FEBS Letters 361 (1995) 238–242 FEBS 15261

Cloning and sequencing of the bacterioferritin gene of *Brucella melitensis* 16M strain

Philippe A. Denoel[a,*], Michel S. Zygmunt[b], Vincent Weynants[a], Anne Tibor[a], Bernadette Lichtfouse[a], Pascal Briffeuil[c], Joseph N. Limet[a], Jean-Jacques Letesson[a]

[a]*Laboratoire de Microbiologie et d'Immunologie, Facultés Universitaires Notre-Dame de la Paix, 61 rue de Bruxelles, B-5000 Namur, Belgium*
[b]*Laboratoire de Pathologie Infectieuse et Immunologie, Institut National de la Recherche Agronomique, 37380 Nouzilly, France*
[c]*Laboratoire de Biologie Quantitative, Facultés Universitaires Notre-Dame de la Paix, B-5000 Namur, Belgium*

Received 13 February 1995

Abstract The 40 N-terminal amino acids of the 20 kDa antigen A2 from *Brucella melitensis* were sequenced and showed important similarities with 4 bacterioferritins. A monoclonal antibody raised against this antigen cross-reacted with *Escherichia coli* bacterioferritin. Hybridization of two sets of degenerate primers with *B. melitensis Hin*dIII-digested genomic DNA identified a 3.8 kb fragment. This fragment was shown to contain a bacterioferritin gene (*bfr*) encoding a 161-amino acid protein. The sequence of the *Brucella* bacterioferritin is 69% similar to that of *E. coli*, and many of the ferroxidase centre and haem-ligation residues are conserved.

Key words: Bacterioferritin; Antigen A2; Brucellosis; *Brucella melitensis*

1. Introduction

Iron is an essential element but can be toxic in excess, and most micro-organisms have developed transport and storage systems for the uptake and the deposition of this element in an intracellular non-toxic form. *Brucella* spp., the agents of brucellosis, are intracellular facultative pathogens for which a correlation between iron metabolism and virulence has been postulated [1,2]. The acquisition of iron by the *B. abortus* siderophore (2,3-dihydroxybenzoic acid) has been the subject of a recent study [3], but no iron storage system has yet been reported. Recently, Zygmunt et al. [4] purified a 20 kDa protein previously described as antigen A2 [5,6]. Six of the seven N-terminal residues of this protein are identical to six of the seven N-terminal residues of *Nitrobacter winogradskyi* bacterioferritin (BFR) [7] and five are identical to five of the seven N-terminal residues of *Escherichia coli* BFR [8]. BFR is an iron storage haemoprotein composed of 24 subunits assembled into a spherical protein shell containing approximately 12 haems per 24 subunits [9–11] and has been detected in many species [12–16].

In this paper, we report cross-reaction of a monoclonal antibody raised against the 20 kDa A2 antigen with *E. coli* BFR, the N-terminal amino acid sequencing of the 40 first residues of the protein (antigen A2), and the cloning of the *bfr* gene of *B. melitensis* 16M strain as well as its complete nucleotide sequence.

*Corresponding author. Fax: (32) (81) 724 420.
E-mail: PDenoel@cc.fundp.ac.be

2. Materials and methods

2.1. Bacterial strain

All whole-cell extracts of *Brucella* strains were prepared by J.-M. Verger and M. Grayon, Institut National de la Recherche Agronomique, Laboratoire de Pathologie Infectieuse et d'Immunologie, Nouzilly, France. Briefly, cell lysis was achieved by boiling in the presence of 1% sodium dodecyl sulfate (SDS).

2.2. Preparation of the 20 kDa antigen

The *B. melitensis* 20 kDa protein (antigen A2) was prepared as previously described [4]. Briefly, a cytoplasmic protein extract from *B. melitensis* was fractionated by HiLoad Superdex 200 chromatography and the fractions containing the 20 kDa protein were separated by SDS-PAGE followed by electroelution of the 20 kDa protein from the gel.

