**Supplementary information**

**Hiding in plain sight: three chemically distinct -helix types**

Shuguang Zhang1, Martin Egli2

1Laboratory of Molecular Architecture, Media Lab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, 02139, USA

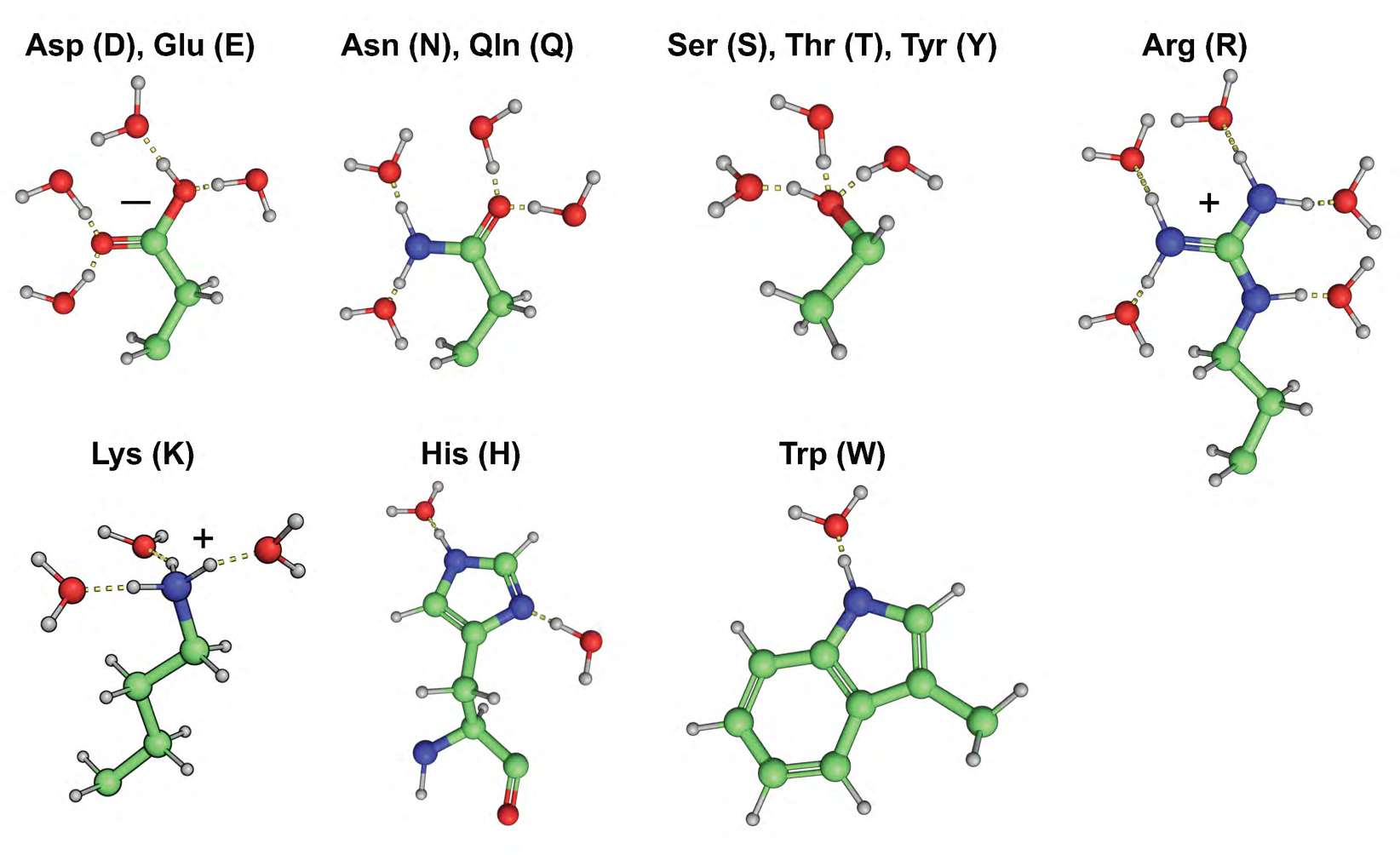
2Department of Biochemistry, Vanderbilt University, School of Medicine, Nashville, Tennessee 37232-0146, USA

Shuguang Zhang’s Email: [Shuguang@mit.edu](mailto:Shuguang@mit.edu)

Martin Egli’s Email: [martin.egli@vanderbilt.edu](mailto:martin.egli@vanderbilt.edu)

Shuguang Zhang [ORCID: 0000-0002-3856-3752](https://orcid.org/0000-0002-3856-3752)

Martin Egli [ORCID: 0000-0003-4145-356X](https://orcid.org/0000-0003-4145-356X)



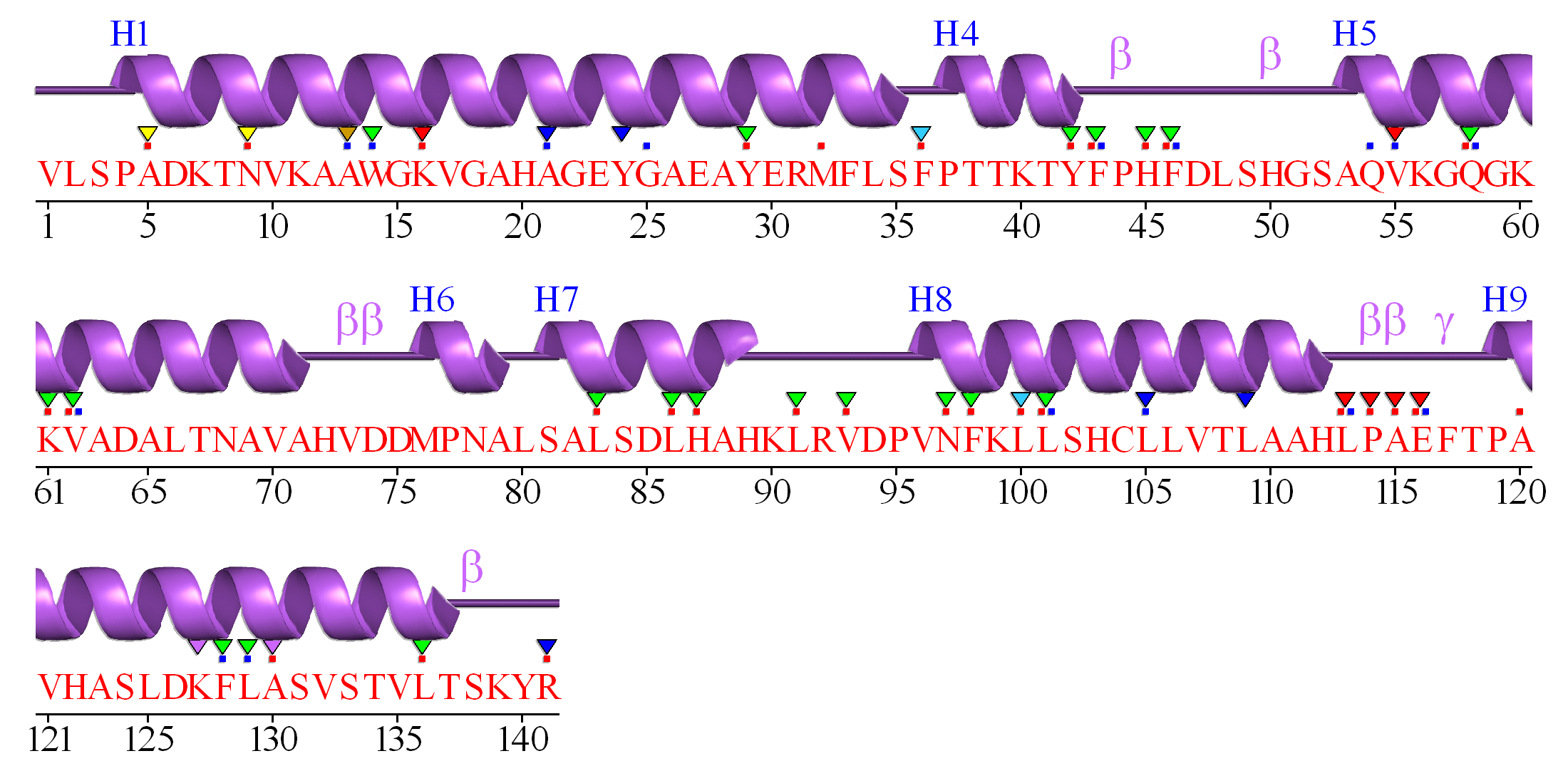
**Figure S1. Seven amino acids that form hydrogen bonds.** Asparatic acid and Glutamic acid each have 4 hydrogen acceptors from the side chain’s oxygen molecules. Asparagine and Glutamine each have 2 hydrogen donors from nitrogen and 2 hydrogen acceptors from oxygen. The –OH group of Serine (S), threonine (T) and tyrosine (Y) can form 3 hydrogen bonds: 2 H-acceptors and 1 H-donor. Arginine has 5 hydrogen donors, Lysine has 3 hydrogen donors, Histidine has 1 donor and 1 acceptor, Tryptophan has 1 hydrogen donor.

