

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	LTQERKREIIIEQFKVHENDTGSPEVQIAILTEQINNLEHLRVHKKDHHSRRGLLMVGKRRRLAYLRNKDVARYEIVEKLG
2	1BKR	KSADALLWCQMKTAGYPNVNIHFTTSWRDGMFALIHKRPDLIDFKLKKSAHYNLQNAFLAEQHLGLTKLLDPEDISVDHPDEKSIITYVVTYYHYFSKM
3	1BM8	QIYSARYSGVDVYEFIHSTGSIMKRKKDDWVNATHILKAANFAKAKRTRILEKEVLKETHEKVQGGFGKYQGTWVPLNIAKQLAEKFSVYDQLKPLDF
4	1BZ4	SGQRWELALGRFWDYLRWVQTLSEQVQELLSSQVTQELRALMDETMKELKAYKSELEEQLTPVAETRARLSKELQAAQARLGADMEDVCGRLVQYRGEVQAMLGQSTEELRVRLASHLRKLRKLLRDADDLQKRLAVYQAG
5	1CDZ	ELPDFFGKHFLLYGEFFPGDERRKLIRYVTAFIGELEDYMSDRVQFVITAQEWDPSEFEEALMDNPSLAFVRPRWIYSCNEKQKLLPHQLYGVVPQA
6	1CQY	TPVMQTVVKNVPTTIGDTVYITGNRAELGSWDTKQYPIQLYDSSNDWRGNVLPAAERNIEFKAFIKSKDGTVKSWQTIQQSWNPVPLKTTSTSSW
7	1DSL	TFRMRIYRDDFRGQMSEITDDCPSLQDRFHLTEVHSLNLEGSWVLYEMPSYRGRQYLLRPGEYRRLDVGAMNAKVGSRLRRVMDFY
8	1DT4	MKDVEIAVPENLVGAILGKGGKTLVEYQELTGCRIQSKGGEFLPGTRNRKVTTITGTPAATQAAQYLITQRI
9	1DXS	SLVSFLTGLGCPNCIEYFTSQGLQSIYHLQNLITIEDLGALKEPQYRMTIWRGLQDL
10	1EAN	GELVRTDSPNFLSSVLPTHWRSNKTLPIAFKVVALGDVPDGLTVTMAGNDENYSAELRNATAAMKNQVARFNDLRFVGRSGRGKSFTLTITVTNPPQVATYHRAIKITVDGP
11	1EOE	ERVVINVSGLRFETQLKTLNQFPDITLLGNPQKRNRYDPLRNEYFDRNRPSFDAILYFYQSGGRLRRPRNVPLDVFSEEKIFELGENAFERYREDEGF
12	1EW4	MNDSEFHRLADQLWLTIEERLDDWDGSDIDCEINGVLTITFENGSKIIINRQEPLHQVWLATKQGGYHFDLKGDEWICDRSGETFWDLLEQAATQQAQAGETVSFR
13	1EYH	HNYSEAEIKVREATSNPWPSSSLMSEIADLTYNVAFSEIMSMIWKRLNDHGKNWRHVYKAMTLMELYIKTGSESVSQCKENMYAVQTLKDFQYVDRDGDQGVNVREKAKQLVALLRDEDRLEEREAHALKTKEKLAQTA
14	1F0M	YTSFNTVDEWLEAIKMGQYKESFANAGFTSFVVSQMMMEDILRVGVTLAGHQKILNSIQVMRAQMNQIQ
15	1GV2	ELIKGPWTKEDQRIKLVQKYGPKRWVIAKHLKGRIGKQCRERWHNHLNPEVKTSWTEEDRIIYQAHKRLGNRWAEIAKLLPGRTDNAIKHNWNSTMR
16	1GYU	MIPSITAYSKNGLKIEFTFERSNTNPSVTVITQASNSTELDMTDFVQAAVPKTFQLQLLSPSSSVPAFNTGTITQVIKVLNPQKQLRMRIKLTYNHKGSAEQDLAEVNNFPPQSWQ
17	1H3Q	MSGSFYFVIVGHNDPVFEMEFLPPGKAESKDDHRLNQFIAHAALDLVDENMWLSNNMYLKTVDKFNEWVFSAFVAGHMRFIMLHDVVRQEDGIKNFFTDVYDLYIKFAMNPFYEPNSPIRSSAFDRKVQFLGKKHLLS
18	1HUF	LSLSDLHRQVSRLVQQESGDCTGKLGRGNVAANKETTFQGLTIASGARESEKVFQTVLSHVANVLTQEDTAKLLQSTVKHNLNNYDLRSVGNNGNSVLSLRSDQMTLQDAKVLEAALRQES
19	1I2T	HRQALGERLYPRVQAMQPAFASKITGMLELSPAQLLLLLASEDSLRLARVDEAMELIIAHG
20	1J27	MKAYLGLYARLETPARSLKEKRALIKPALERLKARFPVSAARLYGLDAWGYEVVGFTLLGNPDAWVEETMRAAARFLAEAGGFQVALEEFRLAEFEL
21	1J74	VKVPRNFRLLLEELEGQKGVGDGTYSWGLEDDMTLRTWTGMIIGPPRTNYENRIYSLKVECGPKYPEAPPSRVFVTKINMNGINNSSGMVDARSIPVLAKWQNSYSIKVVLQELRRLMMSKENMKLPQPEGQTYNN
22	1JB3	ELQRRREEANVLTGTVEEIMNVDPVHHTYSCKVRVWRYLKGKDIVTHEILLDGGKNVIGFGDPLICDNQVSTGDTRIFFVNPAPQYMWPAHRNELMLNSSLMRITLRNLEEVEHCVEEHRKLLA
23	1JMW	MNQQGFVISNELRQQQSELTSTWDLMLQTRINLARSARMMDASNQSSAKTDLQNAKTTLAQAAAHYANFKNMTPLPAMAEASANVDEKYQRYQAALAEIQFLDNGNMDAYFAQPTQGMQNALGEALGNRYARSENLYRQTF
24	1JOS	RSRVAQEIQEIAVILQREVDPRIGMVTVSDVEVSSDLSYAKIFVTFLLFDHDEMAIEQGMGLEKASPYIRSLGKAMRLRIVPEIRFIYDQSLVEGM
25	1K40	NDKVYENVTLGVKAVIEMSSKIQAPPEEYVPMVEVGLALRTLLATVDETIPALPASTHREIEMAEQKLLNSDLGELISKMKLAQQYVMTSLQQEYKQMLTAAHALAVDAKNLLDVIDQARLKML
26	1K6K	MLNQELESLNMAFARAREHRHEFMTVEHLLALLSNPSAREALEACSVDLVALRQELEAFIEQTPVLPASEEERDTQPTLSFQRVLQRAVFHVQSSGRNEVTGANVLVAIFSEQESQAAYLLRKHEVSRLDVVNFISHGT
27	1KU3	ELEKALSKLSEREAMVLKMRGLIDGREHTLEEVGAYFVTRERIRQIENKALRKLKYHES
28	1L2P	TDQLKKAAEAQVIIEQANKRRSQILDEAKAEAEQERTKIVAQAQAEIEAERKAREELRK
29	1LMI	SAYPITGKLGSELTMTDTVGQVVLGWKVSCLKSSTAVIPGYPVAGQVWEATATVNAIRGSVTPAVSQFNARTADGINYRVLWQAAGPDTISGATIPQGEQSTGKIYFDTGPSPTIVAMNNGMEDLLIWEF
30	1LN4	MDLSTKQKQHLKGLAHLPLKPVVLLGSNGLTEGVLAIEQALEHHELIKVKIATEDRETITLVEAIVRETGACNVQVIGKTLVLYRPTKERKISLPLE

Table S1

31	1LOU	MRRYEYVNLNPNLDOSQLALEKEIQRAAENYGARVEKVEELGLRRLAYPIAKDPQGYFLWYQVEMPEDRVNDLARELRIRDNVRRVMVKSQEPP
32	1LU4	ADERLQFTATTLSGAPFDGASLGQKPAVLWFWTPWCPFCNAEAPSLSQVAAANPAVTFVGIATRADVGAMQSFVSKYNLNFNLNDADGVIWARYNPVWPQAFVYFRADGTSTFVNNPTAAMSQDELSGRVAAL
33	1MB1	NQIYSARYSGVDVYEFHSTGSIIMKRKKDDWVNATHILKAANFAKAKRTRILEKEVLKETHEKVQGGFGKYQGTWVPLNIAKQLAEKFSVYDQLKPLF
34	1MHN	LQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVYTYGYNREEQNLSDLLSPICE
35	1MJC	SGKMTGIVKWFNADKGFITPDDGSKDVVFHFSAIQNDDGYKSLDEGQKVSTIESGAKGPAAGNVTSL
36	1NG6	MSLLERLNQDMKLYMKNREKDKLTVVRMVKASLQNEAIKKKSLTEDEELTVLSRELKQRKDSLQFESNANRLDLVDKVKELDILEVYLPEQLSEELRTIVNETIAEVGASSKADMGKVMGAIMPKVKGKADGSLINKLVSSQLS
37	1O3X	AANKLIKEMVQEDQKRMEKISKRVNAIEEVNNVVKLLTEMVMSSHSGGAAAGSSSEDLMKELYQRCERMPTLFRASDTEENDDEALAEILQANDNLTVINLYKQLVRGEEV
38	1O8X	GLDKYLPGIEKLRRGDGEVEVKSAGKLVFFYFSASWCPPARGFTPLQIEFYDKFHESKNFEVFCWDEEEDGFAGYFAKMPWLAVPFAQSEAVQKLSKHFNVESIPTLIGVDADSGDVVTRARATLVKDPEGEQFPWKDA
39	1OZ9	KNRVLVKLKKRVRKDKIEKWAELALSALGLNNVELSVYITDDQEIRELNKTYRKKDKPTDVLSPFMGEEFGGYKILGDVVISQDATERQARELGHSLEEEVKRLIVHGIVHLLGYDHEKGGEEEKFRELENYVLSKLSK
40	1P4P	SKLTRSNGTTLEYSQITDADNATKAVETLKNSIKLEGLSVVGKTTVEIKEGTVTLKREIKDGVKVLNDTAGSNKKTGKWEDSTSTLTISADSKKTKDLVFLTDGTITVQQYNTAGTSLEGSASEIKNLSELKNALK
41	1PGX	ELTPAVTTYKLVLINGKTLKGETTAKAVDAETAFAKQYANDNGVDGVWYD DATKTFTVTMTEVTPVA
42	1Q2Y	MKAVIAKNEEQKDAFYVREEVFVKEQNVPAEEIDELENESEHIVYDGEKPVGAGRWRMKDGYGKLERICVLKSHRSAGVGGIIMKALEKAAADGGASGFILNAQTQAVPFYKKGHYRVLSEKEFLDAGIPHLQMMKD
43	1QAU	NVISVRLFKRKVGGLGFLVKERVSKPPVISDLIRGMAEQSGLIQAGDILAVNDRPLVDLSYDSALEVLRGIASETHVVLIRGPEGFTTHLETTFTGDGTPKTIKIRVTPQ
44	1QZM	SGYTEDEKLNIAKRHLHPKQIERNALKKGLTVDDSAIGIIRYYTREAGVRGLEREISKLRKAVKQLLLDKSLKHIEINGDNLHDYLGVQRF
45	1R69	SISSRVKSKRIQLGLNQAELAQKVGTTQQSIEQLENGKTRPRFLPELASALGVSVDWLLNGT
46	1R6J	GAMDPRTITMHKDSSTGHVGFIFKNGKITSIVKSSAARNGLLTHEKNICEINGQNVIGLKDSQIADILSTSGTVVITIMPFA
47	1R9H	KIDITPKDGGVLKLIKKEGGVVKPTTGTTVKVHYVGTLENGTKFDSSRDGRDQFSLNLRGNVIGKWDLVATMTKGEVAEFTIRSDYGYDAGSPPKIPGGATLIFEVLEFEWSA
48	1SRV	GYQFDKGYISPYFVTNPETMEAVLEDAFILVEKKVSNVRELLPILEQVAQTGKPLLIADVEGEALATLVNKLRTLSVAAVKAPGFGDRRKEMLKDIAAVTGGTVISEELGFKLENATLSMLGRAERVRIKTDETTIVGGK
49	1T3Y	ATKIDKEACRAAYNLVRDDGSAVWVTFKYDGSTVPGEQGAEYQHFIQQCTDDVRLFAFVRFRTGDAMSKRSKFALITWIGENVSGLQRAKTGTDKTLVKEVVQNFAKEFVISRKELEEDFIKSELKKA
50	1TG0	EPEVPFKVVAQFPYKSDYEDDLNFEKDQEIIVTSVEDAEWYFGEYQDSNGDVIEGIFPKSFVAVQG
51	1TIG	INVKEVRLSPTEEHDFTNKLNRARKFLEKGDVKVATIRFKGRAITHKEIGQVRLDRLSEACADIAVETAPKMDGRNMFLVLAPKND
52	1TP6	CAYRREIHHAVAIRDWLAGDSRADALDALMARFAEDFSMVTPHGVVLDKTAGELFRSKGGTRPLGRIEIDGESLLASGVDGATLAYREIQSDAAGRSERLSTVVLHRDDEGRLYWRHLQETFCG
53	1TQG	GSHMEYLGVFVDEKEYLQNLNDTLELEKNPEDMELINEAFRALHTLKGMAGTMGFSSMAKLCHTLENILDKARNSEIKITSDLLDKIFAGVDMITRMVDKIVS
54	1TSF	QGVELIARDWIGLMVEVVEPNHSEVGKGEVVDETQNTLTKIMTEKGLKVAKRGRTFRVWYKGKIMRIKGLDINFRPEDR
55	1TUC	MGPREVMTKKGDILTLLNSTNKDWWKVEVNDRQGFVPAAYVKKLDSGTGKELVLALYDYE
56	1TZV	MKTPRRRMRLAVFKALFQHEFRREDLEQILEEILDETYDKKAKEDARRYIRGIKENLSMIDDLISRYLEKWSLNRSLVVDNRNRLATYELLFEKDIPIEVTIDEAIEIAKRYGTENSGKRVNGILDRIAKEHAPKEKFE
57	1UCS	NKASVVANQLIPINTALTLMIMKAEVVTMPGIPAEIIPKLVGMQVNRVPLGTTLMPPDMVKNYE
58	1VCC	MRALFYKDGKFLTDNNFLNPVSDDNPAYEVLQHVKIPTHLTDVVVEYEQTWEEALTRLFVSGDSKGRRRQYFYGKMHV
59	1WKA	GKLYTLRYEVEGGFIEIATVRPETVFADQAIAPHEDERYRHLGKRARIPLTEVWIPLADPAVEKDFGTGALKVTPAHDPLDYEIGERHGLKPVSVINLEGRMEGERVPEALRGLDRFEARRKAVELFREAGHLVKEEDY
60	1WPA	WIREYPPITSDQQRQLYKRNFDGLQEYKLSQSELEINKELSRDLKELDDYREESEYMAAADEYNRLKQVKSADYKSKKNHCKQLKSKLSHIKKMVG DYDRQKT
61	1WVN	PLGSQTTHELITIPNNLIGCIIIGRQGANINEIRQMSGAQIKIANPVEGSSGRQVITITGSAASISLAQYLINARLS
62	1X6J	MDINNKARIHWACRRGMRELDISIMPFHEHYDSLDDDEKIRIFIRLECCDDPLFNWLMNHGKPADAEMMVRLIQTRNRERGPVAI
63	1XAW	WIREYPPITSDQQRQLYKRNFDGLQEYKLSQSVLEINKELSRDLKELDDYREESEYMAAADEYNRLKQVKSADYKSKKNHCKQLKSKLSHIKKMVG DYDRQKT
64	1Y2F	RKEAVIIMNVAHHGSELNGELLNSIQAGGIFGDMNIYHRHLSPDGSGPALFSLANMVKPGTDFPEMKDFTTPGVTFIMQVPSYGDELQNFKMLQSAQHIADVEGGVVLDDQRRMMTPQKLREYQDIIREVKDANA
65	1YU5	PTKLETFPLDVLVNTAAEDLPRGVDPSRKENHLSDEDFKAVFGMTRSAFANLPLWKQQLNKKKEGLF

