Fasta Format

>NP_001193503.1 E3 ubiquitin-protein ligase TRIM56 [Bos taurus]

MKMRFFSSPCGKAAVDPADRCKEVQQIRD
QHPSKIPVIIERYKGEKQLPVLDKTKFLVPDHV
NMSELVKIIRRRLQLNPTQAFFLLVNQHSMV
SVSTPIADIYEQEKDEDGFLYMVYASQETFGF

>FN640832

CCTGGTAGCTATGGCTTGCCTTTACTAAGA
CCCATCTCAAACAGGCTCAATTATTTTTGGT
TCCAAGGGCCTGAAACATTCTTAAAGAAGC
GAATAGAGAAACACAGGAGCACAGTTTTT
CGCACCAATATCCCTCCAACTTTCCCTTTCT
TCTCCAATGTTAATCCCAGCGTTGTTGCTGT
CCTTGACACCAAGTCTTTTGCACACCTC

A sequence must start with a header line

- Begins with a >
- First "word" is the sequence id
- Rest of line may contain more sequence descriptors

Fasta Format

>gi|31563518|ref|NP_852610.1|
microtubule-associated proteins 1A/1B
light chain 3A isoform b [Homo sapiens]

MKMRFFSSPCGKAAVDPADRCKEVQQIRD
QHPSKIPVIIERYKGEKQLPVLDKTKFLVPDHV
NMSELVKIIRRRLQLNPTQAFFLLVNQHSMV
SVSTPIADIYEQEKDEDGFLYMVYASQETFGF

>FN640832

CCTGGTAGCTATGGCTTGCCTTTACTAAGA
CCCATCTCAAACAGGCTCAATTATTTTTGGT
TCCAAGGGCCTGAAACATTCTTAAAGAAGC
GAATAGAGAAACACAGGAGCACAGTTTTT
CGCACCAATATCCCTCCAACTTTCCCTTTCT
TCTCCAATGTTAATCCCAGCGTTGTTGCTGT
CCTTGACACCAAGTCTTTTGCACACCTC

The header is followed by the
sequence
- May be amino acid or nucleotide
- May be a single line or multiple
  lines

No empty line between sequence
entries