**Hydrophilic class: Water-soluble proteins**

Hemoglobin subunit alpha= HBA1, [P69905](https://www.uniprot.org/uniprot/P69905), [1BAB](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1bab&template=protein.html&r=wiring&l=1&chain=A) (1.5Å), [2W72](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=2w72&template=protein.html&r=wiring&l=1&chain=A) (1.07Å), (142aa)

>sp|[P69905](https://www.uniprot.org/uniprot/P69905)|HBA HUMAN Hemoglobin subunit alpha

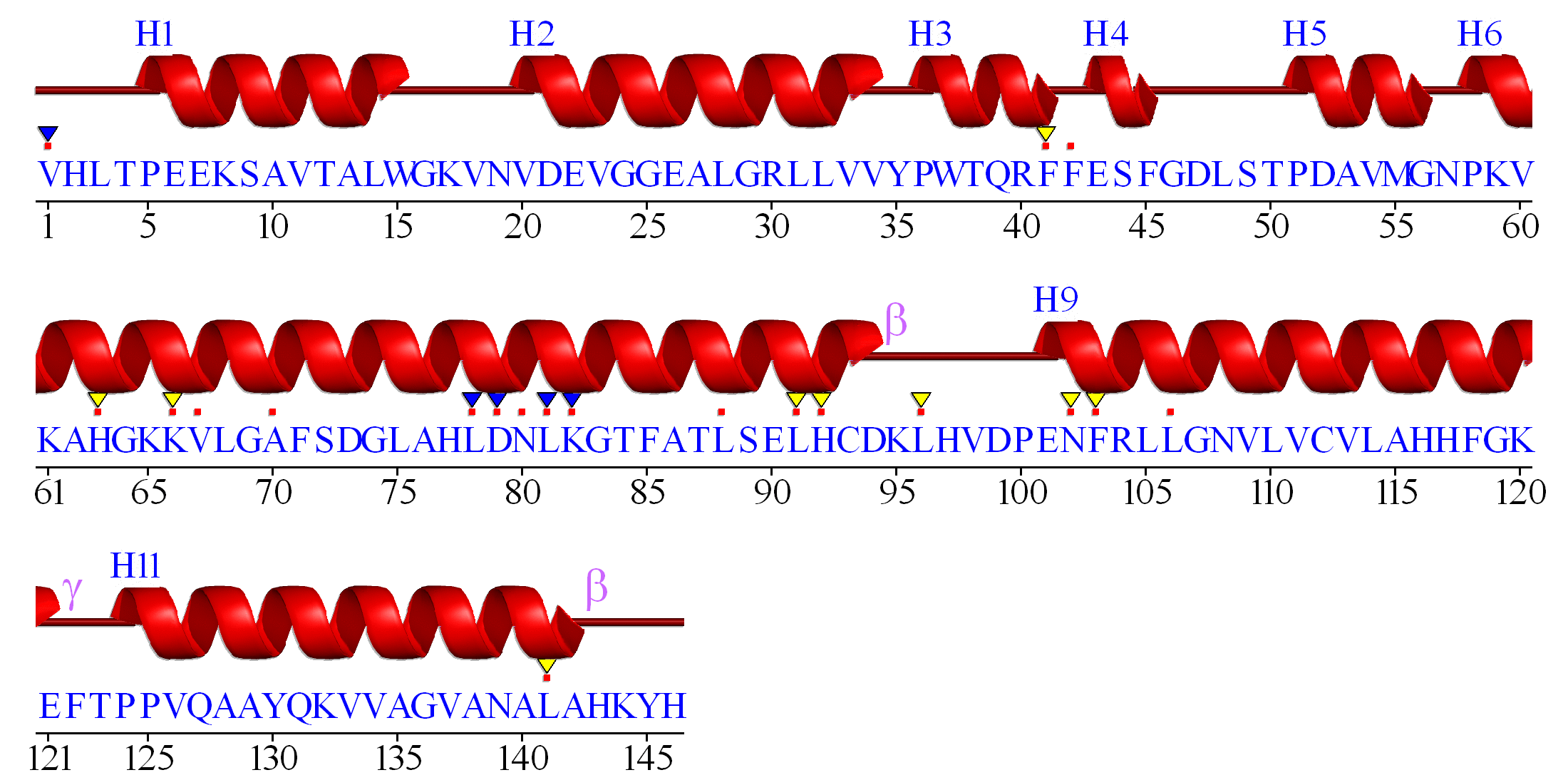
MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR (142aa)

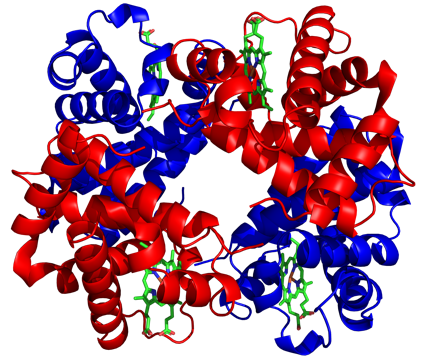
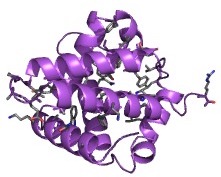
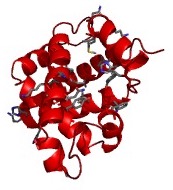


Hemoglobin subunit beta=HBB, [P68871](https://www.uniprot.org/uniprot/P68871), [1BAB](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1bab&template=protein.html&r=wiring&l=1&chain=A) (1.5Å) 1-147aa

>sp|[P68871](https://www.uniprot.org/uniprot/P68871)|HBB HUMAN Hemoglobin subunit beta

MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH (147aa)



**Hemoglobin (1.07Å) Chain A Chain B**

ZUO1 YEAST Zuotin, [P32527](https://www.uniprot.org/uniprot/P32527)

>sp|[P32527](https://www.uniprot.org/uniprot/P32527)|ZUO1 YEAST Zuotin OS=Saccharomyces cerevisiae, 433aa (Type I: H5)

