

# The *hppA* gene of *Helicobacter pylori* encodes the class C acid phosphatase precursor

Renata Godlewska<sup>a,\*</sup>, Janusz M. Bujnicki<sup>b</sup>, Jerzy Ostrowski<sup>c</sup>, Elżbieta K. Jagusztyn-Krynicka<sup>a</sup>

<sup>a</sup>Department of Bacterial Genetics, Institute of Microbiology, Warsaw University, Miecznikowa 1, 02-096 Warsaw, Poland

<sup>b</sup>Bioinformatics Laboratory, International Institute of Molecular and Cell Biology, ks, Trojdena 4, 02-109 Warsaw, Poland

<sup>c</sup>Department of Gastroenterology, Medical Center for Postgraduate Education, Maria Skłodowska-Curie Memorial Cancer Center, and the Institute of Oncology, 02-781 Warsaw, Poland

Received 9 May 2002; revised 30 June 2002; accepted 1 July 2002

First published online 12 July 2002

Edited by Takashi Gojobori

**Abstract** Screening of the *Helicobacter pylori* genomic library with sera from infected humans and from immunized rabbits resulted in identification of the 25 kDa protein cell envelope (HppA) which exhibits acid phosphatase activity. Enzyme activity was demonstrated by specific enzymatic assays with whole-cell protein preparations of *H. pylori* strain N6 and from *Escherichia coli* carrying the *hppA* gene (pUWM192). HppA showed optimum activity at pH 5.6 and was resistant to inhibition by EDTA. Bioinformatics analysis and site-directed mutagenesis of two putative active site residues (D73 and D192) provide further insight into the sequence–structure–function relationships of HppA as a member of the DDDD phosphohydrolase superfamily. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

**Key words:** HppA; Acid phosphohydrolase; *Helicobacter pylori*

## 1. Introduction

*Helicobacter pylori* is a curved, microaerophilic, Gram-negative bacterium that persistently colonizes the gastric epithelium of humans and other primates [1]. The infection persists for decades, resulting in the development of gastric mucosal inflammation, and is a risk factor for the development of peptic ulcer disease and gastric adenocarcinoma. Treatment of *H. pylori* involves expensive combinations of various medicines (antibiotics, proton pump inhibitors) and is not always effective [2].

In a previous study we reported the construction and screening of a *H. pylori* expression library in order to identify potential candidates for subunit vaccine antigens [3]. A number of 114 positive clones were isolated. One of them was predicted to encode a 25 kDa membrane protein of unknown function (92% DNA sequence identity with hypothetical protein HP1285 from *H. pylori* strain 26695). 40% sequence similarity between this protein and acid phosphatase LppC from *Streptococcus equisimilis* [4] prompted us to explore the possibility that this novel antigenic protein is endowed with phosphatase activity.

Many bacterial phosphatases function as scavengers of or-

ganic phosphoesters [5,6] but some of them are also involved in signal transduction and microbial pathogenesis [7,8]. Bacterial non-specific acid phosphohydrolases (NSAPs) are secreted enzymes, produced as soluble proteins associated with the periplasmic space or as membrane-bound lipoproteins. They generally dephosphorylate a broad array of substrates and have optimal catalytic activity at an acidic to neutral pH. Three classes of NSAP, designated A, B and C, have been proposed based on clustering of amino acid sequences, with HP1285 (ortholog of the protein analyzed in this work) grouped with class C [9]. The authors suggest that bacterial NSAPs of classes B and C, as well as related plant phosphatases, belong to a family of enzymes dubbed DDDD phosphohydrolases.

This family was shown to belong to a large superfamily of HAD hydrolases, which also includes P-type ATPases and haloacid dehalogenases [10].

## 2. Materials and methods

### 2.1. Bacteria and growth conditions

*H. pylori* strain N6, kindly provided by Dr. Agnes Labigne (Institut Pasteur) was grown under microaerobic conditions at 37°C on Blood Agar Base No2 (Sigma) containing 10% horse blood. The media for *H. pylori* mutants contained kanamycin (25 µg/ml). *Escherichia coli* strains DH5 (Gibco BRL) and M15 (Qiagen) were grown at 37°C on LB agar or in broth. *E. coli* transformants were grown in the presence of the appropriate antibiotic (ampicillin 100 µg/ml or kanamycin 25 µg/ml).

