



HAL
open science

Drosophila expresses a CD98 transporter with evolutionarily conserved structure and amino acid transport properties

Bruno Reynolds, Pietro Roversi, Robert Laynes, Shubana Kazi, Richard Boyd, Deborah C. I. Goberdhan

► **To cite this version:**

Bruno Reynolds, Pietro Roversi, Robert Laynes, Shubana Kazi, Richard Boyd, et al.. Drosophila expresses a CD98 transporter with evolutionarily conserved structure and amino acid transport properties. *Biochemical Journal*, 2009, 420 (3), pp.363-372. 10.1042/BJ20082198 . hal-00479126

HAL Id: hal-00479126

<https://hal.science/hal-00479126>

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Drosophila* expresses a CD98 transporter with evolutionarily conserved structure and amino acid transport properties**

Bruno Reynolds*, **Pietro Roversi†**, **Robert Laynes***, **Shubana Kazi***, **Richard Boyd***, and **Deborah C. I. Goberdhan*¹**

*Department of Physiology, Anatomy and Genetics, Le Gros Clark Building, University of Oxford, South Parks Road, Oxford, OX1 3QX

†Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, OX1 3ER

¹To whom correspondence should be addressed (deborah.goberdhan@dpag.ox.ac.uk). Tel: +44 1865 282836, Fax: +44 1865 272420

Short title: Evolutionarily conserved properties of *Drosophila* CD98

SYNOPSIS

Mammalian CD98 heterodimeric amino acid transporters consist of a promiscuous single-pass transmembrane glycoprotein, CD98hc, the 'heavy chain', and one of six multipass transmembrane proteins, or 'light chains'. The heterodimeric complexes of CD98hc and the light chains LAT1 or LAT2 specifically promote sodium-independent System L exchange of neutral amino acids, including leucine. CD98hc is also implicated in other processes, including cell fusion, cell adhesion and activation of TOR signalling. Recent reports surprisingly suggested that insects lack a membrane-bound CD98hc, but here we show that *Drosophila* CG2791 encodes a functional CD98hc orthologue with conservation in intracellular, transmembrane and extracellular domains. We demonstrate by RNAi knockdown in *Drosophila* Schneider cells that CG2791 and two *Drosophila* homologues of the mammalian CD98 light chains, Minidiscs (Mnd) and JhI-21, are required for normal levels of System L transport. Furthermore, we show that System L activity is increased by methoprene, an analogue of the developmentally regulated endocrine hormone, Juvenile Hormone, an effect potentially mediated by elevated Mnd expression. Co-expression of CG2791 and JhI-21, but not CG2791 and Mnd, in *Xenopus* oocytes mediates System L transport. Finally, mapping of conserved sequences on to the recently determined crystal structure of the human CD98hc extracellular domain highlights two conserved exposed hydrophobic patches at either end of the domain that are potential protein-protein interaction surfaces. Our data therefore not only show that there is functional conservation of CD98hc System L transporters in flies, but also provide new insights into the structure, functions and regulation of heterodimeric amino acid transporters.

Key words: SLC3A1, SLC7, heterodimeric amino acid transporter, System L, *Drosophila*, growth control.

Abbreviations used: AAT, amino acid transporter; BCH, 2-aminobicyclo[2.2.1]heptane-2-carboxylic acid; CD98hc (also known as 4F2hc, SLC3A2, FRP-1), CD98 heavy chain; LAT (L-type amino acid transporter) 1 and 2; SLC, solute carrier; Mnd, Minidiscs; Gb, Genderblind; PAT (also known as SLC36), proton-assisted amino acid transporter; Slif, Slimfast; IIS, Insulin/IGF signaling; InR, insulin receptor; JH, Juvenile Hormone; RpL32, ribosomal protein L32; S2 cells, Schneider cells; FCS, fetal calf serum; dsRNA, double-stranded RNA; LSC, liquid scintillation counting; Q-RT-PCR, quantitative real time PCR; PBS, phosphate buffered saline.

INTRODUCTION

Transport of amino acids into all eukaryotic cells is essential for maintaining viability and promoting growth. Physiological studies in mammalian cell culture have identified many different amino acid transporter (AAT) 'Systems' [1, 2]. More recently, cloning of the genes involved in these activities gave rise to the revised solute carrier (SLC) transporter classification [3, 4]. However, the complexity of AAT families and the difficulty of developing *in vivo* models, where the activities of individual transporters can be assessed under truly physiological nutrient and growth factor conditions, has until recently limited progress in the functional analysis of these molecules [5].

System L transport activity involves the uptake of leucine and other bulky hydrophobic amino acids independently of sodium ions [6]. It is specifically inhibited under sodium-free conditions by the synthetic amino acid analogue, 2-aminobicyclo[2.2.1]heptane-2-carboxylic acid (BCH; reviewed in [1]). Members of the widely expressed family of CD98 heterodimeric amino acid transporters (reviewed in [7-9]) are important mediators of System L transport. Each CD98 transporter consists of a common single-pass transmembrane N-glycoprotein 'heavy chain' [CD98hc, otherwise known as 4F2hc, SLC3A2 or FRP-1 (9)] associated covalently with a multipass transmembrane 'light chain' (LAT1, LAT2, γ^+ LAT1, γ^+ LAT2, asc1, or xCT in humans). CD98hc requires LAT1 or LAT2, which are members of the SLC7 family [10], for System L transport. Substrate specificity is conferred by the light chain, which forms all or most of the transport pore, while CD98hc has been implicated in trafficking of the light chain to the plasma membrane [11]. These transporters are attracting increasing attention, because there is a growing body of evidence that classical System L transport plays a role in cell growth and in a wide spectrum of human cancers [12]. Indeed, recent studies in cell culture suggest a role for System L transport in leucine-dependent mTOR signalling [13, 14], a process implicated in growth activation in many tumours [15-17].

In addition to regulating growth, work in cell culture has also suggested that CD98hc has roles in cell fusion, cell adhesion and integrin signaling (reviewed in [9]). However, the significance of these functions under physiologically relevant conditions *in vivo* is unclear. Although CD98hc has been shown to be essential for early mouse development [18], these knockout studies have so far failed to provide detailed insights into the critical cellular and developmental functions of this molecule and System L transport.

The genetic tractability of *Drosophila* makes it a particularly powerful system to dissect physiologically relevant roles of AATs [19]. Indeed, several studies have already highlighted important functions for AATs in flies that would be difficult to identify purely through cell culture studies. For example in the SLC7 class [10], the cationic AAT Slimfast [20] and the putative heterodimeric AAT light chain Minidisks (Mnd or CG3297; [21]) seem to play important nutrient-sensing roles in the endocrine larval fat body, which controls growth rates throughout the organism. More recently, we have shown that members of the proton-assisted amino acid transporter (PAT/SLC36) family have uniquely potent cell-autonomous effects on growth when compared to other AATs *in vivo* [22].

Surprisingly, two recent reports have suggested that there is no membrane-associated CD98hc orthologue in *Drosophila* [23, 24], indicating that CD98hc is apparently involved in uniquely vertebrate functions that could not be dissected in simple genetic systems. However, here we report the identification of a functional CD98hc orthologue, CG2791, paving the way for a detailed *in vivo* genetic analysis. We show that *Drosophila* Schneider (S2) cells share characteristic properties of BCH-sensitive System L transport with mammalian cells and that this transport is mediated by CG2791 and a conserved light chain, JhI-21, as well as a second transporter involving the light chain Mnd. Furthermore, sequence comparisons and structural modelling reveal novel conserved features of the CD98hc extracellular domain that may be involved in its interactions with as yet unknown binding partner(s).

EXPERIMENTAL

Drosophila S2 cell culture

Cells were cultured in Schneider's medium (Invitrogen) with 10% heat-inactivated fetal calf serum (FCS) and typically grown to 70% confluence prior to transport analysis.

Uptakes experiments in *Drosophila* S2 cells

Cells were gently washed twice in sodium-free phosphate buffered saline, pH 7.4 (PBS) and then incubated in a further 1 ml of PBS for 30 min. PBS was replaced with choline-substituted PBS containing 50 μM 4,5 ^3H L-Leucine (specific activity: 2.25 TBq/mmol) \pm 10 mM 2-aminobicyclo[2.2.1]heptane-2-carboxylic acid (BCH; Sigma, A7902) or 10 mM unlabelled L-leucine and incubated at room temperature for 3 min (shown in earlier experiments to approximate initial rate conditions). External medium (with labelled extracellular amino acid) was then aspirated and cells quickly washed with 2 ml ice-cold stop solution (PBS + 20 mM cold leucine). Stop solution was removed and cells lysed with 1 ml of 0.1% (w/v) sodium dodecyl sulfate, 0.1% (w/v) NaOH. The radioactivity of a 500 μl aliquot of each sample was measured by liquid scintillation counting (LSC). In this paper, mediated transport is defined as leucine transport under specified conditions minus non-specific leucine uptake and/or binding (see Figure 1). To measure the effect of methoprene on uptake, a comparison between BCH-sensitive leucine uptake of cells preincubated for 36 h in the presence and absence of 50 μM methoprene was made. All transport assays were repeated in triplicate in each experiment, and each experiment was repeated at least three times. Figures show representative data from one experiment.

Efflux measurements in *Drosophila* S2 cells

Washed cells were loaded for 10 min with 50 μM ^3H L-leucine in otherwise amino acid-free PBS, pH 7.4. Following this, the cells were resuspended in an identical medium, but free of radioactivity with or without 10 mM leucine or 10 mM BCH. Efflux rate constants were determined by fitting a mono-exponential decay curve through the experimentally measured cellular ^3H L-Leucine at various times following the start of the washout. Values at the start of the experiment and 90 min later were taken to represent the intracellular values at zero time (Q_0) or at equilibrium (Q_∞) respectively. The data were well fitted by the equation $Q_t = Q_0(e^{-kt})$, where Q_t is the quantity of intracellular amino acid at time t from the start of the washout and k is the efflux rate constant. Samples from the extracellular medium were removed in parallel and counted, allowing the appearance of radioactivity in this compartment to give an independent estimate of efflux. These data fitted quantitatively with the values determined from the cellular amino acid content. Washout medium obtained from cells pre-loaded as above with ^3H L-leucine was analysed by HPLC [25]; in three independent experiments an average of 97% of the radioactivity present in this medium ran on a 5 μm ODS C18 column with the same elution profile as L-leucine, indicating that it was leucine and not a cellular metabolite of this amino acid that was transported out of the cell.