2.3. Monoclonal antibodies (mAbs) and antiserum

BALB/c mice were immunized by 2 intraperitoneal injections of 200 μg of 20 kDa protein in Freund's complete adjuvant given 1 week apart. After 2 months, mice were boosted intraperitoneally with 200 μg of the 20 kDa protein. Three days after the boost injection, spleen cells were fused, as previously described [17], and hybridomas producing anti-20 kDa mAbs were screened by immunoblotting using cytosoluble protein extract as antigen [18]. Among the positive clones, the Z297C3 clone was selected and ascitic fluid was produced in BALB/c mice as previously described [19].

Anti-*E. coli* BFR serum was provided by S.C. Andrews (The Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK).

2.4. N-Terminal sequence analysis

Edman recurrent degradation was performed on an automated stepwise sequencer with an on-line 120A phenylthiohydantoin analyzer (Applied Biosystems). The sample was first desalted by reverse-phase HPLC on a 7-μm 300A Aquapore RP 300 column (Brownlee).

2.5. SDS-PAGE analysis and immunoblotting

Recombinant *E. coli* BFR was provided by S.C. Andrews and was prepared as previously described [20]. Proteins were separated in 16% acrylamide SDS-PAGE slab gels and were then either stained with Coomasie brilliant blue or transferred electrophoretically to 0.2-μm pore-size nitrocellulose for immunoblotting. Immunoblotting with mAb Z297C3 and anti-*E. coli* BFR serum was performed as previously described [21].

2.6. Recombinant DNA techniques

Routine recombinant DNA techniques were performed according to Sambrook et al. [22]. *Escherichia coli XL1-blue* strain (Stratagene) was used as the host for recombinant plasmids. Restriction endonucleases and modifying enzymes (Boehringer) were used as described by the manufacturer.

Brucella melitensis 16M strain genomic DNA was obtained from J.-M. Verger and M. Grayon. Oligonucleotides were synthesized at Eurogentec and were radiolabelled using T4 polynucleotide kinase (Boehringer) and [γ-^{32}P]ATP (Amersham). Southern blotting on positively charged nylon membranes (Amersham) was performed with a Hybaid Vacuum Blotter (Biozym).

SSDI 0014-5793(95)00189-1

Sequencing was performed using the Sequenase 2.0 kit (Pharmacia) according to the manufacturer's instructions.

The nucleotide sequence data reported in this paper were submitted to GenBank and were assigned the accession number U19760.

2.7. DNA and protein sequence analysis

Sequence data obtained from sequencing gels were compiled and analyzed by the DNA Strider 1.2 program [23]. FastA, TFastA, Terminator and Isoelectric programs were used with the Genetics Computer Group Sequence Analysis Software Package version 8.0-OpenVMS or 7.3-UNIX. The GenBank, EMBL and SWISS-PROT nucleic acid and protein sequences databases were used for homology searches. The bacterioferritin sequences were analyzed with the Match-Box software for the simultaneous alignment of several protein sequences [24,25]. Using this approach, a comparison of all 8-residues segments of sequence allows the analysis of similarity between and within groups of sequences according to different statistical thresholds (cut-off distances).

3. Results and discussion

3.1. Western blot analysis of the B. melitensis *20 kDa protein and* E. coli *bacterioferritin*

Immunoblot analysis with the anti-20 kDa mAb (Z297C3) following SDS-PAGE of *B. melitensis* 16M strain whole-cell extracts, purified 20 kDa protein, and *E. coli* recombinant BFR (rBFR) revealed cross-reaction of the anti-20 kDa mAb with *E. coli* rBFR (Fig. 1). In the whole-cell extract, a band of higher molecular mass (38–40 kDa) was detected. This band could consist of dimers of the 20 kDa protein. Though only very weak cross-reaction of anti-*E. coli* BFR serum was observed with the *B. melitensis* 20 kDa protein (data not shown), the above results indicate that the 20 kDa protein (antigen A2) could well be a bacterioferritin, as already suggested by the limited N-terminal sequence data.

