

Utilization of alkaline phosphatase fusions to identify secreted proteins, including potential efflux proteins and virulence factors from *Helicobacter pylori*

James E. Bina^a, Francis Nano^b, Robert E.W. Hancock^{a,*}

^a Department of Microbiology and Immunology, University of British Columbia, Vancouver, B.C. V6T 1Z3, Canada

^b Department of Biochemistry and Microbiology, University of Victoria, Victoria, B.C. V8W 2Y2, Canada

Received 19 December 1996; revised 8 January 1997; accepted 8 January 1997

Abstract

The targeted genomic strategy of random fusions to a partial gene encoding a signal sequence-deficient fragment of bacterial alkaline phosphatase was utilized to screen for secreted proteins in *Helicobacter pylori*. The rationale for targeting extracytoplasmic proteins was based on the hypothesis that most virulence factors and vaccine candidates are secreted or exported proteins. In addition, extracytosolic proteins represent good potential targets for drug intervention since they are in general more accessible to drugs than are cytoplasmically localized proteins. The application of this strategy to *H. pylori* allowed the identification of putative virulence factors and novel targets for drug intervention including four putative antibiotic efflux genes. The strategy used here is rapid and technically simple, relatively inexpensive, adaptable to a wide variety of microbes and genetic systems, and selects for expressed and accessible proteins.

Keywords: *Helicobacter pylori*; Secreted protein; Virulence factor; Alkaline phosphatase

1. Introduction

Currently several bacterial chromosomes have been fully or substantially sequenced including those of *Escherichia coli*, *Haemophilus influenzae* [1], *Helicobacter pylori* (not publicly available), *Mycobacterium leprae*, *Mycobacterium tuberculosis* (not publicly available), *Mycoplasma genitalium* [2] and *Staphylococcus aureus* (not publicly available). One reason for sequencing the genomes of these organisms is to identify genes of therapeutic interest such as viru-

lence factors, vaccine candidates and novel drug targets. An alternative strategy for the identification of genes of potential therapeutic interest is to target extracellular, surface associated or periplasmic proteins. The rationale for targeting this subset of proteins is that most proteins of therapeutic interest are exported or secreted proteins and such proteins are more accessible to drug therapy than are cytoplasmic proteins.

Gene fusion technology based on translational fusions to alkaline phosphatase have been successfully used for the identification of extracellular, surface associated or periplasmic proteins in several bacterial systems [3]. However, the application of gene fusion

* Corresponding author. Tel.: +1 (604) 822 2682; fax: +1 (604) 822 6041; e-mail: bob@cbsd.ca

technology to bacterial species that contain undefined restriction and modification systems, such as *H. pylori*, is more difficult. Attempts at traditional *TnphoA* mutagenesis [6] of *H. pylori* in our lab have not been successful. Therefore we sought to adapt a rapid and simple system for the identification of *H. pylori* secreted proteins in a heterologous system. Recently a gene fusion system based on gene fusion to β -lactamase has been applied to *H. pylori* [4]. In this paper we describe a strategy for the identification of secreted proteins from *H. pylori* based on gene fusions to bacterial alkaline phosphatase, an enzyme that only becomes active after passage across the cytoplasmic membrane.

The approach presented here is based on a set of phagemid vectors that allow the translational fusion of genes to a signal sequence-deficient *E. coli* alkaline phosphatase gene (*'phoA*) [5]. Expression of alkaline phosphatase in these vectors is dependent upon the cloning of an in frame signal sequence upstream of *'phoA*. In these vectors the *'phoA* gene is cloned downstream and in the same orientation as the *lac* promoter which is important for the transcription of genes that contain weak promoters or for the expression of internal genes in polycistronic operons [4]. Expression of alkaline phosphatase activity is detected on solid media containing a colorimetric indicator. The application of this targeted genomics strategy to *H. pylori* allowed the identification of several novel targets for drug intervention and potential virulence factors.

2. Materials and methods

2.1. Bacterial strains, plasmids and growth conditions

Escherichia coli ER1739 (New England BioLabs) and *E. coli* CC118 [6] were used for the construction of the *H. pylori*-alkaline phosphatase fusion library. Plasmids pJDT1, pJDT2 and pJDT3 have been previously described [5]. *H. pylori* National Culture and Tissue Collection (NCTC) strain 11767 was grown as previously described [7]. Recombinant clones were screened for the expression of alkaline phosphatase on XP-PO₄ agar (LB agar, 75 mM sodium phosphate pH 7.2, 75 μ g/ml ampicillin and 90 μ g/ml 5-bromo-4-chloro-3-indolyl phosphate). Plasmid

GW25 was pRK404 containing the *Pseudomonas syringae* pv. *syringae* *oprF::TnphoA*.