Production of dsRNA (double-stranded RNA) and conditions for RNAi knockdown of target genes in *Drosophila* S2 cell culture

DNA fragments of approximately 500 bp in length from the coding sequence of the target genes that included a splice junction were generated by PCR. T7 primer sites were incorporated at the 5' end of both strands. dsRNA was generated with a MEGAscript[®] high yield transcription kit (Ambion). One microgram of PCR product was used in each 20 µl reaction and incubated for 16 h at 37°C. dsRNA transcripts were purified by phenol:chloroform extraction followed by isopropanol precipitation, and resuspended in nuclease-free ddH₂O. Primers used to generate dsRNAs from *Drosophila* Genomics Resource Center clones were: *CG2791* clone RE39378, sense primer 912-931, antisense primer 1411-1392; *Jhl-21* clone LD39658 sense primer 1374-1393, antisense primer 1923-1904; *mnd* clone LD25378 sense primer 930-950, antisense primer 1424-1404. The 5' end of the primer pairs was preceded with the T7 RNA polymerase binding site (GAATTAATACGACTCACTATAGGGAGA).

One ml of 1×10^6 *Drosophila* S2 cells was plated at approximately 70% confluence in medium containing 10% FCS in a six well cell culture dish (Corning), and incubated at 23°C overnight to allow cells to adhere to the base of the well. Medium was aspirated and replaced with 500 µl serum-free medium containing 15 µg of dsRNA in each well and after 1 h incubation at room temperature, 1 ml of medium containing 10% FCS was added, followed by a further 48 h incubation at 23°C. A second 500 µl dose of serum-free medium containing 15 µg of dsRNA was added and the cells incubated for a further 72 h.

Measurement of mRNA levels in *Drosophila* S2 cells

RNA was extracted using the RNeasy minikit (Qiagen). cDNA was synthesized from each RNA preparation using the High Capacity Reverse Transcriptase kit (Applied Biosystems). In all knockdown and methoprene experiments, levels of the knocked down mRNA and of a control ribosomal protein mRNA, *RpL32*, were assessed by quantitative real-time PCR (Q-RT-PCR), using the 7000 sequence detection system from Applied Biosystems and Quantitect[®] primer assays from Qiagen for the following *Drosophila* genes: *CG2791*, *mnd* (CG3297), *Jhl-21* (CG12317), CG13384 and *RpL32* (CG7939).

Uptake experiments in *Xenopus* oocytes

mRNA synthesis, *Xenopus* oocyte preparation and injection was as described previously [26]. Uptakes, initiated by addition of 50 µM ³H L-Leucine to individual oocytes (five oocytes studied separately for each condition in at least three different experiments) were carried out under initial rate conditions (30 min) in choline-buffered sodium-free media with and without 10 mM BCH. At the end of the uptake, individual oocytes were washed, lysed and their radioactivity measured by LSC [26].

Alignments and structural comparisons

In Figure 2 and Figure 6A, the protein sequence for the human CD98 isoform f (accession number NP_001013269) was used because it shares the greatest homology with CG2791. The other isoforms (a-e) all have a longer N terminus of varying lengths between 39-132 residues, but are otherwise 100% homologous with isoform f. Sequences were aligned using the ClustalW programme at EMBL-EBI (<http://www.ebi.ac.uk/clustalw/index.html>) and then processed with the Boxshade programme (http://www.ch.embnet.org/software/BOX_form.html). Percentage identities for Figure 2B were calculated and plotted using GraphAlign (http://darwin.nmsu.edu/cgi-bin/graph_align.cgi) [27].

Evolutionarily conserved properties of *Drosophila* CD98

Transmembrane domains were predicted by the Technical University of Denmark CBS Prediction Server TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) and hydrophathy plots produced by <http://genome.ucsc.edu/cgi-bin/pbGateway> and using [28], e.g. Figure 3B and Supplementary Figure 1. Mapping of the surface sequence conservation of CG2791 on to the structure of the human CD98hc homodimer (PDB ID 2DH3) was performed using the WHISCY server at <http://www.nmr.chem.uu.nl/whiscy> [29]. Figure 6B and Supplementary Figure 3 were drawn using PyMOL [30].

RESULTS

***Drosophila* S2 cells use System L transport to mediate sodium-independent leucine uptake and exchange**

As a first step to determine whether there is a functional homologue of CD98hc in flies, we tested whether S2 cells can take up leucine in a sodium-independent fashion, a process that is predominantly undertaken by CD98hc-containing System L transporters in mammals. Mediated uptake of leucine into S2 cells was measured as described in the Experimental section. Significant levels of sodium-independent leucine uptake were readily detected (Figure 1A). System L transporters in mammalian cells have been shown to be inhibited by 2-aminobicyclo[2.2.1]heptane-2-carboxylic acid (BCH), which competes with leucine for transport [1]. Such inhibition, under sodium-free conditions, is a defining feature of System L. We found that BCH strongly inhibits sodium-independent leucine uptake in S2 cells, suggesting that System L transport also exists in flies.

We also measured efflux of leucine (see Experimental section) from S2 cells pre-loaded with ^3H L-leucine. Efflux of the intact amino acid occurred in control cells (through uncharacterized pathways) but was markedly stimulated by the addition of unlabelled leucine in the extracellular medium (Figure 1B). Such trans-stimulation of amino acid transport is characteristic of heterodimeric amino acid transporter systems in mammals. Furthermore, extracellular BCH also enhanced leucine efflux, again strongly supporting the hypothesis that transport of the amino acid is through System L. These results suggest that System L is the primary sodium-independent transport system for leucine in S2 cells. In addition, the fly System L transporters show properties similar to their mammalian counterparts: they promote a marked exchange of intra- and extra-cellular neutral amino acids, and use BCH as a substrate.

CG2791 encodes the *Drosophila* orthologue of CD98hc.

We performed a sequence search to identify a homologue of human CD98hc in flies. Initial BLAST searches highlighted a range of *Drosophila* gene products with strong sequence homology to CD98hc in the extracellular α -amylase-like domain [31]. Further inspection revealed that the product of just one of the genes identified, CG2791, shows homology with CD98hc (isoform f) along its entire length (17% identity overall; Figure 2 and Table 1). All three major structural domains identified in CD98hc, namely cytosolic domain, TM domain and α -amylase domain [8, 23 and 31] are conserved in CG2791. Interestingly, apart from the transmembrane domain (48% identity), the highest level of conservation between CG2791 and CD98hc is seen within a short intracellular domain, which together with the transmembrane domain has been implicated in integrin binding (46% identity [32-35], while there is only 11% identity in the α -amylase-like region). Three residues in this latter region that are known to be required for amylase activity are present in a second mammalian heterodimeric amino acid transporter heavy chain, rBAT [36], although rBAT has not been shown to have amylase activity. At least one of these residues is absent and one is displaced in both CG2791 and CD98hc (Figure 2; [37]). rBAT also

Evolutionarily conserved properties of *Drosophila* CD98

shares much more limited homology with CG2791 and CD98hc in its transmembrane domain and intracellular C-terminus (Table 1). Critically CG2791 also contains the cysteine residue that covalently links CD98hc to the transporter light chains LAT1 and LAT2 (Cys¹⁰⁹ both in CG2791 and in human CD98hc [this is also present in rBAT]; [7]). Thus, the *CG2791* gene encodes the protein most closely related to CD98hc in flies.

Flies have several transporter light chains related to human LAT1 and LAT2.

System L-like transporters not only require a heavy chain, but also a light chain (LAT1 or LAT2) to transport amino acids [7]. BLAST homology searches with human LAT1 and LAT2 highlighted five fly amino acid transporters that share a significant degree of homology (see Figure 3A; percentage identities to LAT1 and LAT2 respectively are in parentheses): CG1607 (47% and 46%), Mnd ([21]; 49% and 44%), Genderblind (Gb or CG6070 [38]; 44% and 40%), JhI-21 (CG12317; 49% and 44%) and CG9413 (38% and 34%). These molecules share the conserved cysteine that covalently links the light chain to CD98hc (Cys¹⁶⁴ in LAT1), and also contain 12 conserved transmembrane domains in equivalent positions to the human LATs (Figure 3). This contrasts to the more distantly related *Drosophila* putative cationic amino acid transporters that are also identified in a BLAST search with the human LATs. Like human cationic AATs, these molecules appear to have 14 transmembrane domains, and except for CG5535, do not have a cysteine in an equivalent position to the cysteine found in all light chains that is implicated in binding to the heavy chains (Supplementary Figure 1). Thus, flies have up to five transporter light chains that could act as partners for CG2791.

CG2791 and at least two light chains are required for System L transport in *Drosophila*.

To test whether CG2791 and the putative CD98 light chains are involved in System L transport, we performed a series of transport assays in S2 cell culture, using RNAi to reduce expression of these molecules (Figure 4A). We selected the light chain-encoding genes *mnd* and *JhI-21*, since they had both been shown to be expressed in S2 cells [39], an observation that we subsequently confirmed (data not shown). Similar quantities of RNA were isolated from each sample, indicating that the RNAi treatment did not affect the overall viability of S2 cells. Relative levels of *CG2791*, *mnd*, *JhI-21* and a control ribosomal protein transcript (*RpL32*) were measured using quantitative real-time PCR (Figure 4B). In all knockdown experiments there was an approximately 80% selective reduction in levels of the knocked down transcript and 25-30% reduction in BCH-sensitive, sodium-independent leucine transport. By contrast, knockdown of the PAT transporter, CG13384, which we have found reduces proton-assisted transport of alanine by more than 50% (data not shown), had no effect on BCH-sensitive, sodium-independent leucine transport, confirming the specificity of our knockdown experiments (Supplementary Figure 2). Membrane proteins, including mammalian CD98hc, are often reported to be highly stable [40], and this may well explain why we only see a partial knockdown of CD98-like transport activity in our experiments. Interestingly, knockdown of *CG2791* or *mnd* produced a compensatory increase in *JhI-21* mRNA levels, which might also explain why transport was not more drastically reduced in these experiments. Overall, our data show that *CG2791*, Mnd and *JhI-21* play important roles in System L transport in S2 cells and therefore indicate that *CG2791* is indeed the functional orthologue of CD98hc. Based on this and subsequent data we therefore renamed this gene *cd98hc*.

Developmental regulation of System L transporters.

Methoprene, a synthetic analogue of Juvenile Hormone (JH), has previously been shown to regulate the levels of the *Drosophila* transcripts encoding *JhI-21* and *Mnd* in S2 cells [39] and *in vivo*. JH is involved

Evolutionarily conserved properties of *Drosophila* CD98

in the regulation of several key events in fly development and also responds to the nutritionally regulated insulin/insulin-like growth factor signaling (IIS) system [41]. We reasoned that JH-dependent changes in transporter levels might modulate System L transport and therefore tested whether methoprene altered System L transport activity as well as heterodimeric heavy and light chain expression in S2 cells. Treatment of *Drosophila* S2 cells with methoprene for six hours did not produce the substantial increases in *JhI-21* and *mnd* transcripts that have previously been reported [39]; however, a 55% ($n = 9$; $P < 0.002$) increase in *mnd* transcripts was observed. Transport analysis of S2 cells treated with methoprene revealed a 26% ($P < 0.02$) increase in System L transport. These findings suggest that altered expression of Mnd in flies in response to JH affects System L transport activity, an effect that is likely to be important during development and in response to IIS.

