

Table S1

Table S1: List of protein sequences with DisProt/PDB IDs
(the Disprot dataset has the induced fold IDPs highlighted in red and the unfoldable IDPs are in blue)

PDB dataset	#	PDB ID	sequence
	1	1A32	LTQERKREIEQFKVHENDTGSPEVQIAILTEQINNLINEHLRVHKKDHHSRRGLLKMVGKRRRLAYLRNKDVARYREIVEKGL
	2	1BKR	KSAKDALLLWCQMKTAGYPNVNIHFTTSWRDGMFALIHKRPDLIDFDLKKNAHYNLQNAFLAEQHLGLTKLLDPEDISVDHPDEKSIITYVVTYYHYFSKM
	3	1BM8	QIYSARYSGVDVYEFIHSTGSIMKRKKDDVWNATHILKAANFAKAKRTRILEKEVLKETHKEKVGQGGFKYQGTWVPLNIAKQLAEKFSVYDQLKPLDFD
	4	1BZ4	SGQRWELALGRFWDYLRWVQTLSEQVQEELLSSQVTOELRALMDETMKELKAYKSELEEQLTPVAEETRARLSKELQAAQARLGADMEDVCGRLVQYRGEVQAMLGQSTEELRVLASHLRKLRKLLRDADDLQKRLAVYQAG
	5	1CDZ	ELPDFQGGKHFLLYGEFPGDERRKLRVYVAFNGELEDYMSDRVQFVITAQEWDPSEFEEALMDNPSLAFVRPRWIYSCNEKQKLLPHQLYGVVPQA
	6	1CQY	TPVMQTIWKNVPTTIGDVTYITGNRAELGSWDTKQYPIQLYDSSNDWRGNVLPALRNIEFKAFIKSKDGTVKSWQTIQQSWNPVPLKTTSHSSW
	7	1DSL	TFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTEVHSLNVLEGSVWLYEMPSYRGRQYLLRPGEYRRLYLDWGANNAKVGSLRRVMDFY
	8	1DT4	MKDVEIAVPELVGAILGKGGKTLVEYQELTGCRIQSKKGEFLPGRNRKVTITGTPAATQAAQYLITQRI
	9	1DXS	SLVSFLTGLGCPNCIEYFTSQGLQSIYHLQNLTIEDLGAIKPEQYRMTIWRGLQDL
	10	1EAN	GELVRTDSPNFLSSVLPTHWRSNKTLPIAFKVALGDVPDGLTVTMAGNDENYSAELRNATAAMKNQVARFNDLRFVGRSGRGKSFTLTITVFTNPPQVATYHRAIKITVDGP
	11	1EOE	ERVVINVSGLRFETQLKTLNQFPDITLLGNPQKRNRYDPLRNEYFFDRNRPSFDAILYFQSGGRLRRPRNPLDVFSEEIKFYELGENAFERYREDEGF
	12	1EW4	MNDSEFHRLADQLWLTIEERLDDWDGSDIDCEINGVLTITFENGSKIINRQEPHLQVWLATKQGGYHFDLKGDEWICDRSGETFWDLLEQAATQAGGETVSFR
	13	1EYH	HNYSEAEIKVREATSNPWPSSLSMSEIADLTYNVVAFSEIMSMWKRNLNDHGKNRWHVYKAMTLMLEYLKTGSERVSQQCKENMYAVQTLKDFQVYDRDGDQKQGVNVREKAKQLVALLRDEDRLREERAHKTKKELAQTA
	14	1FOM	YTSFNTVDEWLEAIKMGQYKESFANAGFTSFDVVSQMMMEDILRVGVTLAGHQKILNSIQVMRAQMNQIQ
	15	1GV2	ELIKGPWTKKEEDQRVIKLVQKYGPKRWVIAKHLKGRIGKQCRERWHNHLNPEVKKTSWTEEDRIYQAHKRLGNRWAEIAKLLPGRDINAIGNHNSTMR
	16	1GYU	MIPSITAYSKNGLKIEFTFERSNTNPSVTITQASNSTELDMTDFVQAAVPKTFQLQLLSPSSSVPAFNTGTITQVIKVLNPKQQLRMRIKLTYNHKGSAEQDLAEVNNFPPQSWQ
	17	1H3Q	MSGSFYFVIVGHNDNPFVEMEFPPGKAESKDDHRLNQFIAHAALDLVDENMWLSNNMYLKTVDKFNEWVFAFVTAAGHMRIFMLHDVVRQEDGIKNFFTDVYDLYIKFAMNPFYEPNSPIRSSAFDRKVQFLGKHKHLS
	18	1HUF	LSLSDLHRQVSRVQQESGDCTGKLRGNVAANKETTFQGLTIASGARESEKVFQVTLVSHVANVLTQEDTAKLLQSTVKHNLNNYDLRSVNGNSVLSLRSDQMTLQDAKVLLEAALRQES
	19	1I2T	HRQALGERLYPRVQAMQPAFASKITGMLELSPAQLLLLLASEDSLRLARVDEAMELIIAHG
	20	1J27	MKAYLGLYARLETPARSLKEKRALIKPALERLKAARFVSAARLYGLDAWGYEVGFTLLGNPAAWVEETMRAAARFLAEAGGFQVALEEFRLAEFEL
	21	1J74	VKVPRNFRLLEEEGQKGVGDGTVSWGLEDDEDMTLTRWVTGMIGPPRTNYENRIYSLKVECGPKYPEAPPSVRFVTKINMNGINNSSGMVDARSIPVLAKWQNSYSIKVVLQELRRLMMSKENMKLPQPPEGQTYNN
	22	1JB3	ELQRREEEANVLTGTVEEIMNVDPVHHTYSCKVRVWRYLKGKDIVTHEILLDGGKVVIGFGDPLICDNQVSTGDTRIFFVNPAPQYMWPAHRNELMLNSSLMRITLRNLEEVEHCVEEHRKLLA
	23	1JMW	MNQQGFVINSNELRQQQSELTSTWDLMLQTRINLARSARMMDASNQSSAKTDLQLNAKTTLAQAAAHYANFKNMTPLPAMAEASANVDEKYQRYQAAELIQFLDNGNMDAYFAQPTQGMQNALGEALGNYARVSENLYRQTF
	24	1JOS	RSDRVAQEIQEIAVILQREVDPKPRIGMVTVSDVEVSSDLSYAKIFVTFFLDHDHDEMAIEQGMKLEKASPYRSLGKAMRLRIVPEIRFIYDQSLVEGM
	25	1K40	NDKVYENVTLGVKAVIEMSSKIQAPPEEYVPMVKEVGLALRTLATVDETIPALPASTHREIEMAQKLLNSDLGELISKMKLAQQYVMTSLQQEYKQMLTAAHALAVDAKNLLDVIDQARLKML
	26	1K6K	MLNQELESLNMAFARAREHRHEFMTEHLLALLSNPSAREALEACSVDLVALRQELEAFIEQTPVLPASEEERDTQPTLSFQRVLRQAVFHVQSSGRNEVTGANVLAIFSEQESQAAYLLRKHEVSRLDVNVFISHGT
	27	1KU3	ELEKALSKLSEREAAMLKMRKGLIDGREHTEEVGAYFVTRERIRQIENKALRKLKYHES
	28	1L2P	TDQLKKAQAEQVIEQANKRRSQILDEAKAEAEQERTKIVAQAQAEIEAERKRAREELRK
	29	1LMI	SAYPITGKLGSELTMTDVTGQVVLGWKVSCLKSSATVPIPGYPVAGQVWEATVNAIRGSVTPAVSQFNARTADGINRYVWLQAAAGPDTISGATIPQGEQSTGKIYFDVTPSPPTIVAMNNGMEDLLIWEP
	30	1LN4	MDLSTKQKQHLKGLAHLPLKPVVLLGSNGLTEGVLAIEQALEHHELIKVKIATEDRETRKTLVIAVRETGACNVQVIGKTLVLYRPTKERKISLPLE

Table S1

31	1LOU	MRRYEVNIVLNPNDLSQLALEKEIIQRAAENYGARVEKVEELGLRRLAYPIAKDPQGYFLWYQVEMPEDRVNDLARELRIRDNRVRRVMVKSQEPP
32	1LU4	ADERLQFTATTLSGAPFDGASLQGKPAVLWFWTPWCPFCNAEAPSLSQVAAANPAVTFVGIATRADVAMQSFVSKYNLNFNLNDADGVIWARYNPWPQPAFVFRADGTSTFVNNPTAAMSQDELSGRVAAL
33	1MB1	NQIYSARYSGVDVYEFIHSTGSIMKRKDDWVNATHILKAANFAKAKRTRILEKEVLKETHKEVQGGFGKYQGTWVPLNIAKQLAEKFSVYDQLKPLF
34	1MHN	LQQWKVGDKCSAIWSESDGCIYPATIASIDFKRETCVVYTYGYNREEQNLSDLLSPICE
35	1MJC	SGKMTGIVKWFNADKGFITPDDGSKDVFVHFSAIQNDGYKSLDEGQKVSTIESGAKGPAAGNVTSL
36	1NG6	MSLLERLNQDMKLYMKNREKDKLTVVRMVKASLQNEAIKKKDSLTEDEELTVLSRELKQRKDSLQEFNSANRLDLVDKQKELDILEVYLPEQLSEELRTIVNETIAEVGASSKADMGKVMGAIMPKVKGKADGSLINKLVSSQLS
37	1O3X	AANKLIKEMVQEDQKRMEKISKRVNAIEEVNNVVKLLTEMVMSSHSGGAAAGSSEDLMKELYQRCERMRTLFRASDTEEDNEALAEILQANDLNTQVINLYKQLVRGEEV
38	1O8X	GLDKYLPGIEKLRRGDGEVEKSLAGKLVFFYFSASWCPARGFTPLQIEFYDKFHESKNFEVFCWDEEEDGFAGYFAKMPWLAVPFAQSEAVQKLSKHFVNESIPTLIGVDADSGDVTTRARATLVKDPEGEQFPWKDA
39	1OZ9	KNRVLVKLKRKVRKDKIEKWAELALSALGLNNVELSVYITDDQEIRELNKTYYRKKDKPTDVLSPFMGEEFFGGYKILGDVVISQDTERQARELGHSLSEEVKRLIVHGIVHLLGYDHEKGGEEKKFRELENVYLSKLSK
40	1P4P	SKLTRSNGTTLEYSQITDADNATKAVETLKNKSIKLEGLSVGKTTVEIKEGTVTLKREIEKDGKVKVFLNDTAGSNKKTGKWEDSTSTLTISADSKTKDLVFLTDGTITVQQYNTAGTSLEGSASEIKNLSELKNALK
41	1PGX	ELTPAVTTYKLVINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWWTYDDATKFTVTVMTEVPVA
42	1Q2Y	MKAVIAKNEEQKDAFVYREEVFVKEQNPVAAEEIDELENESEHVVYDGEKPVGAGRWRMKDGYGKLERICVLKSHRSAGVGGIIMKALEKAAADGGASGFILNAQTQAVPFYKHKHYRVLSEKEFLDAGIPHLQMMKD
43	1QAU	NVISVRLFKRKVGGGLFVLKERVSKPPVVISDLIRGGAEEQSGLIQAGDIIAVNDRPLVDLSYDSALEVLRGIAETHVVLILRGPEGFTTHLETTFTGDGTPKTIKRVTOP
44	1QZM	SGYTEDEKLNIAKRHLHPKQIERNALKKGELTVDDSAIIGIYYTREAAGRGLEREISKLRKAVKQVLLDLSKSLHIEINGDNLHDYLVGQRF
45	1R69	SISSRVKSKRIQLGLNQAELAQKVGTTQQSIEQLENGKTRPRFLPELASALGVSDWLLNGT
46	1R6J	GAMDPRTITMHDKSTGHVGFIFKNGKITSIVKSSAARNGLLTHEKNICEINGQNVIGLKDSQIADILSTSGTVTITIMPAF
47	1R9H	KIDITPKKDDGGVLLKIKKEGGVVKPTGTTGVKHYVGTLENGTKFSSRRDQDFSNLGRGNVIGKWDLVATMTKGEVAEFTIRSDYGYDAGSPPKIPGGATLIFEVELFEWSA
48	1SRV	GYQFDKGYISPFVVTNPETMEAVLEDAFILIVEKVSINPRELLPIEQVATGKPLLIADVEGEALATLVNKLRTLSVAAVKAPGFGDRRKEMLKDIAAVTGGTVISEELGFKLENATLSMLGRAERVITKDETTIVGGK
49	1T3Y	ATKIDKEACRAAYNLVRDDGSAVWVTFKYDGTSTVPEGEQAEYHQFQCTDDVRLFAVFRFTGDAMSKRSKFALITWIGENVSLQRAKTGTDKTLVKEVVQNAFAKEFVISRKELEEDFIKSELKKA
50	1TG0	EPEVPFKVVAQFPYKSDYEDDLNFEKQDEIIVTSVEDAEWYFGEYQDSNGDVIEGIFPKFSFVAVQG
51	1TIG	INVKEVRLSPTIEEHDFTNKLNRARKFLEKGDVKVATIRFKGRAITHKEIGQRVLDRLSEACADIAVETAPKMDGRNMFLVLAPKND
52	1TP6	CAYRREIHHAAVAIRDWLAGDSRADALDALMARFAEDFSMVTPHGVVLDKTAGELFRSKGGTRPGLRIEIDGESLLASGVGDATLAYREIQSDAAGRSERLSTVVLHRDDEGRLYWRHLQETFCG
53	1TQG	GSHMEYLVGFVDEKEYLQNLNDTLELEKNPEDMELINEAFRALHTLKGMAGTMGFSSMAKLCHTLENILDKARNSEIKITSDLLDKIFAGVDMITRMVDKIVS
54	1TSF	QGVELIARDWIGLMVEVVEVSPNHSEVGIKGEVDEQNTLKMTEKGLKVVAKRGRTRFRVYKGMIRIKGDLINFRPEDR
55	1TUC	MGPREVMTKKGDILTLLNSTNKDWWKVEVNDRQGFVPAAYVKLDSGTGKELVLALYDYQE
56	1TZV	MKTPRRRMRLAVFKALFQHEFRREDELEQILEEILDETYDKAKEDARRYIRGIKENLSMIDDLISRYLEKWSLNRSLVDRNRLRATYELLFEKDIPIEVIDEAIEIAKRYGTENSGKRVNGILDRIAKEHAPKEKFE
57	1UCS	NKASVVANQLIPINTALTLLMMKAEVVTMPGIAEIEPKLVGMQVNRVPLGTTLMPDMVKNYE
58	1VCC	MRALFYKDGKLFDTNDFLNPVSDDNPAYEVLQHVKIPTHLDVYVYEQTWEEALTRLIFVSGSDSKGRRQYFYGKMHV
59	1WKA	GKLYTLRYEVEGGGFIEIATVRPETVFADQAIIVHPEDERYHLLGKRARIPLTEVWIPILADPAVEKDFGTGALKVTPAHDPLDYEIGERHGLKPVSVINLEGRMEGERVPEALRGLDRFEARRKAVELFREAGHLVKEEDY
60	1WPA	WIREYPPITSDQQRQLYKRNFDLGLQEYKLSQSELEINKEKLSRLDKELDDYREESEYMAAADEYNRLKQVKSADYKSKKNHCKQLKSKLSHIKKMVG DYDRQKT
61	1WVN	PLGSQTTHELTIPTNLLIGCIIQRGANINEIRQMSGAQIKIANPVEGSSGRQVITITGSAASISLAQYLINARLS
62	1X6J	MDINNKARHWHACRRGMRELDISIMPFHEHEYDLSDEDEKIRFIRLECCDDPDLFNWLMNHGKPADAELEMMVRLIQTRNRERGPVAI
63	1XAW	WIREYPPITSDQQRQLYKRNFDLGLQEYKLSQSVLEINKEKLSRLDKELDDYREESEYMAAADEYNRLKQVKSADYKSKKNHCKQLKSKLSHIKKMVG DYDRQKT
64	1Y2F	RKEAVIMNVAHHGSELNELLNSIQAGFIFGDMNIYHRHLSPDGSGPALFSLANMVKPGTFDEPMKDFTPGVTIFMQVPSYGEDELQNFKMLQSAQHIADVEVGGVLLDQRRMMTPQKREYQDIIREVKDANA
65	1YU5	PTKLETFFLDVLVNTAAEDLPRGVDP SRKENHLSDEDFKAVFGMTRSAFANLPLWKQQLNKKKEKGLF