2.2. Construction of a *Helicobacter pylori*-alkaline phosphatase fusion library

Chromosomal DNA from *H. pylori* 11637 was prepared by standard methods [8]. *H. pylori* chromosomal DNA was partially restricted with restriction enzyme *Sau3A1* and size fractionated on a continuous 10–40% sucrose gradient. DNA restriction fragments of approximately 2 kb in size were ligated into *Bam*HI digested and phosphatased pJDT1, pJDT2, and pJDT3. The ligation reactions were used for the transformation of either *E. coli* ER1793 or *E. coli* CC118. Alkaline phosphatase expressing transformants were identified on XP-PO₄ agar. The alkaline phosphatase expressing clones were screened for DNA inserts by the slot lysis method [9] and clones containing inserts were screened by DNA sequencing.

2.3. Detection of alkaline phosphatase activity in *Escherichia coli* ER1793

The concentration of phosphate in the growth media that was required to completely inhibit the background phosphatase activity in *E. coli* ER1793 was determined by plating *E. coli* ER1793 and ER1793/pGW25 onto XP agar containing 0, 25, 50 or 100 mM NaPO₄ (pH 7.2). The resulting bacterial colonies were examined following incubation for 24 h at 37°C for the expression of phosphatase activity. Under these conditions, *E. coli* ER1793 produced pale blue colonies on XP agar containing 0 and 25 mM NaPO₄ and white colored colonies on the media containing 50 and 100 mM NaPO₄. *E. coli* ER1793-pGW25 produced dark blue colored colonies under all the tested conditions.

The effect of the inclusion of 75 mM NaPO₄ (pH 7.2) in XP agar on the detection sensitivity for alkaline phosphatase activity in our screening method was determined by the following method. A serial 2-fold dilution of bacterial alkaline phosphatase was made in 50 mM Tris-HCl (pH 8.5), 1 mM EDTA and 2 μ l aliquots from each dilution were spotted onto the surface of both XP-PO₄ and XP agar. Following an 18 h incubation at 37°C the

plates were examined and the highest dilution of BAP that produced a visible blue color on the surface of the agar was recorded. This experiment was performed in triplicate. The results of this experiment showed that there was a 4–8-fold decrease in detection sensitivity for alkaline phosphatase activity on XP-PO₄ agar relative to XP agar, which in practical terms would only be expected to affect the detection of very weakly expressed proteins.

2.4. DNA sequencing and analysis

DNA sequence determination was done with an ABI Model 372 DNA sequencer using the ABI Taq DNA polymerase sequencing kit according to the manufacturer's directions. Sequencing template was prepared by the modified PEG-alkaline lysis method. The DNA sequencing primer used to sequence the *H. pylori*-alkaline phosphatase fusion clones was 5'-ATCACCCGTTAAACGGCGAGCAC-3'. This '*phoA*-specific DNA sequencing primer, which was complementary to the '*phoA* DNA sequence located approximately 30 bp upstream of the unique *Bam*HI site in plasmids pJDT1, pJDT2 and pJDT3, was used to sequence all *H. pylori*-alkaline phosphatase fusion clones. This primer directed sequencing across the *Bam*HI fusion junction and into the inserted *H. pylori* DNA. The resulting DNA sequences were used for homology searches of the National Center for Biotechnology Investigation (NCBI) non-redundant database with the BLASTX program [10]. The default settings for the BLASTX program were used for all the searches.

3. Results and discussion

3.1. Construction of a *Helicobacter pylori*-alkaline phosphatase fusion library

Initial attempts to construct the *H. pylori*-alkaline phosphatase fusion library in *phoA* deficient *E. coli* CC118 [6] were not successful. *E. coli* CC118 proved to be a poor recipient for *H. pylori* DNA; recombinant clones were obtained at low efficiency. These findings are probably a result of *E. coli* CC118 restriction of *H. pylori* DNA [11].