***Drosophila* CD98hc and light chain JhI-21 act together to mediate System L transport in *Xenopus* oocytes**

Our S2 cell RNAi knockdown data indicated that CG2791/CD98hc and the two light chains, Mnd and JhI-21, mediate System L transport in flies. However, some mammalian light chains do not bind to CD98hc to perform their functions. We therefore used the *Xenopus* oocyte system to analyse the effect of expressing *Drosophila* CD98hc and light chains individually and in combination with each other. In multiple experiments, injection of cRNA for each gene alone did not significantly increase System L uptake in oocytes (see Figure 5A). However, co-expression of CD98hc with JhI-21 did produce a strong increase in leucine uptake, the majority of which was inhibited by BCH (61%, see Figure 5B), demonstrating that these two molecules act as a heterodimer to mediate System L transport. In contrast, the combination of CD98hc and Mnd did not produce a detectable increase in leucine uptake, raising the possibility that Mnd requires another cofactor or heavy chain to mediate System L transport in S2 cells.

Inter-species sequence comparisons highlight a novel conserved hydrophobic interaction domain at the surface of CD98hc

The identification of a *Drosophila* CD98hc protein allowed us to identify orthologues of this molecule in other insect species (Figure 6A). Alignment of these sequences with their vertebrate counterparts not only confirmed the conservation of the transmembrane domain and intracellular putative integrin-binding domain, but also more clearly highlighted stretches of sequence within the extracellular α -amylase-like homology domain that are highly conserved over evolution. Mapping of these conserved regions on the X-ray crystal structure of the dimeric human CD98hc (PDB ID 2DH3) [37] was performed with this sequence alignment using the WHISCY server [29]. Figure 6B shows a surface representation of the CD98hc molecule as a dimer [37], painted according to the degree of surface conservation, from red (exposed side chain, most conserved) to blue (surface residues not conserved or any residue with a buried side chain). This analysis revealed that residues K462, S465, H468, G469, L513 and P514 in *Drosophila* CD98hc are conserved across the human, mouse, zebrafish and *Drosophila* sequences (there are conservative changes in some of these positions in other insect and nematode sequences), and N489 is in common between sequences except zebrafish. These residues (depicted in Supplementary Figure 2 and marked on Figure 6A) map to the dimer interface in the crystal structure of the extracellular domain of human CD98hc. Residues that co-ordinate a zinc ion at the human CD98hc dimer interface are not conserved in *Drosophila*. Interestingly these residues were not found to be essential for homodimerisation of the human molecule *in vivo* [37]. Mapping of sequence conservation onto the human CD98hc dimer also identified a second conserved surface region, distal to the dimerisation interface of the CD98hc extracellular domain, and predicted to lie just above the plasma membrane (Figure 6B). *Drosophila* CD98hc residues L187, P189 and Y191 form a small surface-exposed hydrophobic spot that is conserved

Evolutionarily conserved properties of *Drosophila* CD98

across all CD98hc sequences (except the worm CD98hc, which has a charged group in the position of the proline residue). Such hydrophobic exposed regions often represent areas of potential protein-protein interaction, suggesting that this part of CD98hc may be involved in binding to either extracellular or membrane-bound proteins.

DISCUSSION

The mammalian membrane protein CD98hc has been implicated in a variety of cellular functions, including System L amino acid transport, which provides the major route for the sodium-independent uptake of leucine into cells. Here, contrary to recent reports [24], we show that the fly CG2791 protein shares extensive homology to human CD98hc and using RNAi knockdown and expression in the heterologous *Xenopus* oocyte system, we demonstrate that it is the functional orthologue of this molecule.

***Drosophila* CG2791 is structurally related to mammalian CD98hc**

In mammals, System L transport is mediated by heterodimeric amino acid transporters, which contain the heavy chain, CD98hc. The protein encoded by the *Drosophila* CG30359 gene was recently suggested to encode the closest relative of CD98hc, because it shares the greatest similarity within the large extracellular α -amylase-like domain in this molecule [24]. However, unlike the CG2791 protein, it has no conserved intracellular sequences, nor even a transmembrane domain, which is vital for function in mammalian CD98 transporters. Significantly, CG2791 lacks at least one of the three critical residues required for α -amylase function (Figure 2A), like orthologues of CD98hc from all other species (Figure 6A). In contrast, these residues are present in CG30359. Thus, on the basis of sequence similarity alone, CG2791 is the CD98hc orthologue, a conclusion that is supported by our functional data. We speculate that CG30359 encodes a secreted amylase and thus we have not studied this molecule further in the experiments reported here.

Interestingly, apart from the transmembrane region, *Drosophila* and human CD98hc share the highest levels of homology in their intracellular domains, within a region that contains at least part of the putative integrin binding domain [23]. This region is highly conserved in CD98hc orthologues from a range of species, including mouse, zebrafish, mosquito and honey bee (Figure 6A). It will now be interesting to test whether *Drosophila* β -integrin, which does not bind to mammalian CD98hc [23], does interact with fly CD98hc, or whether the sequence conservation in this CD98hc intracellular domain is required for conserved interactions with other, as yet unidentified, molecules.

***Drosophila* CD98 mediates System L transport like its mammalian counterpart**

In mammals, in addition to CD98hc, other heterodimeric transporters have been characterised, namely rBAT [7, 36] and the more recently discovered Collectrin [42]. These molecules have been shown to have a distinct preference for specific light chains, making it important to clarify which of the fly light chains we have studied form heterodimers with CG2791. Interestingly, b⁰+AT, the light chain that binds rBAT in mammals, shows greatest homology to CG9413, raising the possibility that this might be the light chain for an rBAT-related molecule in flies. However, our sequence searches have not identified a second *Drosophila* heavy chain in addition to CD98hc, suggesting that rBAT is not present in flies. Nevertheless, mammals have three 'orphan' light chains, ASC2, AGT1 [7] and arpAT [43], and such light chains might also exist in flies.

Evolutionarily conserved properties of *Drosophila* CD98

RNAi knockdowns of *cd98hc*, *JhI-21* and *mnd* mRNAs all reduced System L transport. However, we were unable to reduce System L transport by more than 25-30%, despite producing more than 80% knockdown of mRNAs in our experiments. This is likely to be at least partly explained by the high stability of CD98 heavy and light chain proteins over the time course of this experiment, since the mammalian counterparts of these molecules appear to persist for many days within the cell [40]. However, our *Xenopus* oocyte data also reveal that Mnd may mediate System L transport via a mechanism that does not involve CD98hc. With two independent System L transport system, knockdown experiments could produce only a partial loss of System L uptake in S2 cells.

Co-expression of CD98hc and JhI-21 in *Xenopus* oocytes clearly demonstrated that these molecules act together as a System L heterodimeric transporter. By contrast, CD98hc and Mnd do not show a similar interaction, suggesting that another unidentified heavy chain or an additional cofactor is required for Mnd to transport leucine in S2 cells. Co-injection of 50 ng total RNA extracted from S2 cells with CD98hc and Mnd, or with any other combinations of light chain and/or heavy chain failed to increase System L transport in oocytes. However, this may merely reflect the fact that there is insufficient mRNA encoding the missing factor(s) in S2 cells to produce a significant level of functional transporter in oocytes. We believe that the most likely explanation for our results is that Mnd interacts with another as yet unidentified heavy chain in flies, although it is always possible that the *mnd* cRNA is not being properly translated in the *Xenopus* oocyte system.

CD98hc and System L transport is developmentally regulated in *Drosophila*

The identification of a CD98hc orthologue in flies has allowed us to initiate an analysis of this molecule's physiological and developmental control, and the regulation of System L transport. Genome-wide studies of gene expression in the fly reveal that *cd98hc* is very highly expressed in early embryos and then transcribed in a broad range of embryonic tissues (<http://www.fruitfly.org/cgi-bin/ex/insitu.pl>). Expression persists throughout development into adulthood, where *cd98hc* transcripts are observed in all parts of the animal ([44]; <http://flyatlas.org/>).

We have also found that developmental regulation of the Mnd light chain by the JH analogue methoprene may modulate System L transport activity. Interestingly, it is known that viable flies mutant for the Insulin Receptor gene (*InR*), which have an extended lifespan, have reduced JH levels [41]. Furthermore, *cd98hc* expression is increased when yeast is added to the diet [45], supporting the link between nutritionally regulated IIS and System L transport components.

The increased lifespan of mutant *InR* flies is reversed by administration of methoprene [41], suggesting that decreased endocrine signalling by JH is involved in the longevity phenotype. Our data may now provide an explanation for this effect, since the reduced System L transport produced by lowering JH levels could decrease TOR signalling in target cells, an effect reversed by methoprene. This might be particularly relevant in the fat body, which plays an important role in the control of lifespan [46] and IIS-dependent growth [47]. In response to TOR-mediated signals, this endocrine gland regulates the sensitivity of other tissues to insulin-like molecules [20]. It seems likely that modulation of System L transport by JH plays a role in this process, an idea that is strongly supported by the fact that the light chain-encoding gene *mnd* controls growth via its effects on the fat body [21].

A new conserved structural feature in the CD98hc extracellular domain

Sequence comparison between CD98hc proteins from invertebrates and vertebrates has highlighted features of the molecule that are conserved from worm to human, and which are therefore presumably required for fundamental CD98hc functions in all higher eukaryotes. In addition to identifying the most conserved residues within the CD98hc dimerisation interface, this analysis also highlighted a second

Evolutionarily conserved properties of *Drosophila* CD98

small hydrophobic conserved surface, on the other side of the molecule that lies just above the plasma membrane. Interestingly, in amylases such as PDB ID 1UOK (an oligo-1,6-glucosidase of *Bacillus cereus*), the central proline residue in this region is replaced by a polar residue: this surface-conserved hydrophobic spot is therefore likely to represent a protein-protein interaction region that is not shared by α -amylases. It will now be important to test whether disruption of this sequence alters the properties of this molecule.

Concluding remarks

The System L transporters and their components that we have identified are under significant developmental and nutritional control in *Drosophila*. In light of System L's role in leucine transport, we propose that this regulation of transport activity may in turn modulate the sensitivity of cells to nutrient-dependent signals mediated by IIS and TOR, particularly in endocrine tissues like the fat body. In addition to the important role for CD98hc in System L transport, this molecule has also been implicated in many other distinct aspects of cell biology. The identification of CD98hc in flies should now allow the physiological significance of a number of these other functions to be fully tested *in vivo*.

Acknowledgements

We thank Clive Wilson for helpful discussion during this work and Susan Lea for access to computing facilities.

Funding

This work was supported by Cancer Research UK [grant numbers C7713/A6174, C19591/A6181 and C19591/A9093].

