

Supplementary information

The structure of the KlcA and ArdB proteins reveals a novel fold and antirestriction activity against Type I DNA restriction systems *in vivo* but not *in vitro*.

Dimitra Serfiotis-Mitsa¹, Andrew P Herbert¹, Gareth A Roberts¹, Dinesh C Soares^{1,2}, John H White¹, Garry W Blakely³, Dusan Uhrin^{1,*}, David TF Dryden^{1,*}

¹EaStChem School of Chemistry, University of Edinburgh, The Kings' Buildings, Edinburgh, EH9 3JJ, UK

²Medical Genetics Section, Molecular Medicine Centre, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK

³School of Biological Sciences, The University of Edinburgh, The King's Buildings, Edinburgh EH9 3JJ, UK

Table S1. Molecular mass of purified proteins determined by size exclusion chromatography and FT-ICR mass spectrometry.

Protein	Predicted mass (Da)	FT-ICR MS (Da)	Size exclusion (Da)
KlcA ₁₃₆	15659.6	Not determined	20700
ArdB _{CFT}	17832.1	17354 [#] + 17827	33000 + 24200
ArdB _{YAF}	17288 (-Met1)	17288 + 17365*	54900 + 27300
ArdB _{YFJ}	17172 (-Met1)	17172 + 17249*	23000

[#] corresponds to full length ArdB_{CFT} minus the four N-terminal residues (MKTL).

* corresponds to a 2-mercaptoethanol adduct (+77 Da).

Figure S1. Residual dipolar coupling data. Fitting of the ¹D_{NH} residual dipolar couplings in the RDC-refined KlcA₁₃₆ structure. Back-calculation of the couplings was performed with REDCAT.

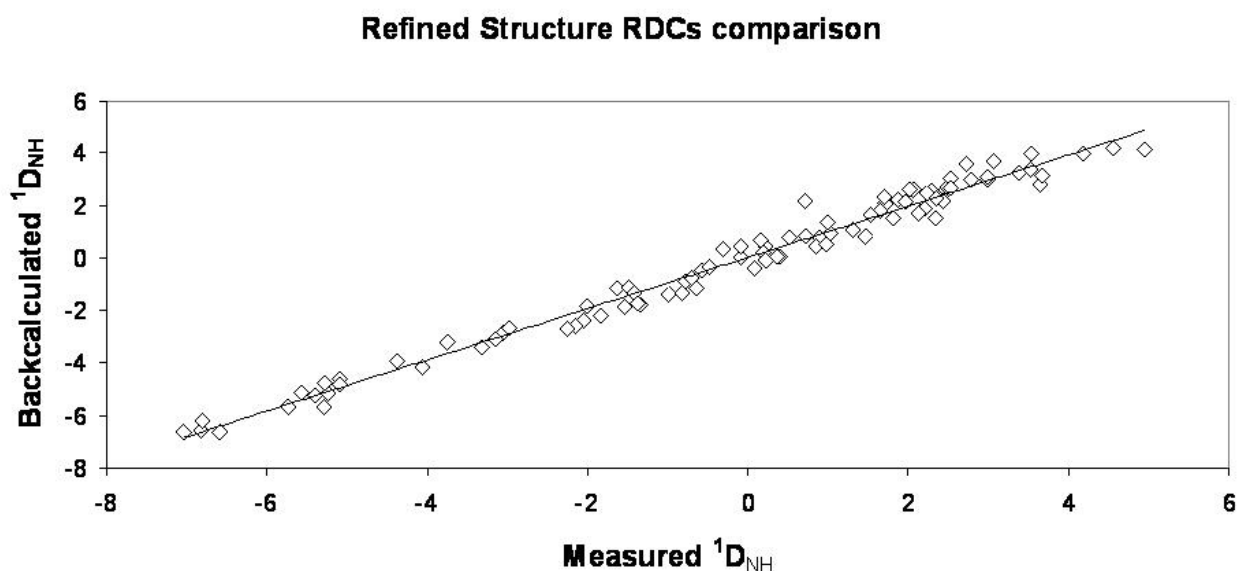


Table S2: NMR-derived restraints and structural statistics of the 20 lowest energy KlcA₁₃₆ structures.

