

## Supplementary Dataset 1:

This dataset contains the precursor sequences of the human proteins (**Part A**), as well as processed active sequences of the precursors (**Part B, p.8**), that were used for the viral genome database search using tBlastn. The Fasta format sequences of the proteins were obtained from Uniprot, and the link for each protein is reported below the sequence. The processed sequences were also obtained from Uniprot.

### PART A. Protein Sequences and Uniprot Links

```
>sp|P01308|INS_HUMAN Insulin OS=Homo sapiens GN=INS PE=1 SV=1
MALWMRLLPLLALLLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAED
LQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN
```

<http://www.uniprot.org/uniprot/P01308#sequences>

```
>sp|P05019|IGF1_HUMAN Insulin-like growth factor I OS=Homo sapiens GN=IGF1
PE=1 SV=1
MGKISSLPTQLFKCCFCDFLKVMTSSSHLFYLAALCLLTFTSSATAGPETLTCGAEALVD
ALQFVCGDRGFYFNKPTGYGSSRRAPQTGIVDECCFRSCDLRRLREMYCAPLKPASARS
VRAQRHTDMPKTKYQPPSTNKNTKSQRKRWPKTHPGGEQKEGTEASLQIRGKKKEQRR
EIGSRNAECRGKKGK
```

<http://www.uniprot.org/uniprot/P05019#sequences>

```
>sp|P01344|IGF2_HUMAN Insulin-like growth factor II OS=Homo sapiens GN=IGF2
PE=1 SV=1
MGIPMGKSMVLVLLTFLAFASCCIAAYRPSETLGGELVDLTLQFVCGDRGFYFSRPAASRVS
RRSRGIVEECCFRSCDLALLETYCATPAKSERDVSTPPTVLPDNFPRYPVGGKFFQYDTWK
QSTQRLRRGLPALLRARRGHVLAKELEAFREAKRHRPLIALPTQDPAHGGAPPEMASNRK
```

<http://www.uniprot.org/uniprot/P01344#sequences>

```
>sp|P01275|GLUC_HUMAN Glucagon OS=Homo sapiens GN=GCG PE=1 SV=3
MKSIIYFVAGLFVMLVQGSWQRSLQDTEEKSRFSASQADPLSDPDQMNEDKRHSQGTFTS
DYSKYLDSTRRAQDFVQWLMNTRNRNNAIKRHDEFERHAEGTFTSDVSSYLEGQAAKEFI
AWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILDNLAARDFINWLIQTKITDRK
```

<http://www.uniprot.org/uniprot/P01275#sequences>

```
>sp|O00253|AGRP_HUMAN Agouti-related protein OS=Homo sapiens GN=AGRP PE=1
SV=1
MLTAAVLSCALLLALPATRGAQMGLAPMEGIRRPDQALLPELPGLGLRAPLKKTTAEQAE
EDLLQEAQALAEVLDLQDREPRSSRRCVRLHESCLGQVPCCDPCATCYCRFFNAFCYCR
KLG TAMNPCSRT
```

<http://www.uniprot.org/uniprot/O00253>

```
>sp|P01189|COLI_HUMAN Pro-opiomelanocortin OS=Homo sapiens GN=POMC PE=1 SV=2
MPRSCCSRSGALLLALLLQASMEVIRGWCLESSQCQDLTTESNLLECIRACKPDLAETPM
FPGNGDEQPLTENPRKYVMGHFRWDRFGRNRSSSSGSSGAGQKREDVSAGEDCGPLPEGG
PEPRSDGAKPGPREGKRSYSMEHFRWGKPVGKRRPVKVYPNGAEDESAAEFPLEFKREL
TGQRLREGDGPDPADGAGAQAADLEHSLLVAAEKKDEGPYRMEHFRWGSPPKDKRYGGF
MTSEKSQTPLVTLFKNAIIKNAYKKGE
```

<http://www.uniprot.org/uniprot/P01189#sequences>

>sp|P41159|LEP\_HUMAN **Leptin** OS=Homo sapiens GN=LEP PE=1 SV=1  
MHWGTLGFLWLPYLFYVQAVPIQKVQDDTKTIKTIVTRINDISHTQSVSSKQKVTGL  
DFIPGLHPILTLSKMDQTLAVYQQILTSMPSRNVIQISNDLENLRDLLHVLAFSKSKCHLP  
WASGLETLDLSLGGVLEASGYSTEVVALSRLQGSQDMLWQLDLSPGC

<http://www.uniprot.org/uniprot/P41159#sequences>

>sp|P01375|TNFA\_HUMAN **Tumor necrosis factor** OS=Homo sapiens GN=TNF PE=1 SV=1  
MSTESMIRDVELAEALPKKTGGPQGSRRCLFLSLFSFLIVAGATTLFCLLHFGVIGPQR  
EEFPRDLSLISPLAQAVRSSRTPSDKPVAVHVVANPQAEGQLQWLNRRANALLANGVELR  
DNQLVVPSEGLYLIYSQVLFKGGQCPSTHVLTTHTISRIAVSYQTKVNLLSAIKSPQRE  
TPEGAEAKPWYEPYIYLGGVFQLEKGDRLSAEINRPDYLDFAESGQVYFGIIAL

<http://www.uniprot.org/uniprot/P01375#sequences>

>sp|P06307|CCKN\_HUMAN **Cholecystokin** OS=Homo sapiens GN=CCK PE=1 SV=1  
MNSGVCLCVMALVAAGALTQPVPADPAGSGLQRAEEAPRRQLRVSQRTDGESRAHLGA  
LLARYIQQARKAPSGRMSIVKNLQNLDPSTRISDRDYMGMWDFGRRSAEEYEYPS

<http://www.uniprot.org/uniprot/P06307#sequences>

>sp|P01350|GAST\_HUMAN **Gastrin** OS=Homo sapiens GN=GAST PE=1 SV=1  
MQRLCVYVLIIFALALAAFSEASWKPRSQQPDAPLGTGANRDLELPWLEQQGPASHHRQL  
GPQGGPHLVADPSKKQGPWLEEEEEAYGWMDFGRRSAEDEN

<http://www.uniprot.org/uniprot/P01350>

>sp|P01160|ANF\_HUMAN **Natriuretic peptides A** OS=Homo sapiens GN=NPPA PE=1 SV=1  
MSSFSTTTVSFLLLLAFQLLQTRANPMYNAVSNADLMDFKNLLDHLEEKMPLEDEVVPP  
QVLSEPNEEAGAALSPLPEVPPWTGEVSPAQRDGGALGRGPWDSSDRSALLKSKLRALLT  
APRSLRRSSCFGGMRDRIGAQSGLGCNSFRYRR

<http://www.uniprot.org/uniprot/P01160#sequences>

>sp|P01178|NEU1\_HUMAN **Oxytocin-neurophysin 1** OS=Homo sapiens GN=OXT PE=1 SV=1  
MAGPSLACCLLGLLALTSACYIQNCPLGGKRAAPDLVDVRKCLPCGPGGKGRFCFGPNICCA  
EELGCFVGTAEALRCQEENYLPSPCQSGQKACGSGGRCAVLGLCCSPDGCHADPACDAEA  
TFSQR

<http://www.uniprot.org/uniprot/P01178#sequences>

>sp|P01185|NEU2\_HUMAN **Vasopressin-neurophysin 2-copeptin** OS=Homo sapiens  
GN=AVP PE=1 SV=2  
MPDTMLPACFLGLLAFSSACYFQNCPRGGKRAMSDLELRQCLPCGPGGKGRFCFGPSICCA  
DELGCFVGTAEALRCQEENYLPSPCQSGQKACGSGGRCAAFVCCNDESCVTEPECREGF  
HRRARASDRSNATQLDGPAGALLLRLVQLAGAPEPFEPAPDAY

<http://www.uniprot.org/uniprot/P01185#sequences>

>sp|P01241|SOMA\_HUMAN **Somatotropin** OS=Homo sapiens GN=GH1 PE=1 SV=2  
MATGSRTSLLLAFLGLLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEA  
YIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQOKSNLELLRISLLLIQSWLEPVQFLR  
SVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHND  
ALLKNYGLLYCFRKDMKVETFLRIVQCRSVEGSCGF