REFERENCES

1. Christensen, H.N. (1990) Role of amino acid transport and countertransport in nutrition and metabolism. *Physiol. Rev.* **70**, 43-77.
2. Christensen, H.N., Albritton, L.M., Kakuda, D.K. and MacLeod, C.L. (1994) Gene-product designations for amino acid transporters. *J. Exp. Biol.* **196**, 51-57.
3. Hediger, M.A., Romero, M.F., Peng, J.B., Rolfs, A., Takanaga, H. and Bruford, E.A. (2004) The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. *Pflugers Arch* **447**, 465-468.
4. Hyde, R., Taylor, P.M. and Hundal, H.S. (2003) Amino acid transporters: roles in amino acid sensing and signalling in animal cells. *Biochem. J.* **373**, 1-18.
5. Hundal, H.S. and Taylor, P.M. (2009) Amino acid transporters: gatekeepers of nutrient exchange and regulators of nutrient signalling. *Am J. Physiol. Endocrinol. Metab.* Epub.
6. Verrey, F. (2003) System L: heteromeric exchangers of large, neutral amino acids involved in directional transport. *Pflugers Arch.* **445**, 529-533.

Evolutionarily conserved properties of *Drosophila* CD98

7. Palacin, M., Nunes, V., Font-Llitjos, M., Jimenez-Vidal, M., Fort, J., Gasol, E., Pineda, M., Feliubadalo, L., Chillaron, J. and Zorzano, A. (2005) The genetics of heteromeric amino acid transporters. *Physiol.* **20**, 112-124.
8. Wagner, C.A., Lang, F. and Broer, S. (2001) Function and structure of heterodimeric amino acid transporters. *Am. J. Physiol. Cell Physiol.* **281**, C1077-1093.
9. Deves, R. and Boyd, C.A.R. (2000) Surface antigen CD98 (4F2): not a single membrane protein, but a family of proteins with multiple functions. *J. Memb. Biol.* **173**, 165-177.
10. Verrey, F., Closs, E.I., Wagner, C.A., Palacin, M., Endou, H. and Kanai, Y. (2004) CATs and HATs: the SLC7 family of amino acid transporters. *Pflugers Arch.* **447**, 532-542.
11. Nakamura, E., Sato, M., Yang, H., Miyagawa, F., Harasaki, M., Tomita, K., Matsuoka, S., Noma, A., Iwai, K. and Minato, N. (1999) 4F2 (CD98) heavy chain is associated covalently with an amino acid transporter and controls intracellular trafficking and membrane topology of 4F2 heterodimer. *J. Biol. Chem.* **274**, 3009-3016.
12. Fuchs, B.C. and Bode, B.P. (2005) Amino acid transporters ASCT2 and LAT1 in cancer: partners in crime? *Seminars in Cancer Biol.* **15**, 254-266.
13. Nicklin, P., Bergman, P., Zhang, B., Triantafellow, E., Wang, H., Nyfeler, B., Yang, H., Hild, M., Kung, C., Wilson, C., Myer, V.E., MacKeigan, J.P., Porter, J.A., Wang, Y.K., Cantley, L.C., Finan, P.M., Murphy, L.O. (2009) Bidirectional transport of amino acids regulates mTOR and autophagy. *Cell* **136**, 521-34.
14. Cohen, A. and Hall, M.N. An amino acid shuffle activates mTORC1. *Cell* **136**, 399-400.
15. Dann, S.G. and Thomas, G. (2006) The amino acid sensitive TOR pathway from yeast to mammals. *FEBS Lett.* **580**, 2821-2829.
16. Easton, J.B. and Houghton, P.J. (2006) mTOR and cancer therapy. *Oncogene* **25**, 6436-6446.
17. Goberdhan, D.C. and Wilson, C. (2003) The functions of insulin signaling: size isn't everything, even in *Drosophila*. *Differentiation* **71**, 375-397.
18. Tsumura, H., Suzuki, N., Saito, H., Kawano, M., Otake, S., Kozuka, Y., Komada, H., Tsurudome, M. and Ito, Y. (2003) The targeted disruption of the *CD98* gene results in embryonic lethality. *Biochem. Biophys. Res. Commun.* **308**, 847-851.
19. Goberdhan, D.C., Ögmundsdóttir, M.H., Kazi, S., Reynolds, B., Wilson, C. and Boyd, C.A.R. (2009) Amino acid sensing and mTOR regulation: inside or out? *Biochem. Soc. Trans.* **37**, 248-52.
20. Colombani, J., Raisin, S., Pantalacci, S., Radimerski, T., Montagne, J. and Leopold, P. (2003) A nutrient sensor mechanism controls *Drosophila* growth. *Cell* **114**, 739-749.
21. Martin, J.F., Hersperger, E., Simcox, A. and Shearn, A. (2000) *minidiscs* encodes a putative amino acid transporter subunit required non-autonomously for imaginal cell proliferation. *Mech. Dev.* **92**, 155-167.
22. Goberdhan, D.C., Meredith, D., Boyd, C.A.R. and Wilson, C. (2005) PAT-related amino acid transporters regulate growth via a novel mechanism that does not require bulk transport of amino acids. *Development* **132**, 2365-2375.
23. Prager, G.W., Feral, C.C., Kim, C., Han, J. and Ginsberg, M.H. (2007) CD98hc (SLC3A2) interaction with the integrin β subunit cytoplasmic domain mediates adhesive signaling. *J. Biol. Chem.* **282**, 24477-24484.
24. Uinuk-Ool, T., Mayer, W.E., Sato, A., Dongak, R., Cooper, M.D. and Klein, J. (2002) Lamprey lymphocyte-like cells express homologs of genes involved in immunologically relevant activities of mammalian lymphocytes. *PNAS* **99**, 14356-14361.
25. Lister, N., Sykes, A.P., Bailey, P.D., Boyd, C.A.R., and Bronk, J.R. (1995) Peptide transport and hydrolysis in isolated loops of small intestine: effects of stereospecificity. *J. Physiol.* **484**, 173-182.

Evolutionarily conserved properties of *Drosophila* CD98

26. Temple, C.S., Stewart, A.K., Meredith, D., Lister, N.A., Morgan, K.M., Collier, I.D., Vaughan-Jones, R.D., Boyd, C.A.R., Bailey, P.D. and Bronk, J.R. (1998) Peptide mimics as substrates for the intestinal peptide transporter. *J. Biol. Chem.* **273**, 20-22.
27. Spalding, J.B. and Lammers, P.J. (2004) BLAST Filter and GraphAlign: rule-based formation and analysis of sets of related DNA and protein sequences. *Nucleic Acids Res.* **32**, W26-32.
28. Gasol E., Jiménez-Vidal M, Chillarón J., Zorzano A., Palacín M. (2004) Membrane topology of system X(c)⁻ light subunit reveals a re-entrant loop with substrate-restricted accessibility. *J. Biol. Chem.* **279**, 31228-36.
29. de Vries, S.J., van Dijk, A.D. and Bonvin, A.M. (2006) WHISCY: what information does surface conservation yield? Application to data-driven docking. *Proteins* **63**, 479-489.
30. DeLano, W.L. The PyMOL Molecular Graphics System (2002) on World Wide Web <http://www.pymol.org>
31. Janecek, S., Svensson, B. and Henrissat, B. (1997) Domain evolution in the α -amylase family. *J. Mol. Evol.* **45**, 322-331.
32. Fenczik, C.A., Zent, R., Dellos, M., Calderwood, D.A., Satriano, J., Kelly, C. and Ginsberg, M.H. (2001) Distinct domains of CD98hc regulate integrins and amino acid transport. *J. Biol. Chem.* **276**, 8746-8752.
33. Zent, R., Fenczik, C.A., Calderwood, D.A., Liu, S., Dellos, M. and Ginsberg, M.H. (2000) Class- and splice variant-specific association of CD98 with integrin β cytoplasmic domains. *J. Biol. Chem.* **275**, 5059-5064.
34. Cai, S., Bulus, N., Fonseca-Siesser, P.M., Chen, D., Hanks, S.K., Pozzi, A. and Zent, R. (2005) CD98 modulates integrin $\beta 1$ function in polarized epithelial cells. *J. Cell Sci.* **118**, 889-899.
35. Kabir-Salmani, M., Fukuda, M.N., Kanai-Azuma, M., Ahmed, N., Shiokawa, S., Akimoto, Y., Sakai, K., Nagamori, S., Kanai, Y., Sugihara, K., *et al.* (2008) The membrane-spanning domain of CD98 heavy chain promotes $\alpha_v\beta_3$ integrin signals in human extravillous trophoblasts. *Mol. Endocrinol.* **22**, 707-715.
36. Bertran, J., Werner, A., Moore, M.L., Stange, G., Markovich, D., Biber, J., Testar, X., Zorzano, A., Palacín, M. and Murer, H. (1992) Expression cloning of a cDNA from rabbit kidney cortex that induces a single transport system for cystine and dibasic and neutral amino acids. *PNAS* **89**, 5601-5605.
37. Fort, J., de la Ballina, L.R., Burghardt, H.E., Ferrer-Costa, C., Turnay, J., Ferrer-Orta, C., Uson, I., Zorzano, A., Fernandez-Recio, J., Orozco, M., *et al.* (2007) The Structure of Human 4F2hc Ectodomain Provides a Model for Homodimerization and Electrostatic Interaction with Plasma Membrane. *J. Biol. Chem.* **282**, 31444-31452.
38. Augustin H., Grosjean Y., Chen K., Sheng Q., Featherstone D.E. (2007) Nonvesicular release of glutamate by glial xCT transporters suppresses glutamate receptor clustering *in vivo*. *J. Neurosci.* **27**, 111-23.38.
39. Dubrovsky, E.B., Dubrovskaya, V.A. and Berger, E.M. (2002) Juvenile hormone signaling during oogenesis in *Drosophila melanogaster*. *Insect Biochem. Mol. Biol.* **32**, 1555-1565.
40. Asano, S., Kameyama, M., Oura, A., Morisato, A., Sakai, H., Tabuchi, Y., Chairoungdua, A., Endou, H. and Kanai, Y. (2007) L-type amino acid transporter-1 expressed in human astrocytomas, U343MGa. *Biol. Pharma. Bulletin* **30**, 415-422.
41. Tatar, M., Kopelman, A., Epstein, D., Tu, M.P., Yin, C.M. and Garofalo, R.S. (2001) A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science* **292**, 107-110.
42. Danilczyk, U., Sarao, R., Remy, C., Benabbas, C., Stange, G., Richter, A., Arya, S., Pospisilik, J.A., Singer, D., Camargo, S.M.R., *et al.* (2006) Essential role for collectrin in renal amino acid transport. *Nature* **444**, 1088-1091.
43. Fernandez, E., Torrents, D., Zorzano, A., Palacín, M. and Chillaron, J. (2005) Identification and functional characterization of a novel low affinity aromatic-preferring amino acid transporter

Evolutionarily conserved properties of *Drosophila* CD98

- (arpAT). One of the few proteins silenced during primate evolution. *J. Biol. Chem.* **280**, 19364-19372.
44. Chintapalli, V.R., Wang, J. and Dow, J.A. (2007) Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nature Genet.* **39**, 715-720.
 45. Gershman, B., Puig, O., Hang, L., Peitzsch, R.M., Tatar, M. and Garofalo, R.S. (2007) High-resolution dynamics of the transcriptional response to nutrition in *Drosophila*: a key role for dFOXO. *Physiol. Genomics* **29**, 24-34.
 46. Hwangbo, D.S., Gershman, B., Tu, M.P., Palmer, M. and Tatar, M. (2004) *Drosophila* dFOXO controls lifespan and regulates insulin signalling in brain and fat body. *Nature* **429**, 562-566.
 47. Hennig, K.M., Colombani, J. and Neufeld, T.P. (2006) TOR coordinates bulk and targeted endocytosis in the *Drosophila melanogaster* fat body to regulate cell growth. *J. Biol. Chem.* **173**, 963-974.

