

Functional evidence for the identification of an *Arabidopsis* clathrin light chain polypeptide

Urte Scheele^a, Susanne E.H. Holstein^{b,*}

^aDepartment of Cell Biology, Center of Anatomy, Hannover Medical School, D-30125 Hannover, Germany

^bCell Biology, Heidelberg Institute of Plant Sciences, D-69120 Heidelberg, Germany

Received 19 December 2001; revised 11 January 2002; accepted 11 January 2002

First published online 21 February 2002

Edited by Gunnar von Heijne

Abstract Clathrin light chains (CLCs) are regulatory subunits of clathrin triskelia. Based on homology searches in *Arabidopsis thaliana* data bases we have identified three putative CLC clones, and have focused on the one with the highest homology to mammalian CLC sequences. Analysis of its sequence has revealed coiled-coil structures within a region that corresponds to the clathrin heavy chain-binding site. In addition there is a stretch of acidic amino acids, which is required for the regulatory function of CLC in clathrin assembly. This putative plant CLC ortholog, expressed in bacteria as a glutathione-S-transferase- and myc-tagged fusion protein, was shown to bind to CLC-free recombinantly expressed mammalian clathrin hubs. In contrast, purified native mammalian triskelia with endogenous CLC did not bind the recombinant putative plant CLC. Based on the conserved sequences between the three *Arabidopsis* candidates it appears that plants, unlike mammals, may have more than two CLCs. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Heterologous binding experiment; Recombinant fusion protein; Clathrin light chain; Clathrin hub; *Arabidopsis thaliana*

1. Introduction

Clathrin-coated vesicles (CCV) are vehicles for receptor-mediated transport processes in all eukaryotic cells. In mammalian cells, CCV are responsible for the endocytotic uptake of nutrients and for receptor down-regulation at the plasma membrane [1]. CCV in mammalian, yeast and plant cells are also involved in the transport of soluble acid hydrolases from the *trans*-Golgi network via a pre-vacuolar/lysosomal compartment to their final destination, the vacuole or lysosome [2–4].

The coat of CCV has two major components, clathrin and adaptor protein (AP) complexes, which constitute the outer and the middle layers surrounding the vesicle membrane, re-

spectively [5,6]. All AP complexes are heterotetrameric and function at different sites within the mammalian cell [7]. The units of the clathrin layer are three-legged structures, termed triskelia, which are built of three clathrin heavy chain (CHC) polypeptides (192 kDa) and three clathrin light chain (CLC) polypeptides [8,9]. All eukaryotes have only one CHC gene, with the exception of humans, who in addition to the ubiquitously expressed CHC17 polypeptide have another isoform of CHC in skeletal muscle cells (CHC22), and which does not bind significantly to CLC [10,11]. Each CHC polypeptide is divided into four regions, namely the carboxy-terminal domain (residues 1550–1675) comprising the trimerization site (residues 1550–1600), the proximal domain (residues 1074–1552), encompassing the CLC-binding site, the distal region (residues 495–1073) and the amino-terminal region (residues 1–494) [12]. The truncated versions of CHCs, the so-called hubs, comprise the carboxy-terminal and proximal domains (residues 1675–1074), and are therefore responsible for three key features of clathrin function: trimerization, CLC-binding and clathrin assembly [13].

In mammals, different genes encode two types of CLC, which are termed CLCa and CLCb and which share 60% similarity on amino acid level. Both types of CLC are expressed in all tissues but at different relative levels [14] and are randomly distributed on the CHC polypeptides [15]. In addition, CLCa and CLCb have isoforms created by alternative splicing in mammalian neuronal cells and contain extra brain-specific inserts [16,17]. On the other hand, yeast and *Drosophila* have only one CLC gene [18,19]. All CLCs investigated so far have molecular masses between 23 and 26 kDa. Similar to the CHC, the CLC can also be described as a linear series of functional domains [20]. Both light chains are able to bind calcium but only CLCb can be phosphorylated by a type II casein kinase [21,22]. Clathrin lattice formation occurs spontaneously in vitro at low pH in the presence of a low calcium concentration, or is intracellularly triggered by AP complexes at physiological pH. In support of their regulatory function to prevent premature association of triskelia in vitro, CHC free of CLC assembles easily in the absence of AP complexes or calcium [13,23]. It was therefore hypothesized that CHC self-assembly is regulated by three amino acid residues, located within the amino-terminal end of the CLC. These amino acids are supposed to control the formation of salt bridges which in turn induce CHC assembly at physiological pH [24].

Although a plant CHC polypeptide, which showed 75% similarity to mammalian CHC17, was unequivocally identified in soybean [25], the situation concerning plant CLC homologs

*Corresponding author. Fax: (49)-6221-546404.

E-mail address: susanne_h.holstein@urz.uni-heidelberg.de (S.E.H. Holstein).

