

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 March 2002 (07.03.2002)

PCT

(10) International Publication Number
WO 02/18595 A2

(51) International Patent Classification⁷: **C12N 15/31**,
C07K 14/21, C12N 15/62, 7/01, 1/21, A61K 48/00, C07K
19/00, 16/12, A61K 39/02, 39/40, C12Q 1/68, G01N
33/569

(71) Applicant (for all designated States except US): **AVEN-
TIS PASTEUR LIMITED** [CA/CA]; Connaught Cam-
pus, 1755 Steeles Avenue, West, Toronto, Ontario M2R
3T4 (CA).

(21) International Application Number: PCT/CA01/01221

(72) Inventors; and

(22) International Filing Date: 28 August 2001 (28.08.2001)

(75) Inventors/Applicants (for US only): **LOOSMORE,
Sheena** [CA/CA]; 70 Crawford Rose Drive, Aurora,
Ontario L4G 4R4 (CA). **WANG, Joe** [CA/CA]; 48 29th
Street, Etobicoke, Ontario M8W 3A8 (CA). **BRADLEY,
Bill** [CA/CA]; Apt#508, 91 Wellesley St. E., Toronto,
Ontario M4A 1H9 (CA). **OCHS, Martina** [DE/CA];
30-100 Vaughan Road, Toronto, Ontario M6C 2M1 (CA).
YANG, Yan-Ping [CA/CA]; Apt 709, 120 Torresdale
Avenue, Toronto, Ontario M2R 3N7 (CA).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/228,295	28 August 2000 (28.08.2000)	US
60/228,294	28 August 2000 (28.08.2000)	US
60/228,296	28 August 2000 (28.08.2000)	US
60/228,441	29 August 2000 (29.08.2000)	US
60/228,439	29 August 2000 (29.08.2000)	US
60/228,442	29 August 2000 (29.08.2000)	US
60/228,438	29 August 2000 (29.08.2000)	US
60/228,443	29 August 2000 (29.08.2000)	US
60/228,773	29 August 2000 (29.08.2000)	US
60/228,742	29 August 2000 (29.08.2000)	US
60/228,440	29 August 2000 (29.08.2000)	US
60/228,512	29 August 2000 (29.08.2000)	US
60/228,511	29 August 2000 (29.08.2000)	US
60/229,474	1 September 2000 (01.09.2000)	US
60/229,465	1 September 2000 (01.09.2000)	US
60/229,475	1 September 2000 (01.09.2000)	US
60/229,478	1 September 2000 (01.09.2000)	US
60/229,803	5 September 2000 (05.09.2000)	US
60/229,805	5 September 2000 (05.09.2000)	US
60/229,804	5 September 2000 (05.09.2000)	US
60/229,806	5 September 2000 (05.09.2000)	US
60/229,740	5 September 2000 (05.09.2000)	US
60/229,811	5 September 2000 (05.09.2000)	US
60/229,809	5 September 2000 (05.09.2000)	US
60/230,250	6 September 2000 (06.09.2000)	US
60/230,214	6 September 2000 (06.09.2000)	US
60/230,252	6 September 2000 (06.09.2000)	US

(74) Agents: **NGUYEN, Thuy, H.** et al.; Smart & Biggar,
900-55 Metcalfe Street, P.O. Box 2999, Station D, Ottawa,
Ontario K1P 5Y6 (CA).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,
ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished
upon receipt of that report
- with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

[Continued on next page]

(54) Title: MORAXELLA POLYPEPTIDES AND CORRESPONDING DNA FRAGMENTS AND USES THEREOF

(57) Abstract: The present invention provides nucleic acids, proteins and vectors for immunization of a host, including humans, against disease caused by infection by a strain of Moraxella, specifically *M. catarrhalis*. The invention includes DNA vaccination methods which employ a vector containing a nucleotide sequence encoding a protein of a strain of Moraxella, specifically *M. catarrhalis*, and a promoter to effect expression of the protein gene product in the host. The nucleotide sequences include those which encode: a major anaerobically induced OMP, rare lipoprotein A proteins, component proteins of the adhesin complex, a potential lipoprotein OMP, a putative oxidoreductase, OprM, collagenase, outer membrane phospholipase A, lipoproteins, OmpA, NADH:ubiquinone oxidoreductase subunit, pilQ fimbriae, invasins protein, a 36 kDa secreted protein, OMP, Outer membrane esterase, pyridine nucleotide transhydrogenase, HecB, a putative virulence factor, heme receptor, ompC, heme receptor, and porin. Modifications are possible within the scope of this invention.



WO 02/18595 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

TITLE OF INVENTION

MORAXELLA POLYPEPTIDES AND CORRESPONDING DNA
FRAGMENTS AND USES THEREOF

REFERENCE TO RELATED APPLICATIONS

5 This application claims the benefit of the
following U.S. Provisional Application Nos: 60/228,294;
60/228,295 and 60/228,296 filed August 28, 2000; 60/228,441;
60/228,439; 60/228,442; 60/228,438; 60/228,443; 60/228,773;
60/228,742; 60/228,440; 60/228,512 and 60/228,511 filed
10 August 29, 2000; 60/229,474; 60/229,465; 60/229,475 and
60/229,478 filed September 1, 2000; 60/229,803; 60/229,805;
60/229,804; 60/229,806; 60/229,740; 60/229,811 and
60/229,809 filed September 5, 2000; 60/230,250; 60/230,214
and 60/230,252 filed September 6, 2000; the contents of
15 which are herein incorporated by reference.

FIELD OF INVENTION

The present invention relates to Moraxella
polypeptides and corresponding DNA molecules, which can be
used to diagnose, prevent and treat Moraxella infection in
20 mammals, such as humans.

BACKGROUND OF THE INVENTION

Moraxella are gram negative bacteria that can be
isolated from the mucous membranes of humans and other warm-
blooded animals. While many Moraxella species are
25 nonpathogenic, several species are associated with a number
of severe diseases. Moraxella *catarrhalis* (also known as
Branhamella catarrhalis) is frequently found in the
respiratory tract of humans and can be isolated from sputa
as part of the normal flora in humans.

M. catarrhalis is of special interest because it is now recognized as an important pathogen in upper and lower respiratory tract infections. While infection occurs most frequently in infants, all ages are at risk. *M. catarrhalis* is a common human pathogen found in the flora of 40-50% of normal school children.

Clinical features of adult patients suspected of *M. catarrhalis* lower respiratory tract infection have been reported increasingly during the last twenty years. Results obtained from good quality sputa in various studies show that from 0.2 to 8.1% of lower respiratory tract specimens have yielded growth of *M. catarrhalis* in adults (Christensen, Jens Jorgen (1999) APMIS 88 :1-36). In approximately one third of the specimens concomitant presence of *H. influenzae* or *S. catarrhalis* was registered (Hagar et al. (1987) Rev. Infect. Dis. 9 :1140 - 1149). A number of studies have also indicated an association between *M. catarrhalis* infection and acute exacerbation of chronic bronchitis. In addition, infection has been linked to chronic obstructive pulmonary disease with or without asthma, emphysema and bronchiectasis. Further still, smaller studies have shown links to recurrent aspiration, coronary artery disease, congestive heart failure, pulmonary- and extra-pulmonary cancer, collagen diseases, leukemia (acute lymphocytic, acute myelogenous, chronic myelogenous, acute myelomonocytic), and diabetes mellitus.

The causal relationship between *M. catarrhalis* and otitis media is of considerable importance. *M. catarrhalis* is the third-highest cause of bacterial otitis media after *Streptococcus catarrhalis* and *Haemophilus influenzae*; and nearly all children experience at least one episode of otitis media during the first year of life. In fact 10% of all children from the United States have at least one attack

by the time they are three months old. Further, approximately 60% have at least one attack in their first year and more than 80% by the age of three (Klein, J.O. 1994. Clin. Infect. Dis. 19 :823-833).

5 Otitis media is the most common illness of early childhood with approximately 70% of all children suffering at least one bout of otitis media before the age of seven. Chronic otitis media can lead to hearing, speech and cognitive impairment in children. It is caused by bacterial
10 infection with *Streptococcus pneumoniae* (approximately 50%), non-typable *Haemophilus influenzae* (approximately 30%) and *Moraxella catarrhalis* (approximately 20%) (Murphy, T. F. 1996. Microbiol. Rev. 60 :267-279; Block, S. L. 1997. Pediatr. Infect. Dis. 16 :449-456). In the United States
15 alone, treatment of otitis media costs between one and two billion dollars per year for antibiotics and surgical procedures, such as tonsillectomies, adenoidectomies and insertion of tympanostomy tubes. Because otitis media occurs at a time in life when language skills are developing
20 at a rapid pace, development disabilities specifically related to learning and auditory perception are documented in youngsters with frequent otitis media.

M. catarrhalis colonizes the respiratory tract and is an important cause of otitis media and sinusitis in
25 infants and children as well as lower respiratory tract infections in adults, especially those with chronic obstructive pulmonary disease (Catlin, B.W. (1990) Clin. Microbiol. Rev. 3 :293-320). Studies using cultures of middle ear fluid obtained by tympanocentesis have shown that
30 it causes approximately 10-15% of cases of otitis media. More rarely, *M. catarrhalis* can cause invasive disease, such as meningitis, bacteremia, and endocarditis (Enright, M. C.

and McKenzie, H. 1997. J. Med. Microbiol. 46 :360-371).
There is also evidence of nosocomial spread of infection.

Many *M. catarrhalis* strains produce β -lactamase which mediates resistance to penicillins. Of increasing
5 concern is the high percentage of antibiotic resistant strains, especially penicillin resistance due to the expression of β -lactamase. In clinical isolates from the US and Europe, 90-95% express β -lactamase, an alarming finding since there was no resistance in 1976 (Doern, G.V. et. al.
10 1996. Antimicrob. Agents Chemother. 40 :2884-6). As new vaccines for pneumococcal disease become available (Mbelle, N. et. al. 1999. J. Infect. Dis. 180 :1171-6), the importance of *Moraxella catarrhalis* as a respiratory pathogen will increase and a vaccine becomes very desirable.

15 *M. catarrhalis* mainly colonizes the respiratory tract and is predominantly a mucosal pathogen. Studies using cultures of middle ear fluid obtained by tympanocentesis have shown that *M. catarrhalis* causes approximately 10-15% of cases of otitis media.

20 The incidence of otitis media caused by *M. catarrhalis* is increasing. As ways of preventing otitis media caused by pneumococcus and non-typable *H. influenzae* are developed, the relative importance of *M. catarrhalis* as a cause of otitis media can be expected to further increase.

25 *M. catarrhalis* is also an important cause of lower respiratory tract infections in adults, particularly in the setting of chronic bronchitis and emphysema (Murphy, T. F. 1996. Microbiol. Rev. 60 :267-279; Enright, M. C. and McKenzie, H. 1997. J. Med. Microbiol. 46 :360-371). *M.*
30 *catarrhalis* also causes sinusitis in children and adults (Tinkelman D.G. et al. 1989. Am J Dis Child 143: 938-942)

and occasionally causes invasive disease (Enright, M. C. and McKenzie, H. 1997. J. Med. Microbiol. 46 :360-371; Meyer, et. al. 1995. Clin. Pediatr. 34 :146-150).

Like other Gram-negative bacteria, the outer
5 membrane of *M. catarrhalis* consists of phospholipids, lipopolysaccharide (LPS), and outer membrane proteins (OMPs). Eight of the *M. catarrhalis* OMPs have been identified as major components. These are designated by letters A to H, beginning with OMP A which has a molecular
10 mass of 98 kDa to OMP H which has a molecular mass of 21 kDa (Murphy, T. F. 1990. Am. J. Med. 88 :415-455).

Recently, Klingman and Murphy purified and characterized a high molecular-weight outer membrane protein of *M. catarrhalis* (Klingman, K. L. and Murphy T. F. 1994.
15 Infect. Immun. 62 :1150-5). The apparent molecular mass of this protein varies from 350 kDa to 720 kDa as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). This protein appears to be an oligomer of much smaller proteins or subunits thereof, of molecular mass
20 about 120 to 140 kDa, and is antigenically conserved among strains of *Moraxella*.

Helminen et al also identified a protein of molecular mass of about 300 to 400 kDa, named UspA, that was reported to be present on the surface of *Moraxella*
25 (Helminen, M. E. et. al. 1994. J. Infect. Dis. 170 :807-872).

In most adults tested, bactericidal antibodies have been identified (Chapman, A. J. et. al. 1985. J. Infect. Dis. 151 :878-882). Strains of *M catarrhalis*
30 present variations in their capacity to resist serum bactericidal activity. In general, isolates from diseased individuals are more resistant than those who are simply

colonized (Hol, C. et. al. 1995. FEMS Immunol. Med. Microbiol. 11 :207-211). Serum resistance could therefore be considered a virulence factor of the bacteria.

The search for suitable vaccine antigens has
5 focused on the surface-exposed outer membrane proteins (OMP) of the bacterium that are relatively well conserved between strains (Murphy, T. F. 1996. Microbiol. Rev. 60 :267-279). A number of *M. catarrhalis* proteins OMPs have been studied in detail, and these proteins have variously been shown to
10 be targets of serum-derived human antibodies (Sethi. et. al. 1995. Infec. Immun. 63 :1516-20 ; Faden, H. et. al. 1992. Infec. Immun. 60 :3824-9), targets of biologically active antibodies in bactericidal assays and animal models. The response of these proteins in children with otitis media has
15 not yet been clearly delineated and information is limited.

In light of these results a protective vaccine against *M. catarrhalis* infection would be of considerable importance. There is not yet an effective vaccine for any human Moraxella infection. It is conceivable that an
20 effective vaccine can be developed using physically or chemically inactivated Moraxella. However, such a vaccine does not have a high margin of safety. In general, safer vaccines are made by genetically manipulating the organism by attenuation or by recombinant means. Accordingly, a
25 major obstacle in creating an effective and safe vaccine against human Moraxella infection has been the paucity of genetic information regarding Moraxella, specifically *M. catarrhalis*.

Accordingly, a need exists for identifying and
30 isolating polynucleotide sequences of *M. catarrhalis* for use in preventing and treating Moraxella infection.

SUMMARY OF THE INVENTION

The present invention provides purified and isolated polynucleotide molecules that encode Moraxella polypeptides. In one embodiment, the amino acid sequences
5 of the polypeptides are set forth in any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54.

The polypeptides may be used in methods to prevent, treat, and diagnose Moraxella infection. In one
10 embodiment, the polynucleotide molecules are DNA whose sequences are set forth in any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53.

The invention also provides polynucleotides
15 encoding fragments derived from such a polypeptide, as welll as homologs, mutants and derivatives of such polypeptides and fragments derived therefrom, which result from the addition, deletion, or substitution of non-essential amino acids as described herein. The invention also provides
20 monospecific antibodies that specifically bind to such polypeptides.

The present invention has wide application and includes expression cassettes, vectors, and cells transformed or transfected with the polynucleotides of the
25 invention. Encompassed within the scope of the invention is:

(i) a method for producing a polypeptide of the invention in a recombinant host system and related expression cassettes, vectors, and transformed or
30 transfected cells;

(ii) a vaccine containing a polypeptide or a polynucleotide of the invention, such vaccines and vaccine vectors being useful for, e.g., preventing and treating Moraxella infection, in combination with a diluent or
5 carrier, and related pharmaceutical compositions and associated therapeutic and/or prophylactic methods;

(iii) a therapeutic and/or prophylactic use of an RNA or DNA molecule of the invention, either in a naked form or formulated with a delivery vehicle, a polypeptide or
10 combination of polypeptides, or a monospecific antibody of the invention, and related pharmaceutical compositions;

(iv) a method for diagnosing the presence of Moraxella in a biological sample, which can involve the use of a DNA or RNA molecule, a monospecific antibody, or a
15 polypeptide of the invention; and (v) a method for purifying a polypeptide of the invention by antibody-based affinity chromatography.

One aspect of the invention provides an isolated nucleic acid molecule comprising a nucleic acid sequence
20 which encodes a polypeptide selected from any one of:

(a) SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54;

(b) at least 12 consecutive amino acids from a
25 polypeptide of (a); and

(c) a polypeptide which is at least 75% identical in amino acid sequence to the corresponding polypeptide of (a) or (b).

In one embodiment, the fragment comprising at
30 least 12 consecutive amino acids from the polypeptide of

(a), or the polypeptide which is at least 75% identical in amino acid sequence to the corresponding polypeptide of (a) or (b), is capable of eliciting an immunogenic response in a mammal against a bacterium of the Moraxella genus.

5 In another embodiment, the homolog/variant which is at least 75% identical to the reference polypeptide retains immunogenicity compared to the corresponding reference polypeptide.

10 The nucleic acid molecule described above may be operably linked to one or more expression control sequences in order to effect production of the polypeptide encoded by the isolated nucleic acid.

15 In another aspect of the invention, the nucleic acid molecule comprises a nucleic acid sequence which is complementary to the nucleic acid molecule described above.

20 The invention also provides nucleic acids which encode a fusion protein which comprises a polypeptide as described above, and a second polypeptide. The second polypeptide may be a heterologous signal peptide, or may have adjuvant activity.

25 Another aspect of the invention provides host organisms comprising an isolated nucleic acid as described above. The isolated nucleic acid is capable of being expressed in the host and there is optionally a second isolated nucleic acid encoding and capable of expressing an additional polypeptide.

30 The host organism may be a eukaryotic cell, a prokaryotic cell, or a virus. Virus hosts include adenovirus, alphavirus, and poxvirus, e.g. a vaccinia or canary pox virus. The host organism may also be a bacterium

such as *E. coli*, *Shigella*, *Salmonella*, *Vibrio cholerae*, *Lactobacillus*, Bacille bilié de Calmette-Guérin (BCG), or *Streptococcus*. Non-toxicogenic *Vibrio cholerae* mutant strains and attenuated *Salmonella typhimurium* strains are
5 useful. The host organisms may be used to produce the polypeptides and fusion proteins of the invention. In one embodiment, the virus hosts and bacterial hosts may be used advantageously as vaccine vectors.

Another aspect of the invention provides
10 oligonucleotides of 5 to 100 nucleotides, which may be used as nucleic acid probes, or oligonucleotides of 10 to 40 nucleotides, which maybe used as primers, which hybridize under stringent conditions to the nucleic acid molecules described above, or to a complementary or anti-sense
15 sequence of those nucleic acid molecules.

Another aspect of the invention provides polypeptides encoded by the nucleic acid sequences described above, as well as fusion proteins comprising the polypeptide and a second polypeptide which may be a heterologous signal
20 peptide or may have adjuvant activity.

Another aspect of the invention provides an antibody against the polypeptides and fusion proteins described above.

Another aspect of the invention provides
25 pharmaceutical compositions comprising nucleic acids, polypeptides, fusion proteins and antibodies of the invention and a pharmaceutically acceptable carrier. The nucleic acid compositions may further include a delivery agent such as bupivacaine, liposome and cationic lipid, and
30 may also include an adjuvant. The polypeptide and fusion protein compositions may further include a delivery agent

such as liposomes, and may also include an adjuvant as well as an additional *Moraxella* polypeptide.

Another aspect of the invention provides a method of detecting *Moraxella* or diagnosing *Moraxella* infection. The method comprises contacting a body fluid of a mammal to be tested, with a component selected from: the nucleic acid, probe, primer, polypeptide, fusion protein or antibody described above. Choice of which component should be used depends on the specific circumstances. For example, if the fluid tested is suspected of containing *Moraxella* antigens, then the antibodies should be used. When very low amounts of *Moraxella* is suspected to be present, then primers should be used to amplify the target *Moraxella* sequences. These choices can be readily made by a person skilled in the art.

Another aspect of the invention provides a method for generating an immune response against *Moraxella*, including generating antibodies specific for *Moraxella*. The method comprises introducing into an animal a composition comprising the nucleic acid, polypeptide, or fusion protein as described above to elicit an immune response. This method may be used advantageously to confer immunoprotection to the animal, or to treat an animal infected with *Moraxella*.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1 shows the nucleotide sequence of the major anaerobically induced outer membrane protein (OMP) gene designated MCA100170 and the deduced amino acid

sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 1 and 2).

Figure 2 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100170.

5 Figure 3 shows the nucleotide sequence of a rare lipoprotein A gene designated MCA100178 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 3 and 4). The protein is encoded on the negative strand of the *Moraxella* genome.

10 Figure 4 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100178.

Figure 5 shows the nucleotide sequence of the gene encoding a component protein of the adhesin complex designated MCA100412 and the deduced amino acid sequence
15 from *Moraxella catarrhalis* (any one of SEQ ID NOs: 5 and 6).

Figure 6 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100412.

Figure 7 shows the nucleotide sequence of the gene encoding another component protein of the adhesin complex
20 designated MCA100413 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 7 and 8).

Figure 8 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100413.

Figure 9 shows the nucleotide sequence of a
25 potential lipoprotein OMP gene designated MCA100414 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 9 and 10).

Figure 10 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100414.

Figure 11 shows the nucleotide sequence of a putative oxidoreductase, internalin B gene designated MCA100527 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 11 and 12).

5 Figure 12 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100527.

Figure 13 shows the nucleotide sequence of the OprM gene designated MCA100564 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs:
10 13 and 14).

Figure 14 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100564.

Figure 15 shows the nucleotide sequence of the collagenase gene designated MCA100567 and the deduced amino
15 acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 15 and 16).

Figure 16 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100567.

Figure 17 shows the nucleotide sequence of the
20 outer membrane phospholipase A gene designated MCA100712 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 17 and 18).

Figure 18 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100712.

25 Figure 19 shows the nucleotide sequence of the lipoprotein gene designated MCA100800 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 19 and 20).

Figure 20 shows the restriction enzyme analysis of the *M. catarrhalis* designated MCA100800.

Figure 21 shows the nucleotide sequence of the OmpA gene designated MCA100854 and the deduced amino acid
5 sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 21 and 22).

Figure 22 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100854.

Figure 23 shows the nucleotide sequence of the
10 NADH:ubiquinone oxidoreductase subunit A gene designated MCA100889 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 23 and 24).

Figure 24 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA100889.

15 Figure 25 shows the nucleotide sequence of the pilQ fimbriae gene designated MCA100992 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 25 and 26).

Figure 26 shows the restriction enzyme analysis of
20 the *M. catarrhalis* gene designated MCA100992.

Figure 27 shows the nucleotide sequence of the invasin protein gene designated MCA101249 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 27 and 28).

25 Figure 28 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101249.

Figure 29 shows the nucleotide sequence of another rare lipoprotein A gene designated MCA101388 and the deduced

amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOS: 29 and 30).

Figure 30 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101388.

5 Figure 31 shows the nucleotide sequence of the gene encoding a 36kDa secreted protein designated MCA101446 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOS: 31 and 32).

10 Figure 32 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101446.

Figure 33 shows the nucleotide sequence and deduced amino acid sequence of a gene designated MCA101580 from *Moraxella catarrhalis* (any one of SEQ ID NOS: 33 and 34).

15 Figure 34 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101580, shown as SEQ ID NO: 33.

20 Figure 35 shows the nucleotide sequence of a rare lipoprotein A gene designated MCA101666 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOS: 35 and 36).

Figure 36 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101666, shown as SEQ ID NO: 35.

25 Figure 37 shows the nucleotide sequence of the OMP gene designated MCA101792 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOS: 37 and 38).

Figure 38 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101792.

Figure 39 shows the nucleotide sequence of the outer membrane esterase gene designated MCA101806 and the
5 deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 39 and 40). The gene is encoded on the negative strand of the *Moraxella* genome.

Figure 40 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101806.

10 Figure 41 shows the nucleotide sequence of the soluble pyridine nucleotide trans-hydrogenase gene designated MCA101862 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 41 and 42). The gene is encoded on the negative strand of the
15 *Moraxella* genome.

Figure 42 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101862.

Figure 43 shows the nucleotide sequence of the lipoprotein gene designated MCA101896 and the deduced amino
20 acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 43 and 44).

Figure 44 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101896.

Figure 45 shows the nucleotide sequence of the
25 HecB gene designated MCA101915 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 45 and 46).

Figure 46 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101915.

Figure 47 shows the nucleotide sequence of the gene encoding a putative virulence factor MviN homolog designated MCA101924 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 47 and 48).

Figure 48 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101924.

Figure 49 shows the nucleotide sequence of the heme receptor gene designated MCA101976 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 49 and 50).

Figure 50 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA101976.

Figure 51 shows the nucleotide sequence of the ompC/heme receptor gene designated MCA102062 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 51 and 52).

Figure 52 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA102062.

Figure 53 shows the nucleotide sequence of the porin gene designated MCA102070 and the deduced amino acid sequence from *Moraxella catarrhalis* (any one of SEQ ID NOs: 53 and 54).

Figure 54 shows the restriction enzyme analysis of the *M. catarrhalis* gene designated MCA102070.

Figure 55 shows construction of expression plasmids for expression of the *Moraxella* ORFs.

Figure 56 shows a purification scheme for the *Moraxella* proteins under native conditions.

Figure 57 shows a purification scheme for the *Moraxella* proteins under denaturing conditions.

Figure 58(A) shows an SDS-PAGE analysis of purified protein MCA100170 (SEQ ID NO:2; 51.2kDa; lane 2, 2ug) and Western blot analysis using horse radish peroxidase-conjugated anti-His tag antibody (lane 4, 2ug purified protein). The molecular weight standards are shown in lanes 1 and 3.

Figure 58(B) shows an SDS-PAGE analysis of purified protein MCA100414 (SEQ ID NO:10; 21.2kDa; lane 2, 2ug) and Western blot analysis using horse radish peroxidase-conjugated anti-His tag antibody (lane 4, 2ug purified protein). The molecular weight standards are shown in lanes 1 and 3.

Figures 58(C) to 58(R) show SDS-PAGE analyses of purified *Moraxella* proteins. Each lane labeled (1) indicates molecular weight standards. The lane labeled (2) indicate the purified protein (2ug), as follows:

- (C) 49.8kDa MCA100567 (SEQ ID NO:16)
- (D) 22.2kDa MCA100854 (SEQ ID NO:22)
- (E) 17.4kDa MCA101249 (SEQ ID NO:28)
- (F) 73.1kDa MCA101915 (SEQ ID NO:46)
- (G) 73.1kDa MCA101976 (SEQ ID NO:50)
- (H) 37.7kDa MCA100800 (SEQ ID NO:20)
- (I) 64.3kDa MCA101806 (SEQ ID NO:40)
- (J) 10.3kDa MCA100178 (SEQ ID NO:4)
- (K) 12.8kDa MCA100413 (SEQ ID NO:8)

(L) 88.9kDa MCA102062 (SEQ ID NO:52)

(M) 53.1kDa MCA100564 (SEQ ID NO:14)

(N) MCA101388 (SEQ ID NO:30)

(O) MCA101792 (SEQ ID NO:38)

5 (P) MCA101896 (SEQ ID NO:44)

(Q) MCA102070 (SEQ ID NO:54)

(R) MCA101446 (SEQ ID NO:32)

Figure 59 show Western blots using antibodies generated against the *Moraxella* proteins. Lanes 1, 4, 7, 10 10, 13, 16, 22, 26, 29 and 33 show mobility of the standard molecular weights (203, 118, 86, 52, 34, 29, 19 and 7.5 kDa). Lanes 2, 5, 8, 11, 14, 17, 20, 24, 27 and 31 show Western-blotted whole cell lysates from stationary phase *M. catarrhalis*. Lanes 3, 6, 9, 12, 15, 18, 21, 25, 28 and 32 15 show Western-blotted whole cell lysates from logarithmic phase *M. catarrhalis*. Lanes 19, 23 and 30 show Western-blotted purified proteins (1 ug).

Lanes 2 & 3: MCA100800 (SEQ ID NO:20)

Lanes 5 & 6: MCA101806 (SEQ ID NO:40)

20 Lanes 8 & 9: MCA100178 (SEQ ID NO:4)

Lanes 11 & 12: MCA100413 (SEQ ID NO:8)

Lanes 14 & 15: MCA102062 (SEQ ID NO:52)

Lanes 17 & 18: MCA100564 (SEQ ID NO:14)

Lanes 19-21: MCA101388 (SEQ ID NO:30)

25 Lanes 23-25: MCA101792 (SEQ ID NO:38)

Lanes 27 & 28: MCA101896 (SEQ ID NO:44)

Lanes 30-32: MCA102070 (SEQ ID NO:54)

DETAILED DESCRIPTION OF INVENTION

Open reading frames (ORFs) encoding *Moraxella* polypeptides been identified from the *M. catarrhalis* genome. Accordingly, these polypeptides and related polypeptides can be used to prevent and treat *Moraxella* infection.

According to one aspect of the invention, isolated polynucleotides are provided which encode *Moraxella* polypeptides whose amino acid sequences are exemplified in any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54. These polypeptides have been identified as having the following putative function *in vivo*:

	<u>SEQ ID NOS</u>	<u>Designation</u>	<u>Putative function</u>
15	1 & 2	MCA100170	major anaerobically induced OMP
	3 & 4	MCA100178*	rare lipoprotein A
	5 & 6	MCA100412	component protein of adhesin complex
	7 & 8	MCA100413	component protein of adhesin complex
20	9 & 10	MCA100414b	lipoprotein OMP
	11 & 12	MCA100527	oxidoreductase, internalin B
	13 & 14	MCA100564	OprM
	15 & 16	MCA100567b	collagenase
	17 & 18	MCA100712	outer membrane phospholipase A
25	19 & 19	MCA100800	lipoprotein
	21 & 22	MCA100854	OmpA/OprF homolog
	23 & 24	MCA100889	NADH:ubiquinone oxidoreductase subunit A
	25 & 26	MCA100992	pilQ fimbriae
30	27 & 28	MCA101249	invasin protein
	29 & 30	MCA101388	rare lipoprotein A

	31 & 32	MCA101446	36kDa secreted protein
	33 & 34	MCA101580	
	35 & 36	MCA101666	rare lipoprotein A
	37 & 38	MCA101792	OMP
5	39 & 40	MCA101806*	Outer membrane esterase
	41 & 42	MCA101862*	soluble pyridine nucleotide transhydrogenase
	43 & 44	MCA101896	lipoprotein
	45 & 46	MCA101915	HecB
10	47 & 48	MCA101924	virulence factor MviN homolog
	49 & 50	MCA101976	heme receptor
	51 & 52	MCA102062b	ompC / heme receptor
	53 & 54	MCA102070	porin

*The gene is encoded on the negative strand of the Moraxella
15 genome.

The term "isolated polynucleotide" is defined as a polynucleotide removed from the environment in which it naturally occurs. "Isolated polynucleotide" is a polynucleotide the structure of which is not identical to
20 that of any naturally occurring polynucleotide. The term therefore covers, for example, (a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecule but is not flanked by both of the coding sequences that flank that part of the molecule in the genome of the
25 organism in which it naturally occurs; (b) a polynucleotide incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a
30 genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e. a gene encoding a fusion protein. Specifically

excluded from this definition are polynucleotides present in mixtures of (i) DNA molecules, (ii) transfected cells, and (iii) cell clones, e.g. as these occur in a DNA library such as a cDNA or genomic DNA library.

5 For example, a naturally-occurring DNA molecule present in the genome of a living bacteria or as part of a gene bank is not isolated, but the same molecule separated from the remaining part of the bacterial genome, as a result of, e.g., a cloning event (amplification), is isolated.

10 Typically, an isolated DNA molecule is free from DNA regions (e.g., coding regions) with which it is immediately contiguous at the 5' or 3' end, in the naturally occurring genome. Such isolated polynucleotides may be part of a vector or a composition and still be defined as isolated in

15 that such a vector or composition is not part of the natural environment of such polynucleotide.

 The polynucleotide of the invention is either RNA or DNA (cDNA, genomic DNA, or synthetic DNA), or modifications, variants, homologs or fragments thereof. The

20 DNA is either double-stranded or single-stranded, and, if single-stranded, is either the coding strand or the non-coding (anti-sense) strand. any one of the sequences that encode the polypeptides of the invention as shown in SEQ ID NO: 1 to 54 is (a) a coding sequence, (b) a ribonucleotide

25 sequence derived from transcription of (a), or (c) a coding sequence which uses the redundancy or degeneracy of the genetic code to encode the same polypeptides.

 By "polypeptide" or "protein" is meant any chain of amino acids, regardless of length or post-translational

30 modification (e.g., glycosylation or phosphorylation). Both terms are used interchangeably in the present application.

Amino acid sequences are provided which are homologous to any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54. As used herein, "homologous amino acid sequence" is any polypeptide which is encoded, in whole or in part, by a nucleic acid sequence which hybridizes at 25-35°C below critical melting temperature (T_m), to any portion of the nucleic acid sequence of any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53. A homologous amino acid sequence is one that differs from an amino acid sequence shown in any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54 by one or more amino acid substitutions, preferably conservative substitutions. Such a sequence also encompasses serotypic variants (defined below) as well as sequences containing deletions or insertions which retain inherent characteristics of the polypeptide such as immunogenicity. Preferably, such a sequence is at least 75% (for example, 78%, 80%, 83%, 85%, 88%), and preferably 90% identical to any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54.

Homologous amino acid sequences include sequences that are identical or substantially identical to any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54. By "amino acid sequence substantially identical" is meant a sequence that is at least 90% (for example 92%, 94%, 95%, 96%, 97%, 98%), and preferably 99% identical to an amino acid sequence of reference. In one preferred embodiment, the homologous sequence differs from the sequence of

reference by a majority of conservative amino acid substitutions.

Conservative amino acid substitutions are substitutions among amino acids of the same class. These classes include, for example, amino acids having uncharged polar side chains, such as asparagine, glutamine, serine, threonine, and tyrosine; amino acids having basic side chains, such as lysine, arginine, and histidine; amino acids having acidic side chains, such as aspartic acid and glutamic acid; and amino acids having nonpolar side chains, such as glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, and cysteine.

Homologous polynucleotide sequences are defined in a similar way. Preferably, a homologous sequence is one that is at least 45%, (for example 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 87%, 90%, 93%, 96%) and preferably 99% identical to the coding sequence of any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53.

Homology is measured using sequence analysis software such as Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705. Amino acid sequences are aligned to maximize identity. Gaps may be artificially introduced into the sequence to attain proper alignment. Once the optimal alignment has been set up, the degree of homology is established by recording all of the positions in which the amino acids of both sequences are identical, relative to the total number of positions.

As used herein, "percent homology" of two amino acid sequences or of two nucleic acids is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA 87:2264-2268, 1990), modified as in Karlin and Altschul (Proc. Natl. Acad. Sci. USA 90:5873-5877, 1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al. (J. Mol. Biol. 215:403-410, 1990). BLAST nucleotide searches are performed with the NBLAST program, score = 100, wordlength = 12, to obtain nucleotide sequences homologous to a nucleic acid of the invention. BLAST protein searches are performed with the XBLAST program, score = 50, wordlength = 3, to obtain amino acid sequences homologous to a reference polypeptide (e.g. any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54). To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res. 25:389-3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g. XBLAST and NBLAST) are used. See <http://www.ncbi.nlm.nih.gov>.

Polypeptides having a sequence homologous to any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54 include naturally-occurring allelic variants, as well as mutants or any other non-naturally occurring variants that retain the inherent characteristics of the reference polypeptide.

As is known in the art, an allelic variant is an alternate form of a polypeptide that is characterized as having a substitution, deletion, or addition of one or more amino acids that does not alter the biological function of the polypeptide. By "biological function" is meant the

function of the polypeptide in the cells in which it naturally occurs, even if the function is not necessary for the growth or survival of the cells. For example, the biological function of a porin is to allow the entry into
5 cells of compounds present in the extracellular medium. Biological function is distinct from antigenic property. A polypeptide can have more than one biological function.

Allelic variants are very common in nature. For example, a bacterial species such as *M. catarrhalis*, is
10 usually represented by a variety of strains that differ from each other by minor allelic variations. Indeed, a polypeptide that fulfills the same biological function in different strains can have an amino acid sequence (and polynucleotide sequence) that is not identical in each of
15 the strains. Such an allelic variation may be equally reflected at the polynucleotide level.

One method of obtaining polynucleotides encoding homologous polypeptides or allelic variants is by polymerase chain reaction (PCR) amplification of genomic bacterial DNA
20 extracted by conventional methods. This involves the use of synthetic oligonucleotide primers matching upstream and downstream of the 5' and 3' ends of the coding sequence. Suitable primers are designed according to the nucleotide sequence information provided in any one of SEQ ID NOs: 1,
25 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53. The procedure is as follows: a primer is selected which consists of 10 to 40 (e.g. 13, 16, 19, 22, 24 and so forth), preferably 15 to 25 nucleotides. It is advantageous to select primers
30 containing C and G nucleotides in a proportion sufficient to ensure efficient hybridization; i.e., an amount of C and G nucleotides of at least 40%, preferably 50% of the total nucleotide content. A standard PCR reaction contains

typically 0.5 to 5 Units of Taq DNA polymerase per 100 μ L, 20 to 200 μ M deoxynucleotide each, preferably at equivalent concentrations, 0.5 to 2.5 mM magnesium over the total deoxynucleotide concentration, 10^5 to 10^6 target molecules, and about 20 pmol of each primer. About 25 to 50 PCR cycles are performed, with an annealing temperature 15°C to 5°C below the true T_m of the primers. A more stringent annealing temperature improves discrimination against incorrectly annealed primers and reduces incorporation of incorrect nucleotides at the 3' end of primers. A denaturation temperature of 95°C to 97°C is typical, although higher temperatures may be appropriate for dematuration of G+C-rich targets. The number of cycles performed depends on the starting concentration of target molecules, though typically more than 40 cycles is not recommended as non-specific background products tend to accumulate.

An alternative method for retrieving polynucleotides encoding homologous polypeptides or allelic variants is by hybridization screening of a DNA or RNA library. Hybridization procedures are well-known in the art and are described for example in Sambrook et al. (Sambrook, J.E.F. et. al. (1989) *Molecular Cloning a Laboratory Manual* 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor N.Y.). Important parameters for optimizing hybridization conditions are reflected in a formula used to obtain the critical melting temperature above which two complementary DNA strands separate from each other. For polynucleotides of about 600 nucleotides or larger, this formula is as follows: $T_m = 81.5 + 0.41 \times (\% \text{ G+C}) + 16.6 \log (\text{cation ion concentration}) - 0.63 \times (\% \text{ formamide}) - 600/\text{base number}$. Under appropriate stringency conditions, hybridization temperature (T_h) is approximately 20 to 40°C, 20 to 25°C, or, preferably 30 to 40°C below the calculated

T_m. Those skilled in the art will understand that optimal temperature and salt conditions can be readily determined.

Hybridizing nucleic acids of the type described herein can be used, for example, as a cloning probe, a
5 primer or a diagnostic probe. Hybridization is typically performed under stringent conditions. Nucleic acid duplex or hybrid stability is expressed as the melting temperature or T_m, which is the temperature at which the hybridizing strands dissociate. This melting temperature is used to
10 define the required stringency conditions. If sequences are to be identified that are related and substantially identical, rather than identical, to the probe sequence, then it is useful to first establish the lowest temperature at which only homologous hybridization occurs with a
15 particular concentration of salt (e.g. SSC OR SSPE). Then, assuming that a 1% mismatching results in a 1°C decrease in the T_m, the temperature of the final wash in the hybridization reaction is reduced accordingly (for example, if sequences having more than 95% identity with the probe
20 are sought, the final wash temperature is decreased by 5%). In practice, the change in T_m can be between 0.5°C and 1.5°C per 1% mismatch. Stringent conditions are achieved for both pre-hybridizing and hybridizing incubations (i) within 4-16 hours at 42°C, in 6 x SSC containing 50% formamide, or (ii)
25 within 4-16 hours at 65°C in an aqueous 6 x SSC solution (1 M NaCl, 0.1 M sodium citrate (pH 7.0)). Typically, hybridization experiments are performed at a temperature from 60 to 68°C, e.g. 65°C. At such a temperature, stringent hybridization conditions can be achieved in 6xSSC,
30 preferably in 2xSSC or 1xSSC, more preferably in 0.5xSSC, 0.3xSSC or 0.1xSSC (in the absence of formamide). 1xSSC contains 0.15 M NaCl and 0.015 M sodium citrate.

Useful homologs and fragments thereof that do not occur naturally are designed using known methods for identifying regions of an antigen that are likely to tolerate amino acid sequence changes and/or deletions. As an example, homologous polypeptides from different species are compared; conserved sequences are identified. The more divergent sequences are the most likely to tolerate sequence changes. Homology among sequences may be analyzed as described above using, as an example, the BLAST homology searching algorithm of Altschul et al., Nucleic Acids Res.; 25:3389-3402 (1997). Alternatively, sequences are modified such that they become more reactive to T- and/or B-cells, based on computer-assisted analysis of probable T- or B-cell epitopes. Yet another alternative is to mutate a particular amino acid residue or sequence within the polypeptide *in vitro*, then screen the mutant polypeptides for their ability to prevent or treat *Moraxella* infection according to the method outlined below.

A person skilled in the art will readily understand that by following the screening process of this invention, it will be determined without undue experimentation whether a particular homolog or fragment of any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54 may be useful in the prevention or treatment of *Moraxella* infection. The screening procedure comprises the steps:

(i) immunizing an animal, preferably mouse, with the test homolog or fragment. The homolog or fragment used for immunization may be a polypeptide or peptide, or it may be a nucleic acid encoding the polypeptide or peptide;

(ii) inoculating the immunized animal with Moraxella, preferably *M. catarrhalis*; and

(iii) selecting those homologs or fragments which elicits an immunogenic response in the animal against a
5 bacterium of the Moraxella genus.

The immunogenic response may be an antibody response, or an immunoprotective response. By "immunoprotection" is meant that there is a reduction in severity of any of the effects of Moraxella infection, in
10 comparison with a control animal which was not immunized with the test homolog or fragment.

In one embodiment, the *M. catarrhalis* complement-mediated, antibody-dependent bactericidal assay may be used as an indicator of how effective a particular Moraxella
15 polypeptide, homolog or fragment is in generating an immune response and conferring protection. The assay is performed essentially as described in Yang et al. FEMS Immunol. Med. Med Microbiol. (1997) 17:187-199. Briefly, a dilution of *M. catarrhalis* is incubated with complement-inactivated
20 antiserum in the presence of a known amount of complement, then either plated and counted (Myers et al. Infection and Immunity (1998) 66(9):4183-4192. Detection of surviving bacteria may be facilitated with the use of alamarBlue (Mountzouros and Howell. J. Clin. Microbiol. (2000) 38:2878-
25 2884).

Polypeptide derivatives are provided that are partial sequences of any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54, partial sequences of
30 polypeptide sequences homologous to any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54, polypeptides

derived from full-length polypeptides by internal deletion, and fusion proteins.

It is an accepted practice in the field of immunology to use fragments and variants of protein immunogens as vaccines, as all that is required to induce an immune response to a protein is a small (e.g., 8 to 10 amino acid) immunogenic region of the protein. Various short synthetic peptides corresponding to surface-exposed antigens of pathogens other than *Moraxella* have been shown to be effective vaccine antigens against their respective pathogens, e.g. an 11 residue peptide of murine mammary tumor virus (Dion et al., 1990. *Virology* 179:474-477), a 16-residue peptide of Semliki Forest virus (Snijders et al., 1991. *J. Gen. Virol.* 72:557-565), and two overlapping peptides of 15 residues each from canine parvovirus (Langeveld et al., *Vaccine* 12(15):1473-1480, 1994).

Accordingly, it will be readily apparent to one skilled in the art, having read the present description, that partial sequences of any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54 or their homologous amino acid sequences are inherent to the full-length sequences and are taught by the present invention. Such polypeptide fragments are at least 10 amino acids long, and preferably are at least 12 amino acids in length; for example 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 amino acids in length, and so forth up to the full-length sequence.

Polynucleotides of 30 to 600 nucleotides of any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53, or their corresponding complements, homologs or

variants, are useful at least as probes for retrieving other homologs or variants, or as a tool for diagnosing *Moraxella* infection. One method for retrieving the partial sequences is by PCR amplification. The parameters outlined above are applied and primers are used which match the sequences upstream and downstream of the 5' and 3' ends of the fragment to be amplified. The template polynucleotide for such amplification is either the full length polynucleotide of any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53, or their corresponding homologs or variants, or a polynucleotide contained in a mixture of polynucleotides such as a DNA or RNA library. As an alternative method for retrieving the partial sequences, screening hybridization is carried out under conditions described above and using the formula for calculating T_m . Where fragments of 30 to 600 nucleotides are to be retrieved, the calculated T_m is corrected by subtracting (600/polynucleotide size in base pairs) and the stringency conditions are defined by a hybridization temperature that is 5 to 10°C below T_m . Where oligonucleotides shorter than 20-30 bases are to be obtained, the formula for calculating the T_m is as follows: $T_m = 4 \times (G+C) + 2 \times (A+T)$. For example, an 18 nucleotide fragment of 50% G+C would have an approximate T_m of 54°C.

Short peptides that are fragments of any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54 or their corresponding homologous sequences, may be obtained directly by chemical synthesis (E. Gross and H. J. Meinhofer, 4 The Peptides: Analysis, Synthesis, Biology; Modern Techniques of Peptide Synthesis, John Wiley & Sons (1981), and M. Bodanzki, Principles of Peptide Synthesis, Springer-Verlag (1984)).

Useful polypeptide derivatives, e.g., polypeptide fragments, are designed using computer-assisted analysis of amino acid sequences. This would identify probable surface-exposed, antigenic regions (Hughes et al., 1992. Infect. Immun. 60(9):3497). Analysis of 6 amino acid sequences contained in any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54, based on the product of flexibility and hydrophobicity propensities using the program SEQSEE (Wishart DS, et al. "SEQSEE: a comprehensive program suite for protein sequence analysis." *Comput Appl Biosci.* 1994 Apr;10(2):121-32), can reveal potential B- and T-cell epitopes which may be used as a basis for selecting useful immunogenic fragments and variants. This analysis uses a reasonable combination of external surface features that is likely to be recognized by antibodies. Probable T-cell epitopes for HLA-A0201 MHC subclass may be revealed by algorithms that emulate an approach developed at the NIH (Parker KC, et al. "Peptide binding to MHC class I molecules: implications for antigenic peptide prediction." *Immunol Res* 1995;14(1):34-57).

Epitopes which induce a protective T cell-dependent immune response are present throughout the length of the polypeptide. However, some epitopes may be masked by secondary and tertiary structures of the polypeptide. To reveal such masked epitopes large internal deletions are created which remove much of the original protein structure and exposes the masked epitopes. Such internal deletions sometimes effect the additional advantage of removing immunodominant regions of high variability among strains.

Polynucleotides encoding polypeptide fragments and polypeptides having large internal deletions are constructed using standard methods (Sambrook, J.E.F. et. al. (1989)

Molecular Cloning a Laboratory Manual 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor N.Y.). Such methods include standard PCR, inverse PCR, and restriction enzyme treatment of cloned DNA molecules. Components for these methods and instructions for their use are readily available from various commercial sources such as Stratagene. Once the deletion mutants have been constructed, they are tested for their ability to prevent or treat Moraxella infection as described above.

As used herein, a fusion polypeptide is one that contains a polypeptide or a polypeptide derivative of the invention fused at the N- or C-terminal end to any other polypeptide (hereinafter referred to as a peptide tail). A simple way to obtain such a fusion polypeptide is by translation of an in-frame fusion of the polynucleotide sequences, *i.e.*, a hybrid gene. The hybrid gene encoding the fusion polypeptide is inserted into an expression vector which is used to transform or transfect a host cell. Alternatively, the polynucleotide sequence encoding the polypeptide or polypeptide derivative is inserted into an expression vector in which the polynucleotide encoding the peptide tail is already present. Such vectors and instructions for their use are commercially available, *e.g.* the pMal-c2 or pMal-p2 system from New England Biolabs, in which the peptide tail is a maltose binding protein, the glutathione-S-transferase system of Pharmacia, or the His-Tag system available from Novagen. These and other expression systems provide convenient means for further purification of polypeptides and derivatives of the invention.

An advantageous example of a fusion polypeptide is one where the polypeptide or homolog or fragment of the invention is fused to a polypeptide having adjuvant

activity, such as subunit B of either cholera toxin or *E. coli* heat-labile toxin. Another advantageous fusion is one where the polypeptide, homolog or fragment is fused to a strong T-cell epitope or B-cell epitope. Such an epitope
5 may be one known in the art (e.g. the Hepatitis B virus core antigen, D.R. Millich et al., "Antibody production to the nucleocapsid and envelope of the Hepatitis B virus primed by a single synthetic T cell site", Nature. 1987. 329:547-549), or one which has been identified in another polypeptide of
10 the invention based on computer-assisted analysis of probable T- or B-cell epitopes. Consistent with this aspect of the invention is a fusion polypeptide comprising T- or B-cell epitopes from any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40,
15 42, 44, 46, 48, 50, 52 and 54 or their corresponding homolog or fragment, wherein the epitopes are derived from multiple variants of said polypeptide or homolog or fragment, each variant differing from another in the location and sequence of its epitope within the polypeptide. Such a fusion is
20 effective in the prevention and treatment of Moraxella infection since it optimizes the T- and B-cell response to the overall polypeptide, homolog or fragment.

To effect fusion, the polypeptide of the invention is fused to the N-, or preferably, to the C-terminal end of
25 the polypeptide having adjuvant activity or T- or B-cell epitope. Alternatively, a polypeptide fragment of the invention is inserted internally within the amino acid sequence of the polypeptide having adjuvant activity. The T- or B-cell epitope may also be inserted internally within
30 the amino acid sequence of the polypeptide of the invention.

The polynucleotides of the invention also encode hybrid precursor polypeptides containing heterologous signal peptides, which mature into polypeptides of the invention.

By "heterologous signal peptide" is meant a signal peptide that is not found in naturally-occurring precursors of polypeptides of the invention.

Polynucleotide molecules according to the invention, including RNA, DNA, or modifications or combinations thereof, have various applications. A DNA molecule is used, for example, (i) in a process for producing the encoded polypeptide in a recombinant host system, (ii) in the construction of vaccine vectors such as poxviruses, which are further used in methods and compositions for preventing and/or treating Moraxella infection, (iii) as a vaccine agent (as well as an RNA molecule), in a naked form or formulated with a delivery vehicle and, (iv) in the construction of attenuated Moraxella strains that can over-express a polynucleotide of the invention or express it in a non-toxic, mutated form.

Accordingly, another aspect of the invention encompasses (i) an expression cassette containing a DNA molecule of the invention placed under the control of the elements required for expression, in particular under the control of an appropriate promoter; (ii) an expression vector containing an expression cassette of the invention; (iii) a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, or a virus expressing a DNA molecule of the invention; as well as (iv) a process for producing a polypeptide or polypeptide derivative encoded by a polynucleotide of the invention, which involves culturing a virus, procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, under conditions that allow expression of the DNA molecule of the invention and, recovering the encoded

polypeptide or polypeptide derivative from the cell or virus culture.

A recombinant expression system is selected from viruses, procaryotic and eucaryotic hosts. Eucaryotic hosts
5 include yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris*), mammalian cells (e.g., COS1, NIH3T3, or JEG3 cells), arthropods cells (e.g., *Spodoptera frugiperda* (SF9) cells), and plant cells. Prokaryotic hosts include: *E. coli*, *Shigella*, *Salmonella*, *Vibrio cholerae*,
10 *Lactobacillus*, Bacille bilié de Calmette-Guérin (BCG), and *Streptococcus*, preferably non-toxicogenic *Vibrio cholerae* mutant strains and attenuated *Salmonella typhimurium* strains. Viruses which may be used to produce the polypeptides and proteins of the invention include
15 adenovirus, alphavirus, and poxvirus, which may be a vaccinia or canary pox virus.

Bacterial and eucaryotic cells are available from a number of different sources including commercial sources to those skilled in the art, e.g., the American Type Culture
20 Collection (ATCC; Rockville, Maryland). Commercial sources of cells used for recombinant protein expression also provide instructions for usage of the cells.

An example of an adenovirus expression system, as well as a method for constructing an adenovirus vector
25 capable of expressing a DNA molecule of the invention, are described in U.S. Patent No. 4,920,209. Poxvirus vectors include vaccinia and canary pox virus, described in U.S. Patent No. 4,722,848 and U.S. Patent No. 5,364,773, respectively. Poxvirus vectors capable of expressing a
30 polynucleotide of the invention are obtained by homologous recombination so that the polynucleotide of the invention is

inserted in the viral genome under appropriate conditions for expression in mammalian cells.

Non-toxicogenic *Vibrio cholerae* mutant strains are known. U.S. Patent No. 4,882,278 describes strains which
5 have a substantial amount of the coding sequence of each of the two *ctxA* alleles deleted so that no functional *cholerae* toxin is produced. WO 92/11354 describes a strain in which the *irgA* locus is inactivated by mutation; this mutation can be combined in a single strain with *ctxA* mutations. WO
10 94/01533 describes a deletion mutant lacking functional *ctxA* and *attRS1* DNA sequences. These mutant strains are genetically engineered to express heterologous proteins, as described in WO 94/19482.

Attenuated *Salmonella typhimurium* strains,
15 genetically engineered for recombinant expression of heterologous proteins or not, are described in WO 92/11361.

Other bacterial strains used as protein expression hosts in the context of the present invention are described for *Shigella flexneri* in High et al., EMBO (1992) 11:1991
20 and Sizemore et al., Science (1995) 270:299; for *Streptococcus gordonii* in Medaglini et al., Proc. Natl. Acad. Sci. USA (1995) 92:6868; and for Bacille Calmette Guerin in Flynn J.L., Cell. Mol. Biol. (1994) 40 (suppl. I):31, WO 88/06626, WO 90/00594, WO 91/13157, WO 92/01796,
25 and WO 92/21376.

In bacteria, the polynucleotide of the invention is inserted into the bacterial genome or remains in a free state as part of a plasmid.

The choice of the expression system depends on the
30 features desired for the expressed polypeptide. For

example, it may be useful to produce a polypeptide of the invention in a particular lipidated form or any other form.

One skilled in the art would readily understand that not all vectors and expression control sequences and
5 hosts would be expected to express equally well the polynucleotides of this invention. With the guidelines described below, however, a selection of vectors, expression control sequences and hosts may be made without undue experimentation and without departing from the scope of this
10 invention.

In selecting a vector, the host must be chosen that is compatible with the vector which is to exist and possibly replicate in it. Considerations are made with respect to the vector copy number, the ability to control
15 the copy number, expression of other proteins such as antibiotic resistance. In selecting an expression control sequence, a number of variables are considered. Among the important variables are the relative strength of the sequence (e.g. the ability to drive expression under various
20 conditions), the ability to control the sequence's function, compatibility between the polynucleotide to be expressed and the control sequence (e.g. secondary structures are considered to avoid hairpin structures which prevent efficient transcription). In selecting the host,
25 unicellular hosts are selected which are compatible with the selected vector, tolerant of any possible toxic effects of the expressed product, able to secrete the expressed product efficiently if such is desired, to be able to express the product in the desired conformation, to be easily scaled up,
30 and to which ease of purification of the final product.

The choice of the expression cassette depends on the host system selected as well as the features desired for

the expressed polypeptide. Typically, an expression cassette includes a promoter that is functional in the selected host system and can be constitutive or inducible; a ribosome binding site; a start codon (ATG) if necessary; a region encoding a signal peptide, e.g., a lipidation signal peptide; a DNA molecule of the invention; a stop codon; and optionally a 3' terminal region (translation and/or transcription terminator). The signal peptide encoding region is adjacent to the polynucleotide of the invention and placed in proper reading frame. The signal peptide-encoding region is homologous or heterologous to the DNA molecule encoding the mature polypeptide and is compatible with the secretion apparatus of the host used for expression. The open reading frame constituted by the DNA molecule of the invention, solely or together with the signal peptide, is placed under the control of the promoter so that transcription and translation occur in the host system. Promoters and signal peptide encoding regions are widely known and available to those skilled in the art and include, for example, the promoter of *Salmonella typhimurium* (and derivatives) that is inducible by arabinose (promoter araB) and is functional in Gram-negative bacteria such as *E. coli* (as described in U.S. Patent No. 5,028,530 and in Cagnon et al., (Cagnon et al., Protein Engineering (1991) 4(7):843)); the promoter of the gene of bacteriophage T7 encoding RNA polymerase, that is functional in a number of *E. coli* strains expressing T7 polymerase (described in U.S. Patent No. 4,952,496); OspA lipidation signal peptide ; and RlpB lipidation signal peptide (Takase et al., J. Bact. (1987) 169:5692).

The expression cassette is typically part of an expression vector, which is selected for its ability to replicate in the chosen expression system. Expression

vectors (e.g., plasmids or viral vectors) can be chosen, for example, from those described in Pouwels et al. (Cloning Vectors: A Laboratory Manual 1985, Supp. 1987). Suitable expression vectors can be purchased from various commercial
5 sources.

Methods for transforming/transfecting host cells with expression vectors are well-known in the art and depend on the host system selected as described in Sambrook, J.E.F. et. al. (1989) Molecular Cloning a Laboratory Manual 2nd ed.
10 Cold Spring Harbor Laboratory Press, Cold Spring Harbor N.Y.

Upon expression, a recombinant polypeptide of the invention (or a polypeptide derivative) is produced and remains in the intracellular compartment, is
secreted/excreted in the extracellular medium or in the
15 periplasmic space, or is embedded in the cellular membrane. The polypeptide is recovered in a substantially purified form from the cell extract or from the supernatant after centrifugation of the recombinant cell culture. Typically, the recombinant polypeptide is purified by antibody-based
20 affinity purification or by other well-known methods that can be readily adapted by a person skilled in the art, such as fusion of the polynucleotide encoding the polypeptide or its derivative to a small affinity binding domain. Antibody-based affinity purification methods are also
25 available for purifying polypeptides extracted from a *Moraxella* strain. Antibodies useful for purifying by immunoaffinity the polypeptides of the invention are obtained as described below.

A polynucleotide of the invention can also be
30 useful as a vaccine. There are two major routes, either using a viral or bacterial host as gene delivery vehicle (live vaccine vector) or administering the gene in a free

form, e.g., inserted into a plasmid. Therapeutic or prophylactic efficacy of a polynucleotide of the invention is evaluated as described below.

Accordingly, another aspect of the invention provides (i) a vaccine vector such as a poxvirus, containing a DNA molecule of the invention, placed under the control of elements required for expression; (ii) a composition of matter comprising a vaccine vector of the invention, together with a diluent or carrier; specifically (iii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a vaccine vector of the invention; (iv) a method for inducing an immune response against *Moraxella* in a mammal (e.g., a human; alternatively, the method can be used in veterinary applications for treating or preventing *Moraxella* infection of animals, e.g., cats or birds), which involves administering to the mammal an immunogenically effective amount of a vaccine vector of the invention to elicit a protective or therapeutic immune response to *Moraxella*; and (v) a method for preventing and/or treating a *Moraxella* infection, which involves administering a prophylactic or therapeutic amount of a vaccine vector of the invention to an infected individual. Additionally, the invention encompasses the use of a vaccine vector of the invention in the preparation of a medicament for preventing and/or treating *Moraxella* infection.

In one embodiment, the vaccine, composition or treatment using the vaccine vector of the invention is free of adjuvant, specifically adjuvants commonly or specifically used in rodents. The vaccine, composition or treatment may be used to treat disorders whose symptoms are caused or aggravated at least in part by *Moraxella* infection, specifically *M. catarrhalis*, and includes such disorders as otitis media, respiratory infection, sinusitis and

pneumonia. Preferred vaccines and pharmaceutical compositions include those formulated for *in vivo* administration to an animal, preferably a human, to confer protection or treatment against disease caused by *M.*

5 *catarrhalis*. Also preferred are compositions formulated as a microparticle, capsule, or liposome.

As used herein, a vaccine vector expresses one or several polypeptides or derivatives of the invention. The vaccine vector may express additionally a cytokine, such as
10 interleukin-2 (IL-2) or interleukin-12 (IL-12), that enhances the immune response (adjuvant effect). It is understood that each of the components to be expressed is placed under the control of elements required for expression in a mammalian cell.

15 The invention also provides a composition comprising several vaccine vectors, each of them capable of expressing a polypeptide or derivative of the invention. The composition may also comprise a vaccine vector capable of expressing an additional *Moraxella* antigen, or a subunit,
20 fragment, homolog, mutant, or derivative thereof; optionally together with or a cytokine such as IL-2 or IL-12.

A general principle is that recognition of a particular antigen is not in itself sufficient to produce an effective immune response. In some cases, a cell-mediated
25 response is appropriate; in others, antibody.

Antigens of microorganisms vary considerably in their accessibility to cells of the immune system. Antigens which normally occur inside a pathogen may become accessible only when the pathogen or an infected cell is killed. Even
30 antigens expressed at the cell surface may present only a limited range of their potential epitopes for antibody binding, depending on their orientation in the membrane.

Protective structures, such as bacterial capsules, further limit the effective recognition of epitopes.

A distinction should be drawn between the overall composition of the immune response, those components of it which are important in the resolution of infection and the components which are responsible for the prevention of re-infection. In many cases, particular elements of the immune response are critically important; for example, cell-mediated immunity in leprosy. Even when considering a particular effector system, the response directed against some antigens is often much more effective than the responses to others. Immune responses to particular microbial antigens have different degrees of relevance to anti-microbial immunity, depending on the nature of the organism, its pathogenicity and the nature of the immune response it initiates.

The primary effectors against extracellular pathogens are antibody and complement. Binding of antibody to receptors on the pathogen can prevent it from attaching to its target cell. Antibody alone, or more effectively in association with complement, opsonizes pathogens for uptake by phagocytes expressing Fc receptors and complement receptors CR1 and CR3. Usually this will lead to intracellular destruction of the pathogen but if the phagocyte is unable to destroy it and is a facultative host cell, then antibody may actually promote the spread of infection. Such an eventuality, however, depends on the dynamic balance between the actions of the humoral and cell-mediated immune responses.

Sometimes effective antibodies must be of the right class to activate appropriate effectors. The important antigens are those involved in evasion of immune

effector mechanisms; that is, pili, fimbriae and capsular antigens which constitute the major antigens of the outer layer of bacteria. Often epitope specificity is important, since it determines whether complement is deposited in a position to damage the outer membrane. There are also numerous protein antigens which can induce an antibody response; however, although the antibody response is partly species-specific and may be diagnostically useful, it is largely irrelevant to immunity. This is most obvious in lepromatous leprosy, where the patients have weak cell-mediated immunity, high levels of specific antibody and tissues heavily infected with bacteria.

In some cases, a particular type of antibody response is mandatory for clearance of the pathogen. This is true of many bacterial infections, where specific antibodies to surface antigens are necessary to neutralize the bacterial defences and opsonize the bacteria for phagocytes.

There are also cases where responses to individual antigens are essential for host immunity. The simplest examples are the toxins produced by the causative agents of diphtheria, tetanus and clostridial enteritis. The damage produced directly by the infectious agent in these diseases is slight by comparison with that produced by the secreted toxins. Consequently, protection against these conditions involves immunization to toxoids. Nevertheless, the immune system must still eradicate the primary site of the bacterial infection if the disease is to be resolved. The target antigens for bactericidal antibodies are extremely diverse and include LPS, capsular polysaccharides and other outer membrane proteins. Virulence factors can also provide good immunogens in a vaccine.

Vaccination methods for treating or preventing infection in a mammal comprises use of a vaccine vector of the invention to be administered by any conventional route, particularly to a mucosal (e.g., ocular, intranasal, oral, gastric, pulmonary, intestinal, rectal, vaginal, or urinary tract) surface or via the parenteral (e.g., subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. Preferred routes depend upon the choice of the vaccine vector. Treatment may be effected in a single dose or repeated at intervals. The appropriate dosage depends on various parameters understood by skilled artisans such as the vaccine vector itself, the route of administration or the condition of the mammal to be vaccinated (weight, age and the like).

Live vaccine vectors available in the art include viral vectors such as adenoviruses, alphaviruses and poxviruses as well as bacterial vectors, e.g., *Shigella*, *Salmonella*, *Vibrio cholerae*, *Lactobacillus*, Bacille bilié de Calmette-Guérin (BCG), and *Streptococcus*.

An example of an adenovirus vector, as well as a method for constructing an adenovirus vector capable of expressing a DNA molecule of the invention, are described in U.S. Patent No. 4,920,209. Poxvirus vectors include vaccinia and canary pox virus, described in U.S. Patent No. 4,722,848 and U.S. Patent No. 5,364,773, respectively. Poxvirus vectors capable of expressing a polynucleotide of the invention are obtained by homologous recombination so that the polynucleotide of the invention is inserted in the viral genome under appropriate conditions for expression in mammalian cells. Generally, the dose of vaccine viral vector, for therapeutic or prophylactic use, can be of from about 1×10^4 to about 1×10^{11} , advantageously from about 1×10^7 to about 1×10^{10} , preferably of from about 1×10^7 to about 1×10^9 .

plaque-forming units per kilogram. Preferably, viral vectors are administered parenterally; for example, in 3 doses, 4 weeks apart. It is preferable to avoid adding a chemical adjuvant to a composition containing a viral vector of the invention and thereby minimizing the immune response to the viral vector itself.

Non-toxicogenic *Vibrio cholerae* mutant strains that are useful as a live oral vaccine are known. U.S. Patent No. 4,882,278 describes strains which have a substantial amount of the coding sequence of each of the two *ctxA* alleles deleted so that no functional *cholerae* toxin is produced. WO 92/11354 describes a strain in which the *irgA* locus is inactivated by mutation; this mutation can be combined in a single strain with *ctxA* mutations. WO 94/01533 describes a deletion mutant lacking functional *ctxA* and *attRS1* DNA sequences. These mutant strains are genetically engineered to express heterologous antigens, as described in WO 94/19482. An effective vaccine dose of a *Vibrio cholerae* strain capable of expressing a polypeptide or polypeptide derivative encoded by a DNA molecule of the invention contains about 1×10^5 to about 1×10^9 , preferably about 1×10^6 to about 1×10^8 , viable bacteria in a volume appropriate for the selected route of administration. Preferred routes of administration include all mucosal routes; most preferably, these vectors are administered intranasally or orally.

Attenuated *Salmonella typhimurium* strains, genetically engineered for recombinant expression of heterologous antigens or not, and their use as oral vaccines are described in WO 92/11361. Preferred routes of administration include all mucosal routes; most preferably, these vectors are administered intranasally or orally.

Other bacterial strains used as vaccine vectors in the context of the present invention are described for *Shigella flexneri* in High et al., EMBO (1992) 11:1991 and Sizemore et al., Science (1995) 270:299; for *Streptococcus*
5 *gordonii* in Medaglini et al., Proc. Natl. Acad. Sci. USA (1995) 92:6868; and for Bacille Calmette Guerin in Flynn J.L., Cell. Mol. Biol. (1994) 40 (suppl. I):31, WO 88/06626, WO 90/00594, WO 91/13157, WO 92/01796, and WO 92/21376.

In bacterial vectors, the polynucleotide of the
10 invention is inserted into the bacterial genome or remains in a free state as part of a plasmid.

The composition comprising a vaccine bacterial vector of the present invention may further contain an adjuvant, preferably an adjuvant appropriate for human or
15 veterinary use and which preferably excludes rodent-specific adjuvants. A number of adjuvants are known to those skilled in the art. Preferred adjuvants are selected as provided below.

Accordingly, another aspect of the invention
20 provides (i) a composition comprising a polynucleotide of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition comprising a therapeutically or prophylactically effective amount of a polynucleotide of the invention; (iii) a method for inducing an immune response
25 against *Moraxella* in a mammal, particularly a mammal susceptible to infection with *Moraxella*, specifically *M. catarrhalis*, by administration of an immunogenically effective amount of a polynucleotide of the invention to elicit a protective immune response to *Moraxella*; and, (iv)
30 a method for preventing and/or treating a *Moraxella* infection, by administering a prophylactic or therapeutic

amount of a polynucleotide of the invention to an infected individual.

Additionally, the invention encompasses the use of a polynucleotide of the invention in the preparation of a medicament for preventing and/or treating *Moraxella* infection. A preferred use includes the use of a DNA molecule placed under conditions for expression in a mammalian cell, especially in a plasmid that is unable to replicate in mammalian cells and to substantially integrate in a mammalian genome.

In one embodiment, the polynucleotide, composition or treatment is free of adjuvant, specifically adjuvants commonly or specifically used in rodents. The polynucleotide, composition or treatment may be used to treat disorders whose symptoms are caused or aggravated at least in part by *Moraxella* infection, specifically *M. catarrhalis*, and includes such disorders as otitis media, respiratory infection, sinusitis and pneumonia. Preferred polynucleotide compositions include those formulated for *in vivo* administration to an animal, preferably a human, to confer protection or treatment against disease caused by *M. catarrhalis*. Also preferred are compositions formulated as a microparticle, capsule, or liposome.

Use of the polynucleotides of the invention include their administration to a mammal as a vaccine, for therapeutic or prophylactic purposes. Such polynucleotides are used in the form of DNA as part of a plasmid that is unable to replicate in a mammalian cell and unable to integrate into the mammalian genome. Typically, such a DNA molecule is placed under the control of a promoter suitable for expression in a mammalian cell. The promoter functions either ubiquitously or tissue-specifically. Examples of

non-tissue specific promoters include the early Cytomegalovirus (CMV) promoter (described in U.S. Patent No. 4,168,062) and the Rous Sarcoma Virus promoter (described in Norton & Coffin, Molec. Cell Biol. (1985) 5:281). An example of a tissue-specific promoter is the desmin promoter which drives expression in muscle cells (Li et al., Gene (1989) 78:243 and Li & Paulin, J. Biol. Chem. (1993) 268:10403). Use of promoters is well-known to those skilled in the art. Useful vectors are described in numerous publications, specifically WO 94/21797.

In one embodiment, polynucleotides of the invention which are used as vaccines encode either a precursor or a mature form of the corresponding polypeptide. In the precursor form, the signal peptide is either homologous or heterologous. In the latter case, a eucaryotic leader sequence such as the leader sequence of the tissue-type plasminogen factor (tPA) is preferred.

As used herein, a composition of the invention contains one or several polynucleotides with optionally at least one additional polynucleotide encoding another Moraxella antigen or a fragment, derivative, mutant, or analog thereof. The composition may also contain an additional polynucleotide encoding a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12) so that the immune response is enhanced. These additional polynucleotides are placed under appropriate control for expression. Advantageously, DNA molecules of the invention and/or additional DNA molecules to be included in the same composition, are present in the same plasmid.

Standard techniques of molecular biology for preparing and purifying polynucleotides are used in the preparation of polynucleotide therapeutics of the invention.

For use as a vaccine, a polynucleotide of the invention is formulated according to various methods outlined below.

One method utilizes the polynucleotide in a naked form, free of any delivery vehicles. Such a polynucleotide is simply diluted in a physiologically acceptable solution such as sterile saline or sterile buffered saline, with or without a carrier. When present, the carrier preferably is isotonic, hypotonic, or weakly hypertonic, and has a relatively low ionic strength, such as provided by a sucrose solution, e.g., a solution containing 20% sucrose.

An alternative method utilizes the polynucleotide in association with agents that assist in cellular uptake. Examples of such agents are (i) chemicals that modify cellular permeability, such as bupivacaine (see, e.g., WO 94/16737), (ii) liposomes for encapsulation of the polynucleotide, or (iii) cationic lipids or silica, gold, or tungsten microparticles which associate themselves with the polynucleotides.

Anionic and neutral liposomes are well-known in the art (see, e.g., Liposomes: A Practical Approach, RPC New Ed, IRL press (1990), for a detailed description of methods for making liposomes) and are useful for delivering a large range of products, including polynucleotides.

Cationic lipids are also known in the art and are commonly used for gene delivery. Such lipids include LipofectinTM also known as DOTMA (N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium chloride), DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane), DDAB (dimethyldioctadecylammonium bromide), DOGS (dioctadecylamidoglycyl spermine) and cholesterol derivatives such as DC-Chol (3 beta-(N-(N',N'-dimethyl

aminomethane)-carbamoyl) cholesterol). A description of these cationic lipids can be found in EP 187,702, WO 90/11092, U.S. Patent No. 5,283,185, WO 91/15501, WO 95/26356, and U.S. Patent No. 5,527,928. Cationic lipids
5 for gene delivery are preferably used in association with a neutral lipid such as DOPE (dioleoyl phosphatidylethanolamine), as described in WO 90/11092 as an example.

Formulation containing cationic liposomes may
10 optionally contain other transfection-facilitating compounds. A number of them are described in WO 93/18759, WO 93/19768, WO 94/25608, and WO 95/02397. They include spermine derivatives useful for facilitating the transport of DNA through the nuclear membrane (see, for example,
15 WO 93/18759) and membrane-permeabilizing compounds such as GALA, Gramicidine S, and cationic bile salts (see, for example, WO 93/19768).

Gold or tungsten microparticles are used for gene delivery, as described in WO 91/00359, WO 93/17706, and
20 Tang et al. Nature (1992) 356:152. The microparticle-coated polynucleotide is injected via intradermal or intraepidermal routes using a needleless injection device ("gene gun"), such as those described in U.S. Patent No. 4,945,050, U.S. Patent No. 5,015,580, and WO 94/24263.

25 The amount of DNA to be used in a vaccine recipient depends, e.g., on the strength of the promoter used in the DNA construct, the immunogenicity of the expressed gene product, the condition of the mammal intended for administration (e.g., the weight, age, and general
30 health of the mammal), the mode of administration, and the type of formulation. In general, a therapeutically or prophylactically effective dose from about 1 μ g to about

1 mg, preferably, from about 10 μ g to about 800 μ g and, more preferably, from about 25 μ g to about 250 μ g, can be administered to human adults. The administration can be achieved in a single dose or repeated at intervals.

5 The route of administration is any conventional route used in the vaccine field. As general guidance, a polynucleotide of the invention is administered via a mucosal surface, e.g., an ocular, intranasal, pulmonary, oral, intestinal, rectal, vaginal, and urinary tract
10 surface; or via a parenteral route, e.g., by an intravenous, subcutaneous, intraperitoneal, intradermal, intraepidermal, or intramuscular route. The choice of administration route depends on the formulation that is selected. A polynucleotide formulated in association with bupivacaine is
15 advantageously administered into muscles. When a neutral or anionic liposome or a cationic lipid, such as DOTMA or DC-Chol, is used, the formulation can be advantageously injected via intravenous, intranasal (aerosolization), intramuscular, intradermal, and subcutaneous routes. A
20 polynucleotide in a naked form can advantageously be administered via the intramuscular, intradermal, or subcutaneous routes.

 Although not absolutely required, such a composition can also contain an adjuvant. If so, a systemic
25 adjuvant that does not require concomitant administration in order to exhibit an adjuvant effect is preferable such as, e.g., QS21, which is described in U.S. Patent No. 5,057,546.

 The sequence information provided in the present application enables the design of specific nucleotide probes
30 and primers that are used for diagnostic purposes. Accordingly, another aspect of the invention provides a nucleotide probe or primer having a sequence found in or

derived by degeneracy of the genetic code from a sequence shown in any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53, or from a corresponding homologous
5 sequence.

The term "probe" as used in the present application refers to DNA (preferably single stranded) or RNA molecules (or modifications or combinations thereof) that hybridize under the stringent conditions, as defined
10 above, to nucleic acid molecules having any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53 or to their corresponding homologous, complementary or anti-sense sequences. Generally, probes are significantly shorter than
15 full-length sequences. Such probes contain from about 5 to about 100, (for example, 10, 12, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60 ... and so forth), preferably from about 10 to about 80, nucleotides. In particular, probes have sequences that are at least 75% (for example at least 80%, 85%, 90%,
20 95%, 96%, 97%, 98%, 99%), and preferably at least 95% homologous to a portion of any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53. Probes may also have sequences that are complementary to such sequences and
25 portions. Probes may contain modified bases such as inosine, methyl-5-deoxycytidine, deoxyuridine, dimethylamino-5-deoxyuridine, or diamino-2, 6-purine. Sugar or phosphate residues may also be modified or substituted. For example, a deoxyribose residue may be replaced by a
30 polyamide (Nielsen et al., Science (1991) 254:1497) and phosphate residues may be replaced by ester groups such as diphosphate, alkyl, arylphosphonate and phosphorothioate esters. In addition, the 2'-hydroxyl group on

ribonucleotides may be modified by including such groups as alkyl groups.

Probes of the invention are used in diagnostic tests, as capture or detection probes. Such capture probes
5 are conventionally immobilized on a solid support, directly or indirectly, by covalent means or by passive adsorption. A detection probe is labelled by a detection marker selected from: radioactive isotopes, enzymes such as peroxidase, alkaline phosphatase, and enzymes able to hydrolyze a
10 chromogenic, fluorogenic, or luminescent substrate, compounds that are chromogenic, fluorogenic, or luminescent, nucleotide base analogs, and biotin.

Probes of the invention are used in any conventional hybridization technique, such as dot blot
15 (Maniatis *et al.*, Molecular Cloning: A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), Southern blot, northern blot, or the sandwich technique (Dunn *et al.*, Cell (1977) 12:23). The latter technique involves the use of a specific capture
20 probe and/or a specific detection probe with nucleotide sequences that at least partially differ from each other.

A primer is a probe of usually about 10 to about 40 nucleotides (for example, 11, 12, 13, 14, 15, 17, 20, 25, 30 or 35 nucleotides) that is used to initiate enzymatic
25 polymerization of DNA in an amplification process (e.g., PCR), in an elongation process, or in a reverse transcription method. Primers used in diagnostic methods involving PCR are labeled by methods known in the art.

As described herein, the invention also
30 encompasses (i) a reagent comprising a probe of the invention for detecting and/or identifying the presence of *Moraxella* in a biological material; (ii) a method for

detecting and/or identifying the presence of Moraxella in a biological material. In one embodiment, the detecting or identifying method involves the following steps: (a) a sample is recovered or derived from the biological material, (b) DNA or RNA is extracted from the material and denatured, and (c) exposed to a probe of the invention, for example, a capture, detection probe or both, under stringent hybridization conditions, such that hybridization is detected. In another embodiment, the detecting or identifying method involves the following steps: (a) a sample is recovered or derived from the biological material, (b) DNA is extracted therefrom, (c) the extracted DNA is primed with at least one, and preferably two, primers of the invention and amplified by polymerase chain reaction, and (d) the amplified DNA fragment is produced.

It is apparent that disclosure of polynucleotide sequences of any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53, their homologs and partial sequences enable their corresponding amino acid sequences. Accordingly, the invention also encompasses substantially purified polypeptides or polypeptide derivatives having an amino acid sequence encoded by a polynucleotide of the invention.

A "substantially purified polypeptide" as used herein is defined as a polypeptide that is separated from the environment in which it naturally occurs and/or that is free of the majority of the polypeptides that are present in the environment in which it was synthesized. For example, a substantially purified polypeptide is free from cytoplasmic polypeptides. One band among a number of bands on an SDS-PAGE gel, for example, would not constitute a "substantially purified polypeptide" since the gel sample contains a number

of polypeptide species. Those skilled in the art would readily understand that the polypeptides of the invention may be purified from a natural source, *i.e.*, a *Moraxella* strain, or produced by recombinant means.

5 The invention encompasses polypeptides, homologs or fragments which are modified or treated to enhance their immunogenicity in the target animal, in whom the polypeptide, homolog or fragments are intended to confer protection against *Moraxella*. Such modifications or
10 treatments include: amino acid substitutions with an amino acid derivative such as 3-methylhistidine, 4-hydroxyproline, 5-hydroxylysine etc., modifications or deletions which are carried out after preparation of the polypeptide, homolog or fragment, such as the modification of free amino, carboxyl
15 or hydroxyl side groups of the amino acids.

 Identification of homologous polypeptides or polypeptide derivatives or fragments which have specific antigenicity is achieved by screening for cross-reactivity with an antiserum raised against the polypeptide of
20 reference having an amino acid sequence corresponding to any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54. The procedure is as follows: a monospecific hyperimmune antiserum is raised against a purified reference
25 polypeptide, a fusion polypeptide (for example, an expression product of MBP, GST, or His-tag systems, the description and instructions for use of which are contained in Invitrogen product manuals for pCDNA3.1/Myc-His(+) A, B, and C and for the Xpress[™] System Protein Purification), or a
30 synthetic peptide predicted to be antigenic. Where an antiserum is raised against a fusion polypeptide, two different fusion systems are employed. Specific antigenicity can be determined according to a number of

methods, including Western blot (Towbin *et al.*, Proc. Natl. Acad. Sci. USA (1979) 76:4350), dot blot, and ELISA, as described below.

In a Western blot assay, the product to be
5 screened, either as a purified preparation or a total *E. coli* extract, is submitted to SDS-Page electrophoresis as described by Laemmli (Nature (1970) 227:680). After transfer to a nitrocellulose membrane, the material is further incubated with the monospecific hyperimmune
10 antiserum diluted in the range of dilutions from about 1:5 to about 1:5000, preferably from about 1:100 to about 1:500. Specific antigenicity is shown once a band corresponding to the product exhibits reactivity at any of the dilutions in the above range.

15 In an ELISA assay, the product to be screened is preferably used as the coating antigen. A purified preparation is preferred, although a whole cell extract can also be used. Briefly, about 100 μ l of a preparation at about 10 μ g protein/ml are distributed into wells of a 96-
20 well polycarbonate ELISA plate. The plate is incubated for 2 hours at 37°C then overnight at 4°C. The plate is washed with phosphate buffer saline (PBS) containing 0.05% Tween 20 (PBS/Tween buffer). The wells are saturated with 250 μ l PBS containing 1% bovine serum albumin (BSA) to prevent non-
25 specific antibody binding. After 1 hour incubation at 37°C, the plate is washed with PBS/Tween buffer. The antiserum is serially diluted in PBS/Tween buffer containing 0.5% BSA. 100 μ l of dilutions are added per well. The plate is incubated for 90 minutes at 37°C, washed and evaluated
30 according to standard procedures. For example, a goat anti-rabbit peroxidase conjugate is added to the wells when specific antibodies were raised in rabbits. Incubation is carried out for 90 minutes at 37°C and the plate is washed.

The reaction is developed with the appropriate substrate and the reaction is measured by colorimetry (absorbance measured spectrophotometrically). Under the above experimental conditions, a positive reaction is shown by O.D. values greater than a non immune control serum.

In a dot blot assay, a purified product is preferred, although a whole cell extract can also be used. Briefly, a solution of the product at about 100 µg/ml is serially two-fold diluted in 50 mM Tris-HCl (pH 7.5). 10 100 µl of each dilution are applied to a nitrocellulose membrane 0.45 µm set in a 96-well dot blot apparatus (Biorad). The buffer is removed by applying vacuum to the system. Wells are washed by addition of 50 mM Tris-HCl (pH 7.5) and the membrane is air-dried. The membrane is 15 saturated in blocking buffer (50 mM Tris-HCl (pH 7.5) 0.15 M NaCl, 10 g/L skim milk) and incubated with an antiserum dilution from about 1:50 to about 1:5000, preferably about 1:500. The reaction is revealed according to standard procedures. For example, a goat anti-rabbit peroxidase 20 conjugate is added to the wells when rabbit antibodies are used. Incubation is carried out 90 minutes at 37°C and the blot is washed. The reaction is developed with the appropriate substrate and stopped. The reaction is measured visually by the appearance of a colored spot, e.g., by 25 colorimetry. Under the above experimental conditions, a positive reaction is shown once a colored spot is associated with a dilution of at least about 1:5, preferably of at least about 1:500.

Therapeutic or prophylactic efficacy of a 30 polypeptide or derivative of the invention can be evaluated as described below. Another aspect of the invention provides (i) a composition comprising a polypeptide of the invention together with a diluent or carrier; specifically

(ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a polypeptide of the invention; (iii) a method for inducing an immune response against *Moraxella* in a mammal, by
5 administering to the mammal an immunogenically effective amount of a polypeptide of the invention to elicit a protective immune response to *Moraxella*; and particularly, (iv) a method for preventing and/or treating a *Moraxella* infection, by administering a prophylactic or therapeutic
10 amount of a polypeptide of the invention to an infected individual. Additionally, the invention encompasses the use of a polypeptide of the invention in the preparation of a medicament for preventing and/or treating *Moraxella* infection.

15 In one embodiment, the polypeptide, composition or treatment is free of adjuvant, specifically adjuvants commonly or specifically used in rodents. The polypeptide, composition or treatment may be used to treat disorders whose symptoms are caused or aggravated at least in part by
20 *Moraxella* infection, specifically *M. catarrhalis*, and includes such disorders as otitis media, respiratory infection, sinusitis and pneumonia. Preferred polypeptide compositions include those formulated for *in vivo* administration to an animal, preferably a human, to confer
25 protection or treatment against disease caused by *M. catarrhalis*. Also preferred are compositions formulated as a microparticle, capsule, or liposome.

As used herein, the immunogenic compositions of the invention are administered by conventional routes known
30 the vaccine field, in particular to a mucosal (e.g., ocular, intranasal, pulmonary, oral, gastric, intestinal, rectal, vaginal, or urinary tract) surface or *via* the parenteral (e.g., subcutaneous, intradermal, intramuscular,

intravenous, or intraperitoneal) route. The choice of administration route depends upon a number of parameters, such as the adjuvant associated with the polypeptide. If a mucosal adjuvant is used, the intranasal or oral route is preferred. If a lipid formulation or an aluminum compound is used, the parenteral route is preferred with the subcutaneous or intramuscular route being most preferred. The choice also depends upon the nature of the vaccine agent. For example, a polypeptide of the invention fused to CTB or LTB is best administered to a mucosal surface.

As used herein, the composition of the invention contains one or several polypeptides or derivatives of the invention. The composition optionally contains at least one additional Moraxella antigen, or a subunit, fragment, homolog, mutant, or derivative thereof.

For use in a composition of the invention, in one embodiment a polypeptide or derivative thereof is formulated into or with liposomes, preferably neutral or anionic liposomes, microspheres, ISCOMS, or virus-like-particles (VLPs) to facilitate delivery and/or enhance the immune response. These compounds are readily available to one skilled in the art; for example, see Liposomes: A Practical Approach, RCP New Ed, IRL press (1990).

In one embodiment, adjuvants other than liposomes and the like are also used and are known in the art. Adjuvants may protect the antigen from rapid dispersal by sequestering it in a local deposit, or they may contain substances that stimulate the host to secrete factors that are chemotactic for macrophages and other components of the immune system. An appropriate selection can conventionally be made by those skilled in the art, for example, from those described below.

