Isolation of *Borrelia burgdorferi* genes encoding homologues of DNA-binding protein HU and ribosomal protein S20

Kit Tilly,¹ John Fuhrman,¹ Jim Campbell¹ and D. Scott Samuels^{2†}

Author for correspondence: Kit Tilly. Tel: +1 406 363 9239. Fax: +1 406 363 9204. e-mail: ktully@nih.gov

Laboratory of Microbial Structure and Function¹ and Microscopy Branch², Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, Hamilton, MT 59840, USA

Linear DNA with covalently closed ends is the predominant form of DNA in the spirochaete *Borrelia burgdorferi*. All bacteria examined to date have small DNA-binding proteins related to the *Escherichia coli* IHF and HU proteins that appear to play roles in DNA compaction and replication, but such proteins had not been isolated from bacteria with linear genomes. We found a single gene in *B. burgdorferi* (named *hbb*) whose product (named Hbb) complements the defects for λ DNA packaging found in *E. coli* strains mutant in the genes for IHF and HU. The sequence of the predicted *B. burgdorferi* protein is similar to those of HU and IHF-like proteins in other bacteria. The gene appears to be in an operon with the order *rpsT*–*hbb*–*orfH*, where the *rpsT* gene is a homologue of the *E. coli* gene encoding ribosomal protein S20 and the *orfH* gene encodes a protein of unknown function. This operon is located upstream of the previously identified *B. burgdorferi* *rho* homologue.

Keywords: Lyme disease, *Borrelia burgdorferi*, integration host factor (IHF), HU protein

INTRODUCTION

Borrelia burgdorferi, the spirochaete that causes Lyme disease, has an unusual genome composed of linear and circular DNA molecules (Barbour & Garon, 1987; Baril *et al.*, 1989; Ferdows & Barbour, 1989). The linear molecules have covalently closed ends (hairpins). The only non-borrelial example of such structures among the prokaryotes is the prophage form of the coliphage N15, which exists as a linear plasmid (Ravin & Shulga, 1970; Svarchevsky & Rybchin, 1984). The left and right telomeres of the 16 kb linear plasmid of *B. burgdorferi* are identical at 16 of the 19 terminal bp (Hinnebusch *et al.*, 1990). The left end also has multiple direct repeats and a large inverted repeat (Hinnebusch & Barbour, 1991; Hinnebusch *et al.*, 1990). These DNA sequences migrate aberrantly through acrylamide gels, suggesting that the DNA is naturally bent (Hinnebusch, 1991). About 250 bp from the left end is a consensus sequence for binding of *Escherichia coli* integration host factor (IHF), a small DNA-binding protein (Hinnebusch, 1991).

IHF was discovered in *E. coli* because it is essential for integration of the bacteriophage λ prophage into the

bacterial chromosome (Kikuchi *et al.*, 1985; Miller *et al.*, 1979; reviewed by Friedman, 1988). Subsequently, it was found to be a heterodimer of the products of the homologous genes *himA* (Friedman, 1988; Mechulam *et al.*, 1985; Miller, 1984) and *hip* (*bimD*) (Flamm & Weisberg, 1985). IHF has a molecular mass of approximately 20 kDa for the heterodimer and binds to specific DNA sites, inducing a sharp (> 90°) bend (Kosturko *et al.*, 1989; Thompson & Landy, 1988). IHF binding affects binding of other proteins to the DNA, either increasing or decreasing levels of transcription, plasmid partitioning and DNA inversion (reviewed by Friedman, 1988).

E. coli has a second small DNA-binding protein, HU, which lacks sequence specificity (reviewed by Drlica & Rouviere-Yaniv, 1987; Schmid, 1990). The protein is also a heterodimer, whose components are encoded by the *hupA* and *hupB* genes (Kano *et al.*, 1985, 1987). These genes are homologous to each other as well as to the *himA* and *hip* genes. The HU protein appears to prefer binding to kinked DNA and four-way junctions, which resemble recombination intermediates (Bonney *et al.*, 1994; Pontiggia *et al.*, 1993), and HU will bend linear substrates. HU is able to compact DNA but *E. coli* does not contain enough HU to condense its entire genome (reviewed by Drlica & Rouviere-Yaniv, 1987). Rather, HU-like proteins have been proposed to be 'DNA chaperones' (Ner *et al.*, 1994; Travers *et al.*, 1994), which curve cellular DNA into configurations favourable for assembly of other

† Present address: Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA.

The GenBank accession numbers for the sequences reported in this paper are L07656 (*rho* gene) and U35673 (the new sequence).

protein-DNA complexes, analogous to the eukaryotic high-mobility group (HMG) proteins.

Other bacteria, phage, plastids and a virus have homologues of IHF and HU (see Oberto & Rouviere-Yaniv, 1996, for a complete listing). Some bacteria, including *E. coli*, appear to have both IHF- and HU-like proteins whereas others have only HU-like proteins, composed of either homotypic or heterotypic dimers. The single gene for a small DNA-binding protein in *Bacillus subtilis* (HBSu) appears to be essential (Micka *et al.*, 1991; Micka & Marahiel, 1992), and the same may be true for the other species with single genes. In contrast, *E. coli* mutants lacking both IHF and HU are viable, although they grow poorly (Kano & Imamoto, 1990).

The crystal structure of the *Bacillus stearothermophilus* HU protein has been determined (Tanaka *et al.*, 1984). It is composed of three clustered α -helices and an arm defined by antiparallel β -sheets. The disordered end of the arm is the DNA-binding domain, recognizing the minor groove of the DNA. Support for this model has come from footprinting and mutagenesis studies of the IHF protein (Lee *et al.*, 1992; Mengeritsky *et al.*, 1993; Yang & Nash, 1989).

