophoresis on a 16% denaturing PAGE. Titration (20 pmol to 500 fmol) was performed with the fluorescein labelled oligonucleotide (17 nucleotides). Immediately after running, the gels were scanned with a Fluorimager 575 (Molecular Dynamics) and the intensities of the bands quantified with the IMAGE QUANT software (Molecular Dynamics). One unit of DNA ligase is defined as the amount of protein that ligates 1 pmol of substrate–template in 30 min at 4 °C.

Nuclease assay

Nuclease contamination was tested by incubating 600 ng of the *Ph* DNA ligase with 600 ng of linear or circular pUC19 DNA and circular single-stranded M13 DNA for 1 h at 37 °C in 10 mM Tris/HCl, pH 7.5, 10 mM $MgCl_2$, 50 mM NaCl and 1 mM dithiothreitol.

Cohesive-end and blunt-end ligation assay

To test ligation of cohesive-end and blunt-end DNA fragments, either *Ph* or *Ec* DNA ligases (100 ng) were incubated at 4 or 16 °C, respectively, for 4 h in a 30-μL reaction mixture containing 50 mM Tris/HCl, pH 7.8, 26 μM NAD^+, 10 mM $MgCl_2$, 10 mM dithiothreitol, 25 μg·mL^{-1} BSA, 300 ng of *Hin*dIII-digested or *Sma*I-digested pUC19 and poly(ethylene glycol) 6000 as indicated in the figure. Products were analysed on 1% agarose gel, containing ethidium bromide (0.4 μg·mL^{-1}).

RESULTS

Cloning of the *Ph* DNA ligase

Two thousand clones from the *Ph* genomic library were screened by PCR applied to pooled colonies from plates, using degenerate primers designed from the alignment of several amino-acid sequences from bacterial DNA ligases [20]. The screening resulted in the identification of four identical positive clones, which contained a 4-kb insert. Complete sequencing of both strands revealed a full length ligase gene with putative promoter and terminator sequences. The DNA ligase gene consists of 2019 bp and it has a GC content of 43.4%. The predicted protein contains 672 amino acids, which is similar to *Ec* (671 amino acids), *Bs* (671 amino acids), *T. filiformis* (667 amino acids), *Ts* (674 amino acids) and *T. thermophilus* (676 amino acids) DNA ligases but shorter than the 731-amino-acid

Table 1. Percentage of arginine and proline residues in NAD^+-dependent DNA ligases and percentage of identity between *Ph* DNA ligase and other NAD^+-dependent DNA ligases. R, arginine; K, lysine; P, proline.

DNA ligase	% P residues	% R residues	R/(R + K)	% identity *Ph*
Ph	4.2	5.7	0.49	100
Ec	4.5	7	0.59	60
Bs	4.3	8.7	0.65	44
Ts	4.9	10.1	0.62	45

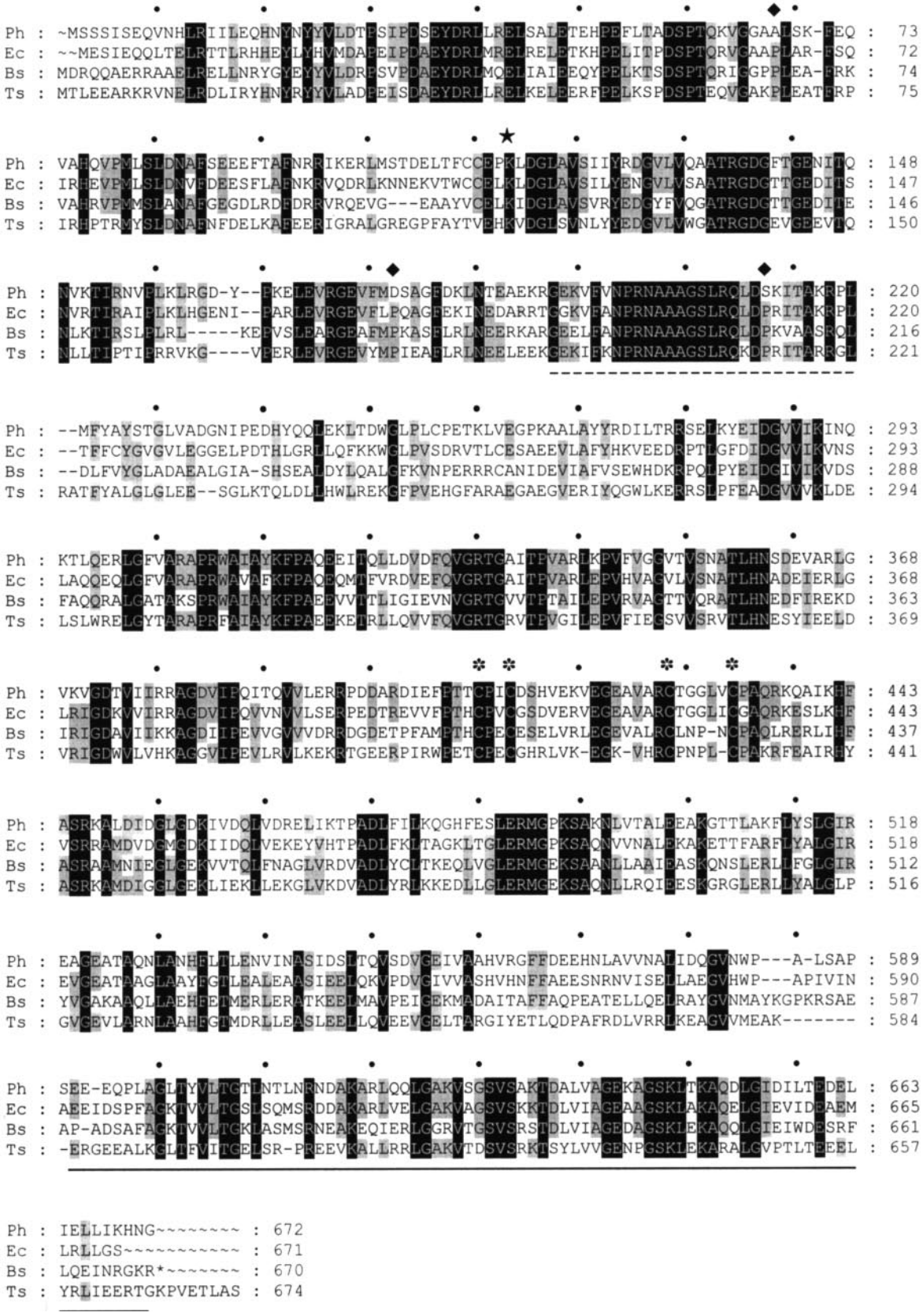

Fig. 1. Amino-acid sequence of the *Ph* DNA ligase sequence and comparison with *Ec*, *Bs* and *Ts* DNA ligases. The PILEUP program [65,66] was used for alignment. Every 10th residue is marked with a dot (.). Residues common to all DNA ligases are boxed in black whereas those common only to 3 of the 4 DNA ligases are boxed in grey. The solid diamonds (◆) indicate the replaced proline residues in *Ph* DNA ligase. The ★ symbol indicates the AMP-binding lysine (K) residue of bacterial DNA ligases. The most conserved motif (containing the NPRNAAAGS motif) in bacterial DNA ligases is shown with a dashed underline. The four universally conserved cysteine residues are shown with * above the alignment and are located in the region between amino acids 408–432 of *Ph* DNA ligase. RF-C homologous DNA binding sequence at the C-terminal is underlined.