Table S1

66	1YZM	GSPLLQIHNITSFIRQAKAAGRMDEVRTLQENLRQLQDEYDQQQT
67	1ZHV	APRIKLKILNGSYGIARLSASEAIPAWADGGGFVSIITRTDELIVCLIDRIPQDVDPGWSCFKQGPFADDETGIIVSVISPLSTNGIGIFVSTFDGDHLLVRSNDLEKTADLLANAGHSLLLEHHHHHH
68	1ZZK	MGPITTTQVTIPKDLAGSIIGKGGQRIKQIRHESGASIKIDEPLGSEDRIITITGTQDQIQNAQYLLQNSVKQYSGKFF
69	2A0J	SLIGEILPLSHIVLDMEVGSKKRLFEEAGLLERESSLSHADVFECLFAREKLGSTLGGQVAIPHGRHAGVKQATGAFIRTPREVGFDPDGKPVSLIFILLVPENATGEHLEVLKLAGKFSQKSIRESLMTVSSAAEEVRAILT
70	2B8I	IRPYMKALIYETLVNLANQDPEQHATIRQNLQEQLDLPFDKQLALYAGALGPASSGKLENHEAISNAVDSSVQLLEI
71	2BK8	GAMVSGQIMHAVGEEGGHVKYVCKIENYDQSTQVTVYFQVGRQLENSEKYEITYEDGVAILYVKDITKLDGTYRCKVNDYGEDSSYAEILFVKGVRE
72	2BZX	PVFAKAIQKRVPCAYDKTALALEVGDIVKVTRMNINGQWEGEVNGRKGLFPFTHVKIFDP
73	2CGQ	AMEEAINATIQRILRTDRGITANQVLVDDLGFDSLKLFLQITELEDEFDIAISFRDAQNIKTGVDVYTSVAVWF
74	2CIU	SGDTQLFNRAVSMVEKNKDIRLLQCDDGITGKERLKAYGELITNDKWTRNRPVSTKKLDKEGRTHHYMRHFHVESKKKIALVHLEAKESKQNYQPDFINMYVDVPGEKRYLIKPKLHPVSN
75	2CJJ	GRPWSAKENKAFERALAVYDKDTPDRWANVARAVEGRTPEEVKKHYEILVEDIKYIESGKVPF
76	2CKX	RPFSSVAEVEALVEAEHLGTGRWRDVKMRAFDNADHRTYVDLKDCKWKTLVHTASIAQQRRGPEVPQDLLDRVLAHAYWSQQ
77	2CWP	MELYDVDEFWKFQMKVGLVKKAEKIKRTKKLIKLVDFGNEERTIVTGADQIPPEELEGKFFIVNLLKPKKFSGVESQGMILAEETDGKVYLPVPEEVPVGARVW
78	2CXC	ITLEELRYISVFHSITGVTAIRCIVDEENRILFLVSEGEAGRAIGRGGRIKLLREALGKNIEVVEYSSDLERIVKNLFPVGKIESINVRERNGVKQVVIKVSDDKGAAGKGGKNVKRARLVLSKLFGEVKVIR
79	2D59	TRPIDGLTDEDIREILTRYKKIALVGASPKPERDANIVMKYLLHEGYDVYPVNPKEYEVLGRKCYPVSLDPDKIEVDLVFKPKLTMEYVEQAIIKKGAKVWVQYNTYINREASKKADEAGLIIVANRCMMREHERLLGEK
80	2E1F	QPVISAEQEQETQVLYGKLVEARQKHANKMDVPPAILATNKILVDMAKMRPTTVENVKRIDGVSEGKAAMLAPLLEVIKHFCQTNVSVQTDLFSS
81	2EA9	MSNTTWGLQRDITPRLGARLVQEGNQLHYLADRASITGKFSDAECPKLDVVFPHFISQIESMLTTGELNPRHAQCVTLYHNGFTCEADTLGSCGYVYIAVYPT
82	2EIF	MVIIMPQTGVNVGSLKVGQYVMIDGVPCIEVDISVSKPGKHGGAKARVVIGIFEKVKKEFVAPTSSKVEVPIIDRRKGQVLAIMGDMVMQIMDLQTYETLELPIEGIEGLEPGGEVEYIEAVGQYKITRVI
83	2END	TRINLTLYSELADQHLMAYRELPVFGAVRKHVANGKRVDRDFKISPTFILGAGHVTFYDQKLEFLRKRQIELIAECLKRGFNIKDTTVQDISDIPQEFRGDYIPHEASIAISQARLDEKIAQRPTWYKYYGKAIYA
84	2F15	QARPTVIRWSEGGKEVFISGSFNNWSTKIPLIKSHNDFVAILDLPEGEHQYKFFVDGQWVHDPSEPVTSQLGTINNLIHVKKSDFEVF
85	2F10	VVMNDIIDVSIPIVAEVDKHPVEVLEILVELGFKPLANPLMRNTVGRKVSLLKQGSKLAGTMDKIVRTLEANGYEVIGLD
86	2F19	HFPGRAPIDAYGNGGFRFADMSHRGSIICIPSGIYIDMTGPVPTQEDISRVLEESDQIEVLLIGTGVELLRLPEELRVLLWEKRISDSTMSTGAAVRTFNVLLAEDRAVAALLFAVE
87	2FO3	YRIQKELHNFLNPPINCTLDVHPNNIRIWIYVGLENTYIANEVYKLIIFPDDYPLKPPIVYFLQKPPKHTHYVSNQDCLSLLGDDYNPSLSISGLVLSIISMLS
88	2FQ3	SKWFLNLEKIHSEIVQSLPEFFTNRIPSKTPEVYMYRNFVMVNSYRLNPNEYFSVTTARRNVSGDAAALFRLHKFLTKWGLINYQV
89	2FZP	SSGLVPRGSTIEYNEILEWVNSLQPARVTRWGGMISTPDAVLQAVIKRSLVESGCPASIVNELIENAHERSWPQGLATLETRQMNRRYYENYVAKRIPGKQAVVVMACENQHMGGDDMVQEPGLVMIFAHGVEEI
90	2GKG	KKILIVESDTALSATLRSALTEGRFTVDETTDGKGSVEQIRDRPDVLVLAVDLVLAGQNGYLCGKLKDDDLKNVPIVIGNPDGFAQHRKLKAHADEYVAKPVDADQLVERAGALIGFPE
91	2H36	KQDLEKIESDIINDWTEADDLDDALDFLMEKVSEFKIKFKDPLKVTEEEYRELLGNYDSSNSVSSNGITIDQYTYDEDDDIMYKLEFTYRKEDNKIYIYEVQGWREK
92	2HAZ	SHMDTPSSPSIDQVEPYSTAQVQFDEPEATGGVPILKYAEWRAVGEEVWHSKWYDAKEASMEGIVTIVGLKPETTYAVRLAALNGKGLGEISAASEFTQPV
93	2HC8	EAIKKLVGLAKATAVIRDGKEIAPVEEAVGDIVVRPEGKIPVDGVVVEGESYVDESMISGEPVVLKSKGDEVFGATINNTGVLKIRATRVGGETLLAQIVKLVEDAMG
94	2HDZ	LPESPKRAEEIWQQSVIGDYLARFKNDRVKALKAMEMTWNNMEKKEKLMWIKKAAEDQKRYERELS
95	2I6V	QEIFYVRLSQVKRDDKVLGYRVSPGKDPVLFESIGLQDGDMAVALNGLDLTDPNVMNTLQFSMNEMTEMSLTVERDGGQHDVYIQF
96	2IGD	MTPAVTTYKLVLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGVWVYDDATKFTTVE
97	2J49	PENYIRAYSMKLNWVDSLEIKPELSYIMPIFYFLNLVAKNPVYARRFFDRSPDFKDFHGEINRFLSVNSIDHIKENEVASAFQSHKYRITMSKTTNLNLLYFLNENESIGGSLIISVINQHLDPNIV
98	2J9V	HHHMFNAKYVAEATGNFITVMDALKLNYNAKQQLHPLLAELLISINRVTRDDFENRSLKLDIWRINKLSIGDTLTETQIRELLFDLELAYKSFYALLD
99	2O0Q	TLIYKILSRAEWDAKAQGRFEGSAVDLADGFIHLSAGEQAQETAAKWFRGQANLVLLAVEAPLEDLKWEEASRGGARFPHLYRPLLVEVTRREADLDDLADGVPQLGDHLAL
100	2O7I	HILNSSPSDRQINQLAQLRGPWEPMVLSLGLSQTDIYRCKANHPHNVSQVVEAFIRWRQRFQKQATFQSLHNLGRAVEVDPSLLHMLE

Table S1

101	2OCH	KETGYDVLGVKPDASDNELKKAYRKMALKFHPDKNPDGAEQFKQISQAYEVLSEKKRQIYDQGG
102	2OUM	ALLEKVDPNKIYTIIDEAAHLVKELATAKFDTEVEVHAKLGIDPRSDQNVRGTVSLPHGGRIEFNRNDKTGAHAPVGKASFPEKLADNIRAFIRALEAHKPEGAKGTFLRSVYVTTTGMGSPVRINPHS
103	2P5K	NKGQRHIKIREITSNEIETQDELVDMLKQDGYKVQTQATVSRDIKELHLVKVPTNNGSYKYSL
104	2PND	GHPTLKTPESTGTWKGDVKIQCIYDPLRGYRQVLVKWLVRHGSDSVTIFLRDSTGDHIQQAQYRGRLLKVSHKVPDVSLLQINTLQMDDRNHYTCEVTWQTPDGNQVIRDKIELRVRK
105	2PST	DDIQFQVVVNHEEQYSIWPEYKEIPQGWRAAGKSGLKKDCLAYIEEVWTDMRPLSLRQHMD
106	2QVO	RIKLLFKEKALEILMTIYESLGGNDVYIQYIASKVNSPHSYVWLIKKFEEAKMVECELEGRTKIIRLTDKGQKIAQQIKSIIDIM
107	2QVT	GYTRFYRSPTASVTLSGLVNVKWDNEQMTPLFKWIGGEQAEELHFCVHIAHSIGPKLNLARTLGTVNSNMDQHWAAQHRHSGATRRTIQDFHLFANDIPNFPDYIKILVPK
108	2RH3	IQVFLSARPPAPEVSKIYDNLILQYSPSKSLQMILRRALGDFENMLADGSFRAAPKSYPIHTAFEKSIIVQTSRMFPVSLIEARNHFDPLGLETARAFGHKLATAALACFFAREKATNS
109	2V75	YTENLKVIVAEKLAGIPNFNEDIKYVAEYVLLIVNGGTVESVVDLASLFDVSVRDLANVVQTAFFALEALQOGESAENIVSKIRMM
110	2V9V	GSPEKILAQIQEHREGLDWQEAATRASLSLEETRKLQLQSMAGQVTLRLRVENDLYAISTERYQAWWQAVTRALEEFHSRYPLRPGLAAREELRSRYFSRLPARVYQALLEEWSREGRLQLAANTVALAGFTPSF
111	2VC8	TDWLGSIVSINCGDSLGVYQGRVSAVDQVSQTISLTRPFHNGVKCLVPEVTFRAGDITELKILEIPGPGDNQ
112	2VO8	TTVGNSTIKVNDEQVQGSFAFEALGIEGLNGDTEVYSAEYLFEYNAEAFILNEITSFNDLSLVKSKVEPVGKVRILVASLGNEIEKDSDLVKVNLTPKISSELEVGLTTALVAGDGNTHDLELSSKEVKINEE
113	2W0G	HKTfVEKYEKQIKHFGMLRRWDDSKYLSDNVHLVCEETANYLVIWCIDLEVEEKALMEQVAHQITVMQFIELEAKSLKVDPRACFRQFFTKIKTADRQYMEGFNDELEAFKERVGRRAKLRIEKAMK
114	2W0I	MPLQPEAQRALQQLKQKMVNYQMKLDERETIELVHTEPTDVAQLPSRVPRDAARYHFFLYKHTEHGDPLESVFYISMPGYKCSIKERMLYSSCKSRLDSVEQDFHLEIAKKIEIGDGAELTAEFLDDEVH
115	2Y6X	PTGLTGNFREDTALISSLEAIALPENDPNKKAQAQAEARKKLNDFALYRRDDSLRSLSSFTMTQALNSLAGHYSSYPNRPLEKLKARLEQEFKQVELALDREAKS
116	2Y9F	MTDIYIPPEGLYFRLGFGASRQVIFARNSPSPDVGLSPVNDQATDQYFSLYGTGEHAGLYAIKSKATGKVLFSRRPAEPYVGGIDGGRYPDNWFKIEPGKTYLSKYFRLVQPSGTALVSRTHLQPYFWNHPQTEVDDQYFTFLFE
117	2YXY	EQKRYSEMTEKEELQOEIAMLTEKARKAEQMGVMNEYAVYERKIAMAKAYMLNPADHFPGEIYEIAGAPGEYFKVRYLKGVFAGWRLKGNGEELPISLLRKP
118	2ZQE	VKEVDLRGLTVAEALLEVDQALEEARALGLSTLRLLHGKGTGALRQAIAREALRRDKRVESFADAPPGEGGHGVTVVALRP
119	2ZRR	GDDRRKKAIEVITELDDLEIDLGNESLRKVLGSYLKKNEGTSVPLVLSRNMIEISNAIKDGVSLNENQSKKLKELMSIS
120	3A4C	RCPEQELRLQRLERLPALARVLRNVFVSERKPALTEMEVVCARMVDSCQTALSPGEMEKLHLVLLAELLPDWLSLHRIRTDITYVKLDKAVDLAGLTARLAHVHVAEGL
121	3ADG	SHVFKSRLQEYAKYKLPVYIEIVKEGSPHKSFLQSTVILDGVRYNLSLPGFFNRKAAEQSAAEVALRELAK
122	3AG7	GPLGSEEIKNIDAKIRKWSGKSGNIRSLSTLQYILWSGSGWKVPVPLMDMIEGNAVRKSYQRALLILHPDKLQQKGASANQYMAEKVFELLQEAWDHFNLTLP
123	3CE7	VVDTDINAVTNYIVGMCQKFLQKGEKVTSSKLEELRTREDRLWDCLDTEFVLDVEEIFDVTVPDEVADNFQTLQEIAFVVSERAKAG
124	3DJN	DMLPEIAAAVGLSSLLRTRGCVSEQRKLVFSRALQDALTDHYKHHWFPEKPSKSGSGYRCIRINHMKMDPIISKVASQIGLSQPQLHRLLPSELTLWVDPYEVSYRIGEDGSCVLYEE
125	3DKM	GRENLVFQGLKYMVPGARVTRGLDWKWRDQDGSPPQEGTVTGELHNGWIDVTWDAGGSNSYRMGAEGKFDLKLAPGYDP
126	3E8V	AKGSVLVTDAGGQPVADATVEFKVYNYAEFYTVATKHTDRSGHASLTAGKGDMLVWASKDGRFGYSKLSFGKDNLKITLDK
127	3EGN	ECISRRELEKGRISREEMETLSVFRSYEPGEPNCRIYVKNLAKHVQEKDLKIFGRYVDFSSETQIRIMFDIRLMKEGRMKGAFIGLPNEKAAAKALKEANGYVLFGKPMVVOFARSARP
128	3FBL	YHKLRILAICEICKTDGIPNIKWGMIAFGEKLLKSYLKMKGASSDMMIAEYINNAISAFSSRTGISQETAQKIADFITSNY
129	3FDR	LQLDKLVNEMTQHYENSVPEDLTVHVGDIVAAPLTNGSWYRARVLGTLENGNLDLYFVDFGDNGDCPLKDLRALRSDFLSLPPQAIIEC
130	3FT9	VQVTFVTQKGSDDPKLVLDIKYTRPGDSLAEVELRHQGSSEWEPLTKKGNVWEVKSSKPLVGPFNFRFMSKGGMRNVFDEVIPTAFSIGKTYKP
131	3GK5	YRSINAADLYENIKAYTVLDVREPFIIFGSIANSINIPISLEKRWKILERDKKYAVICAHNRSAAAVEFLSQLGLNIVDVEGGIQSWIEEGYPVLE
132	3HJ7	KGYWFELPVPALLPLPNGYAIISFEFGEHYPRKQAGNDWVVDPASVSLPLRVTRRRGDRMVLKGTGGTKLKEIFIEAKIPRMERDRWPIVEDADGRILWVPGLKKSFAEQNRGQARYILLQYQAMNS
133	3HNX	HMSYADSSRNAVLNNGRTRLAECRNADGNWVTSELDLDTIGNNDGHFQWGGQNFTEAEDIRFHPKEGAAEQPILRARLDCNGEFHDRDVNLNRIQNVNGLVFQ
134	3I8Z	FAVESIEKKRIRKGRVEYLVKWRGWSPKYNTWEPENILDPRLIAFQNR
135	3ID1	MVRPVVGEIAANSIAEAQIAPGTCLKAVDGIETPDWDAYRLQLVDKIGDESTTITVAPFGSDQRRDVKLDLRHWAPEPKEDPVSSLGIRPRG

