

Clathrin Light Chains LCA and LCB Are Similar, Polymorphic, and Share Repeated Heptad Motifs

T. KIRCHHAUSEN,* P. SCARMATO, S. C. HARRISON, J. J. MONROE, E. P. CHOW, R. J. MATTALIANO,
K. L. RAMACHANDRAN, J. E. SMART, A. H. AHN, AND J. BROSTITUS*

Clathrin Light Chains LCA and LCB Are Similar, Polymorphic, and Share Repeated Heptad Motifs

T. KIRCHHAUSEN,* P. SCARMATO, S. C. HARRISON, J. J. MONROE, E. P. CHOW, R. J. MATTALIANO, K. L. RAMACHANDRAN, J. E. SMART, A. H. AHN, J. BROSIUS*

The clathrin light chains fall into two major classes, LCA and LCB. In an intact clathrin triskelion, one light chain, of either class, is bound to the proximal segment of a heavy chain leg. Analysis of rat brain and liver complementary DNA clones for LCA and LCB shows that the two light chain classes are closely related. There appear to be several members of each class having deletions of varying length aligned at the same position. A set of ten heptad elements, characteristic of α -helical coiled coils, is a striking feature of the central part of each derived amino acid sequence. These observations suggest a model in which the α -helical segment mediates binding to clathrin heavy chains and the amino- and carboxyl-terminal segments mediate interactions with other proteins. They also suggest an explanation for the observed tissue-dependent size variation for members of each class.

RECEPTOR-MEDIATED ENDOCYTOSIS occurs by entrapment of specific macromolecules in coated pits and coated vesicles (1). Clathrin is the main structural component of the lattice covering the cytoplasmic face of these structures (2-4). Assembly of clathrin results in formation of the coat, and this process may be the driving force that leads to vesiculation. Clathrin is a large, soluble protein composed of heavy chains with a molecular size of ~192 kD, deduced from the sequence of the rat complementary DNA (cDNA) heavy chain clones (5) and light chains of about 32 to 38 kD by SDS-polyacrylamide gel electrophoresis (PAGE) (2-4). The unit that assembles into coats has three extended legs, 500 Å in length, splayed out in a pinwheel-

like structure (triskelion). Each of the legs is built from a single heavy chain (3, 4), with a light chain bound to its proximal segment (6-8). The light chains are thought to be important for vesicle uncoating (9), and the multiplicity of light chain species suggests possible other roles in the cycle of coated vesicle assembly and disassembly. Two major classes of clathrin light chains, referred to as LCA and LCB, have been identified (3, 4, 10). Moreover, the existence of several types within each class has been suggested by the detection of LCA and LCB chains of at least two different sizes, whose relative amounts depend on the tissue examined (10). Light chains (LCA or LCB) from tissues other than brain are smaller than their brain counterparts by 2000 to 3000 daltons (10). Brain

also contains about 5 percent of the smaller versions (11). Purified LCA and LCB will compete with each other for the same binding site on the heavy chain (12). The peptide maps of LCA and LCB are different (11-13), however, and monoclonal antibodies to each of the light chains can distinguish chains of each class (6, 14). Taken together, these observations suggest that LCA and LCB are related proteins that are encoded by different genes.

We have analyzed rat brain and liver cDNA clones for LCA and LCB. We find that light chains of the two classes are indeed closely related. We also find evidence for several members of each class, related by deletions of varying length aligned at the same position. A significant common feature of the LCA and LCB amino acid sequences is a set of ten heptad elements, characteristic of α -helical coiled coils. We suggest that these helical segments bind to heavy chains and that the remaining parts of the light chains are sites for interaction with other proteins.

The major species of bovine brain clathrin light chains LCA (~36 kD) and LCB (~33 kD) were isolated by preparative SDS (12 percent)-PAGE from a crude mixture of soluble light chains obtained from a boiled solution of clathrin trimers (11, 15, 16). The electroeluted light chains were succinylated and digested with trypsin, and several tryptic fragments were purified by reversed-phase high-pressure liquid chromatography (17). Three peptides from LCA and two peptides from LCB were selected for determination of amino-terminal amino acid sequences by automated Edman degradation. From this information, five oligonucleotide probes were prepared by chemical synthesis—three for LCA and two for LCB (Fig. 1). These probes were γ -³²P-labeled with T4 kinase and used to identify hybridizing recombinants in a λ gt10 cDNA library derived from rat brain cytoplasmic polyadenylated [poly(A)⁺] RNA (18, 19). Seven independent clones hybridized to LCA probes and six to LCB probes.