### 2.2. Recombinant DNA techniques

Procedures for plasmid DNA isolation and DNA analysis (digestion with restriction enzymes, T4 ligase ligation, Southern blots), agarose gel electrophoresis and transformation of *E. coli* competent cells were carried out as described by Sambrook [11]. DNA sequencing was performed on an ABI Prism 377 automated DNA sequencer (Perkin-Elmer).

The nucleotide sequence of the *hppA* gene has been deposited in the EMBL Nucleotide Sequence Database under accession no. AJ28848.

An *E. coli* strain that overproduces HppA was engineered by cloning the *hppA* coding region into pQE70 under the control of an inducible *lacZ* gene promoter. The 654 bp DNA fragment was PCR-amplified from pUWM192 with the oligonucleotide pair: 5'-GCATCGGTATGTTAAATGCCAAAGAAT and 5'-GGATCCATCTTCCATGTGCCATA containing the recognition sequences (underlined) for *SphI* and *BamHI*. The resulting PCR product was cloned into pQE70 (Qiagen). (His)<sub>6</sub>HppA recombinant protein was prepared from IPTG-induced *E. coli* M15(pREP4, pQE70/hppA) culture and purified by Ni-NTA affinity chromatography as described in the manufacturer's protocol (Qiagen). Purified (His)<sub>6</sub>HppA was used for rabbit immunization.

\*Corresponding author. Fax: (48)-22-5541404.

E-mail address: renatag@biol.uw.edu.pl (R. Godlewska).

Gene replacement was used to construct the *H. pylori* *hpaA* isogenic mutant. Two internal fragments of the gene were amplified by PCR and cloned into the *H. pylori* non-replicable vector pBluescript II SK with an inactivated *Clal* site in MCS. The following oligonucleotides were used: 5'-CCCGGGTCTCGCCATAACAAGA-3', 5'-ATCG-ATCCCCCAAGTGAGCGAAGAAT-3', 5'-ATCGATGCGTGCG-GTTAGAAATGTA-3', 5'-CTTCCCATGTGCCATAAAGAG-3'.

Primers contained *Sma*I and *Clal* restriction site overhangs (underlined) to facilitate cloning of the amplified DNA fragments in the appropriate orientation. A *Clal*-ended kanamycin-resistant cassette carrying the *Campylobacter jejuni* *aph(3')-III* gene (1.4 kb) was inserted into unique *Clal* sites among cloned PCR fragments. *H. pylori* N6 strain, naturally competent for genetic transformation, was used for generation of isogenic *H. pylori* *hpaA* mutants. The obtained kanamycin-resistant transformants were screened by hybridization using cloning vector DNA as a probe to confirm that no vector DNA was integrated into the chromosome via single cross over. Moreover, no 25 kDa protein was detected in the cell extract obtained from *H. pylori* *hpaA* isogenic mutant (data not shown). The lack of *hpaA* gene product is not lethal for bacterial cells. Mutant showed typical *Helicobacter* morphology and no alterations of growth rate.

Point mutations were generated using the Quick-Change site-directed mutagenesis kit following the procedures recommended by the supplier (Stratagene). Plasmid pUWM192 was used as a template for PCR-mediated mutagenesis and the following mutagenic oligonucleotide primers (reverse complements are not shown): HP1285 D73A GCCAGCTGTCATTTTGGCTTTAGATGAAACCG; HP1285 D192A GTGGGCGATACTTTGCATGCTTTTGACGCACATTTTGC.

Mismatches are underlined. Plasmids, pUWM314 containing *hpaA* with D72A mutation and pUWM315 with the D192A mutation, were transformed into *E. coli* DH5 $\alpha$ . The presence of the desired mutations was confirmed by DNA sequencing.

### 2.3. Western blot analysis

SDS-PAGE was carried out as described by Laemmli [12]. Western blot detection of HppA was performed using a 1:750 dilution of anti-HppA rabbit antibodies and goat anti-rabbit IgG conjugated to alkaline phosphatase (Sigma).

### 2.4. Preparation of cellular fractions

Cellular fractions were prepared by the method described by Blaser et al. [13]. The outer membrane fraction was isolated from cell envelopes by the sodium lauroyl sarcosine procedure of Filip [14]. The periplasmic fraction was prepared by the osmotic shock procedure of Ames [15].

