Cloning and Expression in *Escherichia coli* of a Gene, *hup*,
Encoding the Histone-like Protein HU of *Bifidobacterium longum*

Akio TAKEUCHI, Hajime MATSUMURA, and Yasunobu KANO

Department of Molecular Genetics, Institute of Molecular and Cellular Biology for Pharmaceutical Sciences, Kyoto Pharmaceutical University, 1 Shichono-cho, Misasagi, Yamashina-ku, Kyoto 607-8412, Japan

Received October 9, 2001; Accepted November 8, 2001

A genomic DNA library of *Bifidobacterium longum* ATCC15707 was transfected into an *Escherichia coli* strain deficient in both HU and IHF, the growth of which is cold-sensitive because of the deficiency in these proteins. Cold-resistant colonies were selected and the DNA was cloned and sequenced. A polypeptide consisted of 93 amino acids, a predicted molecular mass of 9983 Da with an isoelectric point of 10.35, was deduced from an orf in the middle of the DNA fragment. The amino acid sequence was highly similar to HU family proteins, and 26 aas of N terminal was identical to a histone-like protein, HBl, a HU family protein of *B. longum*. Incapabilities of Mu phage propagation in an *E. coli* mutant deficient in HU or IHF could be suppressed by DNA bearing this orf. These results showed that the orf is a gene *hup* encoding HBl, a histone-like protein HU of *B. longum*.

Key words: *hup* gene encoding histone-like protein HBl; *Bifidobacterium longum*; *Escherichia coli*; DNA-binding proteins HU and IHF; propagation of Mu phage

HU is a histone-like protein that binds DNA nonspecifically. HU is a small, basic, heat-stable protein widely found in prokaryotes, and is expressed highly and constitutively. HU is a major component of nuclear proteins and plays an important role in the formation of higher-order DNA structure, restraining negative supercoils in it. The resulting alteration of DNA structure affects the regulation of initiation of DNA replication, transposition of Mu phage, cell division, site-specific DNA recombination, and transcription. HU generally exists as a homodimer, although it consists of two distinct HU subunits HUα (encoded from *hupA*) and HUβ (encoded from *hupB*) in *E. coli*, *Salmonella typhimurium*, and *Serratia marcescens*.

Integration host factor (IHF), is a HU family histone-like protein highly similar to HU structurally and functionally. A heterodimer consisting of IHFα (encoded from *himA*) and IHFβ (encoded from *himD/him*) binds DNA specifically and bends it sharply. HU compensates for the lack of IHF, in vivo, on cell growth, replication of some oriC-plasmids, and integration of λ phage into the host chromosomal att site.

Bifidobacteria are Gram-positive lactic acid-producing bacteria constituting a major part of the intestinal microflora in humans and other mammals. They have received much attention in the fields of food industry and medical science, although very little knowledge was accumulated on their molecular genetics. Recently, a DNA-binding HU family protein, HBl, was isolated from *B. longum* ATCC15707 and the N-terminal 26 aa sequence was analyzed. HBl bound to DNA with an affinity higher than that of HUα and HUβ, and restrained negative supercoils in DNA much more effectively than *E. coli* HU. Here we will report the DNA structure of a *B. longum hup* gene encoding HBl, the complete amino acid sequence of the protein, and its function in *E. coli*.

Materials and Methods

*Bacterial culture. B. longum* ATCC15707 was cultured anaerobically at 37°C to log phase in Briggs broth with glucose replaced by 2.0% lactose. *E. coli* K12 derivatives (Table 1) were cultured in LB broth (10 g of Bacto tryptone, 5 g of yeast extract, 5 g of NaCl, and 0.1% glucose/liter), and colonies were formed on LB broth containing 1.5% agar. If necessary, drugs (50 μg of Ap and/or 12.5 μg of Tc/ml) were added. Heat-induction and plaque formation of Mucts62, which is a temperature-inducible repressor mutant of Mu phage, were done as described previously. YK1100 and mH3 are derivatives of *E. coli* K12. YK1340 is a HU-deficient derivative (*hupA16 hupB11*) of YK1100, and YK1344 is a Mucts62 lyso-
Table 1. Escherichia coli Strains Used

