

## Supplementary Materials

### Novel BRICHOS-related antimicrobial peptides from the marine worm *Heteromastus filiformis*: transcriptome mining, synthesis, biological activities, and therapeutic potential

Pavel V. Panteleev<sup>1</sup>, Victoria N. Safronova<sup>1</sup>, Shuting Duan<sup>1,2</sup>, Alexey S. Komlev<sup>3</sup>, Ilia A. Bolosov<sup>1</sup>, Roman N. Kruglikov<sup>1</sup>, Tatiana I. Kombarova<sup>4</sup>, Olga V. Korobova<sup>4</sup>, Eugenia S. Pereskokova<sup>4</sup>, Alexander I. Borzilov<sup>4</sup>, Igor A. Dyachenko<sup>5</sup>, Olga V. Shamova<sup>3</sup>, Yu Huang<sup>2</sup>, Qiong Shi<sup>2,6</sup> and Tatiana V. Ovchinnikova<sup>1,7\*</sup>

<sup>1</sup> M.M. Shemyakin & Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117997 Moscow, Russia; p.v.panteleev@gmail.com (P.V.P.); victoria.saf@ibch.ru (V.N.S.); bolosov@ibch.ru (I.A.B.); kruglikov1911@mail.ru (R.N.K.); ovch@ibch.ru (T.V.O.);

<sup>2</sup> Shenzhen Key Lab of Marine Genomics, Guangdong Provincial Key Lab of Molecular Breeding in Marine Economic Animals, BGI Academy of Marine Sciences, BGI Marine, 518081 Shenzhen, China; huangyu@genomics.cn (Y.H.), shiqiong@genomics.cn (Q.S.);

<sup>3</sup> Institute of Experimental Medicine, WCRC "Center for Personalized Medicine" 197022 St. Petersburg, Russia; komlev1420@yandex.ru (A.S.K.), oshamova@yandex.ru (O.V.S);

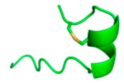
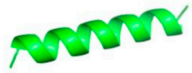
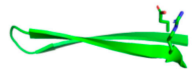


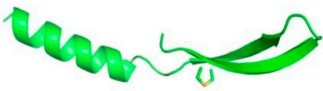

<sup>4</sup> State Research Center for Applied Microbiology & Biotechnology (SRCAMB), 142279 Obolensk, Russia; kombarova@obolensk.org (T.I.K.); korobova@obolensk.org (O.V.K.); pereskokova@obolensk.org (E.S.P.), borzilov@obolensk.org (A.I.B.);

<sup>5</sup> The Branch of M.M. Shemyakin & Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 142290 Pushchino, Russia; dyachenko@bibch.ru (I.A.D.);

