Nucleotide Sequence of the Gene for the Vitamin B_{12} Receptor Protein in the Outer Membrane of *Escherichia coli*

KNUT HELLER† AND ROBERT J. KADNER*

Department of Microbiology, The University of Virginia School of Medicine, Charlottesville, Virginia 22908

Received 10 September 1984/Accepted 5 December 1984

The nucleotide sequence of a 2220-base-pair fragment containing the *btuB* gene of *Escherichia coli* was determined. There was a single open reading frame which was translated into a 614-amino-acid polypeptide, the first 20 amino acids of which comprised a typical leader sequence. The putative mature or processed form had a molecular weight (66400) and a composition in close agreement with that determined for the purified receptor. The distribution of amino acids in the receptor protein was similar to that of other outer membrane proteins, showing a fairly even distribution of charged residues and the absence of extensive hydrophobic stretches. The *btuB451* mutation appears to alter the receptor to eliminate its ability to function in vitamin B_{12} uptake without affecting its ligand binding properties. The sequence of the DNA from this mutant was determined and revealed a leucine-to-proline (C-to-T transition) change in the eighth amino acid of the mature form.

The uptake of vitamin B_{12} in *Escherichia coli* is initiated by the binding of vitamin B_{12} to a receptor protein in the outer membrane. Release of vitamin B_{12} into the periplasm or cytoplasm is dependent on the function of the *tonB* product, which is also necessary for the energy-dependent phases of outer membrane-dependent transport processes (20, 23). The structure of the vitamin B_{12} receptor is of interest for several reasons. Phage BF23, colicin A, and the E. coli colicins bind to the vitamin B_{12} receptor or competitively with each other and with vitamin B_{12} (4, 7). The location on the receptor protein of the binding sites for these ligands and the relationship between these binding sites might help explain the nature of this competitive binding by such different types of molecules. There may also be a binding site on the receptor for the *tonB* product to mediate the energy coupling of the transport process. The receptor does appear to carry out a transport process different from that mediated by the porins, because vitamin B_{12} uptake shows considerable substrate specificity and energy dependence. Comparison of the structure and topology of the vitamin B_{12} receptor with those of the porins and other outer membrane proteins might reveal the structural features common to all outer membrane proteins and necessary for their export and maintenance in that membrane.

To study the structure of the receptor, its gene, *btuB*, was cloned on a 2.3-kilobase (kb) fragment (11). In this study is reported the nucleotide sequence of the fragment and the predicted amino acid sequence of the receptor protein.

A mutant has been described (1, 15) in which vitamin B_{12} transport was completely defective, although the binding of vitamin B_{12} and all other ligands to the receptor was normal in affinity and amount. The receptor functioned normally for transmission of the E. coli colicins and phage BF23, and vitamin B_{12} still protected cells from their lethality. Complementation tests indicated that the mutation in this strain was in *btuB*, and the suggestion was made that this mutation affected the domain on the receptor which was responsible for the response to the *tonB* product (3). The interaction of the *tonB* product with any of the receptor proteins dependent on its presence has not been demonstrated, and it is premature to claim that this *btuB451* mutation affects the TonB binding site, rather than a site involved in vitamin B_{12} translocation across the outer membrane. Whatever the biochemical basis for the phenotype of this mutant, the mutational lesion was determined by nucleotide sequencing. A single amino acid substitution near the amino terminus of BtuB was found.

**MATERIALS AND METHODS**

Bacterial strains and plasmids. Most of the bacterial strains and plasmids were described in the accompanying paper (11). Plasmid pKH3-3 carries a 2.3-kb insert which contains the *btuB* gene and was generated by partial *Sau3A* digestion; this plasmid was the source of DNA for determination of the sequence of the wild-type *btuB* gene. The region carrying the *btuB451* mutation was obtained as follows. Strain RK4775 (RK5173 *btuB451*) was lysogenized with phage λ *darg* (λ *darg* [btuB]) with selection for growth on 5 nM vitamin B_{12}. A phage lysate from pooled lysogens was prepared by thermal induction and lysis by the addition of CHCl_{3}. The lysate was used to infect strain RK4936A [Δ(*arg*-btuB)], which was obtained after excision of *btuB::TnlO*. Arg⁺ transductants were selected at 30°C. All Arg⁺ transductants were sensitive to phage BF23, and approximately 10% were unable to utilize 5 nM vitamin B_{12}, i.e., carried *btuB451*. One of these transductants produced phage upon induction and CHCl_{3} lysis and was used to prepare phage DNA. This phage DNA was digested with *HindIII* and *SalI*, and the appropriate 163-base-pair (bp) fragment was cloned into phage M13mp8.