### 2.5. Phosphatase assays

Whole-cell protein preparations were examined for phosphatase activity with disodium *p*-nitrophenol phosphate (*p*NPP, Sigma) as a substrate [16]. Briefly, the 0.5 ml assay mixture contained 0.1 M sodium acetate, pH 5.6, 5 mM *p*NPP and varying amounts of sample. Incubation was at 37°C for 30 min. The reaction was stopped by addition of 0.5 ml 0.2 M NaOH. Concentration of the released *p*-nitrophenol (*p*NP) was measured at 415 nm in the linear range of the calibration curve. Protein concentrations were determined by the

bicinchoninic acid method (Sigma) with bovine serum albumin as a standard. The phosphatase inhibition test was accomplished in standard reaction mixture in the presence of 15 mM EDTA. Determination of the pH optimum of HppA enzyme was performed with *p*NPP as the substrate in 0.1 M buffers: sodium acetate buffer (pH 4–6), Tris-HCl buffer (pH 7–9) and glycine/NaOH buffer (pH 10).

### 2.6. Protein sequence analysis and structure prediction

The PSI-BLAST algorithm [17] was used to search the non-redundant version of current sequence databases and the publicly available complete and incomplete genome sequences at the NCBI website (<http://www.ncbi.nlm.nih.gov>) using the HppA sequence as a query. The resulting multiple sequence alignment was refined using CLUSTALX [18]. Protein structure prediction was done using the MetaServer available at <http://bioinfo.pl/meta/>, which combines several secondary structure prediction and fold recognition methods [19]. The fold recognition methods compare the query sequence with a library of structures (templates) and return 10 alignments that scored best according to the implemented criterion of compatibility. The results are evaluated by the Pcons server [20], which compares the models and the associated scores and produces a ranking of potential best alignments with related proteins, whose structures are known. The consensus alignment provided by MetaServer served as a basis for homology modeling carried out using the SWISS-MODEL/PROMOD II server [21]. Protein structure evaluation was carried out using the Verify3D program [22], suggesting that the stereochemistry and energetic parameters of the model were acceptable.

## 3. Results and discussion

Our earlier studies have described the construction and screening of *H. pylori* expression libraries. When antiserum raised in rabbits exposed to formalin-killed bacterial cells and when human antisera from infected patients were used to screen *H. pylori* genomic libraries several immunopositive clones were identified [3]. One of them, containing a 1082 bp *H. pylori* DNA fragment inserted into the *Eco*RI site of pBluescript II SK, produced an immunopositive protein with an apparent  $M_r$  of 25 000, as determined by SDS-PAGE and Western blot. Sequence analysis identified a complete 738 bp gene encoding a putative protein with a calculated molecular mass of 28 kDa. Based on the results of the analysis reported herein, we designated his protein HppA (*H. pylori* phosphohydrolase). The *hpaA* gene expression in *E. coli* is probably driven from its own promoter, since cloning of the above-mentioned fragment in the opposite orientation resulted in a product showing cross-reactivity with the same antiserum. To determine subcellular localization of HppA in *H. pylori* N6 and *E. coli*/pUWM192, cell fractionation was carried out. Western blots of *H. pylori* proteins iso-

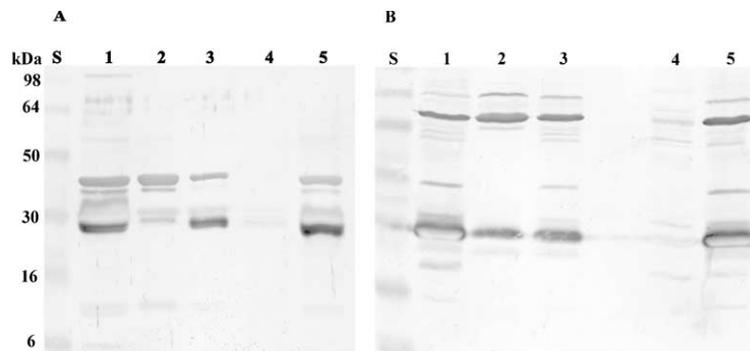


Fig. 1. Western blot detection of HppA protein in cell fractions of *E. coli*/pUWM192 (A) and *H. pylori* N6 (B). The results were obtained from SDS-12% PAGE of whole-cell extract (lane 1), outer membrane fraction (lane 2), cell envelope fraction (lane 3), periplasmic fraction (lane 4) and whole-cell extract without periplasmic proteins (lane 5).

lated from different cell compartments probed with rabbit antiserum to recombinant (His)<sub>6</sub>HppA protein showed that HppA was recovered in the sarcosyl-insoluble fraction. In contrast, analysis of the *E. coli*pUWM192 proteins revealed the presence of HppA in the crude cell envelope fraction when extracted by sarcosyl treatment (Fig. 1). No HppA was detected in the outer membrane fraction of *E. coli*pUWM192.