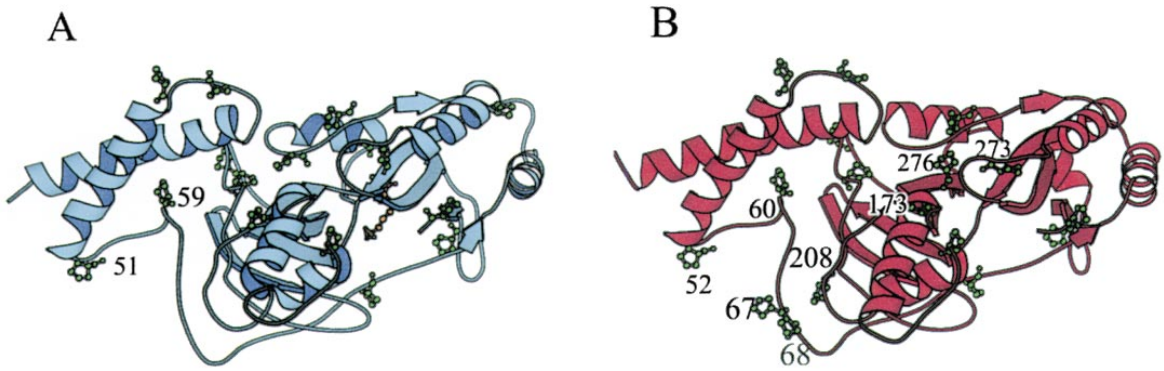

Fig. 2. Schematic view of the 3D structures of the *Ph* and *Bs* DNA ligase N-terminal domains. A and B represent the psychrophilic and thermophilic N-terminal domains. Proline residues are in ball-and-stick representation. Prolines in the region with increased flexibility in the psychrophilic enzyme are indicated. All these P residues are present in a long coil. MOLSCRIPT program was used to make the figure.

ligase of *Zymomonas mobilis* and the 712-amino-acid ligase of *R. marinus*. Sequence alignment revealed that the psychrophilic DNA ligase contains all the conserved boxes that have been previously described in bacterial DNA ligases [20] and shows strong similarity with the other NAD^+-dependent DNA ligases (Table 1). In three of these conserved boxes, proline residues have been replaced in the *Ph* DNA ligase but not in other NAD^+-dependent DNA ligases (Fig. 1, ◆ symbol). Some additional replacements were recorded only in *Ph* DNA ligase and could play a role in the adaptation to low temperatures. Table 1 shows the proportion of arginine and proline residues in the respective DNA ligases. The main sequence features described by Thorbjarnardóttir and colleagues [20], that is the KXDG box, the universally conserved NPRNAAAGS motif, the zinc finger like motif and the RF-C homologous DNA binding sequence, were also found in the *Ph* DNA ligase (Fig. 1).

3D modelling

The sequences of the N-terminal domains of the *Ph* and the *Bs* DNA ligases show 47.5% identity. This allowed us to build a model of the psychrophilic protein. Analysis of the final model using PROCHECK showed that 99.6% of the residues (except Gly and Pro) were in favoured and allowed regions. The folding is similar to that of the *Bs* enzyme, with no secondary structure alterations. Insertions of three and two residues, and a deletion of two residues, all occur in loop regions. As energy minimization tends to favour hydrogen bonds and salt-bridges, the same protocol of energy minimization and simulated annealing was carried out on the crystal structure of the *Bs* enzyme. This procedure also reduced differences due to crystal packing.

Comparison of the model with the *Bs* DNA ligase structure showed that the number of hydrogen bonds had decreased from 299 down to 278, and the number of salt-bridges had decreased from 18 to 13 in the psychrophilic enzyme. There are no major differences in hydrophobic clustering: in the *Bs* DNA ligase, 23 hydrophilic residues are replaced by hydrophobic ones in the *Ph* enzyme. In the latter protein, nine of them are completely buried, while eight of them have solvent accessibilities greater than 60 $Å^2$. Conversely, 23 hydrophilic residues in the psychrophilic enzyme are replaced by hydrophobic residues in the thermophilic protein. In this latter category, 11 residues are completely buried, while only two residues have accessibilities greater than 60 $Å^2$.

The proline (P) content is significantly lower in two regions of the structure. In the *Bs* enzyme, P residues 52, 60, 67, 68 (all in one long coil) and P208 are all on the same side of the enzyme (left side in Fig. 2). In the model of the psychrophilic DNA ligase, there are only two P residues (51 and 59). In the region above the active site in Fig. 2 there are three P residues in the *Bs* DNA ligase (residues 173, 273 and 276), whereas *Ph* DNA ligase contains no P residues in this region. In the region immediately around the active site, both the thermophilic and the psychrophilic proteins contain four P residues.

Overexpression and purification

The coding region of the *Ph* DNA ligase gene was PCR amplified and cloned into the pET28a(+) T7 expression vector which directs the expression of His-tagged proteins. After cell disruption, the DNA ligase was purified in three successive chromatographic steps as described in Materials and methods. Briefly, the first purification step consisted of a Ni^{2+} HiTrap chelating column. Most of the psychrophilic DNA ligase was retained by the column in the presence of 20 mM imidazole and eluted at 250 mM imidazole. After removal of the His-tag by thrombin digestion (Fig. 3, lane 4), the DNA ligase was passed again through a Ni^{2+} HiTrap chelating column to separate the untagged protein from the undigested protein. Finally, the remaining contaminants were removed by a MonoQ HR 5/5 chromatography step. Figure 3 shows the different steps in the *Ph* DNA ligase purification scheme. Purity was also ascertained by silver staining and the enzyme was judged to be more than 99% pure (data not shown). The final yield of the untagged enzyme was 21 mg of protein per litre of culture, representing approximately 3% of total cellular proteins. Protein was tested for contamination by single-stranded/double-stranded exonucleases and endonucleases as described in Materials and methods. Neither linear or circular pUC19, nor circular single-stranded M13 were degraded in the presence of the psychrophilic DNA ligase, indicating the absence of contaminating exonucleases and endonucleases from *Ec* (data not shown).

We used electrospray mass spectroscopy to determine the accurate molecular mass of the purified *Ph* DNA ligase protein (Fig. 4). Two populations were observed: a minor one of 74 265 Da and a main one of 74 592 Da. The mass difference between the two populations is 327 Da, corresponding to the molecular mass of AMP (329 Da) and we thus conclude that the majority of the purified enzyme is in an adenylated

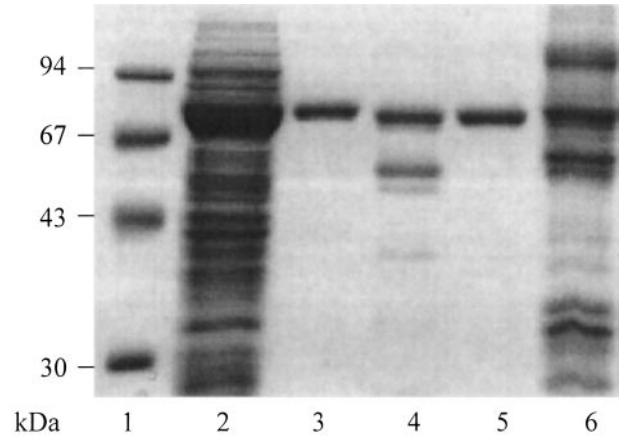

Fig. 3. Purification of the recombinant *Ph* DNA ligase. Electrophoretic analysis of recombinant DNA ligase fractions at different stages of the purification in 8% SDS-polyacrylamide gel. Lane 1, low molecular mass markers (Amersham Pharmacia Biotech); lane 2, total protein extract; lane 3, Ni^{2+} HiTrap chelating eluate; lane 4, Ni^{2+} HiTrap chelating flow-through (after thrombin digestion); lane 5, dialysed MonoQ eluate; lane 6, thrombin.