Abbreviations: CCV, clathrin-coated vesicles; CLC, clathrin light chain polypeptide; CHC, clathrin heavy chain polypeptide; GST, glutathione-S-transferase; AP, adaptor protein complex; EST, expressed sequence tag; IPTG, isopropyl-D-thiogalactopyranoside; PBS, phosphate-buffered saline; TCA, trichloroacetic acid

is less clear. A number of reports have described putative CLC candidates from a number of different plants [26–30], but it has never been conclusively demonstrated that these proteins associate with CHC.

In this study we settle this issue by demonstrating that a putative plant CLC identified by homology searches of the *Arabidopsis* genome interacts specifically with mammalian CHC in a heterologous binding experiment using recombinant triskelion hubs.

2. Materials and methods

2.1. Isolation of the cDNA clone encoding *Arabidopsis* CLC

The expressed sequence tag (EST) clone 284F4T7 obtained from data base screening (*Arabidopsis* Stock Centre, OH, USA) was fully sequenced in both directions using a commercial facility (MWG AG Biotech, Ebersberg, Germany). The missing 211 bp of the 5' end of the cDNA were obtained by PCR with the forward primer 5'-ATGTCTGCCTTTAAGGACGATTCTCC-3' and the reverse primer 5'-CTTCTCCTTCTCCTCAAGTTGAATTGC-3' and mass excised plasmids of the *Arabidopsis* cDNA GenBank CD4-7 as template (ABRC DNA Stock center, OH, USA). To obtain the full-length CLC cDNA clone the PCR fragment in pGEM-T EASY vector (Gibco BRL Life Technologies, Karlsruhe, Germany) (cut by *SacII*, *BamHI*) and the cDNA from the EST (cut by *BamHI*, *EcoRI*) were ligated into Bluescript vector SK, which was opened before with *SacII* and *EcoRI*. Sequences were analyzed using the MacVector Program (Oxford Molecular Group, Accelrys, Cambridge, UK), data base searches were performed using the BLAST algorithm [31] and alignments between two sequences were performed using LALIGN program version 2.0 [32]. Standard cloning procedures were used [33].

2.2. Cloning of the glutathione-S-transferase (GST)–myc–CLC fusion protein

To construct the GST–myc fusion protein the pGEX4T-3 vector (Amersham-Pharmacia Biotech, Freiburg, Germany) was opened with *EcoRI* and *XhoI*. The CLC cDNA sequence was excised from the Bluescript vector by *SacII* and *XhoI*. The c-myc cDNA sequence was cut out with *EcoRI* and *SacII* and the two fragments were ligated into the vector.

2.3. Cloning of (His)₆×-hub from human CHC sequence

The human CHC sequence in the Bluescript vector (KIAA 0034) was digested with *MspI* to obtain the hub region. The insert was cloned into the *EcoRV* site of the PET32a vector (Novagen) as described in [34].

2.4. Purification of (His)₆×-hub fusion protein

Induction occurred at room temperature (RT) for 3 h after addition of 0.5 mM isopropyl- β -thiogalactopyranoside (IPTG). The pellet of 1 l transformed BL21 cells was resuspended in 10 ml lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole, pH 8.0 plus 1% Triton X-100 and one tablet of protease inhibitor cocktail) (Boehringer Complete, Roche Biochemicals, Mannheim, Germany). The resuspended bacteria were sonicated for 5×10 s (Branson sonifier) and subsequently centrifuged in a Beckman 70Ti rotor at 40 000×g for 15 min at 4°C. The resulting supernatant was incubated with 1.5–2 ml packed Ni-NTA agarose beads (Qiagen, Hilden, Germany) that were equilibrated with lysis buffer for 1 h at 4°C on a rotator. After incubation the beads were washed one time with lysis buffer, three times with washing buffer (50 mM NaH₂PO₄, 300 mM NaCl, 20 mM imidazole, pH 8.0) and pelleted through a 10% sucrose/washing buffer cushion after the second washing step to remove contaminations. The bound fusion protein was then eluted with elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 250 mM imidazole, pH 8.0).

2.5. Purification of GST–myc–CLC

0.5 l of Luria Bertani medium was inoculated in a 1:10 dilution from an overnight culture, and cells were grown until an OD_{600 nm} of 0.7 was obtained. Induction occurred for 2 h at RT with a final concentration of 2 mM IPTG. Bacteria were obtained by centrifugation at 4000×g at 4°C for 15 min and frozen at –80°C after washing once with phosphate-buffered saline (PBS). Purification of the GST

fusion protein was performed as described in [35]. Bacteria obtained from 0.5 l cultures were resuspended in 20 ml of buffer A (50 mM Tris, pH 7.5, 100 mM KCl, 5 mM EDTA) plus protease inhibitors (2 mM leupeptin, 0.7 μ M pepstatin, 2 μ g/ml aprotinin, 0.15 mM phenylmethylsulfonyl fluoride) and disrupted by sonication (six bursts of 30 s with 90% power; Bandelin Sonoplus GM70). Centrifugation at 100 000×g for 20 min removed the cell debris. The supernatants were then incubated with 500 μ l packed GSH-Sepharose beads on a rotator at 4°C for 30 min, sequentially washed with 30 ml of buffer A and with 10 ml of buffer B (50 mM Tris–Cl, pH 7.5, 1 mM KCl, 5 mM EDTA). GST fusion proteins were eluted with 5 mM glutathione. Fractions 1–3 were combined and used for the binding studies.