DNA sequence analysis of this first set of clones (Fig. 2) showed that the isolated cDNA's indeed encode LCA and LCB. We

1) LCA	K Q E A E W K E K A I K E L D E W Y A
	AAG CAG GAG GCC GAA TGG AAA GAA AAA GCC ATC AAG GAG CTG GAC GAG TTG TAC GC
	G G G G
2) LCA	V A D E A F Y K Q P F A (M) V I G Y V ? N
	GCC GAC GAG GCC TTC TAC AAA CAA CCC TTC GC
	T T G G
3) LCA	S V L I S L K E A P L V
	GTG CTG ATC TCC CTG AAA GAA GCC CCT CTG GT
	T G G T
4) LCB	L Q E L D A A S K V M E Q E W R
	CTG CAG GAG CTG GAC GCC GCC TCC AAG GTG ATG GAA CAA GAA TGG AG
	G G G C
5) LCB	E K A K K D L E E W N Q
	AAG GCC AAG AAG GAC CTG GAA GAA TGG AAC CA
	T G G T

Fig. 1. Clathrin light chains protein sequence analysis and nucleotide probes. The amino acid sequences of several tryptic fragments from bovine brain clathrin light chains were determined by Edman degradation. Uncertainties in the assignment are indicated. Partially degenerate oligonucleotide probes were designed, based on the pattern of codon usage in rat (28) and prepared by solid-phase chemical synthesis (29).

T. Kirchhausen, P. Scarmato, J. J. Monroe, Department of Anatomy and Cellular Biology, Harvard Medical School, Boston, MA 02115.

S. C. Harrison, Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.

E. P. Chow, R. J. Mattaliano, K. L. Ramachandran, J. E. Smart, Biogen Research Corp., Cambridge, MA 02142. A. H. Ahn and J. Brosius, Department of Genetics and Behavior, College of Physicians and Surgeons, Columbia University, New York, NY 10032.

*To whom correspondence may be sent.

found very high homology between the sequences of the bovine tryptic peptides and the predicted rat sequences (Fig. 3). The agreement includes the correct prediction of the arginine residues at which the tryptic cleavages had occurred. Furthermore, we have also identified partial length bovine brain cDNA clones that encode LCA. The region spanned by the derived amino acid sequence exhibits a perfect match with two of the tryptic fragments obtained from bovine brain LCA and displays better than 90 percent amino acid identity with rat brain LCA (20).

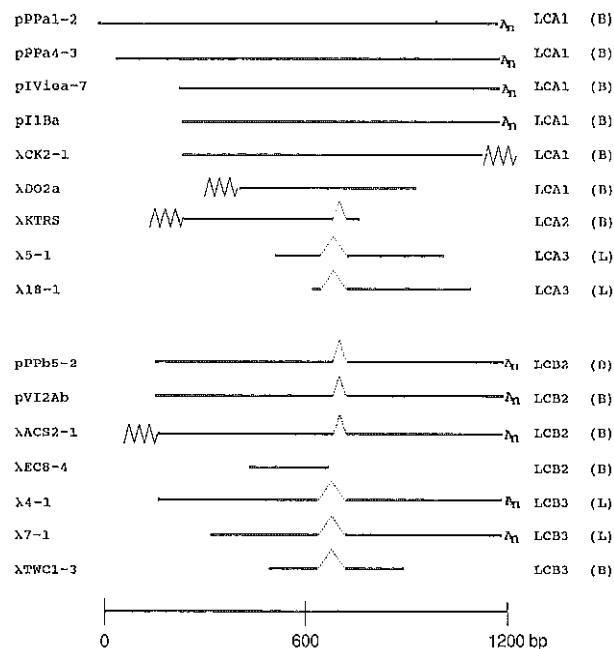
These initial rat brain cDNA clones did not contain the entire light chain sequence (Fig. 2). Additional recombinants coding for light chains were therefore isolated from a search of a plasmid cDNA library from rat brain cerebral cortex poly(A)⁺ RNA; the numbers of independent clones obtained for LCA and LCB were five for each, respectively. We also screened a λ gt10 cDNA library from rat liver cytoplasmic poly(A)⁺ RNA (18), from which we analyzed two independent clones for LCA and two for LCB.