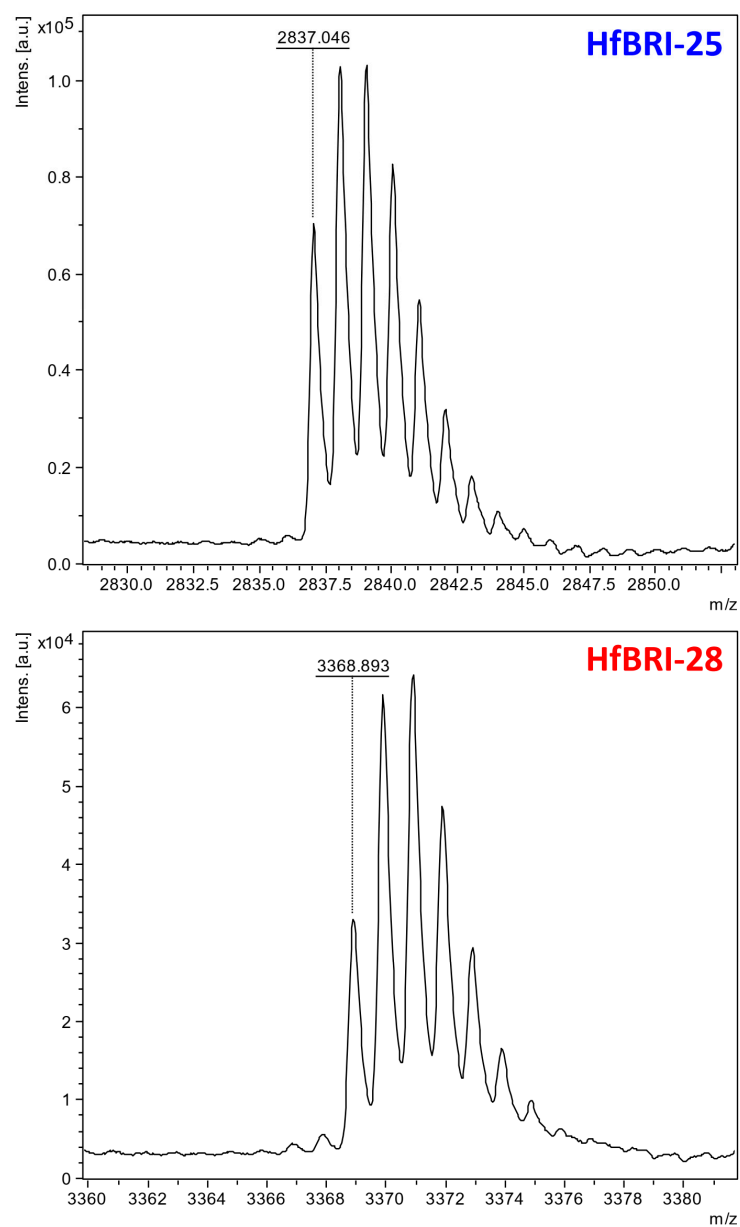
<sup>6</sup> Laboratory of Aquatic Genomics, College of Life Sciences and Oceanography, Shenzhen University, 518057 Shenzhen, China; shiqiong@szu.edu.cn (Q.S.).

<sup>7</sup> Department of Biotechnology, I.M. Sechenov First Moscow State Medical University, 119991 Moscow, Russia