Residual Dipolar Couplings	
¹ D _{NH}	92
Dihedral angle restraints	
Torsion angles (ϕ/ψ)	228
Distance restraints	
Short range ($ i-j \leq 1$)	1117
Medium range ($1 < i-j < 5$)	617
Long range ($ i-j \geq 5$)	604
RMSD (residues 7-142)	
All	1.28
Backbone	0.57
C alpha	0.59
Heavy atoms	1.05
WHATIF Score	-1.432
Ramachandran Plot Statistics (%)	
Most favourable regions	83
Additionally allowed regions	14.8
Generously allowed regions	0.7
Assignment percentages (%)	
Amide	94.2
Backbone	95.9
Side-chain H	90.9
Side-chain other	66.2
Element C	81.5
Element H	91.9
Element N	78.42

Figure S2. Stereoview (cross eye) of the ensemble of the 20 lowest energy structures of KlcA₁₃₆.

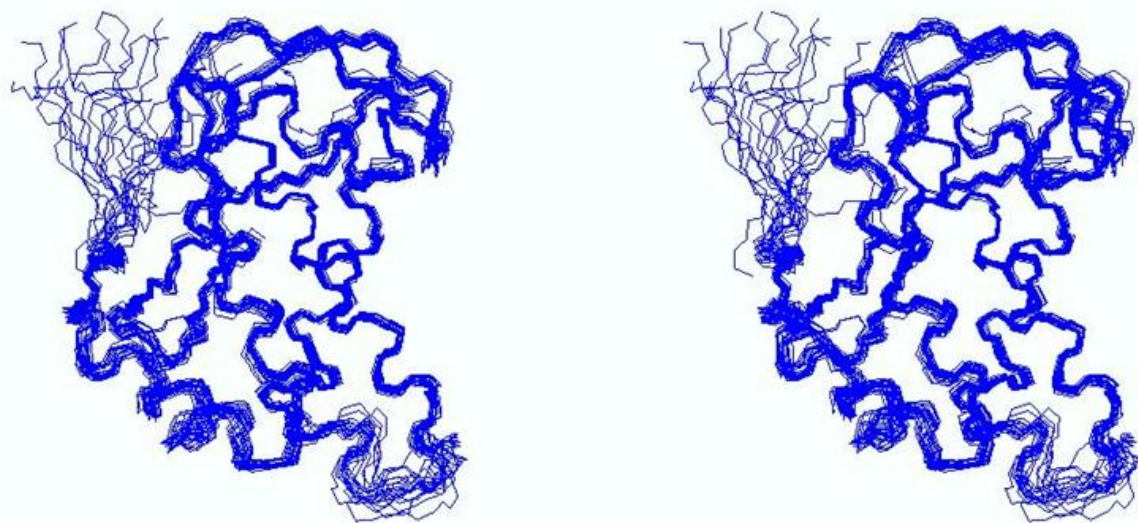
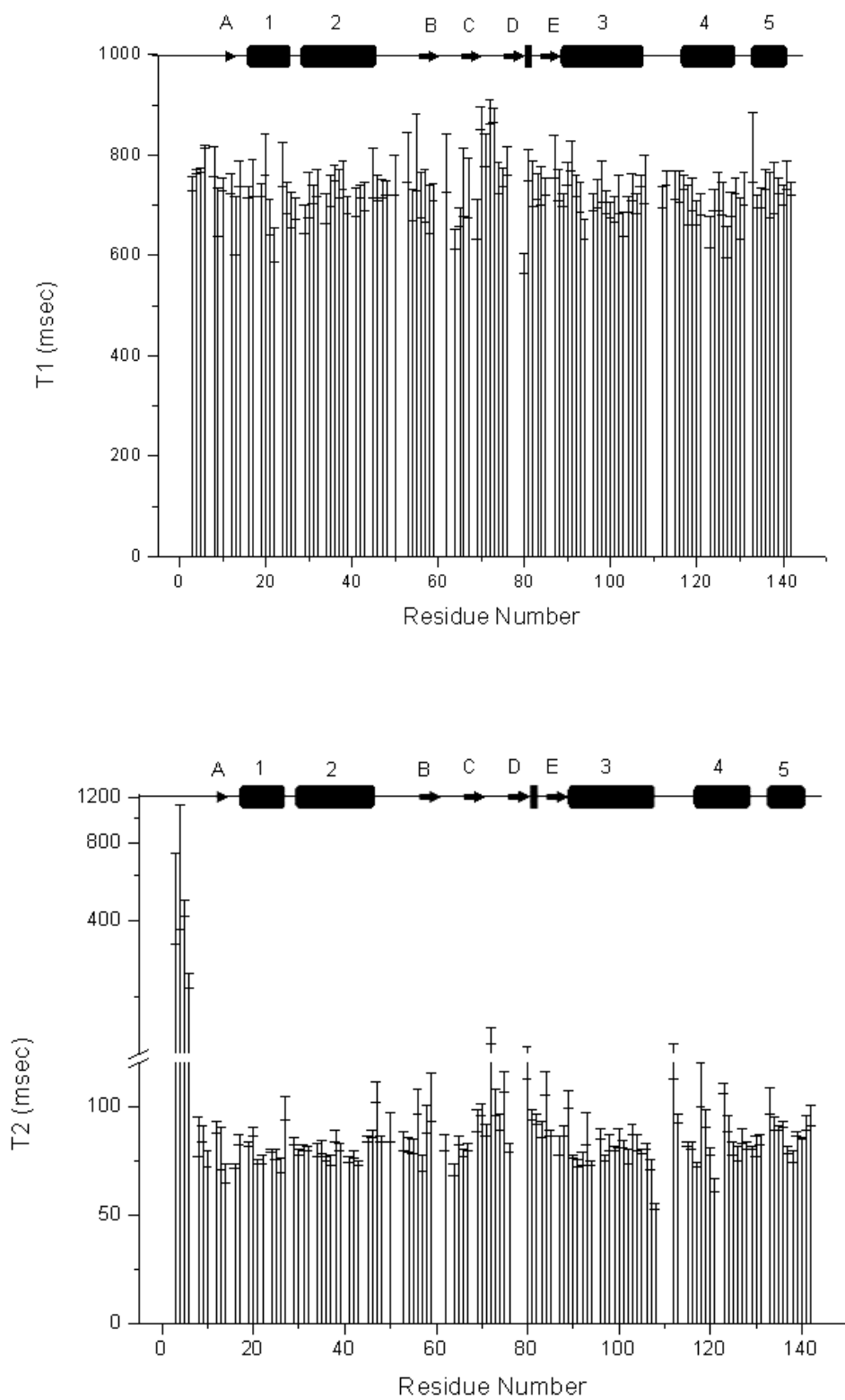