We are interested in the mechanism of replication of linear DNA molecules in borreliae, so we speculated that *B. burgdorferi* might have unusual small DNA-binding proteins, which could participate in telomeric structure, linear plasmid replication or DNA partitioning. We used a previously described selection (Haluzi *et al.*, 1991) to isolate a gene that complemented the λ DNA packaging defects found in *E. coli* *himA* and *hip* mutants (Mendelson *et al.*, 1991). We describe the cloning, sequencing, complementation and transcriptional analysis of this gene, which we name *hbb*. This gene encodes the protein Hbb, which appears to be the HU- or IHF-like protein in *B. burgdorferi*.

METHODS

Bacterial strains. *Escherichia coli* strains used in this study are described in Table 1. The source of DNA for the λ ZAPII library was *Borrelia burgdorferi* isolate B31 (high passage), ATCC 35210.

Genetic methods. *E. coli* strains were grown in L broth (10 g tryptone, 5 g yeast extract and 5 g NaCl l⁻¹), with the addition of 0.2% maltose when used for bacteriophage λ plating or 100 μ g ampicillin ml⁻¹ when carrying plasmids. Transformations and growth of phage stocks in liquid were as described by

Silhavy *et al.* (1984). Phage yield determinations were as previously described by Georgopoulos & Herskowitz (1971).

Library construction and manipulations. *B. burgdorferi* B31 DNA was partially digested with *EcoRI** or *Tsp509I* and size-fractionated by agarose gel electrophoresis. Fragments from about 1 to 8 kb were eluted, pooled, ligated with *EcoRI*-digested λ ZAPII DNA (Stratagene), and plated on *E. coli* strain XL1-Blue as recommended by the manufacturer (Stratagene). Conversion to plasmids via *in vivo* excision was as recommended by the manufacturer (Stratagene).

Molecular analysis. Plaque hybridizations (Sambrook *et al.*, 1989) were performed using Biotrans filters (ICN Biomedicals). Probes were labelled by random priming using a kit from Life Technologies and [α -³²P]dATP (Dupont-NEN) as recommended by the supplier. DNA sequencing was by the chain-termination method (Sanger *et al.*, 1977), using a kit from United States Biochemical and [³²S]dATP α S (Dupont-NEN or Amersham). For primer extension analysis (modification of procedure of Stern *et al.*, 1986), various oligonucleotides were synthesized on a Cyclone Plus Synthesizer (Milligen) and ³²P-labelled using polynucleotide kinase (New England Biolabs) and [γ -³²P]ATP (Dupont-NEN). *B. burgdorferi* RNA was prepared as described by Ausubel *et al.* (1989). Primer (2 ng) was mixed with 5–30 μ g RNA in 50 mM Tris (pH 8.3), 60 mM NaCl, 10 mM dithiothreitol and 1 mM EDTA and heated to 80 °C for 3 min in a Perkin Elmer Cetus DNA Thermal Cycler. The samples were cooled to 45 °C over 45 min and then incubated at 45 °C for 1 h with 30 U avian myeloblastosis virus reverse transcriptase (USB) and 110 U ribonuclease inhibitor (USB) in 8.3 mM Tris (pH 8.3), 10 mM NaCl, 16.7 mM dithiothreitol, 10 mM MgCl₂ and 1.7 mM of all four dNTPs. Ribonuclease A was added to a concentration of 0.6 mg ml⁻¹ and the samples were incubated for 20 min at 37 °C. The samples were diluted, extracted with phenol/chloroform (50:50, v/v), and precipitated with sodium acetate and ethanol. Pellets were resuspended in 10 μ l 0.5 $\times$ Sequenase stop solution (USB) and loaded on 6% (w/v) acrylamide sequencing gels for analysis. Northern blots of 1.4% (w/v) agarose formaldehyde gels were as described by Sambrook *et al.* (1989).

Subcloning of *hbb* and *orfH* genes. The *hbb* gene was subcloned by digesting pKK41 (Fig. 1) with *Bgl*II and *Pst*I (New England Biolabs) and isolating the 650 bp fragment from an agarose gel using NA45 membrane (Schleicher and Schuell). The purified fragment was ligated with *Pst*I and *Bam*HI-digested pBluescript SK(+). The *orfH* fragment was isolated on a 785 bp *Dra*I fragment and ligated with *Sma*I-digested pBluescript SK(+) using similar methods. Each insert was crossed into λ ZAPII by growing a stock of insert-free (blue-plaque-forming) phage in a *rec*⁺ strain carrying the appropriate plasmid. The resultant lysates were plated for individual plaques on plates containing X-Gal and stocks of several white-plaque-

Table 1. *E. coli* strains used

Strain	Relevant genotype	Source
N99	<i>galK</i>	Haluzi <i>et al.</i> (1991)
A5179	N99 <i>hupA16::kan hupB11::cat hip-157</i> Tn10	Haluzi <i>et al.</i> (1991)
A5427	N99 <i>hupA16::kan hupB11::cat ΔhimA82</i> Tn10	Haluzi <i>et al.</i> (1991)
A5477	N99 <i>hupA16::kan hupB11::cat ΔhimA82 hip-157</i>	Haluzi <i>et al.</i> (1991)
A5196	N99 <i>hupA16::kan hupB11::cat</i>	Haluzi <i>et al.</i> (1991)
XL1-Blue	<i>hsdR17 recA1lac⁻/F' proAB lacI^q lacZAM15</i> Tn10	Stratagene