FIGURE LEGENDS

Figure 1. System L amino acid transporters are found in *Drosophila* S2 cells. (A) The uptake of ^3H L-Leucine in sodium-free conditions into S2 cells, which were untreated or exposed to either a large excess of 10 mM BCH (BCH) or a large excess of unlabelled 10 mM leucine (Leu), was measured (nmols leucine per mg protein per min). Mediated System L transport corresponds to additional transport above the dotted line. More than 90% of mediated transport is BCH-sensitive. (B) Leucine efflux from *Drosophila* S2 cells. The graph shows the intracellular ^3H L-Leucine content as a function of time, after cells loaded with ^3H -L-leucine were transferred at zero time to an amino acid-free medium (control; grey data points and dashed curve; efflux rate constant (k [mean \pm s.d.] = $0.19 \pm 0.02 \text{ min}^{-1}$), to a medium with 10 mM leucine (filled circle data points and black curve; $k = 0.59 \pm 0.03 \text{ min}^{-1}$) or to a medium with 10 mM BCH (unfilled data points and dotted curve; $k = 0.54 \pm 0.11 \text{ min}^{-1}$). Leucine efflux is trans-stimulated by leucine and BCH as expected for System L transport.

Figure 2. CG2791 is the fly orthologue of CD98hc. (A) An alignment of the predicted *Drosophila* CG2791 protein with the most closely related isoform of human CD98hc (isoform f [denoted as CD98hc]) and human rBAT sequences. Arrowhead indicates the conserved Cys¹⁰⁹ residue required for covalent linkage to a light chain to form a functional heterodimeric amino acid transporter. The three residues required for amylase catalytic activity (marked with an asterisk) that are conserved in rBAT, but not all present in either CD98hc or CG2791, are boxed within each gene. Identical residues are shaded in black and conservative changes in grey. (B) Percentage identity of conserved sequence in CG2791 aligned against CD98hc. In both Panels A and B, domain architecture is indicated by bars above both the sequence data and graph: the conserved intracellular domain implicated in integrin binding (black) and transmembrane domain (stripes) show much greater conservation than does the α -amylase-like domain (grey).

Figure 3. Orthologues of the System L-like transporter light chains in *Drosophila*. (A) Alignment of the five *Drosophila* molecules [CG1607, Mnd (CG3297), Gb (CG6070), JhI-21 (CG12317), CG9413] with sequence similarity to human LAT1 transporter light chain. The 12 predicted transmembrane domains are marked with striped bars above the sequence. The arrowhead denotes the conserved cysteine required in LAT1 and LAT2 to form a disulphide bridge to CD98hc. (B) Alignment of hydropathy plots for the five putative fly light chains and seven human light chains <http://genome.ucsc.edu/cgi-bin/pbGateway>. Arrow indicates the position of a cysteine present in a conserved position, which was used to align the sequences of all of the light chains shown in this figure and also the cationic amino acid transporters in Supplementary Figure 1, to each other. Each set of sequences is ordered according to similarity to human LAT1. Note the remarkable spatial conservation of TM domains. Vertical dotted lines indicate the positions of the TM domains based on the empirical determination from human xCT [38]. The position of the conserved cysteine required for covalent linkage to heavy chains is indicated by a labeled arrowhead. This is conserved in all human and fly light chains, but only one out of nine cationic amino acid transporters.

Evolutionarily conserved properties of *Drosophila* CD98

Figure 4. CG2791, Mnd and JhI-21 are involved in System L transport. (A) BCH-sensitive mediated uptake of ^3H L-Leucine in sodium-free conditions in control cells and in cells treated with RNAi to knock down *CG2791/cd98hc*, *mnd* and *JhI-21* transcripts. In all knockdown experiments there was a significant reduction of 25-30% compared to untreated cells. (B) Graphs show levels of *CG2791/cd98hc*, *mnd*, *JhI-21* and *RpL32* transcripts relative to untreated cells in multiple experiments, as measured by Q-RT-PCR. There was a significant reduction of roughly 80% for each targeted transcript. For *cd98hc* and *mnd* knockdown, *JhI-21* transcript levels were significantly increased. Experiments were done in triplicate in four different experiments.

Figure 5. CD98hc acts with JhI-21 to mediate System L transport. (A) Injection of cRNA for *Drosophila cd98hc*, *mnd* and *JhI-21* alone in *Xenopus* oocytes did not significantly promote uptake of ^3H L-Leucine when compared to water alone (measured in pmol leucine per oocyte per hour). (B) However, co-injection of *Drosophila cd98hc* heavy chain with *JhI-21* light chain cRNA, strongly enhanced ^3H L-Leucine uptake relative to the water-injected control eggs. Comparison with the level of ^3H L-Leucine uptake in the presence of BCH (relative to water-injected controls) suggests that about 60% of the uptake is BCH-inhibited, consistent with CD98hc and JhI-21 forming a System L heterodimeric transporter. In contrast, co-injection of *cd98hc* with *mnd* cRNAs did not significantly promote System L transport when compared to water alone controls.

Figure 6. Cross-species sequence comparison and sequence conservation mapping on the human CD98hc extracellular domain structure. (A) Sequence alignment of CD98hc orthologues from human (isoform f), mouse, zebrafish, mosquito, fly, honey bee and nematode worm. The conserved transmembrane domain (striped bar), conserved intracellular domain implicated in integrin binding (black bar) and the α -amylase-like domain (grey bar) are marked, as is the conserved cysteine that covalently links CD98hc to CD98 light chains (arrowhead), and the three residues in the extracellular domain that are required for α -amylase catalytic activity (marked by asterices), only some of which are conserved in these molecules (boxed in black; typically only one of these three residues is conserved in identity and position; see also Figure 2A). In *Drosophila* CD98hc, residues K462, S465, H468, G469, L513, P514 and N489 at the CD98hc dimer interface and residues L187, P189, Y191 in a second surface hydrophobic patch have been highlighted from the mapping analysis below as conserved residues that are exposed at the CD98hc surface (red boxes). (B) A surface representation of the human CD98hc extracellular domain dimer from the crystal structure in PDB ID 2DH3 [37] viewed from the cellular side ("bottom view"). The surface is painted from red to blue across orange, yellow and green with decreasing degrees of conservation (residues with buried side chains are all painted blue). Two main sites of surface-exposed, conserved residues are seen in red: the central one localises to the dimer interface, where the N-termini of the two monomers meet. It is likely to interact with the transmembrane helices of the two monomers. The second conserved surface hydrophobic spot is located on each monomer at the opposite end to the dimer interface (arrows), a few Angstroms above the plasma membrane. The image was generated with the WHISCY server [29] and PyMOL [30].

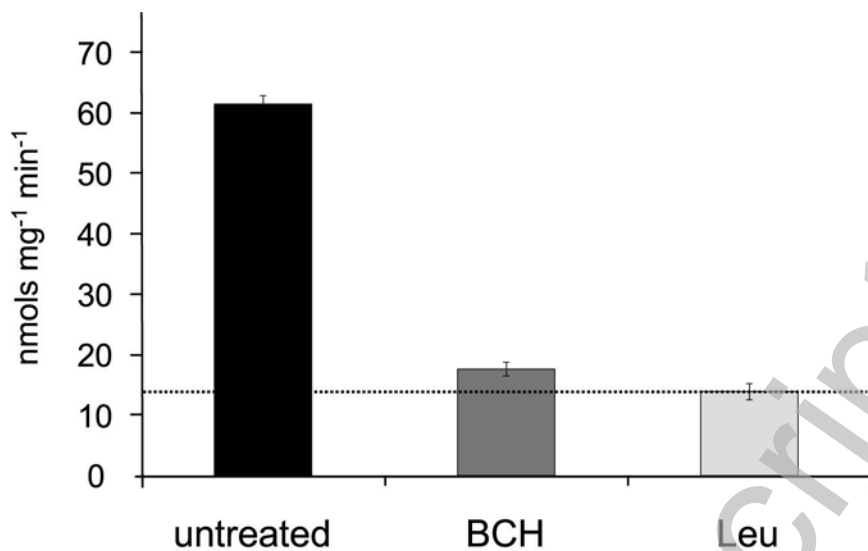
Evolutionarily conserved properties of *Drosophila* CD98

Region	human CD98hc	human rBAT
Novel intracellular domain	46%	21%
Transmembrane domain	48%	26%
Amylase-like domain	11%	13%
Total	17%	15%

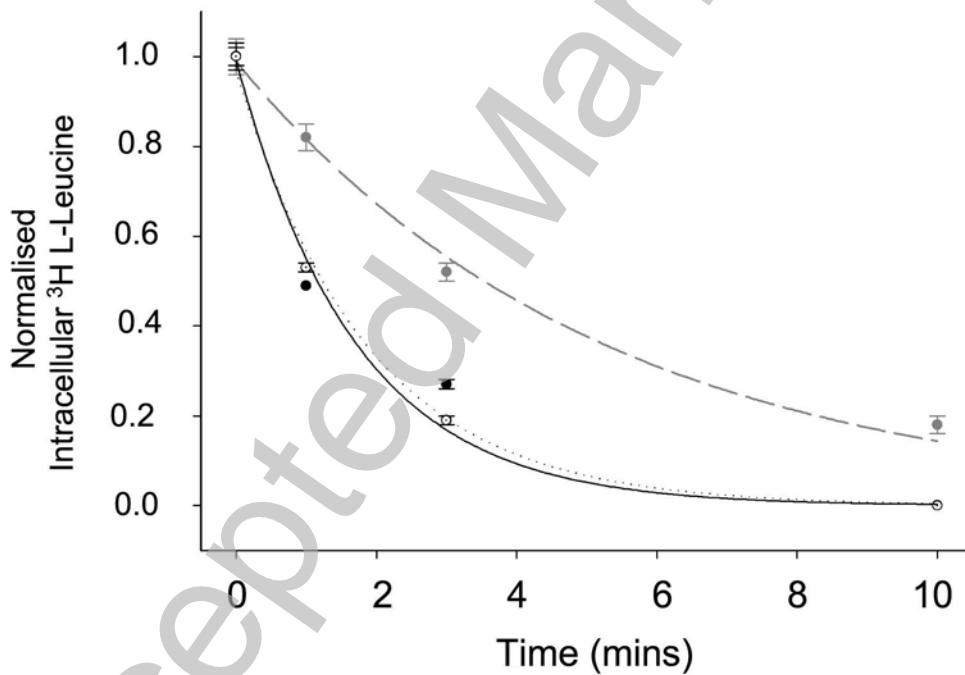
Table 1. CG2791 and CD98hc have multiple conserved structural domains. Percentage identities between the three major structural domains conserved between *Drosophila* CG2791 and the human heavy chain CD98hc. While there are comparable similarities between the amylase domain of CG2791, and those of CD98hc and the second human heavy chain, rBAT [7], CG2791 and CD98hc are much more related similarity to each other in their extracellular domains and extended intracellular domain implicated in integrin binding [23].

