ELSEVIER

Gene 191 (1997) 143–148

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Nucleotide sequence of a *Drosophila melanogaster* cDNA encoding a calnexin homologue

S. Christodoulou [a], A.E. Lockyer [b], J.M. Foster [c], J.D. Hoheisel [d], D.B. Roberts [a,*]

[a] *Genetics Laboratory, Department of Biochemistry, Oxford University, South Parks Road, Oxford, OX1 3QU, UK*
[b] *Department of Zoology, Aberdeen University, Tillydrone Ave., Aberdeen, AB9 2TN, UK*
[c] *Molecular Parasitology Group, Institute of Molecular Medicine, Oxford University, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK*
[d] *Deutsches Krebsforschungzentrum, Im Neuenheimer Feld 506, D-69120 Heidelberg, Germany*

Received 20 September 1996; received in revised form 2 December 1996; accepted 5 December 1996; Received by J. Wild

Abstract

A cDNA which encodes a calnexin (Cnx)-like protein from *Drosophila melanogaster* has been characterized. The deduced amino acid sequence shares several regions of homology with Cnx from other sources with two conserved motifs each repeated four times. The gene was found to be transcribed in various tissues and at all developmental stages. We have mapped the gene at chromosomal position 99A and we have also mapped the related gene coding for *Drosophila* calreticulin at 85E. © 1997 Elsevier Science B.V.

Keywords: Chaperone; Endoplasmic reticulum; Calcium binding; Glycosylation

1. Introduction

Calnexin (Cnx), a calcium-binding membrane protein of the endoplasmic reticulum (ER), is thought to function as a molecular chaperone aiding the folding and oligomeric assembly of nascent polypeptides. Together with calreticulin – which may be its soluble homologue in the lumen of the ER – they recognize specifically the $GlcMan_9GlcNAc_2$ structure on newly synthesized glycoproteins (Ware et al., 1995). Their functional homology is emphasized by the observation that a chimeric protein consisting of calreticulin and the membrane-anchoring segment of Cnx, recognizes a similar spectrum of proteins to that recognized by Cnx alone (Wada et al., 1995). In addition, these two chaperones also retain in the ER incompletely or incorrectly folded secretory proteins. The function of Cnx (and calreticulin) as a constituent of the ER quality control system has been recently reviewed by Hammond and Helenius (1995).

Genes encoding Cnx have been identified in a variety of eukaryotes including mammals, plants and yeasts (Wada et al., 1991; Huang et al., 1993; Parlati et al., 1995a,b) but while disruption of the *Cnx* gene in the yeast *Schizosaccharomyces pombe* is lethal (Jannatipour and Rokeach, 1995; Parlati et al., 1995b), disruption in *Saccharomyces cerevisiae* is not (Parlati et al., 1995a) and *Drosophila* Schneider cells are functionally Cnx deficient (Vassilakos et al., 1996).

Hong and Ganetzky (1996) while investigating the molecular organization of the 14D region of the *Drosophila melanogaster* (*Dm*) polytene chromosome map located a *Cnx* gene in this region. They also reported a partial sequence of the *Drosophila* homologue. Here we present the complete coding sequence of a *Dm Cnx* homologue which we have mapped elsewhere. There are minor differences between the sequence published here and the partial sequence reported in Hong and Ganetzky (1996). These are listed in Fig. 1.

* Corresponding author. Tel. +44 1865 275321; Fax +44 1865 275318; e-mail: dbr@bioch.ox.ac.uk

Abbreviations: *A.*, *Arabidopsis*; aa, amino acid(s); bp, base pair(s); BSA, bovine serum albumin; *Clr*, *Dm* cDNA encoding calreticulin; Cnx, calnexin(s); *Cnx*, *Dm* cDNA encoding a calnexin-like protein; *Dm*, *Drosophila melanogaster*; ER, endoplasmic reticulum; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; RT-PCR, reverse transcriptase-polymerase chain reaction; *S.*, *Saccharomyces*; *Sc.*, *Schizosaccharomyces*; *UTR*, untranslated region(s).