MFSLPTLTSDITVEVNSSATKTPFVRRPVEPVGKFFLQHAQRTLRNHTWSEFERIEAEKN

VKTVDESNVDPDELLFDTELADEDLLTHDARDWKTADLYAAMGLSKLRFRATESQIIKAH

RKQVVKYHPDKQSAAGGSLDQDGFFKIIQKAFETLTDSNKRAQYDSCDFVADVPPPKKGT

DYDFYEAWGPVFEAEARFSKKTPIPSLGNKDSSKKEVEQFYAFWHRFDSWRTFEFLDEDV

PDDSSNRDHKRYIERKNKAARDKKKTADNARLVKLVERAVSEDPRIKMFKEEEKKEKERR

KWEREAGARAEAEAKAKAEAEAKAKAESEAKANASAKADKKKAKEAAKAAKKKNKRAIRN

SAKEADYFGDADKATTIDEQVGLIVDSLNDEELVSTADKIKANAAGAKEVLKESAKTIVD

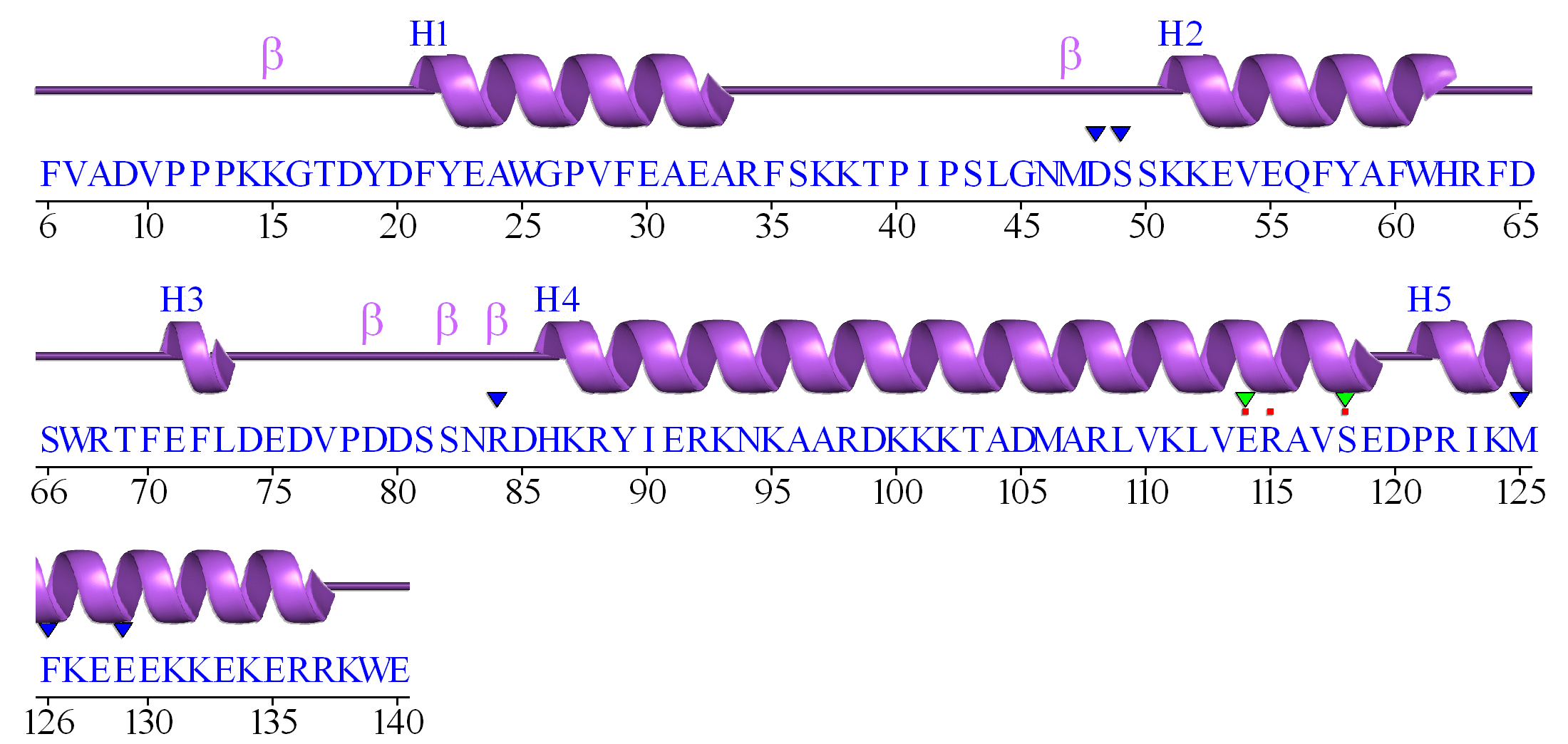
SGKLPSSLLSYFV (The yellow sequence part forms beta-sheet)

Zuotin (166-303aa=135aa)

RIKMFKEEEKKEKERR (3/16=18.75% hydrophobic, 81.25% hydrophilic

FKEEEKKEKERR (1/12=8.3% hydrophobic, 91.7% hydrophilic

HKRYIERKNKAARDKKKT (3/18=16.6% hydrophobic, 83.4% hydrophilic)



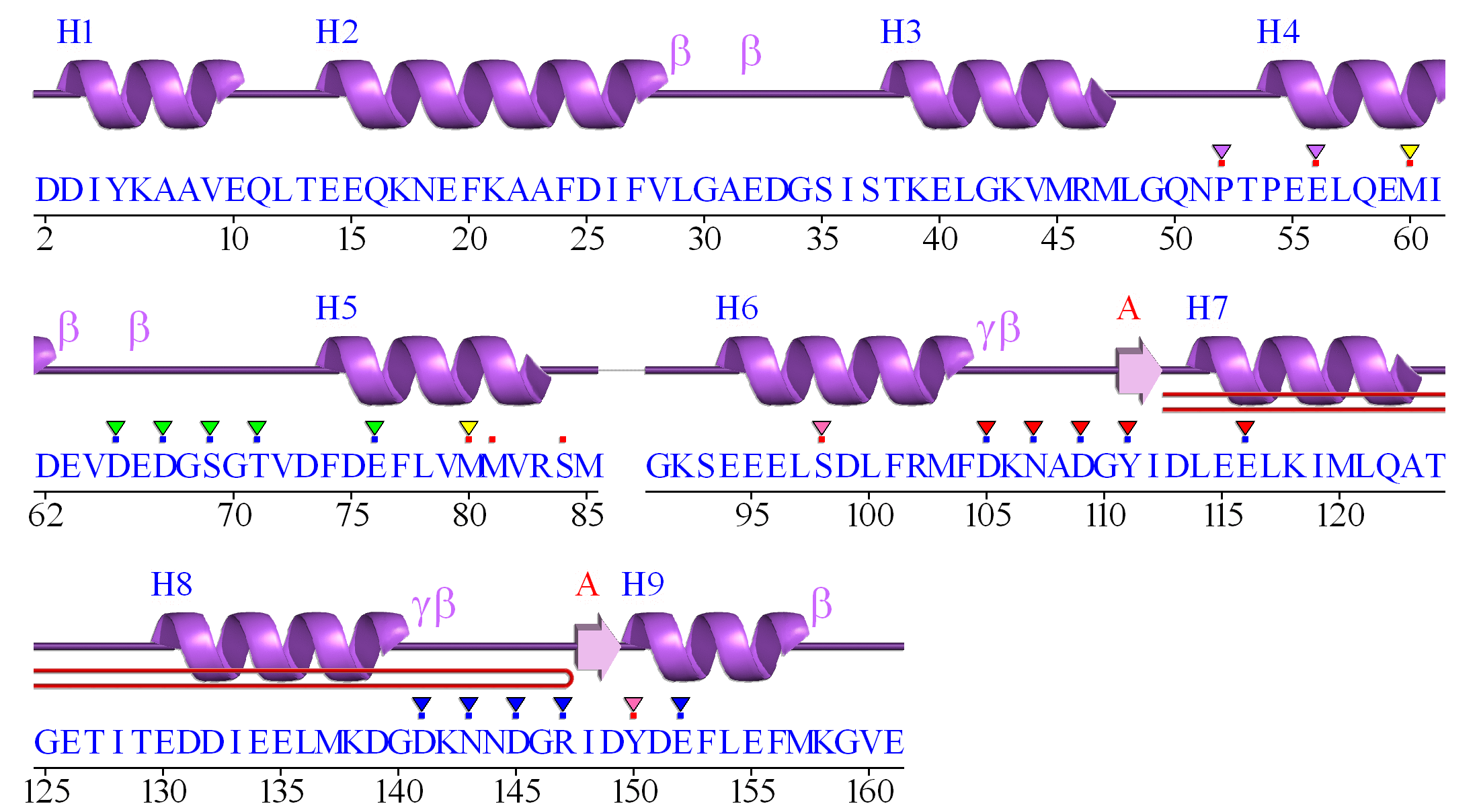
Troponin C= TNNC1, [P63316](https://www.uniprot.org/uniprot/P63316), [4Y99](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=4y99&template=protein.html&r=wiring&l=1&chain=A), 161aa (Type I: H2, H6, H8 )

