SEQUENCING AND CHARACTERIZATION OF *HELICOBACTER PYLORI* PLASMID, pAL202

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

BY

ILSE Y. RICKETS, B.S, M.S.

DENTON, TEXAS

MAY 2004

TEXAS WOMAN'S UNIVERSITY
DENTON, TEXAS

April 8, 2004
Date

To the Dean of the Graduate School:

I am submitting herewith a dissertation written by Ilse Y. Rickets entitled "Sequencing and characterization of *Helicobacter pylori* plasmid, pAL202." I have examined this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy with a major in Molecular Biology.

Dr. Sarah A. McIntire, Major Professor

We have read this dissertation and recommend its acceptance:

Department Chair

Dean of College

Accepted:

Dean of the Graduate School

DEDICATION

For my mom, Albertina Mathilda Susanna Olijfveld.

Suriname, September 27, 1920 –The Netherlands, September 15, 2003.

It has been a blessing and a privilege to be your daughter.

My dad, Eduad Guillaume Rickets

Suriname, October 29, 1912 – March 10, 1982.

Thank you.

My husband, Henry H. Ross.

Thanks for your love and support.

My brothers: Stanley, Rudi, Hubert, John, Percy, and Laurens Olijfveld.

My sisters: Orlanda Nelom-Olijfveld and Elfriede Rickets.

Family is like a blessing from God.

ACKNOWLEDGEMENTS

I would like to thank Dr. Sarah McIntire, my mentor, for her continuous guidance, encouragement, and patience. Thanks to Dr. John Knesek for always being available when needed. A special appreciation is also extended to the other committee members: Dr. Lynda Uphouse, Dr. Nathaniel Mills, and Dr. Daniel Kunz. Thanks for everything! I would like to acknowledge the Multi-Ethnic Biomedical Research Support Program (MBRS) for their financial support, but most importantly for the exposure to scientific research. I will always be grateful for the financial assistance received during graduate school from the Doctoral Fellowship Program of Texas A&M University-Commerce. Finally I would like to thank all of my lab-mates for their support and five wonderful years!

ABSTRACT

ILSE Y. RICKETS

SEQUENCING AND CHARACTERIZATION OF *HELICOBACTER PYLORI* PLASMID, pAL202

MAY 2004

Helicobacter pylori, a Gram-negative, spiral-shaped bacterium, is associated with gastric diseases in humans. We characterize *H. pylori* plasmids to determine their biological role in this microorganism and to test the hypothesis that plasmids are involved in chromosomal rearrangement. In this study, we sequenced and characterized pAL202, a plasmid from an Alaskan strain, AL202, and analyzed additional Alaskan plasmids for conserved genes.

The EZ::TN^{TM}<R6Kγ*ori*/KAN-2> transposon was inserted into pAL202 and transformed into *E. coli pir*$^{+}$ cells. The DNA sequence of pAL202 was determined bidirectionally using transposon-specific primers. The plasmid sequence was constructed from the sequences obtained from these transformants. The plasmid was 12.1 kbp and has 15 open reading frames (ORFs). Submission to GenBank revealed identity to other *H. pylori* plasmids and chromosome. The greatest identity was with pHel4, a plasmid from a European *H. pylori* strain. Sequence identity was found to *H. pylori* plasmid *rep* and *orf2* genes, the IS*607* transposon, and regions of the *cag* PAI of strains J99 and 26695 chromosomes. Observed iterons indicated that pAL202 replicates via the theta type mechanism. Two ORFs showed identity to the microcin genes of *E. coli*, and four to

mobilization genes. A putative *nic* sequence found could serve as a DNA transfer site, but no transconjugants were obtained from mobilization studies. Attempts to transform *H. pylori* cells AL202 and HU71 with single insertion mutants resulted in zero transformants. A comparison study to test for the distribution of pAL202 genes in six other Alaskan plasmids showed that at least one gene was conserved in five of the seven plasmids.

The pAL202 genes were conserved not only among the Alaskan plasmids but also among *H. pylori* plasmids in general. The presence of short *H. pylori* chromosome sequences suggested recombination between plasmid and chromosomal DNA that resulted in dissimilarity of gene order in the plasmids. Recombination in the genome might be crucial for the bacterium to evade the host immune system and therefore contribute to the pathogenicity of the organism

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In 1983 Marshal and Warren first reported the association of *Helicobacter pylori* with gastric inflammation (53, 86). Initially classified as a *Campylobacter* species, rRNA gene analysis revealed enough differences for it to be classified as a species from the new genus *Helicobacter* (49, 55). *H. pylori* is a microaerophilic, curved or spiral-shaped, flagellated, Gram-negative bacterium (53, 86). The bacterium colonizes the stomach of humans and nonhuman primates (26, 29), and is the etiological agent of chronic active gastritis, gastric and duodenal ulcers, gastric adenocarcinoma and mucosa-associated lymphoid tissue (MALT) lymphoma (18, 26) resulting from long-term infection (2).

Transmission of the bacterium occurs orally either by the fecal-oral or the oral-oral routes (17, 26, 55, 80). Infection is established early in childhood and, if no symptoms occur, may persist throughout life. Over half of the world's population is infected with *H. pylori*. Between 20-50% of the adult population in developed countries, and 70-90% in developing countries are infected with the bacterium (17, 26). The lower incidence in developed countries might be due to better sanitation. The high incidence of *H. pylori* in humans and the observation that is has presumably been present for centuries led some researchers to suggest that this bacterium must be part of normal flora (9, 49) and that eliminating it in people that are asymptomatic could lead to other esophageal diseases (8).

There are many virulence factors associated with this organism. *H. pylori* is sensitive to low pH (25, 49), but the bacterium secretes the enzyme urease that hydrolyzes urea to ammonia and carbon dioxide (56). The production of ammonia neutralizes the acid in the bacterium's microenvironment and allows survival in the hostile environment of the stomach (25, 26, 49, 56). Besides neutralization of acid in the stomach, ammonia also serves as a nitrogen source for protein synthesis (31).

The presence of two to six unipolar flagella renders the bacterium motile (26, 67). *H. pylori* mutants in which the flagella genes are disrupted, are unable to colonize and survive in the stomach of gnotobiotic piglets (27). Upon entrance in the stomach, the bacterium maneuvers through the viscous gastric mucous layer to the gastric epithelial cell surface (67). Once *H. pylori* has reached the more neutral environment below the mucus, it adheres tightly to the gastric epithelial cells (11, 16) mediated by multiple surface proteins called adhesins (11).

Other *H. pylori* virulence factors produced are catalase (36), superoxide dismutase (74), phospholipase (24), and the major virulence factors vacuolating cytotoxin protein, VacA, and cytotoxin-associated protein, CagA. VacA is an 87 kDa protein created from the cleavage of a 139 kDa precursor protein (19). The protein causes cytoplasmic vacuolation (51) derived from late endosomes in eukaryotic cells (64) and is capable of inducing apoptosis in the human gastric epithelial cell line AGS (47). CagA is a 128-145 kDa protein (17, 78), encoded by the *cagA* gene present in pathogenic strains of *H. pylori* (15). The size variability of the *cagA* gene among strains is due to the presence of 102 bp repeats (17, 78). The number of the repeats in each gene differs from strain to strain (17),

but always leaves the gene in frame. Covacci *et al.* (17) suggested that repeating sequences in antigenic molecules are common and might be involved in escaping the immune system by creating antigenic diversity. The CagA protein is transported into the gastric epithelial cells by a type IV secretion apparatus (63). The protein is tyrosine phoshorylated in the epithelial cells (78) and could trigger a signal transduction in the host cells (68) followed by induced changes in structure, function, and morphology in the cells (6).

H. pylori strains are divided in two major types, type I and type II (87), depending on the presence of a pathogenicity island called *cag* PAI (15). Pathogenicity islands are large genomic sequences between 10-200 kbp that are present on the genome of pathogenic bacteria, but absent on the genome of non-pathogenic strains of the same organism (33, 35). These structures have direct repeats at the ends, have a different GC percent from the rest of the genome, and possess genes that encode virulence factors (35). The *cag* PAI in *H. pylori* has a GC content of 35% compared to 38-45% of the rest of *H. pylori* genomes and genes deposited in data banks (15). The type I strains express both the *cagA* ($CagA^+$) and the *vacA* ($VacA^+$) genes (87), and are associated with the more severe gastric diseases (17, 87). Type II strains lack the *cagA* gene ($CagA^-$) and express a nonfunctional VacA, and are therefore $VacA^-$ (87). On the chromosome of *H. pylori* strain NCTC 11638, the *cagA* and the *vacA* genes are 300 kbp apart (15), and even though the genes are coexpressed their expression does not depend on each other (2, 15, 18, 87). The presence of the *cag* PAI is also associated with elevated levels of cytokines such as interleukin-8 (IL-8) in *H. pylori* strains (2, 15, 62).

The 40 kbp *cag* PAI is inserted into the 3'-end of the glutamate racemase gene and is flanked by a 31 bp direct repeat, which suggests that the PAI was inserted in the genome by means of chromosomal recombination (15). *H. pylori* presumably acquired the fragment millions of years ago through horizontal gene transfer from an unknown source (15, 78). The *cag* PAI encodes more than 20 genes and can be continuous, or in some strains, divided by the insertion sequence, IS*605*, into a right segment, *cag*I, and a left segment *cag*II (15, 81). The type IV secretion system mentioned above is encoded by the *cag* PAI (63).

The chromosomal DNA of two type I *H. pylori* strains 26695 and J99 have been entirely sequenced and compared (3, 81). Except for some genomic rearrangements, the genomes of these strains are very similar in genomic organization, gene order, and protein prediction (3). Many *H. pylori* strains contain one or more copies or partial copies of the insertion sequences IS*605* (3, 15, 42, 81), IS*606* (3, 15, 42, 81), IS*607* (43), IS*Hp608* (44), and IS*Hp609* (unpublished, AY487825). It is believed that the genomic rearrangement observed in *H. pylori* might be the result of these insertion sequences (42). The presence of insertion sequences, partial and whole, and chromosomal DNA on plasmids (13, 22, 30, 37, 39, 40, 59, 69, 71, 75, 79) suggests that plasmids might be involved in the integration of new DNA into *H. pylori* chromosome and the transfer of DNA between strains (3).

More than 50 % of *H. pylori* isolates examined contain one or more plasmids (22, 39, 59) ranging in size from 1.5 to greater than 148 kbp (21, 22). Several *H. pylori* plasmids have been sequenced and characterized (13, 22, 30, 37, 39, 40, 45, 59, 69, 70,

75, 76, 79), but no common phenotype has been assigned. When considering replication, the plasmid can be categorized into those that replicate in a rolling circle manner and those that replicate by a theta type mechanism. The plasmid pHPK255 has been identified as replicating by the rolling circle mechanism that is common for plasmids in Gram-positive bacteria (45). The replication (Rep) protein of pHP489 showed 57% similarity to the Rep protein of pHPK255 and had two of the three conserved motifs present in the Rep proteins of plasmids that replicate by the rolling circle mechanism (76). This protein was named RepH where the H stands for *Helicobacter* (76). In pHP489, mutations of the RepH protein revealed that the protein was not required for replication in *H. pylori* 489 and that some genomic proteins might be involved in the replication of the plasmid (76). The other characterized plasmids contain tandem repeated sequences, called iterons, which serve as the origin of replication. These plasmids all possess the replication genes *repA* (30, 37, 39, 59, 69, 70, 79), *repB* (40, 75), or both the *repA* and *repB* genes (22), indicative of the theta type replication mechanism.

All the analyzed plasmids exhibit different sizes and restriction digestion profiles. Although identity exists among some of the plasmids, they all reveal unique gene arrangements. Among the theta type plasmids the RepA proteins show a strong conserved middle region, even though the total sizes vary from 19.8 kDa to 29.4 kDa. The iterons composed of four 22 bp repeated sequences have a conserved sequence as well. One additional conserved protein termed ORF2 is found in some, but not all, *H. pylori* theta plasmids. In some plasmids, pHPM8 (70), pHPM180 (59), pHPS1 (22), pHPO100 (unpublished), and pHP51 (75), the *orf2* gene is positioned downstream of the

repA gene. In pKU701 *orf2-like* sequences are observed directly downstream of the *repB* gene. In pHPM8 the *repA* gene and *orf2* genes are cotranscribed (70). Since ORF2 is absent from some plasmids (37, 76), De Ungria *et al.* (22) suggest that ORF2 might be involved in recombination. Quiñones *et al.* (70) suggest that ORF2 might be associated with the replication and/or control of the copy number of plasmids. In pHP51 a 74% similarity between ORF2 and the putative cAMP-induced filamentation (Fic) protein of *Helicobacter* HP1159 (NP_223803) was observed (76). The Fic protein and cAMP are involved in the regulation of cell division via folate metabolism in *Escherichia coli* (46, 83).

Other regions of interest found on some of the *H. pylori* plasmids are: two copies of IS*605* in pHPM186 (13); IS*607* in pHPM179 (79); the tetracycline efflux protein TetA(P)-like and microcin-like proteins in pHPM8 (70) and pHel4 (38); and mob-like proteins in pHel4. The accession numbers or references of the plasmids and strains cited in this study are shown in Table 1.

The plasmids previously characterized in our laboratory were isolated from *H. pylori* strains recovered from patients at Dallas VA Medical Center. Since these strains share a common North Texas origin, a previous project sought to characterize plasmids originating from *H. pylori* isolates from a different geographical region, Alaska. A set of seven strains was shown (41) to carry *repA* and *orf2* sequences similar to pHPM8. The main purpose of this study was to determine the nucleotide sequence of a plasmid from one of these Alaskan strains, AL202 and analyze additional Alaskan plasmids for conserved genes or regions. Conserved genes or regions are genes or regions that are

Table 1. Accession numbers and/or references of *H. pylori* plasmids and strains cited in this study

Plasmid	Accession #		Reference #
pAL202	AY584531		
pHPM8	AF275307		70, 71
pHPM179a			69
pHPM179b			79
pHPM180	U12689		59
pHPM185			30
pHPM186	AF077006		13
pHel1	Z49272		37
pHel4	AF469112		39
pHel5	AF469113		39
pHPS1	AF019894		22
pMCU1	AF019895	& AF055275	21, 22
pMCU2	AF019896	& AF055274	21, 22
pMCU3	AF019897		21, 22
pMCU4	AF019898	& AF055276	21, 22
pMCU5	AF019899	& AF055277	21, 22
pMCU6	AF019900		21, 22
pMCU7	AF019901		21, 22
pHPO100	AF056496		
pHP51	AY267368		75
pHP489	AF027303		76
pKU701	AB078638		40
pHPK255	S84689		45
26695*	AE000511	& AE000537	3, 23, 81
J99*	AE001439	& AE001512	3, 23
CCUG 17874*	AAF80202		

* Chromosome

present on DNA of more than one plasmid or chromosome. Plasmid pAL202 was sequenced using the EZ::TN^{TM}<R6kγ*ori*/KAN-2> transposon (Epicentre). The transposon was randomly inserted within the plasmid and sequencing occurred bi-directionally from primer binding sites within the transposon. The nucleic acid sequence was determined and submitted to the GenBank for comparison to other known sequences. The specific aims of this research were:

1. To determine the sequence of plasmid pAL202 isolated from the Alaskan strain AL202 by using the EZ::TN^{TM}<R6Kγ*ori*/KAN-2> transposon.
2. To analyze the sequence of the plasmid for known sequence identities both to other plasmids and to chromosomal DNA, and determine the presence of possible open reading frames (ORFs).
3. To test for the distribution of DNA sequences present in pAL202 in six other *H. pylori* Alaskan strains.
4. To attempt to transform AL202 cells and HU71 cells (*H. pylori* stain with no plasmid) with single transposon insertion mutants of pAL202 and determine the effect of the insertion in the *repA* gene, the *orf2* gene, or elsewhere in the plasmid.

CHAPTER II

MATERIALS AND METHODS

Materials

Bacterial Strains and Plasmids

H. pylori strains AL202, AL203, AL207, AL208, AL209, AL226, and AL236, were isolated from Alaskan Indians. Douglas E. Berg, Washington University School of Medicine, St. Louis, MO, and Alan J. Parkinson, Centers for Disease Control and Prevention, Anchorage, AK, provided the strains. HU71, an *H. pylori* strain without plasmids, and also provided by D. E. Berg, was used in the transformation studies. The strains were stored at -80°C in freezer medium consisting of trypticase soy broth (Difco), containing 25% glycerol (Sigma) and 10% horse serum (Sigma). *H. pylori* cells were cultured at 37°C under microaerophilic conditions, 5-10% CO_2, on brain hearth infusion (BHI) medium (Difco), supplemented with 10% horse serum. The plasmid DNAs, pAL202, pAL203, pAL207, pAL208, pAL209, pAL226, and pAL236, were isolated from the Alaskan strains.

The EZ::TN™<R6Kγ*ori*/KAN-2> transposon (Epicentre) was randomly inserted into pAL202, pAL208, and pAL236. Electrocompetent *E. coli pir*$^+$ (low copy number) and *pir*-116 (high copy number) cells obtained from Epicentre were transformed with the recombinant plasmids of pAL202. The recombinants of pAL208 and pAL236 were transformed into *pir*$^+$ cells. The genotype of the *pir*$^+$ cells is F$^-$, *mcr*A, Δ(*mrr-hsd*RMA-*mcr*BC), ϕ80d*lac*ZΔM15, Δ*lac*X74, *rec*A1, *end*A1, *ara*D139, Δ*(ara, leu)*7697, *gal*U,

*gal*K, λ⁻, *rpsL*(Str^r), *nup*G, *pir*$^+$(DHFR). The genotype of the *pir*-116 cells is identical to the *pir*$^+$ cells except for containing the *pir*-116 mutation. *E. coli* strains were stored at -80°C with Protect™ Bacterial Preservers (Key Scientific Products) and cultured in NZYM (Bio 101 System) or Luria-Bertani (LB) medium (Bio 101 Sytems).

The plasmid pZErO-2.1™ (Invitrogen) in competent *E. coli* Top10F' (Invitrogen) was obtained from Dr. John Knesek at Texas Woman's University. The genotype of Top10F' is F{*Lac*I^q, Tn10 (Tet^r)} *mcr*A, Δ(*mrr-hsd*RMS-*mcr*BC), ϕ80*Lac*ZΔM15, Δ*lac*X74, *deo*R, *rec*A1, *ara*D139, Δ(*ara-leu*) 7679, *gal*U, *gal*K, *rpsL*(Str^r), *end*A1, *nup*G. The plasmid, pZErO-2.1, was isolated and transformed with the EZ::TN™<R6Kγ*ori*/KAN-2> transposon. The recombinant plasmid, pZErO::TN/KAN, was transformed into *E. coli pir*$^+$ cells and was used as a control in comparative PCR studies.

JA221(pRK24), used as a donor in mating experiments, was provided by David H. Figurski from the College of Physician & Surgeons of Columbia University, Department of Microbiology, New York, New York. The genotype of JA221 is *rec*A1, *lac*Y, *leu*B6, Δ*trp*E5, *hsd*R, with pRK24 providing Ap^r, Tc^r, and Trp$^+$. The recipient used in the mating studies, JC3272, provided by Dr. Sarah McIntire at Texas Woman's University, is *his*, *trp*, *lys*, *rpsL*(Str^r). JC3272 was made Nal^r by plating an overnight culture on LB plates containing 20 μg/ml nalidixic acid (Sigma). The Nal^r derivative was named IR3272.

Media

Brain Heart Infusion (Difco), NZYM (Bio 101 System), and LB media (Bio 101 Systems) were prepared according to the manufacturer's recommendations. SOC medium used in electroporation studies was prepared as described by Sambrook *et al.* (72). BHI was supplemented with 10% horse serum (Sigma). Kanamycin (Amresco) at a concentration of 50 μg/ml was added to the NZYM medium, LB medium, and to BHI-YE [brain heart infusion-yeast extract (EM Science)] agar used in transformation and mating studies. Tetracycline (Amresco) was added to BHI agar to a final concentration of 2 μg/ml to determine antibiotic resistance of *H. pylori* strain AL202. Streptomycin (Amresco) at a concentration of 25 μg/ml, nalidixic acid at 20 μg/ml, kanamycin at 50 μg/ml, and tetracycline at 50 μg/ml were added to M9 minimal medium [200 ml of 5X M9 minimal medium salts (Bio 101 Systems); 10 ml of 20% glucose; 1.0 ml of 1.0 M $MgSO_4$; 1.0 ml of 1.0 M thiamine; histidine, tryptophan, and lysine at 40 mg/L each; H_2O for a total volume of 1.0 L] used in mating studies. Plate medium was solidified with Bacto Agar (Difco) at a concentration of 1.5% (15 g/L).

Enzymes

Restriction endonucleases were acquired from Promega. The restriction endonucleases and other enzymes from various commercial kits were used with buffers provided by, and as recommended by, the suppliers.

Buffers and Reagents

The following reagents were used: TE (10 mM Tris, 1.0 mM EDTA, pH 7.5-8.0); 5X TBE, Tris-Borate-EDTA (Sigma); agarose (SeaKem, FMC); 10 mg/ml ethidium

bromide (Sigma); 3 M sodium acetate; N,N,N',N'-tetramethylethylenediamine, TEMED (Amresco); ammonium persulfate (Amresco); urea (Amresco); Long-ranger™ gel solution (FMC); silane stock [0.5 ml γ-methacryloxypropyltri-methoxysilane (Sigma), in 100 ml of 100% ethanol]; 70%, 95% and 100% ethanol; 100% isopropanol (Fisher), 10% glycerol (Fisher); 5% acetic acid (Fisher); 5X ficoll dye (0.1 M EDTA, 25% ficoll, 0.1% bromophenol blue, 0.1% xylene cyanol); 1 kbp DNA ladder (GIBCO BRL); [α^{32}P]-dCTP (NEN™); Ecolume (Lab Industries); denaturing buffer (3 M NaCl, 0.4 M NaOH); transfer buffer (3 M NaCl, 8 mM NaOH, 2 mM sarkosyl); 5X neutralizing buffer (1.0 M phosphate buffer, pH 6.8); ULTRAhyb hybridization buffer (Ambion); wash buffer (2X SSC, 0.1% SDS); adenosine 5'-triphospate disodium salt (Sigma).

Commercial Kits

The kits used were: The Wizard® *Plus* Miniprep (Promega); QIAGEN Plasmid Midi Kit (QIAGEN Inc.); QIAGEN Large-Construct Kit (QIAGEN, Inc.); Plasmid-Safe™ ATP-Dependent DNase (Epicentre); SequiTherm EXCEL™ II DNA Sequencing Kit-LC (Epicentre); GeneAmp XL PCR Kit (Applied Biosystems); REDTaq™ DNA polymerase without $MgCl_2$ kit (Sigma); BIOLASE Red™ DNA Polymerase kit (BIOLINE); Zymoclean Gel DNA Recovery Kit (ZymoResearch); NEBlot® Kit (New England BioLabs Inc.); The EZ::TN™<R6K*γori*/KAN-2> Insertion Kit (Epicentre) containing the EZ::TN™<R6K*γori*/KAN-2> transposon.

Primers

Labeled primers used for the PCR reaction for sequencing were obtained from LI-COR Inc. The computer program, EuGene (Daniben Systems, Inc.), was used to design primers used in the comparative studies between pAL202 and the other Alaskan plasmids. The primers used in the comparative PCR studies were obtained from Bio-Synthesis Inc. Table 2 shows the sequences and the nucleotide positions of primers used in this study.