Subsequently, *E. coli* ER1793 [12], a *phoA*⁺ strain

that has been shown to be permissive for the cloning of *H. pylori* DNA, was used as the host for the construction of the *H. pylori*-alkaline phosphatase fusion library. Rather than construct a *phoA* deficient strain, we found that the inclusion of 75 mM NaPO₄ in the selection medium completely inhibited the expression of the endogenous phosphatase activity in *E. coli* ER1793 while still allowing the detection of recombinant alkaline phosphatase fusion proteins (see Section 2.3). During the initial library construction alkaline phosphatase positive clones were found at a frequency of approximately 2–3%.

3.2. DNA sequencing and BLASTX analysis of *Helicobacter pylori*-alkaline phosphatase fusion clones

A total of 120 *H. pylori*-alkaline phosphatase fusion library clones were randomly selected for screening by DNA sequencing. Each clone was subjected to a single pass DNA sequencing run and we obtained an average of approximately 350 bp of DNA sequence from each clone. These DNA sequences were used separately for BLASTX homology searches of the NCBI non-redundant nucleotide database. The BLASTX results were screened for ungapped and consistently gapped high-scoring segment pairs that produced P(N) scores of 10⁻⁵ or lower since such pairings are generally considered to be significant. An additional four clones with high high-scoring segment pair scores were also included. Three clones with homology only to cytoplasmic proteins and three clones with homology to proteins encoded on the negative strand relative to '*phoA* were not included in Table 1.

Among the 120 *H. pylori*-alkaline phosphatase fusion library clones that were sequenced 25 clones showed interesting homologies (Table 1). The identified homolog genes included several prime targets for drug intervention or vaccine construction including genes involved in motility (clones 11, 30, and 121), efflux (clone 12, 39, 61 and 200), iron uptake (clones 10 and 158), cell wall (clones 44 and 162), LPS (clone 211) or lipid biosynthesis (clone 182) biosynthesis, permeases (clones 199 and 230), cytochrome (clone 105), and enzymes (clones 3, 26, 52, and 117).

In addition to those clones listed in Table 1, another 88 clones had only weak similarity to sequen-

Table 1

H. pylori-alkaline phosphatase fusion library clones having significant similarity to database sequences

Library clone (accession #) ^a	BLASTX identified gene homolog to the <i>H. pylori</i> -alkaline phosphatase fusion library clones							
	HSP gene (accession #) ^b	Organism ^c	Function ^d	AA residues ^e	% Id. ^f	% Sim. ^g	HSP ^h	P(N) ⁱ
3 (U61522)	<i>folD</i> (D10588)	<i>Ec</i>	dehydrogenase	41	50	70	68	6.5×10^{-6}
10 (U60606)	<i>exbD</i> (U08209)	<i>Hi</i>	siderophore uptake	41	51	80	125	5.5×10^{-9}
11 (U60607)	<i>lafB</i> (L06176)	<i>Vp</i>	motility	101	31	56	92	4.4×10^{-5}
12 (U60608)	<i>helA</i> (U49498)	<i>Lp</i>	efflux	59	38	59	102	2.0×10^{-9}
20 (U61524)	<i>virB4</i> (U28133)	<i>Hp</i>	?	69	48	56	75	4.2×10^{-8}
26 (U61525)	<i>kpnBI</i> (U33094)	<i>Kp</i>	DNA restriction system	64	39	70	90	8.8×10^{-8}
30 (U60610)	<i>motB</i> (M77238)	<i>Bsu</i>	motility	43	53	72	118	2.2×10^{-7}
32 (U60611)	<i>yrbJ</i> (U18997)	<i>Ec</i>	?	40	55	60	92	2.2×10^{-5}
39 (U82393)	<i>tet</i> (M16217)	<i>Sa</i>	tet efflux	62	26	47	66	0.997
44 (U60613)	<i>vanA</i> (M97297)	<i>Ef</i>	vancomycin resistance	52	34	53	86	9.1×10^{-5}
52 (U60615)	<i>T3RE</i> (X06287)	<i>Pl</i>	DNA restriction	83	40	62	74	8.0×10^{-8}
53 (U81608)	HI0325 (U00072)	<i>Hi</i>	?	46	34	58	87	7.2×10^{-5}
57 (U82394)	<i>yjcO</i> (P32713)	<i>Ec</i>	?	46	32	54	65	0.996
61 (U60617)	<i>mtrC</i> (U14993)	<i>Ng</i>	efflux	94	27	48	67	0.94
70 (U60618)	HI1586 (U32779)	<i>Hi</i>	?	53	43	74	91	3.6×10^{-10}
105 (U60620)	<i>cybH</i> (X65189)	<i>Ws</i>	cytochrome	87	54	75	261	1.9×10^{-30}
117 (U60622)	<i>ppsA</i> (D64005)	<i>Ssp</i>	kinase	58	52	76	171	1.1×10^{-20}
121 (U60623)	<i>motB</i> (M77238)	<i>Bs</i>	motility	43	53	72	121	1.0×10^{-9}
158 (U60627)	<i>feoB</i> (U18997)	<i>Ec</i>	iron transport	68	46	66	88	8.4×10^{-10}
162 (U60628)	<i>pbpA</i> (X04516)	<i>Ec</i>	penicillin binding protein	113	32	52	128	9.8×10^{-14}
182 (U60630)	<i>cdh</i> (M11331)	<i>Ec</i>	lipid biosynthesis	89	66	82	269	9.4×10^{-35}
199 (U60633)	<i>yejA</i> (U00008)	<i>Ec</i>	permease	55	31	53	97	1.9×10^{-6}
200 (U60634)	<i>yhiv</i> (U00039)	<i>Ec</i>	efflux	91	34	57	103	3.6×10^{-7}
211 (U81609)	<i>rfbA</i> (L34166)	<i>Sm</i>	LPS biosynthesis	38	37	55	68	0.90
230 (U60638)	HI0325 (U32826)	<i>Hi</i>	hypothetical permease	99	33	55	114	8.0×10^{-18}