Table S1

136	3ID4	MIEPVLENVQPNASAASKAGLQAGDRIVKVDGQPLTQWVTFVMLVRDNPGLSLALEIERQGSPLSLTIPESEKPGNGKAIGFVGIEPKGKASPV
137	3IDW	KYDWFEFFLNCGVDSVNCQRYTINFREQLTEDMMPDINNMSMLRTLGLREGDIVRMKHLDKKFGR
138	3IHO	KWTPPRSPFNLVQETLFHDPWKLIIATIFLNRSTGKMAIPVLWKLEKYPSAEVARTADWRDVSELLKPLGLYLRAKTIVKFSDEYLTQWKYPIELHGIGKYGNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENH
139	3IU5	SMTVDPIAVCHELYNTIRDYKDEQGRLLCELFIKRRNQPDYIEVVSQPIDLMKIQQLKMEYDDVNLITADQFLLFNNAKSYKPSPEYKAACKLWDLYLRTRNEFVQK
140	3JU0	SLTDSKVNAKSLKEYKLTDFGFMHLLVHPNGSKYWRLSYRFEKKQRLLALGVYPAVSLADARQRDEAKLLAAGIDPSAKKQADNKTIQEKR
141	3JVE	APKPLHKVVVCSKLSKKQSELNGIAASLGADYRWSFDETVTHFYQGRPNNDTNEYKSVKRGVHIVSEHWLLDCAQECHKLPESLYPHTYN
142	3K3V	ESSWRYIDTQGGIHGPFPTQMMSQWYIGGYFASTLQISRLGSTPETLGINDIFITLGELMTKLEKYDTPFTTFDKLHV
143	3KT9	MMRVCWLVRQDSRHQRIRLPHLEAVVIGRGPETKITDKKCSRQQVQLKAECNKGYVKVQGVNPTSIDSVVIGKDQEVKLQPGQVLHVMNELYPYVEFEE
144	3L1X	SDAPDEFDRPLMDTLMDPVRLPSGTIMDRSIIHLHLLNSPTDPFNQRTLTESMLEPVPELKEIQAWMREKQNS
145	3LF9	MINEIKKDAQERMDSVEALKNNLSKVRTGGGGTEERRKDLVKIVRGEAEGGRVAVRNIARDAANDLAALGDKDEVNWFDISQALWEIQKLTDAVVKIDEVLAKEKELMEVLEHHHH
146	3LLB	SYIKPLPSGDFIVKALTPVDADFNDFFGSEFSDEEFDTVGGLVMSAFGHLPKRNEVELGEFRFRVLNADSRRLHLLRLSPL
147	3LMO	TSTFDRVATIAETCDIPRETITPESHAIDDLIGDLSLDFLIAFAIDKAFGIKPLEKWTQEVNDGKATTEQYFVLKNLAARIDELVAAKGALE
148	3MX7	MTDLVAVWDVALSDGVHKIEFEHGTTSKGRVVYVDGKEIRKEWMFKLVGKETFYVGAAKTKATINIDAIISGFAYEYTLINGKSLKKYM
149	3NDQ	RAPSTSDSVRLKCREMLAAALRTGDDYIAIGADEEELGSGIEEAIYQEIIRNTDMKYKNRVRSRISNLKDAKNPNLRKNVLCGNIPDLFARMTAEM
150	3NE3	GIVGVGIDLVSIPDAEQVDQPGTVFAETFTPGERRDASDKSSAARHLAARWAAKEAVIKAWSGSRFAQRVLPEDIHREIEVVTDMWGRPRVRLTGAIAEYLADVTIHVSLTHEGDTAAVAILEAP
151	3NPH	ELRSRSTEEEDAVILAVYRQVLGNDHLSQERLTAESLLRGREISVRDFVRAVALSEVYRQKFHNSNPQNRFIELNYKHLGRAPYDQSEIAFHTDLYHQGGYAEINSYIDSVEYTENFGDWVVPYFR
152	3O48	DFWPTLKDAYEPLYPQQLERQQVSEGGPTATIQSRFNAYWGLIKSTDVNDERLGVKILTDIYKEASRRRECLYLTIGCYKLGEYSMAKRYVDTLFEHERNNKQVGALKSMVEDKIQKEENLYFQ
153	3O5E	SPTATVAEQGEDITSKKDRGVKIVKRVNGEETPMIGDKVYVHYKGLSNGKKFDSSHDRNEPVSFSLGKGQVIAWDIGVATMKKGEICHLCKPEYAYGSAGSLPKIPSNATLFFEIELLDFKGE
154	3OBS	SAVSESQLKMMVSKYKRDLTVRETENVITLYKDLKPLVDSYGTGSRELMNLTGTIPVPYRGNTYINIPICLWLLDTYPYNPPICFVKPTSSMTIKTGKHVDANGKIYLPYLHEWKHPQSDLLGLIQVMIVFGDEPPVFS
155	3ONH	KVCRGVIKLSSDCLNKMKLSDFVVLIREKYSYPQDISLLDASNQRLFDYDFEDLNDRTLSEINLNGSIIIFSDEEGDTMIRKAIELFDVDDDELPCNTCSLDPVEVPLIKA
156	3ONJ	SLLSIYESDFKTTLEQAKASLAEAPSQPLSQRNNTLLKHVEQQQDELFDLLDQMDVEVNNSIGDASERATYKAKLREWKKTIQSDIKRPLQSLVDSGD
157	3PDG	MVTIDSPVAGERFEAGKDINISATVKSCTPVSKVEFYNGDTLISSDTPYAKITGAAVGAYNLKAVAVLSDGRRIESPVTPVLVKVIL
158	3QD7	DTLQLDNFLTGTFLDIPLSQPLEFRREGHQHGVLDKLRSGKYPQASLNLLRQPVEECKRMVFSFIQALADGLRNVLIIHGKGRDDKSHANIVRSYVARWLTEFDDVQAYCTALPHHGGSGACYVALRK
159	3RJ2	IVANCEFVNATGKKTTLVNENWAKYCWITWYKFPEKYTLRLYSVDGEMFMHRVTFNATGRYITHLNLHGLEDVLEGLSAPKDAAYARIAAINVSLTNPBGDVHMHYDETEGEQIRSYDAAEFARTLAA
160	3RJP	MYRDLVLPQNRSVNRGDDEISLTKREYDLLNILMTNMNRVMTREELSNVWKYDEAVETNVVDVYIRYLRGKIDIPGKESYIQTVRGMGYVIREK
161	3T46	HQALVDQLHELIANTDLNKLSYLNDAFQKRDILAAHYAKSAIRTKNLDQMTKAKQRLESYNSISNPLHSQNN
162	3V4Q	LAAKEAKLRDLEDSLARERDTSWRLLAEKEREMAEMARMQQQLDEYQELLDIKALDMEIHAYRKLLGEGER
163	3ZHI	KTDTSNRLKQIMAERNLKQVDILNLSIPFQKKFGIKLSKSTLSAYVASVQSPDQNRILYLAKTGLVSEAWLMG
164	3ZZP	QGYTTWYQVEMPEDRVNDLARELRIRDNVRRVMVASTTPGRYEVNVLNPNLDQSQLQNEKEIIRALENYGA
165	4EIH	SLEKHSWYHGPVSRSAEYLLSSLINGSFLVRESESSPQLSISLRYEGRVYHYRINTTADGKVYVTAESRFTSLAELVHHHSTVADGLVTTLHYPAKCN
166	4F25	SGNIFIKNLDKSIDNKALYDTFSAFNGILSCKVVCDENGSKGYGFVHFETQEAERAIEKMNGMLLNDRKVFVGRFKSRKE
167	4FI5	SMTMEELQREINAHGQLVIARQKVRDAEKQYKDPDELNKRITLDREGVAVSIQAKIDELKRLQADRIAT
168	4G3O	RGSENLYFQGQLNAMAHIQEMFPQVPYHLVLDLQLTRSRVEITDNILEGRI
169	4GA2	RRSKADVERYIASVQGSTPSRQKSIKGYFAKLYEKEYDLAKKYICTYINVQERDPAHRFLGLLYELEENTDKAVECYRRSELNPTQKDLVLKIAELCKNDVTDGRAKYWVERAAKLPFGSPAVYKLKEQLLDCEGED
170	4GCO	YINPELAQEEKNKNEYFKGDYPTAMRHYNKVRDPENALYSNRAACLTKLMEFQALRDDCOTCIRLDSKFIKGYIRKAACLVAMREWSKAQRAYEDALQVDPNSNEEAREGVNRNCLR

Table S1

171	4GEI	KIKSEFVWFNDPEKVYSGSERVAGRVIVEVSEVTRVKAVRILASGAVKVLWMQGSQQCAQTSLEYLRVEDTLLEDQPTGENEMVIMRPGNKYEYKGFELPQGGLGTSFKGKYGSVDYVWKAFLDRPSQPTQETKKNFEVVDLV
172	4GMQ	NAVKKNKRVLRGSVKEANYFVEGEASAATIDAVLNDVDLVITIKIDADEIAALAGKLNGLTVADEIKNVWKEEVSRLVGAGKLEKEDIKALVA
173	4HQA	SRKFIANARVENCAVIYCNDGFCELCGYSRAEVMQRPCTCDFLHGPRTORRAAAQIAQALLGAERKVEIAFYRKDGCFLCLVDVVPVKNEGDGAVIMFILNFEVVMK
174	4HTJ	ARGYIVTRDPLRPEEGRRLVEDVARLLQVPSSAFADVEVLGPVTFKVSANVQNVTTEDVEKATVDNKKLEETSGLKILQTVGVSXS
175	4IGV	DLSGKMVKQVEILSDGIVFYEIFRYRLYLISEMSPVNIQGVDLLEGNWGTGVSIVFFKYTIDGKEKTAKDVIDEAEETKSVTFKIVEGDLMELYKTIIIVQVDTKGEHNSVTWTFHYEKLKEDVEEPNTLMNFCIEITKDIETYHLK
176	4IPC	GSHMVGQLSRGAIAIMQKGDNTNIPILQVINIRPITGNSPPRYRLMSDGLNTLSSFLATQLNPLVEEQSSNCVCQIHRFIVNTLKDGRRVILMELEVLKSAEAVGVKIGNPVVYNE
177	4J5Q	EDPEGSRTYVKGDLFACPKTDSLACISEDCRMAGIAVLFKKFGGVQELLNQKKSGEAVLKRDRYIYLLITKKRASHKPTYENLQKSLEAMKSHCLKNGVTDLSMPRIGCGLDRLQWENVSAMIEEVFEATDIKITVYTL
178	4JOX	AREVHIDVNNKTGHTLQLEDKTKLDGGRWRTSPTNVANDQIKTFVAESNGFMTGTEGTIYSINGEAEISLYFDNPFAGSNKYDGHNSKQYEIITQGGSGNQSHVYTIQTSSRYG
179	4KMC	TNWSRRYKANLEKLASGDVNKVAEVVRDLWRRDQERGLSAGEKRLAKARQILVGEALAESTDDAKAETILDEVLAAA
180	4LTT	MLKLNKKSFQKDFDKLLNGFDDSVLNEVILTRKKEPLDPQFQDHALKGKWKPFRECHIKPDVLLVYLKDDLEILLRLGSHSELFEG
181	4MNO	VIRVPLPEGNLQFGVVEQALGAGWMDVRCEDGKIRRCRIPGKLRRRVWIRVGDLVIVQWPVQSDKRGDIVRYTQTQVDWLLRKGITQEFLTG
182	4MQ3	MGIWQMDCTHFDGKIIIVGIIHVESGYIWAQIISQETADCTKAVLQLLSAHNVTETQDNGPNFKNQKMEGVLYNMGVKKHFGIPGNPQSQALVENVNHTLKVWIQKFLPETTSLDNALSLAVHSLNKK
183	4N1L	SSFFPDFGLLLYLEELNKEELNTFKLFLKETMEPEHGLTPWNEVKKARREDLANLMKKYYPGEKAWSVSLKIFGKMNKDLKERAKEENWVSAQL
184	4N6T	ENSLEIEELARFAVDEHNKKENALLEFVRVVAKEQVAVAGTMYLLEAKDGGKKLYEAKVWVKPWENFKELQEFKPV
185	4NPN	HINLVAGQDGSVVQFKIKRHTPLSKLMKAYCERQGLSMRQIRFRDGGPINETDTPAQLEMEDEDIDVF
186	4O7Q	AMESKYKEILLTGLDNITDEELDRFKGLSDEFNIATGKLIHTANRIQVATLMIQNAGAVSAVMKTIRIFQKLNMYLLAKRLQEEKEKVDKQYK
187	4O9L	PHMTVAEDKTFQYRQHHSNFSRIHVLRLPYLSCLTTSQDRLRATYERWGNQDTLLELFTSLRRRNGWVHSLIGALRACELSGLADEVARIYHS
188	4OD6	RFDNELQLRGLSVEAAVEELRAAIAEARALKETPLRVVHGKGMGVLRRTLRDYLTDKNVEFHDAAEQGGHGVITIVNVKR
189	4OIX	MLKRMYYARYGLVQGVGFRKFVQIHAILRGIKGYAKNLPDGSVEVVAEGYEEALSKLLERIKQGPAAEVEKVDYSFSEYKGEFEDFETY
190	4PGO	EILSKKPAGKVVEEVNIMGDKDVIIGTVESGMIGVGFVKVGPSPGIGGIVRIERNREKVEFAIAGDRIGISIEGKIGKVKKGVDLEIYQT
191	4Q2S	IAESLRRLEKYVAGSVKECVAEWCNMPVAVGFDVLSEISYDRMLENCNLLLLTFIYHISLLDSVDDRLSKRMEYISRICLNIDVNDPKVETVHPVLTLTREALLRQSEFFSPIFKRRLVLLRALDGKISEI
192	4Q6L	YDWNGAMQPLVSKMLQADGVTAGSVLLVDSVNNRTNGSLNANEATETLRNALANNGKFTLVSVQQLSMAKQQLGLSPQDSLGRSKAIGARNVGAQYVLYSSASGNVNAPALQMLMLVQTGEIIVSGKGAQQ
193	4QB0	EHSFEEMYRHILRSQGPDAVLYYHMMKDEPVVFTSDGKEYTYPDSLEEEYPPWLTEKEAMNEENRFVTLGGQFYWPVMNHKNKFMAILQHQQ
194	4QCJ	LPAGSALLVVKRGNAGARFLLDQPTTTAGRHPESDIFLDDVTVSRRHAEFRINEGEFEVVDVGSNGTYVNRPRNAQVMQGTGDEIQIGKFLVFLAGPA
195	4QUC	EYVVEKILGKRFVNGRPQVLVKWSGFPNENNTWEPLENVGNCMKLVSDFESEVFRL
196	4RWU	MVKETKLYDLLGVSPSANEQELKKGYRKAALKYHPDKPTGDTKEFKEISEAFEILNDPQKREIYDQYGLEAARSGGPSFGP
197	4W8G	TYDVNICYAPEDREWVINTLVFKLERAGIKTFVNIRDDTPGNFFAENIMDAIENSRTIVVMSPDFFKNNICDKTLQIGLSHQIIPILYRCPCEVPYFLNHMTYLDWCDKDVPRPVFWRNLFRDIRN
198	4YKD	ASATELLQDYMILTRLTKSSQEIQQFAALLHEYRNGASIHEFCINLRQLYGDSTRKFLLLGLRPFPIEKDSQHFENFLETIGVKD
199	4ZC3	LSPKQERFIEEYFINDMNATKAAIAGYSKNSASAIGAENLQKPAIRARIDARLKEI
200	5CVZ	AAATSLVYDTCYVTLTERATTSVQRQSFTLKGMDRAFQVVAFTIQGVSAAPLMYNARLYNPGDTSVHATGVQLMGTVPRTVRLTPRVGQNNWFFGNTEEAETILAIIDLVSSTGANAPSNTVITGCFRLAPSELQSS
201	5CW9	MEMDIRFRGDDEPAYKALREMRQARKFAGTVTVTLIRFRGDDELALEKALKEMIRQARKFAGTVTYTLGDNLEIRITGVPPQVILELVEAIRLAKEFNITVVELVIRITGVPEQVRKEAKEAERLAKEFNITVYTIIRL
202	5DBL	HMIAPGHRDEFDPKPLTGEKEEVPKPGIKNPETGDVVRPPVDSVTYKVPVKGDSIVEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTIITPTLKNPLTGEIISKESKEEITKDPINELTEWGPET
203	5DMA	GWKVGDRANHRKWIGTVSVRGGGDDQELDIAPFSPIGIKRLLAKFAPIEKV
204	"5E10"	GVQVPDVRGQSSADAIATLQNRGFKIRTLQKPDSTIPPDHIGTDPAAANTSVSAGDEITVNVSTGPEQREIPDVSTLTLYAEAVKKLTAAGFRFKQANSPTPELVGKVGITNPPANQTSAITNVVIVGSG
205	5E1Y	SMEILQVALHKRDSGEQLGIKLYRRTDEPGVFILDLLEGLLAQDGRLESSNDRVLAINGHDLKYGTPELAAQIIQASGERVNLTIARPGKPEIEL