Table S1

66	1YZM	GSPLLQIHNITSFIRQAKAAGRMDEVRTLQENLRQLQDEYDQQQT
67	1ZHV	APRIKLIKNGSYGIARLSASEAIPAWADGGGFVSIITRDELSIVCLIDRIPQDVRVDPGWSCFKQGFPAFDETGIVLSVISPLSTNGIGFIVVSTFDGDHLLVRSNDLEKTADLLANAGHSLLLEHHHHHH
68	1ZZK	MGPIITQVITPKDLAGSIIKGGQRIKQIRHESGASIKIDEPLEGSEDRITITGTQDQIQNAQYLLQNSVKQYSGKFF
69	2A0J	SLIGEILPLSHIVLDMEVGSKKRLFEEAGLLERESSLSHADVFECLFAREKLGSTGLGQGVAIPIHGRHAGVKQATGAFIRTPREVGFDPADPGKPVSLIFILLVPENATGEHEVLVSKLAGKFSQKSIRESLMTVSSAEVRAIT
70	2B8I	IRPYMKALIYETLVNLANQDPEQHATIRQNLQEQLDLPFDKQLALYAGALGPASSGKLENHEAISNAVDSVQLLEI
71	2BK8	GAMVSGQIMHAVGEEGGHVKYVCKIENYDQSTQVWYFQVGRQLENSEKYEIYEDGVAILYVKDITKLDGTYRCKVVDYGEDSSYAEELFVKGVRE
72	2BZX	PVFAKAIQKRVPCAYDKTALALEVGDIVKVTMRNINGQWEGEVNKRKGLFPFTHVKIFDP
73	2CGQ	AMEEAINATIQRILRTDRGITANQVLVDDLGFDSLKLFQLITELEDEFDIAISFRDAQNIKTVDVYTSVAWWF
74	2CIU	SGDTQLFNRAVSMVEKNKDIRLLQCDDGITKERLKAYGELITNDKWTRNRPVSTKLDKEGRTHHYMRHFHVESKKKIALVHLEAKESKQNYQPDFINMYVDVPGEKRYLIKPKLHPVSN
75	2CJJ	GRPWSAKENKAFERALAVYDKDTPDRWANVARAVEGRTPEEVKKHYEILVEDIKYIESGKVPF
76	2CKX	RPFVSAEVEALVEAEHLGTGRWRDVKMRAFDNADHRTYVDLKDVKWTLVHTASIAQQRRGPEVPQDLLDRVLAHAHAYWSQQ
77	2CWP	MELYDVDEFWKFQMKVGLVKAIEKIKRTKKLIKLVDFNGEERTIVTGIADQIPPEELGKFFIVNLKPKKFSGVESQGMILAETEDGKYVLIPIVPEEVPVGARVW
78	2CXC	ITLEELRYISVFHSITGVTAIRCIVDEENRILFLVSEGEAGRAIGRGRILKILLREALGKNIEVVEYSSDLERIVKNLFPQVKIESINVRERNGVKQVVIKVEDDKGAAIGKGGKNVVRARLVLKSLFGVEKVVIR
79	2D59	TRPIDGLTDEDIREILTRYKIALVGASPKPERDANIVMKYLLLEHGVDYVVPNPKYEEVLGRKCYPSVLDIPDKIEVDLVFKPKLTMEYVEQAIIKKGAKVWVWQYNTYREASKKADEAGLIIVANRCMMREHERLLGEK
80	2E1F	QPVISAQEQTQVLYGKLVEARQKHANKMDVPPAILATNKILVDMAKMRPTTVENVKRIDGVSEGKAAMLAPLLEVIKHFQCTNSVQTDLFSS
81	2EA9	MSNTTWGLQRDITPRLGARLVQEGNQLHLADRASITGKFSDAECPKLDVVFPHFISQIESMLTTGELNPRHAQCVTLYHNGFTCEADTLGSCGYVYIAVYPT
82	2EIF	MVIIMPQTKQVNVGSLKVGQYVVIDGVPCEIVDISVSPKPKGGAKARVVGIGEFKVKKEFVAPTSSKVEVPIIDRRKQGVLAIMGDMVQIMDLQTYETLELPIPEGIEGEPGGEVEYIEAVGQYKITRVI
83	2END	TRINLTLVSELADQHLMAEYRELPRVFGAVRKHVANGKRVDFKISPTFILGAGHVTFYDLEFLRKRQIEIAECLKRGFNKIDTTVQDISDIPQEFRGDYIPHEASIAISOARLDEKIAQRPTWYKYKAIYA
84	2F15	QARPTVIRWSEGGKEVFISGSFNWSTKIPLIKSHNDFAILDPEGEHQYKFFVDGQWVHDPSEPVTSLQGTINLIHVKKSDFEVF
85	2F10	VVMNDIIVSIPVAEVDKHPEVLEILVELGFKPLANPLMRNTVGRKVSQKQSLAGTMDKIVRTLEANGYEVIGLD
86	2F19	HFPGRAPIDAYNGGFRFADMSHRGSIICIPSGIYDIDMTGPVPTQEDISRVLSESDQIEVLLIGTVELLRLPEELRVLLWEKRISDSTMSTGAAVRTFNVLAEADRAVAALLFAVE
87	2FO3	YRIQKELHNFLNPPINCTLDVHPNNIRIWIYVYGLENTYIYAEVYKLIIFPDDYPLKPPVYFLQKPPKHTHYVYNGDICLSLLGDDYNPSLSISGLVLSIISMLS
88	2FQ3	SKWFLNLEKHSIEVQSLPEFFTNRIPSKTPEVYMYRNFVNSYRLNPNYFVTTARRNVSGDAAALFRLHKFLTKWGLINYQV
89	2FZP	SSGLVPRGSTIEYNEILEWVNSLQPARVTRWGGMISTPDAVLQAVIKRSLVESGCPASIVNELIENAHERSWPQGLATLETQRMNRRYYENYVAKRIPGKQAVVVMACENQHMGGDDMVQEPGLVMIFAHGVEEI
90	2GKG	KKLIVESDTALSATLRSALTEGRFTVDETTDGGKSGVEQIRDRPDLVLAVDLSAGQNGYLICGKLLKDDDLKKNVPIVIGNPDGFAQHRKLAHADEYVAKPVDADQLVERAGALIGFPE
91	2H36	KQDLEKIESDIINDWTEADDLDDALDFLMEKVSEFKIKFDPLKVTEEEYRELLGNYDSSNSVSSNGITDQTYDEDDDIMYKLEFTYRKEDNKIYIEVQGWREK
92	2HAZ	SHMDTPSSPSIDQVEPYSSTAQVQFDEPEATGGVPILKYAEWRAVGEVWWSKWDYAKEASMEGIVTIVGLKPEPTYAVRLAALNGKGLGEISAASEFKTQPV
93	2HC8	EAIKLVGLQAKTAVVIRGKEIAVPVEEAVGDIVVPRGKEIPVDGVVVEGESYVDESMISGEPVVLKSKGDEVFGATINNTGVLKIRATRGGETLLAQIVKLVEDAMG
94	2HDZ	LPESPKRAEIIWQQSVIGDYLARFKNDRVKALKAMEMTWNMEKKEKLMWIKKAAEDQKRYERELS
95	2I6V	QEIFQYVRLSQVKRDKVLGYRVSQKDPVLFESIGLQDGDMAVALNGLDLTDPNVMNTLFQSMNEMTEMSLTVERDGGQHQHDVYIQF
96	2IGD	MTPAVTTYKLVINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGVWYDDATKFTVTE
97	2J49	PENYIRAYSMKLNWVSSLEIYKPELSYIMYPIFYFLNLVAKNPVYARRFFDRFSPDFKDFHGEINRFLSVNSIDHIKENEVASAFQSHKYRITMSKTTNLLLYFLNENESIGGSLIISINQHLDPNIV
98	2J9V	HHHMFNAKYVAEATGNFITVMDALKLNNAKDLQLHPLLAELLISINRVTRDDFENRSLKIDWVIRINKLSIGDITLQIRELLFDELAYKSFYALLD
99	2O0Q	TLIYKILSRAEWDAAKAQGRFEGSAVDLADGFIHLSAGEQAQETAAKWFRGQANLVLLAVEAEPLEDLKWAEASRGGARFPHLYRPLLVEVTRREADLDDLADGVPQLGDHLAL
100	2O71	HILNSSPSDRQINQLAQLRGPWEPMVLSGLSQTDIYRCKANHPHNVSQVQVEAFIRWRQRFGKQATFQSLHNLGRAVEVDPSSLHLMLE

Table S1

101	2OCH	KETGYDVLGVKPDASDNELKKAAYRKMALFKHPDKNPDGAEQFKQISQAYEVLSDKKRQIYDQGG
102	2OUM	ALLEKVDPNKIYTIIDEAAHLVKELATAKFDTEVEHAKLGDIPRRSDQNVRTVSLPHGGRIEFNRDKTGAIHAPVKGASFPPEKLANIRAFIRALEAHKPEGAKGTFLRSVYVTTMTGSPVRINPHS
103	2P5K	NKGQRHIKIREITSNEIETQDELVDMLKQDGYKVTQATVSRDIKELHLVKVPTNNGSYKYSL
104	2PND	GHPTLKTPESTGTWKWGDVQKQIYDPLRGYRQVLVKWLVRHGSDSVTIFLRDSTGDHIQQAKYRGRLLKVSCHKVPGDVSQINTLQMDDRNHYTEVWTQPDGQVIRDKIELRVRK
105	2PST	DDIQFQVVVNHHEEQYSIWPEYKEIPQGWRAAGKSGLKKDCLAYIEEVWTDMPRLSLRQHMD
106	2QVO	RIKLLFKEKALEILMTIYYESLGGNDVYIQYIASKVNSPHSYVWLIKKFEEAKMVECELEGRTKIIRLTDKGQKIAQQIKSIIDIM
107	2QVT	GYTRFYRSPTASVTLSGLVNVKWDNEQMTPLFKWIGGEQAEELHFCVHIAHSIGPKLNLRARTLGTVNSNMDQHWQAHRHSGATRRTIQDFHLFANDIPNFPDYIKKLVPK
108	2RH3	IQVFLSARPPAPEVSKIYDNLILQYSPSKSLQMILRRALGDFENMLADGSFRAAPKSYPIPHATAFEKSIIVQTSRMFVSLIEARNHFDPLGLETARAFGHKLATAALACFFAREKATNS
109	2V75	YTENLKVIVAEKLAGIPNFNEDIKYVAEYVLLIVNGGTVESVVDDELASLFDVSRDLANVVQTAFFALEALQQGESAEINVSKIRMM
110	2V9V	GSPEKILAQIQEHREGLDWQEAATRASLSLEETRKLQSMMAAGQVTLRVDLYAISTERYQAWWQAVTRALEEFHSRYPLRPLGAREELRSRYFSLPARVYQALLEEWSREGRLQLAANTVALAGFTPSF
111	2VC8	TDWLGSIIVSINCSDSLGVYQGRVSAVDQVSTISLTRPFHNGVKCLVPEVTFRAGDITELKLEIPGPGDNQ
112	2VO8	TTVGNSTIKVNDVEQVGSFAFEALGIEGLNGDTEVYSAEYLFYNAEAFILNEITSFNDSLVFKSKEVEPGKVRILVASLNGNEIEKSDLVKVNLTPKISSELEVGLTTLVAGAGDGNTHDLELSSKEVKINEE
113	2W0G	HKTFVKEKYEKQIKHFGMLRRWDDSKYLSDNVHLVCEETANYLVIWCIDLEVEEKCALMEQVAHQITVMQFILELAKSLKVDPRACFRQFFTKIKTADRQYMEGFNDELEAFKERVGRRAKLRIEKAMK
114	2W0I	MPLQPEAQRALQQLKQKMNVIQMKDLDERETIELVHTEPTDVAQLPSRVPRAAARYHFFLYKHTHEGDPLESVFYISMPGYKCSIKERMLYSSCKSRLSDSVEQDFHLEIAKKIEIGDGAELTAEFLDDEVH
115	2Y6X	PTGLTGNFREDTLALISSLREAIAPENDPNKKAQAQAEARKKLNDFALYRRDDSLRSLSSFMQATLNSLAGHYSSYPNRPLEKLRKARLEQEFKQVELALDREAKS
116	2Y9F	MTDIYIPPEGLYFRLGFASRQVIFARNSPSPDVGLSPVNDQATDQYFSLYGTGEHAGLYAIKSKATGKVLFSRRPAEYVGGIDGDRYPDNWFKIEPGKTYLSKYFRLVQVSTGTALVSRTHLQPYFVWHPQTEVFDQYFTFLFE
117	2YXY	EQKRYSEMTEKEELQEIAMLTEKARKAQGMVMEYAVYERKIAMAKAYMLNADPHFPEIYEIAGAPGEYFKVRYLKGVAWGWRLKGNEEEEALPISLLRKP
118	2ZQE	VKEVDLRGLTVAEALLEVDQALEEARALGLSTLRLLHGKGTGALRQAIAREALRRDKRVESFADAPPGEVGGHGVTVVALRP
119	2ZRR	GDDRRKKAIEVITELDDLEIDLGNESLRKVLGSYLKKLNEGTSVPLVLSRNMIEISNAIKKDGVSLENQSKKLKELMSIS
120	3A4C	RCPEQELRLQRLERLPELARVLRNVFVSERKPALTEMVFCARMVDSCQTALSPGEMEKHLVLLAELLDPDWLSLHRIRTDYVKLDAVDLAGLTARLAHVHVAEGL
121	3ADG	SHVFKSRLQEYAQKYKLPVYIEIVKEGSPHKSFLQSTVILDGVRYNLSLPGFFNRKAAEQSAAEVALRELAK
122	3AG7	GPLGSEIKNIDAKIRKWSGKSGNIRSLSTLQYLWVSGSGWKVPLMDMIEGNAVRSYQRALLILHPDKLQQKGASANQYMAEKVFELLQEAWDHFNTLGP
123	3CE7	VVDTDINAVTNYIVGMCQKFLQKGEKVTSSKLEELRTREDRLWDCLDTEVEFLDVEEIFDVTVPDEVADNFQTLQEIADFVVSERAKAG
124	3DJN	DMLPEIAAAVGLFSSLLRTRGCVSEQRKLVFSRALQDALTDHYKHHWFPEKPSKSGSYRCIRINHMDPIISKVASQIGLSQPQLHRLLPSELTLWVDPYEVSYRIGEDGSCVLYEE
125	3DKM	GRENLVYFQGLKYMVPGARVTRGLDWKWRDQDGSPPQEGTVTGELHNGWIDVTDWAGGSNSYRMGAEGKFDLKLAPGYDP
126	3E8V	AKGSVLVTDAGEQPVADATVEFKVYNYAEFYTVATKHTDRSGHASLTAGKGDMLVWASKDGRFGYSKLSFGKDNLKITLDK
127	3EGN	ECISRRELEKGRISREEMETLSVFRSYEPGEPNCRIVYKNAKHVQEKDLKIFGRYVDFSSSETQIRIMFDIRLMKEGRMKQAFIGLPLNEKAAAKLKEANGVYVFGKPMVQFARSARP
128	3FBL	YHKLRLAIKEICKTDGIPNIKWGMIAFGEKLLKSKMAGSASSDMAIEYINNAISAFSSRTGISQETAQKIADFITSNY
129	3FDR	LQLDKLVNEMTQHYENSVPEDLTVHVGDIVAAPLTPNGSWYRARVLTLENGNLDLYFVDFGDNDCPLKDLRALRSDFLSLPFAIEC
130	3FT9	VQVFTVQKGSDDPKLVLDIKYTRPGDSLAEVELRQHSGEWEPLTKKGNVWEVKSPLVGPFFNFRMSKGGMRNVFDEVIPTAFSIGKTYKP
131	3GK5	YRSINAADLYENIKAYTVLDVREPFELIFGSIANSINIPISLREKWKILERDKKYAVICAHGNRSAAAEVFLSQLGLNIVDVEGGIQSWIEEGYPVLE
132	3HJ7	KGYWFELPVPALLPLPNGYAIISFEFGEHYPRKQAGNDWVVDPASVSLPLRVTRRRRGRDMVLKGTGGTKLKEIFIEAKIPRMRERDRWPIVEDADGRILWVPLGLKSAFEAQRNGQARYILLQYQAMNS
133	3HNX	HMSYADSSRNAVLNNGRTLRAECRNADGNWVTSELDDTIIGNNDGHFQWGGQNFTEAEDIRFHPKEGAAEQPILRARLDCNGEFHDDRVDNLRNIQNVNGLRVFQ
134	3I8Z	FAVESIEKKRIRKGRVEYLKWRGWSPKYNTWEPENILDRLLIAFQNR
135	3ID1	MVRPVVGEIAANSIAEAQIAPGTELKAVDGIETPDWDAVRLQLVDKIGDESTTITVAPFGSDQRRDVKLDLRHWAFEPDKEDPVSSLGIRPRG