Computer Software and Lab Equipment

Computer software, EuGene (Daniben Systems, Inc.), was used to design primers, and Vector NTI® Suite 7.0 and 8.0 (InforMax®) was used to analyze pAL202 DNA sequence. The following internet-available software programs were used to further analyze pAL202 DNA sequences and open reading frames (ORFs): BLAST algorithms by the National Center for Biotechnology Information database (4, 5) were used to find identities to other DNA and amino acid sequences; the GC content of pAL202 and other *H. pylori* plasmids was obtained from Vector NTI® Suite 8.0 (InforMax®). Open reading frames were analyzed using several software programs for protein analysis available through the ExPASy Proteomics tools (http://us.expasy.org/): PSORT (http://www.psort.org/); SOSUI (http://sosui.proteome.bio.tuat.ac.jp/sosuiframe0.html); ProtParam (http://us.expasy.org/tools/protparam.html); TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/); and TopPred (http://www.sbc.su.se/~erikw/toppred2/). Conserved domains of proteins were detected

Table 2. Primers used in this study.

Primers 1-8 were sequencing primers; primers 1 and 2 were transposon-specific primers obtained from LI-COR Inc., primers 3-14 were pAL202 primers obtained from Bio-Synthesis Inc. Primers 9-14 were PCR primers used in the comparison study of pAL202 genes.

	Primer Name	Primer Sequence	Position
1	KAN-2 FP-1 Forward	5'-ACCTACAACAAAGCTCTCATCAACC-3'	1907-1931
2	R6KAN-2 RP-1 Reverse	5'-CTACCCTGTGGAACACCTACATCT-3'	52-75
3	202AF	5'-GTCTCTAGCTTTAGCTATGGGG-3'	2142-2163
4	202AR1	5'-GCTTATCCACATCTTGGACTG-3'	1932-1952
5	202BF	5'-CACGATAAAGCGTCACTAAAG-3'	7159-7179
6	202BR	5'-ATACTTCCTCATCATTCTGCCC-3'	8084-8105
7	202CF	5'-ATTCAAAGCCACACTCTAGGC-3'	11658-11678
8	202CR	5'-GTGAAAGTATCTTCACGCGCTG-3'	135-156
9	IR1	5'-GGCTCTGTATGTGCTGAACA-3'	2253-2272
10	IR2	5'-GAAATAATCCAAGAAACCGCTTACCACAGC-3'	2832-2861
11	IR3	5'-TCTGACCATTCGCATGGCGCACC-3'	3142-3164
12	IR4	5'-GGCAGTTCTTAATGATGCCAACTAGGGTG-3'	3718-3746
13	IRMob1	5'-GTAGCGTGTTGTGTGTTTGAAGTAGCG-3'	7693-7719
14	IRMob2	5'-CCAAATCGCTAAACATCTCAACAC-3'	9278-9301

with the NCBI Conserved Domain Search (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml).

The equipment used was: Sephadex-G50 Nick™ Columns (Amersham Biosciences); TURBOBLOTTER™ Rapid Downward Transfer Systems (Schleicher & Schuell); TransPorator™ Plus Electroporator (BTX); GeneAmp® PCR System 2400; LI-COR 4000L Automated DNA Sequencer; BIO-RAD Molecular Imager® FX (BIO-RAD); Packard-Tri-Carb 2100 TR scintillation counter; Shimadzu UV160U spectrophotometer. A 302 nm UV transluminator (VILBER LOURMAT) equipped with an orange filter and the Kodak Digital Science ID camera computer analysis system were used to visualize DNA and record images of agarose gels. The Beckman J2-HS centrifuge was used with the JA-18.8, JA-14, or the JS 7.5 rotors to pellet cells at 4°C, and the Savant Speedfuge HSC10K to pellet cells at room temperature.

Methods

Plasmid Purification

Both *H. pylori* plasmid DNA and *E. coli* plasmid DNA were isolated using the Wizard®*Plus* Miniprep Kit (Promega) or the QIAGEN Plasmid Midi Kit (QIAGEN Inc.). A modification of the Wizard Miniprep procedures (59) included four TE buffer washes of the cells prior to cell lysis. *H. pylori* cells were cultured in four 175 cm^2 Falcon® tissue culture flasks, each containing 100 ml of BHI broth supplemented with 10% horse serum. The cells were grown for 2-5 days at 37°C and 5-10% CO_2. The cells were centrifuged in eight Falcon® 50 ml polypropylene conical tubes, in a Beckman J2-HS centrifuge with a JS- 7.5 rotor at 5,000 rpm, and 4°C. Each cell pellet was suspended

in 5 ml TE buffer; two tubes were combined and centrifuged as above. The above step was repeated after which the cells in each of the two tubes were suspended in 1.0 ml TE buffer. Each suspension was transferred to a1.5 ml microcentrifuge tube (United Laboratory Plastic) and centrifuged in a Beckman J2-HS centrifuge with a JA-18.1 rotor at 9,000 rpm for 15 min at 4°C. Each pellet was suspended in 1.0 ml TE and centrifuged as above. The TE washes preceded the utilization of the Wizard kit that started with suspending each final cell pellet in 200 µl of cell resuspension buffer. To each resuspension, 200 µl of cell lysis buffer was added, followed by careful inversion of each tube four times. Two hundred microliters of neutralization solution was added to each tube followed by inversion as above. The lysates were centrifuged at room temperature for 5 min in a Savant Speedfuge HSC10K at 12,000 rpm. One milliliter of DNA purification resin was added to a mini-column/syringe assembled on a vacuum, followed by the cell lysate. The sample was pulled through the column and then washed with 2 ml of column wash solution. The mini-columns were spun for two min to remove traces of the column wash solution. Plasmid DNA was eluted from the mini-column with 50 µl of H_2O preheated to 65°C. For screening purposes, recombinant plasmid DNA of *E. coli* cells was isolated as above without the TE wash steps.

E. coli recombinant plasmid DNA used in sequencing and PCR reactions, and *H. pylori* plasmid DNA used in PCR reactions was isolated with the QIAGEN Plasmid MidiKit. Cells from 200 ml were centrifuged in a Beckman J2-HS centrifuge with a JS 7.5 rotor, at 6,000 rpm for 15 min at 4°C. The supernatant was discarded and each pellet was resuspended in 4 ml of P1 buffer. The content of two tubes was combined into a

polypropylene bottle to which 4 ml of P2 buffer was added. The mixture was incubated at room temperature for 5 min with gentle shaking. Four milliliters of pre-chilled P3 buffer was added to each bottle and placed on ice for 15 min. Centrifugation followed in a JA-14 rotor at 11,000 rpm for 30 min at 4°C. The supernatant was transferred to clean bottles and centrifuged as above for 15 min. Equilibration of the Qiagen tips with 4 ml of QBT buffer was followed by the addition of the supernatant. The tips were washed twice with 10 ml of QC buffer, after which the plasmid DNA in each tip was eluted with 5 ml of QF buffer. The DNA was precipitated with 5 ml isopropanol, divided into 6 microcentrifuge tubes, and incubated at -20°C for at least 1 hr. After the incubation, the tubes were centrifuged with a JA-18.1 rotor at 11,000 rpm for 30 min at 4°C. Each DNA pellet was washed with 333 µl of 70% ethanol at room temperature. The supernatant was discarded, the pellets were air died for 10 min, and the DNA pellets were redissolved in 50-75 µl of H_2O. In the case of low plasmid yield, plasmid DNA was concentrated with 0.3 M sodium acetate and 95 or 100% ethanol. The plasmid DNA concentration was determined at A_{260} where A_{260} of 1 O.D. = 50 µg/ml.

The QIAGEN Large-Construct Kit was used to isolated plasmid DNA that required treatment with ATP-dependent exonucleases to remove chromosomal DNA contamination, as recommended by the manufacturer. Genomic DNA was also removed from plasmid preps with the use of Plasmid-Safe ATP-Dependent DNase as recommended by the manufacturer.

Transposition and Transformation

The EZ::TN™<R6Kγ*ori*/KAN-2> transposon (Epicentre) was randomly inserted into pAL202 DNA as recommended by the manufacturer. Electrocompetent *E. coli pir*$^+$ (low copy number) and *pir*-116 (high copy number) cells were transformed through electroporation with the recombinants of pAL202. Electroporation was carried out using 40 µl of cells and 1-10 µg DNA according to the manufacturer's procedures in a TransPorator™ Plus Electroporation system (BTX) at 1.5-2.5 kV, 40 µF and 130 Ω in a 1 or 2 mm cuvette. Immediately after electroporation, 960 µl of SOC medium was added to the reaction mixture for a total of 1.0 ml. The mixture was incubated at 37°C for 1 hr followed by plating of serial dilutions on NZYM or LB agar containing kanamycin (50 µg/ml) to select for cells transformed with the recombinant plasmid. The EZ::TN™<R6Kγ*ori*/KAN-2> transposon possesses a kanamycin resistance gene that allowed selection of transformants. The presence of the R6Kγ*ori*, origin of replication, on the transposon allowed the recombinants of pAL202 to replicate in *E. coli pir*$^+$ cells (57). The product of the *pir* gene, the π protein, binds to the R6Kγ*ori* and initiates plasmid replication (57).

Analysis of pAL208 and pAL236 indicated that there could be more than one plasmid present in each bacterial cell of both *H. pylori* strains. As described above, the EZ::TN™<R6Kγ*ori*/KAN-2> transposon was randomly inserted into pAL208 and pAL236 in order to separate the different plasmids present in each cell. Selection of transformants of pAL208 and pAL236 occurred as mentioned above.

Restriction Endonuclease Analysis

One to two units of restriction endonuclease were added to 50-150 ng of plasmid DNA, 2 µl 10X reaction buffer, and H_2O to a final volume of 20 µl. The reaction mixture was incubated overnight at 37° in a water-bath. Digestion reactions were analyzed by agarose gel electrophoresis as described below. The plasmid DNAs isolated from the Alaskan strains were digested with the restriction enzyme *Hin*dIII and the buffer recommended by the manufacturer. Plasmid pAL202 was analyzed further with several restriction enzymes, to determine which restriction enzymes cleaved the plasmid once. Recombinants of pAL202 were digested with two enzymes; *Eco*RI that cleaved once in the plasmid, and *Pst*I that cleaved once in the transposon. The double digestion allowed selection of single inserts of the EZ::TN^{TM}<R6Kγ*ori*/KAN-2> transposon in the recombinant plasmids. All digestion products were analyzed by agarose gel electrophoresis. Only the recombinants of pAL202 that produced one or two bands on the agarose gel were kept for further analysis. The recombinants of pAL208 and pAL236 were digested with *Hin*dIII, which had two restriction sites in the transposon.

Agarose Gel Electrophoresis

A 0.8%-1.0% agarose gel containing 1X TBE was used to confirm the isolation of plasmid DNA, and to analyze restriction endonuclease fragments by electrophoresis. Following electrophoresis, the 40 ml mini or 100 ml standard gels were stained for 20 min in 1X TBE containing ethidium bromide at a final concentration of 1.0 µg/ml. DNA in the gels was visualized with a UV transilluminator (302 nm) and pictures were taken

with the Kodak Digital Science™ID camera computer analysis system equipped with an orange filter (Tiffen 15 orange).

PCR Amplification

Amplification of plasmid DNA fragments for sequencing was performed with the GeneAmp® PCR 2400 thermocycler and the GeneAmp XL PCR Kit (Perkin Elmer) as recommended by the manufacturer. The sequencing reaction was mixed on ice and contained 2.5 μg of plasmid DNA, 8.1 μl of 3.5X SequiTherm sequencing buffer, 2.25 μl of the KAN-2 FP-1 Forward or R6KAN-2 RP-1 Reverse primer (25 pmol/μl), ultrapure H_2O to a volume of 18 μl, and 1.12 μl of SequiTerm EXCEL II DNA polymerase (5 U/μl). Four micoliters of the above reaction was added to each of four 200 μl PCR tubes containing 2 μl of a mixture of dNTPs and one of the four ddNTPs. The PCR amplification that followed consisted of the following program: an initial 5 min at 94°C; 25 cycles of denaturing at 94°C for 5 min, annealing at 55°C for 5 min, and elongation at 72°C for 5 min; an additional 72°C for 10 min; and held ∞ at 4°C. At the end of the PCR reaction 3 μl of stop/loading buffer was added to the reaction in each of the four tubes. The sequencing samples were heated at 95°C for 3 min to denature the DNA and 1.7-2 μl of each sample was applied to the gel.

PCR primers derived from the DNA sequence of pAL202 (Table 2) were used in a comparison study to test for the distribution of DNA sequences present in pAL202 in six other *H. pylori* Alaskan plasmids, pAL203, pAL207, pAL208, pAL209, pAL226, and pAL236. Plasmid DNA of these Alaskan strains, pZErO, and pZErO::TN/KAN were

amplified using the REDTaq Genomic DNA polymerase kit (Sigma) or the BIOLASE Red™ DNA Polymerase kit (BIOLINE). Each reaction contained 50-60 ng of plasmid DNA, 1.0 μl each of a forward and a reverse primer (25 pmol/μl), 4 μl of 10 mM dNTP, 1.5 μl of 50 mM $MgCl_2$, 5 μl of the 10X buffer, 1.0 μl of Red Taq DNA polymerase (1U/μl), and H_2O to a total volume of 50 μl. The PCR amplification started with an initial 5 min at 94°C; 25 cycles of denaturing at 94°C for 1 min, annealing at 55°C for 1 min, and elongation at 72°C for 2 min; an additional 72°C for 10 min; and held ∞ at 4°C. The PCR products were analyzed by electrophoresis on 0.8% agarose gels. PCR products obtained with the same set of primers were gel purified using the Zymoclean Gel DNA Recovery Kit (ZymoResearch) as recommended by the manufacturer and concentrated with 0.3 M sodium acetate and 100% ethanol.

DNA Sequencing

The plasmid DNA of pAL202 recombinants was used to sequence pAL202 using the EZ::TN™<R6K*γori*/KAN-2> transposon specific KAN-2 FP-1 Forward and R6KAN-2 RP-1 Reverse primers (Epicentre) in bi-directional sequencing. The LI-COR 4000L automated sequencer and SequiTherm EXEL™II Long-Read™ DNA Sequencing Kit-LC (Epicentre) were used in the sequencing reactions. From every region both strands were sequenced at least twice. Labeled sequencing primers (Table 2) were designed and obtained from LI-COR to fill gaps and also to remove ambiguities within obtained sequences.

Sequencing plates were washed with ultra-pure H_2O and 100% isopropanol. Three hundred and thirty microliters of silane stock solution and 10 µl of 10% acetic acid were mixed in a microcentrifuge tube and rubbed between the ears of the plates. The plates were allowed to dry for 30 min before plates were sandwiched and the gel solution was poured. The 66 cm denaturing polyacrylamide gels contained 3.6% polyacrylamide and 8 M urea. The gel solution contained 25.2 g urea, 4.8 ml Long Ranger® 50% polyacrylamide, 14.4 ml 5X TBE, and ultra-pure H_2O to a total weight of 67.5 g. Once the urea was dissolved, 400 µl of 10% ammonium persulfate and 40 µl of TEMED was added to the gel solution, which was immediately poured and allowed to polymerize for 2 hr. The gel apparatus was assembled and 125 ml of 1X TBE buffer was added to the bottom buffer tray, and 600 ml of 0.75X TBE buffer to the upper buffer tray. The laser was focused on the sequencing machine while the gel pre-ran for 30 min. In each well 1.7-2.0 µl of each sample was loaded. Initially, the gel was run at 10 V for 15 min after which the voltage was changed to 2000 V. The gel ran overnight or up to 30 frames. Sequencing data were obtained from the Base ImagIR v4.2, a computer software program that is part of the LI-COR sequencer.

DNA Sequence Analysis

Sequences from the individual regions were aligned in the AlignX® program from Vector NTI® Suite 7.0 and 8.0 (InforMax). Consensus sequences from the alignments were submitted to the BLASTN program of NCBI (4, 5) not only for validation of *H. pylori*, but also to confirm that the sequences obtained were not transposon sequences.

Validated sequences were used in the construction of pAL202 using the ContigExpress® from Vector NTI, a program for assembling small overlapping DNA fragments into longer contiguous sequences. The sequence of pAL202 was analyzed with the computer software program Vector NTI® Suite 7.0 (InforMax®). The program was used to find possible open reading frames (ORFs), promoter consensus sequences, and ribosome binding sites. Internet available software was used to find the GC content of the plasmid and repeated sequences. The nucleotide sequence and ORF sequences were submitted to the BLAST program of NCBI (4, 5) to find homologies to known sequences. ORFs were analyzed further with several software programs for protein analysis available through the ExPASy Proteomics tools (http://us.expasy.org/).

Mobilization and Mating Studies

ORF10, ORF11, ORF12, and ORF13 showed some identity to the mobilization genes of *E. coli* indicating that pAL202 could be mobilized if a conjugative plasmid were present. A putative *nic* sequence found in ORF6 (TATCCTG; bp 3379-3385) might be the initiation site for transfer of DNA in the mobilization process. Mating procedures were performed to determine if pAL202 could be mobilized by pRK24, a Tc^r, Kan^s, derivative of RP4. Two recombinants of pAL202, pAL202::TN/KAN21 and pAL202::TN/KAN25 in *E. coli pir*$^+$ cells were used as donors, testing for mobilization of the kanamycin resistance (Kan^r) marker present on the recombinant plasmids. The pRK24 mobilizing strain was JA221 and the final recipient was IR3272. Isolation of the Nal^r IR3272 was described previously.

The first step was to mate JA221(pRK24) with the *pir*$^+$(pAL202::TN/KAN) recombinants, selecting for tetracycline resistance. Two ml of LB broth was inoculated separately with each *pir*$^+$(pAL202::TN/KAN21 or 25), and JA221(pRK24), and incubated overnight at 37°C in a water-bath. The next morning 100 µl of each strain was transferred to 2 ml of fresh broth and incubated for 90 min with shaking at 37°C. For the mating, 200 µl of JA221(pRK24) and 1.8 ml of *pir*$^+$(pAL202::TN/KAN21 or 25) were mixed and incubated for 2 hr at 37°C in a water-bath. Serial dilutions of the mating mixture were plated in 100 µl aliquots on LB agar containing kanamycin (50 µg/ml) and tetracycline (50 µg/ml). The plates were incubated overnight at 37°C. Each transconjugant strain was restreaked on selective media.

In the second step the transconjugants, *pir*$^+$(pAL202::TN/KAN21; pRK24) and *pir*$^+$(pAL202::TN/KAN25; pRK24), were used as donors to test for mobilization of the pAL202::TN/KAN recombinants with IR3272 as the final recipient. The matings were done either as liquid matings or spotted on plates. Each *pir*$^+$ strain, pAL202::TN/KAN21; pRK24 and pAL202::TN/KAN25; pRK24, and IR3272 were grown overnight at 37°C in 5 ml of LB broth. The next morning each strain was transferred to 40 ml of fresh LB broth and grown at 37°C with shaking to an OD_{600} of 0.6. Two hundred microliters of *pir*$^+$(pAL202::TN/KAN21 or 25; pRK24) and 1.8 ml of IR3272 were mixed and incubated for 60 min in a 37°C water-bath. Serial dilutions of the donor strains were plated individually on LB agar to determine the viable count of the strains. After the 60 min incubation period serial dilutions of the mating mixtures were plated in 100 µl aliquots on M9 minimal medium containing histidine, tryptophan, and lysine at 40 mg/L

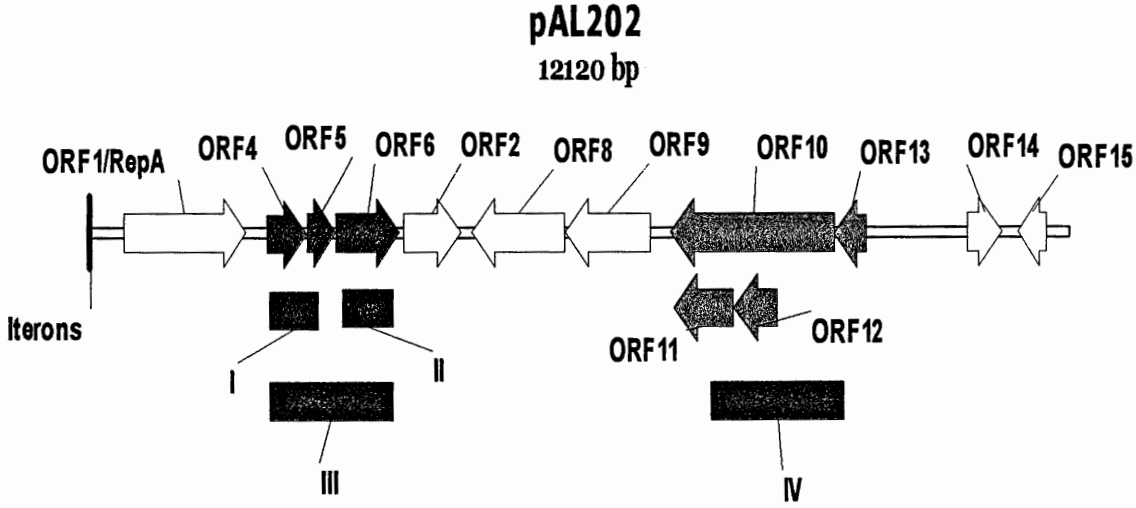


Figure 1. Test for distribution of pAL202 genes in other Alaskan plasmids
Genes in regions I-IV were amplified: region I (ORF4-ORF5), region II (ORF6), region III (ORF4-ORF6), and region IV (ORF10-ORF13, *Mob-like* genes). Region I was amplified with primer set IR1 and IR2, region II with IR3 and IR4, region III with IR1 and IR4, and region IV with IRMob1 and IRMob2 (Table 1).

each, streptomycin (25 μg/ml), nalidixic acid (20 μg/ml), and either kanamycin (50 μg/ml) or tetracycline (50 μg/ml). The plates were incubated at 37° C for 1-3 days.

In a separate plate mating experiment with these same strains, 100 μl of each mating mix was spotted separately on an LB agar plate and incubated overnight at 37°C. The next morning the cells of each mating mixture were scraped from the plates and suspended in 500 μl of TE buffer. Serial dilutions were plated on selective M9 medium as described above followed by incubation at 37°C for 1-3 days.

Test for Distribution of pAL202 Genes in Other Alaskan Plasmids

The test for the distribution of pAL202 genes in other Alaskan plasmids was performed in three parts, two PCR studies, and one hybridization study. The first part was PCR amplification as described previously with plasmid DNA of the Alaskan *H. pylori* strains used as template. PCR primers derived from the DNA sequence of pAL202 (Table 2) were used to test for the distribution of DNA sequences present in pAL202 in six other *H. pylori* Alaskan plasmids, pAL203, pAL207, pAL208, pAL209, pAL226, and pAL236. Genes in four different regions were amplified (Fig. 1); region I (ORF4-ORF5), region II (ORF6), region III (ORF4-ORF6), and region IV (ORF10-ORF13, *Mob-like* genes). Region I was amplified with primer set IR1 and IR2, region II with IR3 and IR4, region III with IR1 and IR4, and region IV with IRMob1 and IRMob2. The PCR amplification reactions were described above.