>sp|P63316|TNNC1, HUMAN Troponin C, slow skeletal and cardiac muscles

MDDIYKAAVEQLTEEQKNEFKAAFDIFVLGAEDGCISTKELGKVMRMLGQNPTPEELQEM

IDEVDEDGSGTVDFDEFLVMMVRCMKDDSKGKSEEELSDLFRMFDKNADGYIDLDELKIM

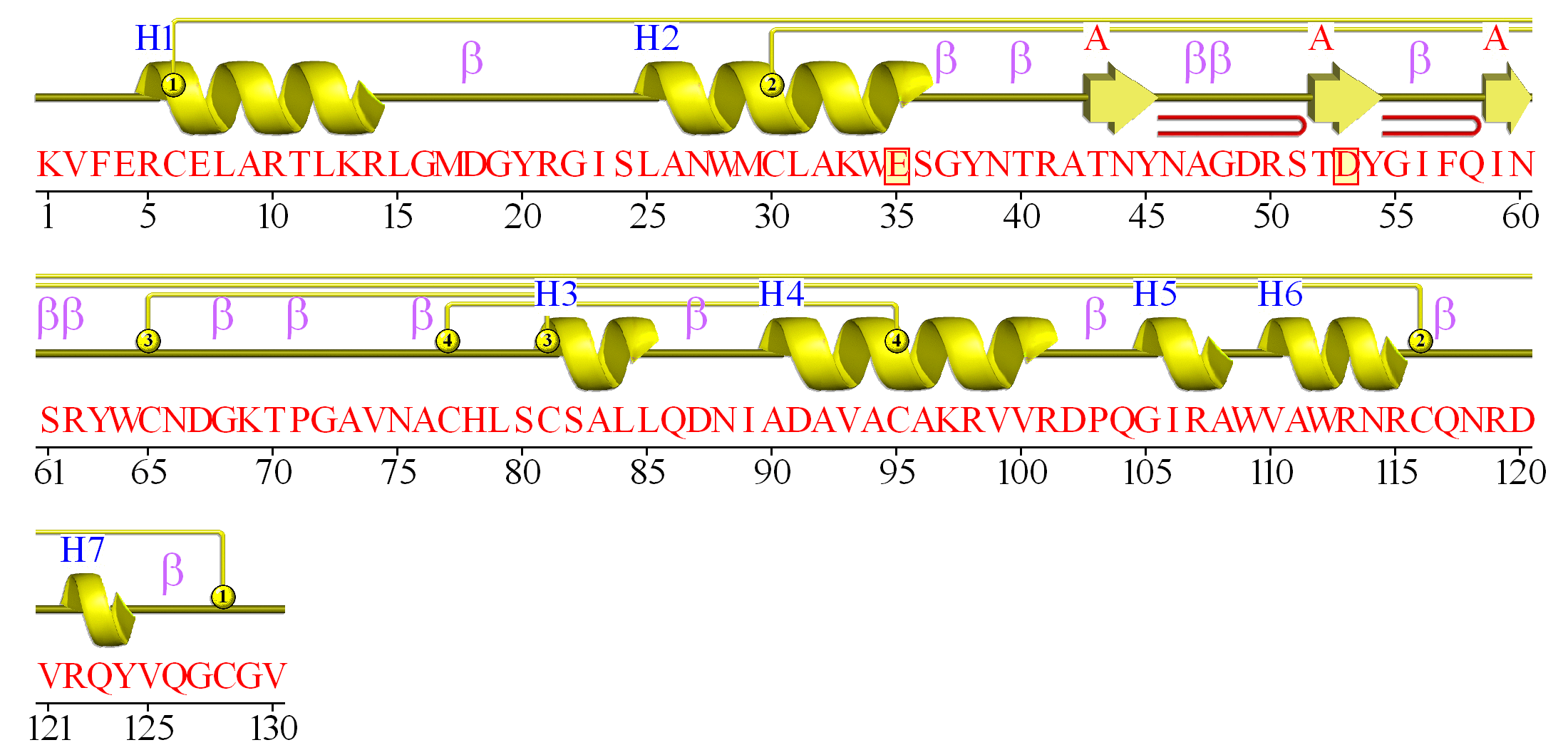
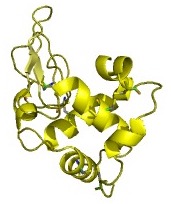
LQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVE



Human Lysozyme C=Lyz, [P61626](https://www.uniprot.org/uniprot/P61626), [2NWD](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=2nwd&template=protein.html&r=wiring&l=1&chain=X), 1.04Å, (19-148aa)

>sp|[P61626](https://www.uniprot.org/uniprot/P61626)|LYSC HUMAN Lysozyme C

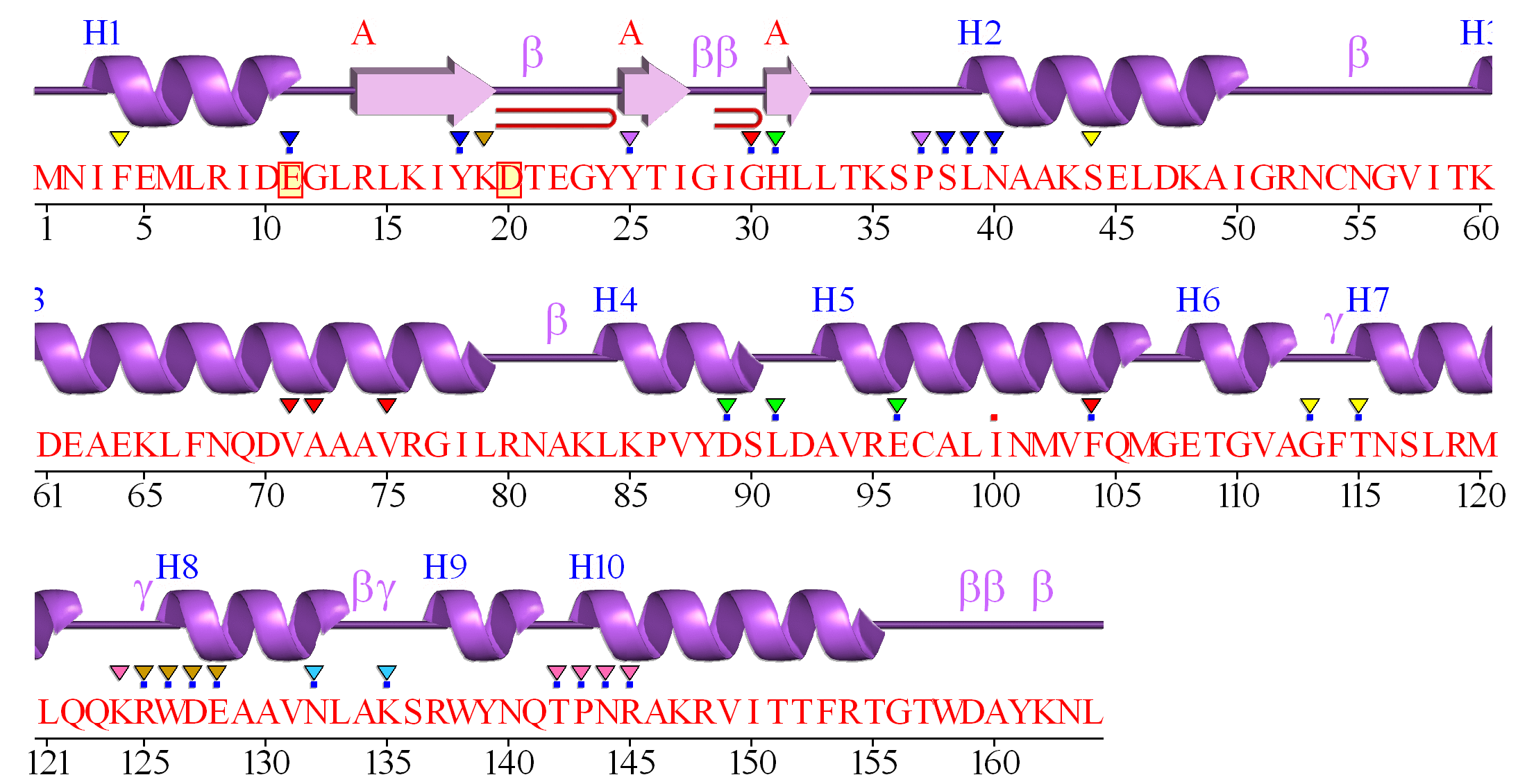
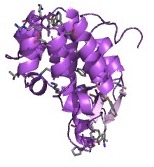
MKALIVLGLVLLSVTVQGKVFERCELARTLKRLGMDGYRGISLANWMCLAKWESGYNTRATNYNAGDRSTDYGIFQINSRYWCNDGKTPGAVNACHLSCSALLQDNIADAVACAKRVVRDPQGIRAWVAWRNRCQNRDVRQYVQGCGV