2.6. Purification of pig brain triskelia

The purification procedure was according to [36].

2.7. Superose 12 gel filtration/FPLC

The GST–myc–CLC fusion protein was transferred into buffer C (100 mM Tris, pH 7.5, 5 mM EDTA) by gel filtration using a PD-10 column (Amersham-Pharmacia Biotech, Freiburg, Germany) and subsequently reduced to a final volume of 850 μ l using a Centricon device (Millipore, Bedford, MA, USA). An aliquot of 250 μ l was diluted 1:1 with buffer C and applied directly to the S12 column. The second aliquot was incubated with 250 μ l of (His)₆×-hubs (250 μ g/ml) for 30 min on ice, spun for 1 min at 10 000×g to remove aggregates and then applied to the S12 column. The third aliquot was incubated with 84 μ l triskelia (1.17 mg/ml) in a total volume of 350 μ l of buffer C for 40 min on ice and also applied to the S12 column. Gel filtration was performed in buffer C with a flow rate of 0.7 ml/min and the fraction size was 0.5 ml. From each gel filtration fraction an aliquot of 200 μ l was precipitated with a final concentration of 10% trichloroacetic acid (TCA), resuspended in 30 μ l sample buffer and subjected to SDS-PAGE.

2.8. SDS-PAGE and Western blotting

SDS gradient gels (10–19%) were prepared as previously described [37]. Proteins were blotted onto nitrocellulose [38] and visualized with the Supersignal West Pico ECL kit (Pierce, Rockford, IL, USA).

2.9. Antibodies used for immunoblots

The ascites containing the monoclonal E910 antibody against c-myc were diluted 1:50 in PBS/5% (w/v) skim milk. All other antibodies were used in a 1:1000 dilution, respectively. The polyclonal anti-GST antibody (Amersham-Pharmacia Biotech, Freiburg, Germany), the polyclonal anti-histidine antibody (BioScience, Göttingen, Germany), the monoclonal anti-CHC (BD Transduction Laboratories, Franklin Lakes, USA) and the monoclonal antibody directed against mammalian CLC, CL57.4 (Dr. R. Jahn, Göttingen, Germany).

3. Results and discussion

Earlier attempts to identify CLCs in plants were solely based upon experiments describing non-unique features of mammalian CLC like calcium-binding, heat resistance and solubility after precipitation with TCA. Based on these criteria polypeptides of 30–60 kDa from carrot, soybean, zucchini and pea cotyledon CCV have been proposed as putative plant light chain polypeptides, reviewed in [39,40].

However, the most important criterion, to be met by light chain orthologs, is their ability to bind to a specific site on CHC. With this in mind, we searched the *Arabidopsis* data bases for sequences which are related to bona fide light chains from non-plant species. We came up with three sequences (AF002109, AF049236, AAD20919) which share similarities of 75% (Fig. 1, Table 1). One of these clones is AF002109, which is identical to the EST clone 28F4T7 and the full-length BAC clone T28M21.22, and they in turn are identical to the At2g40060 gene. Sequence comparison of AF002109 to mammalian CLCa and CLCb as well as to the yeast Clc1p reveals the highest degree of identity to CLCa and the lowest to yeast