Treatment is achieved in a single dose or repeated as necessary at intervals, as can be determined readily by one skilled in the art. For example, a priming dose is followed by three booster doses at weekly or monthly intervals. An appropriate dose depends on various parameters including the recipient (e.g., adult or infant), the particular vaccine antigen, the route and frequency of administration, the presence/absence or type of adjuvant, and the desired effect (e.g., protection and/or treatment), as can be determined by one skilled in the art. In general, a vaccine antigen of the invention is administered by a mucosal route in an amount from about 10 μ g to about 500 mg, preferably from about 1 mg to about 200 mg. For the parenteral route of administration, the dose usually does not exceed about 1 mg, preferably about 100 μ g.

When used as vaccine agents, polynucleotides and polypeptides of the invention may be used sequentially as part of a multistep immunization process. For example, a mammal is initially primed with a vaccine vector of the invention such as a pox virus, e.g., via the parenteral route, and then boosted twice with the polypeptide encoded by the vaccine vector, e.g., via the mucosal route. In another example, liposomes associated with a polypeptide or derivative of the invention is also used for priming, with boosting being carried out mucosally using a soluble polypeptide or derivative of the invention in combination with a mucosal adjuvant (e.g., LT).

A polypeptide derivative of the invention is also useful as a diagnostic reagent for detecting the presence of anti-Moraxella antibodies, e.g., in a blood sample. Such polypeptides are about 5 to about 80 amino acids, (for example, 10, 12, 15, 17, 20, 25, 30 ... and so forth up to 75 amino acids), and are preferably about 10 to about 50

amino acids in length. They are either labeled or unlabeled, depending upon the diagnostic method. Diagnostic methods involving such a reagent are described below.

Upon expression of a DNA molecule of the invention, a polypeptide or polypeptide derivative is produced and purified using known laboratory techniques. As described above, the polypeptide or polypeptide derivative may be produced as a fusion protein containing a fused tail that facilitates purification. The fusion product is used to immunize a small mammal, e.g., a mouse or a rabbit, in order to raise antibodies against the polypeptide or polypeptide derivative (monospecific antibodies). Accordingly, the invention also provides a monospecific antibody that binds to a polypeptide or polypeptide derivative of the invention.

By "monospecific antibody" is meant an antibody that is capable of reacting with a unique naturally-occurring Moraxella polypeptide. An antibody of the invention is either polyclonal or monoclonal. Monospecific antibodies may be recombinant, e.g., chimeric (e.g., constituted by a variable region of murine origin associated with a human constant region), humanized (a human immunoglobulin constant backbone together with hypervariable region of animal, e.g., murine, origin), and/or single chain. Both polyclonal and monospecific antibodies may also be in the form of immunoglobulin fragments, e.g., F(ab)'2 or Fab fragments. The antibodies of the invention are of any isotype, e.g., IgG or IgA, and polyclonal antibodies are of a single isotype or a mixture of isotypes.

Antibodies against the polypeptides, homologs or fragments of the present invention are generated by immunization of a mammal with a composition comprising said

polypeptide, homolog or fragment. Such antibodies may be polyclonal or monoclonal. Methods to produce polyclonal or monoclonal antibodies are well known in the art. For a review, see *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory, Eds. E. Harlow and D. Lane (1988), and D.E. Yelton et al., 1981. *Ann. Rev. Biochem.* 50:657-680. For monoclonal antibodies, see Kohler & Milstein (1975) *Nature* 256:495-497.

The antibodies of the invention, which are raised to a polypeptide or polypeptide derivative of the invention, are produced and identified using standard immunological assays, e.g., Western blot analysis, dot blot assay, or ELISA (Sambrook, J.E.F. et. al. (1989) *Molecular Cloning a Laboratory Manual* 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor N.Y.). The antibodies are used in diagnostic methods to detect the presence of a *Moraxella* antigen in a sample, such as a biological sample. The antibodies are also used in affinity chromatography for purifying a polypeptide or polypeptide derivative of the invention. As is discussed further below, such antibodies may be used in prophylactic and therapeutic passive immunization methods.

Accordingly, the invention also provides (i) a reagent for detecting the presence of *Moraxella* in a biological sample that contains an antibody, polypeptide, or polypeptide derivative of the invention; and (ii) a diagnostic method for detecting the presence of *Moraxella* in a biological sample, by contacting the biological sample with an antibody, a polypeptide, or a polypeptide derivative of the invention, such that an immune complex is formed, and by detecting such complex to indicate the presence of *Moraxella* in the sample or the organism from which the sample is derived.

Those skilled in the art will readily understand that the immune complex is formed between a component of the sample and the antibody, polypeptide, or polypeptide derivative, whichever is used, and that any unbound material
5 is removed prior to detecting the complex. It is understood that a polypeptide reagent is useful for detecting the presence of anti-Moraxella antibodies in a sample, e.g., a blood sample, while an antibody of the invention is used for screening a sample, such as a gastric extract or biopsy, for
10 the presence of Moraxella polypeptides.

For diagnostic applications, the reagent (i.e., the antibody, polypeptide, or polypeptide derivative of the invention) is either in a free state or immobilized on a solid support, such as a tube, a bead, or any other
15 conventional support used in the field. Immobilization is achieved using direct or indirect means. Direct means include passive adsorption (non-covalent binding) or covalent binding between the support and the reagent. By "indirect means" is meant that an anti-reagent compound that
20 interacts with a reagent is first attached to the solid support. For example, if a polypeptide reagent is used, an antibody that binds to it can serve as an anti-reagent, provided that it binds to an epitope that is not involved in the recognition of antibodies in biological samples.
25 Indirect means may also employ a ligand-receptor system, for example, where a molecule such as a vitamin is grafted onto the polypeptide reagent and the corresponding receptor immobilized on the solid phase. This is illustrated by the biotin-streptavidin system. Alternatively, a peptide tail
30 is added chemically or by genetic engineering to the reagent and the grafted or fused product immobilized by passive adsorption or covalent linkage of the peptide tail.

Such diagnostic agents may be included in a kit which also comprises instructions for use. The reagent is labeled with a detection means which allows for the detection of the reagent when it is bound to its target.

5 The detection means may be a fluorescent agent such as fluorescein isocyanate or fluorescein isothiocyanate, or an enzyme such as horse radish peroxidase or luciferase or alkaline phosphatase, or a radioactive element such as ^{125}I or ^{51}Cr .

10 Another aspect of the invention provides a process for purifying, from a biological sample, a polypeptide or polypeptide derivative of the invention, which involves carrying out antibody-based affinity chromatography with the biological sample, wherein the antibody is a monospecific
15 antibody of the invention.

For use in a purification process of the invention, the antibody is either polyclonal or monospecific, and preferably is of the IgG type. Purified IgGs is prepared from an antiserum using standard methods.
20 Conventional chromatography supports, as well as standard methods for grafting antibodies, are described in, e.g., Antibodies: A Laboratory Manual, D. Lane, E. Harlow, Eds. (1988) and outlined below.

Briefly, a biological sample, such as an *M. catarrhalis* extract preferably in a buffer solution, is
25 applied to a chromatography material, preferably equilibrated with the buffer used to dilute the biological sample so that the polypeptide or polypeptide derivative of the invention (i.e., the antigen) is allowed to adsorb onto
30 the material. The chromatography material, such as a gel or a resin coupled to an antibody of the invention, is in either a batch form or a column. The unbound components are

washed off and the antigen is then eluted with an appropriate elution buffer, such as a glycine buffer or a buffer containing a chaotropic agent, e.g., guanidine HCl, or high salt concentration (e.g., 3 M MgCl₂). Eluted
5 fractions are recovered and the presence of the antigen is detected, e.g., by measuring the absorbance at 280 nm.

The invention also provides (i) a composition of matter comprising a monospecific antibody of the invention, together with a diluent or carrier; (ii) a pharmaceutical
10 composition comprising a therapeutically or prophylactically effective amount of a monospecific antibody of the invention, and (iii) a method for treating or preventing a Moraxella infection, by administering a therapeutic or prophylactic amount of a monospecific antibody of the
15 invention to an infected individual. Additionally, the invention encompasses the use of a monospecific antibody of the invention in the preparation of a medicament for treating or preventing Moraxella infection.

In one embodiment, the monospecific antibody,
20 composition or treatment may be used to treat disorders whose symptoms are caused or aggravated at least in part by Moraxella infection, specifically *M. catarrhalis*, and includes such disorders as otitis media, respiratory infection, sinusitis and pneumonia. Preferred antibody
25 compositions include those formulated for *in vivo* administration to an animal, preferably a human, to confer protection or treatment against disease caused by *M. catarrhalis*. Also preferred are compositions formulated as a microparticle, capsule, or liposome.

30 The monospecific antibody is either polyclonal or monoclonal, preferably of the IgA isotype (predominantly). In passive immunization, the antibody is administered to a

mucosal surface of a mammal, e.g., the gastric mucosa, e.g., orally or intragastrically, advantageously, in the presence of a bicarbonate buffer. Alternatively, systemic administration, not requiring a bicarbonate buffer, is carried out. A monospecific antibody of the invention is administered as a single active component or as a mixture with at least one monospecific antibody specific for a different *Moraxella* polypeptide. The amount of antibody and the particular regimen used are readily determined by one skilled in the art. For example, daily administration of about 100 to 1,000 mg of antibodies over one week, or three doses per day of about 100 to 1,000 mg of antibodies over two or three days, are effective regimens for most purposes.

Therapeutic or prophylactic efficacy are evaluated using standard methods in the art, e.g., by measuring induction of a mucosal immune response or induction of protective and/or therapeutic immunity, using, e.g., the *M. catarrhalis* bactericidal assay. Those skilled in the art will readily recognize that the *M. catarrhalis* strain of the model may be replaced with another *Moraxella* strain. For example, the efficacy of DNA molecules and polypeptides from *M. catarrhalis* is preferably evaluated in a model using *M. catarrhalis* strain. Protection is determined by comparing the degree of *Moraxella* infection to that of a control group. Protection is shown when infection is reduced by comparison to the control group. Such an evaluation is made for polynucleotides, vaccine vectors, polypeptides and derivatives thereof, as well as antibodies of the invention.

Adjuvants useful in any of the vaccine compositions described above are as follows.

Adjuvants for parenteral administration include aluminum compounds, such as aluminum hydroxide, aluminum

phosphate, and aluminum hydroxy phosphate. The antigen is precipitated with, or adsorbed onto, the aluminum compound according to standard protocols. Other adjuvants, such as RIBI (ImmunoChem, Hamilton, MT), are used in parenteral
5 administration.

Adjuvants for mucosal administration include bacterial toxins, e.g., the cholera toxin (CT), the *E. coli* heat-labile toxin (LT), the *Clostridium difficile* toxin A and the pertussis toxin (PT), or combinations, subunits,
10 toxoids, or mutants thereof such as a purified preparation of native cholera toxin subunit B (CTB). Fragments, homologs, derivatives, and fusions to any of these toxins are also suitable, provided that they retain adjuvant activity. Preferably, a mutant having reduced toxicity is
15 used. Suitable mutants are described, e.g., in WO 95/17211 (Arg-7-Lys CT mutant), WO 96/06627 (Arg-192-Gly LT mutant), and WO 95/34323 (Arg-9-Lys and Glu-129-Gly PT mutant). Additional LT mutants that are used in the methods and compositions of the invention include, e.g., Ser-63-Lys,
20 Ala-69Gly, Glu-110-Asp, and Glu-112-Asp mutants. Other adjuvants, such as a bacterial monophosphoryl lipid A (MPLA) of, e.g., *E. coli*, *Salmonella minnesota*, *Salmonella typhimurium*, or *Shigella flexneri*; saponins, or polylactide glycolide (PLGA) microspheres, is also be used in mucosal
25 administration.

Adjuvants useful for both mucosal and parenteral administrations include polyphosphazene (WO 95/02415), DC-chol (3 b-(N-(N',N'-dimethyl aminomethane)-carbamoyle) cholesterol; U.S. Patent No. 5,283,185 and WO 96/14831) and
30 QS-21 (WO 88/09336).

Any pharmaceutical composition of the invention containing a polynucleotide, a polypeptide, a polypeptide

derivative, or an antibody of the invention, is manufactured in a conventional manner. In particular, it is formulated with a pharmaceutically acceptable diluent or carrier, e.g., water or a saline solution such as phosphate buffer saline.

5 In general, a diluent or carrier is selected on the basis of the mode and route of administration, and standard pharmaceutical practice. Suitable pharmaceutical carriers or diluents, as well as pharmaceutical necessities for their use in pharmaceutical formulations, are described in
10 *Remington's Pharmaceutical Sciences*, a standard reference text in this field and in the USP/NF.

The invention also includes methods in which Moraxella infection are treated by oral administration of a Moraxella polypeptide of the invention and a mucosal
15 adjuvant, in combination with an antibiotic, an antacid, sucralfate, or a combination thereof. Examples of such compounds that can be administered with the vaccine antigen and the adjuvant are antibiotics, including, e.g., macrolides, tetracyclines, and derivatives thereof (specific
20 examples of antibiotics that can be used include azithromycin or doxycyclin or immunomodulators such as cytokines or steroids). In addition, compounds containing more than one of the above-listed components coupled together, are used. The invention also includes
25 compositions for carrying out these methods, i.e., compositions containing a Moraxella antigen (or antigens) of the invention, an adjuvant, and one or more of the above-listed compounds, in a pharmaceutically acceptable carrier or diluent.

Amounts of the above-listed compounds used in the methods and compositions of the invention are readily determined by one skilled in the art.

Treatment/immunization schedules are also known and readily
5 designed by one skilled in the art. For example, the non-vaccine components can be administered on days 1-14, and the vaccine antigen + adjuvant can be administered on days 7, 14, 21, and 28.

EXAMPLES

10 The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples. These examples are described solely for purposes of illustration and are not intended to limit the scope of the
15 invention. Changes in form and substitution of equivalents are contemplated as circumstances may suggest or render expedient. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitation.

20 Example 1: Amplification and cloning of *Moraxella* open reading frames.

All the open reading frames disclosed were PCR amplified from the ATCC 43617 strain of *Moraxella catarrhalis*. Genomic DNA was obtained from the ATCC
25 *Moraxella* strain using standard techniques (Maniatis et al., Molecular Cloning: A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). PCR amplification was performed using the 5' and 3' primers shown in Table 1 for each of the open reading frames (ORFs).
30 The PCR conditions were standard conditions for primers of

approximately 40 bases in length (see Table 1). The 5' ends of the 5' primers were synthesized to include an Nde I restriction site, while the 3' ends of the 3' primers contained an Xho I restriction site. Once the genes have
5 been amplified, the PCR products were cloned into the plasmid pCR2.1 (Invitrogen) according to the manufacturer's protocol.

The NdeI - XhoI fragment containing the ORF from pCR2.1 was cloned into the pET-24 (Novogen), which was
10 previously restricted with NdeI and XhoI. The pET-24 vector provides the bacteriophage T7 promoter and a C-terminal His-Tag sequence so that the recombinantly expressed protein has a His tag which can be used for purification. After the cloning of each ORF, the 5' and 3' DNA sequence was verified
15 by sequencing these ends. The basic construction scheme is outlined in Figure 55.

In some cases it was more convenient to PCR-amplify the ORF in two parts, a 5' part and 3' part. ORFs amplified in this fashion were then re-assembled by cloning
20 both parts in-frame into pCR2.1.

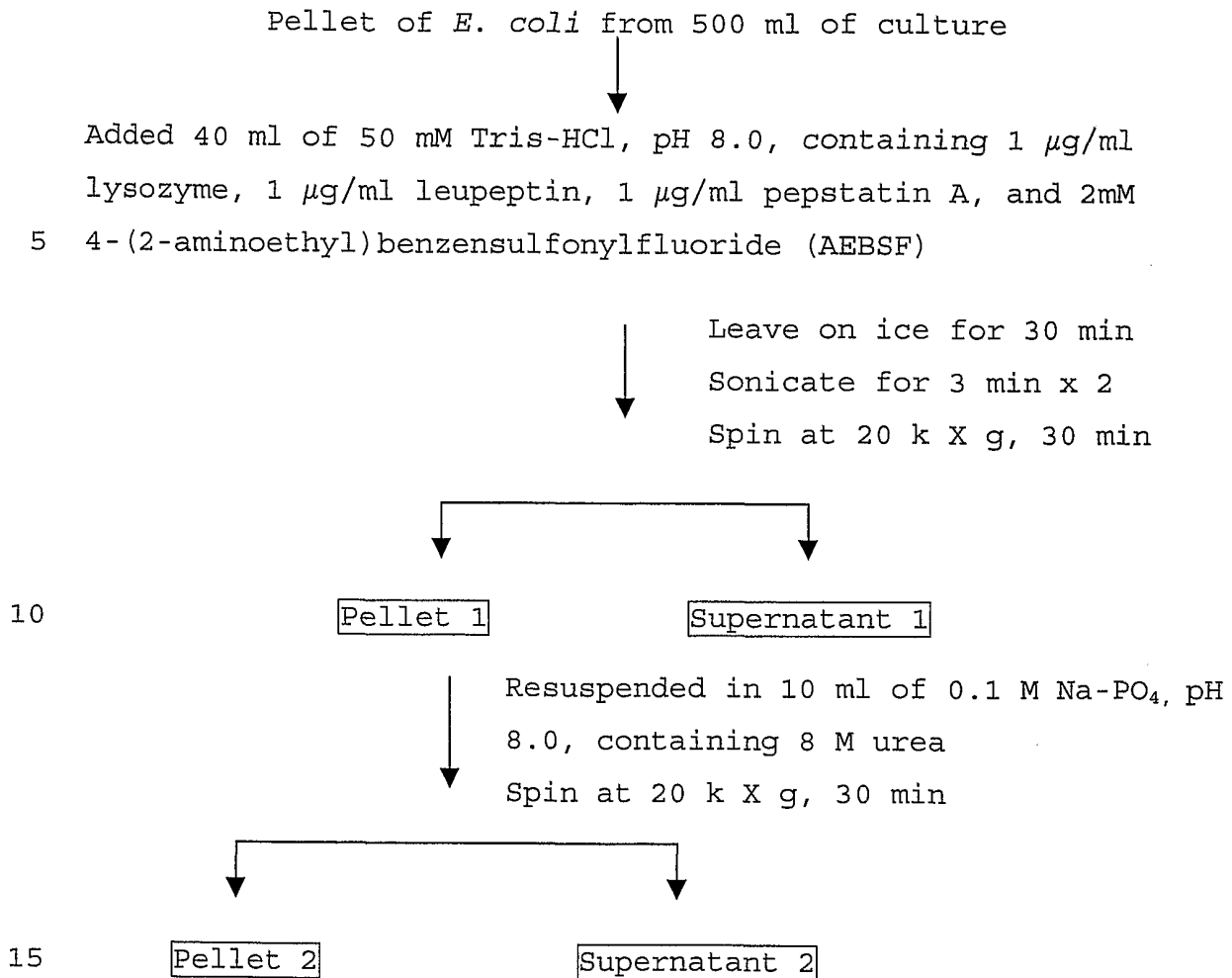
Table 1

Gene	Primer Sequence	Length	SEQ ID NO.
MCA100170	gggaattcatatgagcaatcaagcggacaaagccgcccagcca	43	55
MCA100170	gggatccctcgaggtttggcttgggttttttggcttttgccac	43	56
MCA100412	gggaattcatatgagtcctcaaaaaccgtacaatacacctgc	40	59
MCA100412	gggatccctcgagtccttgggttgagctttttgaacaaagtttt	43	60
MCA100413	gggaattcatatgcaagttgaccatgctcaagaagctcatgat	43	61
MCA100413	gggatccctcgagaagttttgtaacatatctgacattacagcc	43	62
MCA100414	gggaattcatatgtgtgcaaatacaggcacaactggcaatggc	43	63
MCA100414	gggatccctcgagcataccctgcggtgctaaaatcatcagctc	43	64
MCA100800	gggaattcatatgtgcaaaagcgcgattggactgtttgaccgc	43	69
MCA100800	gggatccctcgaggcgtggtattttgatttaattgtccagcgtc	43	70
MCA100992	gggaattcatatgcatatcgatcatgtgtctgttaccctaaaca	43	75
MCA100992	gggatccctcgagtccttaccattttgggtgtaataaaaaatcaa	43	76
MCA101249	gggaattcatatgcaacttttgtgggcaaaagcgtatcggtcat	43	77
MCA101249	gggatccctcgagcatatttgcaccaaagactttcgtttggg	43	78
MCA101896	gggaattcatatgtgcagcaatcaatcaaacgaaccagctgcc	43	85
MCA101896	gggatccctcgagccagcctttgataacaccatctttaaattg	43	86
MCA101806	gggaattcatatgggacgcttgggtgcagattggcaaaaacagt	43	81
MCA101806	gggatccctcgagaaatgccatttgtgcaccacaaaatccacc	43	82
MCA101862	gggaattcatatgtcaaaaaaacccaaaaccaccaagat	40	83
MCA101862	gggatccctcgagaaaaattcggttcatccatttaaggcagc	43	84
MCA101915	gggaattcatatgcagattgcaccagacccaaccaaagcaa	43	87
MCA101915	gggatccctcgagaaactcaatgctggcactcacacccacgac	43	88
MCA102062	gggaattcatatgaatttaaaggataagccaaccgtcatttta	43	91
MCA102062	gggatccctcgagaaatttgaaatttgccccgagcataaaagt	43	92
MCA102070	gggaattcatatggcacctgaattatatggttacgcaaatctt	43	93
MCA102070	gggatccctcgagaaaaatatacagtcgcatagatagttgggtt	43	94
MCA100178	gggaattcatatgaattatgcgacctactatggcgataaaattc	43	57
MCA100178	gggatccctcgagcaatctcctaattgttactttaattttacc	43	58
MCA100527	gggaattcatatgttgacttatgccaaatgcggtgcaacggtt	43	65
MCA100527	gggatccctcgagagcgttttttgggttggaatcaatcacttg	43	66
MCA100854	gggaattcatatggatcatacggtagcgtttaatgcaaccacg	43	71
MCA100854	gggatccctcgagggttttgtgcaacaggtgttggtgcatcttc	43	72
MCA100889	gggaattcatatgattaccatcaaaaaaggcttggatctg	40	73
MCA100889	gggatccctcgaggccttctctgctcaatacgggtcaagttatc	43	74
MCA101446	gggaattcatatgaccgataagaccgacgatgttaagcca	40	79
MCA101446	gggatccctcgagattgaaaacagtgatatcaaatatcacact	43	80
MCA101924	gggaattcatatggttttggatgagtggttttgggtgctggt	40	89
MCA101924	gggatccctcgagaccatgcttaagctgacttgggcgaaaacc	43	90
MCA100712	gggaattcatatgctgacaaaaaagatgccaaacagctt	40	67
MCA100712	gggatccctcgagtagaccatccagtcgttaagcataagttcc	43	68

Example 2: Expression and purification of recombinant proteins from *Moraxella*.

Once the pET-24 plasmids containing the cloned ORFs have been verified as containing the correct sequence, they were introduced into electrocompetent *E. coli* BL21(DE3) cells using a BioRad electroporator. The resulting recombinant *E. coli* strains were then grown for protein expression. A representative isolated colony that expressed the recombinant protein as judged by expression on an SDS-PAGE gel, was then picked for large-scale growth for protein purification. To express the recombinant proteins, the transformed *E. coli* cells were grown at 37°C in appropriate media (i.e. NZCYM) with the appropriate antibiotic (i.e. Kanamycin) to an absorbance at 578nm of approximately 0.3 (A₅₇₈), before inducing expression. For induction, lactose was added to a final concentration of 1% and the cells were allowed to grow for another 4 hours.

For protein purification studies, 500 ml cultures of *E. coli* were grown as above. The cells were then harvested by centrifugation. Different purification techniques were used for proteins expressed in a soluble form or in inclusion bodies. The extraction of *M. catarrhalis* antigens involved using Tris-HCl buffer containing various protease inhibitors. The proteins that were expressed as inclusion bodies required urea for solubilization as shown below.



The soluble proteins were purified under native conditions, and insoluble proteins were purified under denatured conditions. The procedure for purification under native conditions and denatured conditions is shown in Figures 56 and 57, respectively. The procedure from Figure 57 was also used when a Cu⁺² HiTrap affinity column was used.

Results from purification of the ORFs are shown in Table 2.

Table 2. Summary of purified *Moraxella catarrhalis* antigens

Locus ID	SEQ ID NO.	Name	M.W. (kDa)	Solubility	Antibodies produced
MCA100170	2	Major anaerobically Induced protein	51.2	Soluble	Yes
MCA100178	4	Rare lipoprotein homolog	10.3	Inclusion Body	Yes
MCA100412	6	Adhesin component	10.4	Inclusion Body	No
MCA100413	8	Adhesin component	12.8	Inclusion Body	Yes
MCA100414	10	Outer membrane protein	21.5	Soluble	Yes
MCA100527	12	Putative oxido-reductase, Internalin B	24.1	Inclusion Body	N/A
MCA100564	14	OprM protein	53.1	Inclusion Body	Yes
MCA100567	16	Collagenase precursor	49.8	Inclusion Body	Yes
MCA100800	20	Lipoprotein	37.7	Soluble	Yes
MCA100854	22	OmpA, OprF	22.2	Soluble	Yes
MCA100992	26	PilQ fimbrial assembly Protein	48.8	Inclusion Body	N/A
MCA101249	28	Invasion protein	17.4	Inclusion Body	Yes
MCA101388	30	Rare lipoprotein A	15.8	Inclusion Body	Yes
MCA101446	32	36 kDa secreted protein	42.2	Inclusion Body	No
MCA101792	38	Outer membrane protein	20.7	Inclusion Body	Yes
MCA101806	40	Outer membrane esterase	64.3	Inclusion Body	Yes
MCA101862	42	Soluble pyridine nucleotide transhydrogenase	58.3	Inclusion Body	No
MCA101896	44	Putative lipoprotein adhesin	27.7	Soluble	Yes
MCA101915	46	HecB	73.1	Inclusion Body	Yes
MCA101976	50	Heme receptor	100.6	Inclusion Body	Yes
MCA102062	52	ompC, heme receptor	88.9	Inclusion Body	Yes
MCA102070	54	Porin	35.8	Inclusion Body	Yes

Example 3: Immunogenicity of purified recombinant *Moraxella* proteins.

Purified recombinant proteins were used to generate antibodies to the proteins. Proteins (see Table 3) from the ORFs of *Moraxella catarrhalis* were separately used to immunize guinea pigs. Three to four guinea pigs were immunized with 5 µg of purified recombinant protein in complete Freund's adjuvant (CFA). Animals were boosted twice at two weeks intervals with a further 5 µg of the recombinant protein in CFA. After the second boost, animals were bled and the sera were pooled.

Pooled sera were tested using Western blots (see Figure 59 for examples of blots; not all results shown in Table 3 are included in Figure 59) under standard conditions. SDS-PAGE gels were run with 1µg of purified protein and/or whole-cell lysates from stationary phase and logarithmic phase *Moraxella catarrhalis* strain ATCC 43617, grown in brain heart infusion broth. The SDS-PAGE gels were transferred to membranes and probed with the pooled rabbit polyclonal sera. All proteins tested were immunogenic in rabbits and most were able to recognize the native protein in whole-cell lysates (see summary of results in Table 3). The polyclonal anti-sera can also be used to assess functionality of the antibodies in the antibody-dependent bactericidal assay, such as described in Yang et al. FEMS Immunol. Med. Microbiol. (1997) 17:187-199.

Table 3

Antigen	SEQ ID NO.	Antibody binding		
		Purified antigen	Stationary phase lysate	Logarithmic phase lysate
MCA 100170	2	+	+	+
MCA 100414	10	+	+	+
MCA 100567	16	+	-	+
MCA 100854	22	+	+	+
MCA 101249	28	+	-	-
MCA 101915	46	+	-	-
MCA 101976	50	+	-	+
MCA 100800	20	+	+	+
MCA 101806	40	+	+	+
MCA 100178	4	+	+	+
MCA 100413	8	+	+	+
MCA 102062	52	+	+	+
MCA 100564	14	+	+	+
MCA 101388	30	+	-	-
MCA 101792	38	+	-	-
MCA 101896	44	+	+	+
MCA 102070	54	+	-	-

Throughout this application, various references
5 are referred to in parenthesis to describe more fully the
state of the art to which this invention pertains. The
disclosures of these references are hereby incorporated by
reference into the present disclosure.

CLAIMS:

1. An isolated nucleic acid molecule comprising a nucleic acid sequence which encodes a polypeptide selected from any one of:

5 (a) SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54;

(b) a fragment comprising at least 12 consecutive amino acids from a polypeptide of (a), wherein the fragment
10 is capable of eliciting an immunogenic response in a mammal against a bacterium of the Moraxella genus; and

(c) a polypeptide which is at least 75% identical in amino acid sequence to the corresponding polypeptide of (a) or (b), and which is capable of eliciting an immunogenic
15 response in a mammal against a bacterium of the Moraxella genus.

2. A nucleic acid molecule comprising a nucleic acid sequence selected from any one of:

(a) SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19,
20 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53;

(b) a sequence which encodes a polypeptide encoded by any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51
25 and 53;

(c) a sequence comprising at least 38 consecutive nucleotides from any one of the nucleic acid sequences of (a) and (b); and

(d) a sequence which encodes a polypeptide which is at least 75% identical in amino acid sequence to the polypeptide encoded by any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53.

3. A nucleic acid molecule comprising a nucleic acid sequence which is complementary to the nucleic acid molecule of claim 1 or 2.

4. A nucleic acid molecule comprising a nucleic acid sequence which encodes a fusion protein, said fusion protein comprising a polypeptide encoded by a nucleic acid molecule according to claim 1 or 2 and a second polypeptide.

5. The nucleic acid molecule of claim 4 wherein the second polypeptide is a heterologous signal peptide.

6. The nucleic acid molecule of claim 4 wherein the second polypeptide has adjuvant activity.

7. A nucleic acid molecule according to any one of claims 1 to 6, operatively linked to one or more expression control sequences.

8. A host organism comprising an isolated nucleic acid selected from any one of:

(i) any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53;

(ii) a nucleic acid sequence which encodes a polypeptide encoded by any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53;

(iii) a nucleic acid sequence comprising at least 38 consecutive nucleotides from any one of the nucleic acid sequences of (i) and (ii);

(iv) a nucleic acid sequence which encodes a polypeptide which is at least 75% identical in amino acid sequence to the polypeptide encoded by any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53;

(v) a nucleic acid sequence which encodes a polypeptide whose sequence is set forth in any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54;

(vi) a nucleic acid sequence which encodes at least 12 consecutive amino acids from any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54;

wherein the isolated nucleic acid is capable of being expressed and wherein the host organism optionally comprises a second isolated nucleic acid encoding and capable of expressing an additional polypeptide.

9. The host organism of claim 8 wherein the polypeptide encoded by the isolated nucleic acid is part of a fusion protein which further comprises a second polypeptide.

10. The host organism of claim 9 wherein the second polypeptide is a heterologous signal peptide.

11. The host organism of claim 9 wherein the second polypeptide has adjuvant activity.

12. The host organism of any one of claims 8 to 11 wherein each isolated nucleic acid is operatively linked to one or more expression control sequences.

13. The host organism of any one of claims 8 to 12
5 wherein the host further comprises the second isolated nucleic acid and wherein the additional polypeptide is a *Moraxella* polypeptide.

14. The host organism of any one of claims 8 to 13 which is a virus.

10 15. The host organism of claim 14 which is selected from the group consisting of: adenovirus, alphavirus, and poxvirus.

16. The host organism of claim 15 which is a poxvirus.

17. The host organism of claim 16 which is a vaccinia
15 or canary pox virus.

18. The host organism of any one of claims 8 to 13 which is a bacterium.

19. The host organism of claim 18 which is selected from the group consisting of: *E. coli*, *Shigella*, *Salmonella*,
20 *Vibrio cholerae*, *Lactobacillus*, Bacille bilié de Calmette-Guérin (BCG), and *Streptococcus*.

20. The host organism of claim 19 which is a non-toxicogenic *Vibrio cholerae* mutant strain.

21. The host organism of claim 19 which is an
25 attenuated *Salmonella typhimurium* strain.

22. A pharmaceutical composition comprising a nucleic acid according to any one of claims 1 to 7 and a pharmaceutically acceptable carrier.

23. The composition of claim 22 further comprising a delivery agent selected from the group consisting of: bupivacaine, liposome and cationic lipid.

24. The composition of claim 22 or 23 further
5 comprising an adjuvant.

25. An isolated oligonucleotide of 5 to 100 nucleotides which hybridizes under stringent conditions to the nucleic acid molecule of any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37,
10 39, 41, 43, 45, 47, 49, 51 and 53, or to a complementary or anti-sense sequence of said nucleic acid molecule.

26. An isolated oligonucleotide of 10 to 40 nucleotides which hybridizes under stringent conditions to the nucleic acid molecule of any one of SEQ ID NOs: 1, 3, 5,
15 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 29, 31 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53, or to a complementary or anti-sense sequence of said nucleic acid molecule.

27. A polypeptide encoded by a nucleic acid sequence according to any one of claims 1, 2 and 4 to 7.

20 28. A polypeptide comprising an amino acid sequence selected from any one of:

(a) SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54;

25 (b) a fragment comprising at least 12 consecutive amino acids from a polypeptide of (a), wherein the fragment is capable of eliciting an immunogenic response in a mammal against a bacterium of the Moraxella genus; and

(c) a polypeptide which is at least 75% identical in amino acid sequence to the corresponding polypeptide of (a) or (b), and which is capable of eliciting an immunogenic response in a mammal against a bacterium of the *Moraxella* genus.

29. A fusion protein comprising a polypeptide of claim 27 or 28 and a second polypeptide.

30. The fusion protein of claim 29 wherein the second polypeptide is a heterologous signal peptide.

10 31. The fusion protein of claim 29 wherein the second polypeptide has adjuvant activity.

32. A method for producing a polypeptide of claim 27 or 28, or a fusion protein of any one of claims 29 to 31, comprising the step of culturing the host organism of any one of claims 13 to 21.

33. An antibody against the polypeptide of claim 27 or 28, or against the fusion protein of any one of claims 29 to 31.

34. A pharmaceutical composition comprising the polypeptide of claim 27 or 28 and a pharmaceutically acceptable carrier or diluent.

35. The pharmaceutical composition of claim 34 further comprising an additional *Moraxella* polypeptide.

36. A pharmaceutical composition comprising the fusion protein of any one of claims 29 to 31 and a pharmaceutically acceptable carrier or diluent.

37. A pharmaceutical composition comprising the antibody of claim 33 and a pharmaceutically acceptable carrier or diluent.

38. The composition of any one of claims 34 to 37 further comprising liposomes.

39. The composition of any one of claims 34 to 38 further comprising an adjuvant.

5 40. A method of detecting Moraxella infection comprising the step of contacting a body fluid of a mammal to be tested, with a component selected from any one of:

(a) the nucleic acid of any one of claims 1 to 7;

(b) the probe of claim 25;

10 (c) the primer of claim 26;

(d) the polypeptide of claim 27 or 28;

(e) the fusion protein of any one of claims 29 to 31; and

(f) the antibody of claim 33.

15 41. A diagnostic kit comprising instructions for use in detecting the presence of Moraxella, and a component selected from any one of:

(a) the nucleic acid of any one of claims 1 to 7;

(b) the probe of claim 25;

20 (c) the primer of claim 26;

(d) the polypeptide of claim 27 or 28;

(e) the fusion protein of any one of claims 29 to 31; and

(f) the antibody of claim 33.

42. A method for generating antibodies specific for *Moraxella*, comprising introducing into an animal a composition comprising any one of:

(a) the nucleic acid of any one of claims 1 to 7;

5 (b) the polypeptide of claim 27 or 28; and

(c) the fusion protein of any one of claims 29 to 31.

1/189

Figure 1. Sequence of *M. catarrhalis* major anaerobically induced OMP gene designated MCA100170

```

ttaagtgcct tttctaggta ttttaacttt tttgtgcttt ttatttttat tgatggcatt 60
gaagaaaatt acaggcaacc tttctataag gagaacccat atg tct aag cct act 115
                                         Met Ser Lys Pro Thr
                                         1           5

ttg ata aaa aca acc tta att tgt gcc tta agt gca ttg atg ctc agt 163
Leu Ile Lys Thr Thr Leu Ile Cys Ala Leu Ser Ala Leu Met Leu Ser
                        10                15                20

ggg tgt agc aat caa gcg gac aaa gcc gcc cag cca aaa agc agc acg 211
Gly Cys Ser Asn Gln Ala Asp Lys Ala Ala Gln Pro Lys Ser Ser Thr
                        25                30                35

gta gac gct gcc gcc aag aca gca aat gca gat aat gct gcc tca caa 259
Val Asp Ala Ala Ala Lys Thr Ala Asn Ala Asp Asn Ala Ala Ser Gln
                        40                45                50

gaa cat caa ggc gag ctg cct gtc att gat gcc att gtt acg cat gca 307
Glu His Gln Gly Glu Leu Pro Val Ile Asp Ala Ile Val Thr His Ala
                        55                60                65

cca gaa gtt cca cca cct gtt gac cgt gac cac ccc gcc aaa gtg gtg 355
Pro Glu Val Pro Pro Pro Val Asp Arg Asp His Pro Ala Lys Val Val
                        70                75                80                85

gta aaa atg gaa acc gtt gaa aaa gtc atg cgt ctg gca gat ggc gtg 403
Val Lys Met Glu Thr Val Glu Lys Val Met Arg Leu Ala Asp Gly Val
                        90                95                100

gaa tat cag ttt tgg aca ttt ggc ggt caa gtt cca ggg cag atg att 451
Glu Tyr Gln Phe Trp Thr Phe Gly Gly Gln Val Pro Gly Gln Met Ile
                        105                110                115

cgt gtg cgt gaa ggc gac acc atc gaa gtg cag ttc tca aac cac cca 499
Arg Val Arg Glu Gly Asp Thr Ile Glu Val Gln Phe Ser Asn His Pro
                        120                125                130

gat tca aaa atg ccc cat aat gtt gac ttt cac gct gcc aca ggg cct 547
Asp Ser Lys Met Pro His Asn Val Asp Phe His Ala Ala Thr Gly Pro
                        135                140                145

ggc ggc ggg gca gaa gcg tca ttt acc gca ccg ggt cat aca tca acc 595
Gly Gly Gly Ala Glu Ala Ser Phe Thr Ala Pro Gly His Thr Ser Thr
                        150                155                160                165

ttt agt ttt aaa gcc tta cag cct ggt ttg tat gtc tat cac tgt gcg 643
Phe Ser Phe Lys Ala Leu Gln Pro Gly Leu Tyr Val Tyr His Cys Ala
                        170                175                180

gtt gcc cct gtt ggc atg cac att gct aat ggc atg tat ggt ttg att 691
Val Ala Pro Val Gly Met His Ile Ala Asn Gly Met Tyr Gly Leu Ile
                        185                190                195

```

2/189

ttg gtt gaa cca aaa gag ggc ttg cca aaa gta gat aaa gaa tac tat	739
Leu Val Glu Pro Lys Glu Gly Leu Pro Lys Val Asp Lys Glu Tyr Tyr	
200 205 210	
gtc atg caa ggc gac ttt tat acc aaa ggc aaa tat ggc gaa caa ggt	787
Val Met Gln Gly Asp Phe Tyr Thr Lys Gly Lys Tyr Gly Glu Gln Gly	
215 220 225	
cta cag ccc ttt gat atg gaa aaa gcc att cga gaa gat gct gaa tat	835
Leu Gln Pro Phe Asp Met Glu Lys Ala Ile Arg Glu Asp Ala Glu Tyr	
230 235 240 245	
gtt gtc ttt aat ggt tcg gtg ggg gcg ttg act ggt gaa aat gct cta	883
Val Val Phe Asn Gly Ser Val Gly Ala Leu Thr Gly Glu Asn Ala Leu	
250 255 260	
aaa gcc aag gtt ggc gaa act gtt cgc tta ttt gtg ggt aac ggc ggc	931
Lys Ala Lys Val Gly Glu Thr Val Arg Leu Phe Val Gly Asn Gly Gly	
265 270 275	
ccg aat ttg aca tca tca ttc cat gtc att ggt gag att ttt gat aag	979
Pro Asn Leu Thr Ser Ser Phe His Val Ile Gly Glu Ile Phe Asp Lys	
280 285 290	
gtt cac ttt gag ggt ggt aag ggt gaa aac cac aat atc caa acc acg	1027
Val His Phe Glu Gly Gly Lys Gly Glu Asn His Asn Ile Gln Thr Thr	
295 300 305	
cta atc cca gca ggt ggc gct gcc atc act gaa ttt aag gtg gat gtg	1075
Leu Ile Pro Ala Gly Gly Ala Ala Ile Thr Glu Phe Lys Val Asp Val	
310 315 320 325	
ccg ggt gat tat gtc ttg gtt gac cat gcc atc ttc cgt gcc ttt aac	1123
Pro Gly Asp Tyr Val Leu Val Asp His Ala Ile Phe Arg Ala Phe Asn	
330 335 340	
aaa ggg gca ttg ggc ata ctt aag gtg gaa ggt gaa gaa aat cat gag	1171
Lys Gly Ala Leu Gly Ile Leu Lys Val Glu Gly Glu Glu Asn His Glu	
345 350 355	
att tat tca cac aaa caa aca gac gct gtc tat ctg cca gag ggt gcc	1219
Ile Tyr Ser His Lys Gln Thr Asp Ala Val Tyr Leu Pro Glu Gly Ala	
360 365 370	
cca caa gca att gat acc caa gaa gca ccc aaa aca cct gca cct gcc	1267
Pro Gln Ala Ile Asp Thr Gln Glu Ala Pro Lys Thr Pro Ala Pro Ala	
375 380 385	
aac tta caa gag cag att aaa gca ggt aag gca acc tat gac tct aac	1315
Asn Leu Gln Glu Gln Ile Lys Ala Gly Lys Ala Thr Tyr Asp Ser Asn	
390 395 400 405	
tgt gct gct tgt cac caa cct gat ggt aaa ggc gtg cca aac gct ttc	1363
Cys Ala Ala Cys His Gln Pro Asp Gly Lys Gly Val Pro Asn Ala Phe	
410 415 420	

3/189

```

cca ccg ctt gcc aac tct gac tat ctg aac gcc gac cac gct cgt gcc 1411
Pro Pro Leu Ala Asn Ser Asp Tyr Leu Asn Ala Asp His Ala Arg Ala
      425                      430                      435

gcc agc atc gtg gca aat gga ttg tct ggt aag att acc gtc aat ggc 1459
Ala Ser Ile Val Ala Asn Gly Leu Ser Gly Lys Ile Thr Val Asn Gly
      440                      445                      450

aac caa tat gaa agc gtc atg cct gcg att gct ctg agc gac caa cag 1507
Asn Gln Tyr Glu Ser Val Met Pro Ala Ile Ala Leu Ser Asp Gln Gln
      455                      460                      465

att gcc aat gtc atc acc tac acg ctt aac agc ttt ggt aac aaa ggc 1555
Ile Ala Asn Val Ile Thr Tyr Thr Leu Asn Ser Phe Gly Asn Lys Gly
470                      475                      480                      485

ggc caa ctc agt gca gac gat gtg gca aaa gcc aaa aaa acc aag cca 1603
Gly Gln Leu Ser Ala Asp Asp Val Ala Lys Ala Lys Lys Thr Lys Pro
      490                      495                      500

aac tga tcttaaagtt taattaagat atccgctggg tttcattgag catcatggct 1659
Asn

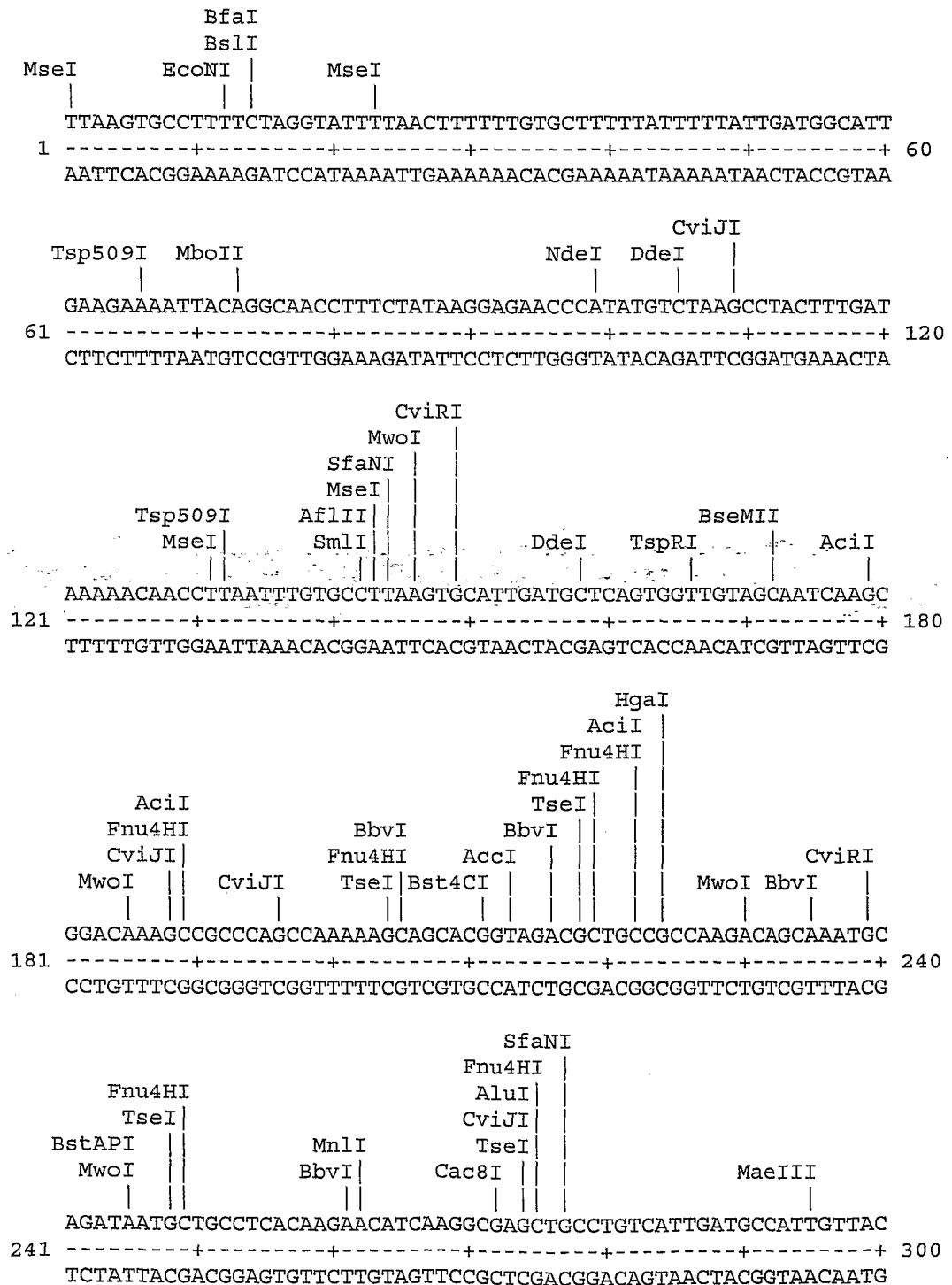
gtcttatgtg gcaagccatg atgtgcttaa tggctttcat gcgttttatg cccaatgatg 1719

cccaaagtgt gtgttttaaaa gccatgccaa aaggtgtatg 1759

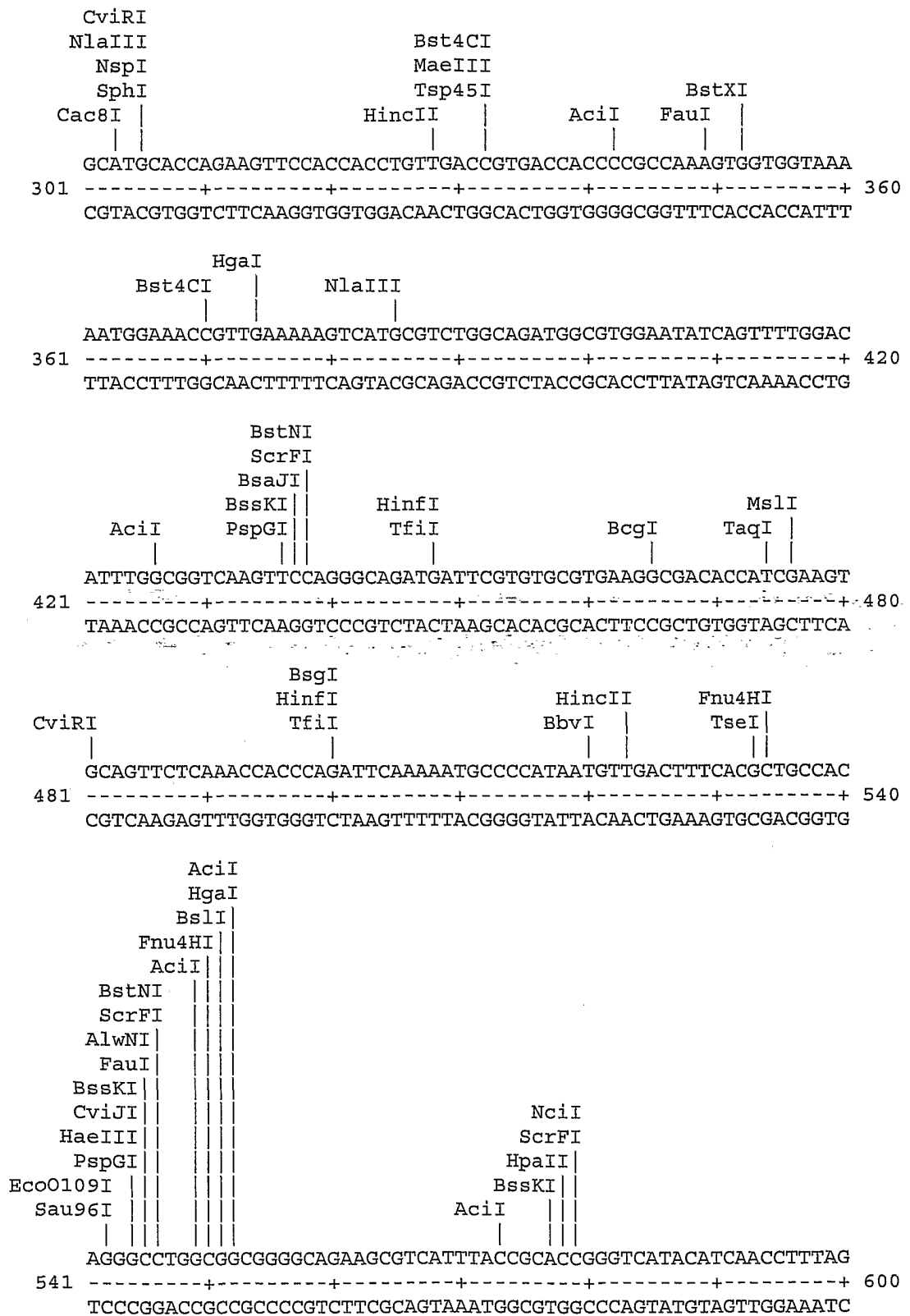
```

4/189

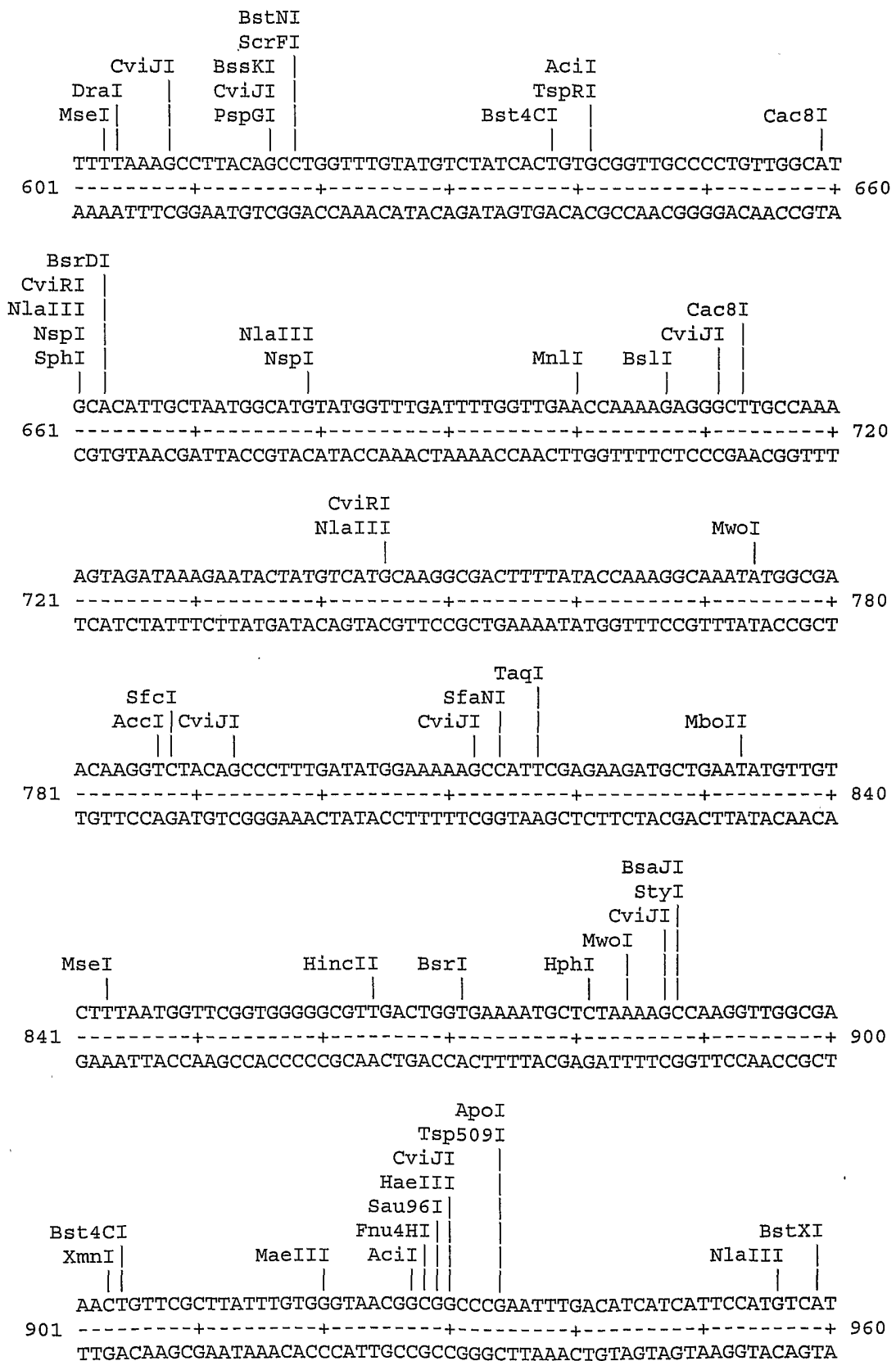
Figure 2. Restriction enzyme analysis of the *M. catarrhalis* major anaerobically induced OMP gene designated MCA100170



5/189



6/189



7/189

HphI MnlI HphI

| | |

961 TGGTGAGATTTTGTATAAGGTTCACTTTGAGGGTGGTAAGGGTGAAAACCACAATATCCA
-----+-----+-----+-----+-----+-----+-----+
ACCACTCTAAAAACTATTCCAAGTGAACTCCCACCATTCCCACTTTTGGTGTATTAGGT 1020

Fnu4HI
 HaeII
 HhaI
 TseI
 HinP1I
 MseI
 ScrFI
 HpaII
 BssKI
 BstF5I
 BspMI
 BbvI
 PflMI
 BslI
 MwoI
 TspRI
 ApoI
 Tsp509I

AACCACGCTAATCCCAGCAGGTGGCGCTGCCATCACTGAATTTAAGGTGGATGTGCCGGG
 1021 -----+-----+-----+-----+-----+-----+-----+-----+ 1080
 TTGGTGCGATTAGGGTCGTCCACCGCGACGGTAGTGACTTAAATTCCACCTACACGGCCC

NlaIII
 MboII
 HincII
 FokI HphI MseI
 TGATTATGTCTTGGTTGACCATGCCATCTTCCGTGCCTTTAAACAAAGGGGCATTGGGCAT
 1081 -----+-----+-----+-----+-----+-----+-----+ 1140
 ACTAATACAGAACCAACTGGTACGGTAGAAGGCACGGAAATTGTTTCCCCGTAACCCGTA

```

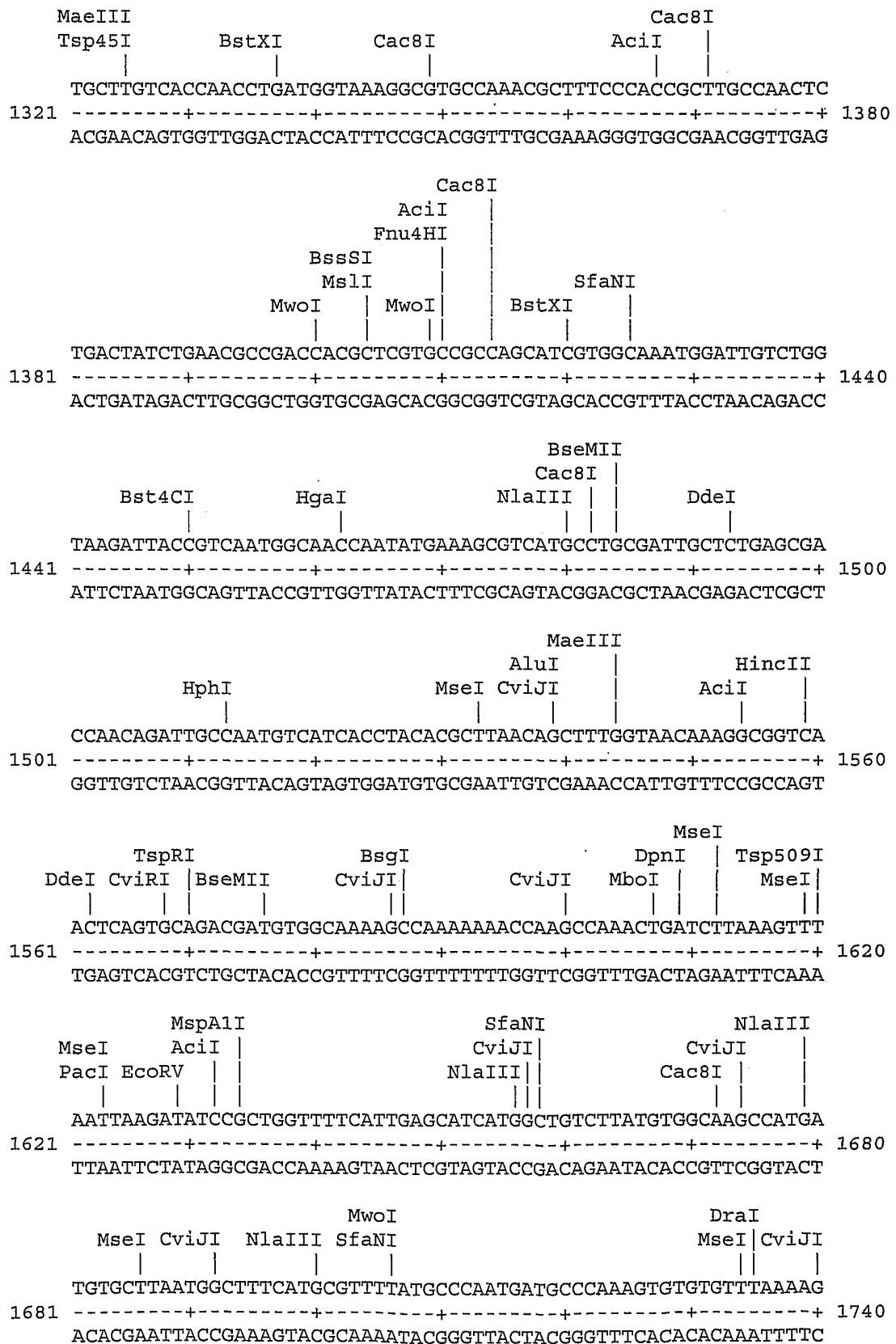
      MboII
MseI    NlaIII
AflII | HphI | AlwNI
SmlI | BspHI | TthIII
|| || |
ACTTAAGGTGGAAGGTGAAGAAAATCATGAGATTTATTACACACAAACAAACAGACGCTGT
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
TGAATTCACCTTCCACTTCTTTTAGTACTCTAAATAAGTGTTGTTTGTCTGCGACA

```

BseSI
 Bsp1286I MfeI
 MnlI NlaIV Tsp509I
 HgaI BanI MwoI BsgI CviRI
 CTATCTGCCAGAGGGTGCCCCACAAGCAATTGATACCCAAGAAGCACCCAAAACACCTGC
 1201 -----+-----+-----+-----+-----+-----+-----+-----+ 1260
 GATAGACGCTCTCCACGGGGTGTTCTGTTAACTATGGGTTCTTCGTGGGTTTTGTGGACG

Fnu4HI
HphI
TseI
MseI
MlyI HinfI
BspMI BspMI BspMI MwoI PleI BbvI Bst4CI
ACCTGCCAACTTACAAGAGCAGATTAAAGCAGGTAAGGCAACCTATGACTCTAACTGTGC
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
TGGACGGTTGAATGTTCTCGTCTAATTTTCGTCCATTCCGTTGGATACTGAGATTGACACG

8/189



9/189

```
      NlaIII
      |
      CCATGCCAAAAGGTGTATG
1741 -----+----- 1759
      GGTACGGTTTTCCACATAC
```

10/189

Figure 3. Sequence of *M. catarrhalis* rare lipoprotein A gene
(SEQ ID NOS: 3 & 4) designated MCA100178

```

ctgctggttt ttggaggatt agatgaatat gagcagctta aagcaaacta acaatgaaca 60
aacgaaaatt gatagctgaa caagtgagcg agcaatcaca atg aaa tac cta ttt 115
                                         Met Lys Tyr Leu Phe
                                         1           5

cta gcc atc gcc ctg ctt gcc act gcg tca caa gcc aat tat gcg acc 163
Leu Ala Ile Ala Leu Leu Ala Thr Ala Ser Gln Ala Asn Tyr Ala Thr
                        10                15                20

tac tat ggc gat aaa ttc cac ggt cgt aaa aca gca agt ggt gag aga 211
Tyr Tyr Gly Asp Lys Phe His Gly Arg Lys Thr Ala Ser Gly Glu Arg
                        25                30                35

ttt gac caa aac gcc atg acc tgt gcc agc aat agc cat aaa ttt ggc 259
Phe Asp Gln Asn Ala Met Thr Cys Ala Ser Asn Ser His Lys Phe Gly
                        40                45                50

aca cgg tta aaa gtt acc aac acc gcc aat gat aaa tca gtg att tgc 307
Thr Arg Leu Lys Val Thr Asn Thr Ala Asn Asp Lys Ser Val Ile Cys
                        55                60                65

cgt gtg aat gac cgt ggc ggt ttt ggt aaa tat ggc acg act tta gat 355
Arg Val Asn Asp Arg Gly Gly Phe Gly Lys Tyr Gly Thr Thr Leu Asp
                        70                75                80                85

tta tca aaa gcc agc ttt gct aag att gca tca ctg tcg caa ggt aaa 403
Leu Ser Lys Ala Ser Phe Ala Lys Ile Ala Ser Leu Ser Gln Gly Lys
                        90                95                100

att aaa gta aca att agg aga ttg tga gatgataaat agactaaaaa 450
Ile Lys Val Thr Ile Arg Arg Leu
                        105                110

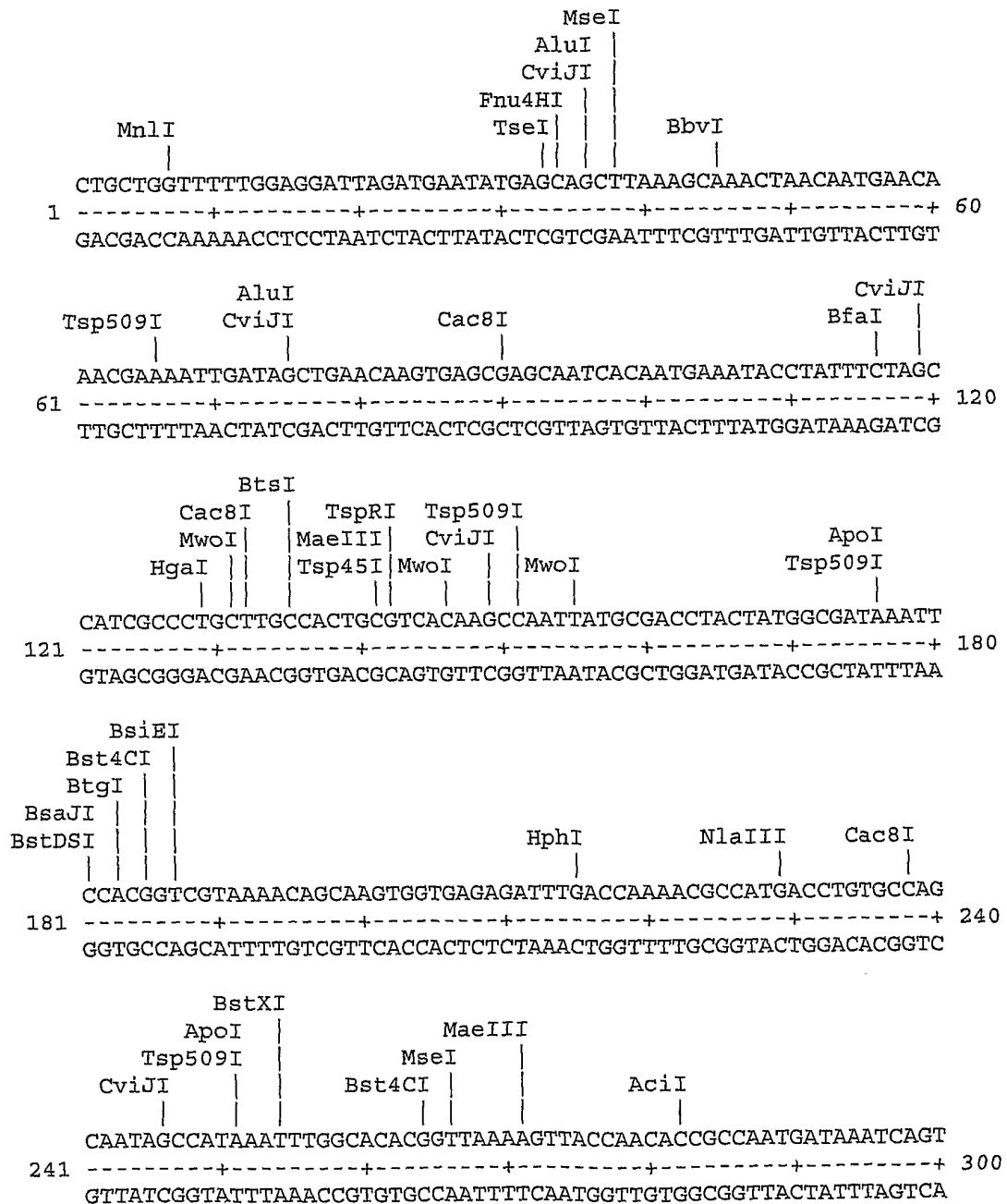
acttaacacc cgaacaagca aaagccgaat ctgaccgttt ttggtcagcg agtgatgatg 510

ctatTTTTTcc gccattaacg 530

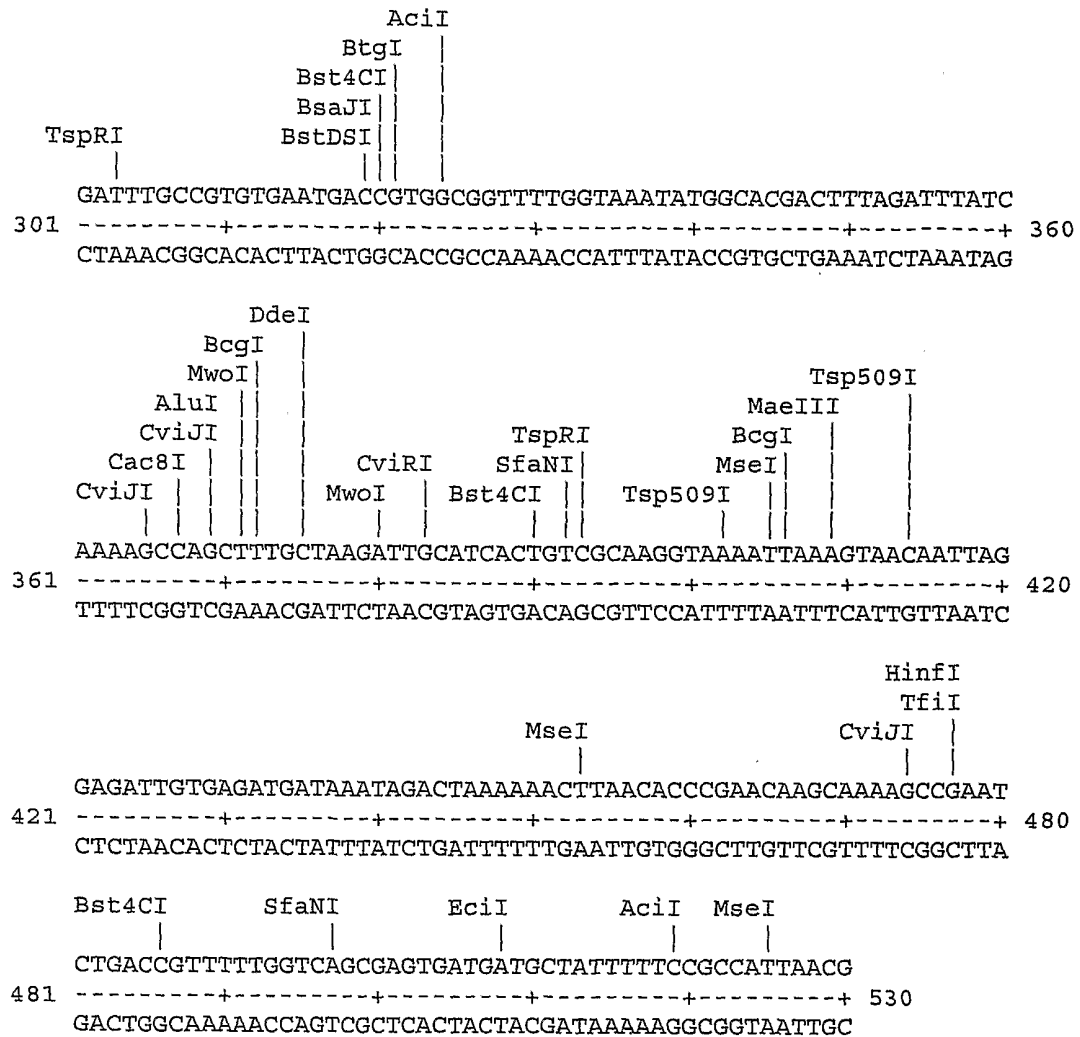
```

11/189

Figure 4. Restriction enzyme analysis of the *M. catarrhalis* rare lipoprotein A gene (SEQ ID NO: 3) designated MCA100178



12/189



13/189

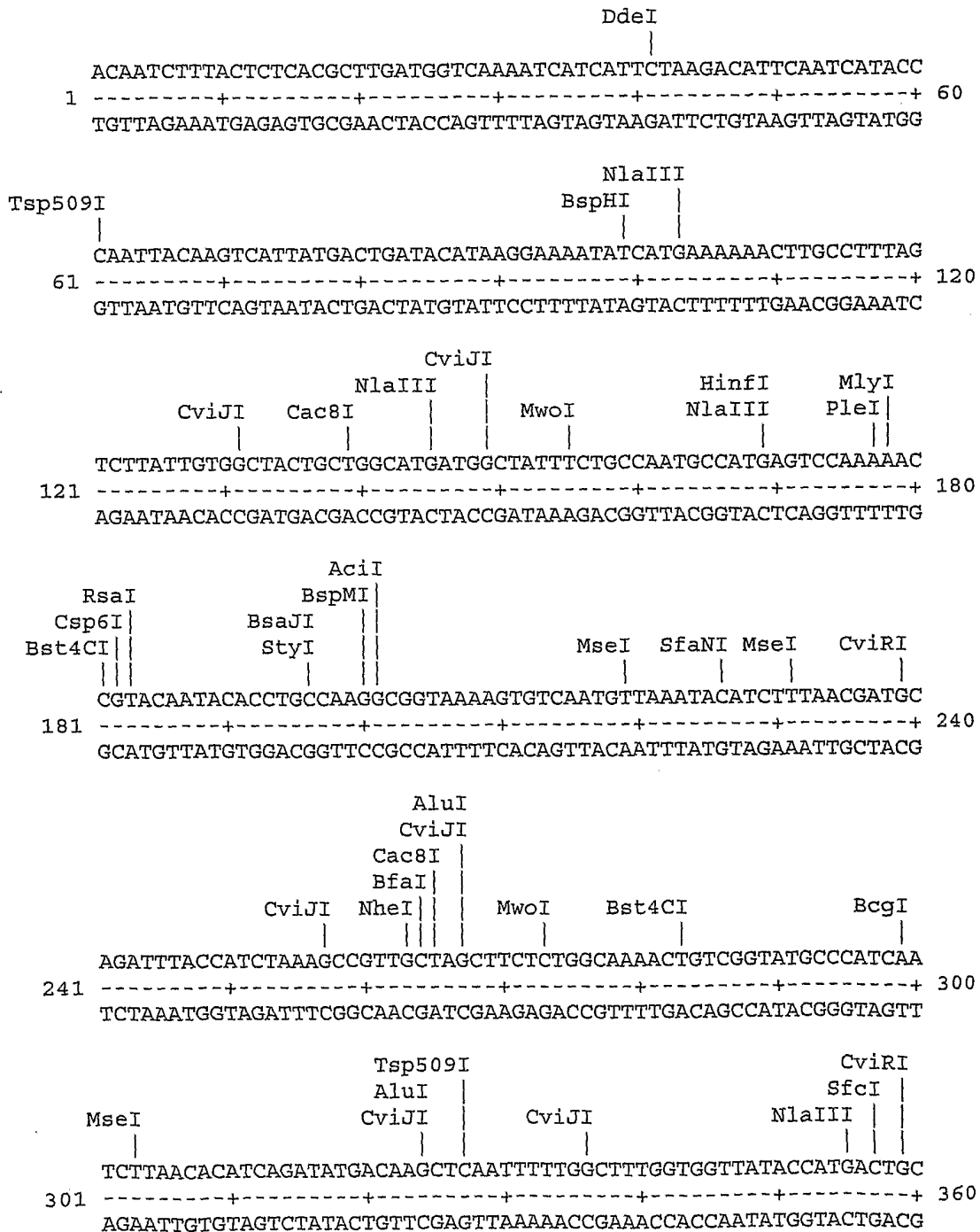
Figure 5. Sequence of *M. catarrhalis* gene encoding a component protein of the adhesin complex (SEQ ID NO: 5 & 6) designated MCA100412

```

acaatcttta ctctcacgct tgatgggtcaa aatcatcatt ctaagacatt caatcatacc 60
caattacaag tcattatgac tgatacataa ggaaaatatc atg aaa aaa ctt gcc 115
                                         Met Lys Lys Leu Ala
                                         1           5
ttt agt ctt att gtg gct act gct ggc atg atg gct att tct gcc aat 163
Phe Ser Leu Ile Val Ala Thr Ala Gly Met Met Ala Ile Ser Ala Asn
                        10                15                20
gcc atg agt cca aaa acc gta caa tac acc tgc caa ggc ggt aaa agt 211
Ala Met Ser Pro Lys Thr Val Gln Tyr Thr Cys Gln Gly Gly Lys Ser
                        25                30                35
gtc aat gtt aaa tac atc ttt aac gat gca gat tta cca tct aaa gcc 259
Val Asn Val Lys Tyr Ile Phe Asn Asp Ala Asp Leu Pro Ser Lys Ala
                        40                45                50
gtt gct agc ttc tct ggc aaa act gtc ggt atg ccc atc aat ctt aac 307
Val Ala Ser Phe Ser Gly Lys Thr Val Gly Met Pro Ile Asn Leu Asn
      55                60                65
aca tca gat atg aca agc tca att ttt ggc ttt ggt ggt tat acc atg 355
Thr Ser Asp Met Thr Ser Ser Ile Phe Gly Phe Gly Gly Tyr Thr Met
      70                75                80                85
act gca gac tat atc gat gcc aag aac tac aat caa gtc ggt att gca 403
Thr Ala Asp Tyr Ile Asp Ala Lys Asn Tyr Asn Gln Val Gly Ile Ala
                        90                95                100
gtc atc acc gac cca aag aac aaa act ttg ttc aaa aac tgc aac cca 451
Val Ile Thr Asp Pro Lys Asn Lys Thr Leu Phe Lys Asn Cys Asn Pro
                        105                110                115
aga taa tttttttact acccaaataa aaaagcatgg ataaccatgc tttttttattg 507
Arg
attgctgatt gattaaatca attcgctggt gtcgtatccg tggttgccgc 557

```

Figure 6. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a component protein of the adhesin complex (SEQ ID NO: 5) designated MCA100412



15/189

```

      PstI      ClaI                      CviRI
      SfaNI     TaqI                      HphI
      |         |                         ||
AGACTATATCGATGCCAAGAACTACAATCAAGTCGGTATTGCAGTCATCACCGACCCAAA
361 -----+-----+-----+-----+-----+-----+ 420
      TCTGATATAGCTACGGTTCCTTGATGTTAGTTCAGCCATAACGTCAGTAGTGGCTGGGTTT

                  CviRI      Tsp509I
                  |         |
GAACAAAACCTTTGTTCAAAAAC TGCAACCCAAGATAATTTTTTTTACTACCCAAATAAAAA
421 -----+-----+-----+-----+-----+-----+ 480
      CTTGTTTTTGAAACAAGTTTTTTGACGTTGGGTTCTATTAATAAAAAATGATGGGTTTATTTTT

                                Tsp509I
                                |
      NlaIII      NlaIII                      MseI
      |         |                         |
AGCATGGATAACCATGCTTTTTTATTGATTGCTGATTGATTAAATCAATTCGCTGTTGTC
481 -----+-----+-----+-----+-----+-----+ 540
      TCGTACCTATTGGTACGAAAAAATAACTAACGACTAACTAATTTAGTTAAGCGACAACAG

                                Fnu4HI
                                |
      BtgI CviRI
      |   |
      BsaJI TseI
      |   |
      BstDSI BciVI
      |   |
GTATCCGTGGTTGCAGC
541 -----+----- 557
      CATAGGCACCAACGTCG
```

16/189

Figure 7. Sequence of *M. catarrhalis* gene encoding a component of the adhesin complex (SEQ ID NOs: 7 and 8) designated MCA100413

```

tttagaaaat tgattattga taaaagtgga agctgagcca aatttataca caaaaccgtt 60
taaatttgag cttttaaata acaacatttg gaggtattct atg aaa atc tta gct 115
                                   Met Lys Ile Leu Ala
                                   1           5

tta gcc cct gcg act gcc ttg gca ctt gtt gga ttc att ggt gca gct 163
Leu Ala Pro Ala Thr Ala Leu Ala Leu Val Gly Phe Ile Gly Ala Ala
                                   10           15           20

cat gct caa gtt gac cat gct caa gaa gct cat gat tct acc gta tta 211
His Ala Gln Val Asp His Ala Gln Glu Ala His Asp Ser Thr Val Leu
                                   25           30           35

aag cgt caa gtt act tat gcc tgc caa tca ggt ggg caa gcc acc gta 259
Lys Arg Gln Val Thr Tyr Ala Cys Gln Ser Gly Gly Gln Ala Thr Val
                                   40           45           50

act tat ggc ttt aac caa caa aga ttg cca act tat gcc agt gct tat 307
Thr Tyr Gly Phe Asn Gln Gln Arg Leu Pro Thr Tyr Ala Ser Ala Tyr
                                   55           60           65

gtt gat ggc aaa aac cgt ttt atg cca ttt aat tta aat cgt tct gac 355
Val Asp Gly Lys Asn Arg Phe Met Pro Phe Asn Leu Asn Arg Ser Asp
                                   70           75           80           85

aat gtt gat agc gtt ttt ggt gat gaa aat aac tac agc ttg atg act 403
Asn Val Asp Ser Val Phe Gly Asp Glu Asn Asn Tyr Ser Leu Met Thr
                                   90           95           100

gat gcc atg aca ttg tat aac tat cac aaa cta ggc att aac atc aca 451
Asp Ala Met Thr Leu Tyr Asn Tyr His Lys Leu Gly Ile Asn Ile Thr
                                   105           110           115

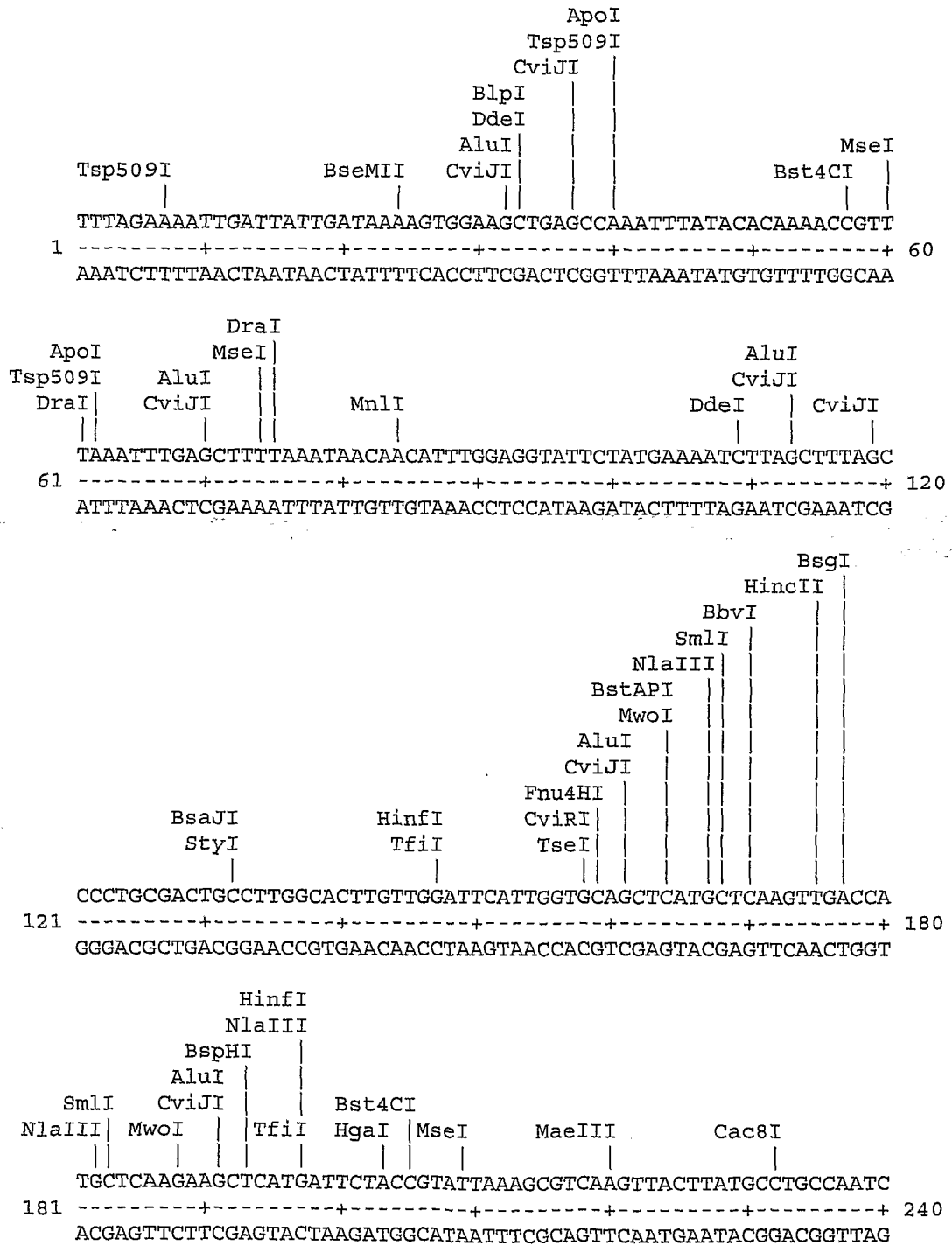
agc cct tct ggt gat ttt gct cat aaa ggc tgt aat gtc aga tat gtt 499
Ser Pro Ser Gly Asp Phe Ala His Lys Gly Cys Asn Val Arg Tyr Val
                                   120           125           130

aca aaa ctt taa cgcttaacat aatattaata aaaataatat taataaggta 551
Thr Lys Leu
                                   135

agctttaatc ccaagtttta atcataaata tcacttgatt ttaacttggg tatttgattt 611

```

Figure 8. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a component of the adhesin complex (SEQ ID NO: 7) designated MCA100413



18/189

BstXI CviJI MaeIII CviJI MseI BsrI
 AGGTGGGCAAGCCACCGTAACCTTATGGCTTTAACCAACAAAGATTGCCAAGCTTATGCCAG
 241 -----+-----+-----+-----+-----+-----+-----+ 300
 TCCACCCGTTTCGGTGGCATTTGAATACCGAAATTGGTTGTTTCTAACGGTTGAATACGGTC

TspRI Bst4CI Tsp509I DraI SmaI
 TGCTTATGTTGATGGCAAAAACCGTTTTATGCCATTTAATTTAAATCGTTCTGACAATGT
 301 -----+-----+-----+-----+-----+-----+-----+ 360
 ACGAATACAACCTACCGTTTTTGGCAAAATACGGTAAATTTAAATTTAGCAAGACTGTTACA

SfaNI SfcI AluI HphI CviJI NlaIII
 TGATAGCGTTTTTGGTGTATGAAAATAACTACAGCTTGATGACTGATGCCATGACATTGTA
 361 -----+-----+-----+-----+-----+-----+-----+ 420
 ACTATCGCAAAAACCACTACTTTTATTGATGTCTGAAGTACTGACTACGGTACTGTAACAT