(a) HU and IHF

α1		α2		β1		β2		DNA-binding arm		β3		α3	
Hbb	MSFSRRPKVT	KSDIVDQIAL	NIKNNLNKLE	KKYIRLVIDA	FFEEKLSNLC	SNNVIEFRSF	GTFEVAKRRKG	RLNARNPQTG	EYVXVLDHHV	AYFRPGKDLK	ERVWGIKG*		
HCj	m.	a.fisl.v.q	----tag.t	.dattat..	vistitd.v.a	kqds.s.ig.	...stqe.aa	..e.v.s..	kti..patr.	.k.kv..n..	.a.akas.kk	kk*	
HTT	aakkt..	.a.l...v.q	----atg.k	lifvkanv..	llakveea.a	ngskvqltg.	a	..tgvk.g.k	.ki..ipaty	pa.k...a..	dk.k*		
HBSu	mn	.telinav.e	----ase.s	.datkav.s	v.dtilda.k	ngdk.qlig.	.n...e.sa	..-kg.....	.eieipask.	pa.k...a..	da.k*		
HRcaIHFB	mi	r.eliak.e	----e.ph.f	qrdvekivnt	l...ileama	rgdrv.l.g.	.a.s.k..da	..-tg...r..	ts.a.dek..	pf.kt...l.r	d.ln.gee*		
HBca	mn	.telinav.e	----tsqls	.datkav..	v.dsitate.r	kgdkvqlig.	.n...e.aa	..-kg.....	.emeipask.	pa.k...a..	da.k*		
EcIHFB	m.	.elierl.t	----qqship	a.tvedavke	ml.hma.t.a	qger..i.g.	.s.sihy.ap	..-tg...k..	dk.elegky.	ph.k...e.r	d.an..y.*		
EcHU-1	mn	.qlii.k..a	----gadis	.aaagraal..	liasvtes.k	egddvalvg.	...a.ke.aa	..-tg.....	keitiaaak.	ps...a..a..	da.n*		
EcIHFA	mal.	.aemseyifd	----x.g.s	.rdakelvel	...irra.e	ngeqvklsg.	.n.dl.dknq	..-pg...k..	.dipitarr.	vt...qk..	s..enaspkd	e*	
HAT	mn	.nel.sav.e	----kag.t	.adaasav..	v.tvq.e.k	nggd.rlag.	.s.s.sr.ea	s-kg...s..	ae.diparn.	pk.sa..gk.	da.ns*		
EcHU-2	mn	.tqlii.v.e	----kae.s	.tgakaales	tlaaite.s.k	egdavlvg.	...k.nh.æ	..-tg.....	kel.iaaan.	pa.vs...a..	da.k*		
HRm	mn	.nel.aav.d	----kag.s	.adassav..	v..tiqge.k	nggd.rlvq.	.n.svs.rea	s-kg...s..	ae.diparn.	pk.ta..g..	da.n*		
HCP	mn	.aelvaav.d	----kag.s	.adassav..	v..svvea.e	kgkvqlvg.	...t.e.aa	..-æg...r.k	.vinipatt.	pv..ka..ef.	dk.nk*		
HRL53	mn	.nel.sav.e	----kag.t	.sdaasav..	v.dvqvæ.k	nkqd.rlag.	.s.t.sh.aa	t-kg...s..	aevdiparn.	pk.t...q..	da.ng*		
HTa	m.g	.i.æiskev.k	----kantt	q.va.t..ks	.ld.iv.ean	ggqk.nlag.	.i..r.tqgp	..-k.....	k.vie.pskkk	fv..asski.	yqq*		
HAPr	mn	.tqlii.f..e	----kad.s	.aqakaale	.tldgvtda.k	egdqvqlig.	...k.nh.aa	..-tg...k..	aelqiaaan.	pa.va...a..	da.k*		
HRcaIHFA	msektl.	rm.lseavfr	----evg.s	rnesaqvlet	vlqhmada.v	rgetvkis..	...s.dkts	..-mg...k..	.e.pispr.	ls...shlm.	d..ægnak*		
HlpA	mn	.qliisk.y	----ytkys	.td.eki.ts	ml.iivdtva	tgekvtilvg.	.s..a.e..a	..-æg...r..	.klflpasri	pt.sv.nffr	nk.nktf*		
TF-1	mn	.telika.q	----dte.t	qsvsvkmlas	.ækittetva	kgdkvqltg.	lnkpva.qa	..-kgf...q	.alelpsvg	vsvk..es..	kaae.l.yed	fak*	
LMW5-AR	.ttkkk.ti.	.qelyslv.a	--dtqlnkal	ier.ftsqkq	liqnalkhnq	eviippgk.	tvvt.kakpa	..-qgh..a..	.piqikakpe	hkavkira..	pvhdmn*		

(b) S20

Bb LRKNASALKR SRQNLKRRIR NVSVKSELKT IEKRCINMIK AGKKDEAIEF KKFVAKKLDLT AARKRIIHKN KAARKKSRLN VLLLK* 85
 Ec .anik..k.. ai.se.arkh .a.rr.mmr. fi.kvyaa.e .d.aa.qka .nemqpv.r q.a.gl....h.an.l taqinkla*