Table S1

206	5EEP	GHPFGTTVTAETLRNTFAPLTQWEDKYRQLIMLGKQLPALPDELKAQAEIAGCENRVWLGYTVAENGKMHFFGDSEGRIVRGLLAVLLTAVEGKTAELQAQSPALFDELGLRAQLSASRSQGLNALSEAIIAAKQV
207	5EFX	GARLQEIYNRMDPRAQTPVPGKPGFGREELLRRKLHHDGCLLWKATGRFKDVLVLLMTDVLVFLQEKDQKYFFPTLDKPSVVSLLQNLIVRDIANQEKGMFLISAAPPEMYEVHTASRDDRSTWIRVIQQSVRTCPSREDFPLI
208	5EJO	AMITDPMDLLRFLDGVQDSTFSLGTVTEIAQKNLPQYNKQTIKNTIKEYAIRSSGKGDLPKRWVIKDAQNWENLRANAN
209	5EQZ	HHHYVEEKKEIDSLMEDVLALVNDSSGGKFQDYDKINELKENLKDIGNAELKEKLLNLQNSFQDKLAAKLAALKAAKNTIENITDKDQDISKRKIWSEAKLVGTVPLLSNTSGNGDKMSKNAVEQIDKVIKFL
210	5EWR	DPRAYFPAPADLVVEIDLVLQQASHYKQIKKGLNEVLKSMNRGLAEFVVLAADTQPLEILLSAPLVAEDKAVPYVFPVSKAALGRACGVSRPVIACAVLRADMSQLRNQITALRTKIEQLL
211	5FD9	SPTKAIVRLREHINLLSKKQSHLRTQITNQENARIFLTKGNVMAKNALKKKKTIEQLLSKVEGTSMESMEQQLFSIESANLNLETMRAMQEGAKAMKTIHSGLDIDKVDETMDEIREQVELGDE
212	5FU5	GPVAGNYVHDFGTANGTSSSFYTIAGNLSTSGKTATYNGKTLTQCLKMETATSISFTAPSAGKLTLVFAEAAATAKVGDGNKVTSASNGIITVDLAQGAHTITKADACNLFYMEYAALEH
213	5GUA	MENVVSAPMPGKVLRLVVRVGDRVRVGGQLLVLEAMKMENEIPSPRDGVVKRILVKEGEYVDTGQPLIELG
214	5H1D	PNVVVTRLTLVCESAPGPITMDLTGDLEALKKETIVLKEGSEYRVKIHFKVNRDIVSGLKYVQHTYRTGVKVDKATFMVGSYGPREEYFELTPVEEAPKGMMLARGTYHNKSFFTDDDKQDHLSEWNL
215	5HBQ	TTISPHDAQELIARGAKLIDIRDADEYLREHIPEADLAPLSVLEQSGLPKLRHEQIIFHDQAGKRTSNNADKLAAIAAPAEIFLLEDGIDGWKKAGLPVAVN
216	5HDW	SEFVATTGDITVSSTFPELSSVHPPHYFFTYRIRIEMSKDALPEKACQLDSRYWRITNAKGDVEEVQGGPVVGEFPIISPGRVYEYTSCTTFTSTSGYMEGYTFHFYFKDKIFNVAIRPFHMACPT
217	5HFD	GSPEFLGEEDIPREPRRIVIHGSGTGLGFNIVGTEDGEGIFISILAGGPADLSGELRKGDQILSVNGVDLRNASAEQAAIALKNAGQVTITIAQYKPEEYSRFEANSRVDSSGRIVTD
218	5HU4	DAVVGSIAPSPVDVNLVFKGTNTANLLAGATTMRSDQVMGKGNYPAGHHMRDESMLFGPIMKVKKGDKIYLTDLLENLEYVTETKTIDETEVSVIDNTKDARITLITCDKPTETTTRKFVAVGELEKTEKLTKELENKYFPSK
219	5I41	SHMNTNMVASELGVSAKTVQRWVKQLNLPARNELGHYSFTAEDVKVLKSVKKQISEGTAIQDIHLP
220	5IJM	SHMSAADFEAAVAYVRSPLKEGPVQLDNAALKQFYSYLKQATEGDTVGSQPWAVQVEARAKWDAWNSCKGMKSEDAKAAVYVRRLLTLRSQGIQWKPG
221	5IWH	GIKSFKHGKLKLLFEKGVTSQVPAQDVDRINDRLQAIDTATEIGELNRQIKLHPLKGDREGYWSITVRANWRITFQFINGDAYILNEDYKLGPEQ
222	5JBM	QKAMITDPMDLLRFLDGVQDSTFSLGTVTEIAQKNLPQYNKQTIKNTIKEYAIRSSGKGDLPKRWVIKDAQNWENLRANAN
223	5JRT	VDFSITQFVRNLGLEHLMIDIFEREQITLRLVLEMGHKELKEIGINAYGHREKLIKGVRLISGQ
224	5KIU	HHQKLSARLRLRQQLDRAAAEVDPVVVKRQEALAAARLKMQEELNAQVEKHKEKLKLEEEKRRQKIEMWDSM
225	5KUX	GIVNVPNCNTTKYQQLARTAVAIYNYHEQAHLTFVENLNCKEQLGEGDYYYITLAATDDAGKKAIEAKIGVSESAGWTGVVEEFKLVGSLEH
226	5L33	HMPEEKAARLFIEALEKGPPELMRKVISPDTRMEDNGREFTGDEVVEYVKEIQKRGQEWHLRRTYKEGNSWRFEVQVDNNGQTEQWEVQIEVRNGRIKRVITITHV
227	5ML3	SAKDERAREILRGFKLNWMNLDAETGKILWQGTEDLSVPGVEHEARVPKILKCKAVSRELNFSSTEQMEKFRLEQKVYFKGQCLEEWFEEFGFVIPNSTNTWQSLIEAAPESQMPASVLTGNVIIETKFFDDLLVSTSRVRLFVY
228	5MR1	ADCQGWLYKKKEGSFSLSNKWKFWVILKGSSLYWYSNQMAEKADGFVNLDPFTVERASECKKHAKHAKISHPQIKTFYFAAENVQEMNVWLNKLGSAVIHQ
229	5NOD	VTSVAMPYSYGSSLEFTKNLQIVGIKEANIEVEVTTAPAGSVEGMVVEQSPRAGEKVDLNKTRVKISYKPK
230	5OHQ	GPWVTTDIQVKVRDITYLDTQVVGGTGVIRSVTGGMCSVYLKDESEKVSISSEHLEPITPTKNKVKVILGEDREATGVLLSIDGEDGIVRMDLDEQLKILNRLFLGKLE
231	5TPJ	MSREEIRKVVVEEMVRKLKQSPEDISKYSPDVRLEVGNITYFEGSEQVTKFWRMLTKFVDRVEVRKVQVDGNHVRVEVEVEWNGKKWTFEVEVEVRNGKIKIRLQVDPEFKVQVNIWNLL
232	5UFT	SMIDDIYNKRILEFAGNMERIGQLAEPDAVATVHSLKCGSTVTYVKMRDGVVTDFAHEVKACALQAASSVMARNVIGATADELRAARDAMYRLKENGPAPEGRFADMKYFEPVRDYKARHASTLLTFDAVADCIRQIEEKA
233	5UMV	SSKIFKNCVIYINGYTKPGRQLQHEMIVLHGGKFLHYLSSKKTVTHIVASNLPKKRIEFANYKVVSPDWIVDSVKEARLLPWQNYSLTS
234	5WD9	APVTELTRLKEYMEDQIAKAKESSSLTAQLKFLENAHTEHFVKMGSLTTIYKGGSEVVDRLKIEIRSLYEEMLELKDCKRDIQIYQ
235	5WF2	AMTSDCQQLSLVQTVTRGSRAFLSREEAQHFVKECGLLNCEAVLELLICHGGTSECQVMRALGAVASLGCTDLLPQEHILLTRPRLQELSAGSPGPVTNKATKILRHFEASC
236	5XUO	AGFGAQHRKMGKSLYLRNEVFAAWIEKVDALVQDELGYSLVELILDDAQDYGIETTQVTFIQAIALGELLRHHGAKPAAVIGQSLGEAASAYFAGGLSLRDATRAICRS
237	5Y3C	PAAASTKVLYYTDRSLTPFLVNIPKRLGDVTLQDFKAAVDRHGSFRYHFKSLDPEFGTVKEEVQDDAIVIPGWEGKIVAWVEE
238	5Y4U	ARLTKLVNAAPVMLFMKGSPPSEPKCGFSRQLVGILREHQVRFGFFDILRDESVRQNLKFSSEWPTFPQLYINGEFGGGLDIIKESLEEDPDFLQHALQS
239	5Y6H	YHEQFLKQNPLAVLGVLRLHKAAPLRLSWNGQLISKLAITPDKLVDFGSAQEDNIAVLKAQHITITAEQGAKEFTVEQLQQSEYLLQPAFITVPPPTLWFVLE
240	5YDE	MADVLSVLRYNQKKEIVKGDVEIFGEFSWPKNVKNYVWGTGKEGQPREYITLDSILFLNNVHLSHPVYVRRRAATENIPVRRPDRKDLLGYLNGEASTSASIDRS

Table S1

241	5YDN	MNYATVNDLCARYTRTRLDILTRPKTADGQPDVAEAEQALADASAFIDGYLAARFVLP LTVVPSLLKRQCCVVAWFYLNESQPTEQITATYRDTVVRWLEQVRDGTKDPG
242	5YRY	PEFMENLVTLSRQGIENLMKLENRKEPWIVVLYAPWCPCQAMEASYDELADKLAGSGIKVAKFRADGDQKEFAKQELQLGSFPTILVFPKNSSRPIKYPSEKRDVESLTSFLEH
243	5YWJ	LLGFYKQYKALSEYIDKKYKLSLNDLAVLDLTMKHCKDEKVLMSQSLKTAMDELDLSRTKLLVSIIRLIEKERLSKVRSSKDERKIYILNDDISKFNALFEDVEQFLN
244	5Z59	WPSAKFIDGRVAFSRMPAERELDEVARDFAVVVLVEDYELPYSLDEWEKRGVEVLHGPIPDFTAPSVQEQLLEILRWIEERVREGKKVLIHCMGGLGRSGTVGVAWLMYSRGLSLREALMEVRRKRPGAVETQEOMEVLKELEERI
245	5ZT3	YPYPVDEIIGGDSVQSIQRRLLGTNNWNP SAHDMQMSRIQAEDLFELKVEIIRKMAGLHPSGDWMGWGARALDNPRATGEEDLARLHQMLDDLQSRNEQSATFWRLVERVRLRA
246	6AR0	PSALAIFTCRPNSHPFQERHVYLDEPIKIGRSVARCRPAQNNATFDCKVLSRNHALVWFDHKTGKFYLDQTKSSNGTFINSQRLSRGSEESPCEILSGDIIQFGVDVTENTRKVTHGCIVSTIKLFLPDGMEARLRS
247	6B26	NSYVALYKFLPQENNDLALQPGDRIMLVDDSNEDWWKGKIGDRVGFFPANFVQVRPGENVWRCCQPFSGNKEQGYMSLKENQICVGVGRSKDADGFIRVSSGKKRGLVPVDALTEI
248	6B8C	VNGAAIAEAYKYGTPYVWGGKDPSGFDSCGFTRYVYLQVTGRDIGGWTVPQESAGTKISVSQAKAGDLLFWGSPGGTYHVAIALGGGQYIHAPQPGESVKVGSVQWFAPDFAVSM
249	6BL5	RVPRHLHAYGVFALPFPMPDPVEWGNWFAGPHPKAFLVSVHPSGPKAGHVYPTDLSDPDSVANVIGMVLGDHGYEADHNVTVTLRAAVPIEYVQQGIEAPLPQDPAVLNAAPQLKLVIKGHYFFDYTR
250	6D1R	KAYRIKKNADFQRIYKKGHSVANRQFVYTCNNKEIDHFRLGISVSKKLGNVLRNKKIRAIENFKVHKSHILAKDIIIVARQPAKDMTTLLQIQNSLEHVLKIAKFVN
251	6DGA	MKDDLDNNLLYRYCGATSPFWRLPLDSNALQLAASEEAVTSHVPLTPEQAAQIRTM SVITSSVTLSSLFSGELVPVHLVGRKVS RKEWAGTASA
252	6E4N	GSHMRVQVSGMSDETTWHTLKDHLRQAGEVTFCKVFSGGRAVVEFVTPEDAARAITEQLQASELEGATLFLR
253	6ELM	SMAEDTGVRELAEEDHGRKSTIALRLWVEDPKLKGPKDNGAIEFTFDLEKETPDEVAQEMIESGFFHESDVKIVAKSIRD RVALIQWRRERIWPA
254	6EW6	DSQIQFTRHASDVLNLRNLRSDILTDVVIVSREQFRAHKTVLMACSGLFSIFTDQLKRNL SVINLDPEINPEGFNILLDFMYTSRNLNREGNIMAVMATAMYLQMEHVVD TARKFIKAS
255	6EXX	KISTLFLNLSAVCNKEFLKYLCHQENIRPFKIQIDGYDENSSTSYSGFIKFRNFEDATRIFNFLNNRLVGGSVITTSWE
256	6F53	LPTAQEVQGLMARYIELVDVDIEAIVQMYADDATVEDPFGQPPHIGREQIAAFYRQGLGGGKVRACL TGPVRASHNGCGAMPFRIEMVWNGQPCAVDAISVMRFEHGRIOQTMQAYWSEVNL SV
257	6FC5	GPMDRYQRKIGCFIQIPNLGRGQLKYVGPVDTKAGMFAGVDLLANIGKNDGSF MGKKYFQTEYPQSGLFIQLQKVASLIEKASI
258	6GPM	GQVIVVRKSDFRPKTYEDLVGKTAVVQIGTTGDIEVSKYDGIKVVRFDKFTDAFLELKRGRADAWLDSATARAFAKPNPDLVSSGVLSS
259	6IPY	TGVTFLVALYDYEAITEDDL SFHKGEKFQILNSSEGDDWWEARSLTTGETGYIPSNYVAPVDSILEHH

Disprot dataset

#	Disprot ID	Sequence
1	DP00070	SKTKEGVVHGVATVAEKTKEQVTN VGGAVVTGVTAVAQKTVEGAGSIAAAT
2	DP00284	TNVLNQGGVEGYDSSSNVIQDSNLDTAINPDDRPLSEIPTDDNPNMSMAEM
3	DP00284	YGVEVKTPARSLSAMAATVAKSSDPADPPIPNDSRILFKEPVSSYKGEYPY
4	DP00088	MFGKLSLDAVPFHEPVMVTIAGIILGGLALVGLITYFGKW TYLWKEWLTS
5	DP00262	SDLACRLLGQSMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN
6	DP00974	MDSSAGPGNISDCSDPLAPASCSPAPGSWNLNLSHVDGNQSDPCGPNRTGLG
7	DP00808	MEDINFASLAPRHGSRPFMGNWQDIGTSNMSGGAFSWGSLWSGIKNFGSTV
8	DP00488	MEAHNQFLKTFQKERHDMKEAEKDEILLMENSRRFVFMPIKYHEIWAAYKK
9	DP00226	EKPSYDTEADPSEGLMNVLKKIYEDGDDDMKRTINKAWVESREKQAREDETF
10	DP00282	NIEGLLHKNNWNIEDNNIKNASLVQIDASNEQSSTSSSMIDAQISNALNAQ
11	DP00623	SNHTFYPWMIAIEGCEPDPTKSKIRSDLTQYGISTDMGKRYS ELAGSLLP
12	DP00807	PPGATPAPPGAAPPPAAGGSPSPADGGSPPPADGGSPPPV DGGSPPPSTH
13	DP00875	KVGS LQYLALALITPKKIKPPLPSVTKLTEDRW NKPQKTGHRGSH TMNGH