Figure S3. Experimental relaxation parameters from top to bottom: T1, T2, $\{^1\text{H}, ^{15}\text{N}\}$ NOE, and S^2 values of KlcA₁₃₆ as a function of residue number.



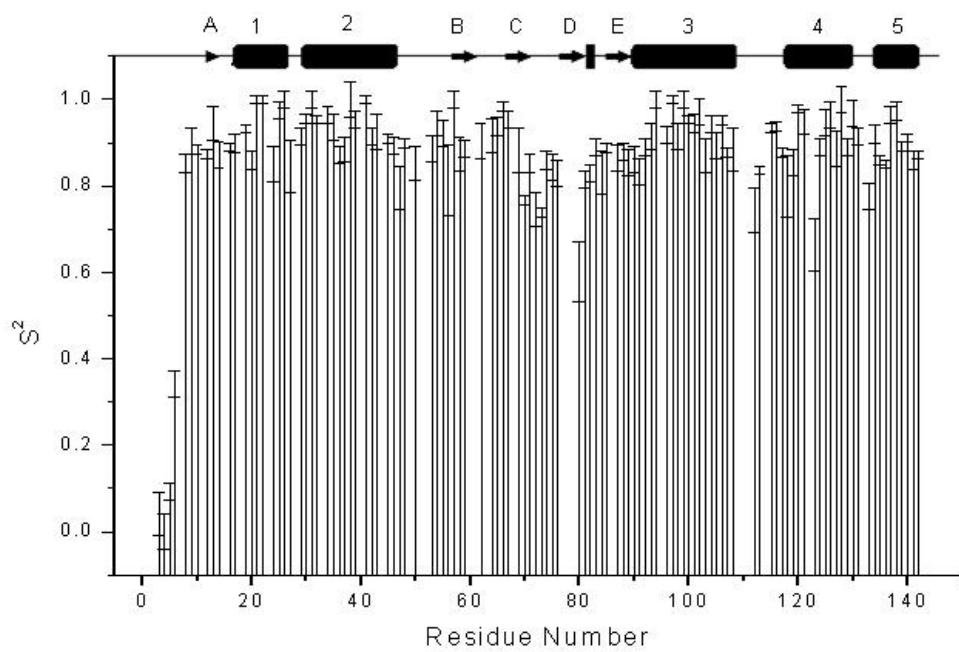
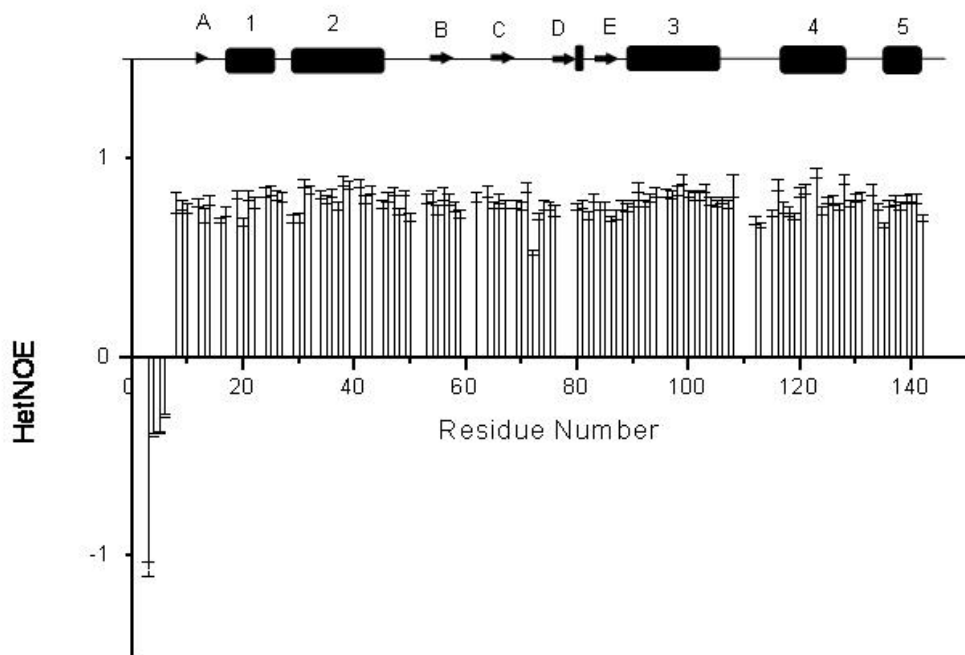


Figure S4. The order parameter S^2 visualised on the KlcA₁₃₆ structure. Blue denotes high flexibility, whereas residues coloured yellow are the most rigid ones. Grey was used for residues for which no data were available.

