Myosin binding protein C: Structural abnormalities in familial hypertrophic cardiomyopathy

Cecily E OAKLEY¹, Brett D HAMBLY¹, Paul MG CURMI², Louise J BROWN²

¹Department of Pathology, University of Sydney, NSW 2006, Australia.

²Initiative for Biomolecular Structure, School of Physics, University of New South Wales, NSW 2052 and Centre for Immunology, St Vincent's Hospital and University of New South Wales, Sydney NSW 2010, Australia.

ABSTRACT

The muscle protein myosin binding protein C (MyBPC) is a large multi-domain protein whose role in the sarcomere is complex and not yet fully understood. Mutations in MyBPC are strongly associated with the heart disease familial hypertrophic cardiomyopathy (FHC) and these experiments of nature have provided some insight into the intricate workings of this protein in the heart. While some regions of the MyBPC molecule have been assigned a function in the regulation of muscle contraction, the interaction of other regions with various parts of the myosin molecule and the sarcomeric proteins, actin and titin, remain obscure. In addition, several intra-domain interactions between adjacent MyBPC molecules have been identified. Although the basic structure of the molecule (a series of immunoglobulin and fibronectin domains) has been elucidated, the assembly of MyBPC in the sarcomere is a topic for debate. By analysing the MyBPC sequence with respect to FHC-causing mutations it is possible to identify individual residues or regions of each domain that may be important either for binding or regulation. This review looks at the current literature, in concert with alignments and the structural models of MyBPC, in an attempt to understand how FHC mutations may lead to the disease state.

Keywords: myosin binding protein C, hypertrophic cardiomyopathy, sarcomeric proteins, immunoglobulin superfamily, mutations.

Introduction

Myosin binding protein C (MyBPC) was first identified as a contaminant in myosin preparations 30 years ago[1], but as its assigned role in the structure and regulation of the sarcomere grows evermore complex, it has come to be of increasing interest to the scientific community. Further importance in gaining an understanding of the MyBPC molecule is illustrated by the identification of numerous disease causing mutations that results in familial hypertrophic cardiomyopathy (FHC). FHC is a serious disorder, especially in children and young adults, that can result in significant morbidity and premature death. One current challenge is to decipher the mechanism through which these mutations in the MyBPC sarcomeric protein can result in severe disease. This review aims to integrate our current understanding about the structure of cardiac MyBPC and its role in the sarcomere.

E-mail: cecily@med.usyd.edu.au

Structure and Function

MyBPC is a modular polypeptide that belongs to the intracellular immunoglobulin superfamily. It is a sarcomeric protein of approximately 137 kDa found in the thick filament of striated muscle. MyBPC is located in the central region of the A-band, known as the C-zone (Fig 1) [2]. In mammalian muscle, seven to nine of the eleven structurally regular transverse C-zone stripes contain MyBPC[3]. The 43 nm spacing[2] of these stripes dictates that only every third level of myosin heads in the Czone is associated with a MyBPC molecule[4]. This, and the number of myosin heads that fall outside the C-zone, limit the number of myosin heads that can interact directly with MyBPC[5]. The core structure of MyBPC is comprised of seven I-class immunoglobulin (IgI) domains and three fibronectin type III (FnIII) domains, numbered from the N-terminus as Motifs 1 to 10 (Fig 1).

Both structural and regulatory roles have been proposed for MyBPC. It is involved in sarcomere assembly by promoting polymerization of thick filaments, via the C-terminal domains binding to specific sites on titin and light meromyosin (LMM)[6]. Indeed MyBPC has been demono-

^{*}Correspondence: Cecily OAKLEY

Tel: 61 2 9351 3059, Fax: 61 2 9351 3429;



Fig 1. Position of MyBPC in the stretched sarcomere and the structure of cardiac MyBPC. The top diagram is a schematic diagram of the sarcomere indicating the location of MyBPC. MyBPC is found in transverse stripes 43 nm apart in the C-Zone, where interaction between the thick and thin filaments occurs. The bottom diagram shows that cardiac MyBPC is comprised of eight IgI domains and three FnIII domains. Motif 0 is cardiac specific as is the extra phosphorylation sites between Motifs 1 and 2. Motifs 7-10 have been found to bind to myosin and titin.



Fig 2. Schematic diagram of the trimeric collar model of cardiac MyBPC. Motifs 5-10 of three cardiac MyBPC molecules assemble in a staggered parallel array around the thick filament. The N-terminal domains (Motifs 0-4) project out from the thick filament to interact with the helical myosin S2 region and possibly the thin filament. Figure modified from Moolman-Smook et al. 2002 [9].



Fig 3. Schematic diagram of the axial model of cardiac MyBPC. Motifs 7-10 are arranged axially along the myosin backbone and are able to interact with titin. The N-terminus reaches out to interact with the myosin crossbridge and/or actin. Figure modified from Squire *et al.* 2003[10].

crossbridges, which modulates muscle contraction[4].

The structural basis for the incorporation of MyBPC into the thick filament is at present poorly understood. Recent data suggest that the formation of a trimer of MyBPC, via the C-terminal half of each molecule, allows the assembly of a collar around the thick filament (Fig 2) [9]. However, an alternative model for MyBPC incorporation into the thick filament has been proposed that does not involve specific interactions between MyBPC molecules (Fig 3)[10].

Isoforms/alignments

Three isoforms of human MyBPC have been identified; fast skeletal, slow skeletal and cardiac[8]. All three isoforms map to different chromosomes indicating that they are not the result of alternative splicing. The computer program, CLUSTAL X[11], was used to create multiple alignments of different isoforms of MyBPC from a range of species. These alignments identified regions of conservation and isoform divergence. The sequence identity, for domains 1-10, across human isoforms is 39.6% (Fig 4). Regions of low identity were found to occur in sequences outside of the IgI or FnIII domains, specifically the sequence that precedes Motif 1, the 1-2 linker, and, in cardiac MyBPC, an insert in Motif 5.

The sequence homology for cardiac isoforms from different species is 46.8% (Fig 4). The cardiac isoform differs from the skeletal isoforms in three major ways. Firstly, the cardiac isoform has an extra IgI motif at the N-terminus (Motif 0)[12]. Secondly, there are three phosphorylation sites between Motifs 1 and 2 in cardiac, compared to one in skeletal[12]. Finally, one of the central IgI domains (Motif 5) contains a proline/charge-rich insert[12]. The cardiac isoform retains most of its unique features when mapped back to the Japanese pufferfish, *Fugu rubripes*, genome sequence[13], which also has Motif 0 and the Motif 5 insert, although one of the extra phosphorylation sites appears to be absent.

Cardiac MyBPC

The cardiac isoform comprises 2% of myofibrillar protein in the heart and has been found to be particularly important during myofibrillogenesis and in regenerating muscle cells[14, 15]. However, despite cardiac MyBPC gene knockout mice showing that MyBPC inclusion is not necessary for the formation of the sarcomere, its absence still resulted in hypertrophy and impaired contractile function[16].

The extent of phosphorylation of MyBPC correlates with increased systolic tension and occurs in response to adrenergic stimulation[17]. Partial extraction of MyBPC from fibres results in an increase in Ca²⁺-sensitivity and velocity of shortening[18, 19].

The cardiac isoform of MyBPC has also been of particular interest because of its link to the heart disease, familial hypertrophic cardiomyopathy (FHC), which is