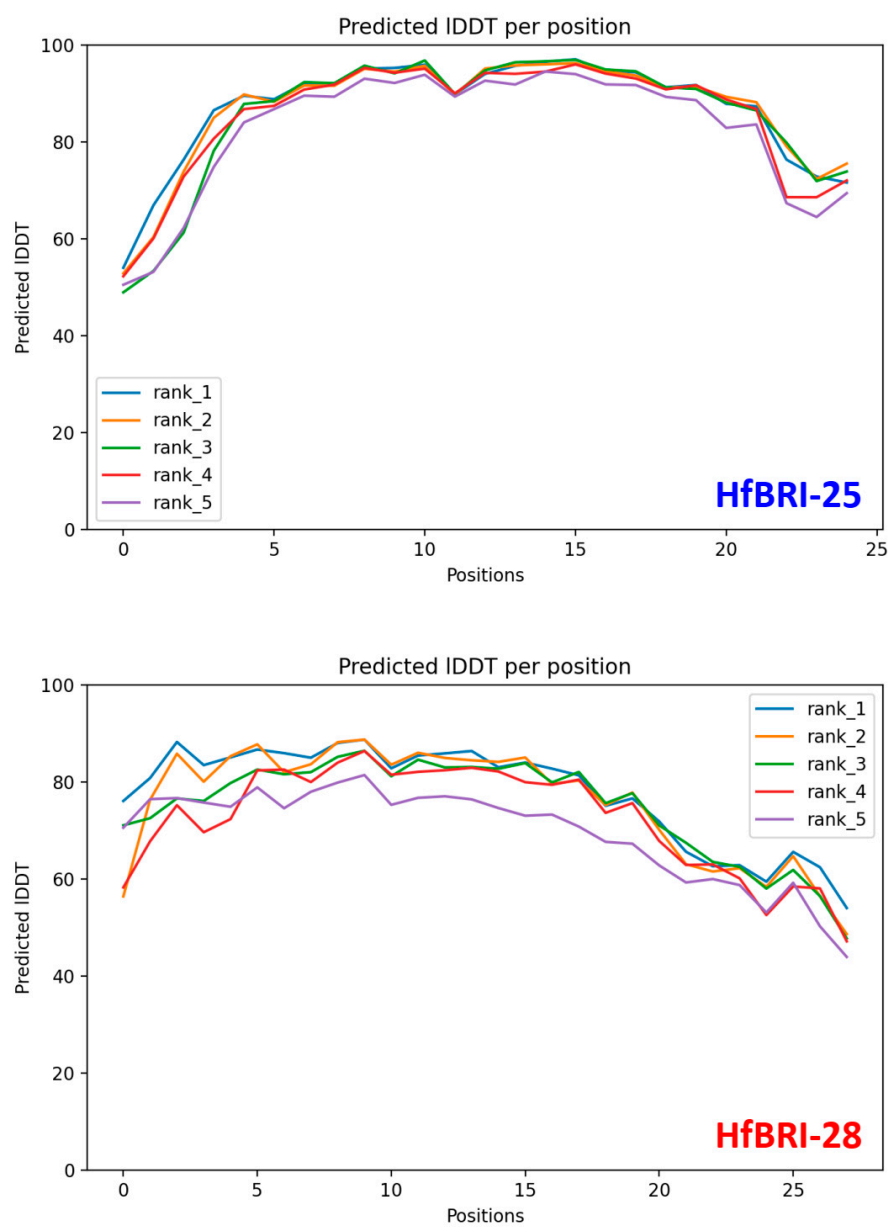
\* Correspondence: ovch@ibch.ru

Peptide	Sequence	Predicted spatial structure	Predicted fold type
HfBRI-18	GRRFISISIGA <b>C</b> VFG <b>L</b> C <b>F</b>		"Rana-box"-like (one disulfide bond)
HfBRI-21	EAKTRTKRHI <b>C</b> ILSI <b>C</b> AWRRR		linear $\alpha$ -helical
HfBRI-20	<b>G</b> ISVSVRVGRVKVTLR <b>I</b> R <b>W</b>		$\beta$ -hairpin
HfBRI-23	<b>R</b> PSI <b>I</b> VTVTVTFHPLRIRV <b>R</b> L		$\beta$ -hairpin
HfBRI-29	RVPR <b>C</b> GKYQ <b>E</b> K <b>C</b> <b>I</b> TR <b>C</b> TKVFLIEL <b>C</b> IRTD		$\beta$ -hairpin + loop (two disulfide bonds)
HfBRI-38a	GWGWVKKTLKKVWVVKKP V <b>C</b> YWT <b>T</b> STYGIPKKV <b>C</b> ILV		$\alpha$ -helix + $\beta$ -hairpin (one disulfide bond)
HfBRI-54	SMD <b>R</b> SM <b>C</b> VW <b>T</b> CRGR <b>F</b> CRW <b>V</b> CP <b>R</b> QTPGPTTAAATPTATPTAT TAG <b>C</b> SNE <b>M</b> VE		$\beta$ -hairpin + random coiled Thr-rich part (two disulfide bonds)
HfBRI-75	G <b>C</b> PIDQRSYR <b>C</b> KVTA <b>K</b> T <b>C</b> QYLV S <b>C</b> DSVAIPGEPGR <b>T</b> AAN <b>C</b> IKHI RGNV <b>K</b> <b>C</b> NE <b>C</b> <b>C</b> KNRQLANEVFP H <b>C</b> DSVS <b>V</b> E <b>C</b>	n.d.	cysteine-rich
HfBRI-91	DTFDPK <b>C</b> PAAQVKYR <b>C</b> KKELLS <b>C</b> QY L <b>Y</b> YET <b>C</b> DAERYRRGRMYAK <b>C</b> LDNS TLHIID <b>E</b> V <b>K</b> <b>C</b> <b>C</b> RE <b>C</b> ADQSQADDF <b>F</b> P <b>H</b> <b>C</b> ETFG <b>E</b> FT <b>C</b> THDMQ	n.d.	cysteine-rich