The longest LCA clone contains 1128 bases, and the longest LCB clone, 982 bases. Northern analysis predicts approximate messenger RNA (mRNA) sizes of 1.3 and 1.2 kb for LCA and LCB, respectively (Fig. 4). Provided that light chain mRNA species have poly(A) tails of 100 to 150 nucleotides, the Northern analysis shows that our sequences lack at most about 120 bases at the 5' end. The LCA sequence has a possible ATG initiation codon at position 116, predicting a polypeptide of 248 amino acids. The LCB sequence, with an ATG at position 140, predicts a smaller polypeptide of 229 amino acids. Even if the start sites were instead at upstream ATG's not yet determined, the maximum estimate for the lengths of LCA and LCB mRNA's would yield polypeptide chains of about 280 residues, somewhat smaller than the size determined by mobility in SDS gels—that is, about 37 and 34 kD for LCA and LCB in rat brain (21).

The translational stop codons of LCA and LCB can be aligned at homologous positions. The LCA clones have a 3' untranslated region of ~270 nucleotides. A putative poly(A) addition signal (AATAAA) is located at a distance of 14 nucleotides from the poly(A) tail. The LCB clones have a 3' untranslated region of similar length and a poly(A) tail, although they lack an apparent poly(A) addition signal. This could be due to either to an anomalous polyadenylation site or to internal priming during the cDNA synthesis.

A computer search (22) performed with

Fig. 2. Schematic representation of sequenced rat brain and rat liver cDNA clones for clathrin light chains LCA and LCB. The top portion of the figure indicates the length and relative alignment of the various sequenced brain (B) and liver (L) LCA cDNA clones derived from λ gt10 (λ) or plasmid (p) libraries in relation to LCA1. The type taken as reference is LCA1, and it does not contain deletions on its sequence. The bottom portion shows a similar alignment of sequenced LCB cDNA clones with respect to LCA1. Observed gaps in the nucleotide sequence are indicated by dotted lines. The deletions, coincident at their 3' ends, correspond to deletions of 12 or 30 amino acids (Fig. 3 and text). Zigzag lines indicate unrelated cDNA portions probably derived by fusion prior to linker addition. Poly(A) tails are marked by A_n. DNA sequence analysis was done on both DNA strands (30).

