INFECTION AND IMMUNITY, Feb. 1995, p. 676–680 Vol. 63, No. 2
0019-9567/95/$04.00+0

Identification of the *Chlamydia trachomatis* RecA-Encoding Gene

D.-J. ZHANG,† H. FAN, G. MCCLARTY, AND R. C. BRUNHAM*

Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba R3E 0W3, Canada

Received 6 September 1994/Returned for modification 13 October 1994/Accepted 7 November 1994

DNA sequencing of the major outer membrane protein (MOMP) gene (*omp1*) from *Chlamydia trachomatis* shows that some strains have a mosaic structure suggestive of homologous recombination between two distinct *omp1* genes. On the basis of this conjecture, we attempted to clone by complementation and sequence the chlamydial *recA* homolog from *C. trachomatis* serovar L_2. Chlamydial genomic DNA was partially restricted with *Xba*I, and fragments of 2 to 4 kb were ligated into pUC19. The recombinant plasmid was electroporated into *Escherichia coli* HB101 (RecA$^-$), and colonies were selected in the presence of methyl methanesulfonate (MMS). A 2.1-kb fragment of *C. trachomatis* DNA in pUC19 conferred relative MMS resistance to *E. coli* HB101. When this recombinant plasmid (pX203) was electroporated into *E. coli* JC14604 (RecA$^-$ *lacZ*), *lac*$^+$ recombinants were isolated. Rabbit polyclonal antibodies produced to purified *E. coli* RecA were immunoreactive in an immunoblot assay with a 35-kDa antigen in RecA$^-$ strains of *E. coli* transformed with pX203. The 2.1-kb insert was cycle sequenced by the dideoxy chain termination method. An open reading frame of 1,056 bp encoding 352 amino acids that had 44% sequence identity with *E. coli* RecA was identified. The finding of a *recA* homolog in *C. trachomatis* suggests that homologous recombination may occur in this organism. The cloned *C. trachomatis* RecA-encoding gene will be useful for the construction of a *recA* mutant once a gene transfer system is developed for chlamydiae.

Chlamydia trachomatis is a dimorphic bacterial pathogen of mammalian cells (13). The organism replicates within an endocytic vacuole which does not fuse with host cell lysosomes. *C. trachomatis* has no known means of genetic exchange, although phage has been identified in some *Chlamydia psittaci* isolates (17). Chlamydiae are presumed to be genetically isolated from other organisms and thus to be relatively stable and slow to evolve.

One *C. trachomatis* gene, *omp1*, which specifies the major outer membrane protein (MOMP), is highly diverse, and MOMP polymorphism determines the serologic classification for the standard 15 *C. trachomatis* serovars (1, 16). Recent genotypic analysis of clinical isolates has demonstrated that *omp1* is much more diverse at the DNA level than was expected from serologic typing (4, 5, 7, 18). The nature of DNA sequence diversity at the *omp1* locus suggests that point mutation and possibly recombination generate variation (10). On the basis of DNA sequencing of an entire *omp1* gene, Lampe et al. (10) suggested that a novel serologic variant of serovar I (termed Ia) may have resulted from recombination between serovar I and serovar H. Subsequently, an unexpectedly high prevalence (19 [32%] of 60 samples) of putative *omp1* recombinants was identified among *C. trachomatis* strains studied in a core group of urban prostitutes in Nairobi, Kenya (4). The DNA sequencing of the *omp1* gene from these infections implied a mosaic structure for MOMP, with variable domains 1 and 2 originating from one serovar and variable domain 4 from a different serovar.

These data suggest that *C. trachomatis* may undergo intergenomic homologous recombination at the *omp1* locus. To determine if *C. trachomatis* has the capacity to undergo homologous recombination, we attempted to exploit heterospecific complementation of RecA$^-$ mutants of *Escherichia coli* with chlamydial DNA in order to identify the *C. trachomatis* *recA* homolog.

Molecular cloning and DNA sequencing. Bacterial strains and plasmids are listed in Table 1. *E. coli* strains were routinely grown in Luria-Bertani (LB) medium. MacConkey's lactose agar was used to detect Lac$^+$ papillae for recombinant *E. coli*. *C. trachomatis* L_2/434/Bu was grown in HeLa cell monolayers in Eagle's minimal essential medium containing 10% fetal calf serum.