In the second part, hybridization studies were performed to confirm the PCR findings. The same PCR primers mentioned above were used to make probes to hybridize to plasmid DNA from the Alaskan plasmids. The NEBlot® Kit (New England

BioLabs Inc.) for preparation of high-specific-activity DNA probes was used to label the probes. To make the probes, 10-20 ng of plasmid DNA (template) in a total volume of 33 µl of H_2O was added to a microcentrifuge tube, boiled for 5 min, placed on ice for 5 min, then centrifuged briefly. To the tube was added; 5 µl 10X labeling buffer, 2 µl 0.5 mM dATP, 2 µl 0.5 mM dTTP, 2 µl 0.5 mM dGTP, 1.0 µl of DNA polymerase 1-Klenow Fragment (5 units), and 5 µl α-dCTP^{32} (3000 ci/mmol, 50 µCi). The reaction was incubated at 37°C for 1 hr. Following incubation, the labeled probes were purified for the removal of the unincorporated nucleotides using Sephadex-G50 Nick columns (Amersham Biosciences) as recommended by the manufacturer. The amount of radioactivity incorporated was determined with a Packard Tri-Carb 2100 TR scintillation counter. Between 40 and 50 million counts were used in each hybridization reaction.

Plasmid DNA was digested with *Hin*dIII and passively transferred to a positively charged Nytran® nylon transfer membrane by using the Turboblotter™ Rapid Downward Transfer Systems (Schleicher & Schuell). The alkaline transfer of the DNA to the nylon membrane occurred as follows: the agarose gel was incubated twice for 30 min in denaturing buffer (3 M NaCl, 0.4 M NaOH) and washed in transfer buffer (3 M NaCl, 8 mM NaOH, 2 mM sarkosyl) for 15 min. The membrane was soaked in H_2O for 15 min. After the transfer apparatus was assembled, the DNA was passively transferred to the nylon membrane overnight. Following transfer, the membrane was washed for 5 min in 1X neutralizing buffer (1.0 M phosphate buffer, pH 6.8). The membrane was added to a hybridization tube containing 10 ml of ULTRAhyb hybridization buffer (Ambion) and was prehybridized for at least 1 hr at 42°C in a Hybridization Incubator Model 310

(Robbins Scientific). The radioactive probe was denatured at 95°C for 5 min, placed on ice for 5 min, and then added to the hybridization tube and incubated overnight. Following hybridization, the hybridization buffer was discarded; the membrane was washed three times with a wash buffer (50 ml of 2X SSC, 0.1% SDS), and finally incubated at 42°C for 30 min with 50 ml wash buffer. The membrane was exposed to a Kodak screen for at least 1.5 hr followed by the detection of hybridization with the BIO-RAD Molecular Imager® FX (BIO-RAD) and analysis with the Quantity One® (BIO-RAD) computer program.

Indications that the plasmid template might have been contaminated with chromosomal DNA led to the third part which consisted of a second PCR amplification, using exonuclease-treated template. To remove chromosomal DNA from the isolated plasmids, plasmid DNA was treated with ATP-dependent exonuclease (Qiagen) and with Plasmid-Safe™ ATP-Dependent DNase as described by the manufacturer (Epicentre). The enzymes degrade linear but not circular, supercoiled DNA. PCR studies were repeated with treated and untreated DNA. A recombinant plasmid of pAL202 (pAL202::TN/KAN25) was used as a positive control, and pZErO-2.1™ (Invitrogen) transformed with the EZ::TN™<R6Kγ*ori*/KAN-2> transposon (pZErO::TN/KAN) and electroporated into the *E. coli pir*$^+$ cells was used as a negative control.

Transformation of AL202 and HU71

Three different single insertion mutants of pAL202 were used in separate attempts to transform both AL202 and HU71, a strain with no plasmid. These single insertion mutants were: pAL202::TN/KAN25, bp 414, in which the transposon was inserted

between *orf*1/*rep*A and its promoter; pAL202::TN/KAN21, bp 4370, in which the transposon was inserted in *orf*2; and pAL202::TN/KAN2, bp10161, in which the transposon was inserted in another location in the plasmid.

In the natural transformation procedure (85), *H. pylori* cells from frozen cultures were plated on cold BHI-YE agar plates and incubated at 37°C under microaerophilic (5-10% CO_2) conditions for 2-5 days. After 2 days growth, the cells were scraped and spread on a cold brain heart infusion-yeast extract (BHI-YE) agar plate in a diameter of 8-10 mm and incubated for 5 hr as mentioned above. Recombinant plasmid DNA (2-10 μg) was spotted on the inoculated area. Following an incubation period of 16-24 hr, the cells were scraped from the plates and suspended in 1ml of TE buffer or BHI-YE broth. Dilutions were plated in 100 μl aliquots on selective (BHI-YE plus 50 μg/ml kanamycin) and non-selective (BHI-YE) medium.

Three different procedures were used in the attempt to transform AL202 and HU71 through electroporation. In all three procedures 10 μg of plasmid DNA resuspended in 10-20 μl of UPH_2O was used in each electroporation attempt. In the first procedure described by Wang *et al.* (85) cells were made electrocompetent by suspension in 10% glycerol. Electroporation was carried out according to the manufacturer's procedures in a TransPorator™ Plus Electroporation system (BTX) at voltages ranging from 1.5 kV to 2.5 kV, 40 μF, 130 Ω, in a 1 or 2 mm cuvette. Following electroporation, the cells were plated immediately on a cold non-selective plate (BHI-YE) and incubated for 12 hr at 37°C under microaerophilic conditions (5-10% CO_2). The 12 h old cells were

scraped and suspended in 1.0 ml of BHI-YE broth, diluted, and plated in 100 µl aliquots on selective and non-selective medium.

In the second procedure, described by Segal *et al.* (73), cells were made electocompetent in electroporation buffer (EPB) [272 mM sucrose, 15% glycerol, 2.43 mM K_2HPO_4, 0.57 mM KH_2PO_4, pH 7]. Electroporation was carried out as described above. Following electroporation, cell volume was brought to 1.0 ml with phosphate-buffered saline (PBS) at pH 7.4. The cells were further diluted and plated in aliquots of 100 µl on selective and non-selective medium and incubated under microaerophilic conditions.

The third procedure was a combination of the above procedures and a method described by Edwards *et al.* (28) in which increased temperature temporarily inactivates the restriction modification complexes of cells. In this procedure AL202 and HU71 cells were first incubated at 50° C for 30 min to inactivate the enzymes, cooled on ice for 2 min, and then made electrocompetent. Electroporation, dilution, and plating of cells were as described for the previous two procedures.

CHAPTER III

RESULTS

Purification and Restriction Enzyme Analysis of Plasmid pAL202

Plasmid DNA of pAL202 was isolated from *H. pylori* strain AL202. Agarose gel electrophoresis (Fig. 2a) of restriction digestions of pAL202 digested with *Eco*RI produced one (12120 bp) fragment; *Hin*dIII produced five fragments (6862, 1791, 1602, 1187, and 678 bp); and *Xba*I produced three fragments (11518, 572, and 30 bp). Some of the *Hin*dIII and *Xba*I fragments were too small to be visualized on the agarose gel, and only after the whole plasmid sequence was determined could the correct sizes of the fragments be determined (Fig. 2b).

Transposition and Selection for Single Insertion

Random insertion of the EZ::TN™<R6Kγ*ori*/KAN-2> transposon (2001 bp) into pAL202 was followed by electroporation into *E. coli pir*$^+$ (low copy number) cells and *pir*-116 (high copy number) cells. NZYM or LB agar containing kanamycin (50 μg/ml) was used to select for cells transformed with recombinant plasmids. Sixty-eight transformants were recovered from the *E. coli pir*$^+$ cells. No transformants were recovered from the electroporation into *E. coli pir*-116 cells. There was one *Eco*RI site on pAL202, and one *Pst*I site on the transposon (Fig. 3a). Plasmid DNA of the transformants was isolated and double digested with *Eco*RI and *Pst*I followed by agarose gel electrophoresis. Single insertion of the transposon in pAL202 resulted in one or two

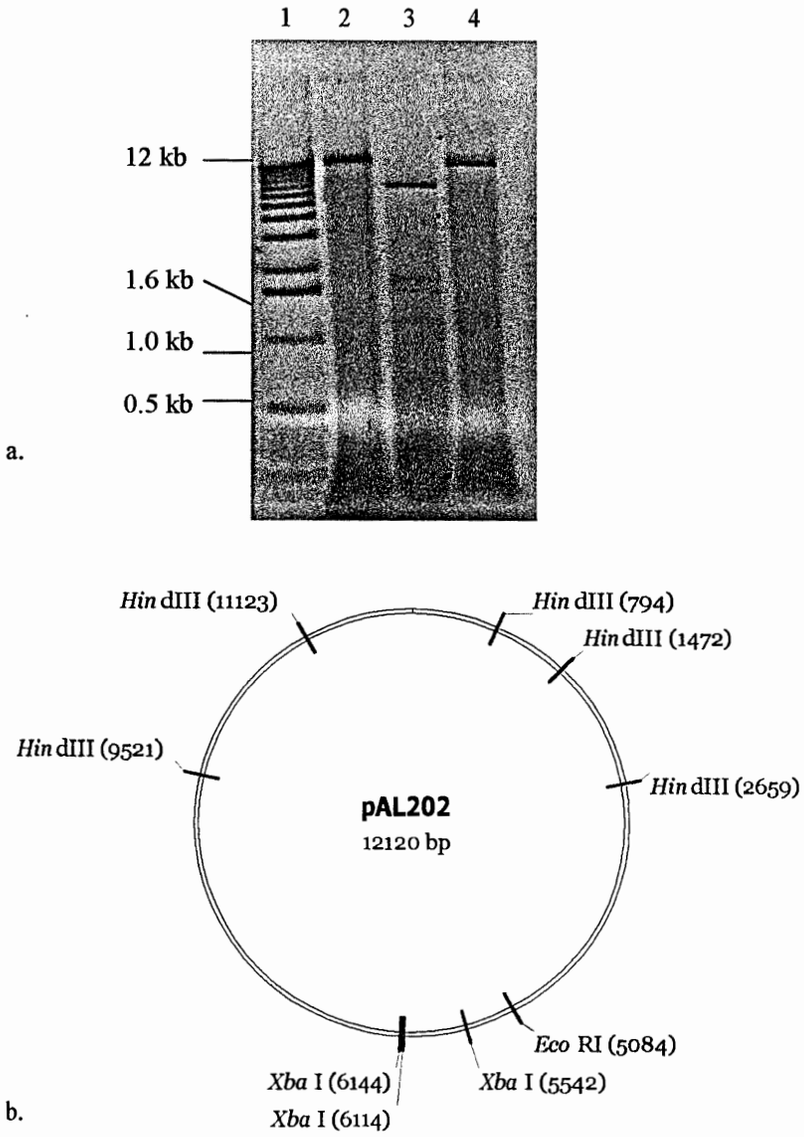
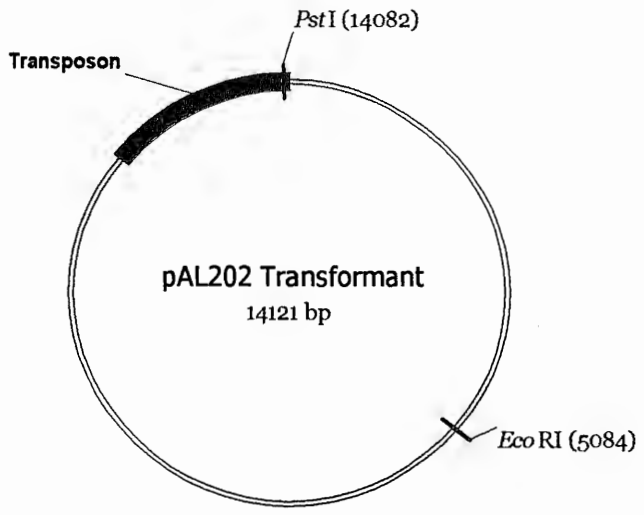


Figure 2. Restriction digestion of pAL202

a. Lane 1, kb ladder standard; lane 2, digestion with *Eco*RI; lane3, *Hin*dIII; lane 4, *Xba*I.

b. A graphic representation of restriction digestion sites on pAL202.

a.



b.

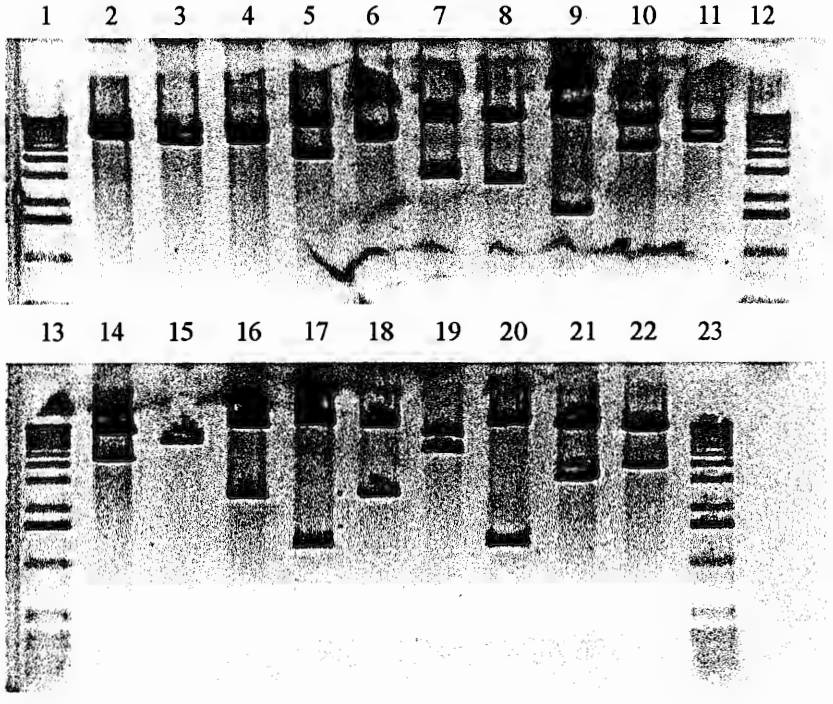


Figure 3. Restriction digestion of pAL202 transformants.
a. A graphic representation of a pAL202 transformant (single insert).
b. Restriction digestion of pAL202 transformants with *Eco*RI and *Pst*I. Lanes 1, 12, 13, and 23, are kb ladder standards. Lanes 2-11 and 14-22 are separate single insertion isolates.

bands on the agarose gel (Fig. 3b). These 19 single insert recombinants were recovered and kept for further analysis.

The EZ::TN™<R6Kγ*ori*/KAN-2> transposon also was inserted into plasmid DNAs of pAL208 and pAL236. Two recombinants were obtained for pAL208, pAL208::TN/KAN8 and pAL208::TN/KAN9, and three for pAL236, pAL236::TN/KAN2, pAL236::TN/KAN5, and pAL236::TN/KAN11. The restriction enzyme *Hin*dIII was used to test for the presence of the transposon in these recombinants. *Hin*dIII has two restriction sites in the transposon, at bp 416 and bp 1969, and produces a 1553 bp fragment on an agarose gel (Fig. 4).

DNA Sequencing

All nineteen recombinant plasmids were used in sequencing reactions in an attempt to construct the plasmid. The KAN-2 FP-1 Forward Primer and the R6KAN-1 Reverse Primer (Table 2) were used bi-directionally to determine the sequence of the recombinant plasmids. The sequence of each strand was determined at least twice. From each individual sequence reaction 200–1000 bp were obtained. The individual sequences were each submitted to the Basic Local Alignment Search Tool (BLAST) program of the National Center for Biotechnology Information (NCBI) (4) to confirm that the sequences obtained were not transposon sequences. Multiple sequences from each individual region were aligned in the AlignX® program from Vector NTI® Suite 7.0 and 8.0 (InforMax). Consensus sequences from the alignment were used in the construction of pAL202 using the ContigExpress® from Vector NTI. Labeled sequencing primers (Table 2) were designed and obtained from LI-COR to fill gaps and also to remove ambiguities within

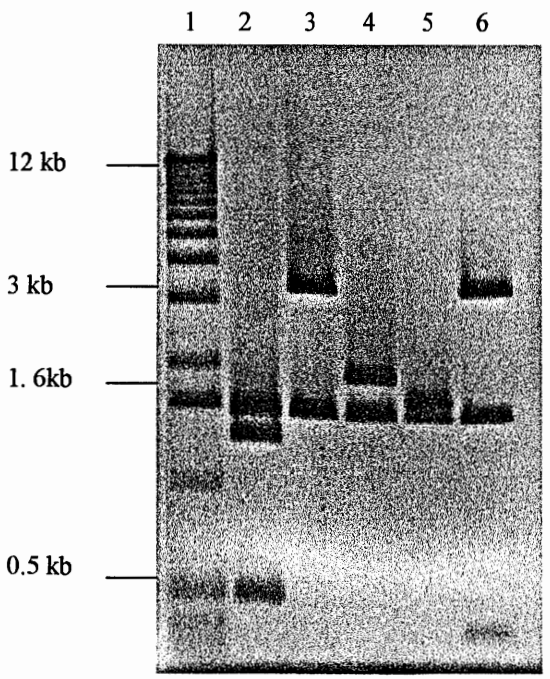


Figure 4. Single transposon insertion of pAL208 and pAL236 digested with *Hin*dIII. Lane 1, kb ladder standard; lane 2, pAL208::TN/KAN8; lane3, pAL208::TN/KAN9; lane 4, pAL236::TN/KAN2; lane 5, pAL236::TN/KAN5; lane 6, pAL236::TN/KAN11.

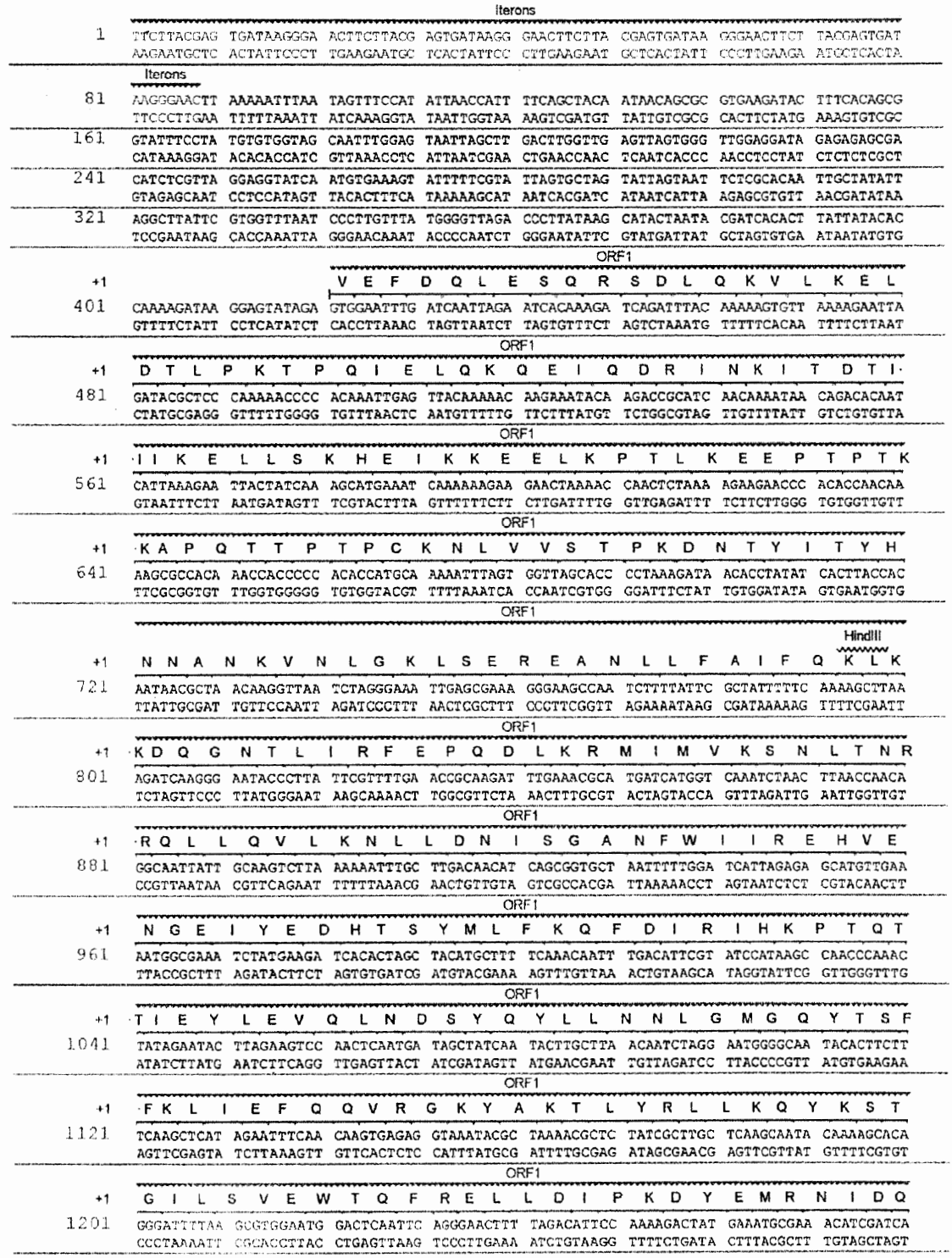
obtained constructs. Based on this sequence determination, *H. pylori* plasmid, pAL202, was established as 12120 bp. Figure 5 shows the complete sequence of pAL202. The open reading frames (ORFs), translated potential proteins, iteron location, and restriction sites are indicated. A total diagram of pAL202 structure is shown in Figure 6. Figure 7 is a diagram showing the transposon insertion sites in relation to each ORF.