H. pylori-alkaline phosphatase fusion library clones and their homolog genes identified by BLASTX homology searches.

^a*H. pylori*-alkaline phosphatase fusion library clone and its GenBank accession number.

^bGene name and accession number of the database gene which was similar to the *H. pylori*-alkaline phosphatase fusion library clone listed in column 1.

^cOrigin of gene listed in column 2. Abbreviations: *Bsu*, *Bacillus subtilis*; *Cf*, *Citrobacter freundii*; *Ec*, *Escherichia coli*; *Ef*, *Enterococcus faecium*; *Hi*, *Haemophilus influenzae*; *Hp*, *Helicobacter pylori*; *Kp*, *Klebsiella pneumoniae*; *Lp*, *Legionella pneumophila*; *Ng*, *Neisseria gonorrhoeae*; *Pl*, bacteriophage P1; *Sa*, *Staphylococcus aureus*; *Sm*, *Serratia marcescens*; *Ssp*, *Synechocystis species*; *Vp*, *Vibrio parahaemolyticus*; *Ws*, *Wolinella succinogenes*.

^dFunction of gene listed in column 2.

^eThe number of amino acids showing similarity in the BLASTX alignment.

^fPercent amino acid identity reported in the BLASTX search results.

^gPercent amino acid similarity reported in the BLASTX search results.

^hHSP score reported in BLASTX results.

ⁱP(N) reported in the BLASTX search results.

ces in the non-redundant database. These clones did not have a sufficient amount of similarity to sequences present in the database to allow a confident prediction of their possible function (data not shown). Additional sequencing in these clones would be required before a prediction of their function could be made. Lastly, seven clones did not produce alignments with any sequences in the NCBI non-redundant database. Whether these sequences represent

novel and unique *H. pylori* genes would be difficult to determine without additional DNA sequence information.

3.3. Confirmation of gene identification for putative *Helicobacter pylori* efflux clones

The major objective for undertaking this study was to identify potential *H. pylori* multiple drug ac-

tive efflux systems. In our initial screen we identified three potential genes from *H. pylori* efflux systems in clones 12, 61 and 200 which were similar to the resistance-nodulation-division (RND) family of bacterial efflux systems. In general, the RND efflux operons contain a conserved membrane fusion (MF) protein upstream of the pump protein. To confirm the initial gene identification we decided to obtain upstream DNA sequence information for each of these clones.

Initial screening suggested that sequence clone 12 contained similarity to an RND pump protein. Therefore, clone 12 was sequenced with the reverse sequencing primer to determine if there was homology to a MF protein in the DNA region upstream from the putative efflux pump protein. BLASTX searches using the resulting DNA sequence showed that there was homology to RND MF proteins upstream of the putative efflux pump protein (data not shown). In addition, the DNA sequence gap between the putative BLASTX identified MF and pump proteins was consistent with the size of the DNA insert in clone 12 and with other RND efflux operons. These results support the initial prediction that clone 12 contains a *H. pylori* active efflux pump-*phoA* fusion protein.

