FOLIA PARASITOLOGICA 46: 81-90, 1999

Lectin-like sequences in genome of *Borrelia burgdorferi*

Nataliia Rudenko, **Maryna Golovchenko** and **Libor Grubhoffer**

Faculty of Biological Sciences, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic, and Institute of Parasitology, Academy of Sciences of the Czech Republic, Branišovská 31, 370 05 České Budějovice, Czech Republic

Key words: *Borrelia burgdorferi*, hemagglutination, genomic library, recombinant, lectin

Abstract. Using degenerative primers designed on the basis of known sequences of lectin genes from different sources a fragment of genomic DNA of *Borrelia burgdorferi* (strain B31) that contained a lectin-like sequence was isolated, cloned and sequenced. The presence of an open reading frame of 268 amino acids (position 1501-2304 bp) and the computer analysis of the predicted amino acid sequence showed 37% of identity and 75% of homology over region of 25 amino acids with the legume lectin proteins, including erythroagglutinating phytohemagglutinin (PHA-E) and leucoagglutinating phytohemagglutinin (PHA-L). The further analysis of the predicted amino acid sequence showed the presence of another two domains (positions 198-211 and 215-226 aa) consisting of the characteristic conserved sequence motifs for legume lectin proteins. Hemagglutinating activity was detected in lysate of *B. burgdorferi* (strain B31) spirochete and the affinity to fetuin was determined in a hemagglutination inhibition test. Hemagglutinating activity was also found in a crude lysate of the recombinant clones carrying the fragment of *B. burgdorferi* genomic DNA. The inhibition of agglutinating activity by fetuin, D-galactosamine and D-mannosamine was determined using the standard procedure of hemagglutination inhibition test with native rabbit red blood cells (RBC).

Lectins are ubiquitous proteins that are probably present in all eukaryotic and many bacterial species as well as in some viruses. They were defined as carbohydrate-binding proteins other than antibodies and enzymes (Kocourek and Hořejší 1981, Barondes 1988). Lectins have been conserved throughout evolution and thus, there is strong evidence that they must have important functions in nature (see Sharon and Lis 1989 for review). It is generally believed that lectins serve primarily as recognition molecules and may also play a role in a transmission of pathogens by invertebrate vectors (Maudlin and Welburn 1988).

Hemagglutinins represent important functional activities of various bacteria. The best characterised are *E. coli* lectins: type 1 fimbriae (mannose-specific), type P fimbriae (galactose-specific) and lectins of oral Actinomycetes*:* S fimbriae (sialic acid-specific) and type 2 fimbriae (galactose-specific) (see Sharon and Lis 1989 for review).

Bacterial surface lectins play a key role in the initiation of infection by mediating bacterial adherence to epithelial cells of the host (e.g. type 1 and type P lectins of *E. coli*). Lectin-containing bacteria may also bind to carbohydrate epitopes on phagocytic cells, such as polymorphonuclear leucocytes or peritoneal macrophages. The lectin-mediated non-opsonic phagocytosis, which has been designated as lectino-phagocytosis, may be of relevance in both vector and host-tissue interactions (Ofek and Sharon 1988).

Spirochetes of *Borrelia* genus are causative agents of serious infective diseases of man and animals. Most of *Borrelia* species are transmitted by argasid and ixodid ticks. They cause Lyme disease (*B. burgdorferi)* and relapsing fevers *(B. duttoni, B. hermsi, B. parkeri,* etc) (for review, see Barbour and Hayes 1986).

It has been already known that *B. burgdorferi* binds to many types of mammalian cells and to extracellular matrix (Szczepanski et al. 1990). Garcia-Monco et al. (1992) reported that host cell galactosylcerebroside is recognised by *B. burgdorferi*, and Isaacs (1994) showed that *Borrelia* spirochete binds host cell proteoglycans. Hemagglutinating properties are not mediated by a lectin, which is strongly inhibited with dextransulfate but not with chondroitinsulfate (Leong et al. 1995). Previously, we have identified (using polyclonal antibodies) and partially characterised the lectin of *B. recurrentis* (Grubhoffer et al. 1993) and demonstrated lectin(s) activity of *B. burgdorferi* spirochete, strain B31, confirmed by Leong et al. (1995). In this paper, we demonstrate for the first time the presence of gene coding for the protein with hemagglutinating activity (designated as B31LEC) similar to the already known lectins (see GenBank Database) in the genome of *B. burgdorferi*. We have isolated, cloned into the plasmid vector and sequenced this gene, determined the domains conserved for the lectin superfamily proteins present in it, showed the strong agglutination activity of that protein against native rabbit erythrocytes and determined its sugar specificity.

MATERIALS AND METHODS

Bacterial strain. *Borrelia burgdorferi*, strain B31 (ATCC 35210) was obtained from Dr. J. F. Anderson (Connecticut Agricultural Experiment Station, New Haven, CN) and is

Address for correspondence: L.Grubhoffer, Institute of Parasitology, Academy of Sciences of the Czech Republic, Branišovská 31, 370 05 České Budějovice, Czech Republic. Phone: ++420 38 777 5403; Fax: ++420 38 47743; E-mail: liborex@paru.cas.cz

routinely passaged in our laboratory. *B. burgdorferi* was grown in BSK-H Complete Medium (Sigma, cat. # B-8291) at 33°C for 5-7 days or until the concentration was approximately 10^9 cells per ml. The cells were collected by centrifugation at 15,000g for 15 min at 4°C. The pellet was washed three times with ice-cold phosphate-saline buffer (pH 7.2), harvested as before and stored at -80°C until use.

Hemagglutination assay. The collected spirochetes were resuspended in TES buffer (10 mM Tris-HCl, 1 mM EDTA, 1% N-lauroylsarcosine, pH 7.2), followed by 3 freeze-thaw steps. The spirochetes were incubated 1 hour at 20°C and then sonicated 3 times for 30 s on ice. The cell debris was collected by centrifugation at 12,000g for 10 min. The supernatant was dialysed against 1 mM Tris-HCl, pH 7.2 using Microdialysis System (BRL, cat. #. 1200MD, USA) for 2 hr, concentrated and then used for hemagglutination test with 2% suspension of rabbit erythrocytes. The test was performed at room temperature. Spirochete lysate (50 µl) containing 50 µg of a total protein was added to 50 µl of agglutination buffer, mixed well and the serial dilution of lysate was made by transferring 1/2 of volume to the next well. Fifty microlitres of rabbit red blood cells were added to each well (Grubhoffer et al. 1991).