DNA and ORF Analysis

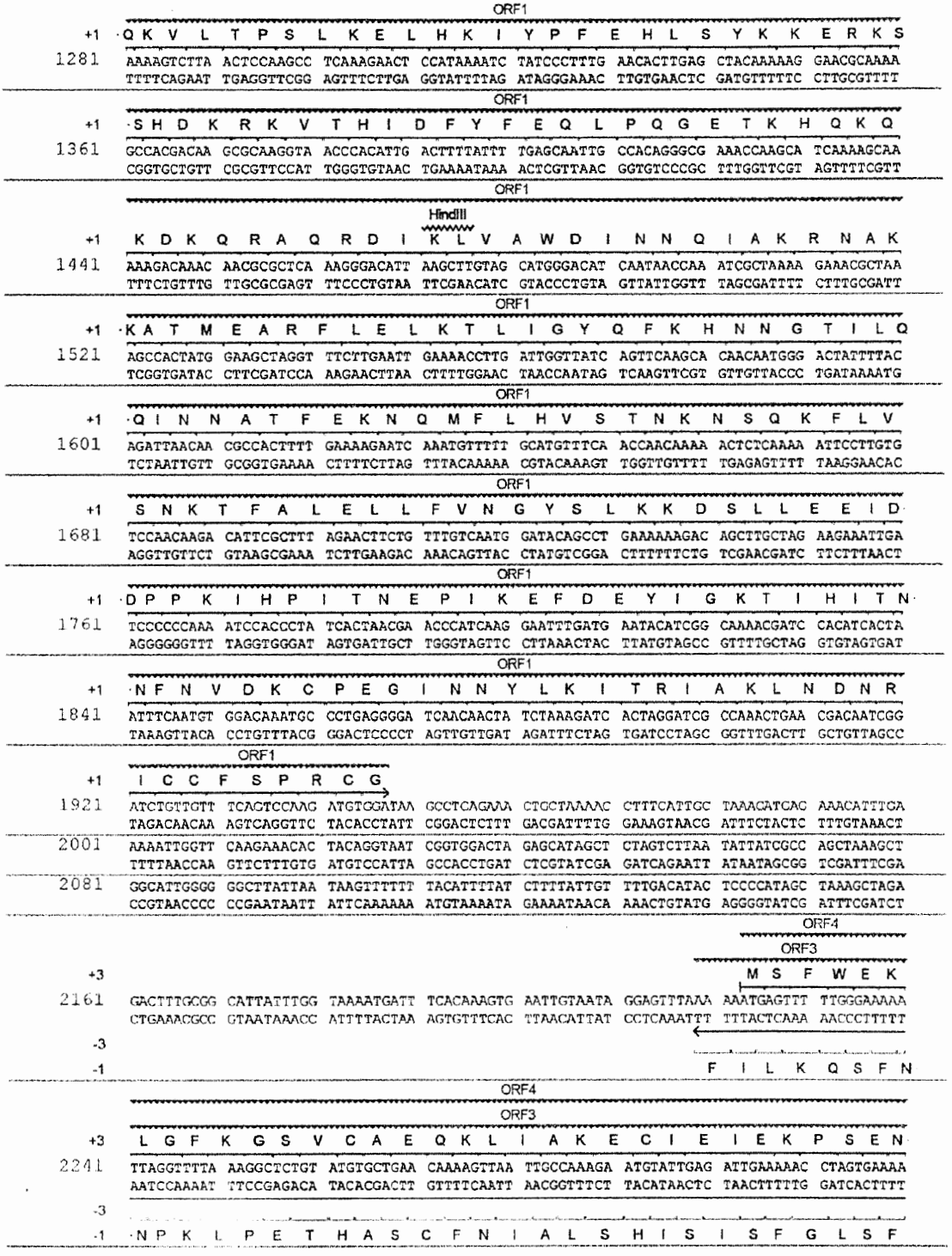
An overall GC content of 34.3% was obtained for pAL202 from the Vector NTI® Suite 8.0 program. This percent was slightly lower than that of *H. pylori* chromosomal DNA of strains J99 (38.9%) and 26695 (39.2%). However, the percent GC of pAL202 was similar to *H. pylori* plasmids with similar open reading frames such as pHel4 (34.3%), pHel5 (34.4%), and pHPM8 (33.4%) (Table 3).

The sequence of pAL202 was analyzed with the computer software of Vector NTI® Suite 7.0 and 8.0 (InforMax). ATG, GTG, and TTG were recognized at putative translation start sites (23, 81). The complete sequence had 15 putative open reading frames (ORFs), with six (ORF1, ORF2, ORF4, ORF5, ORF6, ORF14) on one strand and nine (ORF3, ORF7, ORF8, ORF9, ORF10, ORF11, ORF12, ORF13, ORF15) on the opposite strand. Possible promoter consensus sequences and ribosome binding sites (RBS) were determined by using the consensus sequences of *E. coli* (Table 4). No promoters were found for *orf10* and *orf15*, and no RBS for *orf15*.

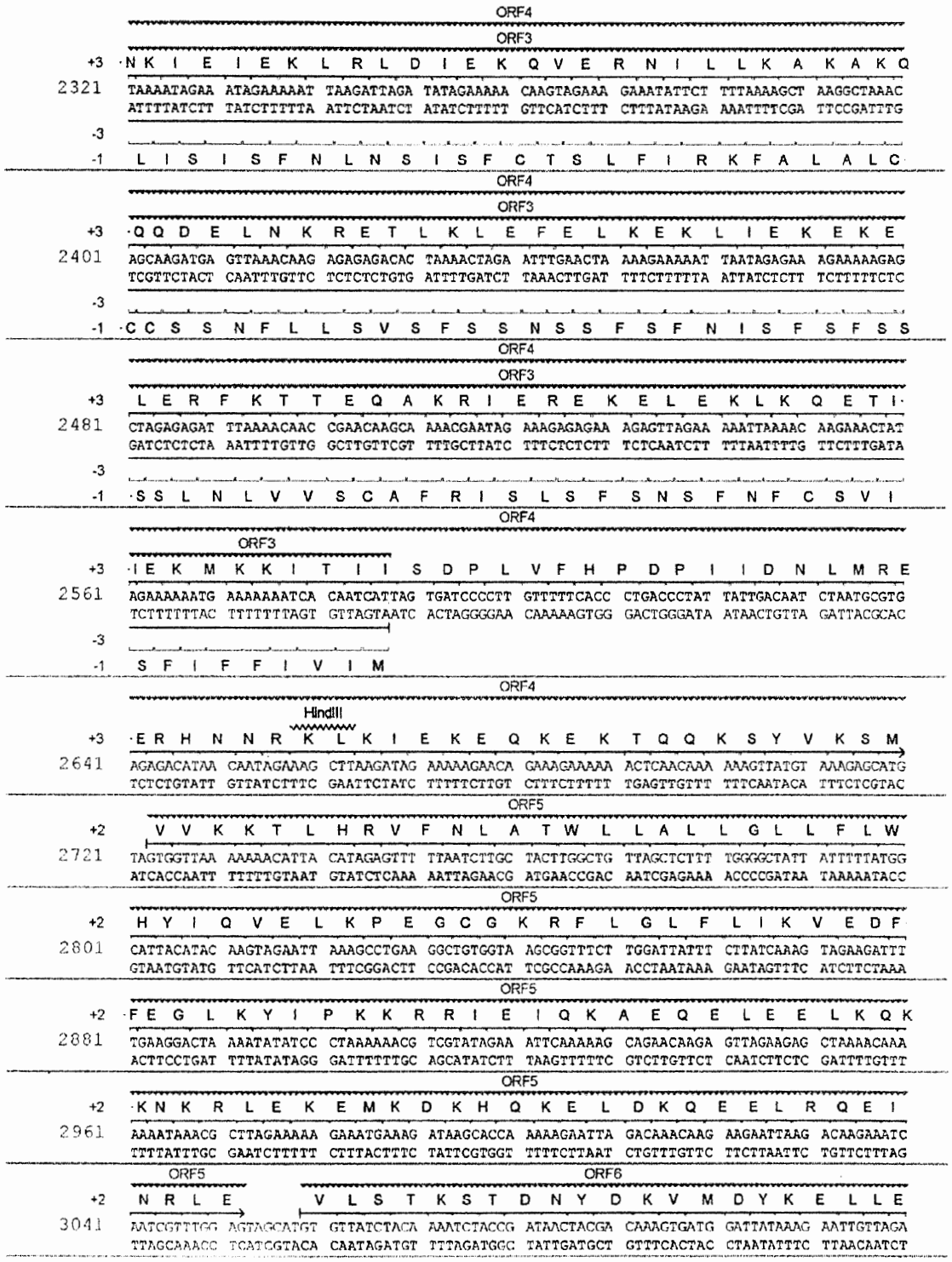
Nucleotide sequences and amino acid sequences were submitted to the BLAST program of NCBI (4, 5) to find homologies to known sequences. The greatest identity between pAL202 and *H. pylori* plasmids was to pHel4 (39) which is 10970 bp and came



