The DOMON domain of Knickkopf is essential for its function during cuticle differentiation in *Drosophila*

Khaleelulla Saheb Shaik[1], L. Aravind[2] and Bernard Moussian[3]

Genetik der Tiere, Interfaculty Institute for Cell Biology, Eberhard-Karls University of Tübingen, Germany

1 Present address: Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India

2

3 Author of correspondence

Abstract

The dopamine monoxygenase N-terminal (DOMON) domain is found in extracellular proteins across ~~all taxa~~. Through sequence analyses, the DOMON domain was proposed to adopt a β-sandwich conformation recognising either a sugar or haem moiety. Four highly conserved residues, triptophan[299], methionine[333], arginine[401] and histidine[437], were ~~assigned as essential for DOMON folding and function.~~ Here we have analysed the role of these amino acids in the DOMON domain of the fruit fly Knickkopf (Knk) protein that is involved in extracellular chitin organisation. Genomic fragments of the *knk* region carrying missense mutations in the codon of either of these residues were tested for their ability to normalise the *knk* mutant phenotype in transgenic flies. Our results confirm the absolute ~~importance~~ of triptophan[299,] methionine[333] and histidine[437] for Knk stability and function. ~~Comparably,~~ arginine[401] plays a rather quantitative role in Knk function and efficiency, as in our rescue experiments some animals survive, while some succumb elimination of this residue. Hence, arginine[401] is needed for Knk function to surpass an activity threshold probably by stabilising the β-sandwich conformation required for ligand binding. Overall, our genetic data support the prediction of these four amino acids of the DOMON domain to mediate the function of the Knk protein during cuticle differentiation in *D. melanogaster*.

Introduction

The dopamine monoxygenase N-terminal domain (DOMON, pfam family PF03351) is present in different bacterial, fungal, animal and plant proteins that act in the extracellular space or within the secretory pathway (ER or Golgi apparatus). It adopts an all β-fold with up to 11 β-strands building a hydrophobic pocket that is proposed to

bind to haem or sugars (Aravind, 2001; Iyer et al, 2007). In parallel, it has been speculated that this domain may mediate ~~homophylic~~ or ~~heterophylic~~ protein-protein interaction (Kapoor et al, 2011). Whereas several DOMON domain proteins have been studied by bioinformatics, a functional characterisation of this domain has not been reported. A DOMON domain protein that has been particularly well characterised at the genetic level is Knickkopf (Knk), a GPI-anchored protein in insects (Chaudhari et al, 2011; Moussian et al, 2006). In the fruit fly *Drosophila melanogaster* and the red flour beetle *Tribolium castaneum* Knk is needed for chitin organisation within the extracellular cuticle (Figure 1). However, the role of the Knk DOMON domain has not been assessed. Indeed, the mutations identified in the *D. melanogaster knk* gene all introduce a premature stop codon and therefore do not provide valuable information on the biochemical function of the protein. In order to elucidate the role of the Knickkopf protein during cuticle differentiation in the *D. melanogaster* embryo, we have analysed the function of conserved amino acids in the Knk DOMON domain. Our results clearly underline the importance of these residues in Knk function supporting their role in the DOMON domain.

Materials & Methods

Site-directed mutagenesis and generation of transgenic flies

A 9501 bps NheI genomic fragment of the *knk* region was isolated from the BAC R39F04 and sub-cloned into the vector XbaI site of the transformation vector pW8 that allows insertion into the fly genome. This fragment was used to introduce base changes by site-directed mutagenesis following the protocol of the manufacturer (Stratagene). The primers used are listed in table 1. After constructs were sequenced to verify mutations, they were sent to Fly-Facility (Clermont-Ferrand) to generate transgenic w^{1118} flies. The presence of introduced mutations in transgenic flies was again verified by sequencing PCR products of the respective gene region.