<table>
<thead>
<tr>
<th>Strain and plasmid</th>
<th>Relevant genotype</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>mH3</td>
<td>hsdR(hsdM&lt;sup&gt;+&lt;/sup&gt;) rpsL trpC 9830</td>
<td>Kano et al. &lt;sup&gt;31&lt;/sup&gt;</td>
</tr>
<tr>
<td>YK1100</td>
<td></td>
<td>Wada et al. &lt;sup&gt;32&lt;/sup&gt;</td>
</tr>
<tr>
<td>YK1340</td>
<td>YK1100 hupA16(Km&lt;sup&gt;+&lt;/sup&gt;) hupB11(Cm&lt;sup&gt;+&lt;/sup&gt;)</td>
<td>Wada et al. &lt;sup&gt;33&lt;/sup&gt;</td>
</tr>
<tr>
<td>YK1344</td>
<td>YK1100 hupA16(Km&lt;sup&gt;+&lt;/sup&gt;) hupB11(Cm&lt;sup&gt;+&lt;/sup&gt;) (Mu&lt;sup&gt;ts62&lt;/sup&gt;)</td>
<td>Goshima et al. &lt;sup&gt;34&lt;/sup&gt;</td>
</tr>
<tr>
<td>YK3552</td>
<td>YK1100 Δ82himA::Tn10 (Mu&lt;sup&gt;ts62&lt;/sup&gt;)</td>
<td>This work</td>
</tr>
<tr>
<td>YK3556</td>
<td>YK1100 Δ3hip::cat (Mu&lt;sup&gt;ts62&lt;/sup&gt;)</td>
<td>This work</td>
</tr>
<tr>
<td>YK2741</td>
<td>YK1100 hupA16(Km&lt;sup&gt;+&lt;/sup&gt;) hupB11(Cm&lt;sup&gt;+&lt;/sup&gt;) Δ82[himA::Tn10]</td>
<td>Kano and Imanoto &lt;sup&gt;35&lt;/sup&gt;</td>
</tr>
<tr>
<td>pFTC1</td>
<td>Tc&lt;sup&gt;+&lt;/sup&gt;</td>
<td>miniF replicon</td>
</tr>
<tr>
<td>pE313</td>
<td>Ap&lt;sup&gt;+&lt;/sup&gt; hip himD&lt;sup&gt;+&lt;/sup&gt;</td>
<td>pBR322 replicon</td>
</tr>
<tr>
<td>pYK20</td>
<td>Ap&lt;sup&gt;+&lt;/sup&gt; hipA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>pBR322 replicon</td>
</tr>
<tr>
<td>pBLHU15</td>
<td>Ap&lt;sup&gt;+&lt;/sup&gt; hup&lt;sup&gt;+&lt;/sup&gt; of B. longum</td>
<td>pBR322 replicon</td>
</tr>
</tbody>
</table>

Electroporation. Tansformation were done by electroporation. Cells were cultured in 5 ml of LB broth, collected as described by Dower et al.,<sup>23</sup> and suspended with 0.25 ml of 10% glycerol, finally. Electroporation was done at 2.5 kV/cm and 25 μF at 200 Ω parallel resistance by a GenePulser supplied by Bio-Rad Laboratories (USA) using 1 μl of DNA and 40 μl of cell suspension.

Heat induction of Mu<sup>ts62</sup> phage. E. coli cells lysogenic for Mu<sup>ts62</sup> were cultured at 30°C and heat-treated for 25 min at 42°C. After being shaken for 120 min at 37°C, cells were centrifuged, the supernatant was treated with chloroform, and pfu were counted on YK1100 by the method of Kano et al.<sup>33</sup>