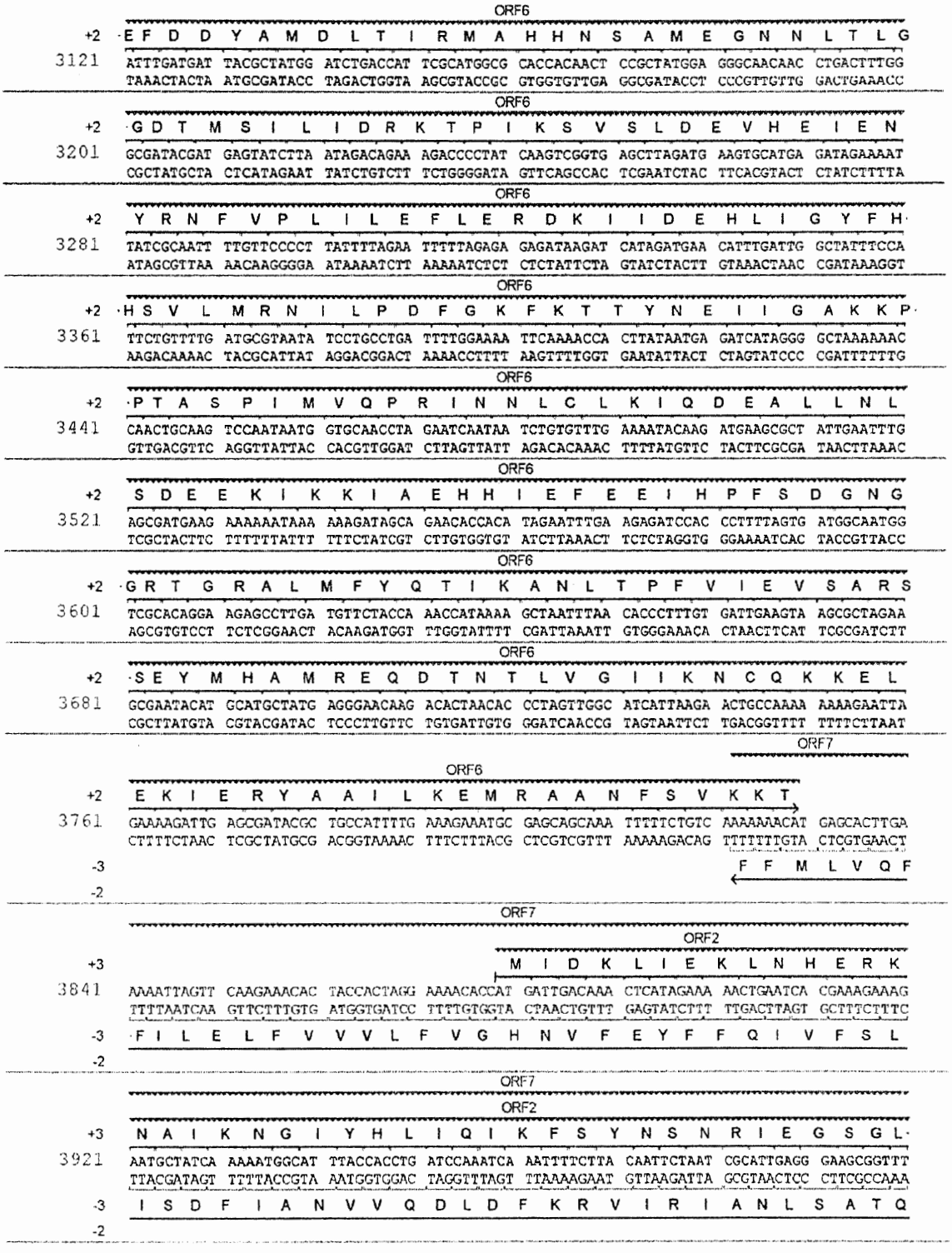
continued on next page



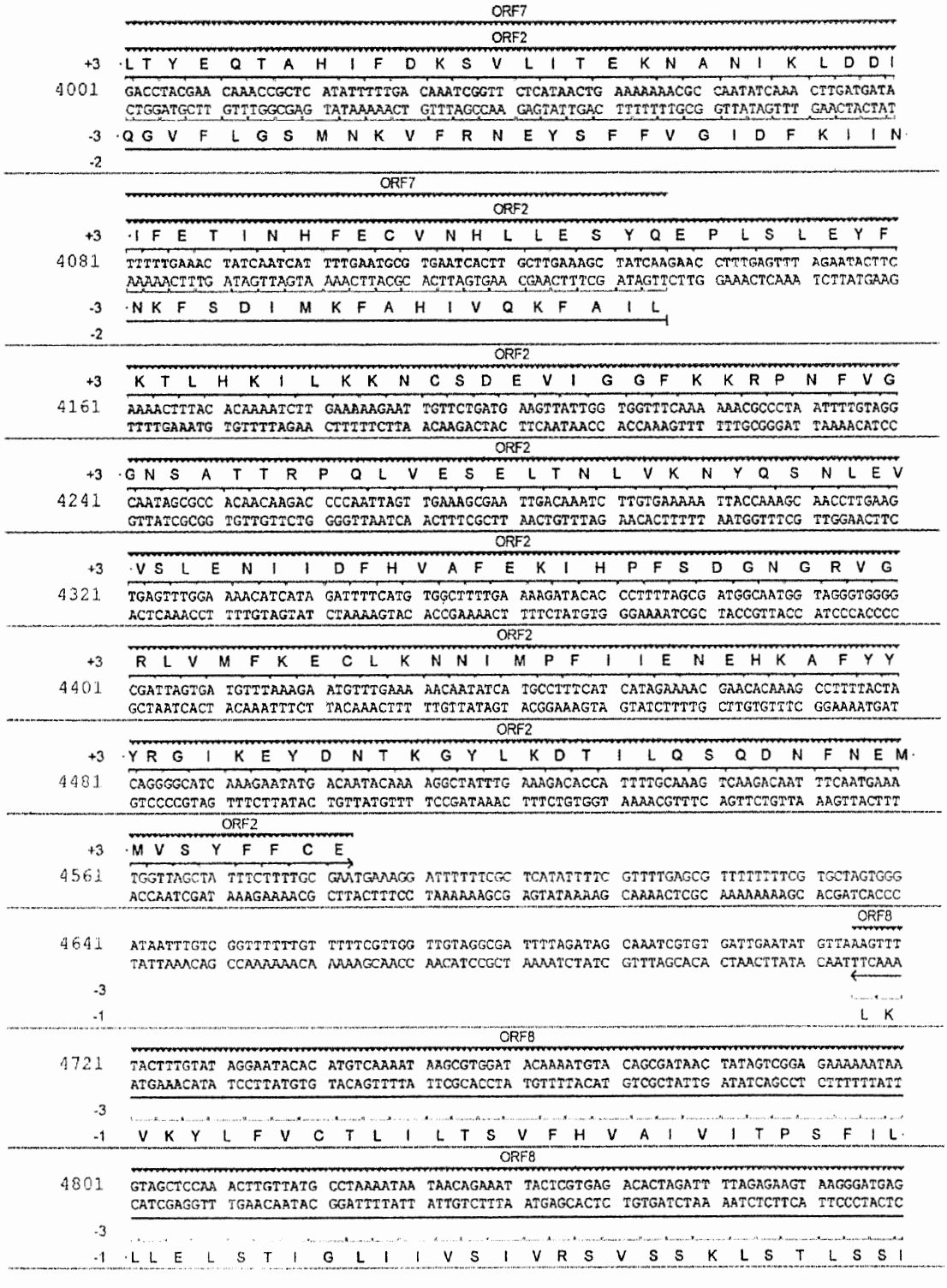
continued on next page



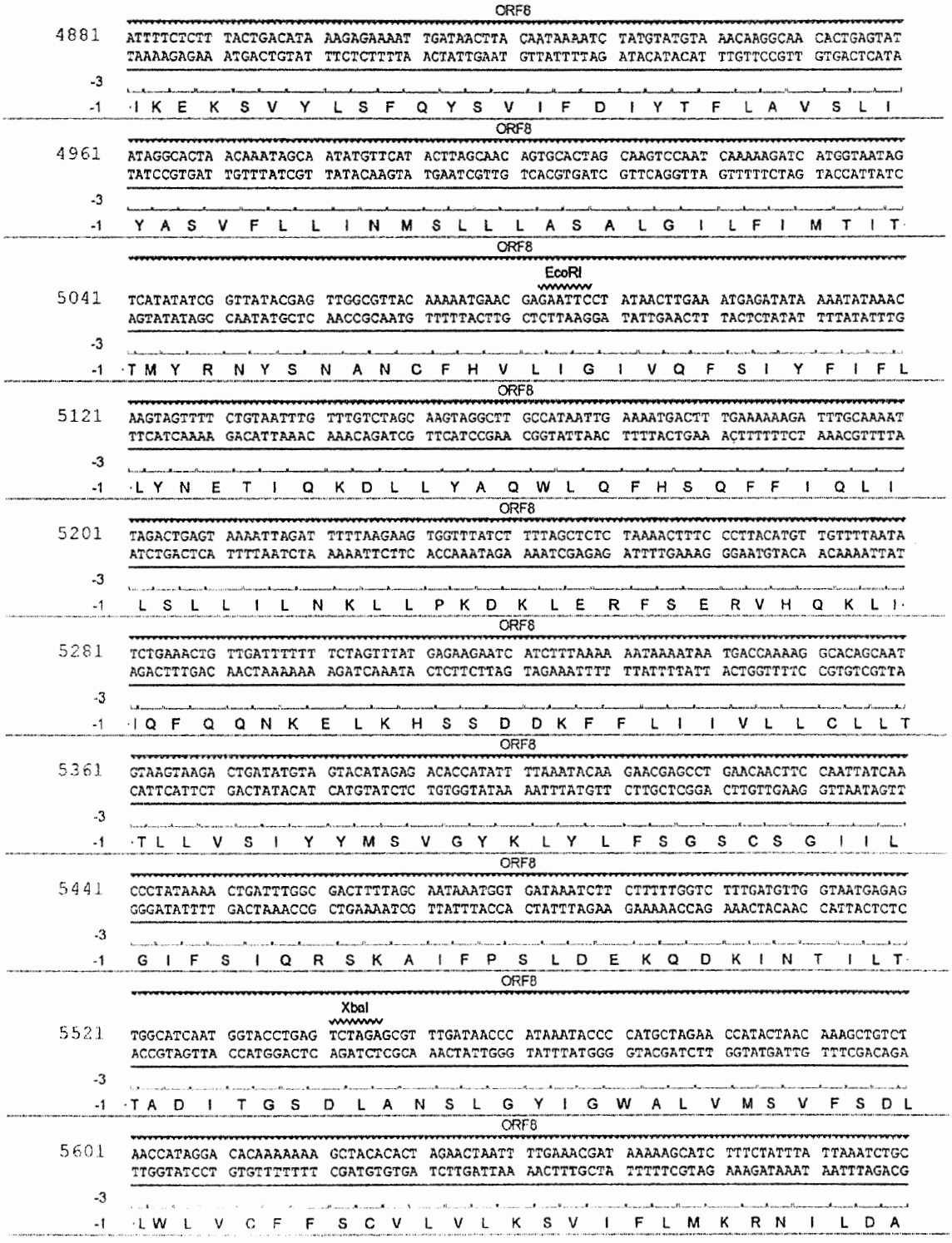
continued on next page



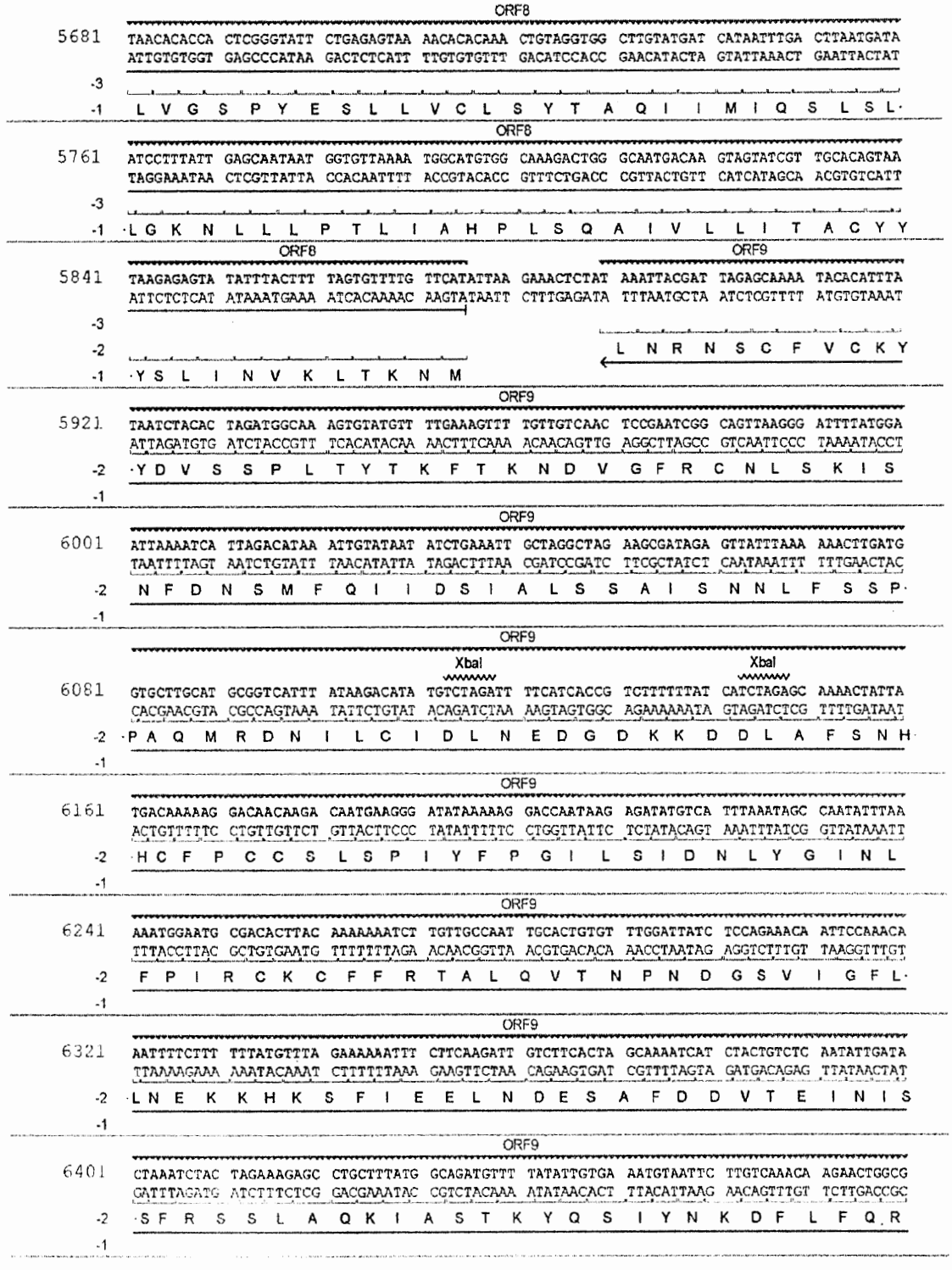
continued on next page



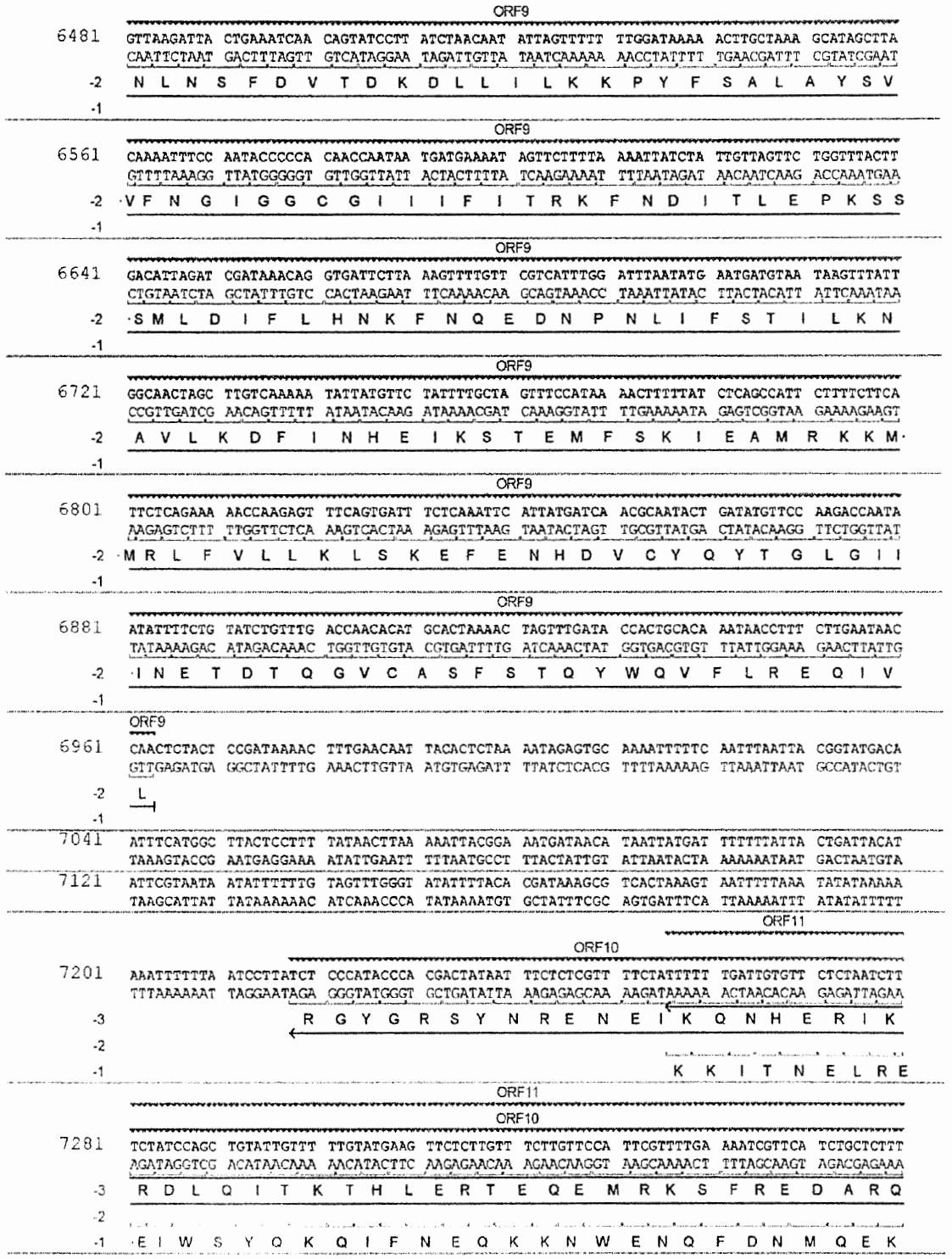
continued on next page



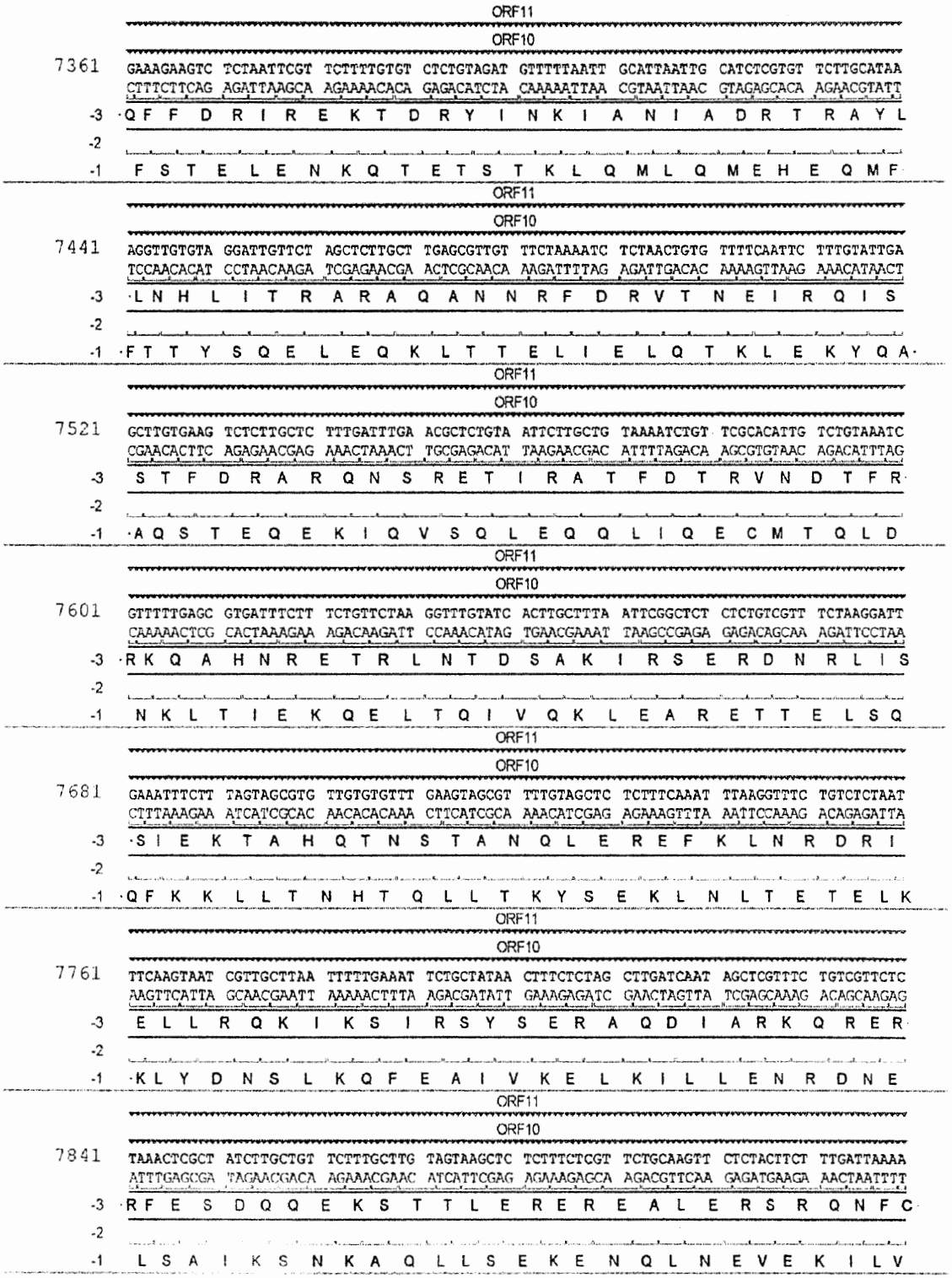
continued on next page



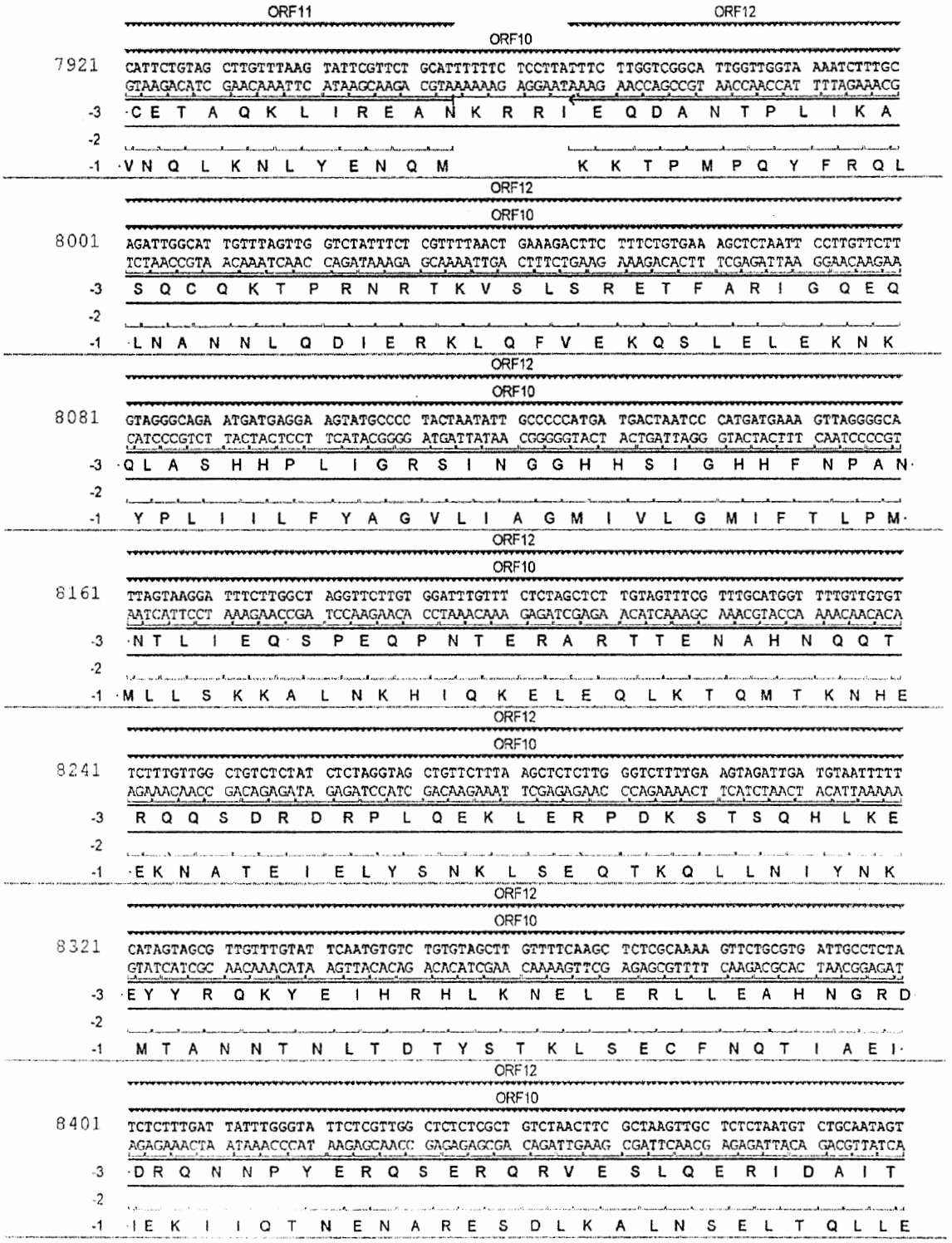
continued on next page



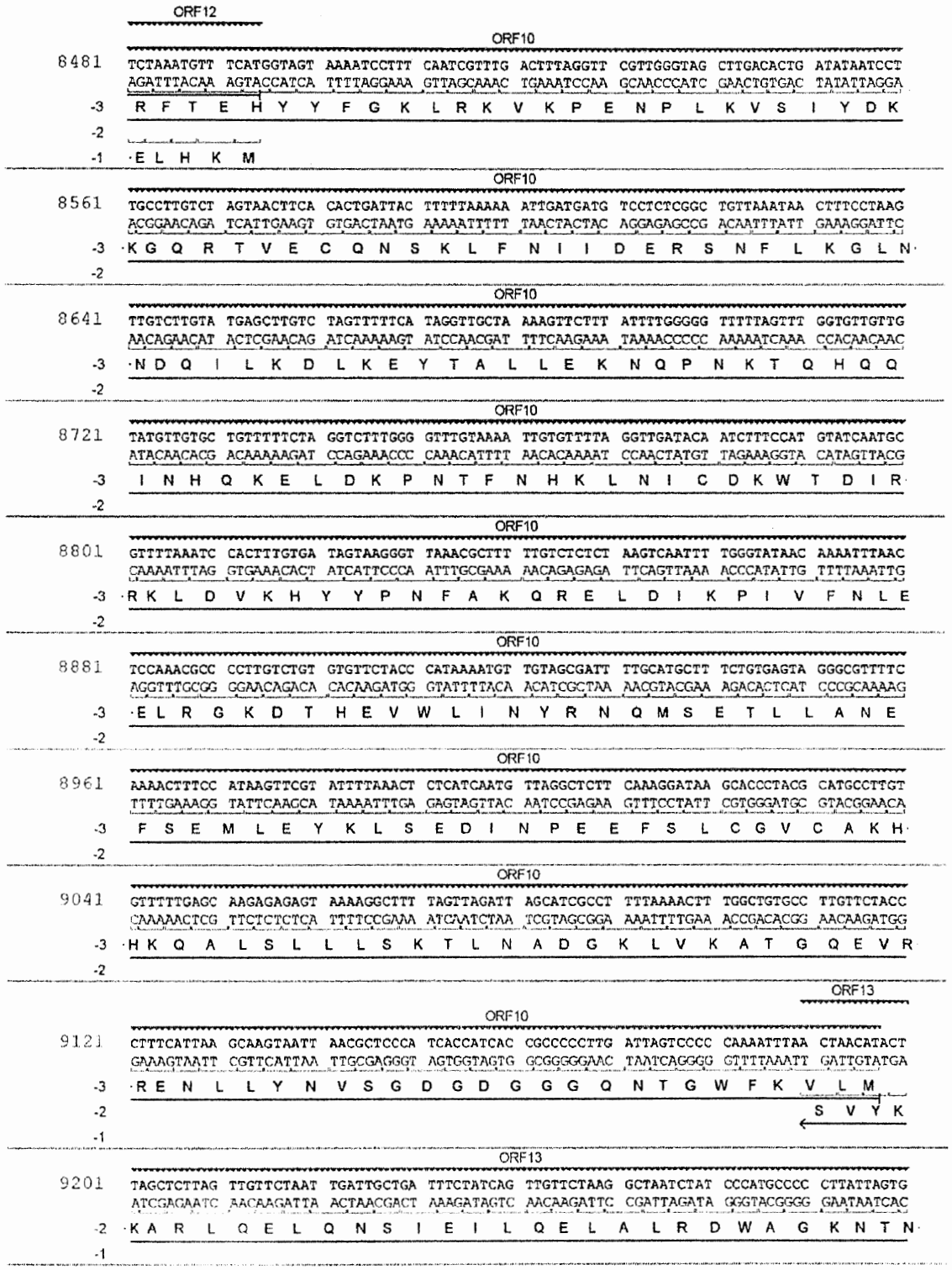
continued on next page



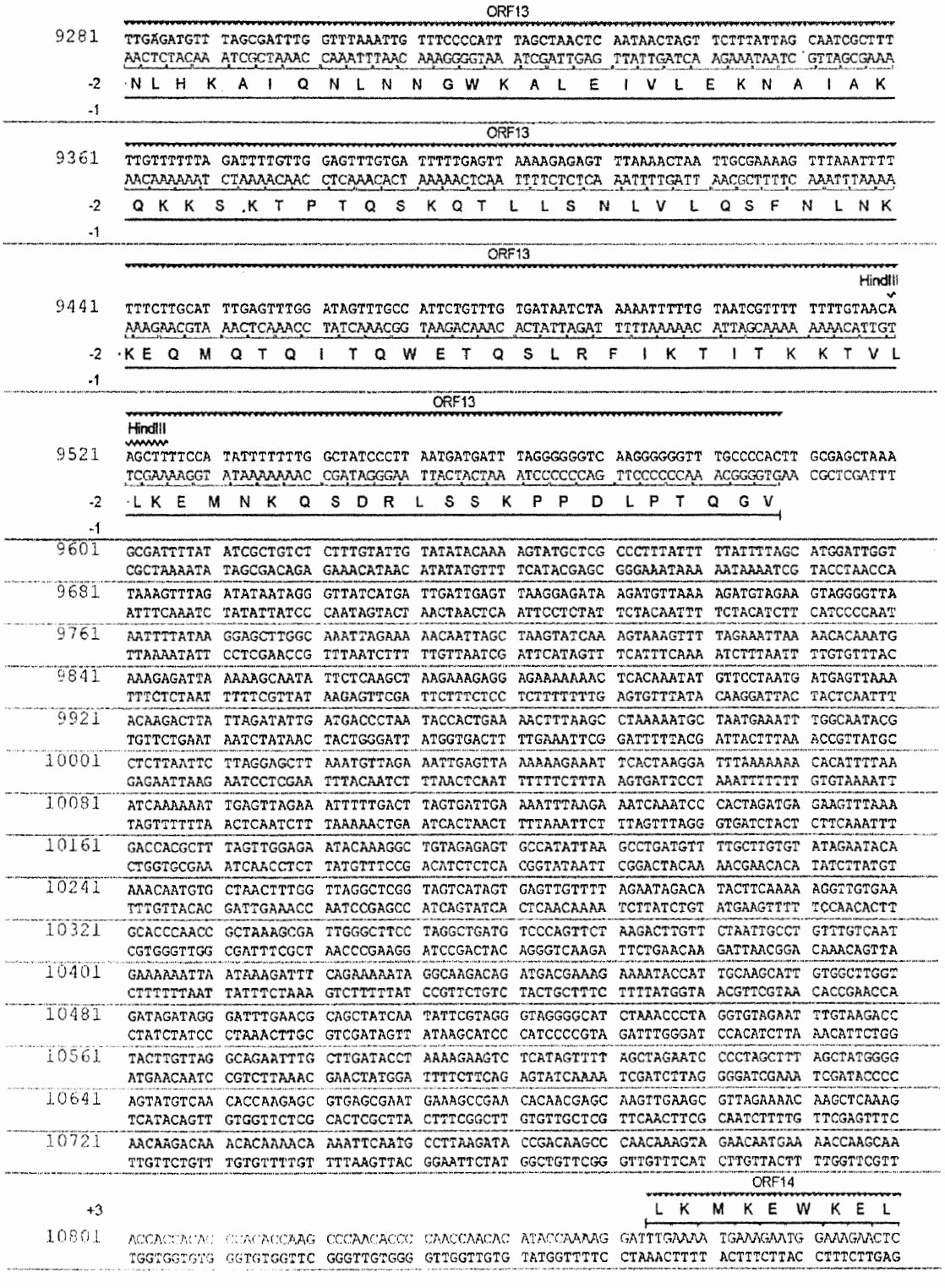
continued on next page



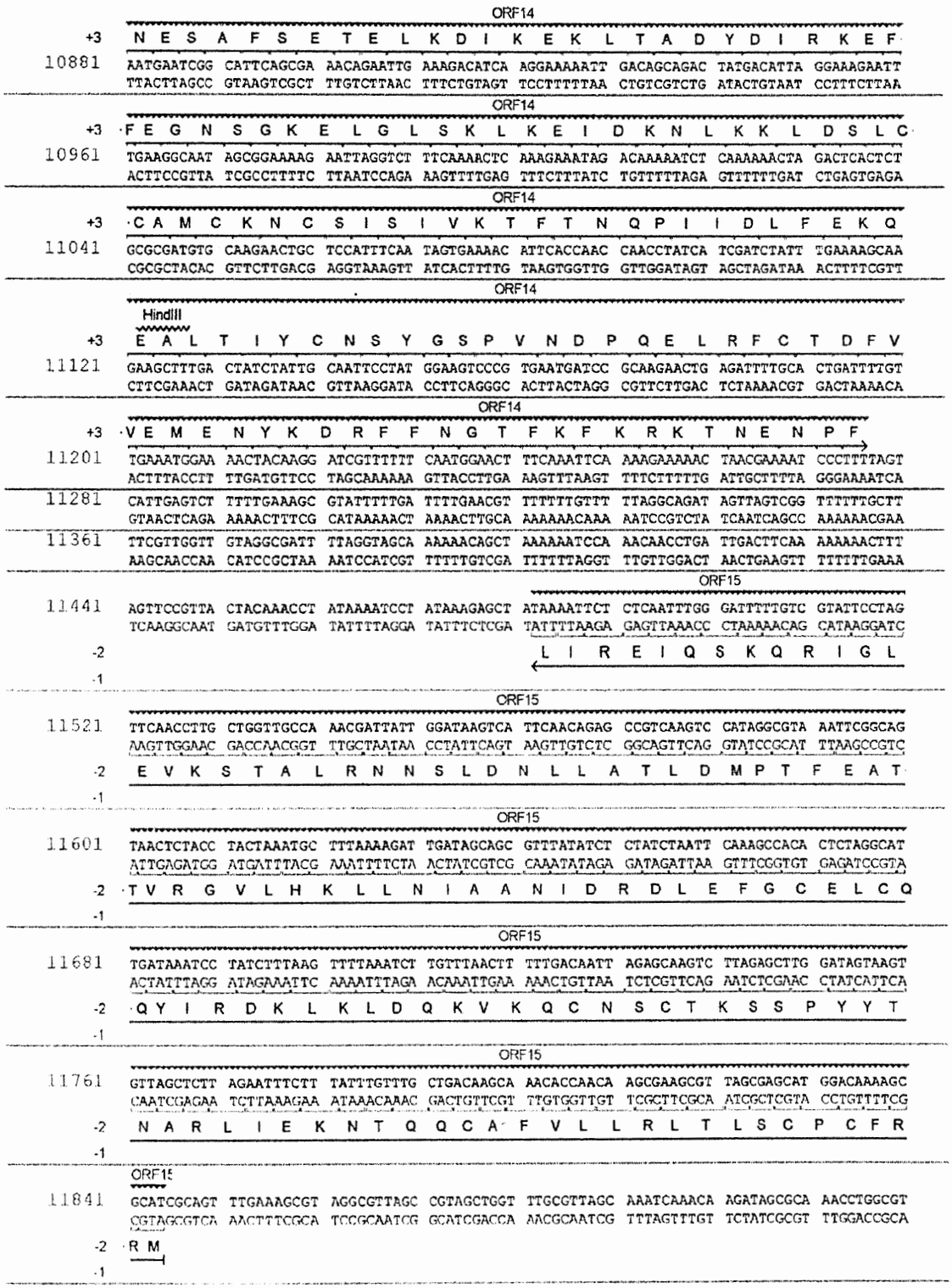
continued on next page



continued on next page



continued on next page



continued on next page

```
11921  TAGCCCAAAA AACCCCTAAA ACTAAAATTC CAAAATATGT AGCGCGTCAT GCGCGTTGTT TTTATTACAT GTTTTAACAA
       ATCGGGTTTT TTGGGGATTT TGATTTTAAG GTTTTATACA TCGCGCAGTA CGCGCAACAA AAATAATGTA CAAAATTGTT

12001  CCATGTTGTT TTTACATGTT TTTACCATGC GCGCGCATGC GAGGGATTTG GGGTTAGAAC CCCCTAAATA CCGAAGCTGT
       GGTACAACAA AAATGTACAA AAATGGTACG CGCGCGTACG CTCCCTAAAC CCCAATCTTG GGGGATTTAT GGCTTCGACA

12081  AGAGTTTCTC ATTTTTGGGT GAAAATGAAA GAATGGGAAC
       TCTCAAAGAG TAAAAACCCA CTTTTACTTT CTTACCCTTG
```

Figure 5. Nucleotide sequence of pAL202.
Iterons are indicated in blue (1-88 bp); restriction digestion sites are indicated by short wavy lines with the name of the restriction digestion enzyme above the lines; putative open reading frames (ORF) are shown as blue arrows in the direction of translation; the amino acid sequences are shown below the ORFs.