Table S1

136	3ID4	MIEPVLENVQPNASAASKAGLQAGDRIVKVDGQPLTQWVTFVMLVRDNPNGKSLALEIERQGSPLSLTIPIESKPGNGKAIGFVGIIEPKGKASPV
137	3IDW	KYDWFEFFLNCGVDSVNCQRYTINFREQLTEDMMPDINNMLRGLTGLREGDIVRMKHLDKKFGFR
138	3IHO	KWTPPRSPFNLVQETLFHDPWKLLIATIFLNRSTGKMAIPVLWKLEKYPYSAEVARTADWRDSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKYGNDYSYRIFCVNEWKQVHPEDHKLKNKYHDWLWENH
139	3IU5	SMTVDPIAVCHELYNTIRDYKDEQGRLLCELFIAPKRRNQPDYEVVSQIPDLMKIQQKLMEEYDDVNLTTADFQLLFNNAKSYKPDSPYKAAACKLWDLYLRTRNEFVQK
140	3JU0	SLTDSKVNAKSLKLEKYLTDGFGMHLLVHPNGSKYWRLSYRFEKKQRLLALGVYPAVSLADARQRREDAKLLAAGIDPSAKKQADNKTIQEKR
141	3JVE	APKPLHKVVVCSKSLKQSELNGIAASLGADYRWVDFEVTTHFYQGRPNNDTNEYKSVKERGVHIVSEHWLLDCAQECHKLPESLYPHTYN
142	3K3V	ESSWRYIDTQQGIHQPFQTTQMSQWYIGGYFASTLQISRLGSTPETLGINDIFITLGEMLTKLEKYDTPFTTFDKLHV
143	3KT9	MMRVCWLVQRDSRHQRIRLPHLEAVVIGRGPETKITDKKCSRQVQLKAECNKGYVVKQVGNPTSIDSVVIGKQDEVKLQPGQVLHMVNELYPYVEFEE
144	3L1X	SDAPDEFDRDPLMDTLPVRLPSGTIMDRSILRHLLNSPTDPFNRTLTESMLEPVPPELKEQIQAWMREKQNS
145	3LF9	MINEIKKDAQERMDKSVKALKNLKVRTGGGGTEERRKDLVKIRGEAEGGRVAVRNIARDAANDLAALGKDEKVNWFDISQALWEIQKLDVAVKIDEVLAAKEKELMEVLEHHHHH
146	3LLB	SYIKPLPSGDFIVKALTPVDADFDFGSEFSDDEEFDVGLVMSAFGHLPKRNEVVELGEFRFRVLNADSRRLHLLRSL
147	3LMO	TSTFDRVATIAETCDIPRETITPESHAIDDLGDLSLDFLDIAFAIDKAFGIKLPLEKWTQEVNDGKATTEQYFVLKNAARIDELVAAKGALE
148	3MX7	MTDLVAVWVVALSDGVHKEIEFHGTTSGKRVVYVDGKEEIRKEWMFKLVGKETFYVGAAKTKATINIDAIISGFAYEYTLINGKSLKYYM
149	3NDQ	RAPSTSDSVRLKCREMLAAALRTGDDYIAGADEEELGSGIEEAIYQEIRNTDMKYKNRVRSRISNLKDAKNPRLKNVLCGNIPDLFARMTADEM
150	3NE3	GIVGVGIDLVSIPDFAEQVDQPGTVFAETFTPGERRDASDKSSAARHAAARWAAKEAVIKAWSGSRFAQRVLPEDIHREIEVVTDMWGRPRVRLTGAIAEYLADVTIHVSLTHEGDTAAVAILEAP
151	3NPH	ELRSRSTEEVDAVILAVYRQVLGNDHLMQSERLTAESLLRGREISVRDFVRAVALSEVYRQFFHNSNPQRNFIELNYKHLGRAPYDQSEIAFHDTLYHQGGYEAIEINSYDSVEYTENFGDWVVPYFR
152	3O48	DFWPTLKDAYEPLYPQQLLELRQVQVSEGGPTATIQSRFNAYWGLIKSTDVNDERLGVKILTDIYEAESRRRECLYLTIGCYKLEGEYSMAKRYVDLFEHERNNKQVGLKSMVEDKIQKEENLYFQ
153	3O5E	SPTATVAEQGEDITSKDRGVKIKRVGNGETPMIGDVKYVYHYGKLSNGKKFVDSHDERNEPFFSLGKGQVIKAWDIGVATMKKGEICHLCKPEAYGSAGSLPKIPSNATLFFEIELLDFKGE
154	3OBS	SAVSESQKMKMVSQYKRDLTVRETIVNITLYKDLKPVLDVSYGTGSRLEMLNLTGTIPVYRGNTYNIPICLWLLDTPYNPPICFVKPTSSMTIKTGKHVDANGKIYLPYLHEWKHPQSDLLGLIQVMVVFGEPPVFS
155	3ONH	KVCRGVIKLSSDCLNKMKLSDFVVLIREKYSYPQDISLLDASNRQLLFDYDFEDLNDRTLSEINLNGNSIILFSDEEGDTMIRKAIELFDVDDDELPCNTCSLPDVEVPLIKA
156	3ONJ	LLISYESDFKTTLEQAKASLAEAPSQPLSQRNNTLKHVEQQQDELFDLLDQMDVEVNNSIGDASERATYKAKLREWKTIQSDIKRPLQSLVDSGD
157	3PDG	MVTIDSPVAGERFEAGKDNISATVSKTPVSKVEFYNGDTLISSDITAPYAKITGAAVGAYNLKAVAVLSDGRRIESPTVPLVKVIVL
158	3QD7	DTLQLDNFLTGTGFLDIPLSQPLEFRREGHQHGVLDKLRSGKYPQQASLNLLRQPVEECKRMVFSFIQALADGLRNLVLIHGKGRDDKSHANIVRSYVARWLTFFDDVQAYCTALPHHGGSGACVVALRK
159	3RJ2	IVANCEFVNATGKKTTLVNWENWAKYCWIWYKFPKYTLRYSVDGEMFMRHRVTFNATGRYITHLNLHGLEDVLEGLAVPKDAAAYARIAAINVSLTNPQGDVHMHYDETEGEQIRSYDAAEFARTLAA
160	3RJP	MYRDLVLPQNRSVNRGDDEISLTKREYDLLNILMTNMNRVMTREELLSNVWKYDEAVETNVVDVYIRYLRGKIDIPGKESYIQTVRGMGYVIREK
161	3T46	HQALVDQLHELIANTDLNKL SYLNLDAFQKRDILAAHYAKSAIRTKNLDQMTKAKQRLESYNSISNPLHSQNN
162	3V4Q	LAAKEAKLRDLEDLAREDTWRLLAEKEREMAEMRARMQQQLDEYQELLDIKALALDMEIHAYRKLLEGEER
163	3ZHI	KTDTSNRLKQIMAERNLKVQDILNLSIPFQKFGIKLSKSTLAYSASVQSPDQNRILYLLAKTLGVSEAWLMG
164	3ZZP	QGYTTWYQVEMPEDRVNDLARELRIRDNVRRVMVASTTPGRYEVNVLNPNLDQSQQLQNEKEIIQRALENYGA
165	4EIH	SLEKHSWYHGPVRSRAAEYLLSSLINGSFLVRESESSPQLSISLRYEGRVYHYRINTTADGKVVYTAESRFSTLAEVLVHHSTVADGLVTLTHYPAPKCN
166	4F25	SGNIFIKNLDKSIDNKALYDTFSAFGNLSCVVCDENGSKGYGFVHFETQEAERAIEKMNGMLLNDRKFVGRFKSRKE
167	4FI5	SMTMEELQREINAHGQLVIARQKVRDAEKQYKDPDELNKRRTLDRREGVAVSIQAKIDELKRLQADRIAT
168	4G3O	RGSENLYFQGLNAMAHIQEMFPQVPYHLVLQDLQLTRSVETITDNILEGRI
169	4GA2	RRSKADVERYIASVQGSTPSRQKSIGFYFAKLYEAKYDLAKKYICTYINVQREDPKAHRFLGLYLEENTDKAVECYRRSELNPTQKDLVLKIAELLCKNDVTDGRAKYWVERAAKLPFGSPAVYKLEQLLDCGEED
170	4GCO	YINPELAQEEKNKNEYFKGDYPTAMRHYNEAVKRDPENALYSNRAACLTKLMEFQRALDDCOTCIRLDSKFIKGYIRKAACLVAMREWSKAQRAYEDALQVDPSEEAAREGVRNCLR

Table S1

171	4GEI	KIKSFEVFNDEPEKVIYSGSERVAGRVIVEVSEVTRVKAVRILASGAVKVLWMQGSQQCAQTSYLRVEDTLLEDDQPTGENEMVIMRPGNKYEYKGFELPQGPLGTSFKGKYGSVDYVWKAFLDRPSQPTQETKKNFEVVDLV
172	4GMQ	NAVKKNRVLRGSVKEANYFVEGEASAATIDAVLNDVDLVITIKDADEIAALAGKLNGLTVADEIKNVWKEEVSRLVGAGKLEKGDIKALVA
173	4HQA	SRKFIIANARVENCAVIYCNDFCELCGYSRAEVMQRPCCTDFLHGPRTRQRAAAQIAQALLGAEERKVEIAFYRKDGSCFLCLVDVVPVKNEGDGAVIMFILNFEVVMKE
174	4HTJ	ARGYIVTRDRPLRPEEGRRLVEDVARLLQVPSAFADVEVLGPAVTFKVSANVQNVTTDEVEKATVDNKDKLEETSGLKILQTVGVS
175	4IGV	DLSGKMVKQVEILSDGIVFIEFRYRLYLISEMSPVNIQGVDLLEGNWGTGVSIVFFKYTIDGKEKTAKDVIDAEDEETKSVTFKIVEGDLMELYKTFIIVQVDTKGEHNSVTWTFHYEKLKEDVEEPNTLMNFCIEITKDIETYHLK
176	4IPC	GSHMVGQLSRGAIAIMQKGDNTNIKPIQVINIRPITGNSPPRYRLMSDGLNLTSSFMLATQLNPLVEEQSSNCVQIHRFIVNTLKDGRRVVILMELEVLKSAEAVGVKIGNPVVYNE
177	4J5Q	EDPEGSRITYVKGDLFACPKTDSLAKHICISEDCRMAGIAVLFKFKFGVQELLNQKQKSGEAVLKRDRYIYLLITKRASHKPTYENLQKSLKAMKSHCLKNGVTDLSMPRIGCGLDRLQWENVSAMIEEVFEATDIKITVYTL
178	4JOX	AREVHIDVNNKTGHTLQLEDKTKLDGGRWRTSPTNVANDQIKTFVAESNGFMTGTEGTIYSINGEAEISLYDNPFAGSNKYDGHNSKSQYEITQGGSGNQSHVYTIQTSSRYG
179	4KMC	TNWSRRYKANLEKLASGDVNKVAEVRDLWRRDQERLSAGEKRMKAKARQILVGELEAESTDDAKAETILDEVLAAA
180	4LTT	MLKLNKKSFKQDFDKLLNGFDDSVLNEVILTRKKEPLDPOFQDHALKKGWPFRECHIKPDVLLVVKDDELILLRLGSHSELFLEG
181	4MNO	VIRVPLPEGNQLFGVVEQALGAGWMDVRCEDGKIRRCRIPKLRRRVWIRVGLDVLIVQPWPVQSDKRGDIVRYTQTQVDWLLRKGKITQEFLTG
182	4MQ3	MGIWQMDCTHFDGKILVGHVESGYIWAQIISQETADCTVAVLQLLSAHNVELQTDNGPNFNKQKMEGLNYMGVKKHFGIPGNPQSQALVENVNTLKVWIIQKFLPETTSLDNALSLAVHSLNKK
183	4N1L	SSFFPDFGLLLYLEELNKEELNFKLFLKEMPEHGLTPWNEVKKARREDLANLMKYYGKAWSVSLKIFGKMNKDLKERAKEEINWSAQL
184	4N6T	ENSLIEELARFAVEDEHNKENALLEFVRVVKAKEQVAVGTMYYLTLEAKDGGKLLYEAKVVVWPWFENKLEQEFKPV
185	4NPN	HINLKVAGQDGSVQFKIKRHTPLSKLMKAYCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDEDITDVF
186	4O7Q	AMESKYKEILLTGLDNITDEELDRFKGLSDEFNIATGKLTANRIQVATLMIQNAGAVSAVMKTIIRIFQKLNMYMLLAKRLQEEKEKVDKQYK
187	4O9L	PHMTVAEDKTFQYRQHNSFRIHVLRILPYLSCLTSDQDRLRATYERWGNQDTLLELFTSLRRRNGWVHSLIGALRACELSGLADEVARIYHS
188	4OD6	RFDNELQLRGLSVEAAVEELRAAIEARALKETPLRVVHGKMGVLRRTLRDYLKTDKNVESFHDAEANQGGHGVITIVNVKR
189	4OIX	MLKRMVYARVYGLVQGVGRKFKVQIHAILRGIKGYAKNLPDGSVEVVAEGYEEALSLLERIKQGPAAEVEKVDYSFSEYKGEFDEFETY
190	4PGO	EILSKKPAGKVVVEEVNIMGKDVIIIGTVESGMIGVFKVKGPSGIGGIVRIERNREKVEFAIAGDRIGISIEGKIGKVKKGDVLEIYQT
191	4Q2S	IAESLRRKKEYVKAQSGKCEVAEWCNMPVAVGFDVLEISYDRMLENCSNLLLLTFIYHISLLDSVDDRLSKRMEYISRICLNIDVNDPKVETVHPVLTLTREALLRQSEFFSPIFKRRLVLLRALDQKISEI
192	4Q6L	YDWNAMQPLVSKMLQADGVTAGSVLLVDSVNNRNTGSLNANAEATETLRNALANNGKFTLVSVQQLSMAKQQLGLSPQDSLGRSKAIGARNVGAQYVLYSSASGNVNAPALQMLMLVQTGEIWSGKGAQQ
193	4QB0	EHSFEEMRHLRSQGFDAVLYHMMKDEPVVFTSDGKEYTYPDSLEEEYPPWLTEKAMNEENRFVTLGGQFYWPMNHKNKFMALQHHQ
194	4QCJ	LPAGSALLVVKRGNAGARFLLDQPTTAGRHPESDIFLDDVTVSRRAEAFRINEGEFEVVDVGLNGTYVNRPRNAQVMQGTGDEIQIGKFRVLVLAGPA
195	4QUC	EYVVEKILGKRFVNGRPQVLVKWSGFPNENNTWEPLENVGNCMKLVSDFESEVFRL
196	4RWU	MVKETKLYDLLGVSPSANEQELKKGYRKAALKYHPDKPTGDTEKFEISEAFEILNDPQKREIYDQYGLEAARSGGPSFGP
197	4W8G	TYDVNICYAPEDREWVINTLVFKLERAGIKTFVNIRDDTPGNFFAENIMDAIENSRTIVMSPDFFKNNICDKTLQIGLSHQIIPILYRCPVYFLNHMTYLDWCDKDVPRVFWRNLFDRIN
198	4YKD	ASATELLQDYMILTRLTKLSSQEQIAALLHEYRNGASIHFCINLRQLYGDSTRKFLLLGLRPFPEKDSQHFENFLETIGVKD
199	4ZC3	LSPKQERFIEEYFINDMNATKAAIAAGYSKNSASAIGAENLQKPAIRARIDARLKEI
200	5CVZ	AAATSLVYDTCYVTLTERATTSFQRQSFPTLKGMDRAFQVVAFTIQGVSAAPLMYNARLYNPGDTSVHATGVQLMGTVPRTVRLTPRVGQNNWFFNGNTEEAETILAIIDLTVSTKGANAPSNTVITGCFRLAPSELQSS
201	5CW9	MEMDIRFRGDDEPAAAYKALREMRQARKFAGTVTLIIRFRGDDELEALEKALKEMIRQARKFAGTVTYTLGNDLEIRITGVPPQVILELVEAIRLAKEFNITVTELVRITGVPEQVRKELAKEAERLAKEFNITVYTI
202	5DBL	HMIAPGHRDEFDPKLPTEKEEVPKPGIKNPETGDDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDLPAGTEKVTREGQKGEKTIPTTLKLNPLTGEIISKESKEEITKDPINELTEWGPET
203	5DMA	GWKVGDRANHRKWIGITVSVRGGDDQELDIAPSPIGIKRLLAKFAPIEKV
204	"5E10"	GVQVPDVRGQSSADAIATLQNRGFKIRTLQKPDSTIPPDHVIGTDPAANTSVSAGDEITVNVSTGPEQREIPDVSTLYAEAVKLTAAAGFRFKQANSPTPELVGKIGTNPANQTSAITNVIIIVGSG
205	5E1Y	SMEILQVALHKRDSGEQLGIKLVRRTDEPGVFLDLEGLLAAQDGRSSNDRVLAINGHDLKYGTPELAAQIIQASGERVNLTIARPGKPEIEL