form. Although the copurification of unadenylated and adenylated forms has previously been described in the literature [26–33,52] the ratio of adenylation in the purified *Ph* DNA ligase is much higher. We must still account for a mass difference of 16 Da between the expected average mass of unadenylated *Ph* DNA ligase (74 249 Da) and the observed one (74 265 Da). N-terminal (GSHMSSSISE) and C-terminal sequencing (KHNG) confirmed the integrity of the recombinant protein, and complete sequencing of the expression construct revealed that this difference cannot be due to a mutation during PCR amplification. The most likely explanation is the oxidation of a *Ph* DNA ligase methionine residue. Indeed, the susceptibility of methionyl residues towards oxidation has long been known but has not been reported for DNA ligases previously.

Characterization of *Ph* DNA ligase

Zn^{2+} binding by Ph DNA ligase. As originally suggested by Thorbjarnardóttir *et al.* [20], the four cysteine residues conserved among all known NAD^+-dependent DNA ligases (see Fig. 1) might form a zinc-finger like structure. Two studies have used atomic absorption spectroscopy to measure the Zn^{2+} content of purified preparations of *Ts* [20] or *Bs* [18] DNA ligases. Although both studies revealed a significant amount of Zn^{2+} in the preparations but a molar ratio of Zn^{2+} to ligase well below a 1 : 1 stoichiometry, the groups reached opposite conclusions. We therefore decided to conclusively test whether *Ph* DNA ligase is a Zn^{2+}-binding enzyme. When *Ph* DNA ligase was assayed for Zn^{2+} binding by the same method, the zinc to protein molar ratio was 0.24. In order to check if this zinc was bound strongly in the zinc finger like structure or was nonspecifically bound, *Ph* DNA ligase was submitted to extensive dialysis against EDTA (see Materials and methods) and then subjected to atomic absorption spectroscopy. After dialysis, the zinc to protein molar ratio decreased to 0.07 (see Discussion).

Catalytic properties. After having modified a fluorimetric assay [51] to obey Michaelis–Menten kinetics (see Materials and methods and Discussion) we proceeded to measure the catalytic properties of *Ph* DNA ligase. Like other NAD^+-dependent DNA ligases [16–19,26], *Ph* DNA ligase requires Mg^{2+} (10 mM) and the ligation reaction is stimulated in the presence of monovalent cations such as NH_4^+ (10 mM) and K^+ (25 mM) (Fig. 5A). Dithiothreitol (10 mM) and BSA (25 $\mu g \cdot mL^{-1}$) also have a stimulating effect (Fig. 5A). The optimum pH value for activity of *Ph* DNA ligase is 8.6 at 4 °C in Tris/HCl buffer (Fig. 5B), which is close to the values found for other NAD^+-dependent DNA ligases. The last part in the investigation of the catalytic properties of the psychrophilic enzyme was to check if it possessed the same substrate specificity properties as the other NAD^+-dependent DNA ligases. NAD^+-dependent and ATP-dependent DNA ligases are able to seal cohesive ends, but under conventional assay conditions, only T4 DNA ligase efficiently catalyses blunt-end ligation. However, in the presence of high concentrations of macromolecules such as poly(ethylene glycol), Ficoll or plasma albumin, NAD^+-dependent DNA ligases are also able to catalyse blunt-end ligation [29–32]. Ligation assays were performed by incubating *Hin*dIII-digested (cohesive assay) or *Sma*I-digested (blunt assay) pUC19 either with *Ec* DNA ligase or with *Ph* DNA ligase. As expected, *Ph* DNA ligase was able to join cohesive DNA fragments (Fig. 6, lane 5). Moreover, it was only able to catalyse blunt-end ligation in the presence of the macromolecular volume excluder poly(ethylene glycol) 6000, as reported for *Ec* DNA ligase [29] and thermophilic *T. thermophilus* HB8 DNA ligase [32]. In the range of 20–25% poly(ethylene glycol) 6000, the blunt-end ligation activity was significant (Fig. 6, lanes 13–14). The ligation products were linear oligomers, due to intermolecular joining events, and no circular forms were observed. In the absence of the polymer, no blunt-end ligation was observed (Fig. 6, lane 9).

Adaptation to cold: comparison between psychrophilic, mesophilic and thermophilic DNA ligases

Thermal denaturation of Ph DNA ligase. Enzymes adapted to low temperatures are characterized by a higher susceptibility to thermal denaturation. To test if *Ph* DNA ligase behaves in the

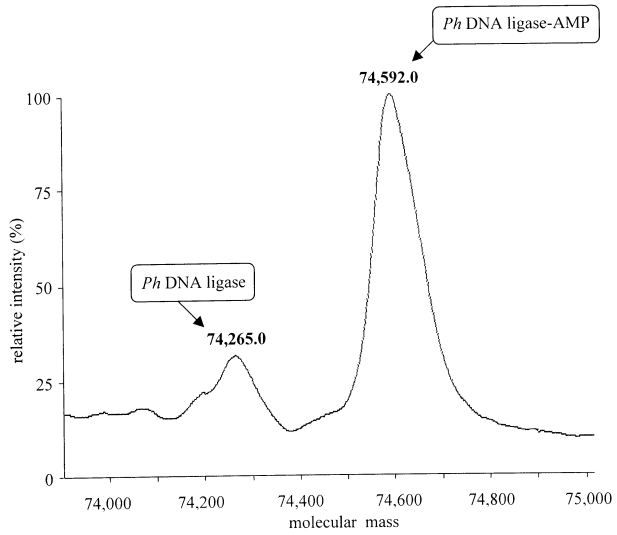

Fig. 4. Electrospray mass spectroscopy showing that *Ph* DNA ligase is constituted by two populations. The minor one consists of the native enzyme (74 265 Da) and the main one corresponds to the adenylated enzyme (74 592 Da).