Endolysin=T4 lysozyme, [P00720](https://www.uniprot.org/uniprot/P00720), [1SX2](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1sx2&template=protein.html&r=wiring&l=1&chain=A), 1.06Å, 164aa

>sp|[P00720](https://www.uniprot.org/uniprot/P00720)|ENLYS BPT4 Endolysin OS=Enterobacteria phage T4

MNIFEMLRIDERLRLKIYKDTEGYYTIGIGHLLTKSPSLNAAKSELDKAIGRNCNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDSLDAVRRCALINMVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLAKSIWYNQTPNRAKRVITTFRTGTWDAYKNL

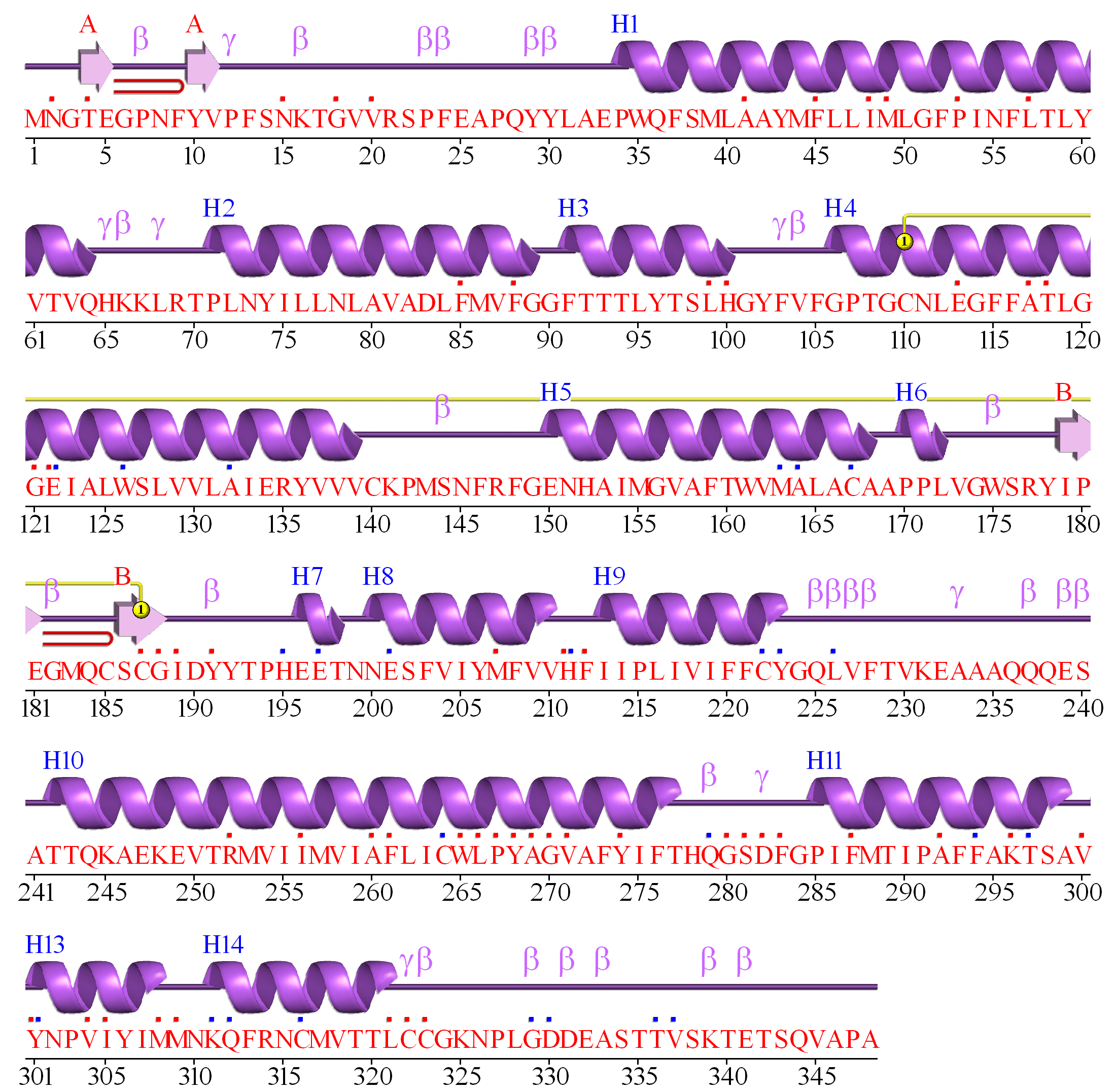
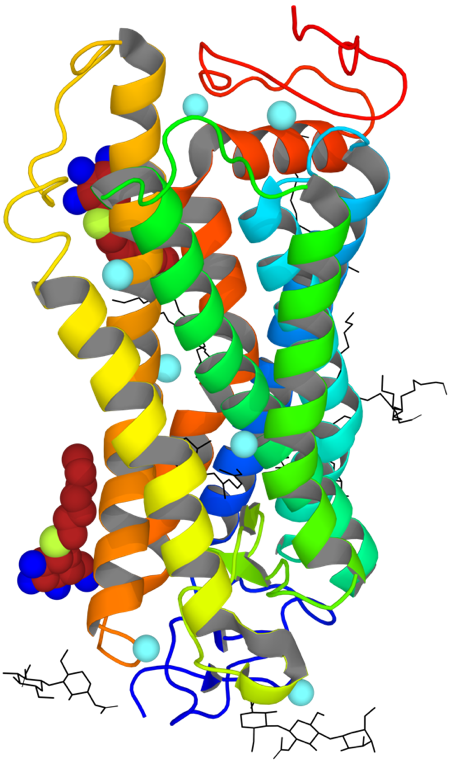
 

**Hydrophobic class: Water-insoluble integral membrane proteins**

Rhodopsin=RHO (Bovine=Cow), [P02699](https://www.uniprot.org/uniprot/P02699), 2.20Å, [1U19](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1u19&template=protein.html&r=wiring&l=1&chain=A), 348aa

>sp|[P02699](https://www.uniprot.org/uniprot/P02699)|OPSD, BOVIN Rhodopsin OS=Bos taurus

MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLYVTVQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGFFATLGGEIALWSLVVLAIERYVVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPPLVGWSRYIPEGMQCSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAAAQQQESATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTIPAFFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQVAPA

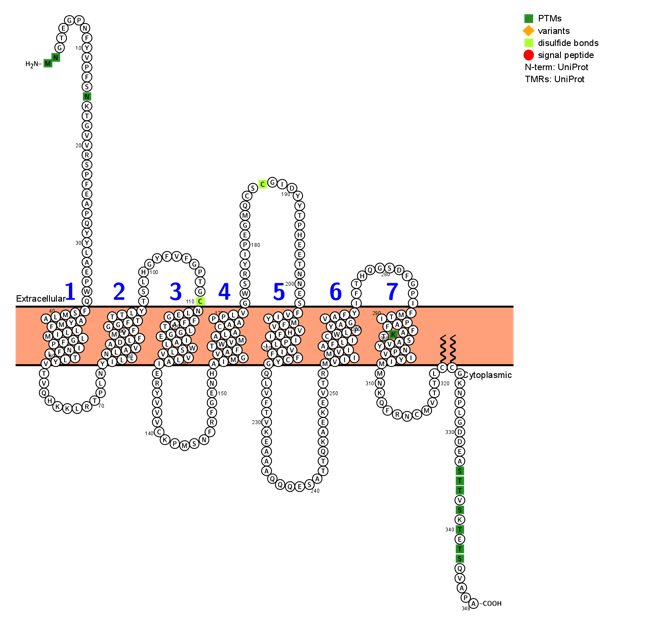
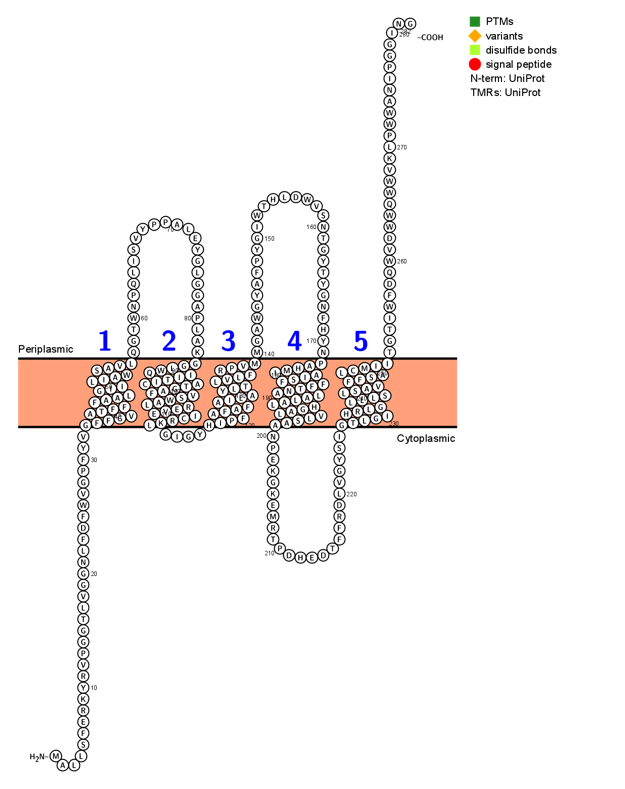
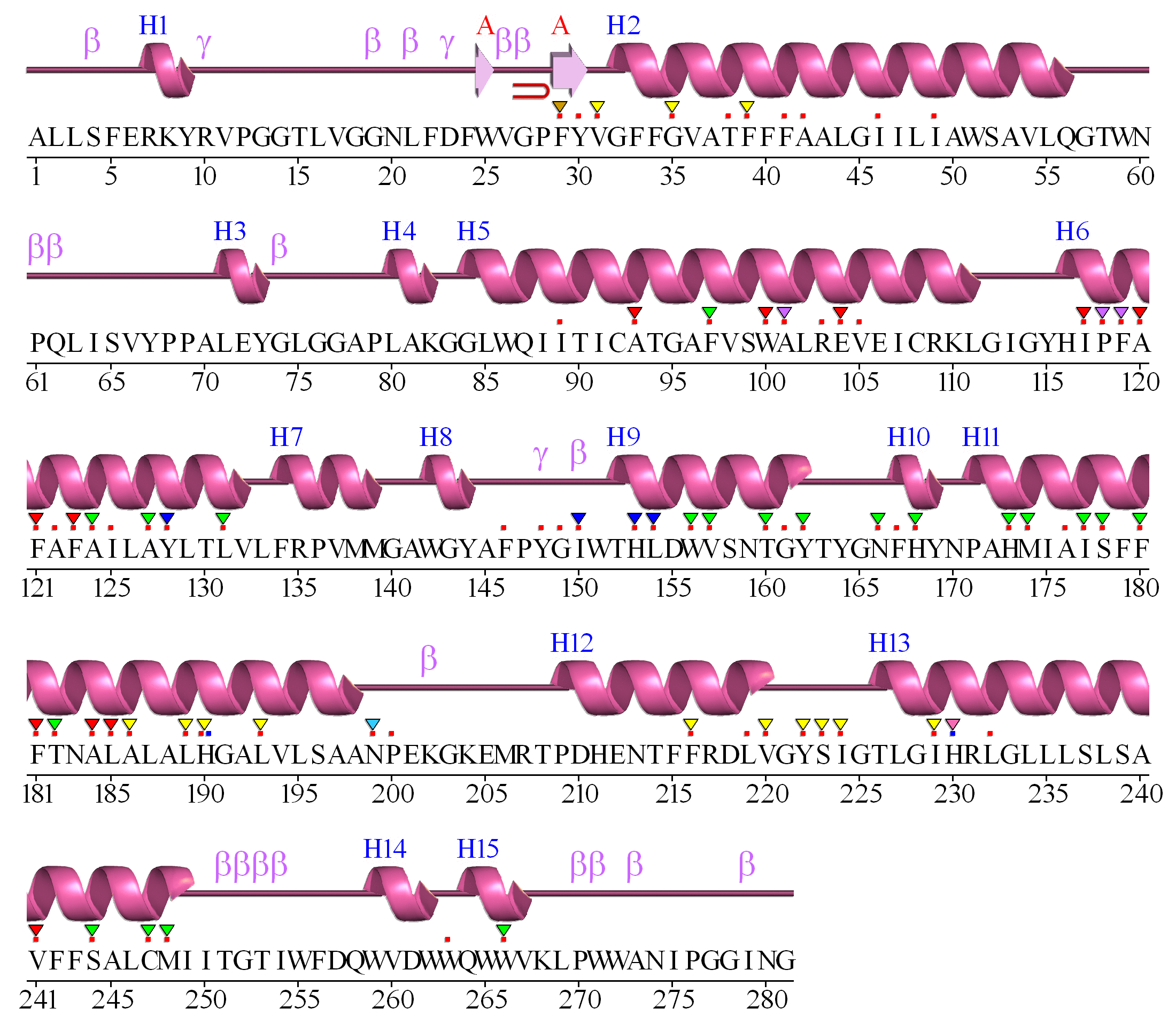


Photo-system Reaction Center

pufL, 1RZH, [P0C0Y8](https://www.uniprot.org/uniprot/P0C0Y8), L-chain, 282aa (Type II)

>sp|[P0C0Y8](https://www.uniprot.org/uniprot/P0C0Y8)|RCEL\_CERSP Reaction center protein L chain OS=Cereibacter sphaeroides OX=1063 GN=pufL PE=1 SV=2

MALLSFERKYRVPGGTLVGGNLFDFWVGPFYVGFFGVATFFFAALGIILIAWSAVLQGTWNPQLISVYPPALEYGLGGAPLAKGGLWQIITICATGAFVSWALREVEICRKLGIGYHIPFAFAFAILAYLTLVLFRPVMMGAWGYAFPYGIWTHLDWVSNTGYTYGNFHYNPAHMIAISFFFTNALALALHGALVLSAANPEKGKEMRTPDHEDTFFRDLVGYSIGTLGIHRLGLLLSLSAVFFSALCMIITGTIWFDQWVDWWQWWVKLPWWANIPGGING

pufM, [P0C0Y9](https://www.uniprot.org/uniprot/P0C0Y9), M-chain, 301aa (Type II)