Cloning of B. longum DNA. Genomic DNA of B. longum ATCC15707 was extracted as described by Wada et al.,<sup>32</sup> then digested with HindIII and inserted into the HindIII site of pBR322. Resulting recombinant plasmids were transfected in mH3, Ap<sup>+</sup> colonies were selected, and then Tc<sup>+</sup> colonies were screened. Each of 100 Ap<sup>+</sup>-Tc<sup>+</sup> colonies was suspended in 20 test tubes each, each containing 5 ml of L-broth, and the plasmid mixture was extracted from these test tubes by rapid alkaline extraction.<sup>29</sup> E. coli YK2741 was transformed with these plasmids, and cold-resistant colonies were selected at 27°C on LB agar plates with 50 μg Ap/ml added. YK1340 was transformed with the plasmids extracted from these cold-resistant colonies, and the resulting Ap<sup>+</sup> cells were then transformed with the miniF plasmid pFTC1<sup>33</sup> to make HU activity in these transformants apparent.

DNA sequencing. A cloned DNA fragment was recovered from pBLHU15 by agarose gel electrophoresis after digestion with HindIII. The DNA fragment was digested with PvuII, PstI, or StuI, and subcloned in the multicloning site of pUC18. Sequencing was done by an ALFred Auto Sequencer and AutoRead Sequencing Kit (Pharmacia Biotech), and surveyed by the DNASIS ver 2.2 program. The ribosome binding region (SD) was searched for using 16S rRNA sequences of B. longum ATCC15707 (Accession Number 173875).

Results

Cloning of 2.3 kb-DNA fragment containing B. longum hup gene

For an E. coli mutant, deletion of both HU<sub>α</sub> and HU<sub>β</sub> is not lethal and it grows normally at 37°C but with a decreased rate at temperatures lower than 30°C compared to the wt strain.<sup>8</sup> E. coli himA or hip/himD mutation, which encodes subunit proteins IHF<sub>α</sub> or IHF<sub>β</sub>, respectively,<sup>25</sup> does not affect the cell growth rate. However, a mutant deficient in HU together with IHF becomes cold sensitive.<sup>14</sup> It was, therefore, expected that the cold sensitivity of HU-IHF-deficient E. coli may be suppressed by HU-family protein HBl of B. longum if it is functional in E. coli. We thus tried to clone a hup gene encoding HBl by selecting cold-resistant transformants of HU-IHF-deficient E. coli after transfection of a genomic library of B. longum ATCC15707. We could obtain three independent Ap<sup>+</sup> transformants of YK2741 with cold-resistant phenotype, and found that the cloned DNA fragments were 2.3-kb in size by agarose gel electrophoresis for all the three plasmids extracted from these transformants, and the same restriction maps were constructed using three restriction enzymes, PvuII, StuI, and PstI. We thus named one of them pBLHU15 (Fig. 1).

Structural analysis of cloned DNA

It has been reported that HU is essential for propagation of Mu phage in E. coli.<sup>3</sup> Therefore, if suppression of the cold-sensitive phenotype and inability of miniF replication of YK2741 were caused by HU protein of B. longum synthesized from pBLHU15, the inability of Mu phage growth in HU-
Fig. 1. Structure of pBLHU15, and Physical and Functional Map of 2.3 kb-DNA Fragment Cloned from *B. longum*.

The numerals under the bar indicates the length of the restriction fragment in bp. A stippled square shows orf93, and an arrow under it indicates the direction of its expression. *E. coli* YK1344 was transformed with pUCBLs, and Mu*cts62* was heat-induced. Pfu were counted and are presented at the right side of this figure. +: pfu higher than $3 \times 10^6$ W/ml; -: pfu lower than $2.6 \times 10^5$ W/ml.

Fig. 2. Nucleotide Sequence of *B. longum hup* Gene and Its Flanking Region, and the Primary Structure of HBl Protein.

Nucleotide sequence was numbered at the left. The predicted aa sequence of HBl is shown below the DNA sequence written in bold letters and numbered at the right. The proposed promoter (−10, −35) and a ribosomal binding sequence were underlined with a single line and a double line, respectively. The regions of dyad symmetry are indicated by horizontal arrows and the following T-rich sequence is shown by a dotted line. Rectangular box shows *Sst* I site.

deficient *E. coli* may also be suppressed by this protein.