A genome DNA library from *C. trachomatis* serovar L_2 was constructed in pUC19 following partial digestion with *Xba*I with DNA fragments in the size range from 2 to 4 kb. Recombinant plasmid was electroporated into competent *E. coli* HB101 by using a 0.1-cm-gap cuvette at 25 μF, 1.8 to 2.0 kV, and 200 Ω. Since RecA-dependent DNA repair is required for growth in the presence of the DNA-alkylating agent methyl methanesulfonate (MMS), clones that suppressed the RecA$^-$ phenotype of HB101 were identified by growth in LB medium containing 2 mM MMS (6). Cultures were incubated overnight at 37°C, and colonies which appeared to be MMS resistant were picked and rescreened in 2 ml of LB broth containing between 2 and 8 mM MMS. Whereas the HB101 host strain was inhibited by 1 mM or more MMS, an isolated *E. coli* clone carrying the recombinant plasmid designated pX203 grew in the presence of 4 mM MMS (data not shown). This recombinant plasmid was used for further study.

To test for recombinational proficiency, pX203 was electroporated into *E. coli* JC14604. This *recA* mutant of *E. coli* harbors a duplication of the *lacZ* region, each copy containing a different *lacZ* gene missense mutation which, in the presence of a functional *recA*, can give rise to Lac$^+$ recombinants via a recombinational event (8). Lac$^+$ phenotypes were scored on MacConkey plates after 48 h of culture. More than 10% of

* Corresponding author. Mailing address: Department of Medical Microbiology, University of Manitoba, Room 543, 730 William Avenue, Winnipeg, Manitoba R3E 0W3, Canada. Phone: (204) 789-3524. Fax: (204) 783-5255. Electronic mail address: brunham@bldghsc.lan1.umanitoba.ca.