PII S0378-1119(97)00025-5

```
GAGGCGAAGAGGAAGCACAGAACACAGTAGGCAGCGCAGCAGAGAAAATGTATAATTTACATTAATTTAATTACAAATAGTTCGTCGTGCCAGATCTCGATTCTTGTTATTCCGCAACAG 120

ACCGCAGTGGGGGAAAACTGTTGAAAAGCACAGCGCGCACAAGTGAGTCCACCACCTGGTTTTCCGCGGATACTAAATCCCAGCCAGGCAGCCAGCGAGCGAGCGAAAAGTGGGTGAAAA 240
CALN1
M  A  W  K  M  G  G  N  R  A  A  T  L  A  L  L  F  A  S  S  L  L  L  L  S  S  A  N  A  A  D  L  D  T  E  S  D  D  F  E   40
TGGCGTGGAAAATGGGAGGAAACCGGGCCGCAACGCTGGCGTTGCTCTTCGCCAGCAGTTTGCTGCTGCTGAGTTCGGCCAATGCCGCCGATTTGGACACGGAGTCGGACGACTTCGAGG 360

D  G  Y  V  E  D  V  Q  E  E  P  A  V  I  G  G  D  E  K  L  A  Y  E  S  P  V  I  D  A  K  K  F  H  F  A  D  H  F  D  D   80
ATGGTTATGTGGAGGATGTGCAGGAGGAGCCCGCCGTCATCGGTGGCGATGAGAAGCTGGCCTACGAGAGCCCGGTTATCGATGCCAAGAAGTTCCACTTCGCGGATCACTTCGACGACG 480
                                                                                                CALN2R
V  E  E  S  R  K  R  W  V  L  S  Q  A  K  K  D  D  I  A  E  E  I  S  K  Y  D  G  I  W  N  W  E  S  P  Q  R  I  V  W  A  120
TGGAGGAGTCCAGGAAGAGGTGGGTGCTGTCGCAGGCCAAGAAGGACGACATCGCCGAGGAGATCTCCAAGTACGACGGCATCTGGAACTGGGAGTCGCCGCAGCGCATCGTTTGGGCTA 600

N  D  L  G  L  V  L  K  S  K  A  K  H  A  A  I  A  A  P  L  R  K  P  F  E  F  K  S  D  K  P  L  V  V  Q  Y  E  V  T  L  160
ACGACTTGGGTCTGGTGCTCAAGTCGAAGGCCAAACACGCGGCCATAGCCGCCCCCCTCCGCAAACCGTTCGAGTTCAAGTCGGACAAGCCGCTGGTGGTGCAGTACGAGGTCACGTTGC 720

Q  E  G  Q  E  C  G  G  S  Y  L  K  L  L  S  A  G  K  D  T  E  Q  L  K  A  F  N  D  K  T  P  Y  T  I  M  F  G  P  D  K  200
AGGAGGGTCAGGAGTGCGGTGGCTCGTATCTGAAGCTTTTGTCCGCCGGAAAAGACACGGAGCAACTAAAAGCGTTCAATGACAAGACACCCTATACCATCATGTTTGGACCGGACAAGT 840
                                                CHO
C  G  N  D  V  K  M  H  F  I  F  R  H  V  N  P  I  N  G  T  I  T  E  K  H  C  N  K  P  K  N  R  L  E  E  P  F  K  D  K  240
GCGGCAACGATGTGAAAATGCACTTCATATTCCGGCACGTCAATCCAATTAATGGCACCATCACCGAGAAGCACTGCAACAAGCCAAAGAACCGTTTGGAAGAGCCATTCAAGGACAAGC 960

L  P  H  L  Y  Q  L  V  V  R  P  D  N  S  F  E  I  R  V  D  H  K  I  I  N  E  G  S  L  L  T  D  F  K  P  P  V  N  P  P  280
TACCCCATCTCTACCAGCTGGTGGTCCGTCCCGATAACAGCTTTGAGATTCGCGTTGACCACAAGATCATCAATGAGGGTTCCCTGCTAACCGATTTCAAGCCACCAGTCAACCCGCCGG 1080