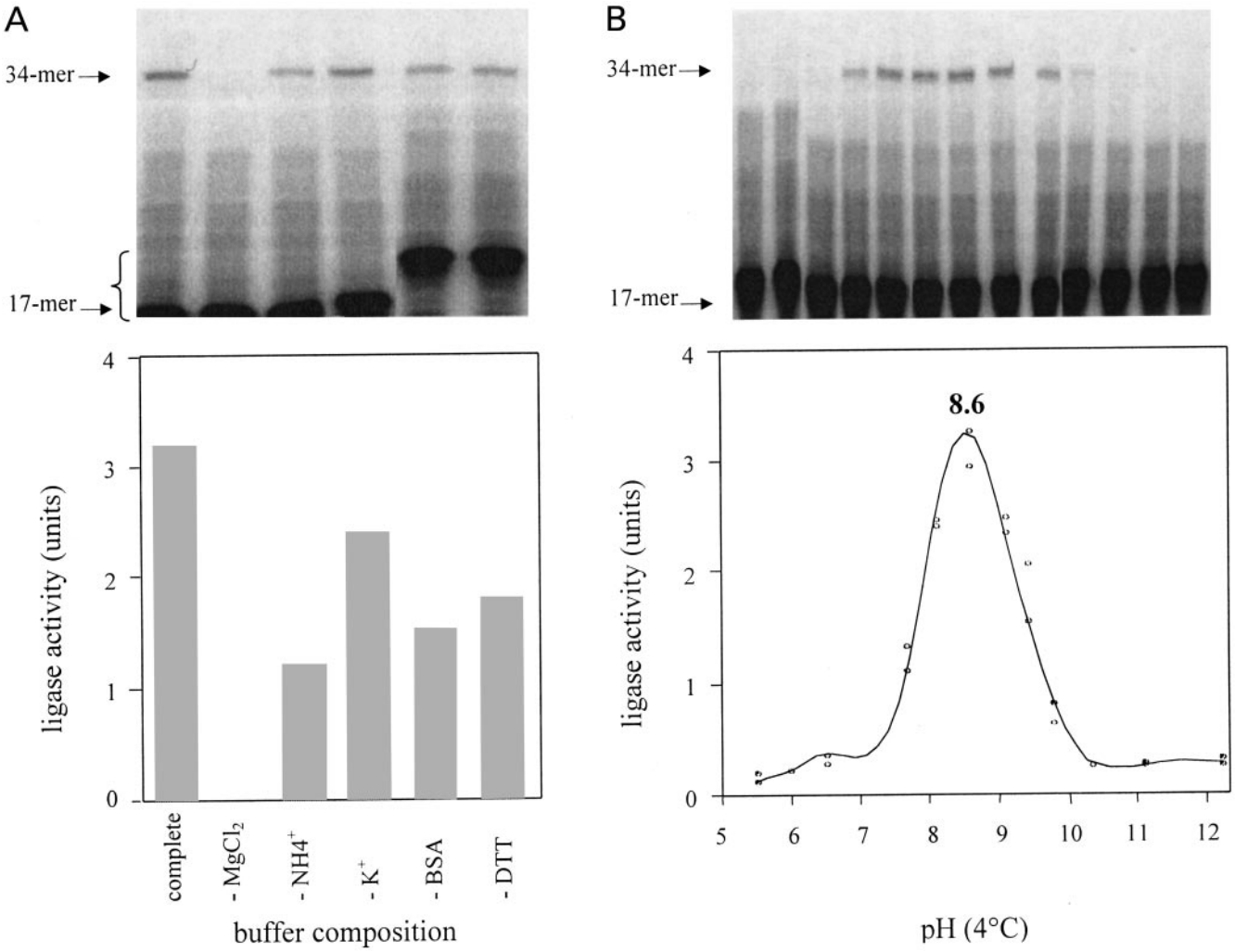

Fig. 5. Determination of the optimal buffer composition and optimal pH for the activity of *Ph* DNA ligase. Optimal buffer composition: (A) *Ph* DNA ligase was incubated either with a complete DNA ligase buffer (50 mM Tris/HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$, 25 mM KCl, 10 mM dithiothreitol (DTT) and 25 $\mu g \cdot mL^{-1}$ BSA) or with a buffer lacking one of these components. The picture is composed of lanes from two gels which explains the different position of the unligated 17mer in the leftmost four vs. the rightmost two lanes. Optimal pH for activity: (B) *Ph* DNA ligase was incubated in various buffers: Bistris (pH 5.5–6.5, 4 °C), Tris/HCl (pH 7.67–9.09, 4 °C) and Caps (pH 9.43–12.23, 4 °C).

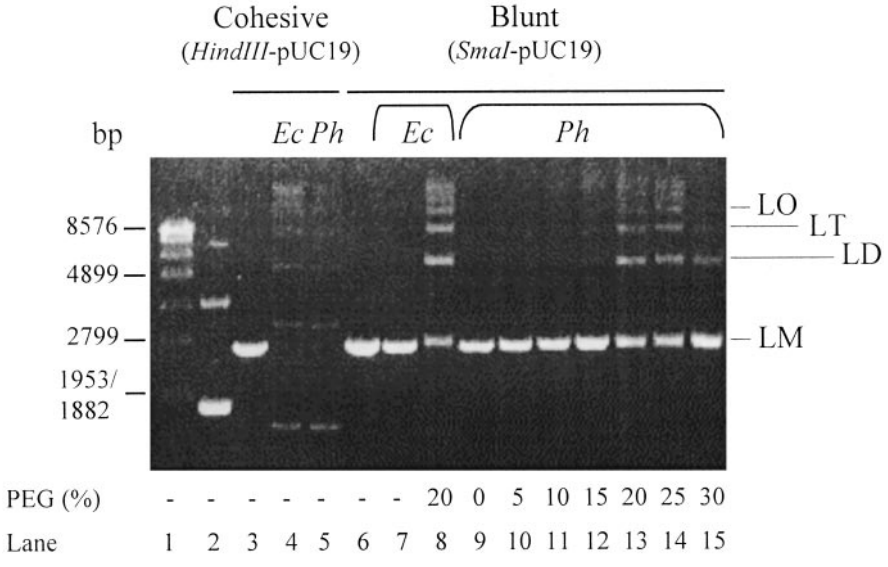

Fig. 6. Substrate specificity in cohesive-end vs. blunt-end ligation. (Cohesive ends) *Hin*dIII-cleaved pUC19 (300 ng) was incubated for 4 h either with 100 ng of *Ec* DNA ligase (New England Biolabs) (lane 4), 100 ng of *Ph* DNA ligase (lane 5) or without enzyme (lane 3) at 16 °C for *Ec* DNA ligase and 4 °C for *Ph* DNA ligase. (Blunt ends) *Sma*I-cleaved pUC19 (300 ng) was incubated either with *Ec* DNA ligase (lanes 7, 8), or *Ph* DNA ligase (lanes 9–15), or without enzyme (lane 6), under the same conditions as indicated above. Poly(ethylene glycol) 6000 (PEG) was added at the indicated concentrations (w/v) in the reaction mixture. Lane 2 shows circular pUC19 and lane 1 is marker VII from Boehringer Mannheim. The products were analysed on 1% agarose containing ethidium bromide (0.4 $\mu g \cdot mL^{-1}$). LM, linear monomer; LD, linear dimer; LT, linear trimer; LO, linear oligomers.

Table 2. Kinetic constants of psychrophilic, mesophilic and thermophilic NAD$^+$-DNA ligases. For the determination of the kinetic constants of each NAD$^+$-dependent DNA ligase the standard assay mixture contained 50 mM Tris/HCl pH 7.8, 26 μM NAD$^+$, 10 mM $MgCl_2$, 10 mM dithiothreitol, 25 μg·mL^{-1} BSA, 10 ng of enzyme and 7.5–100 pmol of 5′-fluorescein labelled 'substrate–template'. The assays were carried out at the indicated temperature and followed by measuring the fluorescence intensity of the ligated product. All experiments were performed in triplicate. V_{max} and $Km_{nicked\text{-}DNA}$ were obtained by Eadie–Hofstee linearization. Abbreviations as in Table 1.