Fig. 3. Alignment of the predicted amino acid sequences of the *B. burgdorferi* Hbb and S20 proteins with the homologues from various other bacteria. Dots indicate identity; dashes indicate deletions. (a) Alignment of sequences of proteins of the HU/IHF family. Sources of protein sequences are as follows: HCj, *Campylobacter jejuni* (Konkel *et al.*, 1994); HTt, *Thermus thermophilus* (Zierer & Choli, 1990); HBSu, *B. subtilis* (Micka & Marahiel, 1992); HRcalHFB, *Rhodobacter capsulatus* (Toussaint *et al.*, 1993); HBca, *Bacillus caldolyticus* (Padas, 1992); EcIHFβ, *E. coli* (Flamm & Weisberg, 1985); EcHU-1, *E. coli* (Kano *et al.*, 1985); EcIHFα, *E. coli* (Miller, 1984); HAt, *Agrobacterium tumefaciens* (Khanaka *et al.*, 1985); EcHU-2, *E. coli* (Kano *et al.*, 1987); HRm, *Rhizobium meliloti* (Laine *et al.*, 1983); HCp, *Clostridium pasteurianum* (Tanaka *et al.*, 1984); HRL53, *Rhizobium leguminosarum* (Khanaka *et al.*, 1985); HTa, *Thermoplasma acidophilum* (DeLange *et al.*, 1981); HApr, *Aeromonas proteolytica* (Giladi *et al.*, 1992); HRcalHFB, *R. capsulatus* (Toussaint *et al.*, 1991); HlpA, *Cryptomonas* Φ plastid (Wang & Liu, 1991); TF-1, *B. subtilis* phage SP01 (Greene *et al.*, 1984); and LMW5-AR, African swine fever virus (Neilan *et al.*, 1993). (b) Alignment of sequences of ribosomal S20 proteins. Bb, *B. burgdorferi*; Ec, *E. coli*.

LMW5-AR, phage and viral proteins that are less similar to the rest, are excluded from the comparison, Gly₆₁ becomes an invariant amino acid that is also found in the Hbb protein. Many of the conserved residues are found in the region predicted to form the DNA-binding arm (Fig. 3). Analysis using the program PILEUP (from the University of Wisconsin Genetics Computer Group package; Devereux *et al.*, 1984) grouped the protein with the 'other' sequences category defined by Oberto & Rouvière-Yaniv (1996), rather than with the IHF or HU protein subunits (data not shown).

The Hbb protein has several differences from the related proteins. (a) There is an N-terminal extension on the sequence, which is also true for five of the related proteins (Fig. 3). The protein may actually initiate at the valine shown at position 9, since that corresponds to the start sites of most of the related proteins. In addition, at least one other *B. burgdorferi* protein, the DnaJ homologue, is predicted to be translated starting with a codon for valine (Tilly *et al.*, 1993), as is the *rpsT* gene (see above). (b) There are five additional amino acids inserted at position 21. The *Neisseria gonorrhoeae* Hip protein also has five amino acids inserted in this region (S. Hill, personal communication). (c) There is a single leucine inserted at position 82. (d) There are several amino acids that are conserved in most of the other HU-like proteins and different in the Hbb protein (e.g. Gly₅₂ and Gly₅₉).

Third, OrfH is not significantly similar to any sequences deposited in the sequence databases. The position of its start codon, adjacent to the stop codon for the *hbb* gene, suggests that the *hbb* and *orfH* genes may be coordinately expressed.

Fourth, the ORF homologous to the *E. coli* rho protein extends 96 amino acids upstream of the position at which the two homologues can be aligned (Tilly & Campbell, 1993). Analysis using the GCWIND program (Shields *et al.*, 1992), which determines the frequencies of A + T versus G + C in the third position of codons, suggests that the protein is initiated at the methionine indicated in Fig. 2 (data not shown).

Complementation

We measured the ability of the *B. burgdorferi* sequences to complement the defects for DNA packaging by two methods (Table 2). The e.o.p. measures the ability of phage carrying complementing genes to form plaques on strains in which those genes are required. Phage yield measurements determine the number of phage released per infected bacterium, a more quantitative parameter of growth. In order to determine which sequences were responsible for the ability of the λ transducing phages to form plaques on the various mutant hosts, we subcloned the *hbb* and *orfH* genes separately in pBluescript SK(+).

Table 2. Phage growth in various strains

Strain	Plasmid*	Phage†	e.o.p.‡	Phage yield§
N99 <i>hip</i> ⁺ <i>himA</i> ⁺	—	λ ZAPII		44
	—	λ bbb <i>orfH</i> ⁺		60
	—	λ bbb ⁺ <i>orfH</i>		44
	pBluescript SK(+)	λ ZAPII	1	76
	pKK55	λ ZAPII	1	124
A5179 <i>hip</i> <i>himA</i> ⁺	—	λ ZAPII		1
	—	λ bbb <i>orfH</i> ⁺		1
	—	λ bbb ⁺ <i>orfH</i>		7
	pBluescript SK(+)	λ ZAPII	$< 2 \times 10^{-7}$	2
	pKK55	λ ZAPII	0.1	26
A5427 <i>hip</i> ⁺ <i>himA</i>	—	λ ZAPII		1
	—	λ bbb <i>orfH</i> ⁺		2
	—	λ bbb ⁺ <i>orfH</i>		9
	pBluescript SK(+)	λ ZAPII	$< 2 \times 10^{-7}$	2
	pKK55	λ ZAPII	ND	26
A5477 <i>hip</i> <i>himA</i>	—	λ ZAPII		1
	—	λ bbb <i>orfH</i> ⁺		2
	—	λ bbb ⁺ <i>orfH</i>		12
	pBluescript SK(+)	λ ZAPII	$< 2 \times 10^{-7}$	4
	pKK55	λ ZAPII	0.2	32

ND, Not determinable because plaques were too small and indistinct to count.

* pKK55 is *rpsT*⁺ *bbb*⁺ *orfH*⁺ and has the *B. burgdorferi* promoters upstream of the three genes.