**Genetic techniques.** Restriction endonuclease digestion, ligation, and transformation were described in the accompanying paper (11). The location of the wild-type allele of the *btuB451* mutation was determined by introducing into strain RK4775 (*btuB451*) derivatives of plasmid pH23-5 from which portions had been removed by restriction endonuclease-generated deletion. The ability of these transformants to give rise to Btu⁺ recombinants was determined.

DNA sequence determination. DNA sequence analysis was performed by the enzymatic method of Sanger et al. (26), using an α³²S-labeled deoxynucleoside triphosphate for l-
RESULTS

DNA sequence of btuB. The nucleotide sequence of the 2220-bp fragment carrying btuB was determined by cloning smaller restriction fragments of the insert into the mp8 or mp9 derivatives of phage M13 (17). The restriction map and regions sequenced are shown in Fig. 1. The nucleotide sequence and its translation are presented in Fig. 2. The sequence of both DNA strands was determined, except for regions of 28 residues (residues 491 to 518), 144 residues (residues 931 to 1075), and 10 residues (residues 1703 to 1712). In these three regions, several restriction fragments were sequenced, and no uncertainties were encountered. Examination of the six possible reading frames within the 2220-bp fragment revealed only one of significant length, extending from residues 195 to 2153. There were two potential ATG initiation codons, located at residues 270 and 312. It is likely that the codon at position 312 is the actual initiation site for the vitamin B12 receptor because it is preceded by a potential Shine-Dalgarno sequence (GTG GATG) (27) and has the sequence characteristics of translation initiation sites (29). This 1842-bp open reading frame encodes a polypeptide with 614 amino acids, the first 20 of which are very similar to the leader peptide (signal sequence) of other outer membrane proteins (18).

The putative leader peptide was followed by a polypeptide with a calculated molecular weight of 66,400, which is in good agreement with the size of the vitamin B12 receptor determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (11, 12, 25). The composition of the polypeptide determined from the DNA sequence was in excellent agreement with the amino acid composition determined for the purified receptor (12). The overall sequence of the protein was typical of other outer membrane proteins, in that it was not appreciably nonpolar (22% charged residues) and that there were only a few long stretches composed only of hydrophobic amino acids. There were no cysteine and only two methionine residues in the mature region. Tryptophan (2.3%), tyrosine (8.0%), histidine (1.8%), and threonine (9.1%) were present more frequently than in total E. coli proteins by at least a factor of two.

The codon usage for the precursor vitamin B12 receptor was very similar to the codon usage pattern of other weakly expressed E. coli proteins (10). All but 4 of the possible 60 codons (only one Cys residue is present) were used. This was in contrast to the situation with the more abundant outer membrane proteins which employ a limited range of codons corresponding to the more abundant tRNA species (6, 13).

The direction of translation of the btuB gene determined from the sequence agreed with that proposed from the properties of Tn1000 insertions, i.e., from the HindIII site to the PstI site (11). The promoter and upstream regulatory regions were carried on the sequenced fragment because the expression of the cloned gene was still subject to repression by vitamin B12 (11, 14). Four regions of palindromic structure were found in the region 5' to the coding segment (24), but their role cannot be assessed until the start site for transcription has been determined.

Sequence of the btuB451 mutation. The btuB451 mutation confers the BtuI (1), or BtuA (15) phenotype, in which the mutant strain lacks vitamin B12 uptake activity, although all other receptor functions remain normal. Plasmid pKH3-5 was able to complement the defect in vitamin B12 utilization in both rec+ and recA strains carrying the btuB451 mutation. Derivatives of this plasmid carrying deletions between restriction sites in btuB were constructed and introduced into the rec+ btuB451 strain. The formation of Btu+ recombinants was determined (Fig. 3). Several independent isolates of each construction were examined. Plasmids in which the HindIII fragment (e.g., pKH35-H1) or both SalI fragments (e.g., pKH35-S1) were deleted did not give rise to Btu+ recombinants. In contrast, plasmids deleted for the smaller SalI fragment (e.g., pKH35-S2) did give Btu+ recombinants. This marker rescue experiment indicates that the wild-type allele of btuB451 resides within the small segment between the HindIII and SalI sites.