To prove that the protein synthesized from the cloned DNA is HU, YK1344 was transformed with pBLHU15. When the transformant was heat-treated at 42°C, a high level of Mucts62 was induced, showing that a HU-like protein was synthesized from pBLHU15 (see Table 3, line 3).

To locate the gene that synthesizes an HU-like protein, the cloned 2.3-kb DNA was digested with *Pvu* II and *Sst* I. The resulting 5 restriction fragments were subcloned in pUC18, generating pUCBL1 to pUCBL6, and YK1344 was transformed with these plasmids. As shown in Fig. 1, Mucts62 was not heat-induced in YK1344 harboring plasmids pUCBL2, pUCBL3, and pUCBL4, which carry DNA fragments derived from the 1.3-kb region at the left side, and pUCBL5 carrying the 1.0-kb region at the right side of the cloned DNA. On the contrary, high pfu was produced after heat-induction of a transformant with pUCBL6 in which the 1.6-kb DNA at the right side was subcloned, as well as pUCBL1 in which the whole region of 2.3-kb DNA was cloned in pUC18. These results suggested that *hup* gene lies straddling the 2nd *Sst* I site from the right end of the fragment. We, then, sequenced the cloned DNA fragment and the orf was searched for in a segment straddling this *Sst* I site.
Figure 2 shows the nt sequence of a region straddling the $Stul$ site (Accession Number AB072446). An open reading frame consisting of 93 aas (orf93) with an initiation codon ATG and a termination codon TGA at positions 193 and 472, respectively, was identified in this region. The orf has a G+C content of 61.3 mol%, similar to that of $B. longum$ genomic DNA. A putative ribosome-binding site (AGAAGG) was located 8 nt upstream from the start codon, five nts of which are complementary to the 3′ terminal region of $B. longum$ 16S rRNA, and is similar to AGGAGG which is one of typical ribosome binding sequences of $E. coli$. Nucleotide sequence, TAGTAT at 66 nt or TATCAT at 63 nt and TTCG-?

The structural gene encoding histone-like protein HBl of $B. longum$ shows especially high, 83 mol% similarity to HU of $E. coli$ and $P. aeruginosa$, respectively. Conservation of aa sequence at the positions 58–75 which is in the arm region (DNA-binding domain) showed especially high, 83% and 67%, identities to $B. subtilis$ $HU$ and $E. coli$ $HU$, respectively. Furthermore, the N-terminal 26 aa sequence of HBl has been reported. The aa sequence deduced from orf93 is identical (Fig. 3) to, and the amino acid composition is almost the same (Table 2) as that of HBl, a structural gene encoding the histone-like protein HBl of $B. longum$ ATCC15707.

HBl compensates for lack of IHF on Mu phage growth

HU compensates for lack of IHF on several physiological functions in vivo (see Introduction). It is, therefore, interesting to know whether HBl can suppress the incapability of heat-induction of Mu ts62 in $E. coli$ deficient in IHF. As shown in Table 3, YK1344 harboring pBLHU15 could induce Mu ts62 to $5.4 \times 10^6$ pfu/ml which is ca. 4,500 fold higher than that seen in YK1344 harboring pBR322 $(1.2 \times 10^6$ pfu/ml) and almost the same level as that seen in YK1344 harboring pYK20 bearing $E. coli$ hupA $(1.0 \times 10^6$ pfu/ml) (lines 1 to 3), showing that HBl complemented the HU deficiency in $E. coli$. Next, $E. coli$ deficient in IHFβ $(YK3556)$ or IHFα $(YK3552)$ harboring plasmids were heat-treated. Both pYK20 and pBLHU15