Fly genetics

To test whether a transgenic *knk* construct is able to normalise the *knk* mutant phenotype, *knk* homozygous mutant embryos were produced that carried the respective transgene. All transgenes used in this work caused comparable red eyes witnessing comparable expression levels of the marker gene *w*+, in turn suggesting that differences in *knk* expression are negligible, as well. In brief, males

~~heterozygous transgenic Knk for the *knk* mutant allele *3-7H* and carrying a transgenic construct were crossed to females heterozygous for the *knk* mutant allele *5C77*. Sta~~tistically, 12,5% (1/8) of the resulting progeny are transheterozygous for $knk^{3\text{-}7H}$ and knk^{5C77} and carry the transgenic wild-type construct. If the wild-type construct were able to rescue the *knk* mutant phenotype, these animals would represent 14,3% (1/7) of total adult flies (1/8 being embryonic lethal). The *curled* (*cu*) mutation that is closely linked to both alleles (0,9 centi-Morgan) and causes curled wings was used to identify these flies. Since in males crossing-over is suppressed and therefore *knk* and *cu* invariably co-segregate, no distortion of segregation statistics is expected through recombination. In other words, presence of curled-wing flies would indicate that the transgenic construct is able to reconstitute Knk function.

Western blotting

For protein extraction, stage 17 embryos were homogenized in PLC buffer containing a cocktail of protease inhibitors. Protein amounts were determined by photometry at 280nm (Nanodrop, Peqlab, Germany). For each probe, about 15 µg protein were separated by SDS-PAGE (7.5% polyacrylamide). Subsequently, proteins were transferred to a nitrocellulose membrane (Whatman) by the semi-dry method. The Odyssey infrared dual-colour detection system (LI-COR Biosciences) was used for protein detection. The dilutions of α-Knk (rabbit) and α-tubulin (mouse, Sigma-Aldrich) primary antibodies were 1:1000 and 1:5000, respectively. The fluorophore-conjugated secondary antibodies, α-rabbit-680 and α-mouse-800 were used at a dilution of 1:7500. Detection was visualised by the Odyssey infrared imaging software (LI-COR Biosciences).

Results & Discussion

The Knickkopf (Knk) protein is composed of three types of domains, an N-terminal signal peptide and a C-terminal signature for GPI-anchor attachment (Figure 2). Two tandem DM13 domains (residues 51-156 and 166-272) follow the N-terminal signal peptide. Amino acids 287 to 442 constitute the DOMON domain. ~~Before~~ the GPI-anchoring signature, we identified a ~~rusticyanin-like~~ domain (residues 511-654) using the HHpred software (Soding et al, 2005). ~~Residues triptophan^{299}, methionine^{333}, arginine^{401} and histidine^{437} of the DOMON domain are highly conserved in the DOMON domains of many different proteins (Iyer et al, 2007). These residues~~

~~stabilise the β-strands 3 (triptophan[299]), 5 (methionine[333]) and 9 (arginine[401]) or are predicted to be involved in substrate i.e. haem or sugar binding (histidine[437]) (Table 2).~~

To test whether these residues are indeed essential for Knk function, we undertook a transgenic approach. First, we identified a recombinant genomic fragment encompassing the *knk* locus to ~~normalise~~ the *knk* mutant phenotype (see Materials and Methods). ~~In other words, flies homozygous mutant for *knk* survive to adulthood if they additionally harbour this transgenic construct.~~ Next, we generated four modified versions of this fragment, each having a missense mutation in either codon for the residues to be tested. Interestingly, the construct carrying the missense mutation to exchange arginine[401] to glycine was able to rescue the *knk* mutant phenotype, albeit efficiency to do so was significantly reduced (Table 2). By contrast, the missense mutations triptophan[299], methionine[333] and histidine[437] failed to rescue the mutant phenotype. Instead, the respective animals died as larvae within the egg case displaying the *knk* mutant phenotype and were unable to hatch.

To verify whether the induced mutations may affect the stability of the resulting protein, we performed Western blot experiments on protein extracts from respective transgenic *knk* mutant larvae (see materials & Methods). It should be noted that in *knk* mutant larvae the Knk protein is not expressed (Moussian et al, 2006). As suggested by rescue experiments, the exchange of arginine[401] to glycine does not severely interfere with Knk function. However, the Western blot experiments showing traces of Knk protein in *knk* mutant larvae harbouring the respective construct suggest that rescue through the Knk[arginine401>glycine] protein is not 100% penetrant. This finding is consistent with reduced ratio of rescued adult *knk* flies harbouring the respective transgenic construct (Table 2). A valine instead of a methionine at the position 333 ~~somehow~~ ~~destabilises~~ the protein. ~~The role of~~ ~~methionine[333]~~ ~~is hence unclear; the inability of the respective protein to normalise the mutant phenotype is either due to reduced protein levels, or to the mutation.~~

