

The amino acid sequence file format must take into account:

- 1) First line should contain the name of the protein and/or some comments if it is necessary. This line is absolutely necessary. It cannot be any longer than 100 characters.
- 2) The following line(s) should contain the sequence in one letter code (uppercase only). Spaces, numbers or carriage returns are also permitted. There are no limitations as to the length of each line. The maximal length is 2000 residues.

These characteristics agree with the Pearson format. The public domain software FORMAT is very useful for formatting sequences. But if you are using a word processor to adjust formats, be careful to save files in TEXT only (ASCII).

The sequences can be handled one by one, or construct a specific file ".FOS" which enables the manipulation of groups of up to 20 sequences.