† Permanent address: Norman Bethune University of Medical Sciences, Changchun, People's Republic of China.

```
CTAGATAAAAAAAGAGTGTAGTTTTTTTGAAATTCTAGCACTTGACTGTTGTCTTTGAAATATTGCAA        68
ACTGACGCCACGAAATATAAGAGGTCGCAGATGGCAACCTCTTAAAGTAGAGTCTTCTAGACAGCGCA       136
                 M   S   V   P   D   R   K   R   A   L   E   A   A          13
      _____    ATG AGC GTT CCC GAC CGA AAA AGG GCT TTG GAA GCG GCC           189
ACAAGGGAGTATAC
 I   A   Y   I   E   K   Q   F   G   A   G   S   I   M   S   L   G          30
ATT GCT TAT ATC GAA AAG CAA TTT GGC GCA GGA TCT ATC ATG AGT TTA GGA          240
K   H   S   S   A   H   E   I   S   T   I   K   T   G   A   L   S          47
AAG CAT TCT TCA GCT CAT GAG ATA TCA ACT ATT AAA ACA GGT GCA TTG TCG          291
 L   D   L   A   L   G   I   G   G   V   P   K   G   R   I   V   E          64
TTG GAT TTA GCC TTA GGA ATA GGC GGG GTT CCT AAA GGA AGA ATT GTA GAG          342
 I   F   G   P   E   S   S   G   K   T   T   L   A   T   H   I   V          81
ATT TTC GGG CCA GAG TCT TCA GGG AAA ACA ACT CTA GCG ACG CAT ATA GTG          393
 A   N   A   Q   K   M   G   G   V   A   A   Y   I   D   A   E   H          98
GCC AAT GCT CAA AAG ATG GGA GGG GTG GCG GCT TAT ATT GAT GCC GAG CAT          444
 A   L   D   P   N   Y   A   A   L   I   G   A   N   I   N   D   L         115
GCC TTA GAC CCG AAT TAT GCT GCG CTT ATT GGA GCA AAT ATT AAT GAT TTA          495
 M   I   S   Q   P   D   C   G   E   D   A   L   S   I   A   E   L         132
ATG ATT TCT CAG CCT GAC TGC GGA GAA GAT GCT TTG AGT ATT GCA GAG CTC          546
 L   A   A   S   G   A   V   D   V   I   V   I   D   S   V   A   A         149
TTA GCG CGT TCT GGA GCT GTC GAT GTG ATT GTG ATT GAC TCG GTA GCA GCA          597
 L   V   P   K   S   E   L   E   G   E   I   G   D   V   H   V   G         166
TTA GTT CCA AAG AGC GAG TTA GAA GGG GAA ATT GGA GAT GTC CAT GTT GGT          648
 L   Q   A   R   M   M   S   Q   A   L   R   K   L   T   A   T   L         183
TTG CAA GCT CGC ATG ATG TCG CAG GCT CTA CGC AAA TTA ACT GCA ACC TTA          699
 A   R   T   N   T   C   A   I   F   I   N   Q   I   R   E   K   I         200
GCA CGA ACC AAT ACT TGT GCC ATT TTC ATT AAC CAG ATT CGG GAG AAA ATA          750
 G   V   S   F   G   N   P   E   T   T   T   G   G   R   A   L   K         217
GGT GTG AGT TTT GGT AAT CCA GAG ACT ACG ACT GGA GGA CGT GCA CTG AAG          801
 F   Y   S   S   I   R   I   D   I   R   R   I   G   S   I   K   G         234
TTT TAT TCT TCG ATT CGT ATC GAT ATT CGT CGT ATT GGC TCC ATA AAG GGA          852
 G   E   N   F   D   I   G   N   R   I   K   V   K   V   A   K   N         251
GGA GAA AAC TTC GAT ATA GGG AAT CGT ATC AAG GTG AAA GTA GCA AAA AAT          903
 K   L   A   P   P   F   R   T   A   E   F   D   I   L   F   N   E         268
AAA TTA GCT CCT CCA TTC CGA ACT GCA GAA TTT GAT ATC TTG TTT AAT GAA          954
 G   I   S   S   A   G   C   I   I   D   L   A   V   E   K   N   I         285
GGG ATT TCT TCC GCA GGG TGT ATC ATT GAT CTC GCT GTC GAA AAA AAT ATT         1005
 I   D   K   K   G   S   W   F   N   Y   Q   D   R   K   L   G   Q         302
ATC GAT AAG AAA GGA TCG TGG TTC AAC TAC CAG GAT CGT AAG TTA GGC CAA         1056
 G   R   E   A   V   R   E   E   L   K   R   N   K   E   L   F   H         319
GGT CGA GAG GCG GTT CGA GAA GAG CTG AAA AGG AAT AAG GAG TTG TTC CAC         1107
 E   L   E   R   R   I   Y   E   S   V   Q   A   S   Q   A   P   A         336
GAG TTG GAA CGC CGT ATT TAT GAG TCT GTA CAG GCT TCA CAA GCT CCA GCG         1158
 A   A   C   V   D   S   E   S   R   E   V   A   E   A   A   K   -         352
GCA GCT TGC GTG GAT TCA GAG TCT CGC GAG GTA GCA GAG GCT GCG AAA TAG         1209

AGGGTGTATTCCCTATCGATTCTAATCAAGGAAAAGAAGTTTTTCTTCTTTTCCTTGGTAGTTTCT        1275
TTTCTCCCCCGTCTTTGTTGTGGTTCTCTACATAGGTTTGGTAATTCCCGGAAGCTGTAGAGCCAA        1341
TTTCATAATTTTTCTTTGAAAGGCTAAAATTTTTTGGTGTAAACTCCACGGATCTTTGGTGTGCGA        1407
CTTCTTTATAGTGTTGAGGTTAGATAAAAGAGACTTCAAAGAGTAAGGCACCTAGAGATCAAGCAT        1473
TTTAACGAGCGAGAAGGCTTTGTGGTTCGCGTTTTCTTGTTAGTGGTTTTTCTCCTTTCTTGTTTC        1539
TCTGCAGGCGGAAGATCATATGTATTTTTTTCTTCGAGAGACGATTCTCCATCGCATTATGATCAT        1605
ACACATCGAACTGTTGCAAGATTAAGAGAGGGAAGTACCTTGTCAGAAGGTCTTTCCATGGAACAA        1671
..........                                                              2140
```