Table S1

206	5EEP	GHPFGTTVTAETLRNTFAPLQWEDKYRQLIMLQKLPALPDELKAQAKEIAGCENRVWLGYTVAENGMHFFGDSEGRVIRGLLAVLLTAVEGKTAELQAOQSPALFDELGLRAQLSASRSQGLNALSEAIIAAAKVQ
207	5EFX	GARLQEIYNRMDPRAQTPVPGKPGFGREELLRRKLIHDGCLLWKATGRFKDVLVLLMTDVLVFLQEKDQKYFFPTLDKPSVSLQNLVLRDIANQEKGMFLISAAPPEMEYEVHTASRDDRSTWIRVIQSSVRTCPSREDFPLI
208	5EJO	AMITDPMDLLRFDGVDSTFSLGTVTEIAQKNLPQYNKQTKNTIKEYAIRSSGKGLPRKWKVIKDAQNWENLRANAN
209	5EQZ	HHHYVEEKKEIDSLMEDVLALVNDSSGGKFKDYDKINELKENLKDIGNAELKEKLLNLQNSFQDKLAAKLAALKAAKNTIENITDKDQDISKRKIWSEAKLVGTVPLLSNTSGNGDKMSKNAVEQIDKVIKFLKE
210	5EWR	DPRAYFPAPADLVEILDVQQASHYKQIKKGLNEVLKSMNRGLAEFVLAADTQPLEILLSAPLVAEDKAVPYVFPVSKAALGRACGVSRPVIACAVLRADMSQLRNQITARTKIEQLL
211	5FD9	SPTKAIVRLREHINLLSKKQSHLRTQITNQENEARIFLTKGNVMAKNALKKKKTIEQLLSKVEGTMESMEQQLFSIESANLLETMRAMQEGAKAMKTIHSGLDIDKVDETMDEIREQVELGDE
212	5FU5	GPVAGNYVHDFTANGTSSSFYTIAGNLSTSGKTATYNGKTLTQCLKMETATSISFTAPSAGKTLVFAEAAATAKVDGNKVTASNGIITVDLAQGAHTITKADACNLFYMEYAALEH
213	5GUA	MENVVSAPMPGKVLRLVVRVGRVVRVGGQLLVEAMKMENEIPSPRDGVVVKRILVKEGEYVDTGQPLIELG
214	5H1D	PNVVVTRLTLVCSAPGPITMDLTGDLEALKKETIVLKEGSEYRVKIHFKVNRDIVSGLKYVQHTYRTGVKVDKATFMVGSYGPREEYEFLLTPVEEAPKGMMLARGTYHNKSFFTDDDDKQDHLSEWNLKIKKEW
215	5HBQ	TTISPHDAQELIARGAKLDIRDAEYLRHEIPEADLAPLSVLEQSGLPKLRHEQIIFHDQAGKRTSNNADKLAIAAPAEIFLLEDGIDGWKKAGLPVAVN
216	5HDW	SEFVATTGDIIVSVSTSFPELSSVHPPHYFFTYRIRIEMSKDALPEKACQLDSRYWRITNAKGDVEEVQGGVVGFEPIISGRVYVYTSCTTFTSTSGYMEGYTFHFYFKDKIFNVAIRPFHMACPT
217	5HFD	GSPEFLGEEDIPREPRRIVIRHGSTGLGFNIVGTEDGEGIFISILAGGPADLSEGLRKGQDQILSVNGVDLRNASAEQAAIALKNAGQVTVIAQYKPEEYSRFEANSRVDSSGRIVTD
218	5HU4	DAVVGSIAPVSDVNLVFKGTNTANLLAGATTMRSDQVMGKGNYPHAGHMRDESMLFGPIMKVKKGDKIYLDLENLEYVTETKTIDETEVSVIDNTDKARITLITCDKPTETTTRKFVAVGELEKTEKLTKELENKYFPSK
219	5I41	SHMNTNMVASELGVSAKTQVRVWVKQLNLPARNELGHYSFTAEDVVKLKSVMKQISEGTAIQDIHLP
220	5IJM	SHMSAADFEAAVAVRSLPKEGVPQLDNAAKLQFYSYLKQATEGDTVTSQPWAVQVEARAKWDASNCSCKGMKSEDAKAAVRRLLTLRSQGIQWKPG
221	5IWH	GIKSFKHKGLKLLFEKGVTSQVPAQDVRINDRLQADTATEIGELNRQIYKHLPLKGDREGYWSITVRANWRITFQFINGDAYILNEDYKLGPEQ
222	5JBM	QKAMITDPMDLLRFDGVDSTFSLGTVTEIAQKNLPQYNKQTKNTIKEYAIRSSGKGLPRKWKVIKDAQNWENLRANAN
223	5JRT	VDFSITQFVRLNGLLEHLMIDIFEREQITLRLVEMGHKELKEIGINAYGHREKLIKVERLISGQ
224	5KIU	HHQKLSARLRALRQRQLDRAAAAVEPDVVVKRQALAAARLKMQEELNAQVEKHKEKQLLEEEKRRQKIEMWDSM
225	5KUX	GIVNVPNCNTTKYQQLARTAVAIYNYHEQAHLTFVENLNCKEQLGEGDYIYITLAATDDAGKAIYEAIGVVESAGWTGVVEEFKLVGSLEH
226	5L33	HMPEEEKAARLFIEALEKGPPELMRKVISPDTRMEDNGREFTGDEVVEYVKEIQKRGQWHLRRTYKEGNSWRFEVQVDNNGQTEQWEVQIEVRNGRIKRVITITHV
227	5ML3	SAKDERAREILRGFKLNWMNLRDAETGKILWQGTEDLSVPGVEHEARVPKIKKCAVSRNLSSTEQMEKFRLEQKVYFKGQCLEWFFFEFGFVIPNSTNTWQSLIEAAPESQMPASVLTGNVIETKFFDDLLVSTSRVLFVY
228	5MR1	ADCQGWLYKKKEKGSFSLSNKWKFWVILKSSLYWYSNQMAEKADGFVNLPDFTVRASECKKHAFKISHPQIKTFYFAAENVQEMNWWLNKLGSAVIHQ
229	5NOD	VTSVAMPYSYGSSLEFTKNLIQVGIKEANIEVEVTTAPAGSVEGMVVEQSPRAGEKVDLNKTRVKISYKPK
230	5OHQ	GPWVTTDIQVKVRDLYLDTQVVGQTVGIRSVTGGMCSVYLKDKSEKVSISSEHLEPITPTKNNKVKVILGEDREATGVLLSIDGEDGIVRMDLDEQLKILNRLFGKLE
231	5TPJ	MSREEIRKVVVEEMVRKLGQSPEDISKYLPDVRLEVGNYTFEGSEQVTKFWRMLTKFVDRVEVRKQVVDGNHVRVEVEWNGKKWTFEVEVEVRNGKIKRIRLQVDPEFKVQNIWNLL
232	5UFT	SMIDDIYNKRILEFAGNMERIGQLAEPDAVATVHSKLCCGTVTVYKMRDGVVTFDAHEVKACALGQASSVMARNVIGATADELRAARDAMYRMLKENGPAPEGRFADMKYFEPVRDYKARHASTLLTFDAVADCIRQIEEKA
233	5UMV	SSKIFKNVINYINGYTKPGRQLHEMIVLHGGKFLHYLSSKKTVTHIVASNLPLKKRIEFANYKVVSPDWIVDSVKEARLLPWQNYSLTS
234	5WD9	APVTELTRLKEYMEDQIAKAKESSSLTAQLKFLENAHTHFVKMGSLTTIYKGGSEVVDRLKIEIRSLYEEMLELKDCKRDIQIYQ
235	5WF2	AMTSDCQQLSLVQTVTRGSRAFLSREEAQHFVKECGLLNCEAVELELICHGGTSECVMRALGAVASLGCDDLPEHILLTRPRLQELSGAGSPVTKATKILRHFEEAC
236	5XUO	AGFGAQHRKMGKSLYLRNEVFAAWIEKVDALVQDELGYSLVELILDADAQDYGIETTQVTFIQAIALGELLRHHGAKPAAVIGQSLGEAASAYFAGGLSLRDAITRAICRS
237	5Y3C	PAAASTKVLYYTDRSLTPFLVNIPKRLGDVTLQDFKAAVDRHGSFRYHFKSLDPEFGTVKEEVQDDAIVIPGWEGKIVAWVEE
238	5Y4U	ARLTKLVNAAVPMFLMKGSPSEPKCGFSRQLVGLREHQVRFQFFDILRDESVRQNLKFFSEWPTFPQLYINGEFQGGLDIIKESLEEDPDFLQHALQS
239	5Y6H	YHEQFLKQNLAVLGVLRDLHKAAPLRLSWNGQLISKLAITPDKLVLDGFSQAEDNIAVLKAQHITITAETQAKAVEFTVEQLQQSEYLLQPAFITVPPPTLWFVLE
240	5YDE	MADVLSVLRQYNIQKKEIVKGDDEVIFGEFSWPKNVKNYVWGTGKEGQPREYITLDSLFLNVLHSHPVVYRRAATENIPVRRPRDKDLLGYLNGEASTSASIDRS

Table S1

241	5YDN	MNYATVNDLCARYTRRLDILTRPKTADGQPPDAVAEQALADASAFIDGYLAARFVLP LTVVPSLLKRCQCVVAFYLNESQPTEQITATYRDTVVRWLEQVRDGKTPG
242	5YRY	PEFMENLVTLSRQGIENLMKLENRKEPWIVVLYAPWCPCQAMEASYDELADKLAGSGIKVAKFRADGDQKEFAKQELQLGSFPTLVFPKNSSRPIKYPSEKRDVESLTSFLEH
243	5YWJ	LLGFYKQYKALSEYIDKKYKLSLNDLAVLDLTMKHCKDEKVLMSFLKTAMDELDLSRTKLLVSRRLIEKERLSKVRSSKDERKIYILNDDISKFNALFEDVEQFLN
244	5Z59	WPSAKFIDGRVAFSRMPAERELDEVARDFDAVVVLVEDYELPYSLDEWEKRGVEVLHGPIPDFTAPSVQEQLLEILRWIEERVREGKKVLIHCMGLGRSGTVGVAWLMSYRGLSLREALMEVRRKRP GAVETQEQMEVLKELEERI
245	5ZT3	YPYVDEIIGGDSVQSIQRRLGNTNWNPSAHDMQMSRIQAEDLFELKVEIIRKMAGLHPSGDWWMGWGARALDNPRATATGEEDLARLHQMLDDLQSRNEQSATFWRLVERVRLRA
246	6AR0	PSALAIFTCRPNSHPFQERHVYLDEPIKIGRSVARCRPAQNNATFDCKVLSRNHALVWFDHKTGKFYLQDTKSSNGTFINSQRLSRGSEESPPEILSGDIIQFGVDVTENTRKVTHGCIVSTIKLFLPDGMEARLRS
247	6B26	NSYVALYKFLPQENNDLALQPGRIMLVDDSNEDWWKKGIGDRVGFPPANFVQVRPGENVWRCCQPFSGNKEQGYMSLKENQICVGVGRSKDADGFIKRVSSGKKRGLVPVDALTEI
248	6B8C	VNGAAIAEAYKYGTPYVWGGKDPGDFDCSGFTRYVYLQVTRDIGGWTVPQESAGTKISVSQAKAGDLLFWGSPGGTYHVAIALGGGQYIHAPQPGESVKVGSVQWVAPDFAVSM
249	6BL5	RVPRHLHAYGVFALPFPMPDPVEWGNWFAGPHPKAFLVSVHPSGPKAGHVYPTDLSDDPSVANVIGMVLGDHGYEADHNVTVTLRAAVPIEVQQGIEAPPLQPDPAVLNAAPQLKLVKIGHYFFDYTR
250	6D1R	KAYRIKKNADFQRIYKGHSVANRQFVYVTCNNKEIDHFRLGISVSKLGNVLRNKKRAIRENFVKVHKSHILAKDIIIVARQPAKDMTTLQIQNSLEHVLKIAKVFN
251	6DGA	MKDDLNDNLLYRYCGATSPFWRLPLDSNALQLAASEEAVTSHVPLTPEQAAQIRTM SVITSSVTLSSLFGLVLPVHLVGRKVSRRKEWAGTASA
252	6E4N	GSHMRVQVSGMSDETTWHTLKDHLRQAGEVTFCKVFSGGRAVVEFVTPEDAARAITELQASELEGATLFLR
253	6ELM	SMAEDTGVRELAEEEDHGRKSTIALRLWVEDPKKLKGGPKDNGAIEFTFDLEKETPDEVAQEMIESGFFHESDVKIVAKSIRDVRVALIQWRRERIWPA
254	6EW6	DSQIQFTRHASDVLNLRNLRSDILTDVVIVVSREQFRAHKTVMACSGLFIYIFTDQLKRNLVSNLDPEINPEGFNILLDFMYTSRNLNREGNIMAVMATAMYQMEHVVDTARKFIKAS
255	6EXX	KISTLFLNLSAVCNKEFLKYLCHQENIRPFKIQIDGYDENSSTSYGFIKFRNFEDATRIFNLRNRLVGGSVITTSWE
256	6F53	LPTAQEVQGLMARYIELVDVGDIEAIVQMYADDATVEDPFGQPPHGREQIAAFYRQGLGGGKVRACL TGPVRASHNGCGAMPFRIEMWVNGQPACVDAISVMRFEHGRQIQTMOAYWSEVNLVS
257	6FC5	GPMDRYQRKIGCFIQIPNLGRGQLKYVGPVDTKAGMFAGVDLLANIGKNDGSMFGKYYFQTEYPQSGLFIQLQKVASLIEKASI
258	6GPM	GQVIVVRKDSDFRPKTYEDLVGKTVAVQIGTTGDIIEVSKYDGKVVRFDKFTDAFLELKRGRADAWLDSATARAFVAKNPDLVISSGVLSS
259	6IPY	TGVTLFAVLYDEAITEDDLSFHKGKFKQLNSSEGDWWEARSLTTGETGYIPSNVYAPVDSILEHH