	Ph DNA ligase			*Ec* DNA ligase			*Ts* DNA ligase	
	4 °C	18 °C	25 °C	18 °C	30 °C	45 °C	45 °C	60 °C
$K_{m,\ nicked\text{-}DNA}$ (μM)	0.165	0.296	0.631	0.179	0.702	2.040	0.236	0.465
V_{max} (pmol ligated·min^{-1})	0.080	0.271	0.374	0.031	0.172	0.257	0.098	0.481
k_{cat} (s^{-1})	0.0100	0.0337	0.0465	0.0038	0.0212	0.0317	0.0125	0.0613
k_{cat}/K_m (s^{-1}·μM^{-1})	0.0603	0.1139	0.0737	0.0214	0.0302	0.0155	0.0530	0.1319

same way, it was incubated at 4 °C, 18 °C and 25 °C for varying periods and then the *Ph* DNA ligase activity was assayed (Fig. 7). The results show that the psychrophilic enzyme is stable at 4 °C with a half-life exceeding 10 h, whereas it decreases to 24 min at 18 °C and 11.5 min at 25 °C.

Kinetic parameters of Ph, Ec and Ts DNA ligases. To evaluate the relative performances of *Ph*, *Ec* and *Ts* DNA ligases, we chose to modify the DNA ligase assay described by Jónsson *et al.* [51]. Most of the DNA ligase assays described in the literature require either a difficult and long substrate preparation, large amounts of radioactivity, or both. We selected the above-mentioned assay for its simplicity, its convenience and its reproducibility. However, the initial amount of the fluorescein labelled substrate described in the test (10 pmol) was too low for application of Michaelis–Menten kinetics. In other words, the measurement of initial velocity requires that less than 20% of the substrate is to be transformed into product after 30 min and this condition was not attained in the initial test. By using 100 pmol of substrate instead of 10 pmol and quantifying the upper bands corresponding to the ligated products, we obtained a sensitive, accurate and reproducible test that allowed us to determine the kinetic parameters of the three DNA ligases. We used the modified DNA ligase assay to measure the kinetic parameters K_m for nicked DNA, and the k_{cat} and K_m/k_{cat} of *Ph*, *Ec* and *Ts* DNA ligases (Table 2). First, the results show that *Ph* DNA ligase had an apparent optimal catalytic efficiency (k_{cat}/K_m) at 18 °C, whereas it was 30 °C for *Ec* DNA ligase, and this optimum was close to 60 °C in the case of *Ts* DNA ligase. Due to the denaturation of the DNA substrate at 75 °C, we were unable to determine the kinetic parameters of *Ts* DNA ligase at this temperature. The thermodependence of the DNA ligase activity was also tested. All the catalytic efficiency values found for *Ph* DNA ligase at 4 °C, 18 °C and 25 °C and *Ts* DNA ligase at 45 °C and 60 °C were higher than the values measured for *Ec* DNA ligase at 18 °C, 30 °C and 45 °C and reflect a temperature adaptation. The higher catalytic efficiency of *Ph* DNA ligase compared to its mesophilic counterpart reflects a cold-adaptation (Table 2). In this case, adaptation to low temperature is reached by an increased k_{cat}. Indeed comparison between *Ph* and *Ec* DNA ligases at 18 °C revealed a 1.7 times lower K_m in *Ec* DNA ligase ($K_{m,Ec} = 0.179$ against $K_{m,Ph} = 0.296$) but a 10 times higher k_{cat} in the case of *Ph* DNA ligase ($k_{cat,Ph} = 0.0337$ against $k_{cat,Ec} = 0.0038$). Concerning the adaptation to high temperatures, K_m for the DNA substrate could be the most important adaptive parameter: comparison at 45 °C between *Ec* and *Ts* DNA ligases revealed a K_m which was almost 10 times lower in the case of *Ts* DNA ligase compared to its mesophilic counterpart ($K_{m,Ts} = 0.236$ against $K_{m,Ec} = 2.040$), whereas the *Ec* k_{cat} was only three times higher than that of *Ts* DNA ligase ($k_{cat,Ec} = 0.0317$ against $k_{cat,Ts} = 0.0125$).

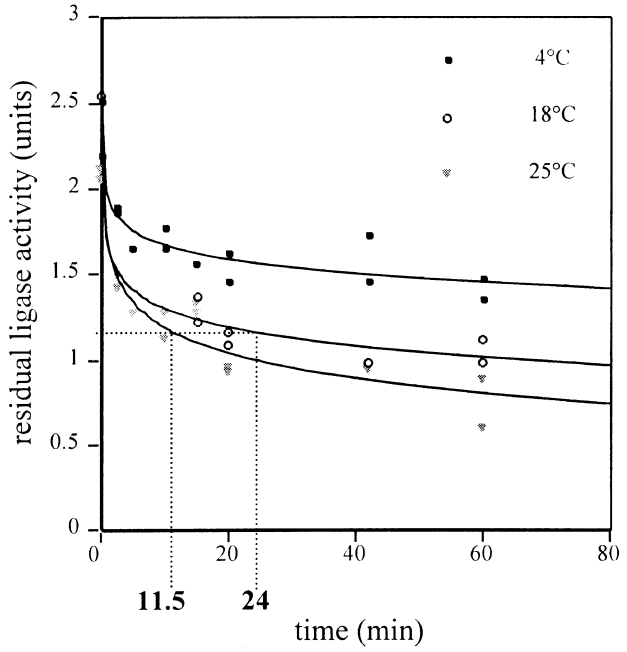

Fig. 7. Thermal denaturation of *Ph* DNA ligase. *Ph* DNA ligase (1 ng) was incubated in 20 mM Tris/HCl, pH 7.1, 50 mM NaCl, 0.1 mM EDTA and 50% glycerol, at 4 °C (■), 18 °C (○) and 25 °C (▼) and the DNA ligase activity was then assayed. The curves were drawn using exponential regression. The half-life of *Ph* DNA ligase is greater than 10 h at 4 °C, whereas it decreases to 24 min at 18 °C and 11.5 min at 25 °C.

DISCUSSION

In this paper, we present the first kinetic comparison between psychrophilic, mesophilic and thermophilic DNA ligases and the cloning, overexpression and characterization of the psychrophilic *Ph* DNA ligase. PCR screening allowed us to isolate the complete *Ph* DNA ligase gene. The encoded protein shows strong identity with other NAD$^+$-dependent DNA ligases and contains the previously described conserved motifs [20]. Analysis of the *Ph* DNA ligase amino-acid sequence underlined two characteristics which might be implicated in the adaptation to cold. First, the number of arginine residues is lower in the

psychrophilic enzyme than in mesophilic and thermophilic counterparts. Moreover, an increased number of arginine residues can be observed in the mesophilic compared with the thermophilic enzymes. Arginine is a stabilizing residue because of the ability of the three nitrogen atoms of its guanidinium group to form five hydrogen bonds with surrounding residues, as well as two salt-bridges with acidic groups. As psychrophilic enzymes are characterized by a more flexible structure, a decreased number of arginine residues can be expected, as is indeed seen. Second, the number of proline residues is lower in *Ph* DNA ligase than in other bacterial DNA ligases. The sidechain of the proline residue is covalently bound to the nitrogen atom of the peptide group, imposing constraints on the rotation about the N–C_α bond of the backbone and thus reducing the flexibility of its environment. In fact an increased number of proline residues has been noticed in some thermophilic enzymes, allowing the retention of a specific 3D conformation at high temperatures [53–58]. Even if the difference in the number of proline residues between the mentioned DNA ligases is not as large as in the case of the enzymes studied so far (reviewed in [59]), it should be noted that the three proline residues missing in *Ph* DNA ligase compared to *Ec* DNA ligase are located in conserved boxes. The rigidity of proline residues reduces flexibility in their vicinity and a decreased number of proline residues is thus consistent with a more flexible structure.