Hemagglutination inhibition test. Different kinds of inhibitors of agglutination were serially diluted in TN Ca^{2+} buffer (50 mM Tris-HCl, 0.15 M NaCl, 20 mM $CaCl_2$, pH 7.0). The inhibition of agglutination was performed on microtitration plates with U- bottomed wells (Grubhoffer et al. 1993). The test was performed at room temperature.

Primer design. Protein sequences of lectins with different sugar specificity from more than 100 organisms were compared, including mammals, other vertebrates, plants, insects and bacteria. Different available databases (GenBank, PIR, EMBL, PDB, etc) were searched for the information on protein and nucleic acid sequences of lectin genes of different organisms. Similar conserved protein regions were identified. The nucleic acid sequences corresponding to the above-mentioned protein were compared and used as a pattern for designing the primers for PCR on genomic DNA of *B. burgdorferi,* strain B31. The five oligonucleotide primers used were synthesised by Lambda-Bio-Med (Prague, Czech Republic), three corresponding to 5' end of the gene (primer A 5'-CACTTCAACCCTCGC-3', primer D 5'-GCCAACACC-ATTGTG-3' and primer E 5'-CCCTTCCAGCCTGGG-3'); and two corresponding to 3' end of the gene (primer B 5'-GGCCACGCACTTAATC-3' and primer C 5'-CCATCC-GCCGCCATG-3').

Purification of *B. burgdorferi* genomic DNA. *B. burgdorferi* genomic DNA was purified using "Wizard Genomic DNA Purification Kit" (Promega, cat. # A1120) according to manufacturer's recommendation, which involves no organic extraction or proteinase digestion.

Polymerase chain reaction on genomic DNA. The PCR reaction was conducted in 0.5 ml Eppendorf tube in a final volume of 100 µl and contained: 50 mM KCl, 10 mM Tris-HCl (pH 9.0 at 25°C), 0.1% Triton X-100, 2.5 mM $MgCl_2$, 0.2 mM for each of the dNTPs, 100 pM of each of designated primers, 3U of Taq DNA polymerase and 100 ng of genomic DNA as a template. All reagents were purchased from Promega Company.

The PCR parameters were: 1 step - 95°C - 5 min, 30 cycles of [94°C - 1 min, 55°C - 1 min, 74°C - 2 min], 1 step - 74°C - 8 min, and the final step - 65°C - 15 min. The PCR was conducted using Progene thermocycler (Techne).

Analysis and purification of PCR product. The PCR products were analysed by agarose gel electrophoresis (0.8% agarose/1×TAE [0.04 M Tris-acetate, 0.001 M EDTA] buffer) and the fragment of interest was purified from low-melted point (LMP) agarose (Sigma). A slice of LMP-agarose with a DNA fragment was weighed. The equal volume of TE (10 mM Tris-HCl, 1 mM EDTA) buffer was added and the tube was incubated for 10 min at 70°C. Thereafter the tube was placed at -80°C for 10 min and thawed at 37°C for 10 min. The procedure was repeated twice, followed by centrifugation at 12,000g for 5 min. The supernatant was transferred to a fresh tube, 1/10 volume of 3 M sodium acetate (pH 5.2) and 2.5 volume of ethanol were added and incubated at -80°C for 10 min. The DNA was collected by centrifugation at 12,000g for 10 min. The pellet was resuspended in a small amount of TE buffer (50-100 µl), incubated at 70°C for 10 min, frozen at -80°C for 10 min and thawed at 37°C. After centrifugation at the same conditions (see above) the pure PCR product appeared in supernatant while the agarose contamination was in a pellet. The PCR product, purified in such a way, is suitable for use in the most demanding molecular biology applications, including restriction analysis, automated or manual sequencing, cloning, hybridisation and PCR.

Genomic southern blot. Thirty micrograms of the genomic DNA was digested separately with various restriction endonucleases (5U/µg DNA) namely: BamHI, BglII, EcoRI, Hind III, Kpn I, Pst I, Pvu II, Xba I and Xho I (Promega) for 3 hr at 37°C in a total volume of 20 µl. The genomic fragments were separated by overnight gel-electrophoresis in 0.7% agarose/1×TAE with no ethidium bromide in the gel. After electrophoresis was completed the gel was stained with ethidium bromide (0.5 µg/ml) and photographed. The gel was washed in depurination (0.25 N HCl) solution for 10 min, denaturation solution (1.5 M NaCl, 0.5 N NaOH) for 30 min and neutralisation solution (1 M Tris-HCl pH 7.2, 2 M NaCl) for 30 min. The capillary blot was set up and the genomic DNA fragments were transferred to the nylon membrane Hybond-A (Amersham, cat. # RPN 303R), using 20×SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0) as a buffer (Southern 1975). The DNA was crosslinked to the membrane by UV light using UV Stratalinker 1800 (Stratagene).

Probe labelling and hybridisation. Highly labelled DNA (up to 10^9 dpm/µg) was obtained using random primering "DNA Labeling Kit" (MBI Fermentas, cat. # K0612) and radioactive alpha ^{32}P-dATP (Amersham). The product (96 bp long) of PCR reaction with designated E and C primers and genomic DNA as template was used as a probe. Hybridisation was carried out overnight at 65°C in NaPi buffer (0.25 M NaH_2PO_4, 0.25 M Na_2HPO_4) containing 7% SDS, 1 mM EDTA and 100µg/ml fragmented salmon sperm DNA. Following hybridisation, the filter was washed 3 times for 15 min in pre-warmed 3×SSC/0.1% SDS and twice for 15 min in 0.3×SSC/0.1% SDS at 65°C. The filter was exposed overnight in PhosphorImmager cassette (Molecular Dynamics Storage Phosphor Screen).

Table 1. Condensed list of the sources of the lectin sequences compared.