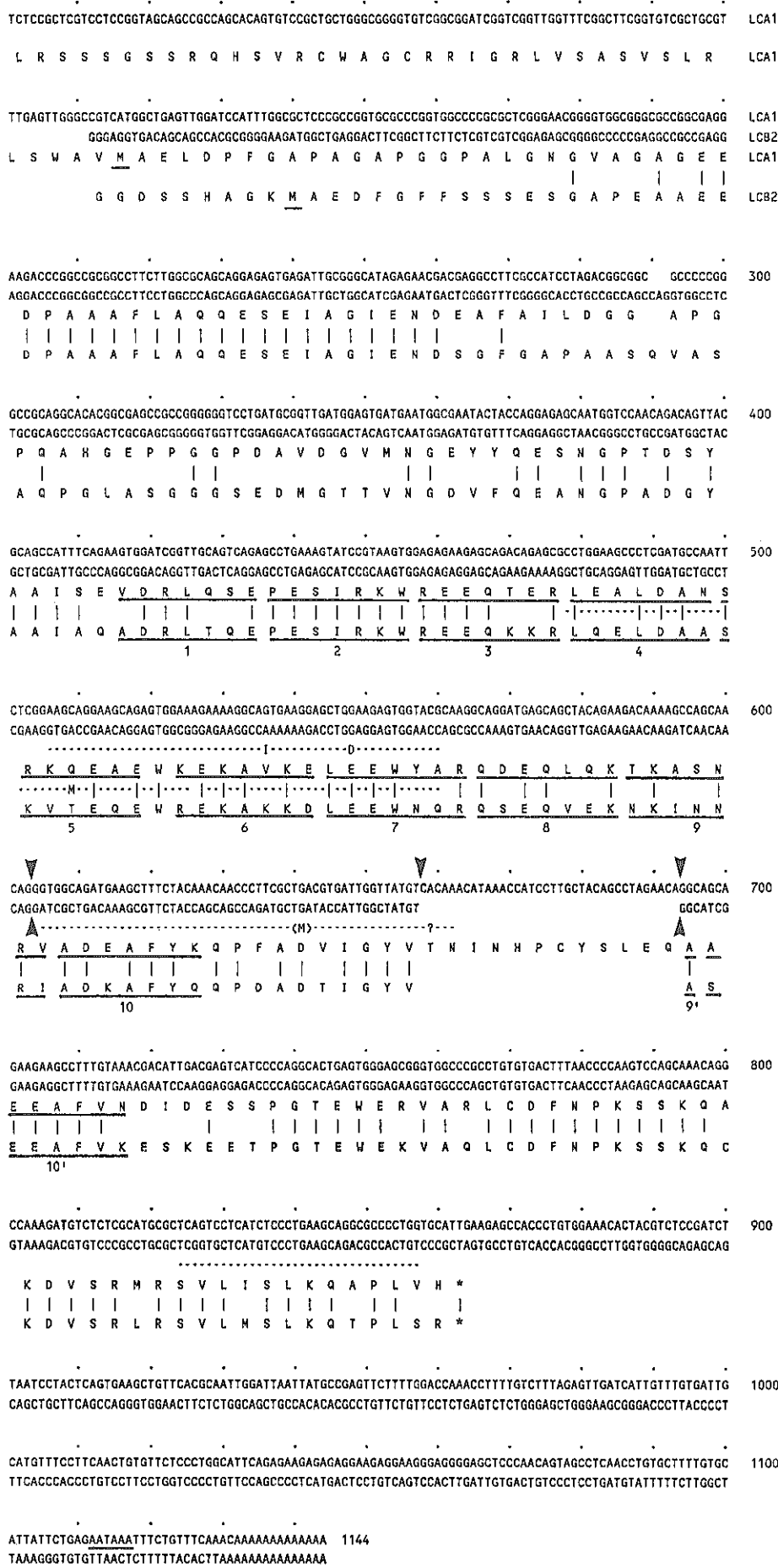


the rat LCA and LCB sequences failed to indicate any significant amino acid homology with rat clathrin heavy chain or with other sequenced proteins that have been published in the National Biomedical Research Foundation protein data bank (August 1986). An optimal alignment of LCA and LCB is shown in Fig. 3. With this arrangement, we detect ~60 percent identity and similarity of better than 80 percent if conservative changes are allowed. Most of the differences are near the amino termini. The extent of DNA homology (coding and 3' noncoding region) of light chains A and B suggests that their genes are related through an ancient gene duplication. In this regard we note that yeast clathrin has only one light chain (23).

For each class of light chain (LCA or LCB), we have found clones that encode proteins of several sizes, and refer to them as LCA1, LCA2, LCA3, LCB2, and LCB3. Of these, LCA1 is the longest, and we regard it as the reference. Representatives of its cDNA were found so far only in the brain cDNA libraries. The LCA2 type, of which only one clone was found, also from brain, is identical to LCA1 except for a deletion of 12 amino acids; the LCA3 (two clones found in liver) is lacking 30 amino acids. The deletions, coincident at their carboxyl termini, are indicated by arrowheads in Fig. 3. Except for these gaps, all LCA clones have identical DNA sequences, including the 3' untranslated regions. There are a few silent point mutations that probably reflect allelic variation. In an analogous fashion LCB2 (found so far only in brain) and LCB3 (one

clone in brain and two in liver) lack stretches of 12 and 30 amino acids when compared to LCA1, at precisely the same positions as the gaps in LCA2 and LCA3, respectively. As with the LCA group, the LCB2 and LCB3 clones show otherwise identical DNA sequences. The junctions that are generated at the 3' ends of all the gaps in LCA and LCB clones are aligned at the same position, 169 bases upstream from their respective stop codons. We consider it unlikely that these in-frame deletions found in the same region of both types of light chain cDNA clones are the result of a cloning artifact. Because of the identity of DNA sequence in the coding and noncoding regions within each class of light chain, we suggest that the size polymorphism derives from differential splicing of the corresponding RNA transcripts. Indeed, Northern analysis of rat RNA shows that mRNA from brain LCA is slightly larger than other LCA's (Fig. 4). The size difference is less discernible for LCB. Other examples of differential splicing within the coding region have been observed in fibronectin, troponin T, and the neural cell adhesion molecule (24). We believe that the multiple cDNA's we find for each class correspond to the multiple species observed on protein gels, but we cannot yet assign a particular cDNA to a particular SDS gel band.