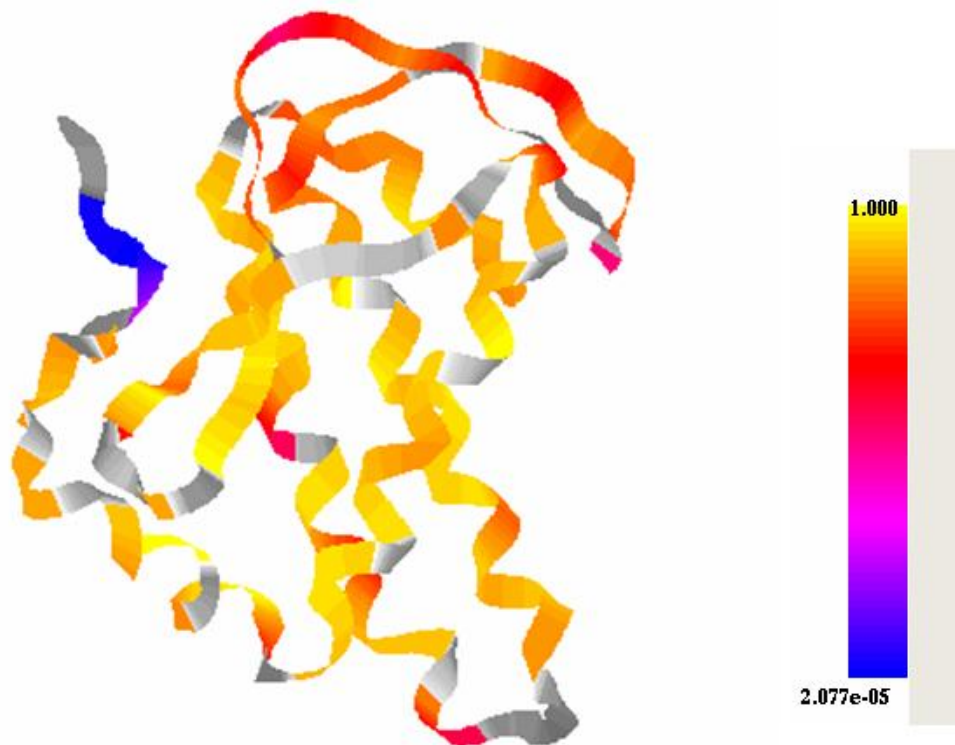
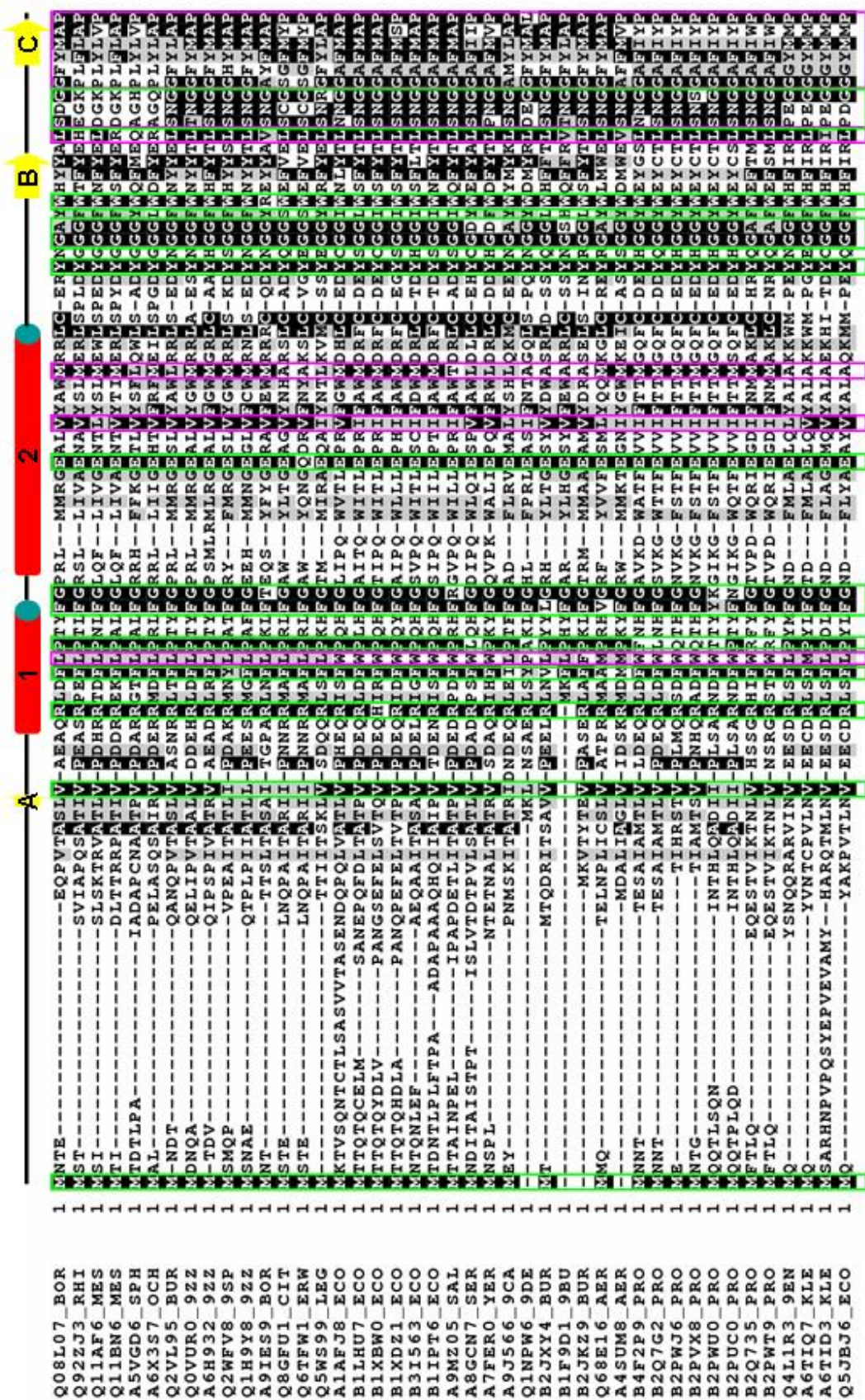


Figure S5. Multiple sequence alignment of 43 non-redundant sequences of ArdB and KlcA using PROMALS3D. The secondary structure (alpha-helices: red; beta-strands: yellow; 3_{10} -helix: purple) from the consensus ensemble of NMR-derived structures is shown above the aligned sequences. Highly conserved amino acid positions that are exposed are shown within a green rectangular box, while highly conserved, buried positions are highlighted within a purple box.