<http://www.uniprot.org/uniprot/P01241#sequences>

>sp|P01258|CALC\_HUMAN **Calcitonin** OS=Homo sapiens GN=CALCA PE=1 SV=2  
MGFQKFSFPLALSILVLLQAGSLHAAPFRSALESSPADPATLSEDEARLLLLAALVQNYVQ  
MKASELEQEEREGSSLDSPRSKRCGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAPGKKR

DMSSDLERDHRPHVSMPQANAN

<http://www.uniprot.org/uniprot/P01258#sequences>

>sp|P01137|TGFB1\_HUMAN **Transforming growth factor beta-1** OS=Homo sapiens  
GN=TGFB1 PE=1 SV=2

MPPSGLRLLLLLLLPLLLWLLVLTGPRPAAGLSTCKTIDMELVKKRRIEAIHQILSKLRLA  
SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI  
YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR  
YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT  
TGRRGDLATIHGMRPFLLLMATPLERAQHLQSSRRRRALDTNYCFSSTEKNCCVRQLYI  
DFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQA  
LEPLPIVYVGRKPKVEQLSNMIVRSCKCS

<http://www.uniprot.org/uniprot/P01137>

>sp|P61812|TGFB2\_HUMAN **Transforming growth factor beta-2** OS=Homo sapiens  
GN=TGFB2 PE=1 SV=1

MHYCVLSAFLILHLVTVLVALSLSTCSTLDMDQFMRKRIEAIHQILSKLKLTPPEDYPEP  
EEVPPPEVISIYNSTRDLLQEKASRRAAACERERSDEEYAKEVYKIDMPPFFPSENAIPP  
TFYRPFYFRIVRFDVSAEKNASNLVKAEFRVFRQLQNPQKARVPEQRIELYQILSKDLTSP  
TQRYIDSKVVKTRAEGEWLSFDVTDVHEWHLHKKDRNLGFKISLHCPCTFVPSNNYIIP  
NKSEELARFAGIDGTSTYTSQDQKTIKSTRKNSGKTPHLLMLLPSYRLESQQTNRK  
KRALDAAYCFRNVQDNCCLRPLYIDFKRDLGWKWIHEPKGYANFCAGACPYLWSSDTQH  
SRVLSLYNTINPEASASPCCVSQDLEPLTILYIIGKTPKIEQLSNMIVKSKCS

<http://www.uniprot.org/uniprot/P61812#sequences>

>sp|P10997|IAPP\_HUMAN **Islet amyloid polypeptide** OS=Homo sapiens GN=IAPP PE=1  
SV=1

MGILKLQVFLIVLSVALNHLKATPIESHQVEKRKCNTATCATQRLANFLVHSSNNFGAIL  
SSTNVGSNTYGKRNAVEVLKREPLNYLPL

<http://www.uniprot.org/uniprot/P10997#sequences>

>sp|Q15848|ADIPO\_HUMAN **Adiponectin** OS=Homo sapiens GN=ADIPOQ PE=1 SV=1

MLLLGAVLLLLLALPGHDQETTTQGGVLLPLPKGACTGWMAGIPGHPGHNGAPGRDGRDG  
TPGKGEKGDPLIGPKGDIGETGVPGAEGPRGFPGIQGRKGEPEGAYVYRSAFSVGLE  
TYVTIPNMPFRFTKIFYNQONHYDGSTGKFFHCNIPGLYFAYHITVYMKDKVKSFLFKDK  
AMLFTYDQYQENNVDAQSGSVLLHLEVGDVWLQVYGEGERNGLYADNDNDSTFTGFLLY  
HDTN

<http://www.uniprot.org/uniprot/Q15848#sequences>

>sp|P06850|CRF\_HUMAN **Corticoliberin** OS=Homo sapiens GN=CRH PE=1 SV=1

MRLPLLVSAGVLLVALLPCPPCRALLSRGPVPGARQAPQHPQPLDFFQPPQSEQPQQPQ  
ARPVLLRMGEEYFLRLGNLNKSPAAPLSPASSLLAGGSGSRPSPEQATANFFRVLLQQLL  
LPRRSLDSPAALAERGARNALGGHQEAPERERRSEEPPISLDLTFHLLREVLEMARAEQL  
AQQAHSNRKLMEIIGK

<http://www.uniprot.org/uniprot/P06850#sequences>

>sp|O00230|CORT\_HUMAN **Cortistatin** OS=Homo sapiens GN=CORT PE=2 SV=1

MPLSPGLLLLLLGSATATAALPLEGGPTGRDSEHMQEAAGIRKSSLLTFLAWWFETSQA  
SAGPLIGEEAREVARRQEGAPPQQSARRDRMPCRNFFWKTFSSCK

<http://www.uniprot.org/uniprot/O00230#sequences>

>sp|P01210|PENK\_HUMAN **Proenkephalin-A** OS=Homo sapiens GN=PENK PE=1 SV=1

MARFLTLCWLLLLLGPGLLATVRAECSQDCATCSYRLVVPADINFLACVMECEGKPLPSLK  
IWETCKELLQLSKPELPQDGTSTLRENSKPEESHLLAKRYGGFMKRYGGFMKKMDELYPM  
EPEEEANGSEILAKRYGGFMKKDAEEDDSLANSDDLKELLETTGDNRRERSHHQDGSNDNEE

EVSKRYGGFMRGLKRSPQLEDEAKELQKRYGGFMRRVGRPEWWMQKRYGGFLKRFAEA  
LPSDEEGESYSKEVPEMEKRYGGFMRF

<http://www.uniprot.org/uniprot/P01210#sequences>

>sp|P05305|EDN1\_HUMAN **Endothelin-1** OS=Homo sapiens GN=EDN1 PE=1 SV=1  
MDYLLMIFSLLFVACQGAPETAVALGAELSAVGENGGKPTSPPPWRLRRSKRCSLMD  
KECVYFCHLDIIWVNTPEHVVPYGLGSPRSKRALENLLPTKATDRENRCQASQKDKKCW  
NFCQAGKELRAEDIMEKDWNHKKGKDCSKLGKKCIYQQLVGRGRKIRRSSEEHLRQTRSE  
TMRNSVKSSFHDPKLGKGPSRERYVTHNRAHW

<http://www.uniprot.org/uniprot/P05305#sequences>

>sp|P20800|EDN2\_HUMAN **Endothelin-2** OS=Homo sapiens GN=EDN2 PE=1 SV=2  
MVSVPPTWCSVALALLVALHEGKGQAAATLEQPASSSHAQGTHLRLRRCSSSWLDKECV  
YFCHLDIIWVNTPEQTAPYGLGNPPRRRRRSLPRRCQCSSARDPACATFCLRRPWTEAGA  
VPSRKSPADVFTGKTGATTGELLQRLRDIISTVKSLFAKRQEQAMREPRSTHSRWRKR

<http://www.uniprot.org/uniprot/P20800#sequences>

>sp|P01588|EPO\_HUMAN **Erythropoietin** OS=Homo sapiens GN=EPO PE=1 SV=1  
MGVHECPAWLWLLLSLPLGLPVLGAPPRILICDSRVLERYLLEAKEAENITGCAEHC  
SLNENITVDPDKVNFYAWKRMEVGGQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQL  
HVDKAVSGLRSLTLLRALGAQKEAISPPDAASAAPLRITITADTFRKLFRVYSNFLRGKL  
KLYTGEACRTGDR

<http://www.uniprot.org/uniprot/P01588#sequences>

>sp|P22466|GALA\_HUMAN **Galanin peptides** OS=Homo sapiens GN=GAL PE=1 SV=3  
MARGSALLLASLAAAALSASAGLWSPAKEKRGWTLNSAGYLLGPHAVGNHRFSKNGL  
TSKRELRPEDDMKPGSFDRSIPENNIMRTIIEFLSFLHLKEAGALDRLLDLPAAASSEDI  
ERS

<http://www.uniprot.org/uniprot/P22466#sequences>

>sp|Q9UBU3|GHRL\_HUMAN **Appetite-regulating hormone** OS=Homo sapiens GN=GHRL  
PE=1 SV=1  
MPSPGTVCSSLLLLGMLWLDLAMAGSSFLSPEHQVQRKESKPPAKLQPRALAGWLRPE  
DGGQAEGAEDLEVRFNAPFDVGIKLSGVQYQHSQALGKFLQDILWEEAKEAPADK