Proteins in which the residues triptophan[299] and histidine[437] are exchanged are expressed at levels comparable to the unchanged wild-type protein indicating that lack of rescue is not due to instable protein (Figure 3). We therefore hypothesise that the substrate or partner binding histidine[437] is crucial for tethering Knk either to chitin or to Retroactive that is needed for correct Knk localisation at the plasma membrane and the extracellular space in *Tribolium castaneum* (Chaudhari et al, 2013). In *D.*

melanogaster, this residue may mediate binding to chitin, in turn recruiting the chitin-binding protein ObstructorA (ObstA) for chitin organisation in the tracheae (Petkau et al, 2012). Alternatively, histidine^{437}, which is pivotal in the haem binding pockets of ethyl benzoate dehydrogenase and cellobiose dehydrogenase that have a DOMON-related domain architecture, may indeed be involved in haem binding. In this scenario, the haem-protein Knk would act in a yet unidentified redox chain reaction to organise chitin. Biochemical binding assays with the mutant Knk versions shall be performed to test these hypotheses. The results would be transferable to other insects as an alignment of the Knk sequences of the distantly related species *D. melanogaster*, *T. castaneum* and *Apis mellifera* (honeybee) reveals that these amino acids are conserved among insects (supplementary figure 1).

Taken together, our data confirm the *in silico* predications that the four amino acids triptophan^{299}, methionine^{333}, arginine^{401} and histidine^{437} are essential residues of the Knk proteins needed for correct chitin organisation.

References

Aravind L (2001) DOMON: an ancient extracellular domain in dopamine beta-monooxygenase and other proteins. *Trends Biochem Sci* **26:** 524-526

Chaudhari SS, Arakane Y, Specht CA, Moussian B, Boyle DL, Park Y, Kramer KJ, Beeman RW, Muthukrishnan S (2011) Knickkopf protein protects and organizes chitin in the newly synthesized insect exoskeleton. *Proc Natl Acad Sci U S A* **108:** 17028-17033

Chaudhari SS, Arakane Y, Specht CA, Moussian B, Kramer KJ, Muthukrishnan S, Beeman RW (2013) Retroactive maintains cuticle integrity by promoting the trafficking of Knickkopf into the procuticle of Tribolium castaneum. *PLoS Genet* **9:** e1003268

Iyer LM, Anantharaman V, Aravind L (2007) The DOMON domains are involved in heme and sugar recognition. *Bioinformatics* **23:** 2660-2664

Kapoor A, Shandilya M, Kundu S (2011) Structural insight of dopamine beta-hydroxylase, a drug target for complex traits, and functional significance of exonic single nucleotide polymorphisms. *PLoS One* **6:** e26509

Moussian B, Tang E, Tonning A, Helms S, Schwarz H, Nusslein-Volhard C, Uv AE (2006) Drosophila Knickkopf and Retroactive are needed for epithelial tube growth

164 and cuticle differentiation through their specific requirement for chitin filament
165 organization. *Development* **133:** 163-171
166 Petkau G, Wingen C, Jussen LC, Radtke T, Behr M (2012) Obstructor-A is required
167 for epithelial extracellular matrix dynamics, exoskeleton function, and tubulogenesis.
168 *J Biol Chem* **287:** 21396-21405
169 Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein
170 homology detection and structure prediction. *Nucleic Acids Res* **33:** W244-248
171
172

173 **Table**

174 Table 1 Primers used in site-directed mutagenesis experiments to exchange amino
175 acids

Residue exchange	Primer
Trp > Gly	GATATGCAGGTGTCC**T/G**GGGAGGTTTTCGGGC
Met > Val	GATGTATCCAGCCAG**A/G**TGATTGGCTCCGATGTCGTG
Arg >Gly	CATTAATACGATCAGTTTTAGG**C/G**GGACGCTCAAGTCAagtcatgtgag
His > Ser	CCAACAACGAGCCCGCCTTT**C/AA/G**CACTTACTATCCCAAGAGCG