† The phages were as follows: λ bbb *orfH*⁺, λ ZAP50; λ bbb⁺ *orfH*, λ ZAP52.

‡ e.o.p. is defined as the ratio of the phage titre on the mutant strain to the titre on the wild-type strain. Plaques on mutant strains were small and indistinct.

§ Mean of two replicates.

Since we suspected that neither gene had a promoter, we crossed the subclones back into λ ZAPII *in vivo* by growing a stock of insert-free (blue-plaque-forming) phage in the plasmid-carrying strains (see Methods). Measuring phage yield per infected mutant bacterium for these phages (Table 2) indicated that the *bbb* gene, encoding the HU homologue, was the only complete gene required for complementation.

The mutant bacterial strains carrying plasmid pKK55 (Fig. 1) complemented λ ZAPII growth and yield about 10-fold (Table 2). These phage yields may have been higher because the bacteria contained the complementing protein at the time of infection, because the level of protein present was higher in plasmid-containing strains, or for both of these reasons. Plasmids derived from λ ZAP56 and λ ZAP41 (pKK56 and pKK41) did not complement in this assay, probably because they do not contain the *B. burgdorferi* promoters required for *bbb* expression (see below).

e.o.p. measurements for λ ZAPII with no insert on strains with the vector or pKK55 (carrying the promoters and the *rpsT*, *bbb* and *orfH* genes) were consistent with the phage yield experiments. Although the plaques were small and indistinct on the complemented strains (a

consequence of their reduced phage yield), the phage made plaques about 10⁶-fold more efficiently in the presence of the *bbb*⁺ plasmid. Qualitatively similar results were obtained in the *himA* mutant strain (A5427) but the plaques were too small and faint to be counted with accuracy.

Bacteriophage Mu requires HU protein in order to form plaques. We assayed the ability of pKK55 to complement *E. coli* *hupA* *hupB* mutants for Mu growth and found that the phage remained unable to form plaques in the plasmid-containing strain. The *B. burgdorferi* protein may fail to allow Mu growth because it lacks an essential function or it may be expressed at improper levels for complementation.

Expression

We used primer extension analysis to look for possible transcription start sites in the region containing the *bbb* gene. We used primers within the *rpsT*, *bbb*, *orfH* and *rbo* genes. The only clear signals that we found were derived from sequences upstream of the *rpsT* gene (Figs 2 and 4). Although there have been no systematic studies to determine the consensus sequences for *B. burgdorferi*

Fig. 4. Primer extension analysis upstream of the *rpsT* gene. Sequencing reactions are shown adjacent to the primer extension (PE). The sequence of the region surrounding the signals is shown. Arrows indicate the nucleotides at the 5' ends of the mRNAs.

promoters, there are often sequences resembling *E. coli* σ^{70} promoters in the appropriate locations upstream of *B. burgdorferi* genes [e.g. *ospA* (Jonsson *et al.*, 1992; Marconi *et al.*, 1993)]. There are overlapping copies of such consensus sequences upstream of the *rpsT* homologue (Fig. 2), at sites spaced appropriately for transcription to start where we observed RNA 5' ends (Fig. 4). Another good match to the consensus is located downstream of the *rbo* gene (data not shown).

Northern blot analysis (Fig. 5) indicated that RNA species of about 0.9 and 1.3 kb hybridize with an *hbb* probe, suggesting that there may be RNA processing or differential termination at the 3' end of the transcript. The larger transcript is long enough to include the *rpsT*, *hbb* and *orfH* genes, whereas the shorter transcript is predicted to lack the *orfH* gene. Consistent with this interpretation,

Fig. 5. Northern blot analysis of *hbb* transcription. The probe was a PCR product internal to the *hbb* gene.

an *orfH* probe did not hybridize with the smaller RNA species (data not shown).

There are potential ribosome-binding sites upstream of the predicted start codons of all of the ORFs (Schwartz *et al.*, 1992; Fig. 2). There is a potential rho-dependent terminator, defined as a C over G bias in the transcribed strand (Alifano *et al.*, 1991), within the *rbo* gene (data not shown). We found no predicted rho-independent terminators in this region.

DISCUSSION

The predominantly linear genomes and long, thin ultrastructures of the borreliae distinguish these bacteria from most, if not all, other genera (reviewed by Barbour & Hayes, 1986). The spirochaete *Borrelia hermsii* has multiple genomes spaced throughout its length (Kitten & Barbour, 1992) and *B. burgdorferi* is likely to have the same genomic arrangement. These features of the borreliae may require unusual or unique proteins for genome compaction, replication and partitioning to daughter cells. A small DNA-binding protein, like the Hbb protein, may play a role in these processes. The sequence of the Hbb protein may reflect binding specificity for sequences or structures peculiar to *B. burgdorferi*.

There are sequence differences between the predicted Hbb protein and the related proteins from other bacteria at several locations (see Results). These differences, how-

ever, are in regions predicted to lie between α -helical or pleated sheet domains so they are unlikely to significantly alter the protein's overall structure. Some of these features may be required for recognition of linear DNA molecules *in vivo*. The sole animal virus in Fig. 3, African swine fever virus, which encodes the LMW5-AR protein (Fig. 3; Neilan *et al.*, 1993), has a linear genome with telomeres homologous to those of the *B. burgdorferi* 16 kb linear plasmid (Hinnebusch & Barbour, 1991).