FIG. 1. DNA sequence of the *C. trachomatis recA* gene from plasmid pX203. The predicted amino acid residues are shown in the single-letter code. The open reading frame starting at bp 150 and ending at bp 1209 contains 1,059 bp and codes for 352 amino acids. The numbers on the right are nucleotide numbers (top) and predicted amino acid residues (bottom). The solid line 5′ to the open reading frame indicates a potential ribosome-binding site, and the solid line 3′ to the open reading frame indicates the sequence of a stable secondary stem and loop structure which is compatible with a rho-independent terminator. The ATGs at codon 1 and codon 27 are in boldface type and may represent alternative start sites in *E. coli.*

more than 1,000 colonies contained Lac^+ papillae. The frequency of *lac* recombination was almost the same as that obtained with plasmid pJB4, which contains the *Francisella novicida recA* gene (2). No Lac^+ papillae appeared when *E. coli* JC14604 was electroporated with pUC19 alone. Thus, pX203 restored recombinational proficiency in *E. coli* JC14604.

Because pX203 exhibited two traits characteristic of RecA, the 2.1-kb *C. trachomatis* DNA insert was sequenced in its entirety using forward and reverse primers in both directions. DNA sequencing was carried out with double-stranded cycle sequencing based on the dideoxy chain termination method as previously described (18). The insert contained 2,140 bp and one major open reading frame of 1,056 bp, encoding a 352-amino-acid protein (Fig. 1). The 2.1-kbp sequence is available in GenBank under accession number U15281. No classic *E. coli* σ^{70} −10 and −35 promoter sequences were identified upstream of the putative ATG initiation codon. A possible ribosome-binding site (GGGAG) was identified 5 bp upstream from the presumed initiation codon. The open reading frame terminates with a TAG stop codon followed by a potential

```
C.trachomatis     MS----------VPDRKRALEAAIAYIEKQFGAGSIMSLGKHSSAHEIST   40
R.prowazekii      ..----------NI.KE..IA..L.Q...SY.K..VMK..QRPNV-D.EA   39
L.pneumophila     .EE-----------NKQK..S..VSQ..R...K..V.RM.DSTVSRD.EA   39
N.gonorrhoeae     ..D-----------.KSK..A..L.Q...S..K.A..KMDGSQQEENLEV   39
B.abortus         ..QNSLRLVEDNSV.KTK..D..LSQIERA..K....R..QNDQVV..E.   50
E.coli            .A----------IDEN.QKAL..ALGQIEKQFGKGSIMRLGEDRSMDVE.   40

                             *       *
C.trachomatis     IKTGALSLDLALGIGGVPKGRIVEIFGPESSGKTTLATHIVANAQKMGGV   90
R.prowazekii      .S..S.G..I...........I...........TL.LI.ES..K..T     89
L.pneumophila     .S..S.G.........L........Y..........TLQVI.EC.....T   89
N.gonorrhoeae     .S..S.G......V..LRR................CLEA..QC..N...   89
B.abortus         VS..S........V..L........Y..........L.TI.E...K..I   100