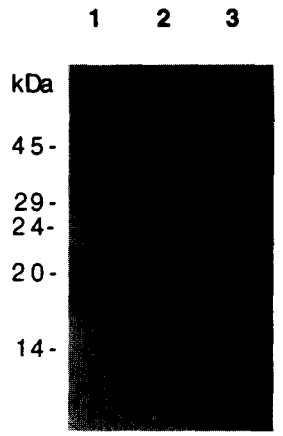

Fig. 1. Immunoblot of *B. melitensis* whole-cell extract (lane 1), purified 20 kDa protein (lane 2) and *E. coli* recombinant bacterioferritin with anti-20 kDa mAb Z297C3. Molecular weight (MW) markers are indicated on the left.

Table 1
Brucella species, biovars and strains tested in immunoblotting with the anti-20 kDa protein mAb

Species	Biovar	Strain	20-kDa protein detected
B. abortus	1	544	+
	1	B19	+
	2	86/8/59	+
	3	Tulya	+
	4	292	+
	5	B 3196	+
	6	870	+
	6	80-236	+
	6	89-43	+
	9	87-46	+
	9	90-64	+
	9	91-135	+
	9	76-299	+
	9	75-60	+
	9	80-133	+
	9	77-9	+
	9	C 68	+
	9	91-28	+
	Rough	45/20	+
B. melitensis	1	16M	+
	1	Rev1	+
	2	63/9	+
	3	Ether	+
	Rough	B115	+
	Rough	H38R	+
B. suis	1	1330	+
	1	S2	+
	2	Thomsen	+
	3	686	+
	4	40	–
	5	513	–
B. canis		RM 6/66	+
B. ovis		BOW 63/290	+
B. neotomae		5K33	+

3.2. Expression of the 20 kDa protein in the genus Brucella

Thirty-four *Brucella* strains, including all six of the known species and all reported biovars, were examined for expression of the 20 kDa protein by SDS-PAGE and immunoblotting, as described above. The anti-20 kDa protein mAb revealed a band in all *Brucella* strains tested except for *B. suis* biovars 4 and 5 (Table 1). These data indicate that the protein is expressed and that the apparent size of the antigen A2 (20-kDa protein) is identical among all strains except for *B. suis* biovars 4 and 5 for which no expression of the protein was detected.

3.3. N-Terminal sequence analysis

The sequence of the first 40 N-terminal residues of the 20 kDa protein was obtained by automated Edman degradation, and is shown in Fig. 2. Sequence similarity searches were performed with the FastA algorithm and showed that this N-terminal sequence is very similar (45–65% identity) to the N-terminal sequences of seven bacterioferritins from the following strains: *Nitrobacter winogradskyi* [7], *E. coli* [8], *Azotobacter vinelandii* [26], *Synechocystis* spp. [13], *Mycobacterium paratuberculosis*, *M. avium*, and *M. leprae* [14–16].

3.4. Cloning of the bacterioferritin gene

Two sets of degenerate primers were derived by back translation of the N-terminal protein sequence, spanning residues 1–9 (sp1d) and 23–31 (sp76d) (Fig. 2).

Southern blot analysis of *B. melitensis* 16M strain *Hin*dIII digested genomic DNA with the radiolabelled sp1d set of primers as probe, revealed hybridization of the primers with a 3.8 kb band. Similar results were obtained with the sp76d probe set. The *Hin*dIII genomic DNA fragments of molecular weight range spanning the fragment of interest were recovered from

```
1                                   10                                  20
M  K  G  E  P  K  V  I  E  R  L  N  E  A  L  F  L  E  L  G
5' ATGAARGGNGARCCNAARGTNATHGA 3' (spld)

                                    30                                  40
A  V  N  Q  Y  W  L  H  Y  R  L  L  N  D  W  G  Y  T  R  L
 5'  AAYCARTAYTGGYTNCAYTAYAGRYA 3' (sp76d)
```

Fig. 2. N-Terminal amino acid sequence of *B. melitensis* 20 kDa antigen A2 aligned with the nucleotide sequences of the corresponding degenerate primers used for cloning (R = A or G; Y = C or T; H = A, C or T; N = A, C, G or T).

an agarose gel and used to construct a partial genomic library of *B. melitensis* 16M strain in the vector pBluescript KS−. *Hin*dIII-digested plasmid DNA from the recombinant clones was analysed by Southern blotting with the radiolabelled spld set of primers as probe. One clone (pCm69) was shown to contain a plasmid which was positive in Southern and dot blotting (data not shown).