The *bbb* locus maps at about 250 kb from the left end of the *B. burgdorferi* chromosome (Casjens & Huang, 1993). The gene organization of the *bbb* region differs from those found to date in other bacteria. The *bbb* gene appears to be in an operon, with the genes encoding ribosomal protein S20 (*rpsT*) and an unknown ORF (*orfH*) flanking the *bbb* gene. Two RNA 5' ends map just upstream of the *rpsT* homologue and we have found no others throughout this region. Since two RNAs hybridize to an *bbb* probe on Northern blots, there is probably alternative transcription termination (or RNA processing) at the 3' ends of the transcripts, so that the portion of the RNA encoding the *orfH* product is absent in more than half of the RNA population. Consistent with this explanation, an *orfH* probe did not hybridize with the smaller RNA species.

There are no predicted rho-independent terminators yet there is a potential rho-dependent terminator within the *rho* gene. Although we have not identified any potential promoters for the *rho* gene, there may be a regulatory mechanism that allows transcriptional read-through when there are low rho concentrations. The *E. coli rho* gene appears to be autoregulated at the level of transcription attenuation, perhaps by a similar mechanism (Matsumoto *et al.*, 1986).

There is no obvious reason why the *bbb* gene should be in an operon with the *rpsT* and *orfH* genes. One of the promoters for the *E. coli hbp* gene is upstream of the adjacent *rpsA* gene, encoding a ribosomal protein (Flamm & Weisberg, 1985), but the other *E. coli* genes and the *B. subtilis* gene for small DNA-binding proteins do not appear to be cotranscribed with other genes.

We suspect that we have identified the only gene for an HU-like protein in *B. burgdorferi*. First, an *bbb* gene probe did not hybridize to any other sequences in *B. burgdorferi* DNA, at low or high stringency (data not shown). Second, in our screen for genes that complement the *E. coli* mutants defective in IHF and HU, we isolated no other gene for an HU-like protein. We isolated one phage capable of growing on the *hupA hupB himA* mutant strain that had an insert composed of the termini of two convergent ORFs (Kit Tilly, unpublished data). The predicted products were homologous to yeast uridine kinase and *E. coli* PriA, a hairpin-binding protein involved in DNA replication initiation. We do not know which sequences in this clone are responsible for complementation and have no explanation for how that complementation works. Neither product resembled HU or IHF. The product of the *bbb* gene is capable of fulfilling all roles of IHF protein required for λ growth, a characteristic of HU-like proteins (Giladi *et al.*, 1992). Third, the predicted

product of the *bbb* gene more closely resembles those derived from bacteria with single genes for small DNA-binding proteins. Unfortunately, phylogenetic analysis did not group the protein with other IHF or HU protein subunits. *B. burgdorferi* has a small genome of around 1000 kbp, so there may be a selection towards limiting the genetic content and eliminating genes for proteins with overlapping functions.

The next step in studying the Hbb protein will be determining its function in *B. burgdorferi* growth and linear plasmid replication. We have made exhaustive attempts to purify the protein from both *B. burgdorferi* cells and *E. coli* cells containing a plasmid carrying the *bbb* gene (pKK55). Although we have purified several small heparin-binding proteins and a number of proteins that bind to DNA fragments containing putative IHF-binding sites, none of these proteins appears to be the *bbb* gene product. Ultimately, we hope to identify ways in which the maintenance of linear DNA molecules in bacteria differs from that of the more common circular molecules.

ACKNOWLEDGEMENTS

We thank J. Swanson, P. Rosa, S. Hill and J. Hinnebusch for critical reading of the manuscript, S. Hill for communication of results prior to publication, S. Morrison for oligonucleotide synthesis, G. Hettrick and R. Evans for help with figure preparation, S. Smaus and C. Smaus for secretarial assistance, and Amos Oppenheim for advice and encouragement. Eldon Walker performed the PILEUP sequence analysis. D. Hogan provided valuable advice on DNA sequencing and primer extension analysis. J.C. and J.F. participated in this work as NIH summer students.

REFERENCES

- Alifano, P., Rivellini, F., Limauro, D., Bruni, C. B. & Carlomagno, M. S. (1991). A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell* **64**, 553–563.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1989). *Current Protocols in Molecular Biology*. New York: Greene Publishing Associates and Wiley-Interscience.
- Barbour, A. G. & Garon, C. F. (1987). Linear plasmids of the bacterium *Borrelia burgdorferi* have covalently closed ends. *Science* **237**, 409–411.
- Barbour, A. G. & Hayes, S. F. (1986). Biology of *Borrelia* species. *Microbiol Rev* **50**, 381–400.
- Baril, C., Richaud, C., Baranton, G. & Saint Girons, I. (1989). Linear chromosome of *Borrelia burgdorferi*. *Res Microbiol* **140**, 507–516.
- Bonnefoy, E., Takahashi, M. & Yaniv, J. R. (1994). DNA-binding parameters of the HU protein of *Escherichia coli* to cruciform DNA. *J Mol Biol* **242**, 116–129.
- Casjens, S. & Huang, W. M. (1993). Linear chromosomal physical and genetic map of *Borrelia burgdorferi*, the Lyme disease agent. *Mol Microbiol* **8**, 967–980.
- DeLange, R. J., Williams, L. C. & Searcy, D. G. (1981). A histone-like protein (HTa) from *Thermoplasma acidophilum*. II. Complete amino acid sequence. *J Biol Chem* **256**, 905–911.

- Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12, 387–395.
- Drlica, K. & Rouviere-Yaniv, J. (1987). Histone-like proteins of bacteria. *Microbiol Rev* 51, 301–319.
- Ferdows, M. S. & Barbour, A. G. (1989). Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. *Proc Natl Acad Sci USA* 86, 5969–5973.
- Flamm, E. L. & Weisberg, R. A. (1985). Primary structure of the *hip* gene of *Escherichia coli* and of its product, the beta subunit of integration host factor. *J Mol Biol* 183, 117–128.
- Friedman, D. I. (1988). Integration host factor: a protein for all reasons. *Cell* 55, 545–554.
- Georgopoulos, C. P. & Herskowitz, I. (1971). *Escherichia coli* mutants blocked in lambda DNA synthesis. In *The Bacteriophage Lambda*, pp. 553–564. Edited by A. D. Hershey. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Giladi, H., Wang, W.-X. & Oppenheim, A. B. (1992). Isolation and characterization of the *bupA* gene coding for HU of *Aeromonas proteolytica*. *Nucleic Acids Res* 20, 4092.
- Greene, J. R., Brennan, S. M., Andrew, D. J., Thompson, C. C., Richards, S. H., Heinrikson, R. L. & Gelduschek, E. P. (1984). Sequence of the bacteriophage SP01 gene coding for transcription factor 1, a viral homologue of the bacterial type II DNA-binding proteins. *Proc Natl Acad Sci USA* 81, 7031–7035.
- Haluzi, H., Goitein, D., Koby, S., Mendelson, I., Teff, D., Mengeritsky, G., Giladi, H. & Oppenheim, A. B. (1991). Genes coding for integration host factor are conserved in gram-negative bacteria. *J Bacteriol* 173, 6297–6299.
- Hinnebusch, B. J., Jr (1991). *Structure and replication of Borrelia burgdorferi linear plasmids*. PhD thesis, University of Texas Graduate School of Biomedical Sciences at San Antonio, USA.
- Hinnebusch, J. & Barbour, A. G. (1991). Linear plasmids of *Borrelia burgdorferi* have a telomeric structure and sequence similar to those of a eukaryotic virus. *J Bacteriol* 173, 7233–7239.
- Hinnebusch, J., Bergström, S. & Barbour, A. G. (1990). Cloning and sequence analysis of linear plasmid telomeres of the bacterium *Borrelia burgdorferi*. *Mol Microbiol* 4, 811–820.
- Jonsson, M., Noppa, L., Barbour, A. G. & Bergström, S. (1992). Heterogeneity of outer membrane proteins in *Borrelia burgdorferi*: comparison of *osp* operons of three isolates of different geographic origins. *Infect Immun* 60, 1845–1853.
- Kano, Y. & Imamoto, F. (1990). Requirement of integration host factor (IHF) for growth of *Escherichia coli* deficient in HU protein. *Gene* 89, 133–137.
- Kano, Y., Yoshino, S., Wada, M., Yokoyama, K., Nobuhara, M. & Imamoto, F. (1985). Molecular cloning and nucleotide sequence of the *HU-1* gene of *Escherichia coli*. *Mol Gen Genet* 201, 360–362.
- Kano, Y., Osato, K., Wada, M. & Imamoto, F. (1987). Cloning and sequencing of the *HU-2* gene of *Escherichia coli*. *Mol Gen Genet* 209, 408–410.
- Khanaka, H., Laine, B., Sautiere, P. & Guillaume, J. (1985). Characterization and primary structures of DNA-binding HU-type proteins from *Rhizobiaceae*. *Eur J Biochem* 147, 343–349.
- Kikuchi, A., Flamm, E. & Weisberg, R. A. (1985). An *Escherichia coli* mutant unable to support site-specific recombination of bacteriophage lambda. *J Mol Biol* 183, 129–140.
- Kitten, T. & Barbour, A. G. (1992). The relapsing fever agent *Borrelia hermsii* has multiple copies of its chromosome and linear plasmids. *Genetics* 132, 311–324.
- Konkel, M. E., Marconi, R. T., Mead, D. J. & Cieplak, W., Jr (1994). Cloning and expression of the *bup* gene encoding a histone-like protein of *Campylobacter jejuni*. *Gene* 146, 83–86.
- Kosturko, L. D., Daub, E. & Murialdo, H. (1989). The interaction of *E. coli* integration host factor and lambda *cos* DNA: multiple complex formation and protein-induced bending. *Nucleic Acids Res* 17, 317–334.
- Laine, B., Belaiche, D., Khanaka, H. & Sautiere, P. (1983). Primary structure of the DNA-binding protein HRm from *Rhizobium meliloti*. *Eur J Biochem* 131, 325–331.
- Lee, E. C., Hales, L. M., Gumpert, R. I. & Gardner, J. F. (1992). The isolation and characterization of mutants of the integration host factor (IHF) of *Escherichia coli* with altered, expanded DNA-binding specificities. *EMBO J* 11, 305–313.
- Marconi, R. T., Konkel, M. E. & Garon, C. F. (1993). Variability of *osp* genes and gene products among species of Lyme disease spirochetes. *Infect Immun* 61, 2611–2617.
- Matsumoto, Y., Shigesada, K., Hirano, M. & Imai, M. (1986). Autogenous regulation of the gene for transcription termination factor rho in *Escherichia coli*: localization and function of its attenuators. *J Bacteriol* 166, 945–958.
- Mechulam, Y., Fayat, G. & Blanquet, S. (1985). Sequence of the *Escherichia coli* *pheST* operon and identification of the *bimA* gene. *J Bacteriol* 163, 787–791.
- Mendelson, I., Gottesman, M. & Oppenheim, A. B. (1991). HU and integration host factor function as auxiliary proteins in cleavage of phage lambda cohesive ends by terminase. *J Bacteriol* 173, 1670–1676.
- Mengeritsky, G., Goldenberg, D., Mendelson, I., Giladi, H. & Oppenheim, A. B. (1993). Genetic and biochemical analysis of the integration host factor of *Escherichia coli*. *J Mol Biol* 231, 646–657.
- Micka, B. & Marahiel, M. A. (1992). The DNA-binding protein HBSu is essential for normal growth and development in *Bacillus subtilis*. *Biochimie* 74, 641–650.
- Micka, B., Grouch, N., Heinemann, U. & Marahiel, M. A. (1991). Molecular cloning, nucleotide sequence, and characterization of the *Bacillus subtilis* gene encoding the DNA-binding protein HBSu. *J Bacteriol* 173, 3191–3198.
- Miller, H. I. (1984). Primary structure of the *bimA* gene of *Escherichia coli*: homology with DNA-binding protein HU and association with the phenylalanyl-tRNA synthetase operon. *Cold Spring Harbor Symp Quant Biol* 49, 691–698.
- Miller, H. I., Kikuchi, A., Nash, H. A., Weisberg, R. A. & Friedman, D. I. (1979). Site-specific recombination of bacteriophage lambda: the role of host gene products. *Cold Spring Harbor Symp Quant Biol* 43, 1121–1126.
- Neilan, J. G., Lu, Z., Kutish, G. F., Sussman, M. D., Roberts, P. C., Yozawa, T. & Rock, D. L. (1993). An African swine fever virus gene with similarity to bacterial DNA binding proteins, bacterial integration host factors, and the *Bacillus* phage SPO1 transcription factor, TF1. *Nucleic Acids Res* 21, 1496.
- Ner, S. S., Travers, A. A. & Churchill, M. E. A. (1994). Harnessing the writhe: a role for DNA chaperones in nucleoprotein-complex formation. *Trends Biochem Sci* 19, 185–187.
- Oberto, J. & Rouviere-Yaniv, J. (1996). *Serratia marcescens* contains a heterodimeric HU protein like *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol* 178, 293–297.
- Padas, P. M. (1992). The DNA-binding protein HU from mesophilic and thermophilic bacilli: gene cloning, overproduction and purification. *Gene* 117, 39–44.
- Pontiggia, A., Negri, A., Beltrame, M. & Bianchi, M. E. (1993). Protein HU binds specifically to kinked DNA. *Mol Microbiol* 7, 343–350.