The sequence of HppA was analyzed using the Structure Prediction MetaServer [19], providing compelling evidence that this protein is remotely related to structurally characterized members of the DDDD superfamily (comprehensive results are available online at the following URL: <http://bioinfo.pl/meta/target.pl?id=5774>). Most of the individual fold recognition servers detected similarity between HppA and the DDDD fold (FSSP: 147.2.3; SCOP: e.21.1 <http://scop.mrc-lmb.cam.ac.uk/scop/>) and the consensus server evaluated this prediction as highly confident. In order to characterize the sequence–function relationships of HppA at the structural level, we attempted to build and evaluate a three-dimensional model. The phosphonoacetaldehyde hydrolase structure [23] was used as the template, since it was suggested as the top match by the consensus server (see Section 2 and the URL given above). The quality of the theoretical model of HppA was evaluated as reasonable (average Verify3D score = 0.29), hence we used it to identify the putative catalytic site (Fig. 2).

To validate the bioinformatics-based assignment of HppA to the HAD superfamily and specifically to the family of non-specific phosphatases, we analyzed the enzymatic activity of the protein. To achieve this, the specific phosphomonoesterase activity of whole-cell fractions from *E. coli*pUWM192, *H. pylori* N6 and the isogenic *hppA* insertion mutant were determined. As shown in Table 1, the cells of *H. pylori* N6/*hppA*::Km possessed enzyme activity at approximately 48%

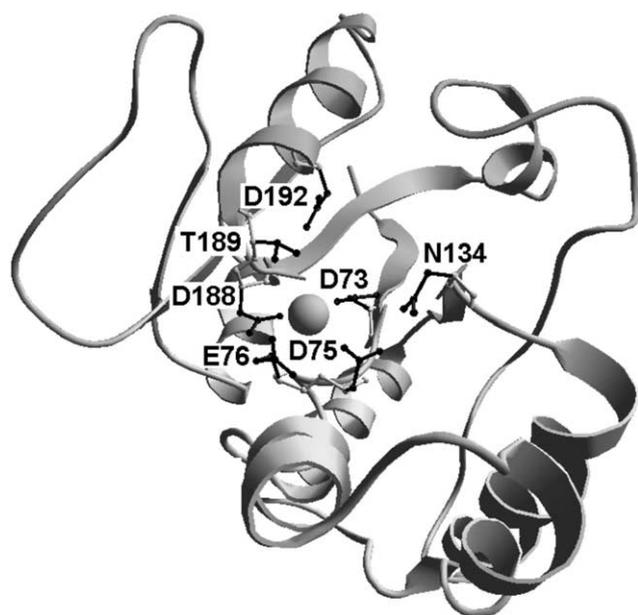


Fig. 2. The predicted three-dimensional structure of HppA and the architecture of the putative phosphatase active site. The putative catalytic/binding residues are shown in ball-and-stick representation and labeled; the metal ion serving as a cofactor is shown as a sphere.

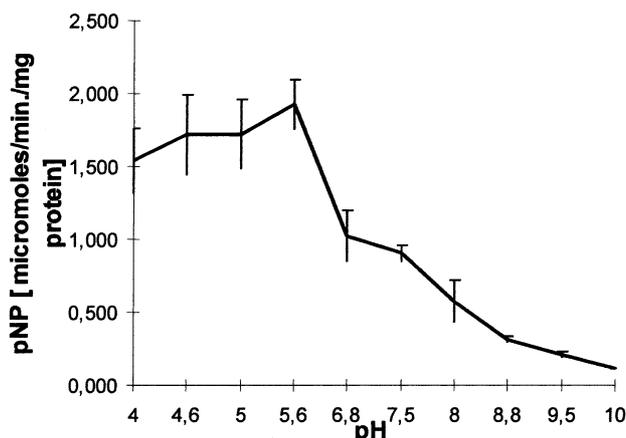


Fig. 3. pNPP hydrolysis activity of *E. coli*pUWM192 whole-cell protein extracts at different pH values.

of the activity level of wild-type cells. A greater difference was observed between two examined *E. coli* strains. *E. coli* pUWM192 was almost four times more active than the *hppA*-free control strain.

In order to validate the structural model and provide further insight into the function of HppA at the molecular level, we carried out alanine mutagenesis of predicted catalytic/metal binding residues D73 and D192.