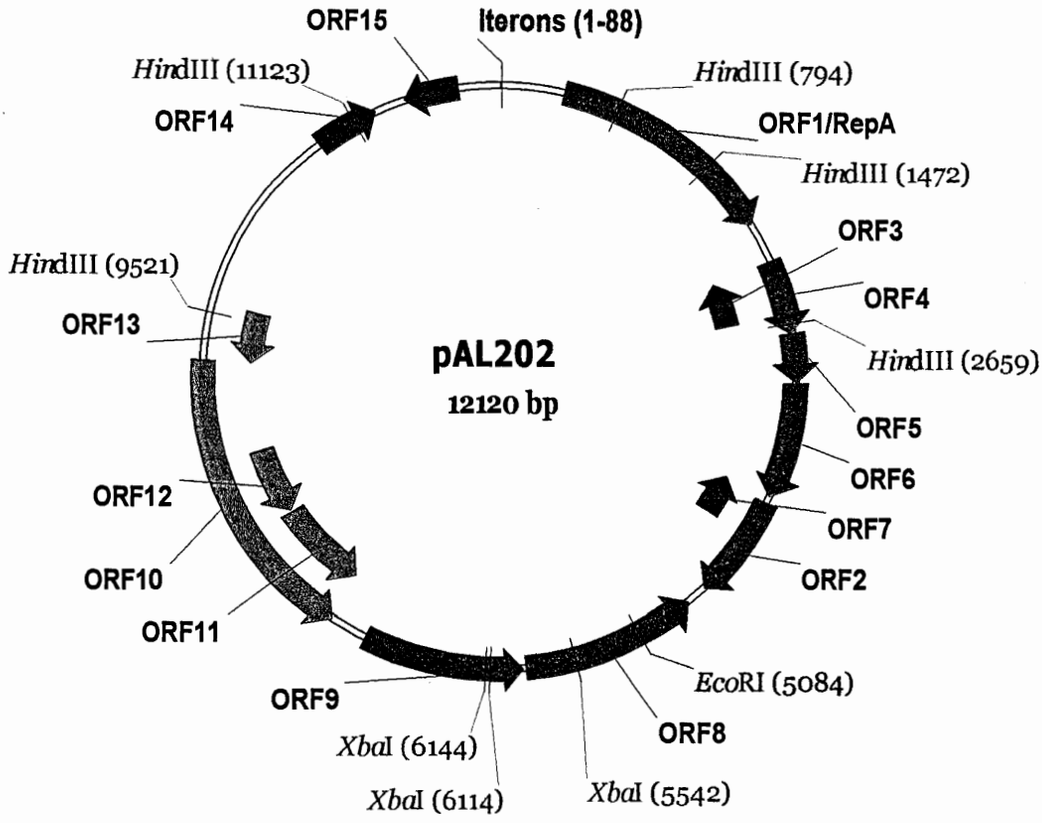


Figure 6. A graphic representation of pAL202.
Arrowheads represent putative open reading frames and the direction of translation. Location of iterons and restriction digestion sites are indicated.

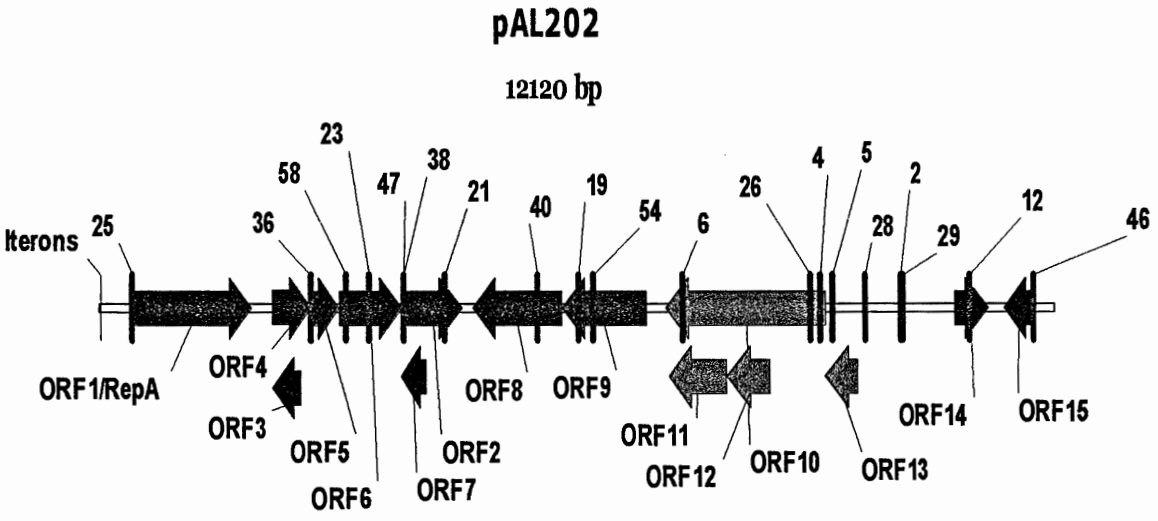


Figure 7. Single insert recombinants of pAL202.
The insertion sites of the transposon in pAL202 are indicated.
The recombinants are: 2, pAL202::TN/KAN2 (bp 10161); 4, pAL202::TN/KAN4 (bp 9135); 5, pAL202::TN/KAN5 (bp 9288); 6, pAL202::TN/KAN6 (bp 7435); 12, pAL202::TN/KAN12 (bp 11049); 19, pAL202::TN/KAN19 (bp 6069); 21, pAL202::TN/KAN21 (bp 4370); 23, pAL202::TN/KAN23 (bp 3421); 25, pAL202::TN/KAN25 (bp 414); 26,pAL202::TN/KAN26 (bp 9016); 28, pAL202::TN/KAN28 (bp 9690); 29, pAL202::TN/KAN29 (bp 10189); 36, pAL202::TN/KAN36 (bp 2716); 38, pAL202::TN/KAN38 (bp 3860); 40, pAL202::TN/KAN40 (bp 11863); 46, pAL202::TN/KAN46 (bp 11863); 47, pAL202::TN/KAN47 (bp 3858); 54, pAL202::TN/KAN54 (bp 6287); 58, pAL202::TN/KAN58 (bp 3140);

Table 3. Percent GC of *H. pylori* plasmid and chromosomal DNA

Plasmid	% GC
pAL202	34.3
pHPM8	33.4
pHPM179a	35.5
pHPM179b	36.2
pHPM180	37.0
pHPM185	33.1
pHPM186	35.9
pHel1	36.1
pHel4	34.3
pHel5	34.4
pHPS1	37.3
pHPO100	36.2
pHP51	35.3
pHP489	33.3
pKU701	38.3
pHPK255	36.8
26695*	38.9
J99*	39.2

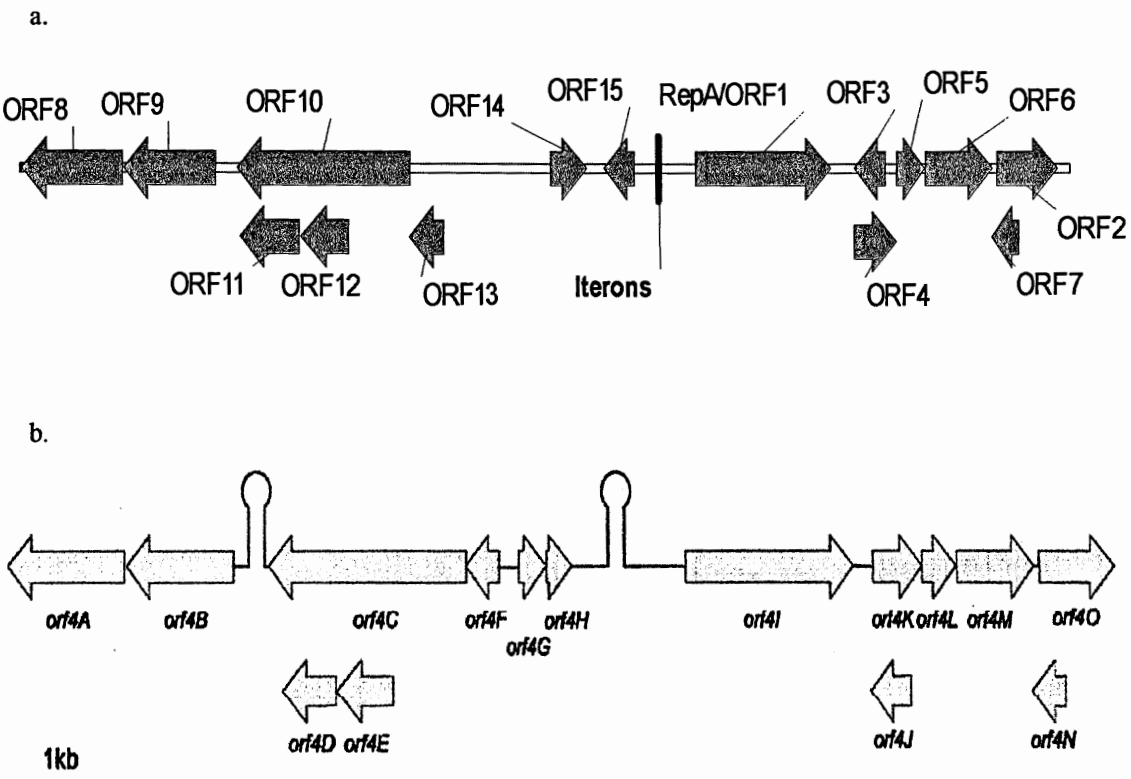
*Chromosomal DNA

Table 4. Putative promoter sequences and ribosome binding sites (RBS) of pAL202's ORFs

ORF	# aa	# bps	Sequence #	Promoter -35 TTGACA	Promoter -10 TATAAT	RSB AGGAGG
1	509	1527	421-1947	TTGTTT (344-349)	TATAAG (365-370)	AGGAGT 410-415
2	235	705	3879-4583	CTGTCA (3816-3821)	AAAAAT (3840-3845)	AGGAAA (3868-3873)
3	123	369	2219-2587	TTGTTA (2648-2653)	AATAAT (2619-2624)	AGGGGA (2594-2599)
4	166	498	2223-2720	TAAAAT (2181-2186)	TAGAGA (2157-2162)	AGGAGT (2210-2215)
5	110	330	2723-3052	GTGAGA (2638-2634)	TAAGAT (2663-2668)	AAGAGC (2712-2717)
6	257	771	3059-3829	TATAAG 10732-10737	CTTATT 10691-10696	TGGAGT (3048-3053)
7	105	315	3822-4136	TTGAAA (4214-4219)	AACAAT (4189-4194)	CAAAGG (4140-4145)
8	387	1161	4715-5875	TTGCCA (5935-5940)	TATAAA (5917-5922)	TAGAGT (5884-5889)
9	358	1074	5890-6963	TTGAAA (7017-7022)	TAGAGT (6994-6999)	CGGAGT (6968-6973)
10	660	1980	7218-9197	-	-	AAGAGC (9203-9208)
11	233	699	7256-7954	TTAAAA (8033-8038)	AACAAT (8009-8014)	AGGAGA (7959-7964)
12	176		7967-8494	TAGACA (8566-8571)	TATCAG (8548-8571)	AGGATT (8503-8508)
13	133	399	9190-9588	TTAACC (9678-9683)	AAATAA (9655-9660)	AGCTCG (9592-9597)
14	141	423	10854-11276	CCGACA (10761-10766)	AACAAT (10782-10787)	AAAAGG (10846-10851)
15	121	121	11482-11844	-	-	-

from a European isolate. An alignment between the two plasmids with the AlignX from Vector NTI Suite 8.0 revealed 83.4% identities between the DNA sequences of the plasmids. The average percent GC of both pAL202 and pHel4 was 34%. The plasmids both had 15 ORFs with 14 positioned and arranged identically. All pAL202 ORFs except ORF14 and ORF15 had sequence and amino acid identity to pHel4. The average identity between the 13 ORFs of pAL202 and pHel4 was 79.5%. The open reading frames in these two plasmids were nearly identically positioned and oriented (Fig 8). The only exceptions were ORF4H in pHel4 on the positive strand, and ORF15 in pAL202 on the negative strand. The greatest difference between these two plasmids in pAL202 was a 2.8 kbp fragment between ORF13 and the beginning of ORF1/RepA, and a 1.7 kbp fragment in pHel4 positioned between the last part of ORF4G and the beginning of ORF4I/RepA.

The search for repeated sequences led to a four times 22 bp tandem repeat, R1, which shows great identity to the iterons of other *H. pylori* plasmids (22, 30, 37, 39, 59, 70, 79). The iterons, nucleotides 1-88, were located upstream of *orf1/repA* and are indicated in Fig. 5 and 6. In pAL202, 71% identity (bp 11275-11507) to one copy of R2, the 232 bp repeat first observed in pHPM180 (59), was found between ORF14 and ORF15. In addition, some of this region (about 300 bp) displayed identity to the right end of *orfB* a transposase gene of *H. pylori* insertion sequence, IS*Hp608*. Short stretches of about 28 bp between *orf1/repA* and *orf3* also showed identity to this same transposase gene. A copy (bp 845-881) of the 36 bp sequence that separates the two R2 repeats in pHPM180 was found in the *repA* gene. One copy (bp 6516-6548) of R3, the 33 bp repeat



D. Hofreuter and R. Haas. 2002. J. Bacteriol. 184: 2755-2766.

Figure 8. Comparison of pAL202 and pHel4 ORFs position and orientation.
a, pAL202 (in pHel4 orientation); b, pHel4. The ORFs in pAL202 and pHel4 are identically positioned and oriented except for ORF14 in pAL202 and ORF4H in pHel4.

that makes up the iterons upstream of *repB* in pHPO100 (unpublished) and pHPS1 (22), was present with 64% identity in *orf9*. A complete copy of a 36 bp repeat, R4, first observed in pHel4 and pHel5 (39), also was located in pAL202. In pAL202 a complete single copy (bp 1994-2029) was present at the 3′end of the *orf1*/*rep*A gene and another copy (bp 3834-3869) with 83% match at the 3' end of *orf7*, between *orf2* and *orf6*. The locations of the repeats in pHel4 corresponded to the locations in pAL202. The locations of the repeated sequences in pAL202 are shown in Figure 9.

ORF1, encoded by nucleotides 421-1947, showed strong identity to the replication protein, RepA, of *H. pylori* plasmids pHel4 (82%), pHel1 (82%), pHPM180 (79%), pHel5 (74%), pHPM8 (72%), pHPM185 (69%), and pHPM179 (65%). There was also some identity to the replication ORFs of plasmids from unrelated species such as *Enterococcus faecalis* (34%), and *Lactobacillus acidophilus* (33%). Identity was detected to the RepB proteins of plasmids from *Campylobacter jejuni* (33%), *Pediococcus acidilactici* (29%), and *Bifidobacterium longum* (28%). During the BLAST search the NCBI Conserved Domain Search indicated a conserved domain for ORF1 which is indicative of an initiator of plasmid replication, RepA, and RepB, which possesses nick-closing (topoisomerase I) like activity and the ability to perform a strand transfer reaction on ssDNA (1, 20, 60, 77). ORF1 was further analyzed using several software programs for protein analysis available on the Internet through the ExPASy Proteomics tools (http://us.expasy.org/). PSORT (http://www.psort.org/) and SOSUI (http://www.sbc.su.se/~erikw/toppred2/) analysis indicated that ORF1 is a soluble cytoplasmic protein. With an instability index of 30.70, ProtParam

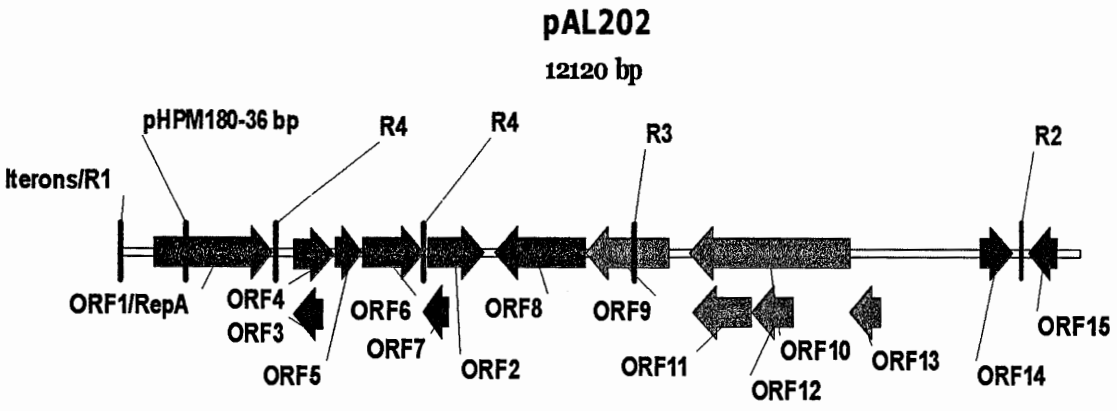


Figure 9. The locations of *H. pylori* plasmid repeated sequences in pAL202. R1, 22 bp tandem repeat called iterons; R2, 71% identity to the 232 bp repeat of pHPM180; 68% identity to the 36 bp separating the 232 bp repeat of pHPM180; R3, 64% identity to a 33 bp repeat of pHPS1; R4, first described in pHel4, a complete copy was present at the 3′end of the *orf1*/*rep*A gene, and a copy with 83% identity between *of2* and *orf6*.

(http://us.expasy.org/tools/protparam.html) classified this ORF as stable. ProtParam predicts that proteins with instability indices above 40 might be unstable. The Vector NTI program analyzed the ORFs of pAL202 and the results are shown in Table 5.

According to the NCBI Conserved Domain Search (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml), ORF2, nucleotides 3879-4583, possessed a central conserved motif HPFXXGNG present in most members of the family of Fic (filamentation induced by cAMP) proteins (46, 83). ORF2 of pAL202 showed great identity to other *H. pylori* plasmid ORF2 proteins: pHPM180 (98%); pHel4 (96%); pMCU2 (97%); pHP51 (87%); pHPS1 (74%); pHPM186 (75%); pHPM8 (72%); and pHPM185 (72%). Thirty percent identity was observed to *H. pylori* chromosomal open reading frame jhp0651 (function unknown) of J99. PSORT and SOSUI designated ORF2 as a soluble cytoplasmic protein and ProtParam recorded an instability index of 30.70 indicating that ORF2 is stable.

ORF3, nucleotides 2219-2587, showed 59% identity to the open reading frame ORF4J of pHel4. This is the only ORF to which ORF3 showed identity. According to PSORT, ORF3 is an inner membrane protein with at least 3 transmembrane domains. TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/) assigned three transmembrane domains to the ORF, and TopPred (http://www.sbc.su.se/~erikw/toppred2/) assigned four. SOSUI, however, described the protein as soluble. ProtParam calculated an instability index of 26.10, and therefore classified ORF3 as stable. The function of this protein is unknown.

Table 5. Analysis of pAL202 ORFs by the Vector NTI program

Name	Length (#aa)	MW[1] (m.w.)	1Microgram (pMoles)	IP[2]	Charge at pH7
ORF1	509	59934	16.685	9.22	15.99
ORF2	235	27489	36.378	6.34	-2.45
ORF3	123	13634	73.344	8.94	3.87
ORF4	166	20020	49.951	9.12	5.97
ORF5	110	13353	74.888	9.33	5.04
ORF6	257	29728	33.638	5.89	-5.42
ORF7	105	12305	81.268	8.28	0.94
ORF8	387	43844	22.808	9.02	9.20
ORF9	385	40982	24.401	6.27	-1.89
ORF10	660	78414	12.753	9.69	30.85
ORF11	233	27945	35.784	5.08	-9.00
ORF12	176	20478	48.834	8.83	3.04
ORF13	133	15270	65.488	10.00	9.86
ORF14	141	16538	60.467	5.72	-1.26
ORF15	121	13867	72.111	9.07	5.79

1-Molecular Weight; 2-Isoelectric Point.

As with ORF3, ORF4 (nucleotides 2223-2720) had identity only to pHel4 open reading frame, ORF4K (59%). PSORT and SOSUI classified the protein as cytoplasmic and soluble. ProtParam computed an instability index of 61.69 for ORF4 and qualifies the protein as unstable.

The segment of pAL202 consisting of ORF3, ORF4, and ORF5 has only been observed in one other *H. pylori* plasmid, pHel4. ORF5, nucleotides 2723-3052, showed 65% identity to ORF4L from pHel4. PSORT described this ORF as an inner membrane protein; SOSUI, TMHMM, and TopPred all assigned one transmembrane domain to it. ProtParam computed an instability index of 57.03 and therefore classified the protein as unstable.

As with ORF2, the conserved domain of the Fic protein family (46, 83) was observed also for ORF6 (nucleotides 3059-3829). This protein had identity to ORF4M of pHel4 (89%), ORF5K of pHel5 (88%), ORF6 of pHMP8 (87%), and to ORF2 of the pMCU plasmids (26-27%), number one, two, and five. ORF6 also exhibited identity to *H. pylori's* chromosomal proteins HP0712 (47%) and HP0713 (42%) of 26695, and JHP0651 of J99 (38%). ORF6 is a cytoplasmic and soluble protein according to PSORT and SOSUI, and according to ProtParam is stable (instability index is 35.38).

ORF7, nucleotides 3822-4136, showed 88% identity to ORF4N of pHel4, and low percentage identity to proteins of unrelated species such as *Staphylococcus* phage *phiN315* (34%), and *Staphylococcus aureus* (32%). PSORT described ORF7 as an inner membrane protein. SOSUI and TopPred assigned one transmembrane domain to the protein. However, two such domains were identified by TMHMM. ProtParam

calculated an instability index of 24.65 and classified the protein as stable. The function of this protein is unknown.

ORF8, nucleotides 4715-5875, showed strong identity to the MccC-like proteins, ORF4A, of pHel4 (85%) and ORF3 of pHPM8 (82%). MccC is a microcin exporter protein and is encoded in the *mccABCD* operon that is part of the *mccABCDEF* region of the *E. coli* plasmid, pMccC7 (32, 34). The *mccABCDEF* region is involved in the synthesis, modification, secretion, and immunity of microcin C7, MccC7, a peptide antibiotic that inhibits protein synthesis (32, 34). ORF8 also showed 45% identity to both the tetracycline resistance protein, tetA(P), of *H. pylori* 26695, and the tetracycline resistance protein homolog of J99. AL202 exhibited no growth in media supplemented with 2 μg/ml tetracycline and thus could not be qualified as tetracycline resistant. Low identity was shown also to the multi-efflux transporter of *Borrelia burgdorferi* (27%), tetA(P) of *Clostridium septicum* (24%), and other transmembrane transport proteins of non-related species. Both TMHMM and SOSUI assigned ten transmembrane regions to ORF8; TopPred assigned only nine. PSORT indicated that the protein is located in the inner membrane. ProtParam calculated an instability index of 34.38 and classified the protein as stable.