Table S1

14	DP00793	ASVPAGGAVAVSAAPGSAAPAAGSAPAAAEKKDEKKEESEESDDDMGFLFD
15	DP00298	SDVEENRTEAPEGTESEMETPSAINGNPSWHLADSPAVNGATGHSSSLDAREV
16	DP00624	GKQRMKEKEQENERKVAQLAEENERLKQEIERTREVEATRRLIDRMVNLHQ
17	DP01080	CQSGSTCVKQNDWYSQCLPGSQVTTTSTTSSSSSTSRATSTRTGGVTSIT
18	DP01048	NRESARRSRYRKAHLKELEDQVEQLKAENSCLLRRLAALNQKYNEANVDNRV
19	DP00061	VPYNEGLGQPQVAPPAPAASPAASSRPQPQNGSSGMGSTVSKAYGASKTFGKA
20	DP00224	MHHQQRMAALGTDKELSDLDFSAMFSPPVSSGKNGPTSLASGHFTGSNVEDR
21	DP00037	RARMESGQYRCLDWWFCWDTPASRDDVEEARRYLRRAAEKPAKLLYEEARTH LH
22	DP00558	MAATEGVGEAAQGGEGPQPAQPPPPHPPPPQQHKEEMAAEAGEAVASPMDDG
23	DP00835	MSQRIQPSASDQQFPGKSDGYEYTVGPKQAITSEASTTYIPRIYSELLFKR
24	DP00028	MSGSSCSQTPSRAIPATRRVVLGDGVQLPPGDYSTTPGGTLFSTTPGGTRIYD
25	DP00148	MQRGNFRNRQRMVVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQAN
26	DP00424	TYLGRSAEPVPLQLPLERLTLD CNEDCGTSGTQGVGSPQILVESPTILESGAKE
27	DP00427	MSERFPNDVDPIETRDWLQAIESVIREEGVERAQYLIDQLLA EARKGGVNVAAGT
28	DP00926	GWQAFKNDATEIIPELGEYEPPELENNTMNAENGGRPPHHPFETKDVSEYS
29	DP00092	ITSFEEAKGLDRINERMPPRRDAMPDANLNSINKALTSETNGTDSNGSNSSNIQ
30	DP00194	DEDTSIMNDEASADNETVMSVIDGDKPDHDDTHPDDEPPQPPRGGRPALRVVK
31	DP00506	KNRKAKAKPVTRGAGAGRQRGQNKERPPVPNPDIPIRKQQRDLYSGLNQRRRI
32	DP00600	QGTRAQKTRA EVTSGGKPPWRQKGTGRARSGSIKPIWRSGGVTF AARPQDHSQKV
33	DP00748	IYARVIQKRVPNAYDKTALALEVGELVKVTINVSGQWEGECNGKRGHFPFTHVRL
34	DP00048	GGKWSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAW
35	DP00224	MHHQQRMAALGTDKELSDLDFSAMFSPPVSSGKNGPTSLASGHFTGSNVEDRSSS
36	DP00175	MPQLNGGGDDL GANDELISFKDEGEQEEKSSENSAERDLADV KSSLVNESETNQ
37	DP00769	EEERPRRQFNKGGRTERDGRNKRGVGRGTRREGEEAAKEEGAAETAQGNQETPAQE
38	DP01015	ELKITETEPGVRGDTATGGTKRAKLETGAVVRVPLFLNEGEIIVDTRRGEYVSRK
39	DP00083	SDPAALKRARNTEAARRSRARKLQRMKQLEDKVEELL SKNYHLENEVARLKKLVGER
40	DP00027	VTLSDAQAKLMQPGVSDINMERVEALKTAIRNGELKMDTGKIADSLIREAQSYLQSK
41	DP00948	PTEPKDKKKKTDEAQPLPQRQKKQPTVTLPAADMDDFSRQLQNSMSGASADSTQA
42	DP00949	DKPTPEEHGPGIDIFTGPKPGEYDLEGGIWEDEDARNFYENLIDLKAFVPAILFKDN
43	DP00287	MPRRAENWDEAEVGAEEAGVEEYGPPEEDGGEESGAEEESGPEESGPEELGAEEEMEAG
44	DP01073	ASTASPAKAESSGVILFGDAFGSGASETPAPQAVSSSSASADLLAGFGGSFMAPST
45	DP00803	MASLATLAAVQPTTLKGLAGSSIAGTKLHIKPARQSFKLNNVRSGAIVAKYGDKSVYF
46	DP00111	GKLGQDQSFESVESYDSCDRLTQSWSSQSFSNLSQVRVPSYDSFDYEDYPAALPNHKPK
47	DP00750	AQSEQRDRFIPSYEETQRILKEQAEAEAAARNDPDLQAAQEENMRKIREMLGINRGGAR
48	DP00600	AQKTRA EVTSGGKPPWRQKGTGRARSGSIKPIWRSGGVTF AARPQDHSQKV NKKMYRG

Table S1

49	DP00348	PPRSISPSALQDLLRTLKSPSSPQQQQVLNILKSNPQLMAAFIKQRTAKYVANQPGMQ
50	DP00249	MFDKNADGYIDLEELKIMLQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVE
51	DP00348	PPRSISPSALQDLLRTLKSPSSPQQQQVLNILKSNPQLMAAFIKQRTAKYVANQPGMQP
52	DP01053	MPKSKELVSSSSSGSDSDSEVDKKLKRKKQVAPEKPVKKQKTGETSRALSSSKQSSSSRD
53	DP00975	YPERDSEAGNEQEERLHVQKRTDGDADYDEYEEDGTTPTDPTAPTAKPRLRGNKP
54	DP00342	MSGGDGRGHNTGAHSTSGNINGGPTGIGVSGGASDGSWSENPNWGGGSGSGIHWGGGSG
55	DP01052	PAPPVLPSPYIDPAKKSFAARMMEKMGWKHGEGLGVNKQGITAPLVAKKTAMSSGVIVQGAE
56	DP00448	SDEIMDLLVQSVTKSSPRALAAREKRSRGNRSLRRTLKSGLDLVQALGLSKGPGLEV
57	DP00719	QEPVQEAPEKGRKRKPRTEPKQVPEPKPVESKKSGKSAKSKEKQKITDTFKVKRKVDRF
58	DP00672	RRSRRVKANDRERNRMHNLNAALDALRSVLPSPDDTKLTKIETLRFAYNYIWALAETLRL
59	DP00824	APQYAPGDEPSYDEDTDDSDKLVENTSITDEYAAIEASLSETFNATAADPGRRLGEGSKP
60	DP00123	ALIVLGGVAGLLFIGLGIFFCVRCRHRRRQAERMSQIKRLLSEKTCQCPHRFQKTCSP
61	DP00977	YPERDSEAGNEQEERLHVQKRTDGDADYDEYEEDGTTPTDPTAPTAKPRLRGNKP
62	DP00835	FRASVSPSSLPRASAFLSQNESSKLSNASNSPGLANSSTATSASANERLQEKQTESLSN
63	DP00014	TVTTHSSDSEVPSRVTEVVVKLFDSDPITVLPPEVSKDNPKFMDTVAEKALQEYRRKSRME
64	DP00278	SPSKDCGSPKYAYFNGCSSPTAPLSPMSPGYKLVTDNRNNSCRNYNKQASEQNWANYSAE
65	DP00389	DALAYGSQGNLNP LINEISKIISAAGNFVKEERAAASLLQLSGNASDFS YGRNSITLTASA
66	DP00592	MSERQGAGATNGKDKTSGENDGQKKVQEEFDIDMDAPETERAAVAIQSQFRKFQKKKAGSQS
67	DP00057	MARYRRSRTSRSPRSRRRRRRSGRRRRSPRRRRRYGSARRSRRSVGGRRRRRYGSRRRRRRRY
68	DP00090	MLAARTGAAGSQISEENTKLRRQSGFSVAGKDKSPKKASENAKDSLSPSGESQLRARQLAL
69	DP00812	MIGKNIKSLRKTDLTQLEFARIVGISRNSLSRYENGTSSTSELIDICQKFNVSYVDIVGE
70	DP00543	MPGLVDSNPAPPESQEKKPLKPCCAPETKAKARDACIIEKGEEHCGLIEAHKECMRALGFKI
71	DP00152	QEREKDIQGEGETAICSEMTADRPVYDESTDDEESENQTDLSGLASAAVSKDDSIISLVDTD
72	DP00220	RYLHTALEGMANPEDPECESEGWLLEKSVPETWKAFLESVKLGKGNQVEAEGEDAGQAPAA
73	DP00718	SPPYYSEKTLQYNKPHEEPSNSLMAIECRVC GDKASGFHYGVHACEGCKGFFRRTIRLKIYD
74	DP00709	MPFFGNTFSPKKTTPRKSASLSNLHSLDRSTREVELGLEYGSPMTMNLAGQSLKFENGQWIAET
75	DP01084	LGCPEKMGDPQPLGPRSAEPQQNPNLGSTGFYGVKSEPTQDTKPQFPRQMPSRNASGGQGSST
76	DP00378	PNLQONQQVLTGLPVMPIQYQVIPQFQTVVDGQQLQFAATGAQVQDGSQGIIPGANQQII
77	DP00064	MSGLFHHRKPREIRAFVMAATRLTKKQLAQAIQNTLPNPRRRKRAKRRAAQVPKPTQAGVSMA
78	DP00877	MAQEQTGRGGGGDDDDIAGSTAAGQERREKLTEETDDLLDEIDDVLENAEDFVRAVYVQKGGQ
79	DP00462	MLSVRTPLATIADQQQLSPLKRLTLADKENTPPTLSSTRVLASKAARRIFQDSAELESKAPT
80	DP00656	MTAAQAKTTKNTAAAAQEAAGAAQPSGLGLDSIGDLSSLLDAPAAQQGSGSPIELDLDLDED
81	DP01038	MADNKSQSFQAGQAAGRAEEKGNVMDKVKDAATAAGASAGTAGQKITEAAGGAVNLVKEKTMGNK
82	DP00026	MAQVINTNSLSLLTQNNLNKQSALGTAIERLSSGLRINSKDDAAGQAIANRFTANIKGLTQAS
83	DP01011	VEVDCSRFPNTTNEEGKDLVCTEDLRPICGTDGVTHSECLLCAYNIEYGTNISKEHDGECREAV

Table S1

84	DP00008	MLWQKSTAPEQAPAPPRPYQGVVRVKEPVKELLRRKRGHTSVGAAGPPTAVVLPHQPLATYSTVGP
85	DP00678	TTQITKTVKGGISETRIEKRIVITGDADIDHDQVLVQAIAKEQHDPMSVTKVWHQETEIAD
86	DP00722	LGGRQGGGAPAGGNIGGGQPQGGWGQPPQPGGNQFSGGAQSRPQQSAPAAPSNPPMDFDDIP
87	DP01034	MSTQQQARALMMRHHQFIKNRQQSMLSRAAAEIGVEAEKDFWTTVQGKPQSSFRTTYDRSNASLS
88	DP00018	KPSACRNLFGPVDHEELTRDLEKHCRDMEEASQRKWNFDQFNHKPLEGKYEWQEEVEKGSLEFYR
89	DP00468	MFQAAERPQEWAMEGPRDGLKKERLLDRHDSGLDSMKDEEYEQMVKELQEIRLEPQEVPRGSEPW
90	DP00948	QASSRSSSRGNSRNSTPGSSRGNSPARMASGGGETALALLLDRLNQLESKVSQKGGQQGGQTV
91	DP01054	MMCGAPSATQPATAETQHIADQVRSQLEEKENKFFPVFKAVSFKSQVVAGTNYFIKVHVGEDEFVHL
92	DP00157	PVQATVHGVDDNKQSQTAIEIVNVKPLASQLAGVENVKDTDLQSDTTVTGNKAGTTDNDNSQTDKT
93	DP00978	TLNCTDLRNATNTSSSWETMEKGEIKNCSFNITTSIRDKVQKEYALFYNLDVVIDNASYRLISCN
94	DP00788	VLDGNKNQFEPISIRSNFQGKHTKFENDELAESTTTKIIDSTDHIRKASSKSKRLDKVGKKGGKK
95	DP01105	ENREIQPPFKPKVCGKAENFDKFFTRGQPVLTTPDQLVIANIDQSDFEFSYVNPQFVHPILQSAV
96	DP00081	EGSIGNGLDLTGILFGNIDSEGRLLQDDDGEGRGGTGFDALRENIQSLSKLGLDSMLLEVIDLKEA
97	DP00747	EVVPPQVLSEPNEEAGAALSPLPEVPPWTGEVSPAQRDGGALGRGPWDSSDRSALLKSKLRALLTAPR
98	DP00179	MNTDQKVSEIFQSSKEKLQGDAAKVSDAFKKMASQDKDGTDADESEKHNYQEYNKLGAGHKKE
99	DP00328	DPNMIPVTSANAATGEITPAQGATGSSGGNTSGVLSTPKAKRAKHPGTEKPRSRSEQPATCPIC
100	DP00753	EVDSTEDIKAFEIFDKEKNGYISASELKHVLTTLGEKLEQEVDLLKEIGVEEGLINVDVFKLITS
101	DP00834	GTKLGSIGSDLGASIKGFKKMSDDPKQDKTSQDADFTAKTIADKQADTNQEQAKTEDAKRHDEQV
102	DP01002	MKKFLPGLLLMGLVACSSNQRMDSYSQPEQSQDLAPKSSTTQFQPQLLSKASSMPLNLLSSSKNGQVSP
103	DP00976	VSLKCTDLKNDNTNSSSGRMIMEKEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSC
104	DP00345	GNGKITQDELSKVVDNYELLKHSKNVTNSLDKLISSVSAFTSSNDSRNVLVAPTSMLDQSLSSLQFARAA
105	DP00539	PQAIPKKRGRKPGSVAAAAAEAKKAVKKESSIRSVQETVLPKIKRKTRETVSIEVKEVVKPLLVLSTLGE
106	DP00042	MPKRKSATKGDEPARRSARLSARPVPKPAKPKKAAAPKAVKGKKAENGDAKAEAKVQAAGDGAGNAK
107	DP00048	MGGKWSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEAQEEEEVGFPV
108	DP00527	MKSTSLFLCSLLLLVTGAIGRKTKEKYSQSEEVVSESFASGPSSGSSDDELVRDKPYGPKVSGGSFGEE
109	DP00752	MQVQVKTLEGETKIYLEQGTSLDLKQISQDMGFEIDMMTLVNGFIAPNTELVTDVITYLSLKLGG
110	DP00333	DAKKPEDWDEEMDGEWEPPIQNPEYKGEWKPRQIDNPDYKGTWIHPEIDNPEYSPDPSIYAYDNFVGLGL
111	DP00017	TSACRSLFGPVDHEELSRELQARLAELNAEDQNRWDYDFQQDMPLRGPGRQLQWTEVDSDSVPFAYRETVQV
112	DP00771	MSDYEEAFNDGNEFDFDVEHFSDEETYEKPKQKDGETTANGKIVTGGNGPEDFQQHEQIRRKTLKE
113	DP00841	LNLDISPQRPRPTLQLPLANDGGSRSPPSESSPQHPTTPARPRHMLGLPSTLFTPRSMESIEIDQKLQEIM
114	DP00719	AAYGGAYGENPCSPCEGFSNGLIESVELRGESAFSGIPNGQWMTQSFDTQIPFSNHCQTQEQUEESH
115	DP00636	MAGPQRATSGLPTRRTTRTQPTRRAGSAIPERQTSTASPAVSTKTAAISRTTLKSPGEPASVLAKRERD
116	DP00343	EQVSHGTQNRPLLRNSLDDL VGPPSNLEGGQSDERALLDQLHTLLSNTDATGLEIDRALGIPELVNQGQAL
117	DP00608	GDGGAFFIEHVAQYPLDMGRKKKMSNALAIQVDSEGIKYDAIARQGQSKDKVIYSKYTDLVPKEVMNADD
118	DP00669	GDIIPDPSKPTPGMADELPTETDYLPPEIYTTTFLPRTIYPQEEMPYDDKPFPSLLSKANDLNAVFEGPA