E.coli            .S..S....I...A..L.M......Y..........TLQVI.A..RE.KT   90

C.trachomatis     AAYIDAEHALDPNYAALIGANINDLMISQPDCGEDALSIAELLARSGAVD   140
R.prowazekii      C.F.........A..KKL.V..DE.I.....T..Q..E..DT.I...GI.   139
L.pneumophila     ..F.........S..QKL.VKVDE.LV....T..Q..E.TDM.V..A...   139
N.gonorrhoeae     C.FV.....F..V..RKL.VKVEE.YL....T..Q..E.CDT.V...GI.   139
B.abortus         C.FV........V..RKL.VHLFN.L....IT..Q..E.TDT.V....I.   150
E.coli            C.F.........I..RKL.VD.DN.LC....T..Q..E.CDA........   140

                                      *       *
C.trachomatis     VIVIDSVAALVPKSELEGEIGDVHVGLGARMMSQALRKLTATLARTNTCA   190
R.prowazekii      M.I...........I...M..AQMASQ..L..........SIN...CIT   189
L.pneumophila     .VI.......T..A.I...M..S.......L..........NIK.S..LV   189
N.gonorrhoeae     MV.V.........A....D...........L.........GHIKK...LV   189
B.abortus         .L.V......T.RA.....M..S.-.....L....V....GSIS.S.CMV   199
E.coli            ...V......T..A.I......S.M..A.......M...AGN.KQS..LL   190

                                                      *
C.trachomatis     IFINQIREKIGVSFGNPETTTGGRALKFYSSIRIDIRRIGSIKGGENFDI   240
R.prowazekii      V......M....M..S.........N.....A.V...........DK.EV-.   238
L.pneumophila     .......M....M..S.......N.....A.V...........K..EI-L   238
N.gonorrhoeae     V...........M..S.......N.......V......T....K..EV-L   238
B.abortus         .......M....M..S.......N.....A.V.L.........ERDEV-V   248
E.coli            .......M....M..........N.....A.V.L......AV.E...VVG   240

C.trachomatis     GNRIKVKVAKNKLAPPFRTAEFDILFNEGISSAGCIIDLAVEKNIIDKKG   290
R.prowazekii      .SQT....V...VS...K..D...MYGS...KE.E....G.KLE..E.S.   288
L.pneumophila     .SETR...V...V....KMT.....Y.....RESE..N.G.QL...E.S.   288
N.gonorrhoeae     ..ETR...I...V.....Q......YG....WE.EL..IG.KNDI.N.S.   288
B.abortus         ..QTR...V........KQV....MYGA.V.KV.ELV..G.KAGVVE.S.   298
E.coli            SETRVKV.KNKIA..FKQAEFQILYGEGINFYGELVDLGVK..L.EKAGA   290

                            *
C.trachomatis     SW-FNYQDRKLGQGREAVREELKRNKELFHELERRIYESVQASQAPAAAC   339
R.prowazekii      ..-.S.NKIRI.....N.KQY..EHPQISN.I.KI.REKSSAITNINLDQ   337
L.pneumophila     A.-YSYKQE.I...K.N..LY..E.PQVAA...QQ.RTEL-LEKKLSVLA   336
N.gonorrhoeae     A.-.S.NGA.I...KDN..VW..E.P.ISD.IDAKIRA-L-NGVEMHITE   335
B.abortus         A.-.S.NSQR......NAKQY..D.P.VAR.I.TTLRQNAGLIAEQFLDD   347
E.coli            WYSYKGEKIGQ.KANATAWLKDNPETAKEI.KKV.ELLLSNPNST.DFSV   340

C.trachomatis     VDSESREVAEAAK  352        Identity (%)
R.prowazekii      TEE----------  340           59
L.pneumophila     SS..DL-FETIDD  348           58
N.gonorrhoeae     GTQDETDGERPEE  348           56
B.abortus         GGP.EDAAGA.EM  360           44
E.coli            D...GVAETNEDF  353           44
```