The NCBI Conserved Domain Search detected a conserved domain in the central region of ORF9, bp 5890-6963, which is shared by different protein families. The first one is the ThiF protein family that contains a repeated domain in ubiquitin-activating enzyme E1 and members of the bacterial ThiF/MoeB/HesA family. The second family is the ThiF dinucleotide-utilizing enzymes involved in the metabolism of the coenzymes

molybdopterin and thiamine. ORF9 showed identity to the HesA/MoeB/ThiF family protein of *Streptococcus pneumonia* (24%) and *Enterococcus faecalis* (43%), ubiquitin-activating enzyme E1 in *Schizosaccharomyces pombe* (41%) and in *Homo sapiens* (27%). HesA is essential for nitrogen fixation in *Anabaena* species, MoeB is involved in the biosynthesis of the organic component of molybdenum-containing prosthetic factors, and ThiF is involved in the synthesis of a precursor of thiamine (32). ORF9 shared identity with the MccB-like ORF4B of pHel4 (92%), ORF4 of pHPM8 (87%), and MccB of *E. coli* (32%). Like MccC, MccB is encoded by the *mccABCD* operon on pMccC7 and acts as a peptide modifier (34). The Gly-X-Gly_X_X-Gly motif, a nucleotide-binding domain, present in MccB (34) was detected in ORF4 of pHPM8 (71), but not in the MccB-like ORF4B of pHel4, or in ORF9 of pAL202. However, the $C\text{-}X_2\text{-}C\text{-}X_n\text{-}C\text{-}X_2\text{-}C$, a metal-binding domain found in MccB, ThiF, MoeB, and HesA proteins was present in ORF9. PSORT classified the protein as an inner membrane protein, however, SOSUI found it to be soluble. Both TMHMM and TopPred detected transmembrane regions, although in different numbers. TMHMM detected one transmembrane domain, but TopPred detected two. ProtParam assigned an instability index of 36.62 for ORF9 and determined that the protein is stable.

NCBI Conserved Domain Search identified a relaxase (14), relaxase/mobilisation nuclease domain (12, 58) in the first half of ORF10, nucleotides 7218-9197. The relaxase/mobilisation proteins are required for horizontal gene transfer during bacterial conjugation. Relaxases are involved in nicking duplex DNA. The conserved domain among the DNA-relaxases, H-X-D-X-D-N/H-X-H-H-a-a-a-N, described by Pansegrau

and Landka (65) was observed in ORF10. The protein showed the greatest identity (80%) to the MobA-like ORF4C of pHel4 and lower identity to mobilization proteins of plasmids from *Pasteurella multocida* (35%), *Aeromonas salmonicida* (34%), and the mobilization protein, MbeA, of *E. coli* plasmid ColE1 (32%). The *mobABC* genes encode proteins necessary for mobilization (48, 58). MobA and MobC are involved in cleavage and separation of double-stranded DNA. MobB has a stimulatory role in the process. A *nic* sequence (origin of transfer) is usually present upstream of *mobA*. In the case of pAL202, no *nic* sequence was observed upstream of *orf10*. A potential *nic* sequence (TATCCTG; bp 3379-3385) conserved among the IncP plasmids such as RP4 (39, 48, 65) was observed in *orf6* and could be a site for transfer initiation. PSORT and SOSUI indicated that ORF10 is cytoplasmic and soluble. ProtParam calculated an instability index of 41.09 and classified ORF10 as unstable.

ORF11, nucleotides 7256-7954, showed 79% identity to the MobD-like ORF4D of pHel4 and 56% to ORF3 of pHPM179. No identity to other *H. pylori* DNA was further observed. Four leucine zipper motifs, L-X(6)-L-X(6)-L-X(6)-L, were detected in ORF11 by the motif finder program 3DinSight (http://www.rtc.riken.go.jp/jouhou/HOMOLOGY/dbsearch/pdb/pdb_seq.html). Only two such motifs were detected in ORF4D. Hofreuter and Haas (39) reported that this motif also is present in the MobD proteins of the colicinogenic plasmids. The NCBI Conserved Domain Search revealed a conserved domain present in serine/threonine protein kinase (signal transduction mechanism). ORF11 is cytoplasmic, soluble, and stable (instability index 39.89).

ORF12, nucleotides 7967-8494, showed 75% identity to the MobB-like ORF4E of pHel4, 29% to ORF18 (*cag* island protein) of J99 and Cag11 of 26695, 28% to Cag-U of *H. pylori* strain CCUG 17874. TMHMM, SOSUI, and TopPred all identified this protein as a membrane protein containing one transmembrane domain. The ProtParam Tool program calculated an instability index of 46.01 for the protein and classified it as unstable.

The only identity of ORF13 to other *H. pylori* proteins was 84% identity to the MobA-like ORF4C of pHel4, and 56% identity to ORF2 of pHPM179. Forty-four percent identity was found to MobC of pSW200, 42% to MobC of plasmid ColK of *E. coli*, and 40% to MbeC of *Enterobacter cloacae*. NCBI Conserved Domain Search program also observed the MobC, bacterial mobilization protein domain within the last third of ORF13. ORF13 consists of 133 amino acids, but the conserved domain aligned only with amino acid residues 84-129. The protein is cytoplasmic and soluble according to PSORT and SOSUI. ProtParam showed a score of 29.47 and classified the protein as stable.

ORF14, nucleotides 10854-11276, showed no identity to known *H. pylori* proteins. Low identity (34%) was present to unrelated species such as *Anopheles gambiae*. This protein is cytoplasmic and soluble according to PSORT and SOSUI, and stable according to ProtParam (instability index, 23.84).

The conserved domain observed for ORF15 (nucleotides 11482-11844) is a putative transposase DNA-binding domain. According to the NCBI Conserved Domain Search, this domain is found at the C-terminus of transposase proteins and contains four

conserved cysteines that might be part of a zinc-binding domain. ORF15 showed 91% identity to OrfB of *H. pylori* IS*607* and 42% to IS*Hp608*. The identity to these insertion sequences was only observed at the right end (C-terminus) of OrfB. The alignment between ORF15 and OrfB of IS*607* was as follows: 103 amino acid residues (316-419) of OrfB (total are 419 amino acids), and 103 amino acid residues (12-115) of ORF15 (total are 121 amino acids) showed 97% identity. Identity was further observed to other transposases of unrelated species such as *Clostridium perfringens*. ORF15 was classified as cytoplasmic, soluble, and stable with a instability index score of 33.74. Table 6 shows the identities between the ORFs of pAL202 and other known ORFs.

Mobilization of pAL202

In the first step of mating, the average numbers of CFU obtained for the transconjugants were 5.8×10^7 for *pir*$^+$(pAL202::TN/KAN21; pRK24) and 9.1×10^8 for *pir*$^+$(pAL202::TN/KAN25; pRK24). This indicated a successful transfer of pRK24 to the *pir*$^+$ cells. The second and third steps resulted only in the transfer of pRK24 to the IR3272 as the average CFUs obtained from M9-Tet plates were 3.8×10^3 for *pir*$^+$ (pAL202::TN/KAN21; pRK24) x IR3272 and 1.9×10^4 for *pir*$^+$ (pAL202::TN/KAN25; pRK24) x IR3272. No transconjugants were recovered on M9-Kan plates indicating that the pAL202 recombinant plasmid was not mobilized.

Test for Distribution of pAL202 Genes in other Alaskan Plasmids

PCR primers (Table 2) derived from the DNA sequence of pAL202 were used in a comparison study to test for the distribution of DNA sequences present in pAL202 in six other *H. pylori* Alaskan plasmids, pAL203, pAL207, pAL208, pAL209, pAL226, and

Table 6. Putative ORFs of pAL202 and identities to known sequences

pAL202's ORF	Main identities to plasmid or chromosomal DNA	%Identities	Proposed function
1	pHel4 (RepA)	82	Plasmid replication
	pHel1 (RepA)	82	
	pHPM180 (RepA)	79	
	pHel5 (RepA)	74	
	pHPS1 (RepA)	74	
	pHPM8 (RepA)	72	
	pMCU3 (RepA)	94	
	pMCU2 (RepA)	93	
	pMCU4 (RepA)	93	
	pHPM186 (RepA)	75	
	pMCU5 (RepA)	71	
	pMCU7 (RepA)	71	
	pMCU1 (RepA)	71	
	pHPM185(RepA)	69	
	pHPM179(RepA)	65	
	pMCU6 (RepA)	65	
2	pHPM180 (ORF2)	98	cAMP induced filamentation
	pHel4 (ORF4O)	96	
	pMCU2 (ORF2)	97	
	pHP51(ORF2)	87	
	pHPO100 (ORF2)	87	
	pHPM8 (ORF2)	89	
	pMCU1 (ORF2)	98	
	pMCU5 (ORF2)	99	
	pHPS1 (ORF2)	81	
	pHPM185(ORF2)	72	
	J99 (JHP0651)	30	
3	pHel4 (ORF4J)	59	Unknown
4	pHel4 (ORF4K)	59	Unknown
5	pHel4 (ORF4L)	65	Unknown
6	pHel4 (ORF4M)	89	cAMP induced filamentation
	pHel5 (ORF5K)	88	
	pHPM8 (ORF6)	87	
	J99 (JHP0651)	38	
	26695 (HP0712)	47	
	pHPM180 (ORF2)	26	
	pMCU1 (ORF2)	26	
	pMCU2 (ORF2)	27	
	pHP51(ORF2)	27	
	pHel4 (ORF4O)	26	
	pMCU5 (ORF2)	26	
	pHPO100 (ORF2)	25	
	26695 (HP0713)	42	
	pHPM8 (ORF2)	26	
	pMCU4 (ORF2)	28	

Table 6 continues:

pAL202's ORF	Main identities to plasmid or chromosomal DNA	%Identities	Proposed function
7	pHel4 (ORF4N)	88	
8	pHel4 (ORF4A/MccC-like)	85	Transporter
	pHPM8 (ORF3)	82	
	26695 (HP1165/tetA(P) homolog)	45	
	J99 (JHP1092/tetA(P) homolog)	45	
	Borrelia burgdorferi	27	
	Clostridium septicum (TetA(P))	24	
9	pHel4 (ORF4B/MccB-like)	92	
	pHPM8 (ORF4)	87	
	E. coli (MccB)	27	
10	pHel4 (ORF4C/MobA-like)	80	Mobilization/Relaxases
	Pasteurella multocida (MbeAy)	35	
	Aeromonas salmonicida	34	
	E. coli (MobA)	32	
11	pHel4 (ORF4D/MobD-like)	79	Mobilization
	pHPM179 (ORF3)	56	
12	pHel4 (ORF4E/MobB-like)	75	Mobilization
	J99 (ORF18/*cag* island protein)	29	
	26695 (Cag11)	29	
13	pHel4 (ORF4F/MobC-like)	84	Mobilization
	pHPM179 (ORF2)	68	
14	*Anopheles gambiae*	34	Unknown
15	IS*607* (ORFB)	91	Transposition
	Clostridium perfringens	57	
	IS*Hp608* (ORFB)	42	

pAL236. This study was performed to determine whether the genes on the Alaskan plasmids are geographically conserved, i.e., were they present on more than one strain. The results of the PCR studies are shown in Table 7a. Results from the hybridization studies performed to confirm the PCR findings are shown Table 7b. The results of the PCR studies and the hybridization studies did not agree. The conflicting results indicated that plasmid DNA isolated for the use of the PCR studies might have been contaminated with chromosomal DNA. Genes that amplified with the PCR studies and did not hybridize indicated that those amplified genes might have been present on the small amount of chromosome present in the plasmid preparations. The results of the second PCR studies using exonuclease-treated DNA are tabulated in Table 7c. Table 8 shows the collective conclusions of the studies.

The results shown in Table 8 indicated that none of the tested regions were present in pAL209 or in any of the pAL236 plasmids. Region IV, the Mob-region, was the only segment present in pAL226. None of the tested regions were present in the two pAL208 plasmids; however, all the regions were present when total plasmid DNA of pAL208 was tested. Since AL208 contained at least two plasmids, these results indicated that region I-IV were not present in the two pAL208 plasmids tested, but probably were present in an additional pAL208 plasmid that was not isolated. All the regions were detected on pAL207, and region I-III on pAL203.

Transformation of AL202 and HU71 with Single Insertion Mutants

Repeated attempts to transform AL202 and HU71 cells resulted in no transformants using a range of 1-10 μg of DNA. Attempts to transform cells using three

Table 7. Comparative studies of Alaskan plasmids

a. PCR

Plasmid	Conserved regions I (*orf4-orf5*)	II (*orf6*)	III (*orf4-ofr6*)	IV (*orf10-orf13*)
pAL202	+	+	+	+
pAL203	+	-	+	+
pAL207	+	+	+	+
pAL209	+	-	-	+
pAL226	+	-	+	+
pAL208::TN/KAN8	+	-	+	+
pAL208::TN/KAN9	+	-	+	-
pAL236::TN/KAN2	+	-	+	-
pAL236::TN/KAN5	+	-	-	-
pAL236::TN/KAN11	+	-	+	-

b. Hybridization

Plasmid	Conserved regions I (*orf4-orf5*)	II (*orf6*)	III (*orf4-ofr6*)	IV (*orf10-orf13*)
pAL202	+	+	+	+
pAL203	-	+	+	-
pAL207	+	+	+	+
pAL209	-	-	-	-
pAL226	-	-	-	+
pAL208::TN/KAN8	-	-	-	-
pAL208::TN/KAN9	-	-	-	-
pAL236::TN/KAN2	-	-	-	-
pAL236::TN/KAN5	-	-	-	-
pAL236::TN/KAN11	-	-	-	-

c. PCR (E vs N)

Plasmid	Conserved regions I (*orf4-orf5*)		II (*orf6*)		III (*orf4-ofr6*)		IV (*orf10-orf13*)	
	E[1]	N[2]	E	N	E	N	E	N
pAL202	+	+	+	+	+	+	+	+
pAL203	-	+	+	+	-	-	-	-
pAL207	+	+	+	+	-	-	+	+
pAL209	-	+	-	-	-	-	-	+
pAL226	-	-	-	-	-	-	+	+
pAL208	+	+	+	+	+	+	+	+
pAL208::TN/KAN8	-	-	-	-	-	-	-	-
pAL208::TN/KAN9	-	-	-	-	-	-	-	-
pAL236	-	+	-	-	-	-	-	-
pAL236::TN/KAN2	-	-	-	-	-	-	-	-
pAL236::TN/KAN5	-	+	-	-	-	-	-	-
pAL236::TN/KAN11	-	-	-	-	-	-	-	-
pZErO::TN/KAN	-	-	-	-	-	-	-	-

1, Exonuclease treated DNA; 2, No exonuclease treatment

Tabe 8. Conserved regions in the Alaskan plasmids

Plasmid	Conserved regions			
	I (*orf4-orf5*)	II (*orf6*)	III (*orf4-ofr6*)	IV (*orf10-orf13*)
pAL202	+	+	+	+
pAL203	+	+	+	-
pAL207	+	+	+	+
pAL209	-	-	-	-
pAL226	-	-	-	+
pAL208	+	+	+	+
pAL208::TN/KAN8	-	-	-	-
pAL208::TN/KAN9	-	-	-	-
pAL236	-	-	-	-
pAL236::TN/KAN2	-	-	-	-
pAL236::TN/KAN5	-	-	-	-
pAL236::TN/KAN11	-	-	-	-

different electroporation procedures were also unsuccessful. After electroporation, a survival rate of only 27% was observed for the AL202 cells, and zero for the HU71 cells even at lower voltage. No transformants per 10 μg of DNA were recovered for either strain.

CHAPTER IV

DISCUSSION

DNA and ORF Analysis

The genomes of *H. pylori* strains 26695 and J99 show high conservation with respect to gene content, gene function, and gene order (23). The genomes have an average GC content of 39%. However, both strains exhibit plasticity regions with different percent GC: eight regions in 26695, and nine regions in J99 (3). In 26695 these regions include the *cag* PAI (35% GC), IS*605* (33% GC), restriction modification systems (33% GC), and genes encoding RNA polymerase subunits (43% GC) (52). Except for the region encoding the RNA polymerase subunits, the percent GC of the other plasticity regions correspond to the percent GC of *H. pylori* plasmids (33%-38%) whose sequences have been determined. The above observation led to the suggestion that plasmids might be involved in the integration of new DNA into *H. pylori* chromosome and in the transfer of DNA between strains (3). Even though several *H. pylori* plasmids have been sequenced and characterized (22, 37, 39, 40, 45, 59, 70, 75, 76), no clearly defined function has been assigned to these structures. The main objective of this study was to characterize another *H. pylori* plasmid, pAL202, and contribute to understanding the biological role of plasmids in the bacterium.

The DNA and ORF sequences of pAL202 were almost identical to pHel4. Since these plasmids originated from two different geographical regions, this identity between the plasmids was surprising since all the plasmids that have been sequenced share some

genes, the *repA* and *orf2* genes for example, but still exhibit their own unique gene content. Due to the near identity of the gene sequences, gene functions, and gene arrangements between pAL202 and pHel4, the argument can be made that these two plasmids derived from the same *H. pylori* plasmid some time ago.

The R1 iterons observed upstream of the *orf1/repA* gene in pAL202 are presumably the binding site for the Rep protein and the initiation site of theta-type replication (22, 59). Veereshlingam (84) confirmed these suggestions showing that in pHPM8 the RepA protein binds to the iterons. The iteron-RepA complexes, however, are strain specific (84). From the strong conserved middle region among the RepA proteins of these theta-type plasmids (41, 70) and the conserved R1 iteron sequences (70) one would not expect the strain-specific binding of the iteron-RepA complexes. It could be that the binding of RepA to the iterons is not strain specific, but other proteins involved in the formation of these complexes are strain specific and recognize the amino acids sequences on RepA that are not common to the other RepA proteins. No studies of RepB binding of *H. pylori* plasmids have been reported. However, Hosaka *et al.* (40) suggested the binding of the RepB protein to the R3 iterons based on the observation that theta-type plasmids are dependent on the binding of Rep proteins to iteron sequences.

Minnis *et al.* (59) and De Ungria *et al.* (22) indicated that R2 might be involved in recombination events. Hofreuter and Haas (39) suggested that these sequences might be hot spots for recombination and site-specific integration. As observed here in pAL202, Hofreuter and Haas (39) report that in pHel4, pHel5, pHPM8, pHPM180, pHPM186, pHPS1, and pHel1 R4 is located also at the 3' end of the *repA* gene. In pHPM8,

pHPM180, and pHPS1, in which *orf2* is positioned immediately downstream of *repA*, the R4 is located between the two genes. However, in plasmids in which *repA* and *orf2* are separated by other genes (pHPM186, pHel4) two copies of R4 are positioned as described here for pAL202. Amplification with primers derived from R4 resulted in 1.7 and 0.8 kbp fragments in pHel4 and four randomly chosen plasmids indicate that duplication of the R4 is common in *H. pylori* plasmids (39). The region between the two R4 sequences in pAL202 was about 1.8 kbp. Based on the above observations, Hofreuter and Haas (39) suggest that R2 and R4 sequences are sites for genetic recombination that might have led to the size variability in *H. pylori* plasmids, and that some cryptic proteins encoded by the plasmids and the chromosome might be involved in specific recombination events. The sequence data for pAL202 supported this suggestion.

The complete sequence of pAL202 had 15 putative ORFs that were submitted to the BLAST program for comparison to known sequences. ORF1/RepA showed strong identity to the replication protein, RepA, of *H. pylori* theta-type plasmids and some identity to replication ORFs of plasmids from unrelated species. ORF2 showed great identity to other *H. pylori* plasmid ORF2 proteins, but also to *H. pylori* chromosomal open reading frame JHP0651 (function unknown) of J99. The central conserved motif HPFXXGNG present in ORF2 is found in most members of the family of Fic proteins that are involved in cell division via the synthesis of folate (46, 83), but the precise molecular function has not been determined. Since it is absent from some *H. pylori* plasmids (37, 76), the suggestion has been made that ORF2 might be involved in recombination (22, 59) or with the replication and/or control of the copy number of

plasmids (70). From the observation that R4 is linked to some *orf2* genes, it is possible that this ORF also might be involved in recombination events between plasmid and chromosome.

The conserved motif HPFXXGNG present in ORF2 and the family of Fic proteins (46, 83) was observed also for ORF6. Hofreuter and Haas (39) observed an extension of the above conserved sequence, PFSDGNGRTGRALMF, in ORF4M of pHel4, ORF5K of pHel5. A similar conserved sequence was seen in ORF6, but not in ORF2 of pAL202. As with ORF2 in pAL202, ORF6 showed identity to chromosomal protein JHP0651, but ORF6 also showed identity to chromosomal proteins HP0712 and HP0713. In pAL202, ORF6 is positioned directly upstream of ORF2 with the genes overlapping by eight bp; the first eight bp in ORF2, and the last eight bp in ORF6. In pHel4, these ORFs are positioned adjacent to each other; however, there is no overlapping of the genes encoding the ORFs. As with ORF2, the function of ORF6 is uncertain.

Open reading frames ORF3, ORF4, and ORF5 showed identity to other ORFs of pHel4. A BLAST search revealed no identity to any other *H. pylori* plasmid or chromosomal proteins, or to proteins of other unrelated species. A comparative study between pAL202 and six other *H. pylori* Alaskan plasmids revealed that the genes encoding these ORFs were also present in three of the six plasmids (see below). In pAL202 the region consisting of these three ORFs and ORF6 was part of the 1.8 kbp region between the R4 repeats. The sequences of these additional Alaskan plasmids has not been determined and since the forward PCR primer-binding site that amplified this region was within *orf3* and the reverse primer-binding site was within *orf6*, it is unknown

whether the R4 repeats were present on these Alaskan plasmids and if this region was positioned between *repA* and *orf2*. The function of ORF3, ORF4, and ORF5 is also unknown.

Two other ORFs of unknown function of pAL202 were ORF7 and ORF14. ORF4N of pHel4 was the only *H. pylori* protein that showed identity to any (ORF7) of these two proteins. Low identity to proteins of unrelated species such as *Staphylococcus* phage *phiN315* and *Anopheles gambiae* were found. Even though a putative promoter and RBS was found for both of the corresponding genes, it is uncertain if these genes are expressed. Also, the observation that *H. pylori* strains 26695 and J99 possess genes that are strain specific indicate that *orf7* could have originated in the chromosome. The *orf14* gene might have been acquired from an unrelated organism.

Presumably, ORF8 and ORF9 could be classified as Microcin-like (Mcc-like) proteins. ORF8 showed strong identity to the MccC-like proteins of pHel4 and pHPM8. In addition ORF8 showed low identity to transmembrane transport proteins of non-related species. Identity to the tetracycline resistance protein tetA(P) homolog of *H. pylori* 26695 and J99 led to the investigation of tetracycline resistance of AL202. AL202 was not tetracycline resistant. The above observation led us to suggest that mutation in the gene encoding ORF8 might have caused loss of tetracycline resistance. The conclusion was that ORF8 is a transmembrane protein of unknown function. Alternatively, this region of pAL202 may represent a recombination event in which acquisition of ORF8 was the result of a plasmid/chromosome recombination.

ORF9 shared identity with the MccB-like proteins and HesA/MoeB/ThiF family proteins. As mentioned in the Results, ORF9 showed identity to numerous proteins of unrelated species such as *Streptococcus pneumonia*, *Enterococcus faecalis*, *Schizosaccharomyces pombe,* and *Homo sapiens*. Such identities made any assumption about the function of ORF9 difficult, since the functions of these proteins differ in each of the organisms. However, both ORF8 and ORF9 are Mcc-like and could therefore be part of the *mccABCDEF* operon acquired from *E. coli* through horizontal gene transfer. Further investigation of these proteins might determine their exact function and their relationship to other proteins encoded by the plasmid.