A  E  I  D  D  P  N  D  H  K  P  E  S  W  D  E  R  E  K  I  P  D  P  T  A  H  K  P  E  D  W  D  E  D  A  P  P  Q  L  P  320
CAGAGATCGACGACCCCAATGACCACAAGCCCGAATCGTGGGATGAGCGAGAGAAGATCCCAGATCCCACTGCCCACAAGCCCGAGGATTGGGATGAGGATGCTCCACCCCAGTTGCCCG 1200

D  T  D  A  V  M  P  N  G  W  L  E  D  E  P  D  M  I  F  D  P  T  A  T  K  P  E  D  W  D  A  E  I  D  G  E  W  E  A  P  360
ATACCGATGCCGTCATGCCCAATGGCTGGCTCGAAGATGAGCCCGACATGATTTTCGATCCAACTGCCACCAAACCCGAGGATTGGGATGCCGAAATCGATGGCGAGTGGGAGGCTCCAT 1320

L  V  D  N  P  V  C  E  K  A  P  G  C  G  K  W  K  A  P  L  I  P  N  P  N  Y  K  G  K  W  R  A  P  M  I  E  N  P  N  Y  400
TGGTGGACAATCCAGTGTGCGAGAAGGCTCCCGGCTGTGGCAAGTGGAAGGCTCCGCTCATCCCCAATCCCAACTACAAGGGCAAGTGGCGTGCCCCAATGATCGAGAACCCCAACTACC 1440

Q  G  K  W  A  P  R  K  I  P  N  P  D  F  F  E  D  L  K  P  F  Q  M  T  P  I  S  A  V  G  L  E  L  W  S  M  S  S  D  I  440
AGGGCAAGTGGGCACCCAGGAAGATTCCCAATCCAGACTTCTTTGAGGATCTGAAGCCCTTCCAAATGACGCCCATCAGCGCTGTGGGTCTGGAACTGTGGTCCATGTCCAGCGATATTC 1560

L  F  D  N  L  I  I  T  D  D  V  E  V  A  R  D  F  A  A  N  S  F  D  I  K  R  R  Y  I  D  R  E  S  D  S  F  V  N  K  V  480
TCTTCGACAACCTAATCATCACGGACGACGTGGAGGTGGCCCGTGACTTTGCCGCCAACAGCTTCGACATCAAGCGTCGCTACATCGATCGTGAATCGGACTCATTCGTGAATAAGGTAG 1680

V  E  L  A  K  A  N  P  S  I  W  G  I  G  L  V  A  I  V  A  L  V  A  L  T  I  Y  C  R  F  G  T  A  K  S  Q  D  S  A  A  520
TCGAGCTAGCCAAGGCCAATCCCTCGATCTGGGGCATTGGCTTGGTGGCCATTGTGGCGCTGGTTGCCCTCACCATCTACTGTAGATTTGGTACCGCTAAGAGTCAGGACTCGGCTGCCA 1800

K  K  A  A  A  E  A  K  K  S  D  D  P  Q  P  D  D  E  P  E  A  E  E  E  S  D  E  R  A  A  G  D  T  S  K  E  S  T  P  L  560
AAAAGGCTGCTGCCGAGGCCAAGAAGTCCGACGATCCTCAGCCCGACGATGAGCCCGAGGCCGAGGAGGAGAGCGACGAAAGGGCCGCTGGCGATACCAGCAAGGAGAGCACACCGCTGT 1920

S  A  S  P  K  K  N  Q  K  S  D  L  D  D  N  E  E  E  S  K  A  A  E  S  R  E  P  A  Q  T  E  E  S  N  T  K  T  R  K  R  600
CCGCTAGTCCCAAGAAGAACCAAAAGTCTGATTTAGATGATAACGAAGAGGAGTCCAAGGCGGCGGAGAGCCGAGAGCCCGCACAGACTGAGGAATCTAACACAAAAACACGCAAGCGTC 2040

Q  A  R  K  E  *                                                                                                          605
AAGCGCGCAAGGAGTAAGCGGGCATCCCCACATGCCGAATCCGAAAACAAACACTGCAACTCTAAAG                                                       2107
```