69 DLA---GRLEIEVNCFRGELSDAAGIVATDFALQLAEEIADTDA---ADALIDRYHF LRGFAAGCHPEAAAIYRAID
 71 QSK---SRFRITGEITFCQGVPAEAGIATIAFASHLSFQYQ-----SEHLSGCGYRFLYAYSADHPAEVEIFQAI
 71 TSK---PRRIICRTNDCYVSNAAAGIATDFVLSHLSFKYQ-----SDRLADGYSRFLFAYRANHAESAEIFQAI
 71 DTD---RRRIESAITDFRCVSAEAGIATDFVLSHLSFKFE-----SDRLAEGYKRFAYSDCHPESSEIFAAN
 75 AGR---PAYRICNSRFECTVSDAAGIATDFAFSHLSFEVE-----DDHIAEAFHRLDIAAGDHPAAKIFCAID
 71 TSK---PRCRFFCEGCEFEVSSDAAGIATDFAFSHLSFRYD-----DDELAEGYGRLYEHAADHAEAAFIQAI
 71 ELT---GRCLCEVDCNGYSELSDAAGIATDFALQLAADNQTDV---ADVLIIDRYHF LRDFAARDHAAEAFIYRAID
 72 VL---GRLRFVDCNGYSEMSADAAGIATDFALQLAEEIADTDA---CDALIDRYHF LREYAGTHAEAAFIYRAID
 73 EHD---GRLREVDCAFCAGLSDAAGIATDFALNQLCAELAGTAD---ADALIDRYHF LAAFASEHAAEAAFIYRAID
 70 EMG---DSRLAVLDCNGFCELSAAGIATDFALNQLTHEAKGEA---QDHLIEQYKLVFEFATTHPESGAIYRAID
 72 ATE---KTRFRVRVACNDFCELSAAGIATDFAYVAVCELANRFG---SDRIVDQYHILRDYAGEHPESAGQIFAAID
 68 HRP---EPMRQWALNWSNETVSSAAGIATDFGIYEI LESFP---DEDLAEHYRILCQYAYELPESCSIMRLIID
 70 LSA---ERFTVSVCNLFCELSAAGIATDFTLNHWIWAHEEGYQHICDMLITQOEK LKLYADQCHAEAGLIYRAID
 70 LSA---ERFTVSVCNLFCELSAAGIATDFTLNHWIWAHADAGYEHICDMLITQOEK LKLYADQCHAEAGLIYRAID
 65 DMD---EPLPIYVCGNYDCVSDAAGIATDFVNLQLCWKTK-----SDKTIQQYIIRLDYIEYHKEAADIYAAID
 88 EPDDDDDETWFNFNAKCNRAEVSPEAGIATDFMTYSHHACRTE-----CYAMTVHYRFLDYALOHPECNAIMRIID
 81 DAD---GDDKWHFLNCGCAEVSPEAGIATDFIYSHHACRTE-----CDAMTEHYRFLDYALOHPESSAIMRIID
 81 EPD---NDETRWLFNCGNDASPEAGIATDFIYSHHACRTE-----CDAMTAHYRFLDYALOHPEAHAILRIID
 76 ESE---HDEKWAFLNCGCAELTCEAGIATDFMTYSHHACRTE-----CDAMTEHYRFLDYALOHPESSAIMHIIID
 85 EVE---NDELWSLFNCGNDASPEAGIATDFMTYSHHACRTE-----CDAMTEHYRFLDYALOHPESSAIMHIIID
 80 EAD---DGGDTSWLFNCGNSALSPPEAGIATDFLEYSHHACRTE---SDLMTAHLHLRDYALTNHPECNAIMHIIID