Table S1

119	DP00862	AVQMQSEKSASDSGEVATEKAPAKQDTKEKSGTETEKGEDGKTGKTQDSSADKETSAAEASEKGTVVTTETAD
120	DP00702	PKSHSRQASTDAGTAGALTPQHVRHSSPASLQLGAVSPGTLTPTGVVSGPAATPTAQHLRQSSFEIPDDV
121	DP00435	FEEAMRFARRSVSDNDIRKYEMFAQTLQQSRGFGSFRFPSSGNQGGAGPSQGSGGGTGGSVYTEDNDDDLYG
122	DP00094	MPTSPAPSPQPKGFRRRAVSEQDAKQAEAVTSPRFIGRRQSLIEDARKERAAAAAAAASSEPGNPLEA
123	DP00768	KISEGWRDEPTDAALGPEPPGPALEIKPMLQAQESPSLPAPGPPEEEEEPIEPQPTIQASSLPPPQDSARL
124	DP00853	MGMLELVGEYWEQLKITVVPVAAAEDDDNEQHEEAAEGEEKEEENGDEDEDEDEDEDDDDDEDEEEEE
125	DP00863	TLREKLMNFLVRHMNEKRELTPVTLEELSRLFGCARPALSRVFQELEREGYIEKHGRRIVLKNPFEHDI
126	DP00055	HHSQSGMDQRQKLQHWIHSCLRKADKNKDNKMNFKELKDFLKELNIQVDDGYARKIFRECDHSQTDLSLEDEEIE
127	DP00806	VTADVVSQAALTDEQKARLADTLAAKSGKTVKLNARVDESIGGMIVKLGSMIDSSIRSKLASLQNAMKEVG
128	DP00159	FIDENLKAADSDPTAPPYDLLVFDYEGSGSEAASLSSLSSESDQDQDYDLNEWGNRFKKLADMYGGGEDD
129	DP00086	MEEFQSDPSVEPPLSQETFSDLWKLLENVLSPLSQAMDMLSPDDIEQWFTEDPGDEAPRMPPEAAPV
130	DP00815	YHPADKDYGLMKIDEPNTPYHNMIGDDEADAYSSEGNEVMTDILAKLAAAEGSEPKYRTREQESSGEEDND
131	DP00086	KKPLDGEYFTLQIRGRERFEMFRELNEALEKDAQAGKEPGGSAHSSHLKSKKGQSTSRRHKLMFKEGPDSD
132	DP00707	GVDPALSDPVLARRLRENRLSAMSRLDEVISKYAMLQDKSEEGERKKRRARLQGTSSHSADTPEASLDSGEGPS
133	DP00146	MARYFRRRKFCRFTAEGVQEIDYKDIATLKNYITESGKIVPSRITGTRAKYQRQLARAIKRARYLSLLPYTDRHQ
134	DP00490	QSVSPMRSVSENSLVAMDVSGQKTRVIDNPTEALSVAVEEGLAWRKGCRLGNHGSPTAPSQSSAVNMALHRSQ
135	DP00539	MVAGMLGLREEKSEDQLQGLKDKPLKFKVKVKKDKKEEKGKHEPVQPSAHHSAEPAEAGKAETSESGSGSAPAVP
136	DP00919	MGGKWSKSSIVGWPPIRERLRAQPAAEERRRAQPAEAGVGAASQDLGRHGALTSNTDITNADEEEVEGFPVRPQ
137	DP00201	NKELDPVQKLFVDKIREYRTKRQTSGGPVDAGPEYQDLDRELKFKLQMYGKADMNTFPNFTFEDPKFEVVEKPKS
138	DP00016	RQNPCGSKACRRFLGPDVSEQLSRDCDALMAGCIEARERWNFDVFTETPLEGDFAWERVRGLGLPKLYLPTGPRR
139	DP00932	MKTIEQKIEQCRKWQKAARERAIAQREKLADPVWRESQYQKMRDLDRIIAKQKERPPASKTRKSAVKIKSRGLK
140	DP01086	LEVVFALDSEIKCEAEKQEEKPAILLSPKEESVLSKPTNAPPLPPVVLKPKQEVKSASQIVNEQRGNAAPAARL
141	DP00015	MTDVETTYADFIASGRTGRRNAIHDLVSSASGNSNELALKLAGLDINKTEGEEDAQRSSTEQSGEAQGEAAKSES
142	DP00970	MDIRKIKLIELVEESGISELISEGEESVRISRAAPASFPVMQYAYAPMMQQAQSNAAAPATVPSMEAPAAE
143	DP00075	PPKTFEKSMMNLQTKIDAKKEQLADARRDLKSAKADAKVMKDAKTKKVVESKKKAVQRLLEEQLMKLEVQATDREENK
144	DP00551	MDPQTAPSRALLLLFLHLAFLGGRSHPLGSPGSASDLETSGLQEQRNHLQGLSELQVEQTSLEPLQESPRPTGVW
145	DP00990	MSFTIPTNLKPLATPKPHLSSSSFAPRSKIVCQENDQQQPKKLELAKVGANAAAALSSVLLSSWSVAPDAAMA
146	DP00797	LPIFTPAKMGATKEEREDTPIQLQELLALETALGGQCVDRQEVAEITKQLPPVVPVSKPGALRRSLSRMSQEAQRG
147	DP00328	AKPGVKKEKTTSGKSSSSGSGSGSGALSSSGSVPPQVTVQSLHTLQGVQVKKDPDAQQQQQQQQQQQQQQAMT
148	DP00469	MQEATSSSLAASSLPSSESSSSSAPHLEIKEGIESDEEIRRVPEFGGEAVGKETSGREGSATGQERTQATVGESQ
149	DP00534	AAPEGGISDVVEKSIKAEQETCAGDPVSGECVAAWDEVEELSAASHARDKKKADGSDPLEEYCKDNPETNECRTYDN
150	DP00223	MALRVTRNSKINAENKAKINMAGAKRVPTAPAAATSKPLRPRTALGDIGNKVSEQLQAKMPMKKEAKPSATGKVIDKK
151	DP00943	DEEDGEPYPESGCVIEPGLSEEDDPAPSRKIHFTAPIQVFSTYSNEDYDRRNEDVDPMAASAEYELEKRVLEL
152	DP00447	DGIRDAMVGLREEMIEKIRTEALMTNDRLEAMARLRNEESEKMAKDTSDVSLNPTSEKLNLLLEGNDSNDLSLED
153	DP00087	TAPITDVSGLDELRLDGEEDVMTPADALDDFDLEMLGDVESPSGMTHDPVSYGALDVDDFEFEQMFTDAMGIDDFG

Table S1

[illegible]

Table S1

189	DP01074	MPPIATRRGQYEPKVQAKLSPDTIPLNPADKTKDPLARADSLHHVSDSQEDDKAAEEPPLSRKRWQNRTRFRKGRRQAPYKHK
190	DP00499	LTAATHWASMDPAVVHPPELNGAAYSRYPPGVSVAPTGIIPAIVEGIVPSAMLSHGLPPVAHPHPAPSPGQTVKPEADRDHTDQDL
191	DP00768	GTPQKTKQEEEGTEPEPELEPKPETAPEETELEEVLPPPEPCVGVKEVAAVTLGPQGTQETALTPPTSLQAQVSVAPAEHSSPRGW
192	DP00032	ADPMVDNRITKKRYRLDHLLEQPEEVLKRLYRAMEVEEENLHALVQAMMQAPERAGGAMTAAGVLDPASGEPGDWTPPRYFNMMF
193	DP00929	MEPVDPRLEPWKHPGSQPKTACTNICYKKCCFHCQVCFITKALGISYGRKKRRQRRRAHQNSQTHQASLSKQPTSQPRGDPGTGPKE
194	DP00869	AQPKSGFSTDRPLWYPGAKAPEYLDGSLVGDYGFDPFLGKPAEYLQYDYGDLQNLAKNLAGDIIGTRTESADVKSTSLQPYSEVF
195	DP00423	TVTAKKASSPSSLTYKEMILKSMPQLNDGKGSSRIVLKYYVKDTFSSKLKTSNFDYLFNSAIKKCVENGELVQPKGPSGIILNKK
196	DP00347	MNLEPPKAEIRSATRVMGGPVTPRKGPVKQQRTRQFKSKPPKGVQGFDDIPGMEGLGTDITVICPWEAFNHLELHELAQYGII
197	DP00157	TGETTSFDEVGDVTSKSIIGKPVQATVHGVDDNKQSQTAIEIVNVKPLASQLAGVENVKDTLQSDTTVITGNKAGTTDNDNSQTDKT
198	DP00423	KASSPSSLTYKEMILKSMPQLNDGKGSSRIVLKYYVKDTFSSKLKTSNFDYLFNSAIKKCVENGELVQPKGPSGIILNKKVKLST
199	DP00260	MPLNVSFNTRNYDLDYDSVQPYFYCDEEENFYQQQQQSELQPPAPSEDIWKKFELLTPPLSPSRRSGLCSPSYVAVTFSLRGDNDG
200	DP01085	ETIGNPTIFGETDTEAQKTFSGTGNIPPNRVVFNEPMVQHSVNRAPPRGVNIQNQANNTPSFRPSVQPSYQPPASYNRHGPIMKNEA
201	DP01014	MSTDVAAAQAQSKIDLTKKKNEEINKKSLDEDEFEFPIDTWANGETIKSNAVQTQNIWEENWDDVEVDDFTNELKAELDRYKRENG
202	DP00461	GSNGADNAHNNAFGGGKNPGIGNTSAGSNGSASSNRGNSNGWSWSNPKHNDGFHSDGSYHITFHGDNNSKPKPGNSGNGRNGNDGA
203	DP00136	AKRKASGPPVSELITKAVAASKERSGVSLAALKKALAAAGYDVEKNNSRIKGLKSLVSKGTLVQTKGTGASGSFKLNKKAASGEAKPK
204	DP00195	MPKRKAEGDAEGDKAKVKDEPQRRSARLSAKPAPPKPEPKPKKAPAKKGEKVPKGGKKGADAGKDGNNPAENGDAKTDQAQKAEGAGDA
205	DP01088	MAKSKEDITYATSQARLSEDAVRVAYEHGSPLEGGKIADSQPDLFSSAHNMPKSGQTTMDSNTSDQSQMORDTQEGGSKEFTTGAPG
206	DP00985	EKNGTLMIDDERILEHLTDEEKSRIVKKVRFRTLGCYPLTGAVESEATSLDIIQEMLLTRTSEQRVIDHDGAGSMEEKRQGYF
207	DP00141	ALAPDSPDNDLRAQQFGISARKPFTLGEVAPVWVPDSQAPNCMKCEARFTTKRRHHCRACGKVFCASSCSLKCKLLYMDRKEARVCV
208	DP00461	MGSNGADNAHNNAFGGGKNPGIGNTSAGSNGSASSNRGNSNGWSWSNPKHNDGFHSDGSYHITFHGDNNSKPKPGNSGNGRNGNDGA
209	DP00997	QDEADELTRRFVAlMDAGEPADSEGAMDAADHRQGIARNHYDCGYEMHTCLGEMYVSDERFTRNIDAAKPLAAYMRDAILANAVRHPT
210	DP00039	MPKRKAEGDAKGDKAKVKDEPQRRSARLSAKPAPPKPEPKPKKAPAKKGEKVPKGGKKGADAGKEGNNPAENGDAKTDQAQKAEGAGDAK
211	DP00191	CLKGTGENYRGNVAVTVSGHTCQHWSAQTPHHTNRTPENFPCKNLDENYCRNPDGKRAPWCHTTNSQVRWEYCKIPSCDSSPVSTEQ LAP
212	DP00631	MTPSTPPRSRGTRYLAQPSGNTSSALMQGQKTPQKPSQNLVPVTPSTTKSFKNAPLLAPPNSNMGMTSPFNGLTSPQRSPPFKSSVKRT
213	DP00864	NKAVQNDSPNSASQPTTGPSAASPASENQNGNLSAPPGGGGPHPHPTPSPHPPSTRITRSQPNHTPAGPPGPSSNPVSNGKETRRSSKR
214	DP00486	PEKVGPKQAEDAAKSITNGSDDGAQPSTSTAQEQDDVLIVDSDEEDSSNNADVSEERSRKRKLDEKENLSAKRSRIEQKEELDDVIALD
215	DP00353	MMNNNGNQVSNLSNALRQVNIIGNRNSNTTDDQSGNINFEFSTGVNNNNNNSSNNNNVQNNNSGRNGSQNNNDNENNIKNTLEQHRQQQQA
216	DP00675	DTKEEGATKKKTQKPDRLERGMKIVPKESEKSKTKPPDATIVVEGVKYQVRKKGKTKSKNTQDGLYHNKNKPQESRKKLEKALLAWAI
217	DP01042	YANEPKVCFYMDDDYNGESLCAAQNSVASIPDKWNRDISSISIPHGLVTVYEDVDFLGASRSFEADVLDVSDKDLIYLDNISAFKIKKA
218	DP00902	PTSIHETVYENDGDPITLLDQIADADEASWFDKIALKKAIEELDERERLIVLYRYRDRQTQSEVASRLGISQVQVSRLEKKILQHIKDKMDG
219	DP01033	MSYRKELEKYRDLDEDKILGALTEELRKLENELELDPDNALLPAGLRQRDQTKQPPTGPFKREELMAHLEQQAKDIKDREDLVFPTGEKR
220	DP00086	MEEPQSDPSVEPPLSQETFSDLWKLPPENNVLSPLSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAAPVAPAPAPAPAPAPSWPL
221	DP00808	GLPTTRPIAPMATGVLGQHTPVTLDPADTQKQKVLPGTAVVVTSPRASLRRAASGRSLRPVASGNWQSTLNSIVGLGVQSLKRRRCF
222	DP00113	MSSSTPFDYPALSEHDEERPQNVQSKSRTAELQAEIDDTVGIMRDNINKVAERGERLTSIEDKADNLAVSAQGFKRGANRVRKAMWYKDLKMK
223	DP00944	MADNEEDVEAEETELTDISGVGSKAESLREAGFESVEDVRGADQSALADVSGIGNALAARIKADVGGLEVESETEAEVEEEGGEEAPDEDV