<http://www.uniprot.org/uniprot/Q9UBU3#sequences>

>sp|P01286|SLIB\_HUMAN **Somatoliberin** OS=Homo sapiens GN=GHRH PE=1 SV=1  
MPLWVFFFVILTLSNSSHCSPPLTLRMRRYADAIFTNSYRKVLGQLSARKLLQDIMSR  
QQGESNQERGARARLGRQVDSMWAEQKQMELESILVALLQKHSRNSQG

<http://www.uniprot.org/uniprot/P01286#sequences>

>sp|P81172|HEPC\_HUMAN **Hepcidin** OS=Homo sapiens GN=HAMP PE=1 SV=2  
MALSSQIWAACLILLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRD  
THFPICIFCCGCCHRSKCGMCKT

<http://www.uniprot.org/uniprot/P81172#sequences>

>sp|P0DN86|CGB3\_HUMAN **Choriogonadotropin subunit beta 3** OS=Homo sapiens  
GN=CGB3 PE=1 SV=1  
MEMFQGLLLLLLLSMGGTASKEPLRPRCRPINATLAVEKEGCPVCITVNTTICAGYCPT  
MTRVLQGVLPALPQVVCNRYRDFVRFESIRLPGCPRGVNPVVSYAVALSCQCALCRRSTTDC  
GGPKDHPLTCDDPRFQDSSSSKAPPPSLPSPSRLPGPSDTPILPQ

<http://www.uniprot.org/uniprot/P0DN86#sequences>

>sp|P0DML2|CSH1\_HUMAN **Chorionic somatomammotropin hormone 1** OS=Homo sapiens  
GN=CSH1 PE=1 SV=1  
MAPGSRTSLLLLAFALLCLPWLQEAGAVQTVPLSRLFDHAMLQAHRAHQLAIDTYQEFEEET  
YIPKDQKYSFLHDSQTSFCFSDSIPTPSNMEETQQKSNLELLLRISLLLLIESWLEPVRFRLR  
SMFANNLVYDTSDDYHLLKDLLEGIQTLMGRLLEDGSRRTGQILKQTYSKFDTNSHNHD  
ALLKNYGLLYCFRKDMDKVETFLRMVQCRSVEGSCGF  
<http://www.uniprot.org/uniprot/P0DML2>

>sp|P0DML3|CSH2\_HUMAN **Chorionic somatomammotropin hormone 2** OS=Homo sapiens  
GN=CSH2 PE=1 SV=1  
MAAGSRTSLLLLAFALLCLPWLQEAGAVQTVPLSRLFDHAMLQAHRAHQLAIDTYQEFEEET  
YIPKDQKYSFLHDSQTSFCFSDSIPTPSNMEETQQKSNLELLLRISLLLLIESWLEPVRFRLR  
SMFANNLVYDTSDDYHLLKDLLEGIQTLMGRLLEDGSRRTGQILKQTYSKFDTNSHNHD  
ALLKNYGLLYCFRKDMDKVETFLRMVQCRSVEGSCGF  
<http://www.uniprot.org/uniprot/P0DML3#sequences>

>sp|P08476|INHBA\_HUMAN **Inhibin beta A chain** OS=Homo sapiens GN=INHBA PE=1  
SV=2  
MPLLWLRGFLASCWIIVRSSPTPGSEGHSAAPDCPSCALAALPKDVPNSQPEMVEAVKK  
HILNMLHLKKRPDVTQVPVKAALLNAIRKLVHGVKVGGENGYVEIEDDIGRAEMNELMEQT  
SEIITFAESGTARKTLHFEISKEGSDLSVERAEVWFLFKVPKANRTRTKVTIRLFQQQK  
HPQGS�DTGEEAAEVGLKGERSELLSEKVV DARKSTWHVFPVSSSIQRLLDQGKSSLDV  
RIACEQCQESGASLVLLGKKKKKEEGEGKKGKGGEGGAGADEEKEQSHRPFLMLQARQS  
EDHPHRRRRRGLCDGKVNICCKKQFFVSFKDIGWNDWIIAPSGYHANYCEGECPSHIAG  
TSGSSLSFHSTVINHYRMRGHS PFANLKS CCVPTKLRPMSMLYDDGQNI IKKDIQNMIV  
EECGCS  
<http://www.uniprot.org/uniprot/P08476#sequences>

>sp|P01229|LSHB\_HUMAN **Lutropin subunit beta** OS=Homo sapiens GN=LHB PE=1 SV=3  
MEMLQGLLLLLLLLLSMGGAWASREPLRPWCHPINAILAVEKEGCPVCITVNTTICAGYCPT  
MMRVLQAVLPPLPQVVCTYRDVRFESIRLPGCPRGVDPVVSFPVALSCRCGPCRRSTSDC  
GGPKDHPLTCDHPQLSGLLFL  
<http://www.uniprot.org/uniprot/P01229#sequences>

>sp|O43612|OREX\_HUMAN **Orexin** OS=Homo sapiens GN=HCRT PE=1 SV=1  
MNL PSTKVSWAAVTLLLLLLLLLPPALLSSGAAQPLPDCCRQKTCSCRLYELLHGAGNHA  
AGILTLGKRRSGPPGLQGRLLQASGNHAAGILTMGRRAGAEPAPRPCLGRRC SAPAA  
ASVAPGGQSGI  
<http://www.uniprot.org/uniprot/O43612#sequences>

>sp|P02818|OSTCN\_HUMAN **Osteocalcin** OS=Homo sapiens GN=BGLAP PE=1 SV=2  
MRALTLLALLALAALCIAGQAGAKPSGAESSKGAAFVSKQEGSEVVKRPRRYLYQWLGAP  
VPYPDPLEPREVCELNPDCDELADHIGFQEAYRRFYGPV  
<http://www.uniprot.org/uniprot/P02818#sequences>

>sp|P01298|PAHO\_HUMAN **Pancreatic prohormone** OS=Homo sapiens GN=PPY PE=1 SV=1  
MAAARLCLSLLLLLSTCVALLLQPLLGAQGAPLEPVYPGD NATPEQMAQYAADLRRYINML  
TRPRYGRHKEDTLAFSEWGS PHAAVPRELSPLDL  
<http://www.uniprot.org/uniprot/P01298#sequences>

>sp|P01270|PTHY\_HUMAN **Parathyroid hormone** OS=Homo sapiens GN=PTH PE=1 SV=1  
MIPAKDMAKVMIVMLAICFLTKSDGKSVKRSVSEIQLMHNLGKHLNSMERVEWLRKKLQ  
DVHNFVALGAPLAPRDAGSQRPRKKEDNVLVESHEKSLGEADKADVNVLTAKKSQ  
<http://www.uniprot.org/uniprot/P01270#sequences>

>sp|P18509|PACA\_HUMAN Pituitary adenylate cyclase-activating polypeptide

OS=Homo sapiens GN=ADCYAP1 PE=1 SV=3

MTMCSGARLALLVYGIIMHSSVYSSPAAAGLRFPGIRPEEEAYGEDGNPLPDFDGSEPPG  
AGSPASAPRAAAAWYRPAGRRDVAHGILNEAYRKVLDQLSAGKHLQSLVARGVGGSLGGG  
AGDDAEPLSKRHS DIGIFTDSYSRYRKQMAVKKYLA AVL GKRYKQ RVKNKGRRIAYL

<http://www.uniprot.org/uniprot/P18509#sequences>

>sp|P01236|PRL\_HUMAN Prolactin OS=Homo sapiens GN=PRL PE=1 SV=1

MNIKGS PWKGS LLL LLL VSN LLL CQSVAPLPICPGGAARCQVTLRDLFDRAVVL SHYIHNLS  
SSEM FSEFDKRYTHGRGFITKAINSCHTSSLATPEDKEQAQQMNQKDFLSLIVSILRSWN  
EPLYHLVTEVRGMQEAPAILSKAVEIEEQTKRLLGEMELIVSQVHPETKENEIYPVWSG  
LPSLQMADEESRLSAYYNLLHCLRRDSHKIDNYLKLKLCRIIHNNNC

<http://www.uniprot.org/uniprot/P01236#sequences>

>sp|P01148|GON1\_HUMAN Progonadoliberein-1 OS=Homo sapiens GN=GNRH1 PE=1 SV=1

MKPIQKLLAGLILLTWCVEGCSSQHWSYGLRPGGKRDAENLIDSFQEIIVKEVGLAETQR  
FECTTHQPRSPRLRDLKGALES LIEEETGQKKI

<http://www.uniprot.org/uniprot/P01148#sequences>

>sp|P04090|REL2\_HUMAN Prorelaxin H2 OS=Homo sapiens GN=RLN2 PE=1 SV=1

MPRLFFFHLLGVCLLLNQFSRAVADSWMEEVIKLCGREL VRAQIAICGMSTWSKRSLSQE  
DAPQTPRPVAEIVPSFINKDTETINMMSEFVANLPQELKLTLESEMQPALPQLQHVPLK  
DSSLLFEFEFKK LIRNRQSEAADSSPSELKYLGLDTHSRKKRQLYSALANKCCHVGCTKRS  
LARFC