ORF10, ORF11, ORF12, and ORF13 made up the Mob-region (58) and showed strong identity to the Mob-like proteins described in pHel4 (39). Lower identity to mobilization proteins of plasmids from unrelated species also was observed for these proteins. The presence of these Mob-like genes, ORF10-ORF13, and a putative *nic* sequence (origin of transfer) on pAL202 led to the investigation of the mobilization of the plasmid. No *nic* sequence was present upstream of the *mobA-like orf10* gene where it was observed in *E. coli*; however, a potential *nic* sequence (TATCCTG; bp 3379-3385) conserved among the IncP plasmids, such as RP4 (39, 48, 65), was observed in *orf6* and could be the site for DNA transfer initiation. Repeated trials using both liquid and plate matings failed to show any mobilization of the pAL202 recombinants used in this study. This led to the conclusion that the putative *nic* sequence found in *orf6* was not the strand transfer initiation site and that pAL202 could not be mobilized by pRK24. Since all the *H. pylori* strains in our laboratory are isolated from patients with some form of gastric

disease these strains are classified as type I strains and therefore assumed to have a *cag* PAI. Hacker and Kaper (35) pointer out that if PAIs contribute to the fitness and pathogenicity of the host organism, then the mobility genes involved in transfer, deletion, or excision will be inactivated or deleted to select for stability. We therefore suggest that mutation in the genes, *orf10-orf13*, encoding these proteins might have caused the loss of their original function.

ORF15 showed identity to the transposase ORF, OrfB, of *H. pylori* IS*607* and IS*Hp608*, and lower percent identity to other transposase proteins of unrelated species. The identity between ORF15 and OrfB of *H. pylori* IS*607* and IS*Hp608* was only observed at the right end (C-terminus) of the transposase. Since neither promoter nor RBS sequences were found for ORF15, this ORF might have been part of a whole insertion sequence and was left on the plasmid after a transposition event. This observation supported the suggestion made by Alm *et al.* (3) that plasmids might be involved in the integration of new DNA into *H. pylori* chromosome and the transfer of DNA between strains.

Test for Distribution of pAL202 Genes in Other Alaskan Plasmids

The five characterized *H. pylori* plasmids from this laboratory were isolated from strains recovered from patients in the North Texas area. All plasmids showed strong conservation of replication genes. Therefore, a study was undertaken to determine whether *H. pylori* plasmids from a different geographical region would also exhibit these replication genes. Johnson (41) tested seven Alaskan strains and determined that the *repA* and *orf2* genes were strongly conserved. Since these Alaskan strains came from a

geographically isolated region, this study examined whether other genes on the Alaskan plasmids were conserved.

As shown in Table 8 the presence of these regions in five of the seven *H. pylori* strains indicated that the genes are conserved among the plasmids in this particular geographical region (Alaskan village). However, the fact that pAL202, an Alaskan strain, was so similar to pHel4, a European strain, and pHPM8, a North Texas strain, indicated that these genes are often conserved among *H. pylori* plasmids but frequent recombination events might lead to the dissimilarity of gene order. Although not shown directly, the inconsistency in data for pAL203 and pAL207 in Table 7 indicated that these strains contained multiple plasmids as did pAL208 and pAL236. This was the first confirmed observation of multiple plasmids in one *H. pylori* strain.

Transformation of AL202 and HU71 with Single Insertion Mutants

An attempt was made to transform AL202 and HU71 naturally and by electroporation. Most *H. pylori* strains are naturally competent for transformation (61, 85) and DNA uptake by natural transformation occurs by means of a type IV secretion system encode by the *comB* operon (38). Tsuda *et al.* (82) report that a higher frequency of transformation of *H. pylori* is obtained through electroporation. Repeated attempts to transform AL202 and HU71 cells naturally and by electroporation resulted in zero transformants for both strains.

Based on analysis of chromosomal DNA, *H. pylori* strains possess common and strain specific type I, type II, and type III restriction-modification (R-M) systems (3, 23, 81). The restriction component degrades foreign DNA (81) and therefore might prevent

transformation with DNA of unrelated strains. The presence of such restriction systems in AL202 and HU71 might explain the failure to transform these strains with recombinants of pAL202. The recombinant plasmid DNA was isolated from *E. coli pir*$^{+}$ cells that are modification proficient and thus the recombinant plasmid DNA should have been modified. When introduced into AL202 and HU71 cells, recombinant pAL202 plasmids could be recognized as foreign and therefore degraded. Alternatively, these strains may be examples of *H. pylori* strains that are not transformable. No control DNA was available to test this hypothesis.

Finally, the observation of border sequences of transposable elements, repeated sequences, chromosomal DNA, and *mob-like* genes on pAL202 indicate recombination events. Characterization of six *H. pylori* plasmids in this laboratory showed that all carry chromosomal and/or transposon remnants, indicating that recombination events occur between chromosome and plasmid. It is suggested that frequent recombination events in the genome might be crucial for the bacterium to evade the host immune system (10). Whether these recombination events between chromosome and plasmid are involved in the evasion of the host immune system is unknown. However, the conservation of *H. pylori* plasmids in clinical isolates argues that the plasmids do increase the fitness of the bacterium and therefore may contribute to the pathogenicity of the organism.

REFERENCES

1. Abeles A. L. 1986. P1 plasmid replication. Purification and DNA-binding activity of the replication protein RepA. J. Biol. Chem. 261: 3548-3555.

2. Akopyants N. S., S. W. Clifton, D. Kersulyte, J. E. Crabtree, B. E. Youree, C.A. Reece, N. O. Bukanov, E. S. Drazek, B. A. Roe, and D. E. Berg. 1998. Analyses of the *cag* pathogenicity island of *Helicobacter pylori*. Mol. Microbiol. 28: 37-53.

3. Alm, R. A., L-S. L. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. de Jonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature 397: 176-180.

4. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

5. Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acid Res. 25: 3389-3402.

6. Amieva, M. R., R. Vogelmann, A. Covacci, L. S. Tompkins, W. J. Nelson, and S. Falkow. 2003. Disruption of the epithelial apical-junctional complex by *Helicobacter pylori* CagA. Science 300: 1430-1434.

7. Bajwa, G. 2003. Effect on *Escherichia coli* of the microcin-like operon from *Helicobacter pylori* plasmid pHPM8. Thesis. Texas Woman's University, Denton, TX.

8. Blaser, M. J. 1997. Not all *Helicobacter pylori* strains are created equal: should all be eliminated? Lancet 349: 1020-1022.

9. Blaser, M. J. 1998. *Helicobacter pylori* and gastric diseases. Br. Med. J. 316: 1507-1510.

10. Blaser, M. J., and J. C. Atherton. 2004. *Helicobacter pylori* persistence: biology and disease. J. Clin. Invest. 113: 321-333.

11. Borén, T., P. Falk, K. A. Roth, G. Larson, and S. Normark. 1993. Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. Science 262: 1892-1895.

12. Bravo-Angel, A. M., V. Gloeckler, B. Hohn, and B. Tinland. 1999. Bacterial conjugation protein MobA mediates integration of complex DNA structures into plant cells. J. Bacteriol. 181: 5758-5765.

13. Burnham, K. D. 1998. Plasmid contains unique *Helicobacter pylori* insertion sequence, IS*605*. Dissertation. Texas Woman's University, Denton, TX.

14. Byrd, D. R., and S. W. Matson. 1997. Nicking by transesterification: the reaction catalyzed by a relaxase. Mol. Microbiol. 25: 1011-1022.

15. Censini, S., C. Lange, Z. Xiang, J. E. Crabtree, P. Ghiara, M. Borodovsky, R. Rappuoli, and A. Covacci. 1996. *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. Proc. Natl. Acad. Sci. USA 93: 14648-14653.

16. Clyne, M., and B. Drumm. 1993. Adherence of *Helicobacter pylori* to primary human gastrointestinal cells. Infect. Immun. 61: 4051-4057.

17. Covacci A., S. Censini, M. Bugnoli, R. Petracca, D. Burroni, G. Macchia, A. Massone, E. Papini, Z. Xiang, N. Figura, and R. Rappuoli. 1993. Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. Proc. Natl. Acad. Sci. USA 90: 5791-5795.

18. Covacci, A., S. Falkow, D. E. Berg, and R. Rappuoli. 1997. Did the inheritance of a pathogenicity island modify the virulence of *Helicobacter pylori*? Trends Microbiol. 5: 205-208.

19. Cover, T. L., and M. J. Blaser. 1992. Purification and characterization of the vacuolating toxin from *Helicobacter pylori*. J. Biol. Chem. 267: 10570-10575.

20. Del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas. 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. 62: 434-464.

21. De Ungria, M. C., D. Tillett, B. A. Neilan, P. T. Cox, and A. Lee. 1998. A novel method of extracting plasmid DNA from *Helicobacter* species. Helicobacter 3: 269-277.

22. De Ungria, M. C. A., T. Kolesnikow, P. T. Cox, and A. Lee. 1999. Molecular characterization and interstrain variability of pHPS1, a plasmid isolated from the Sydney strain (SS1) of *Helicobacter pylori.* Plasmid 41: 97-109.

23. Doig, P., B. L. de Jong, R. A. Alm, E. D. Brown, M. Uria-Nickelsen, B. Noonan, S. D. Mills, P. Tummino, G. Carmel, B. C. Guild, D. T. Moir, G. F. Vovis, and T. J. Trust. 1999. *Helicobacter pylori* physiology predicted from genomic comparison of two strains. Microbiol. Mol. Biol. Rev. 63: 675-707.

24. Dorrell, N., M. C. Martino, R. A. Stabler, S. J. Ward, Z. W. Zhang, A. A. McColm, M. J. Farthing, and B. W. Wren. 1999. Characterization of *Helicobacter pylori* PldA, a phospholipase with a role in colonization of the gastric mucosa. Gastroent. 117: 1098-1104.

25. Dunn, B. E., G. P. Campbell, G. I. Perez-Perez, and M. J. Blaser. 1990. Purification and characterization of urease from *Helicobacter pylori.* J. Biol. Chem. 265: 9464-9469.

26. Dunn, B. E., H. Cohen, and M. J. Blaser. 1997. *Helicobacter pylori.* Clin. Microbiol. Rev. 10: 720-721.

27. Eaton K. A., S. Suerbaum, C. Josenhans, and S. Krakowka. 1996. Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. Infect. Immun. 64: 2445-2448.

28. Edwards, R. A., R. A. Helm, and S. R. Maloy. 1999. Increasing DNA transfer efficiency by temporary inactivation of host restriction. BioTechniques 26: 892-900.

29. Forsyth, M. H., and T. L. Cover. 1999. Mutational analysis of the *vacA* promoter provides insight into gene transcription in *Helicobacter pylori.* J. Bacteriol. 181: 2261-2266.

30. Gao, X. 1999. Characterization of pHPM185 plasmid from *Helicobacter pylori.* Thesis. Texas Woman's University, Denton, TX.

31. Garner, R. M., J. Fulkerson, and H. L.T. Mobley. 1998. *Helicobacter pylori* glutamine synthetase lacks features associated with transcriptional and posttranslational regulation. Infect. Immun. 66: 1839-1847.

32. González-Pastor, J. E., J. L. San Millán, M. Á. Castilla, and F. Moreno. 1995. Structure and organization of plasmid genes required to produce the translation inhibitor Microcin C7. J. Bacteriol. 177: 7131-7140.

33. Groisman, E. A., and H. Ochman. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell 87: 791-794.

34. Guijarro, J. I., J. E. González-Pastor, F. Baleux, J. L. San Millán, M. A. Castilla, M. Rico, F. Moreno, and M. Delepierre. 1995. Chemical structure and translation inhibition studies of the antibiotic Microcin C7. J. Biol. Chem. 270: 23520-23532.

35. Hacker, J., and J. B. Kaper. 2000. Pathogenicity islands and the evolution of microbes. Ann. Rev. Microbiol. 54: 641-679.

36. Hazell, S. L., D. J. Evans Jr., and D. Y. Graham. 1991. *Helicobacter pylori* catalase. J. Gen. Microbiol. 137: 57-61.

37. Heuermann, D., and R. Haas. 1995. Genetic organization of a small cryptic plasmid of *Helicobacter pylori*. Gene 165: 17-24.

38. Hofreuter, D., S. Odenbreit, and R. Haas. 2001. Natural transformation competence in *Helicobacter pylori* is mediated by basic components of a type IV secretion system. Mol. Microbiol. 41: 379-391.

39. Hofreuter, D., and R. Haas. 2002. Characterization of two cryptic *Helicobacter pylori* plasmids: a putative source for horizontal gene transfer and gene shuffling. J. Bacteriol. 184: 2755-2766.

40. Hosaka, Y., R. Okamoto, K. Irinoda, S. Kaieda, W. Koizumi, K. Saigenji, and M. Inoue. 2002. Characterization of pKU701, a 2.5-kb plasmid, in a Japanese *Helicobacter pylori* isolate. Plasmid 47: 193-200.

41. Johnson, S. R. 2001. Characterization of *repA* and *orf2* genes from plasmids isolated from *Helicobacter pylori* Alaskan strains. Thesis. Texas Woman's University, Denton, TX.

42. Kersulyte, D., N. S. Akopyants, S. W. Clifton, B. A. Roe, and D. E. Berg. 1998. Novel sequence organization and insertion specificity of IS*605* and IS*606*: chimaeric transposable elements of *Helicobacter pylori*. Gene 223: 175-186.

43. Kersulyte D., A. K. Mukhopadhyay, M. Shirai, T. Nakazawa, and D. E. Berg. 2000. Functional organization and insertion specificity of IS*607*, a chimeric element of *Helicobacter pylori*. J. Bacteriol. 182: 5300-5308.

44. Kersulyte, D., B. Velapatiño, G. Dailide, A. K. Mukhopadhyay, Y. Ito, L. Cahuayme, A. J. Parkinson, R. H. Gilman, and D. E. Berg. 2002. Transposable element IS*Hp608* of *Helicobacter pylori*: nonrandom geographic distribution, functional organization, and insertion specificity. J. Bacteriol. 184: 992-1002.

45. Kleanthous, H., C. L. Clayton, and S. Tabaqchali. 1991. Characterization of a plasmid from *Helicobacter pylori* encoding a replication protein common to plasmids in Gram-positive bacteria. Mol. Microbiol. 5:2377-2389.

46. Komano, T., R. Utsumi, and M. Kawamukai. 1991. Functional analysis of the *fic* gene involved in regulation of cell division. Res. Microbiol. 142: 269-277.

47. Kuck, D., B. Kolmerer, C. Iking-Konert, P. H. Krammer, W. Stemmel, and J. Rudi. 2001. Vacuolating cytotoxin of *Helicobacter pylori* induces apoptosis in human gastric epithelial cell line AGS. Infect. Immun. 69: 5080-5087.

48. Lanka, E., and B. M. Wilkins. 1995. DNA processing reactions in bacterial conjugation. Annu. Rev. Biochem. 64: 141-169.

49. Lee, A., J. Fox, and S. Hazell. 1993. Pathogenicity of *Helicobacter pylori*: a perspective. Infect. Immun. 61: 1601-1610.

50. Lee, W-K., Y-S. An, K-H. Kim, S-H. Kim, J-Y. Song, B-D. Ryu, Y-J. Choi, Y-H. Yoon, S-C. Baik, K-H. Rhee, and M-J. Cho. 1997. Construction of a *Helicobacter pylori-Escherichia coli* shuttle vector for gene transfer in *Helicobacter pylori*. Appl. Env. Microbiol. 63: 4866-4871.

51. Leunk, R. D., P. T. Johnson, B. C. David, W. G. Kraft, and D. R. Morgan. 1988. Cytotoxic activity in broth-culture filtrates of *Campylobacter pylori*. J. Med. Microbiol. 26: 93-99.

52. Marais, A., G. L. Mendz, S. L. Hazell, and F. Megraud. 1999. Metabolism and genetics of *Helicobacter pylori*: the genome era. Microbiol. Mol. Biol. Rev. 63: 642-674.

53. Marshall, B. 1983. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet i: 1273-1275.

54. Marshall, B. J., and J. R. Warren. 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet i: 1311-1315.

55. Marshall, B. J. 1994. *Helicobacter pylori*. Am. J. Gastroent. 89: S116-S128.

56. Merrell, D. S., M. L. Goodrich, G. Otto, L. S. Tompkins, and S. Falkow. 2003. pH-regulated gene expression of the gastric pathogen *Helicobacter pylori*. Infect. Immun. 71: 3529-3539.

57. Metcalf, W. W., W. Jiang, and B. L. Wanner. 1994. Use of the *rep* technique for allele replacement to construct new *Escherichia coli* host for maintenance of R6Kγ origin plasmids at different copy numbers. Gene 138: 1-7.

58. Meyer, R. 2000. Identification of the *mob* genes of plasmid pSC101 and characterization of a hybrid pSC101-R1162 system for conjugal mobilization. J. Bacteriol. 182: 4875-4881.

59. Minnis, J. A., T. E. Taylor, J. E. Knesek, W. L. Peterson, and S. A. McIntire. 1995. Characterization of a 3.5-kbp plasmid from *Helicobacter pylori*. Plasmid 34: 22-36.

60. Moscoso M., R. Eritja, and M. Espinosa. 1997. Initiation of replication of plasmid pMV158: mechanisms of DNA strand-transfer reactions mediated by the initiator RepB protein. J. Mol. Biol. 268: 840-856.

61. Nedenskov-Sørensen, P., G. Bukholm, and K. Bøvre. 1990. Natural competence for genetic transformation in *Campylobacter pylori*. J. Infect. Dis. 161: 365-366.

62. Nilsson, C., A. Sillén, L. Eriksson, M-L. Strand, H. Enroth, S. Normark, P. Falk, and L. Engstrand. 2003. Correlation between *cag* pathogenicity island composition and *Helicobacter pylori*-associated gastroduodenal disease. Infect. Immun. 71: 6573-6581.

63. Odenbreit, S., J. Püls, B. Sedlmaier, E. Gerland, W. Fisher, and R. Haas. 2000. Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. Science 287: 1497-1500.

64. Papini, E., M. de Bernard, E. Milia, M. Bugnoli, M. Zerial, R. Rappuoli, and C. Montecucco. 1994. Cellular vacuoles induced by *Helicobacter pylori* originate from late endosomal compartments. Proc. Natl. Acd. Sci. USA 91: 9720-9724.

65. Pansegrau, W., and E. Lanka. 1991. Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. Nucleic Acid Res. 19: 3455.

66. Penfold, S. S., A. J. Lastovica, and B. G. Elisha. 1988. Demonstration of plasmids in *Campylobacter pylori*. J. Infect. Dis. 157: 850-851.

67. Porwollik, S., B. Noonan, and P. W. O'Toole. 1999. Molecular characterization of a flagellar export locus of *Helicobacter pylori*. Infect. Immun. 67: 2060-2070.

68. Püls, J., W. Fisher, and R. Haas. 2002. Activation of *Helicobacter pylori* CagA by tyrosine phosphorylation is essential for dephosphorylation of host cell proteins in gastric epithelial cells. Mol. Microbiol. 43: 961-969.

69. Qasem, J. A. A. 1995. Cloning and sequencing analysis of plasmid DNA from *Helicobacter pylori*. Dissertation. Texas Woman's University, Denton, TX.

70. Quiñones M. 2000. Characterization of a 7.8 kbp plasmid from *Helicobacter pylori* strain HPM8. Thesis. Texas Woman's University, Denton, TX.

71. Quiñones M., J. E. Knesek, and S. A. McIntire. 2001. Sequence and gene expression analyses of plasmid pHPM8 from *Helicobacter pylori* reveal the presence of two operons with putative roles in plasmid replication and antibiotic activity. Plasmid 46: 223-228.

72. Sambrook J., and D. W. Russel. 2001. Molecular Cloning. A laboratory manual. Third Edition. Vol. 3. p. A2.3. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

73. Segal, E. D., and L. S. Tompkins. 1993. Transformation of *Helicobacter pylori* by electroporation. BioTechniques 14: 225-226.

74. Seyler, R. W. Jr., J. W. Olson, and R. J. Maier. 2001. Superoxide dismutase-deficient mutants of *Helicobacter pylori* are hypersensitive to oxidative stress and defective in host colonization. Infect. Immun. 69: 4034-4040.

75. Song, J-Y., S-H. Choi, E-Y. Byun, S-G. Lee, Y-H. Park, S-G. Park, S-Y. Lee, K-M. Kim, J-U. Park, H-L. Kang, S-C. Baik, W-K. Lee, M-J. Cho, H-S. Youn, G-H. Ko, D-W. Bae, and K-H. Rhee. 2003. Characterization of a small cryptic plasmid, pHP51, from a Korean isolate of strain 51 of *Helicobacter pylori*. Plasmid 50: 145-151.

76. Song, J-Y., S-G. Park, H-L. Kang, W-K. Lee, M-J. Cho, J-U. Park, S-C. Baik, H-S. Youn, G-H. Ko, and K-H. Rhee. 2003. pHP489, a *Helicobacter pylori* small cryptic plasmid, harbors a novel gene coding for a replication initiation protein. Plasmid 50: 236-241.

77. Spiers, A. J., N. Bhana, and P. L. Bergquist. 1993. Regulatory interactions between RepA, an essential replication protein, and the DNA repeats of RepFIB from plasmid P307. J. Bacteriol. 175: 4016-4024.

78. Stein, M., R. Rappuoli, and A. Covacci. 2000. Tyrosine phosphorylation of the *Helicobacter pylori* CagA antigen after *cag*-driven host cell translocation. Proc. Natl. Acad. Sci. USA 97: 1263-1268.

79. Taylor, T. E. 1997. Unusual structure of a *Helicobacter pylori* plasmid. Thesis. Texas Woman's University, Denton, TX.

80. Thomson, A. B. R. 1996. *Helicobacter pylori*: from infection to cure. Can. J. Gastroent. 10: 167-172.

81. Tomb, J-F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H.G. Khalak, A. Glodek, K. McKenney, L. M. Fitzegerald, N. Lee, M.D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T.R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature 388: 539-547.

82. Tsuda, M., M. Karita, and T. Nakazawa. 1992. Genetic transformation in *Helicobacter pylori*. Microbiol. Immunol. 37: 85-89.

83. Utsumi, R., Y. Nakamoto, M. Kawamukai, M. Himeno, and T. Kamano. 1982. Involvement of cyclic AMP and its receptor protein in filamentation of an *Escherichia coli fic* mutant. J. Bacteriol. 151: 807-812.

84. Veereshlingam, L. 2003. Interaction of pHPM encoded RepA protein from *Helicobacter pylori* with pHPM iteron sequence. Thesis. Texas Woman's University, Denton, TX.

85. Wang, Y., K. P. Roos, and D. E. Taylor. 1993. Transformation of *Helicobacter pylori* by chromosomal metronidazole resistance and by a plasmid with a selectable chloramphenicol resistance marker. J. Gen. Microbiol. 139: 2485-2493.

86. Warren J. R. 1983. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet i: 1273-1275.

87. Xiang Z., S. Censini, P. F. Bayeli, J. L. Telford, N. Figura, R. Rappuoli, and A. Covacci. 1995. Analysis of expression of CagA and VacA virulence factors in 43 strains of *Helicobacter pylori* reveals that clinical isolates can be divided into two major types and that CagA is not necessary for the expression of the vacuolating cytotoxin. Infect. Immun. 63: 94-98.