Table S1

224	DP00016	MSEPADVDVRQNPCGSKACRRLFGPVDSEQLSRDCDALMAGCIQEARERWNFDVFTETPLEGDFAWERVRGLGLPKLYLPTGPRRRGDELGGRR
225	DP00018	ESQDVSGSRPAAPLIGAPANSEDLVDPKTDPSDSQTLAEQCAGIRKRPATDDSSQNKANRTEENVSDGSPNAGSVEQTPKKPGLRRRQT
226	DP00615	SPDADLIEANLLWRQEMGNITRVESENKVILDSFEPLQAEEDEREVSVPAILRRSRKFPRAMPIWARPDYNPPLLESWKDPDYVPPVHGC
227	DP01031	NTYDVMVDANLLMEGGVAQTEPESRVPLDFLEPMAEESLEPSIPSECMPLPRSGFPRALPAWARPDYNPPLVESWRRRPDYQPPTVAGCALP
228	DP00092	SDDELGSEEDGFDGATAARKEVIRNKIRAIIGKMARVSVLREESESVLTLKGLTPTGMLPSGVLGGKQTLQSATVEAIEADEAIKGFSPQHK
229	DP00984	ELSPSLVISLSTGLSLFLGRVFFNFQRENMAKQVPEQNGMSHFEAGDTRAKEYVSLKSNDPVGFNIVDLAWGSIGHIVAYILATASNGYDP
230	DP00984	PSLVISLSTGLSLFLGRVFFNFQRENMAKQVPEQNGMSHFEAGDTRAKEYVSLKSNDPVGFNIVDLAWGSIGHIVAYILATASNGYDPSFF
231	DP00092	DDELGSEEDGFDGATAARKEVIRNKIRAIIGKMARVSVLREESESVLTLKGLTPTGMLPSGVLGGKQTLQSATVEAIEADEAIKGFSPQHKIT
232	DP00552	YNGQSETSQQPETPVFNTLPMMGKASPVSLGVPSEATANNGQQQVQEQRRRINAMLQDYELQRLHSEQLQFEQAQTQQAQVQVPGIQLGTQSQ
233	DP00615	ATCTTRHSDPADLIEANLLWRQEMGNITRVESENKVILDSFEPLQAEEDEREVSVPAILRRSRKFPRAMPIWARPDYNPPLLESWKDPDYVP
234	DP00517	GGGVSPDVCALSDPEGGLTASGRVHEASPTTFQKDALLGSKPNKPSLPSSSQNLGQTEVSKVSETVQEELTPPPQKAAPQGKSKSDPLKKKTDTR
235	DP00092	SDDELGSEEDGFDGATAARKEVIRNKIRAIIGKMARVSVLREESESVLTLKGLTPTGMLPSGVLGGKQTLQSATVEAIEADEAIKGFSPQHKIT
236	DP00491	VVEWTNNTLPLEPCPKEQIIGLSRRIADENGKPRPASFNALEPDFKALSIAVTGSGSCRHLQFIPVAPPSGSSAAPATEVPDRDPEKSSDDVYLH
237	DP00615	RHSDPADLIEANLLWRQEMGNITRVESENKVILDSFEPLQAEEDEREVSVPAILRRSRKFPRAMPIWARPDYNPPLLESWKDPDYVPPVHGC
238	DP00994	ERIETAEKEAYPATKVKVNVKSKQLKTKLEYFDATFVESNTRKNFGRNNSRNFFNRRGGRGARKGNNTANATNSANTVQKNRNDVSNLPSLA
239	DP00082	LEAEIETTDDTDDGNTVSHILNVLKDATPIEDVFSFNYPEGIEGPDIKYKKEHVKYTYGPTFLLQFKDLNVKADAQEWVQSTASKVIPPGMGRGN
240	DP00097	AKGDEPKRSVAFKTKKEVKKVATPKKAAKPKKAAKSPKPKATPVKKAKKKPAATPKKAKKPKVVKVKKPKASKPKKAKTVKPKAKSSAKRASKK
241	DP00136	AKKAGAAKAKKPAGAAKPKKATGTATPKKSTKTKPKAKKPAAGAKKAKSPKAKATKAKKAPKSPAKARAVKPKAAKPKTSKPKAAKPKKTAAK
242	DP00265	KKRKPKGGWNTGGSRYPGGGSPGGNRYPPQGGTWGQPHGGGWGQPHGGSWGQPHGGGWGQPHGGGWGQGGGTHNQWNPSPKPKTNLKHVAGAAAAGAV
243	DP00968	MMDLELPPGLQSQQDMDLIDLWRQDIDLGVSRVDFDSQRQKDYELEKQKKLEKERQEQLQKEQEKAFQAQFQLDEETGEFLPIQPAQHIQTDTSQ
244	DP00024	MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAQGAEPDRAHYNIVTFCKCDSTLRLCVQSTHVDIRTLLEDLLMGTLGIVCPICSQKP
245	DP00720	GGDSNEKGKSSKKRKTETSPSKKANTSGSGFKSKEYISDDSTSSDDEKDNPAKKKSKPPSDGDAKKKKAKSESEPEESEEDSNASDEDEDEASD
246	DP00783	LCKKRPKPGGGWNTGGSRYPGGGSPGGNRYPPQGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGTHGQWNPSPKPKTNMK
247	DP01057	SQRKKSIIYAAKGDGNILDDLNEATKKASDFVTDKTKALADGEKADYVVEKNSETADTLGKEAKAAAYVEEKGEAKAANKAAEFAEGKAGEAKDATK
248	DP01058	VAQRKKSIIYAVKSDGNILDDLNEATKKASDFVTDKTKALADGEKTDYIVIEKTIEANETATEAKKALDYVTEKGEAGNKAEEFVEGKAEEAKNAT
249	DP00953	MAESPTTEAATAGAGAAGPGASSVAGVVGVSOGGGGFPFPLPDVWAAAAAGGAGGPGSLAPLPLPSSAAAHAALLSHWDPTLSSDWGDERTAPQ
250	DP00161	PSPQHQQPPYASAPRQPVQPPAQVPPQHAPHAPQVQPPAYQPPQPEQLQQPVSPQVAPAPQPVHSAPQPAQQAQFQPAEPVAAQPEPVAEPAPVM
251	DP00876	MNNQRKKARNTPFNMLKRERNVSTVQQLTKRFSGLMLQGRGPLKLFMALVAFRLFTIPTTAGILKRWGTIKKSKAINVLRGFRKEIGRMILNLRNRRR
252	DP00699	MDKLELVNDGLNIIDFIQKNQKEIQKTYGRSSIQQPSIKDQTKAWEDFLQCTSGESEQVEGGMSKDDGDDVERRNLEDLSSTSPDTGTIGKRVSNTRDWAE
253	DP01069	MKRLVLVILLSLTLGFSFAGTASADENANRPVNPGEADPEAFRSTPFEATTALGDCRECIAYRTGATTGKGSRRHDDTVSREIKGSSATPGGSEKAGTGRO
254	DP00053	MLEEALAAIQNARDLEELKALKARYLGKGLLTQEMKGLSALPLEERRKRQELNAIKAALEAALEAREKALEEALKEALERERVDVSLPGASLFSGGLH
255	DP01003	MEPVDNPLEPNWHPGSQPKTACNNCYCKRCSYHCLVCFQTKGLGISYGRKKRRQRSSAPPSSEDHQNPPIPKQLPQTRGDQTGSEESKKVKESKETDPPD
256	DP00902	MDVDVKQGQSPIKDQEMKELIRRSQEGDQEARDIEIKNMRLVWSVQRFLNRGYEADDLFQIGCIGLLKSVDKFDLSYDVVKFSTYAVPMIIEIQRFLRD
257	DP00458	LHKDVEEAGALQKNHASVTSQAQSAEAAEAETSCLPTAVEPLSMGSCIEPAEQSQCPGPESSPEVNDAEATGKKENSDAKSSTEEPSENVPKVQDNTETPK
258	DP00070	MDVFMKGLSKAKEGVVAAAEKTKQGVAAAGKTEGVLVYGSKTEGVVHVGVATVAEKTKEQVTNVGGAWVTGVTAVAQKTVEGAGSIAAATGFVKKDLQGLK

Table S1

259	DP00675	SDTKEEGATKKKTQKPDRLERGMKIVPKESEKSKTKPPDATIVEGVGYQVRKKGKTKSKNTQDGLYHNKNKPQESRKKLEKALLAWAIIAIVLFQVTMG
260	DP00466	KKRKPQGGWNTGGSRYPGQSGPGGNRYPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGGTHSQWNKPSKPKTNMKHMAGAAAAAGAVVGG
261	DP00895	MEKFAPEFHGEDANNRATKFLESIKGFTSPKDPKKKDSIISVNSIDIEVTKESPITSNSTIINPTNETDDTAGNKNPYQRKPLVSFKEDPTPSDNPFKSLY
262	DP00864	EPNPEYSTQAPNAKAVQNDNSPASQPTTGPSAASPASENQNGNLSAPPQGGGPHPPHTPSHPPSTRITRSQPNHTPAGPPGPSSNPVSNGETRRSSKR
263	DP00576	RNISQDMTQTSNTLTSEELRKRREAYFEKQQKQQQQQQQDGLSGQSSHPCEPATSSGALGSDLGKACSPFIMFATFTLYLTYELHVIFALHYSSFPL
264	DP00724	QDKKTMMEEDDGNMDELLAVLGYKVRSEMADVAQKLEQLEVMMNSVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSSNAEYDLKAIPGDAILNQ
265	DP00079	MDIKNSPSSLNSPSSYNCSQSILPLEHGSYIPSSYVDSHHEYAMFTYSPAVMNYISIPSNVTNLEGGPGRQTSPNVLWPTPGHLSPLVHRQLSHLYAEPQ
266	DP00718	PIDPESDSFTDLSANISQEMTMVDTEMPFWPTNFGISSVDLSVMEDHSHSFDIKPFTTVDSSISTPHYEDIPFTRTDPVADYKYLKLQEQSAIKVEPA
267	DP01055	MSSLPFVFGAAASSRVVTAAGKGTAEKQESFVDWLLGKITKEDQFYETDPILRGGDVKSSGTSKGKGGTSGKKGTVSIPSKKKNGNGGVFGGLFAKKD
268	DP00473	VNHAETSRNLNIERMKQIYQQLSRYRGLCLLASVPSCVEDLSHPVAEPSSSQAGSMSSAGPRPLPSGPASPKRLEAAEPPGEELSKRARVAELPTPELPSKDA
269	DP00006	MGDVEKGKKIFVQKCAQCHTVEKGGKHTGPNLHGLFGRKTGQAPGFTYTDANKNKGITWKEETLMEYLENPKYIPGTMIFAGIKKTEREDLIAYLKATN
270	DP00673	SKKPGGPGKNRAVNMMLKRGMPRGLSLIGLKRAMLIDGKGPIRFVALLAFFRFTAIAPTRAVLDRWGVNKTAMKHLSSFKELGTLTAINRRSTKQKKR
271	DP00539	VQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGGATTSTQVMVIKRPGRKRKAADPAIPKRGKRKPGSVAAAAAEAKKAVKESSIRSVQETVLPIKKRKTRE
272	DP00646	PGPPTSLVQSLTPPPDFNQCLKNSSGEKFFDSFNNMGSRKNPDALATGEVNPQEQIPGEGFIHMHSYQKPEYASGASAGHRLPGQYHSDKRRLSKASSKARSDD
273	DP00661	MSVYFIHCSYDLRSAEMTEVMNAPSMEIEGLSPRKDLSYQIFPDPSDFDRCCCLKDRPLSIVVEPTGEVEGESGELRWPPPEFLVQEDEQDNCEETTNEKKDQ
274	DP01036	MASRQNNKQELDERARQGETVVPGGTGGKSLAQQLHAEGRSKGGQTRKEQLGTGEGYQEMGRKGGGLSTVEKSGEERAQEEGIGIDESKFRGTNNKNQNNQEDQDK
275	DP00164	MSTESALSYAAILADSEIEISSEKLLTLTNAANVPVENIADIFAKALDGNLKDLLVNFSAAGAAPAGVAGGVAGGEAGEAEAEKEEEEEAKEESDDDMGFGLFD
276	DP00151	TSHVQEEQIEVEETIEAAKAEAEKDEPPSEGEAEFEEGKEKEEAEAEAEFEEGAQEEEEEAEKEEGEGEQQEETKEAEKEEKKDEGAGEEQATKKKD
277	DP00372	MKAKRSHQAVIMSTSLRVSPSIHGYPHFDTSRKKAVGNIFENTDQESLERLFRNSGDKKAEERAKIIFAIQDQVEEKTRALMALKKRTKDKLFQFLKLRKYSIKVH
278	DP01042	KKAVCFYGESNFTGDSLCLSSGEQFDLYRGNYPERKQSHLVNPLNDEVYSIKIPPMQTTVYEDDDYNGKYFVLTEDYTPDDLIIIRMNNKISSMRVSQDEDFICD
279	DP00873	VGQKMTTLQHRSQGDPEDPHDEHYLLATQSKQEQSAKATADRKSASKPIGFEGDLRGQSSDLPERLHSQEVDLGSAAQGDGLMAGNDSEELTALMSRKTALSLFEGK
280	DP00005	MDAQTRRRERRAEKQAQWKAANPLLGVSAKPVNRPILSNLRPKSRVESALNPIDLTVLAEYHKQIESNLQRIERKNQRTWYSKPGERGITCSGRQKIKGKSIPLI
281	DP01025	SFAGLKADADVAALAACSAAFSKHKEFFAKVGLASKSLDDVKKAFYVIDQDKSGFIEDELKFLQNFSPSARALDAETKAFLADGDKDGDGMIGVDEFAAMIKA
282	DP00550	AKDLLKADDIKKALDAVKAEGSFNKKFFALVGLKAMSANDVKKVFAIDADASGFIEEELKFVLKSFAADGRDLDAETKAFLKAADKDGDKIGIDEFETLVHEA
283	DP00084	SDNDDIEVESDEEQPRFQSAADKRAHHNALERKRRDHIKDSFHSRLDSVPSLQGEKASRAQILDKATEYIQYMRKNHHTHQDIDDLKRQNALLEQQVRALEKARSSAQ
284	DP00836	EPYTDENVPEGDSVPEAHWTQLQHS�DTALRRARSAPTAPIAARVTGQTRNITVDPRLFKKRRLHSPRVLFSTQPPPTSSDTLDLDFQAHGTIPFNTRHSKRSSTHPVF
285	DP00341	KRLSVELTSSLFRTEPMAWNETADLGLDIGAQGEALGYRQDDPSYRSFHHGGYQDQALGMDPMMHEMGHHPGADYPVDGLPDLGHAQDLMDGLPPGDSNQLAWFDTDL
286	DP00655	PSSDYCKKKKKKEKNSKEPVIEDILNLESSSTSGSPVSPGLYSKTMPTTYDYPINGTPALSTMTWFGDSPLTEQNCSVLAFSQPPSPDLVDMFQPRSLVDKAMNAGW
287	DP01077	KSNEKKDSGSSKPAFSFGAKPDEKKNDEVSKPAFSFGAKANKEKSEDESASFSGSKPTGKEEGDGAKAAISFGAKPEEQKSSDTSKPAFTFGAQKDNEKKTESSTGK
288	DP00814	MSDAAVDTSSEITTKDLKEKKEVVEEAENGREAPANGNANEENGQEADNEVEDEEEEGGEEEEEGEGDEEAEAAATGKRAAEDDEDDVDTKKQKTDEDD
289	DP00256	MMPANEDKENNIVYTGNSSGINFPQTPAHLKRSHSNILKPPVRLDQLKRDANSNNGNTLYIQGGKEVSPTKRLHTHAQQQGRPLAAKDNNRSKSFIFPETSNSQSKD
290	DP00904	EKEIDEGALYIVSNKKLFKKLKNKNPNLVVCTEGMLDIEDMRAIGVPEKALEGLKKKVEIARKNVERFIEKYKPEKIFVVEDDKDELLYRAKNLYNAEKLDADEILD
291	DP00184	FDRNVPRICGVCGRATGTFHFNAMTCGCGCKGFFRRSMKRKALFTCFPNGDCRITKDNRRHCQACRLKRCVDIGMMKEFILTDEEVQRKREMILKRKEEALKDLSLRPKLS
292	DP00088	AVVPHVHERDAFWEMKEKGEAYKPDHYEIIHMPKNSGAGIVAAFTSTIFGAMIWHIWWLAIVGFAGMIITWIVKSFDEDDVYVYVPAEIEKENQHDFEITKAGLKNGN
293	DP00706	GDSSGFSYSGSSSGSSISARSSSGGSSSGSSSSIAQGGAGSFGKPTGTYSQVSYSSGSSSLQGASGSSSLGSSSSSHSGNSGSHSGSSSSSHSSSSSFQFSSSSFQVG

