File S1

The main characteristic peptides in Ejiao

|  |  |  |
| --- | --- | --- |
| NO | Peptide sequence | SMLES |
| 1 | GPAGPQGPR | [H]NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCC(N)=O)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 2 | GFSGLDGAK | [H]NCC(=O)N[C@@H](CC1=CC=CC=C1)C(=O)N[C@@H](CO)C(=O)NCC(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](CC(O)=O)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](CCCCN)C(O)=O |
| 3 | GVVGPQGAR | [H]NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCC(N)=O)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 4 | GPAGPTGPVGKDGR | [H]NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H]([C@@H](C)O)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N[C@@H](CCCCN)C(O)=O |
| 5 | GPAGPTGPVGKDGR | [H]NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H]([C@@H](C)O)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N[C@@H](CCCCN)C(=O)N[C@@H](CC(O)=O)C(=O)NCC(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 6 | AGVMGPAGSR | [H]N[C@@H](C)C(=O)NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCSC)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](CO)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 7 | SGQPGTVGPAGVR | [H]N[C@@H](CO)C(=O)NCC(=O)N[C@@H](CCC(N)=O)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N[C@@H]([C@@H](C)O)C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 8 | GPSGPQGVR | [H]NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CO)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCC(N)=O)C(=O)NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 9 | SGDRGEAGPA  GPAGPIGPVG | [H]N[C@@H](CO)C(=O)NCC(=O)N[C@@H](CC(O)=O)C(=O)N[C@@H](CCCNC(N)=N)C(=O)NCC(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](C)C(=O)NCC(=O)N  1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H]([C@@H](C)CC)C(=O) |
| 10 | GSEGPQGVR | [H]NCC(=O)N[C@@NHC](CC(O=O)C)N(=1OC)CNC[C[C@@@HH]1](CC(C=OC()ON)[=CO@)@C(H=O](C)N(CC)CC()=CO(=)NO1)CNCCCC[(C=O@)H  N][1CC@(=@OH)N]([CC)@C(@=OH)]N(C[CC@C(@N)H=]O(C)CC(C=NOC)N(NC)C=(N=O)C)(NO[)C=@O@ H](C(C)C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 11 | GHAGLAGAR | [H]NCC(=O)N[C@@H](CC1=CNC=N1)C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 12 | GEAGPAGPAG  PIGPVGAR | [H]NCC(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H  ]1C(=O)N[C@@H]([C@@H](C)CC)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@ |
| 13 | PGPIGPAGAR | [H]N1CCC[C@H]1C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H]@([HC@](C@CHC]N(CC)(CNC)=)CN()=CO(O)N)=COC (=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 14 | GEAGPQGAR | [H]NCC(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCC(N)=O)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 15 | GPAGPSGPPGK | [H]NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CO)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N1CCC[C@H]1C(=O)NCC(=O)N[C@@H](CCCCN)C(O)=O |
| 16 | GRPGAPGPAGAR | [H]NCC(=O)N[C@@H](CCCNC(N)=N)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N[C@@H](C)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 17 | GVQGPPGPAGPR | [H]NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCC(N)=O)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N1CCC[C@H]1C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 18 | GDQGPVGR | [H]NCC(=O)N[C@@H](CC(O)=O)C(=O)N[C@@H](CCC(N)=O)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 19 | GDRGDAGPK | [H]NCC(=O)N[C@@H](CC(O)=O)C(=O)N[C@@H](CCCNC(N)=N)C(=O)NCC(=O)N[C@@H](CC(O)=O)C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCCCN)C(O)=O |
| 20 | VDAASLR | [H]N[C@@H](C(C)C)C(=O)N[C@@H](CC(O)=O)C(=O)N[C@@H](C)C(=O)N[C@@H](C)C(=O)N[C@@H](CO)C(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 21 | GVVGLPGQR | [H]NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N[C@@H](CC(C)C)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N[C@@H](CCC(N)=O)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 22 | SAGISVPGPM  GPSGPR | [H]N[C@@H](CO)C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H]([C@@H](C)CC)C(=O)N[C@@H](CO)C(=O)N[C@@H](C(C)C)C(=O)N1CCC[C@H]1C(=O)NC  C(=O)N1CCC[C@H]1C(=O)N[C@@H](CCSC)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CO)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCCNC(N)= |
| 23 | QGPSGASGER | [H]N[C@@H](CCC(N)=O)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](NCO)C)(CO(=)=OO)N CC(=O)N[C@@H](C)C(=O)N[C@@H](CO)C(=O)NCC(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 24 | GADGSPGKDG  VR | [H]NCC(=O)N[C@@H](C)C(=O)N[C@@H](CC(O)=O)C(=O)NCC(=O)N[C@@H](CO)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N[C@@H](CCCCN)C(=O)N[C@@H](CC(O)=O)C(=O)NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 25 | GQAGVMGFP  GPK | [H]NCC(=O)N[C@@H](CCC(N)=O)C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCSC)C(=O)NCC(=O)N[C@@H](CC1=CC=CC=C1)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCCCN)C(O)=O |
| 26 | ISPGAFTPLVK | [H]N[C@@H]([C@@H](C)CC)C(=O)N[C@@H](CO)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](CC1=CC=CC=C1)C(=O)N[C@@H]([C@@H](C)O)C(=O)N1CCC[C@H]1C(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCCCN)C(O)=O |
| 27 | LDEAEQLALK | [H]N[C@@H](CC(C)C)C(=O)N[C@@H](CC(O)=O)C(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](C)C(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](CCC(N)=O)C(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](C)C(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](CCCCN)C(O)=O |
| 28 | GEAGAAGPAG  PAGPR | [H]NCC(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |
| 29 | GVPGPPGAVGPAGK | [H]NCC(=O)N[C@@H](C(C)C)C(=O)N1CCC[C@H]1C(=O)NCC(=O)N1CCC[C@H]1C(=O)N1CCC[C@H]1C(=O)NCC(=O)N[C@@H](C)C(=O)N[C@@H](C(C)C)C(=O)NCC(=O)N1CCC[C@H]1C(=O)N[C@@H](C)C(=O)NCC(=O)N[C@@H](CCCCN)C(O)=O |
| 30 | QLDGILGER | [H]N[C@@H](CCC(N)=O)C(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](CC(O)=O)C(=O)NCC(=O)N[C@@H]([C@@H](C)CC)C(=O)N[C@@H](CC(C)C)C(=O)NCC(=O)N[C@@H](CCC(O)=O)C(=O)N[C@@H](CCCNC(N)=N)C(O)=O |