>sp|[P0C0Y9](https://www.uniprot.org/uniprot/P0C0Y9)|RCEM\_CERSP Reaction center protein M chain OS=Cereibacter sphaeroides

MAEYQNIFSQVQVRGPADLGMTEDVNLANRSGVGPFSTLLGWFGNAQLGPIYLGSLGVLS

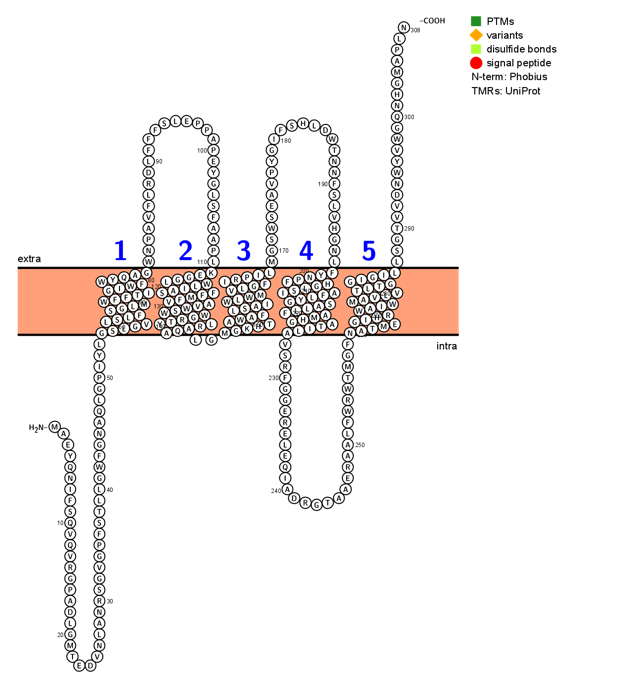
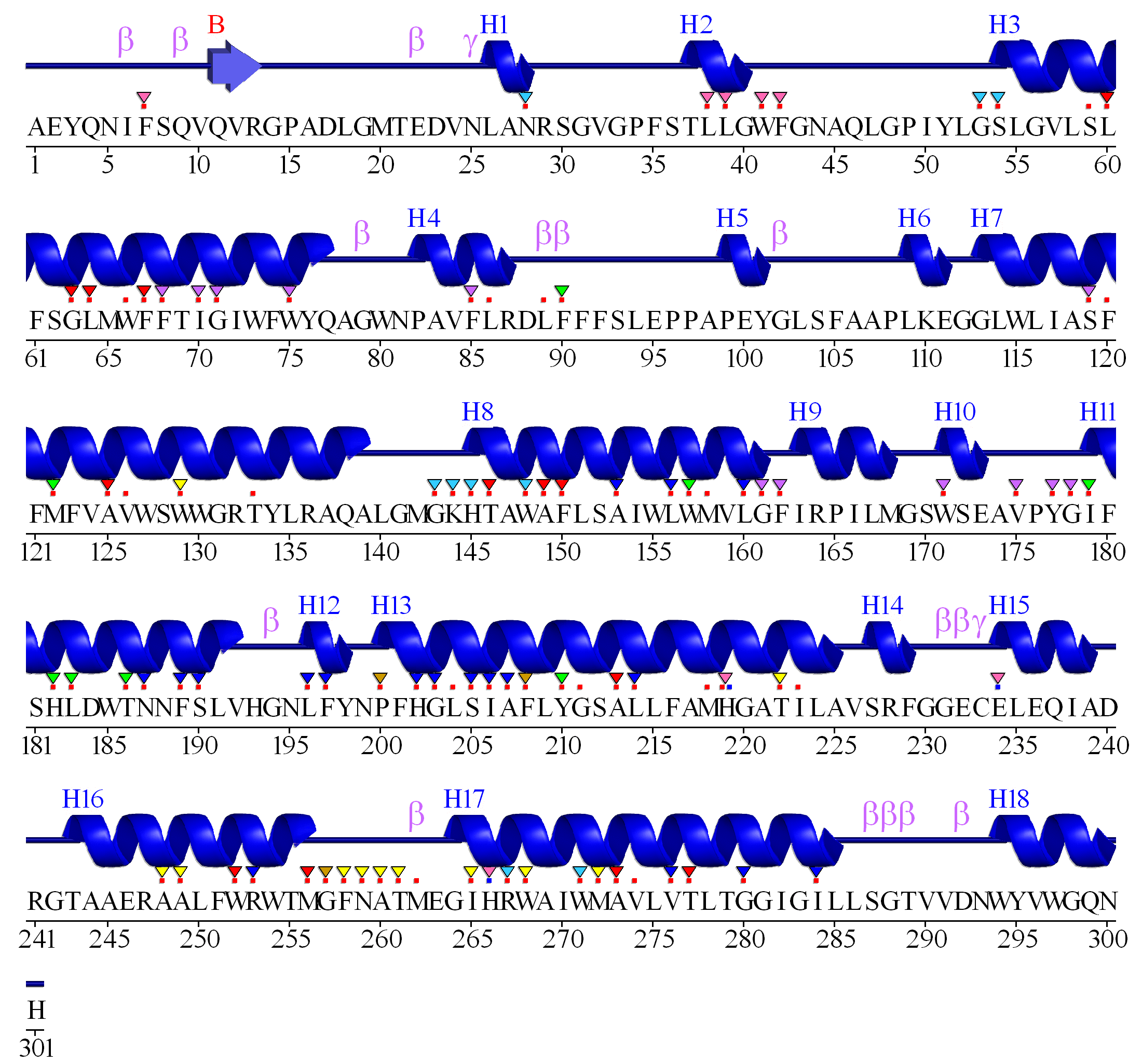
LFSGLMWFFTIGIWFWYQAGWNPAVFLRDLFFFSLEPPAPEYGLSFAAPLKEGGLWLIAS