Accepted Manuscript

A



B



THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20082198

Accepted Manuscript

A

```

CD98hc      1 MSQDT-----EVDMKEVELN- ELEPEKOPMNAAS GAAMSLAGAEKNGLVKI KVAEDEAEAAA AAKFTGLSKEELKVA GSPGWVRI RWALLLLE
CG2791      1 MVQDEKAAHADGAEQMLGSN---GVEDKLAERRDEVKFKGDHONGDAKIDIGTVN--GCKPAETGMSKEELMKYANDPFVWRLRWTFVFC
rBAT       1 MAEDKSKRDSIEMSMKGCQTNNGFVHNEDELEQTPDPGSDTNLKHSS--TRGILGSEQEPDFKGVQPYAGMPKEVLFQFSGQARYRIRPRELFWLT

CD98hc     89 WLQWLGMLAGAVVIVRAPRCRELPAQKWWHTGALYRI G--DLQAFQGHGAGNLAGLKGRLDYLSLTKVKGGLVLP GPHKXNQKDDVAQ-----TD
CG2791     89 WAIWVGMLVGAIIIGAPKCAAPQPIPWYKRGPHAKFA SVETCRPE DVQVAKKLVSA GAIYELLPALTYDVKKPEVEEKI KHLVALYQGS DTRV
rBAT     94 VASVLVLTIAATIAI IALSPKCC---LDWVQEGPMYQLYPRSFKDSNKDGNGLDKGIQDKLDYITALNIKITVWITSFYKSSLKDFRYG---VEVD

CD98hc    176 LIQIDPNFGSKED--FDSLLQSAKKKSI RVLDLTPNYR-----GNSWFFTOVD--
CG2791    184 IEDLTINYAANKNSQLVQDAIANPEKBSAFVWVSGGKALPNN-----WLKVGGRS AWEKVGDN
rBAT    180 FREVDPIFGTMEDE--FENQVAAIHDRLGLKLTIDFI PNHTSDKHIWFQLSRTRTGKYTDYIIMHDCTHENGKTI PPNWLVSVYGNSSVHDFEVRNQ

CD98hc    224 -----T VATKVKDALEFWLQAGVDGFGVRFIENLKDASFLAEWQN-----ITKGFSE-
CG2791    242 YVLSQFEDGYYDLKMNSTIRNEFGAVLKKHLVALGVRGFRKNTKFFALFDSL EDEQISS-----SPKDFSLG
rBAT    273 CYFHQFMKEQDNLNFRNPDVQEIKELIRFWLTKGVDGFSLDAVKFLEAKHLRDEI QVNTQIPDTVTQYSELYHDFTTTQVGMHDLVRSFRQT

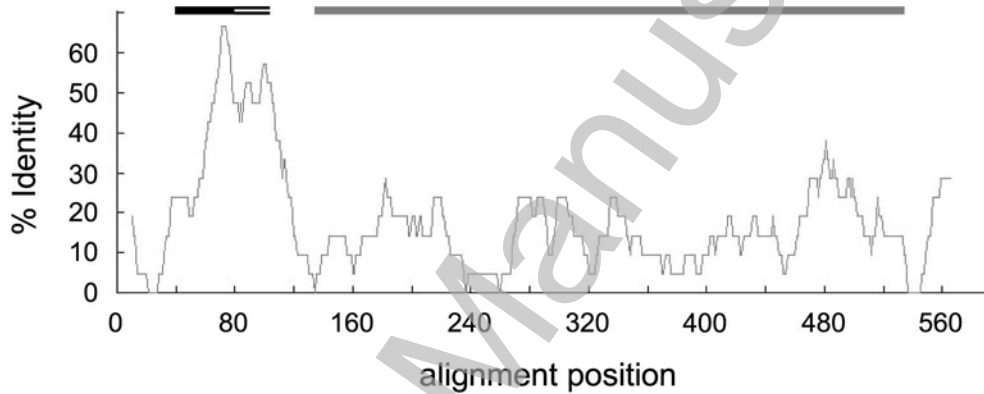
CD98hc    272 -----DRLLIAGTINSDLQQLLSLLESNKDLLTSSYLS D-----SGSTGEHTKSLVITQYLN-ATGNRWCSWSLS--QARLLTSFLPAQLL
CG2791    310 -----PNRYGFYSHRQTTFLSGLGDVLYDYLSTVKN-----ASDEAFSSVAEDVMPQVYQLSGDSGAYGLDLP MYGNFVKVLSK
rBAT    368 MDQYSTPGRYRFMGTEAYAESIDRTVMYGLPFIQEADEFPNYLSMLDVTVSNVYEVITSMENMPEGKMPNMMGGPSSRLLSRIGNQYV

CD98hc    350 RLYQLMFTLPGTVPVFSYQDEI GLDAAALP-----GOPMEAPVWLWDESSFRDIPGAVS-----ANMTVKGQSEDFGSLLSLFRRLSD
CG2791    385 SKPDTSLQKELDNTLAQSGTDSWLNWLA D-----VYVDTTPQDPSALAMFLSLLPVGF-----VAVDAVAYQNVTEETVRIQITS
rBAT    463 NYMNMMLFTLPGTPTIYVGEETGMGNI VAAANLNESYDINILRSKSPQWQNSNAGFSEASNTWLPTNSDYHTVNVVDVQKTQDRSAKLVQDLSL

CD98hc    428 QRSKERSLLHGFDFHAFSAGPELEFSYIRHWQ-ONERFLVVLNFGDVGLSAGLQASD---LPASASLPAKADLLLSIQPGREEGSPLELERIKLEP
CG2791    460 LR-KTASVMHGNLNLYQADP-LVAFSRIKS-GNPGYFVIFNPTLELPQASNFTIPDNL PDKMTVSYFSEQYNNMDSGKVAHAGRVNKKDKVAF
rBAT    558 LHANELLNRRQWFCHLRNDSHYVYVTRREL DGI DRI FIVVLNFGESTLNL-LHN-----MISGLPAKIRIRLSTIN-SADKGSKVDTS GIFL DK

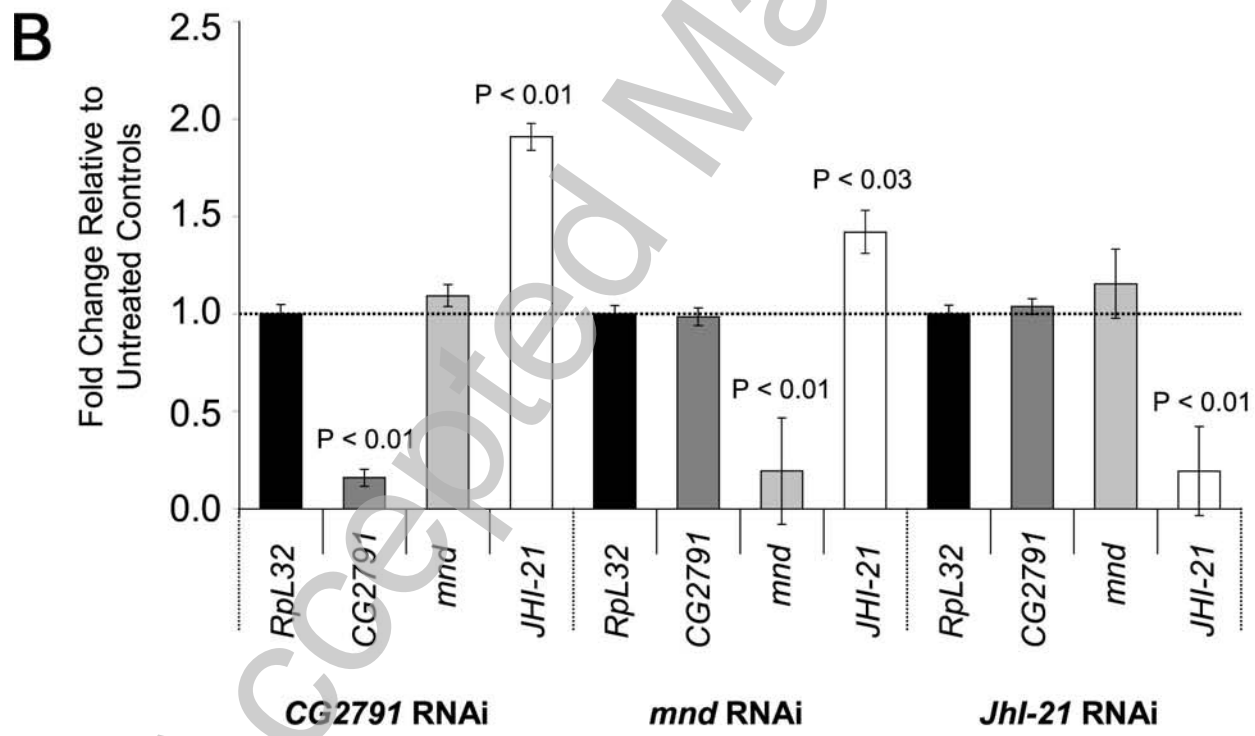
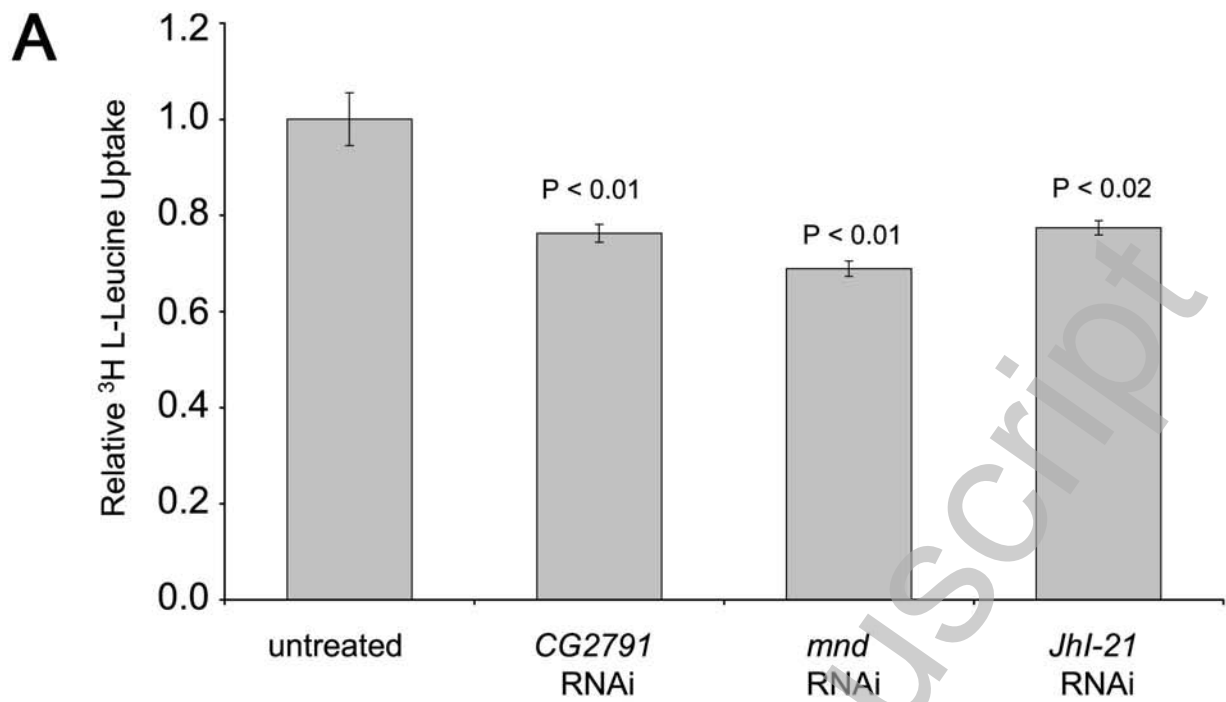
CD98hc    518 HEGLLIRFPYAA-----
CG2791    552 HSSITITVYPVKA E-----
rBAT    643 GEGLI FEHNTKNL LHRQTAFDRDCFVSNRACYSVVLNLI LYTS C
    
```