Clone 61 was initially identified as having a weak similarity to a bacterial membrane fusion protein. To further investigate this clone a DNA sequencing primer was synthesized to extend the original DNA sequence. BLASTX analysis of the resulting DNA sequence (approximately 900 bp) was consistent with the initial prediction that this clone encoded a MF protein. The extended DNA sequence identified high-scoring segment pairs at a much higher confidence level (P(N) of 9.6×10^{-5} for *Alcaligenes eutrophus* *cnrB*, data not shown).

Clone 200 was similar to several RND efflux pumps. Multiple sequence alignments of the putative amino acid sequence of clone 200 and other RND efflux pumps showed that clone 200 shared most of the conserved amino acids with the other efflux pumps. Additional DNA sequencing of the upstream region of this clone revealed the presence of a putative *H. pylori* homologue of the *E. coli* *acrAB* efflux operon including the putative pump protein and an upstream MF protein (data not shown).

3.4. Limitation and benefits of targeted genomics

There are several limitations to the presented strategy; (1) only expressed and translocated proteins will be detected; (2) identification of a gene is inferred from its similarity to database sequences and therefore, novel genes will not be identified; (3) gene identification is dependent upon the gene fusion occurring in a region containing relatively conserved amino acid signatures or residues; (4) poorly expressed proteins may be difficult to detect; (5) *phoA* fusion to significant stretches of hydrophobic amino acids that are present in cytoplasmic proteins can result in alkaline phosphatase activity.

The method presented here is rapid, technically simple, and relatively inexpensive for identifying secreted and exported proteins including, but not limited to, novel drug targets and genes of therapeutic interest. Since protein identification is dependent upon translocation of the alkaline phosphatase fusion proteins, only expressed and accessible proteins are identified. The method is applicable to a wide variety of microbes and genetic systems including bacteria for which there is no genetic system available. The addition of phosphate to the selection medium is an effective method for the inhibition of the endogenous phosphatase activity in *E. coli* while still allowing for the detection of recombinant alkaline phosphatase fusion proteins.

Acknowledgments

This work was supported initially by grants from the Canadian Bacterial Diseases Network and subsequently from the Medical Research Council of Canada.

References

- [1] Fleischmann, R.D., Adams, M.D., White, O. et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269, 496–512.
- [2] Fraser, C.M., Gocayne, J.D., White, O. et al. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270, 397–403.
- [3] Derman, A.I. and Beckwith, J. (1995) *Escherichia coli* alkaline phosphatase localized to the cytoplasm slowly acquires enzy-

- matic activity in cells whose growth has been suspended: a caution for gene fusion studies. *J. Bacteriol.* 177, 3764–3770.
- [4] Odenbreit, S., Till, M. and Haas, R. (1996) Optimized BlaM-transposon shuttle mutagenesis of *Helicobacter pylori* allows the identification of novel genetic loci involved in bacterial virulence. *Mol. Microbiol.* 20, 361–373.
- [5] Mdluli, K.E., Treit, J.D., Kerr, V.J. and Nano, F.E. (1995) New vectors for the in vitro generation of alkaline phosphatase fusions to proteins encoded by G+C-rich DNA. *Gene* 155, 133–134.
- [6] Manoil, C. and Beckwith, J. (1985) *TnphoA*: a transposon probe for protein export signals. *Proc. Natl. Acad. Sci. USA* 82, 8129–8133.
- [7] Exner, M.M., Doig, P., Trust, T.J. and Hancock, R.E. (1995) Isolation and characterization of a family of porin proteins from *Helicobacter pylori*. *Infect. Immun.* 63, 1567–1572.
- [8] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1995) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [9] Sekar, V. (1995) A rapid screening procedure for the identification of recombinant bacterial clones. *Biotechniques* 5, 11–13.
- [10] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- [11] Phadnis, S.H., Westblom, U.T. and Normark, S. (1993) Molecular cloning of *Helicobacter pylori* DNA: important differences between *mcrBC* deletion host strains. *Mol. Microbiol.* 10, 1151.
- [12] Kelleher, J.E. and Raleigh, E.A. (1991) A novel activity in *Escherichia coli* K-12 that directs restriction of DNA modified at CG dinucleotides. *J. Bacteriol.* 173, 5220–5223.