Fig. 1. Nucleotide sequence of *Cnx*, the *Dm* calnexin homologous gene, and its deduced aa sequence. EMBL Data library accession No. X99488. The underlined aa sequence corresponds to the putative transmembrane domain. The single consensus *N*-glycosylation site at N218 is marked CHO. The locations of the two primers (caln1 and caln2R) used in the RT-PCR are indicated by underlining the nt sequence. Our sequence differs from the one presented in Hong and Ganetzky (1996) in the following nt some of which result in aa differences, our sequences in bold: 776, GCTC and **CACG** which changes EL to **DT**; 992, the **C** is missing and 996, CGA instead of **AA** which restores the reading frame but gives **DN** instead of ID; 1022, **C** and T; 1067, **A** and G; 1082, **A** and G; 1105, **A** and T which changes **H** to L; 1145, **ATCCCACT** and ACCCAACT; 1223, **T** and C; 1415, **C** and A; 1499, **CCCTTCCAAATGACGCCCCAT** and CCTTCAAATGACGCCAT; the loss of three Cs (underlined) restores the reading frame but changes the aa sequence **KPFQMTPIS** and KPSNDAIS (mouse sequence PFRMTPFS); 1648, A is missing which results in a change in reading frame up until the end of the partial sequence.

2. Experimental and discussion

2.1. Cloning

A *Dm* cDNA library from 0–8 h embryos, inserted into pNB40 (Brown and Kafatos, 1988) and arrayed as an in situ filter bearing DNA from 19 200 clones (Hoheisel et al., 1991) was screened under low-stringency conditions (hybridization: 0.5 M sodium phosphate pH 7.2, 7% SDS, 1 mM EDTA, 1% BSA at 55°C for 16 h; washing: 6 × SSC at 55°C twice for 30 min) with a ^{32}P-labelled fragment of the rat ER mannosidase cDNA (Bischoff et al., 1990). Four positive clones were selected but only restriction fragments of the cDNA clone 2:37G4 hybridized with the rat sequence at stringent washing conditions (0.2 × SSC, 65°C). An approx. 0.78-kb *Hin*dIII-*Hin*dIII fragment from clone 2:37G4 was subcloned into pUC18 and sequenced. The sequence showed no significant homology to the rat ER clone, and we cannot explain why the sequences hybridized under these stringent conditions. However, the sequence was used to search the GenEMBL and SwissProt databases. This search identified human Cnx and several other mammalian Cnx with significant homology (55% similarity with human Cnx at the amino acid (aa) level and the full sequence 65% similarity and 54% identity). We decided to sequence the entire cDNA in clone 2:37G4 because Cnx plays a role in the processing of

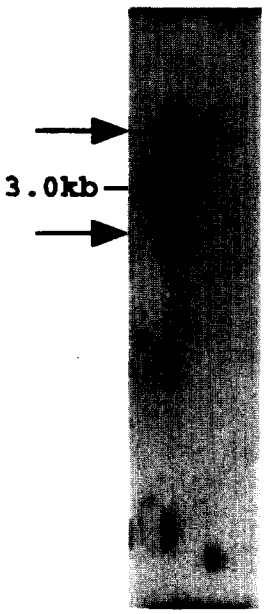

Fig. 2. Northern blot of adult *Drosophila* mRNA probed with the 2:37G4 cDNA clone. 3 μg mRNA (isolated directly from flies using the Quick Prep Micro mRNA purification kit; Pharmacia) was loaded on the gel. The transcript sizes were estimated using RNA standards from Boehringer Mannheim. The two arrows point to the two minor bands of 4.3 and 1.7 kb described in the text.