The amino acid sequences of both classes of light chain have in one segment a striking pattern, typical of α -helices arranged in a coiled coil structure (Figs. 3 and 5) (25). The pattern consists of a series of ten heptads (underlined in Fig. 3) with a sequence



of the format a, b, c, d, e, f, g, where positions a and d, in general, contain hydrophobic or neutral amino acids while the other positions contain most of the polar and charged residues. This distribution of amino acid residues is best seen in the arrangement in Fig. 5. In both light chains the periodicity is broken between the fifth and sixth heptads by the insertion of a tryptophan residue. Such disturbances, known as "skip residues," are found in other α -helical coiled coils. They are expected to produce a local flattening of the superhelix (25). Our assignment of an α -helical structure to this segment is further supported by the virtual absence of proline and glycine residues. There is a proline at the first position of heptad 2. Incorporation of proline into an α -helix introduces a kink and ruptures a hydrogen bond. The helical structure might therefore really begin at this position, and the seven preceding residues could correspond only by coincidence to a consensus heptad. Prolines are present a few residues beyond the carboxyl terminus of the heptad region in all light chain types.

The deletions that distinguish different light chain types fall near the carboxyl-terminal end of the heptad region. In LCA2 and LCB2, the deletion occurs beyond the last heptad. In LCA3 and LCB3, the ninth heptad of the "parent" sequence is interrupted between positions f and g, but the consensus characteristics of the final two heptads are conserved by the sequences fused to it. That is, the join in LCA3 and LCB3 is such that the acceptor end provides the last

Fig. 3. Nucleotide sequences from rat brain and rat liver cDNA clones of clathrin light chains LCA and LCB. The sequences compiled from several independent clones for each light chain type are shown, together with the predicted amino acid sequences. The top line in each row shows the LCA1 sequence; the second line, LCB2. Deletions that yield LCA2, LCA3, and LCB3 are shown by arrowheads. In particular, LCA2 is related to LCA1 by deletion of nucleotides 657 through 693; LCA3 to LCA1 by deletion of nucleotides 604 through 693; and LCB3 to LCB2 by deletion of nucleotides 604 through 657. We have numbered the LCB chain to correspond to the LCA types. It is possible that further search will produce an LCB1, 12 amino acids longer than LCB2. So far, clones for LCA1, LCA2, and LCB2 were found in brain cDNA libraries; clones for LCA3 in liver cDNA libraries, and clones for LCB3 in cDNA libraries from both brain and liver. A schematic map of the clones that were sequenced is given in Fig. 2. The clones span 982 to 1128 bp and include 3' untranslated regions with poly(A) stretches. Possible ATG initiation sites are indicated, and a polyadenylation signal in LCA is underlined. Peptides of bovine brain clathrin light chains that were sequenced by Edman degradation (Fig. 1) are overlaid; amino acids that differ or whose determination was uncertain are indicated. The heptads are indicated by solid lines below the sequence and numbered 1 through 10 (Fig. 5).

residue of a new ninth heptad and a seven-amino acid sequence appropriate for a new tenth heptad (Figs. 3 and 5). At the nucleotide level, the actual junction appears to occur between nucleotides 2 and 3 of the position f codon in heptad 9, and the spliced sequence is such that an arginine at position f is conserved, a valine or isoleucine at position g is changed to alanine, and a new heptad 10' has significant sequence homology to the deleted one. This new heptad is a sequence also present in LCA1, LCA2, and LCB2; but it is separated from the repeat region by a nonperiodic segment that includes a proline. We believe that the remarkable conservation of structure is strong evidence both for the significance of the splicing alternatives and for the importance of the α -helical region itself.

Heptad repeats are signatures of parallel chain α -helical coiled coils. Since there is a single light chain bound to each leg of the triskelion, we propose that the other member of the coiled coil is a segment of the heavy chain. There is indeed a region of the heavy chain sequence, appropriately positioned, that is a candidate for such a structure (5). A series of ten heptads will have a length of 105 Å, on the basis of a residue rise of 1.5 Å. The interaction of light and heavy chain therefore appears to extend along much of the proximal segment of a leg (length ~160 Å) (3, 4). The proposed arrangement (Fig. 6) is consistent with several observations. For example, electron micrographs of clathrin trimers with bound polyclonal antibodies to light chains show antibody decoration along the entire proximal segment (8); hydrodynamic (6) and electron microscopic (8) measurements on purified light chains suggest a highly extended molecule. Circular dichroism measurements indicate about 30 percent α -helix in free light chains (8, 12), with a noncooperative melting transition in guanidinium hydrochloride (8). These observations are consistent with the presence of an extended α -helix rather than a compactly folded, globular structure. Most of the light chain lysines are in the heptad sequences, and Ungewickell has observed their efficient labeling with Bolton-Hunter reagent only when light chains are dissociated from the intact clathrin trimers (8).