BfaI MseI CviJI HphI CviJI
 TAACTATCACAACTAGGCATTAACATCACAAAGCCCTTCTGGTGATTTTGCTCATAAAGG
 421 -----+-----+-----+-----+-----+-----+-----+ 480
 ATTGATAGTGTTTGATCCGTAATTGTAGTGTTTCGGAAGACCACTAAAACGAGTATTTC

AseI MseI SspI SspI
 CTGTAATGTCAGATATGTTACAAAACCTTTAACGCTTAACATAATATTAATAAAAAATAATA
 481 -----+-----+-----+-----+-----+-----+-----+ 540
 GACATTACAGTCTATACAATGTTTTGAAATTGCGAATTGTATTATAATTATTTTATTAT

AseI MseI AluI CviJI MseI MseI SspI SspI
 TTAATAAGGTAAGCTTTAATCCCAAGTTTTAATCATAAATATCACTTGATTTTAACTTGG
 541 -----+-----+-----+-----+-----+-----+-----+ 600
 AATTATTCATTCGAAATTAGGGTTCAAATTAGTATTTATAGTGAAGTAAATTTGAACC

AseI MseI AluI CviJI MseI
 GTATTTGATTT
 601 -----+-----+-----+-----+-----+-----+-----+ 611
 CATAAACTAAA

19/189

Figure 9. Sequence of *M. catarrhalis* gene encoding a potential lipoprotein OMP designated MCA100414

```

taatcaactt acattttgta tttagatgca gaaatatttt ctaatccttag aatttataat 60
aatttcacat ttggtttagcc aaaaagcaaa aggagttttt atg aaa att aaa gca 115
                                         Met Lys Ile Lys Ala
                                         1       5

ttg ggt gtt gtg ctg ttg gca tca agt atg gct ttg gca ggt tgt gca 163
Leu Gly Val Val Leu Leu Ala Ser Ser Met Ala Leu Ala Gly Cys Ala
                        10                      15                      20

aat aca ggc aca act ggc aat ggc aca gga ttt ggt ggt gct aat gtc 211
Asn Thr Gly Thr Thr Gly Asn Gly Thr Gly Phe Gly Gly Ala Asn Val
                        25                      30                      35

aat aag gcg gtg att ggg gct gtg gca ggt gca ctt ggc ggt act gcc 259
Asn Lys Ala Val Ile Gly Ala Val Ala Gly Ala Leu Gly Gly Thr Ala
                        40                      45                      50

att tca aaa gca act ggt ggc gaa aaa aca ggt cgt gat gcc att ttg 307
Ile Ser Lys Ala Thr Gly Gly Glu Lys Thr Gly Arg Asp Ala Ile Leu
                    55                      60                      65

ggg gcg gca gtt ggt gca gca gca ggg gcg tat atg gag cgt caa gca 355
Gly Ala Ala Val Gly Ala Ala Ala Gly Ala Tyr Met Glu Arg Gln Ala
    70                      75                      80                      85

aag cag att gag caa caa atg caa gga acg ggc gtg act gta acc cac 403
Lys Gln Ile Glu Gln Gln Met Gln Gly Thr Gly Val Thr Val Thr His
                        90                      95                      100

gat acc gac acg ggt aat att aat cta act atg cca ggt aat att act 451
Asp Thr Asp Thr Gly Asn Ile Asn Leu Thr Met Pro Gly Asn Ile Thr
                        105                      110                      115

ttt gct cat gat gac gat act tta aac agt gca ttt ttg ggt cgt tta 499
Phe Ala His Asp Asp Asp Thr Leu Asn Ser Ala Phe Leu Gly Arg Leu
                    120                      125                      130

aac cag ctg gct aat acg atg aat cag tat cat gaa aca acg att gtc 547
Asn Gln Leu Ala Asn Thr Met Asn Gln Tyr His Glu Thr Thr Ile Val
                    135                      140                      145

att gta gga cat aca gac tca acg ggt caa gcg gct tat aat caa gag 595
Ile Val Gly His Thr Asp Ser Thr Gly Gln Ala Ala Tyr Asn Gln Glu
    150                      155                      160                      165

ctg tct gag cgt cga gcg gat tca gtg cgt tat tac ttg att aat caa 643
Leu Ser Glu Arg Arg Ala Asp Ser Val Arg Tyr Tyr Leu Ile Asn Gln
                        170                      175                      180

ggc gtt gat cca tat cgt att cag aca gtg ggg tat ggt atg cga caa 691
Gly Val Asp Pro Tyr Arg Ile Gln Thr Val Gly Tyr Gly Met Arg Gln
                        185                      190                      195

```

20/189

```

ccg att gca tcg aat gca acc gaa gca ggt cgt gct caa aat cgc cgt 739
Pro Ile Ala Ser Asn Ala Thr Glu Ala Gly Arg Ala Gln Asn Arg Arg
      200                205                210

gtt gag ctg atg att tta gca ccg cag ggt atg taa aaatgcttgg 785
Val Glu Leu Met Ile Leu Ala Pro Gln Gly Met
      215                220                225

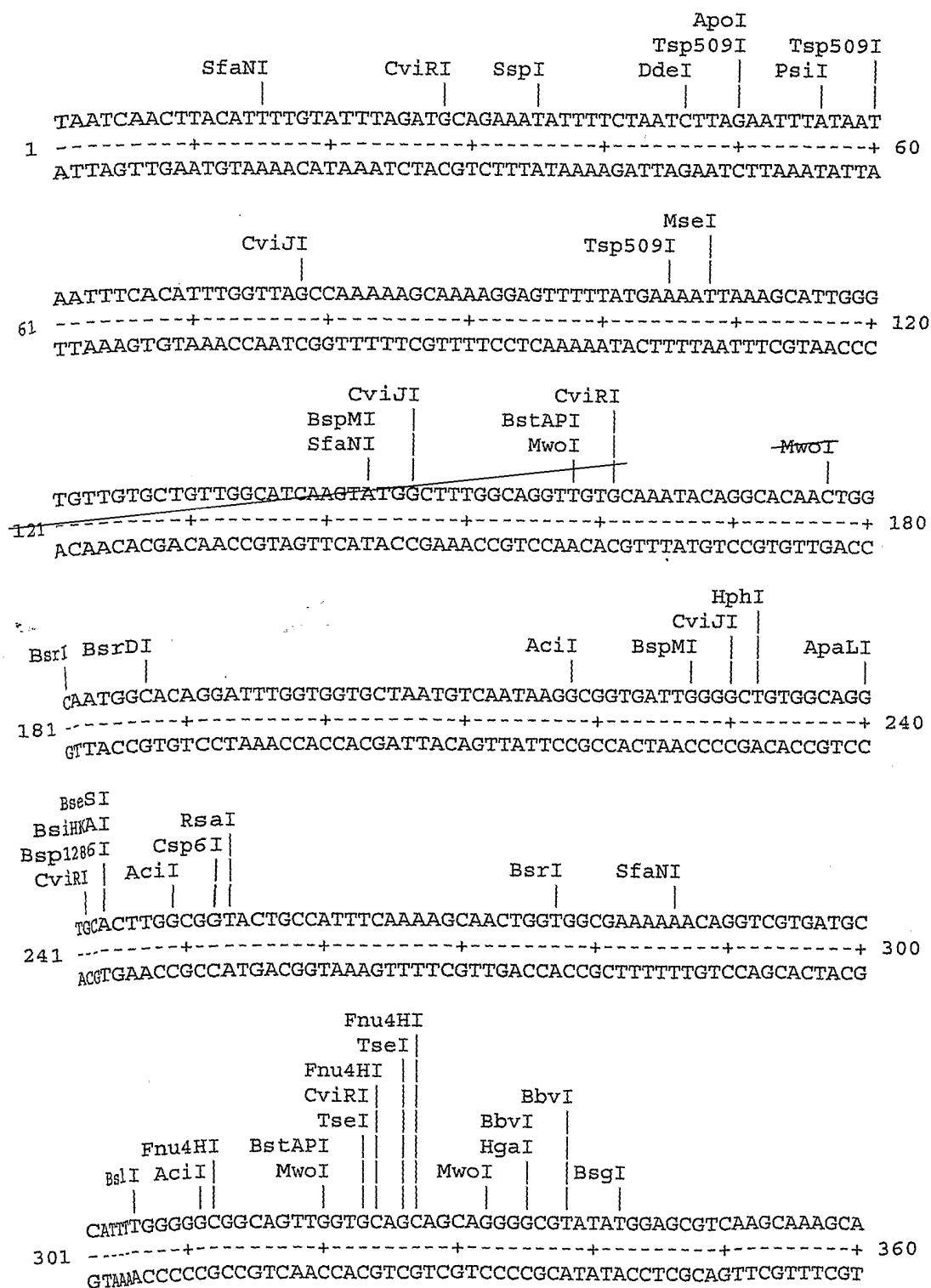
cttatctgac taagataaca tagcgtaa gaatgagata cagataacca tcatacgaat 845

atacacaata tgaatgtata aaatcttagt 875

```

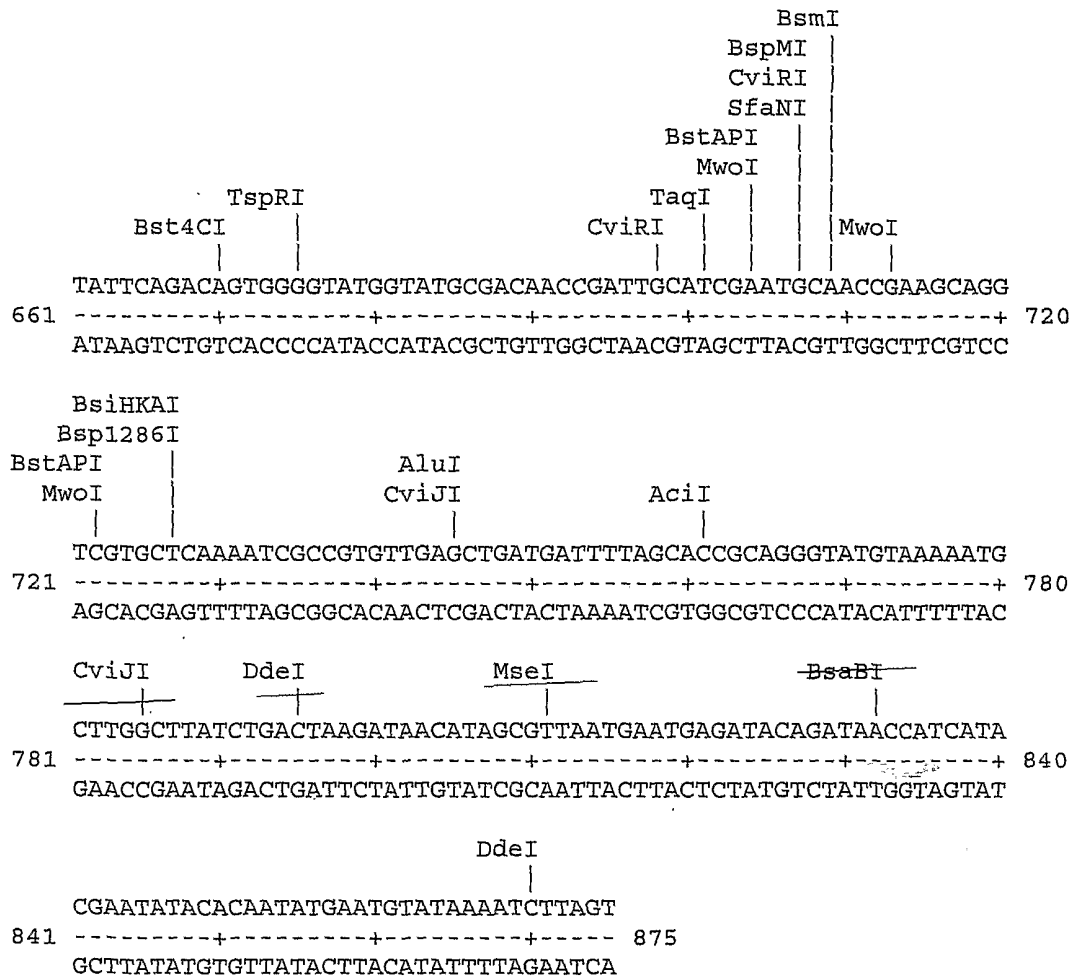
21/189

Figure 10. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a potential lipoprotein OMP designated MCA100414



[illegible]

23/189



24/189

Figure 11. Sequence of *M. catarrhalis* gene encoding a putative oxidoreductase designated MCA100527

aagg	tattat	attac	cttat	tttctt	tctca	taaatt	ttcat	atcaatt	tttaa	catagag	cta						60
acaaaaagct caattagcca ctggataaaa catgacaacc atg cgt tca cat gat												115					
Met Arg Ser His Asp												5					
1																	
gca	att	cgt	gat	ttt	caa	ttc	acc	aaa	gac	agt	cta	aaa	gat	aaa	att	163	
Ala	Ile	Arg	Asp	Phe	Gln	Phe	Thr	Lys	Asp	Ser	Leu	Lys	Asp	Lys	Ile		
				10				15				20					
att	ttg	gtg	aca	ggg	gca	ggc	gat	ggg	ata	ggc	aaa	gtt	gct	gct	ttg	211	
Ile	Leu	Val	Thr	Gly	Ala	Gly	Asp	Gly	Ile	Gly	Lys	Val	Ala	Ala	Leu		
				25				30				35					
act	tat	gcc	aaa	tgc	ggg	gca	acg	gtt	ctt	ttg	ctg	ggg	aaa	act	caa	259	
Thr	Tyr	Ala	Lys	Cys	Gly	Ala	Thr	Val	Leu	Leu	Leu	Gly	Lys	Thr	Gln		
				40				45				50					
tca	aaa	ctt	gaa	gtg	gtt	tat	gat	gaa	att	gag	gca	ctt	ggg	ttg	gcg	307	
Ser	Lys	Leu	Glu	Val	Val	Tyr	Asp	Glu	Ile	Glu	Ala	Leu	Gly	Leu	Ala		
				55				60				65					
ttg	cct	gcc	att	ttg	ccc	atg	gat	ctt	gag	cgg	gca	gat	ttt	gca	cag	355	
Leu	Pro	Ala	Ile	Leu	Pro	Met	Asp	Leu	Glu	Arg	Ala	Asp	Phe	Ala	Gln		
				70				75				80				85	
atg	caa	aag	ctg	gca	ggc	tta	att	gaa	aaa	gaa	ttt	gga	cat	ctt	gat	403	
Met	Gln	Lys	Leu	Ala	Gly	Leu	Ile	Glu	Lys	Glu	Phe	Gly	His	Leu	Asp		
				90				95				100					
ggc	gtg	cta	cat	aat	gca	gcc	att	ttg	ggg	aca	ttg	acg	cca	ctt	gag	451	
Gly	Val	Leu	His	Asn	Ala	Ala	Ile	Leu	Gly	Thr	Leu	Thr	Pro	Leu	Glu		
				105				110				115					
atg	tac	gat	ccg	atc	acc	ttt	gag	cag	gtc	atg	cgg	gtg	aat	atg	aca	499	
Met	Tyr	Asp	Pro	Ile	Thr	Phe	Glu	Gln	Val	Met	Arg	Val	Asn	Met	Thr		
				120				125				130					
gcc	gta	ttc	atg	ctg	acc	caa	agc	cta	att	ccg	cta	ctg	cgt	gca	gcc	547	
Ala	Val	Phe	Met	Leu	Thr	Gln	Ser	Leu	Ile	Pro	Leu	Leu	Arg	Ala	Ala		
				135				140				145					
caa	agc	ggc	tcg	gtg	gta	ttt	aca	agc	agc	agc	gtg	ggg	gca	aga	cca	595	
Gln	Ser	Gly	Ser	Val	Val	Phe	Thr	Ser	Ser	Ser	Val	Gly	Ala	Arg	Pro		
				150				155				160				165	
aga	gcg	ttt	tgg	ggg	gct	tat	gca	ctc	tca	aaa	caa	gcg	gtt	gaa	ggc	643	
Arg	Ala	Phe	Trp	Gly	Ala	Tyr	Ala	Leu	Ser	Lys	Gln	Ala	Val	Glu	Gly		
				170				175				180					
atg	gca	acc	att	ttc	acc	caa	gaa	acc	caa	aac	atg	aca	gca	ctt	cgg	691	
Met	Ala	Thr	Ile	Phe	Thr	Gln	Glu	Thr	Gln	Asn	Met	Thr	Ala	Leu	Arg		
				185				190				195					

25/189

```

ttt aat tgt atc aat cct ggt gca act cgt acc aat atg cgt gca cat 739
Phe Asn Cys Ile Asn Pro Gly Ala Thr Arg Thr Asn Met Arg Ala His
      200                      205                      210

gct ttt cca ggt gag gac cct tat aca ctt aag acg cct gag gat atc 787
Ala Phe Pro Gly Glu Asp Pro Tyr Thr Leu Lys Thr Pro Glu Asp Ile
      215                      220                      225

atg cca gtt tat gtg tat ttg atg aca gat gag gca agc ggt att aaa 835
Met Pro Val Tyr Val Tyr Leu Met Thr Asp Glu Ala Ser Gly Ile Lys
      230                      235                      240                      245

ggt caa gtg att gat tgc caa cca aaa aac gct tag ccgatttagc 881
Gly Gln Val Ile Asp Cys Gln Pro Lys Asn Ala
      250                      255

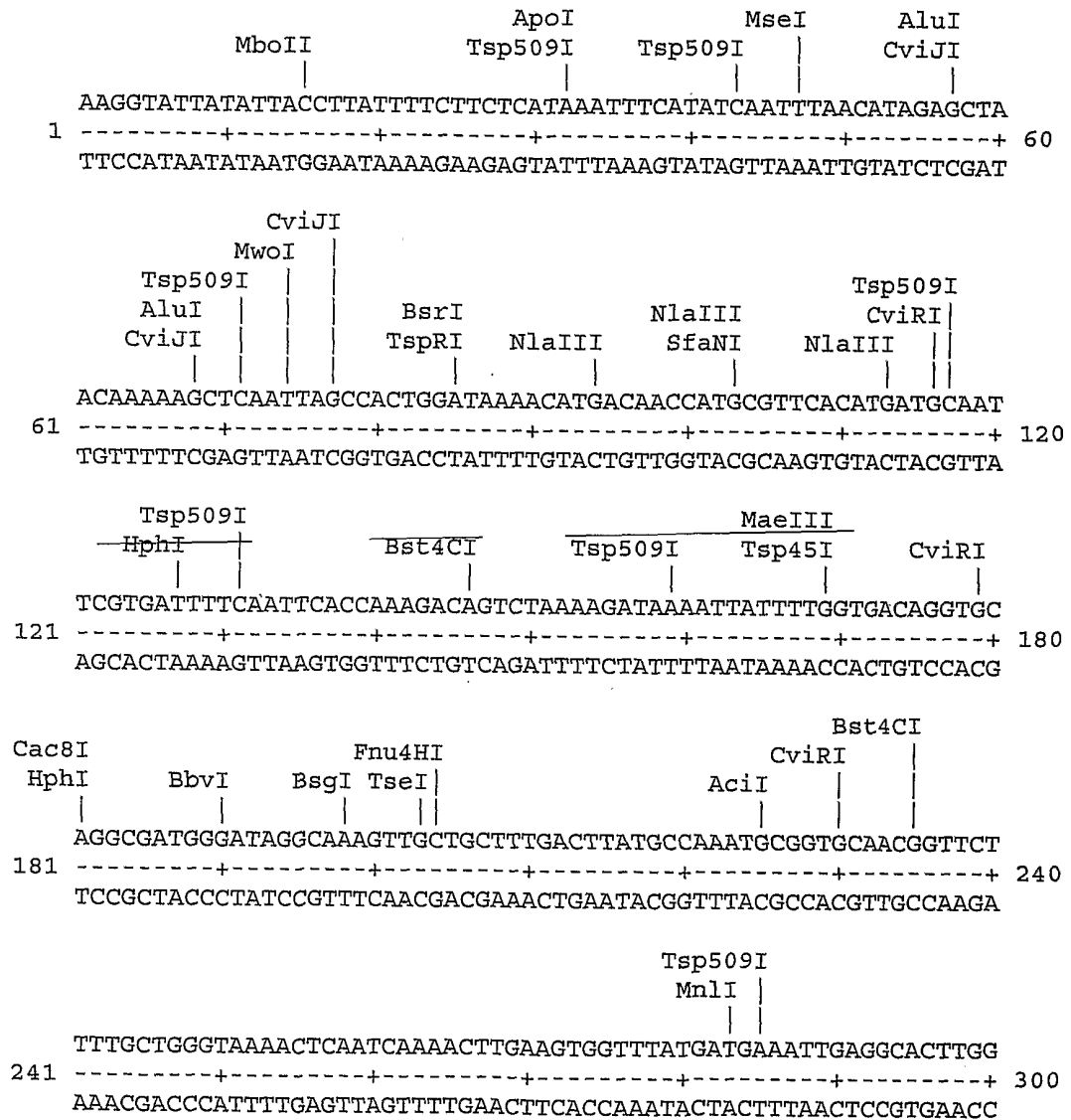
caacacaaat aacatcattg ttattatcaa atcagcctta acctacactc tttaaaccag 941

tcgtggcaaa ttttttaacc aaaacttcac 971

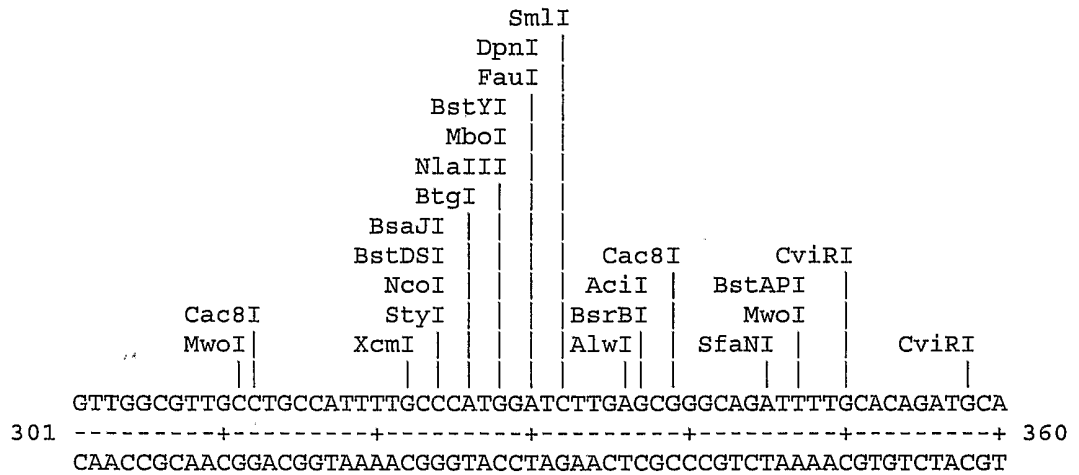
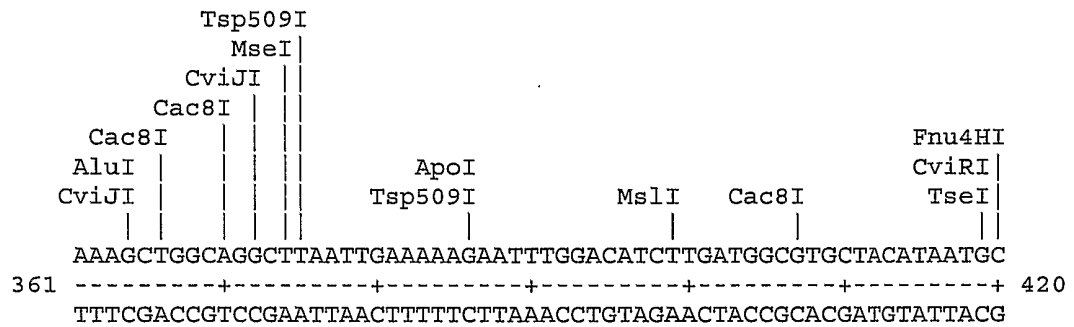
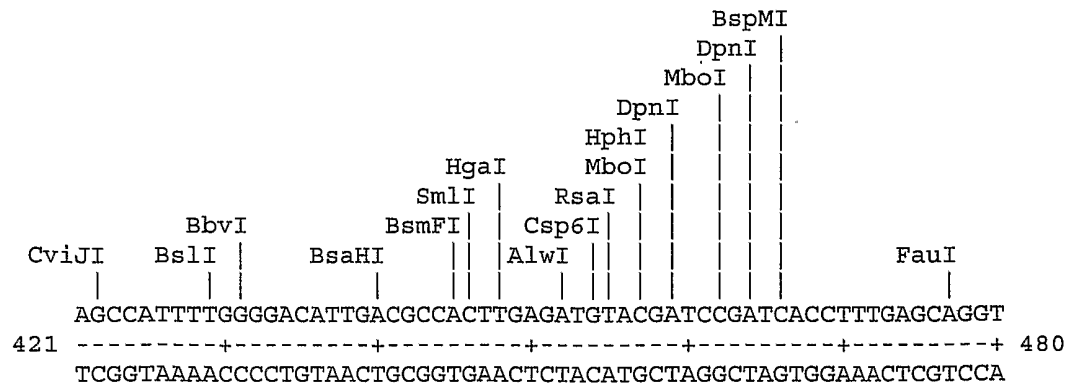
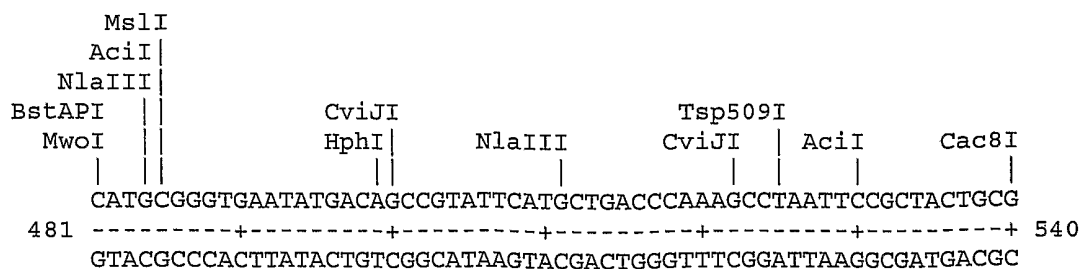
```

26/189

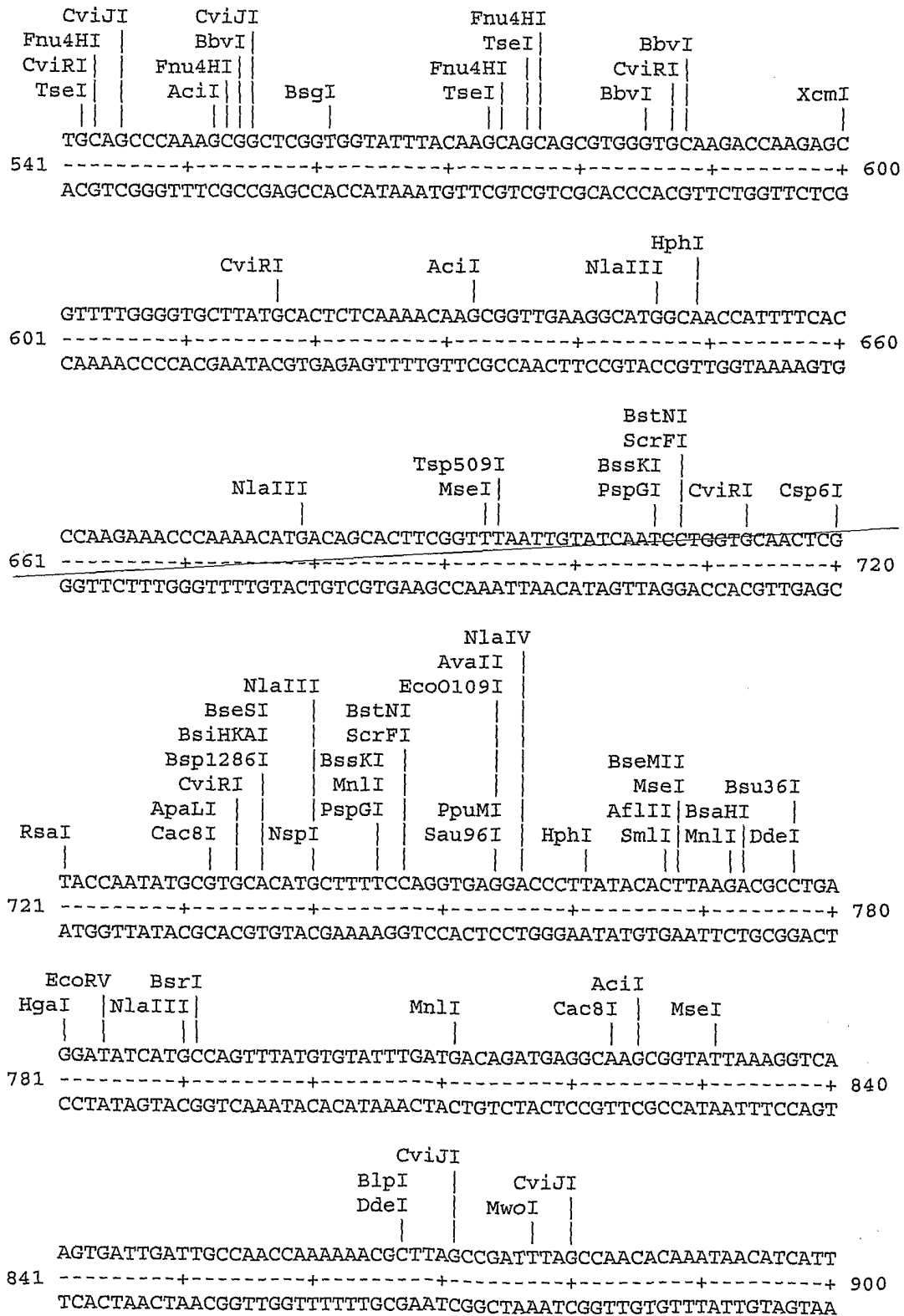
Figure 12. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a putative oxidoreductase designated MCA100527



27/189



28/189



29/189

```

                MseI          DraI          ApoI
          CviJI |          MseI|BsrI      Tsp509I      MseI
            |   |          ||   |         |           |
901  GTTATTATCAAATCAGCCTTAACCTACACTCTTTAAACCAGTCGTGGCAAATTTTAAAC
-----+-----+-----+-----+-----+-----+ 960
    CAATAATAGTTTAGTCGGAATTGGATGTGAGAAATTGGTCAGCACCGTTTAAAAAATTG

    CAAAAC TTCAC
961  -----+-- 971
    GTTTTGAAGTG
```

30/189

Figure 13. Sequence of *M. catarrhalis* OprM gene designated MCA100564

```

cccaatgttt tatgtttggg tgcgtacgct atttccttat aagccaaaag ctcagcctgt 60
aaccatgct aaatctcatg aacatgttta aggattgatt atg aaa atc tct aca 115
                                     Met Lys Ile Ser Thr
                                     1       5
act gca cgt atc ttg acg ctc tca gcg ctg gcg att ggc atg gca gct 163
Thr Ala Arg Ile Leu Thr Leu Ser Ala Leu Ala Ile Gly Met Ala Ala
          10              15              20
tgt agt agc att cca aag aaa att gac aca tcg gca ccg att ttg gcg 211
Cys Ser Ser Ile Pro Lys Lys Ile Asp Thr Ser Ala Pro Ile Leu Ala
          25              30              35
gtg cct aat gtg cct atg aag gat ggc tat caa gtc tat gat gct gac 259
Val Pro Asn Val Pro Met Lys Asp Gly Tyr Gln Val Tyr Asp Ala Asp
          40              45              50
acc atc agt gtg gcg cag gca cca agc gtg gca tca cta cgc tgg cag 307
Thr Ile Ser Val Ala Gln Ala Pro Ser Val Ala Ser Leu Arg Trp Gln
          55              60              65
gaa ttt tat aca aat cct aaa ctt gca gct ttg att gag ctt gct tta 355
Glu Phe Tyr Thr Asn Pro Lys Leu Ala Ala Leu Ile Glu Leu Ala Leu
          70              75              80              85
caa aat aac aaa gac ttg caa tct gcg gtc tta gcc gtg caa tca gcc 403
Gln Asn Asn Lys Asp Leu Gln Ser Ala Val Leu Ala Val Gln Ser Ala
          90              95              100
cgt gct caa tat caa att acc gaa gct ggc agc gtg cca caa gtc ggc 451
Arg Ala Gln Tyr Gln Ile Thr Glu Ala Gly Ser Val Pro Gln Val Gly
          105              110              115
tca aat acc agt gtg aca cgc caa gcg aat aac cgt atc gat gcc aat 499
Ser Asn Thr Ser Val Thr Arg Gln Ala Asn Asn Arg Ile Asp Ala Asn
          120              125              130
gct tca acc aat tat cat gtt ggg ctt gcg atg agt ggt tat gag ctg 547
Ala Ser Thr Asn Tyr His Val Gly Leu Ala Met Ser Gly Tyr Glu Leu
          135              140              145
gat tta tgg ggc aaa gtt gcc agt caa aaa cag caa gca cta cat caa 595
Asp Leu Trp Gly Lys Val Ala Ser Gln Lys Gln Gln Ala Leu His Gln
          150              155              160              165
tat ttg gca acc aac gcc gcc aaa gac gcc gtt cag att tca atc atc 643
Tyr Leu Ala Thr Asn Ala Ala Lys Asp Ala Val Gln Ile Ser Ile Ile
          170              175              180
tca agc gtc gcc caa ggt tac gtt aac tta gct cac gct ttg gct caa 691
Ser Ser Val Ala Gln Gly Tyr Val Asn Leu Ala His Ala Leu Ala Gln
          185              190              195

```

31/189

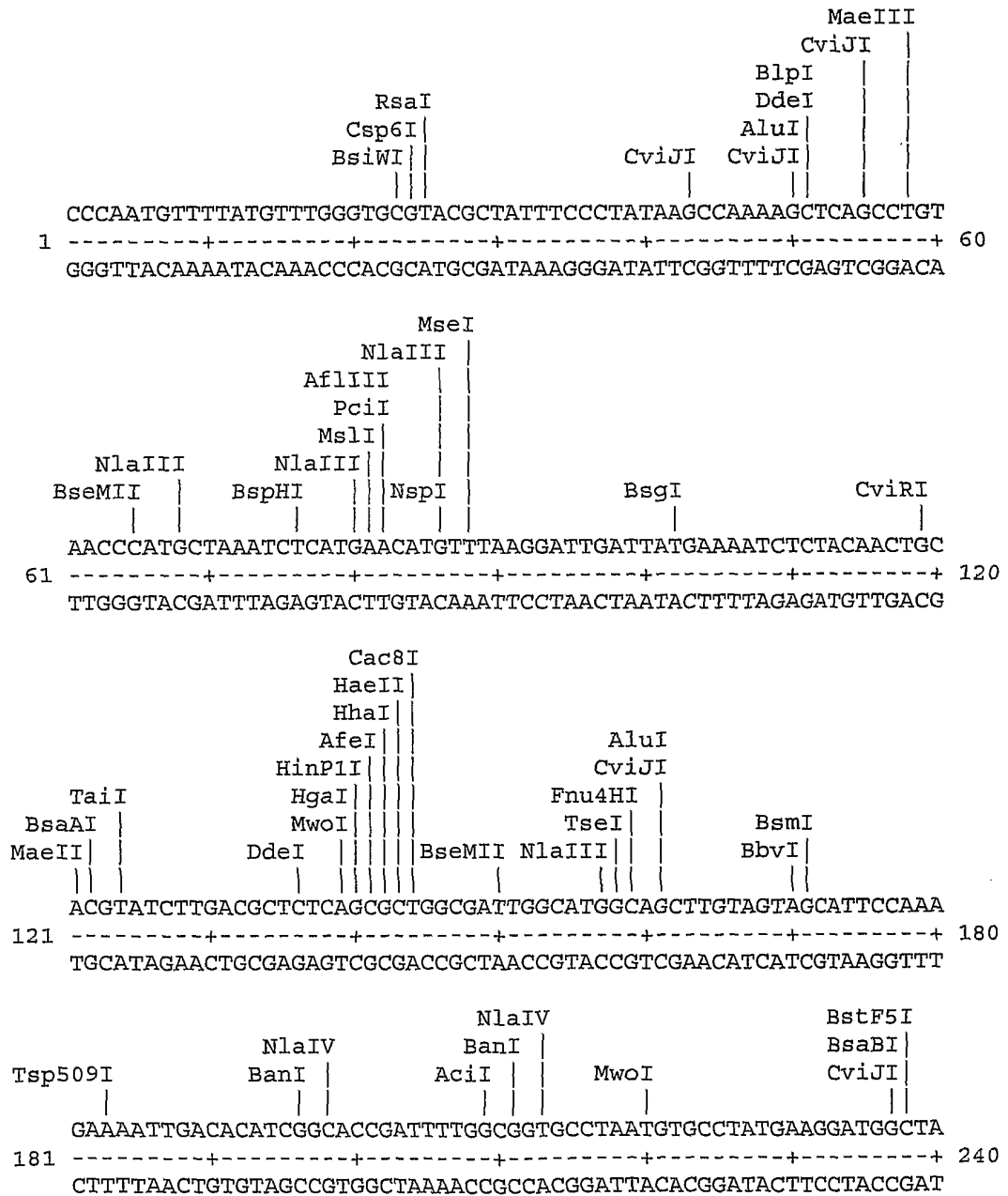
agg cag ttg gct gag cag acg cta aaa acc cgt gaa cat gcg atg atg	739
Arg Gln Leu Ala Glu Gln Thr Leu Lys Thr Arg Glu His Ala Met Met	
200 205 210	
att acc caa aag cgt ttt gaa gcg ggg att gat tct aag tcg cca agt	787
Ile Thr Gln Lys Arg Phe Glu Ala Gly Ile Asp Ser Lys Ser Pro Ser	
215 220 225	
cta caa gca gca agt tca ctt gag tca gca cga ttg gca gta tat gca	835
Leu Gln Ala Ala Ser Ser Leu Glu Ser Ala Arg Leu Ala Val Tyr Ala	
230 235 240 245	
gca gat acc agt atc tta aaa gcc aaa aat gcg tta cag ctg ctg att	883
Ala Asp Thr Ser Ile Leu Lys Ala Lys Asn Ala Leu Gln Leu Leu Ile	
250 255 260	
ggc cgt cct gta cca aac gag cta cta cca gcg ata gat gcc agt atg	931
Gly Arg Pro Val Pro Asn Glu Leu Leu Pro Ala Ile Asp Ala Ser Met	
265 270 275	
cat atg ggt cat att acc aca cag aca ttg ttt agt gca ggt ttg ccc	979
His Met Gly His Ile Thr Thr Gln Thr Leu Phe Ser Ala Gly Leu Pro	
280 285 290	
agc gag ctt tta tat tat cgc cca gat att atg cag gct gag cat cgc	1027
Ser Glu Leu Leu Tyr Tyr Arg Pro Asp Ile Met Gln Ala Glu His Arg	
295 300 305	
cta aaa gca gca ggt gca aat atc aat gtg gca cgc gct gct tat ttt	1075
Leu Lys Ala Ala Gly Ala Asn Ile Asn Val Ala Arg Ala Ala Tyr Phe	
310 315 320 325	
ccg tcg att cgt tta tca tct aat gtg gga ttt agt agt aac agt ttg	1123
Pro Ser Ile Arg Leu Ser Ser Asn Val Gly Phe Ser Ser Asn Ser Leu	
330 335 340	
aat aac tta ttt gaa tca agt gct ttg ggc tgg tct ttt ggg cct gcg	1171
Asn Asn Leu Phe Glu Ser Ser Ala Leu Gly Trp Ser Phe Gly Pro Ala	
345 350 355	
att agc ttg cct atc ttt gat gca ggc agt cgc cgt gcc aat cat gag	1219
Ile Ser Leu Pro Ile Phe Asp Ala Gly Ser Arg Arg Ala Asn His Glu	
360 365 370	
atg gcg caa gtt gct cag cag tcg gca ttg gtg gat tat gaa aaa gct	1267
Met Ala Gln Val Ala Gln Gln Ser Ala Leu Val Asp Tyr Glu Lys Ala	
375 380 385	
att caa aat gcc ttt aaa gaa gtg tcg gat gtt tta gct gag cgt gca	1315
Ile Gln Asn Ala Phe Lys Glu Val Ser Asp Val Leu Ala Glu Arg Ala	
390 395 400 405	
act tta ggc ttg cgt ctt gat gcc cag att cgc ctt cag gat aat tac	1363
Thr Leu Gly Leu Arg Leu Asp Ala Gln Ile Arg Leu Gln Asp Asn Tyr	
410 415 420	

32/189

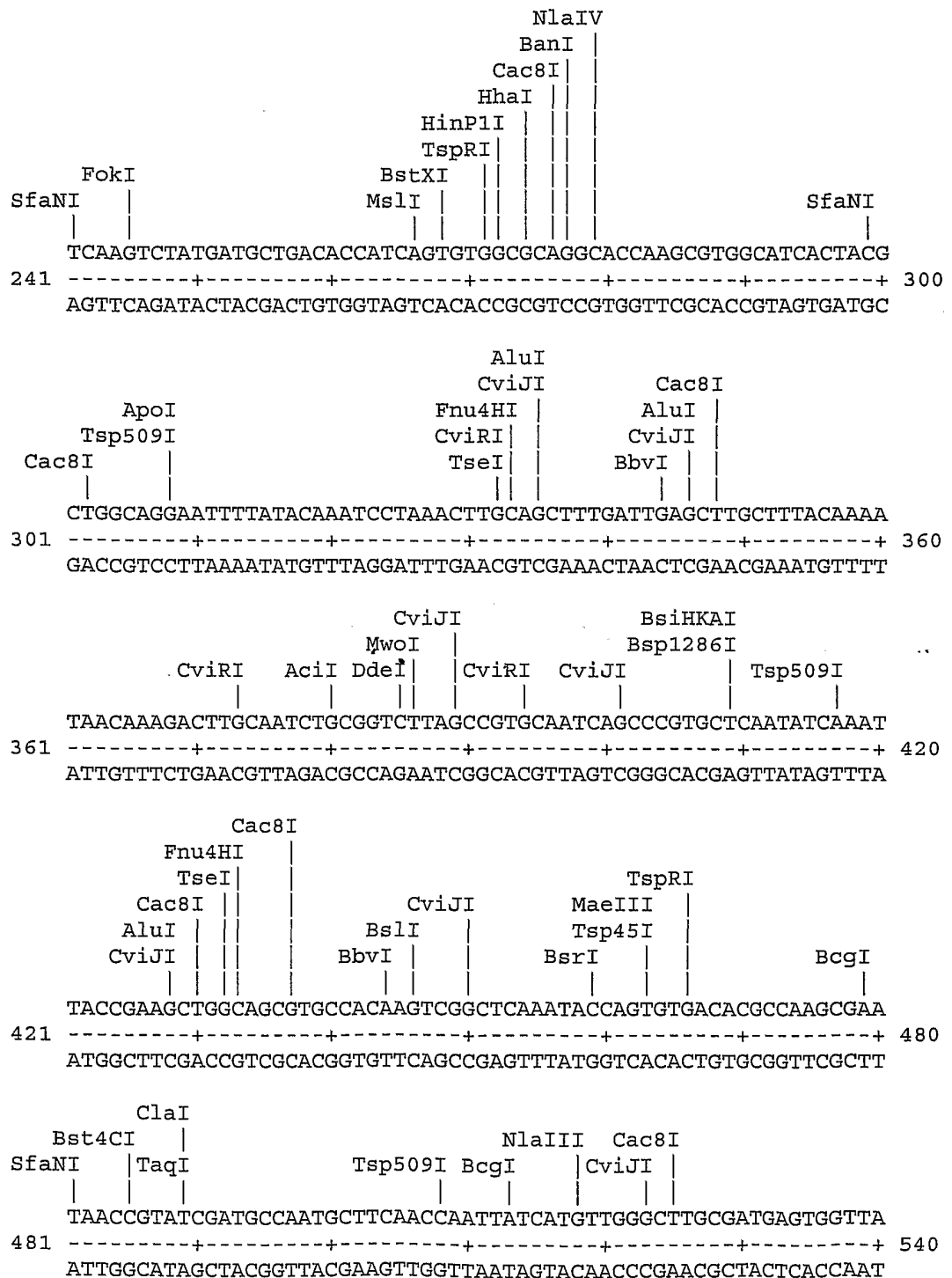
cgt caa act tat gat atc gct tat gca aga ttt cgt tct gga ttg gat	1411
Arg Gln Thr Tyr Asp Ile Ala Tyr Ala Arg Phe Arg Ser Gly Leu Asp	
425 430 435	
aat tat ctg acg gta ctg gac gcc gag cgg tct tta ttt att aat cag	1459
Asn Tyr Leu Thr Val Leu Asp Ala Glu Arg Ser Leu Phe Ile Asn Gln	
440 445 450	
caa aat ata cta cag ctt gaa ctt gcc aag tta gtc agc caa atc cag	1507
Gln Asn Ile Leu Gln Leu Glu Leu Ala Lys Leu Val Ser Gln Ile Gln	
455 460 465	
cta tac caa gca ttg ggc ggc ggt gca agc tta act gct gag caa atc	1555
Leu Tyr Gln Ala Leu Gly Gly Gly Ala Ser Leu Thr Ala Glu Gln Ile	
470 475 480 485	
aca gaa ttt aat cgt cag cgt gaa gcc atg cgt cca gcc atg ctg atg	1603
Thr Glu Phe Asn Arg Gln Arg Glu Ala Met Arg Pro Ala Met Leu Met	
490 495 500	
gat gac gaa aaa tcc caa tag gctgataatc acccaacaac ccaatatacc	1654
Asp Asp Glu Lys Ser Gln	
505	
aaccaagata ttgggttggtt gggcatcaat taccaccggt tagattgtgc agattgaatg	1714
tggttggtta	1724

33/189

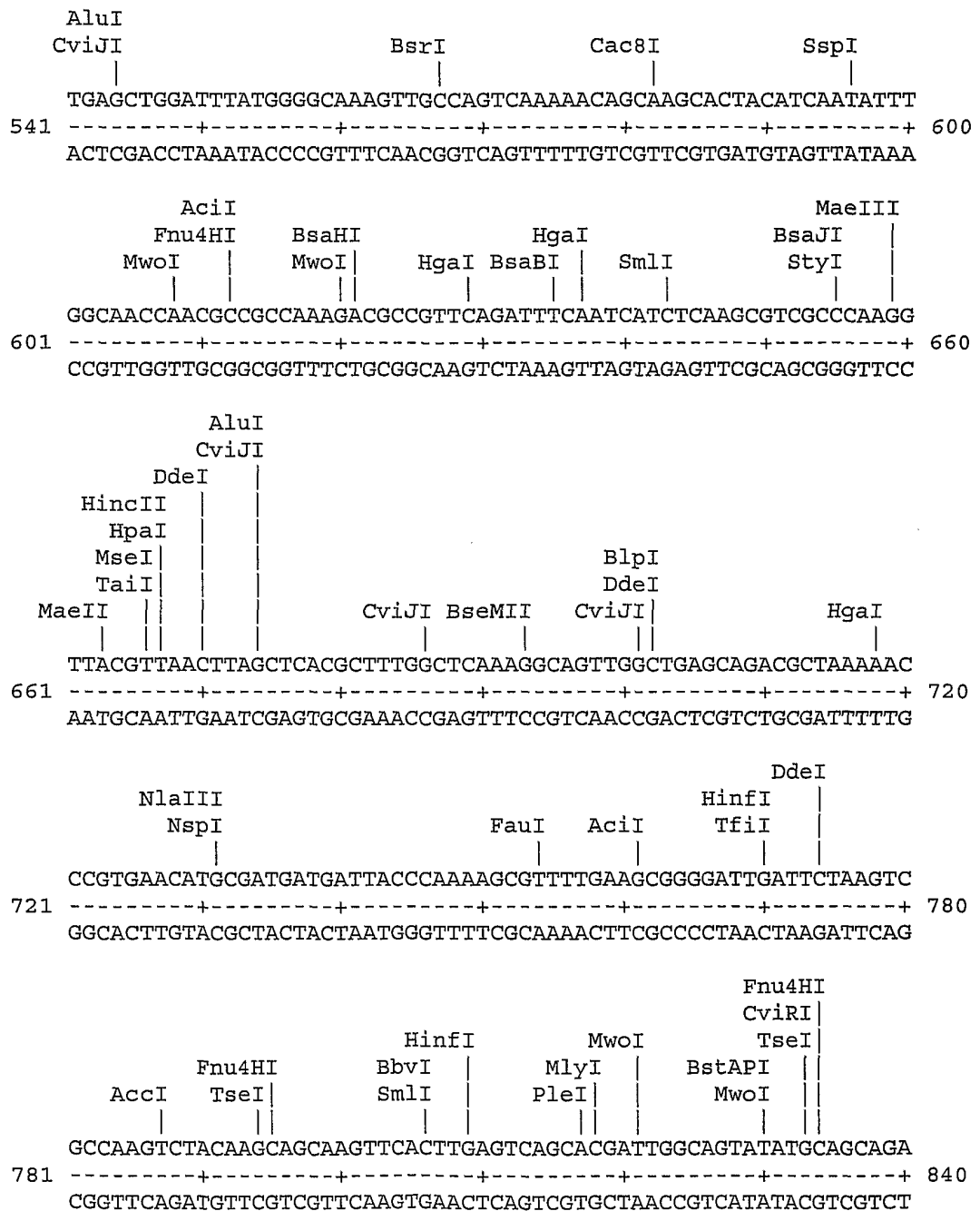
Figure 14. Restriction enzyme analysis of the *M. catarrhalis* OprM gene designated MCA100564



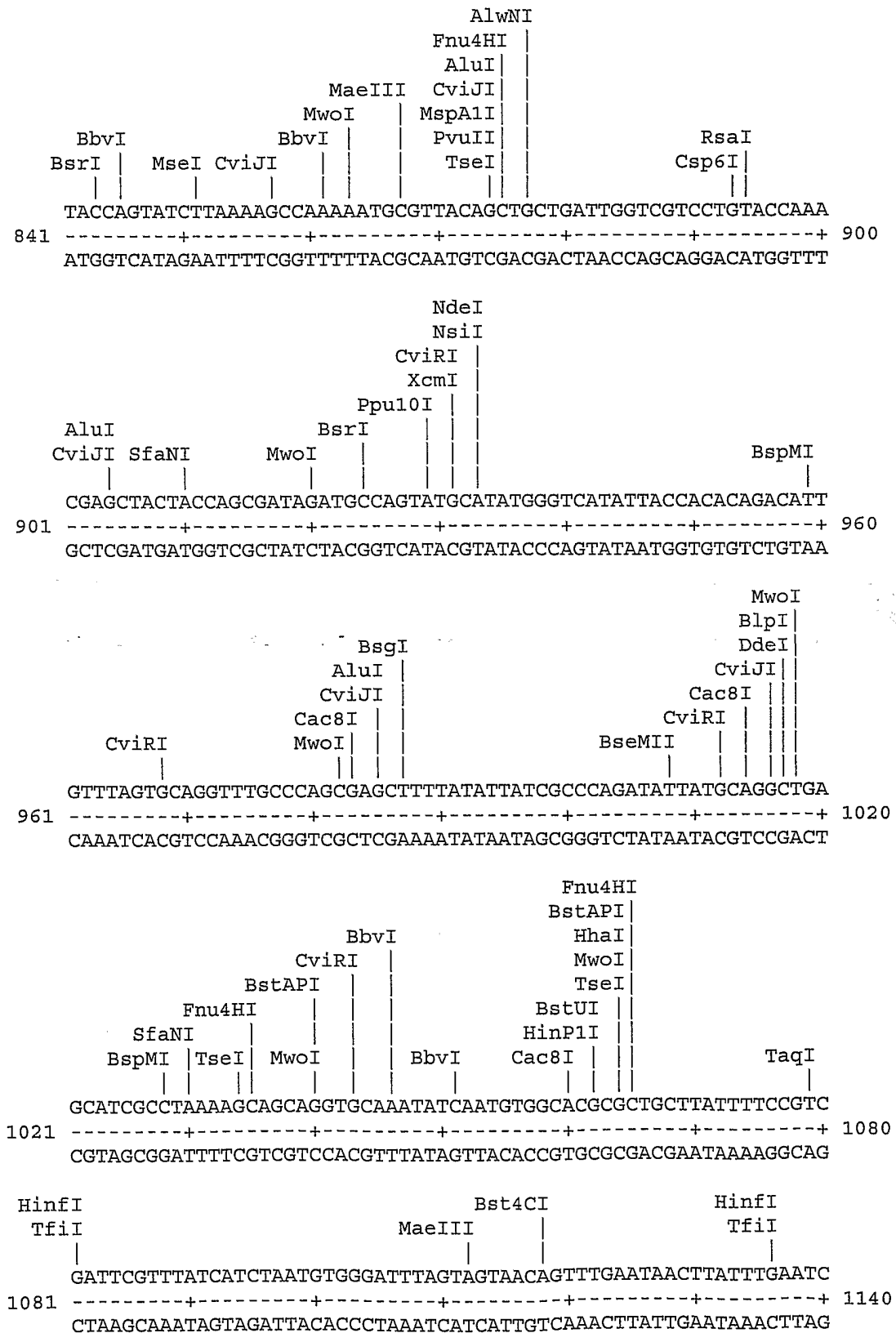
34/189



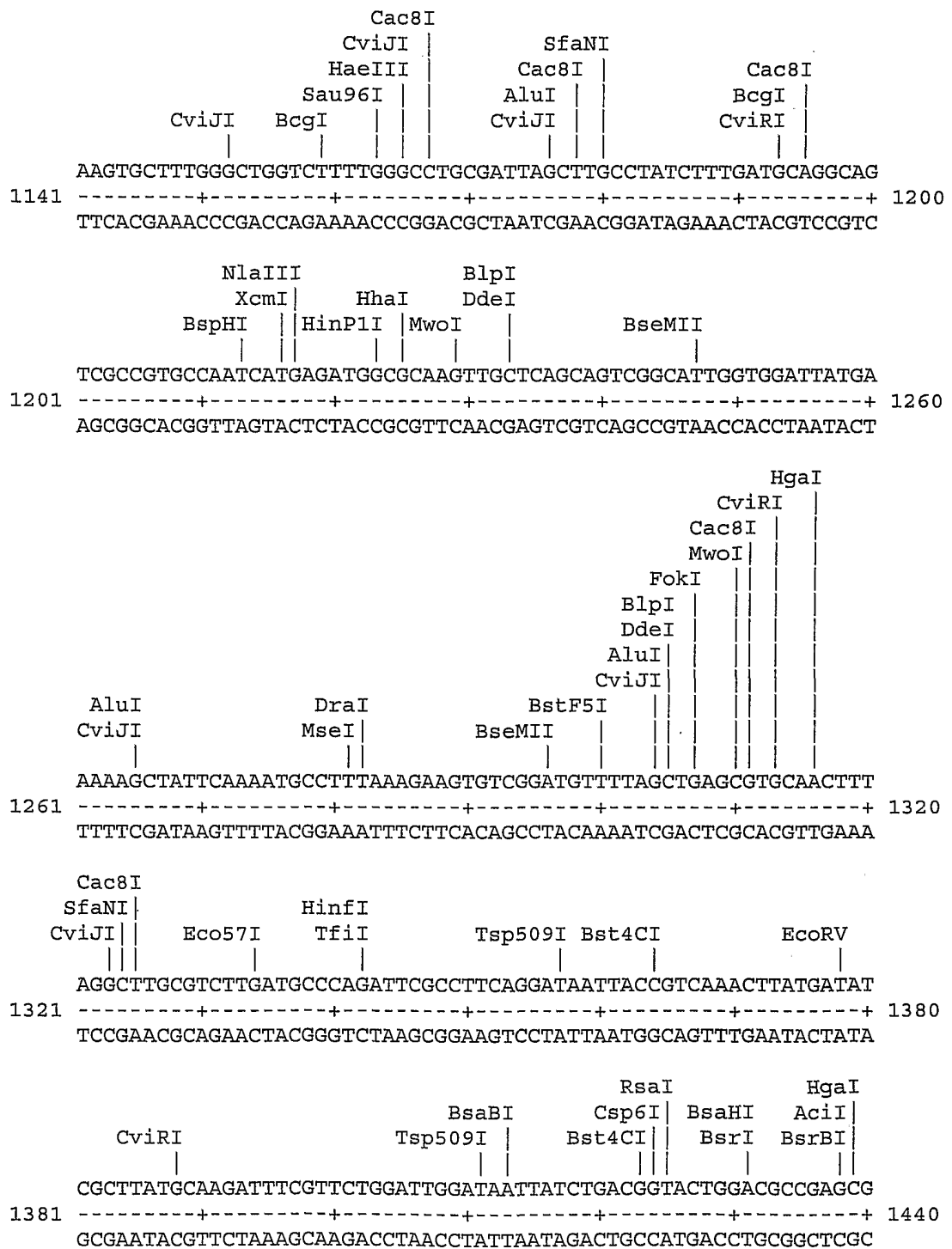
35/189



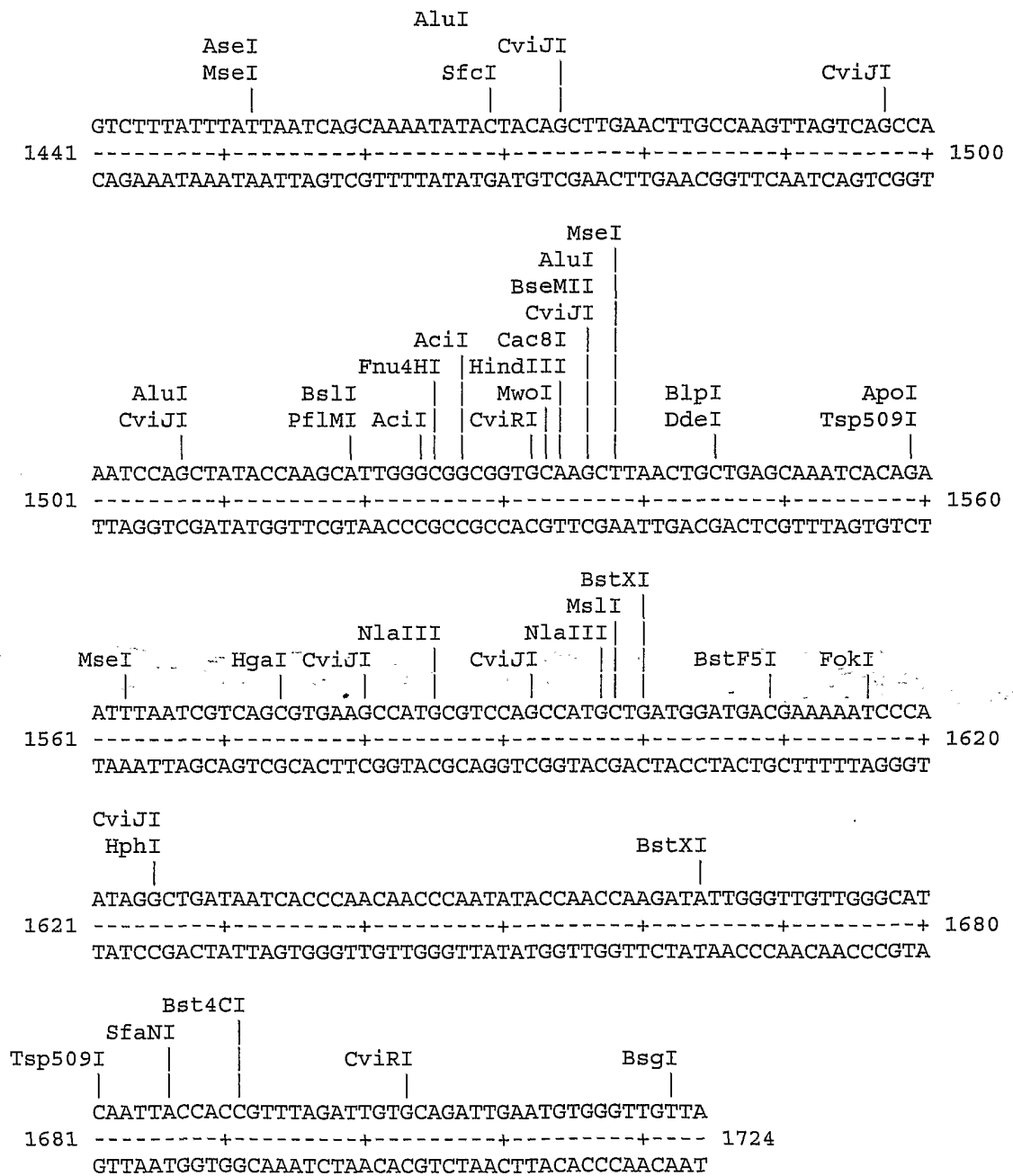
36/189



37/189



38/189



39/189

Figure 15. Sequence of *M. catarrhalis* collagenase gene designated MCA100567

```

taaaattaaa tttaatataa ttttggtgat gatacaaccc taaaagtaga aaaatttgct 60
ataatacaaaa taattttgac cataaaaaag agcatttcatt atg acc tcc cca atc 115
                                     Met Thr Ser Pro Ile
                                     1 5
gcc caa aaa tct gaa ctg ctt tgt cct gct ggc aca ttt aag aac atg 163
Ala Gln Lys Ser Glu Leu Leu Cys Pro Ala Gly Thr Phe Lys Asn Met
                                     10 15 20
caa tat gct ttt gct tat ggg gca gat gcg gtg tat gca ggg caa gcc 211
Gln Tyr Ala Phe Ala Tyr Gly Ala Asp Ala Val Tyr Ala Gly Gln Ala
                                     25 30 35
cgt tac tct ttg cgg gtg cgt aac aac gac ttt gat gaa gaa aat ttg 259
Arg Tyr Ser Leu Arg Val Arg Asn Asn Asp Phe Asp Glu Glu Asn Leu
                                     40 45 50
gca aaa ggc att gcc tac gcc cat tcg ctt ggc aaa aag ttt tat gtg 307
Ala Lys Gly Ile Ala Tyr Ala His Ser Leu Gly Lys Lys Phe Tyr Val
                                     55 60 65
gtg gtg aac att caa gca cac aac gcc aaa tta aaa acc ttt att gaa 355
Val Val Asn Ile Gln Ala His Asn Ala Lys Leu Lys Thr Phe Ile Glu
                                     70 75 80 85
gac att cgc cct gtc att gaa atg aaa cct gac gcg ctt atc atg tcc 403
Asp Ile Arg Pro Val Ile Glu Met Lys Pro Asp Ala Leu Ile Met Ser
                                     90 95 100
gat gct ggc atg att atg atg gtg cga gaa gct ttc ccc aat caa gtc 451
Asp Ala Gly Met Ile Met Met Val Arg Glu Ala Phe Pro Asn Gln Val
                                     105 110 115
att cac ctg tca gta caa gcc aat gcc gtg aac tgg gca acg gtg aag 499
Ile His Leu Ser Val Gln Ala Asn Ala Val Asn Trp Ala Thr Val Lys
                                     120 125 130
ttt tgg cag cag atg ggc att agc cgt gtg att tta agc cgt gaa tta 547
Phe Trp Gln Gln Met Gly Ile Ser Arg Val Ile Leu Ser Arg Glu Leu
                                     135 140 145
tca ctc aaa gaa att gag caa atc atc aaa gaa gtg cct gac atg gag 595
Ser Leu Lys Glu Ile Glu Gln Ile Ile Lys Glu Val Pro Asp Met Glu
                                     150 155 160 165
gtt gaa gtt ttt gtg cat ggg gcg ctg tgc atg gca tat tct ggg cgg 643
Val Glu Val Phe Val His Gly Ala Leu Cys Met Ala Tyr Ser Gly Arg
                                     170 175 180
tgc ttg ctc tct ggc tat atc aac aaa cga gac gcc aac caa ggc act 691
Cys Leu Leu Ser Gly Tyr Ile Asn Lys Arg Asp Ala Asn Gln Gly Thr
                                     185 190 195

```

40/189

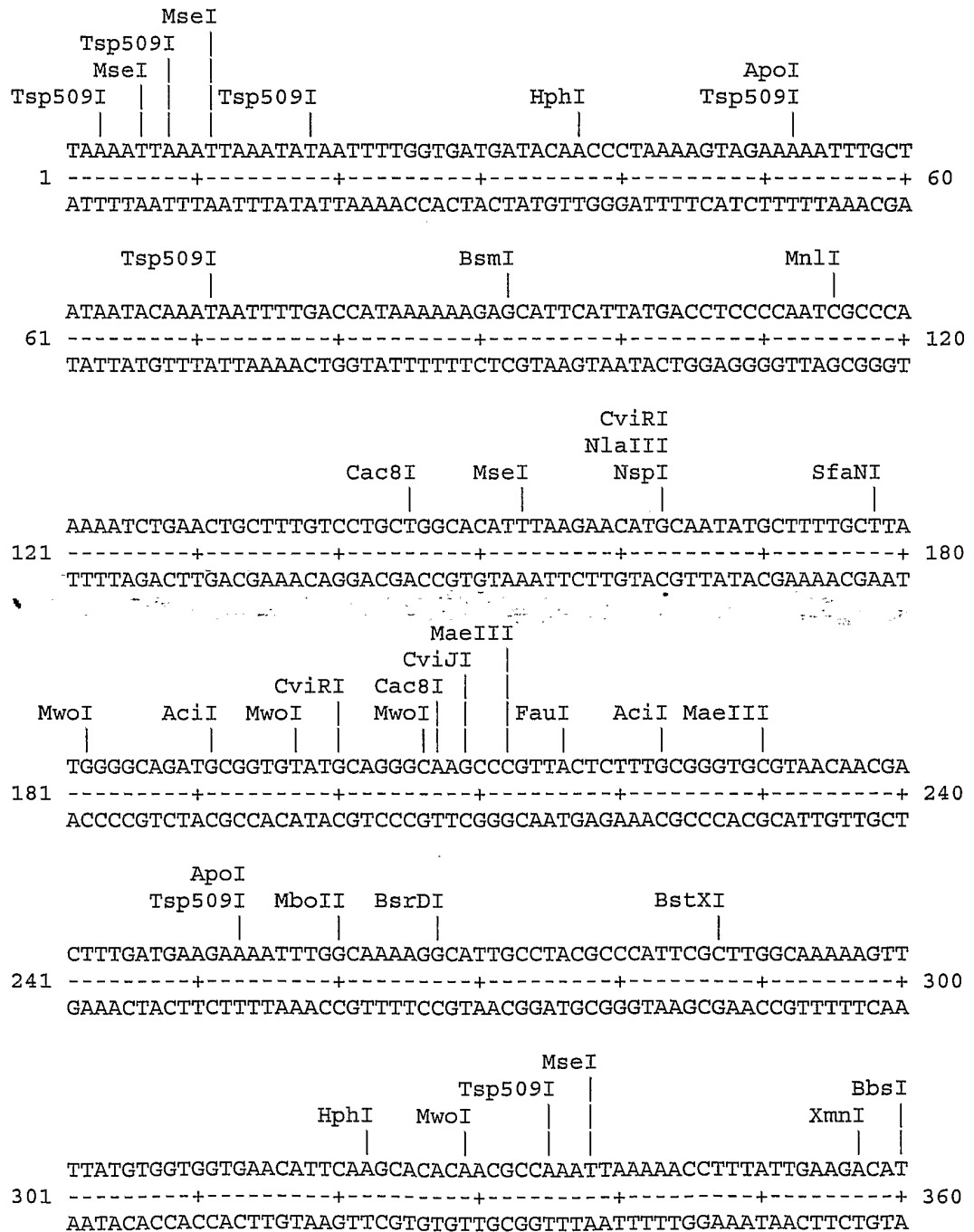
tgc acc aat gcc tgt cgc tgg aat tac aac act tat aaa gcg gtg gag	739
Cys Thr Asn Ala Cys Arg Trp Asn Tyr Asn Thr Tyr Lys Ala Val Glu	
200 205 210	
aat gaa aca ggc gat atc gtc cct gcc aac cca gaa gtg ttc gtc ccc	787
Asn Glu Thr Gly Asp Ile Val Pro Ala Asn Pro Glu Val Phe Val Pro	
215 220 225	
agt caa ggc gag att agc aat gtc aac cta aat ggc gat gtc gtc atg	835
Ser Gln Gly Glu Ile Ser Asn Val Asn Leu Asn Gly Asp Val Val Met	
230 235 240 245	
gac gat gat tat gta gaa aaa ccg tcc gat gag gtt gta ctc att gaa	883
Asp Asp Asp Tyr Val Glu Lys Pro Ser Asp Glu Val Val Leu Ile Glu	
250 255 260	
gaa caa ggt cgc aag ggt gag ctg atg gcg atg tat gaa gac gaa cat	931
Glu Gln Gly Arg Lys Gly Glu Leu Met Ala Met Tyr Glu Asp Glu His	
265 270 275	
ggc act tat atc atg aat tca aaa gac ttg cgt gcg gtg gag ctg gtg	979
Gly Thr Tyr Ile Met Asn Ser Lys Asp Leu Arg Ala Val Glu Leu Val	
280 285 290	
cct gat ttg gtc aat atg ggc gtg cat tct tta aaa att gag ggg cgt	1027
Pro Asp Leu Val Asn Met Gly Val His Ser Leu Lys Ile Glu Gly Arg	
295 300 305	
acc aaa tcg cat tat tat gtc gct cgc acc gct cag gtc tat cgc cgt	1075
Thr Lys Ser His Tyr Tyr Val Ala Arg Thr Ala Gln Val Tyr Arg Arg	
310 315 320 325	
gcc att gat gat gct ttg gca ggc aag cca ttt gat aca gga ctc atc	1123
Ala Ile Asp Asp Ala Leu Ala Gly Lys Pro Phe Asp Thr Gly Leu Ile	
330 335 340	
act gcg ctt gat ggg ctt gcc aat cgt ggt tat acc gaa gga ttt tta	1171
Thr Ala Leu Asp Gly Leu Ala Asn Arg Gly Tyr Thr Glu Gly Phe Leu	
345 350 355	
cgc cgt cat gtg cat tct gat tac caa aat tat gaa tac ggc tca tct	1219
Arg Arg His Val His Ser Asp Tyr Gln Asn Tyr Glu Tyr Gly Ser Ser	
360 365 370	
aag acc gac aat cag cag ttt gtc gct gaa gtg gcg gac atc agc gat	1267
Lys Thr Asp Asn Gln Gln Phe Val Ala Glu Val Ala Asp Ile Ser Asp	
375 380 385	
gag cgt ttg act ttg act gtc aaa aac aag ctg atg gtg ggc gat acc	1315
Glu Arg Leu Thr Leu Thr Val Lys Asn Lys Leu Met Val Gly Asp Thr	
390 395 400 405	
att gaa atc atg act ccg cag ggg aat gtt agc tac acg cta act gag	1363
Ile Glu Ile Met Thr Pro Gln Gly Asn Val Ser Tyr Thr Leu Thr Glu	
410 415 420	

41/189

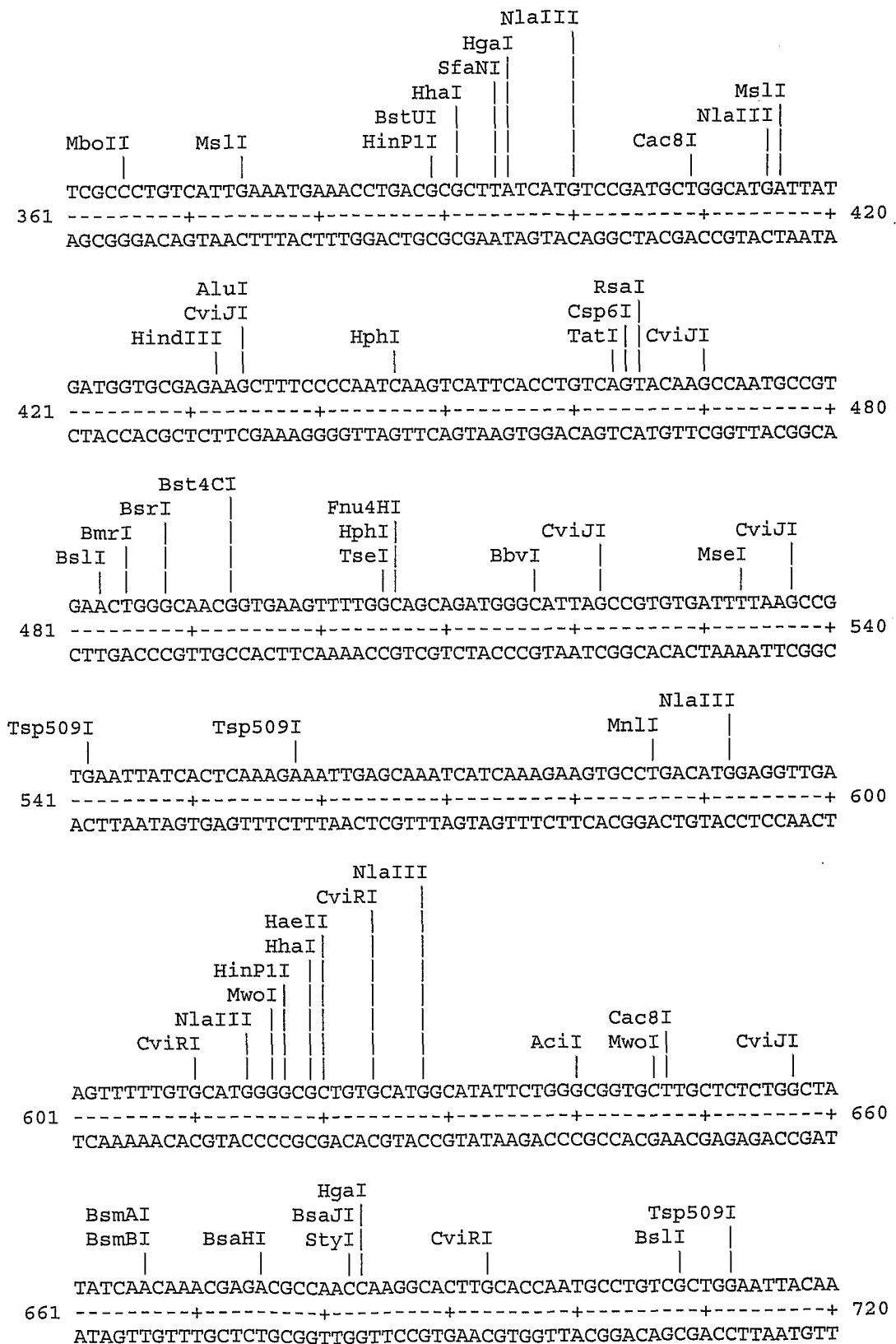
atg tgg gat aaa aaa ggc gag gcg att gat ggt gcg tta ggc tct ggt	1411
Met Trp Asp Lys Lys Gly Glu Ala Ile Asp Gly Ala Leu Gly Ser Gly	
425 430 435	
tggtgtgtgt caa att gcc aat ccc ttt aag gat atg gat aaa gac gtg	1459
Trp Val Cys Gln Ile Ala Asn Pro Phe Lys Asp Met Asp Lys Asp Val	
440 445 450	
cta aaa ttt gcc cta gtg atg aaa agc gtg agt gag gcg att tat agc	1507
Leu Lys Phe Ala Leu Val Met Lys Ser Val Ser Glu Ala Ile Tyr Ser	
455 460 465	
taa gcgggtgattg ttatatgcag ttaaaattta tggatttttg ttctggcata	1560
470	
ggtgcagggc gtctgggtct agaaaatgca ggcattgtgct gtgttgctca	1610

42/189

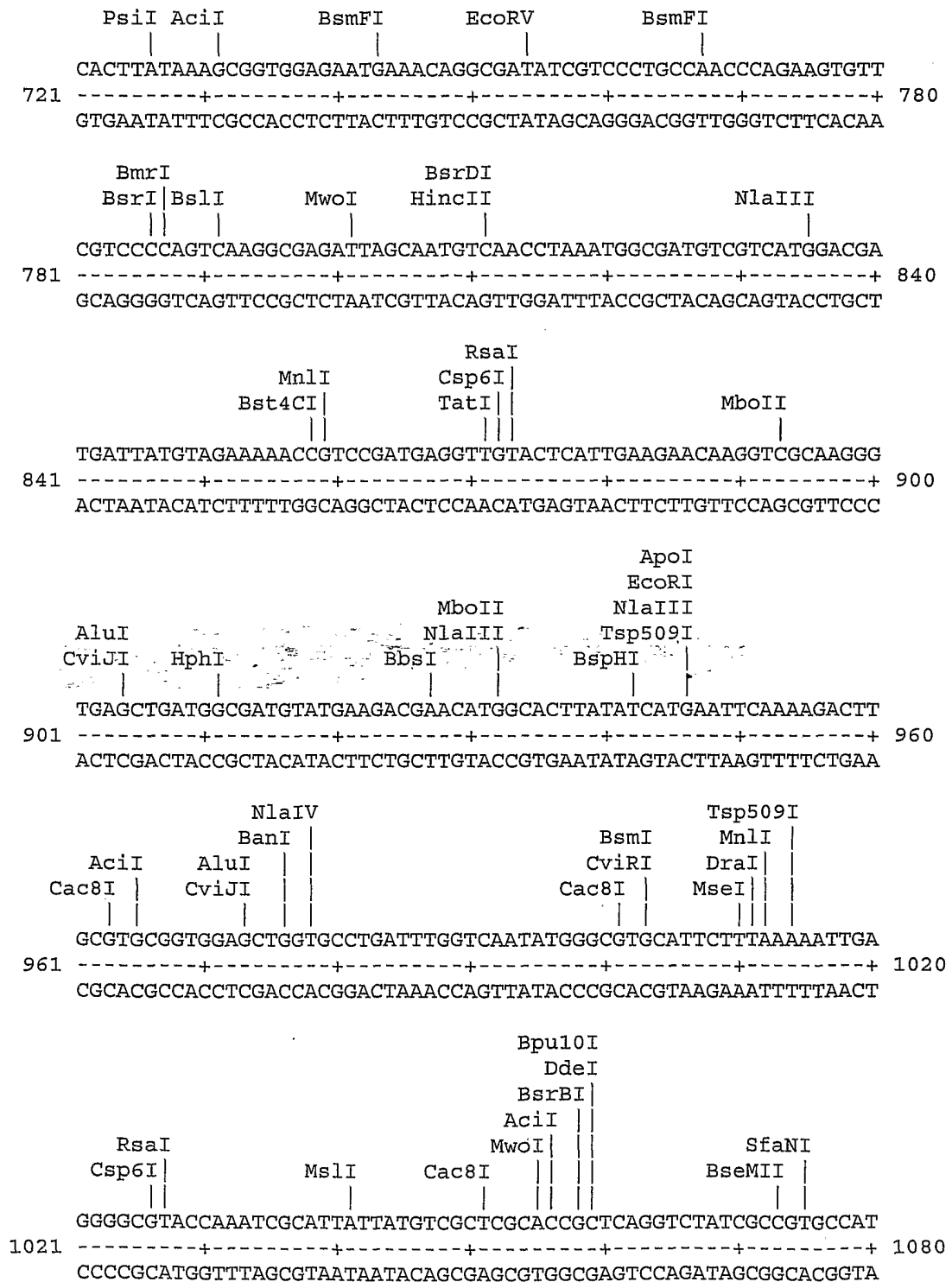
Figure 16. Restriction enzyme analysis of the *M. catarrhalis* collagenase gene designated MCA100567



43/189



44/189



45/189

CviJI
 Cac8I
 MlyI
 PleI
 HinfI
 BtsI
 HhaI
 TspRI
 HinPII
 CviJI
 MwoI
 Cac8I

1081
 TGATGATGCTTTGGCAGGCAAGCCATTTGATACAGGACTCATCACTGCGCTTGATGGGCT
 -----+-----+-----+-----+-----+-----+-----+
 ACTACTACGAAACCGTCCGTTCGGTAAACTATGTCCTGAGTAGTGACGCGAACTACCCGA
 1140

BsmI
 CviRI
 NlaIII
 Tsp509I

1141
 TGCCAATCGTGGTTATACCGAAGGATTTTTACGCCGTCACTGTCATTCTGATTACCAAAA
 -----+-----+-----+-----+-----+-----+-----+
 ACGGTTAGCACCAATATGGCTTCCTAAAAATGCGGCAGTACACGTAAGACTAATGGTTTT
 1200

CviJI
 DdeI
 MwoI
 AclI

1201
 TTATGAATACGGCTCATCTAAGACCGACAATCAGCAGTTTGTGCGTGAAGTGGCGGACAT
 -----+-----+-----+-----+-----+-----+-----+
 AATACTTATGCCGAGTAGATTCTGGCTGTTAGTCGTCAAACAGCGACTTCACCGCCTGTA
 1260

Eco57I
 EclI
 Bst4CI
 AluI
 CviJI

1261
 CAGCGATGAGCGTTTGACTTTGACTGTCAAAAACAAGCTGATGGTGGGCGATACCATTGA
 -----+-----+-----+-----+-----+-----+-----+
 GTCGCTACTCGCAAACCTGAAACTGACAGTTTTTTGTTTCGACTACCACCCGCTATGGTAACT
 1320

HinfI
 NlaIII
 BspHI
 MlyI
 PleI
 BseMII
 AluI
 CviJI
 DdeI
 MnlI
 AclI

1321
 AATCATGACTCCGCAGGGGAATGTTAGCTACACGCTAACTGAGATGTGGGATAAAAAAGG
 -----+-----+-----+-----+-----+-----+-----+
 TTAGTACTGAGGCGTCCCCTTACAATCGATGTGCGATTGACTCTACACCCCTATTTTTTCC
 1380

CviJI
 Tsp509I
 MseI

1381
 CGAGGCGATTGATGGTGCCTTAGGCTCTGGTTGGGTGTGTCAAATTGCCAATCCCTTTAA
 -----+-----+-----+-----+-----+-----+-----+
 GCTCCGCTAACTACCACGCAATCCGAGACCAACCCACACAGTTTAAACGGTTAGGGAAATT
 1440

ApoI
 Tsp509I
 TaiI
 BtrI
 MaeII
 BfaI
 MnlI

1441
 GGATATGGATAAAGACGTGCTAAAATTTGCCCTAGTGATGAAAAGCGTGAGTGAGGCGAT
 -----+-----+-----+-----+-----+-----+-----+
 CCTATACCTATTTCTGCACGATTTTAAACGGGATCACTACTTTTTCGCACTCACTCCGCTA
 1500

```

      BlnI       ApoI
      DdeI       Tsp509I
    AluI|         CviRI
CviJI|AciI     HphI|MseI
      |||        |
TTATAGCTAAGCGGTGATTGTTATATGCAGTTAAAAATTTATGGATTTTTGTTCTGGCATA
1501 -----+-----+-----+-----+-----+ 1560
AAATATCGATTTCGCCACTAACAATATACGTCAATTTTAAATACCTAAAAACAAGACCGTAT

              Cac8I
          BsaHI   BfaI   CviRI   NlaIII
CviRI|       XbaI|BsgI|       NspI
      |         |||      |
GGTGCAGGGCGTCTGGGTCTAGAAAATGCAGGCATGTGCTGTGTTGCTCA
1561 -----+-----+-----+-----+-----+ 1610
CCACGTCCC GCAGACCCAGATCTTTTACGTCCGTACACGACACAACGAGT

```

47/189

Figure 17. Sequence of *M. catarrhalis* outer membrane phospholipase A gene designated MCA100712