<http://www.uniprot.org/uniprot/P04090#sequences>

>sp|Q8WXF3|REL3\_HUMAN Relaxin-3 OS=Homo sapiens GN=RLN3 PE=1 SV=1

MARYMLLLLLLAVVVL T GELWP GAEARAAPYGVRLCGREFIRAVIFTCGGSRRRRSDILAH  
EAMGDTFPDADADED SLAGELDEAMGSSEWLALTKSPQAFYRGRPSWQGT PGVLRGSRDV  
LAGLSSSCCKWGCSEISSL C

<http://www.uniprot.org/uniprot/Q8WXF3#sequences>

>sp|P04808|REL1\_HUMAN Prorelaxin H1 OS=Homo sapiens GN=RLN1 PE=1 SV=1

MPRLF L FHLLEFCLLLNQFSRAVA AKWKDDVIKLCGREL VRAQIAICGMSTWSKRSLSQE  
DAPQTPRPVAEIVPSFINKDTETIIIMLEFIANLPPELKAALSERQPSLPQLQYVPALK  
DSNLSFEFEFKK LIRNRQSEAADSNPSELKYLGLDTHSQKKRRPYVALFEKCC LIGCTKRS  
LAKYC

<http://www.uniprot.org/uniprot/P04808#sequences>

>sp|P00797|RENI\_HUMAN Renin OS=Homo sapiens GN=REN PE=1 SV=1

MDGWRRMPRWGLLLLLLWGSCTFGLPTDTTTFKRIFLKRMP SIRESLKERGVDMARLGPEW  
SQPMKRLTLGNTTSSVILTNYMDTQYYGEIGIGTPPQTFKVVFD TGSSNVVWPSSKCSRL  
YTACVYHKLF DADSSSYKHNGTELT LRYSTGT VSGFLSQDIITVGGITVTQMFGEVTEM  
PALPFMLAEFDGVVGMGFIEQAIGRVTPIFDNIIISQVLKEDVFSFYNNRDSSENSQSLGG  
QIVLGGSDPQHYEGNFHYINLIK TG VVQIQMKGVS VGSSTLLCEDGCLALVDTGASYISG  
STSSIEKLMEALGAKKRLFDYVVKCNEGPTLPDISFHLGGKEYTLTSADYVFQESYSSKK  
LCTLAIHAMDIPPTGPTWALGATFIRK FYTEFDRNNRIGFALAR

<http://www.uniprot.org/uniprot/P00797#sequences>

>sp|P09683|SECR\_HUMAN Secretin OS=Homo sapiens GN=SCT PE=1 SV=2

MAPRPLLLLLLLLLGGSAARPAPPRARRHSDGTF TSELSRLREGARLQRL LQGLVGRSEQ  
DAENSMAWTRLSAGLLCPSGSNM P ILQAWMPLDGTWSPWLP PPGPMVSE P AGAAAEGTLRP  
R

<http://www.uniprot.org/uniprot/P09683#sequences>

>sp|P61278|SMS\_HUMAN **Somatostatin** OS=Homo sapiens GN=SST PE=1 SV=1  
MLSCRLQCALAALSIVLALGCVTGAPSDPRLRQFLQKSLAAAAGKQELAKYFLAELLSEP  
NQTENDALEPEDLSQAAEQDEMRLQLQRSANSNPAMAPRERKAGCKNFFWKFTFTSC  
<http://www.uniprot.org/uniprot/P61278#sequences>

>sp|P40225|TPO\_HUMAN **Thrombopoietin** OS=Homo sapiens GN=THPO PE=1 SV=1  
MELTELLLLVVMLLLLTARLTLSSPAPPACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPV  
LLPAVDVFLGGEWKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRL  
LLGALQSLGTLGTQLPPQGRTTAHKDPNAIFLSFQHLLRGKVRFLMLVGGSTLCVRRAPPTT  
AVPSRTSLVLTNLPNRTSGLLETNFTASARTTGSGLLKWQQGFRAKIPGLLNQTSRSL  
DQIPGYLNRIHELLNGTRGLFPGPSRRTLGAPDISSGTSDTGSLPPNLQPGYSPSPTHPP  
TGQYTLFPLPPTLPTPVVQLHPLLPDPSAPTPTPTSPLLNTSYTHSQNLSQEG  
<http://www.uniprot.org/uniprot/P40225#sequences>

>sp|P01222|TSHB\_HUMAN **Thyrotropin subunit beta** OS=Homo sapiens GN=TSHB PE=1  
SV=2  
MTALFLMSMLFGLTCGQAMSFICPTEYTMHIERRECAVCLTINTTICAGYCMTRDINGKL  
FLPKYALSQDVCTYRDFIYRTVEIPGCPLHVAPYFSYPVALSCKCGKCNTRYSDCIHEAI  
KTYNCTKPQKSYLVGFSV  
<http://www.uniprot.org/uniprot/P01222#sequences>

>sp|P20396|TRH\_HUMAN **Pro-thyrotropin-releasing hormone** OS=Homo sapiens GN=TRH  
PE=1 SV=1  
MPGPWLLLLALALTLNLTGVPGGRAQPEAAQQEAVTAAEHPGLDDFLRQVERLLFLRENIQ  
RLQGDQGEHSASQIFQSDWLSKRQHPGKREEEEEEGVEEEEEEGGAVGPHKRQHPGRRE  
DEASWSVDVTQHKRQHPGRRSPWLAYAVPKRQHPGRRADPKAQRSWEEEEEEEEEREEDL  
MPEKRQHPGKRALGGPCGPQAYGQAGLLGLLDDLSRSQGAEEKRQHPGRRAAWVREPL  
EE  
<http://www.uniprot.org/uniprot/P20396#sequences>

>sp|P01282|VIP\_HUMAN **VIP peptides** OS=Homo sapiens GN=VIP PE=1 SV=1  
MDTRNKAQLLVLLTLLSVLFSQTSAPLYRAPSALRLGDRIPFEGANEPDQVSLKEDIDM  
LQNALAENDTPPYDVSARNARHADGVFTSDFSKLLGQLSAKKYLESLMGKRVSSNISEDPV  
PVKRHSDAVFTDNYTRLRQMAVKKYLNSILNGKRSSEGESPDFPEELEK  
<http://www.uniprot.org/uniprot/P01282#sequences>

>sp|Q02747|GUC2A\_HUMAN **Guanylin** OS=Homo sapiens GN=GUCA2A PE=1 SV=2  
MNAFLLSALCLLGAWAALAGGVTVQDGNFSFSLESVKKLDLQEPQEPQPRVGKLRNFAPIP  
GEPVVPILCSNPNFPEELKPLCKEPNAQEILQRLEETIAEDPGTCEICAYAAGTGC  
<http://www.uniprot.org/uniprot/Q02747#sequences>

>sp|Q16661|GUC2B\_HUMAN **Guanylate cyclase activator 2B** OS=Homo sapiens  
GN=GUCA2B PE=1 SV=1  
MGCRAASGLLPGVAVVLLLLLQSTQSVYIQYQGFRVQLESMMKLSDLAQWAPSPRLQAQ  
SLLPAVCHHPALPQDLQPCASQEASSIFKTLRTIANDDCELCVNVACTGCL  
<http://www.uniprot.org/uniprot/Q16661#sequences>

>sp|P00746|CFAD\_HUMAN **Complement factor D** OS=Homo sapiens GN=CFD PE=1 SV=5  
MHSWERLAVLVLLGAAACAAPPRGRILGGREAEAHARPYMASVQLNGAHLCCGGVLVAEQW  
VLSAAHCLEDAADGKVQVLLGAHLSQPEPSKRLYDVLRAVPHPDSQPDTIDHLLLLLQQL  
SEKATLGPVAVRPLPWQRVDRDVAPGTLCDVAGWIVNHAGRRPDSLQHVLLPVLDRATCN  
RRTHHDGAIATERLMCAESNRRDSCKGDGGPLVCGGVLEGVVTSGRVCGNRKKPGIYTR  
VASAAWIDSVLA  
<http://www.uniprot.org/uniprot/P01024>