CLCa	MAELDPFGVPAGGHALNGVAGEEDPAAAFIAQ—Q—ESETAGIENDEAFAILDGGAPGSQPHGEPPGIP—DAVDGVINGDY
CLCb	—MADDFGFTSS—SESGAPEAAEEDPAAAFIAQ—Q—ESETAGIENDEGFAGAPGSQGLAQPGPASGAS—EDMGATVNGDV
AF002109	MSAEEDDSFVIL—NDDASESVVSCSF—DAIDGFSAPDGLQVEDSVDDFAAPSSDYCA—YNSGCG
AF049236	MAATEETSRLITS—LGFQSQRFDS—FSNFDLSQPEKESDLPQDSSRPET—QSPPS
AAD20919	MAITFDGDFPAQ—THSPSEHEDFGGYDNFSEAOQPPTQHQSGGFSFNGDPASNGYGFAGSSPNHDFSSPFESSVNDANGNGGS
Yeast CLC	—MSEKFPPLDQNDITFNDKQDDITFLKRAEATLGDEFKTEQDDITLFEASPAKDDDETRDFEEQFPDINSANGAVSSDQNGSAIVS
CLCa	YQESNGPTDSYAATISQVDRLO—SEPESTRKWREEQTERLEALDANSRKQEAEMKEKATKELDEWVARQDEQLQRTKANIRAE—
CLCb	FOEANGPADGYAATAQADRLT—QEPESIRKWREEQKRLQELDAASKVMEQEWREKAKKOLEWNRQSEQVEKNKINIRASE—
AF002109	IFGSGNDHDPILPPSEMES—DEGFALREWRNNAIQLEKEKEKEKELKQITEADQYKEEFHKTEVICENNKANREKEKL
AF049236	—INSFDDINDSILPPPSAMEK—EEGFALREWRRI NALRLKEKEKEKEKEMVOITLFAAQYKAEFYSKRNVTIENNKTNREKE—
AAD20919	GGDAIFASDGFILPDINER—EEGFORREWRRI NITLHEKEKEKEKERNQITLFAEDFKKAFYEKROKTEINKTIDNREKEKL
Yeast CLC	SGNDNGEADDSTFEFANQSTESVKEDESEVDQNKQRRRAVEIHEKDLKDEELKKELQDEATKHIDDFYSYNKKKEQOLEDAAKEAEA
CLCa	—EAFVNDIESSPGTEWEKVARLCDENP—KSSQKAK—DVSRRMSVLISLQAPLVH—
CLCb	—EAFVKESKEETPGTEWEKVAQLCDFNP—KSSQCK—DVSRLRSVLMSLQKPLSR—
AF002109	YLENQEKFYAESK—NYWKATAELVPEVPTIEKRAGKKEQDQKPTVSVIQGPKPKGKPTDLIRMQILLVKLHNPPSHKLITSQPPSE
AF049236	—KNQEKFYAEDK—NNWKATAELTPREVPTIENRGNK—KTATITVIOGPKPKGKPTDLIRMQVILVKLHNPPSHKPLSPSPS
AAD20919	YWANQEKFKHEVDK—HYWKATAELTPREVPTIEKRGKQDPD—KKPSVNVIOGPKPKGKPTDLIRMQITFLKLNPPPHMPPPPAKD
Yeast CLC	FLKKRDEFFQDNT—TWDRALQLINQDD—ADLTIGGR—DRSKLKEITLLRLKGNAPAGA—
CLCa	-----213
CLCb	-----210
AF002109	E-----AAAPPKNVPEIKPIEAVTAA-----258
AF049236	-----GADPNVSVSEQVITVEKL-----233
AAD20919	AKDGKDAKGDAGTKGDKGDAKGGKDAKDLKDGKPADPKVTEEKRPSPAKDASVETAKPDAASGEGEKPVAVITEAGTKAE-----338
Yeast CLC	-----233

Fig. 1. Comparison of CLC polypeptide sequences. Alignment of CLC sequences CLCb (accession no. X04853) and CLCa (accession no. X04851) from bovine lymphocytes, yeast (accession no. X52272) and *Arabidopsis*. The CHC-binding region is boxed over the sequences. The highly conserved stretch of 13 amino acids is marked by a line. Identical amino acids between the three plant sequences are in bold letters and the acidic negative residues in the amino-terminal end are underlined.

Clc1p, although the similarities to CLCa and yeast Clc1p are equally high. The yeast Clc1p has been reported by Silveira and colleagues [18] to reveal 18% identity to mammalian CLCs.

Therefore, it has become obvious that the plant CLC candidate is far more identical to mammalian CLCs than the yeast Clc1p. Furthermore, in the CHC-binding region the identity of plant CLC (AF002109) is more similar to mammalian CLCs than to yeast Clc1p. The identity of AF002109 is even higher towards the other two plant CLC candidates (Table 1). Based on these high similarities to non-plant CLCs the clone AF002109 was therefore chosen as being representative of putative plant CLCs. To test whether clone AF002109 is in fact an *Arabidopsis* CLC homolog it was used in binding experiments with CHC. Therefore, we set out to clone its cDNA and to express it in bacteria for binding studies. Since the EST clone lacked 211 bp of the coding sequence from the 5' end (Fig. 2A) we amplified a 323 bp fragment from an *Arabidopsis* cDNA bank which contained the missing 5' end sequence and overlapped with the EST fragment (Fig. 2A,B). The two overlapping fragments were digested with *Bam*HI and the resulting 211 bp amplified frag-

ment was joined with the 620 bp fragment from the EST clone to yield a full-length 774 bp CLC cDNA (Fig. 2B,C). To aid in the purification and the detection of the recombinant protein, the CLC was expressed as a fusion protein with GST, and with the c-myc epitope, thereby giving rise to a predicted molecular weight of 59.4 kDa. On SDS-PAGE, the purified recombinant *Arabidopsis* GST-myc-tagged CLC indeed behaved like a 59 kDa protein (Fig. 2C,D). Antibodies directed against either tag specifically recognize the fusion protein (Fig. 3A).

To determine whether the recombinant plant CLC can associate with clathrin we expressed human recombinant heavy chain hubs that correspond to the human CHC amino acid sequence (positions 1073–1675). The hubs were expressed with an amino-terminal (His)₆×-tag (Fig. 3B). We first confirmed earlier reports [13] that our hub construct readily associates with mammalian light chains (data not shown). To assess binding of plant light chains to these hubs we utilized gel filtration on a Superose 12 column. This method efficiently separates free light chains from clathrin hubs (Fig. 3C). In the absence of hubs the recombinant GST-myc-CLC eluted in fractions 12–15 whereas hubs alone eluted between frac-

Table 1
Identities/similarities of *Arabidopsis* and non-plant CLC amino acid sequences