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Orf93 Number of residues</th>
<th>HBl mol%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>13</td>
<td>13.97</td>
</tr>
<tr>
<td>Arg</td>
<td>6</td>
<td>6.45</td>
</tr>
<tr>
<td>Asp</td>
<td>3</td>
<td>3.22</td>
</tr>
<tr>
<td>Asn</td>
<td>4</td>
<td>4.30</td>
</tr>
<tr>
<td>Glu</td>
<td>6</td>
<td>6.45</td>
</tr>
<tr>
<td>Gln</td>
<td>4</td>
<td>4.30</td>
</tr>
<tr>
<td>Gly</td>
<td>7</td>
<td>7.52</td>
</tr>
<tr>
<td>His</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Ile</td>
<td>4</td>
<td>4.30</td>
</tr>
<tr>
<td>Leu</td>
<td>7</td>
<td>7.52</td>
</tr>
<tr>
<td>Lys</td>
<td>9</td>
<td>9.67</td>
</tr>
<tr>
<td>Met</td>
<td>2</td>
<td>1.07</td>
</tr>
<tr>
<td>Phe</td>
<td>3</td>
<td>3.22</td>
</tr>
<tr>
<td>Pro</td>
<td>3</td>
<td>3.22</td>
</tr>
<tr>
<td>Ser</td>
<td>8</td>
<td>8.60</td>
</tr>
<tr>
<td>Thr</td>
<td>5</td>
<td>5.37</td>
</tr>
<tr>
<td>Tyr</td>
<td>2</td>
<td>2.15</td>
</tr>
<tr>
<td>Val</td>
<td>7</td>
<td>7.52</td>
</tr>
<tr>
<td>Trp</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Cys</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Amino acid composition of Orf93 was expressed as the number of amino acid residues per mol protein and mol%. Mol% of HBl was calculated according to the amino acid numbers reported by Gosho et al. 17) Asp + Asn, **; Glu + Gln, nd; not determined.

IHFβ (YK3556) or IHFα (YK3552) harboring plasmids were heat-treated. Both pYK20 and pBLHU15
stimulated Mucts62 propagation 10-fold and 5-fold, respectively, in YK3556 under the condition where 200-fold stimulation was observed with pE31, which carries wt himD/hip (lines 4 to 7), showing that HBl complemented the IHFβ deficiency in E. coli. Three-fold and 13-fold increases in Mucts62 production were also observed with pYK20 and pBLHU15, respectively, in YK3552 (lines 8 to 11), also showing that HBl complemented the IHFα deficiency. It was concluded from these results that orf93 was a gene hup encoding HU family protein HBl, and the protein was translated and functioned in E. coli.

Discussion

The B. longum hup gene encoding the histone-like HU family protein HBl was cloned and expressed in E. coli. DNA sequence analysis predicted a promoter, a ribosome binding site, a structural gene, and a transcriptional terminator. The ribosome binding site was highly complementary to 16s-rRNAs of B. longum and E. coli. It is, therefore, likely that HBl was translated in E. coli from the authentic initiation codon AUG of the gene on pBLHU15, and could suppress cold-sensitivity of cell growth and incapability of miniF and Mu phage replication caused by deficiency in histone-like proteins of E. coli.

B. longum belongs to the high GC bacteria. In fact, our sequence data revealed that the B. longum hup structural gene consisted of 61.3% GC, which was higher than that seen in E. coli hup (43.8% and 49.1% in hupA and hupB, respectively). In addition, GC usage at the 3rd letter of codons was much higher in B. longum hup (78.7%) than E. coli hup (44.0% and 37.8% in hupA and hupB, respectively). It is imaginable from these observations that a gene with low GC content is translated inefficiently in B. longum.

It is generally believed that Gram-positive bacteria do not possess IHF in contrast to the Gram-negative bacteria in which it plays an important role in the physiology. We previously reported that, in in vitro, HBl binds DNA with much higher affinity and produces negative supercoils more efficiently than E. coli HU heterodimer. The efficient introduction of supercoils in DNA by HBl may cause sharp bending of the chromosomal DNA, as IHF does. This might be a reason why HBl could substitute for lack of IHFα on Mu phage propagation in E. coli more efficiently than E. coli HU.

We reported previously that the level of HBl in B. longum was high and comparable to that of E. coli HU observed in E. coli. It was, therefore, expected that the regulatory region of B. longum hup is useful for construction of a high efficient expression vector of B. longum by fusing a structural gene to it and integrated into a plasmid.

Acknowledgments

We thank Dr. K. Kohno for helpful technical comments.

References