 83 STE---GEYVLFNALNCGATVGEAGIATDFLITYSHHACRTE-----CDAMTEHYRFLDYALTNHPECNAIMRIID
 74 GTE---QDYVLFNALNCGATVGEAGIATDFMAYSHHACRTE---NEFMTEHYRFLREYALOHPESSAIFDLIID
 70 AIE---NRKRLATVDI NCYSGEVSCEAGIATDFVFNALCWKYPQ-----REDFVDLFYKLRDFAFDHPESAEIYAAID
 58 DNA---HEVICPNCFRDLSCEAFGIVVCDIYVSRLSFTN-----PELGRHYHILRTFALDHPESSAIFQAI
 68 AKA---SRVVRWHINCYSDMCAEAGIATDFALCHLAEKTL-----DDAIERYHILREFAVCHVESANILRLAID
 50 EIA---APVSVRWNLSYEGMCEAGIATDFALCHMAECFG-----DETHIEHYHILREFAVCHVESREIFRAID
 64 QSP---ERLAVLRSNDFCEVSDAAGIATDFVLGALCWIEQEEL---REKFATHFYQLRDFALEHEEASVILRAID
 69 GRR---DEMLNIEAM YFSCNSAAGIATDFLYSHLSFHTEGAD---QERFSRLYHSLRDMACEHDEKEFI LAAID
 64 PLK---KEGYTTSINCFCVLTAEAGIATDFAYYMWEMTD-----DDLFRVRLYHSLREYALTLLEHSIIFSAID
 71 DIN---ADTLTFNMEKNEAVSCEAGIATDFILYSIWSFQTK-----SEVMCDRFYQLRDYALOHSESAFI FHLIID
 71 YLP---QETLTFNPHNCNEAVSCEAGIATDFMMYSLWSFQTE---SNTLVDRFYQLRDYAAOHPERSAIFHLIID
 66 DLN---QEELTFNQHNCNEAVSCEAGIATDFMMYSLWSFQTE---SDILVDRFYQLRDYAAOHPERSAIFHLIID
 68 DLN---QEELTFNPHNCNEAVSCEAGIATDFMMYSLWSFQTE---SDTLVDRFYQLRDYAAOHPERSAIFHLIID
 75 NLN---QETLTFNPHNCNEAVSCEAGIATDFMMYSLWSFQTE---SATLVDRFYQLRDYAAOHPERSAIFHLIID
 74 DMI---EITLPMFPHNCNEAVSPEAGIATDFMYSIWSFKTE---SSVLVEYFYQLRDYAMQHPESQAIFHLIID
 74 DML---ETSLQMFNPHNCNEAVSPEAGIATDFLVYSIWSFKTE---SPVLVEYFYQLRDYAIQHPESQAIFRLIID
 69 DCD---RVHMINGENFDRTVSADAAGIATFSLAINRCWTHHDCGNAALTRLYMLRDTQLMNHITFHPECNAIYAAID
 67 DCG---RVHPEVSSENFDRTVSADAAGIATFSLAINRRLLAAHHDSNPALTRLYMLRDAQLMNHITFHPECNAIYAAID
 84 DGD---RFHITNDFNFRDVSADAAGIATFSLAINRRLLWHLHHRGDAALTRHMLRRESQ LWNFI DSHDECAA IYAAID
 66 DGD---RFRHMVNGENFDRTVSADAAGIATFSLVINRQLWLYHDSGDAELTQLYRMRDAQLMWRHIEFHPECNAIYAAID