	Over entire length				In the CHC-binding domain					
	CLCa (20–213)	CLCb (3–210)	Yeast Clc1p (3–232)	AF492356 (1–233)	AAD20919 (1–274)	CLCa (95–158)	CLCb (92–155)	Yeast Clc1p (109–174)	AF492356 (69–132)	AAD20919 (102–164)
AF002109 (1–227)	31.1/64.2									
(7–227)		28.5/55								
(7–236)			23.2/64.2							
(12–255)				55/75						
(1–258)					52/75					
(81–144)						20/58.5	24.6/59.4	18.2/63.6	67.2/87.5	56.3/82.8

The Needleman–Wunsch global alignment was used for the alignment over the entire length of the sequences and the LALIGN program for the best local alignment of the CHC-binding regions. The positions of the aligned regions are given in brackets below the organisms and the identities/similarities are given in percent.

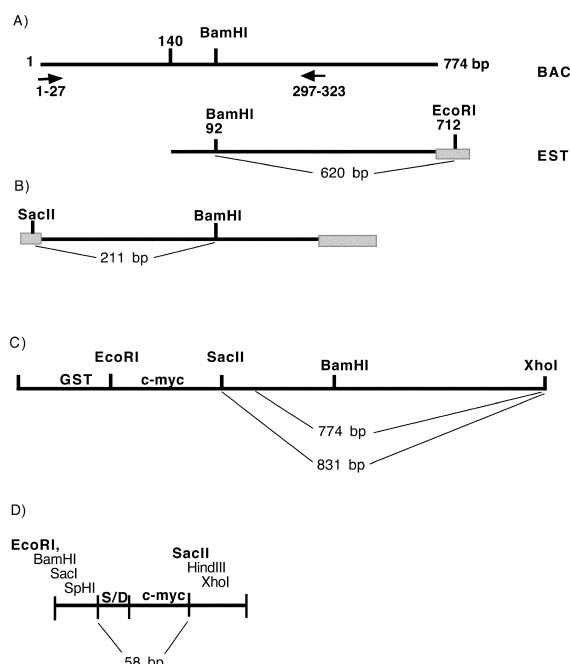


Fig. 2. Cloning strategy of *Arabidopsis* CLC cDNA. A: Upper panel: BAC clone T28M21 containing the genomic sequence of *Arabidopsis* CLC after removing untranslated regions. Arrows indicate the positions of the forward and the reverse PCR primers, respectively. Lower panel: EST clone 28F4T7 in Zip-Lox vector. B: PCR fragment of total 323 bp. C: Full-length cDNA sequence of *Arabidopsis* CLC fused to a myc- and a GST-tag. The sequence of 831 bp contains restriction enzyme sites in addition to the coding region of 774 bp. D: c-myc-tag sequence, fused to the Shine-Dalgarno sequence (S/D) and the given restriction enzyme sites. Dashed boxes indicate the respective vector.

tions 4 and 11. After incubation of hubs with GST-myc-CLC the recombinant light chains co-eluted with the hubs in fractions 5 and 6. In a control experiment, purified intact clathrin triskelia from pig brain with endogeneous light chains were incubated with the GST-myc-CLC fusion protein. Similar to recombinant clathrin hubs, intact clathrin triskelia eluted between fractions 4 and 8, as shown by immunoblotting with antibodies directed against the mammalian heavy and light chains, respectively. In this experiment the GST-myc-CLC were detected only in fractions 10–14, which corresponds to the elution position of the light chains. This result was expected, because in intact native triskelia the light chain-binding sites on the proximal domain of the heavy chain are occupied by endogeneous light chains. These findings clearly demonstrate that binding of the plant CLC does occur to a specific site on the CHC hub. Whether the yeast Clc1p is also able to bind to mammalian hubs remains uncertain. However, it is known that Clc1p is able to trimerize the yeast CHC [41].

The alignment of CLC protein sequences (Fig. 1) revealed acidic residues at the extreme amino-terminal end of CLCb (AEED), CLCa (GEED) and of yeast (KDDD) and also of *Aplysia* CLC (GEVD). This motif is described to inhibit CHC assembly via regulating their high affinity salt bridges [24]. The same motif of three acidic residues can also be observed within a corresponding area in the AF002109 CLC sequence (F5EDD7), and most probably also in the two other clones (T4EET7 and D5DGD8) suggesting a similar or related regulatory function for plant CLCs. Another remarkable feature of non-plant CLCs is their shift to higher molecular weights