FFMFVAVWSWWGRTYLRAQALGMGKHTAWAFLSAIWLWMVLGFIRPILMGSWSEAVPYGI

FSHLDWTNNFSLVHGNLFYNPFHGLSIAFLYGSALLFAMHGATILAVSRFGGERELEQIA

DRGTAAERAALFWRWTMGFNATMEGIHRWAIWMAVLVTLTGGIGILLSGTVVDNWYVWGQ

NHGMAPLN

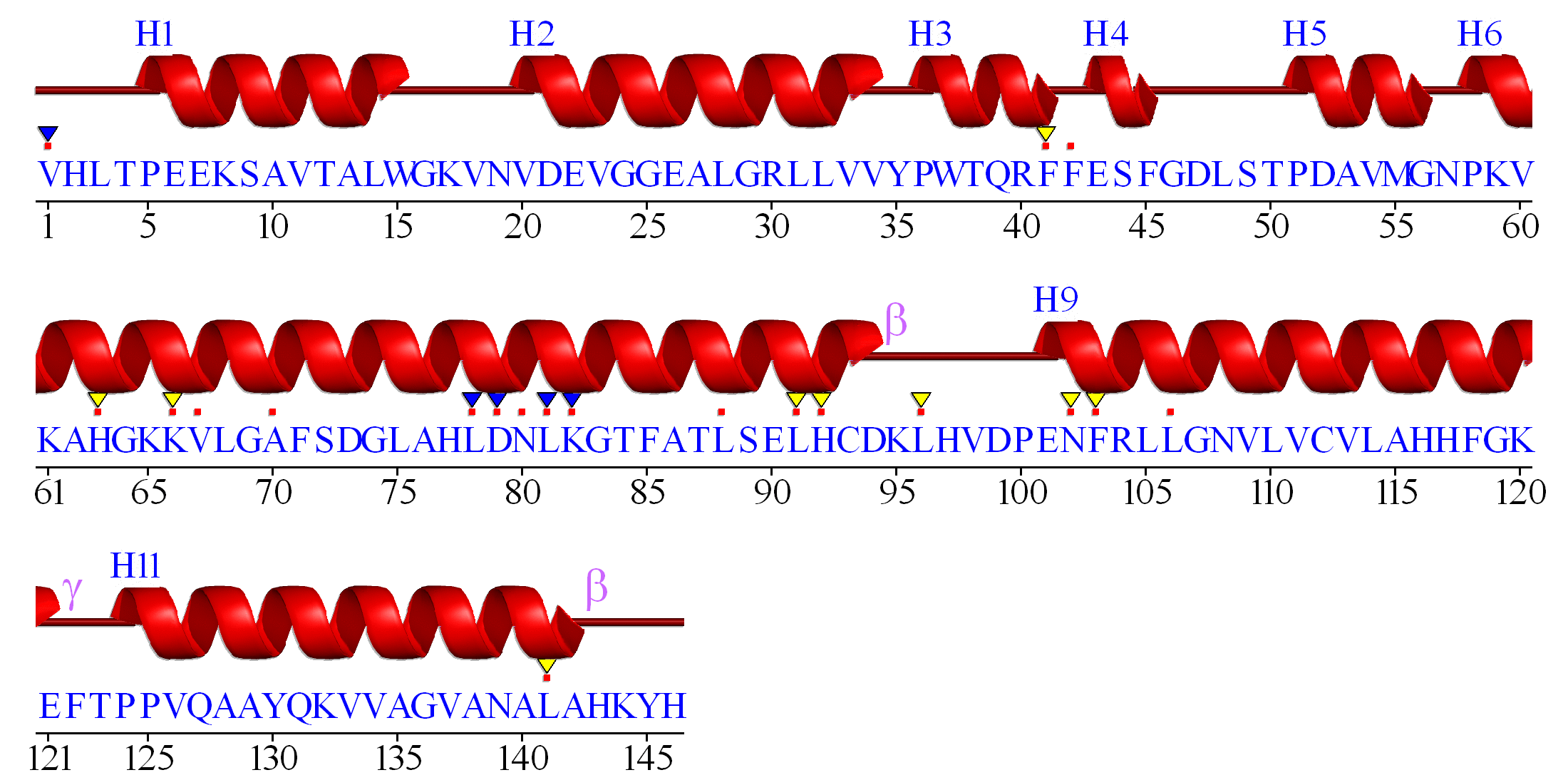
 

**Proteins with amphiphilic alpha-helices**

Hemoglobin subunit beta=HBB, [P68871](https://www.uniprot.org/uniprot/P68871), [1BAB](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1bab&template=protein.html&r=wiring&l=1&chain=A) (1.5Å) 1-147aa (Type III: H6)

>sp|[P68871](https://www.uniprot.org/uniprot/P68871)|HBB HUMAN Hemoglobin subunit beta

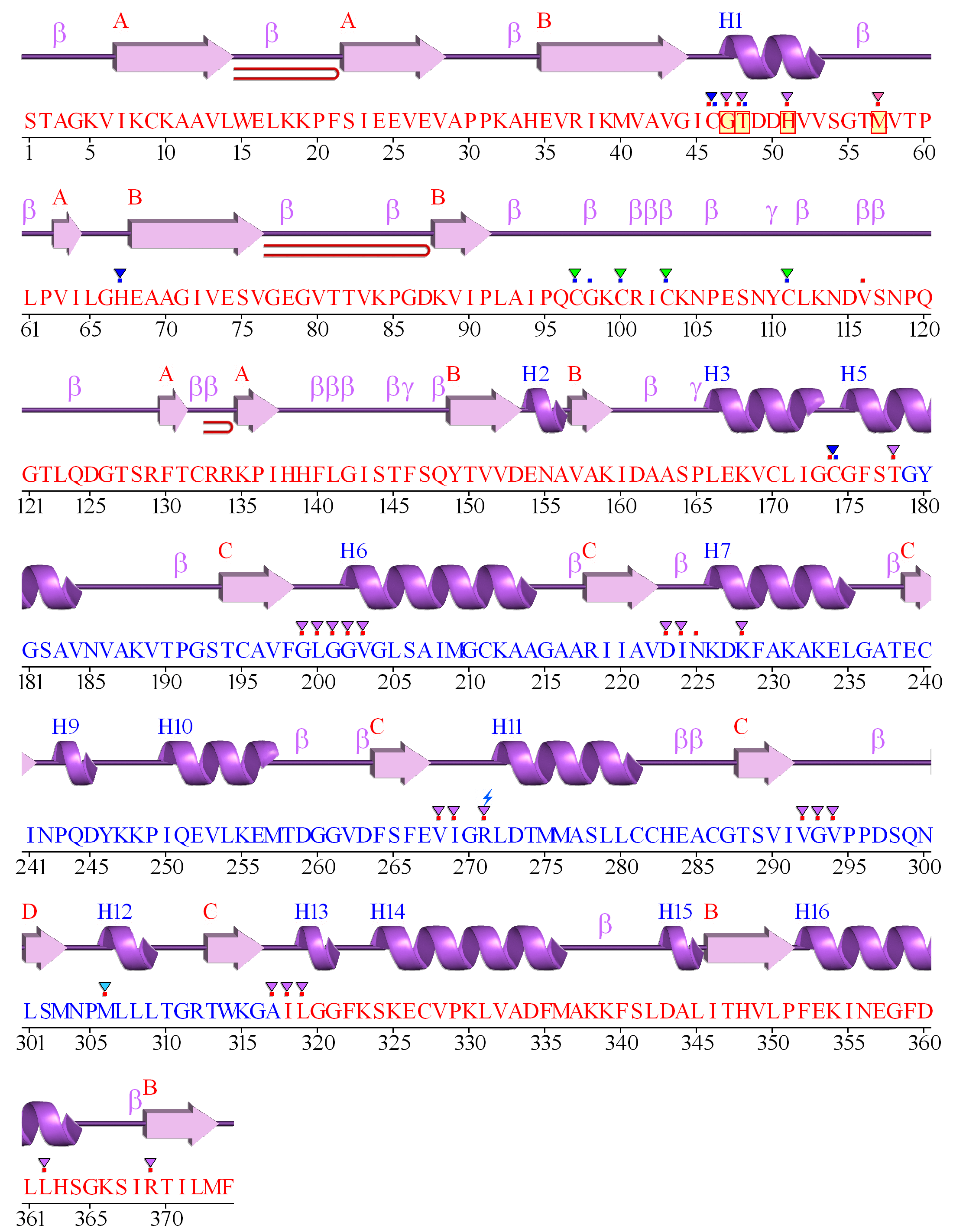
MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH (147aa)



ADH1A =Alcohol dehydrogenase 1A, 1HSO, 375aa (Type III: H16)

>sp|[P07327](https://www.uniprot.org/uniprot/P07327)|ADH1A HUMAN Alcohol dehydrogenase 1A

MSTAGKVIKCKAAVLWELKKPFSIEEVEVAPPKAHEVRIKMVAVGICGTDDHVVSGTMVTPLPVILGHEAAGIVESVGEGVTTVKPGDKVIPLAIPQCGKCRICKNPESNYCLKNDVSNPQGTLQDGTSRFTCRRKPIHHFLGISTFSQYTVVDENAVAKIDAASPLEKVCLIGCGFSTGYGSAVNVAKVTPGSTCAVFGLGGVGLSAIMGCKAAGAARIIAVDINKDKFAKAKELGATECINPQDYKKPIQEVLKEMTDGGVDFSFEVIGRLDTMMASLLCCHEACGTSVIVGVPPDSQNLSMNPMLLLTGRTWKGAILGGFKSKECVPKLVADFMAKKFSLDALITHVLPFEKINEGFDLLHSGKSIRTILMF



CybC=C562, Coiled coils, tetramer cytochrome b562, [P0ABE7](https://www.uniprot.org/uniprot/P0ABE7), [256b](https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=256b&template=protein.html&r=wiring&l=1&chain=A) (Type III: H4)

>sp|[P0ABE7](https://www.uniprot.org/uniprot/P0ABE7)|C562, ECOLX Soluble cytochrome b562 OS=Escherichia coli

ADLEDNMETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYHQKYR