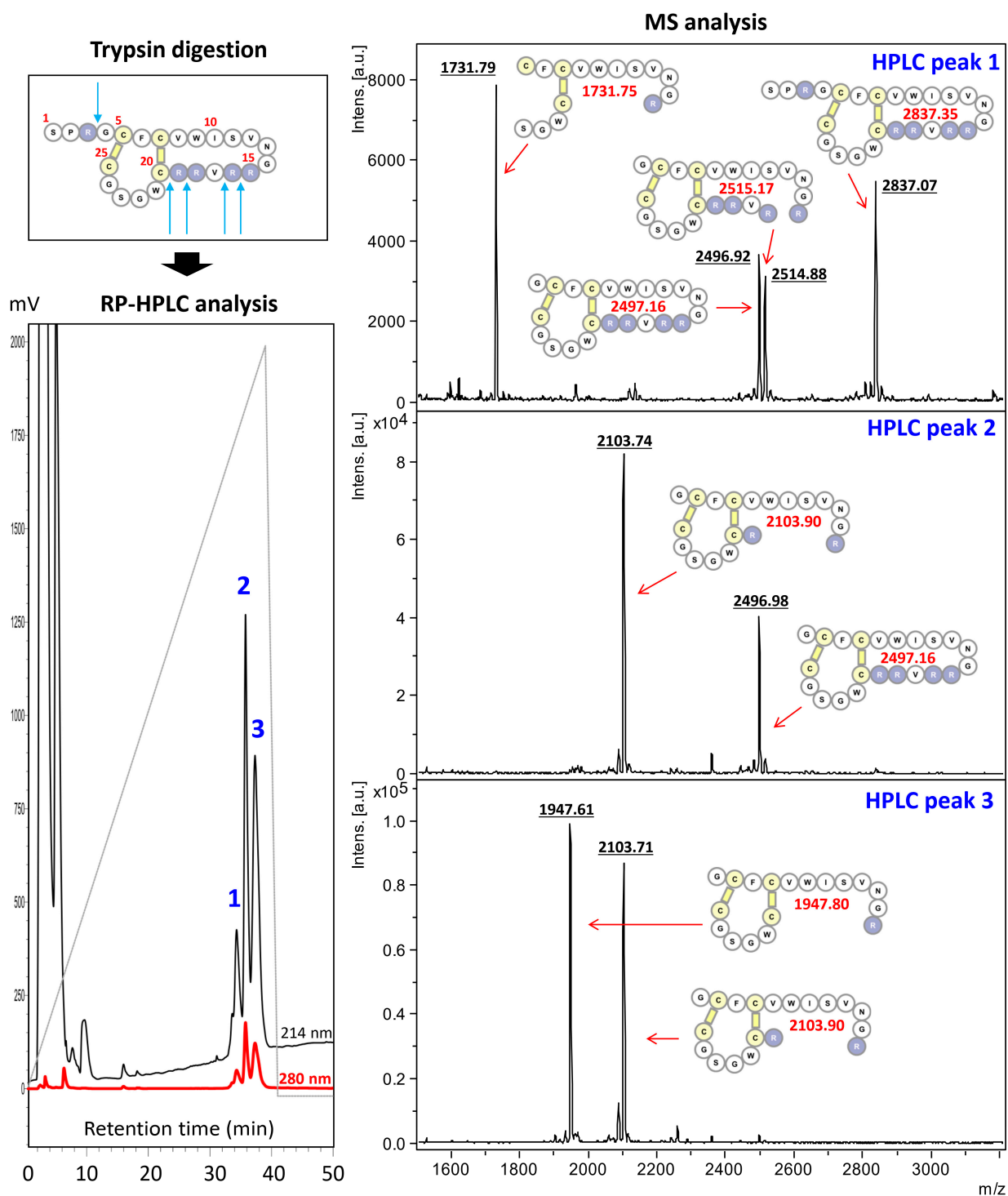
**Figure S1.** Structure analysis of those novel BRICHOS-related peptides from *H. filiformis*. Spatial structures of the peptides were modelled using the AlphaFold algorithm (ColabFold server <https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>). Accuracy of the predicted structures (predicted local distance difference test (pLDDT)) is not indicated. The top ranked models, according to pLDDT scores, were visualized by the PyMOL software and presented as predicted spatial structures.  $\beta$ -Hairpin motif sequences are marked with bold.