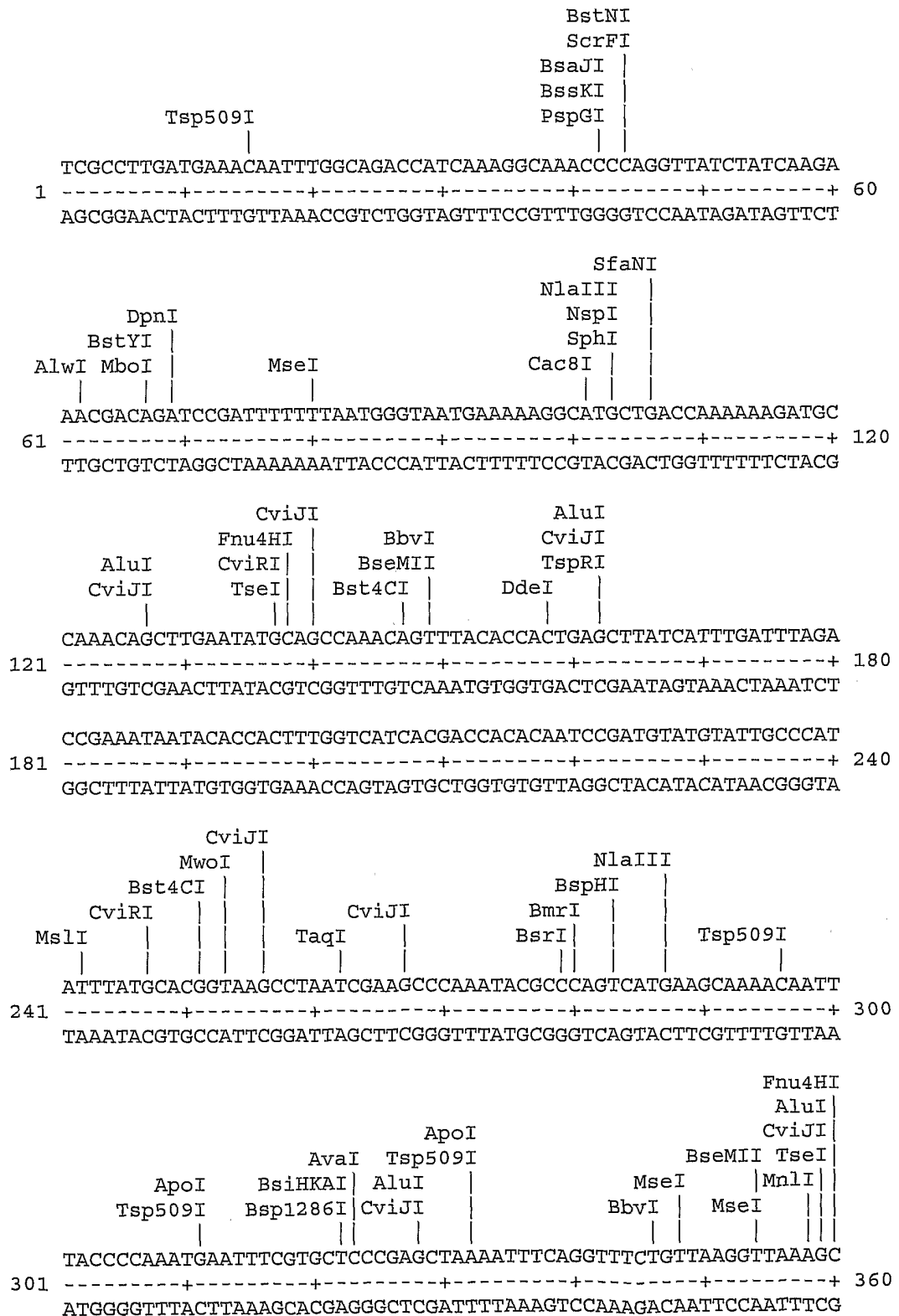
tcgccttgat	gaaacaattt			ggcagaccat			caaaggcaaa			cccaggtta			tctatcaaga			60
aacgacagat	ccgatttttt			taatgggtaa			tgaaaaaggc			atg	ctg	acc	aaa	aaa	115	
										Met	Leu	Thr	Lys	Lys		
										1				5		
gat	gcc	aaa	cag	ctt	gaa	tat	gca	gcc	aaa	cag	ttt	aca	cca	ctg	agc	163
Asp	Ala	Lys	Gln	Leu	Glu	Tyr	Ala	Ala	Lys	Gln	Phe	Thr	Pro	Leu	Ser	
				10					15					20		
tta	tca	ttt	gat	tta	gac	cga	aat	aat	aca	cca	ctt	tgg	tca	tca	cga	211
Leu	Ser	Phe	Asp	Leu	Asp	Arg	Asn	Asn	Thr	Pro	Leu	Trp	Ser	Ser	Arg	
			25					30					35			
cca	cac	aat	ccg	atg	tat	gta	ttg	ccc	ata	ttt	atg	cac	ggg	aag	cct	259
Pro	His	Asn	Pro	Met	Tyr	Val	Leu	Pro	Ile	Phe	Met	His	Gly	Lys	Pro	
		40					45					50				
aat	cga	agc	cca	aat	acg	ccc	agt	cat	gaa	gca	aaa	caa	ttt	acc	cca	307
Asn	Arg	Ser	Pro	Asn	Thr	Pro	Ser	His	Glu	Ala	Lys	Gln	Phe	Thr	Pro	
	55					60					65					
aat	gaa	ttt	cgt	gct	ccc	gag	cta	aaa	ttt	cag	gtt	tct	gtt	aag	gtt	355
Asn	Glu	Phe	Arg	Ala	Pro	Glu	Leu	Lys	Phe	Gln	Val	Ser	Val	Lys	Val	
70					75					80					85	
aaa	gct	gct	gag	gat	tta	tgg	ggg	acg	gat	tca	gat	tta	tgg	ttt	gga	403
Lys	Ala	Ala	Glu	Asp	Leu	Trp	Gly	Thr	Asp	Ser	Asp	Leu	Trp	Phe	Gly	
			90						95					100		
tat	aca	cag	caa	tcg	cac	tgg	cag	att	ttt	aat	gga	aaa	aac	tct	cgt	451
Tyr	Thr	Gln	Gln	Ser	His	Trp	Gln	Ile	Phe	Asn	Gly	Lys	Asn	Ser	Arg	
			105					110					115			
cct	ttt	aga	gta	cat	gac	tac	cag	cca	gag	att	ttc	tta	act	caa	cct	499
Pro	Phe	Arg	Val	His	Asp	Tyr	Gln	Pro	Glu	Ile	Phe	Leu	Thr	Gln	Pro	
		120					125					130				
gta	tac	tca	gac	tta	cca	tgg	gat	ggc	aaa	gtc	cgc	atg	att	ggc	atg	547
Val	Tyr	Ser	Asp	Leu	Pro	Trp	Asp	Gly	Lys	Val	Arg	Met	Ile	Gly	Met	
	135					140				145						
ggg	gcg	gta	cat	cat	tcc	aat	ggg	gaa	agt	gcc	aaa	ctg	tct	cgc	tca	595
Gly	Ala	Val	His	His	Ser	Asn	Gly	Glu	Ser	Ala	Lys	Leu	Ser	Arg	Ser	
150					155					160					165	
tgg	aat	cgt	gct	tat	ttg	atg	gca	ggc	atg	gaa	tgg	aaa	aac	ctg	act	643
Trp	Asn	Arg	Ala	Tyr	Leu	Met	Ala	Gly	Met	Glu	Trp	Lys	Asn	Leu	Thr	
				170					175					180		
gtc	atg	cca	cgc	att	tgg	ggg	cgt	atc	ttt	aaa	gag	ggg	agt	ggc	agc	691
Val	Met	Pro	Arg	Ile	Trp	Gly	Arg	Ile	Phe	Lys	Glu	Gly	Ser	Gly	Ser	
			185					190					195			

48/189

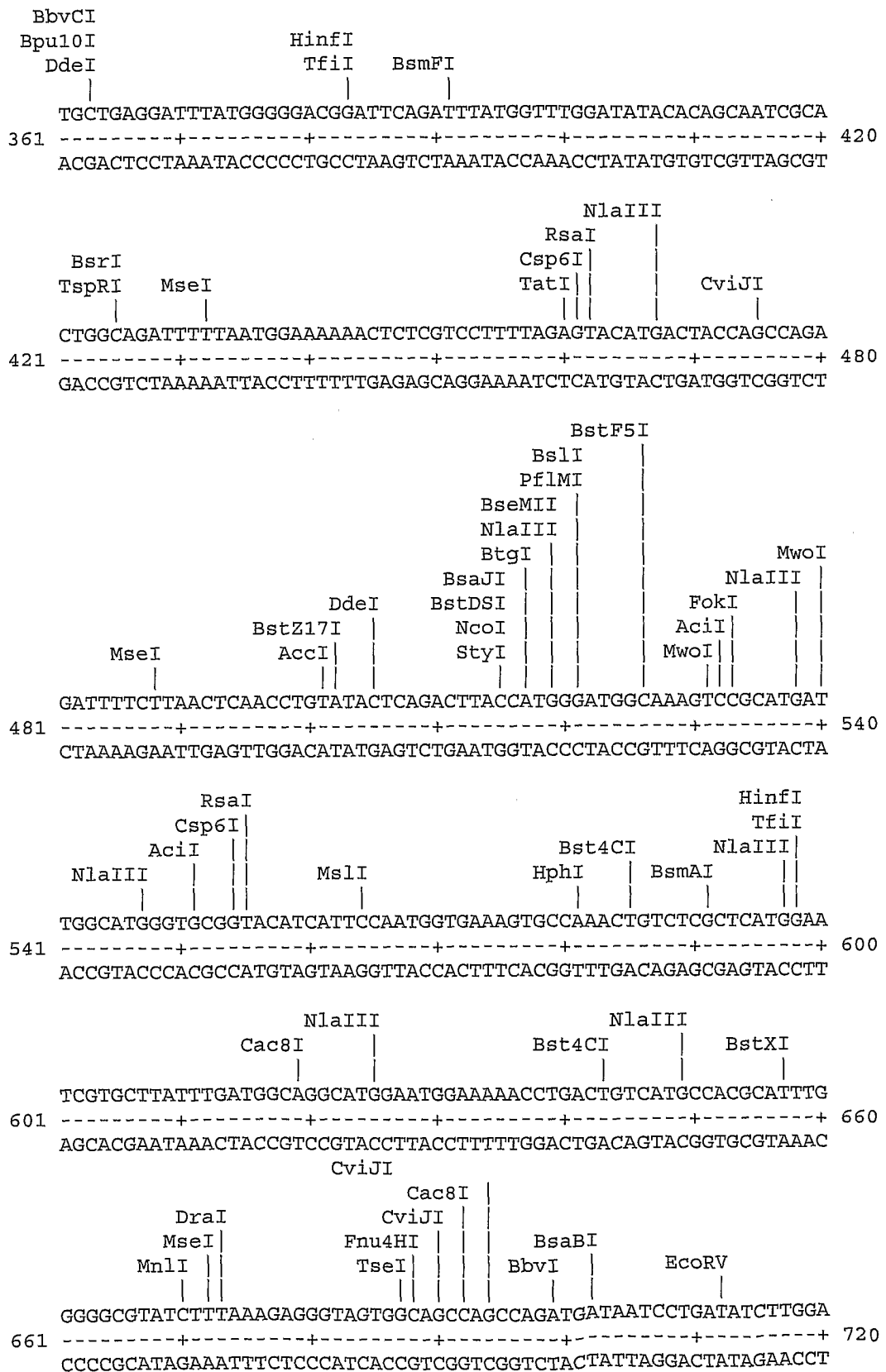
cag cca gat gat aat cct gat atc ttg gac tat tat ggt tat ggt gat	739
Gln Pro Asp Asp Asn Pro Asp Ile Leu Asp Tyr Tyr Gly Tyr Gly Asp	
200 205 210	
gtg cgt ttt tta tat caa cta gaa aat aaa agt aat att tca ggt acg	787
Val Arg Phe Leu Tyr Gln Leu Glu Asn Lys Ser Asn Ile Ser Gly Thr	
215 220 225	
gta cgc tat aat cca cgc tca ggc aaa ggt gcg ttg caa ctt gac tat	835
Val Arg Tyr Asn Pro Arg Ser Gly Lys Gly Ala Leu Gln Leu Asp Tyr	
230 235 240 245	
gtc tat ccg ctt ggt aag gga att agt ggc tat ttt caa ata ttt caa	883
Val Tyr Pro Leu Gly Lys Gly Ile Ser Gly Tyr Phe Gln Ile Phe Gln	
250 255 260	
ggc tat ggg cag tct ttg att gat tat aat cat gag gcg aca agc ttt	931
Gly Tyr Gly Gln Ser Leu Ile Asp Tyr Asn His Glu Ala Thr Ser Phe	
265 270 275	
ggc gtc gga ctt atg ctt aac gac tgg atg ggt cta taa ctcaacggca	980
Gly Val Gly Leu Met Leu Asn Asp Trp Met Gly Leu	
280 285 290	
aatcaatatg gaacacggct tgataaaactt gctaaaaaac cataagcata agcgggcttt	1040
taatcagggc gttttagcta agcctgtatt tgat	1074

49/189

Figure 18. Restriction enzyme analysis of the *M. catarrhalis* outer membrane phospholipase A gene designated MCA100712



50/189



51/189

CTATTATGGTTATGGTGATGTGCGTTTTTTATATCAACTAGAAAATAAAAGTAATATTTTC
721 -----+-----+-----+-----+-----+-----+-----+ 780
GATAATACCAATACCACTACACGCAAAAAATATAGTTGATCTTTTATTTTTCATTATAAAG

RsaI
Csp6I
Bst4CI
RsaI
Csp6I
BslI
Bpu10I
DdeI
CviRI
BseMII
Tth111I
AGGTACGGTACGCTATAATCCACGCTCAGGCAAAGGTGCGTTGCAACTTGACTATGTCTA
781 -----+-----+-----+-----+-----+-----+ 840
TCCATGCCATGCGATATTAGGTGCGAGTCCGTTCCACGCAACGTTGAACTGATACAGAT

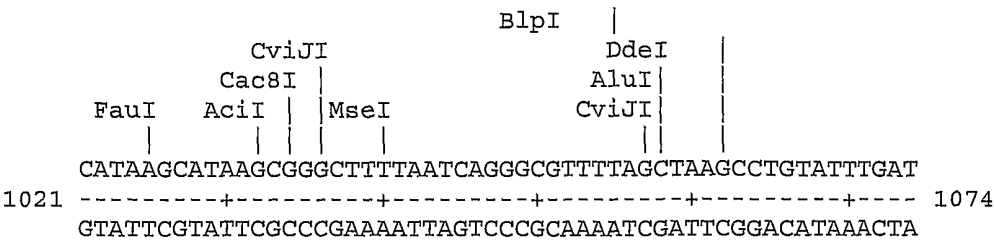
AciI
Tsp509I
CviJI
SspI
CviJI
TCCGCTTGGTAAGGGAATTAGTGGCTATTTTCAAATATTTCAAGGCTATGGGCAGTCTTT
841 -----+-----+-----+-----+-----+-----+ 900
AGGCGAACCATTCCCTTAATCACCGATAAAAGTTTATAAAGTTCCGATACCCGTCAGAAA

NlaIII
BspHI
MnlI
BsaBI
PsiI
AluI
CviJI
HindIII
HgaI
BsaHI
MseI
BsrI
GATTGATTATAATCATGAGGCGACAAGCTTTGGCGTCGGACTTATGCTTAACGACTGGAT
901 -----+-----+-----+-----+-----+-----+ 960
CTAACTAATATTAGTACTCCGCTGTTTCGAAACCGCAGCCTGAATACGAATTGCTGACCTA

BstF5I
FokI
CviJI
GGGTCTATAACTCAACGGCAAATCAATATGGAACACGGCTTGATAAACTTGCTAAAAAAC
961 -----+-----+-----+-----+-----+-----+ 1020
CCCAGATATTGAGTTGCCGTTTGTGACCTTGTGCGGAACTATTTGAACGATTTTTTG

CviJI

52/189



53/189

Figure 19. Sequence of *M. catarrhalis* gene encoding a lipoprotein shown in SEQ ID NOs: 19 & 20 designated MCA100800

```

aatagggtttt gtgtgttttat acagccaata gacattcatt gtcagacca taaaaacatg 60
ctaaactagc ccaaatttag accaattcat gagaattcct atg aaa tca caa aac 115
                                         Met Lys Ser Gln Asn
                                         1           5

aag ctg att gcc gcc atg agg caa ccg act tgg act gtt aca ggt tgc 163
Lys Leu Ile Ala Ala Met Arg Gln Pro Thr Trp Thr Val Thr Gly Cys
              10              15              20

aaa agc gcg att gga ctg ttt gac cgc aag gct gat acc gta gaa acg 211
Lys Ser Ala Ile Gly Leu Phe Asp Arg Lys Ala Asp Thr Val Glu Thr
              25              30              35

gct gaa aaa aca gat gct gca tat tat caa gaa gcc agt gaa gca ctt 259
Ala Glu Lys Thr Asp Ala Ala Tyr Tyr Gln Glu Ala Ser Glu Ala Leu
              40              45              50

gat aaa aat caa aat cgc aac gcc att gaa gca cta aat aac atc cga 307
Asp Lys Asn Gln Asn Arg Asn Ala Ile Glu Ala Leu Asn Asn Ile Arg
              55              60              65

acc ttt tat ccg aca ggt caa tat gca cag cag gca ctg cta gat ctc 355
Thr Phe Tyr Pro Thr Gly Gln Tyr Ala Gln Gln Ala Leu Leu Asp Leu
              70              75              80              85

att tat gcc caa tat aaa gcc aat gat ttt gaa gca gtt cta caa agt 403
Ile Tyr Ala Gln Tyr Lys Ala Asn Asp Phe Glu Ala Val Leu Gln Ser
              90              95              100

aca gaa gag ttt att cac cga tac cca aac agc cgc tcg gta gac tat 451
Thr Glu Glu Phe Ile His Arg Tyr Pro Asn Ser Arg Ser Val Asp Tyr
              105              110              115

gcg tta tat gtc caa ggc gtt act cat atg ggc ggc gct ccc aaa gct 499
Ala Leu Tyr Val Gln Gly Val Thr His Met Gly Gly Ala Pro Lys Ala
              120              125              130

tct cgt ttg gtt aga ttt gat caa tct cat cgt gat gtg act tat cta 547
Ser Arg Leu Val Arg Phe Asp Gln Ser His Arg Asp Val Thr Tyr Leu
              135              140              145

cgc cta gct ttc cga gat ttt caa aac ctg ctt aag cac tat cct gat 595
Arg Leu Ala Phe Arg Asp Phe Gln Asn Leu Leu Lys His Tyr Pro Asp
              150              155              160              165

agt cca tat gtt gcc gat gcc gca caa cgc atg att gcc att tat aat 643
Ser Pro Tyr Val Ala Asp Ala Ala Gln Arg Met Ile Ala Ile Tyr Asn
              170              175              180

gac ttt gct gag cat gaa ttg gcg gca gcg cgt tgg tat gtc aag cgt 691
Asp Phe Ala Glu His Glu Leu Ala Ala Ala Arg Trp Tyr Val Lys Arg
              185              190              195

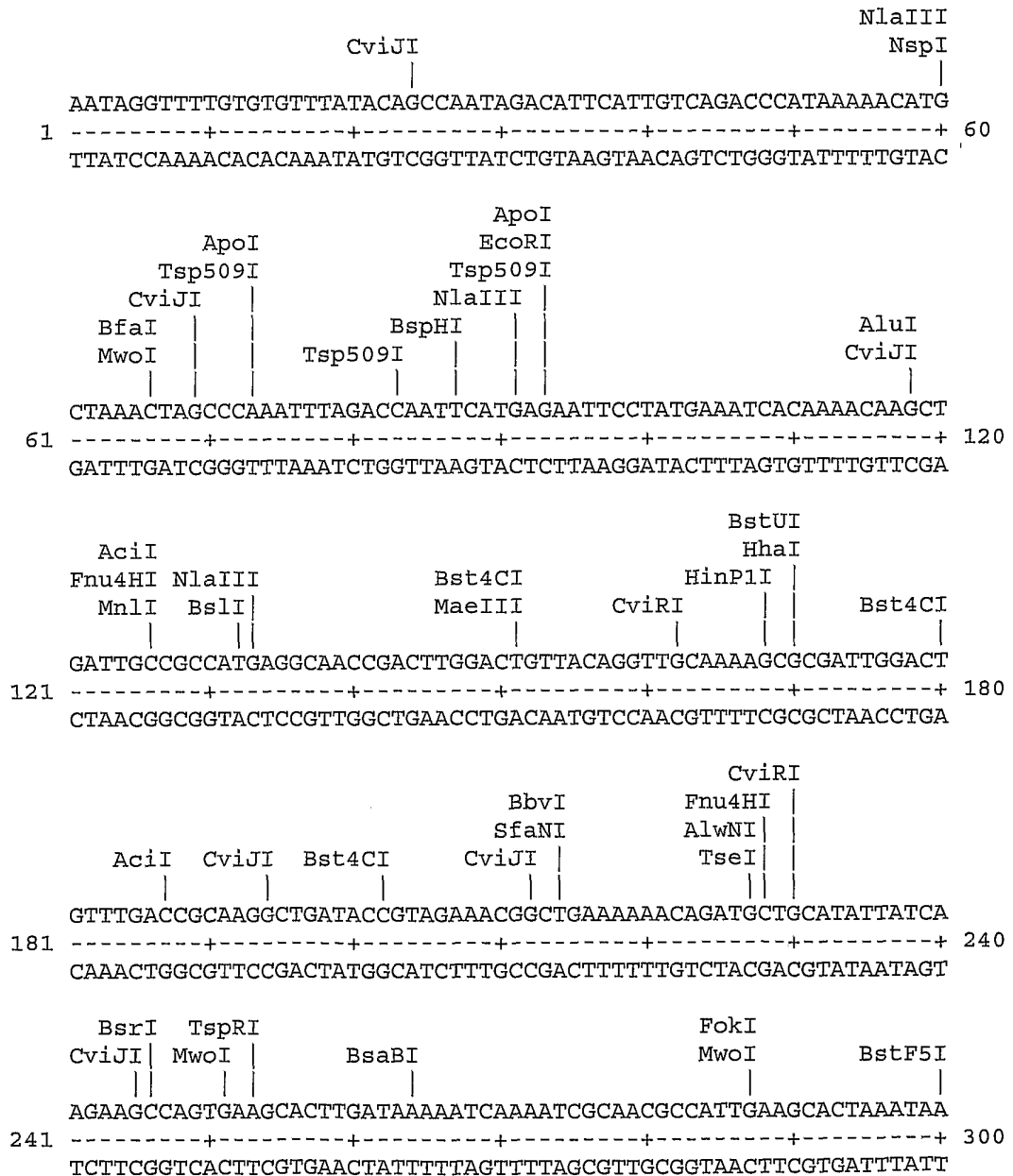
```

54/189

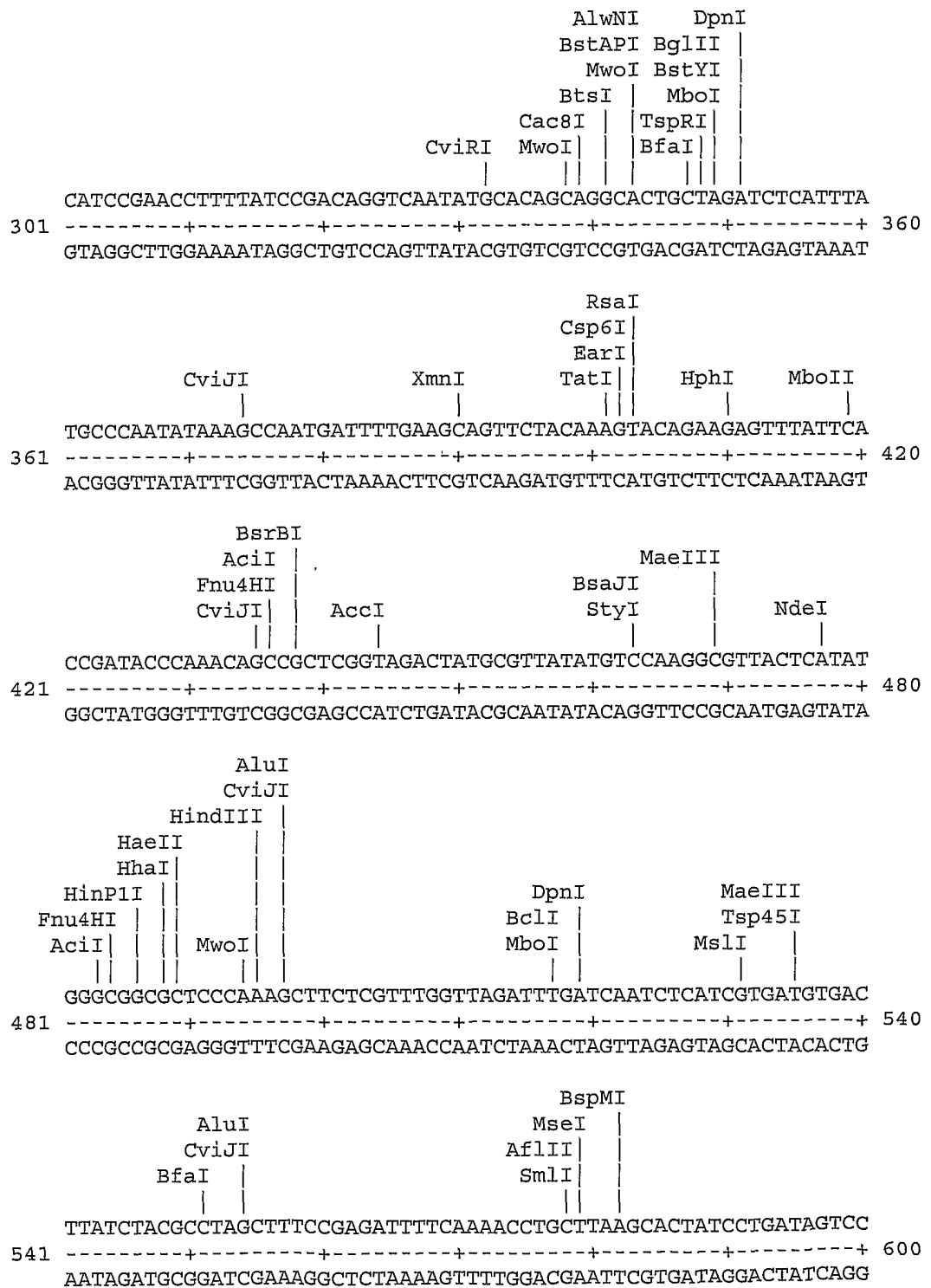
gat gcg atg gtg gct gct gcc aac cga gct aag tgg gtg ttt caa tac	739
Asp Ala Met Val Ala Ala Ala Asn Arg Ala Lys Trp Val Phe Gln Tyr	
200 205 210	
tat cca cag tcc aca ggt gtg cca gaa gct atc gca att ttg gct tat	787
Tyr Pro Gln Ser Thr Gly Val Pro Glu Ala Ile Ala Ile Leu Ala Tyr	
215 220 225	
agc aat caa aag ctt ggc ttg act gag acc gcc aaa cag tac aag cag	835
Ser Asn Gln Lys Leu Gly Leu Thr Glu Thr Ala Lys Gln Tyr Lys Gln	
230 235 240 245	
tta ctc caa att aat tat ccg caa tat ctc aat caa gat ggt tca gta	883
Leu Leu Gln Ile Asn Tyr Pro Gln Tyr Leu Asn Gln Asp Gly Ser Val	
250 255 260	
cgc ctc agt gat acg ggt gcc aag aca cgc aca caa aag gta tta tca	931
Arg Leu Ser Asp Thr Gly Ala Lys Thr Arg Thr Gln Lys Val Leu Ser	
265 270 275	
gca cta agc ttt ggg cgt ttg ggg cgt gtc agc gaa aat ctt gac aac	979
Ala Leu Ser Phe Gly Arg Leu Gly Arg Val Ser Glu Asn Leu Asp Asn	
280 285 290	
tat caa ggt caa tat cat ggc gca aca cgc acc caa atc atc aac caa	1027
Tyr Gln Gly Gln Tyr His Gly Ala Thr Arg Thr Gln Ile Ile Asn Gln	
295 300 305	
gca ggt gcg att agc tta cct gcc aca acc caa gtc cca gaa gaa att	1075
Ala Gly Ala Ile Ser Leu Pro Ala Thr Thr Gln Val Pro Glu Glu Ile	
310 315 320 325	
cct aca gct ggt ggt aca aat cat cgc cct gcc atc agt tta gac ttg	1123
Pro Thr Ala Gly Gly Thr Asn His Arg Pro Ala Ile Ser Leu Asp Leu	
330 335 340	
cca tca gca gag gca gac gct gga caa tta aat caa ata cca cgc	1168
Pro Ser Ala Glu Ala Asp Ala Gly Gln Leu Asn Gln Ile Pro Arg	
345 350 355	
taatgccagt ttttaagcaa gttatcattt gttaattata accttatggt aattataatc	1228
tcaggcgggtg tatggctcgcc tgagattatt acatcaattg gaattttgga catgaat	1285

55/189

Figure 20. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a lipoprotein shown in SEQ ID NO: 19 designated MCA100800



56/189



57/189

NdeI AciI BseMII BlpI NlaIII
SfaNI Fnu4HI PsiI DdeI Tsp509I

ATATGTTGCCGATGCCGCACAACGCATGATTGCCATTTATAATGACTTTGCTGAGCATGA
601 -----+-----+-----+-----+-----+-----+ 660
TATACAACGGCTACGGCGTGTTGCGTACTAACGGTAATATTACTGAAACGACTCGTACT

BstUI
HhaI Fnu4HI
HinPII Fnu4HI TseI DdeI
Fnu4HI TseI CviJI AluI
Fnu4HI AcII SfaNI BbvI TseI CviJI
AcII BbvI BbvI MwoI MwoI

ATTGGCGGCAGCGCGTTGGTATGTCAAGCGTGATGCGATGGTGGCTGCTGCCAACCGAGC
661 -----+-----+-----+-----+-----+-----+ 720
TAACCGCCGTCGCGCAACCATACAGTTCGCACTACGCTACCACCGACGACGGTTGGCTCG

Tsp509I
AluI MwoI
Bst4CI MslI CviJI

TAAGTGGGTGTTTTCAATACTATCCACAGTCCACAGGTGTGCCAGAAGCTATCGCAATTTT
721 -----+-----+-----+-----+-----+-----+ 780
ATTCACCCACAAAGTTATGATAGGTGTCAGGTGTCCACACGGTCTTCGATAGCGTTAAAA

BsaI
BsmAI
CviJI RsaI
BseMII AluI Csp6I
AluI CviJI Bst4CI TatI MaeIII
CviJI HindIII DdeI AcII

GGCTTATAGCAATCAAAGCTTGGCTTGACTGAGACCGCCAAACAGTACAAGCAGTTACT
781 -----+-----+-----+-----+-----+-----+ 840
CCGAATATCGTTAGTTTTTCGAACCGAACTGACTCTGGCGGTTTGTTCATGTTTCGTCAATGA

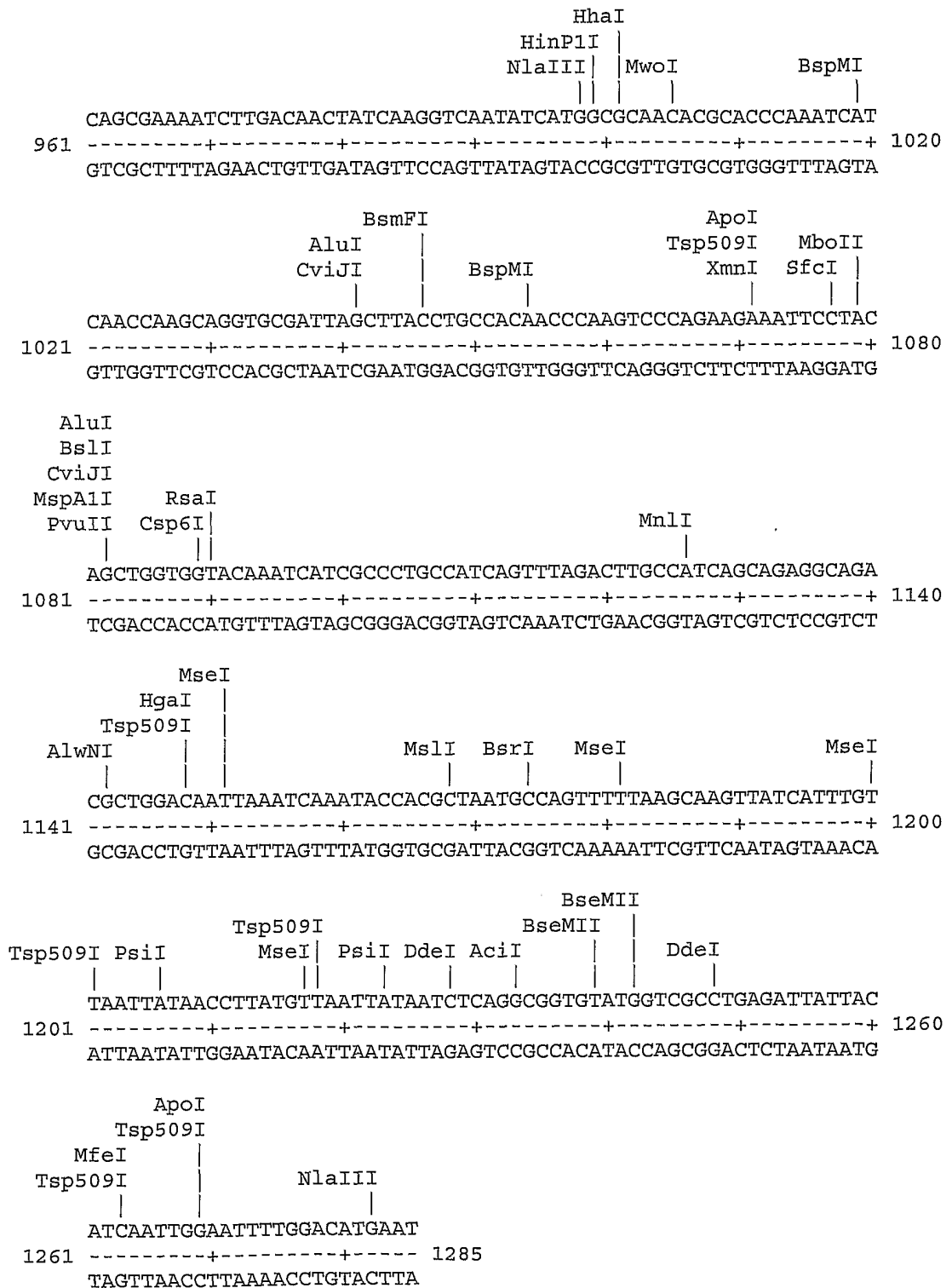
Tsp509I
AseI BanI
MseI RsaI MnlI
Tsp509I AcII Csp6I DdeI TspRI

CCAAAATTAATTATCCGCAATATCTCAATCAAGATGGTTCAGTAGCCTCAGTGATACGGG
841 -----+-----+-----+-----+-----+-----+ 900
GGTTTAATTAATAGGCGTTATAGAGTTAGTTCTACCAAGTCATGCGGAGTCACTATGCC

AluI
CviJI
BseMII HindIII MwoI
NlaIV DdeI

TGCCAAGACACGCACACAAAAGGTATTATCAGCACTAAGCTTTGGGCGTTTGGGGCGTGT
901 -----+-----+-----+-----+-----+-----+ 960
ACG GTTCTGTGCGTG TGT TTTCCATAATAGTCGTGATTCGAAACCCGAAACCCGCACA

58/189



59/189

Figure 21. Sequence of *M. catarrhalis* OmpA gene designated MCA100854

```

tggtagtga gagcttgcag gtcatatccg tgcagctatt gctcaagtct ttggcgtaca 60

agatttaacc attcaaaata ccaatgtaca taccgctacg atg cca gcg gca gaa 115
                                         Met Pro Ala Ala Glu
                                         1           5

tac tta cca gca att ttg ggt ttg atg aaa ggt gta cca aat tca agc 163
Tyr Leu Pro Ala Ile Leu Gly Leu Met Lys Gly Val Pro Asn Ser Ser
                        10                      15                      20

gtt gtg att cat gat cat acg gta cgc ttt aat gca acc acg cca gaa 211
Val Val Ile His Asp His Thr Val Arg Phe Asn Ala Thr Thr Pro Glu
                        25                      30                      35

gat gta gca aaa ctg gta gag ggt gct aaa aat att cta ccc gct gat 259
Asp Val Ala Lys Leu Val Glu Gly Ala Lys Asn Ile Leu Pro Ala Asp
                        40                      45                      50

ttt act gta gaa gca gaa cct gaa ctt gat att aat act gcg gtt gcc 307
Phe Thr Val Glu Ala Glu Pro Glu Leu Asp Ile Asn Thr Ala Val Ala
                        55                      60                      65

gat agt att gaa aca gcg cgt gtt gct att gtt gct ttg ggt gat acg 355
Asp Ser Ile Glu Thr Ala Arg Val Ala Ile Val Ala Leu Gly Asp Thr
                        70                      75                      80                      85

gtt gaa gaa aat gag atg gat att tta atc aat gca tta aat acc caa 403
Val Glu Glu Asn Glu Met Asp Ile Leu Ile Asn Ala Leu Asn Thr Gln
                        90                      95                      100

atc att aac ttt gct tta gac tca acc gaa att ccc caa gaa aat aaa 451
Ile Ile Asn Phe Ala Leu Asp Ser Thr Glu Ile Pro Gln Glu Asn Lys
                        105                      110                      115

gaa atc ttg gat ttg gct gcc gaa aaa tta aag gca gtg cct gaa aca 499
Glu Ile Leu Asp Leu Ala Ala Glu Lys Leu Lys Ala Val Pro Glu Thr
                        120                      125                      130

act ttg cgt atc att ggt cat aca gac act caa ggc acg cat gag tat 547
Thr Leu Arg Ile Ile Gly His Thr Asp Thr Gln Gly Thr His Glu Tyr
                        135                      140                      145

aat caa gat tta tca gaa tct cgt gct gct gct gtt aaa gag tat ttg 595
Asn Gln Asp Leu Ser Glu Ser Arg Ala Ala Ala Val Lys Glu Tyr Leu
                        150                      155                      160                      165

gta tca aaa ggt gtt gct gct gaa cgt ttg aac act caa ggt gca agt 643
Val Ser Lys Gly Val Ala Ala Glu Arg Leu Asn Thr Gln Gly Ala Ser
                        170                      175                      180

ttt gat tat cca gtt gca tca aat gct acc gaa caa ggt cgc ttc caa 691
Phe Asp Tyr Pro Val Ala Ser Asn Ala Thr Glu Gln Gly Arg Phe Gln
                        185                      190                      195

```

60/189

aac cgt cgt att gag ttt gta ctt ttc caa gaa ggt gaa gca att act 739
Asn Arg Arg Ile Glu Phe Val Leu Phe Gln Glu Gly Glu Ala Ile Thr
200 205 210

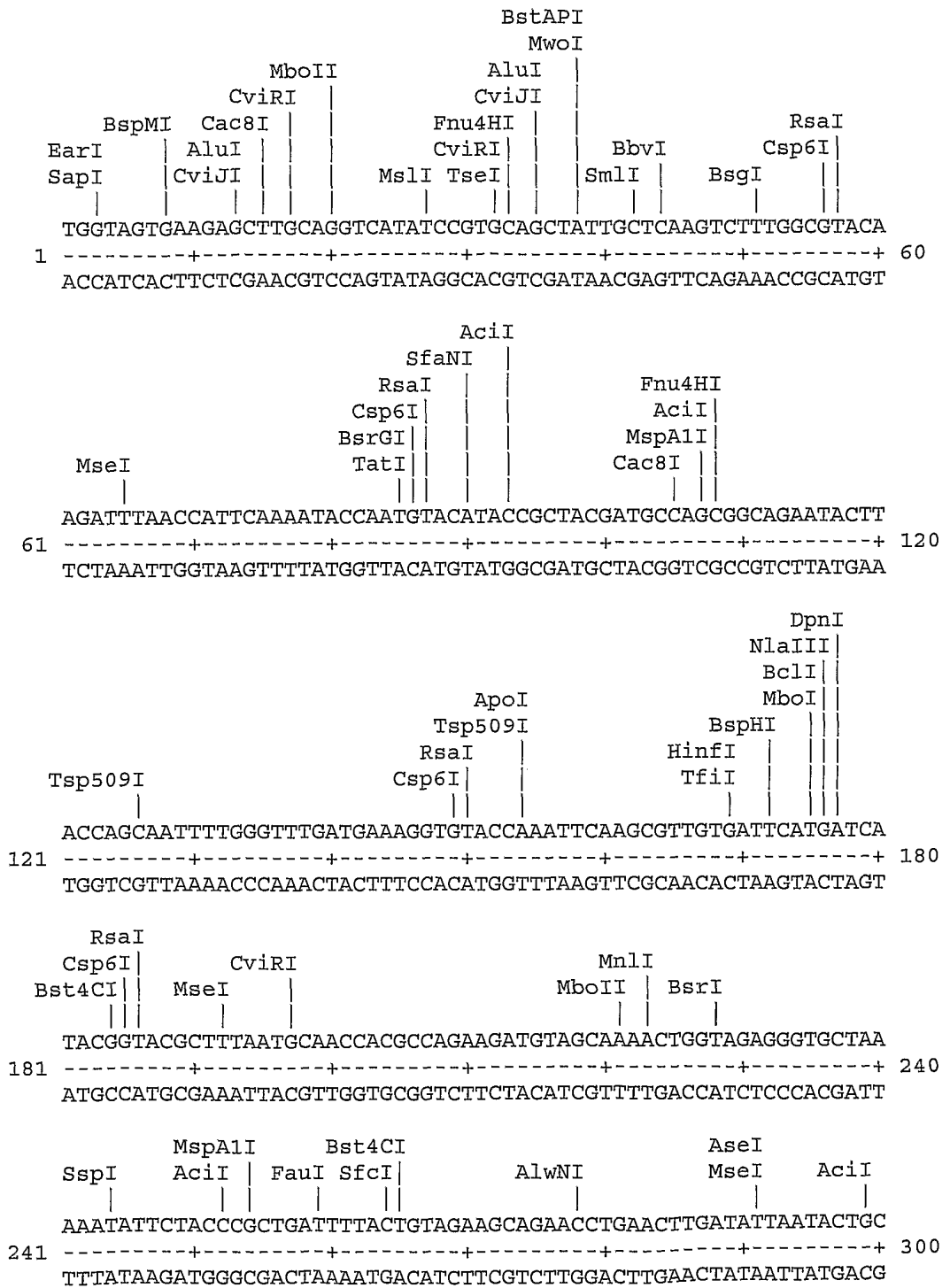
caa gtc ggt cat gct gaa gat gca cca aca cct gtt gca caa aac tga 787
Gln Val Gly His Ala Glu Asp Ala Pro Thr Pro Val Ala Gln Asn
215 220 225

tcattttggtt attgggttatg agtttttagat tgggccaaat gaatgataat ataccaatct 847

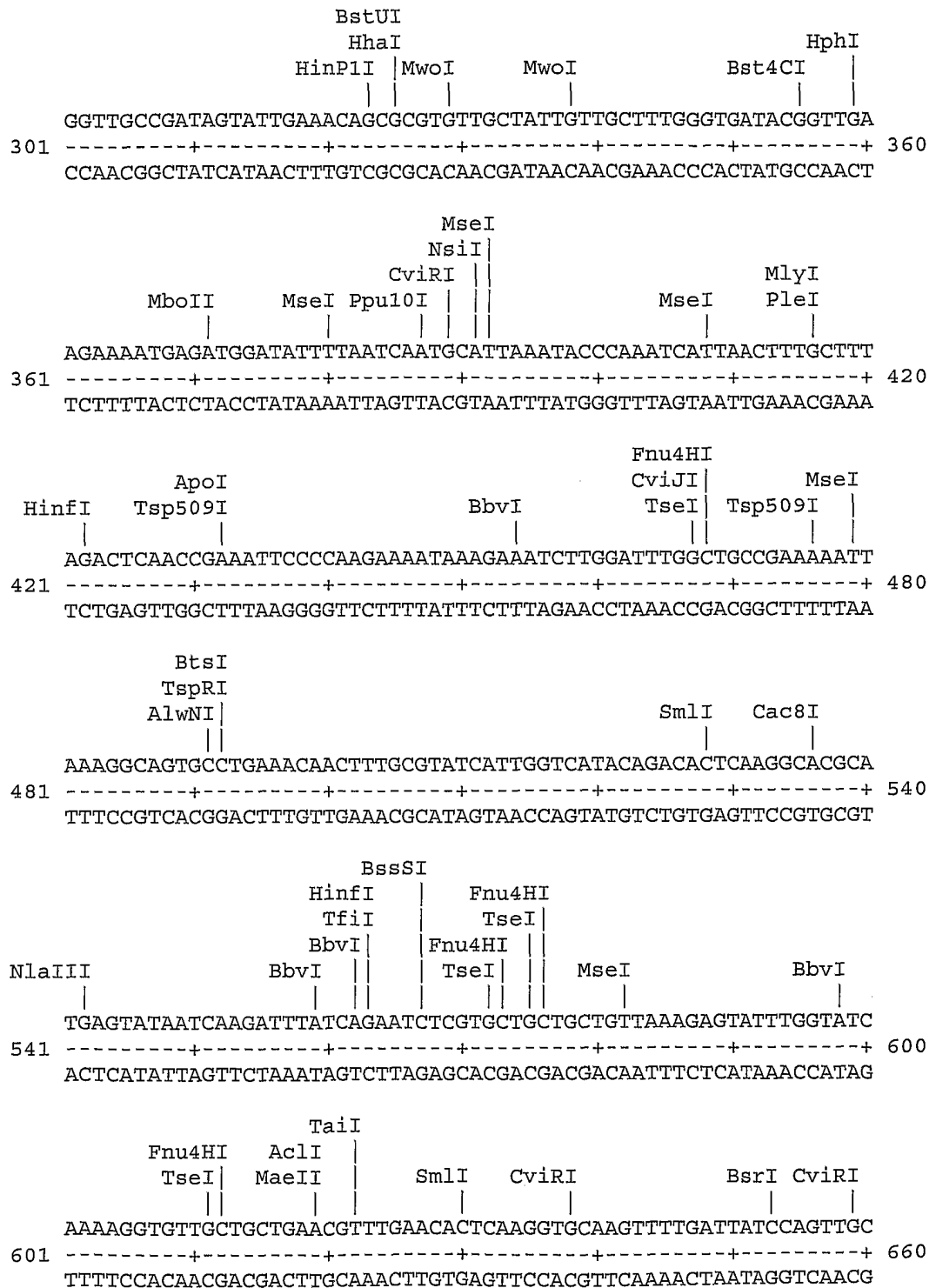
tacaagtact ttttaataacc aaaaccaacc gtaatcaacc caagaaccaa a 898

61/189

Figure 22. Restriction enzyme analysis of the *M. catarrhalis* OmpA gene designated MCA100854



62/189



63/189

```

          SfaNI                      RsaI
      BstAPI |                      Csp6I |
        MwoI |                      TatI |||
          | |                      |||
ATCAAATGCTACCGAACAAGGTCGCTTCCAAAACCGTCGTATTGAGTTTGTACTTTTCCA
661 -----+-----+-----+-----+-----+-----+-----+ 720
TAGTTTACGATGGCTTGTTCAGCGAAGGTTTTGGCAGCATAAATCAAACATGAAAAGGT

          SmlI          NlaIII      CviRI          CviRI
      Tsp509I HphI |      SfaNI |      MwoI |      MboII Eco57I |
          | |      | |      | |      | |      | |      | |
AGAAGGTGAAGCAATTACTCAAGTCGGTCATGCTGAAGATGCACCAACACCTGTTGCACA
721 -----+-----+-----+-----+-----+-----+-----+ 780
TCTTCCACTTCGTTAATGAGTTCAGCCAGTACGACTTCTACGTGGTTGTGGACAACGTGT

          DpnI                      CviJI
      BclI |                      HaeIII
      MboI |                      Sau96I |
          | |                      | |
AAACTGATCATTGTTTGTATGTTATGAGTTTTAGATTGGGCCAAATGAATGATAATATA
781 -----+-----+-----+-----+-----+-----+-----+ 840
TTTGAAGTAGTAAACAATAACCAATACTCAAAATCTAACCCGGTTTACTTACTATTATAT

          RsaI
          ScaI
      Csp6I |
      TatI ||| MseI          Bst4CI          Tsp509I
          | |      | |      | |      | |
CCAATCTTACAAGTACTTTTAATAACCAAAACCAACCGTAATCAACCCAAGAACCAAATT
841 -----+-----+-----+-----+-----+-----+-----+ 900
GGTTAGAATGTTTCATGAAAATTATTGGTTTTGGTTGGCATTAGTTGGGTTCTTGTTTTAA

          XcmI
          |
ACCCATCGGTCATTTGGTTCTTGGGTAGTTT
901 -----+-----+-----+-----+-----+-----+ 931
TGGGTAGCCAGTAAACCAAGAACCCATCAAA

```

64/189

Figure 23. Sequence of *M. catarrhalis* NADH:ubiquinone oxidoreductase subunit A gene designated MCA100889

```

tcgcacggct ttggatgttt ttaggaaact tggtgtccag ccaattaatc accacttggt 60
gaaggctggg atgttttggt cttaatttaa gtatccgctt atg att acc atc aaa 115
                                         Met Ile Thr Ile Lys
                                         1           5
aaa ggc ttg gat ctg cca atc act ggc gag ccc tcc aag cag att agc 163
Lys Gly Leu Asp Leu Pro Ile Thr Gly Glu Pro Ser Lys Gln Ile Ser
                        10                15                20
gag cac acg ccc act aag gtg gcg tta gtc ggc tac gat tat gtc ggc 211
Glu His Thr Pro Thr Lys Val Ala Leu Val Gly Tyr Asp Tyr Val Gly
                        25                30                35
atg cgt ccg acc atg cat gtc aaa gaa ggt gat aag gtc acc aaa gga 259
Met Arg Pro Thr Met His Val Lys Glu Gly Asp Lys Val Thr Lys Gly
                        40                45                50
cag gtg ctt tac gaa gac aaa aaa cga gca ggc gtc ctt tat aca gca 307
Gln Val Leu Tyr Glu Asp Lys Lys Arg Ala Gly Val Leu Tyr Thr Ala
                        55                60                65
cca gtt ggc ggc act gtc gtt gct gtc aat cgt ggc gaa cga cga gta 355
Pro Val Gly Gly Thr Val Val Ala Val Asn Arg Gly Glu Arg Arg Val
                        70                75                80                85
ttt gaa agc ttg gtc atc gcc gtc gat cca aat gca agt gaa gaa atc 403
Phe Glu Ser Leu Val Ile Ala Val Asp Pro Asn Ala Ser Glu Glu Ile
                        90                95                100
act ttt aat gct tat cgc tct gaa gat ttg tca atg ctg gat cgc aag 451
Thr Phe Asn Ala Tyr Arg Ser Glu Asp Leu Ser Met Leu Asp Arg Lys
                        105                110                115
aca gtt gtt gag cag ttg gtc gct tca ggt gaa tgg acc gcc cta cgc 499
Thr Val Val Glu Gln Leu Val Ala Ser Gly Glu Trp Thr Ala Leu Arg
                        120                125                130
act cgt cct ttt agc cgt aca cca gaa att aat agc ata cct tct gcc 547
Thr Arg Pro Phe Ser Arg Thr Pro Glu Ile Asn Ser Ile Pro Ser Ala
                        135                140                145
att ttt gtg acg gca aca gat acc aat cct ttg gct gtt gac cct gcc 595
Ile Phe Val Thr Ala Thr Asp Thr Asn Pro Leu Ala Val Asp Pro Ala
                        150                155                160                165
acc gta tta ttt gaa gaa ata caa gca ttt aat gat ggc ttg gca att 643
Thr Val Leu Phe Glu Glu Ile Gln Ala Phe Asn Asp Gly Leu Ala Ile
                        170                175                180
tta tca acg cta tcc cca aaa acc tat gtc tgt cat ggt gat att gta 691
Leu Ser Thr Leu Ser Pro Lys Thr Tyr Val Cys His Gly Asp Ile Val
                        185                190                195

```

65/189

cct gag acc tta agt aaa acc gct gac ggt aac gat acc atc tat gaa	739
Pro Glu Thr Leu Ser Lys Thr Ala Asp Gly Asn Asp Thr Ile Tyr Glu	
200 205 210	
ggg ttt act ggt gta cat cca gca ggc aac gca ggt acg cac att cac	787
Gly Phe Thr Gly Val His Pro Ala Gly Asn Ala Gly Thr His Ile His	
215 220 225	
ttt tta cac cct ttg gca cga ggc agc tta gtt tgg acg att aat tat	835
Phe Leu His Pro Leu Ala Arg Gly Ser Leu Val Trp Thr Ile Asn Tyr	
230 235 240 245	
caa gat gtc atc gcc att ggt aag cta ttt acc aca ggt cgc att tac	883
Gln Asp Val Ile Ala Ile Gly Lys Leu Phe Thr Thr Gly Arg Ile Tyr	
250 255 260	
aca gag cgt ctg atc agt ttg gca ggt cca gca gtg gcc aat cct cgc	931
Thr Glu Arg Leu Ile Ser Leu Ala Gly Pro Ala Val Ala Asn Pro Arg	
265 270 275	
ttg gtg cgt acc act cga ggt gct gat ttg acc gaa ttg acc caa ggt	979
Leu Val Arg Thr Thr Arg Gly Ala Asp Leu Thr Glu Leu Thr Gln Gly	
280 285 290	
gaa ctc aac ggc gac gat aac cgt gtc att tca ggc tct gtt cta tca	1027
Glu Leu Asn Gly Asp Asp Asn Arg Val Ile Ser Gly Ser Val Leu Ser	
295 300 305	
ggt cgc aca gct gtt ggt gct acg gca ttt ttg ggc cgc ttt cat aat	1075
Gly Arg Thr Ala Val Gly Ala Thr Ala Phe Leu Gly Arg Phe His Asn	
310 315 320 325	
caa gtc tct gta ctt acc gaa ggt cgt gaa cgc cct gca ctt cac ttt	1123
Gln Val Ser Val Leu Thr Glu Gly Arg Glu Arg Pro Ala Leu His Phe	
330 335 340	
ttt aca cca ggg gca aat cgt ttt tca aaa ttg cca att tat att tca	1171
Phe Thr Pro Gly Ala Asn Arg Phe Ser Lys Leu Pro Ile Tyr Ile Ser	
345 350 355	
cag ttt ttt aaa ggc aaa aaa tac gat ttt acg acc agt acc aac ggc	1219
Gln Phe Phe Lys Gly Lys Lys Tyr Asp Phe Thr Thr Ser Thr Asn Gly	
360 365 370	
tcg cct cgt gcc atg gta cca atc ggc tta ttt gaa act gtc atg cct	1267
Ser Pro Arg Ala Met Val Pro Ile Gly Leu Phe Glu Thr Val Met Pro	
375 380 385	
caa gac tat cta cct acc cag cta ctg cgt gct ttg att gtc gaa gat	1315
Gln Asp Tyr Leu Pro Thr Gln Leu Leu Arg Ala Leu Ile Val Glu Asp	
390 395 400 405	
atc atc agt gcg gtt gaa ttg ggt gca ctt gag ctt gat gaa gaa gat	1363
Ile Ile Ser Ala Val Glu Leu Gly Ala Leu Glu Leu Asp Glu Glu Asp	
410 415 420	

66/189

ttg gca ctg tgt act ttc gta tcg cca ggc aaa tat gag ttc ggt gat 1411
Leu Ala Leu Cys Thr Phe Val Ser Pro Gly Lys Tyr Glu Phe Gly Asp
425 430 435

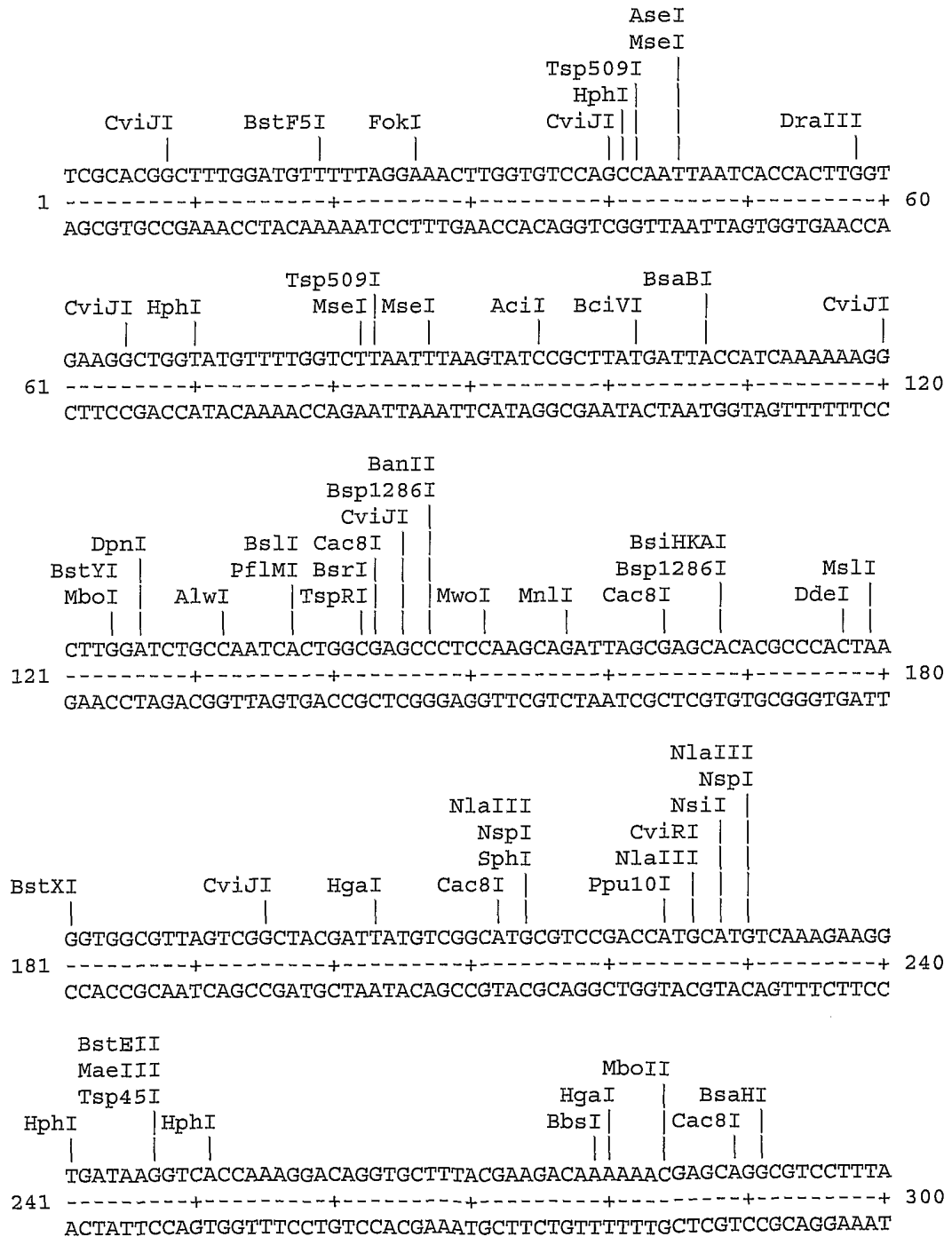
att ttg cgt gat aac ttg acc cgt att gag cag gaa ggc taa 1453
Ile Leu Arg Asp Asn Leu Thr Arg Ile Glu Gln Glu Gly
440 445 450

ccatgaagtt cattcacaat atctttgacc gtatggagcc atcctttacc aaaggtggca 1513

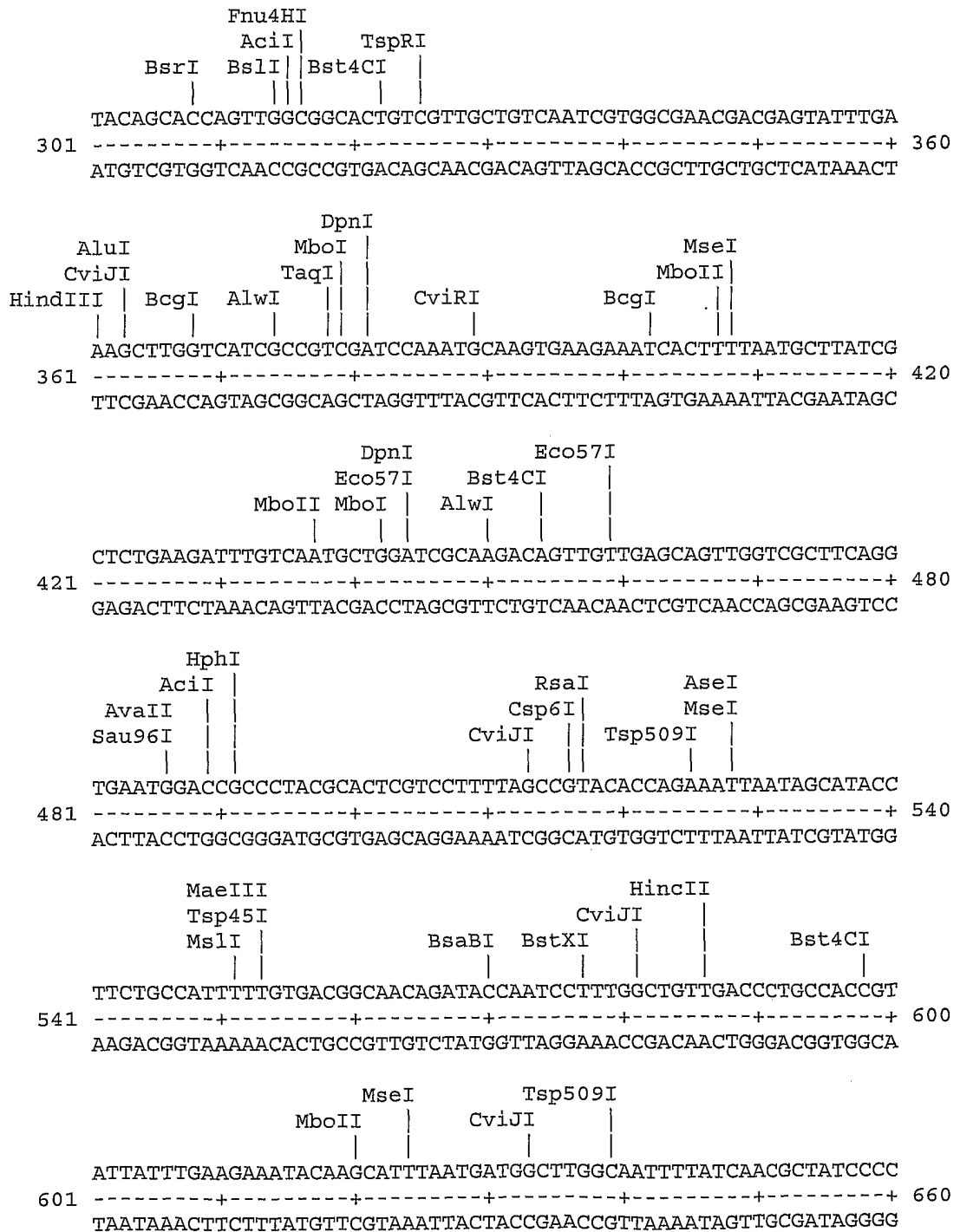
agtacgaaaa atattatgct gtatttgaga tgtttgacac 1553

67/189

Figure 24. Restriction enzyme analysis of the *M. catarrhalis* NADH:ubiquinone oxidoreductase subunit A gene designated MCA100889



68/189



69/189

DdeI
HphI
RsaI
BsaI
BsmAI
MseI
Bst4CI
BseMII
AflII
MspAII
NlaIII
Csp6I
SmlI
AciI
MaeIII

AAAAACCTATGTCTGTCATGGTGATATTGTACCTGAGACCTTAAGTAAAACCGCTGACGG
 661 -----+-----+-----+-----+-----+-----+-----+ 720
 TTTTGGGATACAGACAGTACCCTATAACATGGACTCTGGAATTCATTTGGCGACTGCC

BstF5I
RsaI
Csp6I
BsrI
BsrGI
Cac8I
RsaI
FokI
TatI
BspMI
MwoI
Csp6I

TAACGATACCATCTATGAAGGGTTTACTGGTGTTACATCCAGCAGGCAACGCAGGTACGCA
 721 -----+-----+-----+-----+-----+-----+-----+ 780
 ATTGCTATGGTAGATACTTCCCAAATGACCACATGTAGGTCGTCCGTTGCGTCCATGCGT

DdeI
AluI
CviJI
Tsp509I
BssSI
Fnu4HI
AseI
MnlI
TseI
BbvI
MseI

CATTCACCTTTTACACCCCTTTGGCAGGAGGAGCTTAGTTTGGACGATTAATTATCAAGA
 781 -----+-----+-----+-----+-----+-----+-----+ 840
 GTAAGTGAAAAATGTGGGAAACCGTGCTCCGTCGAATCAAACCTGCTAATTAATAGTTCT

BspMI
DpnI
BclI
AluI
CviJI
HgaI
MboI

TGTCATCGCCATTGGTAAGCTATTTACCACAGGTCGCATTTACACAGAGCGTCTGATCAG
 841 -----+-----+-----+-----+-----+-----+-----+ 900
 ACAGTAGCGGTAACCATTCGATAAATGGTGTCCAGCGTAAATGTGTCTCGCAGACTAGTC

BtsI
TspRI
CviJI
MnlI
TaqI
MwoI
HaeIII
RsaI
AvaI
AvaII
MscI
Csp6I
SmlI
Sau96I
EaeI
XcmI
MnlI
XhoI

TTTGGCAGGTCCAGCAGTGGCCAATCCTCGCTTGGTGCCTACCACTCGAGGTGCTGATTT
 901 -----+-----+-----+-----+-----+-----+-----+ 960
 AAACCGTCCAGGTCGTCACCGGTTAGGAGCGAACCACGCATGGTGAGCTCCACGACTAAA

70/189

```

          BsaJI      StyI              HphI           Bst4CI             AlwNI |
Tsp509I |                |               |               |               |
|        |               |               |               |               |
GACCGAATTGACCAAGGTGAAGTCAACGGCGACGATAACCGTGTCATTTTCAGGCTCTGT
961 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1020
CTGGCTTAAGTGGGTTCCACTTGAGTTGCCGCTGCTATTGGCACAGTAAAGTCCGAGACA


                                AciI
                                Fnu4HI
                    AluI         CviJI |
                    CviJI         CviJI |
                    MspAI         HaeIII |
                    PvuII    MwoI     Sau96I ||
                        |       |       |
TCTATCAGGTGCGCACAGCTGTTGGTGCTACGGCATTTTTGGGCCGCTTTTCATAATCAAGT
1021 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080
AGATAGTCCAGCGTGTGCGACAACCACGATGCCGTAAAAACCCGGCGCAAAGTATTAGTTCA

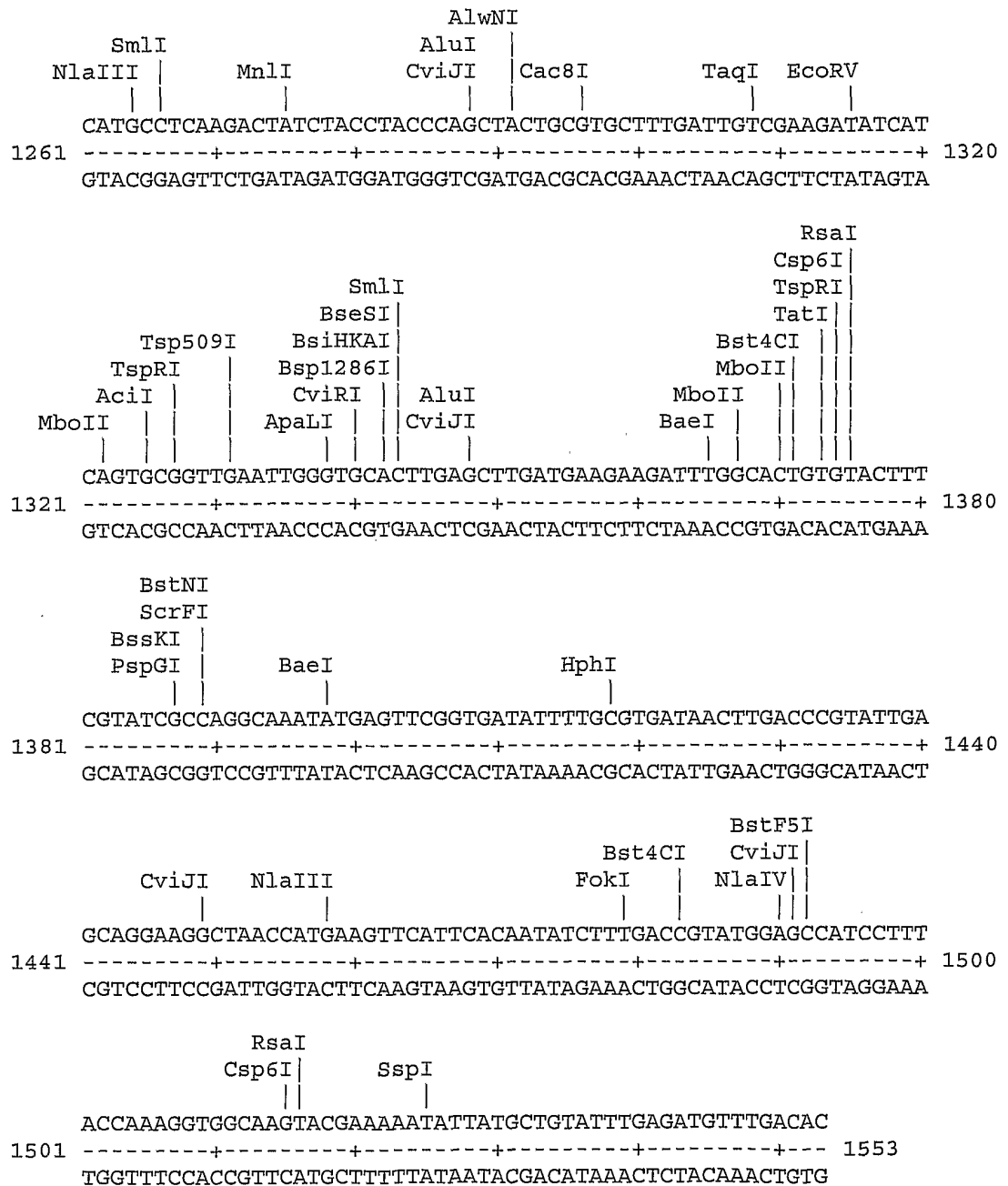

                                BstNI
                                ScrFI
RsaI                               BsaJI |
Csp6I |                           BssKI |
BsmAI |                           PspGI |
TatI  |                           ||   |
            BsgI                   CviRI
            |                       |
CTCTGTACTTACCGAAGGTCGTGAACGCCCTGCACTTCACTTTTTTACACCAGGGGCAAAA
1081 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1140
GAGACATGAATGGCTTCCAGCACTTGCGGGACGTGAAGTGAAAAAATGTGGTCCCCGTTT


                                DraI
                                Tsp509I
Tsp509I |                 Bst4CI  MseI |
|         |               |         |
TCGTTTTTCAAATTGCCAATTTATATTTTACAGTTTTTTTAAAGGCAAAAAATACGATTT
1141 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200
AGCAAAAAGTTTTAACGGTTAAATATAAAGTGTCAAAAATTCCGTTTTTTTATGCTAAA


                                KpnI
                                NlaIV
                                RsaI
                                Csp6I |
                                Acc65I |
                                BanI  |
                                NlaIII |
                                MnlI  |
                                BtgI  |
                                BsaJI |
                                BstDSI |
                                |       |
RsaI                               BssSI |
Csp6I |                             |   |
BsrI  |                             |   |
            Cac8I                   NcoI |
            CviJI                   StyI |
            |                         |   |
TACGACCAGTACCAACGGCTCGCCTCGTGCCATGGTACCAATCGGCCTTATTTGAAACTGT
1201 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260
ATGCTGGTCAATGGTTGCCGAGCGGAGCACGGTACCATGGTTAGCCGAATAAACTTTGACA

```

71/189



72/189

Figure 25. Sequence of *M. catarrhalis* pilQ fimbriae gene designated MCA100992

```

tgacagtcgt acgatacatc tgattgaagc ggtcgctgat acacaagggtg gctattatcg 60

ccgtgatgta aacattcatt ttattcataa gcaatgacac atg gac aaa aaa ttt 115
                                         Met Asp Lys Lys Phe
                                         1           5

tta aaa aat agc ttg gta atc gcc gtg atg ctc ggt gtg gtt cat atg 163
Leu Lys Asn Ser Leu Val Ile Ala Val Met Leu Gly Val Val His Met
                        10                      15                      20

gag gta cgt gct gat act cat atc gat cat gtg tct gtt acc caa aca 211
Glu Val Arg Ala Asp Thr His Ile Asp His Val Ser Val Thr Gln Thr
                        25                      30                      35

gca gag caa cat tat aca ggc aag cct att agc ctt gag ttt gcc gat 259
Ala Glu Gln His Tyr Thr Gly Lys Pro Ile Ser Leu Glu Phe Ala Asp
                        40                      45                      50

att cct gta cga gcg gtt ttt gat att tta gca ggt ttt aca ggt att 307
Ile Pro Val Arg Ala Val Phe Asp Ile Leu Ala Gly Phe Thr Gly Ile
                        55                      60                      65

aat atc atc acc gat gat agt gtg aca ggc agt atg acc att cgc ttg 355
Asn Ile Ile Thr Asp Asp Ser Val Thr Gly Ser Met Thr Ile Arg Leu
                        70                      75                      80                      85

atg aat atc cct tgg gat caa gca ttt gat gtg att ttg caa aca caa 403
Met Asn Ile Pro Trp Asp Gln Ala Phe Asp Val Ile Leu Gln Thr Gln
                        90                      95                      100

aat tta tca gta tta aag cat ggt aat gta tgg ctg atc agc tca aaa 451
Asn Leu Ser Val Leu Lys His Gly Asn Val Trp Leu Ile Ser Ser Lys
                        105                      110                      115

tct atc caa tcc aat cag ccg aca ata acc gaa tac att cgt cta aac 499
Ser Ile Gln Ser Asn Gln Pro Thr Ile Thr Glu Tyr Ile Arg Leu Asn
                        120                      125                      130

tat gca ctt gcc gat gat gtg gca aca ctc atc atg ggc gaa aaa acc 547
Tyr Ala Leu Ala Asp Asp Val Ala Thr Leu Ile Met Gly Glu Lys Thr
                        135                      140                      145

cag cgt ggt aat gtc aat cga acc aat cac cgt gat gtt att tat cct 595
Gln Arg Gly Asn Val Asn Arg Thr Asn His Arg Asp Val Ile Tyr Pro
                        150                      155                      160                      165

gaa gca cca tta tca cgg gcc tat aat acc cag act gat gaa ctt att 643
Glu Ala Pro Leu Ser Arg Ala Tyr Asn Thr Gln Thr Asp Glu Leu Ile
                        170                      175                      180

acc acc gct gtg cgt ggg aca tta ctg tct gag cgt ggc aca gtg acg 691
Thr Thr Ala Val Arg Gly Thr Leu Leu Ser Glu Arg Gly Thr Val Thr
                        185                      190                      195

```

73/189

gtg gat aag cgt acc aat act tta att atc caa gat gtg cct gca agc	739
Val Asp Lys Arg Thr Asn Thr Leu Ile Ile Gln Asp Val Pro Ala Ser	
200 205 210	
gtt gcc aat atc aag gca ctg att gag cgt att gat atc cct gtt gaa	787
Val Ala Asn Ile Lys Ala Leu Ile Glu Arg Ile Asp Ile Pro Val Glu	
215 220 225	
caa gtg atg att gaa gct aga att gtt tct gct aat gaa aat ttt ggg	835
Gln Val Met Ile Glu Ala Arg Ile Val Ser Ala Asn Glu Asn Phe Gly	
230 235 240 245	
cga aag ctt ggc gta agt ttt ggg gcg cat ggt caa aat ggt aaa gtg	883
Arg Lys Leu Gly Val Ser Phe Gly Ala His Gly Gln Asn Gly Lys Val	
250 255 260	
cat tat ggc ggc tct cag gga agt ttg tgg aca atg tgg caa gag ggt	931
His Tyr Gly Gly Ser Gln Gly Ser Leu Trp Thr Met Trp Gln Glu Gly	
265 270 275	
gtg gcg gca ggt ggt cat caa aac agt cat ctg aat gtt gat ttg ggc	979
Val Ala Ala Gly Gly His Gln Asn Ser His Leu Asn Val Asp Leu Gly	
280 285 290	
gta gat aac gcc atg gga cgc att gct ttt ggg ctg tta aat tta cct	1027
Val Asp Asn Ala Met Gly Arg Ile Ala Phe Gly Leu Leu Asn Leu Pro	
295 300 305	
gat atc att tta gat tta gaa ctc tcc gcc atg cag gca gaa aat caa	1075
Asp Ile Ile Leu Asp Leu Glu Leu Ser Ala Met Gln Ala Glu Asn Gln	
310 315 320 325	
ggg gag gtt atc tca aca cca aaa gtc ttg acg gct gat aaa caa acc	1123
Gly Glu Val Ile Ser Thr Pro Lys Val Leu Thr Ala Asp Lys Gln Thr	
330 335 340	
gcg cgt att tca tca ggg cta caa att cct tat caa gaa acc acg cat	1171
Ala Arg Ile Ser Ser Gly Leu Gln Ile Pro Tyr Gln Glu Thr Thr His	
345 350 355	
tca ggt gcc agc aca aca aaa ttt aaa gaa gcg tct tta att ttg gag	1219
Ser Gly Ala Ser Thr Thr Lys Phe Lys Glu Ala Ser Leu Ile Leu Glu	
360 365 370	
gcg aca cca aat atc aca cct gat ggt aaa att ggt ctt aag cta aat	1267
Ala Thr Pro Asn Ile Thr Pro Asp Gly Lys Ile Gly Leu Lys Leu Asn	
375 380 385	
atc aaa aat ggc aat cca gta cca aca ctt ggt cat atc gcc atc caa	1315
Ile Lys Asn Gly Asn Pro Val Pro Thr Leu Gly His Ile Ala Ile Gln	
390 395 400 405	
gaa gat gcc atc gaa acg aat gtc atc atc gaa gat gga caa acc gtc	1363
Glu Asp Ala Ile Glu Thr Asn Val Ile Ile Glu Asp Gly Gln Thr Val	
410 415 420	

74/189

```

gtt ctt ggt ggt atc tat cgc acc agc cag aat gaa ggt gca aat aaa 1411
Val Leu Gly Gly Ile Tyr Arg Thr Ser Gln Asn Glu Gly Ala Asn Lys
      425                      430                      435

gtg cct aga ttg ggt gat atg cct gtg ctt gga cga ttg ttt cgg cat 1459
Val Pro Arg Leu Gly Asp Met Pro Val Leu Gly Arg Leu Phe Arg His
      440                      445                      450

gat tcc aaa gcg cat gat aag tcg gag ctt ttg att ttt att aca ccc 1507
Asp Ser Lys Ala His Asp Lys Ser Glu Leu Leu Ile Phe Ile Thr Pro
      455                      460                      465

aaa ttg gta aga tag ctggcattat attgggtcaat aaaataaatg aaatttttat 1562
Lys Leu Val Arg
470

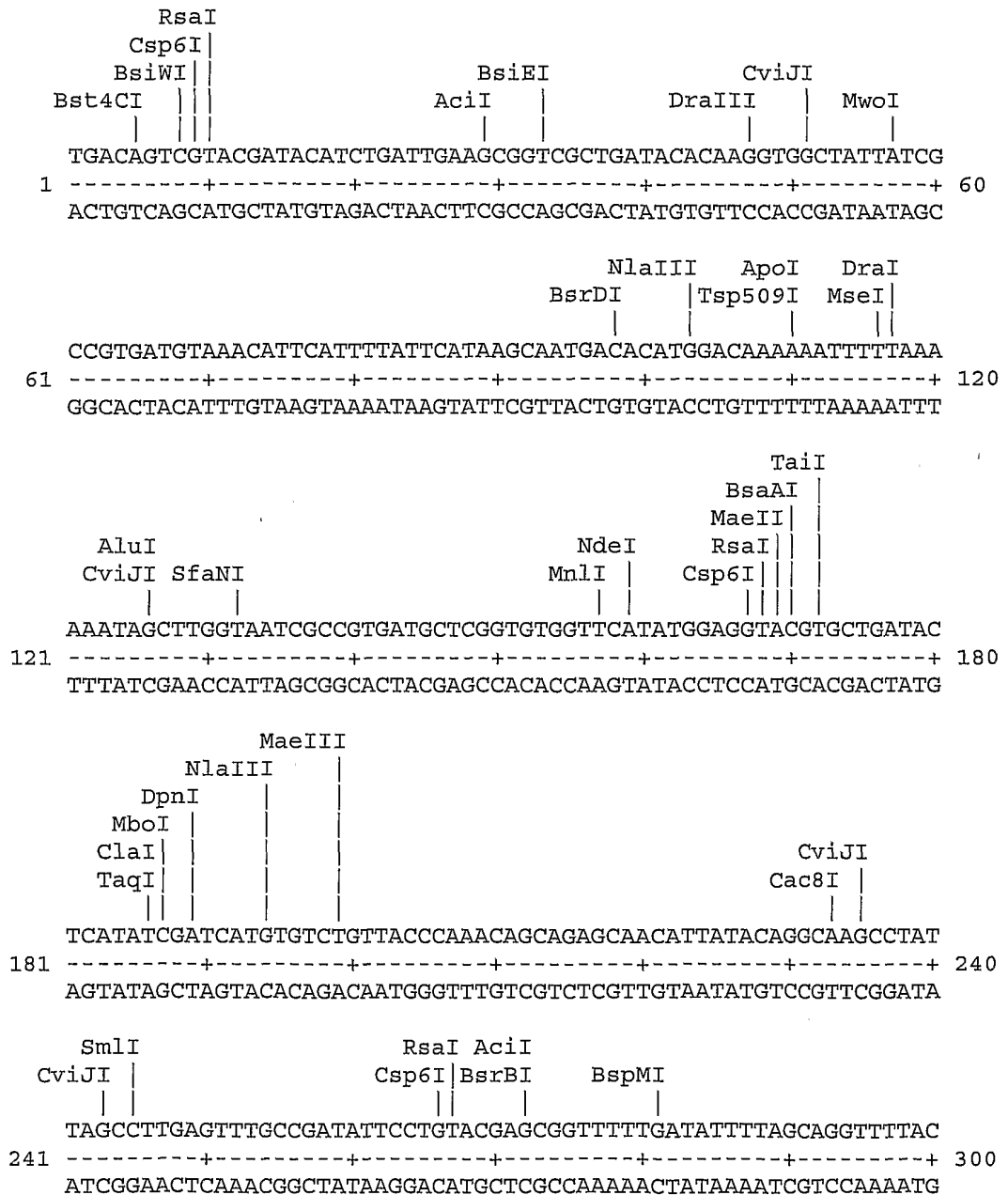
ctaagtaatc ggtatctgaa tgggtcaatac ataaaaatac tggttaacaaa ttgccaacaa 1622

ttgtatgata agataatttt cattcatt 1650

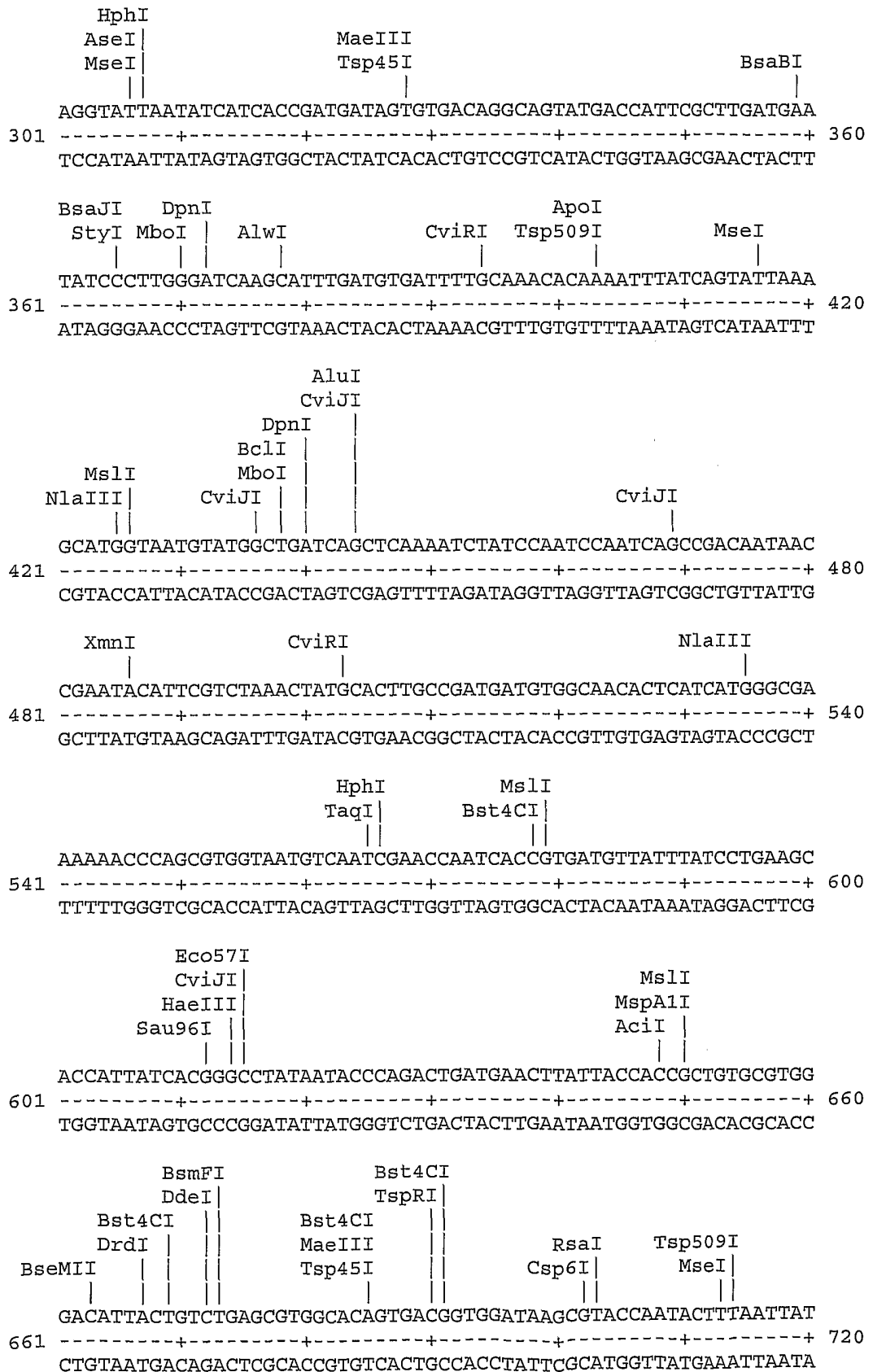
```