FIG. 2. Amino acid alignments of *C. trachomatis* RecA protein and five other bacterial RecA proteins. Dots represent identity with the *C. trachomatis* RecA sequence. Stars represent amino acids known to be essential for homologous recombination.

"hairpin" structure (beginning 30 bp after the stop codon), compatible with a rho-independent terminator whose predicted free energy value is −19.2 kcal. The G+C content of the open reading frame is 44.7%.

When the deduced amino acid sequence was compared with the sequences of known selected RecA proteins, a high degree of identity was observed (Fig. 2). In comparison with the chlamydial RecA, *E. coli* and *Brucella abortus* RecA exhibited 44% sequence identity, *Neisseria gonorrhoeae* RecA showed 56% identity, *Legionella pneumophila* RecA showed 58% identity, and *Rickettsia prowazekii* RecA showed 59% identity. Mutational analysis of *E. coli* RecA has shown that six amino acids

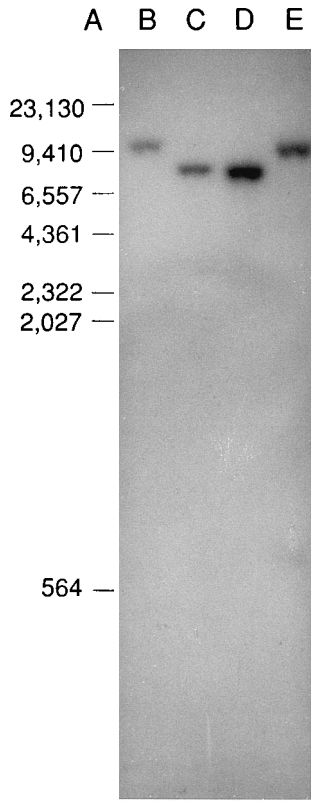

FIG. 3. Southern blot analysis of the *recA* gene in genomic DNA of *C. trachomatis* serovar L_2 (B and C) and *C. trachomatis* serovar C (D and E). Two micrograms of genomic DNA was digested to completion with either *Hin*dIII (C and D) or *Pst*I (B and E), electrophoresed through a 0.75% agarose gel, and then transferred to a nylon membrane. Blots were hybridized with a ^{32}P-labelled probe of the entire *recA* open reading frame. Hybridization in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.5% sodium dodecyl sulfate (SDS) was carried out overnight at 65°C, and then the filter was subjected to high-stringency washing (last washing step was 0.1× SSC–0.1% SDS at 65°C). Lane A, molecular weight markers.

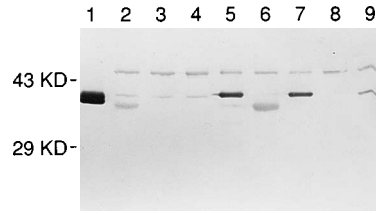

FIG. 4. Immunologic cross-reactivity of the *C. trachomatis* RecA protein expressed in *E. coli* with rabbit polyclonal RecA antiserum. The extracts from different strains of *E. coli* were electrophoresed in a 12% polyacrylamide gel and transferred to a nitrocellulose membrane for immunoblotting. Lane 1, purified *E. coli* K-12 RecA protein; lane 2, *E. coli* HB101 with pX203; lane 3, HB101 alone; lane 4, HB101 with pUC19; lane 5, HB101 with pJB4; lane 6, *E. coli* JC14604 with pX203; lane 7, JC14604 with pJB4; lane 8, JC14604 alone; lane 9, *E. coli* BL21 ($RecA^+$). An antigen of >45 kDa is observed in all *E. coli* extracts and presumably represents rabbit antibody to an unidentified *E. coli* antigen. *E. coli* HB101 has a *recA* point mutation (*recA13*) and expresses an immunoreactive protein product of 37 kDa (lanes 2, 3, and 4), whereas *E. coli* JC14604 has a deleted *recA* gene and expresses no immunoreactive RecA protein (lanes 6 and 8).

(recA1 Gly-161→Asp, recA13 Leu-52→Phe, recA56 Arg-61→Cys, recA1203 Arg-170→Cys, recA1601 Gly-302→Ser, recA1207 Gly-302→Asp, and recA142 Ile-226→Val) are essential to the function of homologous recombination (12). All six positions are conserved in the *C. trachomatis* RecA sequence (marked with a star in Fig. 2).

A nick-translated, α-^{32}P-labelled probe of the entire RecA open reading frame was used for Southern hybridization to *Hin*dIII- or *Pst*I-restricted genomic DNA from *C. trachomatis* serovars L_2 and C. The chlamydial *recA* sequence contains one internal *Pst*I site and no internal *Hin*dIII sites. A single *Hin*dIII band and two *Pst*I bands of similar size for each serovar were observed on Southern hybridization, suggesting that *C. trachomatis recA* is a single-copy gene present in a similar genomic context for both serovars L_2 and C (Fig. 3).

Expression. It has been shown that the sequence requirements for chlamydial promoters differ from those for *E. coli* promoters (11, 15). In pX203, the *recA* open reading frame lies 423 bp downstream of the vector β-galactosidase promoter. To determine if the vector promoter played a role in expression of the chlamydial RecA homolog, the 2.1-kb insert was excised from pX203 with *Xba*I and recloned into pUC19. PCR amplification was used to determine the orientation by employing an internal primer to the *recA* open reading frame (primer 14, 5′-TTG GCT CTA CAG CTT CCG-3′) and either the universal pUC19 primer (5′-GCA ACA AGG GAG TAT ACA TG-3′) or the reverse primer (5′-AGC GGA TAA CAA TTT CAC ACA GGA-3′). Four clones with forward orientation and four with reverse orientation were identified and tested for growth in the presence of 2 mM MMS. All were capable of growth in 2 mM MMS. Thus, the orientation of the insert in pUC19 does not determine the expression of MMS resistance.