hu-fat hu-fat	<pre></pre>
nu-co provide variability in under a second	Can Variant Bio Yakasa Kaya Kata Kaya Wakasa Kaya Kaya Wakasa Kaya Kaya Kaya Kaya Kaya Kaya Kaya K
<pre>sheep Alexa A you for protect A you for the construction of a construction of a</pre>	2 Part of the the second secon
<pre>fu-cc fu</pre>	
Number Notif 1 Proposed Actin Binding Site Notif 1 Back Parker Analog Sign Parker Prove Represented For Parker Prove Represented For Parker Prove Represented For Parker Prove Represented For Parker Prove Parker Park	<pre>Motif 1 Proposed Actin Binding Site Motif 1 Proposed Actin Binding Site Motif 1 Proposed Actin Binding Site Proposed Pr</pre>
Proposed Actin Binding Site Motif 1 Proposed Actin Binding Site Motif 2 Proposed Actin Binding Site Motif 1 Proposed Actin Binding Site Motif 2 Proposed Proposed Actin Proposed Actin Binding Site Motif 2 Proposed Propos	Proposed Actin Binding Site Motif 1 Proposed Actin Binding Site Motif 2 Proposed Ac
<pre>Number 2 September 200000 September 200000 September 200000 September 2000000000000000000000000000000000000</pre>	APARENDERS ANALYSIS AND ALL DEADS - ALL DES ANALONDER AND ALL DEAD PORT AND ALL DEAD AND ALL DEA
me-cc SEVARGADEEVEARTILES/VSEPESSEESVSEPESSEESVSEPESSEEDES/VSEPESSEEVERS/VSEPESSEERES/	PGAPESPUPAATELESYSSPEGS
ch-cc	PARASARAASELDAPVESHORP
<pre>Tu-co R2PEBSTYLEKELEPVLAGENE- Status Moisif 1 cont. * * * * * * * * * * * * * * * * * * *</pre>	P2PESTYLKKSLDFYLALEBLA P2PESTYLKKSLDFYLALEBLA P2PESTYLKKSLDFYLALEBLA MOLIF 1 cont. * (P) * (P) (Phosphorylation sites) STRESN KSANAWYTEHLGKYVLGUNGTTAKOUTUVET PEGEZAKUNGGALEBLASTUTELAGAGUTETIA VKAEDLAB 100 MOLIF 1 cont. * (P) * (P) (Phosphorylation sites) STRESN KSANAWYTEHLGKYVLGUNGTTAKOUTUVET PEGEZAKUNGGALEBLASTUTELAGAGUTETIA VKAEDLAB OLHASTO ASYVLETIHLTDA (PAFTGSUCGVSTUCKES CHPTIC) FEGZEXAUGAGUTELAGAGUTETIA VKAEDLAB OLHASTO ASYVLETIHLTDA (PAFTGSUCGVSTUCKES CHPTIC) FEGZEXAUGAGUELLBASTUTELAGAGUTETIAL SKESTATAGUELG OLHASTO ASYVLETIHLTDA (PAFTGSUCGVSTUCKES CHPTIC) FEGZEXAUGAGUELLBASTUTELAGAGUTETIAL OLHASTO ASYVLETIHLTDA (PAFTGSUCGVSTUCKES CHPTIC) FEGZEXAUGAGUELLBASTUTELAGAGUTETIAL SAGAGUTETIAL OLHASTO ASYVLETIHLTDA (PAFTGSUCGVSTUCKES CHPTIC) FEGZEXAUGAGUEGUELLBASTUTELAGAGUTETIAL OLHASTO ASYVLETHUS LEAANFAAUCUVET WENDSCHPILLWERAYSGE MI FRANTUST LAGAGUTETIAL OLHASTO ASYVLETHUS LEAANFAAUCUVET WENDSCHPILLWERAYSGE MI FRANTUST LAGAGUTETIAL OLHASTO ASYVLETHUS LAANFAAUCUVET WENDSCHPILLWERAYSGE MI FRANTUSCE OLHASTO ASYVLETHUS LAANFAAUCUVET WENDSCHPILLWERAYSGE MI FRANTUSCE P2 FORMUGAUTUT SCHWITT MOLIFAR FTC - EEROVINGUUTUT HOLAS AN ANAWYT MI FRANTUSCE FTC - EEROVINGUUTUT HOLAS ANAWYT MI FRANTUSCE
No.111 1 cont. ************************************	Notif 1 cont. ************************************
hu fat hu-ac bac-ac control of the second secon	DUIL 1 COLL COLL COLL COLL COLL COLL COLL C
hu-cc SCILDLARSTONAS INVLETE LITTOROPATISSIES INCOVER DEPENDENT PLANMETGE - LILEASA PHITELAGGGENESS	OLHOSDIAS MY LEFLINTA OPAPTES IN CONSTRUCT CONFILTY PERGY CELLSA PUTCHAGE (CONSTRUCT PERGY). CONSTRUCT AND TAKEN AND
Gold DLH DS DIA SYVLEFE LIT DAOT SAGE ICCVET DOFTS CHENT THE ALGO COLONG HUT ALGO CHET ST SHEDA OF DES GREGORING ST	QUHRSTEINAS EVILLE HUITTAQTTSAGGY ECTISTERETISCHENT THEALGGEGE - LILENSAFHUTT LAGGEMENTS
Ch-cc Gruco Students of the state of the st	CONTRACTOR NOT TRANSFILLEAMET AGGING USE TO THE DESENTED THE ACCOUNT OF DESENTED THE DESEGNMENT AFTER DESEMPTION AFTER DESEMP
hu-siw CK LGLKHTFE HSTVHTEMOIIKAKDNPAGNYKCEVTY DK DSCSPDLEVH STGTTENIE IRAF 150 () Lotif 2 () Lotif 2 () Lotif 2 Lotif 2	DURETFERENTY TERMOLIKARDNEAGY TY BERDESE DLEY HENDESE DLEY HENDER TOTENTI BERGE TAGES - BOOMAGE DESCRIPTION FOR COMPANY AND ADDRESS AND AD
(P) hu-fst FeVVEEEKKKKKRDDDDLGTPPFIWELLIGGKK-SEVEXTAPOKGITDLGKLGLGGKDEKGAFTGOGLGSAQUVDLGWHT HIMVETSDPDLTW mo-co OGFRROSKLBAPAEEDWHTLGAPP-SETELTAPOKOVTDLGKLGLGGKDEKGSTAFOGLGSAQUVDLGWHT HIMVETSDPDLTW mo-co OGFR	Motif 2 YEEKKKKKKDDDDGGIPPNIWELL GAKK-SEVEKIATONG GITLEGKLGLIAGKUDLAGA VEV MGAAFT GULDPA VOUGG LILLAW SISDPDLTAWFNGG VEL YEEKKKKKKDDDDGGIPPNIWELL GARF-SEVE IAROG UTDLEGKLGLIAGKUDKAG TAFO GUEPA VONG GULLAW SISDPDLTAWFNGG VEL YEEKKKKKKDDDGGULAPAREEVWEIL GARF-SEVE IAROG UTDLEGKLGLIAGKUDKAG TAFO GUEPA VONG GULLAW SISDPDLTAWFNGG VEL YEEK YEEK TAROG UVELAGULAGKUDKAG TAFO GUEPA VONG GULLAW VALLADGA VUNG GULLAU VALLADGA VUNG GULLAU VALLADGA VUNG GULLAU VALLADGA VUNG GULLAU VALLANG VUNG GULLAU VANDAD VALLANG VUNG GULLAU VANDA VANDA VANDAD VALLANG VUNG GULLAU VANDA VAN
hu-fst BVVEEKKKKKKDDDDGIGPESIWELLGAKK-SETERIARGOVITDIGGKLKLGAGKVEKSAAFTENDENAGVDGAN TIAMVEIGDPDITIGW hu-co DFFRROSKLAPAEDWHILGAPP-SETERIARGOVIDIGGKLGAGKEKSTAFGOLDENAGVDGKGHILDTVELADDAAFVEM mo-co DFFRROSKLAPAEDWHILGAPP-SETERIARGOVIDIGGKLGAGKESTAFGOLDENAGVNGGHILDTVELADDAAFVEM ch-co DFFR	TELEKKKKKKKDDDDLGPPETWEILL GARK-SEVEE IAPO GITDL GAL HELKAN VEVESAAT KLUPA GUD GAN ITIAWE ISDPDTLWF NGG 20000000000000000000000000000000000
hu-cc DSFRTPROSKLEAPAGEDYWEILMOAPP-SETERIAPOGVOTDLEGHLENLGGMOCOEKTSAPOKLEAPAGYON GM CHILLYTELAPADAEVGM ch-cc DSFLRTA-NRGOKSBSOPDYWEILMOAPP-SETERIAPOGVIDLEGHLENLGGMOCOEKTSAPOKLEAPAGYON GM CHILLYTELAPADAEVGM hu-siw bv:fs EVKQQ	<pre>FRT PROSKLEAPAEEDVWEILGOAPP- SEVEE IAFOYGUTDIGKLEAKLEAKUSTAFAYOKLEPAYQUSKOHTILLTVELADDIASVWLINGG 401 FRTPROSKLEAPAEEDVWEILGOAPP- SEVEE IAFOYGUTDIGKLEAKUSTAFAYOKLEPAYQUSKOHTILLTVELADDIASVWLINGG 407 ILRTA-NRGDKKSDSOPUDWHILGKAPP- SEVEE IAFOYGUTDIGKLEAKUSTAFAYOKLEPAYQUN GG CITLATEANPPADVWLINGG 407 ILRTA-NRGDKKSDSOPUDWHILGKAPP- SEVEE IAFOYGUTDIGKLEAKUSTAFAYOKLEPAYQUN GG CITLATEANPPADVWLINGG 407 ILRTA-NRGDKKSDSOPUDWHILGKAPP- SEVEE IAFOYGUTDIGHLEAKUSTAFAYOKLEPAYQUN GG CITLATEYANPPADVWLINGG 337 ILRTA-NRGDKKSDSOPUDWHILGKAPP-SEVEI IAFOYGUTDIGHLEAKUSTAFAYOKLEPAYQUN GG CITLATEYANPPADVWLINGG 230 Motif 2 cont. * Motif 3 * * Set VVENKUKIISKAPA-SEVEI IAFOYGITDIGGKLEFYEEPPUITYTPIEOQUNMCGUVEFCCVSEGAQVWLDOVSLITEEDFFA-FHF 5 SG KVIFFSKGAFTALTISOCSLADDAAQCUVGGE CSTELFVEPPUITYTPIEOQUNMCGUVEFCCVSEGAQVWLDOVSLITEEDFFA-FHF 5 SG KVIFFSKGAFTALTISOCSLADDAAQCUVGGE CSTELFVEPPUITYTPIEOQUNMCGUVEFCCVSEGAQVWLDOVSLITEEDFFA-FHF 5 SG KVIFFSKGAFTALTISOCSLADDAAQCUVGGE CSTELFVEPPUITYTPIEOQUNMCGUVEFCCVSEGAQVWLDOVSLITEEDFFA-FHF 5 SG KVIFFSKGGATUTINCSLADDAAACCUVGEE STELFVEPPUITYTPIEOQUNMKGEVEFCCVSEGAQVWLDOVSLITEESFFA-FHF 5 SG KVIFFSKGGATUTINCSLADDAAACCUVGEE STELFVEPPUITYTPIEOQUNMKGEVEFCCVSEGAQVWLDOVSLITEESFFA-FHF 5 SG KVIFFSKGGATUTINCSLADDAAACCUVGEE STELFVEPPUITYTPIEOQUNMKGEVEFCCVSEGAQVWLDGVEITDESFF- FHF 5 SG KVIFFSKGGATUTINCSLADDAAACCUVGEE STELFVEPPUITYTRUSEQUNMKGBVEFCCVSEGAQVWLDGVEITDESFF- FHF 5 SG KVIFFSKGGATUTINCSLADDAACCUVGEE STELFVEPPUITYTRUSEQUNMKGBVEFCCVSEGAQVWLDGVEITDESFF- FHF 5 SG KVIFFSKGGATUTINCSLADAACCUVGEE STELFVEPPUITYTRUSEQUNMKGBVEFCCVSEGAAVWEDGVEITDESFF- FHF 5 SG KVIFFSKGGATUTINCSLADDAACCUVGEE STELFVEPPUITYTRUSEQUNMKGBVEFCCVSEGAAVWEDGVEITDESFF- FHF 5 SG KVIFFSKGGATUTINCSLADAACCUVGEE STELFVEPPUITYTTSLEQUNMKGGVEFCCVSEGAAVWEDGVEITDESFF - FHF 5 SG KVIFFSKGGATUTINCSLADAACCUVGEE STELFVEPPUITYTTSLEQUNMKGGVEFCCVSEGAAVWEDGVEITGESFF - FHF 5 SG KVIFFSKGGATUTINCSLADAACCUVGEE STELFVEPPUITYTTSLEQUNMKGGVEFCCVSEGAAVWEDGVEITGSFF - FHF 5 SG KVIFFSKGGATUTYTNGGSSALEUTOPSK STAGATUYAGAAAAVKEDAAUKEDAATUNTYGGETATISSTGF SAAUAAAXAAAAVKEPAUAAAXAAAAXAA</pre>
Moleco ch-co c	A CALLER A INCIDENCE DURANCE DURANCE AND A DEPOSITION OF A DEP
fu-cc fu-ccc fu-ccc fu-ccc fu-ccc	TLKVNXMDILSKAPS-BEVETAPONGITDLECKLEGPFLMLDDAVGVTGHITLALIVANPADVVMLONG 337 QQ
hu-slw RFVKQQEEEPQVDVWELL[NAKC-SEVEIIAPCKITDL_GALKELKEMMIREKKESAAFA [ILDPAKQVBGGGVFVVELABPKLEV/MY Motif 2 cont. * Motif 3 * Motif 3 * Motif 2 cont. * Motif 3 * Motif 3 * Motif 3 * Motif 2 cont. * Motif 3 * Motif 3 * Motif 4 cont. * Motif 3 * Motif 4 cont. * Motif 4 cont. * Motif 3 HPS-KIVFENGKGLITINKCILADBAAKGUVGGEKSTELFV GPVLITPLEDQUVFVGEVEFECUVSEGAQVMALGGVELTEESTF ch-cc IQVSGSKITFAIGNERLITINGCILADBAAKGUVGGEKSTELFV GPVLITRSLEDQUVMVGEVEFECUVSEGAQVMALGGVELTEESTF ch-cc IQVSGSKITFAIGNERLITINGCILADBAAKGUVGGEKSTELFV GPVLITRSLEDQUVMVGEVEFECUVSEGAQVMALGGVELTEESTF fu-cc IFMPTG-KITFSVGARGTLITINGCILADBAAKGUVGEKSTELFV GPVLITRSLEDQUVMVGEVEFECUVSEGAQVMALGGVELTEESTF fu-cc IFMPTG-KITFSVGARGTLIFINGCATDBESKIVFAEKSTELFV GPVLITRSLEDQUVMVGEVEFECUVSEGANV MEKGGVELTEESTF fu-cc IGVSGSKIFFAGCGELFINGCAGDESKIVFAEKSTELFV GPVLITRSLEDQUVMVGEVEFECUVSEGANV MEKGGVELTEESTF fu-cc IGGCHILIFSDVQEDGGFAQIINGGCEALLVSEGALEVUGSTADLAVGARDAVFKCVSDENKGKVLNGKEVPDSIKVSHGGVELTEESTF Motif 3 cont. Motif 4 CGGCHILITRAALEDAGHALCTSGGQLAELIVDEKLEVVGSTADLAVGARDAVFKCVSDENKGKVLNGKEVPDSIKVSHGGVHLTDID mo-cc KGGKNHLITRAALEDAGHALCTSGGQLAELIVDEKLEVVGSTADLAVGARDAVFKCVSDENKGKVLNGKEVPDSIKVSHGGVHLTDID hu-fst hu-fst hu-fst Du-cc KGGKNHLITRAALEDAGHALTKSGGSSAKELVDEKLEVVGSTADLAVGARDAVFKCVSDENKGKVLNGKEVPDSIKVSHGGVHLTNGKVLNGKVPDSIKVSHGGVHLTDID hu-slw VEGHVILTIEGATKADAAESWATTGOGSSAKELVDEKLEVVGSTADLAVGARDAVFKCVSDENKGKVLNGKEVPDSIKVSHGKVLNGKVPDAELTSH HU-fst hu-fst	<pre>CQ</pre>
Motif 2 cont. * Motif 3 * hu-fst XIPSS-UVFENVGKG ILTINKCTLADDAA YUVKUGE (FTELFV GEPPLITTPLEDOLMNGG VEFCUVSEGAQVMILLGGVELTBEDSTAI hu-cc IQMSGS (IFFSIGAGUTLISCCSLADDAA YUVKUGE (STELFV GEPPLITRELEDOLMNGG VEFCUVSEGAQVMILLGGVELTBEETT mo-cc IQMSGS (IFFSIGAGUTLISCCSLADDAA YUVGGE (STELFV GEPPLITRELEDOLMNGG VEFCUVSEGAQVMILLGGVELTBEETT	Motif 2 cont. * Motif 3 * * SS-KVVFENVGKRILTINKCTLADDAAVQCVVGCEKCFTELV EPVLITTPLEDQLVAVGQVVFGEVSEEGAQVMANDGVEITREEFF 398 SGSK ITFSIGAR TLATSQCSLADDAAVQCVVGCEKCFTELV EPVLITTPLEDQLVAVGQVVFFCCVSEEGAQVMANDGVEITREEFF *
hu-fst hu-fst SUMSGSNIFSSIGARTIJISQCSIADDAAYCCVVGCKCSTLIVVGPVLITRPLEDCUMVGCVFECVSEGAQVMALGVEITREDT mo-cc SUMSGSNIFSSIGARTIJISQCSIADDAAYCCVVGCKCSTLIVVGPVLITRPLEDCUMVGCVFECVSEGAQVVMLKGUFLTREFT 	SS-SWYENWGKARIITINKCILADDAALGUVAGECFELEV GPPULITPLEDQUVGGBVENAVEVSEGAQVNAMGGVELTEEDSFARHFF 396 SGKVIFESGARTITISQCSLADDAALGCUVGECSELEV GPULITRILEDLUVGGGVEFECCVSEGAQVNALGGVELTEETFAVFF 503 SGSKVIFESGARTITISQCSLADDAALGCUVGECSELEV GPULITRILEDQUVMGEVEFECVSEGAQVNALGGVELTEETFAVFF 503 SGSKVIFESGARTITISQCSLADDAALGCUVGESSTELEV GPULITRILEDQUVMGEVEFECVSEGAVNAEGGVELTEETFAVFF 503 SG-KVIFESGARTITISQCSLADDAALGCUVGESSTELEV GPULITRILEDQUVMGEVEFECVSEGAVNAEGGVELTEETFAVFF 503 SG-KVIFESGARTITISQCSLADDAALGCUVGESSTELEV GPULITRILEDQUVMGEVEFECVSEGAVNAEGGVELTEETFAVFF 503 SG-KVIFESGARTITISQCSLADDAALGCUVGESSTELEV GPULITRILEDQUVMGEVEFECVSEGAVNAEGGVELTEETFAVFF 503 SG-KVIFESGARTITISQCSLADDAALGUVAEESSTELEV GPULITRILEDQUVMGEVEFECVSEGAVNAEGGVELTREFFAVFF 503 SG-KVIFESGARTITISQCSLADDAALGUVAEESSTELEV GPULITRILEDQUVMGEVEFECVSEGAVNAEGGVELTREFFAVFF 503 SG-KVIFESGARTITISUGGFACHLUVESSTELEVGENCLEVESPULTRILEDQUVMGEVEFECVSEGAVNAEGGVELTREFFAVFF 503 SG-KVIFESGARTITISUGGFACHLUVESSTELEVGENCLEVGSTADLTVARDQAVFCVSDENVGVELNGKEVVPDSIKVSIGGVENTLTDDVFA 605 SG-LINSSTKEDSGNITVKINGGSVAELLVQENCLEVQSIADLTVARDQAVFCVSDENVGVELNGKEVVPDSIKVSIGGVENTLTDDVFA 605 SG-LINSSTKEDSGNITVKINGGSVAELLVQENCLEVQSIADLTVARDQAVFCVSDENVKGVELNGKEVVPDSIKVSIGFNELTIDDVFA 606 HILLINEATKEDGCHVKVINGGSVAELLVQENCLEVQSIADLTVARDQAVFCVSDENVKGVELNGKEVVPDAINITHIGIHALTIDDVFG 608 SG-LINSSTKEDSGNITVKINGGSVAELVQENCLEVGSIADITVARDQAVFCVSDENVKGVUNGKVENLINGKEVVPDAINITHIGIHALTIDDVFG 609 SG-LINSSTKEDSGNITVKINGGSVAELVQENCLEVGSIADITVARDQAFFCVSDENVKGVUNGKVENGKVANGGVENTINING 700 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-NLSALHENGVENTFORSPF 195VPGGAC-
<pre>store store s</pre>	SGKNTFFSVGARTITISQCSLADDAAYCCVVGEECSTELFVGPPLITREEDGLVWGGPVEECVSEEGAQVWLDGGELTRETFFFFF 500 SGKNTFFAIGMENIITINGCLADDAAYCCVVGEESTELFVGPPLITREEDGLVWGGPVEECVSEEGAQVWLDGGELTRETFFFFF 503 TG-MIFESVGARTITINGCLADDAAYCCVVGEESTELFVGPPLITHFLEDGWVMGGPVEECCVSEEGAQVWLDGGELTRETFFFFF 503 TG-MIFESVGARTITINGCALDDAAYCCVVGEESTELFVGPPULITHFLEDGWVMGGPVEECCVSEEGAQVWLDGGVELTDESFFVFF 503 TG-MIFESVGARTITINGCALDDAAYCCVVGEESTELFVGPPULITHFLEDGWVMGGPVEECCVSEEGAQVWLDGGVELTDESFFVFF 503 TG-MIFESVGARTITINGCALDDAAYCCVVGEESTELFVGPPULITHFLEDGWVMGGPVEECCVSEEGAQVWLDGGEITDESFFVFF 503 TG-MIFESVGARTITINGCATDSSTVTAGDECSTELFVGPPULITAGDEAGEVELCCVSEEGAQVWLDGGVEITDESFFVFF 503 THLIINEALCTSGCGALEFINDESCUTVERGEVSLEVGSTADLTVARSQAVFCEVSDEVVTGKWTNGVEVRESKITISVGFHKLVIDDVPP 503 THLIINEATLEDAGNAVRISGGSLAELIVGEKLEVQSIADLTVARAQAVFCEVSDEVVTGKWLMGKELVPDNIKVSHIGVHKLTIDDVTP 605 TG-LIINESTKEDGGTVKKINGGVSVAELLVQEKLEVQSIADLTVARAQAVFCEVSDEVVTGKWLMGKELVPDNIKVSHIGVHKLTIDDVTP 605 TG-LIINESTKEDGGTVKKINGGVSVAELLVQEKLEVQSIADLTVARAQAVFCEVSDEVVTGWLMGKELVPDNIKVSHIGTHKVSHIGTHKLTIDDVTP 605 TG-LIINESTKEDGGTVKKINGGVSVAELLVQEKLEVQSIADLTVARAQAVFCEVSDEVVRGTWLMGKELVPDRIKVSHIGTHKLTIDDVTP 605 TG-LIINESTKEDGGTVKKINGGVSVAELLVQEKLEVQSIADLTVARAQAVFCEVSDEVVRGTWLMGKELVPDRIKVSHIGTHKLTIDDVTP 605 TG-LIINESTKEDGGTVKKINGGVSVAELLVQEKLEVQSIADLTVARAQAFFCEVSDEVVRGTWLMGKELVPDRIKVSHIGTHKLTIDDVTP 605 TG-LINESTKEDGGTVKKINGGVSVAELLVQEKLEVGSIADLTVARAQAFFCEVSDEVVRGTWLMGKAVPDRIKTHING 605 TG-VITVVRINGGSSALSVDIPLKLITPIDQTVNLGKLELFGSTSDENT 605 TG-VVFDKKLCGSLAGUNKKINFVNIGGOSSALSVDIPLKLITPIDQTVNLGKLELGESSTGFT 706 TG-VVFDKKLCGSLAGUNKKINFFT 705 TVFVPGGAL-SLSALLHEMEVKIDFVPGEPPIHLDCGGSTPD-TIVVVGANLELDVFISGPAPTVNQ 705 TVFVGGALGSLSALLHMEVKIDFVPGEPPIHLDCGGSTPD-TIVVVGANLELDVFISGPAPTVNQ 705 TVFVGGALGSLSALLHMEVKIDFVPGEPPIHLDCGGSTPD-TIVVVGANLELDVFISGPAPTVNQ 705 TVFVGGALGSLSALLHMEVKIDFVPGEPPIHLDCGGSTPD-TIVVVGANLELDVFISGPAPTVNQ 705 TVFVGKKLCETGGVRVETKNRIFTYGGEKEDGGVTVVVNNJGKSLSELFUPJGPPPIHL 705 TVFVGGALGSSALLFEXCUPJGSTVFISGEFVESSCUP 705 TVFVDKKLLCETGGVRVETKNRIFTGGESTVIGSGSVTVVVNNJGKSLSEDGUNKIVVDVDDAPAAPFISNVGEDSCTVVMPPATG
ch-cc flovSGKVIFAJGKKTILINKCSLADDAA'CCVVAEESSFTLEVKEPPLITHPLEDGWVKVGEVSECSVSEGAVVKEDCOVELTREETF- fu-cc FHPTG-NIFFSVGNM_YLTINKCSLADDAA'CCVVGEESTTELFVKEPPLITHPLEDGWVKGGVELECVSEGAVVKEDGVELTREETF- hu-slv FHPTG-NIFFSVGNM_YLTINKCSLADDAA'CCVVGEESTTELFVKEPPLNVTKQLEDTAA'CGVSEGAVVKEDGVELTDESF Motif 3 cont. Motif 4 hu-fst KOGGRHILINEAMEDAGHA'LCTSGGALRELIVESQLEVCDIADLTVKAEGVSDEKVTGKWINGVEVPPSKETTISVGEFKLVTG mo-cc CGGRHILINEAMEDAGHA'LCTSGGALRELIVESKQLEVUQDIADLTVKAEGVSDEKVTGKWINGVEVPPSKETTISVGEFKLVTG mo-cc CGGRHILINEAMEDAGHA'LCTSGGALRELIVESKQLEVUQDIADLTVKAEGVSDEKVTGKWINGKELVPDN LKVSHIGGVKKLTID mo-cc CGGRHILINEATEDAGHA'LCTSGGALRELIVESKQLEVUQDIADLTVKAEGVSDEKVRGWUNGKELVPDN LKVSHIGGVKKLTDD ch-cc CGGRHILINEATEDAGHA'LCTSGGALRELIVESKUEV'GSIADLTVGAKDQAVFCCVSDENVRGWUNGKELVPDN LKVSHIGGVKKLTDD ch-cc CGGRHILINEATEDAGHA'LCTSGGALRELIVESKUEV'GSIADLTVGAKDQAVFCCVSDENVRGWUNGKELVPDN LKVSHIGGVKKLTDD mo-cc CGGRYGLINEATKBOCHTVKINGGSSAKLSVDLPEKLEV'GSIADLTVGAKDQAVFCCVSDENVRGWUNGKELVPDN LKVSHKGEIKKVXING Motif 4 cont. Motif 5 CGCATGESSAUNTELSKVCTVCKGESSAKLSVDLPEKLEVGSIADLTVGAKDQAVFCCVSDENVRGWUNGAKDAVKESVENDER Motif 5 CGCATGESSAUNTELSKVCTVCKGESSAKLSVDLPEKLEFIGSTEN Notif 5 CGCATGESSAUNTELSKVCTVCKGESSAKLSVDLPEKLEFIG CDCATGESSAUNTELSKVCTVCKGESSAKLSVDLPEKLEFIG DEDYTFVDGALSLSAUNTELSKVCTVCKGESSACLSVDLPEKLEFIG CDCATGESSAUNTELSKVCTVCKGESSAUNTELSCCGGSSACLSVDLPEKLESSAUNTVAGNUNGKELKDVVIGGSPATVNCGATGCSSAUNTVIK Motif 5 CGCATGESSAUNTELSKVCTVCKGESSAUNTELSCCGGSSAUNTVIN DEDYTFVDGALSLSAUNTELSKVCTVCKGESSTD-TIVVYGNUNGKLENDVVIGGPATVNCGATGCSSTGGTIGUSS CDCATGESSAUNTELSKVCTSSYVESSACSTD-TIVVYGNUNGKLENDVVIGGSPATVNCGATSSTGGTIGUSSTGGTIGUSS Motif 5 CDCATGESSAUNTUSSYKSGCTISSAUNTVENSINGESSTGD-TIVVYGNUNGKLENDVVIGGSAUNTVCGSSTGGTIGUSSTGGTIGUSSTGGTISSTGGTIGUSSGC CDCC CDCC CDCC CDCC CDCC CDCC CDCC CAFFLINKASTYMULNELVENSINGESSYVESSAGEBEDGENTIKVTNVGEDVASIFLQVVVDPDFPERVENGGESCVVGPPFEDGGPT Motif 5 CDCVYFPDGALSSISAUNTENTSSTGGCISSTGGTISSSGC CDCC CDCC CDCC CDCC CAFFLINKASTSWMUNSCHCGGUVENKGGUCSSTGGCISSSGC CDCC CAFFLINKSSTWMULNFDLGUSSGGU	SGS XIFF A IGNUSTITINCS LADDAA E CVVAEE SFELFVE PLITHPLEDOMVMUGE VEFECEVSEGATV ME DOVELT BEET FFF 439 TG-KIFF SUGNKYLTINNCSLADDAA CCVVGEEKSTELFVE PPVIVKNLEDOMVMGDAVELCEVSEGANVKWE DOVELT BEET KFF 439 ST-KIFF SUGNKYLTINNCSLADDAA CCVVGEEKSTELFVE PPVIVKNLEDDTAYCGEVVELECEVSEGANVKWE NOGVELT BOEST KFF 439 ST-KIFF KKCORIFFINNCOMTDISE YVTAGDEKCSTELFVE PPINVTKOLEDTAYCGEVVELECEVSEGANVKWE NOGVELT BOEST KFF 439 ST-KIFF KKCORIFFINNCOMTDISE YVTAGDEKCSTELFVE PPINVTKOLEDTAYCGEVVELECEVSEGANVKWE NOGVELT BOEST KFF 439 ST-KIFF KKCORIFFINNCOMTDISE YVTAGDEKCSTELFVE PPINVTKOLEDTAYCGEVVELECEVSEGANVKWE NOGVELT BOEST KFF 439 ST-KIFF KKCORIFFINNCOMTOS Y ALLIVE KLEVOSIADLIVGAK DOAVF KCVSDENVRGVENGEKLIVDOSIKVSHIGVENHE VIDUVTPA 605 KHLILISDVVQERGEVQUTNGGVEALIVDEKKLEVOSIADLIVGAKDOAVF KCVSDENVRGVENGEKLVPDNIKKVSHIGVENHE VIDUVTPA 605 KULLINESTKEISGH VVKNTNGGVSVALUVEKKLEVVQSIADLIVGAKDOAVF KCVSDENVRGVENGKELVPDNIKVSHIG VHRLTIDVTPA 605 KULLINESTKEISGH VVKNTNGGVSVALUVEKKLEVVQSIADLIVGAKDOAVF KCVSDENVRGVENGKELVPDNIKVSHIG VHRLTIDVTPA 606 KHLILIEGATKADAAE SVMTTGGOSSAKLSVDIPEKKLEVIGSIADLIVGAKDOAVF KCVSDENVRGVENGKELVPDNIKVSHIG HRLTIDVFG 644 KHLILIEGATKADAAE SVMTTGGOSSAKLSVDIPEKKLEVIGSIADLIVGAKDOAVF KCVSDENVRGVENGAVKNGALVEDAALNITHIG HRLTIDVFG 644 KHLILIEGATKADAAE SVMTTGGOSSAKLSVDIPEKKLEVIGSIADLIVVAGNILLIPUTSGOPATVING ALVGENAAPAPAPAPEDTGDSE 716 NOTIF 4 cont. * Notif 5 (Cardiac specific insert) TFVPBGFAC-NLSALHEMEVNIDEVPROEPPKIHLDCGGSTPD -TIVVVAGNILLIPVISGOPATVING XITQAKABAPAPAPAPEDTGDSE 716 NSFVPSGFAC-NLSALHEMEVNIDEVPROEPPKIHLDCGGSTPD -TIVVVAGNILLIPVISGOPATVING XITGAKABAPAPAPAPEDTGDSE 716 NSFVPSGFAC-NLSALHEMEVNIDEVPROEPPKIHLDCGGSTPD -TIVVVAGNILLIPVISGPPATVUNG XITGEKSTGPTIPYEGSAGNEDSE 716 NOTIF 5 cont. MOTIF 5 cont. NOTIF 6 WVFDKKLLCETGG VRVETKDRSIFING SEVICESSVING XALVVAGNILLELIPVISGPPATVING XINGKELVPESSOFTING XITGSENSS 708 NOTIF 5 cont. NOTIF 6 cont. NOTIF 6 WVFDKKLLCETGG VRVETKDRSIFING SEVICESSOFTINTYNEVGEDVASIFILVVIDVDAPAPAFISNGEDSTVOWE PPAYDGGPVICKIL 804 WVFDKKLLCETGG VRVETKDRSIFING SEVICESSO
fu-cc FHPTG-EXTFSSVGMCRITTINCSLADDAALCCVVCEELSTELEVCEPPUVVVNLEDQMVMCGDEVELECEVSEGAVVAVEKGVSLTIDESFR hu-slw Motif 3 cont. Motif 4 hu-fst GGRHLLINSDVVGEGGGALGLIVGEGCALELIVGEGCUSUUDDADLTVLASEGAVFCEVSEDAVKAVEKGGEIVEGGAVVAVEKGGEIJPGFXS Motif 3 cont. Motif 4 hu-fst GGRHLLINSDVVGEGGGALGLIVGEGCALELIVGEGCUSUUDDADLTVLASEGAVFCEVSEDAVRGWLAGKELVPDSLIXVSHGRVHLTDI ch-cc GGRKHLLINSTLEDAGGVAVRTSGGSIALLVQEXKLEVQSIADLAVGAKDQAVFCEVSEDAVRGWLAGKELVPDSLIXVSHGRVHLTDI ch-cc GGRKHLLINSTLEDAGGVAVRTSGGSIALLVQEXKLEVQSIADLAVGAKDQAVFCEVSEDNRGWLAGKELVPDSLIXVSHGRVHLTDI fu-cc Hu-st WGKKHLLIEGATKADAESSMTGGSSAKLIVQEXKLEVQSIADLAVGAKDQAVFCEVSEDNRGWLAGKELVPDSLIXVSHGRVHLTDI fu-cc DGCKPLFLINSTKEDSGYTVKTNGGSSAKLIVQEXKLEVQSIADLAVGARDQAVFCEVSEDNRGWLAGGSVKPDA INITHIGNIKLTISI fu-cc DGCKPLFLINSTKEDSGYTVKTNGGSSAKLIVQEXKLEVQSIADLAVGARDQAVFCEVSEDNRGWLAGGSVKPDA INITHIGNIKLTISI fu-cc DGCKPLFLINSTKEDSGYTVKTNGGSSAKLIVQEXKLEVQSIADLAVGARDQAVFCEVSEDNRGWLAGGSVKPDA INITHIGNIKLTISI fu-cc DGCKPLFLINSTKEDSGYTVKTNGGSSAKLIVQEXKLEVQSIADLAVGARDQAVFCEVSEDNVGGVYNGGAVKPDA Nu-fst Motif 4 cont. Motif 4 cont. Motif 4 cont. Motif 5 (Cardiac specific ins hu-fst 	TG-MITESVGNMYLTINNCSLADDAA CCVVGERSTTELFVKEPPVIVKNLEDQMMKGDWELECEVSEGANVKWERGVELTRDESFKFR 439 ST-MITESVGNRYLTINNCQMTDDSEYVTAGDEKCSTELFVKEPPVIVKNLEDQTAYCGEVELECEVSEDANVKWERGVETRDESFKFR 392 Motif 3 cont. Notif 4 SHILIFSDVVQEDRGRQVITNGGOCAAELIVSEKQEVQSTADLMVGAKDQAVFKCEVSDEKVTGKWYKNGVEVRPSKNITISHVGFHLLVIDDVRP5 503 SHILINEAMLEDAGHVALCTSGGQALRELIVDEKLEVVQSTADLMVGAKDQAVFKCEVSDENVRGVKLKNGKELVPDSTKKVSHIGVHLLTDDVTPA 609 SHILINESTKEDSGHTVKTNGGSVALLVDEKLEVVQSTADLMVGAKDQAVFKCEVSDENVRGVKLKNGKELVPDSTKKVSHIGVHLLTDDVTPA 609 CQ LIINESTKEDSGHTVKTNGGSVALLVDEKLEVVQSTADLTVARDQAVFKCEVSDENVRGVKLKNGKELVPDSTKKVSHIGTVHLTDVTPA 609 HHLLINEATKEDCGHYKVKTNGESSMALLVDEKLEVVQSTADLTVARDQAVFKCEVSDENVRGVKLKNGKELVPDSTKKVSHIGTVHLTDVTPA 608 HHLLINESTKEDSGHTVKTNGGSVALLVDEKLEVQSTADLTVARDQAVFKCEVSDENVRGVKLKNGKELVPDSTKVSHIGTHLTDVTPA 608 HFLINESTKEDSGHTVKTNGGESSMALLVDEKLEVQSTADLTVARDQAVFKCEVSDENVRGVKJLKNGKELVPDSTKVSHIGTHLTDTVFG 608 HFLINESTKEDSGHTVKTNGGESSMALLVDEKLEVQSTADLTVARDQAVFKCEVSDENVRGVKJLKNGKELVPDSTKISHIGTHLTDTVFG 609 HFLINESTKEDSGHTVKTNGGESSMALLVDEKLETVQSTADLTVARDQAVFKCEVSDENVRGVKJDATNTHIGTHLTDTVFG 544 HILLIEGATKADAAE SVMTTGGQSTALLVDEKLETVQSTADLTVARDQAFFCCVSDENVRGVKJDATNTHIGTHLTDTVKFG 544 HILLIEGATKADAAE SVMTTGQSSTKLSVDI PLKILTPLOTVNLGKALDDVISGPPTVTVGALTVGKKKPAAPAPAPEDFC 710 717VPDGALGSLSALNFLEIVVFVQEPFTHLDCGSGTSENATVVVAGNLLDVSIGGPPTVTVGATTGKKRASAGPPAPAPEDGSDE
Mu-siw Motif 3 cont. Motif 4 hu-fst Motif 3 cont. Motif 4 hu-fst Motif 3 cont. Motif 4 hu-fst Motif 4 Motif 4 hu-fst Motif 4 Motif 4 hu-fst Motif 5 Motif 4 hu-fst Motif 4 Motif 4 hu-fst Motif 4 Motif 4 hu-fst Motif 4 Motif 4 hu-fst Motif 5 Motif 4 hu-fst Motif 4 Motif 4 Motif 4 fu-ce Motif 4 Cont. * hu-fst Motif 4 Cont. * Motif 5 (Cardiac specific instruction) hu-fst DeGOTTEVEGEAC-NI.SALHENEVELEVERDEPPENEDCEGETED - TIVVAGNELEUVESTGEPPEVATURGE (Cardiac specific instruction) hu-fst DeGOTTEVEGEAC-NI.SALHENEVETEPPENEDEPPENEDCEGETED - TIVVAGNELEUVESTGEPPEVATURGE (Cardiac specific instruction) hu-fst Motif 5 Cont. * Motif 5 (Cardiac specific instruction) hu-fst Motif 5 Sont. * Motif 5 (Cardiac specific instruction) hu-fst Motif 5 <t< td=""><td>ST-ENTRAGONILIFIANCONTDISETVIAGDERCSTELLVER PINNTROLEDTIATCGEVELECVSEDAWAWER NGEETPERKS RIE 392 392 Motif 3 cont. Motif 4 BILIESDVVQEDRGR QVITNGGQCEAELIVER QLEVLQDIADLTVASEQAVERCEVSDEKVTGKWYRGVERPSKETTISHVGREHLVIDDVRPS 503 SHHLLINEAMLEDAGHVALCTSGGQALRELIVER QLEVLQDIADLTVASEQAVERCEVSDEKVTGKWYRGVERPSKETTISHVGREHLVIDDVRPS 503 SHHLLINEAMLEDAGHVALCTSGGQALRELIVERKLEVVQSTADLMVGARDQAVERCEVSDENVRGVULNGKELVPDSTKVSHIGVHRLTIDDVTPA 609 CALINESTKEDSGUTVKKNGGSSALLUVERKLEVVQSTADLTVKARDQAVERCEVSDENVRGVULNGKELVPDSTKVSHIGVHRLTIDDVTPA 605 CYLLINESTKEDSGUTVKKNGGSSALLUVERKLEVVQSTADLTVKARDQAVERCEVSDENVRGTHVKNGKELVPDSTKVSHIGVHRLTIDDVTPA 605 MOLIF 4 CALINESTKEDSGUTVKKNGGSVALLVDEKKLEVVQSTADLTVKARDQAVERCEVSDENVRGTHVKNGKEVPDSTKISHIGVHRLTIDDVTPA 605 MOLIF 4 CALINESTKEDSGUTVKKNGGSVALLVDEKKLEVVQSTADLTVKARDQAVERCEVSDENVRGTHVKNGLVDDSTKISHIGVHRLTINLVFG 644 MOLIF 4 CALINESTKEDSGUTVKKNGDSSALLUVERKLEVVDQSPKILLDCSGKTSENAIVVVAGNULLDVSIGGPAPTVINQALTQGNKARDABPAPAPEDTGDSDE 710 YSVPSGFAC-NLSALLHEMEVN IDEVPROEPPKILLDCSGKTSENAIVVVAGNULLDVSIGGPAPTVINQAUTVGNKARABABPAPAPEDAGDSE 710 YSVPSGFAC-NLSALLHEMEVN IDEVPROEPPKILLDCGGSPD - TIVVVAGNULLDVSIGGPAPTVINQAUTVGNKARABABPAPAPEDAGDSE 710 YSVPSGFAC-NLSALLHEMEVN IDEVPROEPPKILLDCGGSPD - TIVVVAGNULLDVSIGGPAPTVVNCKNTKGELMPAPAPEDAGDADE 710 YSVPSGFAC-NLSALLHEMEVNIDEVDREPPKILLDCGGSPD - TIVVVAGNULLDVSIGPAPTVV</td></t<>	ST-ENTRAGONILIFIANCONTDISETVIAGDERCSTELLVER PINNTROLEDTIATCGEVELECVSEDAWAWER NGEETPERKS RIE 392 392 Motif 3 cont. Motif 4 BILIESDVVQEDRGR QVITNGGQCEAELIVER QLEVLQDIADLTVASEQAVERCEVSDEKVTGKWYRGVERPSKETTISHVGREHLVIDDVRPS 503 SHHLLINEAMLEDAGHVALCTSGGQALRELIVER QLEVLQDIADLTVASEQAVERCEVSDEKVTGKWYRGVERPSKETTISHVGREHLVIDDVRPS 503 SHHLLINEAMLEDAGHVALCTSGGQALRELIVERKLEVVQSTADLMVGARDQAVERCEVSDENVRGVULNGKELVPDSTKVSHIGVHRLTIDDVTPA 609 CALINESTKEDSGUTVKKNGGSSALLUVERKLEVVQSTADLTVKARDQAVERCEVSDENVRGVULNGKELVPDSTKVSHIGVHRLTIDDVTPA 605 CYLLINESTKEDSGUTVKKNGGSSALLUVERKLEVVQSTADLTVKARDQAVERCEVSDENVRGTHVKNGKELVPDSTKVSHIGVHRLTIDDVTPA 605 MOLIF 4 CALINESTKEDSGUTVKKNGGSVALLVDEKKLEVVQSTADLTVKARDQAVERCEVSDENVRGTHVKNGKEVPDSTKISHIGVHRLTIDDVTPA 605 MOLIF 4 CALINESTKEDSGUTVKKNGGSVALLVDEKKLEVVQSTADLTVKARDQAVERCEVSDENVRGTHVKNGLVDDSTKISHIGVHRLTINLVFG 644 MOLIF 4 CALINESTKEDSGUTVKKNGDSSALLUVERKLEVVDQSPKILLDCSGKTSENAIVVVAGNULLDVSIGGPAPTVINQALTQGNKARDABPAPAPEDTGDSDE 710 YSVPSGFAC-NLSALLHEMEVN IDEVPROEPPKILLDCSGKTSENAIVVVAGNULLDVSIGGPAPTVINQAUTVGNKARABABPAPAPEDAGDSE 710 YSVPSGFAC-NLSALLHEMEVN IDEVPROEPPKILLDCGGSPD - TIVVVAGNULLDVSIGGPAPTVINQAUTVGNKARABABPAPAPEDAGDSE 710 YSVPSGFAC-NLSALLHEMEVN IDEVPROEPPKILLDCGGSPD - TIVVVAGNULLDVSIGGPAPTVVNCKNTKGELMPAPAPEDAGDADE 710 YSVPSGFAC-NLSALLHEMEVNIDEVDREPPKILLDCGGSPD - TIVVVAGNULLDVSIGPAPTVV
Motif 4 cont. Motif 4 cont. Motif 4 cont. Motif 4 Motif 4 Motif 6 Motif 4 Motif 4 Motif 4 Motif 4 Motif 4 Motif 6 Motif 4 Cont. Motif 6 Motif 4 Cont. Motif 6 Motif 5 Contexprete Structure Structur	Motil 3 cont. Motil 3 cont. Motil 4 cont. Motif 4 cont. Motif 4 cont. Motif 4 cont. Motif 4 cont. Motif 5 cont. Motif 5 cont. Motif 5 cont. Motif 5 cont. Motif 5 cont. Motif 6 cont. Motif 7 Motif 7 Mot
hu-cc mo-cc Mo-ck Mo	AMBLINEAMLEDAGI VALCTSGGQALRELIV QEKLEVYQSTADLAWGAKDQAVFKCVSDENVRGWULNGKELVPDS XIKVSHIGVNLTIDUTPA 609 AMBLINEAMLEDAGI VAVFTSGQQIAELIVQEKLEVYQSTADLAWGAKDQAVFKCVSDENVRGWULNGKELVPDS XIKVSHIGVNLTIDUTPA 605 ACLINEATKEDSGUVAVETSGQQUAELIVQEKLEVYQSTADLAWGAKDQAVFKCVSDENVRGWULNGKELVPDS XIKVSHIGVNLTIDUTPA 605 ACLINEATKEDSGUVAVETSKDGVSVALLVDEKLEVYQSTADLAWGAKDQAVFKCVSDENVRGWULNGKELVPDS XIKVSHIGVNLTIDUTPA 605 ACLINEATKEDSGUVAVETSKDSSALLVDEKLEVYQSTADLTVIARDQAVFKCVSDENVRGWULNGKEVVPDSTIKISHIG INHLTIDUTPA 605 ANDI 1 ACLANDAVE KLEVYQSTADLTVIARDQAVFKCVSDENVRGWULNGKEVVPDSTIKISHIG INHLTIDUTPA 605 ANDI 1 ACLANDAVE KLEVYQSTADLTVIARDQAVFKCVSDENVRGWULNGKEVVPDSTIKISHIG INHLTIDUTPA 606 ANDI 1 ACLANDAVE KLEVYQSTADLTVIARDQAVFKCVSDENVRGWULNGKEVVPDSTIKISGI KUKIKSKUPDSTITISHIG INHLTIDUTPA 605 ANDI 1 ACLANDAVE KLEVYQSTADLTVARDQAVFKCVSDENVRGWULNGKEVVPDSTIKISGI KUKIKSKUPDSTITISHIG INHLTIDUTA 496 Motif 4 Cont. * Motif 5 (Cardiac specific insert) 531 YSFVPEGFAC-NLSALINFLEV KUPVPESEPKI INDCOGSTSENATIVVAGNILRLDVFISGPPATVINQ KATQGNKABAPAPAPEDTADASDASD
mo-cc bocksoline intervention mo-cc bocksoline intervention mo-cc bocksoline intervention mo-cc bocksoline intervention bu-siw bocksoline intervention bu-siw bu-siw bocksoline intervention bu-siw bu-siw bocksoline intervention bu-siw bu-siw bocksoline intervention bu-siw bu-siw bu-siw bu-siw bocksoline intervention bu-siw bu-siw bocksoline intervention bu-siw bu	SHELITIKATLEDAGIYAVRTSGGQSIALIVQEKLEVQSIADLAVGAKDQAVFKCSVSDENVRGVKLKNGKELVPDNIKVSHIGVKHITDDVTPA 605 SKLITINESTKEDGSUTVKTNGGVSVALIVQEKLEVQSIADLTVARDQAVFKCSVSDENVRGVKLKNGKEVVPDENIKISHIGVKHLTDDVTPA 605 SVELTINESTKEDGSUTVKTNGGVSVALIVQEKLEVQSIADLTVARDQAVFKCSVSDENVRGVKLKNGKEVVPDENIKISHIGVKHLTDUTVFG 608 SVELTINESTKEDGSUTVKTNGGVSVALIVQEKLEVQSIADLTVARDQAVFKCSVSDENVRGVKKNGKEVVPDENIKTISKIGVKEVALINKLTG 108 MOTIF 4 * MOTIF 5 VISTVPGGALGSISALINFLEIVFVVPOEPFKIKLDCGGTSENAIVVAKDQAVFKCSVSDENVRGVKKNGLPVQESDLKVVKKGIHKLVIANALTS 496 MOTIF 4 * MOTIF 5 VISTVPGGAC-NLSALINFLEIVFVPOEPFKIKLDCGGSTSENAIVVAGNILLEVSIGPPVTVNG AITQGNKAPAPAPAPEDTGDSDE 710 VISTVPGGAC-NLSALINFLEVKVIDFVPOEPFKIKLDCGGSTPD-TIVVAGNILLEVPISGDPAPTVNG AITQGNKAPAPAPAPEDTGDSDE 710 VISTVPGGAC-NLSALINFLEVKIDFVPOEPFKIKLDCGGSTPD-TIVVAGNILLEVPISGDPAPTVVNG AITQGNKAPAPAPAPEDTGDSDE 710 VISTVPGGAA-NLSALINFLEVKIDFVPOEPFKIKLDCGGSTPD-TIVVAGNILLEVPISGDPTVVNG AVTGGKKELVHOSSE 706 VISTVPGGAA-SLSALINFLEVKIDFVPOEPFKIKLDCGGSTPDTIVVAGNILLEVPISGDPTVVVG AVKGKELVHOSSE 706 VISTVPGGAA-SLSALINFLEVKIDFVPOEPFKIKLDCGGSPD 706 VISTVPGGAA-SLSALINFLEVKIDFVPOEPFKIKLDCGGSTS 706 VISTVPGGAA-SLSALINFLEVKIDFVPOEPFKIKLDCGGSTS 706 VISTVPGGAA-SLSALINFLEVKIDFVPOEPFKIKLDCGGSTS 706
ch-cc KDGKKQLLINESTKEDSGHYUKTNGGGSVALLUQEKLEIGGSLADLIVKARDQAVEKCUSDENVKGTKIKLNGKVVDEDEIKISH GAIHKLITIE hu-slw VEGKHILLIEGATKADAAE SVMTTGGOSSAKLSVDI PLKILTPLTPQTVNLGKLICLGSIS-ENIPGKWTNGLPVQESDILKVVHKGRIHKLVIA Motif 4 cont. * * Motif 5 (Cardiac specific in hu-slw VEGKHILLIEGATKADAAE SVMTTGGOSSAKLSVDI PLKILTPLTPQTVNLGKEICLGSIS-ENIPGKWTNGLPVQESDILKVVHKGRIHKLVIA Motif 4 cont. * * Motif 5 (Cardiac specific in bu-stw VEGKHILLIEGATKADAAE SVMTTGGOSSAKLSVDI PLKILTPLTPQTVNLGKEICLGSIS-ENIPGKWTNGLPVQESDILKVVHKGRIHKLVIA Motif 4 cont. * * Motif 5 (Cardiac specific in bu-scw DEADYSVPEGFAC-NI.SALHHMEVKIDFVPGEPPHILDCSGKISENAIVVVAGNKLELDVFISGPAPTVIKQKAIKGG ch-cc DEADYSVPEGFAC-NI.SALHHMEVKIDFVPGEPPHILDCSGYED-TIVVVAGNKLELDVFISGPAPTVVKQKVHKGEUVHSOR- bu-stw DEGVYFVPGGFA-NI.SALLQEEVKIDFVPRQEPPHILDCCGSYED-TIVVVAGNKLELDVFISGPAPTVVKQKVNKKGEUVHSOR- hu-stw DEGVYFVPGGFA-NI.SALLQEEVKIDFVPRQEPPHILDCCGSYED-TIVVVAGNKLELDVPISGPPTVVKQKVNKKGEUVHSOR- bu-stw DEGVYFAPDANV-TIPAVHVDPPHILDCLGOSPD-TIVVVAGNKLELDVPISGPPTVVKQKVNKKGEUVHSOR- hu-stw DEGVYFAPDANV-TIPAVHVDPPHILDCLGOSPD-TIVVVAGNKLELDVPISGPPFXVWKKGEUVHSOR- Notif 5 cont. Motif 6 	ACULINESTKEDSGNTTVKINGGVSVALIVQEKLEVVQSIADLTVKARQAVFKCSVSDENVKGINLENGKEVVDPEIKISHGGINKLITEDUTPG 608 HFLIINEATKEDCGNVKVKINGGVSVALIVQEKKLEIVQSIADLTVKARQAVFKCSVSDENVKGINLENGKEVVDPEIKISHGRINKLITEDUTPG 608 HFLIINEATKEDCGNVKVKINGGVSVALIVQEKKLEIVQSIADLTVKARQAVFKCSVSDENVKGINLENGLVVDPALINITHGINKLITEDUTPG 608 KHILIIEGATKADAAE SVMTTGGQSSAKLSVDIPEKILITPLTDOTVNLGKEICLKCEIS-ENIPGKWTENGLPVQESDELKVVHKGINKLVARG 496 MOTIF 4 cont. * Motif 5 (Cardiac specific insert) YTFVPGCALGSISALIHEMEVK IDFVPKOEPPKIHLDCGGTPD-TIVVVAGNILLDVSITGEPPTVIKORITCGIKAPAPAPAPAPEDTGDSDE
IU-SIW VGKKILLITEGATKADAGES VALTGGSSAKLSVDL PLKILETING VINLGKELCKCIS = ENIPGKWTKIGLPVQESD LKVVKGETHKUTAL hu-siw Motif 4 cont. * Motif 5 (Cardiac specific ing hu-fst DEGNYTFVPGGALGSISALINFLETING VPKQEPP IILDCSGKTSENAIVVVGNITELDVSITGEPPFVATULGG (Cardiac specific ing hu-fst DEGNYTFVPGGALGSISALINFLETING VPKQEPP IILDCSGKTSENAIVVVGNITELDVSITGEPPFVATULGC (Cardiac specific ing hu-fst DEGNYTFVPGGALGSISALINFLETING VPKQEPP IILDCSGSTGEIP - TIVVVGNITELDVFISGPAPTVVNQ KTVTGKKASAGPH PDAPEDAGAI ch-cc DEADYSFVPEGFAC-NISALINFLEVINFVPREEPP IILDCSGSTP - TIVVVGNITELDVFISGPAPTVVNQ KTVTGKKASAGPH PDAPEDAGAI hu-cs DEADYSFVPEGFAC-NISALINFLEVINFVPREEPP IILDCSGSTP - TIVVVGNITELDVFISGPAPTVVNQ KTVTGKKASAGPH PDAPEDAGAI hu-cs DEADYSFVPEGFAC-NISALINFLEVINFVPREEPP IILDCSGSTP - TIVVVGNITELDVFISGPAPTVVNQ KTVTGKKASAGPH PDAPEDAGAI hu-slw DEGNYFVPGGAL-SISALINFLEVINFVPREEPP IILDCSGSTP - TIVVVGNITELDVFISGPAPTVVNQ KTVTGKKASAGPH PDAPEDAGAI hu-slw DEGNYFVPGGAL-SISALINFLEVINFVGERATISTVVVGNITELDVFISGPAPTVNG KASTGPTETYVNQ KKULLKISNE hu-slw DEGNYFVPGGAL-SISALINFLEVINFVGERATISTVVTVANKILLUFISGPAPTVNTKOK KSGTUNKSKULGKIGKGKTKUSSKVSKULGKTGGTNG KKUKVSGTUNG KASTGPTETYVKYGKKULGKTGGTNG KKUKVSGTUNG KASTGPTETYVKYGKULGKTGGTNG KKUKVSGTUNG KASTGPTETYVKYGKULGKTGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKUGGTNG KKUKVGCKULGKUGGTNG KKUKVGCKULGKTGGTNG KKUKVGCKULGKTGGTNG KKU	Motif 4 cont. * Motif 5 (Cardiac specific insert) SPUPEGFAC-NLSALINTETIVEVPYOEPRIHLOCGGIPSIP * * Motif 5 (Cardiac specific insert) SPUPEGFAC-NLSALINTETIVEVPYOEPRIHLOCGGIP * * * 700 SPUPEGFAC-NLSALINTETIVEVID * * Motif 5 (Cardiac specific insert) YFAPDAANU-TLPAVHID
Motif 4 cont. * * Motif 5 (Cardiac specific innerset) hu-fst DEGVTVVDGVALGSLSALNELETKVEYVEVQEPPIHLCSGKTSENTUVAGNKLELDVSIGEPPPVATULGG	Motif 4 cont. * * Motif 5 (Cardiac specific insert) YTTVPDCVALGELSA LINTLET VEXVPACEPPE II LOCSGYTSENATVVVAGNILELDVSITGEPPPVATUL KGD
hu-fst hu-fst beg0 YTFVPDGYALGSLSANLNFLEIKVCI VPNQEPP I HLDCSGKTSENAIVVVAGNKLELDVSITGEPPPVATVLKGD beAD YSFVPKGFAC-NLSANLHFMEVKIDFVPRQEPP I HLDCSGKTD-TIVVVAGNKLELDVSITGEPPATVING KAITGGNAPABPAPDAPEDIGOS beAD YSFVPKGFAC-NLSANLHFMEVKIDFVPRQEPP I HLDCSGKTD-TIVVVAGNKLELDVPISGPAPTVING KAITGGNAPABPAPDAPEDIGOS beAD YSFVPKGFAC-NLSANLHFMEVKIDFVPRQEPP I HLDCSGKTD-TIVVVAGNKLELDVPISGPAPTVING KAITGGNAPABPAPDAPEDIGOS beAD YSFVPKGFAC-NLSANLHFMEVKIDFVPRQEPP I HLDCSGKTD-TIVVVAGNKLELDVPISGPAPTVING KAITGGNAPABPAPDAPEDIGOS beAD YSFVPKGFAC-NLSANLHFMEVKIDFVPRQPPF I HLDCLGOSPD-TIVVVAGNKLELDVPISGPAPTVNMG KVGKKASAGPH PDAPEDAGAK fu-cc beGD YFFVPGGAA'-SLSANLHFLEVKIDFVPRQPPF I HLDCLGOSPD-TIVVVAGNKLELDVPITGPAPTVNMF KGE STTGFTI YCVSLGGNNRDS hu-stw Motif 5 cont. * Motif 6 VVFDKKLLCETGGNVETKNRSYFTVEGAE KEDEGVTTVVNNPVGEDVALITVKIDVPDAPAAPAFISNVGEDSCTVØWE PPANDGGGPI no-cc VVFDKKLLCETGGNVETKNRSYFTVEGAE KEDEGVTNTVNNPVGEDVALITVKIDVPDAPAAPFISNVGEDSCTVØWE PPANDGGGPI hu-stw ** * * Motif 6 cont. * Motif 7 hu-fst hu-fst hu-fst * * * * * Motif 6 cont. * Motif 7 hu-fst EKKKKSSQRMMELNFEVFTETTYESTMIEGILVEMVFAVNAIGVSQPSNNTKPEMPIAPTSPLLIVEDVDDTTTTLKWRPPRNGGGDDGILVEV ** * * Motif 7 CONSTRUCTION STANLERED SOVENNAIGNAGRSPSASOPENPIGPSEPILLVEDVTDTTTLKWRPPRNGGGDDGILVEV ** * * Motif 7 CONSTRUCTION STANLARD SOVENNAIGSSQPSNNTKPEMPIAPTSPLLIVEDVDDTTTTLKWRPPRNGGGDDGILVEV ch-cc	STEVEGEALGSLSALINFLEIKVETVPAQEPPRIHLDCSGKTSENATVVVAGNILALDVSITGEPPPVATHLKGD VSFVPEGFAC-NLSALHFMEVKIDFVPAQEPPRIHLDCGGIPD-TIVVVAGNILALDVSITGEPPPVATHLKGD VSFVPEGFAC-NLSALHFMEVKIDFVPAQEPPRIHLDCGGIPD-TIVVVAGNILALDVPISGPAFVVHQ KATCGNKAPADPAPEDAPEDTGDSD
hu-cc DEADYSFVPEGFAC-NLSALLHEMEVKIDFVPRQEPPTIHLDCFGETPD-TIVVVAGNKLELDVPISGPAPTVINQKITCGNKAPAGPAPDAPEDTGO mo-cc DEADYSFVPEGFAC-NLSALLHEMEVKIDFVPRQEPPTIHLDCFQSTPD-TIVVVAGNKLELDVPISGPAPTVINQKITCGNKAPAGPHPDAPEDTGO DEADYSFVPEGFAC-NLSALLHEMEVKIDFVPRQEPPTIHLDCFQSTPD-TIVVVAGNKLELDVPISGPAPTVVNQKTVLQKKASAGPHPDAPEDTGO DEADYSFVPEGFAC-NLSALLHEMEVKIDFVPRQEPPTIHLDCFQSTPD-TIVVVAGNKLELDVPISGPAPTVVNQKTVLQKKASAGPHPDAPEDTGO pedpyTFVPDGAAV-SLSALLHEMEVKIDFVPRQEPPTIHLDCGGSPD-TIVVVAGNKLELDVPISGPAPTVVNQKVNKGELVHQSNE-DSLTPSEN hu-stw DEGDYTFVPDGAAV-SLSALNFLEVKIDFVPRQPPTIHLDCMGTAESTUVVAGNKLELDVFISGPAPTVVNTKGELSTGFTFYCYCLGANKRD hu-stw DEGDYTFVPDGAAV-SLSALNFLEVKIDFVPRQEPPTIHLDCMGTAESTUVVAGNKLELDVFISGPAPTVVNTKGESTGFTFYCYCLGANKRD hu-stw Motif 5 cont. * Motif 6 WVFDKKLLCTEGRVAVETKRRSYFTEGAEKEDEGVYTTVKNPVGEDVASIFLQVVDVDPPPEAVNISKGESTVØWEPPANDGGGPTI mo-ccWVFDKKLLCTEGRVAVETKRRSYFTEGAEKEDEGVYTTVKNPVGEDVALITVKVDVPDAPAAPFISNVGEDSCTVØWEPPANDGGPTI fu-ccSTSTSKLLFESEGRVAVEKHEDHCVFIIEGAEKEDEGVRVTVVNNPVGEDVALITVKVDVDPPPEAPKISNGEDSCTVØWEPPANDGGPTI fu-cc	YSFVPEGFAC-NLSA LIHEMEVI IDFVPROEPPKIHLOCPGGIPD-TIVVVGNILALDVPISGDPAFVING KAITQGNKAPABPAPEDAPEDAPED VSFVPEGFAC-NLSA LIHEMEVI IDFVPROEPPKIHLOCPGSTPD-TIVVVGNILALDVPISGDPAFVING KATTQGNKAPABPAPEDAGED VSFVPGGFAC-NLSA LIMEEVI IDFVPROEPPKIHLOCGSTD-TIVVVGNILALDVPISGDPAFVVNG KATTQGNKAPABPAPEDAGADE
mo-cc DEAD ISSUEGORAT-NUSAL LOPLEVE LOPVOEDPPTIELDCGOS DE TIVVVANLALUUVISGDPATVVMOVIVIGAAASDEPIDAELAGA fu-cc DEAD ISSUEGORAT-NUSAL LOPLEVE LOPVOEDPPTIELDCMOENDE LOVDISGDPATVVMOVIVIGAAASDEPIDAELAGA fu-cc DEGDYTVPDGAA-SLSAL LOPLEVE LOPVOEDPPTIELDCMOENDE LOVDISGDPATVVMOVIVIGAAASDEPIDAELAGA hu-slw DEGDYTVPDGAA-SLSAL LOPLEVE LOPVOEDPPTIELDCMOENDE LOVDIVVIGAAASDEPIDAELAGA hu-slw Motif 5 hu-st Motif 6 hu-st Motif 6 hu-st Motif 6 WVFDKKLLCETEGRVRVETKRRSTFTVEGAEKEDEGVYTTVVINPVGEDVASIFLOVVDVDPDPEAVNITSVGEDATLOWEPPANDGGCPTI mo-cc WVFDKKLLCETEGRVRVETKRRSTFTVEGAEKEDEGVYTTVVINPVGEDVALITVVIDVDAPAAPISNVGEDSCTVOWEPPANDGGCPTI fu-cc WVFDKKLLCETEGRVRVETKRRSTFTVEGAEKEDEGVYTVVINPVGEDVANITVVIDVDPAPAAPISNVGEDSCTVOWEPPANDGGCPTI fu-cc	TSFLOGEN UNDER THE ALTER AND ALTER A
fu-cc DEGDTTFVPDGFAT-SLSAKLNFLEVRIDTVPRQDPPRIHLDCMGTAESTIVVVAGNKLRLDVPITGDPAPTVVWTKGERSTTGPTTYCSLGGNWRDD hu-slw DEGDTVFAPDANV-TLPAVHVID	NTFVPDGAT-SLSATINFLEV.IDTVPADDPPRIHLDCMGTASSTIVVVAGNKLELDVPIGDPAPTVVWTKGELSTGPTIFYCVSLGGNWRDSTDQV 648 NVFAPDATNV-TIPAVHVIDPRIILDGLDADNTVVIAGNKLELEIPISGEPPPKAMSKGD
hu-slw DEGDYVFAPDAYNV-TLPAYVHVIDPFKITLDGLDADNTVTVIAGNKLELEIPISGEPPPKAMMS.GD	YVYFAPDAYNV-TLPATVHVID
Motif 5 cont. * Motif 6 hu-fst EVFTTEGKTRIEKRVDCSSVIESAQEDEGKTIKVTNVGEDVASIFLEVVDDPDEAVAITSVGEDWAILWEPPADGGKVT hu-fst EVFTTEGKTRIEKRVDCSSVIESAQEDEGKTIKVTNVGEDVASIFLEVVDDPDEAVAITSVGEDWAILWEPPADGGCVT mo-cc EVFTTEGKTRIEKRVDCSSVIESAQEDEGKTIKVTVVNPVGEDVASIFLEVVDDPDEAVAITSVGEDWAILWEPPADGGCVT mo-cc EVFTKLCETEGNVETKDRSIFTVGAEKEDEGVTVTVNNPVGEDQVNLTVKUDVDDAAAPKISNVGEDSCTVQWEPPADGGCPUT fu-cc	Motif 5 cont. * Motif 6 EVFTTEGATRIEKRVDCSSFVIESAQREDEGRUTIKVUTPYGEDQNASIFLQVDVDPDPEAVAITSVGEDWAILVWEPPMIDGGVPVTGIV 676 WVFDKKLLCETEGAVRVETKNRSIFVIGGAEKEDEGVUTVTVRNPVGEDQUALTVKVDVDPDAPAAPHISNVGEDSCTVQWEPPAIDGG0PVLGVIL 808 WVFDKKLLCETEGAVRVETKNESIFVIGGAEKEDEGVUTVTVRNPVGEDQUALTVKVDVDPDAPAAPHISNVGEDSCTVQWEPPAIDGG0PVLGVIL 804 WVFDKKLLCETEGAVRVETKDRSVFTVEGAEKEDEGVUTVTVRNPVGEDQUALTVKVDVDPDAPAAPHISNVGEDSCTVQWEPPAIDGG0PVLGVIL 804
hu-fstWVFDTKLLCTEGRVRVETKURSITVEGAEKEDEGRVTKVTNPVGEDVASIFLOVDVDDPPEAVNITSVGEDNAILWEPPANDGGRPV hu-ccWVFDKLLCTEGRVRVETKURSITVEGAEKEDEGVVTVVTNPVGEDQVNLTVVIDVDAPAAPKISNVGEDSCTVQWEPPANDGGPV ch-ccSTDSKLLCTEGRVRVETKURSITVEGAEKEDEGVVTVVTNNPVGEDQVNLTVVIDVDAPAAPKISNVGEDSCTVQWEPPANDGGPV fu-ccSTDSKLLFFSEGRVRVEKHEDHCVFIEGAEKEDEGVVRVIVNNAGEDTADINVKVVDVDPPEAPKISNGEDYCTVQW0PPTNDGQPVI fu-ccSTDSKLLFFSEGRVRVEKHEDHCVFIEGAEKEDEGVVRVIVNNAGEDTADINVKVVDVDPPEAPKISNGEDYCTVQW0PPTNDGQPVI fu-cc	
mo-cc WVFDKKLLCITGGVRVETTKDRSVFTVEGAEKEDEGVVTVVKNPVGEDQVNLTVKVDVDPDAPAAPKISNVGEDSCTVQWEPPARDGQDVI ch-cc LSTDSKLLFFSEGRVRVETHEDHCVFIEGAEKEDEGVVRUVVNPVGEDKADITVKVDVDPDFEAPISNIGEDYCTVQWQPPTDGGQDVI fu-cc LSTDSKLLFFSEGRVRVETHEDHCVFIEGAEKEDEGVVRUVVNPVGEDKADITVKVDVDPDFEAPISNIGEDYCTVQWQPPTDGGQDVI hu-slw	WVFDKKLLCETEGVRVETTKDRSVFTVEGAEKEDEGVVTVTVKNPVGEDQVNLTVKVIDVPDAPAAPKISNVGEDSCTVQWEPPATDGGOPVLGYIL LSTDSKLLFSEGVRVEKHEDECVFILGAEKEDEGVRVIVNPVGEDKADITVKVDVPDPPAPAISNUGEDSCTVQWOPPTVDGGOPVLGYIL 1EVKLFTLQVVTRGDGVVHVESTKGCIFTIEGAEKODEGIVSVVRNLAGEDTADITVKVDVPDPPQEVKILSVGESCVVQWDPPLFDGGOPIGVI 1EVKLFTLQVVTRGDGVVHVESTKGCIFTIEGAEKODEGIVSVVRNLAGEDTADITVKVDVPPPPQEVKILSVGESCVVQWDPPLFDGGOPIGVI 53
ch-cc fu-cc fu-cc ch-cc fu-cc ch	LSTDSKLLFRSEGVRVEKHED CVFIIEGAEKEDEGVVRVIVENPVGEDKADITVEVIDVPDEPEAPISNIGEDYCTVQWQPPTYDGGQPVLGYIL IEVKLETLQVVTRGDGVHVESTKGCIFTIEGAEKQDEGINSVVRNLAGEDTADINVVVDVPDPPQEVNILSVGESCVVQMDPPLFDGGQPILGVI 753
fu-cc CAMPIEVKLETLQVVTEGDGGVVESTKGHCIPTTEGAERODEGIXSVVVRILAGEDTADINVKVDVPDPPQEVKILSVGEESCVVQMDPPLEPGGQPII hu-slw	IIV/LETLQVVTEGGGVUVGESTKGCIFTIGAEROBGIVSVVVRILAGEDTADIVVVVDPPOPOLVRILSGESCVVQNDPPLPDGGGPIIGVI
** * * * Motif 6 cont. * * Motif 7 hu-fst ** * * Motif 6 cont. * * * Motif 7 hu-fst RKKKSSRMMCINFEVFTTTYESTMIEGILYEMEVFAVNAIGVSQPSMNTKPFMPIAPTSEPLLIVEDVTDTTTILKWAPPNIGAGGIDGILVEY hu-cc RKKKSSRMMCINFELIGELSHEARMIEGVYEMEVFAVNAIGNSQPSMNTKPFMPIAPTSEPLLIVEDVTDTTTILKWAPPNIGAGGIDGISVEY mo-cc RKKKSSRMMCINFELISELSHEARMIEGVYEMEVFAVNAIGNSRPSPASQPFMPIGPGEPTLAVEOVSTTVSLKWAPPENVGAGELGSSVEY ch-cc RKKKKSIMMMINFELISELSHEARMIEGVYEMEVFAVNAIGNSRPSPASQPFMPIGPGEPTHEVEDVSDTTVSLKWAPPENGGAGELGSSVEY	* * Motif 6 cont. * * Motif 7 KGSQENNKLNFEVETETTEST MIEGILVEM VEAVNAIGVSQPSMNTKPEMPIAPISEPLILIVEDVTDTTTILKMPPPNIGAGGIDG LV:YCLEGS 781
hu-fst ERKKKSSRWMELNFEVFTETTYESTKMIEGILEMRVFAVNAIGVSOPSMNTKPFMPIAPTSEPLHLIVEDVTDTTTTLKWEPPNKIGAGGIDG LVEY hu-cc ERKKKSSRWMOLNFDLICELSHEARMIEGVVYEMRVFAVNAIGMSRPSPASOPFMPIGPOSPTHLAVEOVSTTVSLKWEPPEVGAGGLOG SVEY mo-cc ERKKKSSRWMELNFDLIKELSHEARMIEGVVYEMRITAVNSIGMSRPSPASOPFMPIGPOSPTHLAVEOVSTTVSLKWEPPEVGAGGLOG SVEY ch-cc ERKKKSSRWMELNFDLIKELTYEARMIEGVYYEMRITAVNSIGMSRPSPASOPFMPIGPOSPTHFURVEOVSDTTVALWWEPPENGAGGLOG SVEY	KGSORWMKLNFEVFTETTYESTKMIEGILYEMRVFAVNAIGVSOPSMNTKPFMPIAPTSEPLHLIVEDVTDTTTTLKWRPPNRIGAGGIDGYLVEYGLEGS 781
hu-cc ERKKKSYRWMQLNFDLIQELSHEARMIEGVVYEMRVYAVNAIGMSRPSPASQPFMPIGPPSEPTHLAVEDVSDTTVSLKWRPPERVGAGGLDGYSVEY mo-cc ERKKKSYRWMRLNFDLLRELSHEARMIEGVAYEMRVYAVNAVGMSRPSPASQPFMPIGPGEPTHLAVEDVSDTTVSLKWRPPERVGAGGLDGYSVEY ch-cc ERKKKSYRWMRLNFDLLKELTYEARMIEGVVYEMRIYAVNSIGMSRPSPASQPFMPIAPPSEPTHFTVEDVSDTTVALKWRPPERIGAGGLDGYIVEY	
mo-cc ERKKKKSYRWMRLNFDLLRELSHEARRMIEGVAYEMRVYAVNAVGMSRPSPASQPFMPIGPOEPTHLAVEUVSDTTVSLKWRPPERVGAGGLOGYSVEV ch-cc ERKKKKSYRWMRLNFDLLKELTYEARRMIEGVVYEMRIYAVNSIGMSRPSPASQPFMPIAPPSEPTHFTVEDVSDTTVALKWRPPERIGAGGLOGYSVEV	KSTRWMQLNFDLIQELSHEARMIEGVVYEMRVNAVNAIGMSRPSPASQPFMPIGPSEPTHLAVEDVSDTTVSLKWRPPERVGAGGLDGYSVEYGPEGC 913
	KSYRWMRLNFDLLRELSHEARMIEGVAYEMRVYAVNAVGMSRPSPASOPFMPIGPPGEPTHLAVEDVSDTTVSLKWRPPERVGAGGLDGYSVEYCOEGC 909
fu-cc ERKKKKSYEWMELNFDPYPETTYEAKEMIEGVAYEMEVYAVNSIGMSRHSQASQPFVPVAPTSEPVGLCVDDISDTSIVLKWEPPEEMGSVDLEGYGVEY	EXCHEMELNEDLIKELTHEATENTEQUVENETTAUNSTCMSPOSDASODENDTADDSEDTHETVEDVSDTTVALENBODERTCACCLDCYTVEVENCS 911
hu-s1w E <mark>rkkk</mark> osSRWMRLNFDLCKETTFEP <mark>KK</mark> MIEGVATEVRIFAVNAIGISK <mark>P</mark> SMPSRPFVPLAV <mark>TSPP</mark> TLLTVDSVTDTTVTMR <mark>RPPD</mark> HIGAAGLDCTVLEXC	KKSYRWARLNFDLLKELTYEAKKMIEGVVYEMRIYAVNSIGMSR <mark>PSPASOPFMPIAPPSEPTHFT</mark> VEDVSDTTVALKWRPPERIGAGGLDGVIVEYCKDGS 911 KKSYRWARLNFDPYPETTYEAKRMIEGVAYEMRVYAVNSIGMSRHSOASOPFVPVAPTSEPVGLCVDDISDTSIVLKWRPPERMGSVDLEGYGVEYCKEGT 858
Motif 7 cont. * Motif 8 *	KKSTRMMELNFDLLKELTYEAKEMIEGVVYEMEITAVNSIGMSR <mark>PSPASOPFMPIAPPSEPTHFT</mark> VEDVSDTTVALKWRPPERIGAGGLDGVIVYCKDGS 911 KKSTRMMELNFDPYPETTYEAKEMIEGVAXEMEVYAVNSIGMSRHSQASOPFVPVAPTSEPVGLCVDDISDTSIVLKWRPPERMGSVDLEGXGVEYCKEGT 858 KOSSRMMELNFDLCKETTFEPKEMIEGVAXEVEIFAVNAIGISKP <mark>SMPSRPFVPLA</mark> VTSPPTLLTVDSVTDTTVTMEWEPPDHIGAAGLDGVLEYCFEGT 761
NU-1ST EWVENTEVENCGFTWENLPTGARLLFHVUGVNIAGREPATLAPUTTREIAEPHILEDTILKVGEQUNLVVPGGPBPQVVHTKGARL	AKSTRWMRLNFDLLKELTYEAKEMIEGVVYEMRIAAVNSIGMSRPSPASOPFMPIAPPSEPTHFTVEDVSDTTVALKWRPPERIGAGGLDGVIVYGKDGS 911 IKKSTRWMRLNFDPYPETTYEAKEMIEGVAYEMEVYAVNSIGMSRHSQASOPFVPVAPTSEPVGLCVDDISDTSIVLKWRPPERMGSVDLEGVGVYCKFGT 858 IQSSRWMRLNFDLCKETTFEPKKMIEGVAYEVFIFAVNAIGISKPSMPSRPFVPLAVTSPPTLLTVDSVTDTTVTMWRPPPHIGAAGLDGVVLEYGFEGT 761 Motif 7 cont. * Motif 8 *
mo-cc SEWTPALGCIZERREMLVK0LPTGARLLFVFAINVAGPGGPIVTKEPVTVOELCAPHILGTIORVGEPVNLLTPGGPPEOVTHTEGOPI	KKSTRUMRLINFOLKKLITTEAKUMIEGVNYEMIIAVNSIGMSR <mark>PSPASOPYMPIAPSEPTHFTVE</mark> UVSDTTVALKME PPER IGAGGLOGIVEYGKOG KKSTRUMRLNFDPYPTTYEAKUMIEGVNYEMIVAVNSIGMSRHSQASOPYVVATSPVALVSPVACUDIISDTSIVLKWEPPERMGSVDLEG GVYCKEGT 858 ZOSSRUMRLNFDLCKSTTFEPKMIEGVAYEV IFAVNAIGISKPSMPSRPFVPLAVISPPTLLTVDSVDTTVTIKWEPPCHIGAGLOGIVLEG Motif 7 cont. * Motif 8 * PANTERVENGGFTVNLPTGARILFFVVGVNIAGRTEPATLAOPVTIREIAPPTIREIAPPTIREVEGUNITYTERVGGUNUVFFGGAPLDTSR 886 AL COLTENSILVENT DETARILFFVVGVNIAGRTEPATLAOPVTIREIAPPTIREVATCHVATATORVNIATVATORVENGGAPLDTSR 886