The btuB451 mutation was transferred onto phage λ darg13 by homologous recombination. The 163-bp HindIII-SalI fragment was isolated from a digest of this phage DNA and was cloned into M13mp8. The nucleotide sequence of this fragment was determined in parallel with the corresponding fragment from the wild type (Fig. 4). The mutant differed from the wild type at only a single position. The mutation was associated with a T-to-C transition at residue 394 in Fig. 2. This nucleotide change predicts the substitution of proline for leucine at the eighth amino acid residue of the mature polypeptide (amino acid 28 in Fig. 2).

DISCUSSION

The nucleotide sequence of the 2.2-kb fragment, shown in the accompanying paper to contain the btuB gene (11), allowed prediction of the amino acid sequence of the vitamin B12 receptor protein. The composition of this polypeptide agreed very well with that reported for the colicin E3 receptor by Imajoh et al. (12). The size of the translated...
plypeptide was close to that observed for the vitamin B12 receptor, and the polypeptide started with a typical leader, or signal, sequence. Although the amino-terminal sequence of the mature polypeptide in the outer membrane has not yet been determined, analogy with the sites of leader peptidase action on other secreted proteins suggests that this cleavage occurs after the sequence Ala18-Trp29-Ala20 (18).

The vitamin B12 receptor is similar in several respects to other outer membrane proteins (6, 13, 19, 22). Like them,

FIG. 2. Nucleotide sequence of the btuB region.

FIG. 3. Localization of the btuB451 mutation by marker rescue. Plasmids carrying deletions in the btuB region were generated by cleavage with the indicated restriction enzyme and treatment with DNA ligase. Plasmids deleted for the regions shown were tested for their ability to give rise to Btu" recombinants in a strain carrying btuB451. Restriction enzyme cleavage sites are as follows: C, CiaI; E, EcoRI; H, HindIII; P, PstI; S, Sall.

FIG. 4. Determination of the nucleotide change in the btuB451 mutation. A portion of the sequencing gel obtained with the 163-bp HindIII-SalI fragments from the wild-type (left) and the btuB451 mutant (right) is shown. The region covers nucleotides 369 to 416 on the 3' strand.
There were seven stretches with 13 or more residues interspersed throughout these hydrophobic regions, showing that there are no long nonpolar a-helical segments spanning the bilayer, as is seen with some transmembrane proteins in the cytoplasmic membrane of bacteria or in the membranes of eucaryotic cells or viruses (8).

Chou-Fasman analysis (5) of secondary structure predicted extensive regions of a-structure comprising 55 to 60% of the length of the mature polypeptide (data not shown). The predicted content of a-helical structure was much lower, comprising the signal sequence and 6 to 10% of the mature portion. A high content of b-sheets has been found in other outer membrane proteins (9) and may be a general feature of them.

Considerable amino acid sequence homology exists within the family of porins OmpC, OmpF, and PhoE (13, 19, 22), but much less homology is apparent between the porins and other outer membrane proteins OmpA and LamB (6). Nikaido and Wu (21) have described short regions of partial homology that are located at similar, but not identical, sites within all these proteins. Visual comparison of the BtuB sequence with those of other outer membrane proteins revealed no appreciable overall homology and only weak homology to some of the common consensus regions found by Nikaido and Wu (21).

Examples are presented in Fig. 6. Perhaps the best fit was to homology region b, which was suggested to be involved in the protein export process. Whereas this region lies near the amino terminus of the mature sequence of other outer membrane proteins, in BtuB it was near the carboxyl end. Since the match in these regions is weak, proposals concerning their function would be of questionable significance. The major conclusion is that there are no strongly conserved sequences common to outer membrane proteins.

Since this is, to our knowledge, the first report of the sequence of a tonB-dependent or a vitamin B12-binding protein, it is not possible to use homologies to identify possible functional domains. However, examination of the

FIG. 5. Hydrophathy profile of the BtuB protein. The hydrophathy index was determined at a span setting of 13 with the algorithm presented by Kyte and Doolittle (16). Hydrophobic regions extend above the dotted line.