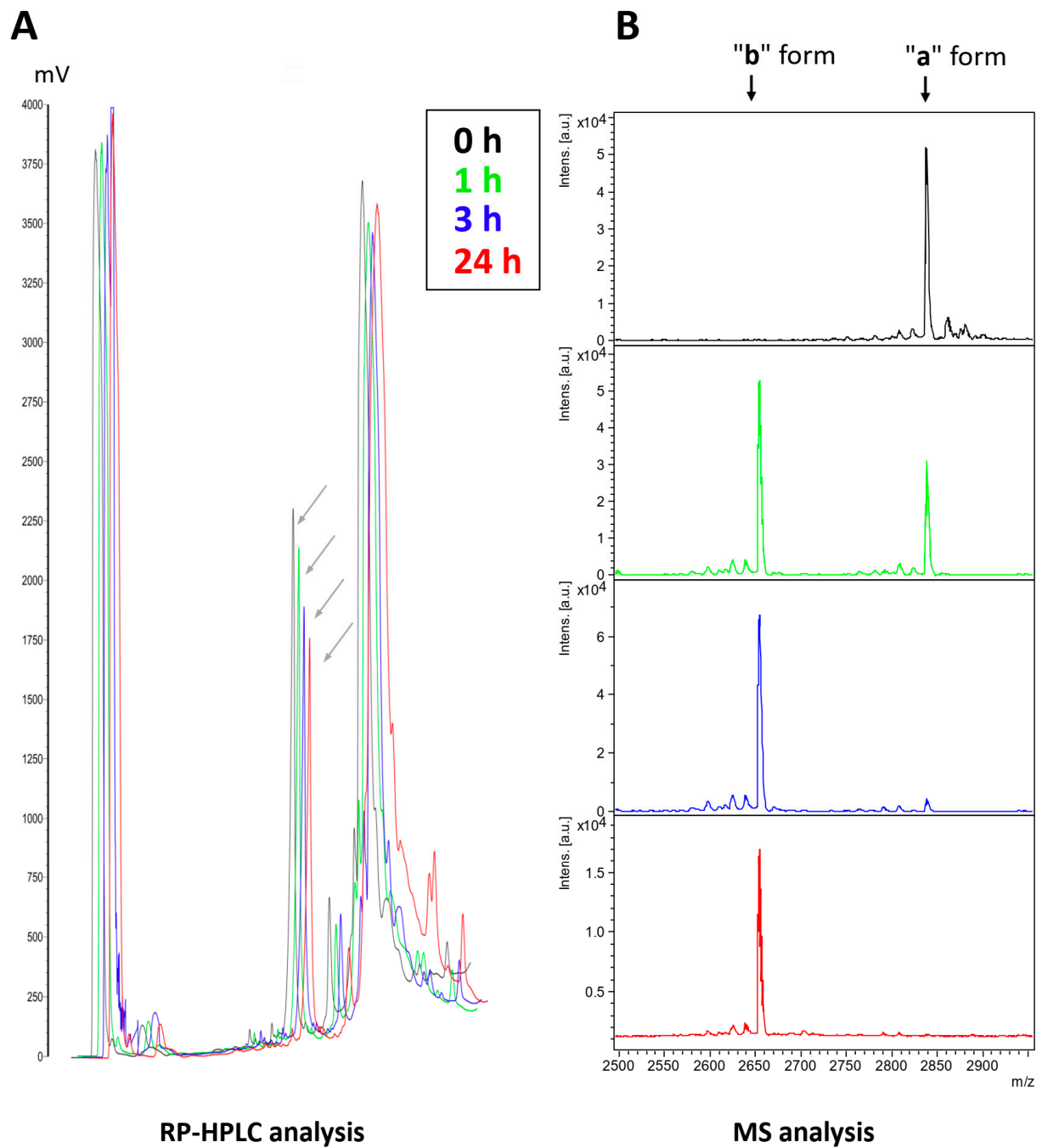
**Figure S2.** MALDI-TOF MS analyses of the obtained peptides HfBRI-25 and HfBRI-28.