As the missing proline residues were located in the N-terminal domain of the protein, a 3D model of this domain was built to investigate the putative role of proline in cold-adaptation. Some interesting characteristics emerged from a comparison between the N-terminal domains of the psychrophilic and thermophilic enzymes. First of all, a decreased number of hydrogen bonds and salt-bridges in the psychrophilic protein was expected, as it renders the structure more flexible. Structural analysis of the psychrophilic model was in concord with this statement. Differences due to crystal packing were minimized by using the same energy minimization and simulated annealing protocol on the thermophilic DNA ligase. It is generally accepted that hydrophobic interactions contribute to the thermostability of proteins [60,61]. When comparing the psychrophilic N-terminal structural model to its thermophilic homologue, no striking differences in the clustering of hydrophobic residues appear in general. However, as observed in the psychrophilic α-amylase [5], citrate synthase [6] and salmon trypsin [62], the degree of exposure to solvent of hydrophobic residues is significantly increased in the *Ph* DNA ligase N-terminal domain compared to its thermophilic counterpart. Indeed, of the 23 hydrophobic residues in the *Ph* enzyme replaced by hydrophilic residues in the *Bs* enzyme, 34.8% have a solvent accessibility greater than 60 $Å^2$. Analogously, in the thermophilic enzyme these residues comprise only 7.1% out of a total of 28. Hydrophobic residues on the outer surface decrease the stability of a protein due to the ordering of water molecules, and the psychrophilic enzyme is clearly destabilized. When compared to the thermophilic X-ray structure, the proline content in the model is lower in two distinct regions. This suggests that the flexibility in these regions may be important. When aligning the sequences of several DNA ligases, it can be seen that in the region to the left of the active site shown in Fig. 2, the Pro content varies from five to six for DNA ligases of thermophilic species such as *R. marinus*, *T. thermophilus* and *Ts*. In mesophilic species there are usually three or four Pro residues, whereas in the psychrophile *Ph* there are only two. There are exceptions, however, such as in *Treponema pallidum* where there is only one Pro residue. The DNA ligase of this organism is about 100 residues longer than the average NAD^+-dependent DNA ligase, and is therefore expected to have distinct structural properties. As for the second region, above the active site shown in Fig. 2, no correlation between Pro content and growth temperature of the organism can be seen.

Alignment of the amino-acid sequences of NAD^+-dependent DNA ligases reveals four universally conserved residues in the region between amino acids 408–432 of the *Ph* DNA ligase. It has been proposed that these cysteine residues could form a zinc finger-like domain, although secondary structure prediction places most of the motif in loops, lacking the α-helix and β-sheet components present in most types of zinc finger [20]. Atomic absorption spectroscopy experiments with *Ph* DNA ligase showed substoichiometric amounts of zinc, as previously observed for both *Ts* and *Bs* DNA ligases [20–63]. Thorbjarnardóttir and coworkers [20] concluded that it was unlikely that the four-cysteine cluster forms a zinc finger, but Timson and Wigley [18] showed that the C-terminal domain of *Bs* contains a zinc binding site, although they observed a zinc to protein molar ratio of only 0.56. As zinc and similar trace metals can bind strongly, but nonspecifically to many large proteins [64], we decided to submit *Ph* DNA ligase to repeated dialysis against EDTA and to measure the Zn^{2+} content. The results show that *Ph* DNA ligase can bind zinc, but the amount detected is much lower than would be expected if the Zn^{2+} was strongly bound in a zinc finger. Conversely, when Barnes *et al.* [50] treated their DNA polymerase III to an exhaustive regime of EDTA dialysis, they found that the protein was able to bind zinc in a 1 : 1 stoichiometry. Our results thus support the conclusions of Thorbjarnardóttir and colleagues [20]. The possibility remains that the structure binds some other cation. Indeed, the existence of an 'iron finger' has been described for the *Clostridium pasteurianum* iron-rubredoxin in which the iron atom is tetrahedrally coordinated by four cysteinyl sulfurs [65].

Recombinant *Ph* DNA ligase shows the same catalytic properties as do other NAD^+-dependent DNA ligases. Indeed, it requires Mg^{2+}, and its activity is stimulated by low concentrations of NH_4^+ and K^+. Moreover, the psychrophilic enzyme efficiently seals cohesive ends but ligates blunt ends only in the presence of 20–25% poly(ethylene glycol). *Ph* DNA ligase is stable at 4 °C but denatures easily with increased temperature. This behaviour has been observed for all known psychrophilic proteins. This susceptibility to thermal denaturation appears to be due to a more flexible structure adopted by psychrophilic enzymes to undergo fast and efficient conformational changes during catalysis [4].

As DNA ligases were now available from psychrophilic, mesophilic and thermophilic organisms, kinetic experiments were performed to investigate whether the parameters k_{cat} or K_m for nicked DNA could be affected by molecular adaptation to extreme temperatures. To do such experiments, the DNA ligase assay described by Jónsson *et al.* [51] was optimized to obtain a convenient and reproducible test. Analysis of the catalytic efficiency shows that all the values measured for *Ph* and *Ts* DNA ligases are higher than those for *Ec* DNA ligase which reflects a temperature adaptation. Careful examination of the data revealed that k_{cat} seemed to be the most important adaptive parameter for *Ph* DNA ligase whereas it could be K_m for nicked DNA in the case of the *Ts* DNA ligase. From the Arrhenius equation $k = Ae^{-Ea/RT}$ (where k is the reaction rate, A is a 'pre-exponential' factor that allows the determination of the activation energy Ea, R is the gas constant (8.31 $J{\cdot}K^{-1}{\cdot}mol^{-1}$) and T is the temperature in Kelvin), one can deduce that a decrease in temperature leads to an

exponential decrease of the reaction rate. Psychrophilic enzymes have to find adaptive mechanisms to counter-balance this fact. In the psychrophilic DNA ligase, an increase of the turnover number of the protein (k_{cat}) compensates for the effect of low temperature on the catalytic rate. Due to their high working temperature, thermophilic enzymes have to cope with significant substrate denaturation. Their adaptation seems to be a decreased K_m for the DNA substrate. The cloning and expression of a psychrophilic DNA ligase will allow us to further study the adaptation to extreme temperatures using techniques such as intrinsic fluorescence, circular dichroism and differential scanning calorimetry. Such experiments have been used to compare psychrophilic *A. haloplanktis* α-amylase with its mesophilic counterparts [66]. They revealed that the heat-labile enzyme was the only one exhibiting a reversible two-state unfolding. This finding argued that the psychrophilic α-amylase has evolved towards the lowest available conformational stability of its native state. As NAD^+-dependent DNA ligases are well conserved, these experiments should reveal to us which conformational parameters are involved in temperature adaptation.

Commercially available DNA ligases have relatively poor activity at temperatures below 15 °C and require long incubation times. Under such conditions, the action of residual nucleases is favoured, which can interfere with the ligation reaction. We have shown here that the main adaptation feature of *Ph* DNA ligase is a 10 times higher k_{cat} at moderate temperatures and therefore this cold-active enzyme constitutes a novel tool in biotechnology. An additional advantage of the cold-active ligase is also the fact that a relatively low inactivation temperature can be used, which does not denature DNA.