>sp|Q9HD89|RETN\_HUMAN **Resistin** OS=Homo sapiens GN=RETN PE=1 SV=1  
MKALCLLLLPVLGLLVSSKTLCSMEEAINERIQEAVAGSLIFRAISSIGLECQSVTSRGDL  
ATCPRGFAVTGCTCGSACGSWDVRAETTCHCQCAGMDWTGARCCRVQP  
<http://www.uniprot.org/uniprot/Q9HD89>

>sp|P02753|RET4\_HUMAN **Retinol-binding protein 4** OS=Homo sapiens GN=RBP4 PE=1  
SV=3  
MKWVWALLLLAALGSGRAERDCRVSSFRVKENFDKARFSGTWYAMAKKDPEGLFLQDNIV  
AEFSVDETGQMSATAKGRVRLNNDVDCADMVGTFTDTEDEPAKFKMKYWGVASFLQKGN  
DHWIVDTDYDYAVQYSCRLNLDGTCADSYFVFSRDPNGLPPEAQKIVRQRQEELCLA  
RQYRLIVHNGYCDGRSERLL  
<http://www.uniprot.org/uniprot/P02753>

>sp|P80188|NGAL\_HUMAN **Neutrophil gelatinase-associated lipocalin** OS=Homo  
sapiens GN=LCN2 PE=1 SV=2  
MPLGLLWLGLALLGALHAQAQDSTSDLIAPPLSKVPLQONFQDNQFQGKQWYVVGLAGNA  
ILREDKDPQKMYATIYELKEDKSYNVTSVLFRRKKKCDYWIRTFVPGCQPGEFTLGNIKSY  
PGLTSYLVRVSTNYNQHAMVFFKKVSNREYFKITLYGRTKELTSELKENFIRFSKSLG  
LPENHIVFPVPIDQCIDG  
<http://www.uniprot.org/uniprot/P80188>

>sp|P05231|IL6\_HUMAN **Interleukin-6** OS=Homo sapiens GN=IL6 PE=1 SV=1  
MNSFSTSAFGPVAFSLGLLLVLPAAFPAPVPPGEDSKDVAAPHRQPLTSSERIDKQIRYI  
LDGISALRKETCNKSNMCESSKEALAENNLNLPKMAEKDGCQSGFNEETCLVKIITGLL  
EFEVYLEYLQNRFESEEQARAVQMSTKVLIQFLQKKAKNLDAITTPDPTTNASLLTKLQ  
AQNQWLQDMTTHLILRSFKEFLQSSLRALRQM  
<http://www.uniprot.org/uniprot/P05231>

>sp|Q14116|IL18\_HUMAN **Interleukin-18** OS=Homo sapiens GN=IL18 PE=1 SV=1  
MAAEPVEDNCINFVAMKFIIDNTLYFIAEDDENLESDFGKLESKLSVIRNLNDQVLFIDQ  
GNRPLFEDMTDSDCRDNAPRTIFIISMYSKQPRGMAVTISVKCEKISTLSCENKIISFK  
EMNPPDNIKDTSKSDIIFQRSVPGHDNKMQFESSYEGYFLACEKERDLFKLILKKEDEL  
GDRSIMFTVQNE  
<http://www.uniprot.org/uniprot/Q14116>

>sp|P09681|GIP\_HUMAN **Gastric inhibitory polypeptide** OS=Homo sapiens GN=GIP  
PE=1 SV=1  
MVATKTFALLLSLFLAVGLGEGKEGHFSALPSLPVGSKAVSSPQPRGPRYAEGTFISD  
YSIAMDKIHQQDFVNWLLAQKGGKNDWKHNITQREARALELASQANRKEEEAVEPQSSPA  
KNPSDELLRDLLIQELLACLLDQTNLCRLRSR  
<http://www.uniprot.org/uniprot/P09681>

>sp|Q8NAU1|FNDC5\_HUMAN **Fibronectin type III domain-containing protein 5**  
OS=Homo sapiens GN=FNDC5 PE=1 SV=3  
MHPGSPSAWPPRARAALRLWLGCVCFALVQADSPSAPVNVTVRHLKANSVVSVDVLEDE  
VVIGFAISQKKDVRMLRFIQEVNTTTRSCALWDLEEDTEYIVHVQAIISIQQSPASEPV  
LFKTPREAEKMASKNKDEVTMKEMGRNQQLRTGEVLIIVVVLFMWAGVIALFCRQYDIK  
DNEPNNNKEKTKSASETSTPEHQGGGLLRSKI  
<http://www.uniprot.org/uniprot/Q8NAU1>

>sp|Q9NSA1|FGF21\_HUMAN **Fibroblast growth factor 21** OS=Homo sapiens GN=FGF21  
PE=1 SV=1  
MDSDETGFEGHSLWVSVLAGLLGACQAHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAH  
LEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDQALYGLHFDPEA  
CSFRELLLEDGYNVYQSEAHGLPLHLPGNKSPhRDPAPRGPAPFLPLPGLPPALPEPPGI  
LAPQPPDVGSSDPLSMVGPSQGRSPSYAS



<http://www.uniprot.org/uniprot/Q9NSA1>

```
>sp|O95750|FGF19_HUMAN Fibroblast growth factor 19 OS=Homo sapiens GN=FGF19
PE=1 SV=1
MRSGCVVHVWILAGLWLVAVAGRPLAFSDAGPHVHYGWGDPRLRLRHLTYSGPHGLSSCFLL
RIRADGVVDCARGQSAHSLLEIKAVLRVVAIKGVHVSRYLRCMGADGKMQGLLQYSEEDC
AFEEEEIRPDGYNVYRSEKHLRPLVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLR
GHLESDMFSSPLETDSMDPFGLVTGLEAVRSPSFEK
http://www.uniprot.org/uniprot/O95750
```

**PART B. Processed Protein Sequences :** The sequences of the processed hormones reported above were obtained from Uniprot and reported with Uniprot naming. For instance for this first code reported >sp|O00253|21-82, O00253 is the ID of the protein at Uniprot and 21-82 shows the corresponding sequence of the active peptide.

```
>sp|O00253|21-82
AQMGLAPMEGIRRPDQALLPELPLGLRAPLKKTTAEQAEEDLLQEAQALAEVLDLQDRE
PR
>sp|Q15848|19-244
ETTTQGGVLLPLPKGACTGWMAGIPGHPHNGAPGRDGRDGTPEKGEKGDPLIGPKG
DIGETGVPGAEGPRGFPGIQGRKGEPEGAYVYRSAFSVGLETYVTIPNMPPIRFTKIFYN
QQNHYDGGSTGKFHCNIPGLYYFAYHITVYMKDVKVSLFKKDKAMLFTYDQYQENNVDQAS
GSVLLHLEVGDQVWLQVYGEGERNGLYADNDNDSTFTGFLLYHDTN
>sp|P01160|124-151
SLRRSSCFGGRMDRIGAQSGLCNSFRY
>sp|P01160|26-123
NPMYNAVSNADLMDFKNLLDHLEEKMPLEDEVVPPQVLSEPNEEAGAALSPLPEVPPWTG
EVSPAQRDGGALGRGPWDSSDRSALLKSKLRALLTAPR
>sp|P01258|85-116
CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAP
>sp|P01258|121-141
DMSSDLERDHRPHVSMQANAN
>sp|P0DN86|21-165
SKEPLRPRCRPINATLAVEKEGCPVCITVNTTICAGYCPTMTRVLQGVLPALPQVVCNYR
DVRFESIRLPGCPRGVNPPVSYAVALSCQCALCRRSTTDCGGPKDHPLTCDDPRFQDSSS
SKAPPSLPSRLPGPSDTPILPQ
>sp|P06307|21-115
QPVPADPAGSGLQRAEEAPRRQLRVSQRTDGESRAHLGALLARYIQQARKAPSGRMSIV
KNLQNLDP SHRISDRDYMGMWDFGRRSAEEYEYPS
>sp|P06307|46-103
VSQRTDGESRAHLGALLARYIQQARKAPSGRMSIVKNLQNLDP SHRISDRDYMGMWDF
>sp|P06307|65-103
YIQQARKAPSGRMSIVKNLQNLDP SHRISDRDYMGMWDF
>sp|P06307|71-103
KAPSGRMSIVKNLQNLDP SHRISDRDYMGMWDF
>sp|P06307|92-103
ISDRDYMGMWDF
>sp|P06307|79-103
IVKNLQNLDP SHRISDRDYMGMWDF
>sp|P06307|86-103
LDPSHRISDRDYMGMWDF
>sp|P01024|23-1663
SPMYSIITPNILRLESEETMVLEAHDAAQGDVPTVTVHDFPGKLVLSSEKTVLTPATNH
MGNVTFTIPANREFKSEKGRNKVTVQATFGTQVVEKVVVLSLQSGYLFIQTDKTIYTPG
STVLYRIFTVNHKLLPVGRTVMVNIENPEGIPVKQDSLSSQNQLGVLP LSWDIPELVNMG
```