The heptad region lies in the middle of the light chain sequence, dividing it into three domains. We do not yet have evidence bearing on the function of the outer two, but we suggest that they may be involved in interactions with proteins other than clathrin—for example, the uncoating adenosine triphosphatase (9). The differences between the primary structures of class A and class B light chains, and further modulation of these

Fig. 4. Northern blot analysis of rat tissues. Poly(A)⁺ RNA (~1 μ g) from rat brain (B), heart (H), kidney (K), liver (L), lung (LU), and testis (T) were hybridized (31) with nick-translated restriction fragments spanning nucleotide positions 225 to 620, and positions 416 to 784 for LCA and LCB, respectively. Nonbrain LCA mRNA's have similar size and are slightly smaller than brain LCA mRNA (~1.3 kb). All LCB mRNA have similar sizes (~1.2 kb). Size markers were 18S and 28S RNA. A faint band of ~3 kb (arrowhead) is also detected with both light chain probes. It appears that the relative abundance of LCA mRNA is higher in liver, while that of LCB mRNA is higher in heart and liver. β -actin [probed with pAc18.1 (32)] of low specific activity was used as internal reference.

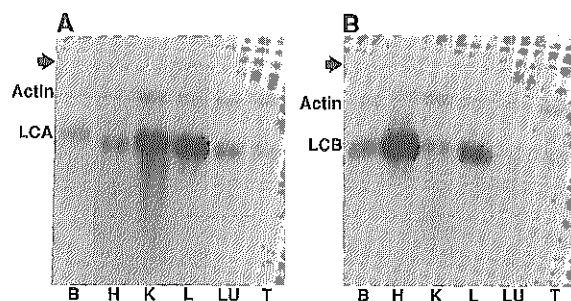
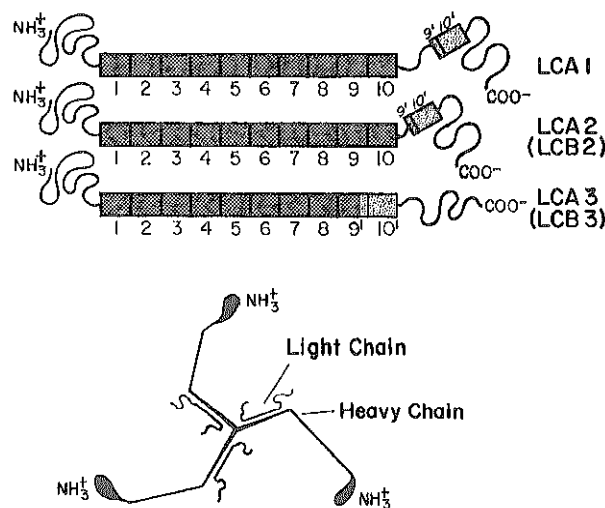


Fig. 5. Periodic features in the amino acid sequence of clathrin light chains LCA and LCB provide evidence for a unit made of ten seven-residue sections or heptads—a pattern characteristic of an α -helical coiled coil (see text). Positions a and d, containing primarily hydrophobic or neutral residues, lie along the interface of the coiled α helices. Negatively charged amino acids are totally excluded from these positions but lysine and arginine can occasionally appear at position a (26). The skip residue tryptophan appears in both chains at a homologous position between the fifth and sixth heptad. The sequences of primed heptads refer to those generated by the amino acid deletions in LCA3 and LCB3.

LCA							LCB								
	a	b	c	d	e	f	g		a	b	c	d	e	f	g
1.	V	D	R	L	Q	S	E ⁻	1.	A	D	R	L	T	Q	E ⁻
2.	P	E ⁻	S	I	R ⁺	K ⁺	W	2.	P	E ⁻	S	I	R ⁺	K ⁺	W
3.	R ⁺	E ⁻	E ⁻	Q	T	E ⁻	R ⁺	3.	R ⁺	E ⁻	E ⁻	Q	K ⁺	K ⁺	R ⁺
4.	L	E ⁻	A	L	D ⁻	A	N	4.	L	Q	E ⁻	L	D ⁻	A	A
5.	S	R ⁺	K ⁺	Q	E ⁻	A	E ⁻ W	5.	S	K ⁺	V	T	E ⁻	Q	E ⁻ W
6.	K ⁺	E ⁻	K ⁺	A	V	K ⁺	E ⁻	6.	R ⁺	E ⁻	K ⁺	A	K ⁺	K ⁺	D ⁻
7.	L	E ⁻	E ⁻	W	Y	A	R ⁺	7.	L	E ⁻	E ⁻	W	N	Q	R ⁺
8.	Q	D ⁻	E ⁻	Q	L	Q	K ⁺	8.	Q	S	E ⁻	Q	V	E ⁻	K ⁺
9.	T	K ⁺	A	S	N	R ⁺	V	9.	N	K ⁺	I	N	N	R ⁺	I
9'.	T	K ⁺	A	S	N	R ⁺	A	9'.	N	K ⁺	I	N	N	R ⁺	A
10.	A	D ⁻	E ⁻	A	F	Y	K ⁺	10.	A	D ⁻	K ⁺	A	F	Y	Q
10'.	A	E ⁻	E ⁻	A	F	V	N	10'.	S	E ⁻	E ⁻	A	F	V	K ⁺