Disprot dataset

#	Disprot ID	Sequence
1	DP00070	SKTKEGVVHGVATVAEKTKEQVTN VGGAVVTGVTAVAQKTVEGAGSIAAAT
2	DP00284	TNVLNQGGVEGYDSSSNVIQDSNLDTAINPDRPLSEIPTDDNPNMSMAEM
3	DP00284	YGVVKTTPARSLSAMAAATVAKSSDPADPPIPNDSRILFKEPVSSYKGEYPY
4	DP00088	MFGKLSLDAVPFHEPVMVTIAGIILGGLALVGLITYFGKWTYLWKEWLTS
5	DP00262	SDLACRLLGQSMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN
6	DP00974	MDSSAGPGNISDCSDPLAPASCSPAGS WLNLSHVDGNQSDPCGPNRTGLG
7	DP00808	MEDINFASLAPRHGSRPFMGNWQDIGT SNMSGGAFSWSGLWSGIKNFGSTV
8	DP00488	MEAHNQFLKTFQKERHDMKEAEKDEILLMENSRRFVFMPIKYHEIWAAYKK
9	DP00226	EKPSYDTEADPSEGLMNVLKKIYEDGDDMKRTINKAWVESREKQAREDETF
10	DP00282	NIEGLLHKNWNIEDNNIKNASLVQIDASNEQSSTSSMIIDAQISNALNAQ
11	DP00623	SNHTFYPMMAIAGECPEDPTKSKIRSDLTQYGGISTDMGKRYSES LAGSLLP
12	DP00807	PPGATPAPPGAAPPAAGGSPSPADGGSPPPADGGSPPPV DGGSPPPPSTH
13	DP00875	KVGSLLQYLALAAALITPKKIKPPLPSVTKLTEDRWNKPKQTKGHRGSHMTMNGH

Table S1

14	DP00793	ASVPAGGAVAVSAAPGSAAPAAGSAPAAAEKKDEKKEEESDDDMGFGLFD
15	DP00298	SDVEENRTEAPEGETESEMETPSAINGNPSWHLADSPAVNGATGHSSSLDAREV
16	DP00624	GKQRMKEKEQENERKVAQLAEENERLKQEIERTREVEATRRLIDRMVNLHQ
17	DP01080	CQSGSTCVKQNDWYSQCLPGSQVTTTSTTSTSSSTSRATSTTRTGGVTSIT
18	DP01048	NRESARRSRYRKAHLKELEDQVEQLKAENSCLLRRLAALNQKYNEANVDNRV
19	DP00061	VPYNEGLGQPQVAPPAPAASPAASSRPQPQNGSSGMGSTVSKAYGASKTFGKA
20	DP00224	MHHQQRMAALGTDKELSDLLDFSAMFSPPVSSGKNGPTSLASGHFTGSNVEDR
21	DP00037	RARMESGQYRCLDWWFCWDTPASRDDVEEARRYLRAAEKPAKLLYEEARTHLLH
22	DP00558	MAATEGVGEAAQGGEPGQPAQPPPQPHPPPPQQHKEEMAAEAGEAVASPMDDG
23	DP00835	MSQRIQPSASDQQFPKSDGYEYTVGPKQAITSEASTTYIPSRYSELLFKR
24	DP00028	MSGSSCSQTPSRAIPATRRVVLGDGVQLPPGDYSTTPGGTLFSTTPGGTRIYD
25	DP00148	MQRGNFRNQRMVVKCFNCGKEGHTARNCRAPRKKGCWKCQKEGHQMKDCTERQAN
26	DP00424	TYLGRSAEPVPLQLPLERLTLDNCNEDCGTSGTQGVGSPQLVESPTILESGAKE
27	DP00427	MSERFPNDVPIETRDWLQAIESVIREEGVERAQYLIDQLLAEARKGGVNVAAGT
28	DP00926	GWQAFKNDATEIIPELGEYEPPELENNKTMNRAENGRPPHPFETKDVSEYS
29	DP00092	ITSFEEAKGLDRINERMPPRRDAMPDANLNSINKALTSETNGTDSNGSNSSNIQ
30	DP00194	DEDTSIMNDEEASADNETVMSVIDGDKPDHDDHDPDEPPQPRGGRPALRVVK
31	DP00506	KNRKAKAKPVRGAGAGGRQGRQNKERPPVPNPDIYEPYPIRKQQRDLYSGLNQRRRI
32	DP00600	QGTRAQKTRAEVTGSGKKPWRQKGTGRARSGSIKSPIWRSGGVTFAARPQDHSQKV
33	DP00748	IYARVIQKRVPNAYDKTALALEVGELVKVTINVSGQWEGECNGKRGHFPFTHVRL
34	DP00048	GGKWSKSSVGVWPTVRERMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAW
35	DP00224	MHHQQRMAALGTDKELSDLLDFSAMFSPPVSSGKNGPTSLASGHFTGSNVEDRSS
36	DP00175	MPQLNGGGDDLGADELISFKDEGEQEEKSENSAERDLADVKSLLVNESETNQ
37	DP00769	EEERPRRQFNKGGRTGERDGRNKRGVGRGTRREGEEAAKEEGAETAQGNQETPAQE
38	DP01015	ELKITETEPGVRGDTATGGTKRAKLETGAVRVPLFLNEGEIIVDTRRGEYVSRKA
39	DP00083	SDPAALKRARNTEAARRSRARKLQRMKQLEDKVEELLSKNYHLENEVARLKKLVGER
40	DP00027	VTLSDAQAKLMQPGVSDINMERVEALKTAIRNGELKMDTGKIADSLIREAQSYLQSK
41	DP00948	PTEPKDKKKKTKDEAQPLPQRQKKQPTVTLPAADMDDFSRQLQNSMSGASADSTQA
42	DP00949	DKPTPEEHGPGIDIFTPGKPGEYDLEGGIWEDEDARNFYENLIDLKAFVPAILFKDN
43	DP00287	MPRRAENWDEAEVGAEEAGVEEYGPPEEDGGEEESGAEEESGPEESGPEELGAEEEMEAG
44	DP01073	ASTASPAKAESSGVIDLFGDAFGSGASETQPAPQAVSSSSASADLLAGFGGSFMAPST
45	DP00803	MASLATLAAVQPTTLKGLAGSSIAGTKLHIKPARQSFKLNVRSGAIVAKYGDKSVYF
46	DP00111	GKLGQDQSFESVESYDSCDRLTQSWSSQSSFNSLQRVPSYDSDYEDYPAALPNHKPK
47	DP00750	AQSEQRDRFIPSYEETQRILKEQAEAEAAARNDPDLOAAQEENMRKIREMLGINRGGAR
48	DP00600	AQKTRAEVTGSGKKPWRQKGTGRARSGSIKSPIWRSGGVTFAARPQDHSQKVNKMYRG

Table S1

49	DP00348	PPRSISPSALQDLLRRTLKSPSSPQQQQVNLNLSNPQLMAAFIKQRTAKYVANQPGMQ
50	DP00249	MFDKNADGYIDLEELKIMLQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGV
51	DP00348	PPRSISPSALQDLLRRTLKSPSSPQQQQVNLNLSNPQLMAAFIKQRTAKYVANQPGMQP
52	DP01053	MPKSKELVSSSSSGSDSDSEVDKLLKRKKQVAPEKPVKKQKTGETSRALSSSKQSSSSRD
53	DP00975	YPERDSAQEGNQEQERALHVQKVRTDGDADYDEYEEDGTTPTPDPTAPTAKPRLRGNKP
54	DP00342	MSGGDGRGHNTGAHSTSGNINGGPTGIGVSGGASDGSWSENPNPVGSGSGIHWGGGSG
55	DP01052	PAPPVLPYIDPAKKSFAARMMEKMGWKHGEGLGVNKGITAPLVAKTAMSSGVIVQGA
56	DP00448	SDEIMDLLVQSVTKSSPRALAARERKRSRGNRSLRRTLKSGLDLVLQALGLSKGPGLEV
57	DP00719	QEPVQEAQKGRKRKPRTEPKQVPEKPKVSEKSKGSAKSEKQEKITDFFVKRKRDRF
58	DP00672	RRRRRVKANDRERNRMHNLNAALDALRSVLPSPDDTKLTKIETLRFAYNYIWAETLRL
59	DP00824	APQYAPGDEPSYEDTDDSDKLVENDTSITDEDYAAIEASLSEFNTAADPGRRLGEGSKP
60	DP00123	ALIVLGGVAGLLLFIPLGIFFCVRCRHRRRQAERMSQIKRLLSEKTCQCPHRFQKTCSP
61	DP00977	YPERDSAQDGNQEKERALLVKVQERYQGNQGDYDEYDQDETTPTPTAQTARPRLRQNGD
62	DP00835	FRASVSPSSLPRASAFLSQNESSSKLSNASNSPGLANSSTATSASANERLQEKQTESLSN
63	DP00014	TVTTHSSDSEVPSRVTEVVVKLFDSDPITVVLPEEVSKDNPKFMDTVAEKALQEYRRKSRME
64	DP00278	SPSKDCGSPKYAYFNGCSSPTAPLSPMSPPGYKLVTDGDRNNSCRNYNKQASEQNWANYSAE
65	DP00389	DALAYGSQGNLNPLINEISKIISAAGNFVKEERAAASLLQLSGNASDFSYGRNSITLTASA
66	DP00592	MSERQGAGATNGKDKTSGENDGQKVKQEEFDIDMDAPETERAAVAIQSFQKFKKAGSQS
67	DP00057	MARYRRSRTRSRSPRSRRRRRRSGRRRSRRRRRYSARRSRRSVGGRRRRYGSRRRRRRRY
68	DP00090	MLAARTGAAGSQISEENTKLRRQSGFSVAGKDKSPKKASENAKSSLSPSGESQLRARQLAL
69	DP00812	MIGKNIKSLRKTDLTLQLEFARIVGISRNSLSRYENGTSSVSTELIDIICQKFNVSVDIVGE
70	DP00543	MPGLVDSNPAPPESQEKPLKPCACPETKKARDACIEKGEEHCGHLIEAHKECMRALGFKI
71	DP00152	QEREKDIKQEGETAICSEMTADRPVYDESTDEESENQTDLSGLASAAVSKDDSIISLDVTD
72	DP00220	RYLHTALEGMANPEDPECESEGWLLEKSVPETWKAFLESVKLKGKNQVEAEGEDAGQAPAA
73	DP00718	SPPYYSEKTLQYNKPHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFFRRTIRLKLID
74	DP00709	MPFFGNTFSPKKTPPRKSASLSNLHSLDRSTREVELGLEYSPTMNLGQSLKFENGQWIAET
75	DP01084	LGCPEKMGDPQLGPRSAEPQQPNLSTGFYGVKSEPTQDTKPFQPRQMPSRNASGGQGSST
76	DP00378	PNLQNGQVLTGLPGVMPNIQYQVQFQTVVDGQQLQFAATGAQVQDQSGGQIQIIPGANQQII
77	DP00064	MSGLFHHRKPREIRAFVMATRLTKKQLAQAIQNTLPNPPRRKRRARRAAQVPKPTQAGVSMA
78	DP00877	MAQEQTKRGGGGDDDDIAGSTAAGQERREKLTTEETDLDLDEIDDVLEENAEDFVRAVYQKGGQ
79	DP00462	MLSVRTPLATIADQQQLSPLKRLTLADKENTPPTLSSTRVLASKAARRIFQDSAELESKAPT
80	DP00656	MTAAQAKTTKNTAAAAQEAAGAAQPSGLGLDSIGDLSSLLDAPAAQQGGSGPIELDLIDLID
81	DP01038	MADNKSQSFQAGQAAGRAEEKGNVMDKVKDAATAAGASAQTAGQKITEAAGAVNLVKEKTGMNK
82	DP00026	MAQVINTNSLSLLTQNNLNKSQALGTAIERLSSGLRINSKADDAAGQAIANRFTANIKGLTQAS
83	DP01011	VEVDCSRFPNTTNEEGKDLVCTEDLRPICGTDGVTHSECLLCAYNIEYGTNISKEHDGECREAV