(32–36 kDa) in SDS-PAGE than their deduced molecular masses (23–26 kDa). This feature is based on the high content of proline and glycine residues within the first 90 amino acids [42]. All three *Arabidopsis* CLC candidates contain an equally high content of these two amino acids. In contrast, yeast Clc1p as well as the plant CLCs reveal the lack of a mammalian type casein kinase II consensus recognition sequence in which two serine residues in CLCb at positions Ser11 and Ser13 become phosphorylated [22]. Moreover, yeast Clc1p and plant CLCs have in common the scattering of serine residues throughout their sequences. While it has been shown for yeast Clc1p that these serines become constitutively phosphorylated in vivo [43] such data are missing for the plant sequences. Whereas plant AF002109 and AF049236 have one cysteine residue each, the yeast Clc1p sequence completely lacks cysteine residues. In this respect, in terms of the number of cysteine residues they more match the non-neuronal CLCa and the shorter form of neuronal CLCa, but not in terms of the positions of these residues, which are located in their carboxy-terminal regions. An additional cysteine residue is found in both forms of CLCb and also in the larger form of neuronal CLCa [44]. Finally, the three plant CLCs do not contain a peptide stretch which is homologous to the brain-specific insert of mammalian CLCs. The yeast Clc1p instead harbors a stretch of 18 amino acids, which shows homology to the brain-specific inserts of both CLCa and

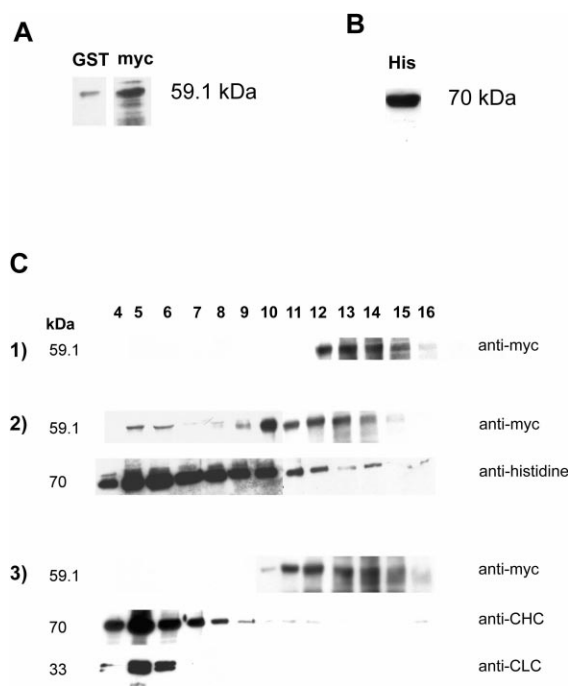


Fig. 3. Binding of GST-myc-CLC and bovine (His)₆×-hub polypeptides. A: Expression of the GST-myc-CLC of *Arabidopsis* as a 59 kDa fusion protein. Binding of either anti-myc antibody or anti-GST antibody identifies the CLC fusion protein. B: The 70 kDa (His)₆×-hub region is recognized by the anti-histidine antibody, specifically. C: *Arabidopsis* CLC-binding to bovine CHC. Superose 12 gel filtration fractions are given on top. The antibodies are given on the right border and the molecular weight of the respective polypeptide on the left border. (1) S12 gel filtration of GST-myc-CLC. (2) Binding of GST-myc-CLC to the (His)₆×-hub region of bovine CHC. (3) Incubation of GST-myc-CLC with native bovine triskelia. A–C: ECL blots.

CLCb. Surprisingly, a segment of 13 residues in the carboxy-terminal region is highly conserved between the three species, although its function is unclear (Fig. 1).

To verify that the putative plant CLC and the other two polypeptides have a structural organization similar to their homologs in animals, we analyzed their primary sequences.