No.	Scientific name	Common name	Accession No.	Specificity
1	*Pseudomonas aeruginosa*	bacterium	93054624	galactose-binding
2	*Geordia cydonium*	marine sponge	93257788	galactose-binding
3	*Cucumaria echinata*	sea cucumber	PIDg1083928	galactose-binding
4	*Stichopus japonicus*	sea cucumber	95204387	calcium-binding
5	*Robinia preudoacacia*	locust tree	96191285	complex
6	*Tulipa* sp.	tulip	96195647	complex
7	*Allium ursinum*	ramson	94039027	mannose-binding
8	*Pisum sativum*	pea	92256807	?
9	*Phaseolus vulgaris*	kidney bean	94002183	?
10	*Allium sativum*	garlic	93008079	mannose-binding
11	*Clivia miniata*	South African tree	94250846	mannose-binding
12	*Caenorhabditis elegans*	nematode	94150718	galactose-binding
13	*Caenorhabditis elegans*	nematode	92348399	beta-galactoside
14	*Bombyx mori*	silkworm	D14168	?
15	*Periplaneta americana*	American cockroach	91302364	lipopolysaccharide
16	*Sus scrofa*	pig	95081129	lactose-binding
17	*Gallus gallus*	chicken	91072353	beta-galactoside
18	*Gallus gallus*	chicken liver	81215504	complex
19	*Cricetulus griseus*	Chinese hamster	95197510	beta-galactoside
20	*Cricetulus longicaudatus*	longtailed hamster	94299546	galactose-binding
21	*Mus spretus*	western wild mouse	94319082	beta-galactoside
22	*Mus musculus*	house mouse	88080093	beta-galactoside
23	*Rattus norvegicus*	laboratory rat	94253130	calcium-binding
24	*Oryctolagus cuniculus*	rabbit	96011642	complex
25	*Canis familiaris*	dog	94075368	beta-galactoside
26	*Bos taurus*	bull	94118330	beta-galactoside
27	*Homo sapiens*	human	PIDg361225	mannose-binding
28	*Homo sapiens*	human	89123203	beta-galactoside
29	*Homo sapiens*	human	94102823	?

Construction of "mini-libraries" from genomic DNA of *B. burgdorferi.* The fragments obtained by the restriction of the genomic DNA with EcoR I and Xba I enzymes that showed the positive signals after hybridisation were selected for the further cloning into plasmid vector pBluescript SK- (Stratagene). After sufficient separation by the agarose gel electrophoresis the DNA fragments of interest were carefully cut out from the gel and the mixture of genomic DNA fragments were isolated from LMP-agarose as described previously. The "mini-libraries" of those parts of *B. burgdorferi* genome were constructed and designed as "EcoR I"- and "XbaI"-mini-libraries for further use.

All ligation and transformation procedures were conducted according to Sambrook et al. (1989).

XL-I Blue cells were selected for transformation. X-gal and IPTG were used for white/blue selection of the recombinants.

Screening of genomic "mini-libraries". Recombinant clones were randomly picked from each mini-library, plated in ten 150-mm plastic Petri dishes (in duplicate) on LB agar/ampicillin. One of each duplicate Petri dishes was covered by nylon membrane. The plates were incubated at 37°C until small colonies appeared and then pre-chilled for at least 30 min at 4°C. The filters were lifted up from the plates and placed for 3 min on 3MM paper, saturated with 10% SDS. Then the filters were denaturated for 5 min and transferred to 3MM paper saturated with neutralisation solution for 3 min. Neutralisation step was repeated once more. Finally the membranes were vigorously washed in 2×SSC to remove the protein debris. The filters were air-dried, DNAs were fixed to the membrane by UV crosslinking procedure and the filters were used for hybridisation with ^{32}P-labelled probe that was carried out under conditions described in "Probe labelling and Hybridisation". The colonies that gave the positive signals in the hybridisation were picked from the second duplicate plate, grown at 37°C overnight in 5 ml LB broth/ampicillin, collected by centrifugation at 5,000g for 10 min. The plasmid DNAs were purified using "Wizard Miniprep DNA Purification Kit" (Promega), and polymerase chain reaction with T3/T7 primers was performed. The PCR products were analyzed by agarose-gel electrophoresis in 0.8% agarose/ 1×TAE buffer, transferred to nylon membrane and the second round of hybridisation with the same probe was conducted.

Sequencing of genomic insert. Nucleotide sequences of the desired clones were determined by the dideoxynucleotide chain termination method (Sanger et al. 1977) with the use of double-stranded pBluescript and Sequenase Version 2.0 DNA Sequencing Kit (USB, cat. # US70770, USA). T3, T7 as well as the specific internal primers were used for sequencing. Analyses of DNA sequences were performed using the GenBank, PIR, EMBL, PDB, and SCOP computer analysis programs.

Construction of expression clones. Clones X3 and X230 were digested by XbaI and clone E200 by EcoRI and XbaI enzymes, the inserts were purified from LMP agarose and

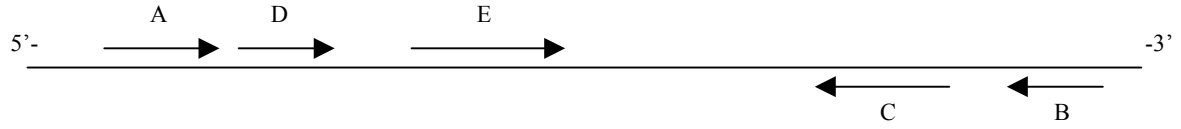

Fig. 1. Schematic location of the designed PCR primers on the "average" lectin gene. The comparison of sequences of different lectin genes revealed the presence of conserved domains, located at about the same position of genes irrespective of the source of the protein, and was used as a pattern for PCR primers design.

```
   1                                                                  30
   C C C  T T C  C A G  C C T  G G G  T C T  C A T  G G C  A A A  A C C
1    P      F      Q      P      G      S      H      G      K      T
  31                                                                 60
  T T A  T C T  C T A  C A A  A A A  A T A  A A A  G A A  A A T  A A A
     L      S      L      Q      K      I      K      E      N      K
  61                                                                 90
  A A T  A A T  A A A  A A A  T T A  G C T  A G G  C A T  G G C  G G C
     N      N      K      K      L      A      R      H      G      G
            96
  G G A  T G G
     G      W     32
```

Fig. 2. Nucleotide and deduced amino acid sequence of E/C PCR product from *Borrelia burgdorferi*. Primers E and C are shaded and underlined.

	S	H	G	K	T	L	S	L	Q	K	I	K	E	N	K	N	N	K	K	L	A	E/C PCR product	
47	S	V	G	R	A	L	Y	T	A	P	I	R	L	W	Q	S	S	S	L	V	A	LEC BOWMI	– 1
73	S	T	G	R	T	L	Y	A	K	P	V	H	I	W	D	M	T	T	G	T	V	LEC ERYCO	– 2
48	S	V	G	R	V	L	Y	Q	T	P	I	H	L	W	D	K	Q	I	D	K	E	LEC ONOVI	– 3
62	S	M	G	R	A	F	Y	S	A	P	I	Q	I	N	D	R	T	I	D	N	L	ARC1 PHAVU	– 4
41	T	V	G	R	A	L	Y	S	L	P	I	H	I	W	D	S	E	T	G	N	V	LEC VICFA	– 5
79	S	L	G	R	A	L	Y	S	T	P	I	H	I	W	D	K	E	T	G	S	V	LEC SOYBN	– 6
50	S	A	G	Q	A	S	Y	S	E	P	V	F	L	W	D	S	T	G	K	A	A	LEC4 GRISI	– 7
65	S	L	G	R	A	F	Y	S	A	P	I	Q	I	W	D	N	T	T	G	T	V	PHAL PHAVU	– 8
66	S	L	G	R	A	F	Y	S	A	P	I	Q	I	W	D	N	T	T	G	A	V	PHAE PHAVU	– 9
45	T	G	G	I	A	R	Y	I	A	P	I	H	I	W	N	C	N	T	G	E	V	LEC1 ULEEU	– 10
38	S	A	G	R	V	L	Y	S	A	P	L	R	L	W	E	D	S	A	V	L	T	LECA DOLLA	– 11

Fig. 3. Alignment of the predicted amino acid sequence of E/C PCR product with sequences of the legume lectin beta-chain proteins (BLOCK BLOO307A).