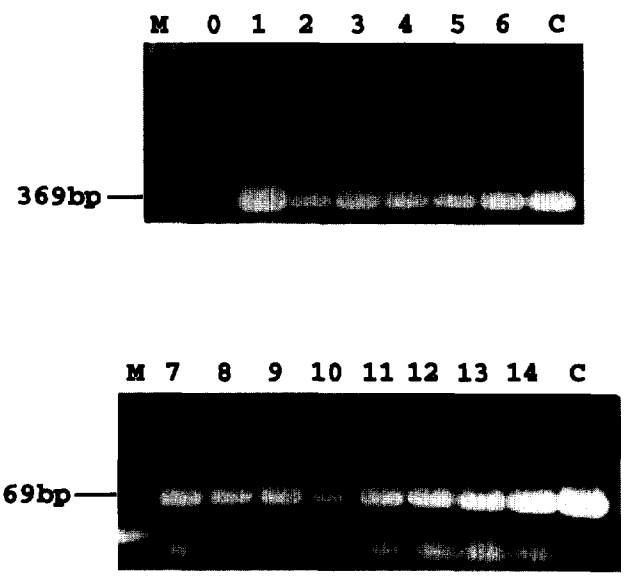

Fig. 3. RT-PCR products. Tissue dissection was performed in Ringer's solution as described in Roberts (1986) and tissues were kept at −40°C until mRNA isolation. First-strand cDNA from all samples was prepared from mRNA (isolated directly from samples using the Quick Prep Micro mRNA purification kit; Pharmacia) using the first-strand cDNA synthesis kit (Pharmacia). The mRNA-cDNA heteroduplex was then used as a template in a standard PCR reaction using the oligos caln1 (5′-CAGACCGCAGTGGGGGA) and caln2R (5′-GGCATCGATAACCGGGC) as primers (see Fig. 1). These amplify a 328-bp fragment from the *Cnx* cDNA. mRNA was isolated from embryos (1), larvae (2), pupae (3), adult flies (4), heads (5), bodies (6), imaginal discs (7), salivary glands (8), fat body (9), gut (10), brain (11), testes (12), oocytes (13) and ovaries (14). M, 123-bp DNA ladder (BRL); 0, no DNA; C, plasmid 2:37G4.

glycoproteins which is in keeping with our interest in the biological significance of *N*-glycosylation in *Dm* (Williams et al., 1991; Foster et al., 1995).

2.2. Sequence analysis

Clone 2:37G4 was sequenced, on both strands, directly from the purified double-stranded DNA using oligodeoxyribonucleotide (oligo) primers prepared as the known sequence advanced. It is 2107 bp long and contains one open reading frame (ORF) (1818 bp), together with 239 bp and 50 bp of 5′ and 3′ non-coding regions, respectively (Fig. 1). The aa sequence deduced from the ORF shares all the features of other Cnx sequences. The ORF predicts a polypeptide of 601 aa (dog 593).

The first in-phase Met residue has an ATG codon in a consensus sequence which fulfils the criteria for a 'strong' start codon, i.e., an A at −3 and a G at +4 (Kozak, 1989). The second ATG codon in the reading frame is 12 bp downstream from the first and shares the same 'strong' start consensus sequence. Neither, however, has an A/C at −4 which is found in a majority of *Drosophila* sequences (Cavener, 1987). There is no reason therefore to suppose that the first ATG is not the start codon.

The N-terminal aa have the features of a typical signal peptide with a positive aa, Arg, followed by a stretch of 15 aa, 12 of which are hydrophobic. This is followed by the potential intra-luminal domain (aa 25–489; dog 21-482) and a typical type I transmembrane domain 490–510 (dog 483-593) with 16 out of 21 aa being hydrophobic. Finally there is a short potential cytoplasmic sequence of 91 aa (dog 90), which is highly charged (27% acidic and 18% basic) and ends with the sequence RQARKE. This C-terminal sequence has an overall similarity to the RKPRRE C-terminal sequence found in mammals. Like its mammalian and *Arabidopsis* counterparts, *Dm* Cnx also has Arg at −3. This is in contrast with the C-terminal motifs shown to retain transmembrane proteins in the ER, all of which have a Lys at −3 (Shin et al., 1991). There is a consensus *N*-glycosylation site at aa positions 218–220 (Fig. 1). Unlike their mammalian counterparts, Cnx from *S. cerevisiae*, *Sc. pombe* and from *Arabidopsis* also possess putative *N*-glycosylation sites within their luminal domains. In *Sc. pombe* this site is actually glycosylated (Jannatipour and Rokeach, 1995); we do not know, however, whether this site is also glycosylated in *Dm*.

2.3. Sequence comparison to other Cnx

In the potential intra-luminal sequence there are two short sequences each of which is highly conserved and is repeated four times in the Cnx and three times in the calreticulins. These sequences are thought to be involved in calcium binding (Tjoelker et al., 1994) since it has