ACKNOWLEDGEMENTS

We thank C. Duez, G. Maga, S. Briand, G. Feller and J.-M. François for advice and scientific discussions. We also thank D. Hasler and N. Gerardin for technical assistance. We acknowledge the 'Institut Français de Recherche et de Technologie Polaire' for generously accommodating our research fellows at the French Antarctic Station J. S. Dumont d'Urville in Terre Adélie. This research has been supported by the E.U. under the form of a network contract *n*° ERBCHCT 940521, a concerted action BIO4-CT95-0017 and a Biotech program BIO4-CT96-0051; by the 'Ministère de l'Education, de la Recherche et de la Formation', concerted action ARC 93/98-170, by the 'Région Wallonne-Direction Générale des Technologies', Convention 1828, by the Swiss NF to Z. O. J. and U. H. (grant 31.43138-98). J. V. B. is indebted to the Fund for Scientific Research-Flanders (project G. 0068-96) and to the University of Gent for a Concerted Research Action (project 12050198).

REFERENCES

1. Feller, G. & Gerday, C. (1997) Psychrophilic enzymes: molecular basis of cold adaptation. *Cell. Mol. Life Sci.* **53**, 830–841.
2. Gerday, C., Aittaleb, M., Arpigny, J.L., Baise, E., Chessa, J.P., Garsoux, G., Petrescu, I. & Feller, G. (1997) Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta* **1342**, 119–131.
3. Gerday, C., Aittaleb, M., Arpigny, J.L., Baise, E., Chessa, J.P., François, J.M., Petrescu, I. & Feller, G. (1999) Cold enzymes: a hot topic. In *Cold Adapted Organisms: Ecology, Physiology, Enzymology and Molecular Biology* (Margessi, R. & Schinner, F., eds), pp. 257–275. Springer-Verlag, Heidelberg.
4. Hochachka, P.W. & Somero, G.N. (1984) Temperature adaptation. In *Biochemical Adaptations* (Hochacka, P.W. & Somero, G.N., eds), pp. 355–449. Princeton University Press, Princeton, NJ.
5. Aghajari, N., Feller, G., Gerday, C. & Haser, R. (1998) Structures of the psychrophilic *Alteromonas haloplanctis* α-amylase give insights into cold adaptation at a molecular level. *Structure* **6**, 1503–1516.
6. Russell, R.J., Gerike, U., Danson, M.J., Hough, D.W. & Taylor, G.L. (1998) Structural adaptations of the cold-active citrate synthase from an Antarctic bacterium. *Structure* **6**, 351–361.
7. Alvarez, M., Zeelen, J.P., Mainfroid, V., Rentier-delrue, F., Martial, J.A., Wyns, L., Wierenga, R.K. & Maes, D. (1998) Triose-phosphate isomerase (TIM) of the psychrophilic bacterium *Vibrio marinus*. Kinetic and structural properties. *J. Biol. Chem.* **273**, 2199–2206.
8. Kim, S.Y., Hwang, K.Y., Kim, S.H., Sung, H.C., Han, Y.S. & Cho, Y.J. (1999) Structural basis for cold adaptation. Sequence, biochemical properties, and crystal structure of malate dehydrogenase from a psychrophile *Aquaspirillium arcticum*. *J. Biol. Chem.* **274**, 11761–11767.
9. Lehman, I.R. (1974) DNA ligase: structure, mechanism, and function. *Science* **186**, 790–797.
10. Lindahl, T. & Barnes, D.E. (1992) Mammalian DNA ligases. *Annu. Rev. Biochem.* **61**, 251–281.
11. Tomkinson, A.E. & Levin, D.S. (1997) Mammalian DNA ligases. *Bioessays* **19**, 893–901.
12. Weiss, B. & Richardson, C.C. (1967) Enzymatic breakage and joining of deoxyribonucleic acid. I. Repair of single-strand breaks in DNA by an enzyme system from *Escherichia coli* infected with T4 bacteriophage. *Proc. Natl Acad. Sci. USA* **57**, 1021–1028.
13. Barker, D.G., White, J.H. & Johnston, L.H. (1985) The nucleotide sequence of the DNA ligase gene (CDC9) from *Saccharomyces cerevisiae*: a gene which is cell-cycle regulated and induced in response to DNA damage. *Nucleic Acids Res.* **13**, 8323–8337.
14. Tomkinson, A.E., Totty, N.F., Ginsburg, M. & Lindahl, T. (1991) Location of the active site for enzyme-adenylate formation in DNA ligases. *Proc. Natl Acad. Sci. USA* **88**, 400–404.
15. Kletzin, A. (1992) Molecular characterisation of a DNA ligase gene of the extremely thermophilic archaeon *Desulfurolobus ambivalens* shows close phylogenetic relationship to eukaryotic ligases. *Nucleic Acids Res.* **20**, 5389–5396.
16. Zimmerman, S.B., Little, J.W., Oshinsky, C.K. & Gellert, M. (1967) Enzymatic joining of DNA strands: a novel reaction of diphosphopyridine nucleotide. *Proc. Natl Acad. Sci. USA* **57**, 1841–1848.
17. Shark, K.B. & Conway, T. (1992) Cloning and molecular characterization of the DNA ligase gene (*lig*) from *Zymomonas mobilis*. *FEMS Microbiol. Lett.* **75**, 19–26.
18. Timson, D.J. & Wigley, D.B. (1999) Functional domains of an NAD^+-dependent DNA ligase. *J. Mol. Biol.* **285**, 73–83.
19. Takahashi, M., Yamaguchi, E. & Uchida, T. (1984) Thermophilic DNA ligase. Purification and properties of the enzyme from *Thermus thermophilus* HB8. *J. Biol. Chem.* **259**, 10041–10047.
20. Thorbjarnardóttir, S.H., Jónsson, Z.O., Andresson, O.S., Kristjánsson, J.K., Eggertsson, G. & Pálsdóttir, A. (1995) Cloning and sequence analysis of the DNA ligase-encoding gene of *Rhodothermus marinus*, and overproduction, purification and characterization of two thermophilic DNA ligases. *Gene* **161**, 1–6.
21. Kim, H.K. & Kwon, S.T. (1998) Cloning, nucleotide sequence, and expression of the DNA ligase-encoding gene from *Thermus filiformis*. *Mol. Cells* **8**, 438–443.
22. Cheng, C. & Shuman, S. (1997) Characterization of an ATP-dependent DNA ligase encoded by *Haemophilus influenzae*. *Nucleic Acids Res.* **25**, 1369–1374.
23. Aravind, L., Tatusov, R.L., Wolf, Y.I., Walker, D.R. & Koonin, E.V. (1998) Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. *Trends Genet.* **14**, 442–444.
24. Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., Overbeek, R., Snead, M.A., Keller, M., Aujay, M., Huber, R., Feldman, R.A., Short, J.M., Olsen, G.J. & Swanson, R.V. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353–358.
25. Burbelo, P.D., Utani, A., Pan, Z.Q. & Yamada, Y. (1993) Cloning of the large subunit of activator 1 (replication factor C) reveals homology with bacterial DNA ligases. *Proc. Natl Acad. Sci. USA* **90**, 11543–11547.