Table S1

84	DP00008	MLWQKSTAPEQAPAPRRPYQGVVRKEPVKELLRRKRGRHTSVGAAGPPTAVVLPHQPLATYSTVGP
85	DP00678	TTQITKTVKGGISETRIEKRIVITGDADIDHDQVLVQAIKEAQHPDMSVTKVWHQETETIADE
86	DP00722	LGGRQGGGAPAGGNIGGGQPQGGWGPQQPQGGNQFSGGAQSRPQQSAPAAPSNEPPMDFDDIP
87	DP01034	MSTQQQARALMMRHHQFIKNRQQSMLSRAAAIEGVEAEKDFWTTVQGKQSSFRTTYDRSNASLS
88	DP00018	KPSACRNLFGPVDHEELTRDLEKHCRDMEEASQRKWNFDQFNHKPLEGKYEWQVEVEKGSLEFYR
89	DP00468	MFQAAERPQEWAMEGPRDGLKKERLLDDRHSGLDSMKDEEYEQMVKELQEIRLEPQEVPRGSEPW
90	DP00948	QASSRSSRSRGNRSRNPSTPGSSRGNSPARMASGGGETALALLLDRLNQLSEKVSQKGGQQGQTV
91	DP01054	MMCGAPSATQPATAETQHIADQVRSQLEEKENKFFVFKAVSFKSQVWAGTNYFIKVHVGDDEFVHL
92	DP00157	PVQATVHGVDDNKQSQTAIEIVNVKPLASQLAGVENVKDTLQSDTTVITGNKAGTTDNDNSQTDKT
93	DP00978	TLNCTDLRNATNTSSSWETMEKGEIKNCSFNITTSIRDKVQKEYALFYNDVVPIDNASYRLISCN
94	DP00788	VLDGNKNQFEPISIRSNFQGKHTKFENDELAESTTTKIDSTDHIRKASSKSKRLDKVGGKGGK
95	DP01105	ENREIQPPFKPKVCGKGAENFDKFFTRGQPVLTPPDQLVIANIDQSDFEFGFSYVNPQFVHPILQSAV
96	DP00081	EGSIGNGLDLTGILFGNIDSEGRLLQDDDDGEGRGGTGDAELRENIGSLSKLGLDSMLLEVIDLKEA
97	DP00747	EVVPPQVLEPNEEAGAALSPLPEVPPWTGEVSPAQRDGGALGRGPWDSRDRSALLKSKLRALLTAPR
98	DP00179	MNTDQQKVSEIFQSSKEKLGQDAKVSADAFKMASQDKDGKTTDAESEKHNYQEYQNKLGAGHKKE
99	DP00328	DPNMIPVTSANAATGEITPAQGATGSSGGNTSGVLSTPKAKRAKHPGTEKPRRSQSEQPATCPIC
100	DP00753	EVDSTEDIKKAFEIFDKENGYISASELKHVLTTLGEKLEQEVDLLKEIGVEEGLINVDVFKLITS
101	DP00834	GTKKLGSGSDLGASIKGFKKAMSDDPEPKDKTSQDADFTAKTIADKQADTNQEQAKTEDAKRHDKQEV
102	DP01002	MKKFLPGLLLMGLVACSSNQRMDSYSQPESSQDLAPKSSTTFQFPQLLSKASSMPLNLLSSKNGQVSP
103	DP00976	VSLKCTDLKNDTNTSSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSC
104	DP00345	GNGKITQDELSKVVDNYELLKHSKNVTNSLDKLISSVSAFTSSNDSRNVLVAPTSMLDQSLSSLQFARAA
105	DP00539	PQAIPKRRGRKPGSVAAAAEAKKKAVKESSIRSQVETVLPKIKRKTRETVSIEVKEVVKPLLVLSTLGE
106	DP00042	MPKRKSATKGDEPARRSARLSARPVPKPAKPKAAAPKAVKGGKAAENGDAKAEAKVQAAGDGAGNAK
107	DP00048	MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEAQEEEEVGFV
108	DP00527	MKSTSLFLCSLLLLVTGAIGRKTKEKYSQSEEVSESFASGSSGSSDDELVRDKPYGPKVGGSGFEE
109	DP00752	MQVQVKTEGETKIYLEQGTSLDLKQISQDMGFEIDMMLVNNGFAPNTELVTDVITYLSLKLGG
110	DP00333	DAKKPEDWDEEMDGEWEPPIQNPYKGEWKPRQIDNPDYKGTWIHPEIDNPEYSPDPSIAYDNFVGLGL
111	DP00017	TSACRSLFGPVDHEELSRELQARLAELNAEDQNRWDYDFQQDMPLRGPGRQLQWTEVSDSVPAFYRETVQV
112	DP00771	MSDYEEAFNDGNENFEDFVEHFSDEETYEKPKQKDGETTANGKTVTGGNGPEDFQQHEQIRKTLKE
113	DP00841	LNLDISPQRPRPTLQLPLANDGGSRSPPSESSQHPPTPARPRHMLGLPSTLFTPRSMESIEIDQKLQEIM
114	DP00719	AAYGGAYGENPCSSCEPGFSSNGLIESVELRGESAFSGIPNGQWMTQSFTDQIPFSFNHCGTQEQEESHSA
115	DP00636	MAGPQRATSGLPTRRITTRQPTRRAGSAIPERQTSTASPAVSTKTAIAISRTRLKSPGEPASVLAKRKERD
116	DP00343	EQVSHGTQNRPLLRNSLDDLVGPPSNLEGGQSDERALLDQLHTLLSNTDATGLEIDRALGIPELVNQGGAL
117	DP00608	GDGGAFPEIHAQYPLDMGRKKKMSNALAIQVDSEKIKYDAIARQGSKDKVIYSKYTDLVPKVMNADD
118	DP00669	GDIPDPGSKPTPPGMADELPTETDYLPPEIYTTFLPRTIYPQEEMPYDDKPPFSLLSKANDLNAVFEQPA

Table S1

119	DP00862	AVQMQSEKSASDSGEVATEKAPAKQDTKEKSGTETEKGKEDGKTGKTDSADKETSASEASEKGTVVTTETAD
120	DP00702	PKSHSRQASTDAGTAGALTPQHVRHSSPASLQLGAVSPGLTPTGVVSGPAATPTAQHLRQSSFEIPDDV
121	DP00435	FEEAMRFARRSVSDNDIRKYEMFAQTLQQSRGFGSFRFPNGQGGAGPSQSGGGTGGSVYTEDNDDDDLYG
122	DP00094	MPTPSAPSPQPKGFRRVSEQDAKQAEAVTSPRFIGRRQSLIEDARKERAAAAAAAAAVASSEPGNPLEA
123	DP00768	KISEGWRDEPTDAALGPEPPGPALEIKPMLQAQESPSLPAPGPPPEEEPIPEPQPTIQASSLPPPQDSARL
124	DP00853	MGMLELVGEYWEQLKITVVPVAAAEDDDNEQHEEKAAEGEEKEEENGDEDEDEDEDEDDDDDEDEEEEEE
125	DP00863	TLREKLMNFLVRHMNEKRELTLPVTLLELSRFGCARPALS RVFQELEREGYIEKHGRRIKVLKNPFEHDR
126	DP00055	HHSGSMDQRQLQHWIHSCLRKADKNKDNKMNFKELKDFLKLNIQVDDGYARKIFRECDHSQTDLSLEDEEEE
127	DP00806	VTADVVSQAALTDEQKARLADTLAAKSGKTVKLNARVDESIGGMIVKLGSMIDSSIRSKLASLQNAMKEVG
128	DP00159	FIDENLKAADSDPTAPPYDSSLVFDYEGSGSEAASLSSLNSESDDQDYDYLNEWGNRFKKLADMYGGGEDD
129	DP00086	MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLSQAMDDLMLSPDDIEQWFTEDP GPDEAPRMPPEAAPV
130	DP00815	YHPADKDYGLMKIDEPNTPYHNMIGDDEDAYS DSEGNEVMTDILAKLAAAEGSEPKYRTREQESSGEEDND
131	DP00086	KKPLDGEYFTLQIRGRERFEMFRELNEALELKAQAQAGKEP GGSRAHSSHLKSKKQGQSTSRRHKKLMFKTEGPDSD
132	DP00707	GVDPALSDPVLARRLRENRLSAMSRLDEVISKYAMLQDKSEEGERKRRARLQGTSSHSADTPEASLDSGEGPS
133	DP00146	MARYFRRRKFCRFTAEGVQEIDYKDIATLKNYITESGKIVPSRITGTRAKYQRQLARAIKRARYLSLLPYTDRHQ
134	DP00490	QSVSPMRSVSENVLVAMDFSGQKTRVIDNPTEALSVAVEEGLAWRKKGCLRLGNHGSPTAPSQSSAVNMALHRSQ
135	DP00539	MVAGMLGLREEKSEDQLQLKDKPLKFKVKVKKDKKEEKGHEPVQPSAHHSAPAEAGKAETSESGSAPAVP
136	DP00919	MGGKWSKSSIVGWP AIRERLRAQPAAEERRRAQPAEAGVGAASQDLGRHGALTSNTDTTNADEEEEVGFVVRPQ
137	DP00201	NKELDPVQKLFVDKIREYTRKQTSGGPVDAGPEYQDLDRELFLKQMYGKADMNTFPNFTFEDPKFEVEK PQS
138	DP00016	RQNPCGSKACRRFLGPDVSEQLSRDCDALMAGCIEARERWNFDVVTETPLEGDFAWERVRGLGLPKLYLPTGPRR
139	DP00932	MKTIEQKIEQCRKWQAARERAIARQREKLADPVWRESQYQKMRDLDLRRIAKQKERPPASKTRKSAVKIKSRGLK
140	DP01086	LEVVFKALDSEIKCEAEKQEEKPAILLSPKEESVLSKPTNAPLPPVVLKPKQEVKSASQIVNEQRGNAAPAARL
141	DP00015	MTDVETTYADFIASGRTGRRNAIHDILVSSASGNSNELAKLAGLDINKTEGEEDAQRSSTEQSGEAQGEAAKSES
142	DP00970	MDIRKIKLIELVEESGISELEISEGEEVSRISRAAPAAAFPMQYAYAAPMMQQAQSNAAAPATVPSMEAPAAAE
143	DP00075	PPKTFEKSMNQLTKIDAKKEQLADARRDLKSAKADAKVMKDAKTKKVVESKKKAVQRLEEQLMKLEVQATDREENK
144	DP00551	MDPQTAPSRALLLFLHLAFLGGRSHPLGSPGSASDLETSGLQEQRNHLQGLSELQVEQTSLEPLQESPRPTGVW
145	DP00990	MSFTIPTNLKPLATKPKHLSSSSFAPRSKIVCQENDQQQPKLELAKVGANAAAALALSSVLLSSWVSVAPDAAMA
146	DP00797	LPIFTPAKMGATKEEREDTPIQLQELLELALGQGCVDREVAEITKQLPPVVPVSKPGALRRSLSRMSQEAQRG
147	DP00328	AKPGVKKKKTTSGKKSSSSGSGSGALSSSGSVPPQVTVQSLHTLQGVQVKKDPDAQQQQQQQQQQQQQAMT
148	DP00469	MQEATSSLAASSLSSSSSSSSSAPHLEIKEGIESDEEIRRVPFEGGEAVGKETS GRESGSATQERTQATVGESQ
149	DP00534	AAPEGGISDVVEKSIKAEQETCAGDPVSGECVAAWDEVEELSAASHARDKKKADGSDPLEEYCKDNPETNECRTYDN
150	DP00223	MALRVRTNRSKINAENKAKINMAGAKRVPTAPAATSKPGLRPRALGDIGNKVSQQLQAKMPMKKEAKPSATGKVIDKK
151	DP00943	DEEDGEPPEYEGSCVEIPGLSEEDPAPSRKIHFSAPIQVFSTYSNEDYDRRNEDVDPMAASA EYELEKRVLEL
152	DP00447	DGIRDAMVGLREEMIEKIRTEALMTNDRLEAMARLRNEESEKMAKDTSEVSLNPTSEKLNLLLEGNSDNDLSLEDF
153	DP00087	TAPITDVS LGDELRLDGEEDVMTPADALDDFDLEMLGDVESPSGMTHDPVSYGALD VDDFFEQMFMTDAMGIDDFGG

Table S1

189	DP01074	MPPIATRGRQYEPKVQQAQLSPDTIPLNPAKTKDPLARADSLHHHVESDSQEDDKAAEPEPLSRKRWQNRTRFRKGRRAQAPYKHK
190	DP00499	LTAATHWASMDPAVVHPELNGAAYSRYPPGVVSVAPTGIPAAVEGIVPSAMLSHGLPPVAHPHPAPSPGQTKPEADRDHTDQL
191	DP00768	GTPQKTKQEEEGTEPEPELEPKPETAPEETELEEVSPPPEPCVGEVAAVTLGPQGTQETALTPPTSLQAQVSVAPAEHSSPRGW
192	DP00032	ADPMVDNRITKKRYRLDHLLEQPEEVLKRLYRAMEVEEENLHALVQAMMQAPERAGGAMTAAGVLDPASGEPGDWTPPRYFNMMF
193	DP00929	MPEVDPRLPEWKHPGSGQPKTACTNICYCKKCCFHCQVCFITKALGISYGRKKRRQRRRAHQNSQTHQASLSKQPTSQPRGDPTGPKE
194	DP00869	AQPKSGFSTDRPLWYPGAKAPEYLDGSLVGDYGFDPFLGLGKPAEYLQYDYDGLDQNLAKNLAGDIIGTRTESADVKSTSLQPYSEVF
195	DP00423	TVTAKKASSPSSLTYKEMILKSMPQLNDGKSSRIVLKYYVKDTFSSKLTSSNFDYLFNSAIKKCVENGELVQPKGSPSGIILNKK
196	DP00347	MNLEPPKAEIRSATRVMGSPVTPRKGPPKFKQRQTRQFKSKPPKGVQGFDDIPGMEGLGTDITVICPWEAFNHLELHQAQYGI
197	DP00157	TGETTSFDEVDGVTSKSIIGKPVQATVHGVDNKKQSQTAIEIVNVKPLASQLAGVENVKDTLQSDTTVITGNKAGTTDNDNSQTDKT
198	DP00423	KASSPSSLTYKEMILKSMPQLNDGKSSRIVLKYYVKDTFSSKLTSSNFDYLFNSAIKKCVENGELVQPKGSPSGIILNKKVKLST
199	DP00260	MPLNVSFNTRNYDLDYDSVQPYFYCDEEENFYQQQQSELQPPAPSEDIWKKFELLTPPLSPSRRSGLCSPSYVAVTFSLRGDNDG
200	DP01085	ETIGNPTIFGETDTEAQTFSGTGNIPPNRVVFNEMPVQHSVNRAPPRGVNIQANNTPSFRPSVQPSYQPPASYRNHGPIMKNEA
201	DP01014	MSTDVAAAQAQSKIDLTKKKNEEINKKLEEDDEFEDFPIDTWANGETIKSNAVQTQTNWEEENWDDVEVDDDFTELKAELEDRYKRENG
202	DP00461	GSNGADNAHNNAFGGKNGPIGNTSAGSNGSASSNRGNSNGWSWSNKPKNDFHSDGSYHITFHGDNNSKPKPGNSGNGRNGNDGA
203	DP00136	AKRKASGPPVELITKAVAASKERSGVSLAALKKALAAAGYDEVKNNSRIKGLKSLVSKGTLVQTKGTGASGSFKLNKKAASGEAKPK
204	DP00195	MPKRKAEGDAEGDKAKVKDEPQRRSARLSAKPAPPKPEPKKAPAKKGEKVPKGGKADAGKDGNNPAENGDQAKTDAQKAEAGGDA
205	DP01088	MAKSKEDITYATSQARLSEDAVRVAYEHGSPLEGGKIADSQPDLFSSAHNMPKSGQTTMDSNTSDQSQMQRDTQEGGSKEFTTGAPG
206	DP00985	EKNGLIMIDDERILEHLDDEEKSRIVKVKVRFRTLGCYPLTGAVESEATSLTDIIQEMLLTRTSEQGRVIDHDGAGSMEEKRQGYF
207	DP00141	ALAPDSPDNDLRAGQFGISARKPFTLGEVAPVWVPSDQAPNCMKCEARFTFKRRRHCRACGKVFCASSCSLCKLLYMDRKEARVCV
208	DP00461	MGSNGADNAHNNAFGGKNGPIGNTSAGSNGSASSNRGNSNGWSWSNKPKNDFHSDGSYHITFHGDNNSKPKPGNSGNGRNGNDGA
209	DP00997	QDEADELRRFVLMADAGEPADSEGAMDAEDHRQGIARNHYDCGYEMHTCLGEMYVSDERFTRNIDAAKPLAAYMRDAILANAVRHPT
210	DP00039	MPKRKAEGDAKGDKAKVKDEPQRRSARLSAKPAPPKPEPKKAPAKKGEKVPKGGKADAGKEGNNPAENGDQAKTDAQKAEAGGDAK
211	DP00191	CLKGTGENYRGNVAVTVSGHTCQHWSAQTPHTRNTPENFPCKNLDENYCRNPDGKRAPWCHTTNSQVRWEYCKIPSCDSSPVSTEQ LAP
212	DP00631	MTPSTPPRSRGTRYLAQPSGNTSSSALMQGQKTPQKPSQNLVPTPSTTKSFKNAPLLAPPNSNMGMTSPFNGLTSPQRSFPFKSSVKRT
213	DP00864	NKAVQNSNPSASQPTTGPSAASPENQNGNLSAPPGGGGPHPPHTPSHPPSTRITRSQPNHTPAGPPGSSNPVSNGETRTRSSKR
214	DP00486	PEKVGPKQAEDAASITNGSDDGAQPSTSTAQEQDDVLIVDSDEEDSSNNADVSEEERSRKRKLEKENLSAKRSRIEQKEELDDVIALD
215	DP00353	MMNNNGNQVSNLSNALRQVNIIGNRNSNTTDDQSNINFEFSTGVNNNNNNSSNNNNVQNNNSGRNGSQNNNDENNINKTLEQHRQQQQA
216	DP00675	DTKEEGATKKTQKPDRLERGMKIVPKESEKSKTKPPDATIVVEGVKYVRKKGKTKSKNTQDGLYHNKPKQESRKKLEKALLAWAI
217	DP01042	YANEPKVCYFYMDDYNGESLCAAQGNVASIPDKWNDRISSISIPHGLVTVYEDVDFLGASRSFEADVLDVSDKLDIYLDNISAFKIKKA
218	DP00902	PTSIHETVYENDGDPTLLDQIADAEASWFDKIALKKAIEELDERERLIVLYRYRDRTQSEVASRLLGISQVQVSRLEKKILQHIKDKMDG
219	DP01033	MSYRKELEKYRDLDEDKILGALTEEELRKLENELELDPDNALLPAGLRQRDQTKQPTGPFKREELMAHLEQQAKDIKDREDLVPTGEKR
220	DP00086	MEEPQSDPSVEPPLSQETFSDLWKLLENVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTAPAPAPAPSWPL
221	DP00808	GLPTTRPIAPMATGVLGQHTPVTLDPADTQKQKVLPGTAVVTRPSRASLRRAASGRSLRPVASGNWQSTLNSIVGLGVQSLKRRRCF
222	DP00113	MSSSTFPDPYALSEHDEERPQNVQSKSRTAELQAEIDDTVGMIRDNINKVAERGERLTSIEDKADNLAVSAQGFKRGANVRVKAMWYKDLKMK
223	DP00944	MADNEEDVEEEEYTELTDISGVGSKAESLREAGSEVEDVRGADQSALADVSGIGNALARIKADVGGLEVESETEAEVEEEGGEEAPDEVD