BtuB contains an appreciable content of charged amino acids (22%) and does not exhibit extensive segments of hydrophobic residues. Charged residues were evenly spaced and provide negative charge along the length of the protein. There were seven stretches with 13 or more uncharged residues. The hydrophobicity profile (16) for the protein confirmed that the protein was not appreciably nonpolar in overall composition, although there were two regions of significant hydrophobic character, one near the amino terminus (residues 80 to 160) and another near the carboxyl terminus (residues 570 to 600) (Fig. 5). The remainder of the protein revealed alternating regions of polar and nonpolar character. It is possible that the polypeptide chain spans the outer membrane bilayer numerous times. Polar or charged residues were interspersed throughout these hydrophobic regions, showing that there are no long nonpolar a-helical segments spanning the bilayer, as is seen with some transmembrane proteins in the cytoplasmic membrane of bacteria or in the membranes of eucaryotic cells or viruses (8).

Chou-Fasman analysis (5) of secondary structure predicted extensive regions of a-structure comprising 55 to 60% of the length of the mature polypeptide (data not shown). The predicted content of a-helical structure was much lower, comprising the signal sequence and 6 to 10% of the mature portion. A high content of b-sheets has been found in other outer membrane proteins (9) and may be a general feature of them.

Considerable amino acid sequence homology exists within the family of porins OmpC, OmpF, and PhoE (13, 19, 22), but much less homology is apparent between the porins and other outer membrane proteins OmpA and LamB (6). Nikaido and Wu (21) have described short regions of partial homology that are located at similar, but not identical, sites within all these proteins. Visual comparison of the BtuB sequence with those of other outer membrane proteins revealed no appreciable overall homology and only weak homology to some of the common consensus regions found by Nikaido and Wu (21). Examples are presented in Fig. 6. Perhaps the best fit was to homology region b, which was suggested to be involved in the protein export process. Whereas this region lies near the amino terminus of the mature sequence of other outer membrane proteins, in BtuB it was near the carboxyl end. Since the match in these regions is weak, proposals concerning their function would be of questionable significance. The major conclusion is that there are no strongly conserved sequences common to outer membrane proteins.

Since this is, to our knowledge, the first report of the sequence of a tonB-dependent or a vitamin B12-binding protein, it is not possible to use homologies to identify possible functional domains. However, examination of the

FIG. 6. Homology between BtuB and other outer membrane proteins. The regions of homology (b and d) in the major outer membrane proteins are those identified by Nikaido and Wu (21). Alignments to provide best fit with these regions are shown and include deletion or insertion of an amino acid. The coordinates are for the processed form of the protein.
distribution of certain amino acids throughout the sequence revealed sections which might be ligand-binding sites. Particularly noteworthy was the case of tyrosine. There were six to 20 amino acid regions which contained at least four Tyr residues (amino acids 241 to 256, 287 to 307, 403 to 425, 455 to 474, 539 to 558, and 592 to 613). In comparison, the porins have roughly the same tyrosine content (13, 19, 22) but, at most, only two tyrosine clusters of the type seen in BtuB. Five of the six regions were very hydrophilic in character and contained at least five charged or hydroxyl-containing residues. One of the clusters was at the carboxy terminus of the protein, which has been implicated in vitamin B₁₂ binding based on the altered vitamin B₁₂ binding and transport properties of the mutant generated by deletion from the PsrI site (11). This altered receptor did allow entry of phage BF23, but the binding of this phage was not blocked by vitamin B₁₂.

The six tyrosine-rich regions were compared for homology at the nucleotide and amino acid levels. Although the clusters were rich in Asp, Gly, and Ser, the longest amino acid sequence shared by any two clusters was only three amino acids long. There was no significant homology at the nucleic acid level. Thus, it is unlikely that these clusters arose by genetic duplication. Experiments to identify binding domains by isolation of mutations in the cloned gene are in progress.

Also of interest is the amino-terminal region of the putative processed protein. This region was affected by the btbU451 mutation, which blocked vitamin B₁₂ uptake but not any other receptor functions (1). One possibility is that this domain is involved in interaction with the tonB product. Comparison of this region with that of other tonB-dependent receptors should prove to be useful. If this domain is involved in interaction with TonB, then this region should face the periplasm. Future studies are directed toward examination of the topology of the BtuB protein in the membrane and its interaction with the tonB product.

ACKNOWLEDGMENTS

We are extremely indebted to M. M. Smith for providing the computer facilities and instruction needed for this study. This study was supported by research grant GM19078 from the National Institute of General Medical Sciences. K.H. was supported by a stipend from the Deutsche Forschungsgemeinschaft.

LITERATURE CITED