Comparison of the translated amino acid sequence to: 1 – lectin (agglutinin) (BMA) from *Bowringia mildibraedii* (LEC BOWMI [Chawla et al. 1993]); 2 – lectin precursor from *Erythrina corallodendron* (coral tree) (LEC ERYCO [Adar et al. 1989]); 3 – lectin from *Onobrychis viciifolia* (common sainfoin) (LEC ONOVI [Kouchalakos et al. 1984]); 4 – arcelin-1 precursor (legume lectin) from *Phaseolus vulgaris* (ARC1 PHAVU [Anthony et al. 1991]); 5 – favin (lectin) from *Vicia faba* (broad bean) (LEC VICFA [Hopp et al. 1982]); 6 – lectin precursor (agglutinin) (SBA) from soybean (LEC SOYBN [Vodkin et al. 1983]); 7 – lectin IV (GS4) *Griffonia simplicifolia* (LEC4 GRISI [Delbaere et al. 1993]); 8 – leucoagglutinating phytohemagglutinin precursor (PHA-L) from *Phaseolus vulgaris* (kidney bean) (PHAL PHAVU [Hoffman and Donaldson 1985]); 9 – erythroagglutinating phytohemagglutinin precursor (PHA-E) DLEC1 *Phaseolus vulgaris* (kidney bean) (French bean) (PHAE PHAVU [Hoffman and Donaldson 1985]); 10 – anti-H (o) lectin I (UEA-I) from *Ulex europaeus* (furze) (LEC1 ULEEU [Konami et al. 1991b]); 11 – lectin from *Dolichos lab lab L.* (field bean) (LECA DOLLA [Gowda et al. 1994]). Numbering of amino acids is as per the published sequences. Identical amino acids are boxed and shaded.

ligated to the previously prepared pGEX-KG vector (Pharmacia). After X-gal/IPTG selection on LB agar plates recombinant clones were grown in LB broth overnight. Overnight cultures of *E. coli* transformed with recombinant pGEX plasmids were diluted 1:20 in 5 ml of fresh medium containing 50 µg/ml ampicillin and grown to midlog phase at 37°C. The IPTG was added to a final concentration of 1mM. Following three hours of growing at 37°C, the cells were pelleted by centrifugation (Guan and Dixon 1991, Smith and Johnson 1988). The cells were resuspended in dd H_2O, boiled for 10 min, chilled on ice, centrifuged for 3 min at 12,000g and the cleared supernatant was used in the agglutination test with the rabbit red blood cells. Bacteria containing pGEX-KG lacking the DNA insert were used as a negative control. The test was conducted according to the standard procedure.

RESULTS

The lysate of *B. burgdorferi* was prepared as described above and used in hemagglutination tests to determine its agglutination activity. Rabbit and mouse red blood cells, both native and treated by trypsin, pronase and neuraminidase were checked in this test, but the use of native rabbit erythrocytes appeared to be the best detection system. The presence of agglutinating activity was detected even in the case when the amount of total protein was as low as 0.39 µg per well. Binding specificity of *B. burgdorferi* hemagglutination activity (HA) was analysed in the hemagglutination inhibition test and the strong inhibition of HA by fetuin was revealed.

After comparison the conserved protein regions of lectin genes of different organisms (Table 1) the theoretical sets of primers were designed for PCR on genomic DNA of *B. burgdorferi* (see "Primer design"). The schematic location of the designed primers on the "average" lectin gene is presented in Fig. 1. Only one PCR reaction with E-C primers gave the product 96 bp long (labelled as "E/C PCR product" for further use). The PCR product was sequenced (Fig. 2) and initially compared with the gene banks using BLAST and ALIGN, with no obvious matches being detected.

The predicted amino acid sequence of E/C PCR product was aligned with the already known sequences of different lectins initially used for our PCR primers design. The presence of the domain similar to one that represents characteristic conserved sequence motif for the legume lectin proteins was revealed (Fig. 3). To determine a copy number of lectin gene in *B. burgdorferi*, Southern blot hybridisation with borrelia DNA digested with various restriction enzymes was performed. The obtained E/C PCR product was used as a probe in hybridisation procedure. The multiple bands were detected according to the restriction enzyme used (Fig. 4).

Fourteen positive signals were obtained after the first round of hybridisation of six hundred recombinant clones from each mini-library with alpha-^{32}P dATP labelled E/C PCR product. The results of the second

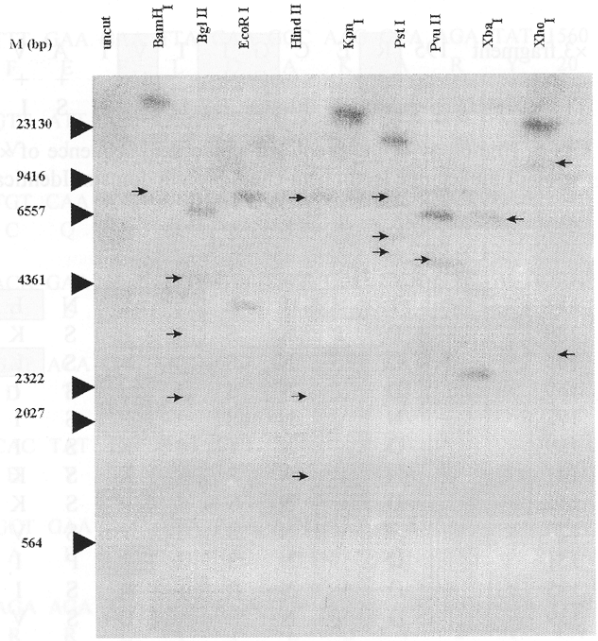

Fig. 4. Genomic Southern blot analysis of *Borrelia burgdorferi* (B31). Multiple bands are visualised in the Southern blot of various restriction enzyme digests of genomic DNA of *B. burgdorferi* hybridised with the E/C PCR product (96 bp) as a probe. M – DNA markers – lambda DNA digested by Hind III; → – presence of the bands on the autoradiogram that are invisible in this figure.