176

177 Table 2 Missense mutations introduced in Knk by site-directed mutagenesis

Original > new residue	rescued flies	protein expression*
wild-type	n=362, 11,9%	++
Trp (hydrophobic) in β-strand 3 > Gly	n=443, 0%	++
Met (ligand binding) in β-strand 5 > Val	n= 447, 0%	+
Arg (polar) in β-strand 9 > Gly	n=428, 9,7%	+
His (ligand binding) in β-strand 11 > Ser	n=433, 0%	++

178 Statistically, 14,3% (1/7) of the adult flies produced by the rescue cross should be
179 rescued *knk* homozygous animals harbouring the respective transgenic construct.
180 The value 11,9% (wild-type construct) is not significantly different from the expected
181 one (Chi^2 test, $T_{X2}<X^2_{1,0.995}$, α = 0,995). The value 9,7% (Arg^{401}>Gly), by contrast, is
182 significantly different from the expected 14,3% suggesting that rescue is not 100%
183 penetrant (Chi^2 test $T_{X2}>X^2$).
184 * see figure 3
185

186 **Figure Legends**

187 Figure 1

188 A) The wild-type larva is spindle-shaped. Its head skeleton is an elaborate melanised
189 cuticular structure (arrow). B) The *knk* mutant larva has an irregular shape, and its
190 head skeleton is condensed.

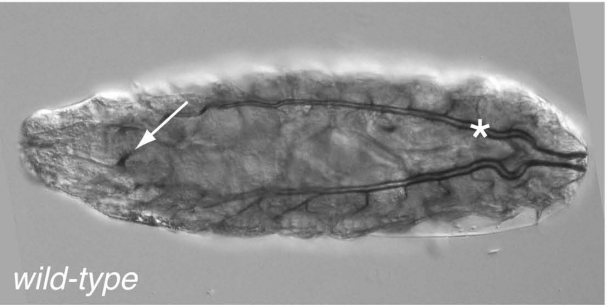
191 Figure 2

192 The Knk protein has 689 residues constituting an N-terminal signal peptide, two
193 DM13 domains, a DOMON domain, a plastocyanin-like domain and a C-terminal GPI
194 anchoring signature.

195 Figure 3

196 Western blot experiments show that mutant versions of Knk are expressed in the
197 embryos harbouring mutated recombinant genomic fragments of the *knk* locus.