75/189

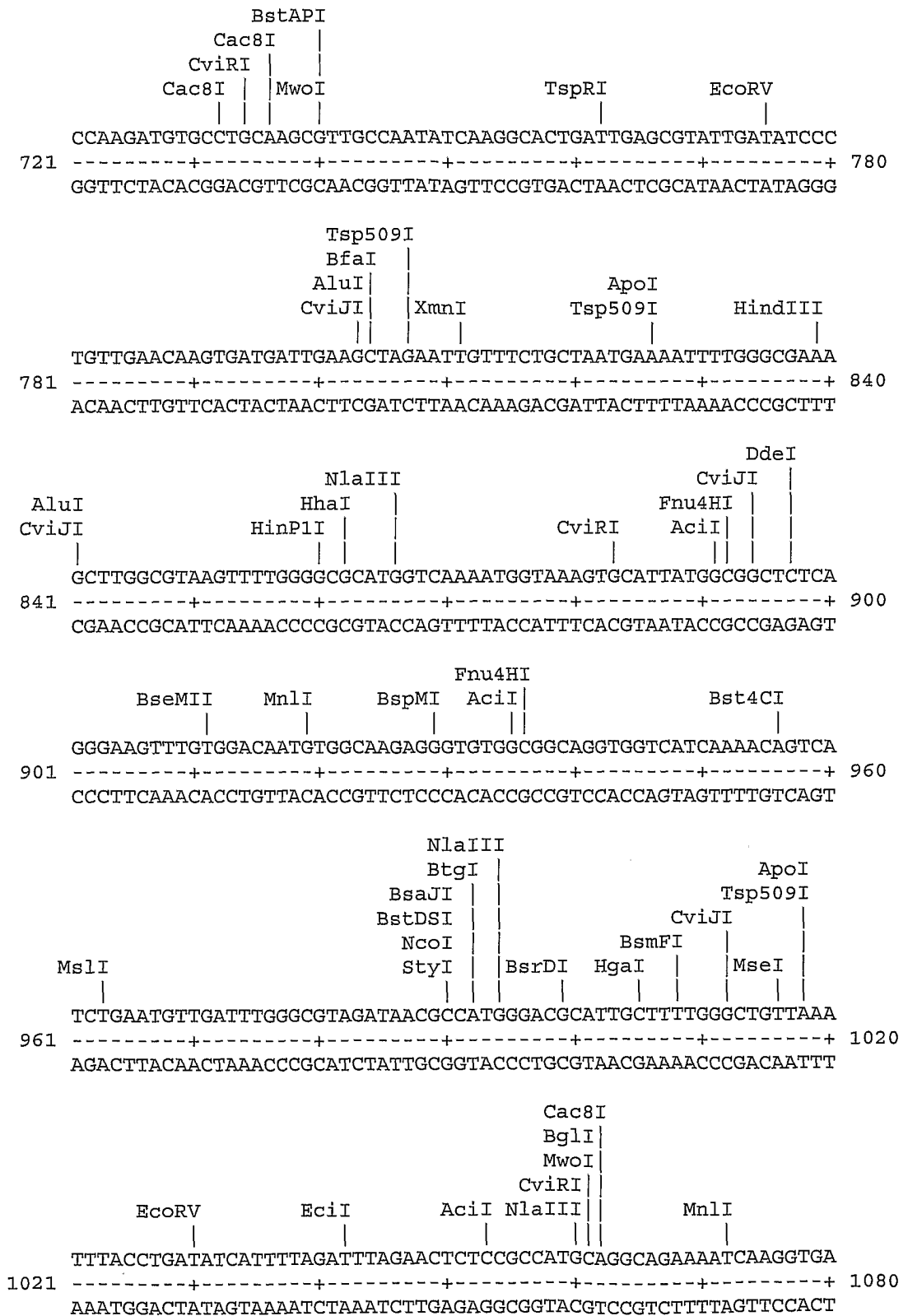
Figure 26. Restriction enzyme analysis of the *M. catarrhalis* pilQ fimbriae gene designated MCA100992



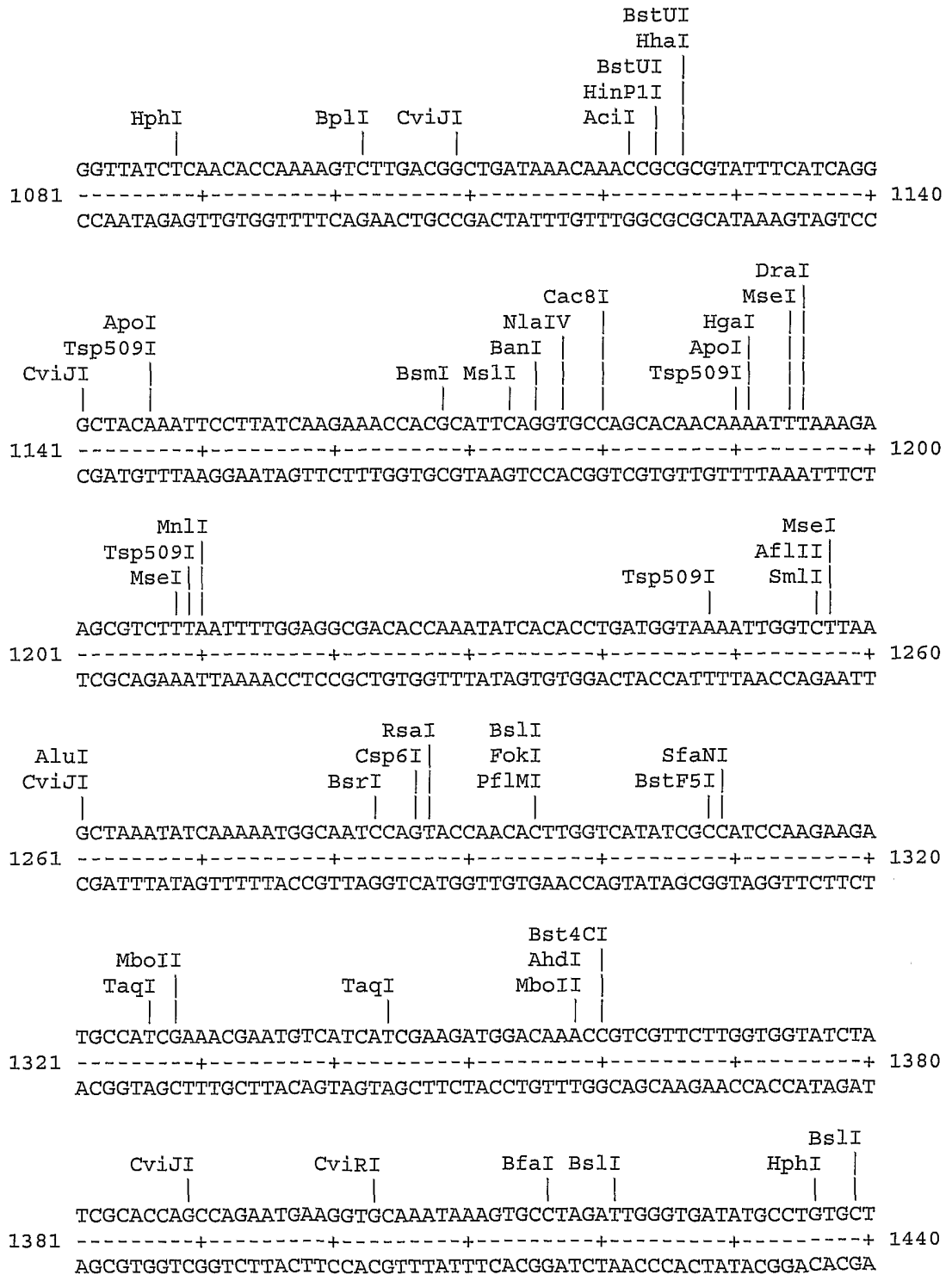
76/189



77/189



78/189



79/189

```

                HinfI      NlaIII      AluI
                NlaIII      HhaI      CviJI
                TfiI      HinPII      |
                |      |      |
TGGACGATTGTTTCGGCATGATTCCAAAGCGCATGATAAGTCGGAGCTTTTGATTTTTAT
1441 -----+-----+-----+-----+-----+-----+ 1500
ACCTGCTAACAAAGCCGTACTAAGGTTTCGCGTACTATTCAGCCTCGAAAATAAAAATA

                Cac8I
                AluI      |
Tsp509I      CviJI      |      ApoI
                |      |      Tsp509I
                |      |      |
TACACCCAAATTGGTAAGATAGCTGGCATTATATTGGTCAATAAAATAAATGAAATTTTT
1501 -----+-----+-----+-----+-----+-----+ 1560
ATGTGGGTTTAACCATTCTATCGACCGTAATATAACCAGTTATTTTATTTACTTTAAAAA

                Tsp509I
                BsrI      |
DdeI      MaeIII      |      |
                |      |      |
ATCTAAGTAATCGGTATCTGAATGGTCAATACATAAAAAATACTGGTAACAAATTGCCAAA
1561 -----+-----+-----+-----+-----+-----+ 1620
TAGATTCATTAGCCATAGACTTACCAGTTATGTATTTTATGACCATTGTTTAACGGTTT

                Tsp509I
                |
CATTGTATGATAAGATAATTTTCATTCATT
1621 -----+-----+-----+-----+ 1650
GTAACATACTATTCTATTAAAAAGTAAGTAA

```

80/189

Figure 27. Sequence of *M. catarrhalis* invasin protein gene designated MCA101249

```

ccatttgtca tgtaaccatg aatgaaaatt gtcaattggt atcaattttc aataacaagg 60

tatcaatgct gggatatctcc tatatatatta gagtttgatt atg att gat gca gac 115
                                         Met Ile Asp Ala Asp
                                         1           5

gga ttt cgt cca aat gtc ggc atc atc ttg gta aat gca cac ggg caa 163
Gly Phe Arg Pro Asn Val Gly Ile Ile Leu Val Asn Ala His Gly Gln
                        10                      15                      20

ctt ttg tgg gca aag cgt atc ggt cat gac agt tgg cag ttt cct caa 211
Leu Leu Trp Ala Lys Arg Ile Gly His Asp Ser Trp Gln Phe Pro Gln
                        25                      30                      35

ggc ggt atc agt cat ggc gaa aca cca cta gat gcg atg tat cgt gaa 259
Gly Gly Ile Ser His Gly Glu Thr Pro Leu Asp Ala Met Tyr Arg Glu
                        40                      45                      50

ctg tat gaa gaa gtg ggg ctg cat cct gag cat gtc aaa gtg ctg gca 307
Leu Tyr Glu Glu Val Gly Leu His Pro Glu His Val Lys Val Leu Ala
                        55                      60                      65

gtc acc cgt gat tgg ttg cgt tat cgc ttg cct aag cgt tat ttg cgt 355
Val Thr Arg Asp Trp Leu Arg Tyr Arg Leu Pro Lys Arg Tyr Leu Arg
                        70                      75                      80                      85

aca ggg cag gag cct ttg tgc att gga caa aaa caa aaa tgg ttt ttg 403
Thr Gly Gln Glu Pro Leu Cys Ile Gly Gln Lys Gln Lys Trp Phe Leu
                        90                      95                      100

tta tat ctg gat gaa gct aat acc aaa tac att cgt ttt gat gtg ggc 451
Leu Tyr Leu Asp Glu Ala Asn Thr Lys Tyr Ile Arg Phe Asp Val Gly
                        105                      110                      115

gta cct gaa ttt gac gac tgg caa tgg gtg agt tat tgg ttt cct ttg 499
Val Pro Glu Phe Asp Asp Trp Gln Trp Val Ser Tyr Trp Phe Pro Leu
                        120                      125                      130

acg ggt gtt gta ccg ttt aaa aga aat gta tat cag cgg gca ttg ctt 547
Thr Gly Val Val Pro Phe Lys Arg Asn Val Tyr Gln Arg Ala Leu Leu
                        135                      140                      145

gag ctg att gat gct gta ccc aaa cga aag tct ttg gtg gca aat atg 595
Glu Leu Ile Asp Ala Val Pro Lys Arg Lys Ser Leu Val Ala Asn Met
                        150                      155                      160                      165

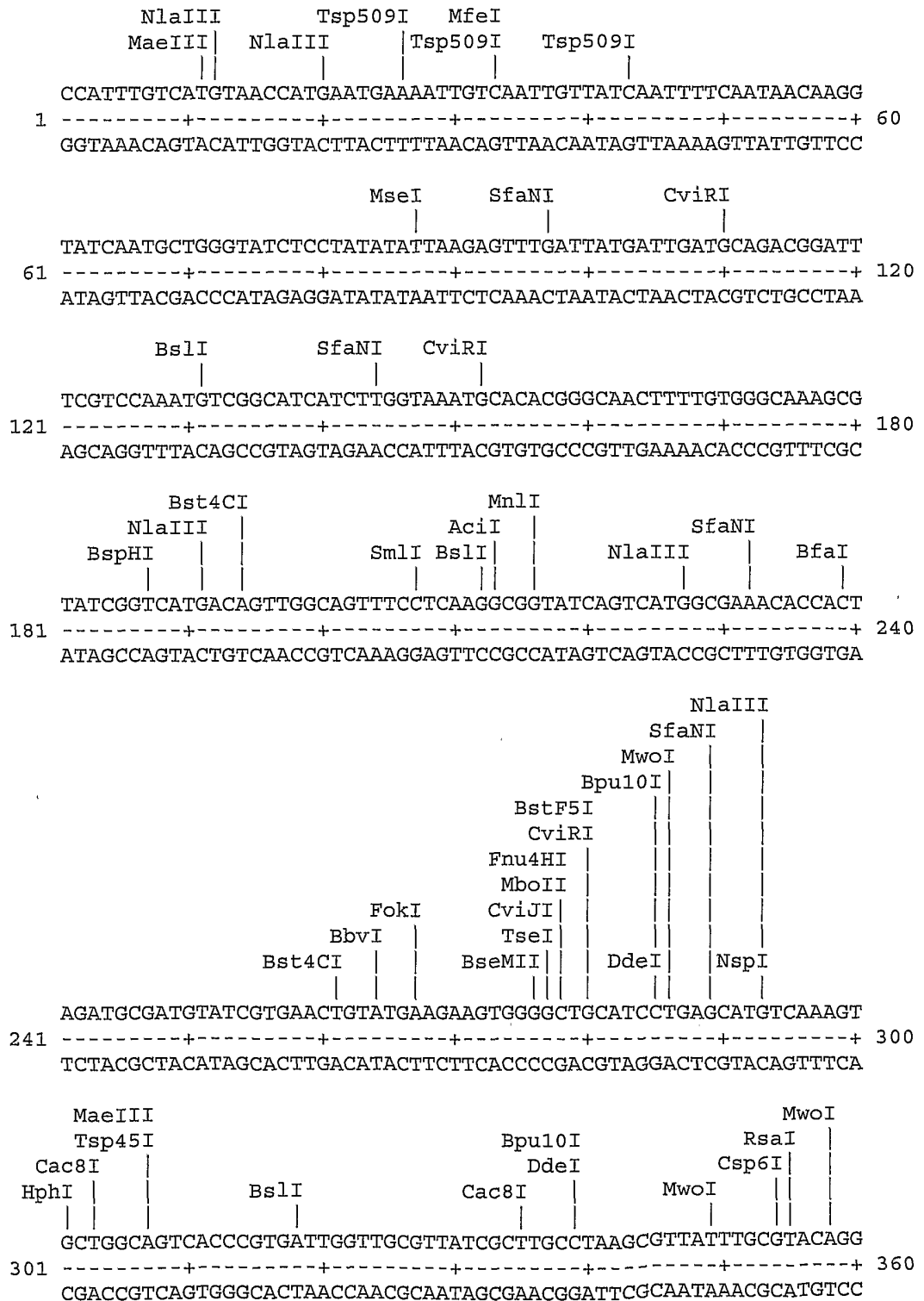
taa taaatatgta ataagcggtta aaaccatcac caaaattggt gatgggttttt 648

tgctatgatt tgagtttttgc gattagcata cttaggcaaa atttgctaaa at 700

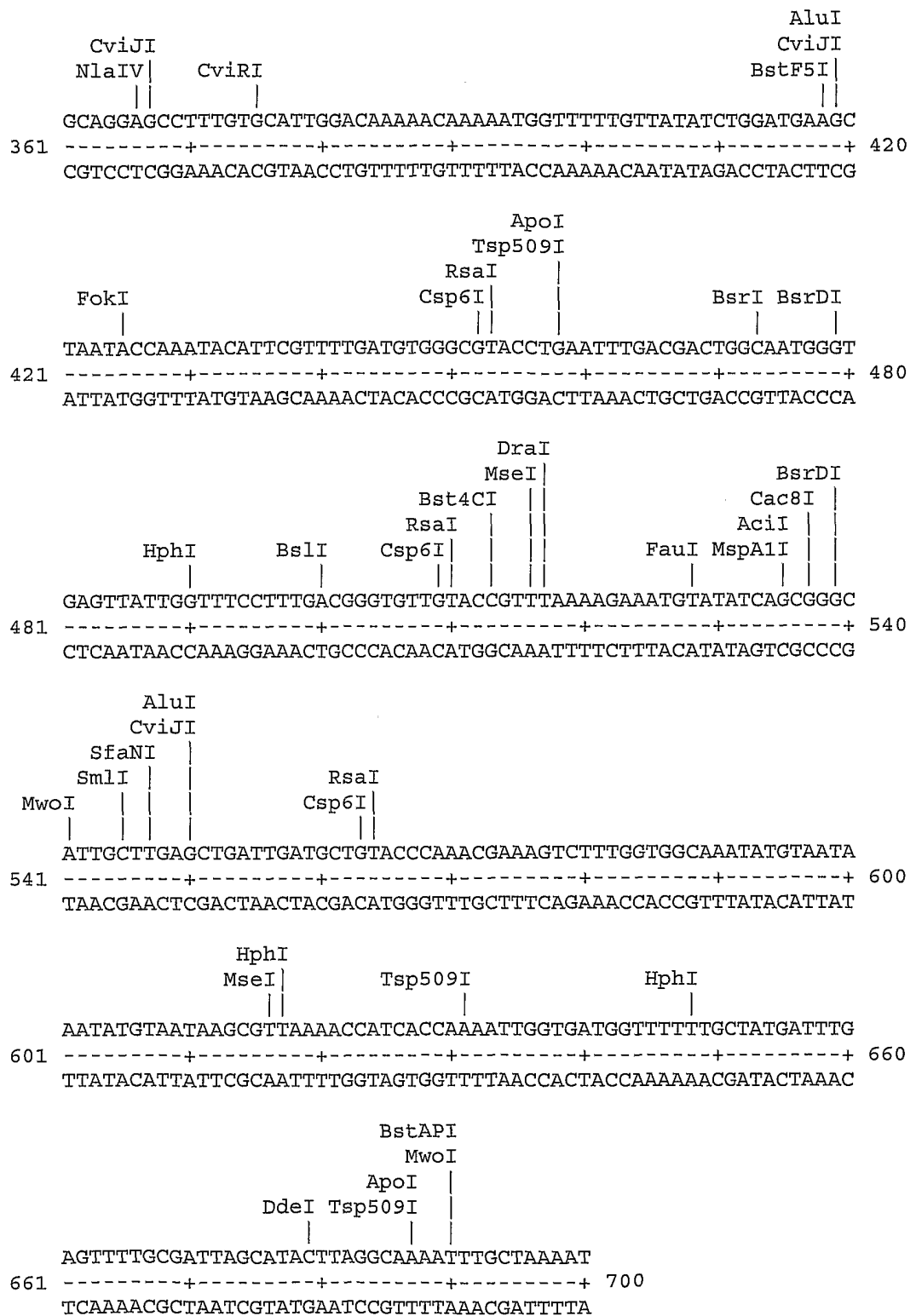
```

81/189

Figure 28. Restriction enzyme analysis of the *M. catarrhalis* invasin protein gene designated MCA101249



82/189



83/189

Figure 29. Sequence of *M. catarrhalis* gene encoding a rare lipoprotein A (SEQ ID NOs: 29 & 30) designated MCA101388

```

tgggtgttgg ctatgatacg ccattatttt tgagatttta gtgaaactgc actgaaatca 60

acataaatcc accaaaggct aaaagcctaa ggatcagatg atg aaa aag att ttt 115
                                         Met Lys Lys Ile Phe
                                         1           5

gta gct gct ttg gcg ctg tct gcc ttg tct gct caa gcg gcg acc gtc 163
Val Ala Ala Leu Ala Leu Ser Ala Leu Ser Ala Gln Ala Ala Thr Val
              10              15              20

acc tac tat ggt gat aag ttt cat ggc aaa aag aca gca agt ggc gag 211
Thr Tyr Tyr Gly Asp Lys Phe His Gly Lys Lys Thr Ala Ser Gly Glu
              25              30              35

aga ttt aac caa aat gca atg acc tgt gcc agt aat agc cat aaa ttt 259
Arg Phe Asn Gln Asn Ala Met Thr Cys Ala Ser Asn Ser His Lys Phe
              40              45              50

ggc acg cat tta aaa gtt acc aac acc gcc aat gat aaa tca gtg att 307
Gly Thr His Leu Lys Val Thr Asn Thr Ala Asn Asp Lys Ser Val Ile
              55              60              65

tgc cga gtg aat gac cgt ggc ggc ttt ggt aaa tat ggc gtg act tta 355
Cys Arg Val Asn Asp Arg Gly Gly Phe Gly Lys Tyr Gly Val Thr Leu
              70              75              80              85

gat ttg tca aaa ggt gca ttt gct aaa att gca cca ctt tct caa ggt 403
Asp Leu Ser Lys Gly Ala Phe Ala Lys Ile Ala Pro Leu Ser Gln Gly
              90              95              100

cgt gca cgg gta atc att gaa aag gtt gat gat tta att gcc caa gac 451
Arg Ala Arg Val Ile Ile Glu Lys Val Asp Asp Leu Ile Ala Gln Asp
              105              110              115

acc caa ggt atc gaa agt agc aca ctg gca caa gac agc cta agc gca 499
Thr Gln Gly Ile Glu Ser Ser Thr Leu Ala Gln Asp Ser Leu Ser Ala
              120              125              130

caa aaa gtg gct gaa gcg tta tca caa gtt gaa aaa gta tca acc gaa 547
Gln Lys Val Ala Glu Ala Leu Ser Gln Val Glu Lys Val Ser Thr Glu
              135              140              145

cac acc act cgc atc act ttg gca agc aat act tca gtg act gaa taa 595
His Thr Thr Arg Ile Thr Leu Ala Ser Asn Thr Ser Val Thr Glu
              150              155              160              165

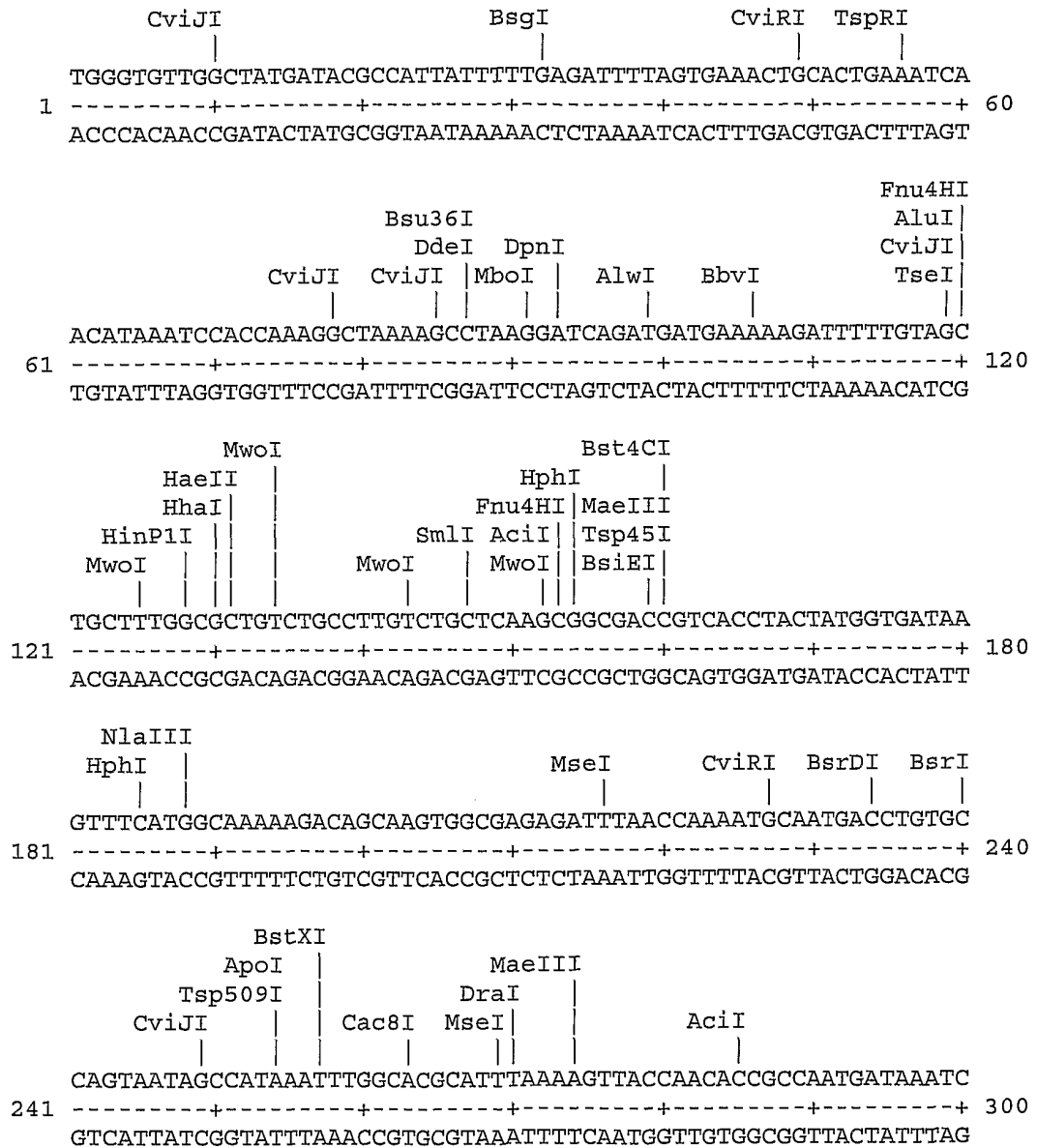
gatgtaaaag cttatgataa aagcgacaga tattgtggtt ggcacaggca aaactgccgt 655

taaagggtgct ttgattatca ctgaatacac aggctatctg 695

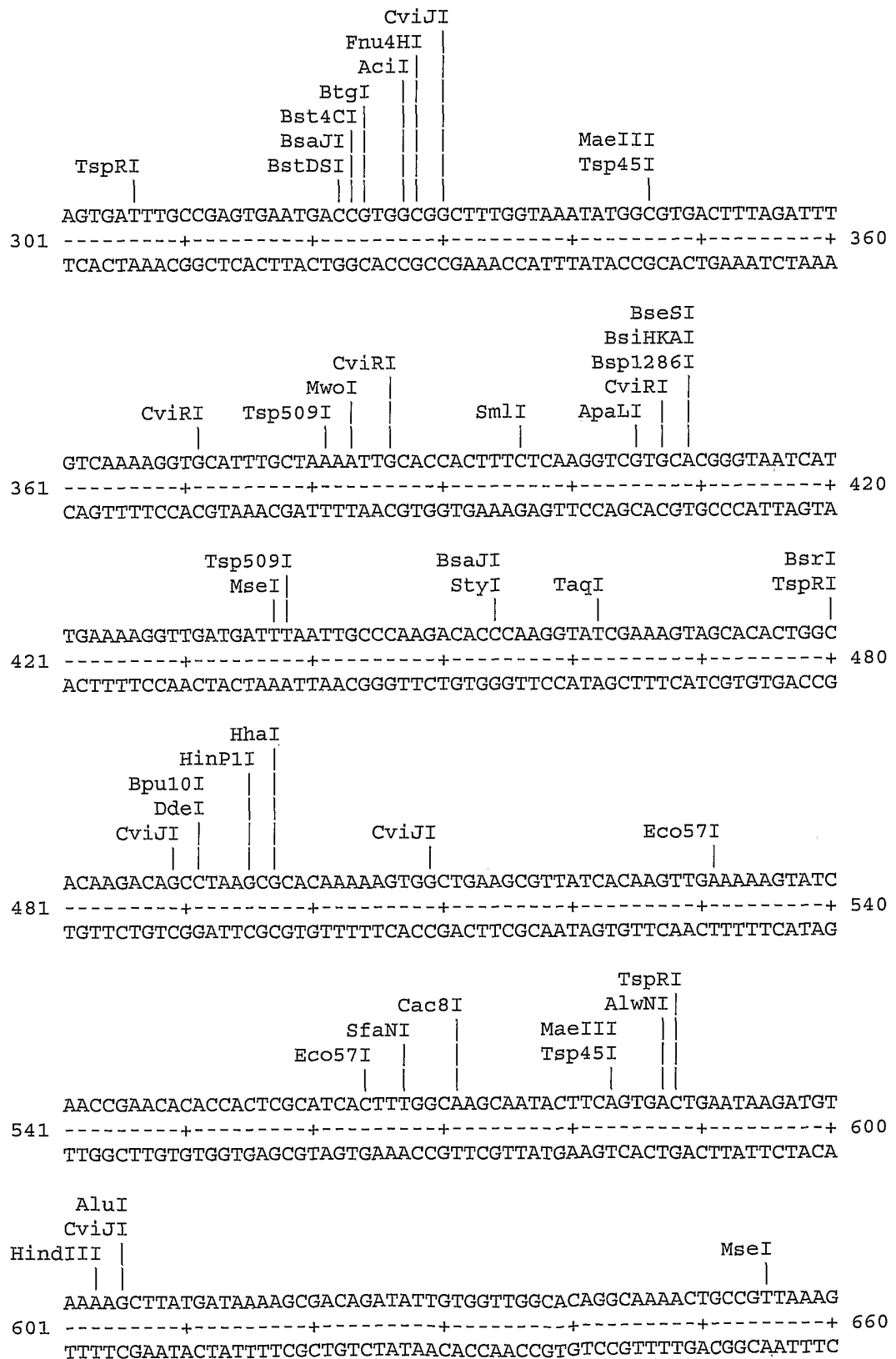
```

84/189

Figure 30. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a rare lipoprotein A (SEQ ID NO: 29) designated MCA101388



85/189



86/189

```
          TspRI   CviJI  
          |       |  
GTGCTTTGATTATCACTGAATACACAGGCTATCTG  
661 -----+-----+-----+----- 695  
CACGAAACTAATAGTGACTTATGTGTCCGATAGAC
```

87/189

Figure 31. Sequence of *M. catarrhalis* gene encoding a 36kDa secreted protein designated MCA101446

```

acaattgatg actttgatga ttatttgtct ttggatgatg ctcataaact gtgattgata 60

acgctcatatc ttaacccccaa ccatttgaca gcccaaaacc atg acc gat aag acc 115
Met Thr Asp Lys Thr
1 5

gac gat gtt aag cca gat acc gat ggc aac cac cca aaa aaa caa aaa 163
Asp Asp Val Lys Pro Asp Thr Asp Gly Asn His Pro Lys Lys Gln Lys
10 15 20

ccc cac aca caa gcc aaa tca aag ccc aga cct ttg gcg tta ttc tta 211
Pro His Thr Gln Ala Lys Ser Lys Pro Arg Pro Leu Ala Leu Phe Leu
25 30 35

tta ttg gca ctc att ggc agt gtg ctt ttg gtg att att ttt ggt ttg 259
Leu Leu Ala Leu Ile Gly Ser Val Leu Leu Val Ile Ile Phe Gly Leu
40 45 50

ata caa agc tat ggc aca ccg caa aac ttg cca aca gct aac gcc att 307
Ile Gln Ser Tyr Gly Thr Pro Gln Asn Leu Pro Thr Ala Asn Ala Ile
55 60 65

ggg agc aca aat gag ccg cca aag ggg gtg att gcc cat ctc aga caa 355
Gly Ser Thr Asn Glu Pro Pro Lys Gly Val Ile Ala His Leu Arg Gln
70 75 80 85

aag att acc gga cct gcc aca ttt agc att caa ggg cgt gtg gag ggc 403
Lys Ile Thr Gly Pro Ala Thr Phe Ser Ile Gln Gly Arg Val Glu Gly
90 95 100

gaa acg gtg aac atc agc acc aaa ctg ccc agt cgt gtt gct gcc att 451
Glu Thr Val Asn Ile Ser Thr Lys Leu Pro Ser Arg Val Ala Ala Ile
105 110 115

tat gtc aaa gag ggt caa gct gtc agg gca ggc gaa cct ttg gtt cag 499
Tyr Val Lys Glu Gly Gln Ala Val Arg Ala Gly Glu Pro Leu Val Gln
120 125 130

ctt atc agt cct gaa att gac gcc aaa aaa cag caa gcc gct gcc atg 547
Leu Ile Ser Pro Glu Ile Asp Ala Lys Lys Gln Gln Ala Ala Ala Met
135 140 145

tta cag tca gcc ctt gcc ttt cag tcg tcc aca gac cgt ggc aca aga 595
Leu Gln Ser Ala Leu Ala Phe Gln Ser Ser Thr Asp Arg Gly Thr Arg
150 155 160 165

caa gaa aat gta gac acc ctg tac gcc aat tgg caa agt gcc aaa gct 643
Gln Glu Asn Val Asp Thr Leu Tyr Ala Asn Trp Gln Ser Ala Lys Ala
170 175 180

cag gca acg ctg gca gag gag agc tat cgc cgt ggt gag tat ttg tat 691
Gln Ala Thr Leu Ala Glu Glu Ser Tyr Arg Arg Gly Glu Tyr Leu Tyr
185 190 195

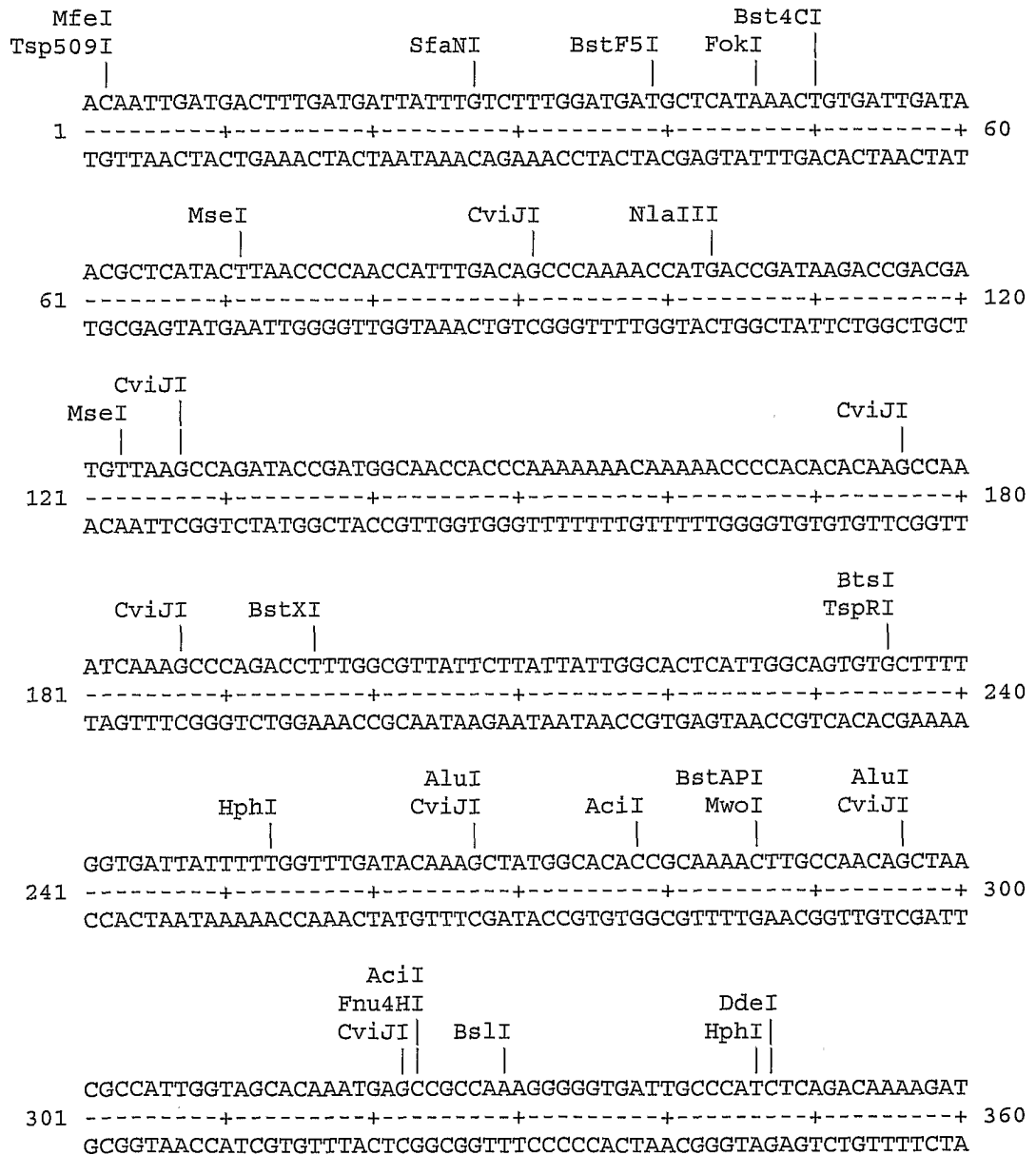
```

88/189

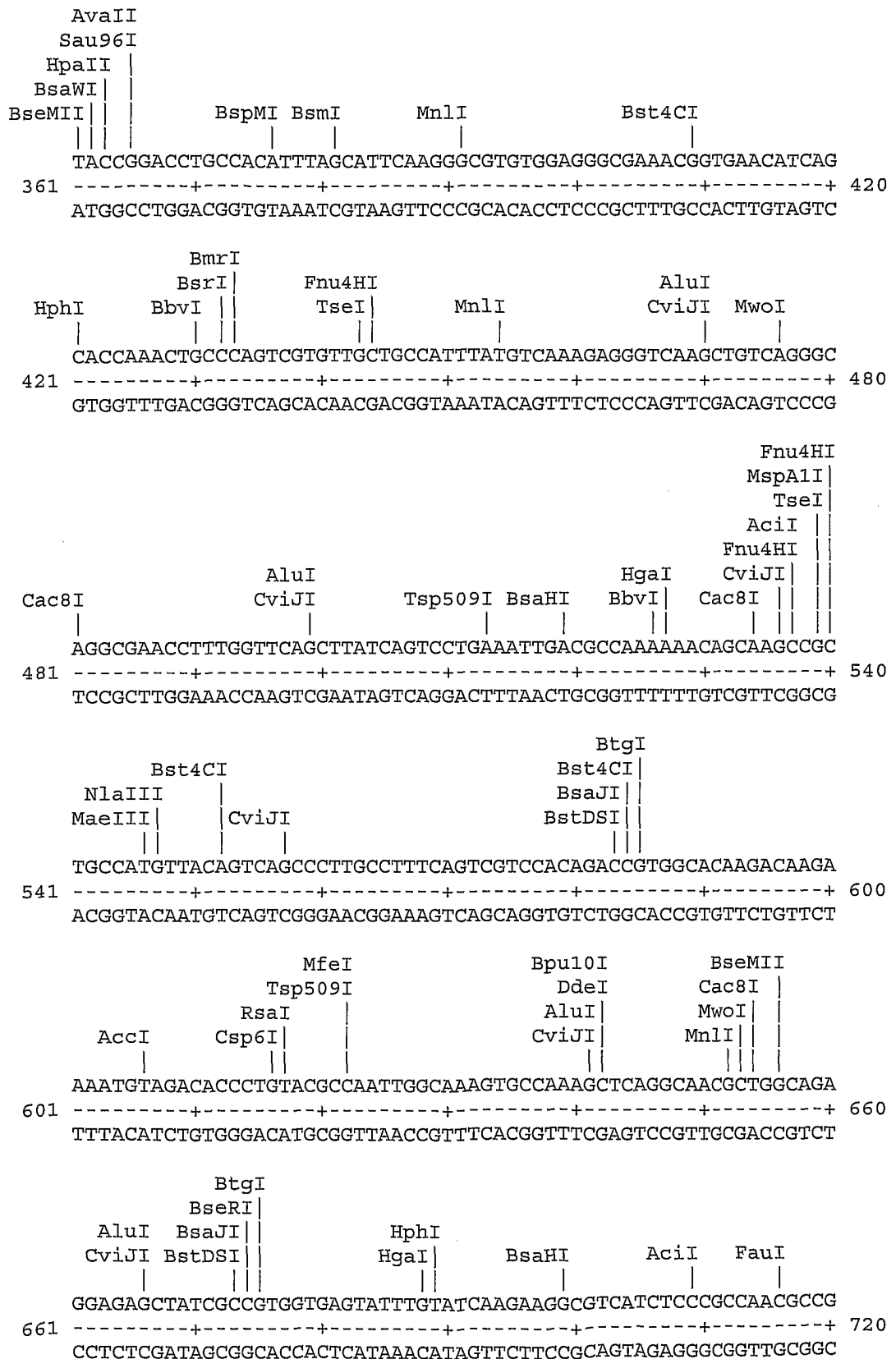
caa gaa ggc gtc atc tcc cgc caa cgc cgt gat gag atg agg gca gcg	739
Gln Glu Gly Val Ile Ser Arg Gln Arg Arg Asp Glu Met Arg Ala Ala	
200 205 210	
aag cag tct tct gcc caa att gcc cat gca gca cag cag cag tat cta	787
Lys Gln Ser Ser Ala Gln Ile Ala His Ala Ala Gln Gln Gln Tyr Leu	
215 220 225	
aaa gcc cag cgt ggt cgc act gat gaa tta aaa tcg tcc gcc gac gcc	835
Lys Ala Gln Arg Gly Arg Thr Asp Glu Leu Lys Ser Ser Ala Asp Ala	
230 235 240 245	
caa gtc caa atc gct cgt gcg gcg gtt gcc gaa gcc gat gcc cta gat	883
Gln Val Gln Ile Ala Arg Ala Ala Val Ala Glu Ala Asp Ala Leu Asp	
250 255 260	
gct gaa acc atg ctt tat gcc ccc att gat ggc acg gtc tca aaa atc	931
Ala Glu Thr Met Leu Tyr Ala Pro Ile Asp Gly Thr Val Ser Lys Ile	
265 270 275	
tac att caa agc agt gaa ctg gtc atg cca gcg gtg cct gta atg agc	979
Tyr Ile Gln Ser Ser Glu Leu Val Met Pro Ala Val Pro Val Met Ser	
280 285 290	
ctg ctt gac aac cac cct acc att acc ata agc atc caa gaa aat cag	1027
Leu Leu Asp Asn His Pro Thr Ile Thr Ile Ser Ile Gln Glu Asn Gln	
295 300 305	
tac agt cat gtc cat gcc ctg ccc aag cta act ggt ttt atc cct gca	1075
Tyr Ser His Val His Ala Leu Pro Lys Leu Thr Gly Phe Ile Pro Ala	
310 315 320 325	
ctg aac aaa acc gcc gat ttt aaa atc gtg cat acc gat gat gag gga	1123
Leu Asn Lys Thr Ala Asp Phe Lys Ile Val His Thr Asp Asp Glu Gly	
330 335 340	
gag ttt gcc acc atc aaa aat act aga cag aca ggc ggt tat gac att	1171
Glu Phe Ala Thr Ile Lys Asn Thr Arg Gln Thr Gly Gly Tyr Asp Ile	
345 350 355	
cat agc ttt aaa acc aaa ctt gaa ccc atc acc ccc att gaa ggg tta	1219
His Ser Phe Lys Thr Lys Leu Glu Pro Ile Thr Pro Ile Glu Gly Leu	
360 365 370	
aaa gtg ggc atg agt gtg ata ttt gat atc act gtt ttg cct taa	1264
Lys Val Gly Met Ser Val Ile Phe Asp Ile Thr Val Leu Pro	
375 380 385	
attttcatcc aaataagtct gatattccatg attgttcgct ttattcgtgc tttttgccaa	1324
gcattttttgc gtgagcttaa ccgcttatca tgtaaccgat ttgatttggc gatggtcatg	1384
atagcaccgc tt	1396

89/189

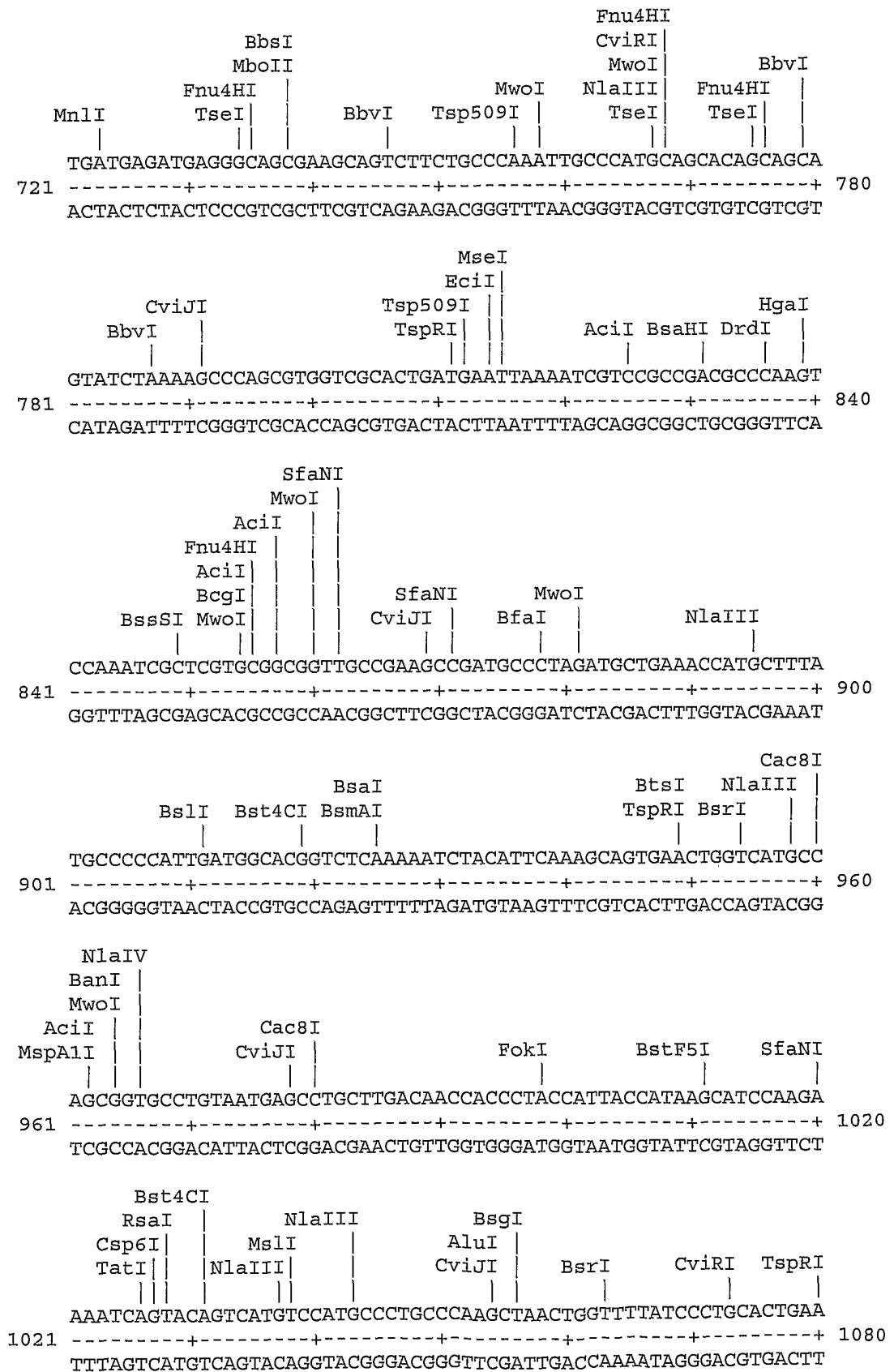
Figure 32. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a 36kDa secreted protein designated MCA101446



90/189



91/189



92/189

```

          AclI      DraI      MnlI
          |        |        |
          |        |        |
    CAAAACCGCCGATTTTAAATCGTGCATACCGATGATGAGGGAGAGTTTGCCACCATCAA
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
    GTTTTGGCGGCTAAAATTTTAGCACGTATGGCTACTACTCCCTCTCAAACGGTGGTAGTT

                                DraI
                                MseI
                                |
          BfaI      AclI      AluI      CviJI      HphI
          |        |        |        |        |
    AAATACTAGACAGACAGGCGGTTATGACATTCATAGCTTTAAAAACCAAACCTGAACCCAT
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
    TTTATGATCTGTCTGTCCGCCAATACTGTAAGTATCGAAATTTGGTTTGAACCTGGGTA

                                FokI
                                TspRI
                                |
          BslI      MseI      MslI      Bst4CI      |
          |        |        |        |        |
    CACCCCATTTGAAGGGTTAAAAGTGGGCATGAGTGTGATATTTGATATCACTGTTTGGC
1201 -----+-----+-----+-----+-----+-----+-----+ 1260
    GTGGGGGTAACTTCCCAATTTTCACCCGTACTCACACTATAAACTATAGTGACAAAACGG

    ApoI
    Tsp509I
    MseI | BstF5I      EcoRV NlaIII
    |    |        |        |
    TTAAATTTTCATCCAAATAAGTCTGATATCCATGATTGTTTCGCTTTATTTCGTGCTTTTTG
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
    AATTTAAAAGTAGGTTTATTTCAGACTATAGGTACTAACAAGCGAAATAAGCACGAAAAAC

                                MseI
                                |
          AluI      NlaIII
          |        |
          |        |
    CCAAGCATTTTTCGCTGAGCTTAACCGCTTATCATGTAACCGATTGATTGGCGATGGT
1321 -----+-----+-----+-----+-----+-----+-----+ 1380
    GGTTCGTAAAAACGCACTCGAATTGGCGAATAGTACATTGGCTAAACTAAACCGCTACCA

    NlaIII      AclI
    |          |
    CATGATAGCACCGCTT
1381 -----+-----+-----+-----+-----+-----+-----+ 1396
    GTACTATCGTGGCGAA

```

93/189

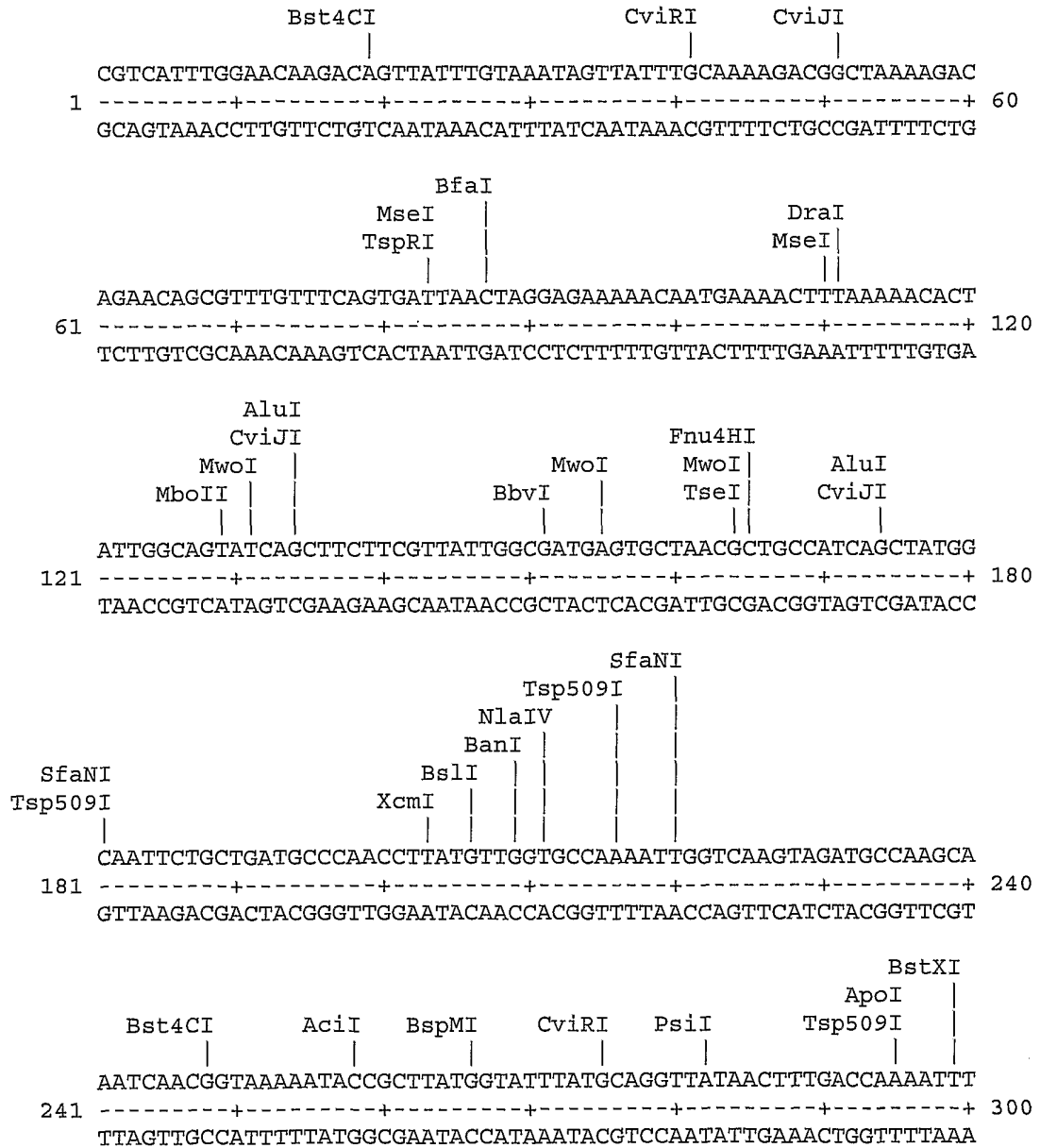
Figure 33. Sequence of *M. catarrhalis* gene (SEQ ID NOs: 33 & 34) designated MCA101580

```

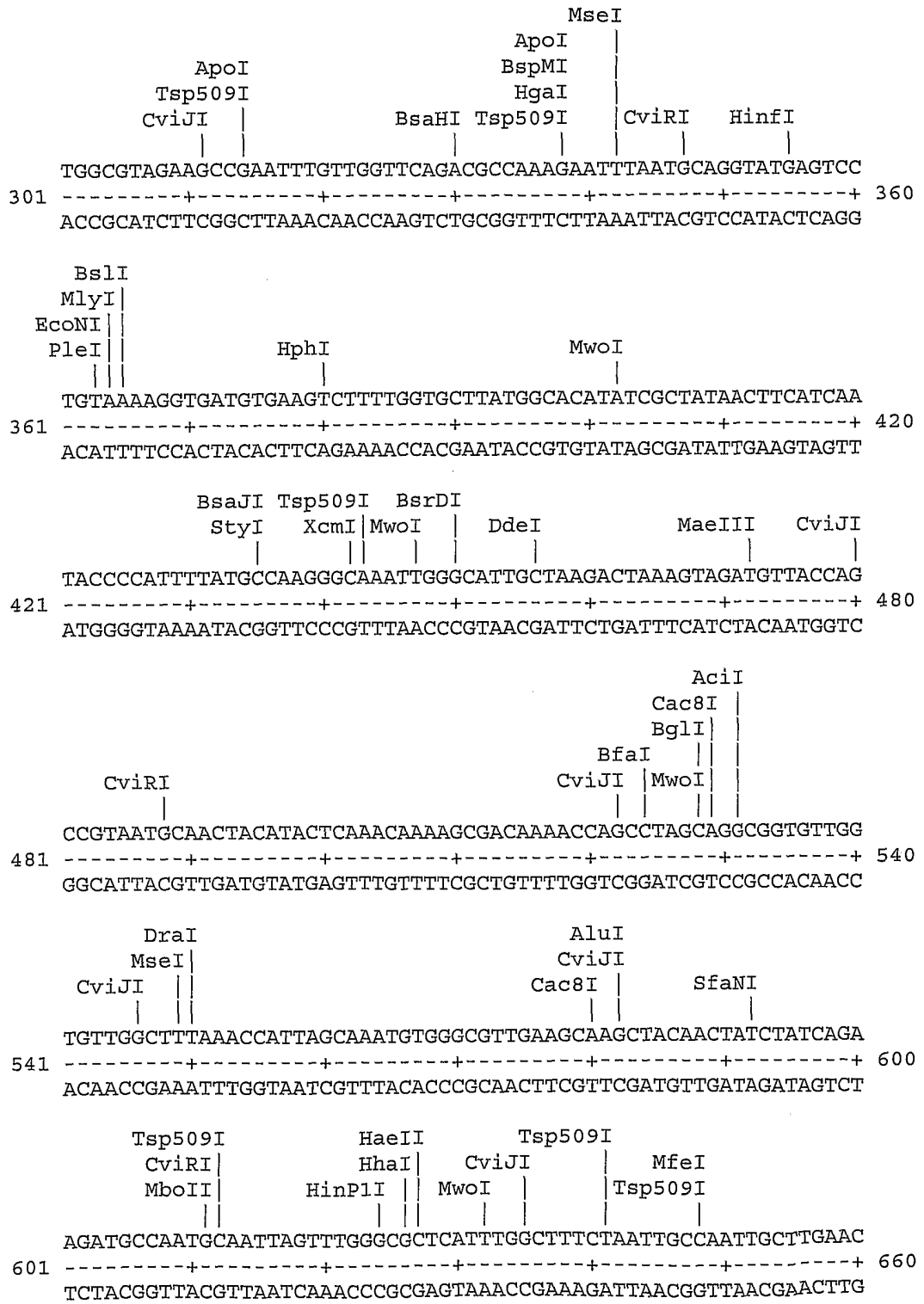
cgtcatttgg aacaagacag ttatttgtaa atagttattt gcaaaagacg gctaaaagac 60
agaacagcgt ttgtttcagt gattaactag gagaaaaaca atg aaa act tta aaa 115
                                         Met Lys Thr Leu Lys
                                         1 5
aca cta ttg gca gta tca gct tct tcg tta ttg gcg atg agt gct aac 163
Thr Leu Leu Ala Val Ser Ala Ser Ser Leu Leu Ala Met Ser Ala Asn
                               10 15 20
gct gcc atc agc tat ggc aat tct gct gat gcc caa cct tat gtt ggt 211
Ala Ala Ile Ser Tyr Gly Asn Ser Ala Asp Ala Gln Pro Tyr Val Gly
                               25 30 35
gcc aaa att ggt caa gta gat gcc aag caa atc aac ggt aaa aat acc 259
Ala Lys Ile Gly Gln Val Asp Ala Lys Gln Ile Asn Gly Lys Asn Thr
                               40 45 50
gct tat ggt att tat gca ggt tat aac ttt gac caa aat ttt ggc gta 307
Ala Tyr Gly Ile Tyr Ala Gly Tyr Asn Phe Asp Gln Asn Phe Gly Val
                               55 60 65
gaa gcc gaa ttt gtt ggt tca gac gcc aaa gaa ttt aat gca ggt atg 355
Glu Ala Glu Phe Val Gly Ser Asp Ala Lys Glu Phe Asn Ala Gly Met
                               70 75 80 85
agt cct gta aaa ggt gat gtg aag tct ttt ggt gct tat ggc aca tat 403
Ser Pro Val Lys Gly Asp Val Lys Ser Phe Gly Ala Tyr Gly Thr Tyr
                               90 95 100
cgc tat aac ttc atc aat acc cca ttt tat gcc aag ggc aaa ttg ggc 451
Arg Tyr Asn Phe Ile Asn Thr Pro Phe Tyr Ala Lys Gly Lys Leu Gly
                               105 110 115
att gct aag act aaa gta gat gtt acc agc cgt aat gca act aca tac 499
Ile Ala Lys Thr Lys Val Asp Val Thr Ser Arg Asn Ala Thr Thr Tyr
                               120 125 130
tca aac aaa agc gac aaa acc agc cta gca ggc ggt gtt ggt gtt ggc 547
Ser Asn Lys Ser Asp Lys Thr Ser Leu Ala Gly Gly Val Gly Val Gly
                               135 140 145
ttt aaa cca tta gca aat gtg ggc gtt gaa gca agc tac aac tat cta 595
Phe Lys Pro Leu Ala Asn Val Gly Val Glu Ala Ser Tyr Asn Tyr Leu
                               150 155 160 165
tca gaa gat gcc aat gca att agt ttg ggc gct cat ttg gct ttc taa 643
Ser Glu Asp Ala Asn Ala Ile Ser Leu Gly Ala His Leu Ala Phe
                               170 175 180
ttgccaattg cttgaacaac ataaaaaaga tcctcgaatt gaggggtcttt tttattggtc 703
agtttatccg taagataagc catcatgata aattatcttt 743

```

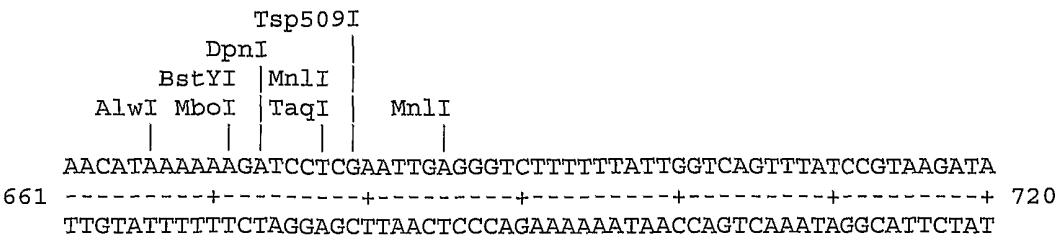
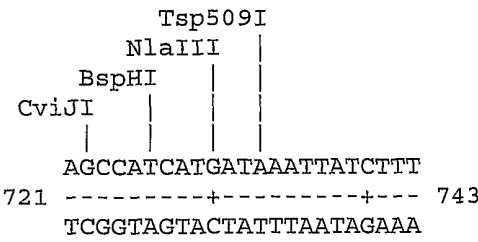
Figure 34. Restriction enzyme analysis of a *M. catarrhalis* gene, as shown in SEQ ID NO: 33 designated MCA101580



95/189



96/189



97/189

Figure 35. Sequence of *M. catarrhalis* gene encoding a rare lipoprotein A (SEQ ID NOS: 35 & 36) designated MCA101666

```

ctgatttgtg cctatttgta catttttgtg tacatatattt ggtcagtatg ttgagatttt 60

ctcaattatt gtcagagtat gcaataactaa aggtgtaacc atg aaa aca gtg agc 115
               Met Lys Thr Val Ser
                   1           5

aaa tct tta ctt gga tta gcg agc att act tta acc gca agt tta agc 163
Lys Ser Leu Leu Gly Leu Ala Ser Ile Thr Leu Thr Ala Ser Leu Ser
                   10           15           20

atg aat gca caa gct ggc aat agc cct gcc aat atc gac aat gta cta 211
Met Asn Ala Gln Ala Gly Asn Ser Pro Ala Asn Ile Asp Asn Val Leu
                   25           30           35

cag cag ctc acc acg cag cat aat aat gcg tca agc ttg gtc aaa tca 259
Gln Gln Leu Thr Thr Gln His Asn Asn Ala Ser Ser Leu Val Lys Ser
                   40           45           50

gtg cgt caa ccc gcc tct ttg gcg ctt aac agt cca atg att gct caa 307
Val Arg Gln Pro Ala Ser Leu Ala Leu Asn Ser Pro Met Ile Ala Gln
                   55           60           65

tcc acc aaa tcc aaa aaa acc gat gaa gat att cct gtg ctt gac cgt 355
Ser Thr Lys Ser Lys Lys Thr Asp Glu Asp Ile Pro Val Leu Asp Arg
                   70           75           80           85

ctg acc gca gtc gct tca agc act gtc aat aag ttc aaa caa aat ggc 403
Leu Thr Ala Val Ala Ser Ser Thr Val Asn Lys Phe Lys Gln Asn Gly
                   90           95           100

atc gct tca tgg tat ggt cgt cag ttt cat ggt cgt aaa act gcc agt 451
Ile Ala Ser Trp Tyr Gly Arg Gln Phe His Gly Arg Lys Thr Ala Ser
                   105           110           115

ggc gaa acc ttt gat atg aat gcc tta acg gca gca cac agt agc ctg 499
Gly Glu Thr Phe Asp Met Asn Ala Leu Thr Ala Ala His Ser Ser Leu
                   120           125           130

cca atg aat tgc cat gtc aga gtg acc aat cgt gat aat ggt aaa tct 547
Pro Met Asn Cys His Val Arg Val Thr Asn Arg Asp Asn Gly Lys Ser
                   135           140           145

gtg gtc gtc aaa atc aac gat cga ccc aaa acc aat cgt gtg ctt gat 595
Val Val Val Lys Ile Asn Asp Arg Pro Lys Thr Asn Arg Val Leu Asp
                   150           155           160           165

ttg tct tat ggt gct gcc caa gcc att ggt atg act ggt aat gtt ggt 643
Leu Ser Tyr Gly Ala Ala Gln Ala Ile Gly Met Thr Gly Asn Val Gly
                   170           175           180

```

98/189

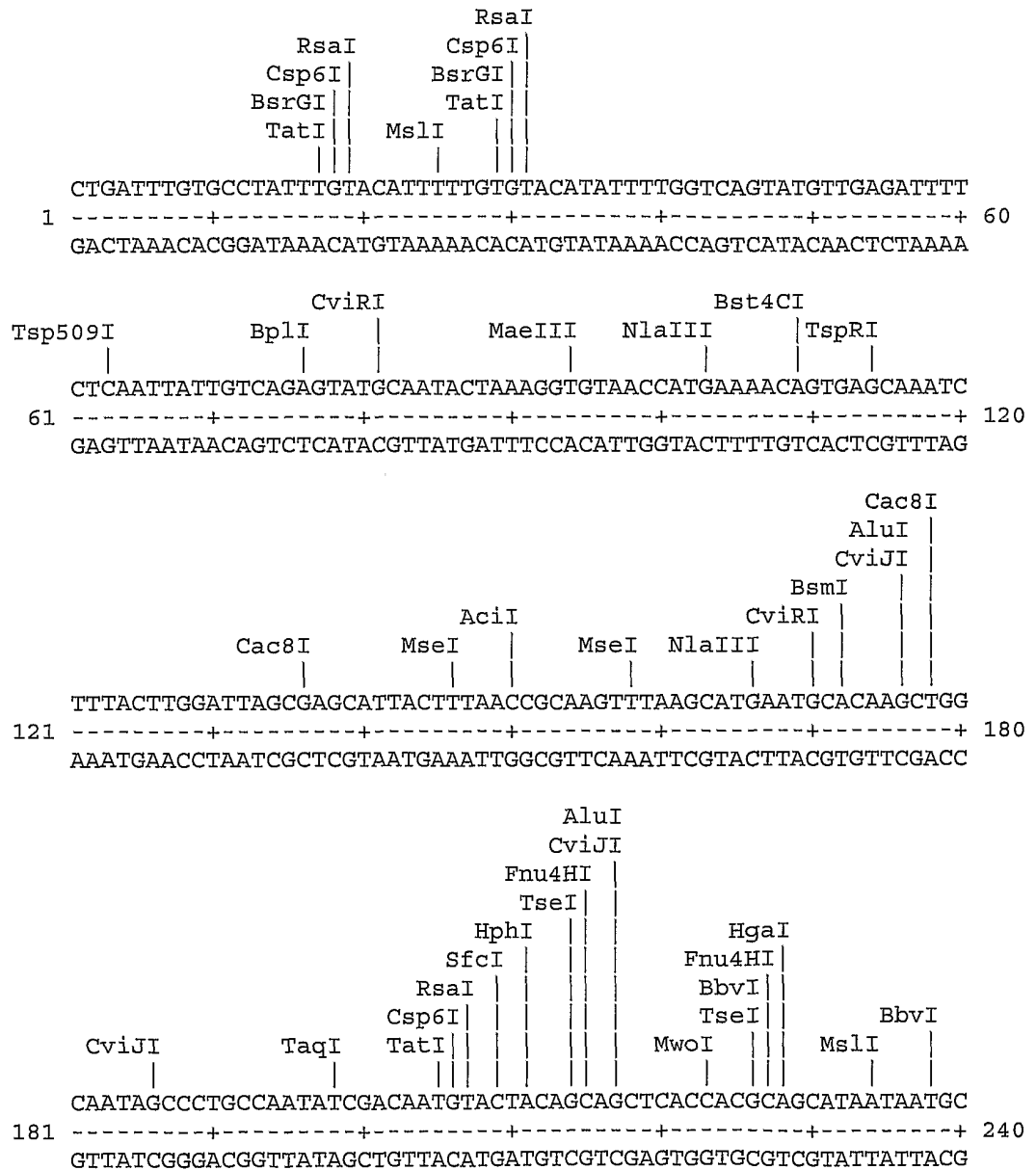
```
aat gta acc att gag cgt att gat taa tcaccttgat tatttatatt      690
Asn Val Thr Ile Glu Arg Ile Asp
          185                      190

aatcattttg attatttatc aagatatcca aaaccagctg ctaaagctgg tttttttgta 750

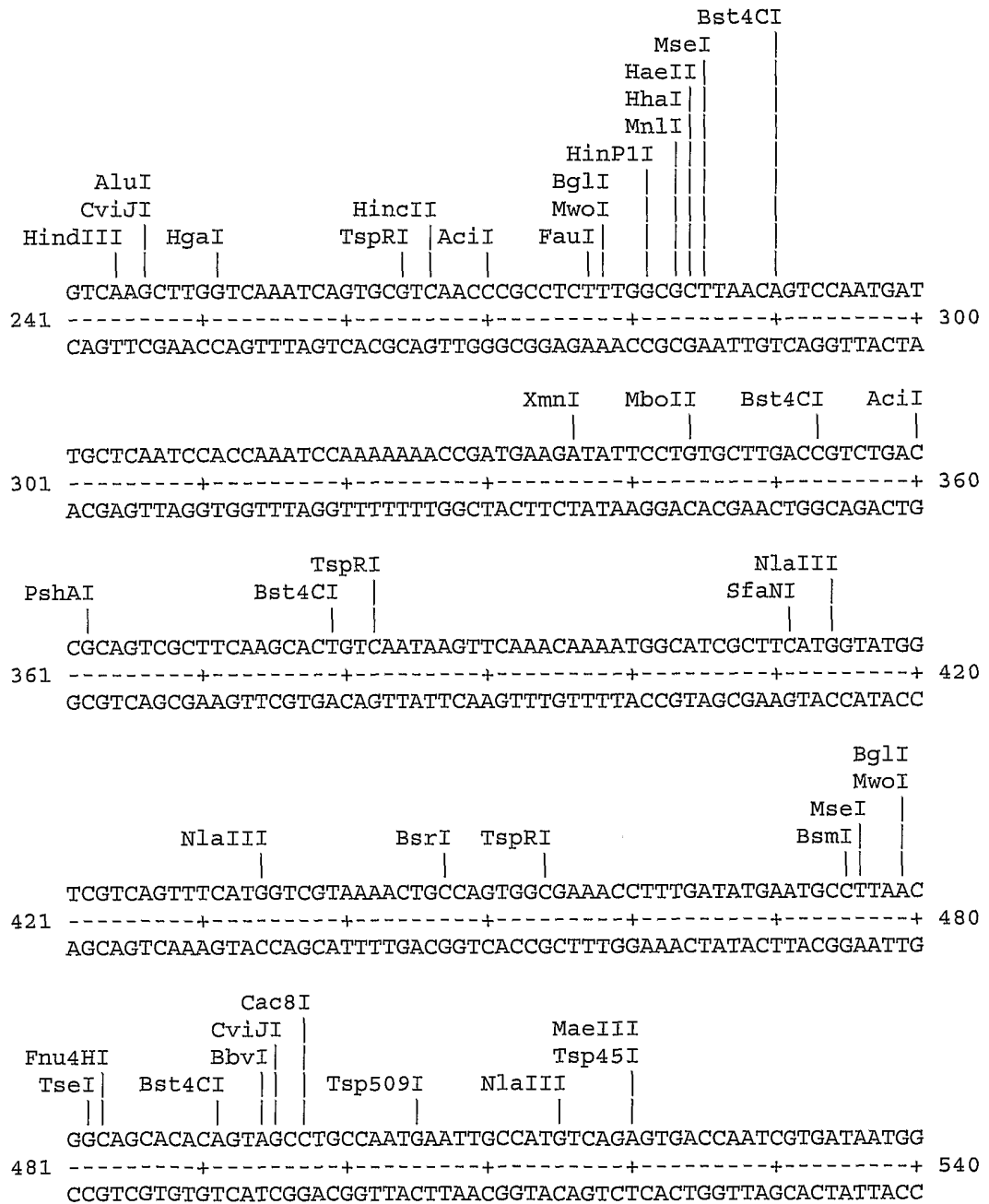
ttaatttttag aaaaacaatt                                           770
```

99/189

Figure 36. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a rare lipoprotein A (SEQ ID NO: 35) designated MCA101666



100/189



101/189

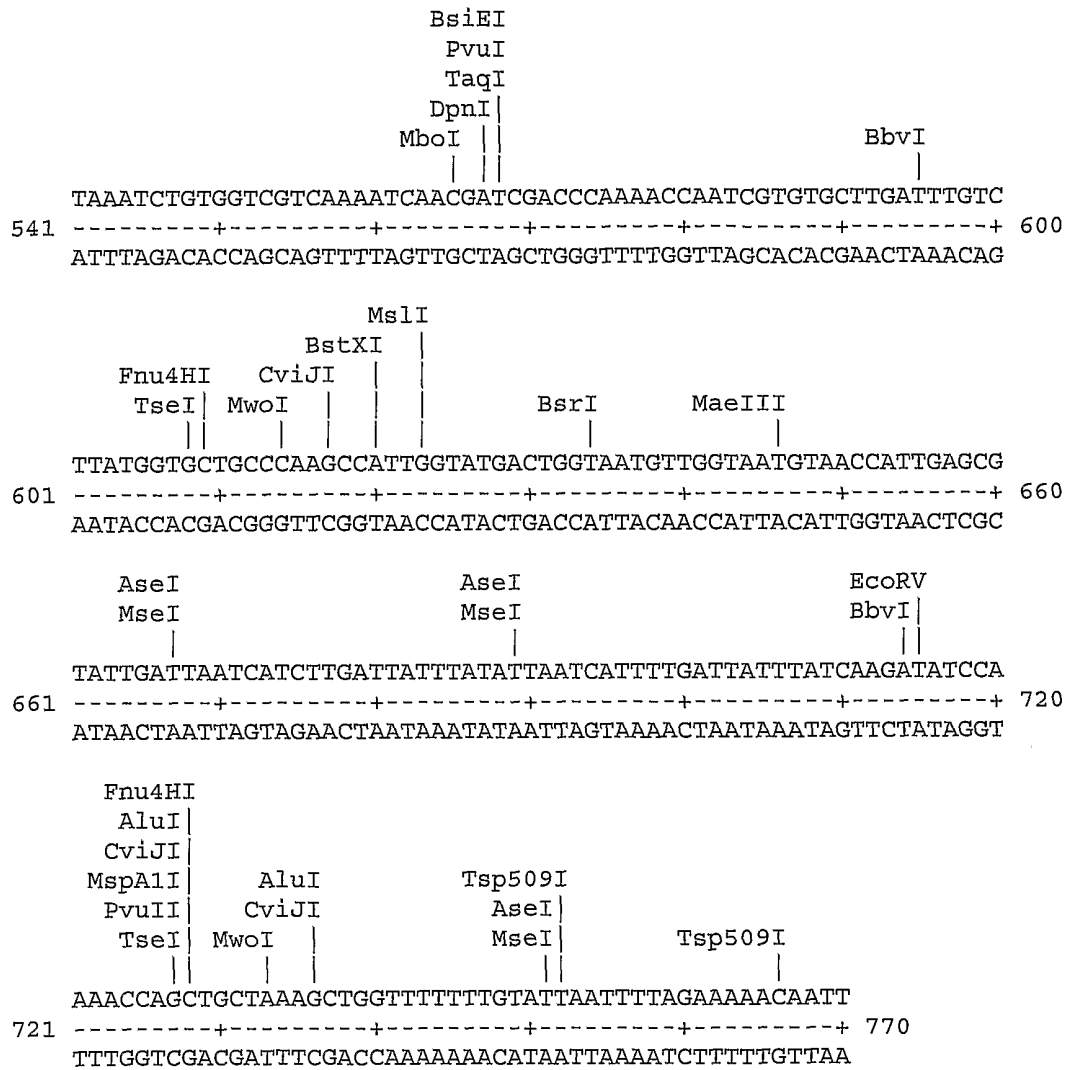


Figure 37. Sequence of *M. catarrhalis* OMP gene designated MCA101792

tgag	ttaatt	acctgt	tatta	attact	ttgca	tcaagt	ttgag	agtggg	gcag	agtttg	ggtgg	60				
tttt	ttttt	tg	tatttt	tattt	ggtatt	tttta	aggaac	gggt	atg	aaa	aag	cta	ctt	115		
									Met	Lys	Lys	Leu	Leu	5		
									1							
ttg	gca	gcc	cta	ttg	aca	agc	act	tca	aca	ctg	gca	atg	gct	caa	atc	163
Leu	Ala	Ala	Leu	Leu	Thr	Ser	Thr	Ser	Thr	Leu	Ala	Met	Ala	Gln	Ile	
				10					15					20		
tcc	acc	gat	ggt	ctg	tat	gta	caa	ggt	aat	gtt	ggc	gta	tca	aaa	tta	211
Ser	Thr	Asp	Gly	Leu	Tyr	Val	Gln	Gly	Asn	Val	Gly	Val	Ser	Lys	Leu	
			25					30					35			
aat	gcc	aag	gag	gac	ggt	gag	aag	gct	aaa	cgc	aac	gca	act	caa	tac	259
Asn	Ala	Lys	Glu	Asp	Gly	Glu	Lys	Ala	Lys	Arg	Asn	Ala	Thr	Gln	Tyr	
		40					45					50				
acc	atc	gcc	gtc	ggt	aaa	gac	atg	ggt	gcg	ata	cgt	ttt	cag	gca	gac	307
Thr	Ile	Ala	Val	Gly	Lys	Asp	Met	Gly	Ala	Ile	Arg	Phe	Gln	Ala	Asp	
	55					60					65					
tat	acc	aac	ttt	ggc	aaa	atg	aaa	aat	aac	ggt	tca	gaa	ggc	aag	cga	355
Tyr	Thr	Asn	Phe	Gly	Lys	Met	Lys	Asn	Asn	Gly	Ser	Glu	Gly	Lys	Arg	
70					75					80					85	
gac	gag	gtg	agg	tat	gaa	gaa	tgg	aaa	gct	tcc	tta	aaa	gct	cat	tca	403
Asp	Glu	Val	Arg	Tyr	Glu	Glu	Trp	Lys	Ala	Ser	Leu	Lys	Ala	His	Ser	
				90					95					100		
ttg	ggc	gtg	tct	gcc	att	tat	gat	ttt	aac	acc	gca	aca	ggg	ctg	acc	451
Leu	Gly	Val	Ser	Ala	Ile	Tyr	Asp	Phe	Asn	Thr	Ala	Thr	Gly	Leu	Thr	
			105					110					115			
ccc	tat	gca	ggc	atg	cgt	gtt	ggt	gcc	aat	cag	cta	aaa	gtt	gac	ggt	499
Pro	Tyr	Ala	Gly	Met	Arg	Val	Gly	Ala	Asn	Gln	Leu	Lys	Val	Asp	Gly	
		120					125					130				
caa	gat	ctt	tat	aat	aag	att	aca	cca	gac	ggt	agg	cta	caa	cta	gta	547
Gln	Asp	Leu	Tyr	Asn	Lys	Ile	Thr	Pro	Asp	Gly	Arg	Leu	Gln	Leu	Val	
	135					140					145					
aaa	gat	atc	aaa	tca	aag	aaa	agc	aac	aaa	gtc	agc	gtg	ggt	gca	gtt	595
Lys	Asp	Ile	Lys	Ser	Lys	Lys	Ser	Asn	Lys	Val	Ser	Val	Gly	Ala	Val	
150					155					160					165	
gca	ggt	gtg	caa	tac	gcc	ttt	aat	cca	cag	ctg	gca	tta	gat	gca	ggt	643
Ala	Gly	Val	Gln	Tyr	Ala	Phe	Asn	Pro	Gln	Leu	Ala	Leu	Asp	Ala	Gly	
			170						175					180		

103/189

gta gaa tat cac tac tta ggt aag att gag agt ctt aaa ctc aat caa 691
Val Glu Tyr His Tyr Leu Gly Lys Ile Glu Ser Leu Lys Leu Asn Gln
185 190 195

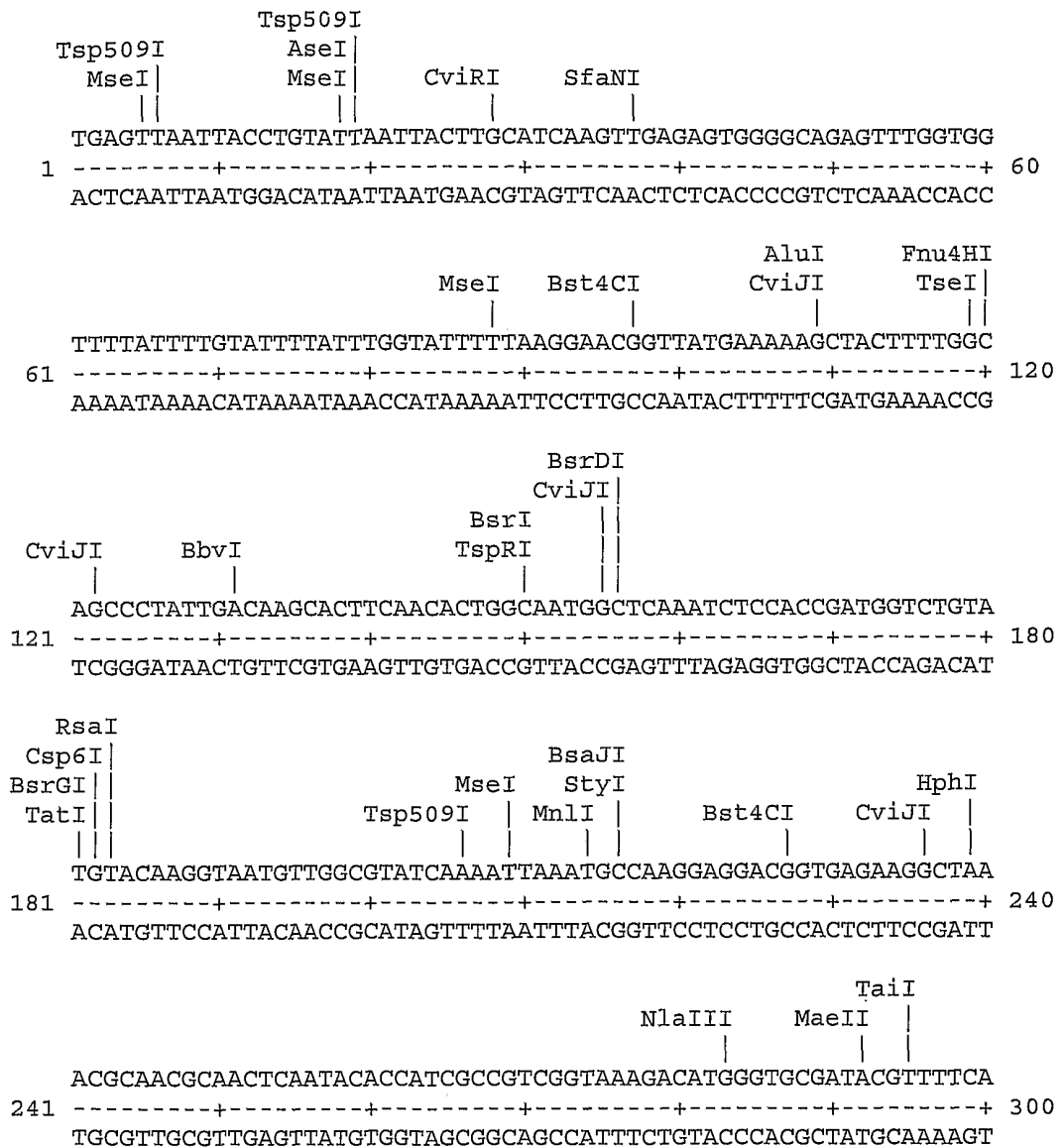
tat ggt gcg aaa gtg ggt ctt cgc cat aac ttt taa aattagcaag 737
Tyr Gly Ala Lys Val Gly Leu Arg His Asn Phe
200 205

ctgctgattt taagcttgct aaaaaaccca tcggtacgat atatcgatgg gtttttgtaa 797

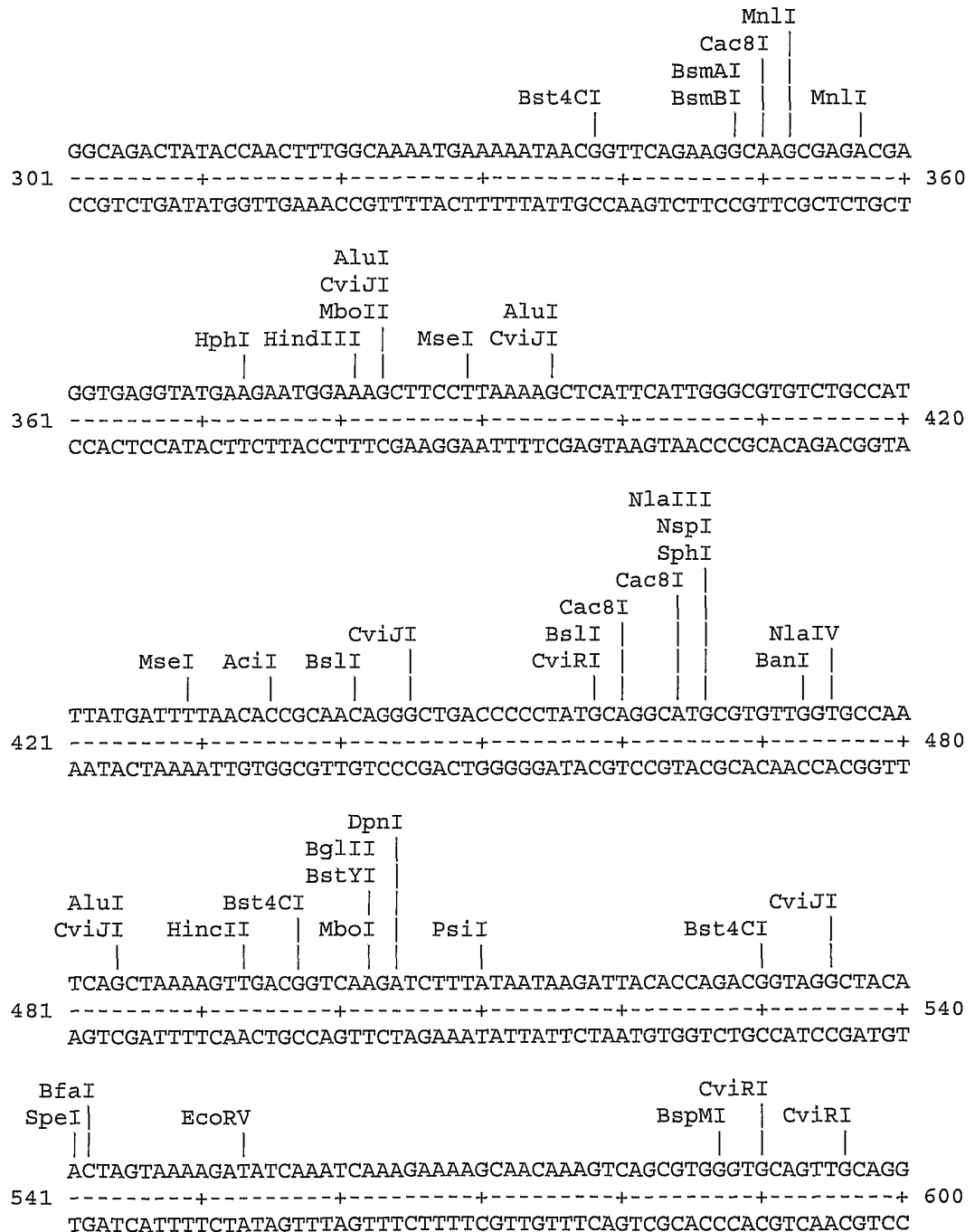
aaagtgataa aaatcagcaa aacagctgac 827

104/189

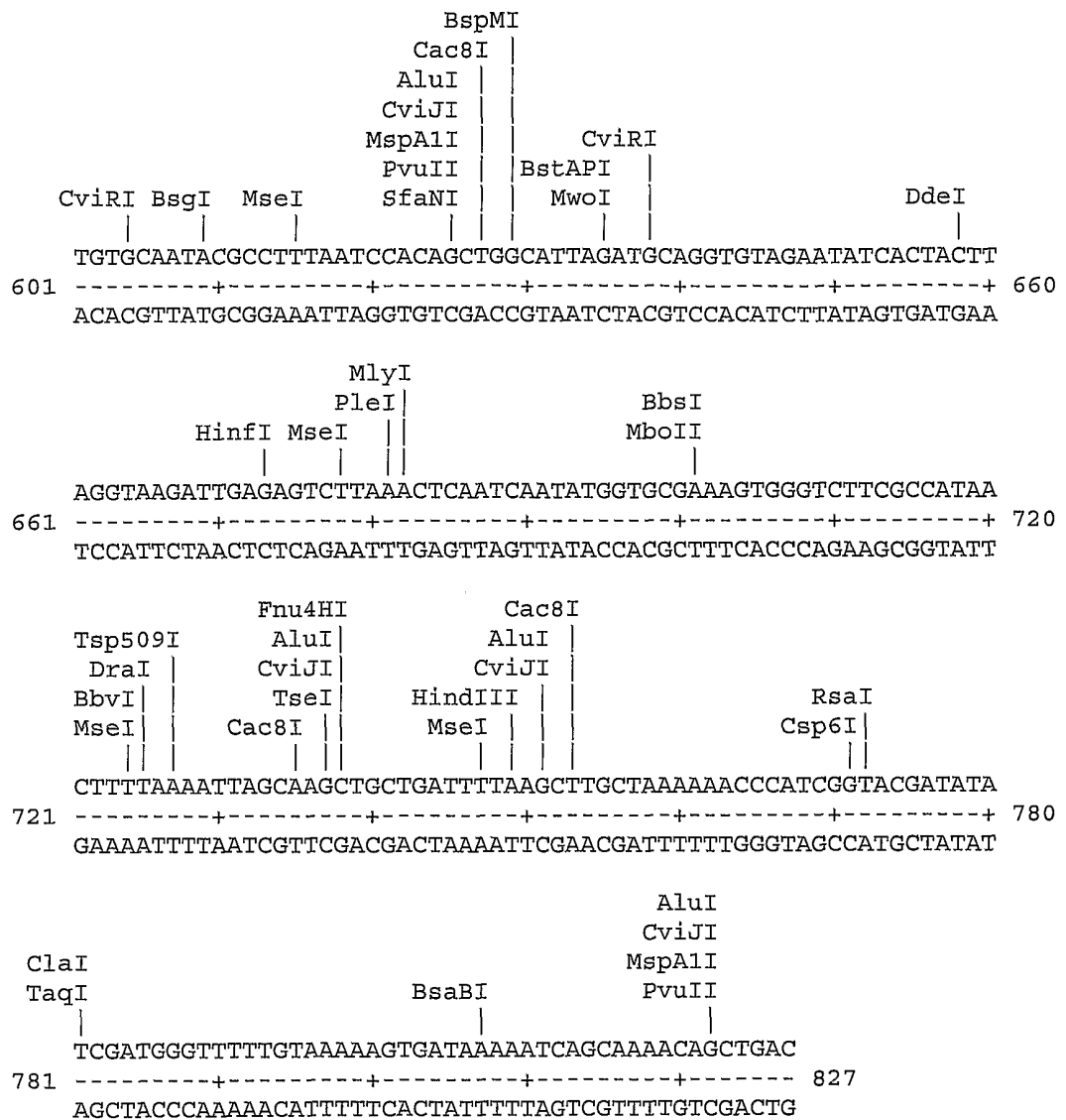
Figure 38. Restriction enzyme analysis of the *M. catarrhalis* OMP gene designated MCA101792