Table S1

224	DP00016	MSEPAADVQRNPGCSKACRRLLFGPVDSEQLSRDCDALMAGCIQEARERWNFDVFTETPLEGDFAWERVRGLGLPKLYLPTGPRRRRDELGGRR
225	DP00018	ESQDVSQSRPAAPLIGAPANSEDLVDPKTDPSDSQTGLAEQCAGIRKRPATDSSSTQNKRRANTEENVSDGSPNAGSVEQTPKPKGLRRRQT
226	DP00615	SPDADLIEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEEDEREVSVAEILRRSRKFPRAMPIWARPDYNPPLLESWKDPDYPVPPVHGC
227	DP01031	NTYDVMVDANLLMEGGVAQTEPESRVPLVDFLEPMAEESDLEPSIPSECMPLRSGFPRALPAWARPDYNPPLVESWRRRPDYQPPTVAGCALP
228	DP00092	SDDELGSEEDGFDGATAARKEVIRNKIRAIGKMARVFSVLRREESESVLTKGLTPTGMLPSGVLGGKQTLQSATVEAIEADEAIKGFSPQHK
229	DP00984	ELSPSLVISLSTGLSLFLGRVFFNFQRENMAKQVPEQNGMSHFAGDTRAKEYVSLKSNDPVGFNIVDVLAWGSIGHIVAYILATASNGYDP
230	DP00984	PSLVISLSTGLSLFLGRVFFNFQRENMAKQVPEQNGMSHFAGDTRAKEYVSLKSNDPVGFNIVDVLAWGSIGHIVAYILATASNGYDPSFF
231	DP00092	DDELGSEEDGFDGATAARKEVIRNKIRAIGKMARVFSVLRREESESVLTKGLTPTGMLPSGVLGGKQTLQSATVEAIEADEAIKGFSPQHKIT
232	DP00552	YNGQSETSQQPETPVFNTPMGMKASPVSLGVPSEATANNGQQQQVQEQRNRINAMLDQYELQRRHLHSEQLQFEQAQTQAAVQVPGIQLGTQSQ
233	DP00615	ATCTTRHSDPADLIEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEEDEREVSVAEILRRSRKFPRAMPIWARPDYNPPLLESWKDPDYP
234	DP00517	GGGVSPDVCALSDPEGGLTASRGRVHEASPTTFQKDALLGSKPNKPSLPSSSQNLGQTEVSKVSETVQEELTPPPQKAAPQGKSKSDPLKKTDR
235	DP00092	SDDELGSEEDGFDGATAARKEVIRNKIRAIGKMARVFSVLRREESESVLTKGLTPTGMLPSGVLGGKQTLQSATVEAIEADEAIKGFSPQHKIT
236	DP00491	VVEWTNNTLPLEPCPKQIIGLSRRIADENKPRPFAFNALPDFKALSIAVTGSGSCRHLQFIPVAPPSPGSSAAPATEVPDRDPEKSSDDVYLH
237	DP00615	RHSDPADLIEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEEDEREVSVAEILRRSRKFPRAMPIWARPDYNPPLLESWKDPDYPVPPVHGC
238	DP00994	ERIETAEKEAYVPATKVKVNSKQLKTKLEYLFDATFVESNTRKNFGRNNSRNFNRRRGGGARKGNNTANATNSANTVQKNRNDVSNLPSLA
239	DP00082	LEAEIETTTDETDDGNTVSHILNVLKDATPIEDVFSFNYPEGIEGDIKYEKHEVYTYGPTFLLQFKDLNVKADAQEWVQSTASKVIPPGMGRGN
240	DP00097	AKGDEPKRSVAFKTKKEVVKVATPKKAAKPKAASKAPSKPKATPVKKAKKKPAATPKAKKPKVVKVVKVASKPKKAKTVKPKAKSSAKRASKK
241	DP00136	AKKAGAAKAKPKAGAACKPKKATGTATPKKSTKTKPKAKKPAAGAKKAKKSPKAKATKAKKAPKSPAKARAVKPKAAKPKTSKPKAAKPKKTAAK
242	DP00265	KKRKPKGGWNTGGSRYPGGSPGGRNRYPPQGGTWGQPHGGGWGQPHGGSWGQPHGGSWGQPHGGWGGQGGGTHNQWKNKPSKPKTNLKHVAGAAAAGAV
243	DP00968	MMDLELPPPLGQSQDMDLIDLWRQDIDLGSVREVFDSQRQKDYELEKQKLEKERQEQLQKEQEKAFQAQFQLDEETGEFLPIQPAQHIQTDTSG
244	DP00024	MHGDTPTLHEYMLDLQPETDLYCYEQLNDSSEEEDEIDGPAQGAEPDRAHYNIVTFCKCDSTLRLCVQSTHVDIRTLLEDLLMGTGIVCPICSQKP
245	DP00720	GGDSNEKGGKSSKRRKTEPSPSKANTSGSGFKSKEYISDDSTSSDDEKNEPAKKSPPSDGDAKKKAKSESEPEESEEDSNASDEDEDEASD
246	DP00783	LCKKRPKPGGGWNTGGSRYPGGSPGGRNRYPPQGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGTHGQWKNKPSKPKTNMK
247	DP01057	SQRKSLIYAAKGDGNDLNEATKASDFVTDKTKALADGEKADYVVEKNSETADTLGKAEKAAAYVEEKGEAANKAAEAEAGKAGEAKDATK
248	DP01058	VAQRKSLIYAVKSDGNLDDLNEATKASDFVTDKTKALADGEKADYVVEKNSETADTLGKAEKAAAYVEEKGEAANKAAEAEAGKAGEAKDATK
249	DP00953	MAESPTTEAATAGAGAAGPAGSSVAGVGVSGSGGGFPPFLPDVWAAAAAGGAGGPGSGLAPLPLPSPAAAHAALLSHWDPDLSDDWDGERTAPQ
250	DP00161	PSPQHQQPPYASAPRQPVQQPPEAQVPPQHAPHAQVQVQPAQYQPEQPLQQPVSPQVAPAPQVHSAPQPAQQAQFQPAEPVAAPQPEPVAEPAPVM
251	DP00876	MNNQRKARNTPFNMLKRRNRVSTVQQLTKRFLSGLMQRGPKLFLMALVAFLRFLTIPTTAGILKRWGTIKKSKAINVLRGFRKEIGRMLNILNRRRR
252	DP00699	MDKLELVNDGLNIIDFIQKNQKEIQKTYGRSSIQQPSIKDQTKAWEDFLQCTSGESEQVEGGMSKDDGDVERRNLEDLSSTSPDTGTIGKRVSNTRDWAE
253	DP01069	MKRLVLVILLSLTLGFSFAGTASADENANRPVNPGEEDPNEAFRSTPFEATTSALGDCRECIAYRTGATTGKGSRRHDDTVSREIKGSSATPGGSEKAGTGRQ
254	DP00053	MLEEALAAIQNARDLEELKALKARYLGKGLLTQEMKGLSALPLEERRKRQELNAIKAALAALEAREKALEEALKEALERERVDVSLPGASLFGSGLH
255	DP01003	MEPVDNLEPNHPSGQPKTACNNCYCKRCSYHCLVCFQTKGLGISYGRKRRRQRSSAPSSSEHDHQNPIPKQLPQTRGDQTSSESSKKVESKTEPDD
256	DP00902	MDVDVKQGSQSPKQKEMKELIRRSQEGDQEARDEIIEKNMRLVWSVQRFLNRGYEADDFQIGCIGLLKSVDFLSDYVVKFSTYAVPMIIEIQRFLRD
257	DP00458	LHKDVEEAGALQKNHASVTSQAQSAEAAEAETSCLPTAVEPLSMGSCIEPAEQSQCPGESSPEVNDAEATGKKENDAKSSTEPESENVPKVQDNTETPK
258	DP00070	MDVFMKGLSKAKEGVAAAETKQGVAAEAGTKQGVLYGSKTKQGVVHVGVATVAEKTKEQVTVNGGAWTVGTAVAQKTVEGAGSIAAATGFVKKDLGK