Rabbit antiserum to commercially available purified RecA from *E. coli* (Pharmacia Ltd., Baie d'Urfe, Quebec, Canada) was prepared by four weekly subcutaneous injections of 200 μg of RecA protein in Freund's incomplete adjuvant. A 1:500 dilution of antiserum was used for immunoblotting with crude soluble extracts of *E. coli* HB101 and JC14604 with and without pUC19, pX203, and pJB4. A wild-type *E. coli* strain (BL21), which contains an intact *recA* gene, was used as an additional positive control (Fig. 4). The antiserum reacted with purified RecA at the approximate molecular mass of 37 kDa (lane 1) and with an antigen of the same molecular mass in wild-type *E. coli* BL21 (lane 9). *E. coli* HB101 and JC14604

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Reference
C. trachomatis L_2/Bu/434		9
E. coli		
HB101	F^- *hsdS20* (r^-_B, m^-_B) *supE44 ara-14 galK-2 lacY1 proA2 rpsL20* (Str^r) *xyl-5 mtl-12 recA13*	3
JC14604	F^- *lac* MS28608II *lacBK1* Δ(*srl-recA*) *hsr supE44*	8
Plasmids		
pUC19	Ap^r	
pX203	Ap^r, 2.1-kb *Xba*I *C. trachomatis* serovar L_2 DNA in pUC19	This study
pJB4	Km^r, *recA* region of *F. novicida* in pBGS19	2

containing pJB4, which specifies the *F. novicida* RecA, also had an antigen of the same molecular mass (lanes 5 and 7, respectively). *E. coli* HB101 and JC14604 containing pX203 (lanes 2 and 6) expressed an immunoreactive protein of approximately 34 kDa. This is presumably the *C. trachomatis* RecA. DNA sequencing shows that the *E. coli* and *C. trachomatis* RecA proteins differ in size by only a single amino acid. The discrepancy in size seen on the immunoblot is unexplained but may represent translational start at the second ATG (methionine) codon (codon 27). Translation from codon 27 yields a predicted protein product of approximately 34 kDa.

In aggregate, these results demonstrate that the *recA* gene from *C. trachomatis* serovar L_2 has been isolated by heterospecific complementation in *E. coli*. To date, *C. trachomatis* genes have been identified through antibody-reactive expression clones (usually as fusion proteins), by degenerate PCR amplification of conserved genes, or with mRNA probes. Palmer and Falkow (14) previously attempted heterospecific complementation in *E. coli* for chlamydial genes involved in proline, leucine, and threonine biosynthesis as well as RecA activity, but were unsuccessful. They attributed failure to absent or poor expression of chlamydial genes in *E. coli*. In general, this has been the experience of subsequent investigators. Our success in cloning the chlamydial *recA* gene by complementation is likely due to the presence of a 5′ upstream DNA sequence that contains an authentic or fortuitous promoter-like sequence capable of recognition by the *E. coli* RNA polymerase.

We were prompted to search for the *C. trachomatis recA* gene because DNA sequencing of selected regions of the *C. trachomatis omp1* gene from clinical samples suggested that homologous recombination might generate mosaic MOMP molecules. The identification of the *recA* gene in *C. trachomatis* supports the notion that this organism can undergo homologous recombination and may contribute to the pathogen's virulence by generating novel antigenic variants of the MOMP. The cloning of the *C. trachomatis recA* gene will be useful in the construction of a *recA* mutant which will be beneficial in genetic manipulations once a gene transfer system for chlamydiae is established.

This work was supported by a grant from the Medical Research Council of Canada (SP 27).