3.5. DNA sequence analysis

The recombinant plasmid pCm69 was used to determine the sequence of the *B. melitensis* bacterioferritin gene (Bm *bfr*). Sequencing was initiated with the sp76d set of primers and was further continued by the primer walking strategy. The nucleotide sequence of the Bm *bfr* gene as well as the amino acid sequence of the encoded protein are shown in Fig. 3. A 486 bp open reading frame (ORF) (positions 209–692) was identified. There is a potential Shine–Dalgarno ribosome-binding site (GGAG) from −12 to −9 bases upstream from the ATG, and a GC-rich region of dyad symmetry identified 20 bases downstream from the stop codon could function as rho-independent transcription terminator [21].

Since *Brucella* genes are often expressed in *E. coli* from their own promoter, revealing a high degree of similarity to published *E. coli* promoters [21], a promoter search was performed using the Staden promoter prediction computer program [27]. There is a weakly predicted promoter at position 162 and a good predicted promoter at position 46. The expression of a number of *E. coli* genes concerned with iron uptake is known to be modulated with iron and this is, in many cases, mediated by the regulatory protein Fur [28]. As for the *E. coli bfr* gene, no iron box resembling the consensus sequence (GATAATGATAATCATTATC) was detected [29]. However, no Fur-like protein has been identified in *Brucella* and another regulation system is then conceivable. Moreover, in *Brucella* an uptake system in which only the siderophore is regulated by the iron available for growth has been suggested [3].

The deduced protein is a 161 amino acids peptide with a predicted molecular mass of 18,647 Da and a calculated pI of 4.53; both values are in agreement with the experimental results obtained by Zygmunt et al. [4]. There is one difference between the N-terminal sequence obtained by Edman degradation and the deduced protein sequence (Glu-13 is in place of Asp-13 in the deduced protein sequence).

3.6. Multiple protein sequence analysis

Similarities between the deduced Bm BFR amino acid sequence and the sequences available in the databases were searched for with the FastA and TFastA algorithms. The Bm BFR sequence shows significant similarity (optimized score >100) to the BFR sequences of *N. winogradskyi* [7], *E. coli* [8], *A. vinelandii* [26], *Synechocystis* sp [13], *M. paratuberculosis*, *M. avium*, and *M. leprae* [14–16].