Table S1

294	DP00070	MDVFMKGLSKAKEGVVAAAEKTKQGVAAEAGKTEGVLYVSGSKTEGVVHGVATVAEKTKEQVTVNGGAVVTVGTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGIL
295	DP00518	MAKQPSDVSSCEDREGGLQPAERPPQLRPGAPTSLQTEPQGNPDGEGDRCPHGSPQGFLAPPASPFPFATRSPLFIVRRSSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTP
296	DP00621	CSELYTEASGYISSLEYPRSYPPDLRCNYSIRVERGLTLHLKFLPEFDIDDHQQVHCPYDQLQIYANGKNIGEFCKGQRPDLDTSSNAVDLLFFTDSESGDSRGWKLRYTTE
297	DP00025	QNSGNQSFEEDEEDKPKYEQQGNIVDIDFDSVPQIHQGNKGNQSFEEDEKDKPKYEHGNNIIDFDSVPHIHGFNKHTEIEEDTNKDKPSYQFGGHNSVDFEEDTLPKV
298	DP00066	MNRGFFNMLGRRFPAPTAMWRPRLRRQAAPMPARNGLASQIQLTTAVSALVIGQATRPQNPRPPRPQKQAPKQPPKPKPKQPKQEKKKKQPAKTKPGKQRQRMALKLEAD
299	DP00200	RVKFSRSADAPAYQQGQNLNELNGRREEYDVLDKRRGRDPEMGGKQRRRNQPEGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALPPR
300	DP00666	KRSKGGKYSVKDKEDTQVDSEARPMKDETFGEYRSLESNDNEEKAFGSSQPSLNGDIKPLGSDDSLADYGGSDVQFNEDGSFIGQYSGKKEKEAAGGNDSSGATSPINPAVALE
301	DP00531	MEEKHHLHHHLFHHKDKAEEGPVDYEKIKHHKLEQIGKLTVAAGAYALHEKHEAKDPEHAHKHKIEEIAAAAVGAGGFATHEHHEKKDAKKEEKKLRGDTTISSKLLF
302	DP00173	MEVTGDAGVPESGEIRTLKPCLLRRNYSREQHGAASCLEDLRSKACDILADKSLTPVTLVLAEDGTIVDDDDYFLCLPSNTKFVALASNEKWAYNNSDGGTAWISQESFDVDET
303	DP00850	HHNKEWLSKPTDKRRLSELTIQIWWDLKGKHEEHGKSNLKGIEIFTKITNDPCAFOITKSLISQYCATRRSQGKIPSSINRDLTCISGMFTALIEALFFGEHPIRGTRKRLKEEKPE
304	DP00625	DLEVIISLGPDPTRLDAKLLDSYSTAATSSSKDVISVATDVSNITITKTSDERLITEGEANQGVATTATVKATESDGNVPKAMTGSIDLKKEGIFDSVGITTIDPEVLKEKPWRQPGA
305	DP00387	MKKRNRLLKNEDFOKVKFHGTSVANRQFVLYTLDPENDELRVGLSVSKKIGNAVMRNRIRLIRQAFLKEEKERLKEKDYIIARKPASQLTYEETKKSQHLFRKSSLYKSSSK
306	DP00205	MKTTLIKVIASVTLFLSMQVYASGHTAHVDEAVKHAAEEAVAHGKEGHTDQLEHAKESLTHAKAASEAGGNTHVGHGIKHLEDAIKHGEEGHVGVATKHQAQEAIEHLRASEHKSH
307	DP00134	APTEERESFLRRGDGRRRGGGGRGQGGRRGGGFGKGNDDHSDRTDNRPNRPREAKGRTTDSGLQIRVDCNNERSVHTKTLQNTSSSEGSRLRTGKDRNQKKEKPDSDVGQQPLVNGVP
308	DP00341	EDKPQDYKRLSVELTSSLFRTEPMAWNETADLGLDGAQGEALGYRQDDPSYRSFHSGGYQGDALGMDPMMEHEMGGHHPGADYPVDGLPDLGHAQDLMGDLPPGDSNQLAWFDTL
309	DP01060	MGSLEMVPMGAGPPSPGGDPDGGDGGNNSQYPSASGSSGNTPTPNDEERESNEEPPPPYEDPYWNGNDRHSDYQPLGTQDQSLYGLQHDGNDGLPPPPYSPRDDSQHIYEEAGRG
310	DP00553	YRKDKRRQEPLRQSPQRGAGAPELGAPEEELAAQLGPTHHECEAGPPHDTRLTALPDYTLTLRRSPDDIPLMTPNTITMIPNSLVGLQTLHPYNTFAAGFNSTGLPHSHSTTRV
311	DP00447	MEKFAPEFHGEDANNRATKFLSEIKGFTSPKDPKKNDSIIVNSTDIEVTKESPITSNSTINPTNETDDNAGNKPNYQRKPLVSFKEDPIPSDNPFKSLYKETIETFDNNEESSY
312	DP01068	DLNTTLPIRQTASIFKQPVTKVTNHPNSNKVKSDPQRMNEQPRQLVWEKRLQGLSASDVTEQIKTMDLPGKQGVGPGSNDETLLSAVASALHTSSAPITGQVSAAEKNPAVWLNTSQ
313	DP00537	MDRSLGWQGNVSPEDRTEAGIKRFLEDTTDDGELSFKVDFSGNASCHPPEAKTWASRPQVPEPRPQAPDLYDDLEFRPPSRPQSSDNQQYFCAPAPLSPSARPRSPWGKLDPYDSSE
314	DP00399	GFVSDSNVVKPDQTFADVLAISQRTTHNTVAVTDGTPHGVLLGLVTQRDYPIDLQTQETKVSMDMTFESKLVTAHQDTKLSEANKIWEKKLNALPIIDDDQHLRYIVFRKDYDRSQVC
315	DP00049	HVKKPYDPSRTEKRESKCPTEGCDGTGHVNGLYPHHRSLSGCPHKDRVPPEILAMHENVLKCPTEGCTGRGHVNSNRNHRSLSGCPIAAAEKLAKAQEKHQSCDVSKSNQASDRVLRP
316	DP00544	MAKSIRSKHRRQMRNVKREHFAKDLDRKLRLASKAQELDLNVVTMKSAAEIKNKPSTSASDADKGMEVDNTKKVKFKTKQNEGDGHYPQWMNQRAVKKQVKVAKLTKKKIGKKIKW
317	DP00871	MDPNTVSSFQDILLRMSKMQLESSEDLNGMITQFESLKYRDSLGEAVMRMGDLHSLQNRNEKWREQLGQKFEIIRWLIIEVVRHKLKITENSFEQITFMQALHLLLEVEQEIRTSFQLI
318	DP00839	PMPGALGASGSSGHELSALGEGGLQSLGNNMHSQMLQGLGAGLGGGLGALTGPGLASLLGSSGPPGSSSSSSSRSSQSAAVTPSSSTSTRATPAPSAPAAASATSPSPAPSSGNGA
319	DP01103	AEASLTSQAQKAKAEKKRRKMSKGLPDKSEEEKSTMAKKLEQKPKGEGIPTAKLKIDEFESNVNEVKDPYPSADFPGDDEEDEPIPLSPRPRPLAELQLKEKAVPIPEASSFFIFSPT
320	DP00020	MSSPDDFETAPAEYVDALDPSMVVVDSGSAAVTAPSDSAAEVKANQNEENTGATAAETSEKVDQTEVEKKDDDDTTEVGVTTTTPSIADTAATANIASTSGASVTEPTDDTAADEKKEQVSG
321	DP00958	MKLPVAQYAPDGVKESFAPIRDDPRYMTTEGRTTGPSDHVLNAGQIDRDKPSEPRTKDGSQLTYLQGLRQLTGLQDDINEFLTGRMELAKNKKKAGADEKRIQEEINQLLDGGDGEDAV
322	DP00629	SGGDGAYHEPTGGGAIEVALDNADIDLETEAHADQDARGWGGESGERWARQVSGGHFVTLHGAERLEEETNDEDVSDIERRIAMRLAERRQEDSATHGDEGRNNGVDHEDDDAAAVAGIGGI
323	DP00787	ATINNAGFESGFSNNWNETDPAAISSDAYSGSKSLKIQGSPARVYQVVDIQPNTETYLTSAYVLGKGQIGVNDLNGLFKNQTFNVSSWTKVTKFTSANTNSLQVFAKHYNNTSDVRFDFNSLIEG
324	DP00012	IRKFSIVQKTLQMNGIEEDSDEPLERRLSLVPDSEGEAILPRISVISTGPTLQARRRQSVLNLMTSHVNGQGNIRKTTASTRKVSLAPQANLTELDIIYSRRLSQETGLEISEEINEEDLKE
325	DP00662	LTNNVLSTAQQVYLRKLGAKPMDENASKIISAGRAKRSIAQPDADAGERFRQLKEQEKRSKKNKAVAKDTVELVEESQSESEEGSDDEEEAREGALASSTTSKPLPEVGQRSSKRSKRKRTV
326	DP00160	HTTEDKISRAGVPRQAQVSLQGDQSENELPRLGGKEDRRVKQSRGEARESYRETGPSRASDARAHLPTGTPLDIDTASESSQDPQDSRRSAEPLLSCKPWQESRKNKAQTRTPLQCTMTEIF
327	DP00629	LSGGDGAYHEPTGGGAIEVALDNADIDLETEAHADQDARGWGGESGERWARQVSGGHFVTLHGAERLEEETNDEDVSDIERRIAMRLAERRQEDSATHGDEGRNNGVDHEDDDAAAVAGIGGI
328	DP00774	QTTDNVDELFDOTQLMLFAKRLSTLDLNEINAADRMATPNRLKHGHVDDDALTFAITFDHGRMVLNDGNGEDIPYSYQREGFADGQLVGEDDPWRFFVWMTSPDGKYRIVVGQWEYREDMALA

Table S1

329	DP00085	PTQAGEGTLSEALLQLQFDDDELGALLGNSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYEAITRLVTGAQRPPDPAPAPLAPGLPGLNGLSGDEDFSSIADMDFSALLSQISS
330	DP00418	RKRRKPGSHTHSASEDNTTNVREQLNQIKNPIEKHGANTVPKIDYENKNSKMSKIRTHNSEVEEDMDKHQQKARFAKQPAYTLVDREKPPNGTPTKHPNWTKNQDNRDLESAQSLNRMEYIV
331	DP01100	QTAPVMPDLKNVKSIGSTENLKHQPGGGKVQIINKKLDLSNVQSKCGSKDNIKHVPGGGSQVIVKPVDSLKVTSCKCSLGNIIHKPGGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGGNK
332	DP00219	MAEAGAGLSETVTETTUTVTTEPENRSLTIKLRKRKPEKKVETSDTVDNEHMGRRSSKCCCIYEKPRAFGESSTESDEEEEGCGHTHCVRGHRKGRRRATLGPTPTTPPPDPSQPPPGPMQH
333	DP01103	EAESLTSQAQKAKAEKKRRKMSKGLPDKSEEEKSTMAKKLEQKPKGEGIPTTAKLKIDEFESNVNEVKDPYPSADFPDGDEEDEPEIPLSPRPRPLAELQLKEKAVPIPEASSFFIFSPTNKIRVL
334	DP00630	MDVFKKGFSIAKEGVGAVEKTKQGVTAAEKTKEGVMYVGAKTENNVQSVTSAEKTKEQANAVSEAVVSSVNTVATKTVEEAENIAVTSQVVRKEDLRPSAPQGEGEASKEKEEVAEEAQSGGD
335	DP00546	MRRRGIDMATEGDVELELETETSGPERPPEKPRKHDSGAADLERVTDYAEKEIQSSNLETAMSVIGDRRSREKQAKQEREKELAKVTIKKEDLELIMTEISRAAAERSLREHMGNVVEALIALTN
336	DP01100	QTAPVMPDLKNVKSIGSTENLKHQPGGGKVQIINKKLDLSNVQSKCGSKDNIKHVPGGGSQVIVKPVDSLKVTSCKCSLGNIIHKPGGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGGNKIE
337	DP00365	PPADGALKRAEELKTQANDYFKAQDYENAKFYSAIELNPSNAIYYGNRSLAYLRTECYGALGDATRAIELDKKYIKGYRRAASNMAKGKFRALRDYETVVKVPHDKDAKMKYQECNKIVKQKA
338	DP00025	GQNSGNQSFEEDTEEDKPKYEQGGNVIDIDFDSVPIHQGNQNGQSFEEDTEKDKPKYEHGGNIIDIDFDSVPHIHGFNKHTIEIEDTNKDKPSYQFGGHNSVDFEEDTLPKVSGQNEGQQTIEEDTTP
339	DP00062	MDTKHFLPLDFSTQVNSSLTSTPTGRGSMAAPSLHPSLGPGLGSPQLHSPISLTSSPINGMGPPFVSISSPMGPHSMVPTTPTLGFSTGSPQLSSPMNPVSSSEDIKPLGLNGVLKVPAPHPSGNMAF
340	DP00224	MHHQQRMAALGTDKELSDLLDFSAMFSPPVSSGKNGPTSLASGHFTGSNVEDRSSSGSWGNGGHPSPSRNYDGTPTYDHMTSRDLGSHDNLSPPFVNSRIQSKTERGSYSSYGRESNLQGCQHSLLGGD
341	DP00534	MATATGLNIATQRVFTVSENRPVCLAGPVHLNNSWNLGSRTTNRMMLQPIKAAPEGGISDVVEKSIKAEQETCAGDPVSGECVAAWDEVEELSAASHARDKKKADGSDPLEEYCKDNPETNECRTYDN
342	DP01104	RKRRRQHGLWFEFGFKVSEASKKRREPLGEDSVGLKPLKNASDGLMDNQNNEWGDEDELETKKFRFEFVPLDLDDQTDHRQWTQQLHDAADLRMSAMAPTPPQGEVDADCMDVNVVRGPDGFTPLMIA
343	DP00341	MATQADLMELDMAMEPDRKAAVSHWQQQSYLDSGIHSGATTAPSLSGKGNPEEEDVDSQVLWEWEQFGSQSFTEQEQADIDGQYAMTRAQVRVRAAMFETLDEGMQIPSTQFDDAHPNTVQRLAEPQMLK
344	DP00698	SVGRQDNMMAQAREAKFAAGGVLVGGGEQDIDEEEPHIEHSGRQSVTFKREMSMSSLADSVSSSVSTSGGTRLTNSLLNLSRLAAKAIKESTAQSSSERNNPPNRPQADSGRKDDQEPKPAQNLDLFVRADV
345	DP00555	MDVFMKGLSMAKEGVVAAAEKTKQGVTEAAEKTKEGVLVYGSKTREGVVGQVASVAEKTKEQASHLGGAVFSGAGNIAATGLVKREEFTDLKPEEVAQEAEEPLIEPLMEPEGESYEDPPQEEYQYEP
346	DP00122	MADERKDEGKAPHWTSASLTEAAHPSPEMKDQGGSGEGLSRANGFPYREEEEGAFGEHGSQGTYSDTKENGINGELTADRETAEEVSARIVQVTAEEAVVLKGEQKEAQHKDQPAALPLAAETVNLPP
347	DP00120	RLEQYTSAVVGNKAAKPAASDLVPAEGVRNIKSMWEKGNVFSPPGTTGTPNKETAGLKVGVSSRINEWLTKTPEGNKSPAPKPSDLRPGDVSGKRNLEWKQSVKPAASSSVTATGKKSETNGLRQFEKEP
348	DP00965	KKDKDGEGAPPAKRARTDQMEVDSGPRKRPSRGGFTDKERQDHRRRKALENKRKQLSAGGKNLSKEEEEELRRLTEEDERRERRRIAGPQVGGVNPLEGGTRGAPGGGFVPSMQGVPESPFTRTGEGLDIRGSQGFP
349	DP00340	MPNFAGTWKMRSSNFDELLKALGVNAMLRLKVAVAAASKPHVEIRQDGDQFYKTTSTTVRTTEINFVGEVGEFEETVDGRKCRSLPTWENENKIHCTQTLLGEDGPKTYWTRELANDELITFGADDVVCRIYVRE
350	DP00794	SSVSSSPSPFGHSAAVSPTFMPRSTQPLTFTTATFAATKFGSTKMKNSGRSNKVARTSPINLGLNVNDLLKQKAISSMSHSLYGLGLGSQQPQQQQQAQPPPPPPPPQQQQQKTSALSPNAKEFIFPNMQGGQSS
351	DP00626	MLFIKPADLREIVTFPLFSDLVQCGFSPAADYVEQRIDLNQLLIQHPSATYFVKASGDSMIDGGISDGLLIVDSAITASHGDVIAAVDGEFTVKKLQRLPTVQLIPMNSAYSPTISSSETLDVFGVVIHVVKAMR
352	DP00070	MDVFMKGLSKAKEGVVAAAEKTKQGVAAEAGKTKEGVLVYGSKTKEGVVHGVAATVAEKTKEQVTNVGGAVVTGTVAAQKTVEGAGSIAATGFVKKDLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEP
353	DP00196	AVPVESKPDKPSGKSGMDAALDDLIDTLGGPEETEEENTTYTGPEVSDPMSSTYIEELGKREVTIPPKYRELLAKKEGITGPPADSSKPIGPDDAIDALSSDFTCGSPTAAGKKTEKEESTVLKAQSAGTVRSAAPPQEK
354	DP00999	CIFEVKEHGKVTGYACLVGDKVMKPAHVKGVIDNADLAKLAFKKSKYDLECAQIPVHMRSADSKYTHEKPEGHYNWHHGAVQYSGGRFTIPTGAGKPGDSGRPIFDNKGRRVVAIVLGGANEGSRTALSVVTWNKDMVTRVT
355	DP00134	FAEDVIQVPRNLVGKVIKNGKLIQEIIVDKSGVVRVRIEAENEKNVPQEEIIMPNSLPSNNSRVGPNAPEEKKHLDIKENSTHFSQPNSTKVQRVLAASSVAVGESQKPELKAWQGMVPFVFGTKDSIANATVLLDYHLN
356	DP00186	MSSQQNQNRQGEQEQGYMEAEEKVNVNAWESTKETLSSTAQAAAEKTAEFRDSAGETIRDLTGQAQEQKGQEFKERAGEKAEETKQRAGEKMDETKQRAAGEMRENAGQKMEEYKQGGKGAEEELRDTAAEKLHQAGEKVKGRD
357	DP00260	MPLNVSFNTNRNYDLDYDSVQPYFYCDEEENFYQQQQQSELQPPAPSEDIWKKFELLPTPLSPSRRSGLCSPSYAVTVPFSLRGDNDGGGGSFSTADQLEMVTELLGGDMVNQSFICDPDETFIKNIIQDCMWSGFSAAK
358	DP00378	TTSGSSGTSNQGGTPQRVSGLQGSDALNIQGNQTSGGSLQAGQKQKEGEQNQQTQQQQLIQPLVQGGQALQALQAAPLSGGQFTTTQAISQETLQNLQLQAVPNSGPIIIRTPTVGPNQGVSWQTLQLQNLQVQNPAQTITLAPMQ
359	DP00378	TNSQGGTPQRVSGLQGSDALNIQGNQTSGGSLQAGQKQKEGEQNQQTQQQQLIQPLVQGGQALQALQAAPLSGGQFTTTQAISQETLQNLQLQAVPNSGPIIIRTPTVGPNQGVSWQTLQLQNLQVQNPAQTITLAPMQVSLGGT
360	DP00092	SDDELGSEEDGFDGATAARKEVIRNKIRAIGKMARVFSVLRREESVLTGLKLTPTGMLPSGVLGGKQTLQSATVEIAEADAIEKGFSPQHKITSFEEAKGLDRINERMPPRRDAMPSDANLNSINKALTSETNGTDSNGSNSSNIQ
361	DP00174	MASSDIQVKELEKRASGQAFELILSPRSKESVPEFPLSPPKKDLGLEIQQKLEAAEERRKSHEAEVLKQLAEKREHEKEVLQKAIENNNFSKMAEKKLTHKMEANKENREAQMAAKLERLREKDKHIEVRKNKESKDADETEAD