REFERENCES

1. **Baehr, W., Y. Zhang, T. Joseph, S. Hua, F. E. Nano, K. D. E. Everett, and H. D. Caldwell.** 1988. Mapping antigenic domains expressed by *Chlamydia trachomatis* major outer membrane protein genes. Proc. Natl. Acad. Sci. USA **85:**4000–4004.
2. **Berg, J. M., K. E. Mdluli, and F. E. Nano.** 1992. Molecular cloning of the *recA* gene and construction of a *recA* strain of *Francisella novicida*. Infect. Immun. **60:**690–693.
3. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41:**459–472.
4. **Brunham, R., C. Yang, I. Maclean, J. Kimani, G. Maitha, and F. Plummer.** 1994. *Chlamydia trachomatis* from individuals in a sexually transmitted disease core group exhibit frequent sequence variation in the major outer membrane protein (omp1) gene. J. Clin. Invest. **94:**458–463.
5. **Dean, D., J. Schachter, C. R. Dawson, and R. S. Stephens.** 1992. Comparison of the major outer membrane protein variant sequence regions of B/Ba isolates: a molecular epidemiologic approach to *Chlamydia trachomatis* infections. J. Infect. Dis. **166:**383–392.
6. **Goldberg, I., and J. J. Mekalanos.** 1986. Cloning of the *Vibrio cholerae recA* gene and construction of a *Vibrio cholerae recA* mutant. J. Bacteriol. **165:** 715–722.
7. **Hayes, L. J., R. L. Bailey, D. C. W. Mabey, I. N. Clarke, M. A. Pickett, P. J. Watt, and M. E. Ward.** 1992. Genotyping of *Chlamydia trachomatis* from a trachoma-endemic village in The Gambia by a nested polymerase chain reaction: identification of strain variants. J. Infect. Dis. **166:**1173–1177.
8. **Keener, S. L., K. P. McNamee, and K. McEntee.** 1984. Cloning and characterization of *recA* genes from *Proteus vulgaris*, *Erwinia carotovora*, *Shigella flexneri*, and *Escherichia coli* B/r. J. Bacteriol. **160:**153–160.
9. **Kuo, C.-C., S.-P. Wang, and J. T. Grayston.** 1973. Effect of polycations, polyanions, and neuraminidase on the infectivity of trachoma-inclusion conjunctivitis and lymphogranuloma venereum organisms in HeLa cells: sialic acid residues as possible receptors for trachoma-inclusion conjunctivitis. Infect. Immun. **8:**74–79.
10. **Lampe, M. F., R. J. Suchland, and W. E. Stamm.** 1993. Nucleotide sequence of the variable domains within the major outer membrane protein gene from serovariants of *Chlamydia trachomatis*. Infect. Immun. **61:**213–219.
11. **Mathews, S. A., and K. S. Sriprakash.** 1994. The RNA polymerase of *Chlamydia trachomatis* has a flexible sequence requirement at the −10 and −35 boxes of its promoters. J. Bacteriol. **176:**3785–3789.
12. **Miller, R. V., and T. A. Kokjohn.** 1990. General microbiology of *recA*: environmental and evolutionary significance. Annu. Rev. Microbiol. **44:**365–394.
13. **Moulder, J. W.** 1991. Interaction of chlamydiae and host cells in vitro. Microbiol. Rev. **55:**143–190.
14. **Palmer, L., and S. Falkow.** 1986. Characterization of cloned genes from *Chlamydia trachomatis*, p. 91–95. *In* L. Leive (ed.), Microbiology—1986. American Society for Microbiology, Washington, D.C.
15. **Sardinia, L. M., J. N. Engel, and D. Ganem.** 1989. Chlamydial gene encoding a 70-kilodalton antigen in *Escherichia coli*: analysis of expression signals and identification of the gene product. J. Bacteriol. **171:**335–341.
16. **Stephens, R. S., R. Sanchez-Pescador, E. A. Wagar, C. Inouye, and M. Urdea.** 1987. Diversity of *Chlamydia trachomatis* major outer membrane protein genes. J. Bacteriol. **169:**3879–3885.
17. **Storey, C. C., M. Lusher, S. J. Richmond, and J. Bacon.** 1989. Further characterization of a bacteriophage recovered from an avian strain of *Chlamydia psittaci*. J. Gen. Microbiol. **70:**1321–1327.
18. **Yang, C., I. Maclean, and R. C. Brunham.** 1993. DNA sequence polymorphism of the *Chlamydia trachomatis* omp1 gene. J. Infect. Dis. **168:**1225–1230.