QWKIRAYYENSPOQVFSTEFEVKEYVLPSPFEVIVEPTEKFYIYNEKGLEVTITARFLYG  
 KKVEGTAFFVIFGIQDGEQRISLPESLKRIPIEDGSGEVVLSRKVLLDGVQNPRAEDLVGK  
 SLYVSATVILHSGSDMVQAERSGIPIVTSPIYQIHFTKTPKYFKPGMPFDLMVFTNPDGS  
 PAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPLSITVVRTKKQELSEAEQATRTMQ  
 ALPYSTVGNSSNYLHLSVLRTELRPGETLNVNFFLLRMDRAHEAKIRYYTYLIMNKGRLK  
 AGRQVREPGQDLVVLPLSITTFDIPSFRLVAYYTLIGASGQREVVADSVWVDVKDSCVGS  
 LVVKSGQSEDRQPVPGQOMTLKIEGDHGARVVLVAVDKGVFVNLKKNKLTQSKIWDVVEK  
 ADIGCTPGSGKDYAGVFSADAGLTFTSSSGQQAQRAELQCPQPAARRRRSVQLTEKRMDK  
 VGKYPKELRKCCEDGMRENPMRFSCQRTRFISLGEACKKVFLDCCNYITELRRQHARAS  
 HGLLARSNLDEDIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLKDSITTW  
 EILAVSMSDDKKGICVADPFVEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYRQNLKLV  
 RVELLHNPAFCSLATTKRRHQQTVTIIPKSSLSVPYVIVPLKTGLQEVEVKAAYVHHFIS  
 DGVRSKLVVPEGIRMNKTVAVRTLDPERLGRGVQKEDI PPADLSDQVPDTESETRILL  
 QGTPVAQMTEDAVIDAERLKHILVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFGLEKR  
 QGALELIKKGYTQQLAFRQPSSAFAAFVKRAPSTWLTAYVVKVFLAVNLI AIDSQVLCG  
 AVKWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKDMALTAFLVLSLQEAKDICEEQV  
 NSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQMGRLKGPLLNKFLTTAKDKNRWED  
 PGKQLYNVEATS YALLALLQKDFDFVPPVVRWLNEQRYGGGGYSTQATFMVFQALAQY  
 QKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLRSEETKENEGFTVTAEGKGQGT  
 SVVTMYHAKAKDQLTCNKFDLKVTIKPAPETEKRPQDAKNTMILEICTRYRGDQDATMSI  
 LDISMMTGFAPDTPDDLKQLANGVDRIYSKYELDKAFSDRNTLIIYLDKVSHEDDCLAFK  
 VHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKEDGKLNKLCRDELRCRAEENCFIGK  
 SDDKVTLEERLDKACEPGVDYVYKTRLVKVQLSNDFDEYIMAIEQTIKSGSDEVQVGGQR  
 TFI SPIKREALKLEEKHYLMWGLSSDFWGEKPNLSYIIGKDTWVEHWPEEDECQDEEN  
 QKQCQDLGAFTESMVVFGCPN  
 >sp|P01189|27-102-**NPP**  
 WCLESSQCQDLTTESNLLECIRACKPDLAETPMFPGNGDEQPLTENPRKYVMGHFRWDR  
 FGRNSSSSGSSGAGQ  
 >sp|P01189|77-87-**Melanotropin gamma**  
 YVMGHFRWDRF  
 >sp|P01189|138-150-**Melanocyte-stimulating hormone alpha**  
 SYSMEHFRWGKPV  
 >sp|P01189|156-176- **Corticotropin-like intermediary peptide**  
 PVKVYPNGAEDESAAEFPLEF  
 >sp|P01189|179-267- **Lipotropin beta**  
 ELTGQRLREGDGPDPADDGAGAQAADLEHSLLVAAEKKDEGPYRMEHFRWGSPPKDKRYG  
 GFMTSEKSQTPLVTLFKNAIIKNAYKKGE  
 >sp|P01189|179-234 **Lipotropin gamma**  
 ELTGQRLREGDGPDPADDGAGAQAADLEHSLLVAAEKKDEGPYRMEHFRWGSPPKD  
 >sp|P01189|217-234-**Melanocyte-stimulating hormone beta**  
 DEGPYRMEHFRWGSPPKD  
 >sp|P01189|237-267- **Beta-endorphin**  
 YGGFMTSEKSQTPLVTLFKNAIIKNAYKKGE  
 >sp|P01189|138-176- **Corticotropin**  
 SYSMEHFRWGKPVGKRRPVKVYPNGAEDESAAEFPLEF  
 >sp|P01189|105-134- **Potential peptide**  
 EDVSAGEDCGPLPEGGPEPRSDGAKPGPRE  
 >sp|O00230|77-105  
 QEGAPPQQSARRDRMPCRNFFWKTFSSCK  
 >sp|O00230|89-105  
 DRMPCRNFFWKTFSSCK  
 >sp|P0DML3|27-217  
 VQTVPLSRLFDHAMLQAHRAHQLAIDTYQEFEEITYIPKDQKYSFLHDSQTSFCFSDSIPT  
 PSNMEETQOKSNLELLRISLIIIWLEPVRFLRSMFANNLVYDTSDDSDYHLLKDLEEG  
 IQTLMGRLEDGSRRTGQILKQTYSKFDTNSHNHDALLKNYGLLYCFRKDMDKVETFLRMV  
 QCRSVEGSCGF  
 >sp|P06850|154-194

SEEPPIISLDLTFHLLREVLEMARAEQLAQQAHSNRKLMEII  
>sp|P0DML2|27-217  
VQTVPLSRLFDHAMLQAHRAHQLAIDTYQEFEEETYIPKDQKYSFLHDSQTSFCFSDSIPT  
PSNMEETQQKSNLELLRISLILLIESWLEPVRFLRSMFANNLVYDTSDDYHLLKDLEEG  
IQTLMGRLEDGSRRTGQILKQTYSKFDTNSHNHDALLKNYGLLYCFRKDMDKVETFLRMV  
QCRSVEGSCGF  
>sp|P01588|28-193  
APPRLICDSRVLERYLLEAKEAENITGCAEHCSLNENITVPDTKVNFYAWKRMEVGGQA  
VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRRLALGAQKEAIS  
PPDAASAAPLRTITADTFRKLFVYSNFLRGKLLKLYTGEACRTGDR  
>sp|P20800|49-69  
CSCSSWLDKECVYFCHLDIIW  
>sp|P05305|53-90  
CSCSSLMDKECVYFCHLDIIWVNTPEHVVPYGLGSPRS  
>sp|P05305|53-73  
CSCSSLMDKECVYFCHLDIIW  
>sp|Q9NSA1|29-209  
HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKP  
GVIQILGVKTSRFLCQRPDQALYGLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHPLG  
NKSPHRDPAPRGPAPRFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSMVGPSQGRSPSYA  
S  
>sp|O95750|25-216  
LAFSDAGPHVHYGWDPIRLRHLYTSGPHGLSSCFRLRIRADGVVDCARGQSAHSLLEIKA  
VALRTVAIKGVHSVRYLCMGADGKMQLLQYSEEDCAFEIIIIRPDGYNVYRSEKHLRPLVS  
LSSAKQRQLYKNRGLPLSHFLPMLPMVPEEPEDLRGHLESDMFSSPLETDSMDPFGLVT  
GLEAVRSPSFEK  
>sp|Q8NAU1|32-212  
DSPSAPVNVTVRHLKANSASVSWDVLEDEVVIGFAISQQKKDVRMLRFIQEVNTTTRSCA  
LWDLEEDTEYIVHVQAISIQQGSPASEPVLFKTPREAEKMASKNKDEVMTKEMGRNQQLR  
TGEVLIIVVVLFMWAGVIALFCRQYDI IKDNEPNNNKEKTKSASETSTPEHQGGGLLRK  
I  
>sp|Q8NAU1|32-143  
DSPSAPVNVTVRHLKANSASVSWDVLEDEVVIGFAISQQKKDVRMLRFIQEVNTTTRSCA  
LWDLEEDTEYIVHVQAISIQQGSPASEPVLFKTPREAEKMASKNKDEVMTKME  
>sp|P01350|22-92  
SWKPRSQQPDAPLGTGANRDLELPWLEQQGPASHHRRQLGPPGPHLVADPSKKQGPWLE  
EEEEAYGWMDF  
>sp|P01350|41-92  
DLELPWLEQQGPASHHRRQLGPPGPHLVADPSKKQGPWLEEEEEAYGWMDF  
>sp|P01350|59-92  
QLGPPGPHLVADPSKKQGPWLEEEEEAYGWMDF  
>sp|P01350|76-92  
QGPWLEEEEEAYGWMDF  
>sp|P01350|79-92  
WLEEEEEAYGWMDF  
>sp|P22466|33-62  
GWTLNSAGYLLGPHAVGNHRSFSDKNGLTS  
>sp|P22466|65-123  
ELRPEDDMKPGSFRSIPENNIMRTIIIEFLSFLHLKEAGALDRLLDLPAASSEDIERS  
>sp|Q9UBU3|24-51  
GSSFLSPEHQRVQQRKESKKPPAKLQPR  
>sp|Q9UBU3|24-50  
GSSFLSPEHQRVQQRKESKKPPAKLQP  
>sp|P09681|52-93  
YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGGKNDWKHNITQ  
>sp|P01275|21-89 **Glicentin**  
RSLQDTEEKSRFSASQADPLSDPDQMNEDKRHSQGTFTSDYSKYLDSTRRAQDFVQWLMN