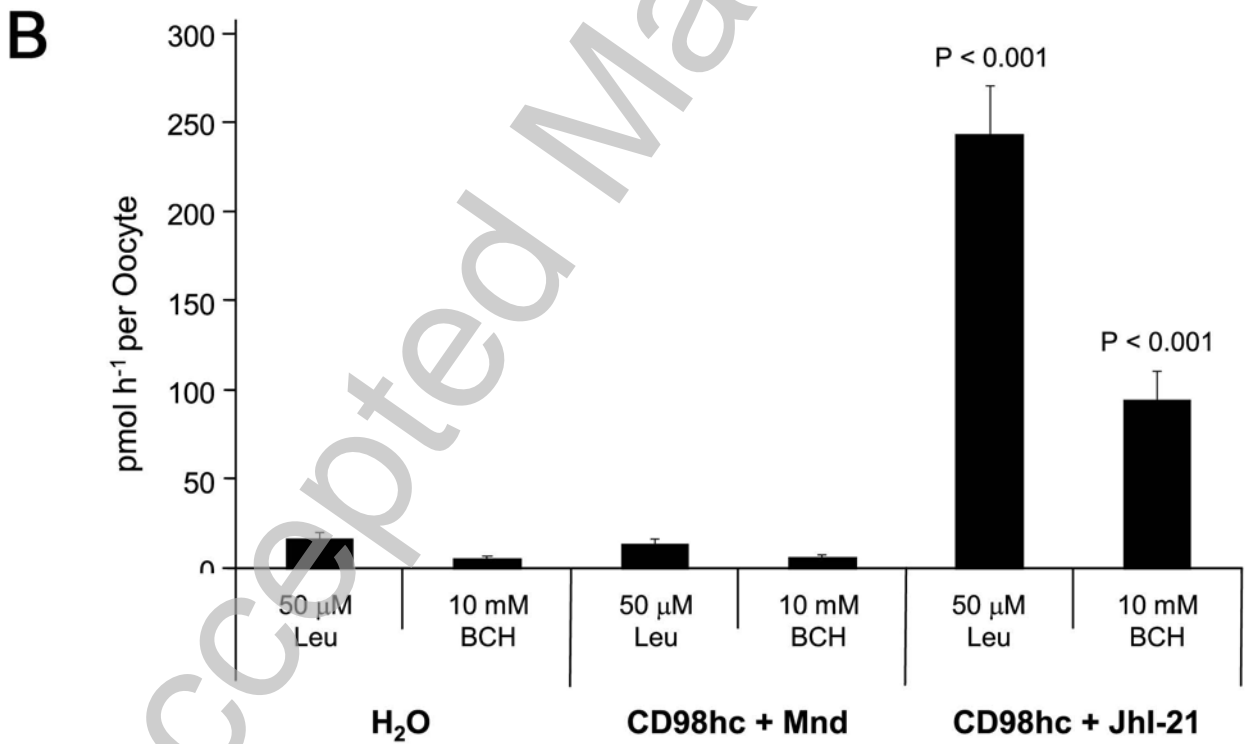
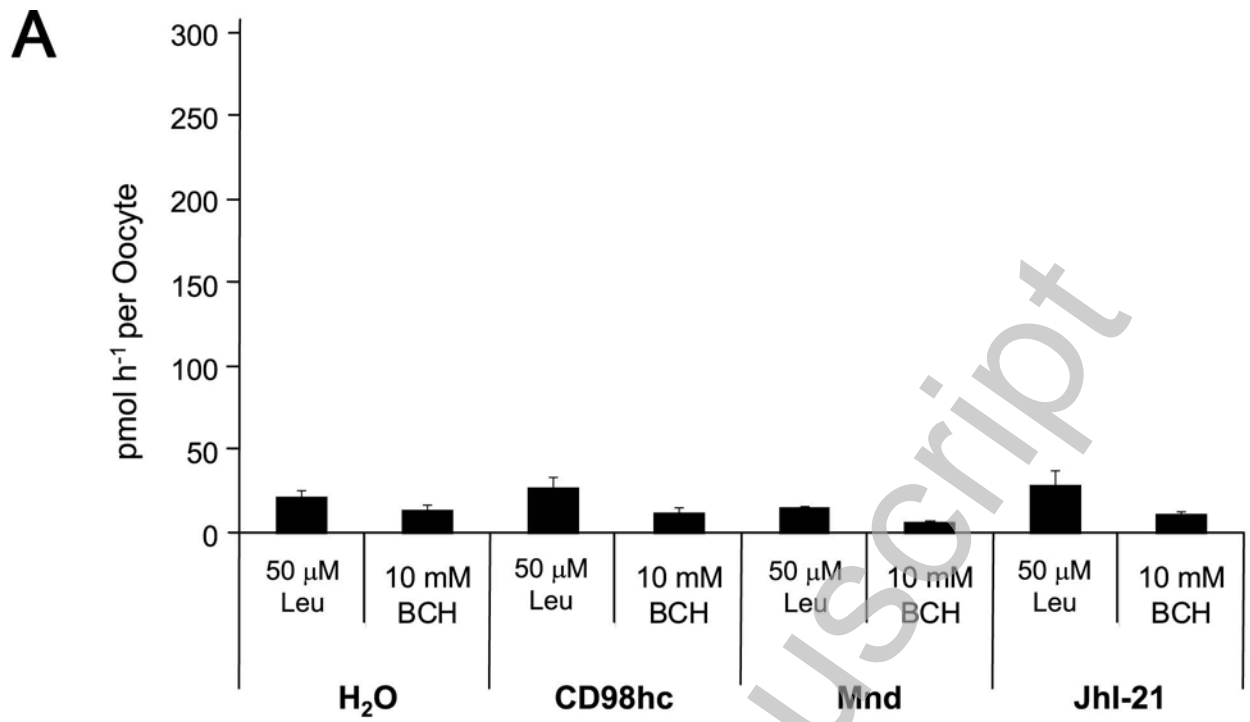
B



THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20082198

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20082198



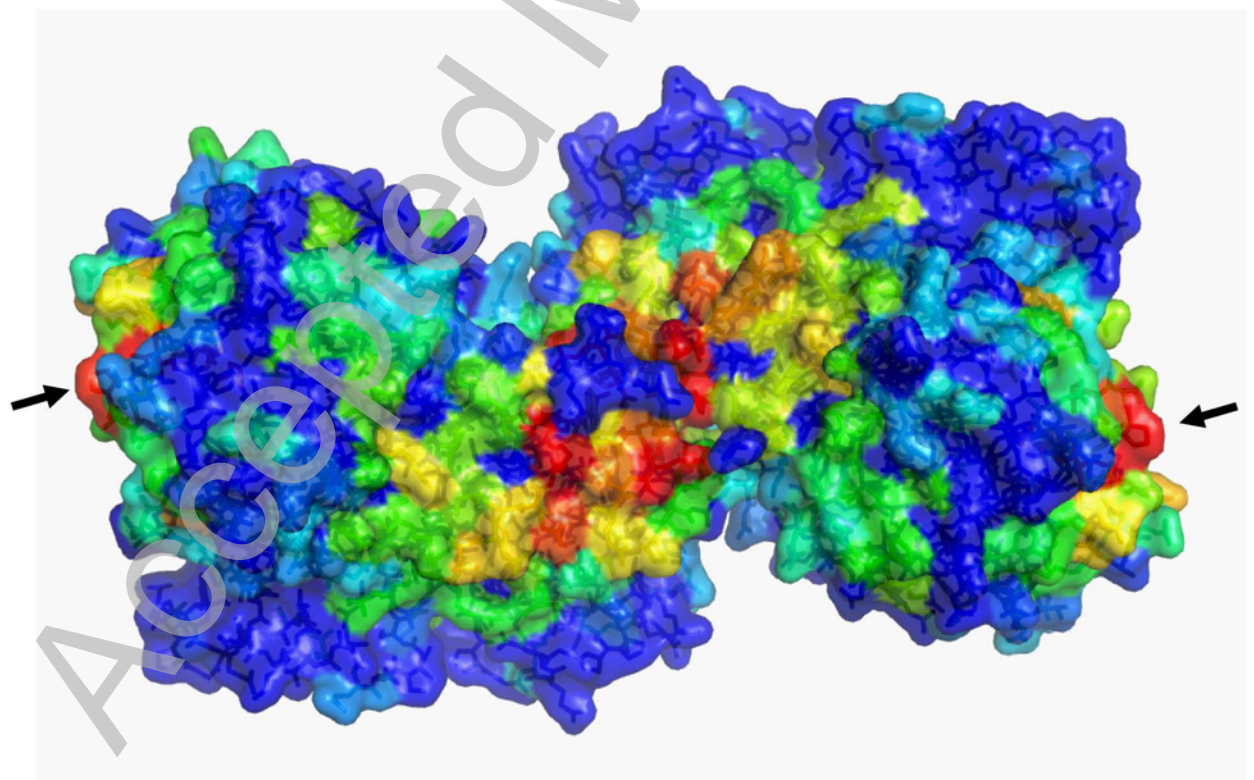


THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20082198

A

Human	1	-----MSQDT EVDKMEVLELNELEPE-----KQP MNAA S-----GAAMSLAGAEKNGLVKIKVAE
Mouse	1	-----MSQDT EVDKMEVLELNELEPE-----KQP MNAA A-----GA AA-----GEKNGLVKIKVAE
Zebrafish	1	-----MSQDT EVDKMEVLELNELEPE-----KQP MTGET-----PTGT-----EKNGLVKIKVAE
Mosquito	1	-----MPRENGTGAADGEEERMLNGAGVEPSGAEPTKEKDKLAQKE-----VEVKFIPF-----SQNGDARDMEL
Fruit Fly	1	-----MVQDEKAAHADGADGEEQMLSGNGVE-----DKLAERR-----DEVKFIKGDHNGDAKIDIGT
HoneyBee	1	MESGRLNLD SVE ANGNAKETAVVATATYKALPDDDTVRQDKIENITMGKPNNEI DDGVHEKMLKDESKISPIKDTTEVKFISENGDTKIDLET
Worm	1	-----MATKDALIAEETGN-----FGGS-----SGDI AVFDKTAEVVNFLELP
Human	50	DEAEAAA AKFTGLSKEELLKLVAGSPGWVRRWALLLFWLWGLMGLAGAVVIVRAPRCRELPAQKWWHTGAYRIG-----DLQAFQGHGA
Mouse	46	DETEAG--VKFTGLSKEELLKLVAGSPGWVRRWALLLFWLWGLMGLAGAVVIVRAPRCRELPAQKWWHTGAYRIG-----DLQAFVGPDA
Zebrafish	43	DA-----EVKFTGLSKEELLKLVAGTAGVWRWALLLFWLWGLMGLAGAVVIVQAPRCRKPPEMHMKGGLVYQIS-----NLDAFS--K
Mosquito	60	ES-----QQALETGMTKEELMKRYANDPFVWRWALLLFWLWGLMGLAGAVVIVQAPRCRKPPEMHMKGGLVYQIS-----NLDAFS--K
Fruit Fly	53	VNGG-----KPAFTGMSKEELMKRYANDPFVWRWALLLFWLWGLMGLAGAVVIVQAPRCRKPPEMHMKGGLVYQIS-----NLDAFS--K
HoneyBee	96	VK-----QALS GMK EELMKRYANDPFVWRWALLLFWLWGLMGLAGAVVIVQAPRCRKPPEMHMKGGLVYQIS-----NLDAFS--K
Worm	40	KP-----GLTKEQLEKRYNDPFVWRWALLLFWLWGLMGLAGAVVIVQAPRCRKPPEMHMKGGLVYQIS-----NLDAFS--K
Human	138	GNIAGLKGRIIDYLSLKVKGVLVGLG--PIHKNOQDDV---AQTDLLQIDPNEFGSKED--FDSLLOQ-----SAKKKSIRVILDLTPNYRGENSWF
Mouse	132	GGIAGLKSHEIYLSLKVKGVLVGLG--PIHKNOQDEI---NETDLQINPTILGSGED--FKDLLQ-----SAKKKSIRVILDLTPNYRGENSWF
Zebrafish	123	NGLKGVEEKDYLSLKVKGVLVGLG--PVHVSQAQDS---SALELTSNPDGSGSE--LTSLLD-----RAHRKGISIVLDTLPNYRGENSWF
Mosquito	147	YGAKGVIYRPADETYVETAAVQ--GKIKQLVSTFG--TKDVKVVVDI--TPNYVTKDALQEQALQ-----KPD SAHSAE VWR--EAGMKP--WL
Fruit Fly	143	LVSAGAIYELPAALTYDVKKPEVE--EKIKHLVLYQ--GSDTRVILLDLPDTPNYVTKDALQEQALQ-----KPD SAHSAE VWR--EAGMKP--WL
HoneyBee	185	VLKEQNI NAI SLASIVKESLTELQTLSDLEALIKAAK--DREQYI LLELDPDTPNYVTKDALQEQALQ-----KPD SAHSAE VWR--EAGMKP--WL
Worm	121	GDFAGISQKIDFLRKLQVTTVYPT--PVIKIKHDEYFNYSYDVVDHNSVDEREGTEQ--FKEIID-----TVHNRAVLMVLDLTPVSTIDLSHPWF
Human	219	F-----TQVDTVATKVRKDALFVWQAAGVDFQMRIDINRYKDASS--FLAEWQNI
Mouse	213	LP-----AQADIVATKVKKEALSSWLDQGVDFGFRFVWGLMNAAPL--YLAEWQNI
Zebrafish	204	N-----NAASVAEKKKACVYMLNKGVCDFISLINDIVNTYA--WSIQAIF
Mosquito	232	SI AE EGS AWKNASNNO-----YVLSQF GADRYDILKMGEPARDKIKATIKALLGGQVRVIRANAKHLINRD--GGDEAPMN
Fruit Fly	227	KVGGNRSAWEKVGDN-----YVLSQF EDGYDILKMSSTIRNEFGAVLKHVLAISVRGFRKNTKFFALFDS--LEDEQ--IS
HoneyBee	278	SVYGS AWE WNEORAQ-----YFHFQNKTOPENYNNPTVYTFSDI LSHLKLSTSGFRANIQYNTEDPN--LHDFPSV
Worm	207	ESRDES KFYIAKPTDPGF NESNFYFFHGANNIKYLGYPSSQNPV--NWKNA D V KSTINGAIQKFLDLGVDGHIHDI SQLAVDSKGRPNHDAVAVV
Human	266	TKGFSEDRLLIAGTNSDQLQI LSLLESNKDILLTSSVYLSDS---GSTGHTKSLVTOYLNATGNRWCS-----WSLSQ-
Mouse	261	TKNLSEDRLLIAGTNSDQLQI VNI LESTS DILLTSSVYLSNS---FTFGRTESLVTFRFNATGSQWCS-----WSVLSQ-
Zebrafish	250	NKTDTGPKKALMGTASLSDVDEITL NQS QV D L L L T R Y P-----AGLSGRQA LMQRECSGHQEA SLG-----WGLGQT
Mosquito	308	LEE A--KSLTHTDYGFWHKEITTYREGLGELFRELVELVHE---QTDDGFLSVTEDI LRREVEVA VING-----TLQMADPLY
Fruit Fly	301	SSPK--DFSLGPNLYGFSYHQRTTFLSGLGDVLYDLSLVKN--ASD--EAFSSVAE DVMPEQVYQLSGDSG-----AYGIDLPMPY
HoneyBee	354	IPT E--ADNYN L VHAYTRDRPENVGILTKWQERINYN---ETENKGFALQDDI GTDMLVEYNE-----K KRL
Worm	302	LEELTKSVQIYVESRDELVDKKI VLFSS LKVDV EDLHVKAT ETGLLHYVIDNS--AND EKK COPHVAKCVHDALNAA YQRHEVAKYTPHWQFSNS E
Human	337	ARLLTSFLPAQLRLYQLMLFTLPG-----TPVFSYDDEIQLDAALPGQMEAPVMLWDES SFPDIPGAVSANMVKG-QSDEPGS
Mouse	332	AGLLADFI PDHLLRLYQLMLFTLPG-----TPVFSYDDELGLQGLPGPAKALMPWNESIFHPRPVS LNMMVKG-QNDEPGS
Zebrafish	322	AGSQASRTPALPKLYQLMLFTLPG-----TPVFSAGDELGLLKA-----GEAMWED-----VNSIAE--ERT A
Mosquito	381	GNIEFDLREGSGPDVARKLRHDIENAYRTIS-----EYAVSAGS GKPWLQRYDRASLEHVP--ASEYTVFHFLLPGVVPFLD-VLTRGNV
Fruit Fly	376	GNFVKVLSK-SKPDTS--LQKEDN-----TLAQSBTDS--WLQWNLADVYVDT PQDPSALAMFLSLLPGVVPVAVD-AVAYQNV
HoneyBee	417	IDIPQNSLFTTADSSINATILKLE-----LSHWL L T S W P R W D I N G K K W S L R H M P A D I A D S L T M L L L P G P V L - K L N D T I S
Worm	397	ASRLASRFESAT A H L S F L Q L T L P G A N S V Y Y Q E Y G L K D A M S K N E F E F K Q M G V M Q W Y P A K G D H H G F S G N D A P I F F P E S D D K L E D N Y N T Q F E T E N S
Human	418	LLSLFRRLSDRSKERLRLDGFHAFSAG-PELFSYIR--HWDQNER--FLVLFNFGDVGLSAGLQAS---DLPASASLPKADL L L S T Q P G
Mouse	412	LLTQFRRLSDRSKERLRLDGFHALSSS-PDLFSYIR--HWDQNER--FLVLFNFRDSDGRSARLGAS---NLBAGISLPSAKL L L S T D S A
Zebrafish	379	VRNFKTLSDLQKMERLRLG EYVNLHNS-PTS LAFLR--LWDSER--FLVAVNWGNSISMTPTND--DLBAEARVRAITDI EKLA--
Mosquito	468	SSESVEITLEKFRATPFCGGLFNAYTDANASAIAYTR--LKSQPG--YFVFLNLA STDQKTFDFSGI--DGLGSELT V L T S Q N Y-----A
Fruit Fly	450	TEETYRQITS LRKTASYMGNLNLQAD-NGLVAFSR--IKSGNPG--YFVFLNPTLEPQASNETIP--DNLDKMTVSYFSEQYNNMMDN
HoneyBee	496	AKDAFATLAKARLDELFVGETMLETIN-KGSVLVTSWLSKSGPG--FLVAVQS AEPIVVDPSIIP--QTGFEVSYLAFSPVLPV-LKNDTIS
Worm	492	ALKIYKRLAKRORDEALVYGETVRDELI NEDVILFSLVYQAVNITATGSTFIVANFNGDKEQKIDPSIEPASKL L P S N K D L N A E I S V I T A N V T
Human	502	R--EEGSPLELERLKEPEPHEGLLRFYPYAA--
Mouse	496	RQSR EEDTSLKLENLNSNPVEGLL LQFPFVA--
Zebrafish	460	---VDSKVKLEKLSGQAVL LSYFPFG--
Mosquito	53	VPNVGKTKVYDAPVPTKAL IATVYPT--
Fruit Fly	533	SGKVAHAGRVNLDLKVAPHSSIIITVYVPKAE
HoneyBee	579	--GETI RTKLPSNKVPLSAKSTIVLTFVPPKN--
Worm	587	DYRVRDHHNFVESPLVLPKQAVL LKL-----

B



THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20082198