26. Modrich, P., Anraku, Y. & Lehman, I.R. (1973) Deoxyribonucleic acid ligase. Isolation and physical characterization of the homogeneous enzyme from *Escherichia coli*. *J. Biol. Chem.* **248**, 7495–7501.
27. Panasenko, S.M., Alazard, R.J. & Lehman, I.R. (1978) A simple, three-step procedure for the large scale purification of DNA ligase from a hybrid lambda lysogen constructed in vitro. *J. Biol. Chem.* **253**, 4590–4592.
28. Ishino, Y., Shinagawa, H., Makino, K., Tsunasawa, S., Sakiyama, F. & Nakata, A. (1986) Nucleotide sequence of the *lig* gene and primary structure of DNA ligase of *Escherichia coli*. *Mol. Gen. Genet.* **204**, 1–7.
29. Zimmerman, S.B. & Pheiffer, B.H. (1983) Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **80**, 5852–5856.
30. Singleton, M.R., Hakansson, K., Timson, D.J. & Wigley, D.B. (1999) Structure of the adenylation domain of an NAD^+-dependent DNA ligase. *Structure* **7**, 35–42.
31. Kaczorowski, T. & Szybalski, W. (1996) Co-operativity of hexamer ligation. *Gene* **179**, 189–193.
32. Takahashi, M. & Uchida, T. (1986) Thermophilic HB8 DNA ligase: effects of polyethylene glycols and polyamines on blunt-end ligation of DNA. *J. Biochem.* (*Tokyo*) **100**, 123–131.
33. Barany, F. & Gelfand, D.H. (1991) Cloning, overexpression and nucleotide sequence of a thermostable DNA ligase-encoding gene. *Gene* **109**, 1–11.
34. Lauer, G., Rudd, E.A., McKay, D.L., Ally, A., Ally, D. & Backman, K.C. (1991) Cloning, nucleotide sequence, and engineered expression of *Thermus thermophilus* DNA ligase, a homolog of *Escherichia coli* DNA ligase. *J. Bacteriol.* **173**, 5047–5053.
35. Luo, J. & Barany, F. (1996) Identification of essential residues in *Thermus thermophilus* DNA ligase. *Nucleic Acids Res.* **24**, 3079–3085.
36. Luo, J., Bergstrom, D.E. & Barany, F. (1996) Improving the fidelity of *Thermus thermophilus* DNA ligase. *Nucleic Acids Res.* **24**, 3071–3078.
37. Pritchard, C.E. & Southern, E.M. (1997) Effects of base mismatches on joining of short oligodeoxynucleotides by DNA ligases. *Nucleic Acids Res.* **25**, 3403–3407.
38. Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
39. Sali, A. & Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.
40. Sali, A. & Overington, J.P. (1994) Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Sci.* **3**, 1582–1596.
41. Sali, A., Potterton, L., Yuan, F., van vlijmen, H. & Karplus, M. (1995) Evaluation of comparative protein modeling by MODELLER. *Proteins* **23**, 318–326.
42. Sali, A. (1995) Comparative protein modeling by satisfaction of spatial restraints. *Mol. Med. Today.* **1**, 270–277.
43. Sanchez, R. & Sali, A. (1997) Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins* **Suppl. 1**, 50–58.
44. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987) Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460.
45. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
46. Roussel, A. & Cambillau, C. (1991) *Silicon Graphics Geometry Partners Directory 86.* Mountain View, California.
47. McDonald, I.K. & Thornton, J.M. (1994) Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
48. Bozzini, M., Zhao, J., Yuan, P.-m., Ciolek, D., Pan, Y.-C., Horton, J., Marshak, D.R. & Boyd, V.L. (1995) Applications using an alkylation method for carboxy-terminal protein sequencing. In *Techniques in Protein Chemistry VI* (Crabb, J., ed.), pp. 229–237. Academic Press, New York.
49. Boyd, V.L., Bozzini, M., Zon, G., Noble, R.L. & Mattaliano, R.J. (1992) Sequencing of peptides and proteins from the carboxy terminus. *Anal. Biochem.* **206**, 344–352.
50. Barnes, M.H., Leo, C.J. & Brown, N.C. (1998) DNA polymerase III of Gram-positive eubacteria is a zinc metalloprotein conserving an essential finger-like domain. *Biochemistry* **37**, 15254–15260.
51. Jónsson, Z.O., Hindges, R. & Hübscher, U. (1998) Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *EMBO J.* **17**, 2412–2425.
52. Brannigan, J.A., Ashford, S.R., Doherty, A.J., Timson, D.J. & Wigley, D.B. (1999) Nucleotide sequence, heterologous expression and novel purification of DNA ligase from *Bacillus stearothermophilus*. *Biochim. Biophys. Acta* **1432**, 413–418.
53. Watanabe, K., Chishiro, K., Kitamura, K. & Suzuki, Y. (1991) Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucosidasius* KP1006. *J. Biol. Chem.* **266**, 24287–24294.
54. Watanabe, K., Masuda, T., Ohashi, H., Mihara, H. & Suzuki, Y. (1994) Multiple proline substitutions cumulatively thermostabilize *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase. Irrefragable proof supporting the proline rule. *Eur. J. Biochem.* **226**, 277–283.
55. Watanabe, K., Kitamura, K. & Suzuki, Y. (1996) Analysis of the critical sites for protein thermostabilization by proline substitution in oligo-1,6-glucosidase from *Bacillus coagulans* ATCC 7050 and the evolutionary consideration of proline residues. *Appl. Environ. Microbiol.* **62**, 2066–2073.
56. Kashiwabara, S., Matsuki, Y., Kishimoto, T. & Suzuki, Y. (1998) Clustered proline residues around the active-site cleft in thermostable oligo-1,6-glucosidase of *Bacillus flavocaldarius* KP1228. *Biosci. Biotechnol. Biochem.* **62**, 1093–1102.
57. Van den Burg, B., Vriend, G., Veltman, O.R., Venema, G. & Eijsink, V.G. (1998) Engineering an enzyme to resist boiling. *Proc. Natl Acad. Sci. USA* **95**, 2056–2060.
58. Igarashi, K., Ozawa, T., Ikawakitayama, K., Hayashi, Y., Araki, H., Endo, K., Hagihara, H., Ozaki, K., Kawai, S. & Ito, S. (1999) Thermostabilization by proline substitution in an alkaline, liquefying α-amylase from *Bacillus* sp. strain KSM-1378. *Biosci. Biotechnol. Biochem.* **63**, 1535–1540.
59. Feller, G., Arpigny, J.L., Narinx, E. & Gerday, C. (1997) Molecular adaptations of enzymes from psychrophilic organisms. *Comp. Biochem. Physiol.* **118**, 495–499.
60. Dill, K.A. (1990) Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155.
61. Doig, A.J. & Williams, D.H. (1991) Is the hydrophobic effect stabilizing or destabilizing in proteins? The contribution of disulphide bonds to protein stability. *J. Mol. Biol.* **217**, 389–398.
62. Smalas, A.O., Heimstad, E.S., Hordvik, A., Willassen, N.P. & Male, R. (1994) Cold adaption of enzymes: structural comparison between salmon and bovine trypsins. *Proteins* **20**, 149–166.
63. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.
64. Cornell, N.W. & Crivaro, K.E. (1972) Stability constant for the zinc-dithiothreitol complex. *Anal. Biochem.* **47**, 203–208.
65. Watenpaugh, K.D., Sieker, L.C. & Jensen, L.H. (1980) Crystallographic refinement of rubredoxin at 1.2 Å degrees resolution. *J. Mol. Biol.* **138**, 615–633.
66. Feller, G., d'Amico, D. & Gerday, C. (1999) Thermodynamic stability of a cold-active α-amylase from the Antarctic bacterium *Alteromonas haloplanctis*. *Biochemistry* **38**, 4613–4619.