To investigate the extent of conserved and unconserved

```
                  -35                          -10      +1
1   AAGATTGCTGGCAACTTATTGTGCCGCGCTGTCTATAATGCCATGAATCA

51  ACGCGGCCGTTGCTGTGGCTGCTTCCCAAATGTCGTAGAAACCATCATAA

                                  -35
101 GGGTGACTGAAGAATACTATCTTCGTCGCAACCAAATGGACGAAAACGTT

    -10      +1
151 TAATTCCAGTTCATGGATCGTGTGCGTTCTCTACGAGATAAATTCGGGAG

 1          M  K  G  E  P  K  V  I  E  R  L  N  D  A
201 TTCATGGAATGAAAGGCGAACCAAAGGTCATCGAGCGGCTTAACGATGCA

15  L  F  L  E  L  G  A  V  N  Q  Y  W  L  H  Y  R  L
251 CTGTTTCTTGAGCTCGGTGCGGTAAACCAGTATTGGCTGCACTACCGTCT
          SacI

32   L  N  D  W  G  Y  T  R  L  A  K  K  E  R  E  E  S
301 TCTCAACGATTGGGGTTACACGCGCCTTGCAAAGAAGGAACGCGAGGAAT

49    I  E  E  M  H  H  A  D  K  L  I  N  R  I  I  F
351 CCATCGAGGAAATGCATCACGCCGACAAGCTGATTAATCGCATTATCTTC

65  F  E  G  F  P  N  L  Q  T  V  S  P  L  R  I  G  Q
401 TTTGAAGGCTTTCCGAACCTCCAGACCGTTTCGCCGTTGCGCATTGGCCA

82   N  V  K  E  V  L  E  A  D  L  K  G  E  Y  D  A  R
451 GAATGTGAAGGAAGTTCTCGAAGCTGACCTCAAGGGTGAATATGACGCTC

99    A  S  Y  K  E  S  R  E  I  C  D  K  L  G  D  Y
501 GCGCTTCGTATAAGGAATCGCGCGAAATCTGCGACAAGCTCGGCGACTAT

115 V  S  K  Q  L  F  D  E  L  L  A  D  E  E  G  H  I
551 GTGTCGAAGCAGCTTTTCGACGAACTTCTGGCCGATGAAGAAGGCCATAT

132  D  F  L  E  T  Q  L  D  L  L  A  K  I  G  E  E  R
601 CGACTTCCTTGAAACCCAGCTCGACCTTCTCGCCAAGATCGGCGAAGAAC

149   Y  G  Q  L  N  A  A  P  A  D  E  A  E  *
651 GCTATGGCCAGCTTAACGCGGCGCCCGCCGACGAAGCTGAGTAAGCCTGT

701 TTCAATCTGTCTTGAAAGGCGGGGGCGCATCGCTCCGGCATTTTCTTTGG
              <<< << <<<<<<  >>>>>> >> >>>
```

Fig. 3. Nucleotide sequence and deduced amino acid sequence of the *bfr* gene of *B. melitensis* and of the corresponding protein. The putative promoters are underlined. The putative ribosome-binding site is indicated in bold and boxed. The dyad symmetry of the putative terminator stem-loop structure is underlined (<<<>>>). The asterisk denotes the termination codon.