Fig. 6. Model for the structure of clathrin light chains. The drawing shows features of clathrin light chains indicated by the derived amino acid sequence. Representations are shown of the longest chain LCA1, its variants LCA2 and LCA3, and the related LCB2 and LCB3. The heptad region that lies in the middle of the sequence divides it into three domains. The segment at the amino terminus of our present sequence lacks positively charged residues and contains 17 and 16 negatively charged amino acids for LCA and LCB, respectively. A splicing deletion, occurring near the end of the ten-heptad region in LCA3 and LCB3, joins part of heptad 9 to the remaining part of a "new" heptad 9' and a complete "new" heptad 10'. The pattern of repeated heptads is characteristic of structures bearing α -helical coiled coils (26). A single light chain is bound to the proximal segment of each leg of the clathrin triskelion (6-8), and we propose that a coiled coil structure mediates this interaction. The overall polarity of the heavy chain is shown (16) (carboxyl terminus at the center). We do not at present have evidence concerning light chain polarity.



differences by the observed deletions, may be important signals for the way in which such other proteins participate in the cycle of clathrin assembly and disassembly. Variation of particular light chain types from tissue to tissue suggests that these signals are

related to functional differences in different types of cells.

Note added in proof: While this paper was in press, Jackson *et al.* (26) have published a similar study of bovine clathrin light chains LCA and LCB. Their sequences correspond

to LCA1, LCA3, LCB2, and LCB3 that we have obtained from rat. If we align corresponding chains, we observe identity at more than 93 percent of the amino acid positions and conservation of the tissue-specific splicing sites. However, from our analyses of the amino acid sequences we come to different conclusions about secondary structure. We find that the repeat of ten heptad motifs, typical of α -helical coiled coils, constitutes a key element in all classes and types of light chains regardless of the species studied. In contrast, Jackson *et al.* have not noticed the heptad elements. Rather, they propose homology of intermediate filaments and clathrin light chains. They suggest that the similarity is highest between amino acid residues 107 and 236 of the bovine light chains (15 percent). In our opinion, the homology is not extensive enough to allow the interpretation that the proteins derive from a common ancestor gene. We propose here that the central segment containing the heptad motifs mediate the binding of light chains to clathrin heavy chain. In a publication in the same issue of *Nature*, Brodsky *et al.* [figure 3 in (27)], using monoclonal antibodies to light chains to block interactions of light chains and heavy chains, conclude that residues 93 to 157 mediate the interaction. This conclusion is precisely consistent with our assignment.