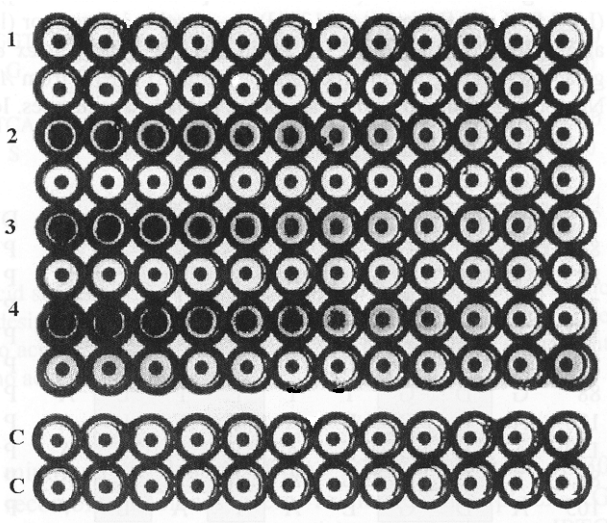

Fig. 5. Presence of agglutination activity in the crude lysate of the recombinant clones carrying the fragments of the *Borrelia burgdorferi* (B31) genomic DNA. 1 – the native plasmid pGEX-KG after IPTG induction (no insert); 2 – X3/pGEX-KG recombinant after IPTG induction; 3 – X230/pGEX-KG recombinant after IPTG induction; 4 – E200/pGEX-KG recombinant after IPTG induction; C – control RBC in agglutination buffer.

×3 fragment	195	R	E	C	G	I	T	V	I	A	V	R	N	L	S	N	S	R	Y	E	F	V	D	G	D	218
		R			G	I		V		+	+	R	+	+			+	R	+	+	F	V	+	G	+	
		R	H	I	G	I	D	V	N	S	I	R	S	I	K	T	T	R	W	D	F	V	N	G	E	179

Fig. 6. Homology of the predicted amino acid sequence of ××3 (195-218aa) to phytohemagglutinin (PHA-L) in the conserved region of the legume lectin beta-chain protein domain. Identical amino acids are boxed and shaded.

198	G	I	T	V	I	A	V	R	N	L	S	N	S	R	×3	
189	G	I	D	I	K	S	V	R	S	K	K	T	A	K	CONA CANEN	– 1
143	G	I	D	V	N	S	I	M	S	L	K	A	V	N	LEC LOTTE	– 2
159	G	I	D	V	N	S	I	R	S	I	K	T	T	R	PHAL PHAVU	– 3
172	G	I	N	V	N	S	I	R	S	I	K	T	T	S	LEC SOYBN	– 4
139	G	I	D	V	N	T	I	K	S	I	S	T	K	S	LEC VICFA	– 5
141	G	I	N	V	N	S	V	K	S	K	I	T	T	P	LEC ONOVI	– 6
145	G	I	D	V	N	S	I	R	S	K	A	A	S	K	LEC BOWMI	– 7
179	G	I	N	V	N	S	T	V	S	V	A	T	T	R	LEC BAUPU	– 8
151	E	I	D	V	N	S	I	R	P	I	A	T	E	S	ARC1 PHAVU	– 9
145	G	I	D	V	N	R	V	K	S	I	N	A	E	R	LEC1 ULEEU	– 10
139	G	I	D	V	N	S	V	D	S	V	K	T	V	P	LECG ARAHY	– 11

Fig. 7. Alignment of the predicted amino acid sequence of ×3 (198-211aa) with sequences of the legume lectin beta-chain proteins (BLOCK BLOO307D).

Comparison of the translated amino acid sequence to: 1 – concanavalin A precursor (CON A) from *Canavalia ensiformis* (horse bean) (CONA CANEN [Cunningham et al. 1975]); 2 – anti-H(O) lectin (LTA) from *Lotus tetragonolobus* (winged pea) (LEC LOTTE [Konami et al. 1990]); 3 – leucoagglutinating phytohemagglutinin precursor (PHA-L) from *Phaseolus vulgaris* (kidney bean) (PHAL PHAVU [Hoffman and Donaldson 1985]); 4 – lectin precursor (agglutinin) (SBA) from soybean (LEC SOYBN [Vodkin et al. 1983]); 5 – favin (lectin) from *Vicia faba* (broad bean) (LEC VICFA [Hopp et al. 1982]); 6 – lectin from *Onobrychis viciifolia* (common sainfoin) (LEC ONOVI [Kouchalakos et al. 1984]); 7 – lectin (agglutinin) (BMA) from *Bowringia mildibraedii* (LEC BOWMI [Chawla et al. 1993]); 8 – lectin precursor from *Bauhinia purpurea* (camel's foot tree) (LEC BAUPU [Kusui et al. 1991]); 9 – arcelin-1 precursor (legume lectin) from *Phaseolus vulgaris* (ARC1 PHAVU [Anthony at al. 1991]); 10 – anti-H (o) lectin I (UEA-I) from *Ulex europaeus* (furze) (LEC1 ULEEU [Konami et al. 1991b]); 11 – galactose-binding lectin precursor (agglutinin) (PNA) from *Arachis hypogaea* (peanut) (LECG ARAHY [Young et al. 1991]). Numbering of amino acids is as per the published sequences. Identical amino acids are boxed and shaded.

215	V	D	G	D	Y	F	F	L	K	D	D	K	×3	
87	V	D	G	L	A	F	F	L	A	P	A	N	LEC1 ABAL	– 1
88	V	D	G	L	A	F	F	L	A	P	A	N	LEC2 ULEEU	– 2
75	A	D	G	L	A	F	I	A	P	P	D	S	LECA DOLLA	– 3
114	A	D	G	L	V	F	F	M	G	P	T	K	LEC ERYCO	– 4
85	T	D	G	L	V	F	F	L	A	P	V	G	LEC LOTTE	– 5
88	G	D	G	I	T	F	F	L	A	P	T	D	LEC ONOVI	– 6
120	A	D	G	F	A	F	F	L	A	P	V	D	LEC BAUPU	– 7
119	A	D	G	L	A	F	F	L	A	P	I	D	LEC SOYBN	– 8
88	T	D	G	L	A	F	F	L	A	P	A	D	LEC2 MEDTR	– 9
105	A	D	G	L	A	F	A	L	V	P	V	G	PHAL PHAVU	– 10

Fig. 8. Alignment of the predicted amino acid sequence of ×3 (215-226aa) with the sequences of the legume lectin beta-chain proteins (BLOCK BLOO307B, **"DGF"** motif).