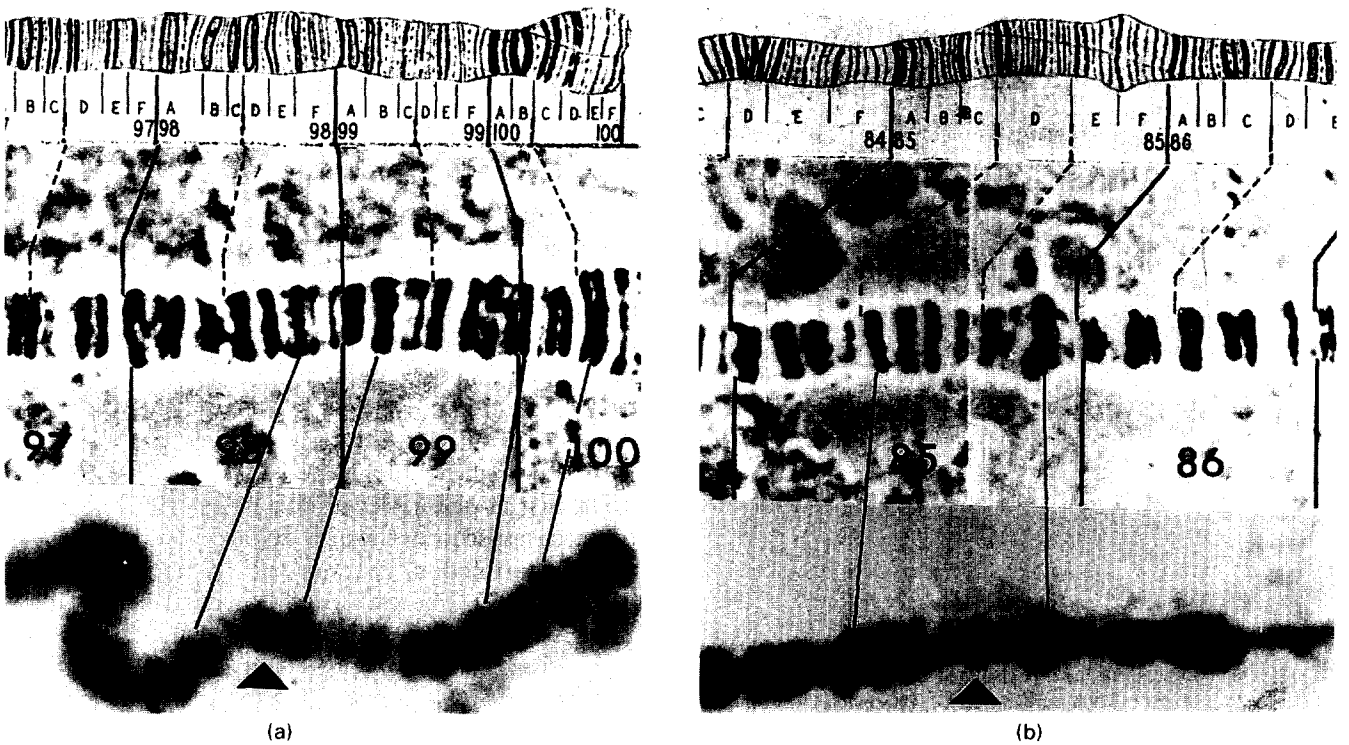

Fig. 4. Chromosomal localization of *Cnx* and *Crl*. Below: polytene chromosome squash probed in situ with the *Dm* cDNA, *Cnx* (left) and *Crl* (right). Above: localization of the in situ positive bands to 99A1-6 (*Cnx*) and 85E1-5 (*Crl*). Adapted from Lefevre (1976). Method is described previously (Foster et al., 1995).

been shown that Cnx bind calcium but do not have the 'E-F hand' motif for binding calcium in calmodulin (Tjoelker et al., 1994 and refs. therein).

Using the alignment programme Gap of Needleman and Wunsch (Genetics Computer Group, University of Wisconsin) the *Dm Cnx* homologue shows 54% identity with the human gene and 36% identity with the deduced *Drosophila* calreticulin sequence (Smith, 1992).

We have modified the previously published consensus sequences by comparing the 24 sequences (four in each of six species: *S. cerevisiae*, *Sc. pombe*, *A. thaliana*, *Caenorhabiditis elegans*, and human, representing the mammals) to produce a consensus sequence in which at least two thirds of the sequences share the same aa at a particular position. The first repeat consensus sequence is P-A-KLP-DWDE and the second repeat consensus sequence is **G-W-P-I-NP**. The aa sequence in bold is either invariant or shows only one difference in 24. There is no evidence for concerted evolution by gene conversion, or by any other mechanism, since repeats within the same species show no greater similarity than the repeats between species. These conserved repeats also occur, but only three times each, in the calreticulin genes.