ch-cc AEWTPALPGLTERTSALIKDLVTGDKLTFRVKAINLAGESGAAIIKEPVTVQEIWQKPKICVPHLKQTLVKKVGETINIMIPPGGKPRPKISWMCDGQTI	IKS FRUMELNEDLIKELTTERAKUMIEGVAVEMETIAVNSIGMSR <mark>PSPASOPYMPIAPSEPTHETVE</mark> DVSDTTVALKWEPPERIGAGGLOGIVEYGKOG KKSTRUMELNEDPYETTYEAKUMIEGVAVEMEVYAVNSIGMSRHSQASOPFVPVAPTSEPVGLCVDDISDTSIVLKWEPPERIGSVDLEGGVEYGKEGT 858 KOSSRUMELNEDLCKSTTFEPKKMIEGVAVEVIFAVNAIGISKPSMPERPEVPLAVTSPPTLITVDSVDTTVIKWEPPERIGSVDLEGGVEYGKEGT Motif 7 cont. * Motif 8 * PANTERVENGGEFVVNLPTGARLIFFVVGNIAGRTEPATLAOPVTIETAPPKILPPHILTOTYIKKVGGULIVVPGGFPRDQVVWTKGGAPLBTSR 886 AALQGITEHTSILVKDLPTGARLIFFVGARNNAGPGGPVTTEPVTVOEILORPLOLPPHILKOTICKVGEPVNLLIPFGGFPRDQVVWTKEGGPLAGEE 1018 PALGGITERSSUVKOLPTGARLIFFVGARNNAGPGGPVTTEPVTVOEILORPLOLPPHILKOTICKVGEPVNLLIPFGGFPROPVTWTEGGPLAGEE 1018
fu-cc EDMLPAVQGLTDRTSLIINLTTGDKLQFRVRAYNAGPSPAATLAPVTIKEIMQAPKILPNNLGTLIRVGGTVNLAIPPQGKPPPKVTWSNNGEPI	IKS RWMRLNFDLKELTTEAKUMI GOVYEMEITAVN SIGMSRPSPASOPYMPIADPSEPTHFTVEDVSDTTVALKW PPERIGAGGLOG IV YGKDGS 911 KKS RWMRLNFDPYPTTYEAKUMI GOVYEWR VYAVNSIGMSRHSQASOPYVPAPTSEPVGLOVDI SDTSIVLKW PPERMGSVDLGG VYYGKDG 858 KQSSRWMELNFDLCKETFFERKMI GOVYEVR IFAVNAIGI KKPSMPSRPFVPLAVTSPTLLTVDSVDTTVTMRW PPERMGSVDLGG VYYGKGG 761 Motif 7 cont. * Motif 8 * PANTEPVERGFTVKNLPTGARILF VVGVNIAGREFERATLAOPVTIREIAEPKIRLPHILKOTVIRKVGEVLNLVVPFGGRPRPQVVWTGGAPLDTSR 886 AALQGITENTSIVKOLPTGARLIF VVGVNIAGREFERATLAOPVTIREIAEPKIRLPHILKOTVIRKVGEVLNLVVPFGGRPRPQVWTGGAPLDTSR 886 PALGGITENTSIVKOLPTGARLIF VVGVNIAGREFEPATLAOPVTIREIAEPKIRLPHILKOTVIRKVGEVNLLPFGGRPRPQVWTTGGAPLDTSR 886 AALQGITENTSIVKOLPTGARLIF VVGVNIAGREFEPATLAOPVTIREIAEPKIRLPHILKOTVIRKVGEVNILLPFGGRPRPQVWTTGGAPLDTSR 886 MOTIF 88500000000000000000000000000000000000
NOVEL 10 TO THE REAL OF THE RE	IKS RWMELNEDLIKELTTEAKUMI GUVYEMI IAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDUSDTTVALKWEPPERIGAGGLOGI UVYCKOG 911 KKS RWMELNEDLYPETTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPEMPIADPSEPUGLCVDI SDTS ULKWEPPERMGSVDLEG GVYCKEGT 858 KQSSRWMELNEDLCKETFEPKMI GUVYEVI FAVNAIGI KSPSMPSRPFVPLAVTSEPULLTUDSVDTTVTMRWEPPHIGAAGLOG VLEVEFEGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGEFUKNLPTGARILEFUVGVNIAGRTEPATLAOPVTIREIAEPKIRLPHILNOTI REVGEQUNLVVPFGKPRPQVVWTGGAPLDTSR 886 AALGGITERTSILVKDLPTGARILFVVAHNMAGPGAPUTTEPTVDVELLOPHILDPHILNOTI REVGEQUNLVVPFGKPRPQVWTTEGGPLAGEE 1018 PALGGITERTSILVKDLPTGARLFVVAHNMAGPGAPUTTEPTVOELLOPHILQTIQKEVGEPVNLLIPFGKPRPQVWTTEGGPLAGEE 1018 PALGGITERTSILVKDLVGALLFVVAHNMAGPGAPUTTEPTVOELLOPHILQTICKEVGEVNLLIPFGKPRPQVWTKEGGPLAGEE 1018 PALGGITERTSILVKDLVGALFFVAHNAGPGAPITEPTVEVELOPHILQTICKEVGEVNLLIPFGKPRPQVWTKEGGPLAGEE 1018 PALGGITERTSILVKDLVGALFFVAHNAGPGAPITEPTVOELLOPHILQTICKEVGEVNLLIPFGKPRPQVWTKEGGPLAGEE 1018 PALGGITERTSILKDLVGEVGALFFVAHNAGPGAPITEPTVOELLOPHILQTICKEVGEVNLLIPFGKPRPQVWTKEGGPLAGEE 1014 PAVGGITDTSILI NLITTGDLOFFVAHNAGPGPADTEPTVOELLOPTING FFICOPHILDPICKTER PTVESENGEPLSTEF 963
MOLLE 0 CONC. MOLLE 9	IKSY RWARLAFDLIKELTYEA KUMI GOVYEMI ILAVNSIGMSRPSPASOPIMPIADPSEPTHFTVEDVSDTTVALKW PPERIGAGGLOGI UVYCKOG 911 KKSY RWARLAFDPYPTTYEA KUMI GOVYEMI ILAVNSIGMSRPSPASOPIMPIADPSEPGLOVDI SDTS ULAW PPERMGSVDLEG GVYCKOG 858 KQSSRWAELAFDLCKETFFERKMI GOVYEMI FAVANIGI KSPSMPSRPFVPLAVTSEPULUTDSVDTTVTMRW PPERMGSVDLEG GVYCKEGT 858 SAUGSTENSILVELTFERKMI GOVYEMI FAVANIGI KSPSMPSRPFVPLAVTSPTLLTVDSVDTTVTMRW PPERMGSVDLEG GVYCKEGT 858 Motif 7 cont. * Motif 8 * PANTEPVER GGFTVENLPTGARILF VVGVNIAGRTEPATLAOPVTIREIAEPFIRLPAHLAQTI REVGEQUNLVVFFGK PROVVWTKEGGPLAGEE PALGETERTSILVEDLPTGARLIF VVGVNIAGRTEPATLAOPVTIREIAEPFIRLPAHLAQTI REVGEQUNLVVFFGK PROVVWTKEGGPLAGEE 1018 PALGETERTSILVEDLPTGARLIF VVGANNAGPGAPUTTEPTVORIL OF PLUCIPHILQTICK VGETVNLLIFFGG FRAQVTWTKEGGPLAGEE 1018 PALGETERTSILVEDLFTGARLIF FVANNAGPGAPUTKE PVTVORIL OF PLUCIPHILQTICK VGETVNLLIFFGG FRAQVTWTKEGGPLAGEE 1014 PALGETERTSALI EDIVTGDLIFFVANNAGPSPAATLAOPVTIREIMOSPHILLPRULKOTICK VGETVNLATFFGG FRAQVTWKEGGPLAGEE 1014 PAVGETDTSILINNITTGDLIFFVANNAGPSPAATLAOPVTIREIMOSPHILLPPUTVKETINNAFFGGFPAPKTENSM OGTILDSKO VANKDIDKTKTTIGLPTDAKIFV VKANNAGASEPYTSOPIKEIPPFRINDEPKIRIPPIKEV DEVILVEFTER VGETVNLATFFGG FRAGENSM OGTILDSKO VANKDIDKTKTTIGEPTDAKIFV VKANNAGASEPYTSOPIKEIPPIKILPPIKILPPIKIKOTI REVGETVNLATPFGG FRAGENSM OGTILDSKO VANKDIDKTKTTIGLPTDAKIFV VKANNAGASEPYTSOPIKE PALGETERTSALK OLVTGDAL FV VKANNAGASEPYTSOPIKE PANTEPPIKER
hu-fst WHYNTSDFDTVFFVNQAARSDSGEVELSVQIENMKDTATIRIRVVERAGPPINVMVKEVWGTNALVEWOAPHDDGNSEIMGWFVORADKKTMEWFNVERN	IKSS RWMRLNEDLIKELTTEAKUMIEGVVTEMIIAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKW PPERIGAGGLOG IVEYGKOG 911 KKSYRWMRLNEDPYPETTYEAKUMIEGVAYEWRIYAVNSIGMSRPSPASOPEMPIADPSEPGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKEG 858 KQSSRWMRLNEDLCKETFEPKMIEGVAYEW IFAVNAIGISKPSMPSRPFVPLVTSPTLTVDSVDTTVTMRWPPHIGAAGLOG VLEYGKEG 761 Motif 7 cont. * Motif 8 * PANTEPVERGEFTVENLPTGARILFEVVGVNIAGRTEPATLAOPVTIREIAEPFNIRLPRHLNQTIKVGEVGLUNUVPFGKPPQVVWTKEGOPLAGES PALGGITERTSILVKDLPTGARILFEVVGVNIAGRTEPATLAOPVTIREIAEPFNIRLPRHLNQTIKVGEVNLLIPFGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLPTGARILFEVTAHNAGPGAPUTTEPTVORILOPFLQLPRHLNQTICKVGEVNLLIPFGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLFTGARLFEVTAHNAGPGAPUTTEPTVORILOPFLQLPRHLNQTICKVGEVNLLIPFGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLFFVAHNAGPGPITKEPTVORILGPPLQLPRHLNQTICKVGEVNLLIPFGKPRPQVTWKEGOPLAGES 1019 PALGGITERTSALIKUVTGARLFEVTANNAGPSPAATLAOPVTIREIMOSPHILLPRUKVGETINIMIPFGKPRPKTSWM.DGGTLDSKO 1016 PAVGGITDTSILINNITTGALOFFVAAYNAAGSSPAATLAOPVTIREIMOSPHIRPHILPPNIRIPHIKQTIKKVGETINIMAFGGPRAGELSTF 963 WANKDIDKTKFTIGLPTDARIFVKAANAAGASEPYYSOPILVKEILEPPNIRIPHILVGYIRKVGETVNLAIPFGKPRPELTWKKDGAEIDKNO 866 Motif 8 cont. MOTIF TSDFDVFFVGAASDSGEVELSVGIENMKDTATIRIKVKKAGPPINVKEVWGTNALVEWOAPKDORSEIMGFVOLADKKTMEWFNVERNHES 991
hu-fst VHVKTSDFDTVFFVRQAASSDSGEVELSVQIENMKDTATIRIRVVEKAGPPINVMVKEVWGTNALVEWQAPKDDGNSEIMGEFVQKADKKTMEWFNVERN hu-cc VSIKNSPTDTILFIRAARKVHSGTVQVTVRIENMEDKATLVLQVVDKPSPPODLKVTDAWGLNVALEWKPPQDVGNTELWGTVQKADKKTMEWFTVLEH	IKS TRAMELINFOLIKELTTEAKUMI GOVTEMETIAVNSTGMSR959ASOPTMPIAPSESPTHFTVEOVSDTTVAL MEDPERIGAGGLOGIVEXGKOG KKSTRAMELINFOLIKELTTEPKAMI GOVATEMEVTAVNSIGMSRHSOASOPTVPAPSESPTHFTVEOVSDTTVAL MEDPERIGAGGLOGIVEXGKOG KOSSRAMELNFOLKETTEPKAMI GOVATEMEVTAVNSIGMSRHSOASOPTVPAPSESPTHFTVEOVSDTTVAL MEDPERIGAGGLOGIVEXGKOG Motif 7 cont. * Motif 8 * PANTESVERGETVNILPTGARLLFVVGVNIAGRTEPATLAOPVTIREIASPPTILTVDSVTDTVDSVMPAPPHIGAAGLOGVLEXGET ALGGITEHESILVKOLPTGARLLFVVGVNIAGRTEPATLAOPVTIREIASPPTILTVDSVDTVOTSVRMPPHIGAAGLOGVLEXGET PANTESVERGETVNLPTGARLLFVVGVNIAGRTEPATLAOPVTIREIASPPTILTUSVENTUSTOVNIAGGAPLDTSR PALOGITEHESILVKOLPTGARLLFVVGVNIAGRTEPATLAOPVTIREIASPPTILTUSVENTUSTOVNIAGGAPLDTSR PALOGITEHESILVKOLPTGARLLFVVGANNAGGAPVTTPUTVOELICOPHILOPTINEVGTIOKVGEPVNLLIPFOGKPHPQVVNTEGOPLAGES 1014 PALOGITEHESILVKOLPTGARLLFVVGANNAGGSGAPVTTEPUTVOELICOPHILOPTINEVGTIOKVGEPVNLLIPFOGKPHPQVVTEGOPLAGES 1014 PALOGITEHESILVKOLPTGARLLFVVGANNAGSSGAAIIKEPVTVOELMGPHICVPHLAQTIVKVGTININIPFOGKPHPKSSNGEJSKD 1016 PAVGGITDSIIINILTTODLOFVARIVAAGSSGAAIIKEPVTVOELMGPHILPPHILAQTIVKVGTININIPFOGKPHPKESSTF 94 VANKDLIDVTFTIGLPTDAKIFVVAINAAGASEPYYSOPILVEITEPPIRTPHLAQTI RVGTINKVGFVNLVFVGGFPHESTF 95 15BFDTVFFVAAASDSGEVELSVOIENMKDTATIRIRVVERAGPPINVWKEVWGTALVEWQAAPLOGNESIMGFVORADKTMEWFNVERNHTS 91 13BFTDTILFIRAARSV GSTVOVARIMEMERATIVLOVOR FOSPPOLAVINEVGVARIAVEWQAAPDGSTEIMGFVORADKTMEWFNVERNHTS 91 122
hu-fst VHVKTSDFDTVFFVRQAALSDSGEVELSVQIENMKDTATIRIRVVSKAGPPINVMVKEVWGTNALVEWQAPKDDGNSEIMGFVQKADKKTMEWFNVERN hu-cc VSIRNSPTDTILFIRAARIVHSGTVQVTVRIENMEDKATLVLQVVDKPSPPQDLRVTDAWGLNVALEWKPPQDVGNTELWGTVQKADKKTMEWFTVLEH mo-cc VSIRNSPTDTILFIRAARITHSGTVQVTVRIENMEDKATLILQIVDKPSPPQDIRIVETMGFNVALEWKPPQDDGNTEIWGTVQKADKKTMEWFTVLEH	IKS TRAMELNFOLKELTTEAKUMI GOVTEMI ILAVNSIGMSR BOASOPHOTADESEPTIFTVEOVSDTTVAL MEDPERIGAGGLOGI VEYGKOG KKSTRAMELNFOPPETTYEAKUMI GOVATEMEVYAVNSIGMSR BOASOPFVPAPTSE VGLOVDI ISDTSIVLKWE PPERMGSVDLEG GOVEYGKEG 858 KOSSRAMELNFOLCKSTTEPKAMI GOVATEMEVYAVNSIGMSR BOASOPFVPAPTSE VGLOVDI ISDTSIVLKWE PPERMGSVDLEG GOVEYGKEG 858 Motif 7 cont. * Motif 8 * PANTEEVEIGGTTVNLPTGARLIFVVGVNIAGR TEPATLAOPVTIREIASPPILLTVDSVDTVDSVDFVGLVKUPPGCK PRPQVWTKGGAPLDTSR 886 AALGGITERSSLIVGLPTGARLIFV VGAINAGGGAPVTTE PVTVQEILOPPLILOPTICKVGEVNLLIPFGCK PRPQVWTKGGAPLDTSR 886 PALGGITERSSLIVGLPTGARLIFVVGAINAGGGAPVTTE PVTVQEILOPPLILOPTICKVGEVNLLIPFGCK PRPQVWTKGGAPLDTSR 886 94LGGITERSSLIVGLPTGARLIFVVGAINAGGGAPVTTE PVTVQEILOPPLILOPTICKVGEVNLLIPFGCK PRPQVWTKEGOPLAGEE 1014 PALGGITERSSLIVGLPTGARLIFVVGAINAGGSGAAIIKEPTVQEILOPPLIQTICKVGEVNLLIPFGCK PRPVTWTEGOPLAGEE 1016 94UGGITDTSIIINITTGDLOFVGAVMAGSSGAAIIKEPTVQEIMGPTITELPPNICUPTIKKVGTVNLINPFGCK PRPVINTEGOPLAGEE 1016 94UGGITDTSIIINITTGDLOFVGAVMAGAGSEPTYSOPILVEILOPPINELKQTVIKVGEVNLLIPFGCK PRPVENKTEGOPLAGEE 1016 94UGGITDTSIIINITTGDLOFVGAVMAGAGSEPTYSOPILVEILOPPINELMQTVIKVGEVNLUPFGCK PRPVENKTEGOPLAGEE 1016 94UGGITDTSIIINITTGDLOFVGAVMAGAGSEPTYSOPILVEILEPPINELPPINELKQTVIRKVGTVNLAKVGFVRENKSNGBGALDSNO 866 Motif 8 cont. Motif 9 15DFDTVFFVQAASDSGEFLSVQIENMEDTATIRIKVG AGPPINVKEVWGTNALVEWQAPPDGNSEIMGFVQAADKKTMEWFNVERNHTS 113NSTDTILIIARAVVISTYQVTVRIENMEDKATIVLOVOF AGPPINVKEVWGTNALVEWPDDGNSEIMGFVQAADKKTMEWFVVLENVHTS 1123 NSFDTJILIIARAVVISTYQVTVRIENMEDKATIVLOVOF ASPPOLITVEKTYGFVALEWKPPDDGNSEIMGFVQAADKKTMEWFVVLEHVRTH 1139 1141 1150
hu-fst VHVXTSDFDTVFFVRQAAXSDSGEVELSVQIENMKDTATIRIRVVEKAGPPINVMVKEVWGTNALVEWQAPKDDGNSEIMGFVQKADKKTMEWFNVIERN hu-cc VSIRNSFDTILFIRAARVHSGTVQVTVRIENMEDKATLVLQVVDRSPPODLRVTDANGLNVALEWKPPQDVGNTELWGVTVQKADKKTMEWFTVLEH mo-cc VSIRNSFDTILFIRAARTHSGTVQVTVRIENMEDKATLIQIVDRSPPODLRVTDANGLNVALEWKPPQDDGNTEIWGVTVQKADKKTMEWFTVLEH ch-cc VGIRNSSTDTILFIRAARTHSGTVQVTVRIENMEDKATLIQIVDRSPPODIRIVETWGFNVALEWKPPQDDGNTEIWGVTVQKADKKTMEWFTVLEH Gh-cc VGIRNSSTDTILFIRKELHHSGAVEVTLQIENMTDTVAITIQIIDRSPPONIRIADVWGFNVALEWKPPDDGNGILFIRVGTVQKADKKTMEWFTVLEH VGIRNSSTDTILFIRKELHHSGAVEVTLQIENMTDTVAITIQIDRSPPONIRIADVWGFNVALEWKPPDDGNGETTGTVGKADKTMEWFTVDHY	KKSTRWMELNFDLIKELTTEAKUMIEGVATEMETIAVNSIGMSRESPASOPTWPIADESEPTHFTVEOVSDTTVALKWEDPEETGAGGLOGIVEYGKOG KKSTRWMELNFDPYETTYEAKUMIEGVATEMEVYAVNSIGMSRESPASOPTWPIADESEPTHFTVEOVSDTTVALKWEDPEETGAGGLOGIVEYGKOG KQSSRWMELNFDLKETTEPKKMIEGVATEMEVYAVNSIGMSRESPASOPTVPAPTSEVGLCVDDISDTSIVLKWEPPEEMGSVDLEGGGVEYGKEGT 858 KQSSRWMELNFDLCKETTEPKKMIEGVATEMEVYAVNSIGMSRESPASOPTVPAPTSEVGLCVDDISDTSIVLKWEPPEEMGSVDLEGGGVEYGKEGT Motif 7 cont. * Motif 8 * PANTERVENGGFTVNLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILQPRULOPRULOPRULOPVNLTGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILQPRULOPRULOPRULOPVNLLIPPGKPRPQVWTKGGAPLDTSR 886 PALGGTTERSSUVKDLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILQPRULOPRULOPRULOPVNLLIPPGKPRPQVWTKGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILQPRULOPRULOPTIKKVGEVNLLIPPGKPRPQVWTKEGGPLAGEE 1018 PALGGTTERSSUVKDLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILQPRULOPRULOPTIKKVGEVNLLIPPGKPRPKISMEDGGTLDSKD 1016 PALQGITERTSLIKDLYTGDKLIFFVRAHNAGPGAPVTTEPVTVQEILQPPTLEPVTVGTIKVGEVNLLIPPGKPRPKISMEDGGTLDSKD 1016 PALQGITERTSLIKDLYTGDKLIFFVRAHNAGPGAPVTIKEINGPTIREINGPFILLPPILKQTIKKVGEVNLLIPPGKPRPKISMEDGGTLDSKD 1016 PALQGITERTSLIKDLYTGDKLIFFVRAHNAGPGAPVTIKEINGPFILLOPPILLPPILKQTIKKVGEVNLNIPFGKPRPKISMEDGGTLDSKD 1016 PALQGITERTSLIKDLYTGDKLIFFVRAHNAGPSGPANTLAOPVTIKEINGPFILLPPILKQTIKKVGEVNLNIPFGKPRPKISMEDGGTLDSKD 1026 Motif 8 cont. Motif 9 15BFDTVFFVRAARSSGESELSVQIENMKDTATIRIKVE AGPPINVKEVWGTANLVEWQAPKDDGNSEIMGFVCAADKKTMEWFVVERNHTS 991 NSFTDTILFIRARSTISGTQVTVRIENMERKATUVLOVDE PSPPOLEVTDANGLVEWGAPRDGORSEIMGFVCAADKKTMEWFTVLENTRH 1119 NSSTDTILFIRARSTISGTQVTVRIENMERKATUVLOVDE PSPOLEVTDANGINVALEWKPPODGGTEIWGTVQAADKKTMEWFTVLENTRH 1121 NSSTDTILFIRARSTISGTQVTVRIENMERKATULVLOVDE PSPOLEVTDANGENVALEWKPPODGGNEIWGFVCAADKKTMEWFTVLENTRH 1121 NSSTDTILFIRKTERKENSKISKDLOVGUENVEDTAATURDTVALTIGTDENGFVALEWFPODGANGLIGTVQAADKKTMEWFTVLENTRHT 1121 NSSTDTILFIRKTERKENSKISKDLOVGUENVEDTAATURDTVALTIGTDENGFVALEWFPODGANGLIGTTUGTVAADKKTMEWFTVLEHTRHT 1121 NSGDTILFIRKTERKENSKISKDLOVGUENVEDTAATURDTVALTIOTDENGFVALEWFPODGANGLIGTVQAADKKTMEWTTVLEHTRHTT
hu-fst VHVXTSDFDTVFFVRQAAXSDSGEVELSVQIENMKDTATIRIRVVSKAGEPPINVMVKEVWGTNALVEWQAPKDDGNSEIMGFVQKADKKTMEWFNVIERN hu-cc VSIKNSFTDTILFIRAARVHSGTVQVTVRIENMEDKATLVLQVVDKPSPPOLRVTDAWGLNVALEWKPPQDVGNTELWGVTVQKADKKTMEWFTVLEH mo-cc VSIRNSFTDTILFIRAARWTHSGTVQVTVRIENMEDKATLILQIVDKPSPPOLRVTDAWGENVALEWKPPQDDGNTEIWGVVQKADKKTMEWFTVLEH ch-cc VGIRNSSTDTILFIRAELHHSGAVEVTLQIENMTDTVAITIQIIDKPGPPONIKLADVWGFNVALEWKPPQDDGNDGUIGYTQKADKKTMEWFTVLEH fu-cc ASVENSEGDTILFIRAELHHSGAVEVTLQIENMTDTVAITIQIIDKPGPPONIKLADVWGFNVALEWKPPDDGNDGNDIGYTQKADKKTMEWFTVDHY fu-cc HSVENSEGDTILFIRAELHSGAVEVTLQIENMTDTVAITIQIDKPGPPONIKLADVWGFNVALEWKPPDDGNAGIIGYTQKADKKTMEWFTVEH inirnsetdTilfirkaESHSGKIDLQVEVVDKFVETASIDIQIDKPGPPOVKIEDVWGENVALEWKPPKDDGNAEITGYTQKADKKSMEWFTVIEH	IKS FRWARLNFDLIKELTTEA KUMIEGVATEMETIAVNSIGMSRESPASOPSWPIADSESPTHFTVEDVSDTTVALKWEDPERIGAGGLOGIVEYGKOG KKS FRWARLNFDPYETTYEAKUMIEGVATEMEVYAVNSIGMSRESPASOPSWPIADSESPTLFTVEDVSDTTVALKWEDPERIGAGGLOGIVEYGKOG KQS SWMRLNFDLYETTYEAKUMIEGVATEMEVYAVNSIGMSRESPASOPSWPIAVTSEPVGLCVDDISDTSIVLKWEPPERIGAGGLOGIVEYGKOG Motif 7 cont. * Motif 8 * PANTERVENGGFTVNLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILGPRLQPFLLYDSVDTVGKPRPPVNLTGGAPLDTSR AALQGLTEHTSILVEDLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILGPRLQPFLLVGTVGKVGEVNLLIPFQGKPRPVVWTKGGAPLDTSR 886 AALQGLTEHTSILVEDLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILGPRLQLPRHLRQTIQKVGEVVNLLIPFQGKPRPVVWTKGGAPLDTSR 886 AALQGLTEHTSILVEDLPTGARLLFFVRAHNAGPGAPVTTEPVTVQEILGPRLQLPRHLRQTIQKVGEVVNLLIPFQGKPRPVTWTKEGQPLAGEE 1018 PALGGTTERTSILVEDLYTGAKLLFFVRAHNAGPGAPVTTEPVTVQEILGPRLQLPRHLRQTIQKVGEVNLLIPFQGKPRPKTWMEGGPLAGEE 1018 PALGGTTERTSLIKDLYTGAKLFFVRAHNAGPGAPVTTEPVTVQEILGPRLUPFNLRQTIQKVGEVNLLIPFQGKPRFKTWMEGGTLBSKD 1016 PAUGGTDETSLIFENSIVKULPTGAKLFFVRAHNAGPGAPVTIKEPTVQEIMGPTIREIMPOTICVENHLRQTIQKVGEVNLLIPFQGKPRFKTWMEGGTLBSKD 1016 PAUGGTDETSLIFENSTLIKDLYTGAKLFFVRAHNAGPGAPATIAQPVTIREIMPOTIEMGAPTURVTVURTUFTIKVGGTVNLAIPFQGKPRFKTWMEDGTLBSKD 1016 PAUGGTDETSLIFENSTLIKDLYTGAKLFFVRANNAGPSPANTLAQPVTIREIMOPTIREIMO PAUGGTDETSLIFENSTLIKDLYTGKFVRANNAGBSPANTLAQPVTIREIMOPTIREIMO NOTIF 8 cont. Motif 9 15BFDTVFVNQAALSDSGEVELSVQIENMKDTATIRIRVVEXAGPPINVKEVMGTNALVEWQAPKDGNSEIMGFVQADKKTMEWENVERNHTS 991 NSTDTILFIARRVFSTQAALSDSGEVELSVQIENMKDTATIRIRVVEXAGPPINVKEVMGTNALVEWQAPKDGNSEIMGFVQADKKTMEWETVLEHYRTH 1123 NSPTDTILFIARRVFSTQAALSDSGEVELSVQIENMERKATILLQIVVEXAGPPINVKEVMGTNALWKEVMGTNALWKFVQDARSEMFFVDAAKTMEWETVLEHYRTH 1121 NSPTDTILFIARRVFSGTVQVTVRIENMERKATILLQIVVEXAGPPINKKEVMGTNALWKFVQDARGTLUKKTVQADKKTMEWFTVLEHYRTH 1121 NSEGDTILFIKKEKSKDLQVGIENVEDTASTLVLUVVD PGPPOVKLEWKPPDDGNAQLIGTVQAADKKTMEWFTVLEHYRTH 1121 NSEGDTILFIKKEKSKSKDLQVGVVCFVETASJLDQUD PGPPOVKLEDVGRVALEWRPPDDGRAQUTGGTLEWETVDEHYVDHYRT 1121 NSECDTILFIKKEKSKSKDLQVVCFVETASJLDQUD PGPPOVKLEDVGRVALEWRPPDAKKTVCLANKKTMENTVLEHYRTH 1126
hu-fst VHV TSDFDTVFFVRQAA.SDSGEVELSVQIENMKDTATIRIRV/CKAPPIN/WKEVNGTNALVEWQAPKDDGNSEING FVQKADKKTMEWFNVFERI hu-cc VSINNSPTDTILETAARVVISGTQVTVRIENMEDKATLIQIVD RSPPQDLRVTDANGLNVALEWKPPDVGNTEINGTVVQKADKKTMEWFTVLEN volce VSINNSPTDTILETAARVVISGTQVTVRIENMEDKATLIQIVD RSPPQDLRVTDANGLNVALEWKPPDDGNTEINGTVVQKADKKTMEWFTVLEN volce VSINNSPTDTILETAARVVISGTQVTVRIENMEDKATLIQIVD (PSPPQDLRVTDANGLNVALEWKPPDDGNTEINGTVVQKADKKTMEWFTVLEN volce VSINNSPTDTILETAARVVISGTQVTVRIENMEDTVAITIQITD (PSPPQDLRVTDANGLNVALEWKPPDDGNTEINGTVVQKADKKTMEWFTVLEN volce VSINNSPTDTILETAARVVISGTNVLIENMTDTVAITIQITD (PSPPQDLRVTDANGINVALEWKPPDDGNAQILGYTVQKADKKTMEWFTVDVD hu-cc VSINNSFDTILETAKELNISGAVSVIQUENVEDTASVIQUENVEDTASVIQUENVEDVNICHVNUWGENVALEWKPPDDGNAQILGYTVQKADKKTMEWFTVDD hu-cc NSNSFDTILETAKEKSKIDLQVQIENVEDTASVIQUENVEDTASVIQUENVEDVNICHVNUWGENVALEWKPPDDGNAAITGKTUQKADKKTMEWFTVDD hu-cc NSNSFDTILETAKEKSKIDLQVQIENVEDTASVIQUENVEDTASVIQUENVEDVNICHVNUWGENVALEWKPPDDGNAAITGKTUQKADKKTMEWFTVDD hu-cc NSNSFDTILETAKEKSKIDLQVQIENVEDTASVIQUENVEDTASVIDU HOKTOD PDPQUKVDKVDKKTVETASIDIQUD HOKTOD HOKTON hu-cc NSNSFDTILETAKEKSKIDLQVKVDKKTVETASIDIQUD HOKTOD HOKTON HOTIF IRALES HSKKKDLQVKVDKKVETASIDIQUD HOKTOD HOKTON HOTIF IRAL SHSKKUDLQVKVDKKVETASIDIQUD HOKTOD HOKTON HOTIF IRAL SHSKKUDLQVKVDKVDKVDKVDKVDKVDKVETASIDIQUD HOKTOD HOKTON HOTIF IRAL SHSKKUDLQVKVDKVDKVDKVDKVDKVDKVDKVDKVDKVDKVDVNGHVALWKPFNDGNAAITGKTIQ HOTIF I NOTIF II HOTIF II	IKS FRAMELNEDLIKELTYEA KUMI GOVAREM TIAVASIGAS RESPASOPENDIADES EPTIFITE DUSDITIVAL WEDPER (AGGLOGIUSING SUB (AGGUOGIUSING SUB (AGGUOGIUS))))))))))))))))))))))))))))))))))))
hu-fst VHV TSDFDTVFFVRQAA.SDSGEVELSVQIENMKDTATIRIRVC KAGPPINVMVKEVNGTNALVEWQAPKDDGNSEIMG FVQKADKKTMEWFNVERI hu-cs VSINSPTDTILFI AAR.VISGTQVTVRIENMEDKATLUQVVD RSPFQDLRVTDANGLNVALEWKPPQDVGNTELWGTVQKADKKTMEWFNVERI VSINSPTDTILFI AAR.VISGTQVTVRIENMEDKATLUQVVD RSPFQDLRVTDANGLNVALEWKPPQDDGNTELWGTVQKADKKTMEWFTVLEN ch-cc VGINSTDTILFI AAR.VISGTQVTVRIENMEDKATLUQVVD (PSPFQDLRVTDANGLNVALEWKPPQDDGNTELWGTVQKADKKTMEWFTVLEN ch-cc VGINSTDTILFI AAR.VISGTQVTVRIENMEDKATLUQVVD (PSPFQDLRVTDANGLNVALEWKPPQDDGNAUGKTILGVTVQKADKKTMEWFTVLEN ch-cc VGINSTDTILFI KAELHNSGAYEVTLQIENMTDTVAITIQID PGPPQDIRVEGNAUKTPFVDGGNALEWKPPRDDGNAUGKTIGVTVQKADKKTMEWFTVIDV fu-cc svinsscotilfi kaelhnsgayevtlQi ENVEDTASVTLQVVD [PGPPEVLKIVDVWGFNVALEWKPPRDDGNAITGVTQKADKKTMEWFTVIDV hu-slw INIRNSECTILFI KAELSNEKKDLQVVLVDKVDKFVTASIDIQID PGPPEVLKIVDVWGFNVALEWKPRKDGNALTGVTQKADKKTMEWFTVIDV hu-slw Motif 9 cont. Notif 10 * hu-fst Notif 10 * *	IKS FRAMELAFDLIKELTTEAKUMI EGVATEMETIAVNSTGMSRPSPAGPENPIAPSEPTIFTYEDVSDTTVALKME PPERIGAGALDG IVEYGKUGG KKS FRAMELAFDLIKELTTEAKUMI EGVATEMEVYAVNSIGMSRHSQASQPFVPVAPTSSPVGLCVDDI SDTSIVLKAM PPERMGSVDLEG GVEYGKGT 858 KQSS FRAMELAFDLIKETTE PKIMIEGVATEMEVTAVNSIGMSRHSQASQPFVPVAPTSSPVGLCVDDI SDTSIVLKAM PPERMGSVDLEG GVEYGKGT 858 Motif 7 cont. * Motif 8 * PANTERVEIGGFTVNLPTGARLLFNVGVNIAGRTEPATLAQPVTIREIAFPHTLFDHLRQTYIRKVGGULUVVPGGFPHPQVWTKGGAPLDTSR 866 AALQGLTEHTSILVIDLPTGARLLFNVGVNIAGRTEPATLAQPVTIREIAFPHTLFDHLRQTYIRKVGGULUVVPGGFPHPQVWTKGGAPLDTSR 866 PAUGGTERTSILVIDLPTGARLLFNVGVNIAGPGGPTVTKEPVTVQELGPPLQLPHLRQTYIKVGGUNUVVPGGFPHPQVWTKGGAPLDTSR 866 940 GUTETSILVIDLYTGARLLFNVGVNIAGPGGPTVTKEPVTVQELGPPLQLPHLRQTYIKVGGUNUVPGGFPHPKISMKDGGTLDSKD 963 964 Motif 8 cont. Motif 9 755 FDTVFFVGAAASDSGEVISVGUIENKKDTATIRIRVE AGPPINVVELGPPLIZVENVGEVGEVALUPFGGFPHFKISMKDGGTLDSKD 991 NSFDTILFI AARTVISGTVQVTRIENMEDKATIVLQVV SPSPPDLIVVEIGFNALEWQAPKDGRSEINGFVQAADKKTMEWFVVERNHTS 991 NSFDTILFI AARTVISGTVQVTRIENMEDKATILIGUV SPSPPDLIVVEIGFNALEWKPPDDGRSEINGFVQAADKKTMEWFVVERNHTS 991 NSFDTILFI RAARVISGTVQVTRIENMEDKATILIGUV SPSPPDLIVVEIGFNALEWKPPDDGRSEINGFVQAADKKTMEWFVVERNHTS 991 NSFDTILFI RAARVISGTVQVTRIENMEDKATILIGUV SPSPPDLIVVEIGFNALEWKPPDDGRSEINGFVQAADKKTMEWFVVERNHTS 991 NSFDTILFI RAARVISGTVQVTRIENMEDKATILIGUV SPSPPDLIVVEIGFNALEWKPPDDGRSEINGFVQAADKKTMEWFVVERNHTS 991 NSFDTILFI RAARVISGTVQVTRIENMEDKATILIGUV SPSPPDLIVVEIGFNALEWKPPDDGRAQIGGTVQAADKKTMEWFTVLEHTHTH 1121 NSSTDTILFI RAARVISGTVQVTRIENMEDKATILIGUV SPSPPDLIVVEIGFNALEWKPPDDGRAQIGGTVQAADKKTMEWFTVLEHTHTH 1121 NSSTDTILFI RAARVISGTVQVDKFVETASIDIGUID PGGPQNI KLEDVMGENVALEWKPPDDGRAQILGTVQAADKKTMEWFTVLEHTHTH 1121 NSSTDTILFI RKENSKGKDLQVGVENTASTDIGUID PGGPQNI KLEDVMGENVALEWKPPDDGRAQILGTVQAADKKTMEWFTVLEHTHTH 1121 NSSTDTILFI RKENSKGKDLQVGVENTASTDIGUID PGGPPQNI KLEDVMGENVALEWKPPDDGRAQILGTVQAADKKTMEWFTVLEHTHTH 1121 NSSTDTILFI RKENSSGKNDLQVKEVETASIDIGUID PGGPPQNI KLEDVMGENVALEWKPPDDGRAQILGTVQAADKKTMENTVEFTNGTS971 NSTDTILFI RKENSKGKNDLQVKEVETASIDIGUID PGFPPNIKTEDVMELTNERVENV
hu-fst VHV TSDFDTVFFVRQAA.SDSGEVELSVQIENMKDTATIRIRVS KAGPPINVMVKEVNGTNALVENQAPKDDGNSEING FVQKADKKTMENFNVERI hu-cc VSILNSPTDTILFI AAR.VHSGTVQVTVRIENMEDKATLUQVVD PSPFQDLRVTDANGLNVALENKPPQDVGNTELWGTVQKADKKTMENFTVLEN volkinsptDTILFI AAR.THSGTQVTVRIENMEDKATLUQVD PSPFQDLRVTDANGLNVALENKPPQDDGNTELWGTVQKADKKTMENFTVLEN ch-cc VGINNSTDTILFI AKELHKSGAVEVTLQIENMTDTVAITIQID PGPFQDLRVTDANGLNVALENKPPQDDGNAULGVTVQKADKKTMENFTVLEN volkinsptDTILFI KAELHKSGAVEVTLQIENMTDTVAITIQID PGPFQNI KLADVNGFNVALENKPPQDDGNAQILGVTVQKADKKTMENFTVLON fu-cc VGINNSTDTILFI KAELHKSGAVEVTLQIENMTDTVAITIQID PGPFQUK KLUDVNGFNVALENKPPRDDGNAQILGVTVQKADKKTMENFTVTQD hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQIENMTDTVAITIQID PGPFQVLKIVDVNGFNVALENKPPRDDGNAQILGVTVQKADKKTMENFTVTQD hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTQD hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTQUD FGPFQVIK EVVNGKNALTEVFNDGKNALTEVTUD hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTQU hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTQU hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTQU hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTQU hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTQU hu-slw NUTKINSECTILFI KAELHKSGAVEVTLQUENVEKVKTVTUD hu-slw Nutkinsettilfi kaelkksga Nutkinsettilfi kaelkeve	IKS TENNELINFOLIKELTTEA KUMI GOVIEMETIAVNSTGASRESPASOPENDIADESE PTETYDUVSDTVAL WEDPEETGAGGLOGTVEYDKOG KKSTENNELINFDPYETTEAKENI GOVIEMEVTAVNSTGASRESOASOPEVDVAPESE PTELYDVSDTVAL WEDPEETGAGGLOGTVEYDKOG KOSSENNELNFDDEKETTE PKKMI GOVIEMEVTAVNSTGASRESOASOPEVDVAPESE PTELYDVSTVIKWEDPEENGSVDLEG GOVIEYDKEG KOSSENNELNFDDEKETTE PKKMI GOVIEWTAVNSTGASRESOASOPEVDVAPESE PTELYDUSDTVTTVIKWE PPENGSVDLEG GOVIEYDKEG Motif 7 cont. * Motif 8 * PANTEDVETGGETVKLPTCARLIF VVGVNTAGRTEPATLAOPTTEETASPENTLEVPERUTYDSVDTVTTVIKWE PPENGAGLOGTVEYDTEG PALOGIEHERSTLVKOLPTCARLIF VVGANNAGGEGAPTTTEPTVVSE LOOPALOLPHILKOTTOKK VGEPVNLLIPFOGR PAPOVWTKGGAPLDTSK PALOGIEHERSTLVKOLPTCARLIF VVGANNAGGEGAPTTTEPTVVSE LOOPALOLPHILKOTTOKK VGEPVNLLIPFOGR PAPOVWTKGGAPLAGE PALOGIEHERSTLVKOLPTCARLIF VVGANNAGGEGAPTTTEPTVVSE LOOPALOLPHILKOTTOKK VGEPVNLLIPFOGR PAPOVWTKGGAPLAGE 1014 PALOGIEHERSTLVKOLPTCARLIF VVGANNAGGEGAPTTTEPTVVSE LOOPALOLPHILKOTTOKK VGEPVNLLIPFOGR PAPLEKKSSKE PALOGIEHERSTLVKOLPTCARLIF VVGANNAGEGAPTTEEVOVOLOOPTIEK MOOPHILKOTTOKK VGEPVNLLIPFOGR PAPLEKKSNEGOTLDSKO 1016 PALOGIEHERSTLVKOLFTCKVATANAGESGAATI KEPTVOSI KOPPTORE LOOPALOLPHILKOTTOKK VGEPVNLLIPFOGR PAPLEKKSNEGOTLDSKO 1016 PAUGGITDSTIIINITTCOKLOFVVATINAGESGAATI KEPTVOSI KOPPTORE MOOPHILKOTTIKK VGKTININTPFOGR PAPLEKKSNEGOTLDSKO 1016 1017 1018 1019
hu-fst VHVITSDFDTVFFVRQAALSDSGEVELSVQIENMKDTATIRIRVE KAGPPINVMVKEVNGTNALVENQAPKDDGNSEING FVQKADKKTMENFNVFER hu-cs VSINSPTDTILFIAARVHSGTVQVTVRIENMEDKATLUQVVD RSPPODENTDANGLNVALENKPPQDVGNTELWGTVQKADKKTMENFTVLEH mo-ce VSINSPTDTILFIAARTHSGTQVTVRIENMEDKATLUQVD RSPPODENTVENGKVALENKPPQDDGNTELWGTVQKADKKTMENFTVLEH vGIRNSSTDTILFI KAELHHSGAVEVTLQIENMTDTVAITIQID RGPPONIKLADVNGFNVALENKPPQDDGNAGILGVTQKADKKTMENFTVLEH vGIRNSSTDTILFI KAELHHSGAVEVTLQIENMTDTVAITIQID RGPPONIKLADVNGFNVALENKPPQDDGNAGILGVTQKADKKTMENFTVLEH hu-se VSINSEDTILFI KAELHHSGAVEVTLQIENMTDTVAITIQID RGPPOVIKLADVNGFNVALENKPPRDDGNAGILGVTQKADKKTMENFTVGDQ hu-se Notif 9 cont. Notif 10 * hu-fst CTVSDLIVGNEYTFRVTENIGGLSDSPGVSKNTARILKIGITF PFEKEHDFRMAPKFLTPLDRVVVGSAALNGAVNGP KPVVNMKNMEMEIREL hu-c VVSELIGNGYFRVFSHNNGSSDKAATTEPVFIPPOFTFPPGITTEPPNKKALDFSEAPSFTQVINKSVIAGTAMLCAVIGSPKPISWFKNGLDLGET ch-cc CVVSELIGNGYFRVFSHNNGSSDKAATTEPVFIPPOFTFPPKKALDFSEAPSFTQPLANKSVIAGTNTLSIAVIGFPKPISWFKNGLDLGET	IKSYEWWELNFDLIKELTTEAKUMIEGVATEMUIGAVEKTEMUTAVNSIGMSRESDASOPHVPIAPESEPTHFTVEUVSDTVALWEPPERIGAGALOGIVEXTKOGS SAMELNFDLYETTEPKKMIEGVATEMUTAVNSIGMSRESDASOPFVPIAPESEPTHFTVEUVSDTVALWEPPERIGAGALOGIVEXTKOGS 80 KOSSEWWELNFDLYETTEPKKMIEGVATEMUTAVNSIGMSRESDASOPFVPIAPESEPTHFTVEUVSDTVALWEPPERIGAGALOGIVEXTKOGS 80 Motif 7 cont. * Motif 8 * PANTERVEIGGETVNLPTGARLEFVVGNIAGRTEPATLAOPVTIBEIASPPILAVDSVDTVDSVDTVJKWEPPERIGAGALOGIVEXTGF 866 PALGGLEERSMIVKDLPTGARLEFVVGNIAGRTEPATLAOPVTIBEIASPPILAVDSVDTVDSVKOFVNLLIPFOGKPEPOVWTKGGAPLDTSR 867 806 806 806 806 806 806 806 806
hu-fst VHV TSDFDTVFFVRQAA SDSGEVELSVQIENMKDTATIRIRVE KAGPPINVMVKEVNGTNALVENQAPKDDGNSEING FVQKADKKTMENFNVERN hu-cc VSINSPTDTILFI AAR VHSGTVQVTVRIENMEDKATLUQVVD RSPPODENTDANGLNVALENKPPODVGNTELNGGTVQKADKKTMENFTVLEN mo-cc VSINSPTDTILFI AAR THSGTVQVTVRIENMEDKATLUQVVD RSPPODENTDANGLNVALENKPPODOGNETENG VTVGKADKKTMENFTVLEN vGIRNSSTDTILFI KAELHHSGAVEVTLQIENMTDTVAITIQID RPGPPOIR KLADVNGENVALENKPPODDGNADIGVTVQKADKKTMENFTVLEN vGIRNSSTDTILFI KAELHHSGAVEVTLQIENMTDTVAITIQID RPGPPOIR KLADVNGENVALENKPPODGNADIGVTVQKADKKTMENFTVLEN vGIRNSSTDTILFI KAELHHSGAVEVTLQIENMTDTVAITIQID RPGPPOIR KLADVNGENVALENKPPPODGNADIGVTVQKADKKTMENFTVUEN tu-cc ASVENSEGDTILFI KAELHHSGAVEVTLQIENMTDTVAITIQID RPGPPOIR KLADVNGENVALENKPPRDDGNETIGVTVQKADKKTMENFTVEDQ hu-si Notif 9 cont. Notif 10 * hu-fst CTVSDLIVGNE IFFVYTENIGGLSDSPGVSKNTARILKTGITF PFEKEHDFRMAPKFLTPLDRVVVAG SAALNGAVNGP KPVVNMKNMENEIRET hu-cc Notif 9 cont. Notif 10 * hu-fst CTVSDLIVGNE IFFVYTENIGGLSDSPGVSKNTARILKTGITF PFEKEHDFRMAPKFLTPLDRVVVAG SAALNGAVNGP KPVVNMKNMENEIRET hu-cc Notif 9 cont. Notif 10 * hu-fst CTVSDLIVGNE IFFVYTENIGGLSDSPGVSKNTARILKTGITF PFEKEHDFRMAPKFLTPLINKSVIAGVTANLCAVNGS PKP ISVFKNGLDGEN mo-cc VVSLLIGNGYFFVSHNVGSSDRAATTEPVFIPPOITY PPOITY PPOITY SPARSTQPINKSVIAGVTNSVIAGVTANLCAVNGSPKP ISVFKNGLDGEN mo-cc VVSLLIGNGYFFVSHNVGSSDRAATTEPVFIPPOITY PPOITY PPNK KALDFSEAPSTQPINKSVIAGVTNTLS AVNGSPKP ISVFKNGLDGEN mo-cc VVSLLIGNGYFFVFSHNVGSSDRAATTEPVFIPPOITY PPNK KALDFSEAPSTQPINKSVIAGVTNTLS AVNGSPKP ISVFKNGLDGEN mo-cc VVSLLIGNGYFFVFSHNVGSSDRAATTEPVFIPPOITY PPNK KALDFSEAPSTQPINKSVIAGVNTLS SVNGSPKP ISVFKNGLDGEN mo-ccc VVSLLIGNGYFFVFSHNVGSSDRAATTEPVFIPPOITY PPNK KAL	IKSTENNELINFOLIKELTTERAKUMI GOVIEMETIAVNSTGMSRESPASOPENPIAPSEEPTHFTVEOVSDTTVAL WEIPPERIGAGALOGIVEXTKOGS S11 KKSTENNELINFDPYETTYEAKUMI GOVAEMEVYAVNSIGMSRESOASOPEVPARTSEVGLOVDI ISDTSIVLKWEIPPERIGAGALOGIVEXTKOGS S25 SUMELINFOLCKSTTEPKKMI GOVAEMEVYAVNSIGMSRESOASOPEVPARTSEVGLOVDI ISDTSIVLKWEIPPERIGAGALOGIVEXTKOGS S25 SUMELINFOLCKSTTEPKKMI GOVAEMEVYAVNSIGMSRESOASOPEVPARTSEVGLOVDI ISDTSIVLKWEIPPERIGAGALOGIVEXTKOGS Motif 7 cont. * Motif 8 * PANTEEVEICGETVVNLPTGARLIFVVGANNAGRTEPATLAOPVTIBEIASPPILAVDSVDTIVTSGPAPHILGAGALOUVEKGEPRILA PALOGITERSSILVID PTGARLIFVVGANNAGRTEPATLAOPVTIBEIASPPILAVDSVDTIVKSGLAUVUPFGKEIPPOVVNTKGGAPLDTSR S44CGUTERRSILVID LPTGARLIFVVGANNAGRTEPATLAOPVTIBEIASPPILAUVOSVDTIVKVGEPVNLLIPFGKEIPPOVVNTKGGAPLDTSR S44CGUTERRSILVID LPTGARLIFVVGANNAGGAPUTTPUTVOSI LOOPPLOLPHILQTIOKVGEPVNLLIPFGKEIPPOVVNTKGGAPLDTSR PALOGITERSSILVID LPTGARLIFVVGANNAGGAPUTTPUTVOSI LOOPPLOLPHILQTIOKVGEPVNLLIPFGKEIPKKSSMEDGGTLSSKD 1016 PALOGITERTSILIKUUTDALIFVVGANNAGSSGAAIIKEPTVVGEIMGPHILVPHLAQTIVKVGTINIMPFGCKEIPKKSSMEDGGTLSSKD 1016 PAUGGITDTSIIINILTTCDLIOFVGANNAGSSEPTYSOPILVEITEPPTIREINOTIIPPKIKVGTINIMPFGCKEIPKKSSMEDGGTLSSKD 1016 MOTIF 8 cont. Motif 9 TSDFDTVFVQAASSDSGEELSVQIENMKDATIRIRVVSKAGPPUTIKEIMONTVKEWGTALVEWORADKKTMEWENVERSNERSIST 1020 NAKDLID VKTVIKIENMEDKATILILIVVSKAGPPINKEUMGVALWEVDEDGNSEIMGFVQAADKKTMEWENVERNIHTS 1121 NARRUNSCT VQVTVRIENMEDKATILILUVDE AGPPINKLAUVEWQTALEWKPPODGSKEIMGFVQAADKKTMEWETVLEHYRTH 1123 NSETDTILFI RAARTISCT VQVTVRIENMEDKATILUGVOF PSPPOLIKUTUAMGINVALEWKPPODGSKEIMGFVQAADKKTMEWETVLEHYRTH 1121 NSETDTILFI KAARTISCT VQVTVRIENMEDKATILILUVDE AGPPINKLAUVEWQTALEWKPPODGSKEIMGFVQAADKKTMEWETVLEHYRTH 1121 NSETDTILFI RAARTISCT VQVTVRIENMEDKATILLUVDE AGPPINKLAUVEWGEVALEWKPPODGSKEIMGFVQAADKKTMEWETVLEHYRTH 1121 NSETDTILFI KAESKSKKDLOVVINKEVETASIDIGITD PGPPOIK LUVDWGEVALEWKPPDDGSKAALSKYPVQADAKTMEWTVLEHYRTH 1121 NSETDTILFI KAESKSKKDLOVVINKEVETASIDIGITD PGPPOIK LEDVGENVALEWKPPDDGSKAALGAVGSPIPHISWETVLEHYRTH 1121 NSETDTILFI KAESKSKKDLOVVINKEVETASIDIGITD PGPPOIK LEDVGENVALEWKPPDDGSKAALSKYPPOVYGAD
hu-fst VHVITSDFDTVFFVRQAALSDSGEVELSVQIENMKDTATIRIRVE KAGPPINVMVKEVNGTNALVENQAPKDDGNSEINGFVQKADKKTMENFNVERV VSINNSPTDTILFIRAARVHSGTVQVTVRIENMEDKATLUQVD RSPPODLRVTDANGLNVALENKPPQDVGNTELNGGTVQKADKKTMENFTVLEN vor vsinnsptdtilfiraarthsgtvurgenvaltagtugenvaltensptvgkadkktmenfvler chcc VSINNsptdtilfiraarthsgtvurgenvaltagtugenvaltagtugenvaltensptvgkadkktmenfvler vsinnsptdtilfiraarthsgtvurgenvaltagtugenvaltagtugenvaltagtugenvaltagtugenvaltensptvgkadkktmenfvler ducc chcc VSINNSptdtilfiraarthsgtvurgenvaltagtugen	IKSY RAMELNFDLIKELTTEA KUMI GOVIEME ILAVNSIGMSRESASOPENDIADESEPTETTEDVSDTVALK MEDPERIGAGALOG IVEYGKOG KKSY RAMELNFDPYPTTYEAKEMIEGVALEMEVKAVNSIGMSRESASOPENDIADESEPTETVOVSDTVALK MEDPERIGAGALOG IVEYGKOG KQSS RAMELNFDLYETTEPKKMI GOVALEMEVKAVNSIGMSRESASOPENDIADESEPTETVOVSDTVALK MEDPERIGAGALOG IVEYGKOG Motif 7 cont. * Motif 8 * PANTERVEIGGTVVNLPTGARLLFVVGVNIAGRTEPATLAOPVTIREIASPPTILTVDSVDTVDSVDTVJKMR PPHIGAAGLOG VLEYGKEG RALGUTEHTSILVKOLPTGARLLFVVGVNIAGRTEPATLAOPVTIREIASPPTILTVDSVDTVDSVDVNIAGGAPLDTSR RALGUTEHTSILVKOLPTGARLLFVVGANNAGPGGPTVTKEVVVGEILOPPHILAOTIOKVGEVNLLIPPGKPRPQVWTKGGAPLDTSR PALDGITERSKILVKOLPTGARLLFVVGANNAGPGGPTVTKEVVVOEILOPPHILAOTIOKVGEVNLLIPPGKPRPVVWTKGGAPLDTSR PALDGITERSKILVKOLPTGARLLFVVGANNAGPGGPTVTKEVVVENLOPPHILAOTIOKVGEVNLLIPPGKPRPVIXTEGOPLAGE 1014 PALPGITERTSALKOLVTGDKLIFVVGANNAGPGGPTVTKEVVVENLOPPHILAOTIOKVGEVNLLIPPGKPRPXISSMEDGGTLDSKD 1016 PALPGITETSALKOLVTGDKLIFVVGANNAGPGGPTVKEVVVENLOPPHILAOTIOKVGEVNLLIPPGKPRPXISSMEDGGTLDSKD 1016 PALPGITETSALKOLVTGVKAVNAGFGGPTVKEVFVVENLOPPHILAOTIOKVGEVGVNLLIPPGKPRPXISSMEDGGTLDSKD 1016 PALPGITETSILINUTTGDKLIFVVGANNAGPGGPTVKEVFVVENLOPPHILAOTIOKVGEVNLVVFGKPRPXISSMEDGGTLDSKD 1016 PAUGGITDSTILINITTGDKLOFVVGANNAGSSENTYSOPTIEN 1017 101
hu-fst VHV TSEFDTVFFVRAALSDSGEVELSVQIENMKDTATER RVVE CAGPENUVVKEVNGTNALVEWAAPKDEGSEING FVOLADKTMEWFNVEREN hu-cs VSINSPTDTILFI AARVVEGTKOVTVRIENMEDKATLUJUVD RSPPODLRVTANGENVALEWKPPODGSTEING FVOLADKTMEWFNVERE VSINSPTDTILFI AARVVEGTKOVTVRIENMEDKATLUJUVD RSPPODLRVTANGENVALEWKPPODGSTEING FVOLADKTMEWFTVLEH hu-cs VSINSPTDTILFI AAREVISGTKOVTVRIENMEDKATLUJUVD RSPPODIR LADVEGFVALEWKPPODGSTEING FVOLADKTMEWFTVLEH hu-cs VSINSPTDTILFI KARLENSGANEVTOINMEDVALTUGIVD RSPPODIR LADVEGFVALEWKPPODGSTEING FVOLADKKTMEWFTVIEH hu-siw ININSSTDTILFI KARLENSGANEVTOINMEDVALSVILOVUD PGPPEVIR LADVEGFVALEWKPPODGSTEING FVOLADKKTMEWFTVIEH hu-siw ININSSTDTILFI KARLENSGANEVTOINMEDVALSVILOVUD PGPPEVIR LADVEGFVALEWKPPEDDGSTIGTVOKADKKTMEWFTVIEH hu-fst TVSDLIVGNEVFFVGSSDKAATTEPVIPPEPERGTEPPEKENDEGSANEVTOFADKKSMEWFTVIEH hu-fst TVSDLIVGNEVFFVGSSDKAATTEPVIPPEPGITEPPKYALDFSEAPSTOPINWINSVIAGSALNAVGHPKPVUMMENMEMEIRET mo-cc VVSELIIGNCYFFVSSNLGLSSTAATTEPVIPPEPGITEPPKYALDFSEAPSTOPINKSVIAGSNAILCAN GSPEPISWFNGLDGGEI ch-cc CVVSELIIGNCYFFVSSNLGSSDKAATTEPVIPPEPGITEPKYALDFSEAPSTOPINKSVIAGSNAILSSVIGSPEPISWFNGLDGGE hu-siw ATTELVIGNEVFFVSSNLGSSDKAATTEPVIPPEPGITEPKYALDFSEAPSTOPINKSVIAGSNAILSSVIGSPEPISWFNGLDGSE hu-siw ATTELVIGNEVFFVSSNLGSSDATMTESAVIADCHTPPSTPFISTPLVNKSVIAGSNAILSSVIGSPEPISWFNGLDGSE hu-siw ATTELVIGNEVFFVSSNLGSSDATMTESAVIADCHTPPSTPFISTPLVNKSVIAGSNAILSSVIGSPEPISMFNGLDGSE hu-siw ATTELVIGNEVFFNSSNLGSSDATMTESAVIADENTESAVIADFNOVIGSPEPISMFNGUSSDKAATSEVVGCD- 1142	IKSTENNELINFOLIKELTTEAKUMIEGVNEMTIGVNEMTIGVNEMTIGVNEMTIGVNEGSENGEPROVINDESEPTIFTVEDVSDTVALKWEPPERGAGGLOGIVEKTADG KASTENNELINFDPYETTYEAKUMIEGVNEMKVAVNSIGMSRESASOPFVPVAPTSEPVGLCVDDISDTSIVLKWEPPERGAGGLOGIVEKTADG KOSSENNELINFDICKSTTEPKKMIEGVNEMKVAVNSIGMSRESASOPFVPVAPTSEPVGLCVDDISDTSIVLKWEPPERGAGGLOGIVEKTAG 858 KOSSENNELINFDICKSTTEPKKMIEGVNEWTIAGNESATEVNIGSKEPSMPERPVLLVTSSPUTLTVDSVTDTVTMEMEPPERGAGGLOGIVEKTAG 858 Motif 7 cont. * Motif 8 * PANTERVENGGFTVNNLPTGARLLFFVRAINAGPGAPVTTEPTVVQELLOPPLLVTSPTLLVDSVTDTVTMEMPPPHGARAFDQVWTKGGAPLDTSR 886 PALOGLTEHRSHLVKDLPTGARLLFFVRAINAGPGAPVTTEPTVVQELLOPPLLQPHLLQTIOKVGEPVNLLIPPGKPRPQVWTKGGAPLDTSR 886 PALOGLTEHRSHLVKDLPTGARLLFFVRAINAGPGAPVTTEPTVVQELLOPPLLQPHLLQTIOKVGEPVNLLIPPGKPRPKTKEGOPLAGEE 1018 PALOGLTEHRSHLVKDLPTGARLLFFVRAINAGPGAPVTTEPTVVQELLOPPLLQPHLLQTIOKVGEPVNLLIPPGKPRPKTKEGOPLAGEE 1018 PALOGLTEHRSHLVKDLPTGARLFFVRAINAGPGSPANTLAOPVTIREIMGPHTLLPPLLQTIVKVGETNILTPFGKPRPKTKEGOPLAGEE 1018 PALOGLTEHRSHLVKDLPTGARLFFVRAINAGPGSPANTLAOPVTIREIMGPHTLLPPLLQTIVKVGETNILTPFGKPRPKTKSMEDGGTLDSKD 1016 PALOGLTDTSLIINITTGDLUFFVRAINAGPSSANTLAOPVTIREIMGPHTLLPPLIKQTIIRVGGAVNLVIPPGKPRPKTKSMEDGGTLDSKD 1026 Motif 8 cont. Motif 9 TSBFDTVFFVQAARSDSGETSVQIENNKDTATIRIKVG AGPPINVKKVWGTNALVEWGAPNLLWFPGKPRPKTKMENGAGELDRND 866 Motif 8 cont. Motif 9 158FDTVFFVQAARSDSGETSVQIENNKDTATIRIKVG AGPPINVKKVWGTNALVEWGAPNDGNSEIMGFVVAARMEMETVLEHYMEN 119 NSFTDTILFI RARKVISGKVDLQVVRIENMEDKATULQUV SAGPPINVKKVWGTNALVEWGAPNDGNSEIMGFVVAARMEMETVLEHYMEN 1119 NSFTDTILFI RARKVISGKVDLQVVRIENMEDKATULQUVD PSPPQDIRIVETWGFVALEWKPPQDGGNSEIMGFVVAARMEWFTVLEHYMEN 1121 NSGTDTILFI KAELHISGARSVTLGIENMEDTATIRIKVG AGPPINVKKVWGTNALEWKPPQDGGNSEIMGFFVAARMEWFTVLEHYMEN 1121 NSGTDTILFI KAELHISGARSVTLGIENMEDTATIRIKVG SAGPPONT ALDFSSPOOLSTURGVAARMEPPONGAALGETVGAADKKMEWFTVLEHYMEN 1121 NSGTDTILFI RARKVISGKVDLQVVRIENMEDKATULQUPPON NSGTDTILFI RARKVISGKVDLQVVRIENTEMETATILFGITPFPENENCHLADVGGTVALEWKPPONGAALGEVGENGETAGTVORADKTMEWFTVLEHYMEN 1121 MOTIF 9 cont. MOTIF 9 cont. MOTIF 9 cont. MOTIF 9 CONT. MOTIF 9 CONT. MOTIF
hu-fst VHVITSDFDTVFFVRQAALSDSGEVELSVQIENMKDTATIRIRVDE KAGPPINVMVKEVNGTNALVEWQAPKDDGNSEINGFVQKADKKTMEWFNVFER hu-cc VSINSPTDTILFI AARAVISGTQVTVRIENMEDKATLUQVD ADSPPQDLRVDANGLWALEWKPPDDVGNTEWGTVQKADKKTMEWFNVFER ch-cc VSINSPTDTILFI AARAVISGTQVTVRIENMEDKATLUQVD (PSPPQDLRVDENGENVALEWKPPDDGNTEWGTVQKADKKTMEWFTVLEN ch-cc VGINSSTDTILFI AKREKNISKIDLQVLVURISKADLUQVD (PSPPQDLRVDENGENVALEWKPPDDGNALIGYTVQKADKKTMEWFTVLEN ch-cc VGINSSTDTILFI KKEKKSKUDLQVLVURISKADLUQVD (PSPPQDLRVDENGENVALEWKPPDDGNALIGYTVQKADKKTMEWFTVDD fu-cc VGINSSTDTILFI KKEKKSKUDLQVLVURISKETASIDIQID (PGPPVIKILDVWGENVALEWKPPDDGNALIGYTVQKADKKTMEWFTVDD) hu-siw ININNSEDTILFI KKEKKSKUDLQVKUNKFVETASIDIQID (PGPPVIKILDVWGENVALEWKPPDDGNALIGYTVQKADKKTMEWFTVDD) hu-siw Notif 9 cont. Motif 10 * hu-fst TVSDLVGKEYFFVSGNMGSFDSAAATTEPVFIPPGTTEPPFKKALDFSEAPSTQLVNKSVAAGKAALCAVGSPKPISWFNGLDGEL cvvsELIGNGYFFVSGNMGSFDSAAATTEPVFIPPGTTEPPKKALDFSEAPSTQLNNSVIAGYAALCAVGSPKPISWFNGLDGEL ch-cc VVSELIGNGYFFVSENNGGSSKAAATEPVFIPPGTTEPPKKALDFSEAPSTQLNNSVIAGYAALSSV GSPKPISWFNGLDGGEL hu-siw ATTELVIGNEYFFVSENNGGSSCKAAATEPVFIPPGTTEPPSKKALDFSEAPSTQLNNSVIAGYAALSSV GSPKPISWFNGLDGEL hu-siw ATTELVIGNEYFFVSENNGGSSCKAATTEPVFIPPGTTEPPKKALDFSEAPSTQLNNSVIAGYNAILS SV GNPKPISWFNMENTIS hu-siw ATTELVIGNEYFFVSENNGGSSCKAATTEPVFIPPGTTEPPSKKALDFSEARKTSFJVVNSVIAGYNAILS SV GNPKPISWFNMENTIS hu-siw ATTELVIGNEYFFVSENNGGLSEDATMTESAVIADCKTTEPPCT HP10HSEARKTSFJVVNSVIAGYNAILS SV GNPKPITWMENVAIVDISG hu-siw ATTELVIGNEYFFVFSENNGGLSEDATMTESAVIADGKIFNPVEDFDFSEAPKTPJVNTYAIAGYNAILS SV GNPKPITWMENVAIVDI Motif 10 * hu-fst TNYGVULLIKEKPCPFEGGTVCKAVELGEALACKLEVVFQ- 1142 hu-fst TNYGVULLIKEKPCPFEGGTVCKAVELGEALACKLEVVFQ- 1274	<pre>KKSTRWMELNFDLKKLTTEAKUMIEGVXEMIIAVNSTGMSRPSPASOPNYLAPSEPTHFTYEOVSDTVALKWEDPERGAGGLOGIVEYGKUG KKSTRWMELNFDPYETTYEAKUMIEGVXEWRVAVNSIGMSRHSQASOPFVPVAPTSEPVGLCVDISDTSIVLKWEDPERGAGGLOGIVEYGKUG 858 KQSSRWMELNFDLCKSTTEPKKMIEGVXEWRVAVNSIGMSRHSQASOPFVPVAPTSEPVGLCVDISDTSIVLKWEDPERGAGGLOGIVEYGKUG 858 KQSSRWMELNFDLCKSTTEPKKMIEGVXEWRIFAVNAIGIKSPSMPERPVPLAVTSPFVLLVDSVDTVJKWEDPERGAGGLOGIVEYGKUG 858 Motif 7 cont. * Motif 8 * PANTERVENGGFTVNLPTGARLLEFVRAHNAGPGAPVTTEPVTVQELQPRLQPFNLLQPTIQKVGEVNLLIPPGKPRPQVWTEGGAPLDTSR 846 AALQGLTERSHLVGLPTGARLLEFVRAHNAGPGAPVTTEPVTVQELQPRLQPFNLLQPTIQKVGEVNLLIPPGKPRPQVWTEGGAPLDTSR 846 AALQGLTERSHLVGLPTGARLLEFVRAHNAGPGAPVTTEPVTVQELQPRLQPFNLLQPTIQKVGEVNLLIPPGKPRPQVWTEGGAPLAGE 910 GLTERSHLVGLPTGARLLEFVRAHNAGPGAPVTTEPVTVQELQPRLQPFNLLQTIQKVGEVNLLIPPGKPRPKTSMSDGGTLDSKD 1016 PALQGLTERSHLVGLPTGARLEFVRAHNAGPGAPVTTEPVTVQELQPRLLQPFNLLQTIVKVGEVNLLIPPGKPRPKTSMSDGGTLDSKD 1016 PALQGLTERSHLVGLPTGARLEFVRAHNAGPGSPANTLAOPVTIREIMOPPTIREIMOPTILERVGTVNLAFVGGTNNLAFPGKPRPKTSMSDGGTLDSKD 1016 PAVQGLTDRTSLITENITGDKLOFVRAYNAGPSPANTLAOPVTIREIMOPPTIREIMOPTINEVGTVNLAFPGKPRPKTSMSDGGTLDSKD 1016 Motif 8 cont. Motif 9 TSBFDTVFFVAAALSDSGFLSVQIENMKDTATIRIRVVE KAGPPINVKEVGENALVEWQAPKDGNSELMGFVQADKKTMEWENVPERNHTS 991 NSFTDTILIFARNUFSGTQVTVRIENMEKATLILQIVO PSPPOLEVTDANGINALVEWQAPKDGNSELMGFVQADKKTMEWFVVERNHTS 991 NSFTDTILIFARSTVGVGADKKTMATTIRIRVVE KAGPPINVKEVUGENALVEWQAPKDGNSELWGFVQADKKTMEWFTVLEHYRTH 1123 NSFTDTILFI AARNUFSGTQVTVRIENMEKATLILQIVO PSPPOLEVTDANGINVLEWKPPDDGNAALGETUGTVQADKKTMEWFTVLEHYRTH 1121 NSFTDTILFI AARNUFSGKDLQVVVRIENMEKATLILQIVO PSPPOLEVTDANGINVLEWKPPDDGNAALGETUGTVQADKKTMEWFTVLEHYRTH 1121 NSFTDTILFI KAELHSGAVEVTDIENMEKATLILQIVO PSPOLEVTDANGINVLEWKPPDDGNAALGETUGTVQADKKTMEWFTVLEHYRTH 1121 NSFTDTILFI KAELHSGKVDLQVTVRIENMEKATLILQIVG PSPOLEVGTVQADKKTMENTULEHYRTH 1121 NSFTDTILFI KAELHSGAVEVTDIENMEKATLILQIVGFYALENDPFFYNTHAUHTPPKDDGNAALGETUGTVQADKKTMEWFTVLEHYRTH 1224 DLIVGNEFFFYDGAATTSGKVDLQVKTVFYTAGITFPFYTREGITYEPPNYKALDFSEAFSFQPLNKSVLAGVATLSSVGGPPNISKTKNDLDGEDAFF 1224 OVITTEI KTEKKNSKSKDLQVVGFSSRAATSEPYTPPTFFFFYTFFFFFYF</pre>
hu-fst VHV TSDFDTVFFV QAALSDSGETELSVQIENMKDTATIRIRVC KAGPPINVMVKEVNGTNALVENQAPKDDGNSEING FVQKADKKTMENFNVERI hu-cst VSINSPTDTILET AARAVISGTKQVTVRIENMEDKATLUQVD RSPPQDLRVDANGLNVALENKPPQDVGNTENGSTVQKADKKTMENFTVLEN mo-cc VSINSPTDTILET AARAVISGTKQVTVRIENMEDKATLUQVD RSPPQDLRVDANGLNVALENKPPQDDGNTENGGVTVQKADKKTMENFTVLEN ch-cc VGINSSTDTILET AARAVISGTKQVTVRIENMEDKATLUQVD (PSPPQDLRVDENGKTENKPPDDGNAULENKPPQDDGNAULGNTVQKADKKTMENFTVLEN ch-cc VGINSSTDTILET AARAVISGTKQVTVRIENMEDKATLUQVD (PSPPQDLRVDENGKTENKPPDDGNAULENKPPQDDGNAQILGTVQKADKKTMENFTVDG) hu-sc VSINSEGDTILET KAKESKIDLQVQLENNEDTASVILQVD (PGPPEVLKIVDVNGENVALENKPPRDGCNAITGTTQKADKKTMENFTVDQ) hu-slw ININNSEDTILFTRKENSKKIDLQVKUKKPETASIDIQID PGPPEVLKIVDVNGENVALENKPRDGNAATGTGTVGKADKKTMENFTVDQ) hu-slw Notif 9 cont. Notif 10 * Notif 10 * hu-sc CVVSELIGNGVFFNVFSONNVGFSDRAATTEPVFPPFPGTTEPPKKHADF SAPSTQFLANKSIAGNATLCAV GSPEPT SNFNGLDLGET mo-cc CVVSELIGNGVFFNVFSONNVGFSDRAATTEPVFPPFPFPGTTEPPKK HADF SAPSTQFLANKSIAGNATLCAV GSPEPT SNFNGLDLGET hu-sc VVSELIGNGVFFNVFSONNGGSDRAATTEPVFPPFPFPFPFPFFNVALDFSAPSTQFLANKSIAGNATLCAV GSPEPT SNFNGLDLGET hu-sc VVSELIGNGVFFNVFSONNGGSDRAATTEPVFPFPFPFPFPFFNK ALDFSEAPSTQFLANKSIAGNATLS SV GSPEPT SNFNKGLDLGET hu-sc VVSELIGNGVFFNVFSONNGGSDRAATTEPVFPFPFPFPFPFNK ALDFSEAPSTQFLANKSIAGNATLS SV GSPEPT SNFNKGLDLGET hu-sc Motif 10 *	<pre>HKSTRWMLINFDLIKELTYEA GMNIEGVVEMMITEAVNSIGMSRFSPASOPENPINOSSEPTHETVEDVSDTTVALWRPPEIGAGGLOGIUCYUKGG 911 KKSTRWMLINFDLYKLTYEA GMNIEGVAYEMITEAVNSIGMSRFSOASOPFVPLAVESEPTHETVEDVSDTTVALWRPPEIGAGLOGIUCKYCKGG 858 KMKLINFDLCKTTFEPKMIEGVAEVKIFAVNAIGISKFSWPSOASOPFVPLAVESPTULVISVDTTVTMEWRPPEIGAGLOGVUKKGG 858 KMKLINFDLCKTTFEPKMIEGVAEVKIFAVNAIGISKFSWPSOEVPLAVESPFVLAVISSPTULVISVDTTVTMEWRPPEIGAGLOGVUKKGG 858 KMKLINFDLCKTTFEPKMIEGVAEVKIFAVNAIGISKFSWPSOEVPLAVISSPTULVISVDTTVTMEWRPPEIGAGLOGVUKKGGF 858 KMKLINFDLCKTTFEPKMIEGVAEVKIFAVNAIGISKFSWPSOEVPLAVISSPTULVISVDTTVTMEWRPPEIGAGLOGVUKKGGF 858 KMKLINFDLCKTTFERKUNEGALERVVAINAGRAPUTTEPATLAOPVTIERIASPKIELOTIIKVSEOLILVVPFGGPRPOVUTKGGAPLDTSR 866 PALGGLEERSMLVDLPTGALLERVVAINLAGEGAPUTTEPUVOEILOPPLICPTIKVSEOVIALIPFGGPRPOVUTKGGPLAGE 1014 PALGGLEERSMLVDLPTGALLERVVAINLAGEGAPUTTEPUVOEILOPPLICPTIKVSEOVIALIPFGGPRPOVUTKGGPLAGE 1014 PALGGLEERSMLVDLPTGALLERVVAINAGPGAPUTTEPUVOEILOPPLICPTIKVSEOVIALIPFGGPRPROVUTKGGPLAGE 1014 PALGGLEERSMLVDLPTGALLERVVAINAGPSPAATLAOPVTIEKUPIVOEILOPPLILPPLICUVKVSETTININIPFGGPRPROVUTKEGOPLAGE 1014 PALGGLEERSMLVDLIPTGALERVVAINAGPSPAATLAOPVTIEKUPIVOEILOPPLILPPLIKUVETININIPFGGPRPELTWKNDGALDSKV 1016 PAUGGLDETSILINILTGOLOFTVAINMAGPSPAATLAOPVTIEKUPIVELMEPHILOPPLIRVGVINALIPPGGPRPELTWKNDGALDSKV 1016 PAUGGLDETSILINILTGOLOFTVAINMAGPSPAATLAOPVTIEKUPIVELMEPHILOPPLIRVAINALTIKVSETNIALIPPGGPRPELTWKNDGALDSKV 1016 PAUGGLDETSILINILTGOLOFTVAINMAGPSPAATLAPVTVELKAPTVIENTIENDEPHILOPVIIRVGEAVNLUPFGGPRPELTWKNDGALDSKV 1016 PAUGGLDETSILINILTGOLOFTVAINMAGPSPAATLAPVIERNIHTS 991 NSPTDTILFIAARTVSGTQVTVRIENMEDKATLVLQVVD PSPODLEVTDANGINALEWRPODEONTINGTVOKADKKTMEWFVVERNIHTS 991 NSPTDTILFIAARTVSGTQVTVRIENMEDKATLVLQVD PSPPOLEVTDANGLAVALEWRPODEONGTINGTVOKADKKTMEWFVVERNIHTS 991 NSSTDTILFIAARTVSGTQVTVRIENMEDKATLVLQVD PSPPOLEVTDANGUNALEWRPODEONGTINGTVOKADKKTMEWFVVERNIHTS 991 NSSTDTILFIAARTVSGTQVVVRIENDKATLVLQVD PSPPOLEVTDANGUVALEWRPODEONGTINGTVOKADKKTMEWFVVERNIHTS 991 NSSTDTILFIAARTVSGTQVVRIENDKATLVLGVUENDAVGTVALUVDAUNDVRALEWRPODEONGTINGTVOKADKKTMEWFVVERNIHTS 991 NSSTDTILFIAARTVSGAVUUGUVVVERNENNENTIIDIDIDID</pre>