Whole-cell extracts from cultures of D73A and D192A mutants were analyzed for specific phosphatase activity (Table 1). Mutants possessed activity at approximately the same level as negative control (*E. coli*pBluescript II KS). Western blot analysis using specific anti-HppA antibodies showed that *E. coli* containing wild-type and mutated versions of the *hppA* gene produced comparable amounts of HppA protein (data not shown). Thus, aspartic residues in positions D73 and D192 of HppA protein are required for phosphatase activity, thereby validating the structural model.

The whole-cell protein preparations from *E. coli*pUWM192 were used as a source of HppA phosphatase to study some functional properties of the enzyme. When pNPP was used as substrate, the HppA enzyme showed a pH optimum around 5.5, retaining more than 50% of the maximal activity in the pH range from 4 to 7 (Fig. 3). Enzyme activity was not inhibited by 15 mM EDTA (data not shown).

In this study we provided evidence for the fact that in accordance with in silico conducted analyses, the *hppA* gene product exhibits acid phosphatase activity and is a class C NSAP. Mutagenesis of the predicted catalytic residues, common with other phosphatases from the NSAP family, abolished the enzymatic activity of HppA.

Enzyme activity elaborated by *E. coli*pUWM192 exceeded

Table 1  
Specific acid phosphatase activity of *H. pylori* and *E. coli* whole cells as measured by the release of pNP from pNPP

Strain	pNP released (μmol/min/mg protein)
<i>H. pylori</i> N6	0.92 ± 0.33
<i>H. pylori</i> N6/ <i>hppA</i> ::Km	0.49 ± 0.10
<i>E. coli</i> pUWM192	1.92 ± 0.42
<i>E. coli</i> pBluescript II SK	0.44 ± 0.13
<i>E. coli</i> pUWM314	0.12 ± 0.02
<i>E. coli</i> pUWM315	0.13 ± 0.05

Average activity of five independent samples for each clone is shown with standard deviation.

the activity of the control strain only about four times, although the gene was cloned into a high copy number vector (pBluescript II KS), whereas the *S. equisimilis lppC* gene (cloned into pACYC184) product was 14.8 times more active than the lppC-free control [4]. Both analyzed genes were expressed in *E. coli* from their authentic promoters. The divergence of these results may reflect a difference in affinity of the *E. coli* RNA polymerase for the individual gene promoters. Even though the knowledge concerning *H. pylori* promoter consensus sequences is incomplete, it is documented that not all *H. pylori* promoters are efficiently recognized by *E. coli* transcription machinery [24]. Consistent with that, we were unable to demonstrate phosphatase activity of *E. coli* pUWM192 by zymographic techniques (data not shown).

The *hpaA* coding sequence contains two putative ATG start codons located at a distance of 45 nucleotides. Although the sequence of the HppA N-terminus has not been determined, membrane localization of the protein strongly suggests that the first ATG codon is recognized as a translation start.

The production of the cytoplasm-located HppA seems to be toxic for *E. coli* cells. Induction of the recombinant gene expression by IPTG was found to lead to rapid growth cessation (data not shown).

The present results indicate that in *H. pylori* HppA protein is localized to the outer membrane. Computational analysis of the signal peptide revealed that the putative peptidase cleavage site (LNAK↓AC) differs strongly from the consensus sequence of the lipoprotein signal cleavage site (LV(A,G)↓C) [25]. One potential reason for this discrepancy could be differential localization of HppA of *H. pylori* and *E. coli*. On the other hand, numerous attempts at *H. pylori* cell fractionation revealed problems with obtaining clear fractions. Some proteins which are cytoplasmic in other species may leak out of the *Helicobacter* cell and become associated with the surface [26,27]. Moreover, the inner and outer membranes of *H. pylori* interact tightly, preventing absolute separation of fractions [28].

We also failed to overproduce HppA as a fusion with GST or His tags added to the C-terminus of the mature protein. Constructs were genetically correct as verified by DNA sequencing. In this case IPTG induction did not influence the cell growth rate, nevertheless recombinant proteins were not produced. The cloned *hpaA* heterologous gene differs markedly in codon bias from that of *E. coli*. Codon usage in *H. pylori* is characterized by a relative lack of heterogeneity among genes, and in particular an absence of differentiation between highly expressed genes and others [29]. In contrast, *E. coli* has a set of major codons that occur in highly expressed genes and a set of rare codons in genes expressed at a low level [30]. Within the *hpaA* gene, 15.5% of the amino acids are encoded by codons used by *E. coli* at a frequency of <1%. Especially usage of the AGG/AGA, CUA, AUA, CGA, CCC cluster of codons, which consist of almost 6% of *hpaA* codons and presence of lysine residues constitutes

13% of HppA, being three times as numerous as in average *E. coli* proteins, could create a translation problem, and reduce the quantity and quality of synthesized protein [31].