Comparison of the translated amino acid sequence to: 1 – lectin I (seed lectin (LAA-1) from *Laburnum alpinum* (Scotch laburnum) (LEC1 LABAL [Konami et al. 1991c]); 2 – anti-H(O) lectin II (UEA-II) from *Ulex europaeus* (furze) (LEC2 ULEEU [Konami et al. 1991a]); 3 – lectin from *Dolichos lab lab* (field bean) (LECA DOLLA [Gowda et al.1994]); 4 – lectin precursor from *Erythrina corallodendron* (coral tree) (LEC ERYCO [Adar et al. 1989]); 5 – anti-H(O) lectin (LTA) from *Lotus tetragonolobus* (winged pea) (LEC LOTTE [Konami et al. 1990]); 6 – lectin from *Onobrychis viciifolia* (common sainfoin) (LEC ONOVI [Kouchalakos et al. 1984]); 7 – lectin precursor from *Bauhinia purpurea* (camel's foot tree) (LEC BAUPU [Kusui et al. 1991]); 8 – lectin precursor (agglutinin) (SBA) from soybean (LEC SOYBN [Vodkin et al. 1985]); 9 – truncated lectin2 precursor (gene lec2) from *Medicago truncatula* (barrel medic) (LEC2 MEDTR [Bauchrowitz et al. 1992]); 10 – leucoagglutinating phytohemagglutinin precursor (PHA-L) from *Phaseolus vulgaris* (kidney bean) (PHAL PHAVU [Hoffman and Donaldson 1985]). Numbering of amino acids is as per the published sequences. Identical amino acids are boxed and shaded.

```
1501 ATG GCT GTT TTT GTT TCA AGA AAG TCG CGT TTT GAA GAA TTA CAA GCC AAG GCA AGA TAT 1560
       M   A   V   F   V   S   R   K   S   R   F   E   E   L   Q   A   K   A   R   Y  20

     TTT GGT TGG TTG AAG CAT ATG AAA ACA TTT GTT ATT ATT GGA CTT AGT AAT TTA GGC ATT 1620
       F   G   W   L   K   H   M   K   T   F   V   I   I   G   L   S   N   L   G   I  40

     CAC TTA CTT GAA GAT TTA AGC AGG CTT GAT TGT CAA ATT ATT ATT ATA GAT ACA TCT AAA 1680
       H   L   L   E   D   L   S   R   L   D   C   Q   I   I   I   I   D   T   S   K  60

     GAG CTT ATT GAA GAA TAT GAT GTG ATA TCT ACA GAA AGC TTT GTT GTT GAG CAA TTC ACT 1740
       E   L   I   E   E   Y   D   V   I   S   T   E   S   F   V   V   E   Q   F   T  80

     AAA AAT GCT TTG AAA AGA ATA ATT CCA GTA GAT ACA GAC GCT GTT GTT ATT GAT TTT GAT 1800
       K   N   A   L   K   R   I   I   P   V   D   T   D   A   V   V   I   D   F   D 100

     GAT GAT CTT GGC AAA AGT GCT CTT GTT ACT CAC TAT TGT AAT CTT TTA GGT TTG AAA GAA 1860
       D   D   L   G   K   S   A   L   V   T   H   Y   C   N   L   L   G   L   K   E 120

     ATA TGC GTT AAG ACA GAA AAT AGA GAT GAT GCT GAA ATC TTA AAA ACT CTT GGG GCA ACA 1920
       I   C   V   K   T   E   N   R   D   D   A   E   I   L   K   T   L   G   A   T 140

     AAA ATT ATA TTT CCA AGT AAA GAT GCC GCC AGA AGA TTA ACT CCA TTA TTA GTA TCT CCA 1980
       K   I   I   F   P   S   K   D   A   A   R   R   L   T   P   L   L   V   S   P 160

     AAT CTT TCA ACT TAT AAT ATT ATT GGG TAT GAT ATT ATT GTG GCT GAA ACT GTT ATT CCC 2040
       N   L   S   T   Y   N   I   I   G   Y   D   I   I   V   A   E   T   V   I   P 180

     AAA AGA TAT GTT GGT AAA ACT CTT TTT GAA GCC GAT CTT AGA AGA GAA TGT GGG ATT ACA 2100
       K   R   Y   V   G   K   T   L   F   E   A   D   L   R   R   E   C   G   I   T 200

     GTT ATT GCT GTT AGA AAT TTA AGT AAT TCT AGG TAT GAA TTT GTT GAT GGC GAT TAT TTT 2160
       V   I   A   V   R   N   L   S   N   S   R   Y   E   F   V   D   G   D   Y   F 220

     TTT TTA AAA GAT GAT AAA ATT GTA ATT TGT GGT AAA CCA GAT AGC ATT GAA AAT TTT ACA 2220
       F   L   K   D   D   K   I   V   I   C   G   K   P   D   S   I   E   N   F   T 240

     AAT AAT AAA GAT TTA ATT AAA GAT TTA ATT TCA GGC TCT AAA GAG GAT GAA AAT TTA AAT 2280
       N   N   K   D   L   I   K   D   L   I   S   G   S   K   E   D   E   N   L   N 260

     AAA GAT GCT GAG AAA AAA TCT AGA                                               2224
       K   D   A   E   K   K   S   R                                               268
```

Fig. 9. The partial nucleotide sequence and the predicted amino acid sequence of X3 clone carrying the fragment of the genomic DNA of *B. burgdorferi* (B31). Amino acids are numbered by designating the first methionine as amino acid 1. Nucleotide numbers are shown at the right of each line. The numbers of amino acid residues starting from the first Met are given at the right of each line under nucleotide numbers. One-letter symbols of amino acid residues were used.

screening showed the strong positive signals only in four cases: clone E200 (3250 bp insert) from "EcoR mini-library" and X3 (2350 bp insert), X139 (2350 bp insert), X167 (2350 bp insert) clones from "XbaI minilibrary". The inserts from X3 and E200 recombinants were used as the probes for cross-hybridisation. Using X3 PCR product as a probe, only one strong positive signal was obtained in the case of E200 clone from "EcoR I mini-library" and the strong positive signals were obtained in the case of X3, X139, X167 and the additional X230 clone (2250 bp insert) from "XbaI minilibrary". Using E200 PCR product as a probe resulted in the only positive signal – E200 from "EcoR I mini-library" and X3, X139, X167 and X230 clones from "XbaI minilibrary". Inserts from E200, X3 and X230 recombinant clones were re-cloned into pGEX-KG expressing vector (Pharmacia) with the strong IPTG inducible "tac" promoter. The expressive agglutination was received in all three cases with E200, X3 and X230 recombinants (Fig. 5). The strong inhibition of agglutination by fetuin, D-galactosamine and D-mannosamine was observed when the crude lysates of these recombinants were used in hemagglutination inhibition test. D-glucuronic acid showed some weaker inhibition activity as well.