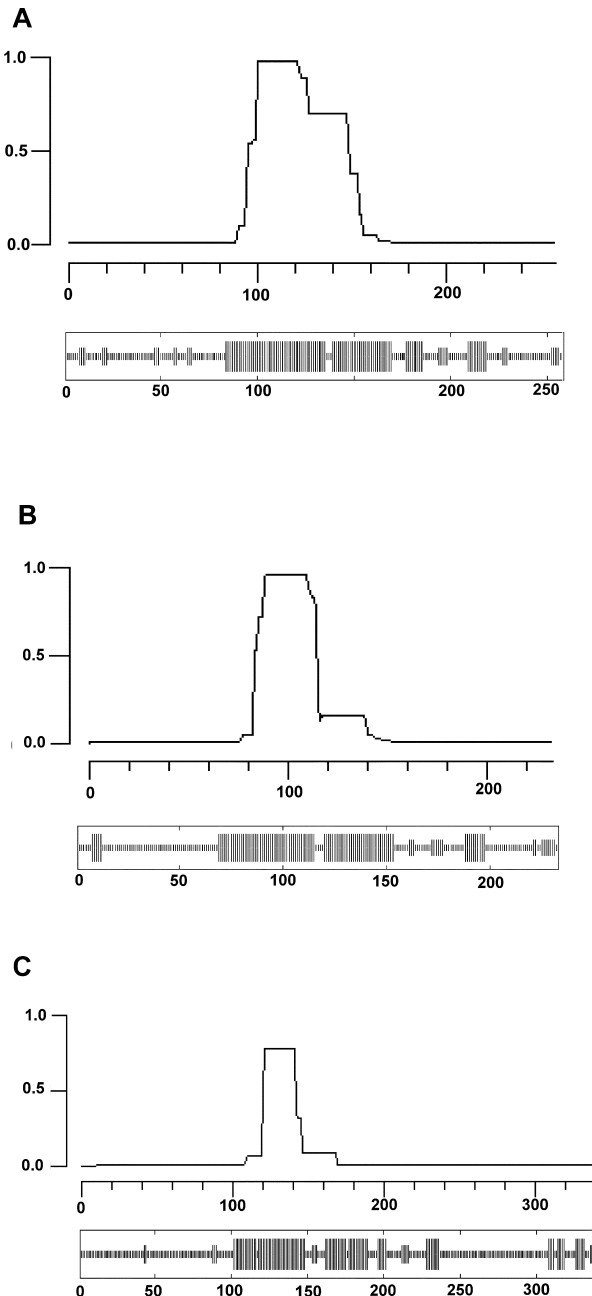


Fig. 4. Prediction of the secondary structure of *Arabidopsis* CLC sequences AF002109 (A), AF049236 (B) and AAD20919 (C). Upper panels: coiled-coil structure prediction using a window size of 21 amino acids and the MTK matrix in the Coil program [45]. Note that a high probability of coiled-coil structure corresponds to the central part of each polypeptide and correlates with the respective consensus secondary structure. Lower panels: full vertical lines indicate the positions of α -helices along the sequences, while half vertical lines indicate β -sheet formations, the smallest vertical lines indicate random coil formation. Amino acid residues in each panel are given on the abscissa. Secondary structure prediction was performed using the GOR4 program [47].

For this purpose we used an algorithm that predicts secondary structures and coiled-coil interactions [45], and which has been successfully used for the prediction of two coiled-coil regions of mammalian CLCs: one connecting the amino-terminal coiled-coil domain to the hsc70-binding site while the second coiled-coil domain correlates with the CHC-binding domain [46]. We found that a propensity for coiled-coils exists between the residues 101–148 in AF002109, 85–115 in AF049236 and 121–146 in AAD20919. It therefore extends throughout the α -helical region in the plant sequence that corresponds to the CHC-binding site (Fig. 4). Using the same parameters, the helical regions of the mammalian and the yeast Clc1p appear to be divided into two domains (data not shown), a feature also shared by all three plant CLCs.

Based on the high homologies between the three *Arabidopsis* CLC candidates, and the fact that they share the same type and location of protein interaction domain we had chosen the one harboring the most extensive coiled-coil region for the functional proof of CHC-binding. Our biochemical data strongly support the notion that the *Arabidopsis* gene At2g40060 does indeed encode a CLC polypeptide. This CLC gene is located on chromosome II and corresponds to a protein of 258 amino acids with a deduced molecular mass of 28.8 kDa which is slightly acidic with a pI of 4.9. The genes of the two other candidates are located on chromosome IV (AF049236), encoding a 26.5 kDa protein, and on chromosome 2 (AAD20919), encoding a 37.2 kDa protein.

In addition, data base searches have also revealed CLC orthologs in other plant species. So far, three clones from soybean and two clones each from *Medicago* and tomato can be identified, while rice, wheat and maize show only one clone, respectively. The clones from the dicotyledons all show the highest degree of homology to AF002109 (76–89%), and the clones from the monocotyledons are slightly more homologous to one of the other *Arabidopsis* clones (79%) compared with a homology to AF002109 of around 76%.

In summary, although we have only proved for one *Arabidopsis* polypeptide the function of a CLC, it is reasonable to assume that plants, unlike yeast and mammals, have probably more than two genes encoding CLC.

Acknowledgements: We are grateful to Drs. Klaus Weber and E. Ungewickell for valuable advice and want to thank Dr. D.G. Robinson for helpful comments on the manuscript. We thank Dr. R. Jahn for the gift of the antibody CL57.4. This work was supported by the Deutsche Forschungsgemeinschaft (TP A7, SFB 523) to S.E.H.H.

References

- [1] Marsh, M. and McMahon, H.T. (1999) *Science* 285, 215–220.
- [2] Bassham, D.C. and Raikhel, N.V. (2000) *Curr. Opin. Cell Biol.* 12, 491–495.
- [3] Odorizzi, G., Cowles, C.R. and Emr, S.D. (1998) *Trends Cell Biol.* 8, 282–288.
- [4] Schmid, S. (1997) *Annu. Rev. Biochem.* 66, 511–548.
- [5] Hirst, J. and Robinson, M.S. (1998) *Biochim. Biophys. Acta* 1404, 173–193.
- [6] Kirchhausen, T. (1999) *Annu. Rev. Cell Dev. Biol.* 15, 705–732.
- [7] Rohn, W.M., Rouille, Y., Waguri, S. and Hoflack, B. (2000) *J. Cell Sci.* 113, 2093–2101.
- [8] Kirchhausen, T. (2000) *Annu. Rev. Biochem.* 69, 699–727.
- [9] Ungewickell, E. and Branton, D. (1981) *Nature* 239, 420–422.
- [10] Brodsky, F.M. (1997) *Trends Cell Biol.* 7, 175–179.
- [11] Liu, S.-H., Towler, M.C., Chen, E., Chen, C.-Y., Song, W., Apodaca, G. and Brodsky, F.M. (2001) *EMBO J.* 20, 272–284.