## References

- [1] Warren, J.R. and Marshall, B. (1983) *Lancet* i, 1273–1275.
- [2] Dunn, B.E., Cohen, H. and Blaser, M.J. (1997) *Clin. Microbiol. Rev.* 10, 720–741.
- [3] Lazowska, I., Trzeciak, L., Godlewska, R., Hennig, E., Jaguszyn Krynicka, K., Popowski, J., Regula, J. and Ostrowski, J. (2000) *Digestion* 61, 14–21.
- [4] Malke, H. (1998) *Appl. Environ. Microbiol.* 64, 2439–2442.
- [5] Beacham, I.R. (1979) *Int. J. Biochem.* 10, 877–883.
- [6] Thaller, M.C., Schippa, S., Bomci, A., Cresti, S. and Rossolini, G.M. (1997) *FEMS Microbiol. Lett.* 146, 191–198.
- [7] Reilly, T.J., Baron, G.S., Nano, F.E. and Kuhlenschmidt, M.S. (1996) *J. Biol. Chem.* 108, 545–552.
- [8] Galperin, M.Y. and Koonin, E.V. (1999) *Genetica* 106, 159–170.
- [9] Rossolini, G.M., Schippa, S., Riccio, M.L., Berlutti, F., Macaskie, L.E. and Thaller, M.C. (1998) *Cell. Mol. Life Sci.* 54, 833–850.
- [10] Aravind, L. and Koonin, E.V. (1998) *Trends Biochem. Sci.* 23, 469–472.
- [11] Sambrook, J. and Russell, D.W. (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [12] Laemmli, U.K. (1970) *Nature* 227, 680–695.
- [13] Blaser, M.J., Hopkins, J.A., Berka, R.M., Vasil, M.L. and Wang, W.L. (1983) *Infect. Immun.* 42, 276–284.
- [14] Filip, C., Fletcher, G.L.W.J. and Earhart, C.F. (1973) *J. Bacteriol.* 175, 966–972.
- [15] Ames, G.F., Prody, C. and Kustu, S. (1984) *J. Bacteriol.* 160, 1181–1183.
- [16] Kier, L.D., Weppelman, R. and Ames, B.N. (1977) *J. Bacteriol.* 130, 399–410.
- [17] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) *Nucleic Acids Res.* 25, 3389–3402.
- [18] Thompson, J., Gibson, T., Plewniak, F., Jeanmougin, F. and Higgins, D. (1997) *Nucleic Acids Res.* 25, 4876–4882.
- [19] Bujnicki, J.M., Elofsson, A., Fischer, D. and Rychlewski, L. (2001) *Protein Sci.* 10, 352–361.
- [20] Lundstrom, J., Rychlewski, L., Bujnicki, J. and Elofsson, A. (2001) *Protein Sci.* 10, 2354–2362.
- [21] Guex, N. and Peitsch, M.C. (1997) *Electrophoresis* 18, 2714–2723.
- [22] Luthy, R., Bowie, J.U. and Eisenberg, D. (1992) *Nature* 356, 83–85.
- [23] Morais, M., Zhang, W., Baker, A., Zhang, G., Dunaway-Mariano, D. and Allen, K. (2000) *Biochemistry* 39, 10385–10396.
- [24] Spohn, G., Beier, D., Rappuoli, R. and Scarlato, V. (1997) *Mol. Microbiol.* 26, 361–372.
- [25] Pugsley, A.P. (1993) *Microbiol. Rev.* 57, 50–108.
- [26] Doig, P., Austin, J.W. and Trust, T.J. (1993) *J. Bacteriol.* 175, 557–560.
- [27] Spiegelhalter, C., gerstenecker, B., Kersten, A., Schiltz, E. and Kist, M. (1993) *Infect. Immun.* 61, 5315–5325.
- [28] Doig, P. and Trust, T.J. (1994) *Infect. Immun.* 62, 4526–4533.
- [29] Lafay, B., Atherton, J.C. and Sharp, P.M. (2000) *Microbiology* 146, 851–860.
- [30] Zhang, S. and Zubay, G.E.G. (1991) *Gene* 105, 61–72.
- [31] Wada, K., Wada, Y., Ishibashi, F., Gojobori, T. and Ikemura, T. (1992) *Nucleic Acids Res.* 20, 2111–2118.