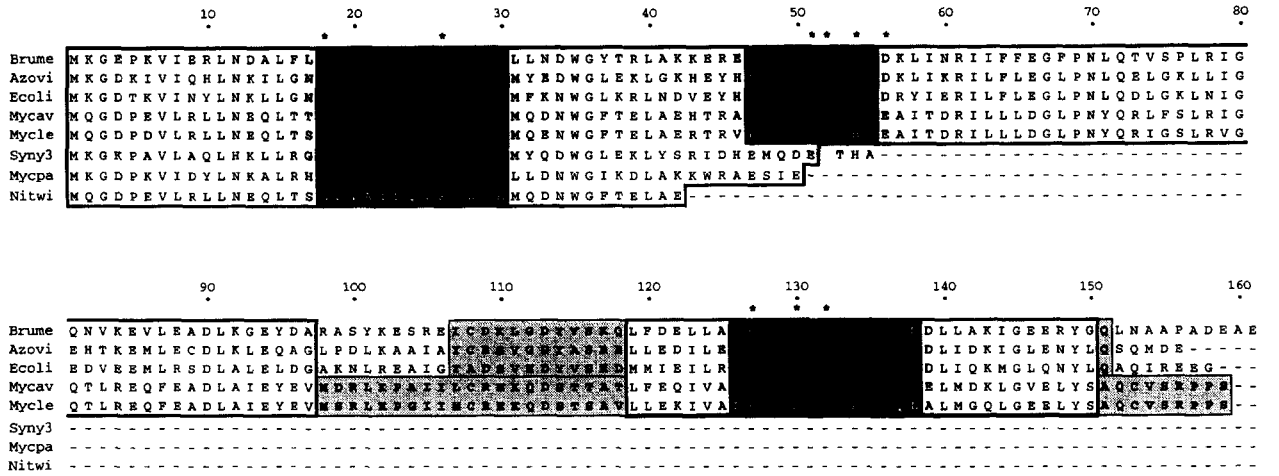

Fig. 4. Simultaneous multiple alignment of BFR sequences of *B. melitensis* (Brume), *A. vinelandii* (Azovi), *E. coli* (Ecoli), *M. avium* (Mycav), *M. Leprae* (Mycle), *Synechocystis* sp. (Syny3), *M. paratuberculosis* (Mycpa) and *N. winogradskyi* (Nitwi). Boxes delineate the matching regions. Black boxes outline matching regions in the whole set of sequences (when sequences are complete) at a severe statistical threshold (400). Open boxes outline matching regions at a tolerant statistical threshold (520). Shaded boxes outline matching regions within groups of sequences (statistical threshold = 520). *Residues mentioned in the text.

sequence regions between Bm BFR and these proteins, simultaneous multiple alignment was performed using the Match-Box package [24,25] with the Blosum 62 amino acid substitution matrix [30]. At a relatively severe statistical threshold (cut-off distance = 400), three boxes (black) are highlighted (Fig. 4). These boxes correspond to the most similar regions in all sequences. These regions enclose several totally conserved amino acids and some of them are involved in the binuclear metal binding site [11]. For instance, Glu-51, Glu-127, and His-130 are fully conserved, and Glu-18 is replaced by a Gln in the *M. avium* sequence. Two of the three residues directly involved in the haem binding site [11] (residues 26 and 52) are also enclosed in these boxes. The third residue involved in the haem binding site, Asp-56, is enclosed in statistically similar regions delineated at a tolerant threshold (open boxes, cut-off distance = 520). Residues of the pseudo-twofold symmetry related subunits described by Frolow [11] as having van der Waals contacts (<4 Å) with haem are enclosed in boxes (except for Arg-45, Asp-56, and Leu-71) and are identical or very similar.

The region comprising residues 98–118 appears to be the most variable, as it can only be delineated at a very tolerant threshold (cut-off distance = 585). This 98–118 variable region corresponds to the end of the C helix, a loop and the beginning of the D helix in the *E. coli* BFR [11,29], and does not contain residues of apparent importance for haem or iron binding. However, it is noteworthy that, within this highly variable region, Asp-113 is the only residue to be fully conserved.

The relevance of *Brucella* BFR as an antigen for purposes of diagnosis has not yet been explored in detail. However, antigen A2 was originally identified with sera from infected ruminants [4,6]. Therefore, BFR is an immunodominant B-cell antigen which could be used in a new brucellosis diagnostic test.

Over-expression of the Bm BFR will enable us to provide functional evidence for the role of the protein and to estimate the number of haem per molecule. Regarding both the proposed involvement of BFR in the safe storage of iron and the importance of iron in the *Brucella*–macrophage interaction [2], and considering its identification and proposed significance in *M. avium*, another intracellular pathogen, it would be of interest to determine the role of the protein in the virulence of *Brucella*.

Acknowledgements: We would like to thank Simon Andrews for stimulating helpful discussions and good advice. We are very grateful to D. Prozzi, X. De Bolle and D. Gilson who provided encouragement and help for the realization of this work. We are grateful to J.-M. Verger and M. Grayon for providing *Brucella* strain extracts. Philippe Denoel holds a specialization bursary from the IRSIA/FRIA (no. 910622).