105/189



106/189



107/189

Figure 39. Sequence of *M. catarrhalis* outer membrane esterase gene designated MCA101806

```

aacgactaag gccaaaagtc ttcataactca aaatatcgcc caaagaacgc tcaaagcggc 60
gggtcataca taccgccctt tatcaatggc gattgcactg atg ggc tta agc agc 115
                                         Met Gly Leu Ser Ser
                                         1                               5

atc gcc acc gcc caa gaa ttt agc caa acc gta ttt ttt ggt gat agc 163
Ile Ala Thr Ala Gln Glu Phe Ser Gln Thr Val Phe Phe Gly Asp Ser
                        10                               15                               20

cta act gac aca gga cgc ttg gtg cag att ggc aaa aac agt ctg gcg 211
Leu Thr Asp Thr Gly Arg Leu Val Gln Ile Gly Lys Asn Ser Leu Ala
                        25                               30                               35

tct tct gtt ttt aat caa gcc caa cca tcg ttt acg acc aat acc gat 259
Ser Ser Val Phe Asn Gln Ala Gln Pro Ser Phe Thr Thr Asn Thr Asp
                        40                               45                               50

cca gtt tgg tca agc ata ttg gca aaa tct tat gga cat acc gcc gat 307
Pro Val Trp Ser Ser Ile Leu Ala Lys Ser Tyr Gly His Thr Ala Asp
                        55                               60                               65

gcc aat gat gga aca acc tta act ggt aca aac tat gcc gtt ggt ggt 355
Ala Asn Asp Gly Thr Thr Leu Thr Gly Thr Asn Tyr Ala Val Gly Gly
                        70                               75                               80                               85

gca aga act aaa gaa gat gtg gtc aaa aat gca cct gtt cct ttt ttc 403
Ala Arg Thr Lys Glu Asp Val Val Lys Asn Ala Pro Val Pro Phe Phe
                        90                               95                               100

acc att cct tta ttt acc atc cca tca gca caa acc caa atc aat cgc 451
Thr Ile Pro Leu Phe Thr Ile Pro Ser Ala Gln Thr Gln Ile Asn Arg
                        105                               110                               115

tac ctg acc tta aat aat cat caa gcc gac ccc aaa gca ttg tat act 499
Tyr Leu Thr Leu Asn Asn His Gln Ala Asp Pro Lys Ala Leu Tyr Thr
                        120                               125                               130

gtt tgg aca ggt gcc aat gat ttg ttt gag gca gcc aaa gca cca acc 547
Val Trp Thr Gly Ala Asn Asp Leu Phe Glu Ala Ala Lys Ala Pro Thr
                        135                               140                               145

cag ttg caa gca gcc gaa atc att acc acc gct gcc aat gac caa gcc 595
Gln Leu Gln Ala Ala Glu Ile Ile Thr Thr Ala Ala Asn Asp Gln Ala
                        150                               155                               160                               165

aat ttg gtt ggg cag ctt ggg caa gca ggt gca aaa cac att tta gta 643
Asn Leu Val Gly Gln Leu Gly Gln Ala Gly Ala Lys His Ile Leu Val
                        170                               175                               180

ccc agt ctt cct gat gtt ggc gtc acg cca gaa tac gcc caa gat ccg 691
Pro Ser Leu Pro Asp Val Gly Val Thr Pro Glu Tyr Ala Gln Asp Pro
                        185                               190                               195

```

108/189

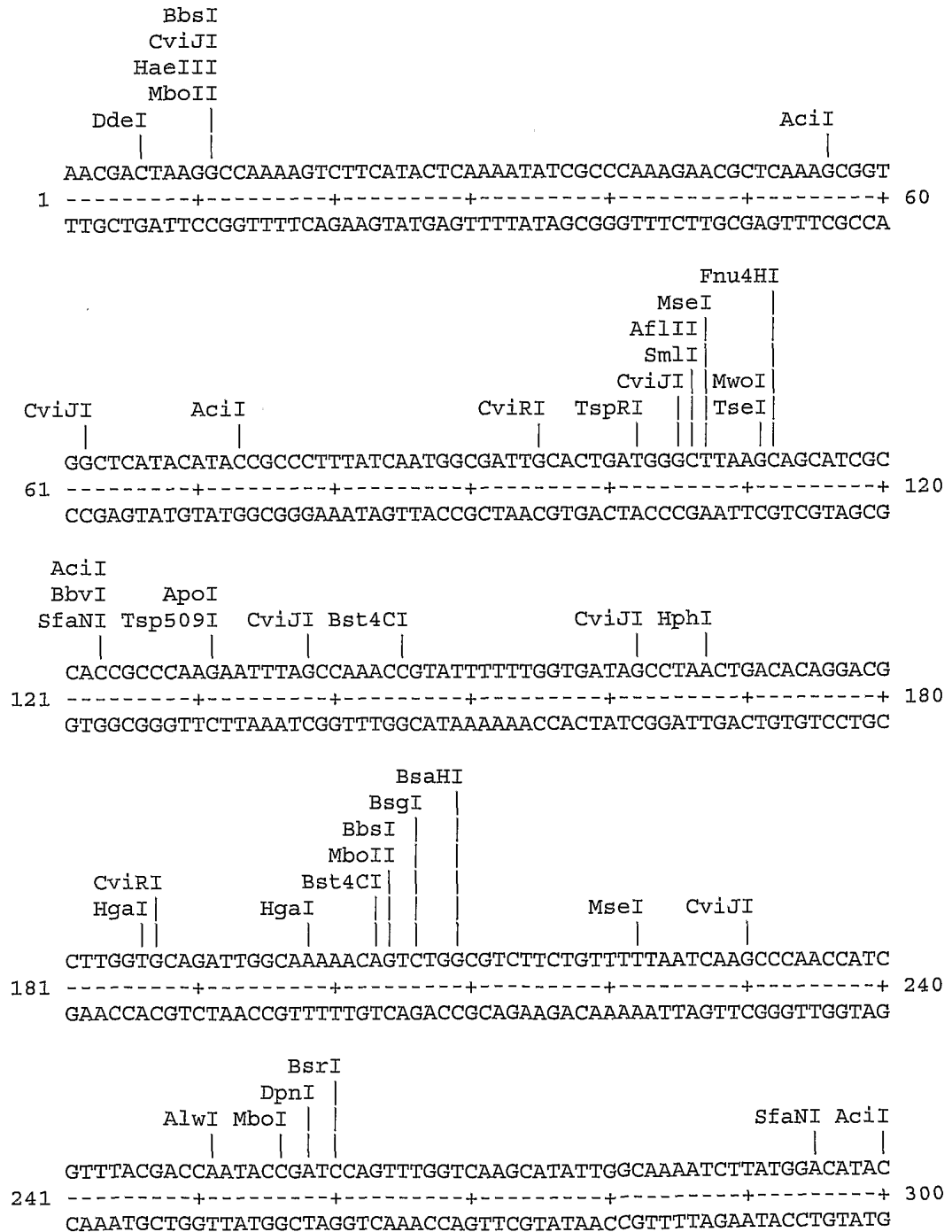
acc aaa tct gcc aca gca tca ttg tct gct cat atc tac aac caa act	739
Thr Lys Ser Ala Thr Ala Ser Leu Ser Ala His Ile Tyr Asn Gln Thr	
200 205 210	
tta tat caa agt tta aac aac caa acc acc aat gtc att gct gcc aat	787
Leu Tyr Gln Ser Leu Asn Asn Gln Thr Thr Asn Val Ile Ala Ala Asn	
215 220 225	
acc ttt gcc cta ctc aaa gag gcg gtt agt aat cca gca gga ttt ggg	835
Thr Phe Ala Leu Leu Lys Glu Ala Val Ser Asn Pro Ala Gly Phe Gly	
230 235 240 245	
ttt agt aat gtg acc gaa cct gct tgc caa aac cta gat gcc aat tta	883
Phe Ser Asn Val Thr Glu Pro Ala Cys Gln Asn Leu Asp Ala Asn Leu	
250 255 260	
tca tcg tta gca tgt aaa cca agc gac tgg caa caa acc gct gcg aat	931
Ser Ser Leu Ala Cys Lys Pro Ser Asp Trp Gln Gln Thr Ala Ala Asn	
265 270 275	
gct aat gaa act tat gcg ttt gct gat aaa att cac cct tca ggg cgt	979
Ala Asn Glu Thr Tyr Ala Phe Ala Asp Lys Ile His Pro Ser Gly Arg	
280 285 290	
acg cat cgc att ttg gca caa tat tat cgc tct att atc gat agc ccc	1027
Thr His Arg Ile Leu Ala Gln Tyr Tyr Arg Ser Ile Ile Asp Ser Pro	
295 300 305	
aac caa atg ggt caa cta ccc cag cat ctg att aaa cat ggt caa caa	1075
Asn Gln Met Gly Gln Leu Pro Gln His Leu Ile Lys His Gly Gln Gln	
310 315 320 325	
agc cag caa caa ttg gta cgc cgc tta gac agt ctt aac aat cat caa	1123
Ser Gln Gln Gln Leu Val Arg Arg Leu Asp Ser Leu Asn Asn His Gln	
330 335 340	
caa tcg cta tgg att gat ggc agc atc agc aat caa tcg ctt gat aca	1171
Gln Ser Leu Trp Ile Asp Gly Ser Ile Ser Asn Gln Ser Leu Asp Thr	
345 350 355	
cgc aat aaa ata gac cgc cct acc atc aaa ctt gga ctt gat att gca	1219
Arg Asn Lys Ile Asp Arg Pro Thr Ile Lys Leu Gly Leu Asp Ile Ala	
360 365 370	
cga cct aac cag cat aca ggg gca tat tta acc tac caa caa caa gat	1267
Arg Pro Asn Gln His Thr Gly Ala Tyr Leu Thr Tyr Gln Gln Gln Asp	
375 380 385	
tat cat cta agc gat tct ata acc gct gat gtc aaa caa aca ggc att	1315
Tyr His Leu Ser Asp Ser Ile Thr Ala Asp Val Lys Gln Thr Gly Ile	
390 395 400 405	
ggg ttg tat cat cgc cat gat ttt gat cac tta cgc atc agt gcc aac	1363
Gly Leu Tyr His Arg His Asp Phe Asp His Leu Arg Ile Ser Ala Asn	
410 415 420	

109/189

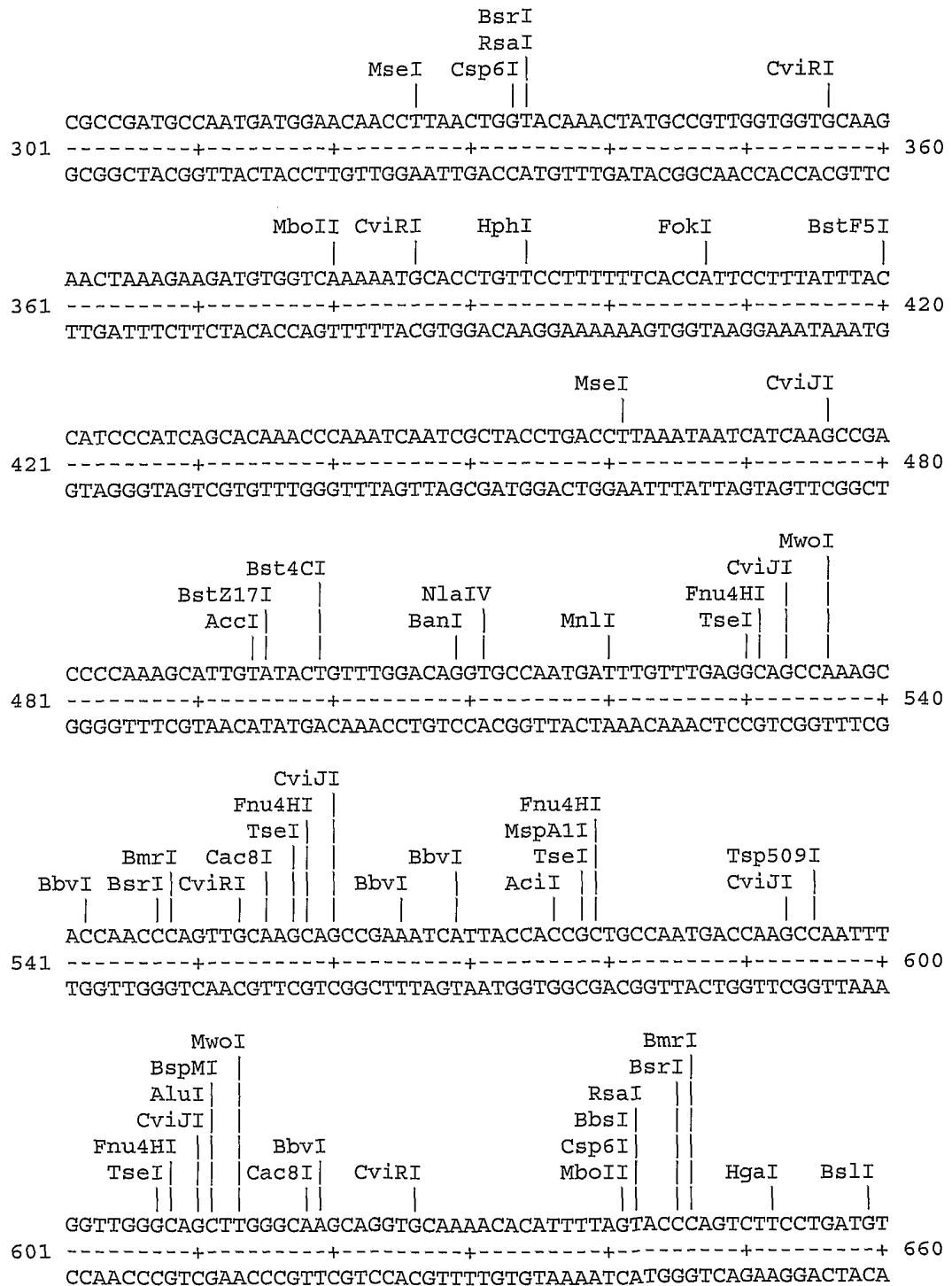
ctt ggg gtt gac cat tta agt aca caa agc tta cgc caa atc atg tgg	1411
Leu Gly Val Asp His Leu Ser Thr Gln Ser Leu Arg Gln Ile Met Trp	
425 430 435	
gac ggt gag aat cgc aac cat caa gca cat gcc aca ggt aag cgt cga	1459
Asp Gly Glu Asn Arg Asn His Gln Ala His Ala Thr Gly Lys Arg Arg	
440 445 450	
tat gca agt gta caa ggc agc tat gga ttt acc caa aac aat gtt acc	1507
Tyr Ala Ser Val Gln Gly Ser Tyr Gly Phe Thr Gln Asn Asn Val Thr	
455 460 465	
tat cgt cct tat gtc ggt att cat gct caa gat atc aag cag aat cgc	1555
Tyr Arg Pro Tyr Val Gly Ile His Ala Gln Asp Ile Lys Gln Asn Arg	
470 475 480 485	
ttc ttt gaa aat caa cca aat cta tca act gct cta agt ttt aag ctg	1603
Phe Phe Glu Asn Gln Pro Asn Leu Ser Thr Ala Leu Ser Phe Lys Leu	
490 495 500	
cct gac cat cag tct ttg caa gct gat att ggc gtt aat att gat tat	1651
Pro Asp His Gln Ser Leu Gln Ala Asp Ile Gly Val Asn Ile Asp Tyr	
505 510 515	
gcc atg aat gat aaa cta aac ctt ttg gca gga ttg ggt tat cag cat	1699
Ala Met Asn Asp Lys Leu Asn Leu Leu Ala Gly Leu Gly Tyr Gln His	
520 525 530	
gag ttt aaa gac aat aac aaa act gtt gag aca gct gta tta tcc aac	1747
Glu Phe Lys Asp Asn Asn Lys Thr Val Glu Thr Ala Val Leu Ser Asn	
535 540 545	
cgt gat tat cat cgc agc ttt gtc aca acc gtg ccc ttt gat aaa aag	1795
Arg Asp Tyr His Arg Ser Phe Val Thr Thr Val Pro Phe Asp Lys Lys	
550 555 560 565	
cac aca acg cac gca cat tta ggt gca aca ctt gct ttg ggt aat aac	1843
His Thr Thr His Ala His Leu Gly Ala Thr Leu Ala Leu Gly Asn Asn	
570 575 580	
act cat ttg aat gct ggc ata caa gcc aat cac caa gat cat gat aca	1891
Thr His Leu Asn Ala Gly Ile Gln Ala Asn His Gln Asp His Asp Thr	
585 590 595	
aaa gtt ggt gga ttt gtg ggt gca caa atg gca ttt taa ccaatcaatt	1940
Lys Val Gly Gly Phe Val Gly Ala Gln Met Ala Phe	
600 605 610	
tttttgtgac agccattcat ttaaataaag acgataatct ctcatcgtct ttattgttat	2000
ttttgtctgc ttgtttttat caatgcaatg tgtcagcgtt tatcattga	2049

110/189

Figure 40. Restriction enzyme analysis of the *M. catarrhalis* outer membrane esterase gene designated MCA101806



111/189



112/189

MaeIII DpnI
 Tsp45I BstYI
 BsaHI AlwI MboI SfaNI
 TGGCGTCACGCCAGAATACGCCAAGATCCGACCAAATCTGCCACAGCATCATTGTCTGC
 661 -----+-----+-----+-----+-----+-----+-----+ 720
 ACCGCAGTGCAGTCTTATGCGGGTTCTAGGCTGGTTTAGACGGTGTCGTAGTAACAGACG

DraI
 PmeI Fnu4HI
 MseI BbvI BsrDI TseI
 TCATATCTACAACCAAACCTTTATATCAAAGTTTAAACAACCAAACCAATGTCAATTGC
 721 -----+-----+-----+-----+-----+-----+-----+ 780
 AGTATAGATGTTGGTTTGAAATATAGTTTCAAATTTGTTGGTTTGGTGGTTACAGTAACG

MnlI AciI
 TGCCAATACCTTTGCCCTACTCAAAGAGGCGGTTAGTAATCCAGCAGGATTTGGGTTTAG
 781 -----+-----+-----+-----+-----+-----+-----+ 840
 ACGGTTATGGAAACGGGATGAGTTTCTCCGCCAATCATTAGGTCGTCCTAAACCCAAATC

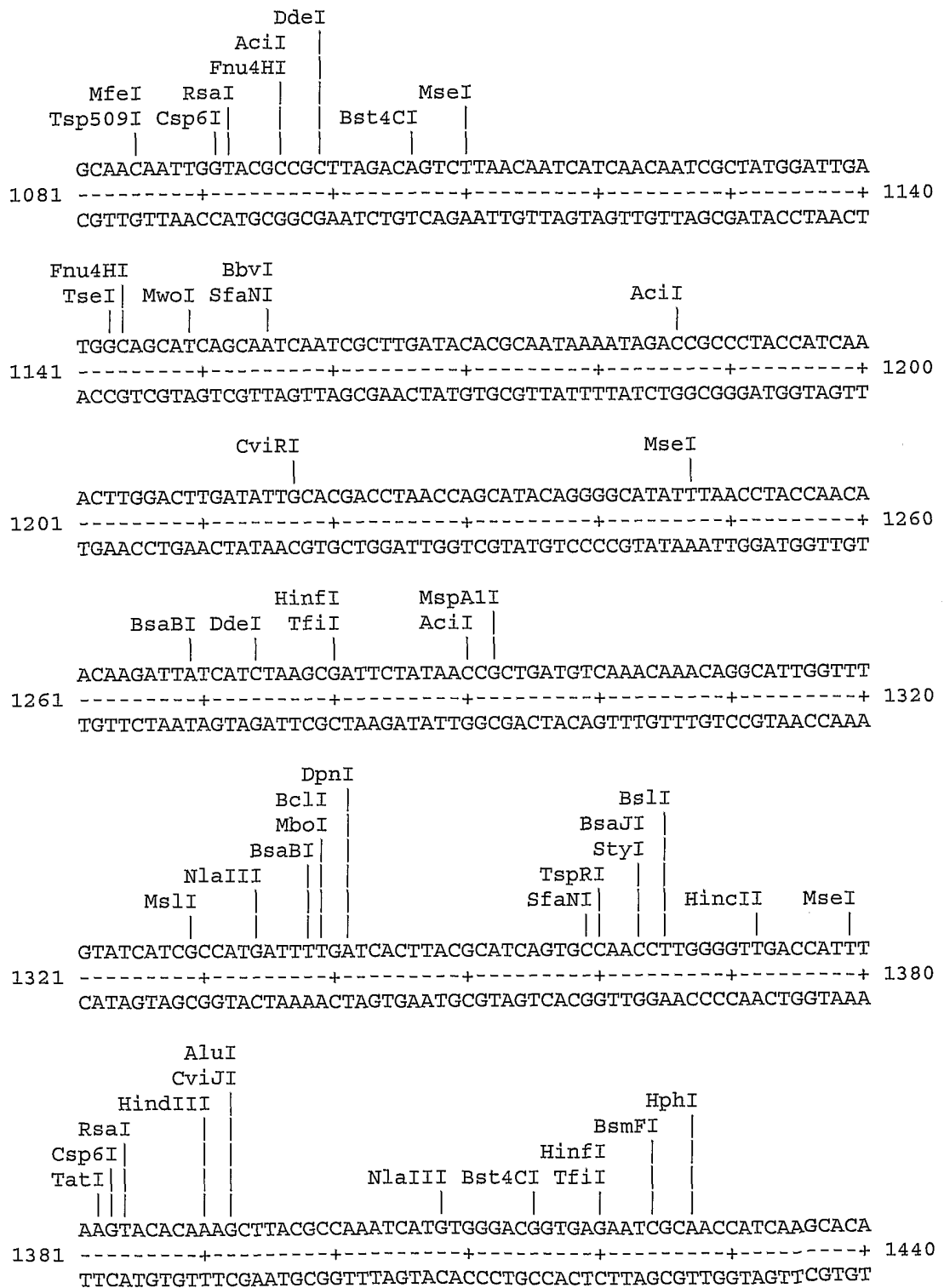
BspMI
 MaeIII SfaNI NlaIII
 Tsp45I Cac8I BfaI Tsp509I NspI
 TAATGTGACCGAACCTTGCTTGCCAAAACCTAGATGCCAATTTATCATCGTTAGCATGTAA
 841 -----+-----+-----+-----+-----+-----+-----+ 900
 ATTACACTGGCTTGGACGAACGGTTTTGGATCTACGGTTAAATAGTAGCAATCGTACATT

Fnu4HI ApoI
 MspAII Tsp509I
 BbvI TseI HphI
 BstXI BsrI AciI MwoI BsmI Eco57I
 ACCAAGCGACTGGCAACAAACCGCTGCGAATGCTAATGAACTTATGCGTTTGCTGATAA
 901 -----+-----+-----+-----+-----+-----+-----+ 960
 TGGTTCGCTGACCGTTGTTTGGCGACGCTTACGATTACTTTGAATACGCAAACGACTATT

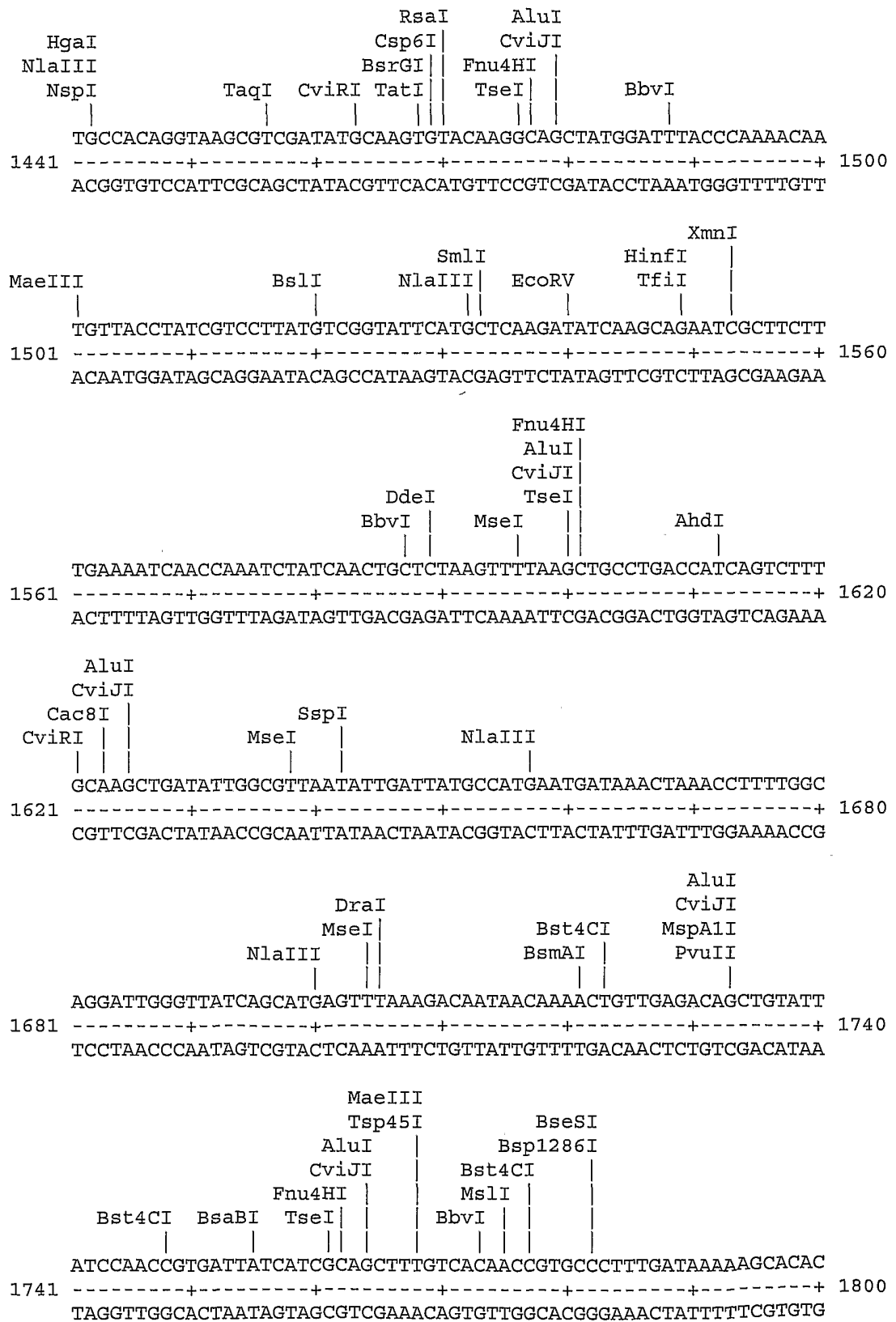
RsaI
 Csp6I
 BsiWI SfaNI SspI ClaI
 AATTCACCTTCAGGGCGTACGCATCGCATTTTGGCACAATATTATCGCTCTATTATCGA
 961 -----+-----+-----+-----+-----+-----+-----+ 1020
 TTAAGTGGGAAGTCCCGCATGCGTAGCGTAAAACCGTGTATAATAGCGAGATAATAGCT

SfaNI HincII Cac8I
 CviJI BstXI HincII AlwNI MseI NlaIII CviJI
 TAGCCCCAACCAAATGGGTCAACTACCCAGCATCTGATTAAACATGGTCAACAAAGCCA
 1021 -----+-----+-----+-----+-----+-----+-----+ 1080
 ATCGGGGTTGGTTTACCCAGTTGATGGGGTCGTAGACTAATTTGTACCAGTTGTTTCGGT

113/189



114/189



115/189

BstAPI
 Cac8I MwoI MslI CviRI MwoI MslI BsmI
 AACGCACGCACATTTAGGTGCAACACTTGCTTTGGGTAATAACACTCATTTGAATGCTGG
 1801 -----+-----+-----+-----+-----+-----+-----+ 1860
 TTGCGTGCGTGTAATCCACGTTGTGAACGAAACCCATTATTGTGAGTAACTTACGACC

BseSI
 BsiHKAI
 Bsp1286I
 CviRI
 NlaIII
 BspHI
 DpnI
 CviJI
 HphI
 MboI
 BaeI
 MwoI
 CATACAAGCCAATCACCAAGATCATGATACAAAAGTTGGTGGATTTGTGGGTGCACAAAT
 1861 -----+-----+-----+-----+-----+-----+-----+ 1920
 GTATGTTTCGGTTAGTGGTTCTAGTACTATGTTTTCAACCACCTAAACACCCACGTGTTTA

DraI
 SwaI
 MaeIII
 Tsp45I
 CviJI
 MseI
 Tsp509I
 MseI
 GGCATTTTAACCAATCAATTTTTTTGTGACAGCCATTCATTTAAATAAAGACGATAATCT
 1921 -----+-----+-----+-----+-----+-----+-----+ 1980
 CCGTAAAAATTGGTTAGTTAAAAAACACTGTTCGGTAAGTAAATTTATTTCTGCTATTAGA

CviRI BsrDI
 CTCATCGTCTTTATTGTTATTTTTGTCTGCTTGTTTTTATCAATGCAATGTGTCAGCGTT
 1981 -----+-----+-----+-----+-----+-----+-----+ 2040
 GAGTAGCAGAAATAACAATAAAAAACAGACGAACAAAAATAGTTACGTTACACAGTCGCAA

TATCATTGA
 2041 ----- 2049
 ATAGTAACT

116/189

Figure 41. Sequence of *M. catarrhalis* pyridine nucleotide transhydrogenase gene designated MCA101862

```

tctcaattta gcttgatcgc ttaatggact taccacaaaa acgcactcgc cagacattta 60
ggatcgttgc aaccaacaaa accatgatta aggattatatt atg tca aaa aaa ccc 115
                                         Met Ser Lys Lys Pro
                                         1           5

aaa acc acc caa gat acc gac cca gtc gcc ttg att gaa acc gct cca 163
Lys Thr Thr Gln Asp Thr Asp Pro Val Ala Leu Ile Glu Thr Ala Pro
                        10                      15                      20

caa aca aaa aaa tct caa act ttt gag cat gta gag aac gat caa acc 211
Gln Thr Lys Lys Ser Gln Thr Phe Glu His Val Glu Asn Asp Gln Thr
                        25                      30                      35

cat cct agt ctc atc agt cct tta gat ggt cag cgt gtg ggt ttg gat 259
His Pro Ser Leu Ile Ser Pro Leu Asp Gly Gln Arg Val Gly Leu Asp
                        40                      45                      50

cgc tcc aat aaa aac aaa ata aag cac aat tat gaa tat gat gcc atc 307
Arg Ser Asn Lys Asn Lys Ile Lys His Asn Tyr Glu Tyr Asp Ala Ile
                        55                      60                      65

gtc att ggt gca ggc cct gct ggt gaa gcg gct gca atg aaa ctt gct 355
Val Ile Gly Ala Gly Pro Ala Gly Glu Ala Ala Ala Met Lys Leu Ala
                        70                      75                      80                      85

aag tct ggg cat cgt gtt gcc gtc atc gat ccg cgt gac caa gtc ggt 403
Lys Ser Gly His Arg Val Ala Val Ile Asp Pro Arg Asp Gln Val Gly
                        90                      95                      100

ggt aat tgt gca cat gtt ggt act att cca agt aag gca ctg cgt cag 451
Gly Asn Cys Ala His Val Gly Thr Ile Pro Ser Lys Ala Leu Arg Gln
                        105                      110                      115

tct gtt ttt aat att att aac ttt cgc cga gat cct ttg ttt aaa caa 499
Ser Val Phe Asn Ile Ile Asn Phe Arg Arg Asp Pro Leu Phe Lys Gln
                        120                      125                      130

aat aaa gat agc tat caa gtt cca cta agt aaa gtc tta agt cgt gcg 547
Asn Lys Asp Ser Tyr Gln Val Pro Leu Ser Lys Val Leu Ser Arg Ala
                        135                      140                      145

cgt gaa gtt atc cgt cat cag gtc aat acg cat aag ctt ttt tat gag 595
Arg Glu Val Ile Arg His Gln Val Asn Thr His Lys Leu Phe Tyr Glu
                        150                      155                      160                      165

cgt aac caa att caa gtc att cat ggt tgg ggc agt ttt att gat gca 643
Arg Asn Gln Ile Gln Val Ile His Gly Trp Gly Ser Phe Ile Asp Ala
                        170                      175                      180

cat acg cta gag gtg tac atc aat gat cag gag cct acc aaa acc att 691
His Thr Leu Glu Val Tyr Ile Asn Asp Gln Glu Pro Thr Lys Thr Ile
                        185                      190                      195

```

117/189

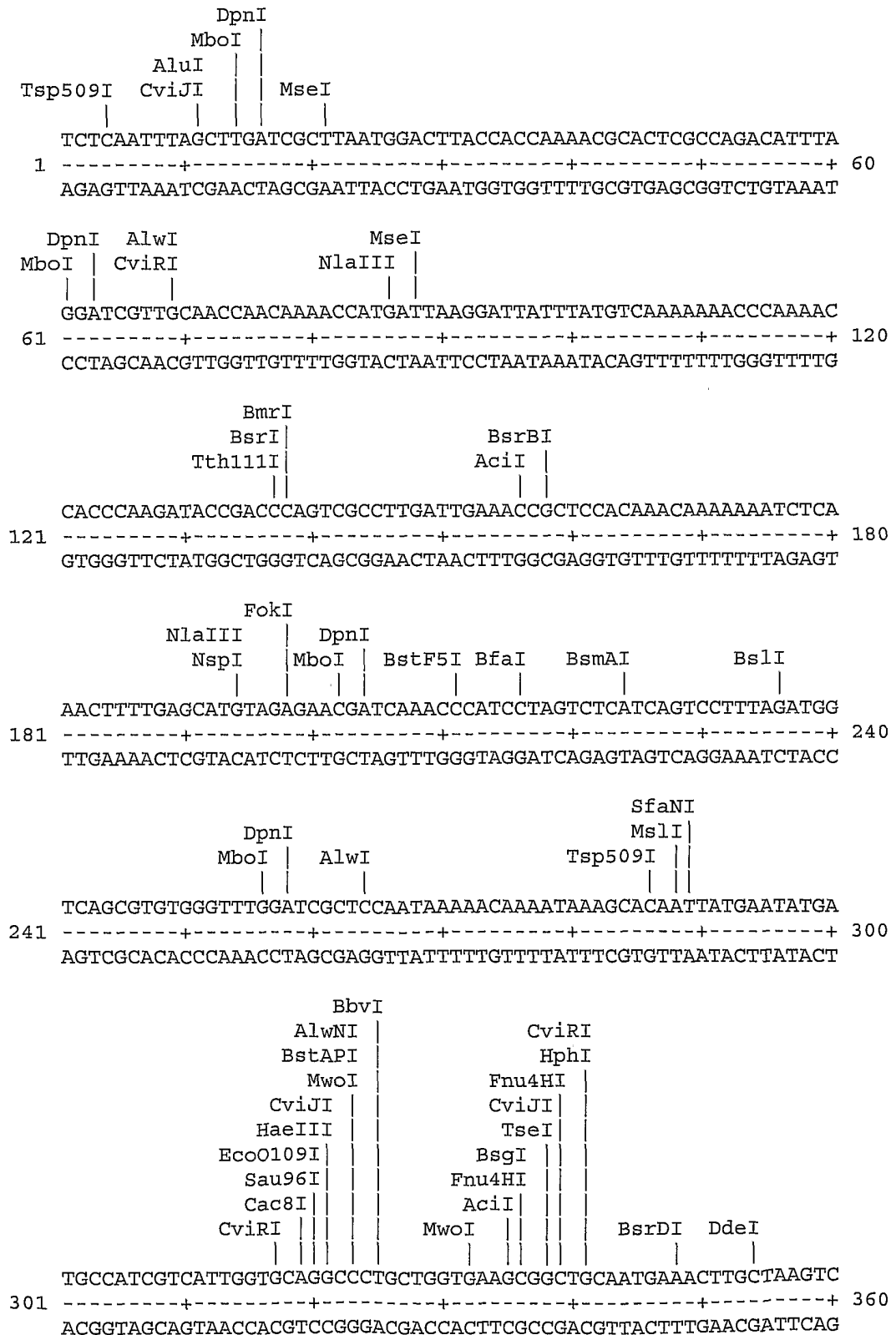
acc ttt aat cag gca atc att gcg gtt ggt tct cgt cct tat cgc cct	739
Thr Phe Asn Gln Ala Ile Ile Ala Val Gly Ser Arg Pro Tyr Arg Pro	
200 205 210	
gag att tta gat ttc gat cat cca cgc gtc ttt gac tct gat aaa atc	787
Glu Ile Leu Asp Phe Asp His Pro Arg Val Phe Asp Ser Asp Lys Ile	
215 220 225	
cta caa atg gat tat att gcc cgc aaa atc atc atc tat ggt gct ggt	835
Leu Gln Met Asp Tyr Ile Ala Arg Lys Ile Ile Ile Tyr Gly Ala Gly	
230 235 240 245	
gtc att ggc tgt gag tat gcc tct att ttt aca ggt tta ggc tat gtg	883
Val Ile Gly Cys Glu Tyr Ala Ser Ile Phe Thr Gly Leu Gly Tyr Val	
250 255 260	
gtt gat ttg att aat aat aaa gat cag ctt ttg agc tac tta gat aaa	931
Val Asp Leu Ile Asn Asn Lys Asp Gln Leu Leu Ser Tyr Leu Asp Lys	
265 270 275	
gaa att tct gat gca cta agc cat gac ttc cgt caa ttt ggc gta cgc	979
Glu Ile Ser Asp Ala Leu Ser His Asp Phe Arg Gln Phe Gly Val Arg	
280 285 290	
att cgt aac cac gaa gaa atc gat cac ctt gaa acg cat gat gac tgc	1027
Ile Arg Asn His Glu Glu Ile Asp His Leu Glu Thr His Asp Asp Cys	
295 300 305	
gtg gta tta cat cta aaa agc ggt aaa aaa atc aaa gcc gac gcg atc	1075
Val Val Leu His Leu Lys Ser Gly Lys Lys Ile Lys Ala Asp Ala Ile	
310 315 320 325	
tta tgg tgt aat ggc cgc tcg ggt aat acc gat agc cta aac cta gaa	1123
Leu Trp Cys Asn Gly Arg Ser Gly Asn Thr Asp Ser Leu Asn Leu Glu	
330 335 340	
gcg gtt ggg ctt acc gca aat agc cgt ggt cag ctt gaa gtg gac aaa	1171
Ala Val Gly Leu Thr Ala Asn Ser Arg Gly Gln Leu Glu Val Asp Lys	
345 350 355	
acc tac cgt acc agc ata cca cat att tat gct gtg ggc gat gtc atc	1219
Thr Tyr Arg Thr Ser Ile Pro His Ile Tyr Ala Val Gly Asp Val Ile	
360 365 370	
ggc tgg cca tca ttg gct tcg gca gcc tac gac caa ggg cgt aat gca	1267
Gly Trp Pro Ser Leu Ala Ser Ala Ala Tyr Asp Gln Gly Arg Asn Ala	
375 380 385	
gca ggg ttt atg gta ggt gat gaa gat gct gag ttt gtc aat agc gtc	1315
Ala Gly Phe Met Val Gly Asp Glu Asp Ala Glu Phe Val Asn Ser Val	
390 395 400 405	
cca act ggc att tat acc att cct gaa atc tca agc att ggt aaa act	1363
Pro Thr Gly Ile Tyr Thr Ile Pro Glu Ile Ser Ser Ile Gly Lys Thr	
410 415 420	

118/189

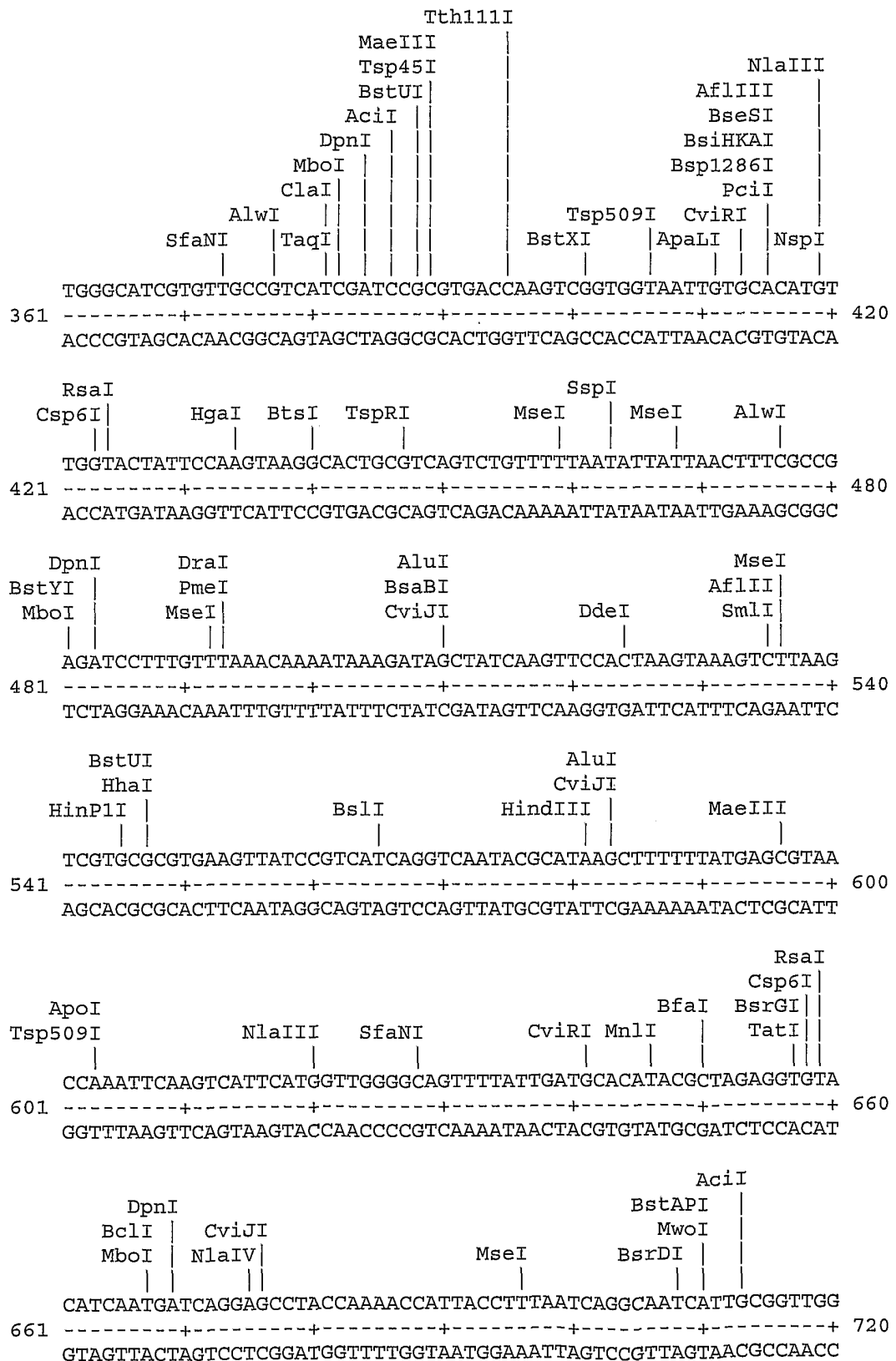
gaa gct gag ctg acc gcc gaa aaa aca cct tat gaa gtt ggg cag gct	1411
Glu Ala Glu Leu Thr Ala Glu Lys Thr Pro Tyr Glu Val Gly Gln Ala	
425 430 435	
ttt ttt aag cac ttg gca cgc tca cag att att ggc gag cgt tca ggt	1459
Phe Phe Lys His Leu Ala Arg Ser Gln Ile Ile Gly Glu Arg Ser Gly	
440 445 450	
gtc tta aaa atc cta ttt cac cgt gaa aca ctt gag ctt ctt ggt atc	1507
Val Leu Lys Ile Leu Phe His Arg Glu Thr Leu Glu Leu Leu Gly Ile	
455 460 465	
cat tgc tat ggt aac cat gct tct gaa att att cat att ggg cag gca	1555
His Cys Tyr Gly Asn His Ala Ser Glu Ile Ile His Ile Gly Gln Ala	
470 475 480 485	
gtc atg aaa tgt ggt cat acg ctt gaa tat ttt atc aat acc aca ttt	1603
Val Met Lys Cys Gly His Thr Leu Glu Tyr Phe Ile Asn Thr Thr Phe	
490 495 500	
aat tat cca acc atg gca gag gcc tat cgt gtt gct gcc tta aat ggg	1651
Asn Tyr Pro Thr Met Ala Glu Ala Tyr Arg Val Ala Ala Leu Asn Gly	
505 510 515	
atg aac cga att ttt tag ttggcatggt ccataccaat taaactaaac	1699
Met Asn Arg Ile Phe	
520	
attaatcaaaa ccaaaaatga ctgatgaata catcagctgt ttttggtttt aatcactttt	1759
tatttgattg	1769

119/189

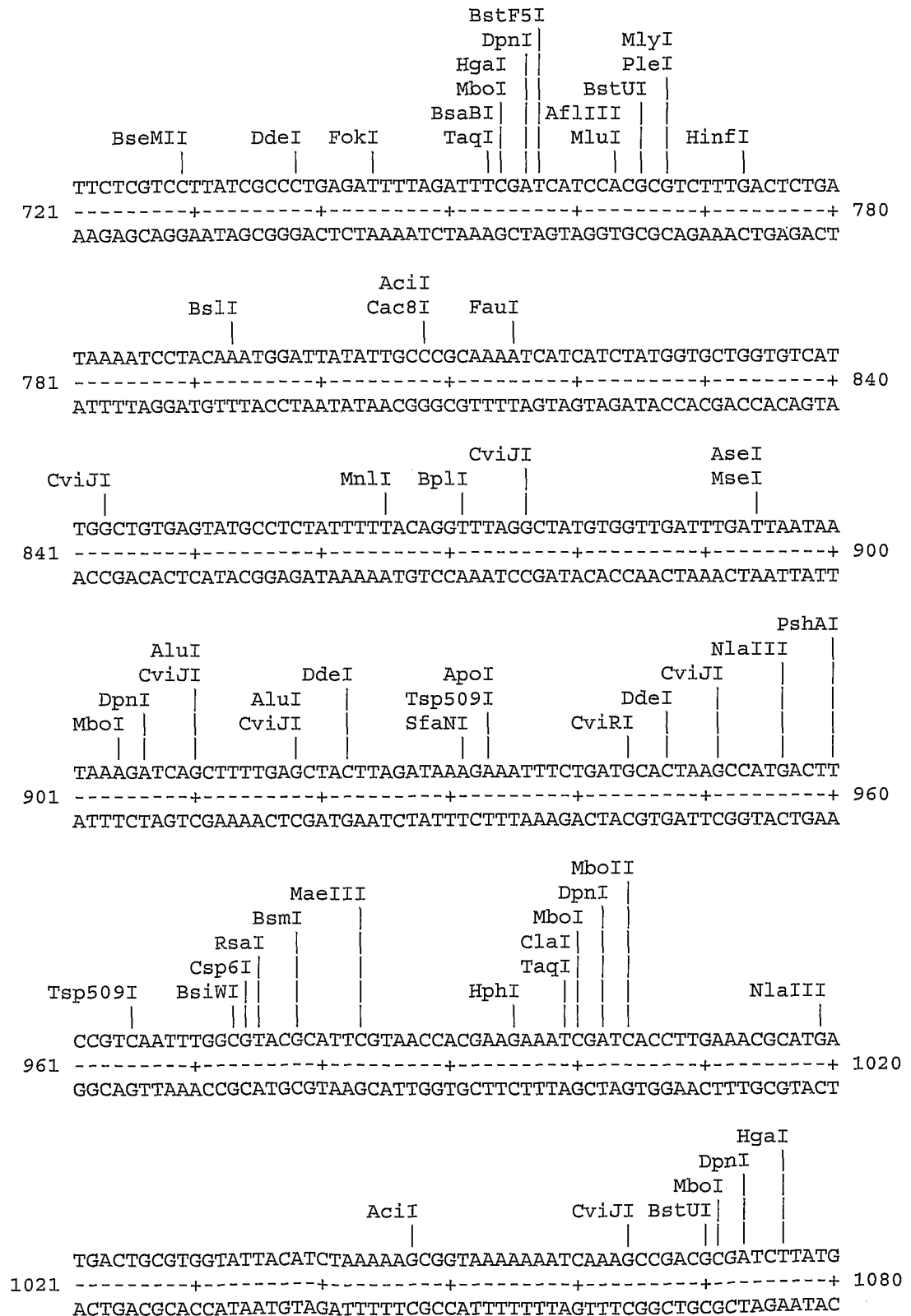
Figure 42. Restriction enzyme analysis of the *M. catarrhalis* pyridine nucleotide transhydrogenase gene designated MCA101862



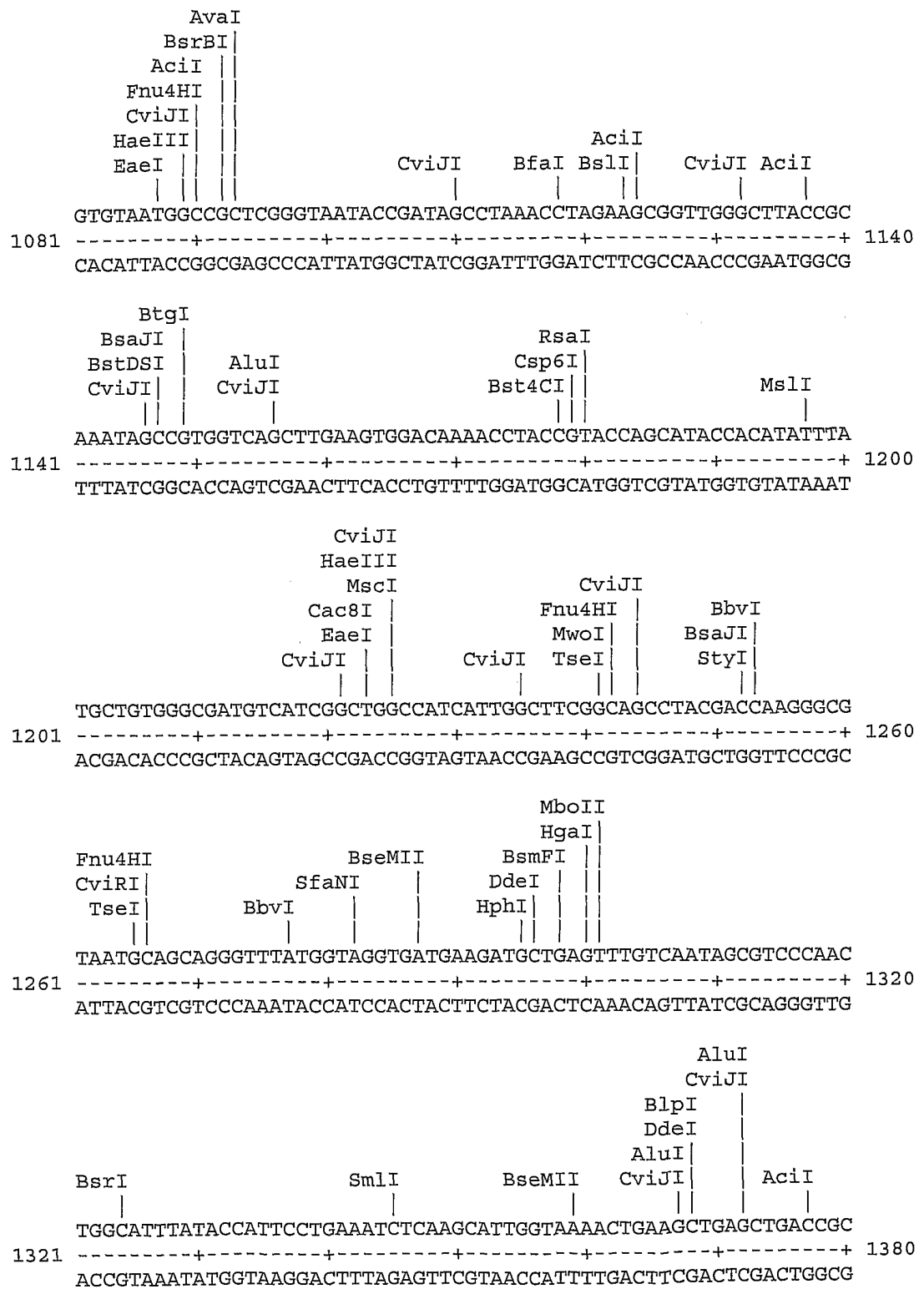
120/189



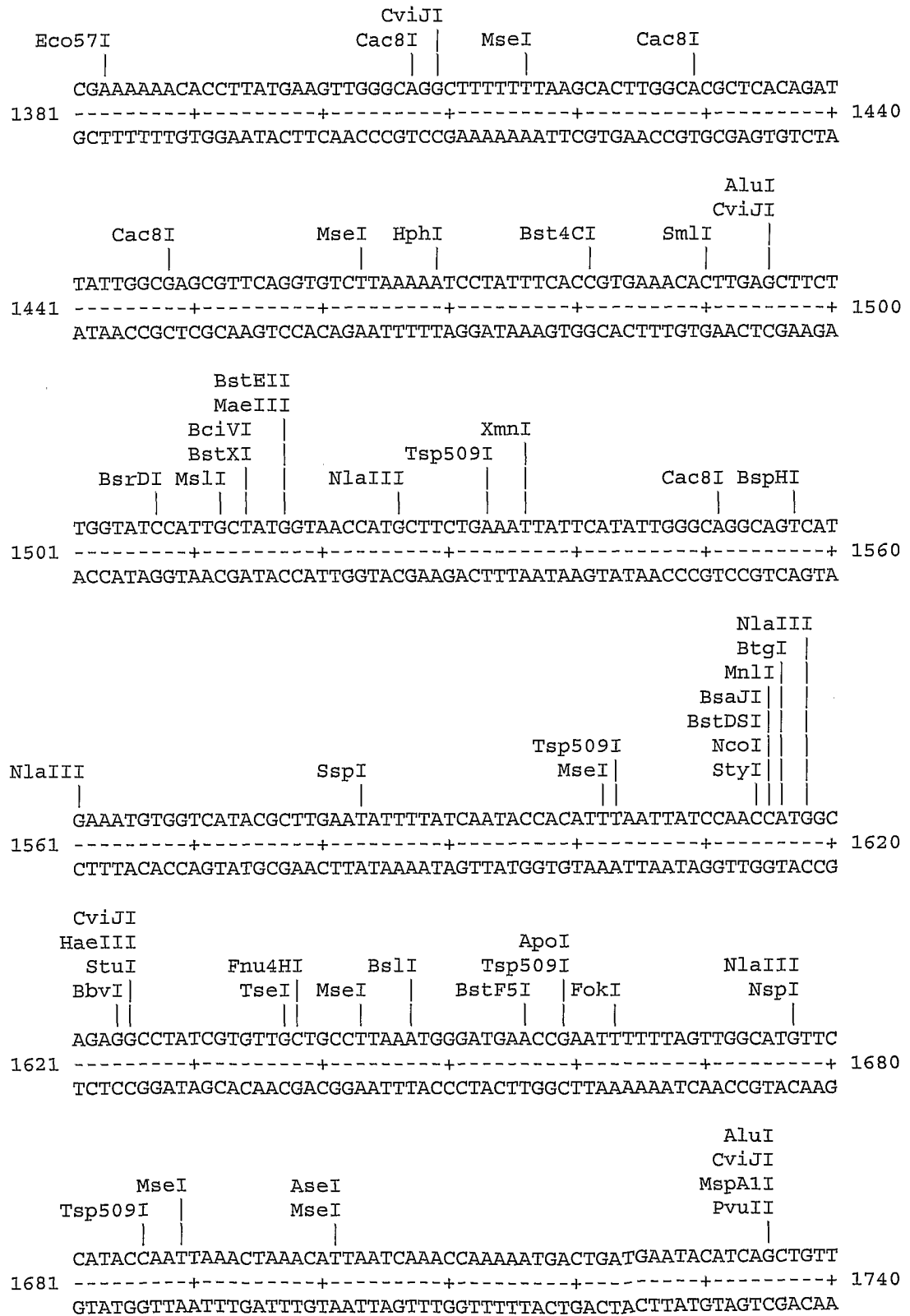
121/189



122/189



123/189



124/189

MseI
|
TTTGGTTTAAATCACTTTTATTGATTG
1741 -----+-----+----- 1769
AAACC AAAATTAGTGAAAAATAAACTAAC

125/189

Figure 43. Sequence of *M. catarrhalis* gene encoding a lipoprotein shown in SEQ ID NOS: 43 & 44 designated MCA101896

```

tgctttaata tcttttatgc tttaatagat aagtttttat agcttttttc aataacttaag 60

cgatttttga acatataaatt ttatttttgta aggtcatttt atg aat ttt ggt aaa 115
                                   Met Asn Phe Gly Lys
                                   1 5

att ttt ggt att tgt gca ctg gca tct ggc atc gca ttg gca ggc tgc 163
Ile Phe Gly Ile Cys Ala Leu Ala Ser Gly Ile Ala Leu Ala Gly Cys
                                   10 15 20

agc aat caa tca aac gaa cca gct gcc ata tct gat aca gct gca cag 211
Ser Asn Gln Ser Asn Glu Pro Ala Ala Ile Ser Asp Thr Ala Ala Gln
                                   25 30 35

act atc aag gtt ggc gtc atg gca ggt cct gag caa gct gtg gca gag 259
Thr Ile Lys Val Gly Val Met Ala Gly Pro Glu Gln Ala Val Ala Glu
                                   40 45 50

gta gca ggt caa gtc gcc aaa gaa aaa tac aac ctg acc gtt gaa ttg 307
Val Ala Gly Gln Val Ala Lys Glu Lys Tyr Asn Leu Thr Val Glu Leu
                                   55 60 65

gtt gag ttt aat gac tat gcc atg cca aac tca gcc gtc tca aaa ggt 355
Val Glu Phe Asn Asp Tyr Ala Met Pro Asn Ser Ala Val Ser Lys Gly
                                   70 75 80 85

gaa ctt gac gcc aat gcc atg cag cac aaa ccc tat ctt gaa aaa gac 403
Glu Leu Asp Ala Asn Ala Met Gln His Lys Pro Tyr Leu Glu Lys Asp
                                   90 95 100

agc caa gaa aaa ggc cta aat aac ttg gtc atc gtc ggc aac acc ttt 451
Ser Gln Glu Lys Gly Leu Asn Asn Leu Val Ile Val Gly Asn Thr Phe
                                   105 110 115

gta tac cca ttg gca ggt tat tca acc aaa atc aag aca tta aat gag 499
Val Tyr Pro Leu Ala Gly Tyr Ser Thr Lys Ile Lys Thr Leu Asn Glu
                                   120 125 130

cta aaa gat ggt gca acc atc gcc gtt cca aat gat ccc tca aac tta 547
Leu Lys Asp Gly Ala Thr Ile Ala Val Pro Asn Asp Pro Ser Asn Leu
                                   135 140 145

gct cgt gca tta att tta ctt gaa aaa caa ggc tta att aaa tta aaa 595
Ala Arg Ala Leu Ile Leu Leu Glu Lys Gln Gly Leu Ile Lys Leu Lys
                                   150 155 160 165

gac aac acc aac cta ttc tca acc aca ctt gat atc gta gaa aat cca 643
Asp Asn Thr Asn Leu Phe Ser Thr Thr Leu Asp Ile Val Glu Asn Pro
                                   170 175 180

aaa aaa ttg gtc atc aaa gaa gtg gat acc tca gtt gct gct cgt gca 691
Lys Lys Leu Val Ile Lys Glu Val Asp Thr Ser Val Ala Ala Arg Ala
                                   185 190 195

```

126/189

```

att gac gat gtg gac ttg gca gtg gta aat aac aac tat gca ggt caa 739
Ile Asp Asp Val Asp Leu Ala Val Val Asn Asn Asn Tyr Ala Gly Gln
      200                      205                      210

gta ggt tta aca gcc agt gaa aat ggc gtt ttt gtt gaa gat aaa gac 787
Val Gly Leu Thr Ala Ser Glu Asn Gly Val Phe Val Glu Asp Lys Asp
      215                      220                      225

tcg cct tat gtc aat atc atc gtc gct cgt gct gac aat aaa gac tct 835
Ser Pro Tyr Val Asn Ile Ile Val Ala Arg Ala Asp Asn Lys Asp Ser
      230                      235                      240                      245

aag gcc atc caa gac ttt gtg aaa gcc tat caa acc gat gaa gtg gaa 883
Lys Ala Ile Gln Asp Phe Val Lys Ala Tyr Gln Thr Asp Glu Val Glu
      250                      255                      260

gct gaa gcc aaa aag caa ttt aaa gat ggt gtt atc aaa ggc tgg taa 931
Ala Glu Ala Lys Lys Gln Phe Lys Asp Gly Val Ile Lys Gly Trp
      265                      270                      275

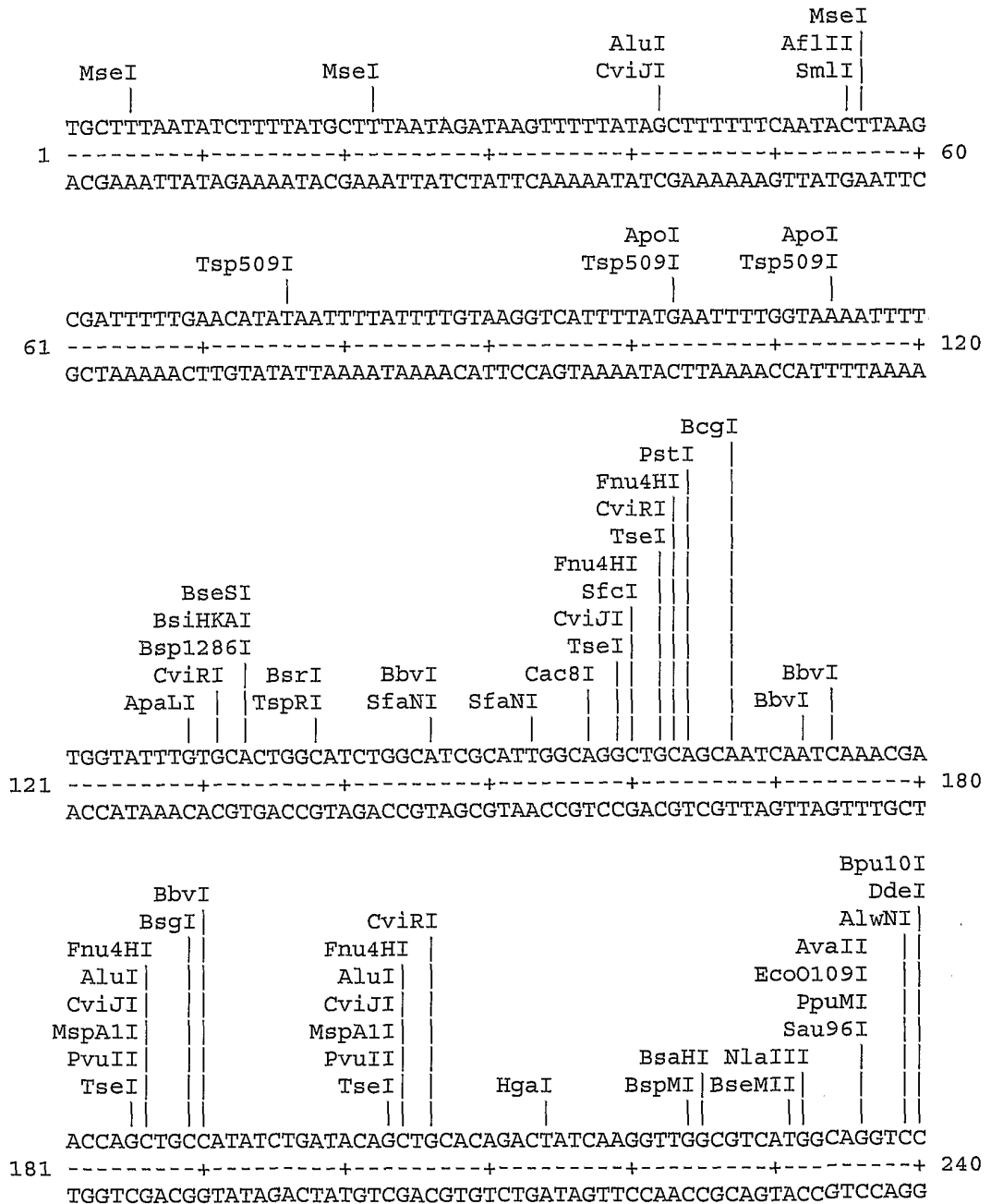
tctgaccgat ctgaccataa aaatcctcca accgttggag gattttttatt gaaccatcaa 991

agccatactt taacgcttag tatcttgaac tgcattacgc tcatcaattt 1041

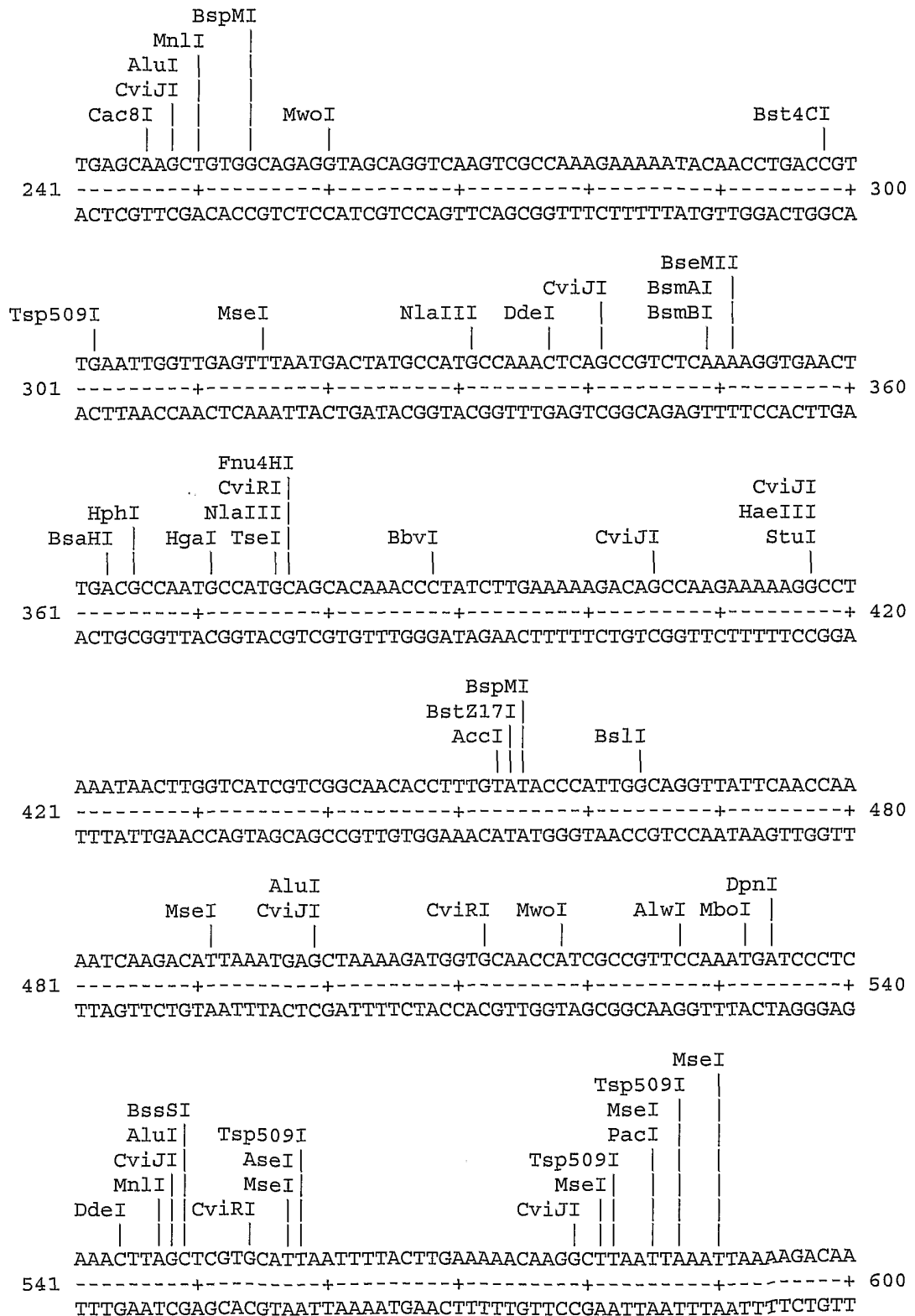
```

127/189

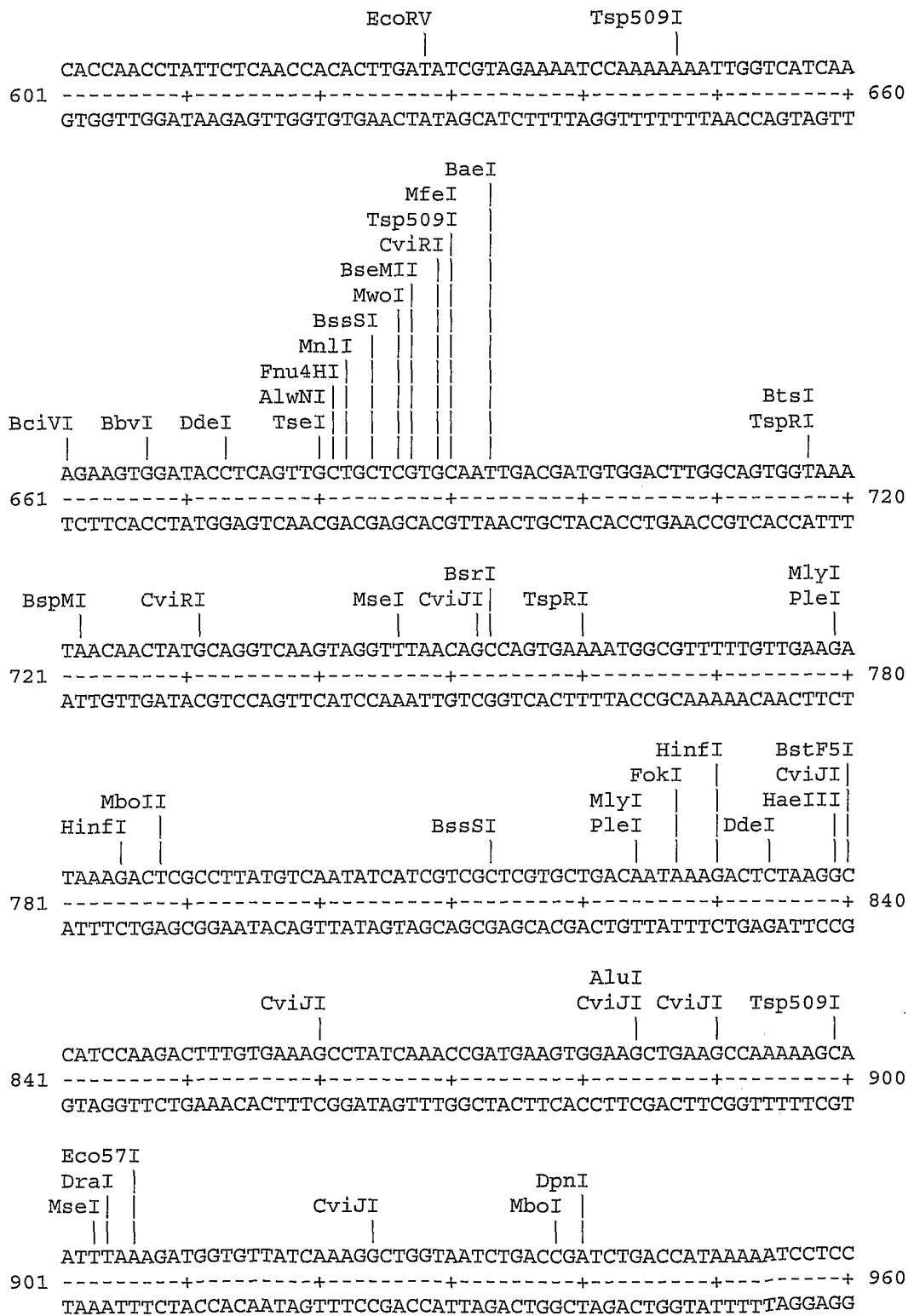
Figure 44. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a lipoprotein shown in SEQ ID NO: 43 designated MCA101896



128/189



129/189



•

```

      MnlI
      BslI
      PflMI
      Bst4CI
      MnlI
      CviJI
      MseI
      DdeI
      AACCGTTGGAGGATTTTTATTGAACCATCAAAGCCATACTTTAACGCTTAGTATCTTGAA
961  -----+-----+-----+-----+-----+-----+-----+
      TTGGCAACCTCCTAAAAATAACTTGGTAGTTTCGGTATGAAATTGCGAATCATAGAACTT

      CviRI
      Tsp509I
      CTGCATTACGCTCATCAATTT
1021 -----+-----+-----+-----+-----+-----+
      GACGTAATGCGAGTAGTTAAA

```

131/189

Figure 45. Sequence of *M. catarrhalis* HecB gene designated MCA101915

```

gtaagtataa atatcttctg atatgatgac aacatctttc tataatgtta caaaaatatt 60

gcttcctatg tcactcataa gtcaaggaca ggacaaaccc atg cgt ttt ata cct 115
Met Arg Phe Ile Pro
1 5

ttt gct gta act ctt tta tct ttt act ggc tcc tcc atc ttt tct gtg 163
Phe Ala Val Thr Leu Leu Ser Phe Thr Gly Ser Ser Ile Phe Ser Val
10 15 20

gcc aaa gct cag att gca cca gac cca acc aaa gca aat ctt tct gca 211
Ala Lys Ala Gln Ile Ala Pro Asp Pro Thr Lys Ala Asn Leu Ser Ala
25 30 35

ata acc gat gat cgt cgt cag gcg gct ctt gca cat tta gca aga caa 259
Ile Thr Asp Asp Arg Arg Gln Ala Ala Leu Ala His Leu Ala Arg Gln
40 45 50

gat ttg gca gca aca gat gac gac atg ggt gtc gat atc aat caa acc 307
Asp Leu Ala Ala Thr Asp Asp Asp Met Gly Val Asp Ile Asn Gln Thr
55 60 65

aat atc aac cag cca aat tgg caa gat gac aag aat ggc gat tca tgt 355
Asn Ile Asn Gln Pro Asn Trp Gln Asp Asp Lys Asn Gly Asp Ser Cys
70 75 80 85

ttt ttg gtt gat gag att gac ttt gtg att gag ggt gga gaa gat aaa 403
Phe Leu Val Asp Glu Ile Asp Phe Val Ile Glu Gly Gly Glu Asp Lys
90 95 100

cat att cta ggt atg acg ccg agc aat tta cgc agc ttg cta tcg ccc 451
His Ile Leu Gly Met Thr Pro Ser Asn Leu Arg Ser Leu Leu Ser Pro
105 110 115

ctg ttg gac cgg aca aag cca aca tat gct ctt aat cac tgc atc aac 499
Leu Leu Asp Arg Thr Lys Pro Thr Tyr Ala Leu Asn His Cys Ile Asn
120 125 130

aat cac aac cta tct tta att gta gat att gcc cat aat gaa tta cta 547
Asn His Asn Leu Ser Leu Ile Val Asp Ile Ala His Asn Glu Leu Leu
135 140 145

aag cgt ggt tat tta act agc agt att agc att gaa gag cag gat tta 595
Lys Arg Gly Tyr Leu Thr Ser Ser Ile Ser Ile Glu Glu Gln Asp Leu
150 155 160 165

tca act aaa aaa ctc act tta aca gtc cat gct ggc aag gtg aaa aag 643
Ser Thr Lys Lys Leu Thr Leu Thr Val His Ala Gly Lys Val Lys Lys
170 175 180

gtg ata ttg gcc gat gct agc aaa acg cca act tat gtc aaa gcc gcc 691
Val Ile Leu Ala Asp Ala Ser Lys Thr Pro Thr Tyr Val Lys Ala Ala
185 190 195

```

132/189

att cct ctt aag ttc aat caa gcc ttt cgt ctg tca tat tta gaa caa	739
Ile Pro Leu Lys Phe Asn Gln Ala Phe Arg Leu Ser Tyr Leu Glu Gln	
200 205 210	
ggg cta gat aat tta aaa cgc atc gat cca acg gca acc gtg cag att	787
Gly Leu Asp Asn Leu Lys Arg Ile Asp Pro Thr Ala Thr Val Gln Ile	
215 220 225	
ata cca agc agc tcc agt aat gtc gcc aat gat att gac att aac acc	835
Ile Pro Ser Ser Ser Ser Asn Val Ala Asn Asp Ile Asp Ile Asn Thr	
230 235 240 245	
aca aac ctt ggt ttt agt gat tta ctc att agg atg gac cgc agt caa	883
Thr Asn Leu Gly Phe Ser Asp Leu Leu Ile Arg Met Asp Arg Ser Gln	
250 255 260	
agc gct gta ttt ggc att aat att gat aac agt tta tcc aaa gat tat	931
Ser Ala Val Phe Gly Ile Asn Ile Asp Asn Ser Leu Ser Lys Asp Tyr	
265 270 275	
ggc aat tat ctg att tct gct aat gtt cgt gct aac aac ctc atg cat	979
Gly Asn Tyr Leu Ile Ser Ala Asn Val Arg Ala Asn Asn Leu Met His	
280 285 290	
ctt aac gat gaa tgg aat ctt tct gcc aac tat cca ttg gcg cgc ctg	1027
Leu Asn Asp Glu Trp Asn Leu Ser Ala Asn Tyr Pro Leu Ala Arg Leu	
295 300 305	
att gac gct gcc caa aat gat ctt ggg gtt cag gtg ggc aag gat agg	1075
Ile Asp Ala Ala Gln Asn Asp Leu Gly Val Gln Val Gly Lys Asp Arg	
310 315 320 325	
cag gtt aat tat cac gct tcc ttg act att cct tat gga ttg tat aaa	1123
Gln Val Asn Tyr His Ala Ser Leu Thr Ile Pro Tyr Gly Leu Tyr Lys	
330 335 340	
ttt tct gcc aca cac agc cac cac caa tac cag caa ttt ctt gaa ggg	1171
Phe Ser Ala Thr His Ser His His Gln Tyr Gln Gln Phe Leu Glu Gly	
345 350 355	
ctt cat gcg ccc ttg act tat cat ggc acc agc aag aca agc tca att	1219
Leu His Ala Pro Leu Thr Tyr His Gly Thr Ser Lys Thr Ser Ser Ile	
360 365 370	
gga ttg tct cgt tta ttg cac cgt gat ggc aat caa aag act gaa ggt	1267
Gly Leu Ser Arg Leu Leu His Arg Asp Gly Asn Gln Lys Thr Glu Gly	
375 380 385	
tat ata aag gtt aat cac aaa cga agc agt aat tat att gat gat gtt	1315
Tyr Ile Lys Val Asn His Lys Arg Ser Ser Asn Tyr Ile Asp Asp Val	
390 395 400 405	
aat ctt gaa gtg caa aac cgc cgc aca aca gga tat aac att ggc atc	1363
Asn Leu Glu Val Gln Asn Arg Arg Thr Thr Gly Tyr Asn Ile Gly Ile	
410 415 420	