- Ravin, V. K. & Shulga, M. G. (1970). Evidence for extrachromosomal location of prophage N15. *Virology* **40**, 800–807.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Schmid, M. B. (1990). More than just “histone-like” proteins. *Cell* **63**, 451–453.
- Schwartz, J. J., Gazumyan, A. & Schwartz, I. (1992). rRNA gene organization in the Lyme disease spirochete, *Borrelia burgdorferi*. *J Bacteriol* **174**, 3757–3765.
- Shields, D. C., Higgins, D. G. & Sharp, P. M. (1992). GCWIND: a microcomputer program for identifying open reading frames according to codon positional G + C (communication). *CABIOS* **8**, 521–523.
- Silhavy, T. J., Berman, M. L. & Enquist, L. W. (1984). *Experiments with Gene Fusions*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Stern, A., Brown, M., Nickel, P. & Meyer, T. F. (1986). Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell* **47**, 61–71.
- Svarchevsky, A. N. & Rybchin, V. N. (1984). Characteristics of plasmid properties of bacteriophage N15. *Mol Genet Mikrobiol Virusol* **5**, 34–39 (in Russian).
- Tanaka, I., Appelt, K., Dijk, J., White, S. W. & Wilson, K. S. (1984). 3-Å resolution structure of a protein with histone-like properties in prokaryotes. *Nature* **310**, 376–381.
- Thompson, J. F. & Landy, A. (1988). Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. *Nucleic Acids Res* **16**, 9687–9705.
- Tilly, K. & Campbell, J. (1993). A *Borrelia burgdorferi* homolog of the *Escherichia coli* rho gene. *Nucleic Acids Res* **21**, 1040.
- Tilly, K., Hauser, R., Campbell, J. & Ostheimer, J. (1993). Isolation of *dnaJ*, *dnaK*, and *grpE* homologues from *Borrelia burgdorferi* and complementation of *Escherichia coli* mutants. *Mol Microbiol* **7**, 359–369.
- Toussaint, B., Bosc, C., Richaud, P., Colbeau, A. & Vignais, P. M. (1991). A mutation in a *Rhodobacter capsulatus* gene encoding an integration host factor-like protein impairs *in vivo* hydrogenase expression. *Proc Natl Acad Sci USA* **88**, 10749–10753.
- Toussaint, B., Delic-Attree, I., De Sury D’Aspremont, R., David, L., Vincon, M. & Vignais, P. M. (1993). Purification of the integration host factor homolog of *Rhodobacter capsulatus*: cloning and sequencing of the *hip* gene, which encodes the beta subunit. *J Bacteriol* **175**, 6499–6504.
- Travers, A. A., Ner, S. S. & Churchill, M. E. A. (1994). DNA chaperones: a solution to a persistence problem? *Cell* **77**, 167–169.
- Wang, S. & Liu, X.-Q. (1991). The plastid genome of *Cryptomonas phi* encodes an hsp70-like protein, a histone-like protein, and an acyl carrier protein. *Proc Natl Acad Sci USA* **88**, 10783–10787.
- Yang, C.-C. & Nash, H. A. (1989). The interaction of *E. coli* IHF protein with its specific binding sites. *Cell* **57**, 869–880.
- Zierer, R. & Choli, D. (1990). The primary structure of DNA binding protein II from the extreme thermophilic bacterium *Thermus thermophilus*. *FEBS Lett* **273**, 59–62.

Received 12 February 1996; revised 16 April 1996; accepted 19 April 1996.