The complete sequence of the XbaI/XbaI fragment of genomic DNA from the recombinant clone X3 was obtained using T3, T7 and the internal primers.

Sequence analysis and amino acid translation were performed using the programs from the 1993-96 DNASTAR. The comparison of the obtained sequence with the data of PIR, EMBL, Swiss Protein, and GenBank databases were made by FASTA program. The nucleotide sequence analysis revealed the presence of an open reading frame of 268 amino acids (position 1501-2304 bp) containing the sequence that was 37% identical and 75% homologous to the conserved region of proteins belonging to the clusters 341-33 and 341-35, i.e. PHA-E and PHA-L respectively and contained the potential glycosylation site -**SNS** (Fig. 6). The further structural analysis of the deduced protein sequence ("SCOP: Structural Classification of Proteins") added another five amino acids to the previous 24 amino acids, confirming the homology to the above mentioned groups of proteins.

Alignment of the predicted amino acid sequence with the sequences of already known lectins revealed the presence of at least two domains with high percentage of similarity to legume lectin beta-chain proteins. The fragment from 198 to 211 amino acids showed significant homology to the legume lectin proteins from block BLOO307D with the "**GIN**" motif (Fig. 7), while the fragment from 215 to 226 amino acids indicated the similarity to the legume lectin proteins from block BLOO307C with the "**DGF**" motif (Fig. 8).

DISCUSSION

Our results confirmed the existence of hemagglutination activity in a crude lysate of the spirochete *B. burgdorferi* that was previously described by Grubhoffer et al. (1993) and showed the strong inhibition of this activity by fetuin. As a result of PCR with the degenerative primers a 96bp fragment of genomic DNA was received, isolated and sequenced. The alignment of the predicted amino acid sequence of this fragment indicates the presence of the domain that consists of the characteristic conserved sequence motif for the legume lectins. The Southern blot analysis of the genomic DNA of *B. burgdorferi* with the above-mentioned DNA fragment as a probe indicated the presence of multiple bands according to the restriction enzyme used. These bands may be derived from the related sequences that are present in *B. burgdorferi* genome and confirm the existence of at least two (or more) related genes coding proteins with hemagglutination activity. We isolated and sequenced the XbaI/XbaI fragment (2304 bp) of the genomic DNA of *B. burgdorferi*. The recombinant clone that carried the 2304bp genomic insert revealed the strong agglutination activity against native rabbit erythrocytes. The same strong inhibition of agglutination by fetuin as in the case with native borrelia was observed. In addition to fetuin, the strong inhibition of agglutination was also observed with D-galactosamine and D-mannosamine. The nucleotide sequence analysis of the genomic insert revealed the presence of an open reading frame of 268 amino acids (position 1501-2304 bp). The calculated molecular size of the predicted protein is 30324 Da and pI=5.44. GenBank searching for the plain homology of the protein sequence showed that it contained the sequence that was 37% identical and 75% homologous to the conserved region of proteins belonging to the clusters 341-33 and 341-35, i.e. PHA-E and PHA-L respectively and contained the potential glycosylation site -**SNS**. The alignment of the predicted amino acid sequence with already known protein sequences of lectins indicates two domains (198-211 aa and 215-226 aa) consisting of the conserved sequence motifs characteristic for legume lectin proteins. Two signature patterns specific to legume lectins were described previously (Lis and Sharon 1986, Sharon and Lis 1990). The first pattern, i.e. [LIV]-[STAG]-**V**-[EQV]-[FLI]-**D**-[ST] or (XXVXXDX) is located in the C-terminal section of the beta chain and contains a conserved aspartic acid residue important for the binding of calcium and manganese. The second pattern, i.e. [LIV]-X-[EDQ]-[FYWKR]-**V**-X-[LIV]-**G**-[LF]-[ST] or (XXXX**V**XX**G**XX) is located in the N-terminal of the alpha chain. The analysis of predicted 268 amino acid sequence of our recombinant revealed the presence of 10 amino acids (position 26-35, – "MKTF**V**II**G**LS") in the N-terminal region of protein that corresponds to the second developed signature pattern specific to legume lectins. As a result of GenBank searching for the plain homology of our 268 aa protein with the other proteins we found that it showed 94% identity with the protein from *Borrelia burgdorferi* that was designed as "conserved hypothetical protein", giI2688659 (AE001172). Fifteen amino acids were missing from the C-terminal of the predicted protein in the comparison with giI2688259, because the fragment of genomic DNA that we isolated was restricted by XbaI site from both ends. From the other side our 268aa protein contained 23 additional amino acids on the N-terminal that were absent in the conserved hypothetical protein.

In conclusion, we have isolated, cloned, sequenced and partially characterised the gene coding the protein with hemagglutinating activity from the genome of *B. burgdorferi* B31 that we designated as B31LEC for the further use (Fig. 9). The sequence was submitted to the GenBank and the submission was accepted under the AF028001 NCBI accession number (giI3138928).

Acknowledgements. This work was financially supported by the Grant VŠ96066 of the Ministry of Education of the Czech Republic and the Grant A6087601 of the Grant Agency of the Academy of Sciences of the Czech Republic. The authors are very grateful to Mrs. I. Švecová, L. Štifterová and H. Žemličková for the skilled technical assistance, Mgr. P. Svoboda for the assistance in the computer analysis of the obtained data and RNDr. L. Jindrák for the useful discussion and research advice during this work.

REFERENCES

ADAR R., RICHARDSON M., LIS H., SHARON N. 1989: The amino acid sequence of *Erythrina corallodendron* lectin and its homology with other legume lectins. FEBS Lett. 257: 81-85.

ANTHONY J.L., VONDER HAAR R.A., HALL T.C. 1991: Nucleotide sequence of a genomic clone encoding arcelin, a lectin-like seed protein from *Phaseolus vulgaris*. Plant Physiol. 97: 839-840.

BARBOUR A.G., HAYES S.F. 1986: Biology of *Borrelia* species. Microbiol. Rev. 50: 381-400.

BARONDES S.H. 1988: Bifunctional properties of lectins: lectins redefined. Trends Biochem. Sci. 12: 480-482.

BAUCHROWITZ M.A., BARKER D.G., NADAUD I., ROUGE P., LESCURE B. 1992: Lectin genes from the legume *Medicago truncatula*. Plant Mol. Biol. 19: 1011-1017.