**Figure S3.** Accuracy of the predicted structures (pLDDT) for both peptides HfBRI-25 and HfBRI-28.



**Figure S4.** RP-HPLC and MALDI-TOF MS analyses of the tryptic digests of HfBRI-25. The experimentally measured monoisotopic  $[M+H]^+$  values of the major RP-HPLC fractions pointed by the numbers (colored in blue) matched well the corresponding calculated values (colored in red).



**Figure S5.** Stability of HfBRI-25 in serum (RP-HPLC and MS data). **(A)** RP-HPLC profiles (Abs 214 nm) of fractions containing HfBRI-25 and its shortened 23-residue form after 0, 1, 3, and 24 h incubation in 25% buffered human serum. Fractions containing mixed target peptides (HfBRI-25 ("a" form) and its shortened 23-residue "b" form) are marked with grey arrows. **(B)** MALDI-TOF MS analysis of the target RP-HPLC peaks containing HfBRI-25 and its shortened 23-residue form after 0, 1, 3, and 24 h incubation in 25% buffered human serum.

**Table S1.** A list of predicted non-BRICHOS-related AMPs from the *H. filiformis* RNA-seq data

Predicted AMP #	Sequence
1	LISGMRGLGVEIAKNVILAGVKSVTIHDTGN
2	GRVEKFIQKSERRIFTNFH
3	YVPKVVGlyTVTVTYGGKPIKKSPFKVEVSLAFDSSKVYATGRG
4	LDIWDKVHEALHTLIKLAKKGLTAKHS
5	LVTNPKRVQTGIDKKACNCLLLKVN
6	LWEIGKIIGQMWR
7	RALVTAVHSVPGVASLPAGSIKAALTVVQRYPQCLGYNWGP
8	RHIKIWFQNRRMKWKK
9	LSVKVAPTIFGTVTK
10	EAAPLARGASNLSKRASITTKASTLTCKA
11	RGASNLSKRASITAKKASTLTCKASLAKQA
12	LVMERKWMIRARHIAGK
13	FTKNCIVLFPGHPIVHGRENARGFRNLDTLF
14	AVTPKVATKDSKKPS
15	GKHITGLQMGSNKGASQAGTGFRPRQV
16	RKYRNAHKKRLQKKEYRNLOKIVPSVAEKS
17	EHGRRLERLAKGFFPSSEKACPAFLRHK
18	KLLRLQRKIAQQKRKKY
19	LSRNYRKRIERLKKLKKKRPVVRW
20	GLARGLHECAKALDK
21	VKLKRKAKALISRMKSNISK
22	LAGVTKSLQDLASRFAGL
23	MAKTILKRLFRVYAHY
24	GSGPVRSSCLRGSSHSPSAVHHHCISEAIHKQ
25	IRGKSQSVKSPVKGKLPSIRTSPVKGETSYVKNNP
26	FFQTSWNKLGNEEDRQCLKGMKTKGARFGTAL
27	GFGSALFSIGRKLIPKILP
28	TTILQSVGKQFAKNIFGILQSIKAREEGS
29	RLLTRIVKFIWRFVRRLL
30	GCRQALGRARRRLRNLRW
31	IKVGINGFGRIGRLV
32	EDLKKIIQKIKSHYK