133/189

act	cat	cag	cat	cat	ctt	cac	caa	ggt	ggt	tat	ttg	tat	gcc	aac	ttg	1411
Thr	His	Gln	His	His	Leu	His	Gln	Gly	Gly	Tyr	Leu	Tyr	Ala	Asn	Leu	
			425					430					435			
gat	tat	aag	caa	gga	aca	ggg	gca	cta	aaa	gca	aag	cct	gca	cct	gaa	1459
Asp	Tyr	Lys	Gln	Gly	Thr	Gly	Ala	Leu	Lys	Ala	Lys	Pro	Ala	Pro	Glu	
		440					445					450				
gag	cat	att	tat	gat	gct	ttt	ggg	cga	caa	cta	ccc	agt	gag	ggt	ttt	1507
Glu	His	Ile	Tyr	Asp	Ala	Phe	Gly	Arg	Gln	Leu	Pro	Ser	Glu	Gly	Phe	
	455					460					465					
gca	aaa	gca	cca	att	tgg	tca	cta	tac	acc	agt	ttc	caa	aag	cca	ttt	1555
Ala	Lys	Ala	Pro	Ile	Trp	Ser	Leu	Tyr	Thr	Ser	Phe	Gln	Lys	Pro	Phe	
470					475				480						485	
gtg	tta	aac	aac	cca	aat	gat	act	gaa	caa	gcg	caa	gat	aaa	gca	aca	1603
Val	Leu	Asn	Asn	Pro	Asn	Asp	Thr	Glu	Gln	Ala	Gln	Asp	Lys	Ala	Thr	
				490					495					500		
gca	aat	aaa	ccc	att	tac	aca	tcc	att	cct	ctc	act	tat	act	gct	aga	1651
Ala	Asn	Lys	Pro	Ile	Tyr	Thr	Ser	Ile	Pro	Leu	Thr	Tyr	Thr	Ala	Arg	
			505					510					515			
tta	caa	gcc	caa	tat	gcc	aaa	caa	ctt	cct	gta	cca	tct	gat	tta	ttc	1699
Leu	Gln	Ala	Gln	Tyr	Ala	Lys	Gln	Leu	Pro	Val	Pro	Ser	Asp	Leu	Phe	
		520					525				530					
tat	ttg	ggt	ggc	aga	tat	agc	att	aaa	ggc	att	aaa	gag	gga	aat	tat	1747
Tyr	Leu	Gly	Gly	Arg	Tyr	Ser	Ile	Lys	Gly	Ile	Lys	Glu	Gly	Asn	Tyr	
	535					540					545					
tta	tca	gga	gaa	cat	ggc	ttt	agc	cta	tcc	caa	gaa	ctt	gct	tgg	caa	1795
Leu	Ser	Gly	Glu	His	Gly	Phe	Ser	Leu	Ser	Gln	Glu	Leu	Ala	Trp	Gln	
550					555					560					565	
ttg	cca	tta	cag	aat	ctt	aat	cag	cac	ttt	agt	aca	aat	gca	aac	agc	1843
Leu	Pro	Leu	Gln	Asn	Leu	Asn	Gln	His	Phe	Ser	Thr	Asn	Ala	Asn	Ser	
				570					575					580		
gcc	cag	ttg	tat	gca	agt	att	gat	caa	ggc	tat	gcc	tat	gga	aaa	aac	1891
Ala	Gln	Leu	Tyr	Ala	Ser	Ile	Asp	Gln	Gly	Tyr	Ala	Tyr	Gly	Lys	Asn	
			585					590					595			
act	ctt	aat	aat	caa	cgc	cat	atc	ttg	gct	gga	gcg	gtt	ggt	atg	agg	1939
Thr	Leu	Asn	Asn	Gln	Arg	His	Ile	Leu	Ala	Gly	Ala	Val	Gly	Met	Arg	
		600					605					610				
tat	tat	ttt	caa	ggc	agt	caa	gat	cca	aga	att	caa	gaa	aca	caa	aat	1987
Tyr	Tyr	Phe	Gln	Gly	Ser	Gln	Asp	Pro	Arg	Ile	Gln	Glu	Thr	Gln	Asn	
	615					620					625					
ggc	ttg	act	cat	ttc	aaa	gaa	tca	aat	act	tat	cta	aat	aac	aca	cca	2035
Gly	Leu	Thr	His	Phe	Lys	Glu	Ser	Asn	Thr	Tyr	Leu	Asn	Asn	Thr	Pro	
630					635					640					645	

134/189

```
acc aca gct cat ttg gat ata ttc att gga aaa gga att aag acg cct 2083
Thr Thr Ala His Leu Asp Ile Phe Ile Gly Lys Gly Ile Lys Thr Pro
                      650                      655                      660

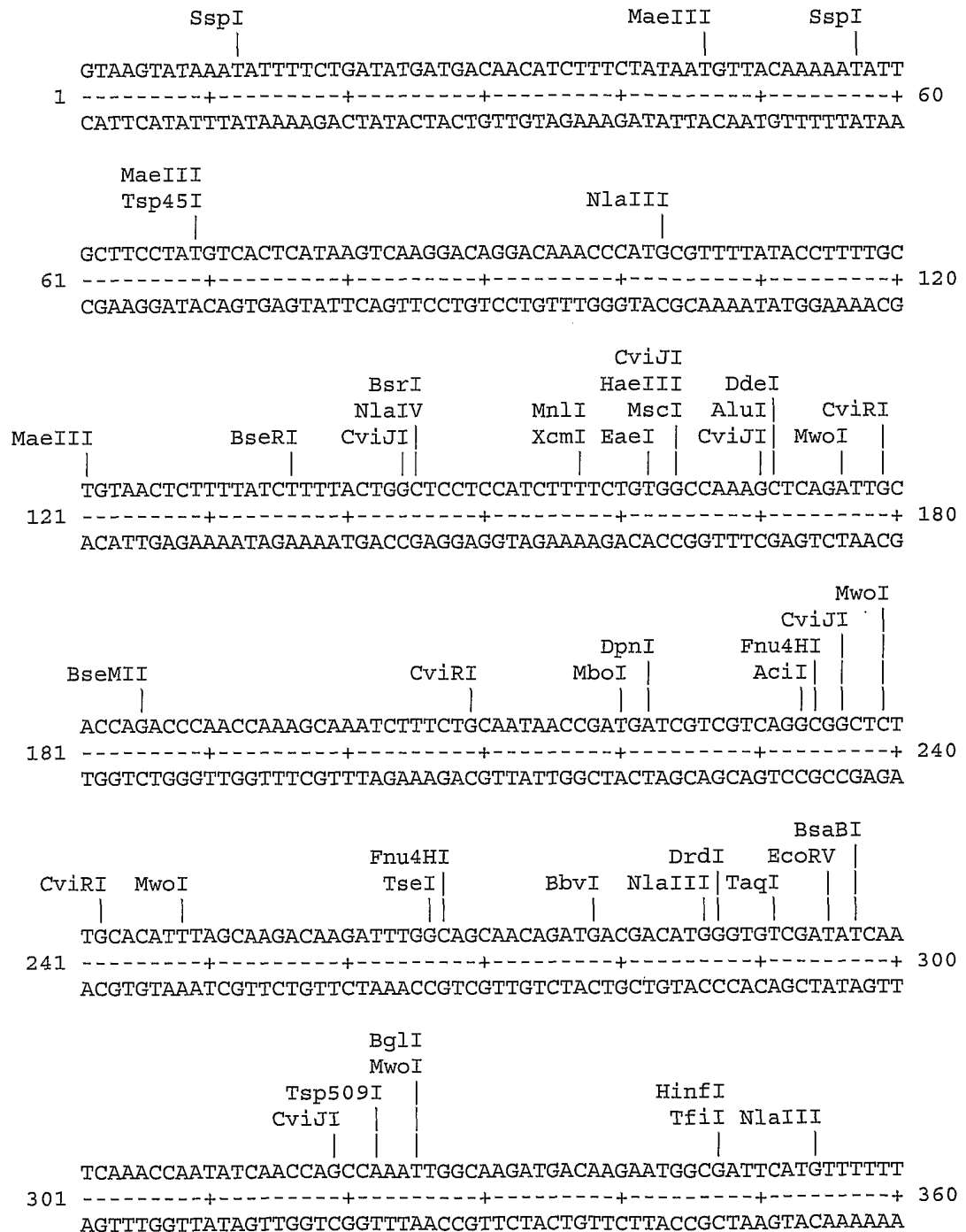
gaa ttt atg aag aaa gaa acc gtc gtg ggt gtg agt gcc agc att gag 2131
Glu Phe Met Lys Lys Glu Thr Val Val Gly Val Ser Ala Ser Ile Glu
                      665                      670                      675

ttt taa gtcattattat tttttgtaat ggagtgggaag atgactcact ttgcattatc 2187
Phe

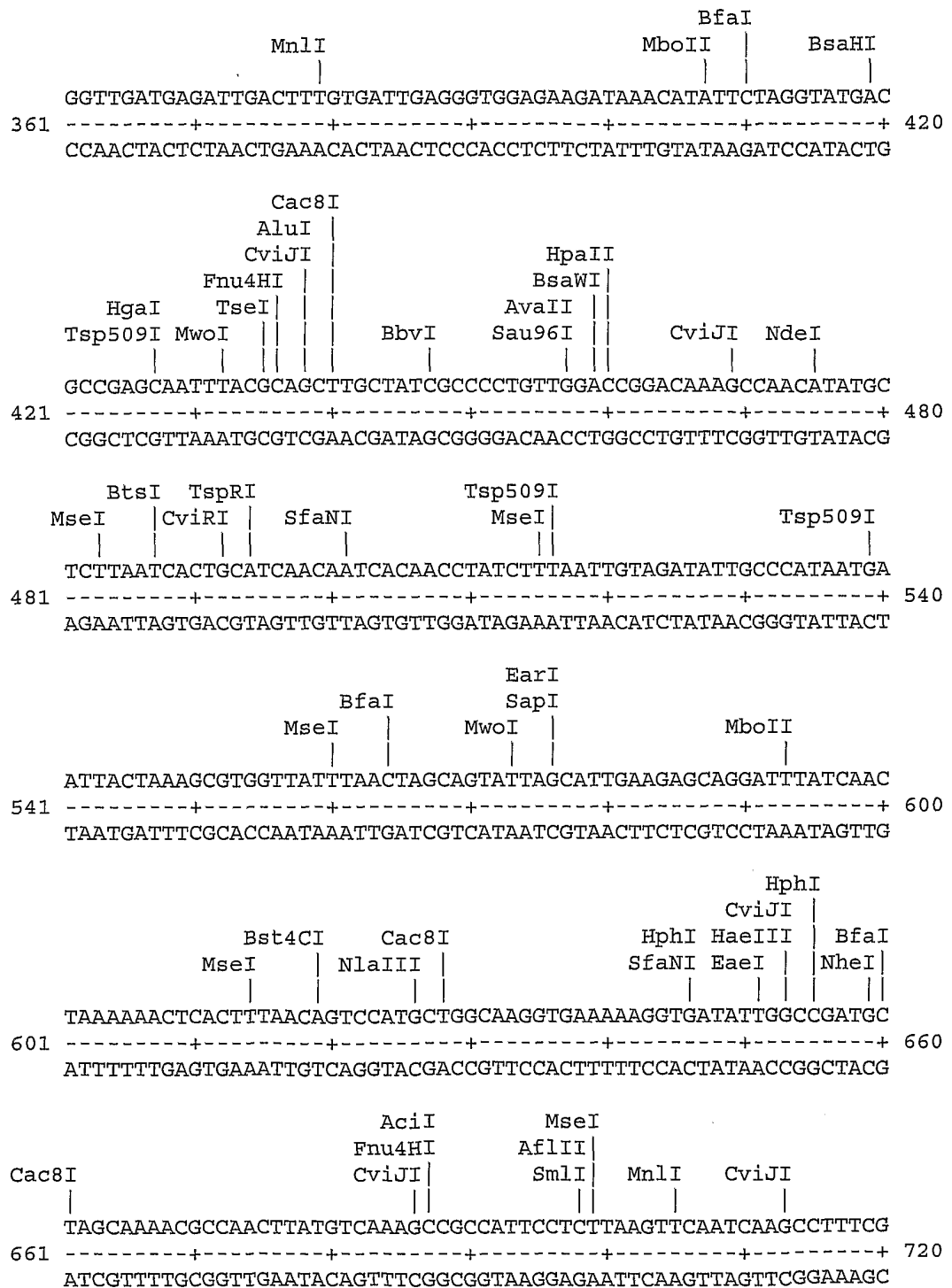
taagtgagtc attaaggcctt cttcgataat gacgcgcacg ggtacgccat 2237
```

135/189

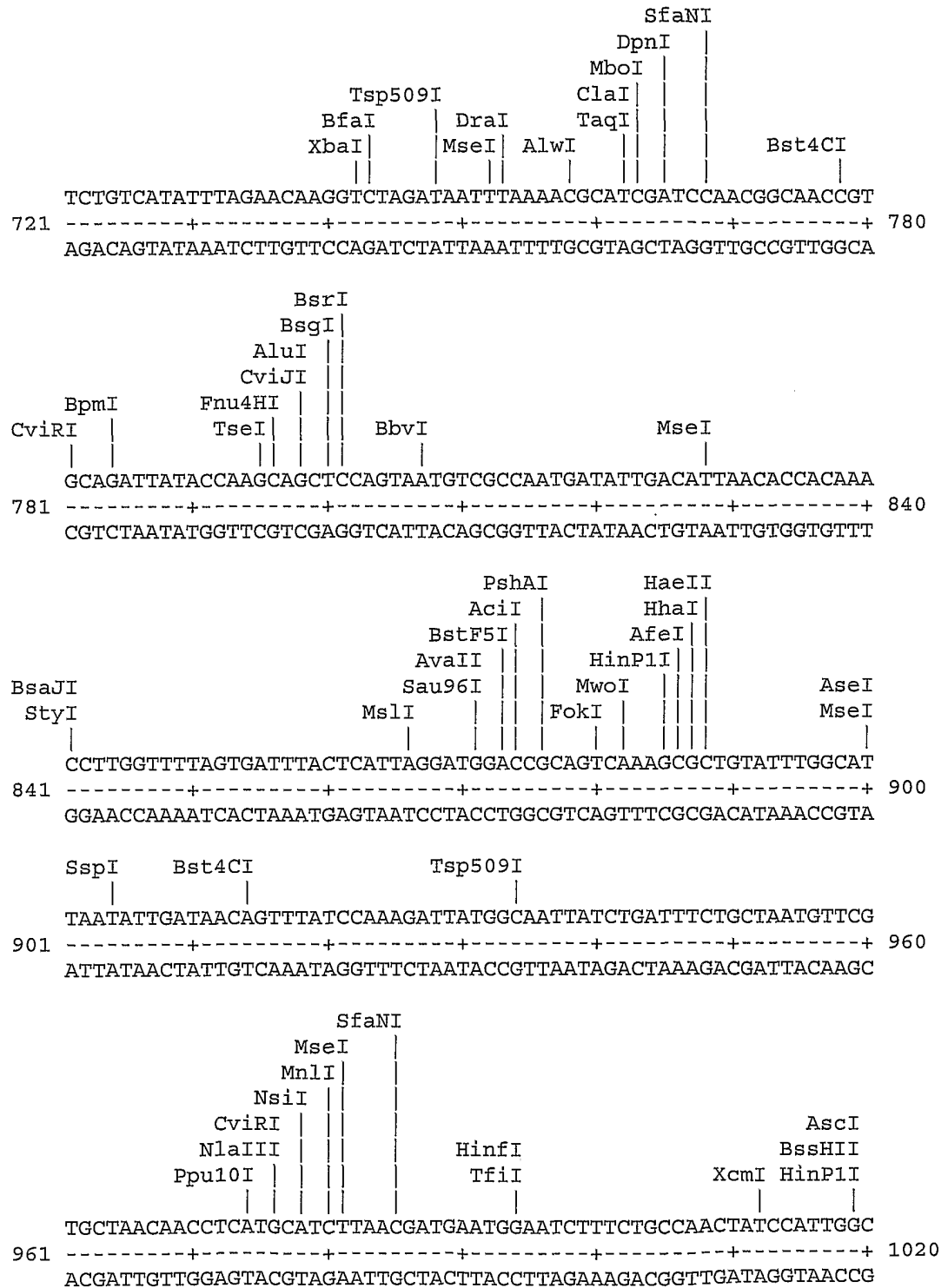
Figure 46. Restriction enzyme analysis of the *M. catarrhalis* HecB gene designated MCA101915



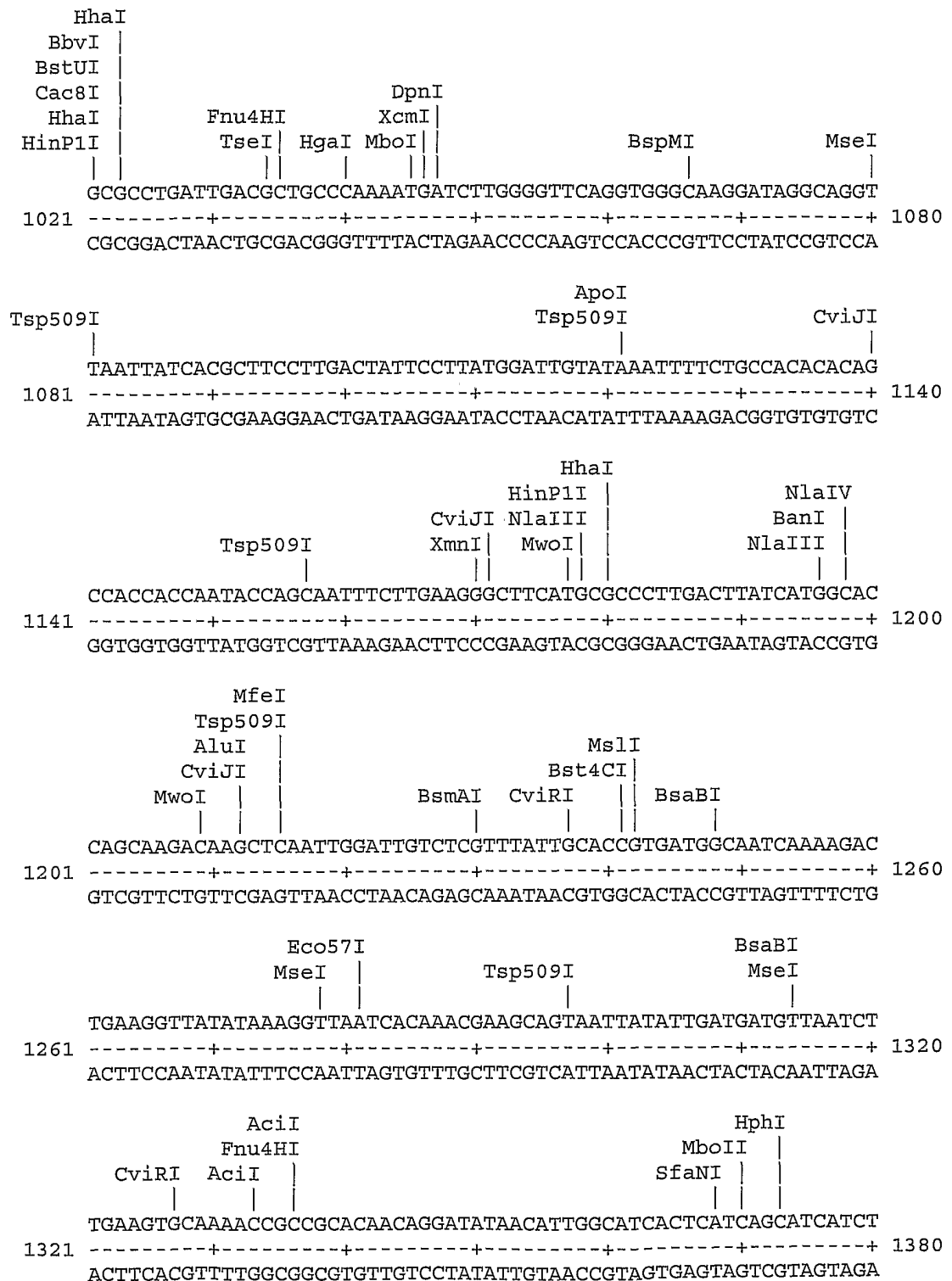
136/189



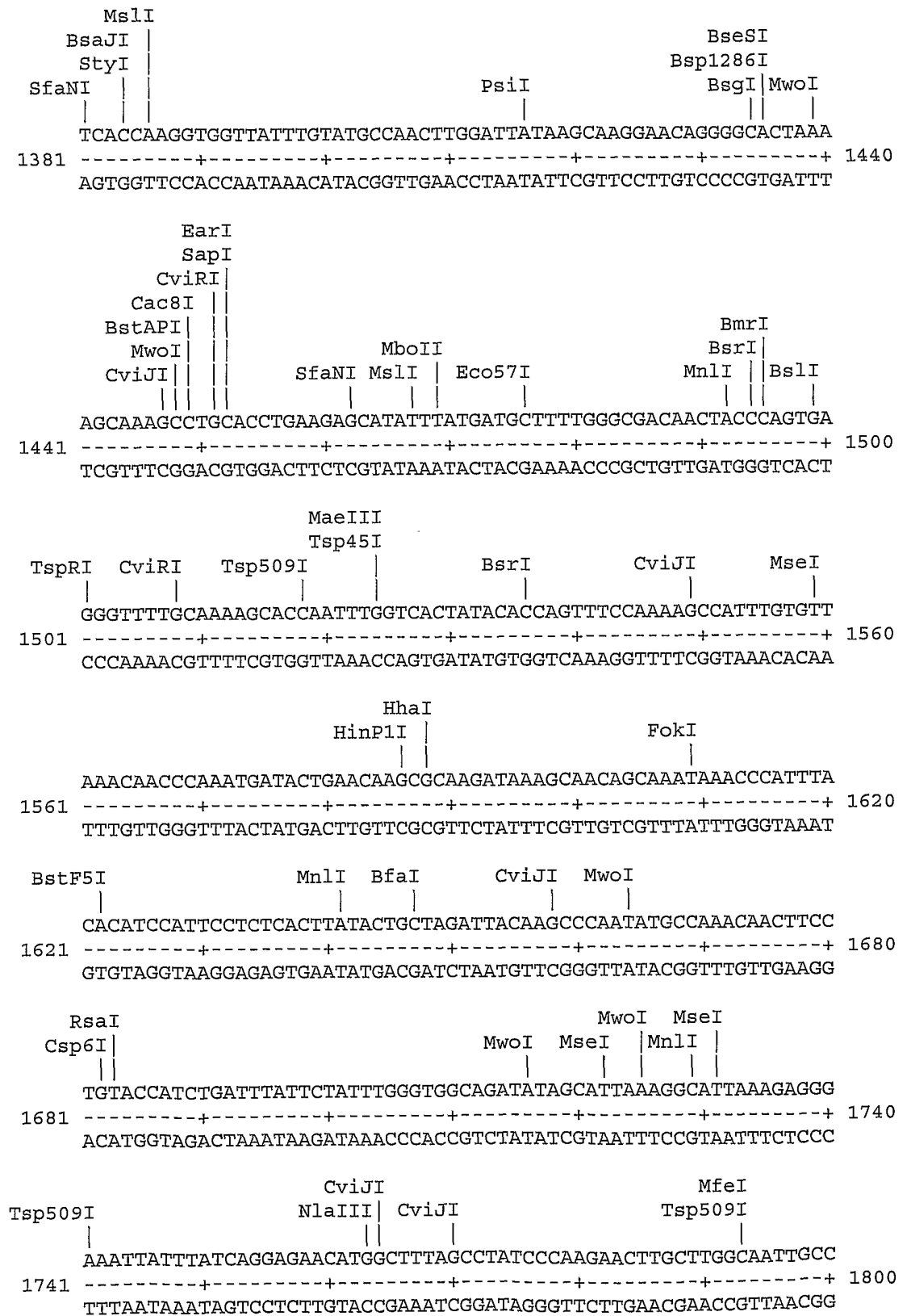
137/189



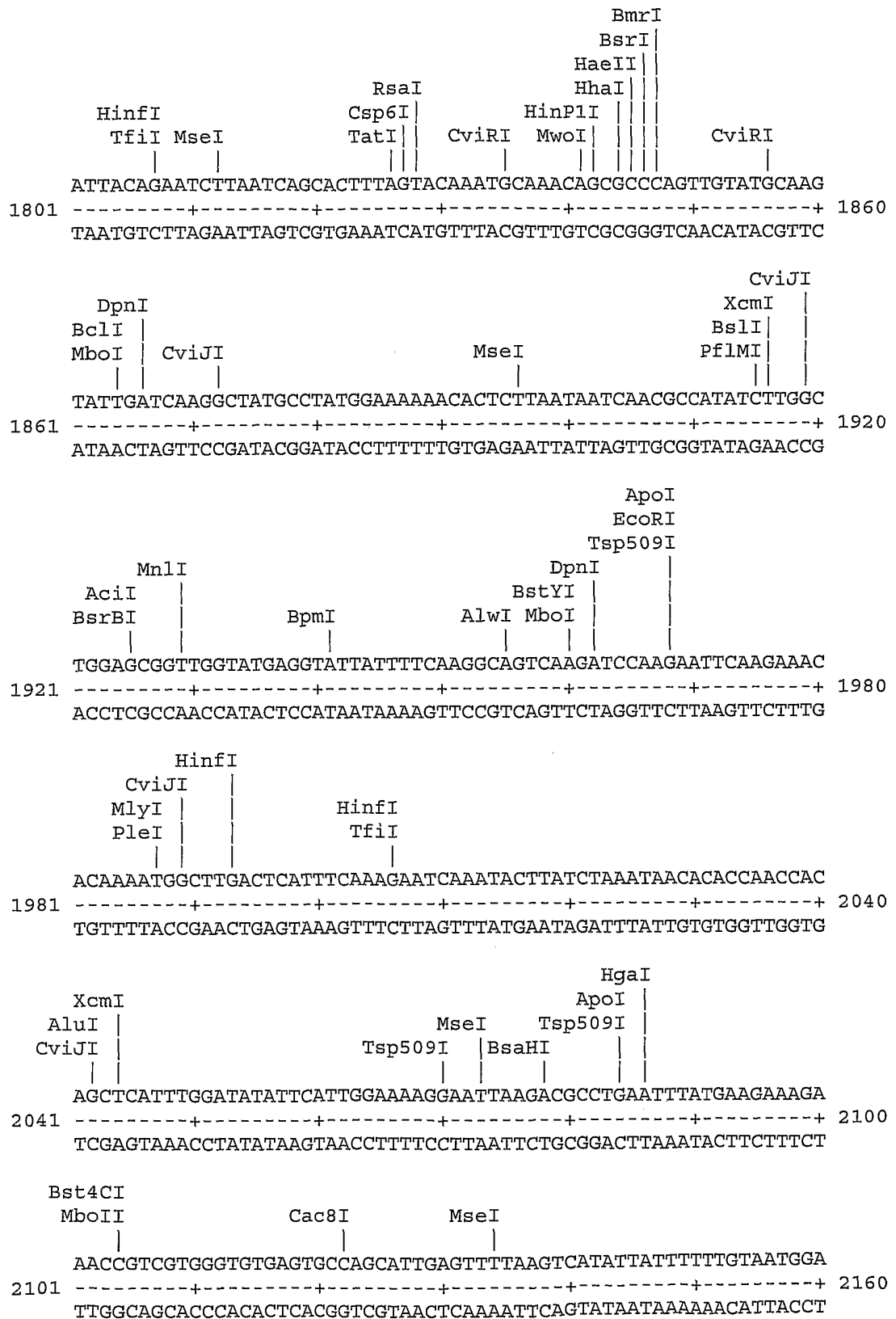
138/189



139/189



140/189



```

                                     CviJI
                                     MlyI
                                     PleI
MlyI                               CviRI       MboII
PleI HinfI MboII                 DdeI HinfI MseI           TaqI
|   |   |   |   |   |   |   |   |   |   |   |   |
GTGGAAGATGACTCACTTTGCATTATCTAAGTGAGTCATTAAGGCTTCTTCGATAATGAC
2161 -----+-----+-----+-----+-----+-----+-----+ 2220
CACCTTCTACTGAGTGAAACGTAATAGATTCACTCAGTAATCCGAAGAAGCTATTACTG

                                     SfaNI
HhaI   RsaI
BstUI  Csp6I
HinP1I HgaI
|   |   |   |   |   |   |
GCGCATCGGTACGCCAT
2221 -----+----- 2237
CGCGTAGCCATGCGGTA

```

142/189

Figure 47. Sequence of *M. catarrhalis* gene encoding a putative virulence factor designated MCA101924

```

caaaattttt atcactttta ttgataaagt atccaagaaa actgagcaaa ttagtgaaaa 60
accaagtaaa ataacacatt tttgataagt ttatattaag atg gca aaa tct cgg 115
                                         Met Ala Lys Ser Arg
                                         1           5

ctc ttt cgc tcc acc gtc att gtc agt tcc atg acc atg ctt tct cgt 163
Leu Phe Arg Ser Thr Val Ile Val Ser Ser Met Thr Met Leu Ser Arg
                        10                      15                      20

att tta ggt tta atc aga gac atg gtt ttg atg agt gtt ttt ggt gct 211
Ile Leu Gly Leu Ile Arg Asp Met Val Leu Met Ser Val Phe Gly Ala
                        25                      30                      35

ggt ggg ttg atg gat gcg ttt tta gtt gcc ttt aaa atc ccc aat ttt 259
Gly Gly Leu Met Asp Ala Phe Leu Val Ala Phe Lys Ile Pro Asn Phe
                        40                      45                      50

tta cgc cgc ctg ttt gcc gaa ggg gca ttt tcc caa gcc ttt gtt cct 307
Leu Arg Arg Leu Phe Ala Glu Gly Ala Phe Ser Gln Ala Phe Val Pro
                        55                      60                      65

gta tta act gag tac aaa gaa gca cgc act tta act gag gtg caa ctg 355
Val Leu Thr Glu Tyr Lys Glu Ala Arg Thr Leu Thr Glu Val Gln Leu
                        70                      75                      80                      85

ctg att gct cgt gtc tct ggc gta tta tcc ttg gtg cta ctt gta ttg 403
Leu Ile Ala Arg Val Ser Gly Val Leu Ser Leu Val Leu Leu Val Leu
                        90                      95                      100

acg gtg gcg gtt atc ttt ttg gca cct caa gtg gtg tat tta ttc gca 451
Thr Val Ala Val Ile Phe Leu Ala Pro Gln Val Val Tyr Leu Phe Ala
                        105                      110                      115

cca ggt ttt gcc gat gac ccc caa aag ttt gac act acc gcc aag ctg 499
Pro Gly Phe Ala Asp Asp Pro Gln Lys Phe Asp Thr Thr Ala Lys Leu
                        120                      125                      130

tta cgc ttg act ttc cct tat tta tta ttc gtc tcc atg acg gca ttt 547
Leu Arg Leu Thr Phe Pro Tyr Leu Leu Phe Val Ser Met Thr Ala Phe
                        135                      140                      145

ttt ggt agc att cta caa agt tat gga aga ttt gca gca ccc gcc ttt 595
Phe Gly Ser Ile Leu Gln Ser Tyr Gly Arg Phe Ala Ala Pro Ala Phe
                        150                      155                      160                      165

gct cct gta tta cta aat ctt tgt atg att ggc ggt aca tta ata tta 643
Ala Pro Val Leu Leu Asn Leu Cys Met Ile Gly Gly Thr Leu Ile Leu
                        170                      175                      180

```

143/189

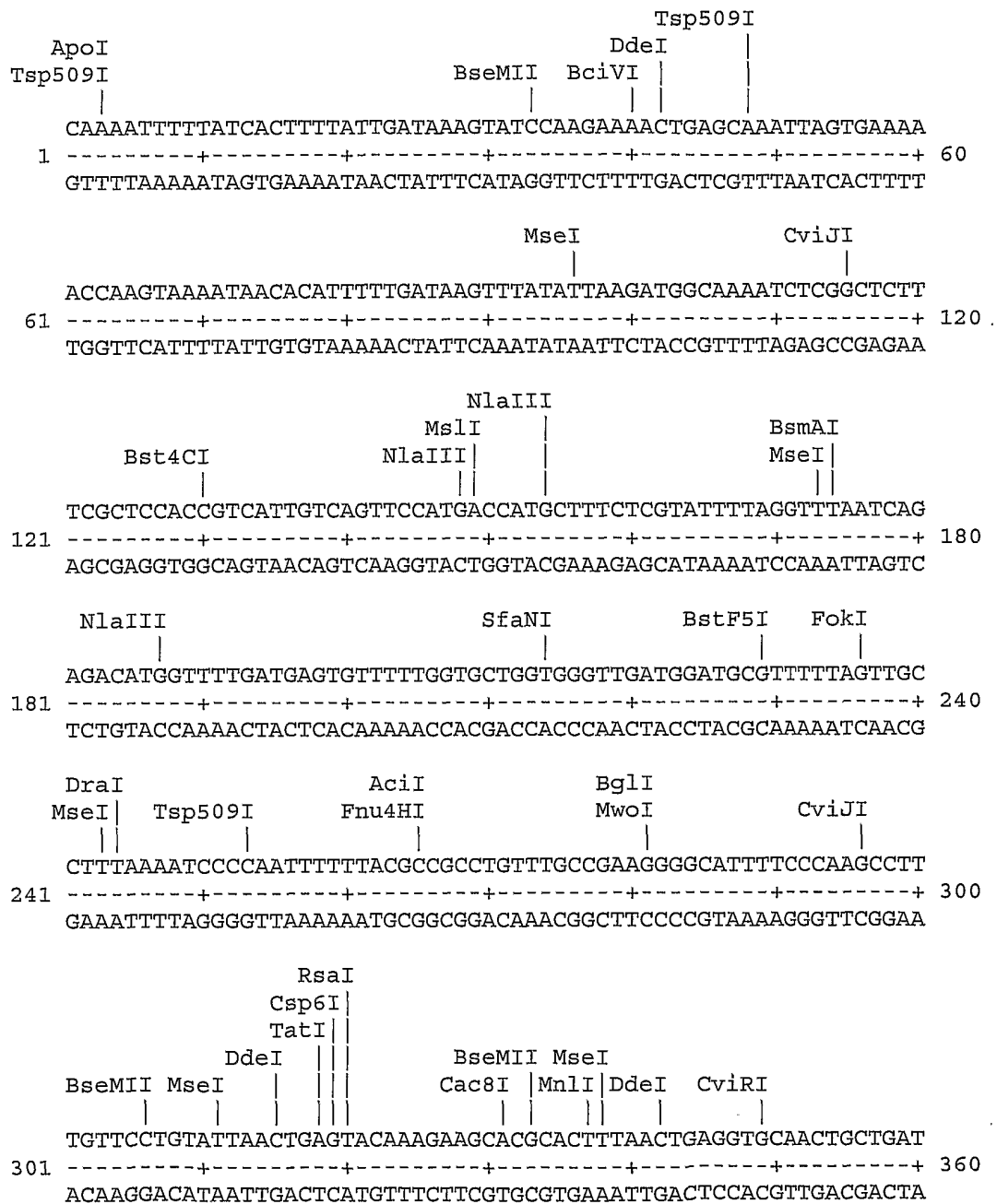
gca cca atg ctt gat aag cca atc atg gca ttg ggt tat gct gtc gcc	691
Ala Pro Met Leu Asp Lys Pro Ile Met Ala Leu Gly Tyr Ala Val Ala	
185 190 195	
att tca ggc att ttg cag ctg ctc ata cag ttg cca cag ctt tgg caa	739
Ile Ser Gly Ile Leu Gln Leu Ile Gln Leu Pro Gln Leu Trp Gln	
200 205 210	
caa aaa ctg ctt att cca ccc agt att agc ttt aaa gat gaa ggc gtt	787
Gln Lys Leu Leu Ile Pro Pro Ser Ile Ser Phe Lys Asp Glu Gly Val	
215 220 225	
agg cgt att tta aag ctc atg ctg cct gcc att ttt ggg gta tcc gtc	835
Arg Arg Ile Leu Lys Leu Met Leu Pro Ala Ile Phe Gly Val Ser Val	
230 235 240 245	
acc cag att aac tta tta tta agc acc gta ttt gcc tca tta atg att	883
Thr Gln Ile Asn Leu Leu Leu Ser Thr Val Phe Ala Ser Leu Met Ile	
250 255 260	
gca ggt tca gta tca tgg atg tat gct gcc gaa cgc tta agt gaa ttg	931
Ala Gly Ser Val Ser Trp Met Tyr Ala Ala Glu Arg Leu Ser Glu Leu	
265 270 275	
cca tta gga ctc att ggt gtg gcg att ggt acg gtg att tta cca agc	979
Pro Leu Gly Leu Ile Gly Val Ala Ile Gly Thr Val Ile Leu Pro Ser	
280 285 290	
cta tca tca agt cgt gcc aag tca gat gat gaa aat ttt cgc caa acg	1027
Leu Ser Ser Ser Arg Ala Lys Ser Asp Asp Glu Asn Phe Arg Gln Thr	
295 300 305	
cta gac tgg gct gcc cgc tta att ctt gtc atc ggt ata cca gcc agt	1075
Leu Asp Trp Ala Ala Arg Leu Ile Leu Val Ile Gly Ile Pro Ala Ser	
310 315 320 325	
ctt gcg ata ttt gtt tta tca gat att ttg atg gat gct ttg ttt gtg	1123
Leu Ala Ile Phe Val Leu Ser Asp Ile Leu Met Asp Ala Leu Phe Val	
330 335 340	
cgt ggt caa ttt acc cac caa gat gcc ttg atg agt ggt att gct cta	1171
Arg Gly Gln Phe Thr His Gln Asp Ala Leu Met Ser Gly Ile Ala Leu	
345 350 355	
aag tct tta tca gga ggg atc tta gga ttt atg ctc att aaa att ttt	1219
Lys Ser Leu Ser Gly Gly Ile Leu Gly Phe Met Leu Ile Lys Ile Phe	
360 365 370	
gca cct gca ttt ttt gca aat caa gac acc aaa acg cct gtt aaa att	1267
Ala Pro Ala Phe Phe Ala Asn Gln Asp Thr Lys Thr Pro Val Lys Ile	
375 380 385	
ggc atc att tca gta ttt gct aat atg att ttt agc gtc gta ttt att	1315
Gly Ile Ile Ser Val Phe Ala Asn Met Ile Phe Ser Val Val Phe Ile	
390 395 400 405	

144/189

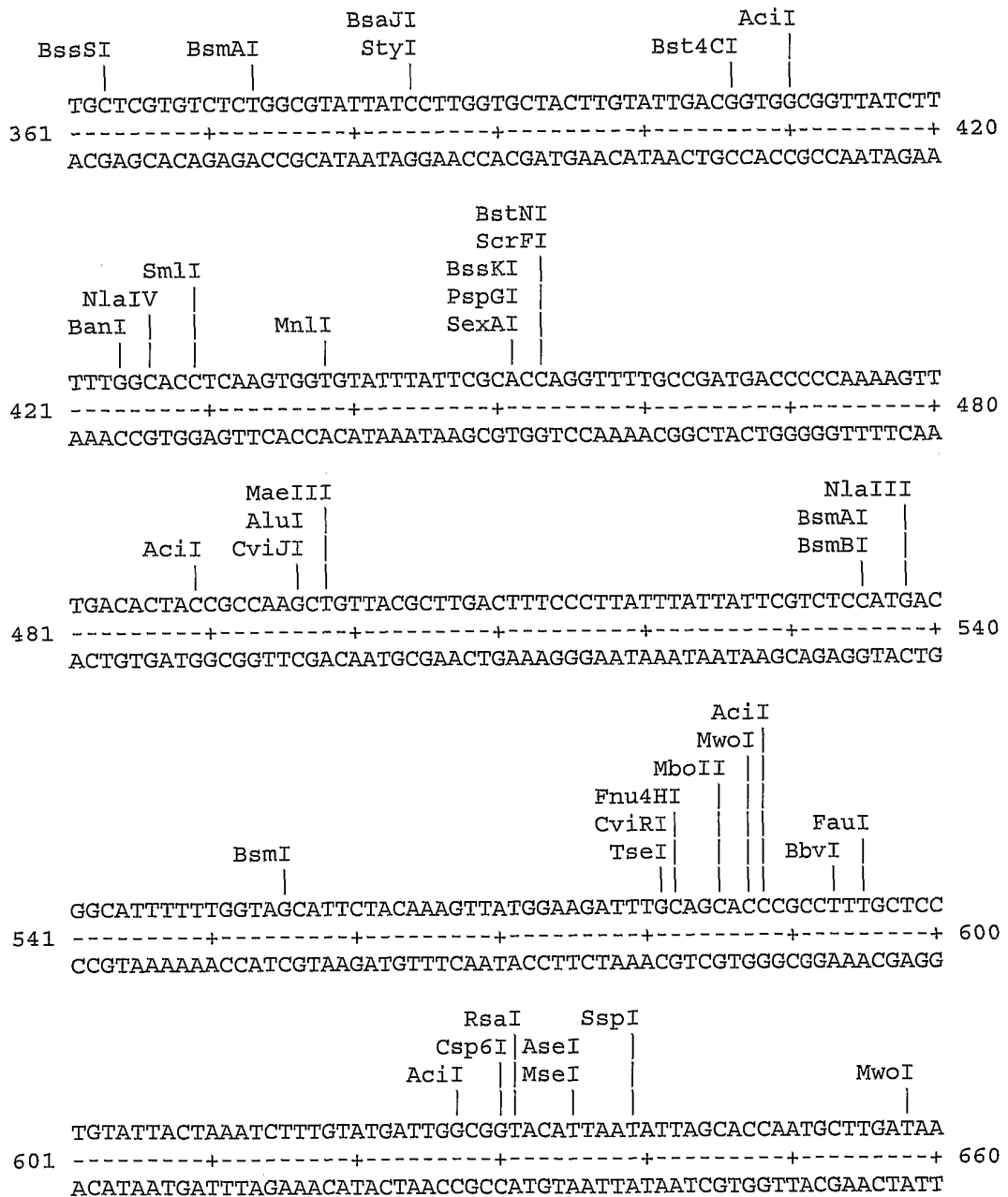
ggg cta ttt tat tta atg aaa ctg cct tta cat ggt gga ttg gcc ttg	1363
Gly Leu Phe Tyr Leu Met Lys Leu Pro Leu His Gly Gly Leu Ala Leu	
410 415 420	
gca acc aca gca gca agt ttt gtt aac gca ggt cta ctc tat gcc gtt	1411
Ala Thr Thr Ala Ala Ser Phe Val Asn Ala Gly Leu Leu Tyr Ala Val	
425 430 435	
tta cat cgc cgt ggt atc tat cgt ttt ggt act cac tgg aaa aag atg	1459
Leu His Arg Arg Gly Ile Tyr Arg Phe Gly Thr His Trp Lys Lys Met	
440 445 450	
ata cga caa ttt gtg att gcc aat att gcc atg gtg att agc tta tct	1507
Ile Arg Gln Phe Val Ile Ala Asn Ile Ala Met Val Ile Ser Leu Ser	
455 460 465	
gtc att ttg ccc tta tat cca att gat gct aat cac tgg ctc aaa att	1555
Val Ile Leu Pro Leu Tyr Pro Ile Asp Ala Asn His Trp Leu Lys Ile	
470 475 480 485	
gcc atc ttg ctg ggc ata tgc ggt ttt ggg gca atc att tat ggt atc	1603
Ala Ile Leu Leu Gly Ile Cys Gly Phe Gly Ala Ile Ile Tyr Gly Ile	
490 495 500	
aca cta ctc atc act ggt ttt cgc cca agt cag ctt aag cat ggt tga	1651
Thr Leu Leu Ile Thr Gly Phe Arg Pro Ser Gln Leu Lys His Gly	
505 510 515	
gctgtcatta tatattctac tgtcattata catgataccc tgccgaacca ttgacataag	1711
caatacaaaag cacacaaaaa aaattacttg	1741

145/189

Figure 48. Restriction enzyme analysis of the *M. catarrhalis* gene encoding a putative virulence factor designated MCA101924



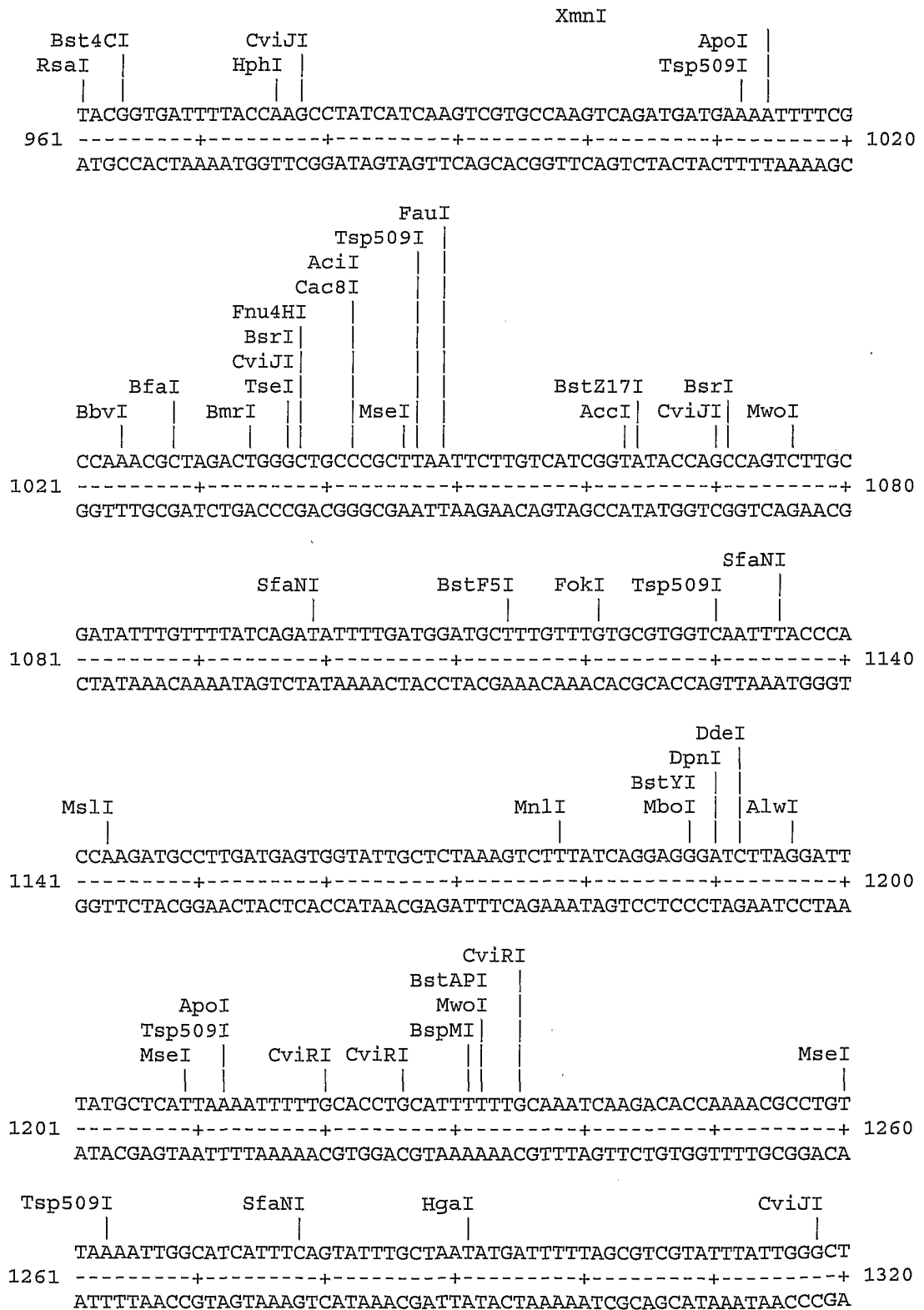
146/189



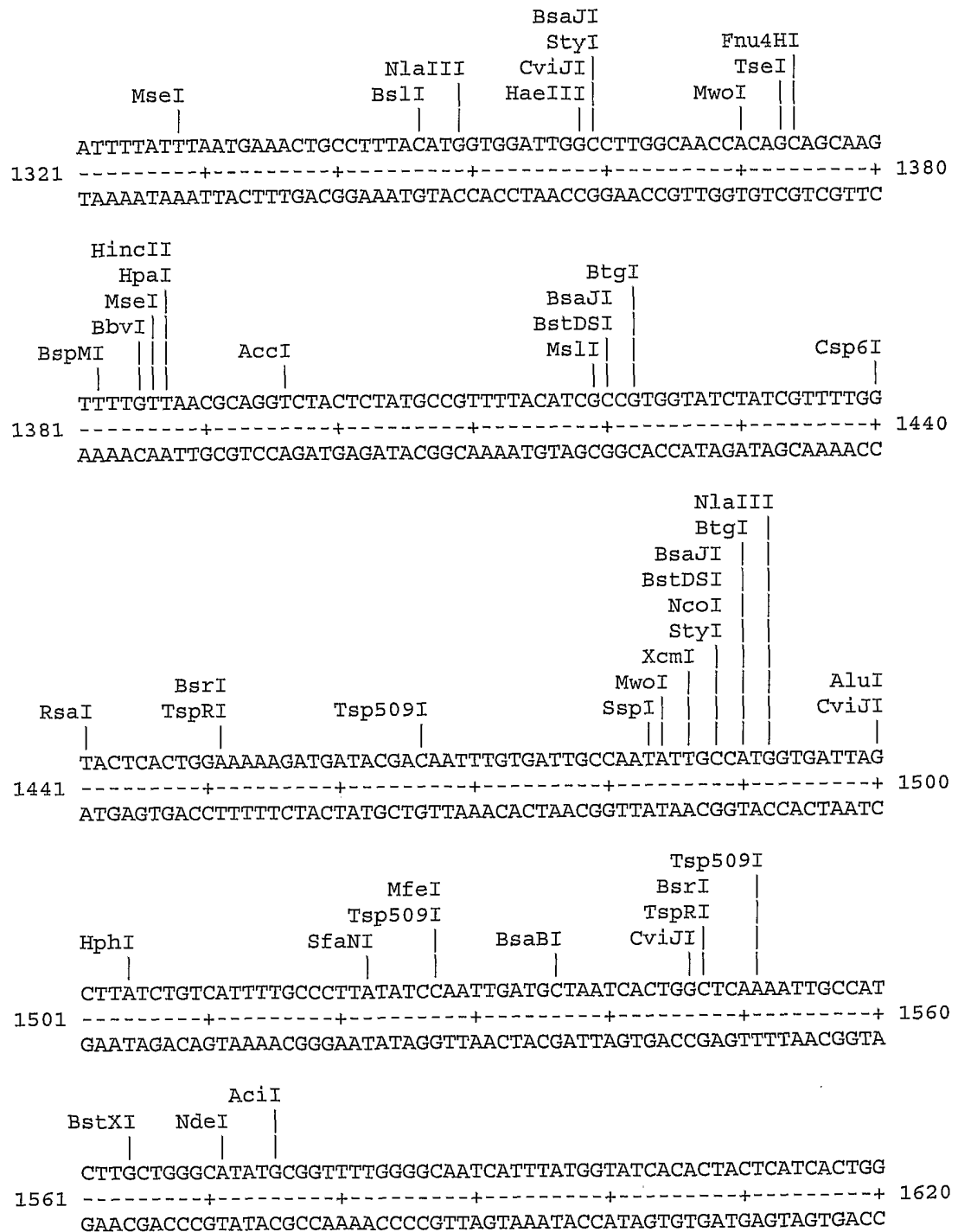
147/189

Fnu4HI
 AluI
 CviJI
 MspAII
 PvuII
 TseI
 Fnu4HI
 CviRI
 TseI
 Bst4CI
 BbvI
 CviJI
 NlaIII
 BbvI
 661
 GCCAATCATGGCATTGGGTTATGCTGTGCGCCATTTTCAGGCATTTTGCAGCTGCTCATACA
 -----+-----+-----+-----+-----+-----+-----+-----+ 720
 CGGTTAGTAGCCGTAAACCAATACGACAGCGGTAAAGTCCGTAAAACGTCGACGAGTATGT
 BstXI
 AluI
 CviJI
 BmrI
 BsrI
 AluI
 CviJI
 MseI
 DraI
 721
 GTTGCCACAGCTTTGGCAACAAAACTGCTTATTCCACCCAGTATTAGCTTTAAAGATGA
 -----+-----+-----+-----+-----+-----+-----+-----+ 780
 CAACGGTGTGCGAAACCGTTGTTTTTGACGAATAAGGTGGGTCATAATCGAAATTTCTACT
 AluI
 Cac8I
 CviJI
 Fnu4HI
 DraI
 MseI
 BbvI
 MwoI
 NlaIII
 TseI
 MaeIII
 HphI
 BslI
 Tsp45I
 BciVI
 781
 AGGCGTTAGGCGTATTTTAAAGCTCATGCTGCCTGCCATTTTGGGGTATCCGTCACCCA
 -----+-----+-----+-----+-----+-----+-----+-----+ 840
 TCCGCAATCCGCATAAAATTTTCGAGTACGACGCGGTAACCAACCCATAGGCAGTGGGT
 AseI
 BspMI
 CviRI
 NlaIII
 MseI
 MseI
 Bst4CI
 MseI
 MnlI
 BbvI
 841
 GATTAACCTTATTATTAAAGCACCGTATTTGCCTCATTAATGATTGCAGGTTTCAAGTATCATG
 -----+-----+-----+-----+-----+-----+-----+-----+ 900
 CTAATTGAATAATAATTCGTGGCATAAACGAGTAATTACTAACGTCCAAGTCATAGTAC
 Tsp509I
 Fnu4HI
 TseI
 BstF5I
 FokI
 MseI
 AflIII
 SmlI
 MlyI
 PleI
 HinfI
 MslI
 Csp6I
 901
 GATGTATGCTGCCGAACGCTTAAGTGAATTGCCATTAGGACTCATTGGTGTGGCGATTGG
 -----+-----+-----+-----+-----+-----+-----+-----+ 960
 CTACATACGACGGCTTGCGAATTTCACTTAACGGTAATCCTGAGTAACACACCGCTAAC

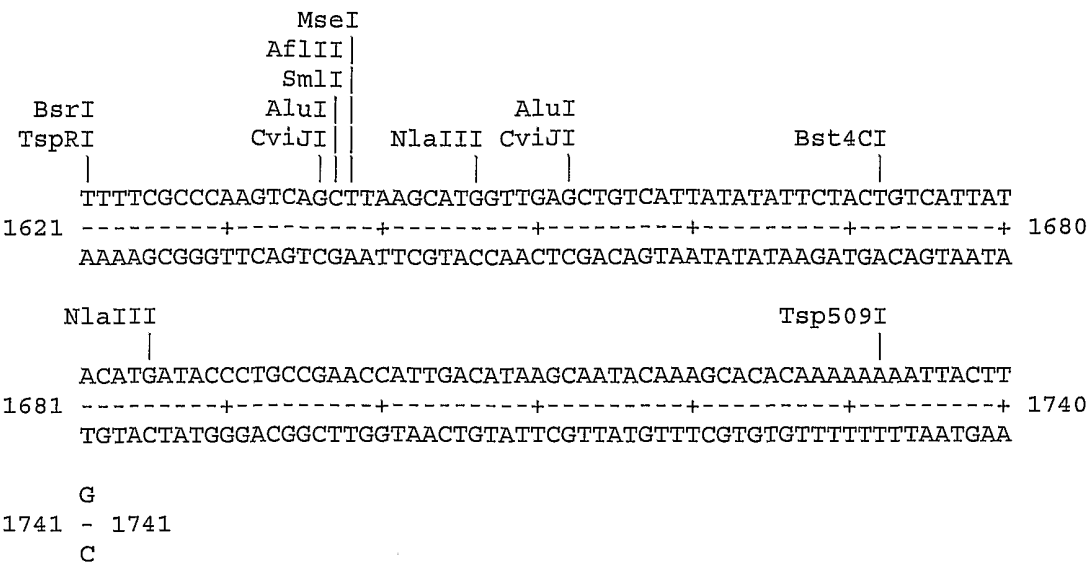
148/189



149/189



150/189



151/189

Figure 49. Sequence of *M. catarrhalis* heme receptor gene designated MCA101976

```

atgctcaagt atcttttctt attaaatcat tggtattttt taaatatgct atataataat 60

gtgcttttatt taattatggt atatcataac cgagggtatc atg ata aaa aaa cca 115
                                         Met Ile Lys Lys Pro
                                         1           5

ctt gtt tgt gcg ata tcg gcc acc ttt gcg atg cca gcg gta gca gat 163
Leu Val Cys Ala Ile Ser Ala Thr Phe Ala Met Pro Ala Val Ala Asp
                        10                15                20

aat acc aag ctg ggt gaa gag cca acc acc acc tta aag ggt gta ttg 211
Asn Thr Lys Leu Gly Glu Glu Pro Thr Thr Thr Leu Lys Gly Val Leu
                        25                30                35

gta agc tcg caa acg aac caa aat aca ggt ttt gta tct aat gat tca 259
Val Ser Ser Gln Thr Asn Gln Asn Thr Gly Phe Val Ser Asn Asp Ser
                        40                45                50

aaa caa tcc agt gat ctt acg ctt tca aaa gat aaa tta aaa tat cgt 307
Lys Gln Ser Ser Asp Leu Thr Leu Ser Lys Asp Lys Leu Lys Tyr Arg
                        55                60                65

tcg gca acc ttg ggc aat gcg tta agt ggt gag ctt ggt att cat agt 355
Ser Ala Thr Leu Gly Asn Ala Leu Ser Gly Glu Leu Gly Ile His Ser
                        70                75                80                85

aac cct ttt ggt ggc ggt tca tct gca cct gtt gtg cga ggg caa gag 403
Asn Pro Phe Gly Gly Gly Ser Ser Ala Pro Val Val Arg Gly Gln Glu
                        90                95                100

ggt gtg cgt ctt aag att tta caa aat gga act gat gtg att gat gtg 451
Gly Val Arg Leu Lys Ile Leu Gln Asn Gly Thr Asp Val Ile Asp Val
                        105                110                115

tca tca ata tcg cct gat cat gtt gtg gcg acc gat aca ctt tta gcg 499
Ser Ser Ile Ser Pro Asp His Val Val Ala Thr Asp Thr Leu Leu Ala
                        120                125                130

tct aaa gtt gag ctt gtt cgt ggt gct gat acg ctg tta tat ggc ttg 547
Ser Lys Val Glu Leu Val Arg Gly Ala Asp Thr Leu Leu Tyr Gly Leu
                        135                140                145

gca tcg cca gct ggt gtg att aat gtt gtt gat gac cgt atc ccg aat 595
Ala Ser Pro Ala Gly Val Ile Asn Val Val Asp Asp Arg Ile Pro Asn
                        150                155                160                165

cgt atg cct agt ggt gct atc cat gac aaa atc gaa ggc gag acg atg 643
Arg Met Pro Ser Gly Ala Ile His Asp Lys Ile Glu Gly Glu Thr Met
                        170                175                180

```

152/189

ctt cga tat aac aca aac aac cat gaa aag ctt gca act gca ggg gtg	691
Leu Arg Tyr Asn Thr Asn Asn His Glu Lys Leu Ala Thr Ala Gly Val	
185 190 195	
agc ttt ggg gta gga gat cgc att gcg gtt cgg gtg gag ggc tta aag	739
Ser Phe Gly Val Gly Asp Arg Ile Ala Val Arg Val Glu Gly Leu Lys	
200 205 210	
cga gag gct gat gac tat caa gtt ccc cat ttt cag gca gat cgc atg	787
Arg Glu Ala Asp Asp Tyr Gln Val Pro His Phe Gln Ala Asp Arg Met	
215 220 225	
tta gat tat gtg cca ggt agt gca aat aac tct acc gtt ggc atg att	835
Leu Asp Tyr Val Pro Gly Ser Ala Asn Asn Ser Thr Val Gly Met Ile	
230 235 240 245	
ggc gtg tct tat att cat gat aat ggg cat atc ggt gct tct tat agc	883
Gly Val Ser Tyr Ile His Asp Asn Gly His Ile Gly Ala Ser Tyr Ser	
250 255 260	
cac cgt aaa gat cgt tat ggt atc cca ggg cat atc cac tgc gac agc	931
His Arg Lys Asp Arg Tyr Gly Ile Pro Gly His Ile His Cys Asp Ser	
265 270 275	
caa cga gag cat ttt atc aaa tgg cat aat atc aca aaa tcc aat tat	979
Gln Arg Glu His Phe Ile Lys Trp His Asn Ile Thr Lys Ser Asn Tyr	
280 285 290	
tat tta ccc att tat cct cat ttg atg gag gat tca gat att gat gac	1027
Tyr Leu Pro Ile Tyr Pro His Leu Met Glu Asp Ser Asp Ile Asp Asp	
295 300 305	
aat cct cat acg cat tgc cgc cac aac cac gaa gac cat atc ggt gag	1075
Asn Pro His Thr His Cys Arg His Asn His Glu Asp His Ile Gly Glu	
310 315 320 325	
cat aat ccc acg ggc gtg ccc atc aat cat gaa cat cat tcg cct tgg	1123
His Asn Pro Thr Gly Val Pro Ile Asn His Glu His His Ser Pro Trp	
330 335 340	
att gat atg aaa acc aat cgc tac gac att cgt ggc gag gta tat cgg	1171
Ile Asp Met Lys Thr Asn Arg Tyr Asp Ile Arg Gly Glu Val Tyr Arg	
345 350 355	
cct att caa ggt ttg gat aaa att aag cta agc tta act tat gca gat	1219
Pro Ile Gln Gly Leu Asp Lys Ile Lys Leu Ser Leu Thr Tyr Ala Asp	
360 365 370	
tat tat cat gat gaa aaa gat gct ggc aat gag caa gac cca aac aat	1267
Tyr Tyr His Asp Glu Lys Asp Ala Gly Asn Glu Gln Asp Pro Asn Asn	
375 380 385	
cac aaa cct tct gag cgt gat aca acg gtg gat aag ggt cat gcc agc	1315
His Lys Pro Ser Glu Arg Asp Thr Thr Val Asp Lys Gly His Ala Ser	
390 395 400 405	

153/189

tct att ttt aca aaa aaa ggc gtt aat ggt cgc ttg gag tta tat cat	1363
Ser Ile Phe Thr Lys Lys Gly Val Asn Gly Arg Leu Glu Leu Tyr His	
410 415 420	
aca ccg acc aaa cgc tta tct ggg gta ttg ggt att ggg tat caa acc	1411
Thr Pro Thr Lys Arg Leu Ser Gly Val Leu Gly Ile Gly Tyr Gln Thr	
425 430 435	
caa aaa tct gca gca gga gag gcg tat ttg cca agc tat ttt caa tca	1459
Gln Lys Ser Ala Ala Gly Glu Ala Tyr Leu Pro Ser Tyr Phe Gln Ser	
440 445 450	
gaa gca gaa tgg caa aaa gcc caa agt caa aac att aac caa tat cgt	1507
Glu Ala Glu Trp Gln Lys Ala Gln Ser Gln Asn Ile Asn Gln Tyr Arg	
455 460 465	
cct tac tta tta gtt cca aat acc aat aaa agc ctt ggt att ttt gga	1555
Pro Tyr Leu Leu Val Pro Asn Thr Asn Lys Ser Leu Gly Ile Phe Gly	
470 475 480 485	
ctt gag caa cta aag cta aat caa atg act ttt aag gtg gcg atg cgt	1603
Leu Glu Gln Leu Lys Leu Asn Gln Met Thr Phe Lys Val Ala Met Arg	
490 495 500	
cat gaa aga caa aaa aca cca att gaa tat gac cag cat tta ctt gac	1651
His Glu Arg Gln Lys Thr Pro Ile Glu Tyr Asp Gln His Leu Leu Asp	
505 510 515	
cat gct ttg cag tat ttt tta agt aaa gca cag cta aaa gca cct gat	1699
His Ala Leu Gln Tyr Phe Leu Ser Lys Ala Gln Leu Lys Ala Pro Asp	
520 525 530	
cat cct gat ttg acg aca tat aaa caa cat gcc acc tct tat gct ggt	1747
His Pro Asp Leu Thr Thr Tyr Lys Gln His Ala Thr Ser Tyr Ala Gly	
535 540 545	
agt gcc tta tgg gat att acg cca aat cat cga ttg tca ttg acc tac	1795
Ser Ala Leu Trp Asp Ile Thr Pro Asn His Arg Leu Ser Leu Thr Tyr	
550 555 560 565	
tca cat aac gaa cgc att cca tcg ccg atg gag ctg tat tat caa ggc	1843
Ser His Asn Glu Arg Ile Pro Ser Pro Met Glu Leu Tyr Tyr Gln Gly	
570 575 580	
gga cat ttg gcg acc agc tct ttt gag cat ggc aat aaa aac ttg gtc	1891
Gly His Leu Ala Thr Ser Ser Phe Glu His Gly Asn Lys Asn Leu Val	
585 590 595	
aaa gaa aaa tcg gat aat tat gag ctg ggt ttt atg cat aca gca gat	1939
Lys Glu Lys Ser Asp Asn Tyr Glu Leu Gly Phe Met His Thr Ala Asp	
600 605 610	
aaa gtc agc tat aaa gca agc act tac tat agc aat ttt gat aac tat	1987
Lys Val Ser Tyr Lys Ala Ser Thr Tyr Tyr Ser Asn Phe Asp Asn Tyr	
615 620 625	

154/189

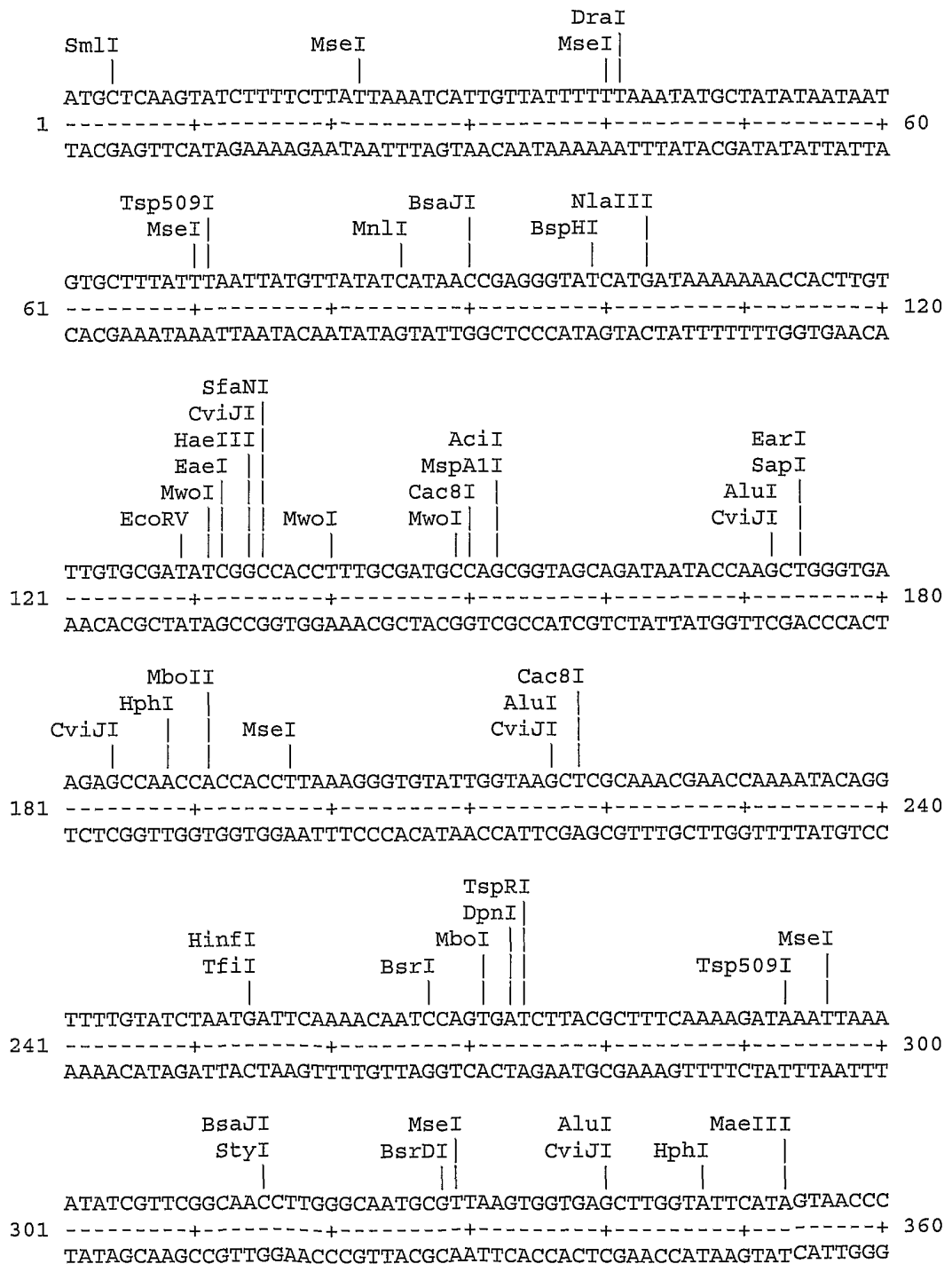
att ttt aat gag acc att gcc aaa gaa ggc aat tta tac atc aga cgc	2035
Ile Phe Asn Glu Thr Ile Ala Lys Glu Gly Asn Leu Tyr Ile Arg Arg	
630 635 640 645	
tat aat cag acg acg gct aag ttt tat ggt gtg gag ggt tca tta act	2083
Tyr Asn Gln Thr Thr Ala Lys Phe Tyr Gly Val Glu Gly Ser Leu Thr	
650 655 660	
tac cag cca aat gcc aat cac agt gtg atg ttt ttt ggt gat atg gtg	2131
Tyr Gln Pro Asn Ala Asn His Ser Val Met Phe Phe Gly Asp Met Val	
665 670 675	
caa ggt aaa att ggt gca tta tct gat att aaa ggt aag ctt gta tat	2179
Gln Gly Lys Ile Gly Ala Leu Ser Asp Ile Lys Gly Lys Leu Val Tyr	
680 685 690	
gct ggc aga aag tgg gtt tat ttt gat gat gat atc aag gat atg acc	2227
Ala Gly Arg Lys Trp Val Tyr Phe Asp Asp Asp Ile Lys Asp Met Thr	
695 700 705	
gtt gat gat aat ggt gat tat gat gcc gat ggt ggt ttg act tgt gcc	2275
Val Asp Asp Asn Gly Asp Tyr Asp Ala Asp Gly Gly Leu Thr Cys Ala	
710 715 720 725	
tta aaa acg cct gaa caa tgg gga caa att aac gat aat aat gat tgt	2323
Leu Lys Thr Pro Glu Gln Trp Gly Gln Ile Asn Asp Asn Asn Asp Cys	
730 735 740	
agt aca acc att aat gtc tat aaa aac ggt aca aca acc tca ggc gag	2371
Ser Thr Thr Ile Asn Val Tyr Lys Asn Gly Thr Thr Thr Ser Gly Glu	
745 750 755	
gag gat tac gac cgt ttg gca cga aat cca act tac gca ccc aga gta	2419
Glu Asp Tyr Asp Arg Leu Ala Arg Asn Pro Thr Tyr Ala Pro Arg Val	
760 765 770	
ccg ccc agc cgc ttg ggc att cgt tgg caa ggg tat ttt ggt gat cat	2467
Pro Pro Ser Arg Leu Gly Ile Arg Trp Gln Gly Tyr Phe Gly Asp His	
775 780 785	
tgg tct gcc aat gca gaa ttt aac cat gtg ttt gca caa aat aaa gtt	2515
Trp Ser Ala Asn Ala Glu Phe Asn His Val Phe Ala Gln Asn Lys Val	
790 795 800 805	
gcc acc tca acg gtt gcc att aaa cct caa ttc aag cag cca gaa ggt	2563
Ala Thr Ser Thr Val Ala Ile Lys Pro Gln Phe Lys Gln Pro Glu Gly	
810 815 820	
tgc caa cgc cat gag agt cat tgc cga atc agc gac tat ggc agt gat	2611
Cys Gln Arg His Glu Ser His Cys Arg Ile Ser Asp Tyr Gly Ser Asp	
825 830 835	
aac aac cct ttg atg atg cag cca aga tat atc aca gaa aac aaa acg	2659
Asn Asn Pro Leu Met Met Gln Pro Arg Tyr Ile Thr Glu Asn Lys Thr	
840 845 850	

155/189

gca gga tat aat ttg ctc aat gtt ggc tta gat tat aac aat gca tat	2707
Ala Gly Tyr Asn Leu Leu Asn Val Gly Leu Asp Tyr Asn Asn Ala Tyr	
855 860 865	
cgt aat gtt gat tat acg ctg tca att cgt gcg aat aat tta ctt aat	2755
Arg Asn Val Asp Tyr Thr Leu Ser Ile Arg Ala Asn Asn Leu Leu Asn	
870 875 880 885	
gaa caa atc tac att cac aac tca ttt ttg ccg ttt gta ccg cag atg	2803
Glu Gln Ile Tyr Ile His Asn Ser Phe Leu Pro Phe Val Pro Gln Met	
890 895 900	
ggg cgt aat ctg act tta ggt ttg acg gct aaa ttt tga tagtacatcg	2852
Gly Arg Asn Leu Thr Leu Gly Leu Thr Ala Lys Phe	
905 910	
accatagctt caaatacatt tgatatcaag gacgatatca aatgtatttg accaaagaag	2912
ggcgttatct tgataaaagt tgggcgaatg	2942

156/189

Figure 50. Restriction enzyme analysis of the *M. catarrhalis* heme receptor gene designated MCA101976



157/189

BslI BsgI | AcilI | CviRI | MnlI | MnlI | HgaI | MseI | AflIII | SmlI |
 TTTTGGTGGCGGTTTCATCTGCACCTGTTGTGCGAGGGCAAGAGGGTGTGCGTCTTAAGAT
 361 -----+-----+-----+-----+-----+-----+-----+ 420
 AAAACCACCGCCAAGTAGACGTGGACAACACGCTCCCGTTCTCCACACGCAGAATTCTA

 NlaIII
 DpnI |
 BclI |
 MboI |
 TTTACAAAATGGAAGTGTGATTGATGTGTCATCAATATCGCCTGATCATGTTGTGGC
 421 -----+-----+-----+-----+-----+-----+-----+ 480
 AAATGTTTTACCTTGACTACACTAACTACACAGTAGTTATAGCGGACTAGTACAACACCG

 HgaI | AluI |
 BsiBI | CviJI |
 GACCGATACACTTTTAGCGTCTAAAGTTGAGCTTGTTTCGTGGTGCTGATACGCTGTTATA
 481 -----+-----+-----+-----+-----+-----+-----+ 540
 CTGGCTATGTGAAAATCGCAGATTTCAACTCGAACAAGCACCACGACTATGCGACAATAT

 AluI
 CviJI
 MspAI
 PvuII
 SfaNI
 CviJI | Cac8I | AseI | Bst4CI | HinfI | BciVI |
 MwoI | MseI | BcgI | TfiI |
 TGGCTTGGCATCGCCAGCTGGTGTGATTAATGTTGTTGATGACCGTATCCCGAATCGTAT
 541 -----+-----+-----+-----+-----+-----+-----+ 600
 ACCGAACCGTAGCGTTCGACCACACTAATTACAACAACACTACTGGCATAGGGCTTAGCATA

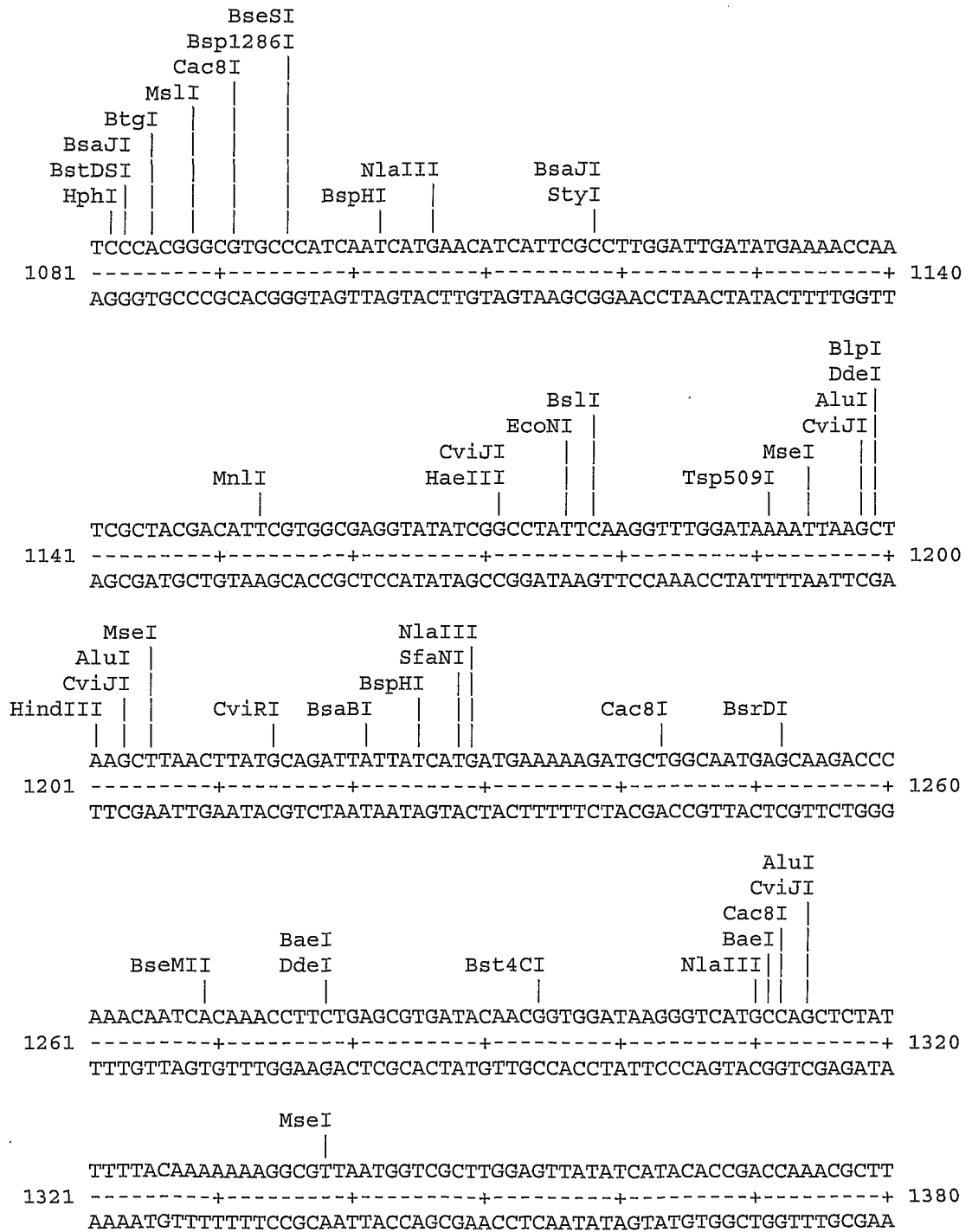
 SfaNI
 BsmAI |
 BsmBI |
 BfaI | BcgI | NlaIII | TaqI | TaqI |
 GCCTAGTGGTGTCTATCCATGACAAAATCGAAGGCGAGACGATGCTTCGATATAACACAAA
 601 -----+-----+-----+-----+-----+-----+-----+ 660
 CGGATCACCACGATAGGTACTGTTTTAGCTTCCGCTCTGCTACGAAGCTATATTGTGTTT

 CviRI
 Cac8I |
 AluI |
 CviJI | PstI | AluI | BsrDI |
 HindIII | SfcI | CviJI | HphI | MboI | AcilI |
 NlaIII |
 CAACCATGAAAAGCTTGCAACTGCAGGGGTGAGCTTTGGGGTAGGAGATCGCATTGCGGT
 661 -----+-----+-----+-----+-----+-----+-----+ 720
 GTTGGTACTTTTCGAACGTTGACGTCCCCACTCGAAACCCCATCCTCTAGCGTAACGCCA

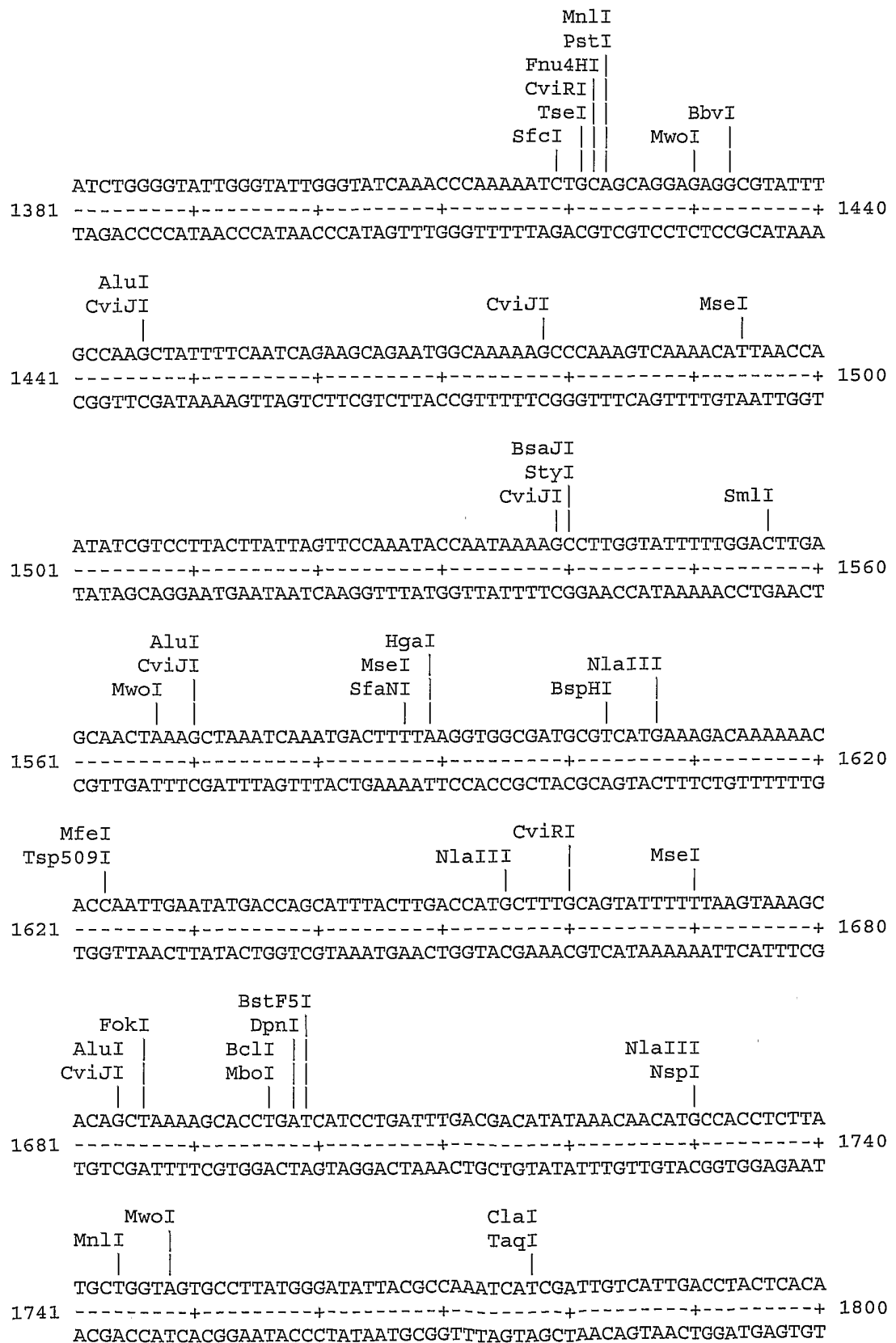
158/189



159/189



160/189



161/189

[illegible]

Bst4CI | SfaNI | HphI | MseI |
 TATGACCGTTGATGATAATGGTGATTATGATGCCGATGGTGGTTTGACTTGTGCCTTAAA
 2221 -----+-----+-----+-----+-----+-----+-----+ 2280
 ATACTGGCAACTACTATTACCACTAATACTACGGCTACCACCAAACGAACACGGAATTT

 Tsp509I | MseI | BsmFI | RsaI | Csp6I | AseI |
 AACGCCTGAACAATGGGGACAAATTAACGATAATAATGATTGTAGTACAACCATTAATGT
 2281 -----+-----+-----+-----+-----+-----+-----+ 2340
 TTGCGGACTTGTTACCCCTGTTTAATTGCTATTATTACTAACATCATGTTGGTAATTACA

 RsaI | Bsu36I | MnlI | BseRI |
 Csp6I | DdeI | BsmII | Bst4CI |
 Bst4CI | MnlI | MnlI | BsiEI |
 CTATAAAAACGGTACAACAACCTCAGGCGAGGAGGATTACGACCGTTTGGCACGAAATCC
 2341 -----+-----+-----+-----+-----+-----+-----+ 2400
 GATATTTTTTGCCATGTTGTTGGAGTCCGCTCCTCCTAATGCTGGCAAACCGTGCTTTAGG

 BsmI |
 BglI |
 MwoI |
 BstXI |
 AciI | AciI |
 RsaI | Fnu4HI |
 Csp6I | CviJI |
 AACTTACGCACCCAGAGTACCGCCAGCCGCTTGGGCATTTCGTTGGCAAGGGTATTTTGG
 2401 -----+-----+-----+-----+-----+-----+-----+ 2460
 TTGAATGCGTGCGTCTCATGGCGGGTCGGCGAACCCGTAAGCAACCGTTCCTATAAAACC

 DpnI | ApoI |
 BclI | Tsp509I |
 MboI | HphI | CviRI | MseI | NlaIII | CviRI |
 TGATCATTGGTCTGCCAATGCAGAATTTAACCATGTGTTTGCACAAAATAAAGTTGCCAC
 2461 -----+-----+-----+-----+-----+-----+-----+ 2520
 ACTAGTAACCAGACGGTTACGTCTTAAATTGGTACACAAACGTGTTTTATTTCACGGTG

 CviJI |
 Fnu4HI |
 TseI |
 MnlI |
 Bst4CI | MseI | Tsp509I | BbvI | HinfI |
 NlaIII |
 CTCAACGGTTGCCATTAAACCTCAATTCAAGCAGCCAGAAGGTTGCCAACGCCATGAGAG
 2521 -----+-----+-----+-----+-----+-----+-----+ 2580
 GAGTTGCCAACGGTAATTTGGAGTTAAGTTCGTCGGTCTTCCAACGGTTGCGGTACTCTC

163/189

MwoI
 HinfI
 TfiI
 MlyI
 BsrDI
 PleI
 MwoI
 BtsI
 TspRI
 SfaNI
 CviJI
 Fnu4HI
 CviRI
 TseI
 2581 TCATTGCCGAATCAGCGACTATGGCAGTGATAACAACCCTTTGATGATGCAGCCAAGATA
 -----+-----+-----+-----+-----+-----+-----+ 2640
 AGTAACGGCTTAGTCGCTGATACCGTCACTATTGTTGGGAACTACTACGTCGGTTCTAT

BbvI
 Tsp509I
 DdeI
 CviJI
 PsiI
 Ppu10I
 2641 TATCACAGAAAACAAAACGGCAGGATATAATTTGCTCAATGTTGGCTTAGATTATAACAA
 -----+-----+-----+-----+-----+-----+-----+ 2700
 ATAGTGTCTTTTGTGTTTGCCGTCCTATATTAAACGAGTTACAACCGAATCTAATATTGTT

NsiI
 CviRI
 Tsp509I
 Tsp509I
 MseI
 2701 TGCATATCGTAATGTTGATTATACGCTGTCAATTCGTGCGAATAATTTACTTAATGAACA
 -----+-----+-----+-----+-----+-----+-----+ 2760
 ACGTATAGCATTACAATAATATGCGACAGTTAAGCACGCTTATTAAATGAATTACTTGT