TKRNRNNIA  
>sp|P01275|21-50 **Glicentin-related polypeptide**  
RSLQDTEEKSRFSASQADPLSDPDQMNE  
>sp|P01275|53-89 **Oxyntomodulin**  
HSQGTFTSDYSKYLDSTRRAQDFVQWLMNTKRNRNNIA  
>sp|P01275|53-81 **Glucagon**  
HSQGTFTSDYSKYLDSTRRAQDFVQWLMNT  
>sp|P01275|92-128 **Glucagon-like peptide 1**  
HDEFERHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG  
>sp|P01275|98-128 **Glucagon-like peptide 1(7-37)**  
HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG  
>sp|P01275|98-127 **Glucagon-like peptide 1(7-36)**  
HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR  
>sp|P01275|146-178 **Glucagon-like peptide 2**  
HADGSFSDEMNTILDNLAARDFINWLIQTKITD  
>sp|P01275|GLUC\_HUMAN Glucagon OS=Homo sapiens GN=GCG PE=1 SV=3  
MKSIFYFVAGLVFVMLVQGSWQSLQDTEEKSRFSASQADPLSDPDQMNEKDRHSQGTFTS  
DYSKYLDSTRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFI  
AWLVKGRGRDRDFPEEVAIVEELGRRHADGSFSDEMNTILDNLAARDFINWLIQTKITDRK  
>sp|P01148|24-92  
QHWSYGLRPGGKRDAENLIDSFQEIIVKEVGQLAETQRFECTTHQPRSPLRDLKGALESLE  
EEETGQKKI  
>sp|P01148|37-92  
DAENLIDSFQEIIVKEVGQLAETQRFECTTHQPRSPLRDLKGALESLEEEETGQKKI  
>sp|P01229|21-141  
SREPLRPWCHPINAILAVEKEGCPVCITVNTTICAGYCPTMMRVLQAVLPPLPQVVCTYR  
DVRFESIRLPGCPRGVDVSVFVALSCRCGPCRRSTSDCGGPKDHPLTCDHPQLSGLLF  
L  
>sp|P01185|32-124  
AMSDLELRQCLPCGPGGKRCFGPSCCADELGCFVGTAEALRCQEENYLPSPCQSGQKA  
CGSGGRCAAFVCCNDESCVTEPECREGFHRA  
>sp|P01185|126-164  
ASDRSNATQLDGPAGALLLRLVQLAGAPEPFEPAPDAY  
>sp|P80188|21-198  
QDSTSDLIAPPLSKVPLQONFQDNQFQGKQWVVLGAGNAILREDKDPQKMYATIYELKE  
DKSYNVTSVLFRKKKCDYWIRTFVPGCQPGFTLGNIKSYPLTSYLVRVSTNYNQHAM  
VFFKQVSNREYFKITLYGRTEKELTSELKENFIRFSKSLGLPENHIVFPVPIDQCIDG  
>sp|P01178|32-125  
AAPDLVDRKCLPCGPGGKRCFGPNICCAEELGCFVGTAEALRCQEENYLPSPCQSGQKA  
CGSGGRCAVLGLCCSPDGCHADPACDAEATFSQR  
>sp|P01236|29-227  
LPICPGGAARCVTLRDLFDRAVLSHYIHNLSSSEMSEFDKRYTHGRGFITKAINSCHT  
SSLATPEDKEQAQMNQKDFLSLIVSILRSWNEPLYHLVTEVRGMQEAPAEILSKAVEIE  
EQTKRLLEGMELIVSQVHPETKENEIYPVWVSLPSLQMADEESRLSAYYNLLHCLRRDHS  
KIDNYLKLKLCRIIHNNNC  
>sp|P01270|32-115  
SVSEIQLMHNLGKHLNSMERVEWLRKKLQDVHNFVALGAPLAPRDAGSQRPRKKEDNVLV  
ESHEKSLGEADKADVNLTKAKSQ  
>sp|P09683|28-54  
HSDGTFTSELSRLREGARLQRLQGLV  
>sp|P01286|32-75  
YADAIFTNSYRKVLGQLSARKLLQDIMSROQGESNQERGARARL  
>sp|P61278|89-116  
SANSNPAMAPRERKAGCKNFFWKTFTSC  
>sp|P61278|103-116  
AGCKNFFWKTFTSC  
>sp|P61812|21-302