Figure 1



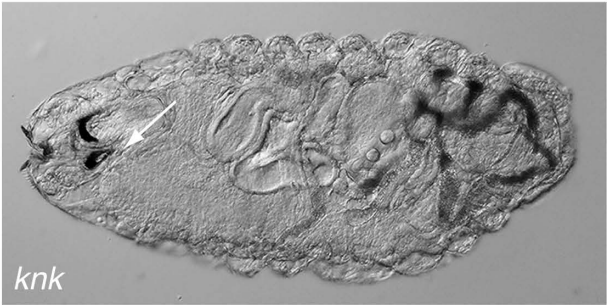


Figure 2

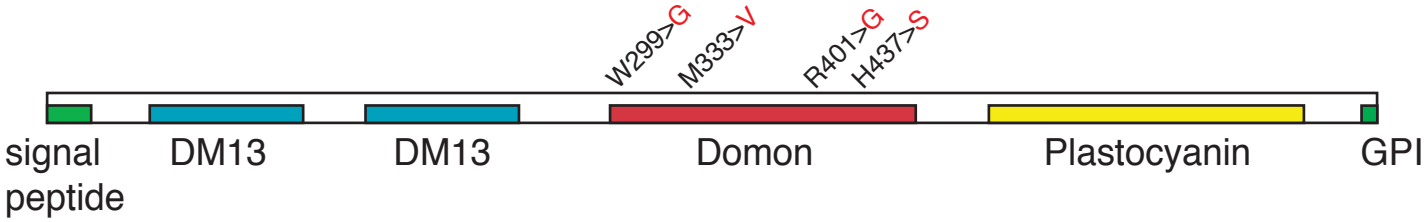
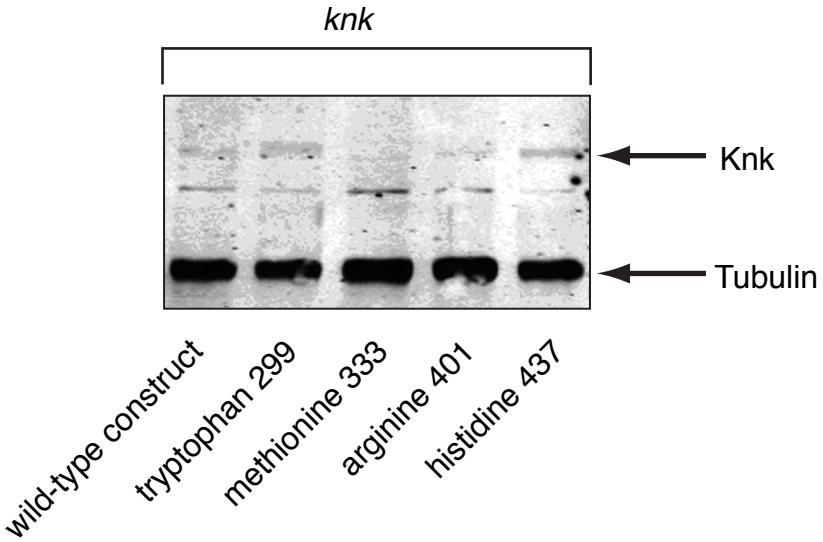


Figure 3



Supplementary Figure 1.