AciI
 RsaI
 Csp6I
 MwoI
 BslI
 2761 AATCTACATTCAAACTCATTTTTTGCCGTTTGTACCGCAGATGGGGCGTAATCTGACTTT
 -----+-----+-----+-----+-----+-----+-----+ 2820
 TTAGATGTAAGTGTGAGTAAAAACGGCAAACATGGCGTCTACCCCGCATTAGACTGAAA

ApoI
 Tsp509I
 CviJI
 RsaI
 Csp6I
 TatI
 TaqI
 AluI
 CviJI
 EcoRV
 2821 AGGTTTGACGGCTAAATTTTGATAGTACATCGACCATAGCTTCAAATACATTTGATATCA
 -----+-----+-----+-----+-----+-----+-----+ 2880
 TCCAAACTGCCGATTTAAACTATCATGTAGCTGGTATCGAAGTTTATGTAAACTATAGT

EcoRV
 BslI
 2881 AGGACGATATCAAATGTATTTGACCAAAGAAGGGCGTTATCTTGATAAAAGTTGGGCGAA
 -----+-----+-----+-----+-----+-----+-----+ 2940
 TCCTGCTATAGTTTACATAAACTGGTTTCTTCCGCAATAGAACTATTTCAACCCGCTT

TG
 2941 -- 2942
 AC

164/189

Figure 51. Sequence of *M. catarrhalis* ompC/heme receptor gene designated MCA102062

```

taattacttg cttaaattgt atctcaactg attgtaattt ttcatatatg ttatataata 60

cttcaacaaa tgttatagta tatcatatat gaagggttacc atg ata aaa aaa ccg 115
                                         Met Ile Lys Lys Pro
                                         1           5

ctt gcc tgt gcc ata ttg gca act ttt tca atg cca atg ctg gca gag 163
Leu Ala Cys Ala Ile Leu Ala Thr Phe Ser Met Pro Met Leu Ala Glu
                      10                      15                      20

gcg aat tta aag gat aag cca acc gtc att tta gat ggc gtt tcg atc 211
Ala Asn Leu Lys Asp Lys Pro Thr Val Ile Leu Asp Gly Val Ser Ile
                      25                      30                      35

acc tct tta gct gac caa aat aca gag ttt ggc gtt aat cat tca aaa 259
Thr Ser Leu Ala Asp Gln Asn Thr Glu Phe Gly Val Asn His Ser Lys
                      40                      45                      50

aca gtc agt ggc atc aca gtt tca aaa gag caa cta caa caa cga gca 307
Thr Val Ser Gly Ile Thr Val Ser Lys Glu Gln Leu Gln Gln Arg Ala
                      55                      60                      65

acc acc cta ggc gat gcc ttg gca ggt gag ctt ggc gtt cat tct aac 355
Thr Thr Leu Gly Asp Ala Leu Ala Gly Glu Leu Gly Val His Ser Asn
                      70                      75                      80                      85

cat ttt ggg ggc ggt gcc tca gcc ccc atc att cgt ggg cag gag ggt 403
His Phe Gly Gly Gly Ala Ser Ala Pro Ile Ile Arg Gly Gln Glu Gly
                      90                      95                      100

aaa cgc ctg aaa atc cta caa aac ggt tca gag gtt gtg gac atg tct 451
Lys Arg Leu Lys Ile Leu Gln Asn Gly Ser Glu Val Val Asp Met Ser
                      105                      110                      115

ggg ttg tcg cca gac cat gcc ata gcg gtg gac acc aca ctg gca aaa 499
Gly Leu Ser Pro Asp His Ala Ile Ala Val Asp Thr Thr Leu Ala Lys
                      120                      125                      130

cag gtg gag att gtg cga ggc tct ggt gcc ttg ttg tac gcc tct ggc 547
Gln Val Glu Ile Val Arg Gly Ser Gly Ala Leu Leu Tyr Ala Ser Gly
                      135                      140                      145

aac tca gca ggc gtg gtc aat gtc gtt gat gac aaa ata ccc agc aaa 595
Asn Ser Ala Gly Val Val Asn Val Val Asp Asp Lys Ile Pro Ser Lys
150                      155                      160                      165

ttg ccc agc aaa tta caa ggt gat gtg acg gtg cgt ctt agc agt gcc 643
Leu Pro Ser Lys Leu Gln Gly Asp Val Thr Val Arg Leu Ser Ser Ala
                      170                      175                      180

```

165/189

aac cgt gaa aaa tta atc acc gcc agt gcc gaa gcc cca ctg gga gag	691
Asn Arg Glu Lys Leu Ile Thr Ala Ser Ala Glu Ala Pro Leu Gly Glu	
185 190 195	
cat gtg gca gtg cgt gtt gca ggg ctg tcc aaa caa gca gca gac tat	739
His Val Ala Val Arg Val Ala Gly Leu Ser Lys Gln Ala Ala Asp Tyr	
200 205 210	
aaa acg cca cgc ttt gac cgc cat gtc ttt aac aaa aaa cat gaa gat	787
Lys Thr Pro Arg Phe Asp Arg His Val Phe Asn Lys Lys His Glu Asp	
215 220 225	
gat aac act cag cca gaa ttc atc tat aaa gac acc tta aag cat ctg	835
Asp Asn Thr Gln Pro Glu Phe Ile Tyr Lys Asp Thr Leu Lys His Leu	
230 235 240 245	
cca gac agc cat gcc aaa tca aac gca gga acg ctt ggc gtg tca tgg	883
Pro Asp Ser His Ala Lys Ser Asn Ala Gly Thr Leu Gly Val Ser Trp	
250 255 260	
gtt ggc aat caa ggc ttt ttg ggg gca tgc gtg agc tta cgc cga gac	931
Val Gly Asn Gln Gly Phe Leu Gly Ala Ser Val Ser Leu Arg Arg Asp	
265 270 275	
aaa tat ggc ctg ccc aac cat tca cat gaa tat gaa gaa tgt agc gtg	979
Lys Tyr Gly Leu Pro Asn His Ser His Glu Tyr Glu Glu Cys Ser Val	
280 285 290	
cat ggc att tct cag tcc gcc tta caa tac aag cca tat ttg cgt ttg	1027
His Gly Ile Ser Gln Ser Ala Leu Gln Tyr Lys Pro Tyr Leu Arg Leu	
295 300 305	
tat cct ttt ttg atg gaa aat gat gac tta gag ttt gac aat gca ggg	1075
Tyr Pro Phe Leu Met Glu Asn Asp Asp Leu Glu Phe Asp Asn Ala Gly	
310 315 320 325	
ctt gaa tgc cat aca cat gat gac cac gac cac gag cac gac cat gct	1123
Leu Glu Cys His Thr His Asp Asp His Asp His Glu His Asp His Ala	
330 335 340	
cat gac cac gag cac gac cac gag cac gac cat ggc aaa cct tgg att	1171
His Asp His Glu His Asp His Glu His Asp His Gly Lys Pro Trp Ile	
345 350 355	
gat ttg aaa atg aag cgt tat gat gtg caa ggg caa atc aat gcg ccg	1219
Asp Leu Lys Met Lys Arg Tyr Asp Val Gln Gly Gln Ile Asn Ala Pro	
360 365 370	
ttt gct ggc att gat aaa atc cga gcc agc atg ggc aaa gtg gat tat	1267
Phe Ala Gly Ile Asp Lys Ile Arg Ala Ser Met Gly Lys Val Asp Tyr	
375 380 385	
cat cat gat gag ata gat ggg ggt gag aag acc agc ttt ttt gat aat	1315
His His Asp Glu Ile Asp Gly Gly Glu Lys Thr Ser Phe Phe Asp Asn	
390 395 400 405	

166/189

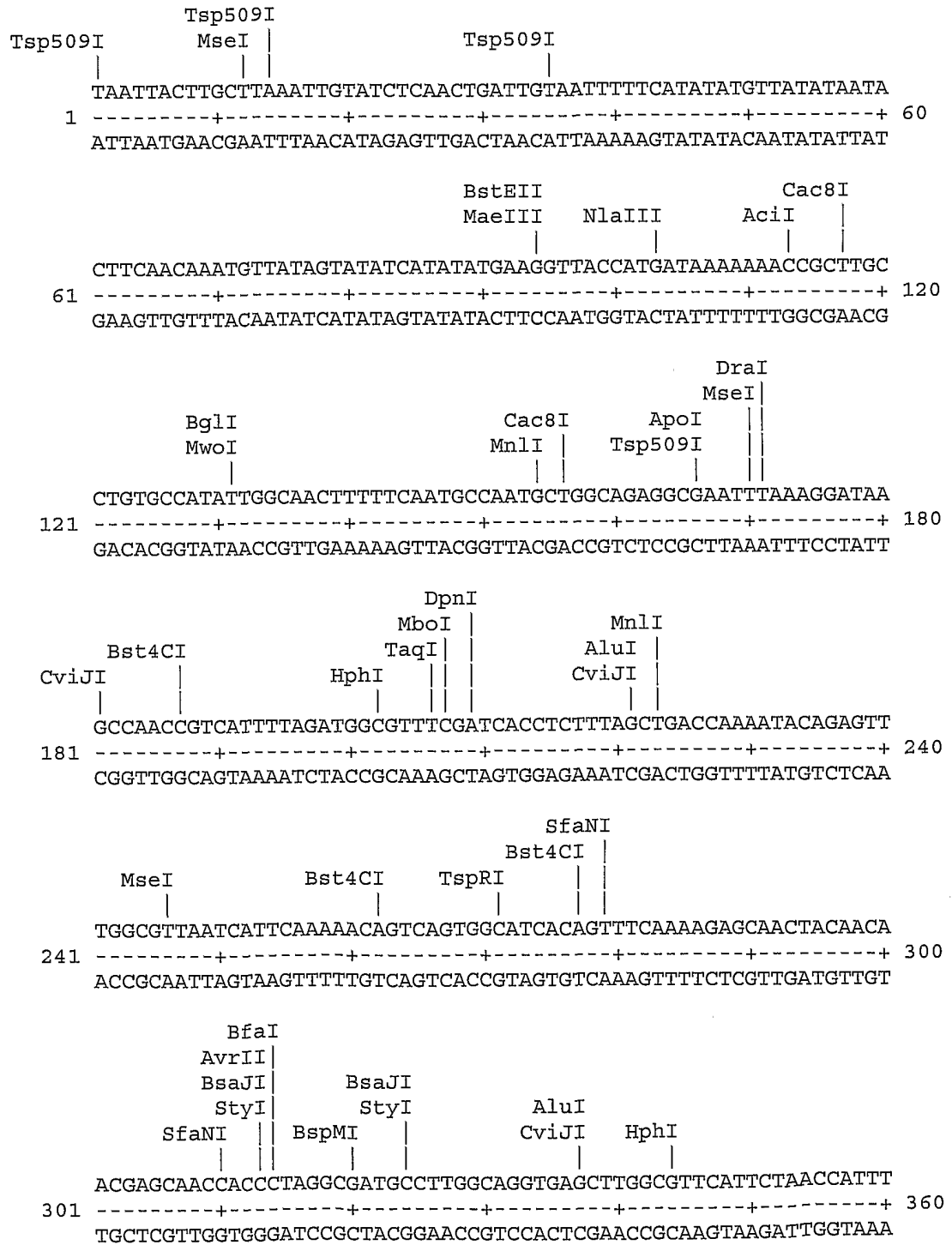
caa gcc aat gtg tgg cgt ctg gaa gcc tca cat acc ccc att cat acg	1363
Gln Ala Asn Val Trp Arg Leu Glu Ala Ser His Thr Pro Ile His Thr	
410 415 420	
ccg atg ggc aag ttt agc ggg gtg ttt ggg gta ggt tat ctc acc tca	1411
Pro Met Gly Lys Phe Ser Gly Val Phe Gly Val Gly Tyr Leu Thr Ser	
425 430 435	
aaa aac agc gga ctt gtg cca cct cgt tat gaa gat ggc aat aaa caa	1459
Lys Asn Ser Gly Leu Val Pro Pro Arg Tyr Glu Asp Gly Asn Lys Gln	
440 445 450	
gac acc caa aac atc ttg cac aat aat aaa acc aaa aca ggc agt gtg	1507
Asp Thr Gln Asn Ile Leu His Asn Asn Lys Thr Lys Thr Gly Ser Val	
455 460 465	
ttt tgg ttt gaa gaa tac aaa ccc aat gac aag ctg acc gtt gac gcc	1555
Phe Trp Phe Glu Glu Tyr Lys Pro Asn Asp Lys Leu Thr Val Asp Ala	
470 475 480 485	
gcc gct cgc att gag aaa caa acc atc acc atg gat tat gat aaa gac	1603
Ala Ala Arg Ile Glu Lys Gln Thr Ile Thr Met Asp Tyr Asp Lys Asp	
490 495 500	
gcc att tat cag agc tta aac tta ggc tta gca acc gct cat gaa cca	1651
Ala Ile Tyr Gln Ser Leu Asn Leu Gly Leu Ala Thr Ala His Glu Pro	
505 510 515	
gac ata cgc ttt aaa cga ttg ctg gac agc ggt act tta aac ccc aaa	1699
Asp Ile Arg Phe Lys Arg Leu Leu Asp Ser Gly Thr Leu Asn Pro Lys	
520 525 530	
aaa caa acc gca cgc tct tat gct gtt ggg acg cat tta caa tta acg	1747
Lys Gln Thr Ala Arg Ser Tyr Ala Val Gly Thr His Leu Gln Leu Thr	
535 540 545	
ccc aaa cat aaa tta tcg ctg aat ctg tcg cat caa gaa cgc ctg cca	1795
Pro Lys His Lys Leu Ser Leu Asn Leu Ser His Gln Glu Arg Leu Pro	
550 555 560 565	
aat gct cag gaa ttg tat gct cac ggc atg cac ctt gcc acc aac tcg	1843
Asn Ala Gln Glu Leu Tyr Ala His Gly Met His Leu Ala Thr Asn Ser	
570 575 580	
ttt gaa att ggc aac cgc ttt tta aac aaa gaa aaa tcc aac aac att	1891
Phe Glu Ile Gly Asn Arg Phe Leu Asn Lys Glu Lys Ser Asn Asn Ile	
585 590 595	
gat ttg ggg ctg aca ttt caa ggt gat aaa tgg gat tat cgt ctt ggg	1939
Asp Leu Gly Leu Thr Phe Gln Gly Asp Lys Trp Asp Tyr Arg Leu Gly	
600 605 610	
ggc tat cat tat gat ttt gat aac tat gtg ttt tta caa aca ttg tcg	1987
Gly Tyr His Tyr Asp Phe Asp Asn Tyr Val Phe Leu Gln Thr Leu Ser	
615 620 625	

167/189

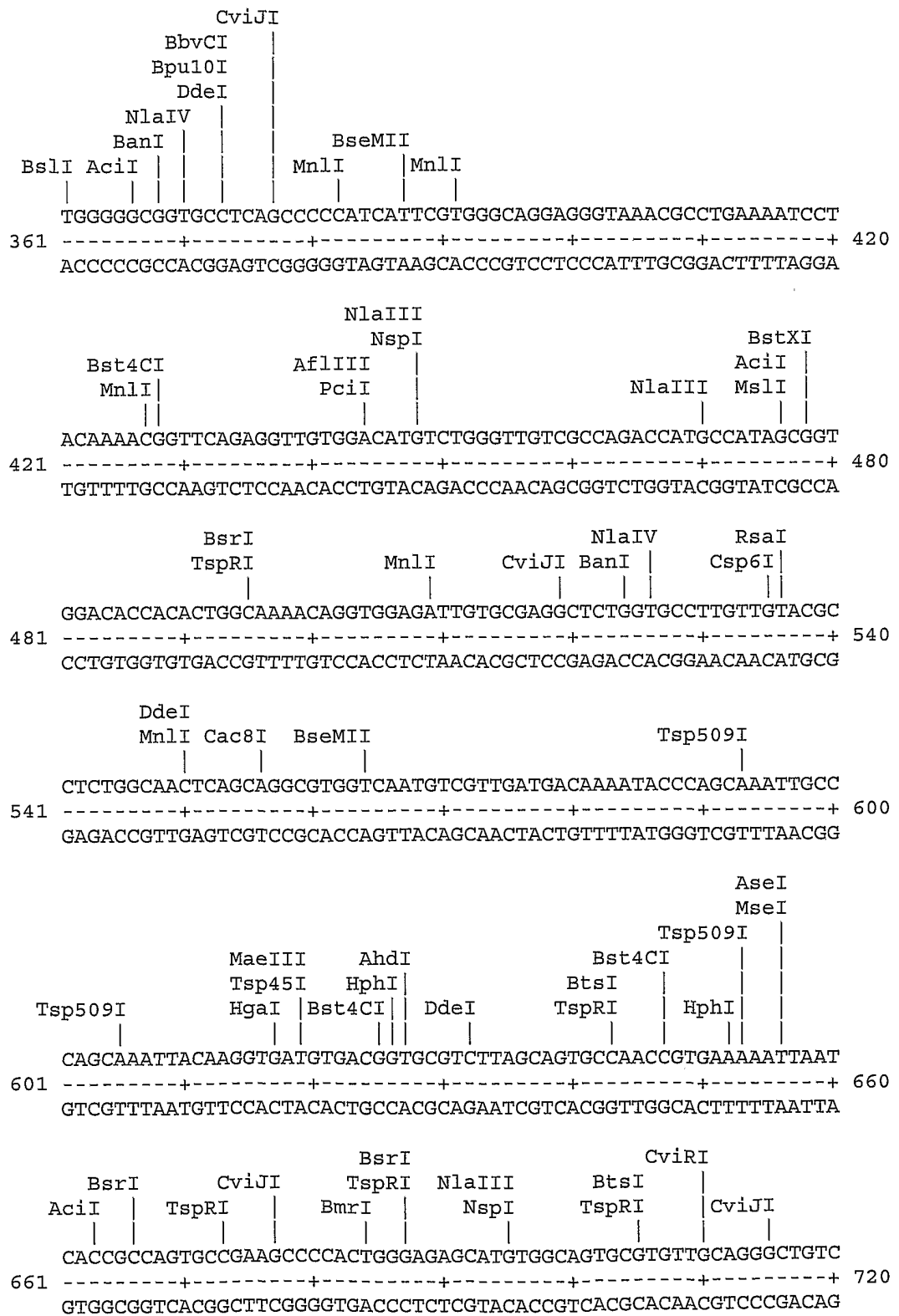
cag tat aag caa ggt ttg cgt ggc atg cgt cat gat aaa gac tta aaa	2035
Gln Tyr Lys Gln Gly Leu Arg Gly Met Arg His Asp Lys Asp Leu Lys	
630 635 640 645	
acc gca cgc tat gaa caa gca gcg gcg aaa ttt tat gga ttt gat gtc	2083
Thr Ala Arg Tyr Glu Gln Ala Ala Ala Lys Phe Tyr Gly Phe Asp Val	
650 655 660	
aac atc ggt tat cag att aat gat gta tat cat gtg gcg tta ttt ggt	2131
Asn Ile Gly Tyr Gln Ile Asn Asp Val Tyr His Val Ala Leu Phe Gly	
665 670 675	
gat tat att cgt ggc aag ctc acc aat ttg cct gac aaa aag ggc aga	2179
Asp Tyr Ile Arg Gly Lys Leu Thr Asn Leu Pro Asp Lys Lys Gly Arg	
680 685 690	
acc gat gcg tat ggc aac cgt cct ctc atc aaa cag cca gac agt cat	2227
Thr Asp Ala Tyr Gly Asn Arg Pro Leu Ile Lys Gln Pro Asp Ser His	
695 700 705	
acg cca aga ctg cca cca aaa cgc ctt ggc atg aaa tta acc gcc aat	2275
Thr Pro Arg Leu Pro Pro Lys Arg Leu Gly Met Lys Leu Thr Ala Asn	
710 715 720 725	
gtt aat gca aat tgg tca ggg ttt ttg gaa tat cgc cat acc ttt aaa	2323
Val Asn Ala Asn Trp Ser Gly Phe Leu Glu Tyr Arg His Thr Phe Lys	
730 735 740	
caa gat aaa ttg gcg aat ttt gaa cgc cca acc cca gct cat aac ttg	2371
Gln Asp Lys Leu Ala Asn Phe Glu Arg Pro Thr Pro Ala His Asn Leu	
745 750 755	
gtg aat ttg ggg ctt aac tat cag cac aag cca agc cat caa gca ggc	2419
Val Asn Leu Gly Leu Asn Tyr Gln His Lys Pro Ser His Gln Ala Gly	
760 765 770	
tcg gtt cag gta ttt ttt aat gct aac aat cta tta aac gat aaa gtc	2467
Ser Val Gln Val Phe Phe Asn Ala Asn Asn Leu Leu Asn Asp Lys Val	
775 780 785	
ttt gct cat gag aca ttt ttc cca gac atg cca caa atg ggg cga aac	2515
Phe Ala His Glu Thr Phe Phe Pro Asp Met Pro Gln Met Gly Arg Asn	
790 795 800 805	
ttt atg ctc ggg gca aat ttc aaa ttt tga caaatcattt aactaaaggc	2565
Phe Met Leu Gly Ala Asn Phe Lys Phe	
810 815	
aatcaatata tgcttgact aaactattat attgattaaa tgtgatattc attaaaaatc	2625
tgcatcaca attaaggaat t	2646

168/189

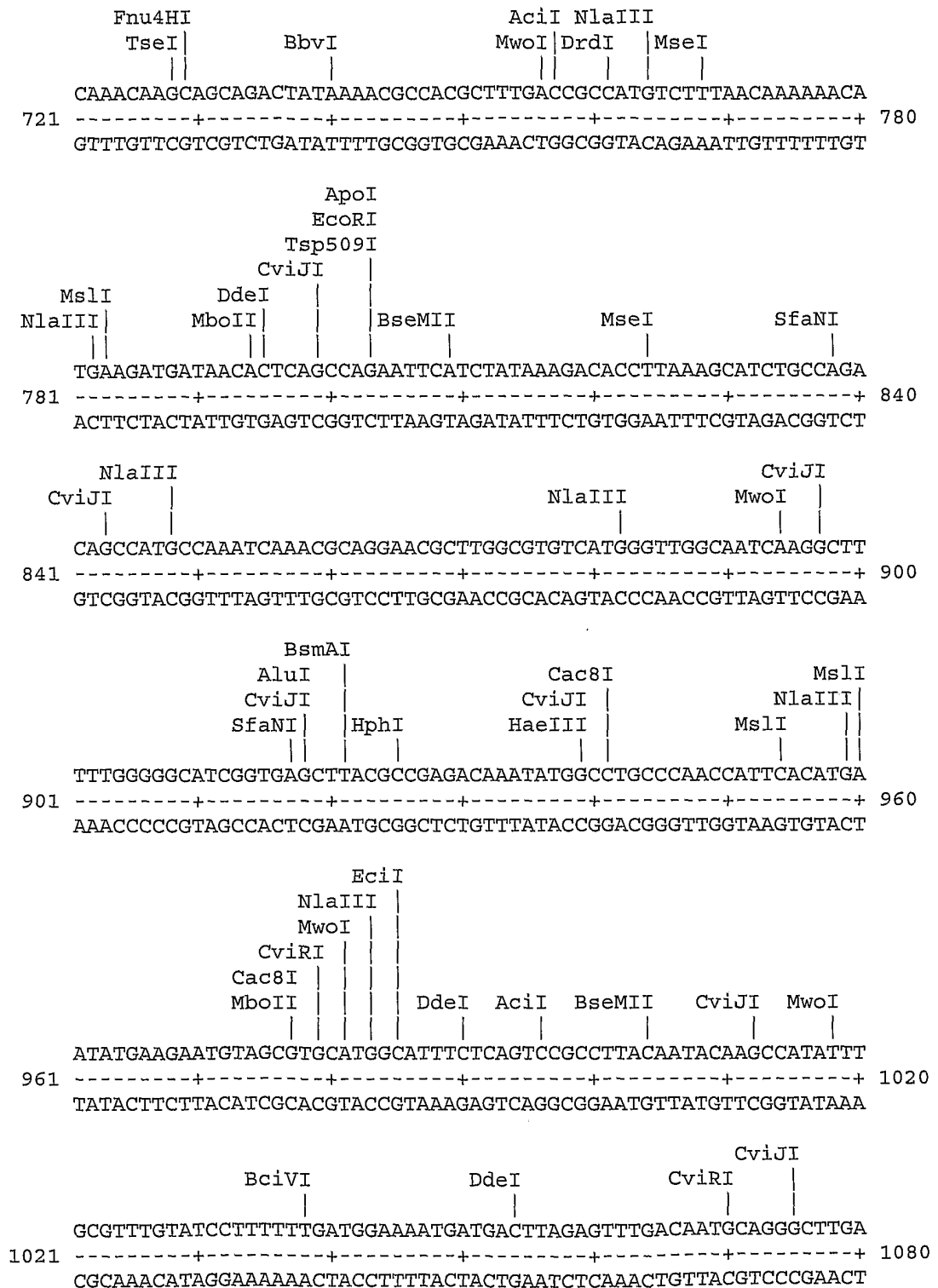
Figure 52. Restriction enzyme analysis of the *M. catarrhalis* ompC/heme receptor gene designated MCA102062



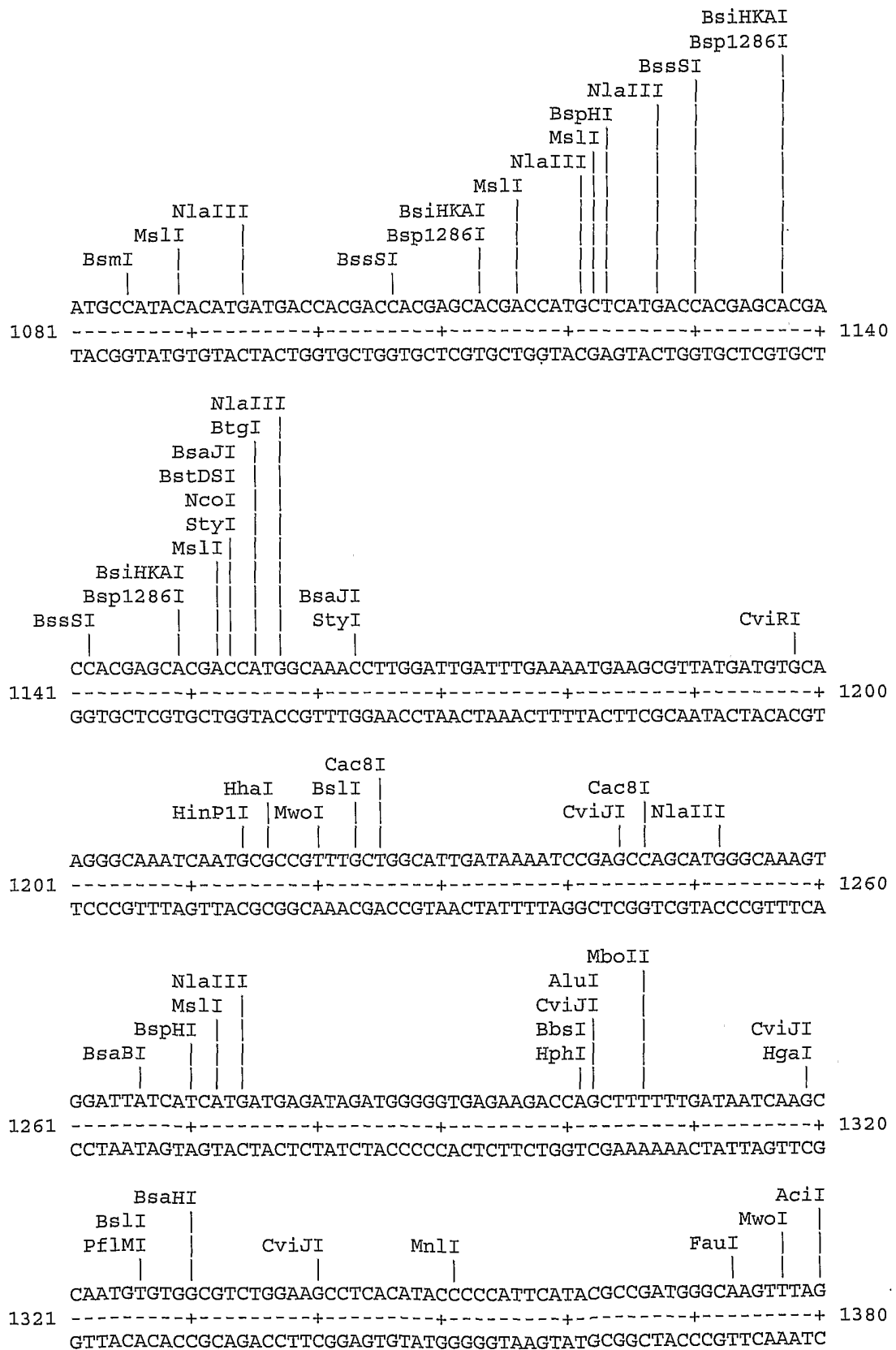
169/189



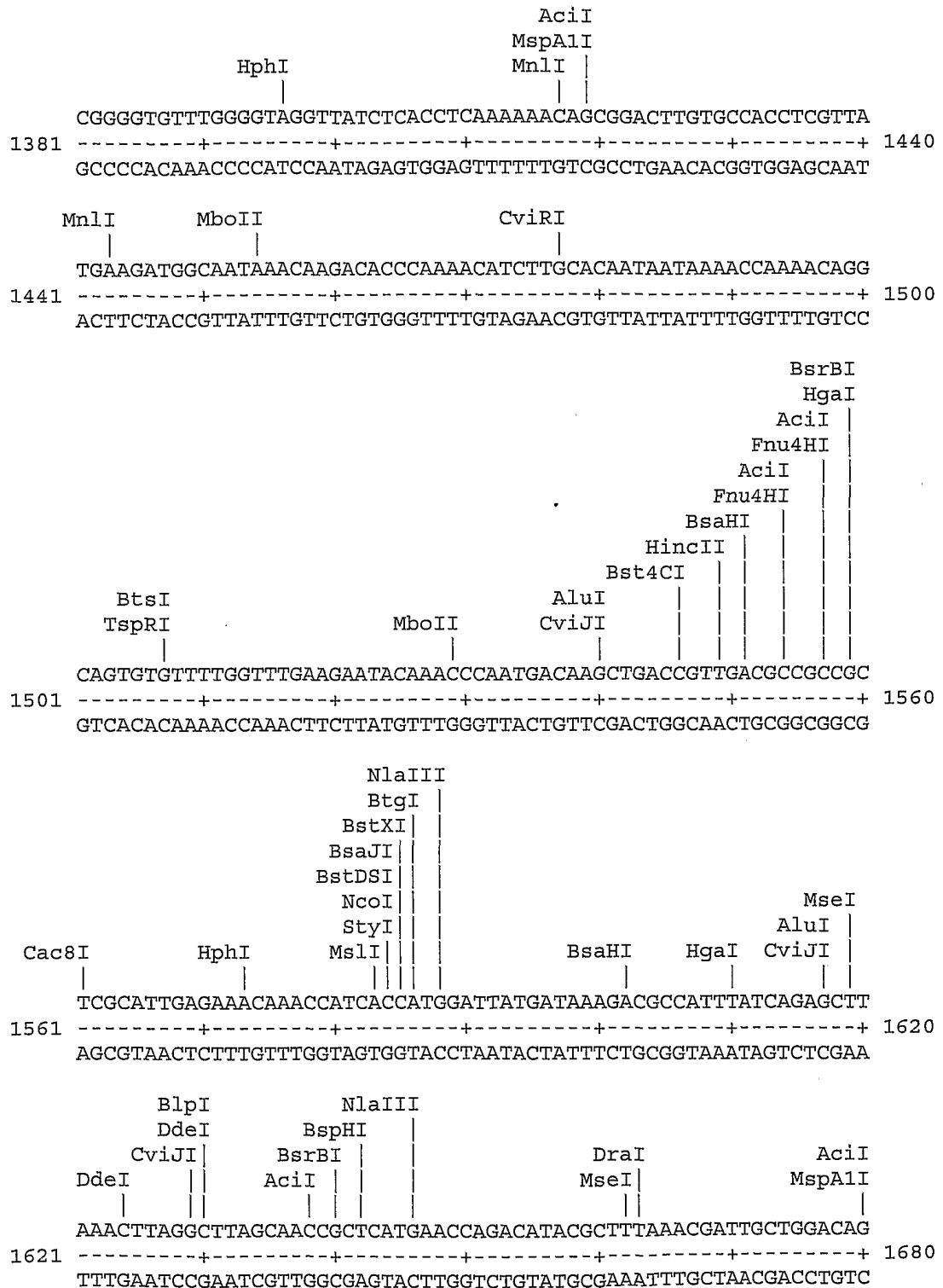
170/189



171/189



172/189



173/189

RsaI DraI
 Csp6I MseI
 Cac8I
 Tsp509I
 HgaI
 1681 -----+-----+-----+-----+-----+ 1740
 CGGTACTTTAAACCCCAAAAAACAAACCGCACGCTCTTATGCTGTTGGGACGCATTTACA
 GCCATGAAATTTGGGGTTTTTTGTTTGGCGTGCGAGAATACGACAACCCTGCGTAAATGT

MseI
 BsmFI
 Tsp509I
 HinfI
 TfiI
 SfaNI
 Cac8I
 Bpu10I
 DdeI
 1741 -----+-----+-----+-----+-----+ 1800
 ATTAACGCCCAAACATAAATTATCGCTGAATCTGTTCGCATCAAGAACGCCTGCCAAATGC
 TAATTGCGGGTTTGTATTTAATAGCGACTTAGACAGCGTAGTTCTTGCGGACGGTTTACG

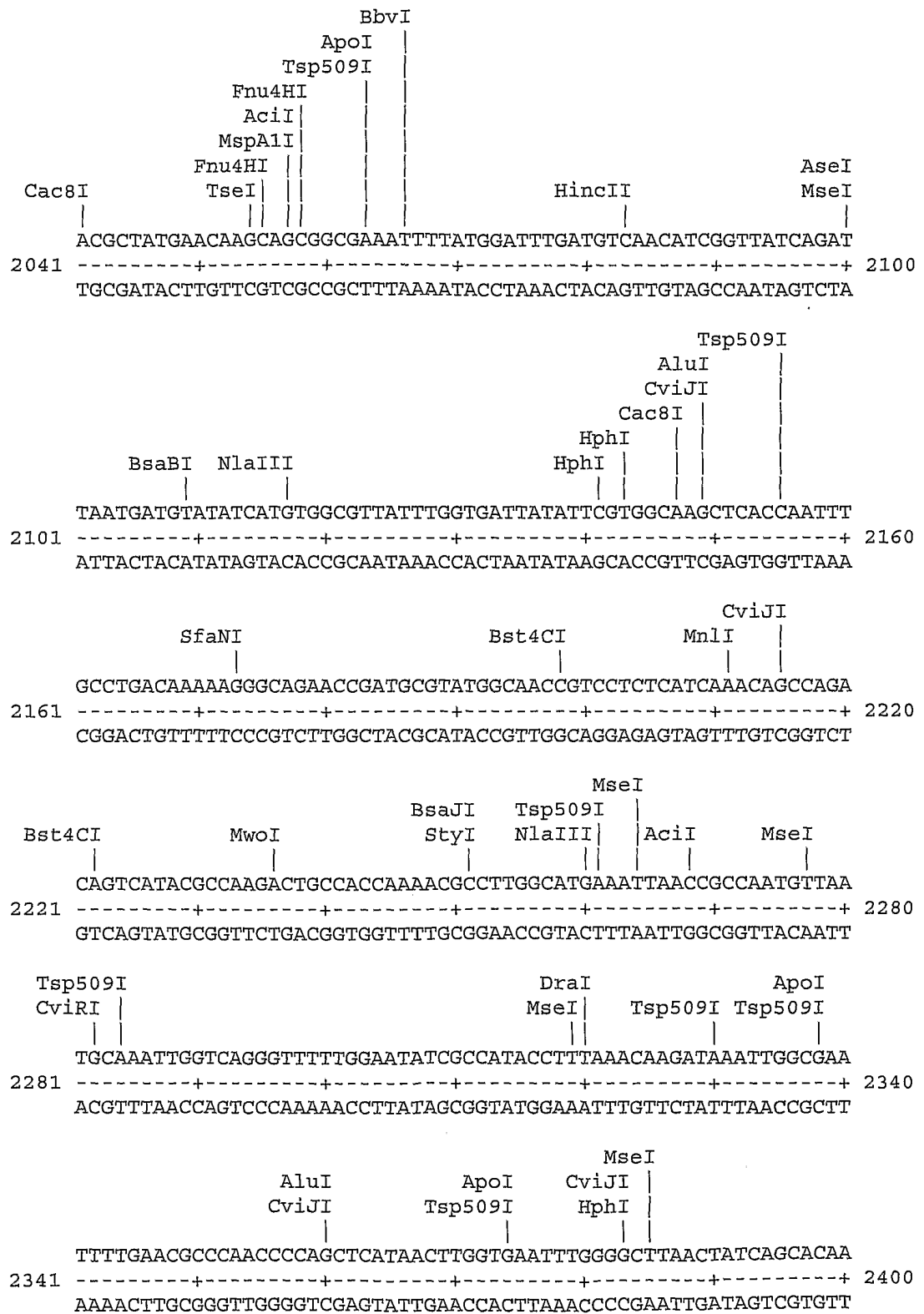
CviRI
 NlaIII
 NspI
 SphI
 Tsp509I BseMII Cac8I Tsp509I AcII
 1801 -----+-----+-----+-----+-----+ 1860
 TCAGGAATTGTATGCTCACGGCATGCACCTTGCCACCAACTCGTTTGAAATTGGCAACCG
 AGTCCTTAACATACGAGTGCCGTACGTGGAACGGTGGTTGAGCAAACCTTAACCGTTGGC

DraI
 MseI
 CviJI
 1861 -----+-----+-----+-----+-----+ 1920
 CTTTTTAAACAAAGAAAAATCCAACAACATTGATTTGGGGCTGACATTTCAAGGTGATAA
 GAAAAATTTGTTTCTTTTAGGTTGTTGTAACATAACCCCGACTGTAAAGTTCCACTATT

HphI
 CviJI
 1921 -----+-----+-----+-----+-----+ 1980
 ATGGGATTATCGTCTTGGGGGCTATCATTATGATTTTGATAACTATGTGTTTTTACAAAC
 TACCCTAATAGCAGAACCCCGATAGTAATACTAAAACCTATTGATACACAAAAATGTTTG

NlaIII
 BspHI
 NlaIII
 NspI
 SphI
 HgaI
 BstAPI
 MwoI
 MwoI
 Cac8I
 MwoI
 MseI
 AcII
 1981 -----+-----+-----+-----+-----+ 2040
 ATTGTCGCAGTATAAGCAAGGTTTGGCGTGGCATGCGTCATGATAAAGACTTAAAAACCGC
 TAACAGCGTCATATTCGTTCCAAACGCACCGTACGCAGTACTATTTCTGAATTTTGGCG

174/189



175/189

```

      CviJI      CviJI
      |          |
CviJI |          | Cac8I |          | MseI          | MseI
      |          |          |          |          |
GCCAAGCCATCAAGCAGGCTCGGTTCAGGTATTTTAAATGCTAACAATCTATTAAACGA
2401 -----+-----+-----+-----+-----+-----+ 2460
      CGGTTCGGTAGTTTCGTCCGAGCCAAGTCCATAAAAAATTACGATTGTTAGATAATTTGCT

      NlaIII
      |
      BspHI
      |
      BsmAI |          |
      |          |          |
TAAAGTCTTTGCTCATGAGACATTTTCCAGACATGCCACAAATGGGGCGAAACTTTAT
2461 -----+-----+-----+-----+-----+ 2520
      ATTCAGAAACGAGTACTCTGTAAAAAGGTCTGTACGGTGTTTACCCCGCTTTGAAATA

      ApoI
      |
      Tsp509I
      |
AvaI |          | ApoI |          | MseI          | TatI
      |          |          |          |          |          |
GCTCGGGGCAAATTTCAAATTTTGACAAATCATTTAACTAAAGGCAATCAATATATGCTT
2521 -----+-----+-----+-----+-----+ 2580
      CGAGCCCCGTTTAAAGTTTAAAACTGTTTAGTAAATTGATTCCGTTAGTTATATACGAA

      RsaI
      |
Csp6I |          | MseI          | MseI          | BsmI          | Tsp509I | MseI
      |          |          |          |          |          |          |          |
GTAATAACTATTATATTGATTAAATGTGATATTCATTAAAAATCTGCATTACAAATTAA
2581 -----+-----+-----+-----+-----+ 2640
      CATGATTTGATAATATAACTAATTTACACTATAAGTAATTTTAGACGTAAGTGTTAATT

Tsp509I
|
GGAATT
2641 ----- 2646
CCTTAA

```

176/189

Figure 53. Sequence of *M. catarrhalis* porin gene designated MCA102070

```

gcatagatga aactcaggct aaaagttcag cgataagcca aacttgatgg ctgaaaatct 60
atgataacca acctaggaaa cttcaatcaa ggagagatgt atg aaa gca caa ttt 115
                                         Met Lys Ala Gln Phe
                                         1                               5

aca acc cta aaa gtc atc ggc gga cta ttt gta ggt ggc atg atg cta 163
Thr Thr Leu Lys Val Ile Gly Gly Leu Phe Val Gly Gly Met Met Leu
                               10                               15                               20

aat gca cat gca gca cct gaa tta tat ggt tac gca aat ctt gtc gtt 211
Asn Ala His Ala Ala Pro Glu Leu Tyr Gly Tyr Ala Asn Leu Val Val
                               25                               30                               35

tct gct aat cat aaa aaa aca acc aac aag aca acc aat gtc agt aca 259
Ser Ala Asn His Lys Lys Thr Thr Asn Lys Thr Thr Asn Val Ser Thr
                               40                               45                               50

tca aca aca gac cgc ccc tat ctg tat agc agc gga tca cgc att ggc 307
Ser Thr Thr Asp Arg Pro Tyr Leu Tyr Ser Ser Gly Ser Arg Ile Gly
                               55                               60                               65

ctt aga ggg tct gaa aag tta aat gat gat tat gag gtt ttg tac aag 355
Leu Arg Gly Ser Glu Lys Leu Asn Asp Asp Tyr Glu Val Leu Tyr Lys
                               70                               75                               80                               85

ttg gag tac cgt cta gaa aat gat ggt gat tta cgc aat gaa aaa atc 403
Leu Glu Tyr Arg Leu Glu Asn Asp Gly Asp Leu Arg Asn Glu Lys Ile
                               90                               95                               100

aag cag cca gat ggt act gaa aag act gtt gct aaa aca cgc aat ttt 451
Lys Gln Pro Asp Gly Thr Glu Lys Thr Val Ala Lys Thr Arg Asn Phe
                               105                               110                               115

gag gct cgt gat tca tgg att ggc gtg aag cat aaa aag tat ggt acc 499
Glu Ala Arg Asp Ser Trp Ile Gly Val Lys His Lys Lys Tyr Gly Thr
                               120                               125                               130

atc aag gcg ggt cgt atg ttg tct ttg gat cca tat gtg cgt tat act 547
Ile Lys Ala Gly Arg Met Leu Ser Leu Asp Pro Tyr Val Arg Tyr Thr
                               135                               140                               145

gcc tat ttg gcg tca ggc gta gat ggg gtg cgt acc aat aat acc att 595
Ala Tyr Leu Ala Ser Gly Val Asp Gly Val Arg Thr Asn Asn Thr Ile
                               150                               155                               160                               165

gca tac gaa tca cca aaa atc aaa gat gtt agt ttt cag gcg atg tac 643
Ala Tyr Glu Ser Pro Lys Ile Lys Asp Val Ser Phe Gln Ala Met Tyr
                               170                               175                               180

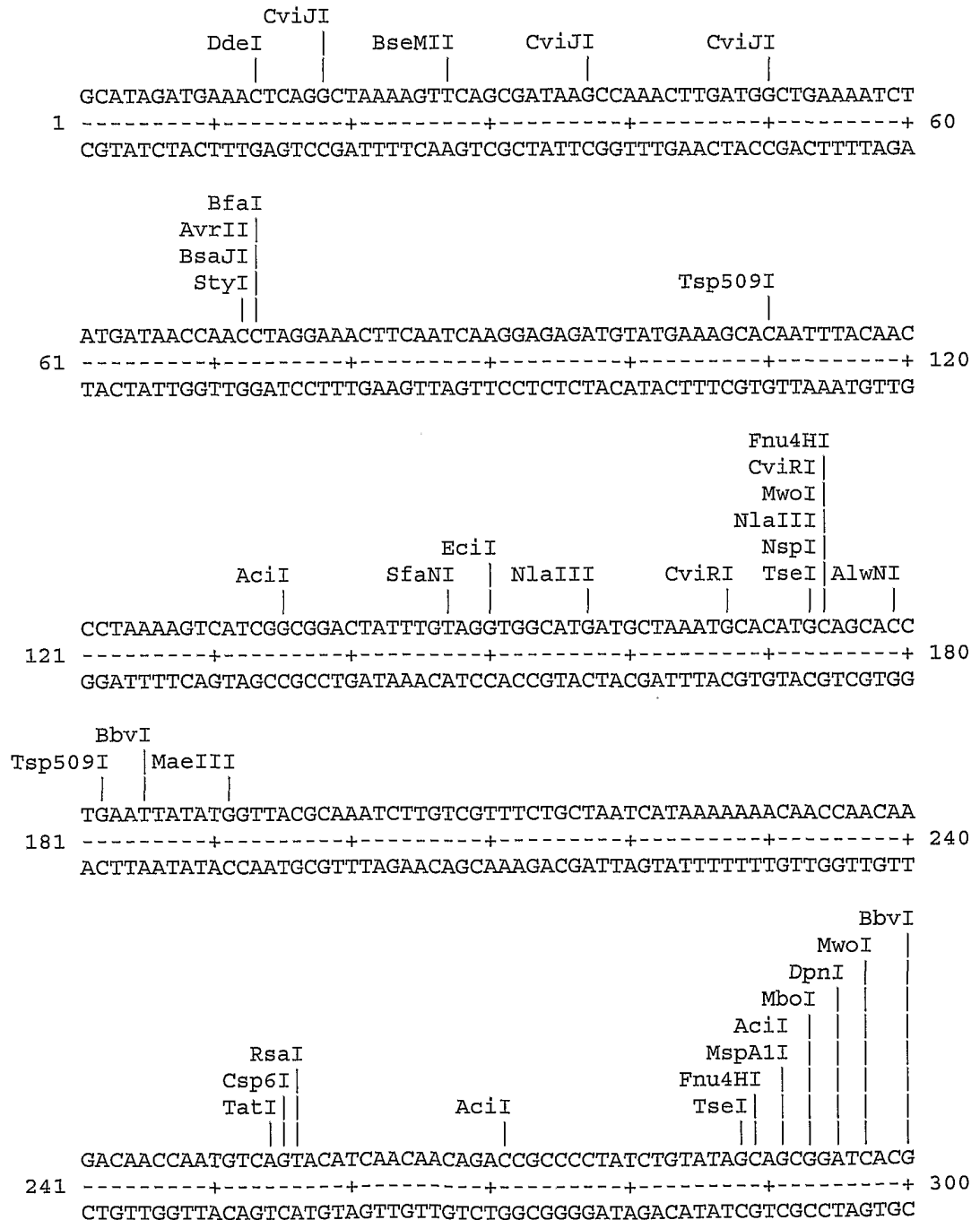
```

177/189

atc tta gat gaa aat aaa gag aca gat acc att gat cgt gat ggt tat	691
Ile Leu Asp Glu Asn Lys Glu Thr Asp Thr Ile Asp Arg Asp Gly Tyr	
185 190 195	
tca ttg ctt gtc aag aaa aac aca gat aca tat aat gtt ggt gca gct	739
Ser Leu Leu Val Lys Lys Asn Thr Asp Thr Tyr Asn Val Gly Ala Ala	
200 205 210	
tat gcc tat ttt ggt aaa gca aaa acc tct tat ggc aag att aac tat	787
Tyr Ala Tyr Phe Gly Lys Ala Lys Thr Ser Tyr Gly Lys Ile Asn Tyr	
215 220 225	
act gcg cgt gtg aca ggt aat tat aaa att aat gag gat tat aaa gta	835
Thr Ala Arg Val Thr Gly Asn Tyr Lys Ile Asn Glu Asp Tyr Lys Val	
230 235 240 245	
ggg ggt atc tat cag cat gtc ggc tat gcc aat gac gac agt gcc aaa	883
Gly Gly Ile Tyr Gln His Val Gly Tyr Ala Asn Asp Asp Ser Ala Lys	
250 255 260	
aat aat aca gaa caa gct gtt ggc gtg gct ttg cag cat ttt aaa gac	931
Asn Asn Thr Glu Gln Ala Val Gly Val Ala Leu Gln His Phe Lys Asp	
265 270 275	
aaa tgg act cat acc gca cac atg aat ctt gtg aat aat cct agt ggg	979
Lys Trp Thr His Thr Ala His Met Asn Leu Val Asn Asn Pro Ser Gly	
280 285 290	
aaa aag ggc gat ggg ttt gag ctg atc gga gca att gac cgt gat att	1027
Lys Lys Gly Asp Gly Phe Glu Leu Ile Gly Ala Ile Asp Arg Asp Ile	
295 300 305	
tct aaa aat gtc tct gca ggc atg gat atc acc tat ggt aac ttt aac	1075
Ser Lys Asn Val Ser Ala Gly Met Asp Ile Thr Tyr Gly Asn Phe Asn	
310 315 320 325	
tac gca aca gaa aaa gaa tct tac atc aac cca act atc tat gcg act	1123
Tyr Ala Thr Glu Lys Glu Ser Tyr Ile Asn Pro Thr Ile Tyr Ala Thr	
330 335 340	
gta tat ttt taa ccaattcatt gttgggtaat taagtaagat gtgataaagt	1175
Val Tyr Phe	
345	
caatttacga caaggcgaca gattgccaat ctactcaata aaagtaccac aatgaaaaac	1235

178/189

Figure 54. Restriction enzyme analysis of the *M. catarrhalis* porin gene designated MCA102070 .



179/189

DdeI
 CviJI
 HaeIII
 MnlI
 AlwI
 MseI
 MnlI
 RsaI
 Csp6I
 BsrGI
 TatI

CATTGGCCTTAGAGGGTCTGAAAAGTTAAATGATGATTATGAGGTTTTGTACAAGTTGGA
 301 -----+-----+-----+-----+-----+-----+ 360
 GTAACCGGAATCTCCAGACTTTTCAATTTACTACTAATACTCCAAAACATGTTCAACCT

BfaI
 XbaI
 Bst4CI
 RsaI
 Csp6I
 BsrDI
 HphI
 CviJI
 Fnu4HI
 TseI
 BbvI
 RsaI
 Csp6I

GTACCGTCTAGAAAATGATGGTGATTACGCAATGAAAAATCAAGCAGCCAGATGGTAC
 361 -----+-----+-----+-----+-----+-----+ 420
 CATGGCAGATCTTTTACTACCACTAAATGCGTTACTTTTTTAGTTCGTCTCGGTCTACCATG

Bst4CI
 Tsp509I
 MnlI
 BssSI
 CviJI
 TfiI
 NlaIII
 HinfI

TGAAAAGACTGTTGCTAAAACACGCAATTTTGAGGCTCGTGATTTCATGGATTGGCGTGAA
 421 -----+-----+-----+-----+-----+-----+ 480
 ACTTTTCTGACAACGATTTTGTGCGTTAAAACTCCGAGCACTAAGTACCTAACCGCACTT

FauI
 KpnI
 NlaIV
 RsaI
 Csp6I
 Acc65I
 BanI
 AciI
 DpnI
 NlaIV
 BamHI
 BstYI
 AlwI
 MboI
 NdeI

GCATAAAAAGTATGGTACCATCAAGGCGGGTCGTATGTTGTCTTTGGATCCATATGTGCG
 481 -----+-----+-----+-----+-----+-----+ 540
 CGTATTTTTCATACCATGGTAGTTCGCCCCAGCATAACAGAAACCTAGGTATACACGC

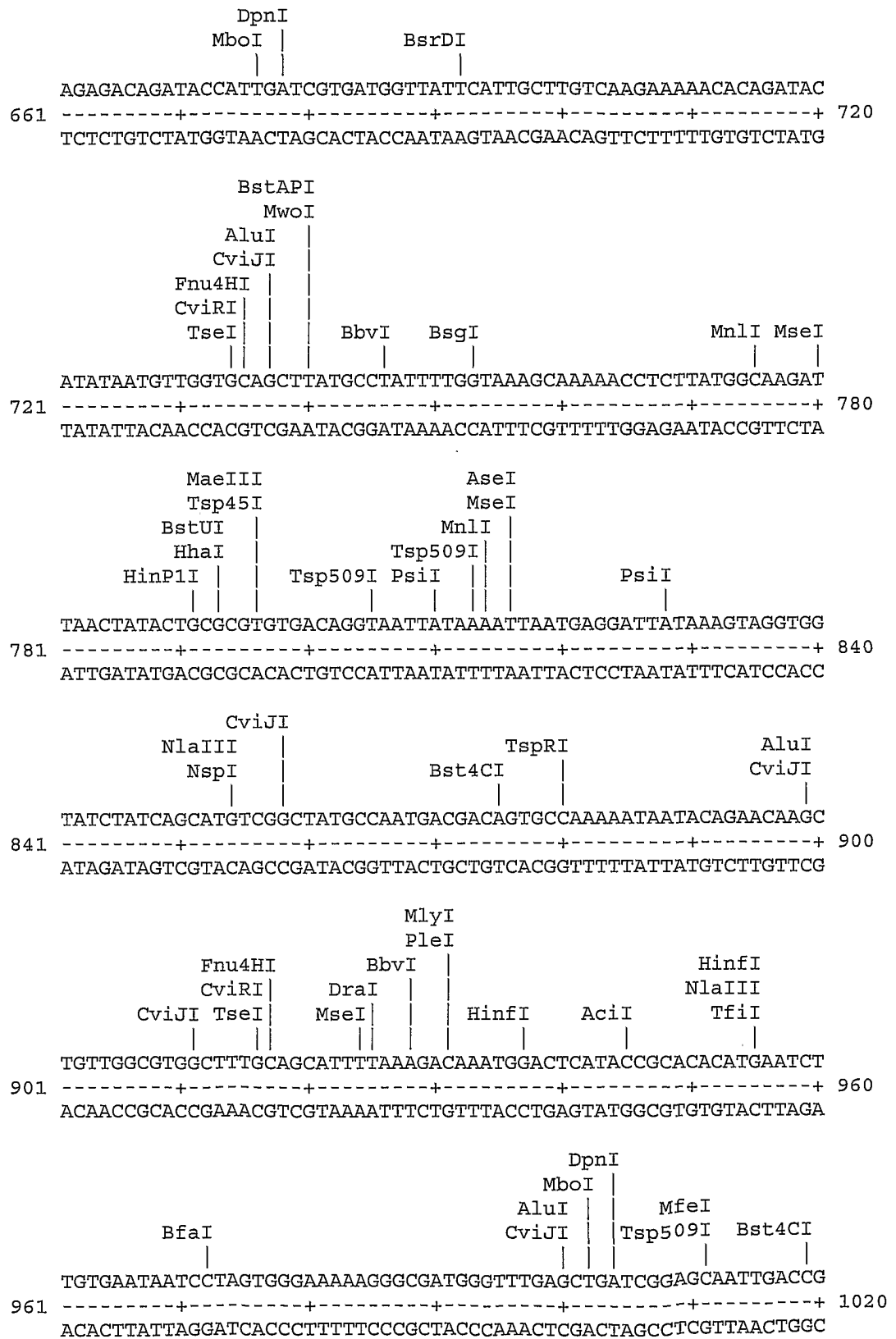
BsaHI
 HgaI
 BglI
 MwoI
 RsaI
 Csp6I
 BsrDI
 HphI
 CviRI

TTATACTGCCTATTTGGCGTCAGGCGTAGATGGGGTGCGTACCAATAATACCATTCGATA
 541 -----+-----+-----+-----+-----+-----+ 600
 AATATGACGGATAAACCGCAGTCCGCATCTACCCACGCATGGTTATTATGGTAACGTAT

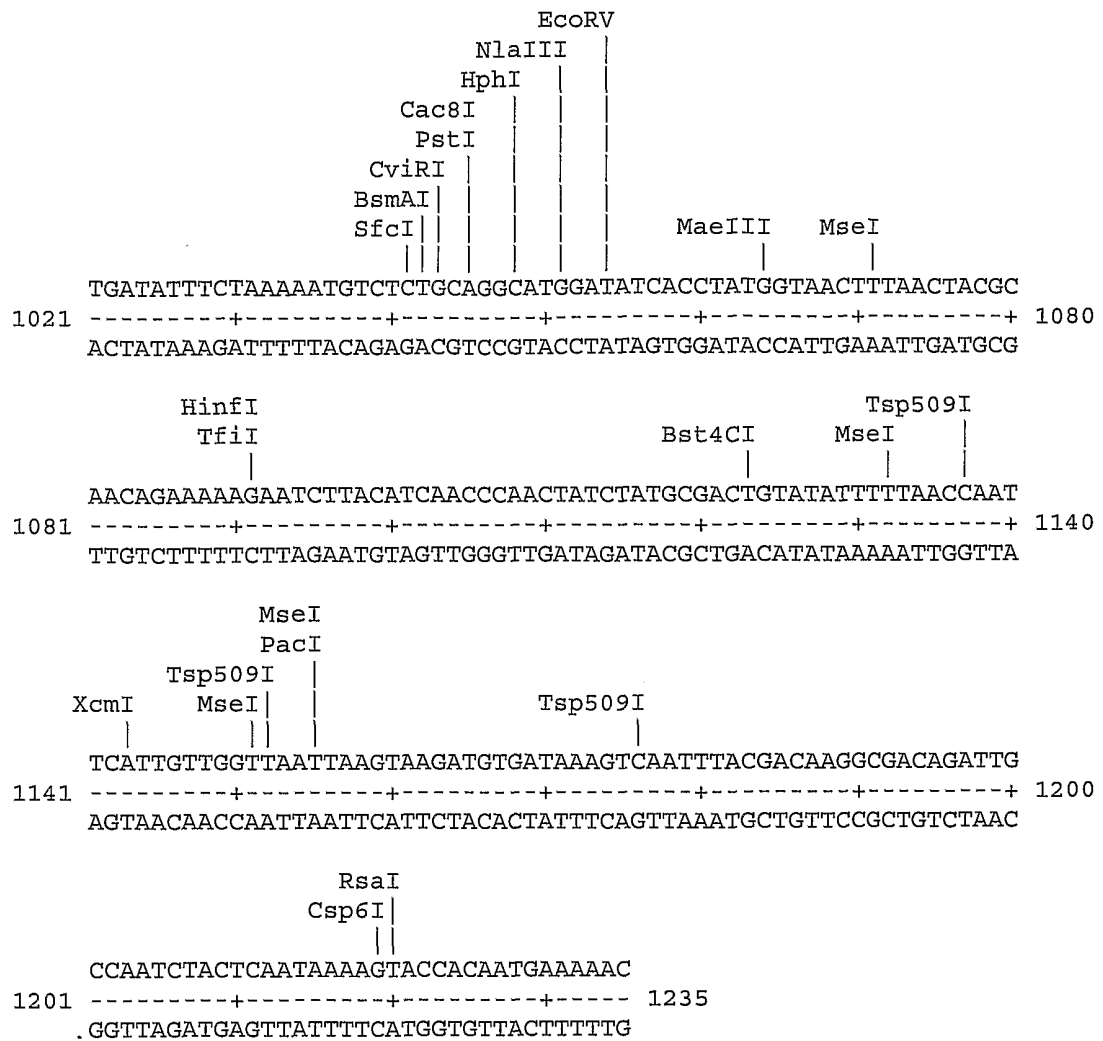
BsaBI
 RsaI
 Csp6I
 BsrGI
 TatI
 HinfI
 TfiI
 DdeI
 BsmAI

CGAATCACCAAAAATCAAAGATGTTAGTTTTTCAGGCGATGTACATCTTAGATGAAAATAA
 601 -----+-----+-----+-----+-----+-----+ 660
 GCTTAGTGTTTTTAGTTTTCTACAATCAAAAGTCCGCTACATGTAGAATCTACTTTTTATT

180/189

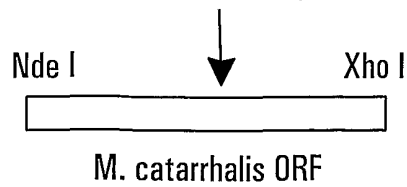


181/189

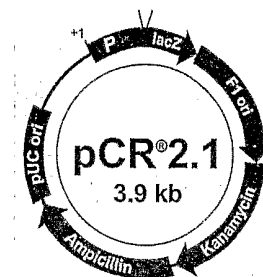


182/189

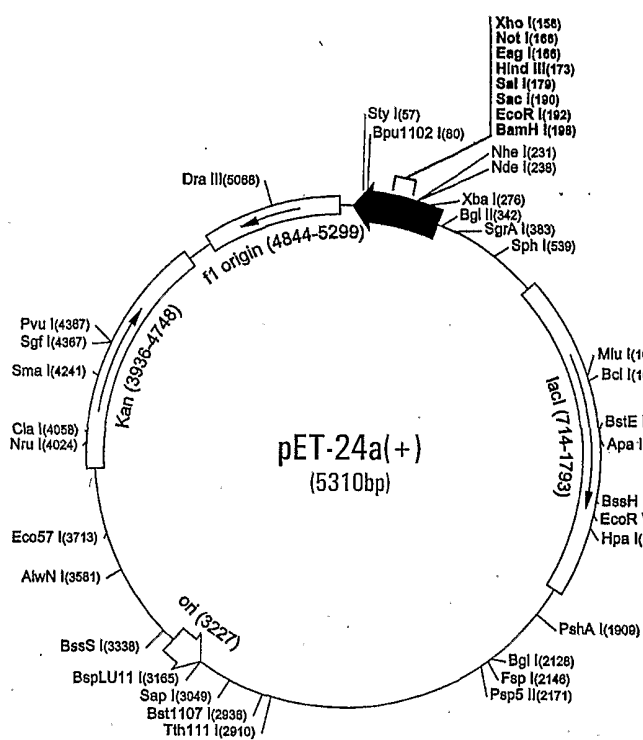
Genomic DNA--PCR with 5' and 3' primers



insert



Nde I/Xho I



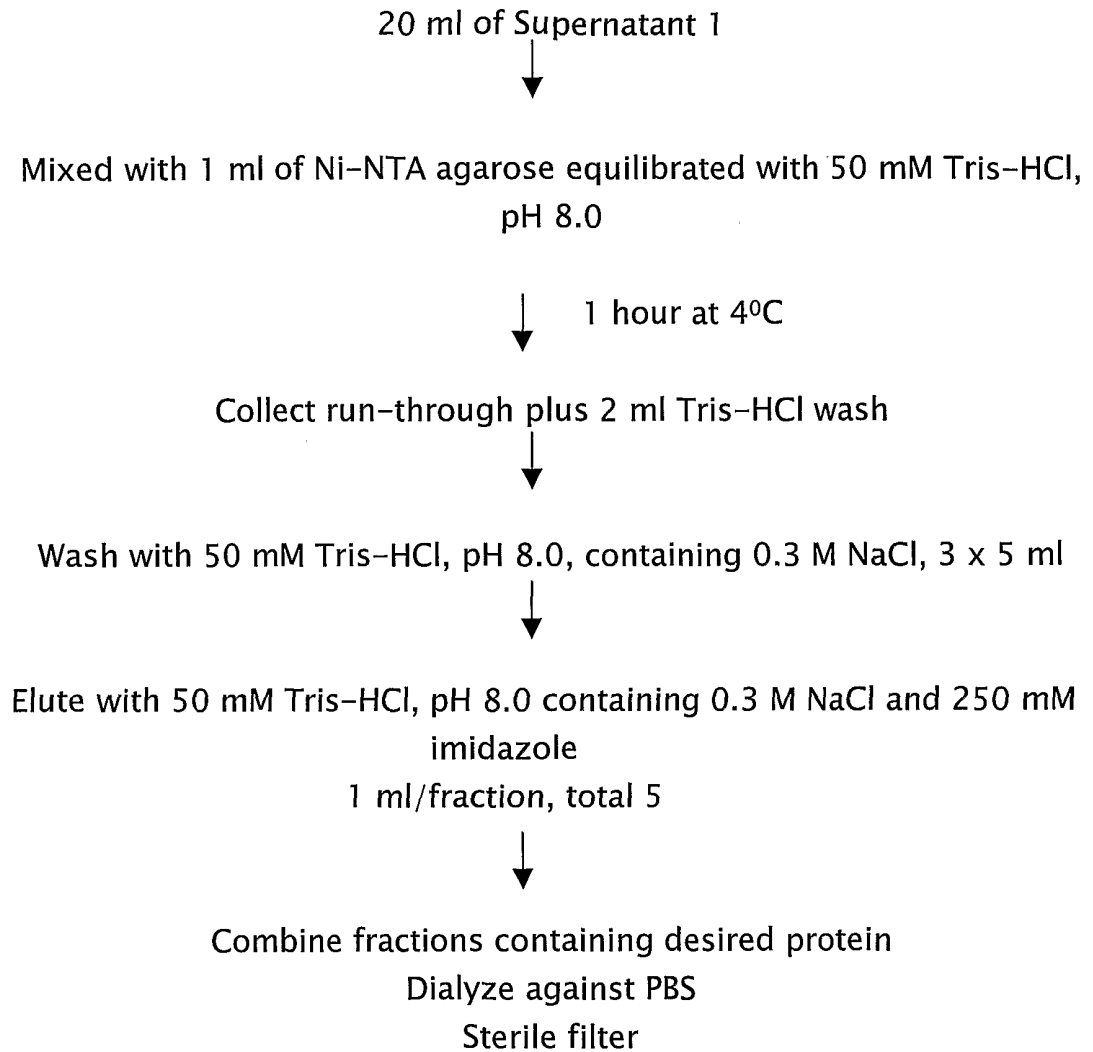
Nde I

Xho I

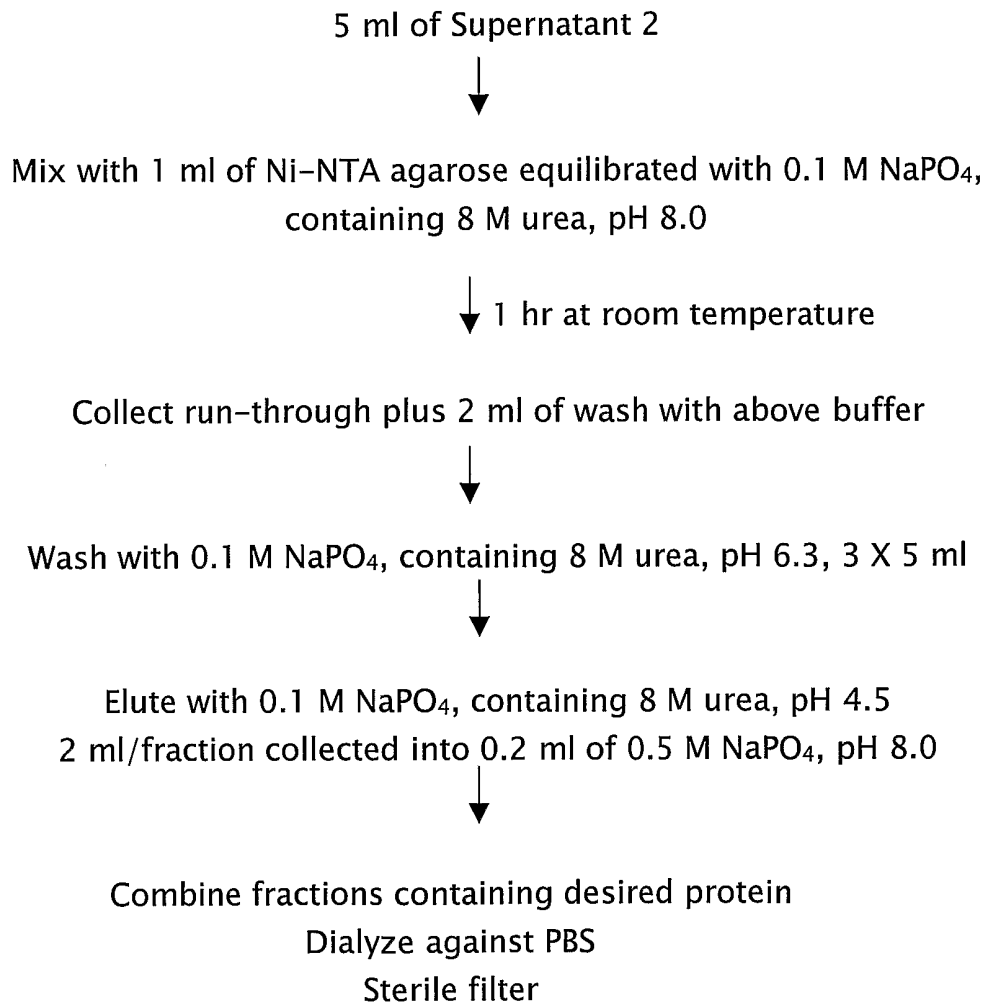
Moraxella ORF
Expression
plasmid

FIG. 55

183/189

Figure 56. Purification under Native conditions

184/189

Figure 57: Purification under Denaturing conditions

185/189

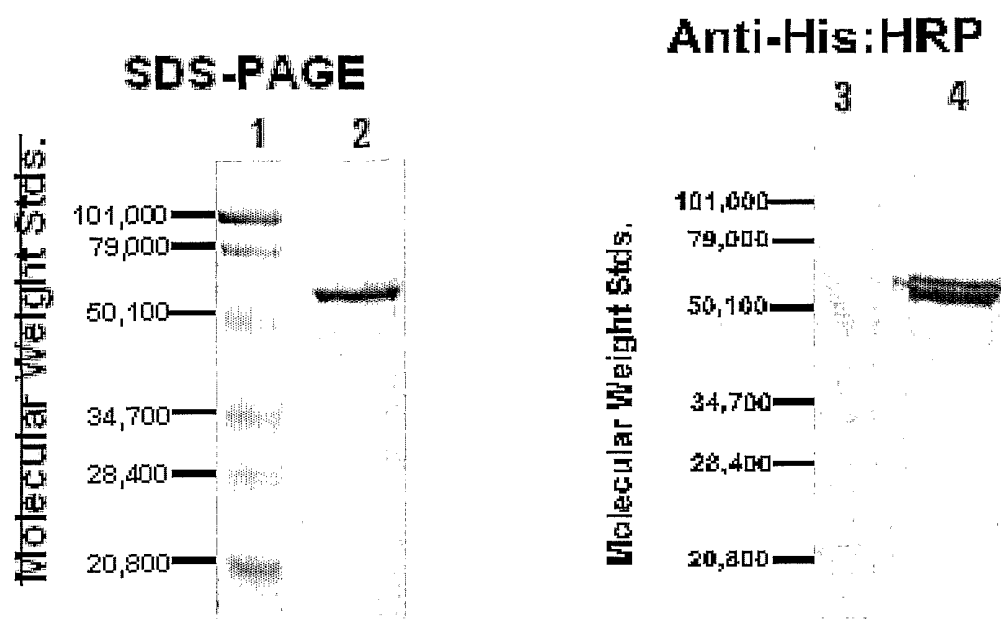


Fig. 58a

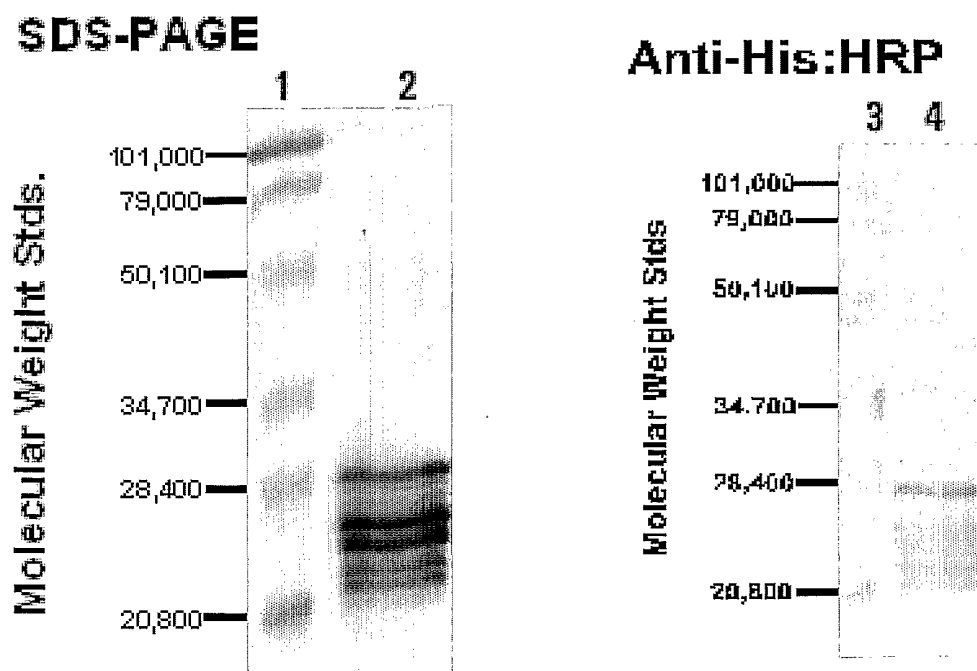


Fig. 58b

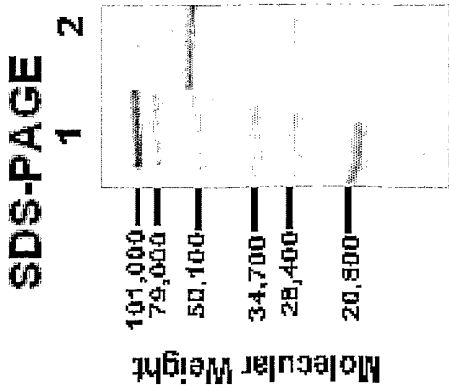


Fig. 58(c)

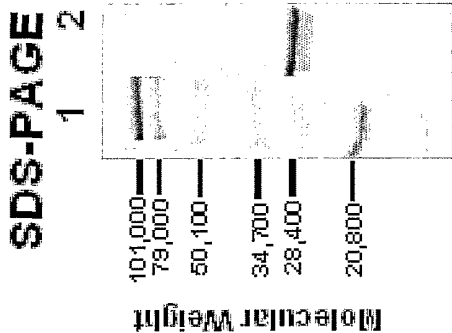


Fig. 58(d)



Fig. 58(e)

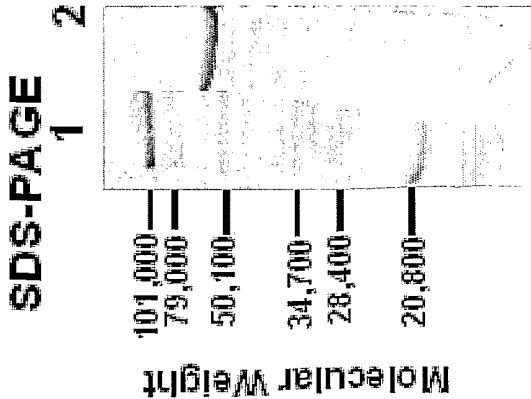


Fig. 58(f)

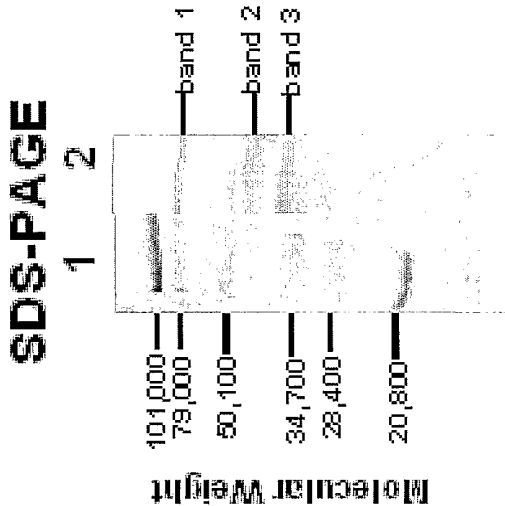


Fig. 58(g)

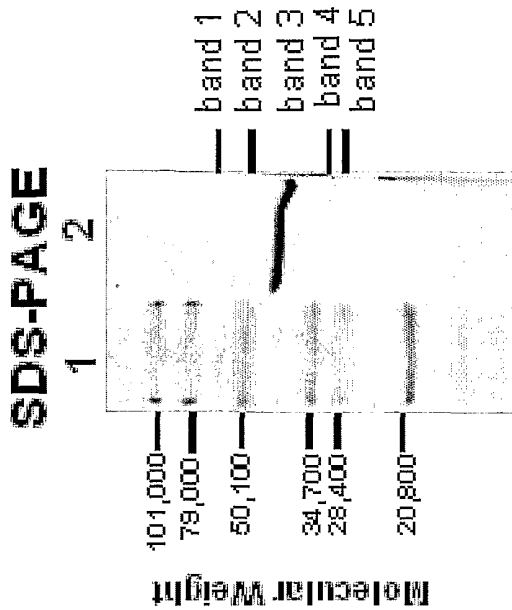
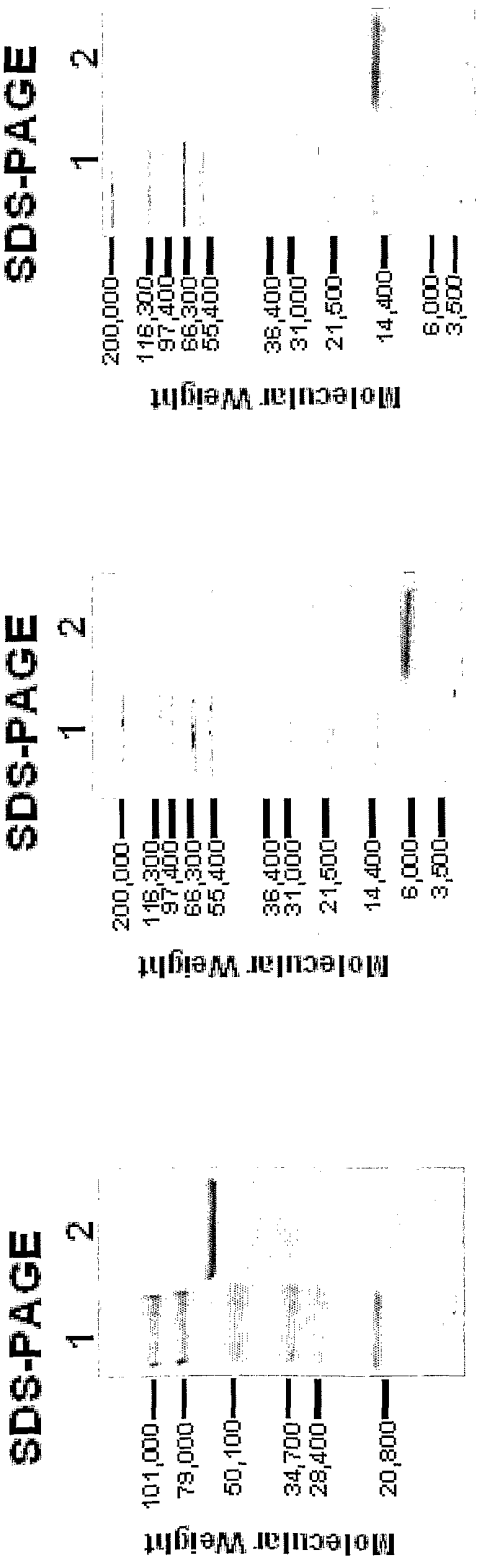


Fig. 58(h)



188/189

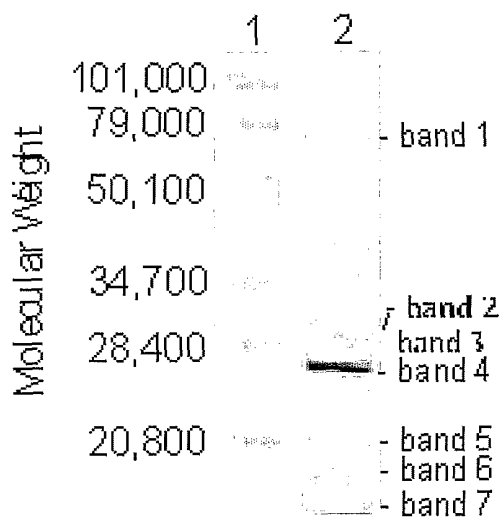
SDS-PAGE

Fig. 58(o)

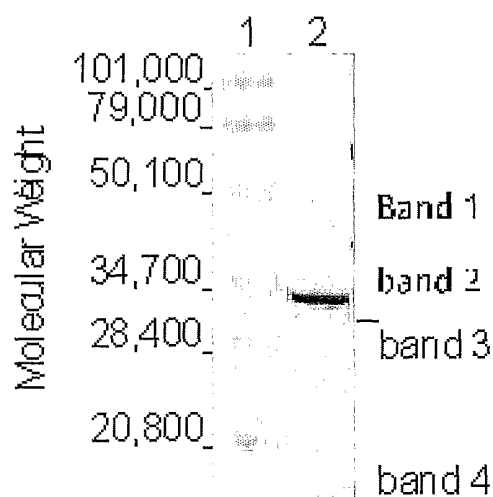
SDS-PAGE

Fig. 58(p)

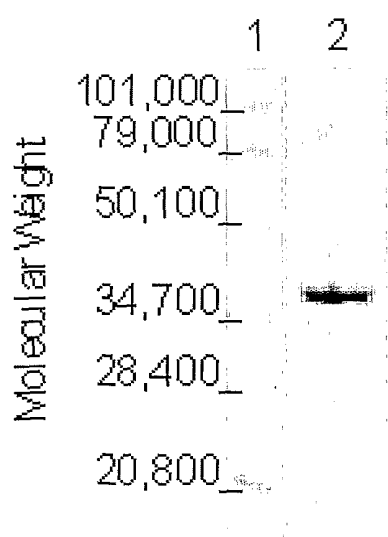
SDS-PAGE

Fig. 58(q)

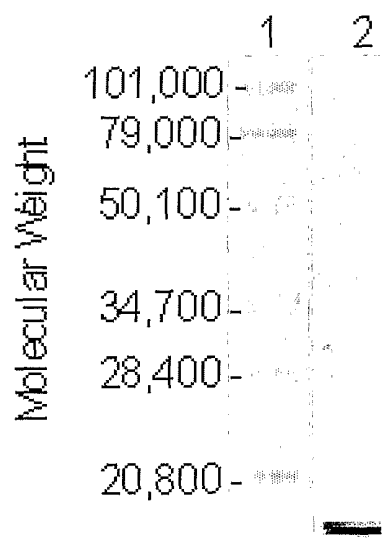
SDS-PAGE

Fig. 58(r)

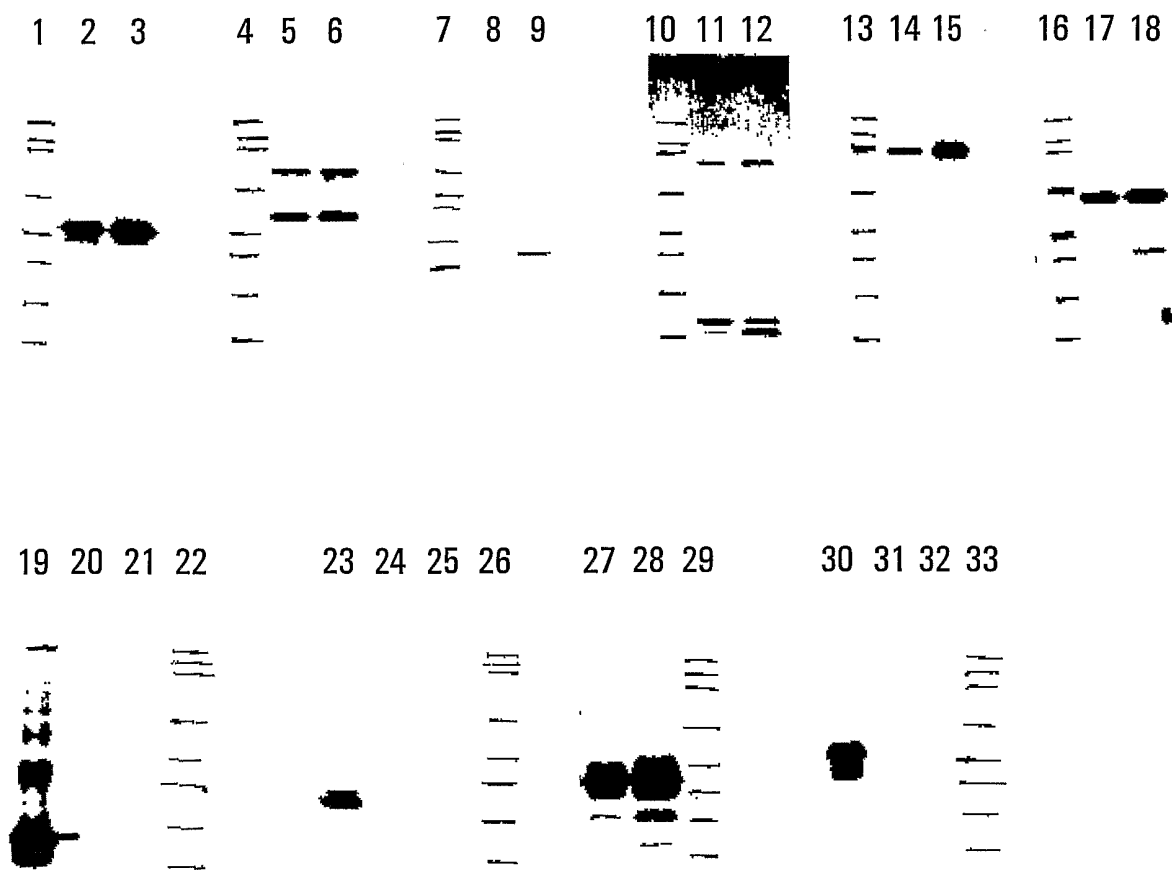


Fig. 59