Table S1

294	DP00070	MDVFMKGLSKAKEGVVAAAEKTKQGVAAEAGKTEGVLVYSGSKTEGVVHGVATVAEKTKEQVTVNGGAVVTGVTVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGIL
295	DP00518	MAKQPSDVSSCEDREGGQLQPAERPPQLRPGAPTSLQTEPQGNPDGEGDRCPHGSPQGLAPPASPGFPATRSPFLIFVRRSSLLSRSSSYGFSFDTDRSPAPMSCDKSTQTP
296	DP00621	CSESELYTEASGYISSLEYPSPYDPLRCNYSIRVERGLTLHLKFLPEFDIDDHQVQHCPYDQLQIYANGKNIGEFCKGQRPPDLTSSNAVDLLFFTDSEGDSRGWKLRYTTE
297	DP00025	QNSGNQSFEEDEEDKPKYEQGGNVDIDFDSVPHIQHGQKNGNQSFEEDTEKDKPKYEHGGNIIDFDSVPHIHGFNKHTEIEEDTNKDKPSYQFGGHNSVDFEEDTLPKV
298	DP00066	MNRGFFNMLGRRPFPAPTAMWRPRRRRQAAMPARNGLASQIQQLTTAVSALVIGQATRPNQNRPPRRPQKQAPKQPPKPKPKQEKKKKQPAKTKPKQRQRMALKLEAD
299	DP00200	RVKFSRSADAPAYQQGQNLNELNGRREEYDVLDKRRGRDPMEGGKQRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGDGLYQGLSTATKDTYDALHMQUALPPR
300	DP00666	KRSKGGKYSVKDKEDTQVDSEARPMKDETFGEYRSLESDNEEKAFGSSQPSLNGDIKPLQSGDDSLADYGGSDVQFNEGDSFIGQYSKGKEEAAGGNDSSGATSPINPAVALE
301	DP00531	MEEEHKHHHHLFHHKDKAEEGPVYDYEKEIKHHKLEQIGKLTVAAGAYALHEKHEAKDPEHAHKHKEIEEIAAAAAGAGGFATHEHHEKDAKKEEKKLRGDTTISSKLLF
302	DP00173	MEVTGDAGVPESGEIRTLKPCLLRRNYSREQHVAASCLEDLRSKACDILADIKSLTPVTLVAEDGTIVDDDDYFLCLPSNTKFFVALASNEKWAYNNSDGGTAWISQESFDVDET
303	DP00850	HHNKEWLSKPTDKRRLSELTAQWDLKGGHEEHGKSNLKGIEIFTKITNDPCAFQITKSLISYQCATRRSQGKIPSSINRDLTCISGMFTALIEAEFFGEHPHPIRGTKRLEKEEKE
304	DP00625	DLEVIISLGPDPTRLDKLLDSYSTAATSSSKDVISVATVSNITTKTSDERLITEGEANQGVATTVKATESDGNVPKAMTGSIDLKEGIFDSVGTITIDPEVLKEKPPWRQPGA
305	DP00387	MKKRNLRLKKNEDFQKVFKHGTSVANRQFVLYTLDPENDELRLVGLSVSKKIGNAVMRNRIRLIRQAFLKEEKERLKEKDYIIARKPASQLTYEETKSLQLHLFRKSSLYKSSSK
306	DP00205	MKTTLIKVIAASVTLFLSMQVYASGHTAHVDEAVKHAAEAVAHGKEGHTDQLEHAKESLTHAKAASEAGGNTHVGHGKIHLEDAIKHGEEGHVGVATKHAQEAIEHLRASEHKSH
307	DP00134	APTEEERESFLRRGDRRRGGGGRGQGGRRGGGFKGNDHDSRTDNRPNRPREAKGRTRDGSGLQIRVDCNNERSVHTKTLQNTSSSEGSRLRTGKDRNQKKEKPDSDVGGQQPLVNGVP
308	DP00341	EDKPQDYKRLSVELTSSLFRTPEMAWNETADLGLDIGAQGEALGYRQDDPSYRSFHSGGYQDQDALGMDPMMEHEMGGHHPGADYPVDGLPDLGHAQDLMDGLPPGDSNQLAWFDTL
309	DP01060	MGSLEMVPMGAGPPSPGGDDPDGDDGGNNSQYPSASGSSGNTPPPNDEERESNEEPPPPYEDPYWNGNDRHSDYQPLGTQDQSLYGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRG
310	DP00553	YRKDKRRQEPLRQSPQRGAGAPELGAAPPEELAAQLGPTTHECEAGPPHDTRLTALPDYTLTLRRSPDDIPLMTPNTITMIPNSLVGLQTLHPYNTFAAGFNSTGLPHSHSTRV
311	DP00447	MEKFAPEFHGEDANNRATKFLSESIGKFTSPKDPKKSISVNSTDIEVTKLESPISTNIPTNETDDNAGNKPNYQRKPLVSFKEDPISDNPFSLYKTIETFDNNEESSY
312	DP01068	DLNTLPIRQTASIFKQPVTKVTNHPSNKVKSDPQRMNEQPRQLFVLEKLSASDVSTEQIKTMDLPGKLVGQVPGSNDETLTSAVASALHTSSAPITGQVSAAVEKNPAVWLNTSQ
313	DP00537	MDRSLGWQGNVSPEDRTEAGIKRFLDITDDGELSFKVDFSGNASCHPPEAKTASRPQVPEPRPQAPDLYDDDLFRPPSRPQSSDNQQYFCAPAPLSPSARPRSPWGKLDPYDSSE
314	DP00395	GFVSDSNVVKPQTFADVLAISQRTHTHTVAVTDDGTPHGVLLGLVTRDYPIDLTQTETKVSMDMTPFSKLVTAHQDTKLEANKIWEKKNLALPIIDDDQHLRYIVFRKYDRSQVC
315	DP00049	HVKKPYDPSRTEKRESKCPTEGCDGTGHVNGLYPHHRSLSGCPHKDRVPEILAMHENVLKCPTPGCTGRGHVNSNRNSHRSLSGCPAAAAEKAKAQEKHQSDVSKSNQASDRVLRP
316	DP00544	MAKSIRSKHRRQMRNVKREHFAKDLDRKRLASKAQELDLNVTMKSAAEIKNKPSTASADADKGMVDNTKVKFKKTQQNEDGHYPQWMNQRAVKKQKVKVAKLTKKKIGKIKIW
317	DP00871	MDPNTVSSFQDILLRMSKMLESSSEDLNGMITQFESLKYRDSLGEAVMRMGDLHSLQNRNEKWREQLGQKFEIIRWLIIEVVRHKLKITENSFEQITFMQALHLLLEVEQEIRTFSFQLI
318	DP00839	PMPGALGASGSSGHELALGGEGGLQSLGNSMHSQMLQIGPAGLGGGLGALTGPGLASLLGSSGPPGSSSSSSRSQSAAVTPSSSTSTRATPAPSAPAAASATSPSPAPSSGNGA
319	DP01103	AEAESLTAQKAKAEKRRKMSKGLPDKSEEEKSTMAKKLEQKPKGEGIPTAKLKIDEFESNVNEVKDPYPSADFPGDDEEPEIPLSPRPRPLAELQLKEKAVPIPEASSFFIFSPT
320	DP00020	MSSPDDFETAPAAYVDALDPSMVVVDSGSAAVTAPSDSAAEVKANQNEENTGATAAETSEKVDQTEVEKKDDDDTTEVGVTTTTTSPSIADTAATANIASTSGASVTEPTDDTAADEKKEQVSG
321	DP00958	MKLPVAQYAPDGVKESFAPIRDPRYMTTEGRITGPDHVLNAGQIDRDKPSEPRTKDGSQLTYLQGLRQLTGLQDDINEFLTRMELAKNKKKAGADEKRIQEEINQLDGGDGEDAV
322	DP00629	SGGDGAYHEPTGGGAIEVALDNADIDLETEAHADQDARGWGGESGERWARQVSGGHFVTLHGAERLEEETNDEDVSDIERRIAMRLAERRQEDSATHGDEGRNNGVDHEDDDAAAVAGIGGI
323	DP00787	ATINNAGFESGFSNWNEDTAPAAISSDAYSGSKSLKIQGSPARVYQVVDIQNPTEYTLASVYLGKGQIGVNDLNGLFKNQTFNVSSWTKVTKTFTSANTNSLQVFAKHNNYNTSDVRFDFNSLIEG
324	DP00012	IRKFSIVQKTLQMNIGEEEDSDEPLERRLSLVPDSEGEAILPRISVISTGPTLQARRRQSVLNLMTSHVNVGQNIHRKTTASTRKSVLAPQANLTELDIYSRRLSQETGLEISEEINEEDLKE
325	DP00662	LTNNVLSAQVQVYRLKGGAKPNMDENASKIISAGRAKRSIAQPDADAGERFRQLKEQEKRSKKNKAVAKDVELVEESQSESEEGSDDEEEEAAREGALASSTSKPLPEVGGRRSKRSKRKRTV
326	DP00160	HTTEDKISRAGVPRQAQVSLFQGDQSENELPRLGGKEDRRVKQSRGEARESYRETGPSRASDARAAHLPTGTLPIDITASESSQDPQDSRRSAEPLLSCKPQWQESRKNKAQTRTPLQCTMTEIF
327	DP00629	LSGGDGAHYHEPTGGGAIEVALDNADIDLETEAHADQDARGWGGESGERWARQVSGGHFVTLHGAERLEEETNDEDVSDIERRIAMRLAERRQEDSATHGDEGRNNGVDHEDDDAAAVAGIGGI
328	DP00774	QTTDNVDELFDTLMLFAKRLSTLDLNEINAADRMAQTPNRLKHGHVDDDALTFAIPTHDRGMVLDNGDNGEDIPYSYQREGFADGQLVGEDDPWRFVWMTSPDGKYRIVVGGWEYREDMALA

Table S1

329	DP00085	PTQAGEGTLSEALLQLQFDDDELGALLGNSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYEAITRLVTAQRRPPDPAPAPLGLPGLNGLSGDEDFSSIIDMDFSAALLSQISS
330	DP00418	RKRRKPGSHTHSASEDNNTNVRQELNLIQKINPIEKHGANTVPIKDYENKNSKMSKIRTHNSEVEEEDMDKHQQKARFAKQPAYTLVDREKPPNGTPTKHPNWTKQDNRDLESAQLNRMEYIV
331	DP01100	QTAPVPMPDLKNVSKIGSTENLKHQPGGGKVQIINKKLDLSNVQSKCGSKDNKIHVPGGGSVQIVYKPVLDLKVTSKCGSLGNIHHKPGGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGGNK
332	DP00219	MAEAGAGLSETVTTETTVTTPEPENRSLTIKLRKRKPEKKEVETSDTDVNEHMGRSSKCCCIYEKPRAFGESSTESDEEEEGCGHTHCVRGHRKRRRRATLGPPTPTPPQPPDPSQPPPGPMQH
333	DP01103	EAEESLTAQKAKAEKRRKMSKGLPDKSEEEKSTMAKLEQKPKGEGIPTTAKLKIDEFESNVNEKDPYPSADFPDDEDEFEIPLSPRPRPLAELQLKEKAVPIPEASSFFIFSPTNKIRVL
334	DP00630	MDVFKKGFSAKEGAVGAVETKQGVTEAAEKTKEGVMYVGAKTENVVQSVTSAEKTKEQANAVSEAVVSSVNTVATKTVEEAENIAVTSGVVRKEDLRPSAPQGEGEASKEEVEAAEQSGGD
335	DP00546	MRRRGEIDMATEGDVELELETETSGPERPPEKPRKHDSGAADLERVTDYAEKEIQSSNLETAMSVIGDRRSREKQAKQEREKELAKVTIKKEDLELIMTEISRAAAERSLREHMGNVVEALIALTN
336	DP01100	QTAPVPMPDLKNVSKIGSTENLKHQPGGGKVQIINKKLDLSNVQSKCGSKDNKIHVPGGGSVQIVYKPVLDLKVTSKCGSLGNIHHKPGGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGGNKIE
337	DP00365	PPADGALKRAEELKTQANDYFKAKDYENAKFYSAQIELNPSNAIYYGNRSLAYLRTECYGALGDATRAIELDKKYIKGYRRAASNMAKGFRAALRDYETVVKVPHDKDAKMKYQECNKIVKQKA
338	DP00025	GQNSGNQSFEEDEEDKPKYEQGGNVIDIDFDSVPIHQGNKGNQSFEEDEKDKPKYEHGGNIIDIDFDSVPHIHGFNKHTEIIEEDTNKDKPSYQFGGHNSVDFEEDTLPKVSGQNEGQQTIEEDTTP
339	DP00062	MDTKHFLPLDFSTQVNSSLTSTPTGRGSMAPSLHPSLPGGIGSPQLHSPITLSSPINGMGPPFSVSISSPMGPHSMSVPTTTLGFSTGSPQLSSPMNPVSSSEDIKPPGLNGVLKVPAPHPSGNMASF
340	DP00224	MHHQQRMAALGTDKELSDLLDFSAMFSPVSSGKNPTSLASGHFTGSNVEDRSSSGSWGNGGHPSPSRNYGDGTPYDHMTSRDLGSHDNLSPPFVNSRIQSKTERGSYSSYGRESNLQGCQSLGGD
341	DP00534	MATIATGLNIATQRVFTVSENRPVCLAGPVHLNNSWNLGSRTTNRMMKLQPIKAPEGGISDVVEKSIKAEQETCAGDPVSGECAVAWDEVEELSAASHARDKKAADGSDPLEEYCKDNPETNECRTYDN
342	DP01104	RKRRRQHQQLWFPEFGKVFSEASKKRREPLGEDSVGLKPLKNASDGLMDDNQNQEWGDEDELETKKFRFEFVPLDLDLDDQDHRQWTTQQLDAADLRMSAMAPTPPQGEVDADCMDVNVNRGPDGFTPLMIA
343	DP00341	MATQADLMELDMAMEPDRKAAVSHWQQQSYLDSGIHSGATTAPSLSGKNPEEEDVDSQVLWEQGFQSFSTQEQVADIDGQYAMTRAQRVRAAMFPETLDEGMQIPSTQFDDAHPNTVQRLAEPQMLK
344	DP00698	SVGRQDNMQAREAKFAAGGVLVGGGEQDIDEEEPHSGRQSVTFKREMSMSLADSVSSVSTSGGTRLNLSLLNLSRLAAKAIKESTAQSSSERNNPPNRPQADSGRQDDQEPKPAQNLDLDFVRADV
345	DP00555	MDVFMKGLSMAKEGVVAAAEKTKQGVTEAAEKTKEGVLVYGSKTREGVVQGVASVAEKTKEQASHLGGAVFSGAGNIAAATGLVKREEFPTDLKPEEVAQEAEEPLIEPLMEPEGESYEDPPQEEYQYEP
346	DP00122	MADERKDEGKAPHWTSASLTAEAHHPHSPENKDDQGGSGEGLSRANSAGFPYREEEEGAFGEHGSQGTYSDTKENGINGELTADRETAEEVSARIVQVTAEAVALKGEQEKQHKDQPAALPLAAEETVNLPP
347	DP00120	RLEQYTSAVVGNKAAKPAKPAASDLVPAEAGVRNIKSMWEKGNVFSPPGTTGTPNKETAGLKVGVSSRINEWLTKTPEGNKSPAPKPSDLRPGDVSGKRNLEKQSVKPAASSKVTATGKKSETNGLRQFEKEP
348	DP00965	KKDKDGEGAPPKARARTDQMEVDSGPRKRPSRGGFTDKERQDHRRRKALENKRKQLSAGGKNLSKEEEELRRLTEEDERRERRIAGPQVGGVNPLEGGTRGAPGGGFVPSMQGVPESPFTRTGEGLDIRGSQGF
349	DP00340	MPNFAGTWKMRSSNFDELLKALGVNMLRKAVAASAAPHVEIRQDGDQFYKTTSTVTRTEINFVGGEGFEEETVDGRKCRSLPTWENENKIHCTQTLEGGDPKTYWTRELANDELITFGADDVCTRIVRE
350	DP00794	SSVSSSPFPFGHSAAVSPTFMPRSTQPLTFTTATFAATKFGSTMKNSGRSNKVRTSPINLGLNVNDLLKQKAISSMSHLYGLGLGSQQPQQQQPAQPPPPPPPPQQQQQKTSALSPNAKEFIFPNMQGQSS
351	DP00626	MLFIKPADLREIVTFPLFSDLVQCFFSPAADYVEQRIDLNQLLIQHPSATYFVKASGDSMIDGGISDGLLIVDSAITASHGDVIAAVDGEFTVKKLQRPTVQLIPMNSAYSPTISSSETLDVFGVVIHVKAMR
352	DP00070	MDVFMKGLSKAKEGVVAAAEKTKQGVAAEAGKTKEGVLVYGSKTKEGVVHGVATVAEKTKEQVTVNGGAVVGTAVTAQKTVEGAGSIAATGFVKKQDLGKNEEGAPQEGILEMPVDPDNEAYEMPSEEGYQDYEP
353	DP00196	AVPVESKPKDPSGKSGMDAALDDLITLGGPEETEENTTYTGPVSDPMSSTYIEELGKREVTIPPKYRELLAKKEGITGPPADSSKPIGPDDAIDALSSDFTCGSPTAAGKTEKEESTEVLKAQSAGTVRSAAPPQEK
354	DP00999	CIFEVKHEGKVTGYACLVGDKVMKPAHVKGVIDNADLAKLAFKKSKEYLECAQIPVHMRSASKYTHEKPEGHYNWHHGAQVYSGGRFTIPTGAGKPGDSGRPIFDNKGRRVVAIVLGGANEGSRITALSVVTWNKDMVTRVT
355	DP00134	FAEDVIQVPRNLVGVIGKNGKLIQEIIVDKSGVVRVRIEAENEKNVQEEIIMPNSLPSNNSRVGNAPAEKHLIDIKENSTHFSQPNSTKVQVRLVASSVAVGESQKPELKAWQGMVPFVFGTKDSIANATVLLDYHLN
356	DP00186	MSSQQNQNRQGEQEQGYMAAEKVVNAWESTKETSSTAQAAAETAEFRDSAGETIRDLTGQAQEQKQEFKERAGEKAEETKQRAQEKMDETKQRAQEMRENAGQKMEEYKQKQKGAEEELRDTAAEKLHQAGEKVKGRD
357	DP00260	MPLNVSFTRNRYDLDYDVSQPYFYCDEEENFYQQQQSELQPPAPSEDIWKKFELLTPPLSPSRRSGLCSPSYAVTVPFSLRGDNDGGGGFSTADQLEMVTELLGGDMVNSQFICDPDETFIKNIHQDCMWSGFSAAK
358	DP00378	TTSGSSGTSNQQTQPRVSGLQGSDALNIQQNTSGGSLQAGQKQKEGEQNTQQQQLIQPQLVQGGQALQALQAAPLSGQFTTTQAISETLQNLQLQAVPNSGPIIIRTPVGPNGQVSWQTLQNLQVQNPQAQITLAPMQ
359	DP00378	TNSQGGTPQRVSGLQGSDALNIQQNTSGGSLQAGQKQKEGEQNTQQQQLIQPQLVQGGQALQALQAAPLSGQFTTTQAISETLQNLQLQAVPNSGPIIIRTPVGPNGQVSWQTLQNLQVQNPQAQITLAPMQVSLGQT
360	DP00092	SDDELGSEEDGFDGATAARKEVIRNKIRAIGKMARVSVLREESESVLTKLGLTPTGMLPSGVLGGKQTLQSATVEIAEADAEIKGFSQPKHITSFEEAKGLDRINERMPPRRDAMPDANLNSINKALTSETNGTDSNGSNSSNIQ
361	DP00174	MASSDIQVKELEKRASGQAFELILSPRSKESVPEFPLSPPKKDLSEIEIQKLEAAEERRKSHAEVLKQLAEKREHEKEVLQKAIENNNFSKMAEELTHKMEANKENREAQMAAKLERLREKDKHIEVVRKNKESKDADETEAD