Alignment of the *Apis mellifera* (Amel), *Tribolium castaneum* (Tcas) and *Drosophila melanogaster* Knk proteins. The DOMON domain is highlighted in yellow. The conserved amino acids studied in this work are red.

```
CLUSTAL 2.1 multiple sequence alignment
Amel            MADVTKRWTKDTKAHVFVIILTYVICSVYGCSEYDRLDGDEYKGKCIGKLNTYHHEVSGE 60
Tcas            ---------------MLLLFVLLLASSGSCQDSE----GPYRGKHIGKLNSYHHEVGGD 40
Dmel            MALSSRVLCLRLQHHLLLVILTIFLLCPAASAQNEEED-GPYRGKYLGKLNSYHHQVSGD 59
                                :::::. .: .  .. : :    . *:** :****:***:*.*:
Amel            VYAVDENTLLLINFKYDGNGDDTFFFAGASTRAGPHGFIVPDEWGKTNVLDKYFNKNLTL 120
Tcas            VYAVDDYTLLLTSFSYDGNGADTFFWAGASNRPGPQGFIVPDEYGKTNVLARYLQKDFTL 100
Dmel            VYAVNEYTFLIVGFNYDGNGADTFFWSGASNRPGPQGFIVPDEYGKTNILDRYHNKDFTL 119
                ****:: *:*: .*.***** ****::***.*.**:*******:****:* :* :*::**
Amel            NLPDGKTIADIKWFAIYDLGSQNVFGDVIFPDEFRAPAPATISQLKKYSHGVSSEPIEII 180
Tcas            TLPDNKKITDIKWFAIYDLLSQNTFGDIYIPEEFEPPTVQKIPQLASKSHSVSSGAVEII 160
Dmel            TLPDRKKITEIKWLAVYDLSSQNNFGDVYIPEEFDPPMSQLGGTFSKRSHNVSSSSVEIL 179
                .*** *.*::***:*:*** *** ***: :*:** .*       :  . **.*** .:**:
Amel            DSKTIKIPNFKYDGEGKDTYFWVGDGPQPSTRGTKVPDEYGYLDSLHAYRGEDVIIQLPG 240
Tcas            DAKRIRLNDFRYNGQAKRAHFWVGVGAQPVSKGHKVPDENGYLDPLRAYKGETITLELPG 220
Dmel            DSKTIRIKDFTYDGRGKRTFFWTGVGPQPSSRGSKLPDERGYLDPIRQYNKETIELELPG 239
                *:* *:: :* *:*..* :.**.* *.** ::* *:*** ****.:: *. * : ::***
Amel            DMTIFNINWLSIFDVETASNFGSVIIPKELNVPPSLVKVIKHINSLPNCVQLHKRFQVSW 300
Tcas            DLTIFNIAWFSVFDLETNENLGSIVIPDGLNVPPSLVKVIPHEHHLPNCLQLHKDFQVSW 280
Dmel            DKTIFDIDWISVYDVADNENYGHVLFNDKLNVPPSLVKVTPFEFSLPNCRQLHKDMQVSW 299
                * ***:* *:*::*:   .* * ::: . **********  .   **** **** :****
Amel            EIFGPQITIQLAGDIGENEYMAFGLSGSEVSSQMEGADVAVAHLDSIRGYVVDYNITAKA 360
Tcas            DIFGPQITIQMAGQVDENSYMAFGMSGSKEQSQMLGSDVTVAYLDGYRGYATDYNITALT 340
Dmel            EVFGPQITFQLSGQVGGNDYMSFGISGSDVSSQMIGSDVVVAYIDDIRGYTVDYNITSLA 359
                ::******:*::*::. *.**:**:***. .*** *:**.**::*. ***..*****: :
Amel            PCGKVLGQYKGVCRDELFGGLDNNQLHTAFKENGITVINYRRTLNSSDPGDKEYPTDRSV 420
Tcas            PCVKVLGQYKGVCKDELVGGQDSNQMHTATRENGINIITFRRALISPDPGDQEYPTEGSI 400
Dmel            PCVQVLGQNKGVCRDDVVGGLDSFQLNTYSRKDGINTISFRRTLKSSDDGDKEIFLDRSN 419
                ** :**** ****:*::.** *. *::*  :::**. *.:**:* *.* **:*   : *
Amel            YVVWALGRLNENKEPTFHDLYLKNDLKLELGRKEPEDACMDFTENNEQLVVP-WEKVKIY 479
Tcas            YIIWAMGRLDHNKEPTFHDIYPKTNISIELQPKEPSSNCFSFTRSDTQLREP-WVKGQIY 459
Dmel            YVIWAFGPLDSNNEPAFHTYYPKSDIVIDFNTTEPVNDCFAFTKRAETTNPPVWERTRIT 479
                *::**:* *: *:**:**  * *.:: :::  .** . *: **.       * * : :*
Amel            GRSIRTFKATIGPSGGKKGYQGITGQPSTGLAWYIEGRLVPELYLRRGLTYNFRVFGGNN 539
Tcas            DKTIRTFKAYLGPSGGRRGYQGTTGQTSTTLAWYVNGLLAPELWLRRGLTYAFRVYGGNN 519
Dmel            DATVRTFNAYLGPSGGLRGYQGLTNHVSSGLAWYINGYMIPELYLKRGLTYTFKVRGGNN 539
                  . ::***:* :***** :**** *.: *: ****::* : ***:*:***** *:* ****
Amel            PHSSNLYHPLIITDELHGGYDKLSDAAQSNIRVLAGVEFTRRGQPRPTAVGPLCLSKHN- 598
Tcas            PHSAEFYHPLIITDEPHGGFDRLSDEAQAKIRVLAGVEYSRRGRPRPTAAGPLCLAKHQD 579
Dmel            PHSPEHYHPLVITDDPQGGYDRLSDAKQSEIRVLAGVEFTRRGRPKPTAAGPLCLSRYPQ 599
                ***.: ****:***: :**:*:***  *::********::***:*:***.*****:::
Amel            GRDRRRDDDFVTFKQFNRTLINTCEPGEGGNLEITPNSTWPDIVYYNSFTHANMGWKIHI 658
Tcas            SNDRRLDGDFPTFKKFNRSLVYSCEEGEPGVLEFTPNSSWPDIVYYNSFTQSNMGWKIHI 639
Dmel            NSDRRLDDNFPTFKKFNRSLITECVEGEPALLEITPNITWPDTVYYNSFTHGNMGWKIHI 659
                . *** *.:* ***:***:*:  *  ** . **:*** :*** *******:.********
Amel            VDSYSSSNAVVRQLS--LILGITAVFLHFL- 686
Tcas            IDSFSSRTNGISKMCGNIFVLVSILFVKQFV 670
Dmel            VDSYTNLKSGSMGLSWSLCIILLPWLVLQN- 689
                :**::. .     :.  : : :   ::
```