References

[1] Waring, W.S., Elberg, S.S., Schneider, P. and Green, W. (1953) J. Bacteriol. 66, 82–89.
[2] Baldwin, C.L., Jiang, X. and Fernandes, D.M. (1993) Trends Microbiol. 1, 99–104.
[3] Lòpez-Goni, I., Moriyòn, I. and Neilands, J.B. (1992) Infect. Immun. 60, 4496–4503.
[4] Zygmunt, M.S., Gilbert, F.B. and Dubray, G. (1992) J. Clin. Microbiol. 30, 2662–2667.
[5] Stemshorn, B., Nielsen, K. and Samagh, B. (1981) Can. J. Comp. Med. 45, 77–81.
[6] Schurig, G.G. (1982) Vet. Immunol. Immunopathol. 3, 311–324.
[7] Kurokawa, T., Fukumori, Y. and Yamanaka, T. (1989) Biochim. Biophys. Acta 976, 135–139.
[8] Andrews, S.C., Smith, J.M.A., Guest, J.R. and Harrison, P.M. (1989) Biochem. Biophys. Res. Commun. 158, 489–496.
[9] Yariv, J., Kalb, J., Sperling, R., Bauminger, E.R., Cohen, S.G. and Ofer, S. (1981) Biochem. J. 197, 171–175.
[10] Andrews, S.C., Smith, J.M.A., Guest, J.R. and Harrison, P.M. (1990) Biochem. Soc. Trans. 18, 658–659.
[11] Frolow, F., Kalb (Gilboa), A.J. and Yariv, J. (1994) Nature Struct. Biol. 1, 453–460.
[12] Andrews, S.C., Arosio, P. ,Bottke, W. ,Briat, J.F. , von Darl, M. ,Harrison, P.M., Laulhere, J.P., Levi, S., Lobreaux, S. and Yewdall, S.J. (1992) J. Inorg. Biochem. 47, 161–174
[13] Laulhere, J.P., Laboure, A.M., Van Wuytswinkel, O., Gagnon, J. and Briat, J.F. (1992) Biochem. J. 281, 785–793.
[14] Brooks, B.W., Young, N.M., Watson, D.C., Robertson, R.H., Sugden, E.A., Nielsen, K.H. and Becker, S.A.W.E. (1991) J. Clin. Microbiol. 29, 1652–1658.
[15] Pessolani, M.C.V., Smith, D.R., Rivoire, B., McCormick, J., Hefta, S.A., Cole, S.T. and Brennan, P.J. (1994) J. Exp. Med. 180, 319–327.

[16] Inglis, N.F., Stevenson, K., Hosie, A.H.F. and Sharp, J.M. (1994) Gene 150, 205–206.
[17] Cloeckaert, A., de Wergifosse, P., Dubray, G. and Limet, J.N. (1990) Infect. Immun. 58, 3980–3987.
[18] Zygmunt, M.S., Martin, J.C. and Dubray, G. (1990) FEMS Lett. 70, 263–268.
[19] Limet, J.N., Plommet, A.M., Dubray, G. and Plommet, M. (1987) Ann. Inst. Pasteur Immunol. 138, 417–424.
[20] Andrews, S.C., Smith, J.M.A., Hawkins, C., Williams, J.M., Harrison, P.M. and Guest, J.R. (1993) Eur. J. Biochem 213, 329–338.
[21] Tibor, A., Weynants, V., Denoel, P., Lichtfouse, B., De Bolle, X., Saman, E., Limet, J.N. and Letesson, J.J. (1994) Infect. Immun. 62, 3633–3639.
[22] Sambrook J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
[23] Marck, C. (1988) Nucleic Acids Res. 16, 1829–1836.
[24] Depiereux, E. and Feytmans, E. (1991) Protein Eng. 4, 603–613.
[25] Depiereux, E. and Feytmans, E. (1992) CABIOS 8, 501–509.
[26] Andrews, S.C., Findlay, J.B.C., Guest, J.R., Harrison, P.M., Keen, J.N. and Smith, J.M.A. (1991) Biochim. Biophys. Acta 1078, 111–116.
[27] Staden, R. (1984) Nucleic Acids Res. 12, 521–538.
[28] Hantke, K. (1981) Mol. Gen. Genet. 182, 288–292.
[29] Andrews, S.C., Harrison, P.M. and Guest, J.R. (1989) J. Bacteriol. 171, 3940–3947.
[30] Henikoff, S. and Henikoff, J.G. (1992) Proc. Natl. Acad. Sci. USA 89, 10915–10919.