<pre>hu-fst VHV TSDFDTVFFVRQAA.SDSGEVELSVQIENMKDTATIRIRVD CAGPPINVMVKEVNGTNALVEWQAPKDDGNSEING FVQKADKKTMEWFNVFERI hu-cc VSINSPTDTILFI AARAVISGTKQVTVRIENMEDKATLUQVD RSPPQDLRVTDANGLNVALEWKPPQDVGNTENGGTNEWGTVQKADKKTMEWFTVLEN ch-cc VGINSSTDTILFI AARAVISGTKQVTVRIENMEDKATLUQVD (PSPPQDLRVTDANGLNVALEWKPPQDDGNTEINGGTVQKADKKTMEWFTVLEN ch-cc VGINSSTDTILFI KAELHSGAYEVTLQIENMTDTVAITIQID PGPPQDIRICHTGHVVALEWKPPQDDGNAUGKTINGUTVQKADKKTMEWFTVLEN ch-cc VGINSSTDTILFI KAELHSGAYEVTLQIENMTDTVAITIQID PGPPQDIRIVETGFVNALEWKPPQDDGNAGILGVTVQKADKKTMEWFTVUEN fu-cc ASVINSEGDTILFI KTEKKISKIDLQVKUDKVDKSVTLQVD LGGPPEVLKIVDVNGFNVALEWKPPQDDGNAGILGVTVQKADKKTMEWFTVIDU hu-slw ININSEGDTILFI KTEKKISKIDLQVKUDKVDKFVFTASIDIQID PGPPPUKLIVDVNGFNVALEWKPPQDDGNAGTGTTQTVQKADKKTMEWFTVIDU fu-cc Notif 9 cont. Notif 10 * hu-fst TVSDLVGNEYFFVFSQNNGGSDRAATTEPVFTPPIPGTTEPPNYHALDFSAPSTQPLNNSVIAGYAALCQAVGSPPPINVMKNNMENEIRE hu-cc CVVSELIGNGYFFVFSQNNGGSDRAATTEPVFTPPIPGTTEPPNYHALDFSAPSTQPLNNSVIAGYNATLCQAVGSPPPINFWCDLGGE ch-cc VVSELIGNGYFFVFSQNNGSSDRAATTEPVFTPPIPGTTEPPNYHALDFSAPSTQPLNNSVIAGYNATLCQAVGSPPPIFWKNDLDGE fu-cc TVSDLIMGNEYFFVFSNNGGSDRAATTEPVFTPPIPGTTEPPNYHALDFSAPSTQPLNNSVIAGYNATLCQAVGSPPPIFWKNDLDGE fu-cc TVSDLIMGNEYFFVFSNNGGSDRAATTEPVFTPPIPGTTEPPNYHALDFSAPSTQPLNNSVIAGYNATLCQAVGSPPPIFWKNDLDGE fu-cc TVSDLIMGNEYFFVFSNNGGSDRAATTEPVFTPPIPGTTEPPNYHALDFSAPSTQPLNNSVIAGYNATLSGVGIPPFYFWNDLDGG fu-cc TVSDLIMGNEYFFVFSNNGGSDRAATTEPVFTPPIPGTTEPPNYHALDFSAPSTQPLNNSVIAGYNATLSGVGIPPFYFWNDLDGG fu-cc MSKQGULTEIRKPCFDGGTVCTATNAATDATTELCCAVGENEVFF0- 1142 hu-fst ITTYQGVLTINIRPSPFTAGTTCTANNELGEALAECKLEVFVF0- 1270 mo-cc MFSKQGULTEIRKPCFDGGTVCTATNLQGAQCECRLEVVF0- 1270 mFSKQGULTEIRKPTPDGGTVCTATNLQGACCECRLEVVF0- 1270 ch-cc MFSKQGULTEIRKPTPDGGTVCTATNLGEACECRLEVVF0- 1270 ch-cc MFSKQGULTEIRKPTPDGGTVCTATNLGEACECRLEVVF0- 1270 ch-cc MFSKQGULTEIRKPTPDGGTVCTATNLGEACECRLEVVF0- 1270 ch-cc MFSKQGULTEIRKPTPDGGTVCTATNLGEACECRLEVVF0- 1270 ch-cc MFSKQGULTEIRKPTPDGGTVCTATNLGEACECRLEVVF0- 1270 ch-cc MFSKQGULTEIRKPTPDGGTVCTATNLGEACCECRLEVVF0- 1270 ch-cc MFSKQGULTEIRKPTPDGGTVCTATNLQGEACECRLEVVF0- 1270 ch-cc MF</pre>	IKSYMAMILNEDLEKELTYEARDNIEGVYEMETIAVNSIGASGAPENGENDEVDENDEVELTYEDVEDTVALAWEDPERIGAGLEG IVEXCHOS 911 KKSYMAMILNEDLEKELTYEARDNIEGVAEMENVAVNSIGASGAPENDEVDEVDAPESEPVGLVDISDTSIVLAWEDPERIGAGLEG VVEXCHOS 911 KKSYMAMILNEDLEKELTFERKAMIGVAEVAENSIGASGAPENDEVDEVDAPESEPVGLVDISDTSIVLAWEDPERIGAGLEG VVEXCHOS 911 KGSYMAMILNEDLEKELTFERKAMIGVAEVAENTAANISIGASPENDEVDEVDEVDEVDEVDEVDEVDEVDEVDEVDEVGEVDENDEVVEXWEDPERIGAGLEG VVEXCHOS 916 Motif 7 cont. * Motif 8 * PANTEPVERCGETVNNEPTGALLFYVVGVNIAGREPATLADEVTIREINA PPHILEGTIRVGGOLAUVPEGGEPHOVTNKEGOPLAGE 1014 PALGELEERSMLVTDLPTGALLFYVAINLAGEGAPUTTE VVOEILGEPLIOLPHILEGTIRVGGVENULTPFGCEPHOVTNKEGOPLAGE 1014 PALGELEERSMLVTDLPTGALLFYVAINLAGESGAAIKEPVITUGILGEPLIOLPHILEGTIRVGFVLLIPFGCEPHEVWINGGOLAGEE 1014 PALGELTERSALIDUVTOD UFFVAINLAGESGAAIKEPVITUGILGEPLIOPHILEGTIRVGFVLLIPFGCEPHEXWKNGGOLAGEE 1014 PALGELTERSALIDUVTOD UFFVAINLAGESGAAIKEPVITUGILGEPLICHPLLGTVIKVGFTINNTPFGCEPHEXWKNGGELSSEF 963 VANKDLIDETKETIGLEDTAA IFVVAINAAGSEPATLADEVITELIGOPPLICHENGTVORVGEVNILLIPFGCEPHEXWKNGGELSSEF 963 VANKDLIDETKETIGLEDTAAIFVVAINAAGSEPHYSOPILVKEITERPONTIKELGOPVILLIPPLEVTINKUGAPLINET 1019 NSPTDTILFIAARSTISGTVOTVRIENMEDKATIRIKVE AGEPINVMKEVAGTNALVENGAPKILAPPCOXPREDVINENKEGAEIDNO 1019 1019 1011 101 101 101 101
ch-cc skikk ksynmelnfdllkeltyeakrmiegvyremeitavnsigmsr <mark>ps</mark> pasopfmpia <mark>p</mark> septhftvedvsdttvalkm <mark>pper</mark> igagelocyivey	KKYRWMALNFDLLRELSHEARAMIEGVAYEMRVYAVNAVGMSRPSPASOPFMPIGPPGEPTHLAVEDVSDTTVSLKWRPPERVGAGGLBGS SVEVCCEGC 909
fu-cc ERKKKKSYRWMELNFOPYPETTYEAKRMIEGVAVEMRVYAVNSIGMSRHSQASQPFVPVAPTSEPVGLCVDDISDTSIVLKWRPPERMGSVDLEGYGVEV	KSYRWMELNFDLLKELTYEAKRMIEGVVYEMEITAVNSIGMSR <mark>PSPASOPFMPIAPPSEPTH</mark> FTVEDVSDTTVALKWEPPERIGAGGLDGTIVEYCKDGS 911
hu-siw ErkkkOSSRWMLNFDLCKETTFEPKKMIEGVATEVEIFAVNAIGISKPSMPSRPFVPLAVISPTLLIVDSVTDTTVTMKWRPPDHIGAAGLDGVLEV	KKSTRMMELNFDLLKELTYEAKEMIEGVVYEMEITAVNSIGMSR <mark>PSPASOPFMPIAPPSEPTHFT</mark> VEDVSDTTVALKWRPPERIGAGGLDGVIVYCKDGS 911 KKSYRMMELNFDPYPETTYEAKEMIEGVAYEMEVYAVNSIGMSR <mark>H</mark> SQASOPFVPVAPTSEPVGLCVDDISDTSIVLKWPPPERMGSVDLEGVGVYCKFGT 858
Motif 7 cont. * Motif 8 *	IKKSTRWMRLNFDLLKELTYEAKEMIEGVVYEMRIYAVNSIGMSR <mark>PSPASQPFMPIAPPSEPTHFT</mark> VEDVSDTTVALKWRPPERIGAGGLDGVIVYCKDGS 911 IKKSYRWMRLNFDPYPETTYEAKEMIEGVAYEMEVYAVNSIGMSR <mark>H</mark> SQASQPFVPVAPTSEPVGLCVDDISDTSIVLKWEPPERMGSVDLEG GVEYCKEGT 858 IKQSSRWMRLNFDLCKETTFEPKKMIEGVAYEVFIFAVNAIGISKPSMPSRPFVPLAVTSPPTLLTVDSVTDTTVMRWEPPDHIGAAGLDGVLEYCFEGT 761
hu-ist EEWVPANTEPVERCGFTVKNLFTGARILFRVVGVNIAGRTEPATLAGPVTIREIAEPPKIKLPRHLRGTVIREVGEQLNLVVPFQGAPRP0VVNTKGGAPI	KKSTRWMRLNFDLLKELTYEAKEMIEGVVYEMRIAAVNSIGMSRPSPASOPFMPIAPPSEPTHFTVEDVSDTTVALKWRPPERIGAGGLDG IVEYGKDGS 911 KKSTRWMRLNFDPYPETTYEAKEMIEGVAYEMEVYAVNSIGMSRHSQASOPFVPVAPTSEPVGLCVDDISDTSIVLKWRPPERMGSVDLEGVGVYEKFGT 858 KOSSRWMRLNFDLCKETTFEPKKMIEGVAYEVFIFAVNAIGISKPSMPSRPFVPLAVTSPPTLLTVDSVTDTTVTMWRPPPHIGAAGLDGVVLEYGFEGT 761 Motif 7 cont. * Motif 8 *
hu-cc SEWVAALQGLTEHTSILVKDLPTGARLLFEVRAHNMAGPGAPVTTTEPVTVQEILQEPRLQLPRHLRQTIQKKVGEPVNLLIPFQGKPRPQVTWTKEGQPI	KKSERWARLAFDILKELTTEAKUMIEGVVTEMEITAVNSIGMSRPSPASOPFMPIADPSEPTHFTVEDVSDTTVALKAPPERIGAGGLOGIVEYGKOGS 911 KKSERWARLAFDPYPETTEAKAMIEGVAFEN VVAVNSIGMSRBSORSOPFVPAPTSEPVGLOVDISDTSIVLKAPPERMGSVDLEGGVEYGKGG 858 KOSSHWARLAFDICKETFFERKMIEGVAFEN FAVNAIGISKPSMPSRPFVPLAPTSPTLLTVDSVDTTVTMRAPPHIGAAGLOGVEYGFGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGGTVKNLPTGARILFTVVGVNIAGREPATLAOPVTIREIAEPFKIRLPRHLROTYIRKVGEQLNLVVPFGGPPPVVVTRGGAPLDTSR 886
mo-cc. SEWTPALOGLTERSMLVRDLPTGARLLFRVRAHNVAGPGGPIVTKEPVTVOEILORPRLOLPRHLROTIOKKVGEPVNLLIPFOGKPRPOVTWTKEGOPI	IKSYRWMELNFOLLKELTTEAKUMIEGVVEMEIIAVNSIGMSR <mark>959ASOPPMPIAPSSPTHFIV</mark> EUVSDTTVALKME DPERIGAGGLOGIVESTK OGS 911 CKSYRWMELNFDPYPETTEAKEMIEGVAIEMEVEAVNSIGMSRHSOASOPFVPAPTSSPVGLCVDDISDTSIVLKWEPPERMGSVDLEGGVESTGKEG 858 KWMELNFDLCKSTTEPKKMIEGVAFEVFIFVNAIGISKPSMPSRPFVPLAVTSPFTLTVDSVDTTVIKWEPPERMGSVDLEGGVESTGKEG 858 KWMELNFDLCKSTTEPKKMIEGVAFEVFIFVNAIGISKPSMPSRPFVPLAVTSPFTLTVDSVDTTVIKWEPPERMGSVDLEGGVESTGKEG 90 Motif 7 cont. * Motif 8 * Motif 8 * MOTIF 8 PANTEPVERGETVKILPTGALLFFVVGVNIAGRTEPATLAOPVTIEFASPFTLTVDKUMENTGGALLDVPFGGKPEPOVWTKGGALLDTSS 886 AALOGIEFHTSILVKOLPTGALLFFVVGNIAGRTEPATLAOPVTIEFIASPFTLTOTOKIKGEPVNLTGGGFDFUTTEGOPLAGES 1018
mo-cc SEWTPALGELTERSMLVADLPTGARLERVAARNVAGPGGPTVTKEPVTVEILCEPRLCLPRALAGTICKAVGEPVNLLTPFGRPRPQVTWTKEGQPT	KKSYRWARLAFDLLKELTYEAKUMI GUVYEMI I AVNSIGMSRPSPASOPFMPIADPSEPTHFTVEDVSDTTVALKW PPERIGAGGLDG IVEYGKDGS 911 KKSYRWARLAFDPYPETTEAKAMI GUAYEMI VIAVNSIGMSRHSOASOPFVPIAPTSEPVGLCVDI SDTSIVLKW PPERMGSVDLGG GVEYGKGT 858 KOSSHWARLAFDLCKETFEPKMI GUAYEVI FAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVDTVTMRW PPHIGAAGLDG VLEYGFGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGGTVKNLPTGARILFRVVGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLRQTI KNUGEQLNLVVPFQG PAPQVVWTKGGAPLDTSR 886 AALOGITEHTSILVKOLPTGARILFFVRAHNAGGGAPUTTT PUTVOEILOFPKIQLPHHLQTICKTVGEPVNLIFFQGFPAPGVTWTKGGAPLDTSR 886
CH-CC AFETTAL DOLTER TSALTED UTCOMENT FOUND THAT A CESCA T THE DUTUGE THAT DE TOUDE THAT UNTUGET TAT THE DE TOUDE THAT THE DE TOUDE TOUDE THAT THE DE TOUDE TOUDE TOUDE THAT THE DE TOUDE THAT THE DE TOUDE TOUDE THAT THE DE TOUDE	IKS FRUMELNEDLIKELTTEAKUMIEGVAVEMEILAVNSIGMSR <mark>PSPASOPYMPIAPSEPTHETVE</mark> DVSDTTVALKWEPPERIGAGGLOGIVEYGKOG KKSFRUMELNEDPYPETTYEAKUMIEGVAVEWEVIAVNSIGMSRHSQASOPFVPVAPTSEPVGLCVDDISDTSIVLKWEPPERMGSVDLEGGVEYGKEGT 858 KOSSFMMELNEDLIKETTEPKMIEGVAVEVIFAVNAIGISKPSMPBRPFVPLAVTSPPTLLTVDSVDTTVIKWEPPHIGAGGLOGIVEYGKEGT Motif 7 cont. * Motif 8 * "PANTERVENGGETVNLPTGARLIFFVGAINNAGPGAPVTTEPVTVOEILGPPLICEPHILGYTIKEVGGLALVVPFGGFPRDQVVWTGGAPLBTSR 886 AALQGITEHTSILVKOLPTGARLIFFVGAINNAGPGAPVTTEPVTVOEILGPPLICEPHILGYTIKEVGSUNLVVPFGGFPRDQVVWTGGAPLBTSR 886 AALQGITEHTSILVKOLPTGARLIFFVGAINNAGPGGPTVTEPVTVOEILGPPLICEPHILGYTIKEVGFVNLLIFFGGFPAPGVVWTKEGGPLAGEE 1018
ch-cc AEWTPALPGLTERTSALIKDLVTGDKLYFRVKAINLAGESGAAIIKEPVTVOEIMORPKICVPRHLROTLVKKVGETINIMIPFOGRPRPKISWMKDGOTI	IKSYRWMRLNFDLIKELTTEAKUMIEGVVEMENIAVNSIGMSR959ASQPFMPIAPS55PTHFTVEUVSDTTVALKWEPPERIGAGGLOGIVEYGKOG SKSYRWMRLNFDPYPETTEAKUMIEGVATEMEVYAVNSIGMSRHSQASQPFVPAAPS5PVGLEVDISDTSIVLKWEPPERMGSVDLEG GVEYGKEGT SSSRWMELNFDLCKTTFEPKMIEGVAYEVRIFAVNAIGISKPSMPSRPFVPLAVTS9PVGLEVDISDTSIVLKWEPPERMGSVDLEG GVEYGKEGT Motif 7 cont. * Motif 8 * PANTEEVEIGGTTVNLPTGARLLFVVGANNAGRTEPATLAOPVTIREIASPPTIREIASPPTIREVGEJALVVPFGGKPRPQVVWTKGGAPLDTSR SALGGIFHTSILVKDLPTGARLLFVVGANNAGPGGAPVTTEPVTVQELGPRLQLPHLRQTIGKVGEPVNLLEPGGKPRPQVWTKGGAPLDTSR PALGGIFHTSILVKDLPTGARLLFVGANNAGPGGAPVTTEPVTVQELGPRLQLPHLRQTIGKVGEPVNLLEPGGKPRPQVWTKGGAPLDTSR 1014
Ch-CC REWTPALPGLTERTSALIRDLVTGDALIFRVRAINLAGESGAAIIKEPVTVQEIMQRPAICVPRHLRQTLVRAVGETINIMIPFQGAPRPKISWMRDGQTI	IKS TRAMELNFOLLKELTTEA KUMI GOVYEME I KAVNSIGMSR 959ASOPFMPIAPSSEPTHFTVEOVSDTTVAL AWEDPERIGAGGLOGI VESTKOGS SKSTRAMELNFOPYETTEAKEMI EGVATEMEVEAVNSIGMSR 30ASOPFVPAAPSSEPTHEVEOVSDTTVAL AWEDPERIGSVDLEG GVESTKOGS SSRAMELNFOLCKSTTEPKAMI EGVATEMEVEAVNSIGMSR 30ASOPFVPAAPSSEVGLCVDI ISDTSIVLKWEDPERIGSVDLEG GVESTKEG SSRAMELNFOLCKSTTEPKAMI EGVATEMEVEAVNSIGMSR 30ASOPFVPAAPSSEVGLCVDI ISDTSIVLKWEDPERIGSVDLEG GVESTKEG Motif 7 cont. * Motif 8 * PANTEPVEIGGTTVKLEPTGAR LEFVUSVNIAGR TEPATLAOPVTI REIASPPKIRLEGTHI RVGSLALVVPFGG FRADEVWIT GGAPLDTSR 886 AALOGITEHTSILVKOLPTGAR LEFVUSVNIAGR TEPATLAOPVTI REIASPPKIRLEGTNI RVGSLALVVPFGG FRADEVWIT GGAPLDTSR 886 PALOGITEHTSILVKOLPTGAR LEFVUSANNAG PGGAPUTTEPTVVDEI LOPPKILOPTI RVGSLAVVPFGG FRADEVWIT GGAPLAGES 1018
	IKS REMELNEDLIKELTYEAKEMI GVVYEMI I AVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKE PERIGAGGLOG IVEYGKOGS 911 KKSYRMELNEDDIYETTEAKEMI GVAYEMI I AVNSIGMSRBSOPEVPVAPTSEPVGLOVDI SDTSIVLKE PERMGSVDLEG GVEYGKGT 858 KOSSEMBLNEDLIKETTEEPKMI GVAYEVI FAVNAIGI SKPSMPSRPFVPLAVTSPTLLTVDSVTDTVIMER PPDHIGAAGLOG VLEYGFGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGETVKLPTGARLLEFVVGVNIAGRTEPATLAOPVTI BEI ASPPKIRLPRHLROTYI KVGEQLALVVPFGGPRAPQVVWTKGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARLLEFVVAHNAGPGGPIVTKE PVTVQEI LOFPRLOLPHLROTI KKVGEPVNLLIPFGGPRAPQVWTKGGPLAGEE 1018 PALQGLTERSMLVKDLPTGARLLEFVVAHNAGPGGPIVTKE PVTVQEI LOFPRLOLPHLROTI KKVGEPVNLLIPFGGPRAPQVWTKGGPLAGEE 1014 PALQGLTERSMLVKDLPTGARLLEFVVAHNAGFGGPIVTKE PVTVQEI LOFPRLOLPHLROTI KKVGEPVNLLIPFGGPRAPQVTWTKEGGPLAGEE 1014
FUER DAUGOT POPTOT T TRUE TECHT OF DUE AVANACIDE DA MET A COMPTENT A DAVA TO THE TANK THE TAN	IKKS RWMRLNFDLKELTTEAKUMIEGVVTEMEILAVNSIGMSRPSPASOPFMPLADPSEPTHFTVEDVSDTTVALKW PPERIGAGGLOGIVEYGKOG IKKS RWMRLNFDPYPETTYEAKUMIEGVATEWR VIAVNSIGMSRBSOPFVPAPTSEPVGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKEGT 858 INGSSRWMELNFDLCKETFFERKMIEGVATEVPIFAVNAIGISKPSMPSRPFVPLAVTSEPVGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKEGT Motif 7 cont. * Motif 8 * PANTEPVERGEFTVENLPTGARILFFVVGVNIAGRTEPATLAOPVTIREIAEPFNIRLPRHLNQTIKVGEQLNLVVPFQKPRPQVVWTGGAPLDTSN 74ALGGITEHTSILVKDLPTGARILFFVVGNNAGPGAPVTTEPVTVQEILOPPLQIPHLNQTICKVGEPVNLLIPFQGKPRPQVWTKEGGPLAGEE 1018 PALGGITEHTSILVKDLPTGARLLFFVRAHNNAGPGAPVTTEPVTVQEILOPPLQIPHLNQTICKVGEPVNLLIPFQGKPRQVWTKEGGPLAGEE 1018 PALGGITERTSALJKDLVTDLFFVRAHNAGPGAPVTKEPVTVQEILOPPLQUPHLNQTICKVGEPVNLLIPFQGKPRQVWTKEGGPLAGEE 1014 PALGGITERTSALJKDLVTDLFFVRAHNAGPGAPITKEPVTVQEILOPPLQUPHLNQTICKVGETNININIPPGKPRSVMKEGGPLAGEE 1014
funce FOWL PAVOGLED BEST LIENT TECHLOF VERY NMAGPS PAATLAOPVELET MORPETLI, PENTROT, I REVED TVNLAIPFOCE PROKUTWSKNGED	IKS RWMRLNFDLKELTTEAKUMI EGVYEME ILAVN SIGMSRPSPASOPYMPLADPSEPTHFTVEDVSDTTVALKW PPERIGAGGLOG IVYYGKOG 911 KKS RWMRLNFDPYPTTYEAKUMI EGVAFEW VYAVNSIGMSRHSQASOPYVPAPTSEPGLCVDI SDTSIVLKW PPERMGSVDLEG GVYGKEGT 858 KQSSRWMELNFDLCKETFFERKMI EGVAFEW IFAVNAIGI KKPSMPSRPFVPLAVTSPTLLTVDSVDTTVTMRW PPERMGSVDLEG GVYGKEGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGEFTVKNLPTGARILF VVGVNIAGREFEPATLAOPVTIREIAEPKIRLPHILNOTIINKVGEQLNLVVPFGGRPRPQVVWTGGAPLDTSR 886 AALQGITENTSILVKDLPTGARLFFVRAHNMAGGAPVTTEPTVOELGBPRINGTIGKVGEPVNLLIPFGGRPRPQVWTTEGGPLAGEE 1018 PALGETERSSLVKDLPTGARLFFVRAHNMAGGGAPVTTEPTVOELGPRINGTIGKVGEPVNLLIPFGGRPRPQTWTEGGPLAGEE 1018 PALGETERSSLVKDLPTGARLFFVRAHNMAGGGAPVTTEPTVOELGPRINGTIGKVGEPVNLLIPFGGRPRPQTWTEGGPLAGEE 1018 PALGETERSSLVKDLPTGARLFFVRAHNMAGGGAPVTTEPTVOELGPRINGTIGKVGEPVNLIPFGGRPRPVTWTEGGPLAGEE 1018
TU-CC PERMENY CONTRACTOR AND TO THE THE TO THE THE TO THE THE TO THE	IKS FRAMELNFOLLKELTTERAKUMI GOVYEMI ILAVNSIGMSR ^D SPASOPTMPIAPSEPTHFTVEOVSDTTVALKW PPERIGAGGLOGIVEYGKOG 911 KKS FRAMELNFOPPETTYEAKUMI GOVYEMI VAVNSIGMSR ^H SQASOPFVPVAPTSEVGLOVDI ISDTSIVLKW PPERMGSVDLEG GOVYEKGG 858 KOSSRMELNFOLCKSTTFEPKMI GOVYEVI FAVNAIGISKPSMPERPFVPLAVTSPPTLIVDSVDTTVIKW PPERMGSVDLEG GOVYEKGF 761 Motif 7 cont. * Motif 8 * PANTERVEI GEFTVNLPTGARILFFVGVINAGRTEPATLAOPVTIREIAPPKI FLIPHILAOTIVKVGEVALLVVPGG FRAQVUVTGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARILFFVGAINMAGPGAPVTTEFVTVQEI LOPHILQLPRHINQTICKVGEVNLLIPFGG FRAQVUVTGGAPLDTSR 816 PALOGLTEHTSILVKDLPTGARILFFVGAINMAGPGAPVTTEPVTVQEI LOPHILQUPHLIAOTICKVGEVNLLIPFGG FRAQVUVTGGAPLAGE 1018 PALOGLTEHTSILVKDLPTGARILFFVGAINLAGFGGFIVTKEPVTVQEI LOPHILQUPHLIAOTICKVGEVNLLIPFGGFPAPKTSKMEDGGTLASE 1014 PALOGLTEHTSILVKDLYTGANGEVGAPVINEGGFIVTKEPVTVQEI LOPHILQUPHLIAOTICKVGEVNLLIPFGGFPAPKTSKMEDGGTLASE 1014
A A A A A A A A A A A A A A A A A A A	IKSY RWARLAFDLIKELTTEAKUMI GOVYEME ILAVASIGASRS PASOPYMPIAPSE PTHFTVEOVSDTTVAL AV PPERIGAGGLOGI VEYGADGS SAMARLAFDPYPETTEAKUMI GOVYEMEVANNSIGASR SASOPYVPAPTSE VGLOVDI ISDTSIVLKWAPPERAGSVDLEG GOVYGKGG SAMARLAFDLCKSTTEPKUMI GOVYEWEVANNSIGASR SASOPYVPAPTSE VGLOVDI ISDTSIVLKWAPPERAGSVDLEG GOVYGKGG Motif 7 cont. * Motif 8 * PANTERVEIGGTTVNLPTGARLEFVVGNIAGRTEPATLAOPVTIREIASPPILTVDSVDTVTSOFAPVNLLTPGGCFRPQVWTKGGAPLDTSR SALGGITERSILVGLPTGARLEFVVGANNAGGGAPVTTEPTVOGI LOPPLICEPTICKVGGVNLLVPFGGCFRPQVWTKGGAPLDTSR PALGGITERSLVKULPTGARLEFVVGANNAGFGAPVTTEPTVOGI LOPPLICEPTICKVGGVNLLIPFGGCFPRQVWTKGGAPLDTSR PALGGITERSLVKULPTGARLEFVVGANNAGFGGAPVTTEPTVOGI LOPPLICEPTICKVGGVNLLIPFGGCFPRQVWTKGGAPLDTSR PALGGITERSLVKULPTGARLEFVVGANNAGFGGPIVTKEPVVQEI LOPPLICEPTICKVGTVKVGFVNLLIPFGGCPRPQVWTKEGGPLAGE 1014 PALGGITETSLIKULVTDLEFVVGANNAGFGSGAAI IKEPTVOGI LOPPLICEPTLEVGTICKVGCTININTFGGCPRPKISSMEDGGTLSKD 1016 PALGGITETSLIKULVTDLEVGVVGANNAGFGSPAATINGPTTEPTVOGI LOPPLICEPTLEVGTINKVGCTININTFGGCPRFST SAGGITERSLIKULVTDLEFGCKFSALLFSKANGGTLSKD SAGGITESSLIKULFTCLEFFKSALLFSKANGFFANGEFSANGFFSF SAGGITERSLIKULFGCKFSKANGFSANGFSANGFSANGFSANGFSANGFFS SAGGITERSLIKULFGCKFSKANGFSSANGFSANGFSANGFSANGFSSANGFSF SAGGITESSLIKULFGCKFSKANGFSSAN
	IKSY RWMRLNFDLIKELTTEAKUMI GOVYEME ILVNSIGMSR 959A30PFWP1APSSPTHFTVEUVSDTTVLKWE PPERIGAGGLOGI VEYGKOG SAMARLNFDPYETTEAKUMI GOVYEMEVAVNSIGMSR 90A30PFVPAPSSPTHFTVEUVSDTTVLKWE PPERIGSVDLEG GOVYGKGG SAMARLNFDLCKTTFEPKMI GOVYEWEVAVNSIGMSR 90A30PFVPAPSSPVGLOVDI ISDTSIVLKWE PPERIGSVDLEG GOVYGKGG SAMARLNFDLCKTTFEPKMI GOVYEWEVAVNSIGMSR 90A30PFVPAPSSPVGLOVDI SDTSIVLKWE PPERIGSVDLEG GOVYGKGG Motif 7 cont. * Motif 8 * PANTERVENGGTTVNLPTGARLFFVVGVNIAGR TEPATLAOPVTI BEIASPPI ILVOUSTVKUGVALVOUSVDLGVEVGKGF SALGGITEHTSILVGVGVFVGVNIAGR TEPATLAOPVTI BEIASPPI ILVOUSTVKVGVNILVOUSVDUVTKGGAPLDTSR 886 SALGGITEHTSILVGLPTGARLFFVVGVNIAGR TEPATLAOPVTI BEIASPPI ILVOUSTI CKVGCIMUVVPGGK PPOVVNTKGGAPLDTSR 886 SALGGITEHTSILVGLPTGARLFFVVGVNIAGR TEPATLAOPVTI BEIASPPI ILVOUSTI CKVGCIMUVVPGGK PPOVVNTKGGAPLDTSR 94LOGITEHTSILVGUVGVAVAVAGGGAPVTTTPUTVOEI LOPPI LOPPI LOPVI LIPOGK PPOVVNTKGGAPLDTSR 94LOGITEHTSILVKDLPTGARLFFVVGI NAGR 950A3I IK PVTVOEI LOPPI LOPPI LOVIT KVGTI NIM PFGK PPRKISSME 00GTLDSK0 1016 PAUGGITENTSILI KULVTDUKUVRAVAMGESGAATI KEPTVOEI MGP HILVOHILMVGTI KKVGT NIMI PFGK PRKISSME 00GTLDSK0 1016 PAUGGITENTSILI KULVTDUKUVRAVAMGESGAPVTTEGVAPTLAOPVTI KOOP HILVOHILMVGT KVGT NIMI PFGK PRKISSME 00GTLDSK0 1016 PAUGGITENTSILI KULVTDUKUVRAVAMGESGAATI KEPTVOEI MGP HILVOHILMVGT KVGT NIMI PFGK PRKISSME 00GTLDSK0 1016 PAUGGITENTSILI KULVTDUKUVRAVAMGESGAPTUTENT MGP PTI BEIASPP TI LIPONI TE KVGT TIMI PFGK FRKISSSF 953 1016 PAUGGITENTSILI KULVTDUKUVRAVAMGESSAATI KEPTVOEI MGP PTI BEIASPT 953 1016 PAUGGITENTSILI KULVTDUKUVRAVAMGESSAATI KEPTVOEI MGP PUTIENT MGP FOKTAVIATI FOKTAVIATI FOKTAN FFST 1016 PAUGGITENTSILI KULVTDUKUVRAVAMGESSAATI KEPTVOEI MGP PUTIENT MGP FOKTAVIATI FFOKTAN FFST 1016 10
	IKSTRWMLNFDELKELTTEAKUMIEGVVTEMIIAVNSIGMSRPSPASOPFMPIADPSEPTHFTVEDVSDTTVALKWPPERIGAGGLOGIVEYGKOGS 911 KKSTRWMELNFDELKETTEEKKMIEGVATEMIYAVNSIGMSRHSOASOPFVPAPTSEPVGLCVDISDTSILLKWPPERMGSVDLEGGVEYGKGT 858 QSSRWMELNFDELCKETTEEPKMIEGVATEVIFAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVTDTVIMKWPPPHIGAAGLOGVLEYGFGT 761 Motif 7 cont. * Motif 8 * PANTEEVERGETVKLIPTGAKLLFFVVGVNIAGRTEPATLAOPVTIREIASPPKIRLPRHLRQTIIKVGEQLALVVPFQGPERPQVVWTKGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGAKLLFFVVGANNAGPGAPIVTKEPVTVQEILOPKLQLPHLRQTIQKKVGEPVNLLIPFQGFPRQVTWTKEGQPLAGEE 1018 PALQGLTERTSALIKDLVKGLPTGAKLLFFVKANNAGPGGPIVTKEPVTVQEILOPKLQLPHLRQTIQKKVGEPVNLLIPFQGFPRQVTWTKEGQPLAGEE 1014 PALPGLTERTSALIKDLVKGLFFVKANNAGPGAPIVTKEPVTVQEILOPKLQUPHLRQTIQKKVGEPVNLLIPFQGFPRQVTWTKEGQTLAGEE 1014
ANALY ANY ANY ANY ANY ANY ANY ANY ANY ANY AN	IKSYRWALAFDILKELTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPFMPIADPSEPTHFTVEDVSDTTVALKAPPERIGAGGLOGIVEYGKOG IKKSYRWALAFDILKELTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPFMPIADPSEPUGLCVDI SDTSILAVA PPERMGSVDLEG GVEYGKGT 858 QSSWMELAFDILKETFEPKMI GUVYEVI FAVAAIGISKPSMPSRPFVLAVTSPTLLTVDSVDTVTMAVAPPHIGAGGLOGVLEYGFG Motif 7 cont. * Motif 8 * PANTEPVERGETVINLPTGARLLFRVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLNQTI INVGEOLALVPPGG PRPQVWIKGGAPLDTSR 886 TALGGITEHTSILVKDLPTGARLLFRVAGNAGGGPUTTEPUTVQEILGPRLOLPHLRQTICKVGEVNLLIPFGGPRAPQVWIKGGAPLDTSR 1018 PALOGITEHTSILVKDLPTGARLLFRVARNVAGFGGPLVIKEPVTVQEILGPRLOLPHLRQTICKVGEVNLLIPFGGPRAPQVWIKGGPLAGEE 1018 PALOGITEHTSILVKDLPTGARLLFRVARNVAGFGSTVIKEPVTVQEILGPRLOLPHLRQTICKVGEVNLLIPFGGPRAPQVWIKEGOPLAGEE 1014 PALOGITEHTSILVKDLYTGARLLFRVARNVAGFGSAAIKPVTVEPVTVQEILGPRLOLPHLRQTICKVGEVNLLIPFGGPRAPQVWIKEGOPLAGEE 1014 PALOGITEHTSILIKNDLPTGARLLFRVARNVAGFGSAAIKPVTVEPVTVQEILGPRLOLPHLRQTICKVGEVNLLIPFGGPRAPQVWIKEGOPLAGEE 1014 PALOGITEHTSILIKDLTGDKLOFVARNVAGFGSAAIKPVTVEFVTVQEILGPRLOLPHLRQTICKVGEVNLLIPFGGPRAPQVWIKEGOPLAGEE 1014 PALOGITEHTSILIKALFVARNAGFGSAAIKPVTVEFVTVQEILGPRLAQTICKVGEVNLLIPFGGFRAPVENKEGOPLAGEE 1014 PALOGITEHTSILIKVKKANKAGFGFRAVINGGSGAAIKFFYTVEFVTVQEILGPRLAQTICKVGEVNLLIPFGGFRAPVENKEGOPLAGEE 1014 PALOGITEHTSILIKVKKANKAGFGFRAVINGGFFFFYTVEFFFYTVEFFFYTVEFFFFFFFFFFFFFFFFFF
husely EDUTUANED TO FILE TITGE DE ANTE THE ACARE AND A CARE AND A	IKSYRWALAFDILKELTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPFMPIADPSEPTHFTVEDVSDTTVALKWPPERIGAGGLOG IVEYGKOG IKKSYRWALAFDILKELTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPFMPIADPSEPTGLCVDI SDTSILKW PPERMGSVDLCG GVEYGKGT 858 VSSRWBLAFDICKETTFEPKMI GUVYEVI FAVANIG SKPSMPSRPFVPLAYTSPTLLTVDSVDTVDKWRPPHIGAAGLOG VLEYGFG Motif 7 cont. * Motif 8 * PANTEPVER GETVKALPTGARLLFFVVGVNIAGRTEPATLAOPVTI REIAPPKIRLPRHLRQTI IKVGEQIALVVPFQG PRPQVWIKGGAPLDTSR 886 AALGGLTEHTSILVKDLPTGARLLFFVVGVNIAGRTEPATLAOPVTI REIAPPKIRLPRHLRQTI KKVGEQVNLLIFFQG PRPQVWIKGGAPLDTSR 886 PALGGLTERSMLVKOLPTGARLLFFVRAHNAGFGGPIVTKE PVTVQEI LOPPLOLPHLRQTI KKVGEPVNLLIFFQG PRPQVWIKGGAPLAGE 1018 PALGGLTERSMLVKOLPTGARLLFFVRAHNAGFGGPIVTKE PVTVQEI LOPPLOLPHLRQTI KKVGEPVNLLIFFQG PRPQVWIKGGAPLAGE 1014 PALFGLTETSALIKOUVTOKLIFFVRAHNAGFGGPIVTKE PVTVQEI LOPPLOLPHLRQTI KKVGEVNLLIFFGG PRPQVWIKGGDISSKD 1016 PALVGLTDRTSALIFUVTOKINT KVANINAGFSGAAT IKPUTVEN KOGPVNI LIPPGK PRPVTVKEGOFLAGE 1016
AND ADDRESS OF A DESCRIPTION OF A DESCRIPT	IKSY RWMELNFDLIKELTTEAKUMI GOVYEMEI IAVNSIGMSR 959A30PFWPIAPPSSPTHFTVEUVSDTTVAL WEDPERIGAGGLOGI VEYGKOG KKSYRWMELNFDPYETTEAKUMI GOVYEMEVIAVNSIGMSR 30A30PFVPVAPTSEPVGLOVDI ISDTSIVLKWEDPERMGSVDLEG GOVYGKGG KSYRWMELNFDLCKTTEPKKMI GOVYEWI FAVNAIGI KFSMPSR PFVPLAVTSPFTLIVDSVDTTVIK WEDPERMGSVDLEG GOVYGKGG S88 WWELNFDLCKTTEPKKMI GOVYEWI FAVNAIGI KFSMPSR PFVPLAVTSPTLIVDSVDTTVIK WEDPERMGSVDLEG GOVYEYGKGG Motif 7 cont. * Motif 8 * PANTEPVERGSTVKLPTCAKILF VVGAINNAGPGGAPVTTEPTVOEI LOPPKILROTTI KVGGLALVVPFGG FRAPOVWTKGGAPLDTSR 886 AALGGLEHRSILVKDLPTCAKILF VVGAINNAGPGGAPVTTEPTVOEI LOPPKILPTI KVGGLALVVPFGG FRAPOVWTKGGAPLDTSR 816 PALGGLEHRSILVKDLPTGAKILF VRAINNAGPGGPVVKE PVVQEI LOPPKILPTI KVGGLALVVPFGG FRAPOVWTKEGOPLAGE 1018 PALGGLEHRSILVKDLPTGAKILF VVGAINNAGPGGPVVKE PVVQEI LOPPKILPTI KVGGVKI VGEVNILLIPFOG FRAPOVWTKEGOPLAGE 1014 PALGGLEHRSLIVKDLPTGAKILF VVGAINNAGPGSPATIAPVVQEI LOPPKILPTVEI KVGFTININTPFGG FRAPISTSI PAVGGTDSTII INITTOKLOFVGAINNAGPSSAATI KEPTVOEI KOPPHILVPHILKQTI KVGKTININTPFGG FRAPISSSTF 963
bu-cly Charty MUDI TO WART TO I DTDA TO UNA CASE DE VISCANT UNTER DATA TO THE REAL OF A TO	IKSY RWARLAFDLIKELTTEAKUMI GOVYEME ILAVASIGASRS PASOPINPI APSEPTIFTVEOVSDTTVAL AVEPPERIGAGGLOGI VEYGADGS SAMARLAFDPYPETTYEAKUMI GOVYEMEVANVSIGASR SASOPIVPVAPIS PVGLOVDI ISDTSIVLKWE PPERIGSVDLEG GOVYGKGG SAMARLAFDDICKSTTEPKAMI GOVYEWEVAVANSIGASR SASOPIVPVAPIS PVGLOVDI ISDTSIVLKWE PPERIGSVDLEG GOVYGKGG Motif 7 cont. * Motif 8 * PANTERVEIGGTTVNLPTGARLEFVVGNIAGR TEPATLAOPVTI REIASPPILLTVDSVDTVTVTMEN PPERIGSVDLEG GOVYGKGAPLDTSR 886 AALGGITEHTSILV NLPTGARLEFVVGNIAGR TEPATLAOPVTI REIASPPILLOVINI RIVGEONILLTVPGG FRADVWTKGGAPLDTSR 886 PALOGITEHTSILV NLPTGARLEFVVGANNAGG GAPVTTEPTVVGEL LOPPLICEPTICKVGEDVNLLIPFGG FRADVWTKGGAPLDTSR 886 PALOGITERSSLVNDLPTGARLEFVVGANNAGFGGAPVTTEPTVVGEL LOPPLICEPTICKVGEDVNLLIPFGG FRADVWTKGGAPLDTSR 1014 PALOGITERSSLVNDLPTGARLEFVVGANNAGFGGPIVTKE PVTVQEL LOPPLICEPTICKVGCTVNLAFDG FRADVWTKEGOPLAGE 1014 PALOGITETSLIKDLVTDLIFFVGANNAGFGSGAAI INE PVTVQEL LOPPLICEPTICKVGTVNLAFDGGFLAGFAPLTSR 1016 PALOGITETSLIKDLVTCDLIFFVGANNAGFGSAPATING STADIAOPVTER FACTOR KVGCTVNLAFDCGTNAFTGGFLAGE 1016 PALOGITETSLIKDLVTCDLIFFVGANNAGFGSAPATING STADIAOPVTER FACTOR KVGCTVNLAFTGGFLAGE 1016 PALOGITETSLIKDLVTCDLIFFVGANDAFFGFAPTDEDTETSLAGPTTER FYTVORTI LOPPLICTUR VGCTVNLAFFGGFLAGE 1016 PALOGITETSLIKDLVTCDLIFFVGANDAFFGFAPTDEFFATAGPTTER FYTVORTI LOPPLICATION VGCTVNLAFFGGFLAGE 1016 PALOGITETSLIKDLVTCDLIFFVGANDAFFGFATTEPTVORTI LOPPTTER FYTVORTI LOPPLICATION VGCTVNLAFFGGFLAGE 1016 PALOGITETSLIKDLYTCDLIFF
The of the second secon	IKS HUMELNFOLLKELTTERAKUMI GOVTEME ILAVNSIGMSRESPASOPEMPIADESEPTHFTVEDVSDTTVALKMEDPERIGAGGLOGI VEXTKOGS KKS HUMELNFOPYETTYEAKUMI GOVTEME VANNSIGMSRESPASOPEMPIADESEPTHFTVEDVSDTTVALKMEDPERIGAGGLOGI VEXTKOGS SS HUMELNFOLCKETTEPKKMI GOVTEME VANNSIGMSRESPASOPEVPIADESE VGLCVDI ISDTSIVLKWEPPERIGAGLOGI VEXTKOGS Motif 7 cont. * Motif 8 * PANTERVENG GETVNLPTGARILFEVGUNIAGRIEPATLAOPVTIREIAEPPTILTVDSVDTVKEGOLALVVPFGGFREQVVNTGGAPLDTSR SALOGLTEHTSILVKDLPTGARILFEVGUNIAGRIEPATLAOPVTIREIAEPPTILGPHILOTIGKVGEVNLLIPFGGFREQVVNTGGAPLDTSR SALOGLTEHTSILVKDLPTGARILFEVGUNIAGFGGFIVTKEPVTVOEILOFPHILQTICKVGEVNLLIPFGGFREQVVNTEGGPLAGEE 1018 PALOGLTEHTSILVKDLPTGARILFEVGUNIAGFGGFIVTKEPVTVOEILOFPHILQTICKVGEVNLLIPFGGFPERSTANGOGTLDSKD 1019 PALOGLTEHTSILVKOLPTGARLLFEVGUNIAGFGGFIVTKEPVTVOEILOFPHILQTICKVGEVNLLIPFGGFPERSTANGOTLDSKD 1016 PALOGLTEHTSILVKOLPTGARLFEVGUNIAGFGGFIVTKEPVTVOEILOFPHILQTICKVGEVGLINNTFGGPERSTANGEGTLDSKD 1016 PALOGLTEHTSILVKOLPTGARLFEVGUNIAGFGGFLANGEVGEVGEVGUNIAGTINKVGEVGEVGEVGEVGEVGEVGEVGEVGEVGEGTLAGES 1016 1019 PALOGLTEHTSLIKDLVTDDLTFVKANNAGFGGFLANGEVGEVGEVGEVGEVGEVGEVGEVGEVGEVGEVGEVGEVG
	IKS HUMELNFOLLKELTTERAKUMI GOVJEMETIAVNSIGMSR ^B SPASOPFMPIAPSEPTHFTVEOVSDTTVALKMEDPERIGAGGLOGIVEXCKOG KKSTRMMELNFOPYPTTTERKMIEGVATEMEVVAVNSIGMSR ^B SPASOPFVPAPTSEPVGLOVDI ISDTSIVLKWEPPERMGSVDLEG GOVIKKGT 858 (285 RUMELNFOLCKSTTEPKKMIEGVATEVE IFAVNAIGISKPSMPERPFVPLAVTSPPTLIVUSVITTVIKWEPPERMGSVDLEG GOVIKGT Motif 7 cont. * Motif 8 * PANTERVENG GSTVNLPTGARILFEVGNIAGRTEPATLAOPVTIEEIASPPTIFLEPKILOOTIKKVGCIALVVPGGFRAPQVWTKGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARILFEVGAINAGPGGPTVTTEPVTVQEILOPHILQOTICKVGEPVNLLIPFGGFPAPQVWTKEGGPLAGEE 1018 PALOGLTEHTSILVKDLPTGARILFEVGAINAGPGGPTVTKEPVTVQEILOPHILQOTICKVGEPVNLLIPFGGFPAPQVWTKEGGPLAGEE 1018 PALOGLTEHTSILVKDLPTGARILFEVGAINNAGPGGPTVTKEPVTVQEILOPHILQOTICKVGEPVNLLIPFGGFPAPQTWTKEGGPLAGEE 1019 PALOGLTEHTSILVKOLPTGARLLFEVGAINNAGPGGPTVTKEPVTVQEILOPHILQOTICKVGEPVNLLIPFGGFPAPKISMEDGGTLBSKD 1016 PALOGLTEHTSILVKOLPTGARLFEVGANNAGPGGPTVTKEPVTVQEILOPHILQOTICKVGEPVNLLIPFGGFPAPKISMEDGGTLBSKD 1016
hu-slw EDWIVANKOLIDATKFTITGLPTOARIFVRVKAVNAAGASEPAY SOPILVKEILEPPAIRIPRILKOT IRVGEAVNLVIPFOGNPRPELTWKKOGAEJ	IKSY RWMRLNFDLIKELTTEAKUMI GOVYEME ILVNSIGMSR 959A30PFWP1APSSPTHFTVEUVSDTTVLKWE PPERIGAGGLOGI VEYGKOG SAMARLNFDPYETTEAKUMI GOVYEMEVYAVNSIGMSR 30A30PFVPAPSSPTHFTVEUVSDTTVLKWE PPERIGSVDLEG GOVYGKGG SAMARLNFDLCKTTFEPKMI GOVYEWEVYAVNSIGMSR 30A30PFVPAPSSPVGLOVDI ISDTSIVLKWE PPERIGSVDLEG GOVYGKGG SAMARLNFDLCKTTFEPKMI GOVYEWEVYAVNSIGMSR 30A30PFVPAPSSPVGLOVDI ISDTSIVLKWE PPERIGSVDLEG GOVYGKGG Motif 7 cont. * Motif 8 * PANTERVENG GSTVNLPTGARLFFVVGVNIAGR TEPATLAOPVTI BEIASPPI ILVDISVDTVIS VDEVVGVGGAPLOST ALGGI TEHTSILV GOVYGK VGAPNAGGAPVTTFEPTVORI LOOPHILOOTI KVGGI MILVVPGGK PROVVNT GGAPLDTSR 886 ALGGI TEHTSILV GLOVG FRAULFFVVGVNIAGR TEPATLAOPVTI BEIASPPI ILVDI VIGI VGV VGLOK VGE VNLLI PFOGK PROVVNT GGAPLDTSR PALOGI TEHTSILV GUV GANNAGGGAPVTTFEPTVORI LOOPHILOOTI OK VGE PVNLLI PFOGK PROVVNT GGAPLDTSR 1014 PALOGI TERSMI VKDLPTGARLFFVVGI NIAGR 30A31 IK PVTVOZI LOOPHILOOTI KVGT INIM PFOGK PROVVNT GGAPLAGE 1014 PALOGI TETSILI KULVTODKI VFVGANNAGGSGAATI KE PVTVOZI LOOPHILOOTI KVGT INIM PFOGK PROVVNT GGAPLAGES 1016 PALOGI TETSILI KULVTODKI VFVGANNAGGSFAATLAOPVTI BEI GOPHI DE VLI PVTVOZI OK VGE VNLLI PFOGK PROVVNT GGAPLASES 1016 PALOGI TETSILI KULVTODKI VFVGANNAG SGAATI KE PVTVOZI LOOPHILOOTI KVGT INIM PFOGK PROVSKI SST 963 1016 PALOGI TENSILI KULVTODKI VFVGANNAG SGAATI KE PVTVOZI MORPHILI VGT I KVGT INIM PFOGK PROVSKI SST 963 1016 PALOGI TENSILI KULVTODKI VFVGANNAG SFAATLAOPVTI BEI GOPHI TENG FINICATI KVGT INIM PFOGK PROVSKI SST 957 1016 PALOGI TENSILI KULVTODKI VFVGANNAG SFAATLAOPVTI FOG PPI TI FINICATI FINICATI FINICATI FINICATI FINICASSI FIN
hu-siw Edwivankdlidktkftitglptdakifvkvavnakgasepkyksopilvkeiteppkiripkhlkotyirkogeavnlvipfogkpreltwikdgaei	IKSYRWALAFDILKELTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPFMPIADPSEPTHFTVEDVSDTTVALKAPPERIGAGGLOGIVEYGKOG IKKSYRWALAFDILKELTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPFMPIADPSEPTGLOVDI SDTSILAVA PPERMGSVDLEG GVEYGKGT 858 QSSWMELAFDILKETFEPKMI GUVYEVI FAVAAIGISKPSMPSRPFVLAVTSPTLLTVDSVDTVTMAVAPPHIGAGGLOGVLEYGFG Motif 7 cont. * Motif 8 * PANTEPVERGETVKALFTGARLLFRVGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLNQTI KKUGQLALVPFGG PAPQVWIKGGAPLDTSR 886 TALGGITEHTSILVKDLPTGARLLFRVGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLNQTI KKUGQLALVPFGG PAPQVWIKGGAPLDTSR 1018 PALGGITEHTSILVKDLPTGARLLFRVGANNAGFGAPUTTEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKGGAPLAGE 1018 PALGGITERSMLVKDLPTGARLLFRVGANNAGFGAPUTTEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKEGOPLAGE 1018 PALGGITERSMLVKDLPTGARLLFRVGANNAGFGAPUTTEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKEGOPLAGE 1014 PALGGITERSALIKUVITGALLFVKANNAGFGAPATIKEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKEGOPLAGE 1014 PALGGITERSALIKUVAGNALFVKANNAGFGAPATIKEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKEGOPLAGE 1014 PALGGITERSALIKUVAGNAKANAGFGAPATIKEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKEGOPLAGE 1019 PALGGITERSALIKUVAGNASAFKKANAGFGAPATIKEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKEGOPLAGE 1014 PALGGITERSALIKUVAGNAKANAGFGAPATIKEPVTVQEILGPRLOLPHLNQTI KKUGPVALLIPFGG PAPQVWIKEGOPLAGE 1014 PALGGITERSALIKUVAGNAKANAGFGAPATIKEPVTVQEILGPRLOPILLPTICAKVGAVALLIPFGG PAPVVKKEGOPLAGE 1014 PALGGITERSALKKANAGFGAPATIKEPVTVQENAGAPVIKEGAPVIKEGAPVAKANAGFGAPATIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVAKANAGPGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVAKANAGFGAPVIKEGAPVIKEGAPVIKEGAPVIKEGAPVIKEGAPVAKANAGFGAPVIKEGAPVI
NU-SIW EUWIVAADILUIININETITGEFEDATIVWVAAVAADASEFIISEPILUAEIIEFETATEFETATEFULAVIENUEVIIIEGEFETEIWADGAS	IKS RWMELNEDILKELTTEAKUMI GUVTEMEI LAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDUSDTTVALKWEPPERIGAGGLOGI UVEYGKOG 911 KKSYRWMELNEDIYPETTYEAKUMI GUVTEMEI LAVNSIGMSRPSPASOPEMPIADPSEPUGLCVDI SDTSIVLKWEPPERMGSVDLEG GVEYGKEG 858 KQSSRWMELNEDICKETFEPKMI GUVTEVI FAVNAIGI KSPSMPSRPFVPLAVTSPTLTVDSVDTVDKWEPPHIGAAGLOG VLEYGKEG 761 Motif 7 cont. * Motif 8 * PANTEPVER GGFTVENLPTGARILFEVUGVNIAGREPATLAOPVTIREIAEPKIRLPRHLNQTI RKVGEQLNLVVPFQG PROVVVTKGGAPLDTSR 886 TAALQGITEHTSILVKDLPTGARILFEVUGVNIAGREPATLAOPVTIREIAEPKIRLPRHLNQTI KKVGEPVNLLIPFQG PROVVVTKGGAPLDTSR 886 PALQGITEHTSILVKDLPTGARILFEVUGVNIAGREPATLAOPVTIREIAEPKIRLPRHLNQTI KKVGEPVNLLIPFQG PROVVVTKGGAPLDTSR 886 PALQGITEHTSILVKDLPTGARILFEVUGVNIAGREPATLAOPVTIREIAEPKIRLPRHLNQTI KKVGEPVNLLIPFQG PROVVTKEGOPLAGE 1018 PALQGITEHTSILVKDLPTGARLFEVUGVNIAGESGAAI KEPVTVQEI LOPHLQLPHLRQTICKKVGETVNIAIPFQG PROVVKKEGOPLAGE 1014 PALOGITEHTSILINUTTODKLOFFVAANNAGPGAPATLAOPVTIREIMOOPHILPRILPRILPRICOVGETNININIPFQG PRAFKISWM.DOGTLDSK0 1016 PAVQGITDTSILINUTTODKLOFFVAANNAGPSAATLAOPVTIREIMOOPHILPNILOVILLPROGTVKEGTVNIAIPFQG PRAFKISWM.DOGTLDSK0 1016
Motif 8 cont. Motif 9	IKSYRWMELNFDLIKELTTEAKUMIEGVATEMETIAVNSIGMSRESPASOPYNPIADESEPTHFTVEDVSDTTVALKWEDPERIGAGGLOGIVEYGKOG KKSYRWMELNFDPYETTYEAKUMIEGVATEMEVYAVNSIGMSRESPASOPYNPADESEPTHFTVEDVSDTTVALKWEDPERIGAGGLOGIVEYGKOG KQSSRWMELNFDLCKSTTFEPKKMIEGVATEMEVYAVNSIGMSRESPASOPYNPADTSEPVGLCVDISDTSIVLKWEPPERIGAGLOGIVEYGKEG Motif 7 cont. * Motif 8 * PANTEFVERIGGFTVNLPTGARILFFVVGNIAGRTEPATLAOPVTIREIAEPPTIREPHILAOTIVKSOCIALVVPGGFREQVVNTGGAPLDTSR AALQGLTEHTSILVKDLPTGARILFFVVGHNAGPGGPTVTKEPVTVQEILOEPRLQUPHLIQOTIVKVGGVNLLIFFQGFREQVVNTGGAPLDTSR PANTEFVERIGFTVNLPTGARILFFVVGHNAGPGGPTVTKEPVTVQEILOEPRLQUPHLIQOTIVKVGEVNLLIFFQGFPREQVVNTGGAPLDTSR 1018 PALOGLTEHTSILVKDLPTGARILFFVVGHNAGPGGPTVTKEPVTVQEILOEPRLQUPHLIQOTIVKVGEVNLLIFFQGFPREVTKEGGPLAGEE 1018 PALOGLTEHTSLIKULVTGDLIFVVGANNAGPSGPTVTKEPVTVQEILOEPRLQUPHLIQOTIVKVGETNINTFFQGFPREVTKEGGPLAGEE 1016 PAVQGLTEHTSLIKULVTGDLIFVVGANNAGASEFYSSOPILAPPTIREIMPPFTFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
Motif 8 cont. Motif 9	IKSY RWMELNFDLIKELTTEAKUMI GOVYEMEI IAVNSIGMSR 959ASOPTMPIAPSSEPTHFTVEOVSDTTVAL WEDPERIGAGGLOGI VEYGKOG KKSYRWMELNFDPYETTEAKUMI GOVYEMEVIAVNSIGMSR 90ASOPFVPVAPTSEPVGLOVDI ISDTSIVLKWEDPERMGSVDLEG GOVYGKGG KOSSRWELNFDLCKSTTE PKKMI GOVYEW IFAVNAIGISKPSMPSR PFVLAVISPFULAVISPTLAVDSVDTTVIK WEDPERMGSVDLEG GOVYGKGG Motif 7 cont. * Motif 8 * PANTEPVERGSTVKLPTCAKILF VVGANNAGCGAPVTTEPTAVETAPPTIRE ASPMILINGTI IKVGGLAUVVPGGK PROVWIKGGAPLDTSR 886 AALGGIEHRSILVKOLPTCAKILF VVGANNAGCGAPVTTEPTVOEI LOOPKILOPHILGOI OKVGEPVNLLIPFGGK PROVWIKGGAPLDTSR 816 PALGGIEHRSILVKOLPTCAKILF VVGANNAGCGAPVTTEPTVOEI LOOPKILOPHILGOI OKVGEPVNLLIPFGGK PROVWIKEGOPLAGE 1018 PALGGIEHRSILVKOLPTCAKILF VVGANNAGCGAPVTTEPTVOEI LOOPKILOPHILGOI OKVGEPVNLLIPFGGK PROVWIKEGOPLAGE 1018 PALGGIEHRSILVKOLPTCAKILF VVGANNAGCGFD POTVERI WOOPKICH VGTOK VGEPVNLLIPFGGK PROVWIKEGOPLAGE 1018 PALGGIEHRSILVKOLPTCAKILF VVGANNAGCGFD POTVERI MOOPHILAOPHILKOTI OKVGEPVNLLIPFGGK PROVWIKEGOPLAGE 1018 PALGGIEHRSILVKOLPTCAKILF VVGANNAGCGFD POTVERI MOOPHILE VGTOK VGEPVNLLIPFGGK PROVKIKEGOPLAGE 1018 PALGGIEHRSILVKOLPTCAKILF VVGANNAGGSGAPUTTEPTVOEI KOOPHILOPHILKOTI KVGGTINKINTEGOPLAGES 1018 PALGGIEHRSILVKOLPTCAKIKFKIKSSEN PARTAGON PATERIM OFFICI VGTOK VGEPVNLLIPFGGK PROVKIKSSES PANGELGIEFSILVKOLPTCAKIKKKKKESSEN PANTAGON PANGEN PANTERING PANTAGON VGTOK VGEPVNLLIPFGGK PROVKIKSSES PANGELGIEFSILVKOLPTCAKIKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
	IKS RWWELNEDLIKELTTEAKUMI EGVAREWRYAANSIGKSRESPASOPENPIADESEPTHETVEDVSDTVALKWEPPERGAGGLOG IVEYGKOG 911 KKS RWWELNEDPYPETTYEAKUMI EGVAREWRYAANSIGKSRESPASOPENPIADESEPTHETVEDVSDTVALKWEPPERGAGDOG IVEYGKOG 858 KQSSRWWELNEDLIKETFEPKIMI EGVAREWRIGANEWRIGASERSPERGENDESEVGLOVDI SDTSIVLKWEPPERGSVDLEG GVEYGKEGT 858 Motif 7 cont. * Motif 8 * PANTEPVERGEFTWENLETGARILE VVGVNIAGREEPATLAOPVTIREIAEPFIREPHLEDTIINVGEQLNLVVPFGGPRPQVVWTKGGAPLDTSR 886 AALQGITENTSIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPFIREPHLEDTIINVGEQLNLVVPFGGFPRQVWTKGGAPLDTSR 886 AALQGITENTSIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPFIREPHLEDTIINVGEQLNLVVPFGGFPRQVWTKGGAPLDTSR 886 AALQGITENTSIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPFIREPHLEDTIINVGEQLNLVVFGGFPRQVWTKGGAPLDTSR 886 AALQGITENTSIIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPFIREPHLEDTIINVGEQLNLVVFGGFPRQVWTKGGAPLDTSR 886 AALQGITENTSIIVKOLPTGARLEFVANNAGGSGFIVTKEPTVOELGAPRICQTPHLACTICKVGEPVNLLIPFGGFPRQVTWTKEGGPLAGEE 1018 PALGGITENTSLIVKOLPTGARLEFVANNAGBSPATLAOPVTIREIMORPHICTICKVGEVALLIPFGGFPRKISMKDGGTLDSKD 1016 PAVGGITDTSIIINLTTGDKLQFFVANNAGBSPATLAOPVTIREIMORPHILEPPINLACTIRKVGTVNLAIPFGGFPRKISMKDGGTLDSKD 1016 PAVGGITDTSIINKLTTGDPLOFVANNAGBSPATLAOPVTIREIMORPHILEPPINLACTIRKVGTVNLAIPFGGFPRKISMKDGGTLDSKD 866 VANKDLDVKKFTIGLPTDAKIEV VKAVNAGASEPYTSSDFJGGPIVKEIFFPHLKQTIRKVGTVILAIPFGGFPRELVKKDGAELDSND 866
	KKSTRWMELNFDLIKELTTEAKUMIEGVATEMETIAVNSIGMSRESPASOPTWPIAPSEPTHFTVEOVSDTTVALKWEDPERIGAGGLOGIVEYGKOG 911 KKSTRWMELNFDPYPTTTEAKUMIEGVATEMEVYAVNSIGMSRESPASOPTWPIAPSEPTHFTVEOVSDTTVALKWEDPERIGSVDLEG GVEYGKEGT 858 KQSSRWMELNFDLCKSTTFEPKKMIEGVATEMEVYAVNSIGMSRESPASOPTVPIAVTSSPVGLCVDISDTSIVLKWEPPERIGSVDLEG GVEYGKEGT 761 Motif 7 cont. * Motif 8 * PANTERVENGGFTVNLPTGARILFFVVGNIAGRTEPATLAOPVTIREIASPPTIALVDSVIDTVOKKVEPVNLLSPFGKPRPQVWTKGGAPLDTSR 886 AALGGLTEHTSILVKDLPTGARILFFVVGNIAGRTEPATLAOPVTIREIASPPTIALVDSVIDTVGKVGEVNLLIPFGKPRPQVWTKGGAPLDTSR 886 PANTERVENGGFTVNLPTGARILFFVVGNIAGRTEPATLAOPVTIREIASPPTICLOPRILOPTIGKVGEVNLLIPFGKPRPQVWTKGGAPLDTSR 886 PALGGLTERSSILVKDLPTGARILFFVVGANNAGPGGPTVTKEPVTVQEILGPRLOLPRILRQTICKVGEVNLLIPFGKPRPQVWTKEGGPLAGEE 1018 PALGGLTERSSILVKDLYTGALVFFVVGANNAGPGSPANTLAOPVTIREIASPPTICVPHLIQTICKVGEVNLLIPFGKPRPKISMEDGGTLSKD 1016 PAVGGLTERTSLIKDLYTGDKLEFVVGANNAGPGSPANTLAOPVTIREIMOPTIREIMGPTICVPHLIQTICKVGEVNLLIPFGKPRPKISMEDGGTLSKD 1016 PAVGGLTDRTSLIIRNLTGDKLOFVGANNAGPSPANTLAOPVTIREIMPPKILPPKILVFVGANNLVPFGKPRPKISMENGEFLSKFF 963 VANKDLIDKTKFTTGLPTDAKTFVVGAVNAGASEPYTSOPILVKEIMEPKIRFPKILKYTYNKVGAVNLVPFGKPRPKIKMEDGGLEDSKD 966 Motif 8 cont. Motif 9
hus fat. UUUT TO FOR FOR TO A DECEMBER OVATENT OF THE ADDRESS TO T	IKSY RWMELNFOLLKELTTERAKUMI GOVYEME ILVNSIGMSR BOASOPHVPIAPSEPTHFTVEOVSDTVAL WEDPERIGAGGLOGIVEYGKOG KKSYRWMELNFDPYPTTYEAKUMI GOVYEMEVAVVASIGMSR BOASOPFVPVAPTSEVGLOVDI SDTSIVLKWEDPERIGSVDLEG GOVYGKGG 858 KOSSRWMELNFOLCKSTTEPKKMI GOVYEWEVAVVASIGMSR BOASOPFVPVAPTSEVGLOVDI SDTSIVLKWEPPERIGSVDLEG GOVYGKGG 858 Motif 7 cont. * Motif 8 * PANTERVEIGGTVVNLPTGARLEFVVGVNIAGR TEPATLAOPVTIREIASPPILIVOSVDTVJKWEPPHIGAAGLOGVUSKGG 886 AALGGITERSSLVKULPTGARLEFVVGVNIAGR TEPATLAOPVTIREIASPPILIOPHILQTICKVGCLALUVPFGCKPRPOVWTKGGAPLDTSR 886 846 846 846 846 846 847 847 847 847 847 847 847 847
hu-fst WHYNTSDFDTVFFVNQAARSDSGEVELSVQIENMKDTATIRIRVVERAGPPINVMVKEVWGTNALVEWOAPHDDGNSEIMGWFVORADKKTMEWFNVERN	IKSS RWMRLNEDLIKELTTEAKUMIEGVVTEMIIAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKW PPERIGAGGLOG IVEYGKOG 911 KKSYRWMRLNEDPYPETTYEAKUMIEGVAYEWRIYAVNSIGMSRPSPASOPEMPIADPSEPGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKOG 911 KKSYRWMRLNEDLIKETFEPKIMIEGVAYEWRIFAVNAIGKSRPSPEPDPLTYDSPTLTVDSVDTTVTMRWPPERMGSVDLEG GVEYGKEG 858 (VSSRWMRLNEDLIKETFEPKIMIEGVAYEWRIFAVNAIGISKPSMPSRPFVPLAVTSPTLTVDSVDTTVTMRWPPHIGAAGLOG VLEYGFG 761 Motif 7 cont. * Motif 8 * PANTEPVERGEFTVENLPTGARILFEVVGVNIAGRTEPATLAOPVTIREIAEPKIRLPRHLNGTIKKVGEVNLLPFGGKPRPQVWTKEGGPLAGEE 1018 PALGGITERTSILVKDLPTGARILFEVRAHNMAGPGAPUTTEPVTVQEILOPHLQIPHLNGTICKVGEVNLLIPFGGKPRPQVWTKEGGPLAGEE 1018 PALGGITERTSILVKDLFTGARLLFEVRAHNMAGPGAPUTTEPVTVQEILOPHLQIPHLNGTICKVGEVNLLIPFGGKPRPQVWTKEGGPLAGEE 1018 PALGGITERTSILVKDLFTGARLLFEVRAHNMAGPGAPUTTEPVTVQEILOPHLQIPHLNGTICKVGEVNLLIPFGGKPRPQVWTKEGGPLAGEE 1018 PALGGITERTSILIKNLTTGDKLOFFVANYAMAGPSPAATLAOPVTIREIMOSPHILLPHLNGTIKKVGETINIMIPFGKPRPKTSSMGLOGTLDSKO 1016 PAVGGITDTSILINNLTTGDKLOFFVANYAMAGPSPAATLAOPVTIREIMOSPHILPPILLPHLNGTIKKVGETINIMIPFGKPRPKTSSMGLOGALDSKN 666 MOTIF 8 cont. MOTIF 8 cont. NOTIF 9 TSDFDVYFFVGAASDSGEVELSVGIENMKDTATIRIKVEKAGPPINVMKEVMGTNALVEMOAPKDORSEIMGFVOLADIKKTMEWFNVMERNHES 991
hu-fst VHV <mark>RTSDFDTVFFVNOAAL</mark> SDSGEVELSVOIENNKDTATIRIRVVEKAGPPINVMVKEVNGTNALVEWOAPKDDGNSEIMCVFVQKADKKTMEWFNVFERN	IKS RWMRLNFDLKELTTEAKUMI EGVARM VAVNSIGMSRPSPASOPYNPIAD PSEPTHFTEOVSDTTVALKW PPERIGAGGLOG IVEYGKOG KKS RWMRLNFDPYPTTYEAKUMI EGVARM VYAVNSIGMSRBSASOPYVPAPTSEPGLCVDDI SDTSIVLKW PPERMSSVDLEG GVEYGKEGT 858 KQSSRWRLNFDLCKETFFPKMI EGVARM VFIFAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVDTTVTMRW PPERMSSVDLEG GVEYGKEGT Motif 7 cont. * Motif 8 * PANTEPVER GEFTWENLPTGARLIF VVGWI AGREEPATLAOPVTIREI ASPETTEPHLEGTVI REVGEQLNIVVPFGG PROVVWTRGGAPLDTSR 886 AALQGITENTSILVKOLPTGARLIF VVGWI AGREEPATLAOPVTIREI ASPETTEPHLEGTVI REVGEQLNIVVPFGG PROVVWTRGGAPLDTSR 886 AALQGITENTSILVKOLPTGARLIF VVGWI AGREEPATLAOPVTIREI ASPETTEPHLEGTVI REVGEQUNIVVPFGG PROVVWTRGGAPLDTSR 886 AALQGITENTSILVKOLPTGARLIF VVGWI AGREEPATLAOPVTIREI ASPETTEPHLEGTVI REVGEQUNIVVFGG PROVVWTRGGAPLDTSR 886 AALQGITENTSILVKOLPTGARLIF VVGU AGNAGGED VTEPVTOEL GAPRI OLOPHIL ROTI REVGEVALLPFGG FRAPOTWITEGOPLAGE 1018 PALGITENTSLIVKOLPTGARLIF VVGU AGNAGGED VTEPVTOEI LOPHILOTI REVGETVILLIPFGG FRAPOTWITEGOPLAGE 1018 PAUGUTENTSLI NULTGOKLFF VVGU AGNAGGESPATLAOPVTIREI MEDEPTICEL VRGU KVGETNIMI PFGG FRAPOTWITEGOPLAGE 1018 PAUGUTENTSLI NULTGOKLFF VVGU AGNAGGESPATLAOPVTIREI MEDEPTICE HUGPFICVPRILLPGT KEVGETVNILLPFGG FRAPOTWITEGOPLAGE 1018 PAUGUTENTSLI NULTGOKLFF VVGU AGNAGGESPATLAOPVTIREI MEDEPTICE VVGU KVGETNIMI PFGG FRAPOTWITEGOPLAGE 1016 PAVGGITDTSLI NULTGOKLFF VVGU AGNAGGESPATLAOPVTIREI MEDEPTICE VVGU KVGETNIMI PFGG FPKKSWKDGGTLDSKD 1016 1016 1016 1016 1016 1016 1016 1017
hu-fst VHV TSDFDTVEV OAA SDSGE ELSVOIENNKDTATIR RVF KAGPPINVAVKEVAGNALVSKOAPKDECKSEIMG EVOKADKKIMEWRNVERA	IKSS RWARLAFDLIKELTTEAKUMI EGVATEMETIAVNSTGASRPSPASOPFAPIADPSEPTHFTVEDVSDTTVALKAF PPERIGAGGLOG IVEYGKOG 8911 KKSKRWARLAFDPYPTTTEAKUMI EGVATEMEVYAVNSIGASRHSQASOPFVPYAPTSEPVGLCVDDI SDTSIVLKAF PPERMGSVDLEG GVEYGKEGT 858 KQSSRWARLAFDLCKETTEPKAMI EGVATEV IFAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVDTTVINKAFPERMGSVDLEG GVEYGKEGT 761 Motif 7 cont. * Motif 8 * FANTERVEIGGFTVKALPTGARILFRVGAMINAGPGAPVTTEPVTVORILGPRILPPHILAGYI IRVGGCALLVVPFGG PRPQVVNT GGAPLDTSR 886 AALQGITEHTSILVKDLPTGARILFRVGAMINAGPGAPVTTEPVTVORILGPRILPPHILAGYI IRVGGCALLVVPFGG PRPQVVNT GGAPLDTSR 886 AALQGITEHTSILVKDLPTGARILFRVGAMINAGPGAPVTTEPVTVORILGPRILPPHILAGYI IRVGGCALLVVPFGG PRPQVVNT GGAPLDTSR 886 PALGGITEHTSILVKDLPTGARILFRVAHNAGPGAPVTTEPVTVORILGPRILPPHILAGYI IRVGGCALLPPGG PRPQVVNT GGAPLDTSR 886 AALQGITEHTSILVKDLPTGARILFRVAHNAGPGAPVTTEPVTVORILGPRILPPHILAGYI IRVGGCALLPPGG PRPQVTWT EGOPLAGEE 1014 PALGGITERSMLVKDLUTGDKLFFV KAINAGPGAPVTTEPVTVQEIGOPRICVPRILVGTIGKVGEPVNLLIPFGG PRPQVTWT EGOPLAGEE 1016 PAVGGITDTSLIINLTTGD KLOFFV KAINAGPSPARILAGPVTIREIMSPHILPPINLLPNLAGTI KRVGTVNIALPFGG PRPKISMKDGGTLDSKD 1016 PAVGGITDTSLIINLTTGD KLOFFV KAINAAGPSPARILAGPVTIREIMGPFILPPNIRIPPINLAGYI IRVGAVNILVPFGG PRPKISMKDGGTLDSKD 963 VANKDLIDTKKFTIGLPTDAKIFV KAVNAAGASEPYYS SOPILVKEIFPPNIRIPPINLAGYI REVGAVNILVPFGG PRPKISKKDGAELSSTF 963 Motif 8 cont. Motif 9 TSEFDTVFV QAARSDSGEYELSVOIENNKDTATITRKVCEAGPPINVMKEVNGTALVENQAPKDDGNSEINGFVGADKKTMENNVFERNHTS 991
hu-fst VHVTSDFDTVFFVROAASDSGEVELSVQIENMKDTATIRIRVVEKAGPPINVMVKEVMGTNALVEMOAPKDDGNSEIMGVFVQKADKKTMEWENVVERN	IKS HEWELNFOLKELTTEAKUMI GOVTEMETIAVESTGKSRESPASOPENPIADESETTETVOUSDTVALKWEDPERIGAGGLOGIVESTGKOG KKS HEWELNFOLKETTEPKKMI GOVTEMETVAVESTGKSRESPASOPENPIADESETTETVOUSDTVALKWEDPERIGAGGLOGIVESTGKOG KSS HEWELNFOLSKITTEPKKMI GOVTEMETVAVESIGKSRESPASOPENPIADESE PULLVOUSTIVLKWEDPERIGAGLOGIVESTGKOG Motif 7 cont. * Motif 8 * PANTERVENG GEFTVENLPTGARLLEFVENINGERTEPATLAOPUTIERIAEPHILEOTYIKEVGOLKUVUPGGEPHEQUVUTGGAPLDTSR AALOGLTEHTSILVEDLPTGARLLEFVENINGEGEPUTTEPVTVOEILOEPHILOOTIGKVGEVENLLIPFOGEPHEQUVUTGGAPLDTSR BALOGLTEHTSILVEDLPTGARLLEFVENINGEGEPUTTEPVTVOEILOEPHILOOTIGKVGEVENLLIPFOGEPHEXTEGOPLAGEE 1018 PALOGLTEHTSILVEDLPTGARLLEFVENINGEGEPUTTEPVTVOEILOEPHILOOTIGKVGEVENLLIPFOGEPHEXTEGOPLAGEE 1019 PALOGLTEHTSILVEDLUTTOLIFVENINGEGEPUTTEPVTVOEILOEPHILOOTIGKVGEVENLLIPFOGEPHEXTEMEDGOTLBSKD 1016 PAUGUTEHTSILVEDLUTTOLIFVENINGEGEPATLAOPVTIEFINDEPTIELENDEPTILEOTIGKVGEVENLLIPFOGEPHEXTEMEDGOTLBSKD 1016 PALOGLTEHTSILVEDLPTGARLLEFVENINGEGEPATLAOPVTIEFINDEPTILOOTIGKVGEVENLLIPFOGEPHEXTEGOPLAGEE 1018 PALOGLTEHTSILVEDLPTGARLLEFVENINGESGAAIIKEPTVOEILOOPTIELENDEPTILEOTIGKVGEVENLLIPFOGEPHEXTEGOPLAGEE 1019 PALOGLTEHTSLIKUVUTTOLIFVENINGESGAAIIKEPTVOEIMOOPTIEIMOOPTIEVENINGEVENEVENEGEPHEXTEMEDOOTLBSKD 1016 PAVOGLTDHTSLIINNITGOKLOFVENINGESGAAIIKEPTVISEIMOOPTIEIMOOPTIEVENINGEVENINGEFENINGENENENE PATUGUTTETSLIKUVUTTOLIFVENINGESGAAIIKEPTVENENENENENENENENENENENENENENENENENENEN