LSTCSTLDMQDFMRKRIEAIHQILSKLKLTSPPEDYPEPEEVPPEVISIYNSTRDLLQE  
 KASRRAAACERERSDEEYYAKEVYKIDMPPFFPSENAIPPTFYRPFYFRIVRFDVSAMEKN  
 ASNLVKAEFRVFRQLQNPVKARVPEQRIELYQILKSKDLTSPTQRYIDSKVVKTRAEGEWLS  
 FDVTDVAHEWLHHKDRNLGFKISLHPCCTFVPSNNYIIPNKSEELEARFAGIDGTSTYT  
 SGDQKTIKSTRKKNNGKTPHLLMLLPSYRLESQQTNRKRR  
 >sp|P61812|303-414  
 ALDAAYCFRNVQDNCLRLPLYIDFKRDLGWKWIHEPKGYNANFCAGACPYLWSSDTQHSR  
 VLSLYNTINPEASAPCCVSDLEPLTILYYIGKTPKIEQLSNMIVKSCCKS  
 >sp|P01137|30-278  
 LSTCKTIDMELVKKRIEAIHQILSKLRLASPPSQGEVPPGGLPEAVLALYNSTRDRVA  
 GESAEPEPEPEADYYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVL  
 LSRAELRLLRLKLVKVEQHVELYQKYSNNSWRYLNSRLLAPSDSPEWLSFDVTGVVRQWLS  
 RGGEIEGFRLSAHSCDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQ  
 HLQSSRHR  
 >sp|P01137|279-390  
 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK  
 VLALYNQHNPGASAAPCCVPALEPLIVYVGRKPKVEQLSNMIVRSCCKS  
 >sp|P20396|25-242  
 QPEAAQQEAVTAAEHPGLDDFLRQVERLLFLRENIQRLQGDQGEHSASQIFQSDWLSKRQ  
 HPGKREEEEEEGVEEEEEEGGAVGPHKRQHPGRREDEASWSVDVTQHKRQHPGRRSPWL  
 AYAVPKRQHPGRRADPKAQRSWEEEEEEEEREEDLMPEKRQHPGKRALGGPCGPQAYG  
 QAGLLGLLDDLRSQGAEEKRQHPGRRAAWVREPLEE  
 >sp|P01375|57-233  
 GPQREEFPRLDLSISPLAQAVRSSRTPSDKPVAVHVVANPQAEGLQWLNRRANALLANG  
 VELRDNQLVVPSEGLYLIYSQVLFKQGQCPSTHVLLTHTISRIAVSYQTKVNLLSAIKSP  
 CQRETPEGAEAKPWYEPYIYLGGVFQLEKGDRLSAEINRPDYLDFAESGQVYFGIIAL  
 >sp|P40225|22-353  
 SPAPPACDRLVLSKLLRDSHVLHSRLSQCEVHPLPTPVLLPAVDFSLGEWKTQMEETKA  
 QDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGALQSLGTLPPQGRTTA  
 HKDPNAIFLSFQHLLRGKVRFLMLVGGSTLCVRRAPPTTAVPSRSLVLTNLNLPNRTSG  
 LLETNFTASARTTGSGLLKQWQGFRAKIPGLLNQTSRSLDQIPGYLNRIHELLNGTRGLF  
 PGPSRRTLGAADISSGTSDTGSLPPNLQPGYSPSPHPPTGQYTLFPLPPTLPTPVVQLH  
 PLLPDPSPAPTPTSPLLNNTSYTHSQNLSQEG  
 >sp|Q02747|22-115  
 VTVQDGNFSFSLESVKKLDLQEPQEPVGVKLRNFAPIPGEPVVPILCSNPNFPEELKPL  
 CKEPNAQEILQRLEEIAEDPGTCEICAYAAGTGC  
 >sp|Q02747|101-115  
 PGTCEICAYAAGTGC  
 >sp|Q16661|89-112  
 FKTLRTIANDDCELCVNVACTGCL  
 >sp|Q16661|97-112  
 NDDCELCVNVACTGCL  
 >sp|P81172|60-84  
 DTHFPICIFCCGCCHRSKCGMCKT  
 >sp|P81172|65-84  
 ICIFCCGCCHRSKCGMCKT  
 >sp|P10997|34-70  
 KCNTATCATQRLANFLVHSSNNFGAILSSSTNVGSNTY  
 >sp|P05019|49-118  
 GPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSRRAPQTGIVDECCFRSCDLRRLLEY  
 CAPLKPAKSA  
 >sp|P08476|311-426  
 GLECDGKVNICCKKQFFVSFKDIGWNDWIIAPSGYHANYCEGECPSHIAGTSGSSLSFHS  
 TVINHYMRGHSFPANLKSCCVPTKLRPMSMLYDDGQNI IKKDIQNMIVEECGCS  
 >sp|P05231|30-212  
 VPPGEDSKDVAAPHRQPLTSSERIDKQIRYILDGISALRKETCNKSNMCESSKEALAENN  
 LNLPKMAEKDGCFFQSGFNEETCLVKIITGLLEFEVYLEYLQNRFEESSEEQARAVQMSTKV

LIQFLQKKAKNLDAITTPDPTTNASLLTKLQAQNQWLQDMTTHLILRSFKEFLQSSLRAL  
RQM  
>sp|P01308|25-54  
FVNQHLCGSHLVEALYLVCGERGFFYTPKT  
>sp|P01308|90-110  
GIVEQCCTSICSLYQLENYCN  
>sp|P01344|25-91  
AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVSRRSRGIVEECCFRSCDLALLETYC  
ATPAKSE  
>sp|P41159|22-167  
VPIQKVQDDTKTLIKTIVTRINDISHTQSVSSKQKVTGLDFIPGLHPILTLSKMDQTLAV  
YQQILTSMPSRNVIQISNDLENLRDLLHVLAFSKSKCHLPWASGLETLDLGGVLEASGYS  
TEVVALSRLQGSLLQDMLWQLDLSPGC  
>sp|O43612|34-66  
QPLPDCCRQKTCSCRLYELLHGAGNHAAGILTL  
>sp|O43612|70-97  
RSGPPGLQGRLLQASGNHAAGILTM  
>sp|P02818|52-100  
YLYQWLGAPVYPDPLEPRREVCELNPDCDELADHIGFQEAAYRRFYGPV  
>sp|P18509|82-129  
DVAHGILNEAYRKVLDQLSAGKHLQSLVARGVGGSLGGGAGDDAEPLS  
>sp|P18509|132-169  
HSDGIFTDSYSRYRKQMAVKKYLAAVLGKRYKQRVKNK  
>sp|P18509|132-158  
HSDGIFTDSYSRYRKQMAVKKYLAAVL  
>sp|P01298|30-65  
APLEPVYPGDNATPEQMAQYAADLRRYINMLTRPRY  
>sp|P01298|69-88  
HKEDTLAFSEWGSPhAAVPR  
>sp|P01210|25-97  
ECSQDCATCSYRLVRPADINFLACVMECEGKLP SLKIWETCKELLQLSKPELPQDGTSTL  
RENSKPEESHLLA  
>sp|P01210|114-133  
MDELYPMEPEEEANGSEILA  
>sp|Q9HD89|19-108  
KTLCSMEEA INERI QEVAGSLIFRAISSIGLEQCQSVTSRGLATCPRGFAVTGCTCGSAC  
GSWDVRAETTCHCQCAGMDWTGARCCRVQP  
>sp|P04090|25-53  
DSWMEEVIKLCGRELVRAQIAICGMSTWS  
>sp|P04090|162-185  
QLYSALANKCCHVGCTKRSLARFC  
>sp|P04808|23-53  
VAAKWKDDVIKLCGRELVRAQIAICGMSTWS  
>sp|P04808|163-185  
PYVALFEKCLIGCTKRSLAKYC  
>sp|Q8WXF3|26-52  
RAAPYGVRLCGREFIRAVIFTCGGSRW  
>sp|Q8WXF3|119-142  
DVLAGLSSSCCKWGCSKSEISSL  
>sp|P02753|19-201  
ERDCRVSSFRVKENFDKARFSGTWYAMAKKDPEGLFLQDNIVAEFSVDETGQMSATAKGR  
VRLNNDVDCADMVGTFTDTEPAKFKMKYWGVAFLQKGNDDHWIVDTDYDYAVQYSC  
RLLNLDGTCADSYFVFSRDPNGLPPEAQKIVRQRQEELCLARQYRLIVHNGYCDGRSER  
NLL  
>sp|P00797|67-406  
LTLGNTTSSVILTNYMDTQYYGEIGIGTPPQTfKVVFDTGSSNVVWPSSKCSRLYTACVY  
HKLFDASDSSSYKHNGTELTLYSTGTVSGFLSQDIITVGGITVTQMFGEVTEMPALPFM

LAEFDGVVGMGFIEQAIGRVTPIFDNIISQGVLKEDVFSFYNNRDSSENSQSLGGQIVLGG  
SDPQHIEGDFHYINLIKTVWQIQMKGVSVGSSTLLCEDGCLALVDTGASYISGSTSSIE  
KLMEALGAKKRLFDYVVKCNEGPTLPDISFHLGGKEYTLTSADYVFQESYSSKKLCTLAI  
HAMDIPPPTGPTWALGATFIRKFYTEFDRRNNRIGFALAR  
>sp|P01222|21-132  
FCIPTEYTMHIERRECAVCLTINTTICAGYCMTRDINGKLFLPKYALSQDVCTYRDFIYR  
TVEIPGCPLHVAPYFSYPVALSCKCGKCNTDYSDCIHEAIKTNCTKPKQSY  
>sp|P01282|81-122  
HADGVFTSDFSKLLGQLSAKKYLESLMGKRVSSNISEDPVPV  
>sp|P01282|81-107  
HADGVFTSDFSKLLGQLSAKKYLESLM  
>sp|P01282|125-152  
HSDAVFTDNYTRLRKQMAVKKYLNSILN  
>sp|P01178|NEU1\_HUMAN Oxytocin OS=Homo sapiens GN=OXT PE=1 SV=1  
CYIQNCPLG  
>>sp|P01185|NEU2\_HUMAN Arg-vasopressin 20 - 28, OS=Homo sapiens GN=AVP PE=1  
SV=2  
CYFQNCPRG  
>sp|P01210|100-104 **Met-enkephalin**  
YGGFM  
>sp|P01210|107-111 **Met-enkephalin**  
YGGFM  
>sp|P01148|24-33 **Gonadoliberin-11**  
QHWSYGLRPG