CHAWLA D., ANIMASHAUN T., HUGHES R.C., HARRIS A., AITKEN A. 1993: *Bowringia mildibraedii* agglutinin composition, primary structure and homologies with other legume lectins. Biochim. Biophys. Acta 1202: 36-46.

CUNNINGHAM B.A., HEMPERLY J.J., HOPP T.P., EDELMAN G.M. 1979: Favin versus Concanavalin A: circularly permuted amino acid sequences. Proc. Natl. Acad. Sci. U.S.A. 76: 3218-3222.

CUNNINGHAM B.A., WANG J.L., WAXDAL M.J., EDELMAN G.M. 1975: The covalent and three-dimensional structure of concanavalin A. II. Amino acid sequence of cyanogen bromide fragment F3. J. Biol. Chem. 250: 1503-1512.

DELBAERE L.T.J., VANDONSELAAR M., PRASAD L., QUAIL J.W., WILSON DAUTER Z. 1993: Structure of the lectin IV of *Griffonia simplicifolia* and its complex with the Lewis b human blood group determinant at 2.0Å resolution. J. Mol. Biol. 230: 950-965.

GARCIA-MONCO J., FERNANDEZ-VILLAR B., ROGERS R., SZCZEPANSKI A., WEELER C., BENACH J. 1992: *Borrelia burgdorferi* and other related spirochetes bind to galacto-cerebroside. Neurology 42: 1341-1348.

GRUBHOFFER L., UHLÍŘ J., VOLF P. 1993: Functional and structural identification of a new lectin activity of *Borrelia recurrentis* spirochetes. Comp. Biochem. Physiol. 105B: 535-540.

GRUBHOFFER L., VEREŠ J., DUSBÁBEK F. 1991: Lectins as the molecular factors of recognition and defence reaction of ticks. In: F. Dusbábek and V. Bukva (Eds.), Modern Acarology. Academia, Prague and SPB Academic Publishing bv, The Hague, Vol. 2, pp. 381-388.

GOWDA L.R., SAVITHRI H.S., RAJAGOPAL RAO D. 1994: The complete primary structure of a unique mannose/glucose-specific lectin from field bean (*Dolichos lab lab*). J. Biol. Chem. 269: 18789-18793.

GUAN K.L., DIXON J.E. 1991: Eukaryotic proteins expressed in *Escherichia coli:* an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. Anal. Biochem. 192: 262-267.

HOFFMAN L.M., DONALDSON D.D. 1985: Characterisation of two *Phaseolus vulgaris* phytohemagglutinin genes closely linked on the chromosome. EMBO J. 4: 883-889.

HOPP T.P., HEMPERLY J.J., CUNNINGHAM B.A. 1982: Amino acid sequence and variant forms of favin, a lectin from *Vicia faba*. J. Biol. Chem. 257: 4473-4483.

HULÍNSKÁ D., VOLF P., GRUBHOFFER L. 1992: Study of *Borrelia burgdorferi* glycoproteins and surface carbohydrates by lectin conjugates. Zentralbl. Bacteriol. Microbiol. Hyg. A, 276: 473-480.

ISAACS R. 1994: *Borrelia burgdorferi* bind to epithelial cell proteoglycan. J. Clin. Invest. 93: 809-819.

KOCOUREK J., HOŘEJŠÍ V. 1981: Defining a lectin. Nature 290: 188.

KONAMI Y., YAMAMOTO K., OSAWA T. 1990: The primary structure of the *Lotus tetragonolobus* seed lectin. FEBS Lett. 268: 281-286.

KONAMI Y., YAMAMOTO K., OSAWA T. 1991a: Purification and characterisation of a new type lactose-binding *Ulex europaeus* lectin by affinity chromatography. Hoppe-Seyler's Z. Biol. Chem. 372: 95-102.

KONAMI Y., YAMAMOTO K., OSAWA T. 1991b: The primary structure of the two types of the *Ulex europaeus* seed lectin. J. Biochem. 109: 650-658.

KONAMI Y., YAMAMOTO K., TOYOSHIMA S., OSAWA T. 1991c: The primary structure of the *Laburnum alpinum* seed lectin. FEBS Lett. 286: 33-38.

KOUCHALAKOS R.N., BATES O.J., BRADSHAW R.A., HAPNER K.D. 1984: Lectin from sainfoin (*Onobrychis viciifolia* Scop*.).* Complete animo acid sequence. IS Biochemistry 23: 1824-1830.

KUSUI K., YAMAMOTO K., KONAMI Y., OSAWA T. 1991: cDNA cloning and expression of *Bauhinia purpurea* lectin. J. Biochem. 109: 899-903.

LEONG J.M., MORRISSEY P.E., ORTEGA-BARRIA E., PEREIRA M.E.A., COBURN J. 1995: Hemagglutination and proteoglycan binding by the Lyme disease spirochete, *Borrelia burgdorferi.* Infect. Immun. 63: 874-883.

LIS H., SHARON N. 1986: Lectins as molecules and as tools. Annu. Rev. Biochem. 55: 33-37.

MAUDLIN I., WELBURN S.C. 1988: The role of lectins and trypanosome genotype in the maturation of midgut infections in *Glossina morsitans.* Tropenmed. Parasitol. 39: 56-58.

OFEK L., SHARON N. 1988: Lectinophagocytosis: a molecular mechanism of recognition between cell surface sugars and lectins in the glycoprotein of bacteria. Infect. Immun. 50: 539-547.

SAMBROOK J., FRITSCH E.F., MANIATIS T. 1989: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. pp. 1.63-1.67, 1.74-1.84.

SANGER F.S.N., COULSON A.R. 1977: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74: 5463.

SHARON N., LIS H. 1989: Lectins. Chapman and Hall, London, 127 pp.

SHARON N., LIS H. 1990: Legume lectins - a large family of homologous proteins. FASEB J. 4: 3198-3208.

SMITH D.B., JOHNSON K.S. 1988: Single-step purification of polypeptides in *Escherichia coli* as fusions with glutathtione S-transferase. Gene 67: 31-40.

SOUTHERN E.M. 1975: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

SZCZEPANSKI A., FURIE M.B., BENACH J.L., LANE B.P., FLEIT H.B. 1990: Interaction between *Borrelia burgdorferi* and endothelium *in vitro*. J. Clin. Invest. 85: 1637-1647.

VODKIN L.O., RHODES P.R., GOLDBERG R.B. 1983: cA lectin gene insertion has the structural features of a transposable element. Cell 34: 1023-1031.

YOUNG N.M., JOHNSTON R.A.Z., WATSON D.C. 1991: The amino acid sequence of peanut agglutinin. Eur. J. Biochem. 196: 631-637.

Received 18 June 1998

Accepted 3 February 1999