hu-fst VHVTTSDFDTVFFVHQAARSDSGEVELSVQIENMKDTATIRIRVVERAGPPINVMVKEVMGTNALVEWQAPKDGNSEIMGVFVQKADKKTMEWFNVVERN	IKS HWWELNFOLKELTTERAKUMIEGVATEWETIAVNSIGMSRESPASOPNETADESEPTIFTVEDVSDTTVALKWEDPERIGAGALOGIVEYGKOG 911 KKS HWWELNFOPYETTTERKEMIEGVATEWEVYAVNSIGMSRESPASOPNETADESEPTIFTVEDVSDTTVALKWEDPERIGAGALOGIVEYGKOG 858 KQSS HWWELNFOLCKETTEPKKMIEGVATEWEVIAVNSIGMSRESPASOPVEVAPTSSPVGLCVDISDTSIVLKWEPPERIGAGALOGIVEYGKOF 858 Motif 7 cont. * Motif 8 * PANTERVEIGGEFTVNLPTGARILFFVVGNIAGRTEPATLAOPVTIREIAEPFILTPHILKOTIKKVGQLKIVVEFGGFREVVUTTGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARILFFVVGHNAGPGGPTVTKEPVTVQEILOPHLROTICKVGEVNLLIPFGGKPLPQVWTTGGAPLDTSR 886 PALGGTEHTSILVKDLPTGARILFFVVGHNAGPGGPTVTKEPVTVQEILOPHLROTICKVGEVNLLIPFGGKPLPQVTWTEGGPLAGEE 1018 PALGGTEHTSILVKDLYTGARLFFVVGHNAGPGGPTVTKEPVTVQEILOPHLROTICKVGEVNLLIPFGGKPLPKSKMMGDGFLBSKD 1016 PALGGTEHTSLIKUVTGLYTGARLFFVVGHNAGPGGPTVTKEPVTVQEILOPHLROTICKVGEVNLLIPFGGKPLPKSKMMGDGFLBSKD 1016 PAVGGTDFTSLIKDLYTGALFFVVGAVNAGASEPYTSCEILOPPTIREIMEPPKILVFMUGTVNLAIPFGGKPLPKSKMEMEDGFLBSKD 1016 PAVGGTDFTSLIKDLYTGALFFVVGAVNAGASEPYTSCEILOPPTIREIMEPPKILVFMUGTVNLAIPFGGKPLFKSKMEMEDGFLBSKD 1016 TSBFDTVFVFQAALSDSGEKULSVGIENKKDTATIRIKVEKAGFFVSAKFVGFVNLAIPFGCKPLFVTVSKDGELSKD 103 TSBFDTVFVFVFQAALSDSGEVUSTURINTITIRIKVEKAGFFVSAKFVGFVNLAIPFGCKPLFVTVSKDGELSKD 104 STSSTDTVFVFVFVGAALSDSGEVUSTURINTITIRIKVEKAGFFVSAKFVGFVNLAIPFGCKPLFVTVSKDGELSKD 104 STSSTDTVFVFVFVGAALSDSGEVUSTURINTITIRIKVESAGFFVSSDFVSKDFVGFVNLAIPFGCKPLFVTVSKDGELSKD 104 STSSTDTVFVFVGAALSDSGEVUSTURINTITIRIKVESAGFFVSSDFVSKDGFUSTINTIFFGFFFFFVFVGFVGFVGFVNLAIPFGFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
hu-fst VHV TSEPTVECAA SDSGEVELSVOIENMKDTATERRVUE KAGPPINVAVKEVNGTALLEWOAPKDEGNSEING FVOLADEKTMEWENVERN	IKS RWARLAFDLIKELTTEAKUMI EGVATEMETIAVNSTGASRPSPASOPTAPIADPSEPTHFTVEDVSDTTVALKAF PPERIGAGGLOG IVEYGKOG 8911 KKS RWARLAFDPYPTTTEAKUMI EGVATEMEVTAVNSIGASRHSQASOPFVPVAPTSSPVGLOVDI SDTSIVLKAF PPERMGSVDLEG GVEYGKEGT 858 KQSSRWRLAFDLCKTTFEPKMIEGVATEV IFAVNAIGISKPSMPSRPFVPLAVTSSPVGLOVDI SDTSIVLKAF PPERMGSVDLEG GVEYGKEGT 761 Motif 7 cont. * Motif 8 * PANTERVEIGGFTVNLPTGARLLFRVGVINIGRTEPATLAOPVTIREIASPMTLEPHILKOTI INVGGUALVVPGGFPRDQVVVTGGAPLDTSR 886 AALQGITEHTSILVKDLPTGARLLFRVGVINIGRTEPATLAOPVTIREIASPMTLEPHILKOTI INVGGUALVVPGGFPRDQVVVTGGAPLDTSR 886 PALGGITEHTSILVKDLPTGARLLFRVGVINIGGTEPATLAOPVTIREIASPMTLEPHILKOTI INVGGUALVVPGGFPRDQVVVTGGAPLDTSR 886 PALGGITEHTSILVKDLPTGARLLFRVGVINIGGTEPATLAOPVTIREIASPMTLEOTI INVGGUALVVPGGFPRDQVVVTGGAPLDTSR 886 PALGGITEHTSILVKDLPTGARLLFRVGVINIGGTEPATLAOPVTIREIASPMTLLOPPILLOPTIREVGGUALVVPGGFPRDQVVVTGGAPLDTSR 886 PALGGITEHTSILVKDLPTGARLLFRVAHNAGPGGPIVTKEIGOPVIVCELOPPILLOPRILKOTI KVGGUALAFPGGFPRKEGPLAGEE 1014 PALGGITEHTSILVKDLVTGDKLFFVKAINAGPGGPIVTKEIGOPVILLEPVIVQEI KAPPILKOTI KVGGUALAFPGGFPRKISSMKDGGTLDSKD 1016 PAVGGITDKTSILINLTGDKLGFVVAINAGPSPARILAOPVTIREIMGPVILLPPINLLPUTVKVGGTVALAFPGGFPRKISSMKDGGTLDSKD 1016 PAVGGITDKTSILINLTGDKLGFVVAINAGPSPARILAOPVTIREIMGPPILKOTI KKVGTVALAFPGGFPRKVVSSNGBELSSTF 963 VANKDLIDTKFTITGLPTDAKIFVVAINAAGASEPYTSOLIOPVILVEIFEPPNRILPPNTRIPFHKVGTVRNILPFGGFPRKVVSSNGBLSISTF 963 Motif 8 cont. Motif 9 TSBFDTVFVVGAARSSGEVILSVOIENNKDTATTRIRVVSFAGGPPINVMVKEVNGTALVENGAARSFDGFFVSGAARKTERVFVERNHTS 991
hu-fst VHV TSDFDTVFFV QAARSDSGEVELSVQIENNKDTATIRIRVVERAGPPINVMVKEVWGTNALVEWQAPKDDGNSEIMG FVQKADKKTMEWENVVERA	IKSS HWMELNFDLIKELTTEAKUMIEGVAVEMEILAVNSIGMSRPSPASOPYNPIADPSEPTHFTVEUVSDTVALKWEDPSEIGAGGLOGIVEYGKOG KKSSHWMELNFDPYPTTYEAKUMIEGVAVEWEVIAVNSIGMSRHSQASOPYVPADTSEVGLOVDISDTSVLKWEPPEMGSVDLEGGVEYGKGT 858 KQSSHWMELNFDLCKTTFEPKMIEGVAVEVIFAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVDTTVTMEWEPPHIGAAGLOGVLYGEG Motif 7 cont. * Motif 8 * FANTEPVEIGFTVALLFTVALTURGTEPATLAOPVTIEFIASPFTLEPHILOTYINKVGEQLILVVPFOGPAPEOVWTEGAPLDTSR 886 AALQGITEHTSILVKOLFTGARLIFTVAHNAGPGAPVTTEPVTVOFILGPRILOTYINKVGEQLILVVPFOGPAPEOVWTEGAPLDTSR 886 AALQGITEHTSILVKOLFTGARLIFTVAHNAGPGAPVTTEPVTVOFILGPRILOTYINKVGEQLILVVPFOGPAPEOVWTEGAPLDTSR 886 AALQGITEHTSILVKOLFTGALLFTVAHNAGPGAPVTTEPVTVOFILGPRILOTYINKVGEQLILVVFTGGPPAPETAGAPLDTSR 886 AALQGITEHTSILVKOLFTGALLFTVAHNAGPGAPVTTEPVTVOFILGPRILOTYINKVGEQUNUVEFGGPPAPETAGAPLDTSR 886 AALQGITEHTSILVKOLFTGALLFTVAHNAGPGAPVTTEPVTVOFILGPRILOTYINKVGEVATINNIFFGGPPAPETAGPIAGE 911 920 920 920 920 920 920 921 921 921 921 921 921 921 921
hu-fst VHVETSDFDTVFFVEQAAESDSGEVELSVQIENMKDTATIRIRVVEKAGPPINVMVKEVWGTNALVEWOAPKDDGNSEIMGVFVQKADKKTMEWENVVERN	IKSS RWMRLNEDLIKELTTEAKUMIEGVVTEMEILAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKW PPERIGAGGLOGIVEYGKOG 911 KKSYRWMRLNEDPYPETTYEAKUMIEGVATEWR VYAVNSIGMSRPSPASOPEMPIADPSEPGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKOG 858 KQSSRWMRLNEDLCKETFEPKMIEGVATEWR IFAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVDTTVTMRW PPERMGSVDLEG GVEYGKEGT 858 Motif 7 cont. * Motif 8 * PANTEPVERGEFTVENLPTGARILFEVGATEWRIFGANAGEPATLAOPVTIREIAEPFIRLPAHLNQTIKVGEVGLUNUVPFGKPRQVWTKEGGPLAGEE 1018 PALGETERTSILVKDLPTGARILFEVGATMAGEGPLOTTEPTVORILOPFILEDHLEQTICKVGEVNLLIPFGKPRQVWTKEGGPLAGEE 1018 PALGETERTSILVKDLPTGARLLFEVGATMAGEGPLOTKEPVTVQEILOPHLQLPHLEQTICKVGETVNLLIPFGKPRQVWTKEGGPLAGEE 1018 PALGETERTSILVKDLFTGARLLFEVGATMAGEGPLOTKEPVTVQEILOPHLQTICKVGETVNLLIPFGKPRQVTWKEGGPLAGEE 1018 PALGETERTSILVKDLFTGARLLFEVGATMAGEGPLOTKEPVTVQEILOPHLQTICKVGETVNLLIPFGKPRPQVWKKEGGPLAGEE 1018 PALGETERTSILVKDLFTGARLFFVGATNAGESGATIKEPVTVQEILOPHLQTICKVGETVNLAIPFGKPRPVTKEGGPLAGEE 1014 PALGETERTSILINUTTDLIFTVAINAGESGATIAOPVTIREIMONPHILEPNLQTIKKVGETVNLAIPFGKPRPVTKEGGPLAGEE 1014 PALGETERTSILINUTTGDLIFTVAINAGESGATIAOPVTIREIMONPHILEPNLQTIKKVGETVNLAIPFGKPRPKTWSNGEGPLAGEE 1014 PALGETERTSILINUTTGDLIFTVAINAGESGATIAOPVTIREIMONPHILEPNLQTIKKVGETVNLAIPFGKPRPKTWSNGEGPLAGEE 1014 PALGETERTSILSUKULFTVAINAGESGATIAOPVTIREIMONPHILEPNLQTIKKVGETVNLAIPFGKPRPKTWSNGEGPLAGEE 1014 PALGETERTSILSUKULFTVAINAGESGATIAOPVTIREIMONPHILEPNLQTIKKVGETNALTFFGKFREVTVSNGEPLSTF 963 866 MOTIF 8 cont. MOTIF 8 cont.
hu-fet UHUETSDEDTVEFUEOAAESDSGEWETSVOTENMKDTATTRERVELAGDDTNUAUKEUWCTNALVEWOADEDCNSETWCKEUWOADETSCH	IKS TRAMELNFOLLKELTTEA KUMI GOVTEMETIAVN STGMSRESPASOPEMPIAPSESPTHFTVEOVSDTTVAL MEDPERIGAGGLOGIVEYGKOG KKSTRAMELNFOPYETTEAKEMI GOVTEMEVEAVNSIGMSRESOASOPEVPAAPSESPTHFTVEOVSDTTVAL MEDPERIGAGGLOGIVEYGKOG KOSSRAMELNFOLKSTTEPKAMI GOVTEMEVEAVNSIGMSRESOASOPEVPAAPSESPTHFTVEOVSDTVAL MEDPERIGAGGLOGIVEYGKOG Motif 7 cont. * Motif 8 * PANTEEVERGETVKLEPTGARLEFVVGVNIAGRTEPATLAOPTIBEIASPPTILTVDSVTDTVDSVMEMPPHIGAAGLOGVLESGEVEGT ALOGITEHTSSILVGLEPTGARLEFVVGVNIAGRTEPATLAOPTIBEIASPPTILTOTVSTKAVGVNILTVDSVDTVTGGAPLOSS PALOGITEHTSSILVGLEPTGARLEFVVGVNIAGRTEPATLAOPTIBEIASPPTILTOTVSTKAVGTIMINOPTGSPHPQVVNIKGGAPLDTSR 886 ALOGITEHTSSILVGLEPTGARLEFVVGVNIAGRTEPATLAOPTIBEIASPPTILTOTVSTKAVGTIMINOPTGSPHPQVVNIKGGAPLDTSR PALOGITEHTSSILVGLEPTGARLEFVVGANNAGGAPUTTEFUTVOEILOPHILOTIOKVGSTVMSTGOFVNLLIPFGSPHPQVVNIKGGAPLDTSR 1014 PALOGITEHTSSILVGLEPTGARLEFVVGANNAGGSGAPUTTEFUTVOEILOPHILOTIOKVGSTVMIATGOSPHQEKPKSSMEDGGTLSKD 1016 PALOGITEHTSSILVGLEVGVANVAGGSGAPUTTEIDTVVOEILOPHILOTIOKVGSTVMIATPGCKPHPKIKSSNGEDGTLSKD 1016 PALOGITEHTSSILVGLEVGVANVAAGSSGAATIKEPTVOEIMGPPTIREINOPTIREINOPTIREKGEVKUSTINKAPGCKPHPKIKSSNGEDSSTF963 VANKDLIDVTKTIGLEPTDAKIFVVANVAAGASEPTYSOPTIREINOPTIREINOPTIREINOPTIREVGETVMLATPGCKPHPKKSSNGEDSSTF963 VANKDLIDVTKTUGLEVGVANVAAGASEPTYSOPTIREINOPTIREINOPTIREVGETVMLATPGCKPHPKKSSNGEDSSTF963 1016 MOTIF 8 cont. MOTIF 8 1017 10170000000000000000000000000000
	KKSTRWMELNFDLIKELTTEAKUMIEGVATEMETIAVNSIGMSRESPASOPTWPIAPSEPTHFTVEOVSDTTVALKWEDPERIGAGGLOGIVEYGKOG 911 KKSTRWMELNFDPYPTTTEAKUMIEGVATEMEVYAVNSIGMSRESPASOPTWPIAPSEPTHFTVEOVSDTTVALKWEDPERIGSVDLEG GVEYGKEGT 858 KQSSRWMELNFDLCKSTTFEPKKMIEGVATEMEVYAVNSIGMSRESPASOPTVPIAVTSSPVGLCVDISDTSIVLKWEPPERIGSVDLEG GVEYGKEGT 761 Motif 7 cont. * Motif 8 * PANTERVENGGFTVNLPTGARILFFVVGNIAGRTEPATLAOPVTIREIASPPTIALVDSVIDTVOKKVEPVNLLSPFGKPRPQVWTKGGAPLDTSR 886 AALGGLTEHTSILVKDLPTGARILFFVVGNIAGRTEPATLAOPVTIREIASPPTIALVDSVIDTVGKVGEVNLLIPFGKPRPQVWTKGGAPLDTSR 886 PANTERVENGGFTVNLPTGARILFFVVGNIAGRTEPATLAOPVTIREIASPPTICLOPRILOPTIGKVGEVNLLIPFGKPRPQVWTKGGAPLDTSR 886 PALGGLTERSSILVKDLPTGARILFFVVGANNAGPGGPTVTKEPVTVQEILGPRLOLPRILRQTICKVGEVNLLIPFGKPRPQVWTKEGGPLAGEE 1018 PALGGLTERSSILVKDLYTGALVFFVVGANNAGPGSPANTLAOPVTIREIASPPTICVPHLIQTICKVGEVNLLIPFGKPRPKISMEDGGTLSKD 1016 PAVGGLTERTSLIKDLYTGDKLEFVVGANNAGPGSPANTLAOPVTIREIMOPTIREIMGPTICVPHLIQTICKVGEVNLLIPFGKPRPKISMEDGGTLSKD 1016 PAVGGLTDRTSLIIRNLTGDKLOFVGANNAGPSPANTLAOPVTIREIMPPKILPPKILVFVGANNLVPFGKPRPKISMENGEFLSKFF 963 VANKDLIDKTKFTTGLPTDAKTFVVGAVNAGASEPYTSOPILVKEIMEPKIRFPKILKYTYNKVGAVNLVPFGKPRPKIKMEDGGLEDSKD 966 Motif 8 cont. Motif 9
	KKSERWARLAFDLIKELTTEAKUMIEGVAVEMETIAVNSIGMSRESPASOPYNPIADPSEPTHFTVEUVSDTTVALKWEDPERIGAGGLOGIVEYGKUGS KKSERWARLAFDPYPTTTEAKUMIEGVAVEWEVIAVNSIGMSRESPASOPYNPIADPSEPTHFTVEUVSDTTVALKWEDPERIGAGALOGIVEYGKUGS 858 KQSERWARLAFDLIKETTEPKIMIEGVAVEVIFAVNAIGISKPSMPBRPFVPLAVTSPTLLTVDSVDTTVIKKWEDPERIGAGLOGIVEYGKGT Motif 7 cont. * Motif 8 * FANTERVEIGGFTVKALPTGARILFRVGVINIGREFENIAOPVTIEFIAPPTIEDHILGYTIRVGGCALVVPFGGPRPQVVWTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGREFENIAOPVTIEFIAPPTIEDHILGYTIRVGGCALVVPFGGPRPQVVWTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGREFENIAOPVTIEFIAPPTIEDHILGYTIRVGGCALVVPFGGPRPQVWTTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGREFENIAOPVTIEFIAPPTIEDHILGYTIRVGFCININUPPGGPRPQVWTTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGGPOTTTEPVTVOFICGPRLOUPHLROTOKVGEPVALLPFGGRPRQVVWTTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGANAGPGGPIVTKEPVTVOFICGPRLOUPHLROTOKVGEPVALLPFGGRPRQVWTTGGAPLAGES 1018 PAUGGITERSULVKOLPTGARLFRVKANNAGPGAPVTIEFVTVOFICGPRICVPRILROTOKVGEPVALLPFGGRPRPVTWTEGGPLAGES 1019 PAUGGITERTSALIKDLVTGDKLFFVKANNAGPGSPATLAOPVTIEFIMEMORPTICVPRILVGTVKVGTVNIAIPFGGRPRKVTWSNGEPLSSTF 963 VANKDLIDTKFTIGLPTDAKIFVKANNAGASEPYTYSOFILVEFIEFPFIRIPPILVGTVIRFGGAVNIVFFGGRPRFUTWSNGEPLSSTF 964 Motif 8 cont. Motif 9
MOLLE 0 CONC. MOLLE 9	IKSY RWARLAFDLIKELTYEA KUMI GOVYEMI ILAVNSIGMSRPSPASOPIMPIADPSEPTHFTVEDVSDTTVALKW PPERIGAGGLOGI UVYCKOG 911 KKSY RWARLAFDPYPTTYEA KUMI GOVYEMI ILAVNSIGMSRPSPASOPIMPIADPSEPGLOVDI SDTS ULAWA PPERMGSVDLEG GVYCKOG 858 KQSSRWARLAFDLCKETFFERKMI GOVYEMI FAVANIGI KSPSMPSRPFVPLAVTSEPULUTDSVDTTVTMRW PPERMGSVDLEG GVYCKEGT 858 SOMMELAFDLCKETFFERKMI GOVYEMI FAVANIGI KSPSMPSRPFVPLAVTSEPULUTDSVDTTVTMRW PPERMGSVDLEG GVYCKEGT 858 Motif 7 cont. * Motif 8 * PANTEPVER GGFTVENLPTGARILF VVGVNIAGRTEPATLAOPVTIREIAEPFIRLPRHLNQTI REVGEQUNLVVPFGK PRPQVWTKEGGPLAGEE PALGETERTSILVKOLPTGARLLF VVGVNIAGRTEPATLAOPVTIREIAEPFIRLPRHLNQTI REVGEQUNLVVFFGK PRPQVWTKEGGPLAGEE 1018 PALGETERTSILVKOLPTGARLLF VVGVNIAGRSEPATLAOPVTIREIAEPFIRLPRHLNQTI KEVGEVNILLPFGK PRPQVTWTKEGGPLAGEE 1018 PALGETERTSILVKOLPTGARLLF FVANNAGEGPLATEPVTVQEILGPHLQTICK VGETVNILLPFGK PRPQVTWTKEGGPLAGEE 1014 PALGETERTSALIKOLVTGDLI FVVANNAGESGAAI KE PVTVQEILGPHLQTICK VGETVNILLPFGK PRPQVTWTKEGGPLAGEE 1014 PALGETERTSALIKOLVTGDLI FVVANNAGESBAATLAOPVTIREIMOSPHILLPRULKQTI REVGETVNLUPFGK PRPKVTWSINGEPLSSTF 963 WANKDLIDKTKFTIGLPTDAKIFV VKANNAGESPAATLAOPVTIREIMOSPHILLPPUR REVGETVNLUPFGK PRPKVTWSINGEPLSSTF 963 WANKDLIDKTKFTIGLPTDAKIFV VKANNAGESPAATLAOPVTIREIMOSPHILLPPUR REVGEVNULVFPGK PRPKVTWSINGEPLSSTF 963 866
Motif 8 cont. Motif 9	IKSS RAMELNEDILKELTTEAKUMI GUVTEMEI LAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDUSDTTVALKAPPERIGAGGLOGI UVYCKDGS 911 IKKSYRAMELNEDILKELTTEAKUMI GUVTEMEI LAVNSIGMSRBSOPSVPUPAPTSEPUGLCVDI SDTS ULIKAPPERIGAGGLOGI UVYCKDGS 911 IKKSYRAMELNEDILKETTEPKEMI GUVTEMEI GAVEV IFAVANIGISKSRBSOPSVPUPAPTSEPUGLCVDI SDTS ULIKAPPERIGSUDLGG GVYCKGG 761 Motif 7 cont. * Motif 8 * PANTEPVER GETVENLPTGARILFEVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLRQTI KVGSQLNLVVPFQG PRPQVVWTKGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFEVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLRQTI KVGSQLNLVVPFQG PRPQVVWTKGGAPLDTSR 886 PALQGITEHTSILVKOLPTGARILFEVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLRQTI KVGSQLNLVVPFQG PRPQVWTKEGOPLAGEE 1018 PALQGITEHTSILVKOLPTGARILFEVUGANNAGGGAPUTTEPVTVQEILOPPKIRLPRHLRQTI KVGSQVNLLIPFQG PRPQVTWTKEGOPLAGEE 1018 PALQGITEHTSILVKOLPTGARILFEVUGANNAGGSGADI KEPVTVQEILOPPHLQTI KVGSQVNLLIPFQG PRPQVTWTKEGOPLAGEE 1018 PALQGITEHTSILVKOLPTGARILFEVUGANNAGGSGADI KEPVTVQEILOPPHLQTI KVGSQVNLLIPFQG PRPQVTWTKEGOPLAGEE 1014 PALQGITEHTSILVKOLPTGARILFEVUGANNAGSGADI KEPVTVQEILOPPHLQTI VKVGSQVNLLIPFQG PRPQVTWTKEGOPLAGEE 1014 PALQGITEHTSILIKOLFTGARILFEVUGANNAGGSGADI KEPVTVQEILOPPHLQTI VKVGSQVNLLIPFQG PRPQVTWTKEGOPLAGEE 1014 PALQGITEHTSILIKOLFTGARLFEVUGANNAGGSGADI KEPVTVQEILOPPHULQTI KVGSQVNLLIPFQG PRPQVTWTKEGOPLAGEE 1014 PALQGITEHTSILIKOLFTGARLFEVUGANNAGGSGADI KEPVTVQEILOPPHICATI VKVGSVGEVNLLIPFQG PRPVTVSKEGOPLAGEE 1014 PALQGITEHTSILIKOLFTGARLFEVUGANNAGGSGADI KEPVTVQEILOPPHICAPVIKEGOPLAGEN 1016 PAVQGITDTSILIENLTTGDKLOPVTANKAGASEPYTYSQUPVIKEINGPHIKGPVIKKSGANIKTIPFGG PRPVENKKEGAPLAGEN 1000 VANKOLIDTKFTITGLPTDAKIFVKAVNAGASEPYTISQUPVIKEILPPUTIKEIKGPVILLPPUGAPNELTINKKEGAPLAGENDKOG 866
	IKSSTRUMELINFULKELTTEAKUMIEGVAVEMETIAVNSTGMSRPSPASOPYMPIADPSEPTHFTVEDVSDTTVALKWEDPSETGAGGLOGIVEYGKUGS SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUST SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUSTUNDER SAUGUST SAUGUSTUNDER SAUGUSTUNDER S
hu-slw EDWIVANKOLIDETKFTITGLPTDAKIFVKVKAVNAAGASEPKYYSQPILVKEILEPPKIRIPRHLKQTYIRRVGEAVNLVIPFQGKPRPELTWKKDGAEI	IKSS RAMELNEDLIKELTTEAKUMI GUVTEMEI LAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKAPPERIGAGGLOGI IVEYGKOG KKSS RAMELNEDDIKETTEEAKUMI GUVTEMEI LAVNSIGMSRHSOASOPEVPUAPTSEPVGLOVDI SDTS IVLKAPPERIGAGGLOGI IVEYGKOG NOSS RAMELNEDLIKETTEEPKIMI GUVTEMEI FAVNAIGISKPSMPSRPFVLAVTSPTLITVUSVTUTVINKAPPERIGAGGLOGI VEYGKOG Notif 7 cont. * Notif 8 * PANTEPVER GETVENLPTGARILEFVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLROTI INVGEOLNLVVPFOG PRPOVVNIK GGAPLDTSR RALOGITEHTSILVKOLPTGARILFFVUANNAGPGAPUTTEPVTVOEILOPPKIRLPRHLROTI KAVGEVNLLIPFOG PRPOVVNIK GGAPLDTSR PALOGITEHTSILVKOLPTGARILFFVUANNAGPGAPUTTEPVTVOEILOPPKIRLPRHLROTI KAVGEVNLLIPFOG PRPOVVNIK GGAPLDTSR PALOGITEHTSILVKOLPTGARILFFVUANNAGPGAPUTTEPVTVOEILOPPKIRUPTOKI KUGEVNLLIPFOG PRPOVVNIK GOPLAGEE 1018 PALOGITEHTSILVKOLPTGARILFFVUANNAGPSGAAI KEPVTVOEILOPPKIRUPTIOK KUGEVNLLIPFOG PRPOVVNIK EGOPLAGEE 1014 PALOGITERTSALIKOUVTONLIFFVUANNAGPSGAAI KEPVTVOEILOPPKIRUPTIVKVGAUNVNIKUGATINNIPFOG PRPOVINKEGOPLAGEE 1014 PALOGITERTSALIKOLPTGARILFFVUANNAGPSGAAI KEPVTVOEILOPPKIRUPTIVEVIKUKUS VINLAIPFOG PRPOVINKEGOPLAGEE 1014 PALOGITERTSALIKOLPTGARILFFVUANNAGPSGAAI KEPVTVOEILOPPKIRUPTIVKVSKINGUGTINNAGPSGAPLOSE 1015 PALOGITERTSALIKOLVTONKING VINAGESGAAI KEPVTVOEILOPPKIRUPTIVKVSKINGUTIVKKSKAPPKILIPFOG PROVINKEGOPLAGEE 1014 PALOGITERTSALIKOLVTONKING VINAGESGAAI KEPVTVOEILOPPHICOPILVKKSKOFTINING POKTNIKEGOPLAGEE 1014 PALOGITERTSALKKOLVTONKING VINAGESGAAI KEPVTVOEILOPPHICOVINKKSKOFTININGESSKAI SEG 1016 PAVOGITORTSALIFIKVANNAGPSGANIKEPVTVOEILOPPHICOVINKKSKOFTINING POKTNIKSKANGOGTISBS 1016 PAVOGITORTSALIFIKVANNAGPSGANIKEPVTVENEN 1016 PAVOGITORTSALIFIKVANNAGPSGANIKEPVTVENEN 1016 PAVOGITORTSALFKOKSKANGESSKAIKS 1016 PAVOGITORTSALFKOKSKANGESSKAIKS 1016 PAVOGITORTSALFKOKSKANGESSKAIKSKANGESSKAIKS 1016 1017
fu-cc EDWLPAVQGLTDRTSLIINLTTGDKLQFVRAYNMAGPSPAATLAQPVTIREIMQRPTILLPRNLRQTIIRRVGDTVNLAIPFQGKPRPKVTWSKNGEPI	IKS RWARLAFDLIKELTTEAKUMI EGVATEMETIAVASIGASRESPASOPENPIADESEPTIFTEOUSDITVALKAFDPERIGAGGLOGIVENDGS 911 KKS RWARLAFDPYPTITEAKUMI EGVATEMEVVAVASIGASRESPASOPENPIADESEPTIFTEOUSDITVALKAFDPERMGSVDLEGGVENDGS 858 KOSS RWARLAFDLICKTIFEPKMI EGVATEV IFAVANIGISKESMESPEPPIADISEVGLOVDI SDISIVLKAFDERMGSVDLEGGVENDGF Motif 7 cont. * Motif 8 * PANTERVENGEFTWALLPTGARLIFTVGAVNIGREFEPATLAOPVTIENIASPENTERDALISEVGLOVDI INTEGORAGLOGIVENDGS AALQGITENTSIVKOLPTGARLIFTVGANAGEGPTVTEPVIDELGEPRILEDIENIASPENTERDALISEVGLORVGEVALLIPFGGEPRAVENDGGEDINTEGORAGE PANTERVENGETTALIFTVGANLEFTVGANLEFTVGAPLIGESEVGLOGIVENDELGEPRILEDIENIASTIKVGEGLAUVVENGGERAFLDTSR 886 AALQGITENTSIVKOLPTGARLIFTVGANAGEGEPTTEPVIDELGEPRILEDIENIASTIKVGEGLAUVVENGGERAFLDTSR 886 AALQGITENTSIVKOLPTGARLIFTVGANAGEGEPTTEPVIDELGEPRILEDIENIASTIKVGEGLAUVVENGGERALIFTEGATAGEDISS 886 AALQGITENTSIVKOLPTGARLIFTVGANAGEGEPTVENEDIENIASPENTERDALISTIKVGEGLAUVVENGGERAFLDTSR 886 AALQGITENTSIVKOLPTGARLIFTVGANAGEGEPTVENEDIENIASPENTERDALISTIKVGEGLAUVVENGGERAFLEGENIAGEDISS 886 AALQGITENTSIVKOLPTGARLIFTVGANAGEGEPTVENEDIENIASPENTERDENIENIASTIKVGEGLAUVVENGGENVENEDIENIASE PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASPENTERDENIENIASTIKS PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASPENTERDENIENIASTIKS PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASPENTERDENIENIASTIKS PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASPENTERDENIENIASTIKS PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASTIKS PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASPENTERDENIENIASTIKS PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASTIKS PALGTENTSIVKOLPTGARLIFTVGANGEGEPTVENEDIENIASTIKS PALGTENTENIKKOLFTGENENIASTIKS PALGTENTENIK
	IKSYRWARLAFDILKELTYEAKEMI GOVYEMI HAVNSIGMSRPSPASOPFMPIADPSEPTHFTVEDVSDTTVALKW PPERIGAGGLDG IVEYGKDGS IKKSYRWARLAFDILKELTYEAKEMI GOVYEMI HAVNSIGMSRHSOASOPFVPIAPTSEPVGLOVDI SDTSIVLKW PPERMGSVDLGG GVEYGKGT 858 QSSRWARLAFDILKETFEPKMI GOVYEVI FAVAIGSKPSPSPSPPPIADTSPTLLTVDSVDTVTMRW PPERMGSVDLGG GVEYGKGT Motif 7 cont. * Motif 8 * PANTEPVERGETVKLPTGARLLFFVGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLROTIIKVGGUNLVVPFGKPRPQVWIKGGAPLDTSR 886 TALGGTERTSLIVKDLPTGARLLFFVGANNAGFGAPITTEPVTVQEILOPPKIQLPHLROTICKVGEPVNLLIFFGGPRAGTAGE PALGGTERSMLVKDLPTGARLLFFVGANNAGFGGPITTEPVTVQEILOPPKIQLPHLROTICKVGEPVNLLIFFGGPRAGTAGE 1014 PALGGTERSMLVKDLPTGARLLFFVGANNAGFGGPITTEPVTVQEILOPPKIQTOCKVGEPVNLLIFFGGPRAFQVWIKGGAPLAGE 1014 PALGGTERSMLVKDLPTGARLLFFVGANNAGFGGPITTEPVTVQEILOPPKIQTOCKVGEPVNLLIFFGGRPRQVWIKGGAPLAGE 1014 PALGGTERSMLVKDLPTGARLLFFVGANNAGFGGPITTEPVTVQEILOPPKICUTURTUGTUCKVGEPVNLLIFFGGRPRQVWIKGGPLAGE 1014 PALGGTERSMLVKDLPTGARLLFFVGANNAGFGGPITTEPVTVQEILOPPKICUTURTUGTUCKVGEPVNLLIFFGGRPRGVTWKEGGPLAGE 1014
	IKSS RWALNEDLIKELTYEAKUMI GUVYEMI I AVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKW PPERIGAGGLOG IVEYGKOGS 911 KKSYRWARLNEDPIPETTEAKAMI GUAYEMI I AVNSIGMSRBSOPEVPUAPTSEPUGLOVDI SDTSIVLKW PPERMGSVDLEG GVEYGKGT 858 KOSSRWARLNEDLIKETFEPKMI GUAYEVI FAVNAIGISKPSMPSRPFVLAVTSPTLLTVDSVDTTVIKW PPERMGSVDLEG GVEYGKGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGETVENLPTGARILFRVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLRQTI KVGEQINLVPFQG PRPQVWIKGGAPLDTSR 886 TALQGITEHTSILVKDLPTGARILFRVARNAGFGAPUTTEPVTVQEILGPRIQLPHILQTICKVGEPVNLLIPFQG PRPQVWIKGGAPLDTSR 886 PALQGITEHTSILVKDLPTGARILFRVARNAGFGAPUTTEPVTVQEILGPRIQLPHILQTICKVGEPVNLLIPFQG PRPQVWIKGGPLAGEE 1018 PALQGITERTSALLKVUTDLVTGARILFFVRARNAGFGAPUTTEPVTVQEILOPHILQTICKVGEPVNLLIPFQG PRPQVWIKGGPLAGEE 1014 PALPGITERTSALLKOULPTGARILFFVRARNAGFGAPUTTEPVTVQEILOPHICQTICKVGEPVNLLIPFQG PRPQVWIKGGPLAGEE 1014
	IKSS RWMRLNEDDLIKELTTEAKUMIEGVVTEMEIEAVNSIGMSRPSPASOPFMPIADPSEPTHETVEDVSDTTVALKWEPPERIGAGGLOGIVEYGKOGS 911 KKSYRWMRLNEDDYPETTEAKUMIEGVATEWEVIAVNSIGMSRBSOPFVPIAPTSEVGLCVDDISDTSIVLKWEPPERMGSVDLEGGVEYGKGG SSRWMRLNEDLCKETFFEPKMIEGVATEWEIFAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVDTVTMRWEPPHIGAAGLOGVLEYGFG Motif 7 cont. * Motif 8 * PANTEPVERGEFVENLPTGARILFEVVGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLNOTIGKVGEPVNLLIPFOGFPRPQVWTKGGPLAGEE 1018 PALOGLTEMTSILVKDLPTGARILFEVTAHNVAGPGGPIVTEPVTVOEILOPPKIRLPRHLNOTIGKVGEPVNLLIPFOGFPRPQVWTKEGGPLAGEE 1018 PALOGLTEMTSILVKDLPTGARILFEVTAHNVAGPGGPIVTKEPVTVOEILOPPKIRLPRHLNOTIGKVGEPVNLLIPFOGFPRPQVWTKEGGPLAGEE 1018 PALOGLTEMTSILVKDLPTGARILFEVTAHNVAGPGGPIVTKEPVTVOEILOPPKIRLPRHLNOTIGKVGEPVNLLIPFOGFPRPQVWTKEGGPLAGEE 1014 PALOGLTEMTSILVKDLPTGARILFEVTAHNVAGFGGPIKEPVTVOEILOPPKIRLPRHLNOTIGKVGEFVNLLIPFOGFPRPQVWTKEGGPLAGEE 1014
FUECO EDUCATION AND A CONTRACTOR AND A C	IKS RWMRLNFDLKELTTEAKUMI GOVYEMI HAVNSIGMSRPSPASOPYMPLADPSEPTHFTVEOVSDTTVALKW PPERIGAGGLOG IVEYGKOG KKS RWMRLNFDPYPTTYEAKUMI GOVYEMRVYAVNSIGMSRHSQASOPYVPAPTSEPVGLOVDI SDTSIVLKW PPERMGSVDLEG GVEYGKEGT 858 KQSSRWMELNFDLCKETFFERKMI GOVYEV IFAVNAIGISKPSMPSRPFVPLAVTSPTLLTVDSVDTTVTMRW PPERMGSVDLEG GVEYGKEGT Motif 7 cont. * Motif 8 * PANTEPVERGEFTVKNLPTGARILF VVGVNIAGREFEPATLAOPVTIREIAEPFNIRLPRHLNOTVIRKVGEVINLVVPFGKPRPQVVVTGGAPLDTSR 886 AALQGITENTSILVKOLPTGARLFFVRAHNAGEGAPVTTEPTVOELGAPRIAQTICKVGEVNLLPFGKPRPQVVVTGGAPLDTSR 886 AALQGITENTSILVKOLPTGARLFFVRAHNAGEGAPVTTEPTVOELGAPRIAQTICKVGEVNLLPFGKPRPQVVVTGGAPLDTSR 940GUTENTSILVKOLPTGARLFFVRAHNAGEGAPVTTEPTVOELGAPRIAQTICKVGEVNLLPFGKPRPQVVTTEGGPLAGE 1018 PALGETENTSLIKOLVTGOKLFFVRAINLAGESGAAIIKEPVTVOELGAPRICATICVPRHLRQTIVKVGFTNINIPFGKPRPVKISMKDGGTLDSKD 1016
IU-CC EDWLPAVOGLIDETSLIIENLITGDKLOFEVRAYNMAGPSPAATLAOPVIIREIMORPHILLPENLEOTLIREVGDTVNLAIPFOGEPRPKVTWSENGEPI	IKS RWARLAFDLIKELTTEAKUMI EGVATEMETIAVASIGASRESPASOPENPIADESEPTIFTEOUSDITVALKA PPERIGAGGLOGIVEYGKOG KKS RWARLAFDPYPTITEAKUMI EGVATEMEVYAVASIGASRESPASOPENPIADESEVGLOVDI SDISIVLKA PPERMGSVDLEG GVEYGKEGT 858 KOSS RWARLAFDLICKTIFEPKMIEGVATEV IFAVANIGISKPSMPSRPFVPLAVISPTLLTVDSVDITVIKAN PPERMGSVDLEG GVEYGKEGT Motif 7 cont. * Motif 8 * TRANEEVE GGFTWALPTGARILFRVGVINGRTEPATLAOPVTIETIASPHTIRDENINGGULIVUPGGEPREQUVUTGGAPLDISR 886 AALQGITEHTSILVKOLPTGARILFRVGVINGRTEPATLAOPVTIETIASPHTIRDENINGGULIVUPGGEPREQUVUTGGAPLDISR 886 AALQGITEHTSILVKOLPTGARILFRVGVINGGGEPTTEPVIVQEILGPRILOPPICTURGVGEVALLFPGGEPROVUTGGAPLDISR 886 AALQGITEHTSILVKOLPTGARILFRVGVINGGEOFUTTEPVIVQEILGPRILOPPICTURGVGEVALLFPGGEPROVUTTEGOPLAGES 1018 PALGGITERTSALIKDIVTGDKLFFVAANNAGPGGPTVTEPVIVQEILGPRILOPPICKVGEVALLPFGGEPVALLPFGGEPLAGES 1016
IU-CC PERLEMAY GENTRETS LITENLITED LOF WAAI MAGES PAATLA PVILE IMORPHILLPHNLHOTLIER VGBTVNLAIPFOGEPEPKVTWSENGEPI	KKS RWARLNFDLKELTTEAKUMI GOVYEMI ILAVNSIGMSRPSPASOPYMPIAPSEPTHFTVEOVSDTTVALKW PPERIGAGGLOGIVYYGKOG 911 KKS RWARLNFDPYETTYEAKUMI GOVYEWRVYAVNSIGMSRHSQASOPFVPVAPTSEVGLCVDDI SDTSIVLKWR PPERMGSVDLEG GVYGKGGT 858 KQSSRWRLNFDLCKSTTFEPKMI GOVYEV IFAVNAIGI KPSMPERPFVPLAVTSPPTLTVDSVDTTVIKWR PPERMGSVDLEG GVYGKGFG 761 Motif 7 cont. * Motif 8 * PANTEFVER GEFTVNLPTGARLLFRVGVINGRTEPATLAOPVTIREIAEPFTILFPHIROTYIRKVGGUNLVVPGGFPHQVWTKGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARLLFRVGHMAGPGAPVTTEFVTVQEILOPPLQLPHRLROTICKVGEPVNLLIPFGGKPHQQVWTKGGAPLDTSR 886 PANTEFVER GEFTVNLPTGARLLFRVGHMAGPGAPVTTEFVTVQEILOPPLQLPHRLROTICKVGEPVNLLIPFGGKPHQQVWTKGGPLAGEE 1018 PALGGTEHTSILVKDLPTGARLLFFVRAHNAGPGGPTVTKEPVTVQEILOPPLQLPHRLROTICKVGEPVNLLIPFGGKPHQQTWTKEGGPLAGEE 1018 PALGGTEHTSILVKDLYTGARLFFVRAHNAGPGGPTVTKEPVTVQEILOPPLQTPHLROTICKVGEPVNLLIPFGGKPHQQTWTKEGGPLAGEE 1018 PALGGTEHTSALIKDLVTGDUTGARDTFFVRAHNAGPGGPLOVER
	IKSY RWMELNFDLIKELTTEAKUMI GOVYEMEI IAVNSIGMSR 959A30PFWPIAPPSSPTHFTVEUVSDTTVAL WEDPERIGAGGLOGI VEYGKOG KKSYRWMELNFDPYETTEAKUMI GOVYEMEVIAVNSIGMSR 30A30PFVPVAPTSEPVGLOVDI ISDTSIVLKWEDPERMGSVDLEG GOVYGKGG KSYRWMELNFDLCKTTEPKKMI GOVYEWI FAVNAIGI KFSMPSR PFVPLAVTSPFTLIVDSVDTTVIK WEDPERMGSVDLEG GOVYGKGG S88 WWELNFDLCKTTEPKKMI GOVYEWI FAVNAIGI KFSMPSR PFVPLAVTSPTLIVDSVDTTVIK WEDPERMGSVDLEG GOVYEYGKGG Motif 7 cont. * Motif 8 * PANTEPVERGSTVKLPTCAKILF VVGAINNAGPGGAPVTTEPTVOEI LOPPKILROTTI KVGGLALVVPFGG FRAPOVWTKGGAPLDTSR 886 AALGGLEHRSILVKDLPTCAKILF VVGAINNAGPGGAPVTTEPTVOEI LOPPKILPTI KVGGLALVVPFGG FRAPOVWTKGGAPLDTSR 816 PALGGLEHRSILVKDLPTGAKILF VRAINNAGPGGPVVKEPVVQEI LOPPKILPTI KVGGLALVVPFGG FRAPOVWTKEGOPLAGE 1018 PALGGLEHRSILVKDLPTGAKILF VVGAINNAGPGGPVVKEPVVQEI LOPPKILPTI KVGGVKILPFGG FRAPOVWTKEGOPLAGE 1014 PALGGLEHRSILVKDLPTGAKILF VVGAINNAGPGGPVVKEPVVQEI LOPPKILVPHLKQTI KVGKTININTFGG FRAPKSISWKDGGTLSSK 1016 PALGGLEHRSLIKDLVTCOLVFVVAINNAGPSGAPATIKEPVVQEI LOPPKILVTYHLKQTVKVGKTININTFFGG FRAPKSISWKDGGTLSSK 1016
IG-21* PURTANIADIDITATICITIONETRATIVAVAANAAGASERIISOPILVAEILEPPETKIPKILAUTIIKEVGEAVALVIPPOGEPEPEDIWKADGAEI	IKS RWMELNEDLIKELTTEAKUMI GUVYEMI IAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDUSDTTVALKWEPPERIGAGGLOGI UVYCKOG 911 KKS RWMELNEDLYPETTYEAKUMI GUVYEMI IAVNSIGMSRPSPASOPEMPIADPSEPUGLCVDI SDTS ULKWEPPERMGSVDLEG GVYCKEGT 858 KQSSRWMELNEDLCKETFEPKMI GUVYEVI FAVNAIGI KSPSMPSRPFVPLAVTSEPULLTUDSVDTTVTMRWEPPHIGAAGLOG VLEVEFEGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGEFUKNLPTGARILEFUVGVNIAGRTEPATLAOPVTIREIAEPHIRLPHILQTI RVGEQLNLVVPFQKPRPQVWTKEGOPLAGES PALGGITERTSILVKDLPTGARILF VKARNVAGPGOPUTTEPTVDVRILOPHIRQTICKVGEPVNLLIPFQKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLPTGARILF VKARNVAGPGOPUTTEPTVVRILQPHIRQTICKVGETVNLIPFQKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLFTGARILF FVARNVAGPGGPINKE PVTVQEILOPHILQTICKVGETVNLIPFQKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLFTGARLFFVARNVAGPGGPINKE PVTVQEILOPHILQTICKVGETVNLAIPFQKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLFTGARLFFVARNVAGPGGPINKE PVTVQEILOPHILQTICKVGETVNLAIPFQKPRPQVWTKEGOPLAGES 1014 PAUGGITDTSILINLTTGDLIFFVARNAGPSPARTAOPTIREIMORPHICOPHILDPILKQTICKKVGETVNLAIPFQKPRPKVTKEGOPLAGES 104 PAUGGITDTSILINLTTGDLIOFVARNAGPSPARTAOPTIREIMORPHICOPHILPPILKQTIKKVGETVNLAIPFQKPRPKVTWSNGGFLSKS 104 PAUGGITDTSILINLTTGDLIOFVARNAGPSPARTAOPTIREIMORPHICOFVARNAGESPARTACOPTILS 104 PAUGGITDTSILINNLTGDLIOFVARNAGPSPARTAOPTIREIMORPHICOFVARNAGESPARTACOPTILS 104 PAUGGITDTSILINNLTGDRIGTVARNAGPSPARTAOPTIREIMORPHICOFVARNAGESPARTAOPTIREIMORPHILDFURVENT 104 PAUGGITDTSILINNLTGDRIGTVARNAGPSPARTAOPTIREIMORPHILDFURVENT 104 PAUGGITDTSILINNLTGDRIGTVARNAGPSPARTAOPTIREIMORPHICOFVARNAGENLSSTF 963
	IKSYRWMRLNFDLIKELTTEAKUMIEGVATEMETIAVNSIGMSRPSPASOPYNPIADPSEPTHFTVEUVSDTVALKWEDPERIGAGGLOGIVEYGKOG 911 KKSYRWMRLNFDPYPTTTEAKUMIEGVATEMEVYAVNSIGMSRHSQASOPFVPVAPTSEVGLCVDDISDTSIVLKWEPPERMGSVDLEGGVEYGKEGF 858 KQSSRWMELNFDLCKSTTFEPKKMIEGVATEVIFAVNAIGISKPSMPERPFVPLAVTSPPTLTVDSVDTVTTWRPPERMGSVDLEGGVEYGKEGF Motif 7 cont. * Motif 8 * PANTEFVERGEFTVNLPTGARILFFVVGNIAGRTEPATLAOPVTIREIAPPKIRLPPHILAOTIGKVGEVNLLIPFOGFPRQVVWTGGAPLDTSR 886 AALQGLTEHTSILVDLPTGARILFFVVGNIAGRTEPATLAOPVTIREIAPPKIRLPPHILAOTIGKVGEVNLLIPFOGFPRQVVWTGGAPLDTSR 886 AALQGLTEHTSILVDLPTGARILFFVVGANNAGPGAPVTTEPVTVOEILORPHLAOTIGKVGEVNLLIPFOGFPRQVVWTGGAPLDTSR 940 940 940 940 940 940 940 940
Motif 8 cont. Motif 9	IKSS RWMELNEDILKELTTEAKUMI GUVYEMEI LAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDUSDTTVALKWEPPERIGAGGLOG IVEYGKOG 911 KKSYRWMELNEDILKELTTEAKUMI GUVYEMEI LAVNSIGMSRBSOPEVPUAPTSEPUGLCVDI SDTSIVLKWEPPERMGSVDLEG GVEYGKGG 911 KSSRWMELNEDILKETFEPKMI GUVYEMI FAVNAIGISKPSMPSRPFVPLAVTSPTLTVDSVDTVDKWEPPERMGSVDLEG GVEYGKGG 761 Motif 7 cont. * Motif 8 * PANTEPVERGETVENLPTGARILFEVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLROTIVKVGEVILLIPFOGEPRPOVWVTKGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARILFEVUGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLROTIGKVGEPVNLLIPFOGEPRPOVWVTKGGAPLDTSR 886 PALQGLTEHTSILVKDLPTGARILFEVUGVNIAGRTEPATLAOPVTIREIVUGVIGI LOPPKICAPHLROTICKVGEVNLLIPFOGEPRPOVWVTKGGAPLDTSR 886 PALQGLTEHTSILVKDLPTGARILFEVUGANAUGPGAPUTTEPVTVOEILOPPKICAPHLROTICKVGEVNLLIPFOGEPRPOVTWTKEGOPLAGEE 1018 PALQGLTEHTSILVKDLPTGARILFEVUGANAUGPGAPUTKEPVTVOEILOPPKICAPHLROTICKVGETVNLAIPPGGEPROVTWTKEGOPLAGEE 1018 PALQGLTEHTSILVKDLPTGARILFEVUGANAUGPGAPUTKEPVTVOEILOPPKICAPHLROTICKVGETVNLAIPPGGEPROVTWKEGOPLAGEE 1014 PALQGLTEHTSILIKDLFTGARLFEVUGANAUGPGPITKEPVTVOEILOPPKICAPHLROTICKVGETVNLAIPFGGEPRAGEPSTEF 963 VANKDLIDKTKTITGLPTDAKIFFVGANAAGASEPSYSSOPILVKEITIFPKIROPPKIRLPPKIRVGTVNLAIPFGGPRAGEPSTEF 963
	KKSERWARLAFDLIKELTTEAKUMIEGVAVEMETIAVNSIGMSRESPASOPYNPIADPSEPTHFTVEUVSDTTVALKWEDPERIGAGGLOGIVEYGKUGS KKSERWARLAFDPYPTTTEAKUMIEGVAVEWEVIAVNSIGMSRESPASOPYNPIADPSEPTHFTVEUVSDTTVALKWEDPERIGAGALOGIVEYGKUGS 858 KQSERWARLAFDLIKETTEPKIMIEGVAVEVIFAVNAIGISKPSMPBRPFVPLAVTSPTLLTVDSVDTTVIKKWEDPERIGAGLOGIVEYGKGT Motif 7 cont. * Motif 8 * FANTERVEIGGFTVKALPTGARILFRVGVINIGREFENIAOPVTIKEIASPHTIRPHILGYTIKVGGCLAUVUPGGFPRDQVVNTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGREFENIAOPVTIKEIASPHTIRPHILGYTIKVGGCLAUVUPGGFPRDQVVNTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGREFENIAOPVTIKEIASPHTIRPHILGYTIKVGFQLAUVUPGGFPRDQVVNTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGGEOFVTKEVGPVILQUPALLOTIKVGFVILLIPFOGPPRDQVVNTGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVGVINIGGEOFVTKEVGPVIVQEIGOPRICUPALLOTIKVGFVNLLIPFOGPPRDVVNTEGOPLAGEE 1018 PALOGITERSMLVKOLPTGARLIFRVGVINIGGEOFVTKEVGPVIVQEIGOPRICUPALLOTIKVGFVNLLIPFOGPPRDVIVTKEGOPLAGEE 1018 PALOGITERSMLVKOLPTGARLIFRVKANNAGPGAPATIAOPVTIKEIMOOPKILPPULLPUTIKKVGTVKIAIPFOGPPKILSPCS 1016 PAVQGITDATSILIRNITTODKIPTVKAVNAGPSPAATLAOPVTIKEIMOPPIIKEIMOPPIRILPPURVUTFIGGPPKKISMKDGGTLDSKD 1016 PAVQGITDATSILIRNITTODKIPTVKAVNAGPSPAATLAOPVTIKEIMOPPIRIPPKRIPPIRIKVGTVIRAIPPOGKPRKVTWSNGEPLSSTF 963 VANKDLIDTKKFTIGLPTDAKIFVVKAVNAGASEPYTYSEIDPPKRIPPKRIPPKRIPPKRIPPKRIPPELTKKDGAELDKNO 866
	IKSY RWMELNFDLIKELTTEAKUMI GOVYEME ILVNSIGMSRESPASOPINPI APSEPTIFTYEOVSDITVAL WEDPERIGAGGLOGIVEYGKOG KKSYRWMELNFDPYPTTYEAKUMI GOVYEMEVAVNSIGMSRESOASOPIVPVAPISEVGLOVDI ISDISIVLKWEDPERIGSVDLEG GOVYGKGG S88 KOSSRWMELNFDLCKSTTEPKKMI GOVYEWEVAVNSIGMSRESOASOPIVPVAPISEVGLOVDI ISDISIVLKWEPPERIGSVDLEG GOVYGKGG Motif 7 cont. * Motif 8 * PANTERVEGGETVNLPTGARLEFVVGVNIAGRTEPATLAOPVTIREIASPPILAVOTSPPILAVDSVDTVJKWEPPERIGSOVUKTGGAPLDTSR 886 AALGGIEHENSILVKOLPTGARLEFVVGVNIAGRTEPATLAOPVTIREIASPPILEOPHILROTIKKVGGLAUVVPFGGFPEROVWTKGGAPLDTSR 886 PALOGLEHENSILVKOLPTGARLEFVVGVNIAGRTEPATLAOPVTIREIASPPILAOPVILEFVGVICKVGGLAVLVVPFGGFPEROVWTKGGAPLDTSR 886 941 GGITERSALVKOLPTGARLEFVVGVNIAGRGAPVTTFEUTVOEILOPHILROTICKVGGTVNLLIPFGGFPEROVWTKGGAPLDTSR 940 GGITERSALVKOLPTGARLEFVVGANNAGGSAPATTEPOTVOEILOPHILOTICKVGTVGTVGTVGFPARLESPPILS 940 GGITERSLIKULVTCDKLYFVVGANNAGSSGAAIIKEPVTVOEIMGPPITERUGPHLAQTIKKVGTININIPFGGFPERSTSSMEDGGTLSKD 1016 940 GGITERTSLIKULTCDKLYFVVGANNAGSSGAAIIKEPVTVOEIMGPPITREINGPPIREVENUGTINKVGGTVNLATFGGFPERSTSST VANKDLIDVTKFTIGLPTDAKIFVVGANNAGASEPYYSOPILVKEITEPPIREINGPIREVGTIRKVGGANLVVPFGGFPELTWKSDGALSSKD 1016 MOTIF 8 cont. MOTIF 9 1016 10 cont. MOTIF 9 1017 101
hu-fst VHVETSDFDTVFFVEQAAESDSGEVELSVQIENMKDTATIRIRVVEEAGPPINVMVKEVWGTNALVEWQAPKDDGNSEIMGWFVOKADKKTMEWFNVVERN	IKSS RWMRLNEDLIKELTTEAKUMIEGVVTEMIIAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKW PPERIGAGGLOG IVEYGKOG 911 KKSYRWMRLNEDPYPETTYEAKUMIEGVAYEWRIYAVNSIGMSRPSPASOPEMPIADPSEPGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKOG 911 KKSYRWMRLNEDLIKETFEPKIMIEGVAYEWRIFAVNAIGISKPSMPSRPFVPLAVTSPTLTVDSVDTTVTMRWPPHIGAAGLOG VLEYGKOG 761 Motif 7 cont. * Motif 8 * PANTEPVERGEFTVENLPTGARILFEVVGVNIAGRTEPATLAOPVTIREIAEPFNIRLPAHLNQTIKVGEVNLLPFOGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLPTGARILFEVRAHNMAGPGAPUTTEPTVORILOPPLIQLPHLRQTIKKVGEVNLLIPFOGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLPTGARLLFEVRAHNMAGPGAPUTTEPTVORILOPPLQLPHLRQTIKKVGEVNLLIPFOGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLFTGARLLFEVRAHNMAGPGAPUTTEPTVORILOPPLQLPHLRQTICKVGEVNLLIPFOGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILVKDLFTGARLLFEVRAHNAGPGAPUTKEPTVORILOPPLQLPHLRQTICKVGEVNLLIPFOGKPRPQVWTKEGOPLAGES 1018 PALGGITERTSILIKNLTTGDRLOFFVANYAMAGPSPAATLAOPVTIREIMONPHILLPPINLQTILKVGETINIMIPFOGKPRPUTWKEGOPLAGES 1014 PALVGLTDTSLIINNLTTGDRLOFFVANYAMAGPSPAATLAOPVTIREIMONPHILPPINLPUTIKEVGETINIMIPFOGKPRPELTWKKDGAELDKNO 866 MOTIF 8 cont. MOTIF 8 cont. NOTIF 9 TSDFDVYFFVGAALSDSGEVELSVGIENMKDTATIRIKVEKAGPPINVKEVWGTNALVEWOAPKDORSEIMGFVORADKKTMEWFNVERNHTS 991
hu-fst VHVETSDFDTVFFVEQAAESDSGEVELSVQIENMKDTATIRIRVVERAGPPINVMVKEVWGTNALVEWOAPKDDGNSEIMGEFVOKADKKTMEWFNVERN	IKSS RWMRLNEDDLKELTTEAKUMIEGVVTEMIIAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKW PPERIGAGGLDG IVEYGKDGS 911 KKSYRWMRLNEDDYPETTYEAKUMIEGVVTEMRIVAVNSIGMSRPSPASOPEMPIADPSEVGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKGG 911 KKSYRWMRLNEDDLKETFEPKMIEGVAREVVIAVNSIGMSRBSOPEVPAPTSEVGLCVDISDTSIVLKW PPERMGSVDLEG GVEYGKGG 761 Motif 7 cont. * Motif 8 * PANTEPVERGEFTVENLPTGARILFEVVGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLNGTIGKVGEPVNLLIPFGGRPAPQVVWTKGGAPLDTSR 886 FAALGGITEHTSILVKDLPTGARILFEVTAHNNAGPGAPVTTEPTVORILGPRLQIPHLNGTIGKVGEPVNLLIPFGGRPAPQVWTKEGGPLAGEE 1018 PALGGITEHTSILVKDLPTGARLLFEVTAHNNAGPGAPVTTEPTVORILGPRLQIPHLNGTIGKVGEPVNLLIPFGGRPAPQVWTKEGGPLAGEE 1018 PALGGITEHTSILVKDLPTGARLLFEVTAHNAGPGAPVTTEPTTVGEILGPRLQIPHLNGTIGKVGEPVNLLIPFGGRPAPQVWTKEGGPLAGEE 1018 PALGGITEHTSILVKDLPTGARLLFEVTAHNAGPGAPVTKEPVTVQEILGPRLQPRLQTICKVGETNINAIPFGGRPAPGVWTKEGGPLAGEE 1014 PALGGITEHTSILIKNLTTGDKLQFFVTANINAGPSPAATLAOPVTIREIMOSPHILLPPNLRQTIKVGETNILAIPFGGRPAPKTENSMIEGGLESKEN VANKDLIDKIKFTVAINNAGPSPAATLAOPVTIREIMOSPHILPPNIRIPHLKQTIKKVGETNINAIPFGGRPAPELTWKKDGAEIDKNQ 866 MOTIF 8 cont. MOTIF 9 10500000000000000000000000000000000000
hu-fst WHYRTSDEDTVFFV ROAASDSGEVELSVOLENMKDTATIRIRUV RAGPPINUMUKEUWGTNALVEWCAPROBONSEIMGEVEVENINUMERN	IKS TRAMELINFOLIKELTTEAKUMI GOVIEMETIAVNSIGMSRESASOPINDIADESEDTIFTVEOVSDITVAL AMEDDEE IGAGGLOGI VEYGKOG KKSTRAMELINFDILKELTTEAKUMI GOVIEMEVIAVNSIGMSRESOASOPIVDIADESEDTIFTVEOVSDITVAL AMEDDEE IGAGGLOGI VEYGKOG KSTRAMELINFDILKEITTE PKAMI GOVIEMEVIAVNSIGMSRESOASOPIVDIADESEDTIFTVEOVSDITVAL AMEDDEE IGAGGLOGI VEYGKOG SSRAMELINFDILKEITTE PKAMI GOVIEMEVIAVNSIGMSRESOASOPIVDIAVISPPTLLVIDSVDITVIKAMEPPEMGSVDLEG GOVIEVGEG Motif 7 cont. * Motif 8 * PANTERVEGGETVKLEPTGARLEF VVGVNIAGREPATLAOPVIERIASPPTILEVESYDDILKTOSVDITVIKAMEPPEMGSVDLEG GOVIEVGEG PALOGIEFHSSILVGLEPTGARLEF VRAHNAGPGGAPVITTEPTVOEILOOPHILOTIOK VGEPVNLLEPGGEPAPQVWITGGAPLDTSR PALOGIEFHSSILVGLEPTGARLEF VRAHNAGPGGPIVTKEPVVQEILOOPHILOTIOK VGEPVNLLIPFOGRPAPQVWITGGAPLDTSR PALOGIEFRSLIVGLEPTGARLEF VRAHNAGPGGPVVKEPVVQEILOOPHILOTIOK VGEPVNLLIPFOGKPAPQVWITGGAPLDTSR PALOGIEFRSLIVGLEPTGARLEF VRAHNAGPGGPVVKEPVVQEILOOPHILQTIOK VGEPVNLLIPFOGKPAPQVKITEGOPLAGE 1014 PALOGIEFRSLIVGLEPTGARLEF VRAHNAGPGGPVVKEPVVQEILOOPHILQTIOKVGTVKGEPVNLLIPFOGKPAPKISSING PALOGIEFRSLIVGLEPTGARLEF VRAHNAGPGAPVTERIMOPVIERIMOPHILOTIOKVGEPVNLLIPFOGKPAPKISSING PALOGIEFRSLIVGLEPTGARLEF VRAHNAGPGAPVTERIMOPVIERIMOPVIERIMOPVIENG PALOGIEFRSLIVGLEPTGARLEF VRAHNAGPGAPVIERIMOPVIERIMOPVIERIMOPVIENG PALOGIEFSIIINILTODILOVGUVAINAGASEPYYSOPILVKEITEPVINGEPVIERIMOPTIREWGENVENGEN VANKDLIDTTGLEOFTGARLEFVKAINAGBGAPVITERIMOPVIERIMOPVIERIMOPVIERIMOPTIREWGENVENGEN MOTIF 8 cont. MOTIF 9 MOTIF 8 cont. MOTIF 9 MOTIF 8 cont. MOTIF 9 MOTIF 8 CONT. MOTIF MOTIF 9 MOTIF 8 CONT. MOTIF MOTIF 9 MOTIF 9
	KKS RWARLNFDLIKELTTEAKUMIEGVAVEMEILAVNSIGMSRESPASOPYNPIADESEPTHFTVEUVSDTVALKWEDPERIGAGGLOGIVEYGKUGS KKSRWARLNFDPYETTYEAKUMIEGVAVEMEVAVNSIGMSRESPASOPYNPADESEVGLOVDISDTSVLKWEDPERMGSVDLEGGVEYGKGT 858 KQSSRWARLNFDLCKETTEPKMIEGVAVEVIFAVNAIGISKPSMPBRPFVPLAVTSPTLLTVDSVDTTVIKWEPPHIGAGGLOGIVEYGKGT Motif 7 cont. * Motif 8 * PANTERVEIGGFTVNLPTGARLLFRVGAUNIGREFPATLAOPVTIERIASPMTIEDHLEGYTIKVGGCLUVVPGGFPHQVVWTGGAPLDTSR AALQGLTEHTSILVKDLPTGARLLFRVGAUNAGPGGPTVTKEVVOELOGPRLQLPRHLRGTICKVGEVNLLIPFGGKPHQVVWTGGAPLDTSR 886 AALQGLTEHTSILVKDLPTGARLLFRVGAUNAGPGGPTVTKEVTVOELOGPRLQLPRHLRGTICKVGEVNLLIPFGGKPHQVVWTGGAPLDTSR 846 PALGGLTEHTSILVKDLPTGARLLFRVFAUNAGPGGPTVTKEVTVOELOGPRLQLPRHLRGTICKVGEVNLLIPFGGKPHQVWTKEGGPLAGEE 1018 PALGGLTEHTSILVKDLYTGARLFRVFAUNAGPGGPTVTKEVTVOELOGPRLQPRLVPHLRGTICKVGEVNLLIPFGGKPHQVWTKEGGPLAGEE 1018 PALGGLTEHTSILVKDLYTGARLFRVFAUNAGPGGPTVTKEIMOPTIERIMS PAUGGLTDETSLINDLYTGARLFRVFAUNAGPGGPTVTKEIGPPTVOELOGPRLCVFHLRGTICKVGETVILIPFGGKPHQVKTKEGGPLAGEE 1019 PAUGGLTDETSLINDLYTGARLFFVKAUNAGPGFPATLAOPVTIERIMGPFILDENLAGTICKVGETVILATPFGGKPHQVKTKEGGTLDSKD 1016 PAVGGLTDKTSLINNLTGDRLFFVKAUNAGPSPATLAOPVTIERIMGPFILDEPNIRIFYFICKVGTVNLATPFGGKPHQVKTKSNGGTLDSKT 963 WANKDLIDTKKFTITGLPTDAKIFVKAUNAGAGSFPATLAOPVTIERIMPFNRIPPNIRIFYFIKFKGEAVNILVFFGGKPHQVKTKSNGGLEDKNG 866
IN SIN SUNTAINDIDAINE I I OUF IDANIE VERVINAGABERI OUF IDANIE FRI AFRIDAVILI ANVOLAVADI I FEURALINA DOAD	IKS RWMELNEDLIKELTTEAKUMI GOVYEMI IAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKWEPPERIGAGGLOGI VEYGKOG KKS RWMELNEDPYPETTYEAKUMI GOVYEMI VAVNSIGMSRBSORSOPEVPYAPTSEVGLOVDI SDTSIVLKWEPPERIGAGGLOGI VEYGKOG SSRWMELNEDLIKETFEPKIMI GOVYEWI FAVNAIGI KSPSMPSRPFVPLAVTSEVGLOVDI SDTSIVLKWEPPERIGAGLOGI VEYGKOG Motif 7 cont. * Motif 8 * PANTEPVERGETVENLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPPI IRVGEQLNLVVPFQGEPROVVTKEGAPLDTSR SR AALQGITERTSIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPPI IRVGEQLNLVVPFQGEPROVVTKEGAPLDTSR SR AALQGITERTSIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPPI IRVGEVGVICKVGEVNILLPFQGEPROVVTKEGAPLDTSR SR AALQGITERTSIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPPI IRVGEVGVICKVGEVNILLPFQGEPROVVTKEGAPLDTSR SR AALQGITERTSIVKOLPTGARILE VVGVNIAGREEPATLAOPVTIREIAEPPI IRVGEVGVICKVGEVNILLPFQGEPROVVTKEGAPLAGE 1018 PALGGITERTSIVKOLPTGARLEFVKAINNAGEGAPUTKEPVTVQEILGPHLOPHILAOITOK VGEVNILLPFQGEPROVVTKEGAPLAGE 1018 PALGGITERTSIIVKOLVTGOKLEFVKAINLAGESGAAIIKEPVTVQEILGPHLOPHILAOITOK VGEVNILLPFQGEPREVIKKEGGILGSE 1016 PAVQGITDTSIIINLTTGDLOFVKAINNAGESPANTLAOPVTIREIMORPHILPNLKOITIKKVGEVNILAPFQGEPREVIKKEGESTF 963 VANKELTVENETETTALEVKAINAGESGEPUTKEDI VKOPTI KEINGEPI VVOETU ILPNLKOTI KEVGTVALATEPGGEPREVIKTEGOPI SETF
fu-cc EDWLPAVQGLTDRTSLIIRNLTTGDKLQFRVRAYNMAGPSPAATLAQPVTIREIMQRPKILLPRNLRQTLIRRVGDTVNLAIPFQGKPRPKVTWSKNGEPI	IKS RWARLAFDLIKELTTEAKUMI EGVATEMETIAVASIGASRESPASOPENPIADESEPTIFTEOUSDITVALKAFDPERIGAGGLOGIVENDGS 911 KKS RWARLAFDPYPTITEAKUMI EGVATEMEVVAVASIGASRESPASOPENPIADESEPTIFTEOUSDITVALKAFDPERMGSVDLEGGVENDGS 858 KOSS RWARLAFDLICKTIFEPKMI EGVATEV IFAVANIGISKESMESPEPPIADISSPULAVINGSVDITVALKAFPERMGSVDLEGGVENDGF Motif 7 cont. * Motif 8 * PANTERVENGEFTWALLFRUKGENTEFTARTAFTERTARDVTIKEIASPEPTIEDAFTERTARGEOLINUVFROKERROVUNTEGGALDISS 886 AALQGITENTSILVKOLFTGARLIFFVRAHNAGEGAPVITEPVIVORILGERRIASPENTERHEIASPENTERHEIASTIKVGEOLINUVFROKERROVUNTEGGALDISS 886 AALQGITENTSILVKOLFTGARLIFFVRAHNAGEGAPVITEPVIVORILGERRIASPENTERHEIASPENTERHEIASTIKVGEOLINUVFROKERROVUNTEGGALDISS 886 AALQGITENTSILVKOLFTGARLIFFVRAHNAGEGAPVITEPVIVORILGERRIASPENTERHENDIGTIGKVGEPVALLIPFGSERROVUNTEGGALDISS 886 AALQGITENTSILVKOLFTGARLIFFVRAHNAGEGAPVITEPVIVORILGERRIASPENTERHENDIGTIKVGEVANDISTIKTEGOLINUVFROGENTERS 886 AALQGITENTSILVKOLFTGARLIFFVRAHNAGEGAPVITEPVIVORILGERRIASI 940 940 940 940 940 940 940 940
CD-CC REWTPALPGETERTSALIRDEVTGDREMTRVKAINLAGESGAAIIKEPVTVOEIMORPRICVPRHEROTEVREVGETINIMIPFOGRPRPKISWMRDGOTE	IKS TRAMELNFOLLKELTTEA KUMI GOVYEME ILAVNSIGMSR 959ASOPFMPIAPSS 57HFTVEOVSDTTVAL AM PPPERIGAGGLOGIVSTOKOGS 911 GKSTRAMELNFOPYETTEAKEMI EGVATEMEVEAVNSIGMSR 30ASOPFVPAAPSS FVGLOVDIISDTSIVLKWE PPERMGSVDLEG GVEYGKEGT 858 KOSSRAMELNFOLCKITTE PKAMI EGVATEV IFAVNAIGISK 9MPSR 9FVPLAVIS 9FVLLVDSVIDTTVIK AM PPPHIGAAGLOGIVLSYGFEGT 761 Motif 7 cont. * Motif 8 * PANTEPVE GGTTVKLPTGARLLF VVGVNIAGR TEPATLAOPVTIREIASPPKIRLPTKIRVGGLAUVVPFGGK PRPQVWTKGGAPLDTSR 886 AALGGITEHTSILVKOLPTGARLLF VVGANNAGGGAPUTTEPTVVDEILGPKILGPKILKOTIGK VGEPVNLLIPFGGK PRPQVWTKGGAPLDTSR 816 PALGGITEHRSILVKOLPTGARLLF VRANNAGGGAPUTTEPTVVDEILGPKILOPTIGNING VGEPVNLLIPFGGK PRPQVWTKGGAPLDTSR 1014
ING-CC DENTERLUGITERKSELVEDETGARDER VRAHNVAGPGGPTVTKEPVTVDETLORPRIOTIOREVGEPVNLLTPFOGEDEDAVTWTKEGOPT	IKSS RWMELNEDILKELTTEAKUMI GUVTEMEILAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKWEPPERIGAGGLDG IVEYGKDGS 911 IXKSYRWMELNEDIYPETTEAKUMI GUVTEMEILAVNSIGMSRHSOASOPEVPUAPTSEPVGLCVDI SDTSIVLKWEPPERMGSVDLEG GVEYGKGG 858 IVSSRWMELNEDICKETFEPKMI GUVTEV IFAVNAIGISKPSMPSRPFVLAVTSPTLLTVDSVDTVTMRWEPPHIGAAGLDG VLEYGFGT 761 Motif 7 cont. * Motif 8 * PANTEPVERGETVENLPTGARILFRVVGVNIAGREPATLAOPVTIREIAEPPKIRLPRHLRQTYIRKVGSQLNLVVPFQG PRPQVVWKGGAPLDTSR 886 AALQGITEHTSILVKOLPTGARILFRVAGNDAGGAPUTTEPVTVOEILORPHIGALDEPHIRQTSIKVGFVERULTPFQG PRPQVTWKGGAPLDTSR 886
THE TAXABLE IN TAXABLE INTENTING INTENTING INTENTING INTENTION TAXABLE INTENTION INTENTION IN TAXABLE INTENTING INTENTION INTENTION INTENTING INTENTION INTENTIA INTENTIA INTENTIA INTENTIA INTENTIA INTENTIA INTENTIA INTEN	IKSTRUMELNFOLIKELTTEAKUMIEGVNYEMEIIAVNSIGMSR <mark>959ASQPFWPIAPSSPTHFIVE</mark> DVSDTTVALKWE PPER IGAGGLOGIVEYGKOG SANNELNFDPYETTEAKUMIEGVNYEMEVAVNSIGMSRHSQASQPFVPAPISPVGLCVDDISDTSIVLKWEPPERMGSVDLEG GVYGKEGT SSSRWMELNFOLCKITTEPKKMIEGVNYEVRIFAVNAIGISKPSMPSRPFVPLAVISPPTLLTVDSVDTVJKWEPPERMGSVDLEG GVYGKGT Motif 7 cont. * Motif 8 * PANTEFVERGETVNLPTGARLEFVVGVNIAGRTEPATLAOPVTIEEIASPPTIREPMEROTOFVCVGCLAUVVPFGCKPEPQVVVTKGGAPLDTSR 886 ALOGIEHISILVKDLFTGARLEFVVGVNIAGRTEPATLAOPVTIEEIASPPTIREPMEROTOFVCVGCLAUVVPFGCKPEPQVVVTKGGAPLDTSR 886
THE REAL PROPERTY AND A DRAWN AND A	KKSTRWARLNEDLLKELTYEAKEMIEGVVYEMIIAVNSIGMSRPSPASOPEMPIADPSEPTHETVEDVSDTTVALKWEPPERIGAGGLDG IVEYGKDGS 911 KKSTRWARLNEDPIPETTEEAKEMIEGVAYEWIIAVNSIGMSRHSOASOPEVPVAPTSEPVGLCVDDISDTSIVLKWEPPERMGSVDLEGGVEYGKGGT 858 KQSSRWARLNEDLCKETTFEPKMIEGVAYEVIFAVNAIGISKPSMPSRPFVPLAVTSPPTLLTVDSVDTVVMRWEPPHIGAAGLDGVLEYGFGGT Motif 7 cont. * Motif 8 * PANTEPVERCGTVKNLPTGARLLFRVVGVNIAGRTEPATLAOPVTIREIAEPPKIRLPRHLROTYIRKVGQLALVVPFQCKPRPQVVWTKGGAPLDTSR 886
IN TO PRIME A DECORTANIA A DAMANGA DE	KKSTRWMRLNFDLLKELTYEAKEMIEGVVTEMEIAVNSIGMSRPSPASOPFMPIAPPSEPTHFTVEDVSDTTVALKWPPERIGAGGLDG IVEYGKDGS 911 KKSYRWMRLNFDEYPETTYEAKEMIEGVAYEM VVAVNSIGMSRHSOASOPFVPVAPTSEPVGLCVDDISDTSIVLKWPPERIGAGGLDG IVEYGKEGT 858 KOSSRWMRLNFDLCKETTFEPKMIEGVAYEV IFAVNAIGISKPSMPSRPFVPLAVTSPPTLLTVDSVTDTTVTMRWPPDHIGAGLDG VLEYGFEGT 761 Motif 7 cont. * Motif 8 * ADNTEDVEGTVUNUFGABLIEGVVGUNUGEFERATIADVTSETAEDEVFDIDEVLEVENTEVENTEVENTEVENTEVENTEVENTEVENTEV