Q08L07_BOR
 Q92ZJ3_RHI
 Q11AF6_MES
 Q11BN6_MES
 A5VGD6_SPH
 A6X3S7_OCH
 Q2VL95_BUR
 QOVU0_92Z
 A6H932_92Z
 Q2WHY8_92Z
 Q1H9Y8_92Z
 A9IES9_BOR
 Q8GFU1_CIT
 Q6TFW1_ERW
 Q5WS99_LEG
 A1AFJ8_ECO
 B1LHU7_ECO
 B1XDW0_ECO
 B1XDZ1_ECO
 B3I563_ECO
 B1IPT6_ECO
 A9MZ05_SAL
 A8GCN7_SER
 A7FER0_YER
 A9J566_9CA
 Q1NPW6_9DE
 B2JXY4_BUR
 B1F9D1_9BU
 B2JKZ9_BUR
 Q68E16_AER
 A4SUM8_AER
 B4F2P9_PRO
 B2Q7G2_PRO
 B2PWJ6_PRO
 B2PVX8_PRO
 B2PWU0_PRO
 B2PUC0_PRO
 B2Q735_PRO
 B2PWT9_PRO
 Q4L1R3_9EN
 A6TIQ7_KLE
 A6TID3_KLE
 Q5JBJ6_ECO