

Sample EMBL Format

ID AB000263 standard; RNA; PRI; 368 BP.

XX

AC AB000263;

XX

DE Homo sapiens mRNA for prepro cortistatin like peptide, complete cds.

XX

SQ Sequence 368 BP;

acaattggccc.....60

Gactttaagg 368

//



Sample Fasta format protein sequence

>gi|13959637|sp|C9PTU8|VSP3_BOTJA Venom serine proteinase A precursor
MVLIRVIANLLILQLSNAQKSSELVIGGDECNITEHRFLVEIFNSSGLFCGGTLIDQEWVLAAHCDMRN
MRIYLGVHNEGVQHADQQRRFAREKFFCLSSRNYTKWDKDIMLIRLNRPVNNSEHIAPLSLPSNPPSVGS
VCRIMGWGTITSPNATFPDVPHCANINLFNYTVCRGAHAGLPATSRTLCAVLQGGIDTCGGDSGGPLIC
NGTFQGIVSWGGHPCAQPGEPALYTKVFDYLPWIIQSIIAGNTTATCPP



An example sequence in FASTQ format is:

@SEQUENCE_ID
GTGGAAAGTTCCTTAGGGCATGGCAAAGAGTCAGAAATTTGAC
+
FAFFADEDGDBGEGGBCGGHE>EEBA@@@=



Protein Structure File Format

- PDB - the PDB file format is used to store both sequence information, but more importantly stores 3-dimensional structure information. This information can be used to visualize the crystal structure of a protein. PDB files are simply text files, thus can be viewed with a text editor, and often have the file extension '.pdb'.