Region (Motif)	Mutation	Phenotype	Reference
0	T59A	Mild, elderly onset	[58]
1	D228N	Rare form of midventricular hypertrophy	[88]
1	Y237S	Mild hypertrophy	[42]
Linker 1-2	H257P	No specific information	[27]
Linker 1-2	E258K	Delayed onset	[26, 70]
Linker 1-2	G278E	No specific information	[27]
Linker 1-2	G279A	No specific information	[27]
Linker 1-2	R326Q	Elderly onset. Also seen in healthy controls [54, 55]	[42, 52, 53, 55, 56]
Linker 1-2	L352P	No specific information	[27]
2	E451Q	Delayed onset, incomplete penetrance, long life expectancy	[26]
3	R495Q	Delayed onset, incomplete penetrance, long life expectancy	[26, 52]
3	R502Q	Delayed onset, incomplete penetrance, long life expectancy	[26, 27]
	R502W		
5	R654H	Mild hypertrophy, elderly onset	[89]
5	R668H	Death of one son?	[42]
5	N755R	Severe phenotype	[64]
6	R810H	NYHA II - homozygote	[70]
6	K811R	No specific information	[27]
6	R820Q	Burnt out phase in elderly-left ventricular dysfunction/dilation	[70, 71]
6	A833V	Mild hypertrophy. Lacks definite proof of being disease causing	[27, 42]
	A833T		
7	Р873Н	Mild / moderate	[70]
7	V896M	May be disease causing or modifying [27, 54]	[42, 78]
7	N948T	Dilated, severe	[56]
8	R1002Q	Elderly onset mild	[58]
10	A1194T	No specific information	[27]
10	1253 Insert	Early onset, one instance of sudden death.	[77]
10	A1255T	No specific information	[27]

Tab 1. Summary of FHC point mutations in cardiac MyBPC

Fig 4. Multiple alignments of MyBPC isoforms from different species. hu = human, mo = mouse, ch = chick, fu = fugu, fst = fast skeletal, cc = cardiac, slw = slow skeletal. The alignment is coloured where 60% of the sequences have the same residue or the residues have similar properties (with the exception of proline and glycine which are always coloured yellow and grey, respectively); cyan = hydrophobic, red = positive, violet = negative, green = hydrophilic, blue = aromatic. Asterix above the sequence indicate residues mutated in the human MyBPC isoform that result in FHC. The alignment was built using the program CLUSTAL X[11].

caused by the expression of abnormal contractile proteins in the heart muscle. To date, mutations in ten sarcomeric proteins have been identified as causes of FHC[20]. The mechanism by which MyBPC mutations cause sarcomeric dysfunction at present poorly understood, but has provided a substantial stimulus to efforts to understand the basic biochemistry and physiology of this important molecule.

Familial hypertrophic cardiomyopathy

FHC is a clinically and genetically heterogeneous disorder characterised macroscopically by increased left ventricular mass in the absence of any apparent loading stress, and histologically by myofibrillar and myocyte disarray and fibrosis[20]. Clinically, the extent of hypertrophy determines the level of impaired cardiac function. Associated fibrosis can lead to arrythmogenesis and sudden death. Indeed FHC is the most common cause of sudden cardiac death in young athletes[21]. The prevalence of FHC is believed to be about 0.2%, or 1 in 500[22, 23]. However despite its relatively high prevalence, the mortality rate for FHC-related deaths in an unselected population was estimated to be only 1%[24]. This mortality rate reflects the wide range of severity associated with this disease.

The cardiac MyBPC gene (*MYBPC3*) was the fourth gene identified associated with FHC[25]. Various studies put the percentage of FHC patients with cardiac MyBPC mutations at between 20-45%[26, 27], making it the second most common cause of FHC. Three types of mutations in *MYBPC3* that result in FHC have been identified; truncations, point mutations and insertions.

The majority of MYBPC3 FHC mutations generate a frame-shift in the coding sequence, which results in the premature termination of translation of the C-terminus of MyBPC. Patients with these truncated forms of MyBPC usually have a mild phenotype, delayed age of onset and favourable prognosis[26, 28], probably since the patients still have a normal copy of the MyBPC gene on the other allele. A similar phenotype is observed in mouse models of FHC with MyBPC C-terminal truncations[29-31]. Truncation mutations are thought to cause haploinsufficency, since the absence of the C-terminus of MyBPC appears to result in a failure of the mutant MyBPC to incorporate into the sarcomere[32]. However this has been disputed and Flavigny et al. proposed that the truncated MyBPC acts via a dominant negative mechanism, possibly as a poison peptide[33].

In contrast, a number of missense mutations in *MYBPC3* have been identified, some of which result in a severe FHC phenotype. Investigations of the consequences of some of these point mutations have been particularly informative in defining the structure and function of the domain in which they are found. Tab 1 lists the current FHC-causing point mutations identified in MyBPC. The precise mechanism by which many of these mutations cause FHC remains unsolved, but in some cases the degree to which these mutations may affect the structure and function of MyBPC can be inferred from sequence comparisons.

Immunoglobulin superfamily domain structure

Sequence alignments of individual modules from a number of species and proteins have identified key conserved residues for the classification of IgI and FnIII domains (Fig 5 and 6). The immunoglobulin fold is composed of two β -sheets comprising approximately 5 anti parallel β strands in each sheet. The size and number of β -strands and the conformation of the links between each strand are used to further classify the Ig molecules into sets. The intermediate-set (I-set) of immunoglobulin domains was so named because it has characteristics of both the variable (V) and the constant (C) sets (Fig 5). In an IgI domain, one β -sheet is composed of strands A, B, E and D and the other of strands A, G, F, C and C'[34]. The alignment of all IgI domains within MyBPC again emphasises the presence of a cardiac-specific insert in Motif 5 of MyBPC. The domain boundaries of IgI domains are generally defined as beginning at a conserved proline residue and end approximately 100 amino acids later with a hydrophobic-X-hydrophobic sequence defining the C-terminus[35].

The second class of protein modules found in MyBPC are fibronectin type III domains (FnIII). Similar to the IgI domain, FnIII is also composed of two β -sheets that fold into a β -sandwich, although FnIII conforms to an s-type topology, with one sheet containing strands A, B and E and the other C, C', F and G[36, 37] (Fig 7). FnIII motifs are uniquely proline rich and, like most globular proteins, fold to create a hydrophobic core. Based on the alignment of multiple FnIII modules, the N-terminal boundary is defined by a PXPP motif but the C-terminal boundary is more difficult to decipher. Generally it is positioned at the residue following the last β -strand, making the domain approximately 100 amino acids long.

Interactions between β -sandwich domains are generally thought to occur via their sheets, while interactions between these domains and their ligands are thought to occur via their loops[38]. Additionally, often two consecutive domains are involved in ligand binding[38]. One study of module-module interactions, in the FnIII rich protein fibronectin, found that the highly conserved proline residues positioned at the domain-domain boundaries may be present to prevent aggregation in this multi-modular protein[39]. In contrast, computational modelling of FnIII domains suggests that, as proline side-chains can form low-energy interfaces for protein contacts, FnIII modules may interact via the BC and EF loops[37].

Correlation between the structure of these homologous domains and the specific location of FHC mutations has allowed, in many cases, the prediction of how FHC point mutations in MyBPC may affect the structure of its motifs and thereby alter their stability, binding or function. MyBPC will be discussed in three parts; the N-terminal region containing Motifs 0-2, the central region containing Motifs 3-6 and the C-terminal region containing Motifs 7-10.