2.4. RNA analysis

It seems that the 50 bp of non-coding sequence downstream from the TAA stop codon do not represent the complete 3′ *UTR* of the *Cnx* mRNA as there is no recognisable polyadenylation signal present in this sequence. We presume that the sequence was cloned from an internal run of poly(A) before the poly(A) tail (see Brown and Kafatos, 1988). In agreement, the major mRNA band detected in Northern analysis of adult flies with the cDNA 2:37G4 probe is an approx. 3.0-kb band, nearly 1 kb larger than the sequence reported here (Fig. 2). Hong and Ganetzky (1996) also reported an approx. 3.0-kb transcript corresponding to Cnx. Our Northern blot identified two additional minor bands at 4.3 and 1.7 kb. These minor transcripts could be the result of cross-hybridization to either calreticulin or to other *Cnx*-like *Drosophila* genes. In mammals, Cnx isoforms have been described both in the ER of specific tissues (Ohsako et al., 1994) and in different cellular compartments (Gilchrist and Pierce, 1993). However, our chromosome mapping data (see Section 2.5) are consistent with *Cnx* being a single copy gene. It is interesting to speculate that the 1.7-kb band seen on our Northern could be the result of alternative splicing of exons as has recently been reported for the pre-messenger RNA encoding mouse calcitonin and calcitonin-related peptide (Lou et al., 1996).

RT-PCR was used to test for the presence of the *Cnx* mRNA species in various larval and adult tissues and also during different developmental stages. The primers used in the PCR resulted in the amplification of a 328-bp

fragment from the *Cnx* cDNA (Fig. 3). We found the *Cnx* mRNA to be present in all tissues and at all developmental stages examined. We anticipate the presence of the message to correlate with presence of the polypeptide consistent with its key role in cellular functions.

2.5. *Chromosome mapping of* Cnx

The chromosomal location of *Cnx* was determined by in situ hybridization to squashes of polytene chromosomes from salivary glands of *Dm*. The *Cnx* sequence was located to a single site (99A1-6) on the right arm of chromosome 3 (Fig. 4). As mentioned in Section 1, Hong and Ganetzky (1996) localized the *Cnx* gene in the 14D region. We were unable to detect a signal in that chromosomal region in any of our squashes probed with cDNA 2:37G4. Thus, the results from our in situ hybridization support the existence of a single *Cnx*-like sequence in the *Dm* genome (if more than one copy is present then they must be very close together).

We have also localized the related calreticulin sequence to 85E1-5 on the right arm of the same chromosome (Fig. 4), using the pUC/D2 plasmid described in Smith (1992) as a probe. Although these two sequences (*Cnx* and *Clr*) show considerable homology they did not cross-hybridize in our preparations.

3. Conclusions

We have described the nt sequence of *Cnx*, a *Dm* cDNA encoding a Cnx homologue. The deduced aa sequence of Cnx shares 54% identity with the human protein and contains two conserved motifs each of which is repeated four times throughout the central region. The gene is constitutively expressed at the RNA level and localizes at 99A1-6 on the 3 chromosome.

Acknowledgement

This work was supported by a grant from the Wellcome Trust to D.B.R. We wish to thank Dr. Joyce Bischoff for providing the ER mannosidase clone, Dr. M. Smith for providing the calreticulin clone and Dr. Yong Quing Zhang for confirming the cytological positions of the probes in the in situ hybridization preparations.

References

Bischoff, J., Moremen, K. and Lodish H.F. (1990) Isolation, characterization and expression of a cDNA encoding a rat liver endoplasmic reticulum α-mannosidase. J. Biol. Chem. 265, 17110–17117.

Brown, N.H. and Kafatos, F.C. (1988) Functional cDNA libraries from *Drosophila* embryos. J. Mol. Biol. 203, 425–437.

Cavener, D.R. (1987) Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. Nucleic Acids Res. 15, 1353–1361.

Foster, J.M., Yudkin, B., Lockyer, A.E. and Roberts D.B. (1995) Cloning and sequence analysis of *GmII*, a *Drosophila melanogaster* homologue of the cDNA encoding murine Golgi α-mannosidase II. Gene 15, 183–186.

Gilchrist, J.S.C. and Pierce, G.N. (1993) Identification and purification of a calcium-binding protein in hepatic nuclear membranes. J. Biol. Chem. 267, 22542–22549.

Hammond, C. and Helenius, A. (1995) Quality control in the secretory pathway. Curr. Opin. Cell Biol. 7, 523–529.

Hoheisel, J., Lennon, G.G., Zehetner, G. and Lehrach, H. (1991) Use of high coverage reference libraries of *Drosophila melanogaster* for relational data analysis. J. Mol. Biol. 220, 903–914.

Hong, C.-S. and Ganetzky, B. (1996) Molecular characterization of neurally expressing genes in the *para* sodium channel gene cluster of *Drosophila*. Genetics 142, 879–892.

Huang, L., Franklin, A.E. and Hoffman, N.E. (1993) Primary structure and characterization of *Arabidopsis thaliana* calnexin-like protein. J. Biol. Chem. 268, 6560–6566.

Jannatipour, M. and Rokeach, L.A. (1995) The *Schizosaccharomyces pombe* homologue of the chaperone calnexin is essential for viability. J. Biol. Chem. 270, 4845–4853.

Kozak, M. (1989) The scanning model for translation: an update. J. Cell Biol. 108, 229–241.

Lefevre Jr., G. (1976) The polytene chromosomes. In: Ashburner, M. and Novitski, E. (Eds.), The Genetics and Biology of *Drosophila melanogaster*, Vol 1a. Academic Press, London, pp. 31–66.

Lou, H., Gagel, R.F. and Berget, S.M. (1996) An intron enhancer recognized by splicing factors activates polyadenylation. Genes Dev. 10, 208–219.

Ohsako, S., Hayash, Y. and Bunick, D. (1994) Molecular cloning and sequencing of calnexin-t. J. Biol. Chem. 269, 14140–14148.

Parlati, F., Dominguez, M., Bergeron, J.J.M. and Thomas, D.Y. (1995a) *Saccharomyces cerevisiae CNE1* encodes an endoplasmic reticulum (ER) membrane protein with sequence similarity to calnexin and calreticulin and functions as a constituent of the ER quality control apparatus. J. Biol. Chem. 270, 244–253.

Parlati, F., Dignard, D., Bergeron, J.J.M. and Thomas, D.Y. (1995b) The calnexin homologue *cnx1*$^+$ in *Schizosaccharomyces pombe*, is an essential gene which can be complemented by its soluble ER domain. EMBO J. 14, 3064–3072.

Roberts, D.B. (1986) Basic *Drosophila* care and techniques. In: Roberts, D.B. (Ed.), *Drosophila*: a Practical Approach. IRL Press, Oxford, pp. 1–38.

Shin, J., Dunbrack, R.L., Lee, S. and Strominger, J.L. (1991) Signals for retention of transmembrane proteins in the endoplasmic reticulum studied with CD4 truncation mutants. Proc. Natl. Acad. Sci. USA 88, 1918–1922.

Smith, M.J. (1992) Nucleotide sequence of a *Drosophila melanogaster* gene encoding a calreticulin homologue. DNA Sequence 3, 247–250.

Tjoelker, L.W., Seyfried, C.E., Eddy, R.L., Byers, M.G., Shows, T.B., Calderon, J., Schreiber, R.B. and Gray, P.W. (1994) Human, mouse and rat calnexin cDNA cloning: identification of potential calcium binding motifs and gene localization to human chromosome 5. Biochemistry 33, 3229–3236.

Vassilakos, A., Cohen-Doyle, M.F., Peterson, P.A., Jackson, M.R. and Williams, D.B. (1996) The molecular chaperone calnexin facilitates folding and assembly of class I histocompatibility molecules. EMBO J. 15, 1495–1506.

Wada, I., Rindress, D., Cameron, P.H., Ou, W.-J., Doherty II, J.J., Louvard, D., Bell, A.W., Dignard, D., Thomas, D.Y. and Bergeron, J.J.M. (1991) SSRa and associated calnexin are major calcium

binding proteins of the endoplasmic reticulum membrane. J. Biol. Chem. 266, 19599–19610.

Wada, I., Imai, S.I., Kai, M., Sakane, F. and Kanoh, H. (1995) Chaperone function of calreticulin when expressed in the endoplasmic reticulum as the membrane-anchored and soluble forms. J. Biol. Chem. 270, 20298–20304.

Ware, F.E., Vassilakos, A., Peterson, P.A., Jackson, M.R., Lehrman, M.A. and Williams, D.B. (1995) The molecular chaperone calnexin binds $Glc_1Man_9GlcNA_{c2}$ oligosaccharide as an initial step in recognizing unfolded glycoproteins. J. Biol. Chem. 270, 4697–4704.

Williams, P.J., Wormald, M.R., Dwek, R.A., Rademacher T.W., Parker, G.F. and Roberts, D.B. (1991) Characterisation of oligosaccharides from *Drosophila melanogaster* glycoproteins. Biochim. Biophys. Acta 1075, 146–153.