REFERENCES AND NOTES

1. J. L. Goldstein, R. G. W. Anderson, M. S. Brown, *Nature (London)* **279**, 679 (1979); B. M. F. Pearce, *Trends Biochem. Sci.* **5**, 131 (1980); S. C. Harrison and T. Kirchhausen, *Cell* **33**, 650 (1983).
2. B. M. F. Pearce, *J. Mol. Biol.* **97**, 93 (1975); R. A. Crowther, J. T. Finch, B. M. F. Pearce, *ibid.* **103**, 785 (1976).
3. T. Kirchhausen and S. C. Harrison, *Cell* **23**, 755 (1981).
4. E. Ungewickell and D. Branton, *Nature (London)* **289**, 420 (1981).
5. T. Kirchhausen *et al.*, in preparation.
6. T. Kirchhausen, S. C. Harrison, P. Parham, F. Brodsky, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2481 (1983).
7. E. Ungewickell, E. Unanue, D. Branton, *Cold Spring Harbor Symp. Quant. Biol.* **57**, 723 (1982).
8. E. Ungewickell, *EMBO J.* **2**, 1401 (1983).
9. S. L. Schmid, W. A. Braell, D. M. Schlossmann, J. E. Rothman, *Nature (London)* **311**, 228 (1984); J. E. Rothman and S. L. Schmid, *Cell* **46**, 5 (1986).
10. B. M. F. Pearce, *J. Mol. Biol.* **126**, 803 (1978); F. M. Brodsky and P. Parham, *ibid.* **167**, 197 (1983).
11. N. Holmes, J. S. Biermann, F. M. Brodsky, D. Bharucha, P. Parham, *EMBO J.* **3**, 1621 (1984).
12. F. K. Winkler and K. K. Stanley, *ibid.* **2**, 1393 (1983).
13. M. J. Geisow and R. D. Burgoyne, *Nature (London)* **301**, 432 (1983); C. E. Creutz and J. R. Harrison, *ibid.* **308**, 208 (1984).
14. F. M. Brodsky, *J. Cell Biol.* **101**, 2047 (1985).
15. M. P. Lisanti *et al.*, *Eur. J. Biochem.* **125**, 463 (1982).
16. T. Kirchhausen and S. C. Harrison, *J. Cell Biol.* **99**, 1725 (1984).
17. M. W. Hunkapiller, E. Lujan, F. Ostrander, L. E. Hood, *Methods Enzymol.* **91**, 227 (1983); A. N. Glazer, R. J. DeLange, D. S. Sigman, *Laboratory Techniques in Biochemistry and Molecular Biology*, T. S. Work and E. Work, Eds. (Elsevier, New York, 1976), vol. 4, p. 78; K. S. Huang *et al.*, *Cell* **46**, 191 (1986).
18. I. Moechetti, R. Einstein, J. Brosius, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7221 (1986).
19. A. Ullrich, C. H. Berman, T. J. Dull, A. Gray, J. M. Lee, *EMBO J.* **3**, 361 (1984); T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
20. P. Scarmato, J. Brosius, T. Smith, T. Kirchhausen, in preparation.
21. J. Keen, personal communication.
22. D. J. Lipman and W. R. Pearson, *Science* **227**, 1435 (1985).
23. G. Payne and R. Schekman, *ibid.* **230**, 1009 (1985).
24. S. E. Schwarzbaur, J. W. Tamkun, I. R. Lemischka, R. O. Hynes, *Cell* **35**, 421 (1983); T. A. Cooper and C. P. Ordahl, *J. Biol. Chem.* **260**, 11140 (1985); K. E. M. Hastings, E. A. Bucher, C. P. Emerson, Jr., *ibid.*, p. 13699; B. A. Murray *et al.*, *J. Cell Biol.* **103**, 1431 (1986).
25. A. D. McLachlan and J. Karn, *Nature (London)* **299**, 226 (1982).
26. A. P. Jackson, H. F. Seow, N. Holmes, K. Drickamer, P. Parham, *ibid.* **326**, 154 (1987).
27. F. M. Brodsky *et al.*, *ibid.*, p. 203.
28. R. F. Grantham, *Nucleic Acids Res.* **9**, 1 (1981).
29. L. J. McBride and M. H. Caruthers, *Tetrahedron Lett.* **24**, 245 (1983); R. B. Pepinsky *et al.*, *J. Biol. Chem.* **261**, 4239 (1986).
30. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
31. B. P. Wallner *et al.*, *Nature (London)* **320**, 77 (1986).
32. U. Nudel *et al.*, *Nucleic Acids Res.* **11**, 1759 (1983).
33. We thank M. Brownstein, NIMH, for the rat brain plasmid library; D. Elstein, S. Frucht, D. Harvey, T. Smith, E. Tobias, and students of the Cold Spring Harbor Laboratory Advanced Molecular Cloning course for help at various stages of cloning; R. Tizard for making available his modifications to the DNA sequencing protocol; W. Gilbert for computer time; L. Wachter for DNA synthesis; and C. Beepot and L. M. Oliphant for secretarial assistance. Supported in part by NIH grant RO1 GM 36548-01 (T.K.), AHA grant-in-aid 86119 (T.K.), NSF grant DMB 85-02920 (S.C.H.) and NIMH grant MH 38819 (J.B.) and by a contribution from the Foundation for Medical Research (T.K.), an Established Investigatorship of the American Heart Association (T.K.), and an Irma T. Hirsch career scientist award (J.B.).

19 December 1986; accepted 16 March 1987