N-terminal region (Motifs 0-2)

Within Motifs 0-2 there may be up to three binding sites. The phosphorylation-dependent binding of the Motif 1-2 Cell research, 14 (2), Feb 2004

E

Shoote

Cecily E OAKLEY et al

	A	A'	в	С	Cardiac	specific	insert	C'	D	
Sheets						~		_		
hu-ti-5	SSKTTLAARILTK	-RSMTVYEG	ESARFSCDTDG	EPVPTVTW	LR		к	GQVLSTS	ARHQVTTTK	60
hu-cc-5	PRQEPPKIHLDCPGRIP	-DTIVVVAG	NKLRLDVPISG	DPAPTVIW	QKAITQGNKAPAR	PAPDAPEDTG	DSDEWVFDK	KLLCETH	GRVRVETTK	92
ch-sk-5	PKQEPPKIHLDCSGKA	ENTIVVVAG	NKVRLDVPISG	EPAPTVTW	KRG		D	QLFTATE	GRVHIDSQA	65
ch-h-2	PRIRLPRQLF	-QVYVRHVG	EAVNLLIPFQG	KPQPQVTW	T <mark>K</mark>			DNQPLD	SRVNIRNTD	55
hu-cc-8	<mark>PRLQLP</mark> RHLF	-QTIQKKVG	EPVNLLIPFQG	KPRPQVTW	T <mark>K</mark>			EGQPLAC	EEVSIRNSP	55
ch-sk-8	<mark>P</mark> KIRLPRHLF	-QTYIRRVG	EQVNLVIPFQG	KPRPQVTW	S <mark>R</mark>			EGGALP-	AEVQTRTSD	55
hu-h-2	<mark>P</mark> KIRVPRHLF	-QTYIRQVG	ETVNLQIPFQG	KPKPQATW	T <mark>H</mark>			NGHALDS	SQRVSMRTGD	55
hu-telo	PHVKPYFSKTI	-RDLEVVEG	SAARFDCKIEG	YPDPEVVW	FK		D	DQSIRES	SRHFQIDYDED	58
rb-telo	PHVKPYFSKTI	-RDLEVVEG	SAARFDCKIEG	YPDPEVVW	FK		D	DQSIRES	SRHFQIDYDED	58
tu-telo	<mark>P</mark> HVK <mark>PYFT</mark> KTI	-LDMDVVEG	SAARFDCKVEG	YPDPEVMW	F <mark>K</mark>		D	DNPVKES	SRHFQIDYDEE	58
ch-h-4	PPKFTGPI	-TDRATTRG	YSTHLFCSVRG	FPQPKIIW	MK		N	KMEIREI	PKYIAMIEQ	54
hu-h-4	PSFTQPI	-ADHTSTPG	YSTQLFCSVRA	SPKPKIIW	MK		N	KMEIQGI	PKYRALSEQ	53
hu-cc-10) <mark>PSFTQPI</mark>	-VNRSVIAG	YTAMLCCAVRG	SPKPKISW	F <mark>K</mark>		N	GLDLGEI	ARFRMFSKQ	53
ch-sk-10)PQFLTPL	-VDRSVVAG	YTVTLNCAVRG	HPKPKVTW	L <mark>K</mark>		N	SVEIGAL	PKFLSRHGL	53
hu-cc-4	KLEVYQSI	-ADLMVGAK	DOAVFKCEVS-	DENVRGVW	L <mark>K</mark>		N	GKELVPI	SRIKVSHIG	54
ch-sk-4	KQLEVLQVM	-ADLTVKAS	EQAVERCEVS-	DEKVTGRW	FR		N	GVEVKPS	KRIHISHNG	55
hu-cc-2	EKKSTAFQKKLE	-PAYQVSKG	HKIRLTVELA-	DHDAEVKW	L <mark>K</mark>		N	GQEIQMS	G-SKYIFESIG	58
ch-sk-2	PKKSEAFIRKLD	-PAYQVDKG	NKIKLVVELS-	DPDLPLKW	Y <mark>K</mark>		N	GQLLKPS	TKYVFENVG	57
hu-cc-3	PVLITRPI	-EDQLVMVG	QRVEFECEVS-1	EEGAQVKW	L <mark>K</mark>		D	GVELTRE	ETFKYRFKKDG	55
ch-sk-3	PDPEQLE	-LSKQVVVG	DRVVLEAEVS-	EEGAQVMW	L <mark>K</mark>		D	GVDVTRI	DAFKYRFKKDG	54
hu-cc-1	PIGLEVMRE	-QDGEVTVG	GSITFSARVAG	ASLLKPPVVKW	FKG		KWV	DLSSKV	QHLQLHDSYDRAS	65
ch-sk-1	PEQLFLSK	-QNVMVESG	RDVTVSARVAG	AALPCAPAVKW	FKG		KWA	ELGDKS	RCRLRHSVDDD	63
hu-cc-0	-MPEPGKKPVSAFSKK	-RSVEVAAG	SPAVFEAETER	AGVKVRW	RG			GSDISAS	NKYGLATEG	61
IgI		G						U	.+U	

hu-ti-5	YKSTFEISSVQASDEGNYSVVVENSEGKQEAEFTLTIQK	99
hu-cc-5	DRSIFTVEGAEKEDEGVYTVTVKNPVGEDQVNLTVKVIDV	132
ch-sk-5	DLSSFVIESAERSDEGRYCITVTNPVGEDSATLHVRVVD	104
ch-h-2	KDTIFFIRTAQRSDSGKYQLSVRINGAEDKAILDRIVIER	95
hu-cc-8	TDTILFIRAARRVHSGTYQVTVRIENMEDKATLVLQVVDK	95
ch-sk-8	VDSVFFIRSAARPLSGNYEMRVRIDNMEDCATLRLRVVER	94
hu-h-2	QDSILFIRSAQRSDSGRYELTVRVEDLEAKAVIDILVIEK	96
hu-telo	GNCSLIISDVCGDDAKYTCKAVNSLGEATCTAELIVETME	99
rb-telo	GNCSLIISDVCGDDAKYTRKAVNSLGEATCTAELIVETME	99
tu-telo	GNCSLTISEVCGDDDAKYTCKAVNSLGEATCTAELLVETMG	99
ch-h-4	GVCSLEIRKPSPFDAGVYTCKAVNPLGEASVDCKLDV	91
hu-h-4	GVCTLEIRKPSPFDSGVYTCKAINVLGEASVDCRLEVKASA	94
hu-cc-10	GVLTLEIRKPCPFDGGIYVCRATNLQGEARCECRLEVRVPQ	94
ch-sk-10	GVLSLLIRRPGPFDGGTYGCRAVNEMGEATTECRLDVRVPQ	94
hu-cc-4	RVHKLTIDDVTPADEADYSFVPEGFACNLSAKLHFMEVKI	93
ch-sk-4	RFHKLVIDDVRPEDEGDYTFIPDGYALSLSAKLNFLEIKVEYV	97
hu-cc-2	AKRTLTISOCSLADDAAYOCVVGGEKCSTELFV	90
ch-sk-2	LKRILTIHKCSLADDAAYECRVNDEKCFTEVFVKEPPVTVVR-	99
hu-cc-3	ORHHLIINEAMLEDAGHYALCTSGGOALRELIV	88
ch-sk-3	KKHFLIINEAELSDSAHYKIMINGGESEAELSVEEKQLEVLQD	97
hu-cc-1	KVYLFELHITDAQPAFTGSYRCEVSTKDKFECSNFNLTV	104
ch-sk-1	KVHTFELTITKVAMGDRGDYRCEVTAKEQKDSCSFSIDVEAPR	106
hu-cc-0	TRHTLTVREVGPADQGSYAVIAGSSKVKFDLKVIEAEKAE	101
IgI	U.U.U.D.G.Y.UU.U.U.U.	

F

G

Fig 5. Multiple alignments[11] of IgI domains from human (hu), chick (ch), rabbit (rb) and turkey (tu) muscle proteins including; titin (ti), cardiac MyBPC (cc), skeletal MyBPC (sk), MyBPH (h) and telokin (telo). The alignment is coloured where 60% of the sequences have the same residue or the residues have similar properties (with the exception of proline and glycine which are always coloured yellow and grey, respectively); cyan = hydrophobic, red = positive, violet = negative, green = hydrophilic, blue = aromatic. The β -strands are indicated by lines across the top of the alignment and are based on the known structure of m5 from titin[83] and Motif 5 from cardiac MyBPC[59]. The last line indicates the residues required for the classification of IgI domains. This figure highlights the cardiac-specific insert present in Motif 5 of cardiac MyBPC.



Fig 6. Multiple alignments [11] of FnIII domains from human (hu), and chick (ch) muscle proteins; titin (ti), cardiac MyBPC (cc), skeletal MyBPC (sk) and MyBPH (h). The alignment is coloured where 60% of the sequences have the same residue or the residues have similar properties (with the exception of proline and glycine which are always coloured yellow and grey, respectively); cyan = hydrophobic, red = positive, violet = negative, green = hydrophilic, blue = aromatic. The last line indicates the residues required for the classification of FnIII domains. The strands are indicated by lines across the top of the alignment and are based on the known structure of A71 from titin[37].

linker to S2 of myosin is now well established[4], however recent data have suggested that Motif 0 and the Motif 0-1 linker may bind to myosin and/or actin[5, 10, 33, 40]. The strength of the evidence for each of these interactions is variable, and the physiological consequences remain the subject of investigation.

Despite Motif 0 being unique to the cardiac isoform of MyBPC, its function and specific role in heart muscle is still unclear. There is a surprising lack of FHC-associated mutations in this motif. Only one patient with a mild FHC phenotype has been identified. This mutation lies between strands D and E in a position of low sequence conservation[26].

A possible interaction between Motif 0 and some part of the myosin crossbridge has been based on data from a mutant MyBPC knock-in mouse model[5]. In this mouse, MyBPC is missing both the linker between Motifs 0 and 1, and Motif 1, although the Motif 1-2 linker could still be phosphorylated. There was an increase in Ca²⁺ sensitivity to force production in this mutant mouse, similar to that seen previously in cardiomyocytes depleted of MyBPC [18]. These data are consistent with a low affinity interaction between Motif 0 and some part of the crossbridge, or actin (discussed below). Truncation of MyBPC in the genetically manipulated mouse was thought to prevent Motif 0 from reaching out to interact with the myosin head. In contrast, sequence comparison between Motif 0 and myomesin (a myosin binding protein) suggests that Motif 0 contains a novel putative LMM binding site[33]. Whether this interaction occurs and what its function would be is unclear.

To further complicate interpretation of the function of this region, an actin binding site at the N-terminus of MyBPC has also been proposed. Co-sedimentation assays showed a low affinity of F-actin for all isoforms of MyBPC [8, 41], indicating that any interaction with actin is unlikely to be via a cardiac-specific region. Homology modelling with the essential light chain of myosin identified the Pro-Ala rich "linker" preceding Motif 1 as the likely candidate for binding actin[10]. However, no FHC mutations have been found in this linker to support its importance. Recently, fragments of MyBPC containing Motif 0 were shown to bind to actin, with the suggestion that it contributes to the weak binding state by shifting the binding of the N-terminus of MyBPC between actin and myosin[40].

Motif 0 and the 0-1 linker are highly unlikely to be able to bind to the myosin crossbridge, the myosin backbone and actin simultaneously. It is possible that the N-terminus of MyBPC cycles through different binding partners, but a more detailed investigation will be required before binding partner/s or roles for Motif 0 and the 0-1 domain linker can be positively assigned. Motifs 1 and 2 and the phosphorylatable linker that connects them, possess numerous FHC mutations of varying penetrance, stressing the importance of this region. The mutation Y237S in Motif 1 results in a mild FHC phenotype but shows a strong disease association[42]. This residue is highly conserved across all isoforms and species of MyBPC and is a core residue for defining both the IgI and FnIII domain folds. Located in strand F, Y237 is predicted to point into the hydrophobic core of the domain and to form stabilising hydrogen bonds with residues in strand G. Therefore, when mutated, these interactions may be disrupted, leading to a decreased stability of the domain and the resulting disease state.

The linker between Motifs 1 and 2 is of particular significance due to two additional phosphorylation sites (S284 and S304) in the cardiac isoform, including one present in a cardiac-specific LAGGGRRIS sequence (S284)[12]. All three phosphorylation sites can be phosphorylated by Ca²⁺/ calmodulin-dependent (CaM-II) kinases, including the endogenous CaM-II-like kinase that co-purifies with MyBPC [43, 44]. The first phosphate must be added by a CaM-II kinase to residue S284, to make the other phosphorylation sites accessible[45, 46]. Upon adrenergic stimulation, cAMP-dependent protein kinase can phosphorylate the other two sites (S275 and S304). Interestingly, *in vivo*, mono- and di-phosphorylated intermediates occur, which could result in more subtle changes in MyBPC's function [46].

The Motif 1-2 linker region is believed to regulate muscle contraction by specifically binding to the subfragment-2 (S2) region of myosin[4, 7, 47]. Phosphorylation of cardiac MyBPC releases the S2 region, which is thought to allow the myosin crossbridges to reach out and interact more efficiently with actin, increasing force generation and systolic tension[48-50]. Flexibility of the myosin head, both as a whole, and within its two domains, is a necessary requirement for efficient force generation[51-53]. The extent of phosphorylation required to abolish the interaction with S2 remains unclear[4].

Several FHC mutations are found in the Motif 1-2 phosphorylatable linker. Two of these mutations, in glycine residues G278E and G279A, occur within the conserved, cardiac-specific LAGGGRRIS sequence, in close proximity to two of the phosphorylation sites[27]. Glycine is unique in that it has no side chain and therefore can adopt phi and psi angles in all four quadrants of the Ramachandran plot[54]. If it is replaced with another residue it may permit a change in the three dimensional structure of the region and/or possibly obstruct access for the kinase. Notably, the residue at position 279 is an alanine (A) in the normal murine heart.

Another mutation in this region, R326Q, is more

controversial. It is a highly conserved residue and is associated with incomplete penetrance and is associated with onset FHC[42, 55, 56]. However, it has also been found in healthy controls[57, 58], suggesting that it may be a neutral polymorphism. These data suggest that any sarcomeric dysfunction associated with this mutation may be minimal.

Central region (Motifs 3-6)

Both IgI Motifs 3 and 4 have, as of yet, no experimentally defined roles, but may be required for the flexibility of the N-terminal region for its interactions with either the S2 region of myosin or the actin filament. Three FHC mutations have been identified in arginine residues in Motif 3 (R495Q, R502Q and R502W). Despite being highly conserved in MyBPC, they are not essential residues for defining the IgI fold and mutation results in a favourable disease prognosis[26, 27, 55].

The first detailed structural study of an isolated domain of cardiac MyBPC, and two FHC related point mutations, was performed by NMR and circular dichroism on the Motif 5 domain [59]. In addition to possessing the key characteristics of the IgI set, the structure exhibited a novel feature. An additional ten amino acids beyond the predicted N-terminus of Motif 5 were required for the stability of the isolated domain, with these 10 residues forming an integral part of the isolated Motif 5 fold. It is not clear whether this unusual packing occurs in the full length MyBPC, or if these 10 residues actually form a short linker between domains 4 and 5, as the sequence alignments would predict. Electron micrographs of isolated cardiac MyBPC revealed over half of the protein molecules were V-shaped, with arms of 22±4.5 nm[60], implying a point of flexibility occurs between Motifs 4 and 5[9]. Electron micrographs of skeletal MyBPC also showed V-shaped molecules[61]. An atomic resolution structure of Motifs 4 and 5 together is needed to determine the true nature of this unique "linker".

The second feature of the NMR structure of isolated Motif 5 is the 28 residue cardiac-specific insertion in the CD loop. This loop was unstructured, highly dynamic and pointing away from the domain's surface[59]. This proline/charge rich insert, while always present in the cardiac isoform, varies greatly in its sequence and length (Fig 5), and has been identified as a possible target binding region for an, as yet, unidentified ligand. One suggestion is that this insert forms an SH3 domain recognition sequence, perhaps binding the Calmodulin class -II (CaM-II) like kinase that co-purifies with cardiac MyBPC[12, 43, 62]. However, only the human cardiac isoform conforms to the PXXP–PXXP sequence usually required for SH3 target recognition. Additionally, the proposed target CaM-II kinase does not contain an identifiable SH3 domain[63]. Nevertheless, the Motif 5 insert could bind a different class of kinase.

Recently, a structural model of the MyBPC molecule has been developed, based on data from a yeast two-hybrid assay, using the Motif 5 sequence as "bait". Motif 5 was chosen due to its unique cardiac insert and the identification of several FHC related mutations. In conjunction with deletion mapping studies, the yeast two hybrid assay surprisingly identified Motif 8 as a preferential binding partner in a screen of $>7 \times 10^6$ clones[9]. Additional yeast two-hybrid assays revealed an interaction between Motifs 7 with 10. An interaction between Motifs 6 and 9 is also predicted to occur. As a result of these studies, a model of the myosin filament was developed, in which three MyBPC molecules formed a 'collar' around the myosin thick filament, stabilized by intermolecular interaction between Motifs 5-7 of each molecule and Motifs 8-10 of the next molecule, thus forming a staggered parallel arrangement (Fig 2)[9].

To date, three FHC-associated mutations have been identified in Motif 5. The N755K point mutation in Motif 5 exhibits a severe phenotype and lies in the highly conserved position 1 of a Type 1 β -turn connecting the F and G strands of the IgI domain. It was predicted that this highly conserved β -turn is stabilized by hydrogen bonding between N755 and G758 at position 4 of this β -turn [64]. Circular dichroism and NMR have confirmed that this N755K mutant is unstable and largely unfolded compared to the wild-type Motif 5, due to the loss of several key interactions [59, 65]. The NMR study suggested that tight packing by adjacent domains may partially stabilize the mutant Motif 5 in a folded conformation, to allow it to withstand modest mechanical stress. However, the severe phenotype of this mutation suggests significant stress within the sarcomere is likely to lead to partial or complete loss of structure and, thus, function. Furthermore, the N755K mutation leads to a weakened interaction with its binding partner Motif 8, where a 10 fold decrease in the affinity of Motif 5 for 8 was measured in the presence of the mutation[9].

A second FHC mutation in Motif 5, R654H, exhibits a milder phenotype, and results in a smaller decrease in the binding affinity between Motif 5 and Motif 8 (2 fold decrease)[9]. The R654H mutation did not appear to affect the stability or overall fold of the module, in good agreement with the location on an exposed protein surface [56]. Residue R654 is located on the CFGA' face of Motif 5. CFGA' together with the negatively charged cardiac-specific insert, results in a highly negative surface [59]. This surface on Motif 5 may be the binding site for the positively charged surface on Motif 8. Thus, the R654H



Fig 8. Solved structure of a FnIII pair from human fibronectin[69]. The homologous positions of FHC mutations found in Motifs 6 (left domain) and 7 (right domain) are superimposed as spheres. The sequential order of the strands is given by their labels: A, A', B, etc.

mutation, although it does not affect the protein fold or stability, is likely to affect an interaction with Motif 8.

A third mutation in Motif 5, R668H, was recently found [42]. This residue, located in strand B on the surface of Motif 5, is not a key residue for the IgI fold, although it is

a highly conserved residue in MyBPC. Substitution of a polar positive histidine for a polar positive arginine would not be predicted to cause significant structural instability of the fold of the domain, but it may impact on Motif 5's ability to bind a ligand. However, this mutation is not located on the negatively charged surface proposed to be a target for Motif 8, and is thus unlikely to interfere with Motif 8 binding.

No specific functions or sarcomeric interactions have so far been assigned to the Motif 6 fibronectin domain. Based on its location between Motifs 5 and 7, the trimeric collar model of MyBPC[9] suggests that Motif 6 may bind to Motif 9. This specific interaction has not been directly tested, although a construct of Motifs 5-7 binds to a construct of Motifs 8-10 with higher affinity than Motif 5 alone[9]. The demonstration of a specific interaction between Motifs 6 and 9 and a definition of their binding interface will be important for interpretation of FHC mutations in Motif 6.

FnIII modules Motifs 6 and 7 are reasonably homologous to similar FnIII domain pairs found in titin and fibronectin[63-69]. This homology includes a short linker region between the domains. FnIII domain pairs from titin have been shown to possess several highly conserved residues that are found in loops, in particular the BC loop. These residues point towards the module-module interface and interact via electrostatic charges[37]. The EF loop at the "bottom" of one domain is predicted to form a salt bridge and hydrogen bonds with the BC loop at the "top" of the following domain (Fig 7 and 8). Residue A833 is a mutational hot spot at the end of strand E in Motif 6 (Fig 8)[27, 42]. It is possible that mutations in this conserved amino acid may disrupt the bonds between the EF loop of Motif 6 and the BC loop of Motif 7, resulting in incorrect packing and assembly of MyBPC.

Another region of mutational activity reported in 3 patients occurs in strand C at the conserved residues R810 and K811 (Fig 8)[27, 70]. However, the reported mutations (R810H and K811R) do not alter the charge of the residue and are thus less likely to affect the stability of the FnIII fold. Of relevance is the clinical observation that these mutations appear to result in a mild phenotype, since one heterozygote had only moderate hypertrophy, and severe hypertrophy was only seen in a very unusual homozygote[70].

The proposed mutation R820Q (Fig 8) in Motif 6 has led to some uncertainty concerning the wildtype amino acid sequence of human cardiac MyBPC. Numerous entries in the NCBI database for wild-type human cardiac MyBPC show a glutamine at position 820, compared to one entry submitted by Niimura *et al.* in 1997 (number AAC04620) which has an arginine. On the other hand, MyBPC sequence alignments show that all isoforms and species other than human cardiac have either an arginine or a lysine residue at position 820, suggesting that the glutamine in the wild-type human cardiac sequence may not be correct. The proposed mutation R820Q is located in strand C' in Motif 6 and is not predicted to be part of the Motif 6-7 interface. The associated FHC phenotype appears to be mild, usually with first presentation being late in life[67, 68]. The confusion in defining R820Q as a FHC mutation or a neutral polymorphism makes the prediction of possible structural defects difficult.

C-terminal region (Motifs 7-10)

Motifs 7 to 10 of the C-terminal region of MyBPC bind the backbone region of the myosin thick filament[47, 72]. The primary myosin and titin binding regions of MyBPC are localised to Motif 10 and Motifs 8-10, respectively [73, 74]. These three C-terminal domains are the minimal requirement for incorporation into the A-band of the sarcomere[6], with Motif 7 improving the targeting of MyBPC to the C-Zone[75]. The trimeric collar model[9] also proposes specific interactions between C-terminal domains of adjacent MyBPC molecules.

A large majority of FHC mutations result in the premature termination of translation of the C-terminus of MyBPC, thus eliminating the titin and/or myosin binding sites[26, 32, 64, 76, 77]. This probably results in minimal or no incorporation of these truncated mutant proteins into the sarcomere[32].

Three missense mutations are located in Motif 7 (Fig 8). The first, P873H[60], changes a proline that is highly conserved across species and isoforms of MyBPC and is at the position that defines the N-terminal boundary in FnIII domains. The mutation P873H is unique, being found in only one patient with a mild phenotype. Modelling predicts that this mutation may affect the formation of strand A or the stability of the linker between Motifs 6 and 7.

It is unclear whether a second mutation, V896M (Fig 8)[42, 78], located in strand B of Motif 7 is disease causing, disease modifying or represents a neutral polymorphism [27, 57]. Patients identified with the V896M mutation in several studies were found to be double heterozygous, with these patients also possessing known FHC mutations in the myosin heavy chain gene[27, 42]. Furthermore, a clinically unaffected relation was found to be homozygous for the mutation. The V896M variant has also been found in control subjects [57]. Thus this mutation may not cause disease, but it may increase the severity of the disease in the presence of other FHC mutations. V896 is a conserved residue, predicted to point into the core of the FnIII domain. The replacement of the valine residue by the longer methionine may affect domain stability, but, as they are both non-polar residues, it is unlikely to result in incorrect folding.

The final mutation found in Motif 7 is N948T (Fig 8) [56]. This mutation, found in only one patient, occurs in a conserved position and results in a severe disease

phenotype. From homology modelling this mutation is expected to be located in a non-classical β -type turn connecting strands F and G[37]. This is in a similar position to the N755K mutation found in Motif 5, although the asparagine in Motif 5 forms part of a tight, type-1 β turn. Therefore, N948 is likely to be critical for domain folding and its loss could result in a poorly folded domain unable to help target MyBPC to the A band of the sarcomere. Additionally, the mutation may disrupt the interaction between Motif 7 and its proposed binding partner, Motif 10.

Only one FHC mutation has been identified in Motif 8. The mutation R1002Q[58] is located in strand C, but, while it is conserved across species and isoforms of MyBPC, it is not a key IgI folding residue. Motif 8 is the potential binding partner of Motif 5 in the trimeric collar model, in which a negative face on Motif 5 is thought to bind to a positive region on Motif 8[9]. This mutation results in the loss of a positively charged amino acid but this loss probably has little impact on the Motifs 5/8 interaction, consistent with the disease phenotype for this mutation being mild. The binding faces for this proposed interaction require elucidation before any definite conclusions can be made.

Motif 9 contains no known FHC mutations. The trimeric collar model predicts that it may interact with Motif 6, although this has not been directly tested[9]. Its contribution to myosin and titin binding has only been evaluated as part of the C-terminal 3 or 4 domains[6, 75]. Thus, the functional importance of Motif 9 remains unclear and may be clarified by future structural studies.

The binding of the C-terminal Motif 10 domain to the myosin filament has been extensively studied and nine key residues involved in binding to the filament have been identified by mutagenesis/sedimentation assays[6, 73, 79]. When these residues were positioned onto a model of cardiac Motif 10, the myosin binding faces were found to be located on two surfaces, the first formed by strands B and E and the second by strand C of the IgI fold. A unique insertion-mutation associated with FHC is a 6 amino acid duplication of residues 1248-1253[77]. Functional and spectroscopic characterization of the isolated wildtype and the insertion-mutant of Motif 10, using a myosin binding sedimentation assay, circular dichroism spectroscopy and molecular modeling, revealed that the structure of Motif 10 is minimally perturbed and the mutation does not interfere with myosin binding[80]. Molecular modelling positioned the six amino acid duplication-mutation on the surface shown to not be involved in myosin binding[80]. In addition, the close location of the single FHC point mutation in a conserved hydrophobic residue, A1255T[27], is also unlikely to interfere with myosin binding. These results suggest that both the unique 6 amino acid 1248-1253 duplication mutation and the A1255T mutation may affect some other function of Motif 10, possibly its binding to titin[78] or an alteration in an interaction that may occur with Motif 7 or the adjacent Motif 9.

A second FHC point mutation in Motif 10, also resulting in a change from a hydrophobic alanine to a polar threonine at position 1194[27], is located at the end of strand A preceding the A-B loop. Modelling suggests that replacement of this residue with the more bulky threonine could alter the interaction with strand F. This residue is also located close to two identified key myosin-binding residues[79]. Therefore, this FHC mutation, unlike the others found in Motif 10, may disrupt binding to the myosin filament.

Models for the assembly of MyBP-C onto the thick filament

There are currently two models for the arrangement of MyBPC in the sarcomere. The trimeric collar model proposed by Moolman-Smook *et al.* (2002)[9], has three staggered MyBPC molecules forming a ring around the thick filament. The collar is thought to be stabilized by specific interactions that have been demonstrated between Motifs 5 and 8, and Motifs 7 and 10 (Fig 2). Motifs 0 to 4, forming the N-terminal half of MyBP-C are predicted to have sufficient length to reach out from the thick filament and interact with the myosin crossbridge and/or the actin filament[5, 10].

This model raises several issues. Firstly, the geometry of this model needs further verification. The diameter of the myosin thick filament is 13-15 nm[82], resulting in a circumference of approximately 41-47 nm. The trimeric collar model proposes that the length of the ring that wraps around the myosin thick filament is 9 immunoglobulin domains. The longest diameter of these domains is 3.4-3.9 nm [83, 84], resulting in the largest ring possible being 35-36 nm, 5-12 nm shorter than the measured circumference of the myosin thick filament. No current model of MyBPC function suggests that the C-terminal motifs are subjected to high degrees of stretch. Alternatively, the MyBPC linker regions between domains may provide additional length to the collar-forming C-terminal half of the model. However, the only linker of significant length in this half of MyBPC is located between Motifs 9 and 10, and stretching of this linker may result in mismatch between interacting domains in the trimeric collar model (Fig 2).

Secondly, how does the titin molecule fit into this model? Titin is the molecular ruler of the sarcomere[85] and three strands are bound axially to the thick filament (Fig 2). The C-terminal Motifs 7-10 of MyBPC bind to both titin and the LMM region of the myosin thick filament. It is unclear how titin interacts with Motifs 7-10 in this model, given that the trimeric collar model would orient them as being close to mutually perpendicular[74].

An alternative model of MyBPC has the C-terminal motifs running parallel to the myosin backbone (Fig 3)[10]. This more recent model was developed primarily to explain differences between the observed length of MyBPC in experimental X-ray diffraction patterns and the predicted spatial arrangement of MyBPC in the sarcomere. MyBPC has a longer periodicity than the myosin filament repeat [86, 87]. This model accounts for this discrepancy by allowing the N-terminal domain of MyBPC to interact with neighbouring actin filaments in defined muscle states, thereby shortening the overall length of MyBPC running parallel with the myosin filament. Additionally, with the MyBPC molecules arranged axially they could more readily interact with titin, in contrast to the trimeric collar model. However, apart from the X-ray diffraction patterns, there is no direct experimental evidence that currently supports this model and no explanation for the experimental findings of Motif 5 binding to 8 and Motif 7 binding to 10. Clearly, more experimental data is required to clarify the manner in which MyBPC binds to the thick filament.

Conclusion

Analysing the effect of a mutant protein on muscle cell architecture, physiology, and biochemistry should provide an insight into the FHC disease process. A first step in this process is a careful examination of the sequence of MyBPC, with an extensive comparison of both the sequence and the predicted individual domain structures. The analysis presented here has clarified possible structural consequences of FHC mutations, laying a framework for the design of future structural and physiological studies to directly test these predictions. Most notably the predictions presented here emphasise the correlation between disease phenotype and the extent of conservation of the associated mutant residue/s.

ACKNOWLEDGEMENTS

CEO is a recipient of an Australian Postgraduate Award. This work was supported by a National Heart Foundation of Australia Grant-in-Aid.

REFERENCES

- Starr R, Offer G. Polypepetide chains of intermediate molecular weight in myosin preparations. FEBS Lett 1971; 15(1):40-4.
- 2 Craig R, Offer G. The location of C-protein in rabbit skeletal muscle. Proc R Soc Lond [Biol] 1976; **192(1109)**:451-61.
- 3 Bennett P, Craig R, Starr R, Offer G. The ultrastructural location

of C-protein, X-protein and H-protein in rabbit muscle. J Muscle Res Cell Motil 1986; **7(6)**:550-67.

- 4 Gruen M, Prinz H, Gautel M. cAPK-phosphorylation controls the interaction of the regulatory domain of cardiac myosin binding protein C with myosin-S2 in an on-off fashion. FEBS Lett 1999; **453(3)**:254-9.
- 5 Witt CC, Gerull B, Davies MJ, Centner T, Linke WA, Thierfelder L. Hypercontractile properties of cardiac muscle fibers in a knockin mouse model of cardiac myosin-binding protein-C. J Biol Chem 2001; 276(7):5353-9.
- 6 Gilbert R, Kelly MG, Mikawa T, Fischman DA. The carboxyl terminus of myosin binding protein C (MyBP-C, C-protein) specifies incorporation into the A-band of striated muscle. J Cell Sci 1996; **109 (Pt 1)**:101-11.
- 7 Gruen M, Gautel M. Mutations in β-myosin S2 that cause familial hypertrophic cardiomyopathy (FHC) abolish the interaction with the regulatory domain of myosin-binding protein-C. J Mol Biol 1999; 286(3):933-49.
- 8 Yamamoto K, Moos C. The C-proteins of rabbit red, white, and cardiac muscles. J Biol Chem 1983; **258(13)**:8395-401.
- 9 Moolman-Smook J, Flashman E, de Lange W, et al. Identification of novel interactions between domains of Myosin binding protein-C that are modulated by hypertrophic cardiomyopathy missense mutations. Circ Res 2002; 91(8):704-11.
- 10 Squire JM, Luther PK, Knupp C. Structural evidence for the interaction of C-protein (MyBP-C) with actin and sequence identification of a possible actin-binding domain. J Mol Biol 2003; 331(3):713-24.
- 11 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997; 25(24):4876-82.
- 12 Gautel M, Zuffardi O, Freiburg A, Labeit S. Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? EMBO J 1995; 14 (9):1952-60.
- 13 Aparicio S, Chapman J, Stupka E, *et al*. Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. Science 2002; **297(5585)**:1301-10.
- 14 Offer G, Moos C, Starr R. A new protein of the thick filaments of vertebrate skeletal myofibrils. Extractions, purification and characterization. J Mol Biol 1973; 74(4):653-76.
- 15 Saad AD, Obinata T, Fischman DA. Immunochemical analysis of protein isoforms in thick myofilaments of regenerating skeletal muscle. Dev Biol 1987; 119(2):336-49.
- 16 Harris SP, Bartley CR, Hacker TA, *et al.* Hypertrophic cardiomyopathy in cardiac myosin binding protein-C knockout mice. Circ Res 2002; **90(5)**:594-601.
- 17 Winegrad S. Cardiac myosin binding protein C. Circ Res 1999; 84(10):1117-26.
- 18 Hofmann PA, Hartzell HC, Moss RL. Alterations in Ca²⁺ sensitive tension due to partial extraction of C-protein from rat skinned cardiac myocytes and rabbit skeletal muscle fibers. J Gen Physiol 1991; 97(6):1141-63.
- 19 Hofmann PA, Greaser ML, Moss RL. C-protein limits shortening velocity of rabbit skeletal muscle fibres at low levels of Ca²⁺ activation. J Physiol 1991; **439**:701-15.
- 20 Chung MW, Tsoutsman T, Semsarian C. Hypertrophic cardiomyopathy: from gene defect to clinical disease. Cell Res

2003; 13(1):9-20.

- 21 Maron BJ. Triggers for sudden cardiac death in the athlete. Cardiol Clin 1996; **14(2)**:195-210.
- 22 Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults. Circulation 1995; **92(4)**:785-9.
- 23 Miura K, Nakagawa H, Morikawa Y, *et al.* Epidemiology of idiopathic cardiomyopathy in Japan: results from a nationwide survey. Heart 2002; **87(2)**:126-30.
- 24 Maron BJ, Olivotto I, Spirito P, *et al.* Epidemiology of hypertrophic cardiomyopathy-related death: revisited in a large nonreferral-based patient population. Circulation 2000; **102(8)**:858-64.
- 25 Carrier L, Hengstenberg C, Beckmann JS, *et al.* Mapping of a novel gene for familial hypertrophic cardiomyopathy to chromosome 11. Nat Genet 1993; **4(3)**:311-3.
- 26 Niimura H. Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. N Engl J Med 1998; **338(18)**:1248-57.
- 27 Richard P, Charron P, Carrier L, et al. Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. Circulation 2003; 107(17):2227-32.
- 28 Charron P, Dubourg O, Desnos M, *et al*. Genotype-phenotype correlations in familial hypertrophic cardiomyopathy. A comparison between mutations in the cardiac protein-C and the β -myosin heavy chain genes. Eur Heart J 1998; **19(1)**:139-45.
- 29 Yang Q, Sanbe A, Osinska H, Hewett TE, Klevitsky R, Robbins J. A mouse model of myosin binding protein C human familial hypertrophic cardiomyopathy. J Clin Invest 1998; 102(7):1292-300.
- 30 McConnell BK, Jones KA, Fatkin D, et al. Dilated cardiomyopathy in homozygous myosin-binding protein-C mutant mice. J Clin Invest 1999; 104(9):1235-44.
- 31 Yang Q, Sanbe A, Osinska H, Hewett TE, Klevitsky R, Robbins J. *In vivo* modeling of myosin binding protein C familial hypertrophic cardiomyopathy. Circ Res 1999; **85(9)**:841-7.
- 32 Rottbauer W, Gautel M, Zehelein J, *et al.* Novel splice donor site mutation in the cardiac myosin-binding protein-C gene in familial hypertrophic cardiomyopathy. Characterization Of cardiac transcript and protein. J Clin Invest 1997; **100(2)**:475-82.
- 33 Flavigny J, Souchet M, Sebillon P, *et al.* COOH-terminal truncated cardiac myosin-binding protein C mutants resulting from familial hypertrophic cardiomyopathy mutations exhibit altered expression and/or incorporation in fetal rat cardiomyocytes. J Mol Biol 1999; **294(2)**:443-56.
- 34 Harpaz Y, Chothia C. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. J Mol Biol 1994; 238(4):528-39.
- 35 Politou AS, Gautel M, Joseph C, Pastore A. Immunoglobulintype domains of titin are stabilized by amino-terminal extension. FEBS Lett 1994; 352(1):27-31.
- 36 Main AL, Harvey TS, Baron M, Boyd J, Campbell ID. The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions. Cell 1992; 71(4):671-8.

- 37 Goll CM, Pastore A, Nilges M. The three-dimensional structure of a type I module from titin: a prototype of intracellular fibronectin type III domains. Structure 1998; 6(10):1291-302.
- 38 Bork P, Holm L, Sander C. The immunoglobulin fold. Structural classification, sequence patterns and common core. J Mol Biol 1994; 242(4):309-20.
- 39 Steward A, Adhya S, Clarke J. Sequence conservation in Ig-like domains: the role of highly conserved proline residues in the fibronectin type III superfamily. J Mol Biol 2002; 318(4):935-40.
- 40 Kulikovskaya I, McClellan G, Flavigny J, Carrier L, Winegrad S. Effect of MyBP-C Binding to Actin on Contractility in Heart Muscle. J Gen Physiol 2003; 122(6):761-74.
- 41 Moos C, Mason CM, Besterman JM, Feng IN, Dubin JH. The binding of skeletal muscle C-protein to F-actin, and its relation to the interaction of actin with myosin subfragment-1. J Mol Biol 1978; 124(4):571-86.
- 42 Morner S, Richard P, Kazzam E, *et al.* Identification of the genotypes causing hypertrophic cardiomyopathy in northern Sweden. J Mol Cell Cardiol 2003; 35(7):841-9.
- 43 Hartzell HC, Glass DB. Phosphorylation of purified cardiac muscle C-protein by purified cAMP-dependent and endogenous Ca²⁺-calmodulin-dependent protein kinases. J Biol Chem 1984; 259(24):15587-96.
- 44 Schlender KK, Bean LJ. Phosphorylation of chicken cardiac Cprotein by calcium/calmodulin-dependent protein kinase II. J Biol Chem 1991; 266(5):2811-7.
- 45 Mohamed AS, Dignam JD, Schlender KK. Cardiac myosin-binding protein C (MyBP-C): identification of protein kinase A and protein kinase C phosphorylation sites. Arch Biochem Biophys 1998; 358(2):313-9.
- 46 McClellan G, Kulikovskaya I, Winegrad S. Changes in cardiac contractility related to calcium-mediated changes in phosphorylation of myosin-binding protein C. Biophys J 2001; 81(2):1083-92.
- 47 Starr R, Offer G. The interaction of C-protein with heavy meromyosin and subfragment-2. Biochem J 1978; 171(3):813-6
- 48 Kunst G, Kress KR, Gruen M, Uttenweiler D, Gautel M, Fink RH. Myosin binding protein C, a phosphorylation-dependent force regulator in muscle that controls the attachment of myosin heads by its interaction with myosin S2. Circ Res 2000; 86(1): 51-8.
- 49 Weisberg A, Winegrad S. Alteration of myosin cross bridges by phosphorylation of myosin-binding protein C in cardiac muscle. Proc Natl Acad Sci USA 1996; 93(17):8999-9003.
- 50 Kulikovskaya I, McClellan G, Levine R, Winegrad S. Effect of extraction of myosin binding protein C on contractility of rat heart. Am J Physiol 2003; 285(2):H857-65.
- 51 Brown LJ, Klonis N, Sawyer WH, Fajer PG, Hambly BD. Independent movement of the regulatory and catalytic domains of myosin heads revealed by phosphorescence anisotropy. Biochemistry 2001; 40(28):8283-91
- 52 Baumann BA, Hambly BD, Hideg K, Fajer PG. The regulatory domain of the myosin head behaves as a rigid lever. Biochemistry 2001; 40(26):7868-73
- 53 Palm T, Sale K, Brown L, Li H, Hambly B, Fajer PG. Intradomain distances in regulatory domain of myosin head in prepower and postpower stroke states: fluorescence energy transfer. Biochemistry 1999; 38(40):13026-34

- 54 Hovmoller S, Zhou T, Ohlson T. Conformations of amino acids in proteins. Acta Crystallogr D Biol Crystallogr 2002; 58(Pt 5): 768-76.
- 55 Maron BJ, Niimura H, Casey SA, *et al.* Development of left ventricular hypertrophy in adults in hypertrophic cardiomyopathy caused by cardiac myosin-binding protein C gene mutations. J Am Coll Cardiol 2001; **38(2)**:315-21.
- 56 Daehmlow S, Erdmann J, Knueppel T, *et al.* Novel mutations in sarcomeric protein genes in dilated cardiomyopathy. Biochem Biophys Res Commun 2002; **298(1)**:116-20.
- 57 Jaaskelainen P, Kuusisto J, Miettinen R, *et al.* Mutations in the cardiac myosin-binding protein C gene are the predominant cause of familial hypertrophic cardiomyopathy in eastern Finland. J Mol Med 2002; **80(7)**:412-22.
- 58 Niimura H, Patton KK, McKenna WJ, *et al.* Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly. Circulation 2002; **105(4)**:446-51.
- 59 Idowu SM, Gautel M, Perkins SJ, Pfuhl M. Structure, stability and dynamics of the central domain of cardiac myosin binding protein C (MyBP-C): implications for multidomain assembly and causes for cardiomyopathy. J Mol Biol 2003; **329(4)**:745-61.
- 60 Hartzell HC, Sale WS. Structure of C protein purified from cardiac muscle. J Cell Biol 1985; 100(1):208-15.
- 61 Swan RC, Fischman DA. Electron microscopy of C-protein molecules from chicken skeletal muscle. J Muscle Res Cell Motil 1986; 7(2):160-6.
- 62 Redwood CS, Moolman-Smook JC, Watkins H. Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy. Cardiovasc Res 1999; **44(1)**:20-36.
- 63 Hudmon A, Schulman H. Structure-function of the multifunctional Ca²⁺/calmodulin-dependent protein kinase II. Biochem J 2002; **364(Pt 3)**:593-611.
- 64 Yu B, French JA, Carrier L, *et al*. Molecular pathology of familial hypertrophic cardiomyopathy caused by mutations in the cardiac myosin binding protein C gene. J Med Genet 1998; **35(3)**: 205-10.
- 65 Singh L, Oakley C, Brown LJ, Hambly B. Structural effects of hypertrophic cardiomyopathy mutations on sarcomeric proteins. Biophys J 2003; 84:452.
- 66 Leahy DJ, Aukhil I, Erickson HP. 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. Cell 1996; 84(1):155-64.
- 67 de Pereda JM, Wiche G, Liddington RC. Crystal structure of a tandem pair of fibronectin type III domains from the cytoplasmic tail of integrin $\alpha 6\beta 4$. EMBO J 1999; **18(15)**:4087-95.
- 68 Sharma A, Askari JA, Humphries MJ, Jones EY, Stuart DI. Crystal structure of a heparin- and integrin-binding segment of human fibronectin. EMBO J 1999; 18(6):1468-79.
- 69 Leahy DJ, Erickson HP, Aukhil I, Joshi P, Hendrickson WA. Crystallization of a fragment of human fibronectin: introduction of methionine by site-directed mutagenesis to allow phasing via selenomethionine. Proteins 1994; **19(1)**:48-54.
- 70 Nanni L, Pieroni M, Chimenti C, et al. Hypertrophic cardiomyopathy: two homozygous cases with 'typical' hypertrophic cardiomyopathy and three new mutations in cases with progression to dilated cardiomyopathy. Biochem Biophys Res Commun 2003; **309(2)**:391-8.

- 71 Konno T, Shimizu M, Ino H, *et al.* A novel missense mutation in the myosin binding protein-C gene is responsible for hypertrophic cardiomyopathy with left ventricular dysfunction and dilation in elderly patients. J Am Coll Cardiol 2003; 41(5):781-6.
- 72 Moos C, Offer G, Starr R, Bennett P. Interaction of C-protein with myosin, myosin rod and light meromyosin. J Mol Biol 1975; 97(1):1-9.
- 73 Okagaki T, Weber FE, Fischman DA, Vaughan KT, Mikawa T, Reinach FC. The major myosin-binding domain of skeletal muscle MyBP-C (C protein) resides in the COOH-terminal, immunoglobulin C2 motif. J Cell Biol 1993; **123**(3):619-26.
- 74 Freiburg A, Gautel M. A molecular map of the interactions between titin and myosin-binding protein C. Implications for sarcomeric assembly in familial hypertrophic cardiomyopathy. Eur J Biochem 1996; 235(1-2):317-23.
- 75 Gilbert R, Cohen JA, Pardo S, Basu A, Fischman DA. Identification of the A-band localization domain of myosin binding proteins C and H (MyBP-C, MyBP-H) in skeletal muscle. J Cell Sci 1999; **112(Pt 1)**:69-79.
- 76 Bonne G, Carrier L, Bercovici J, *et al.* Cardiac myosin binding protein-C gene splice acceptor site mutation is associated with familial hypertrophic cardiomyopathy. Nat Genet 1995; **11(4)**: 438-40.
- 77 Watkins H, Conner D, Thierfelder L, *et al.* Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. Nat Genet 1995; **11(4)**: 434-7.
- 78 Moolman-Smook JC, De Lange WJ, Bruwer EC, Brink PA, Corfield VA. The origins of hypertrophic cardiomyopathy-causing mutations in two South African subpopulations: a unique profile of both independent and founder events. Am J Hum Genet 1999; 65(5):1308-20.
- 79 Miyamoto CA, Fischman DA, Reinach FC. The interface between MyBP-C and myosin: site-directed mutagenesis of the CX myosin-binding domain of MyBP-C. J Muscle Res Cell Motil 1999; 20(7):703-15.
- 80 Brown LJ, Singh L, Sale KL, *et al.* Functional and spectroscopic studies of a familial hypertrophic cardiomyopathy mutation in Motif X of cardiac myosin binding protein-C. Eur Biophys J 2002; **31(5)**:400-8.
- 81 Furst DO, Vinkemeier U, Weber K. Mammalian skeletal muscle C-protein: purification from bovine muscle, binding to titin and the characterization of a full-length human cDNA. J Cell Sci 1992; 102(Pt 4):769-78.
- 82 Squire J, Cantino M, Chew M, *et al.* Myosin rod-packing schemes in vertebrate muscle thick filaments. J Struct Biol 1998; 122(1-2):128-38.
- 83 Pfuhl M, Pastore A. Tertiary structure of an immunoglobulinlike domain from the giant muscle protein titin: a new member of the I set. Structure 1995; **3(4)**:391-401.
- 84 Improta S, Politou AS, Pastore A. Immunoglobulin-like modules from titin I-band: extensible components of muscle elasticity. Structure 1996; 4(3):323-37.
- 85 Whiting A, Wardale J, Trinick J. Does titin regulate the length of muscle thick filaments? J Mol Biol 1989; **205(1)**:263-8.
- 86 Huxley HE, Brown W. The low-angle x-ray diagram of vertebrate striated muscle and its behaviour during contraction and rigor. J Mol Biol 1967; **30(2)**:383-434.

- 87 Sjostrom M, Squire JM. Cryo-ultramicrotomy and myofibrillar fine structure: a review. J Microsc 1977; **111(3)**:239-78.
- 88 Andersen PS, Havndrup O, Bundgaard H, *et al*. Myosin light chain mutations in familial hypertrophic cardiomyopathy: phenotypic presentation and frequency in Danish and South African

populations. J Med Genet 2001; **38(12)**:E43.

89 Moolman-Smook JC, Mayosi B, Brink P, Corfield VA. Identification of a new missense mutation in MyBP-C associated with hypertrophic cardiomyopathy. J Med Genet 1998; 35(3):253-4.