- [12] Smith, C.J. and Pearse, B.M.F. (1999) *Trends Cell Biol.* 9, 335–338.
- [13] Liu, S.-H., Wong, M.L., Craik, C.S. and Brodsky, F.M. (1995) *Cell* 83, 257–267.
- [14] Acton, S.L. and Brodsky, F.M. (1990) *J. Cell Biol.* 111, 1419–1426.
- [15] Kirchhausen, T., Harrison, S.C., Parham, P. and Brodsky, F.M. (1983) *Proc. Natl. Acad. Sci. USA* 80, 2481–2485.
- [16] Jackson, A.P., Seow, H.-F., Holmes, N., Drickamer, K. and Parham, P. (1987) *Nature* 326, 154–159.
- [17] Stamm, S., Casper, D., Dinsmore, J., Kaufmann, C.A., Brosius, J. and Helfman, D.M. (1992) *Nucleic Acids Res.* 20, 5097–5103.
- [18] Silveira, L.A., Wong, D.H., Masiarz, F.R. and Schekman, R. (1990) *J. Cell Biol.* 111, 1437–1449.
- [19] Vasyukevich, K. and Bazinet, C. (1999) *DNA Cell Biol.* 18, 235–241.
- [20] Brodsky, F.M., Hill, B.L., Acton, S.L., Näthke, I., Wong, D.H., Ponnambalam, S. and Parham, P. (1991) *Trends Biochem. Sci.* 16, 208–213.
- [21] Näthke, I., Hill, B.L., Parham, P. and Brodsky, F.M. (1990) *J. Biol. Chem.* 265, 18621–18627.
- [22] Hill, B.L., Drickamer, K., Brodsky, F.M. and Parham, P. (1988) *J. Biol. Chem.* 263, 5499–5501.
- [23] Ungewickell, E. and Ungewickell, H. (1991) *J. Biol. Chem.* 266, 12710–12714.
- [24] Ybe, J.A., Greene, B., Liu, S.-H., Pley, U., Parham, P. and Brodsky, F.M. (1998) *EMBO J.* 17, 1297–1303.
- [25] Blackbourn, H.D. and Jackson, A.P. (1996) *J. Cell Sci.* 109, 777–787.
- [26] Balusek, K., Depta, H. and Robinson, D.G. (1988) *Protoplasma* 146, 174–176.
- [27] Cole, L., Coleman, J.O.D., Evans, D.E., Hawes, C.R. and Horsley, D. (1987) *Cell Rep.* 6, 227–230.
- [28] Coleman, J.O.D., Evans, D., Hawes, C., Horsley, D. and Cole, L. (1987) *J. Cell Sci.* 88, 35–45.
- [29] Lin, H.-B., Harley, S.M., Butler, J.M. and Beevers, L. (1992) *J. Cell Sci.* 103, 1127–1137.
- [30] Wiedenhoefft, R.E., Schmidt, G.W. and Palevitz, B.A. (1988) *Plant Physiol.* 86, 412–416.
- [31] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) *J. Mol. Biol.* 215, 403–410.
- [32] Huang, X. and Miller, W. (1991) *Adv. Appl. Math.* 12, 373–381.
- [33] Sambrook, J., Fritsch, E.F. and Maniatis, Z. (1989) *Molecular Cloning, A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [34] Scheele, U., Kalthoff, C. and Ungewickell, E. (2001) *J. Biol. Chem.* 276, 36131–36138.
- [35] Smith, D.B. and Johnson, K.S. (1988) *Gene* 67, 31–40.
- [36] Keen, J.H., Willingham, M.C. and Pastan, I.H. (1979) *Cell* 16, 303–312.
- [37] Holstein, S.E.H., Ungewickell, H. and Ungewickell, E. (1996) *J. Cell Biol.* 135, 925–937.
- [38] Towbin, H., Staehelin, T. and Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* 76, 4350–4354.
- [39] Beevers, L. (1996) *Int. Rev. Cytol.* 167, 1–35.
- [40] Robinson, D.G., Hinz, G. and Holstein, S.E.H. (1998) *Plant Mol. Biol.* 38, 49–76.
- [41] Huang, K.M., Gullber, L., Nelson, K.K., Stefan, C.J., Blumer, K. and Lemmon, S.K. (1997) *J. Cell Sci.* 11, 899–910.
- [42] Scarmato, P. and Kirchhausen, T. (1990) *J. Biol. Chem.* 265, 3661–3668.
- [43] Chu, D.S., Pishvae, B. and Payne, G.S. (1999) *Mol. Biol. Cell* 10, 713–726.
- [44] Parham, P., Brodsky, F.M. and Drickamer, K. (1989) *Biochem. J.* 257, 775–781.
- [45] Lupas, A., Van Dyke, M. and Stock, J. (1991) *Science* 252, 1162–1164.
- [46] Näthke, I.S., Heuser, J., Lupas, A., Stock, J., Turck, C.W. and Brodsky, F.M. (1992) *Cell* 68, 899–910.
- [47] Garnier, J., Gibrat, J.-F. and Robson, B. (1996) *Methods Enzymol.* 266, 540–553.