

Cn3D: Editing Alignments Using Sequence and Structure Information

Paul A. Thiessen, David L. Wheeler, Stephen H. Bryant

National Center for Biotechnology Information (NCBI) • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

Cn3D has long been NCBI's structure and structure alignment viewer. With version 4.1, Cn3D now includes a suite of alignment tools that lets the user edit and create multiple alignments using a variety of built-in algorithms that operate on both sequence and structure information. Cn3D's alignment model is well suited to structure-based sequence alignments, and guides the user towards defining a protein family's conserved core, wherein each member sequence contains all the sites that are homologous in both sequence and structure. The program's built-in algorithms assist the discovery of conserved sites and the accurate alignment of new member sequences, both with and without known structure. Sequence and structure conservation can be visualized directly, helping the user to gauge where strong or weak homology exists, and even where errors in alignment are more likely. Cn3D does not automate alignment creation, since the user must still use some judgment in validating sequence and structure conservation; however, it provides many tools for analyzing sequence and structure relationships in different ways, and for selectively combining this information by building up an alignment iteratively. Alignments can be annotated to highlight important conserved residues, and exported for use in other applications.

www.ncbi.nlm.nih.gov

Key Cn3D Features

- visualizing and annotating 3D structures interactively
- visualizing sequence and structure alignments
- importing sequences and structures
- coloring alignments based on sequence conservation to spot regions with poor sequence homology
- superpositioning structures based on sequence alignments to show blocks with poor spatial correspondence
- editing alignments both manually and using several algorithms:
 - pair-wise BLAST
 - BLAST sequence against a PSSM (e.g., PSI-BLAST)
 - block alignment
 - threading (sequence to structure alignment)

www.ncbi.nlm.nih.gov/Structure/CN3D/cn3dinstall.shtml

Cn3D's Alignment Model

Aligned blocks have fixed lengths common to all aligned sequences.

Blocks are separated by unaligned loops of variable length.

Blocks typically correspond to secondary structure elements.

Blocks can have no internal gaps.

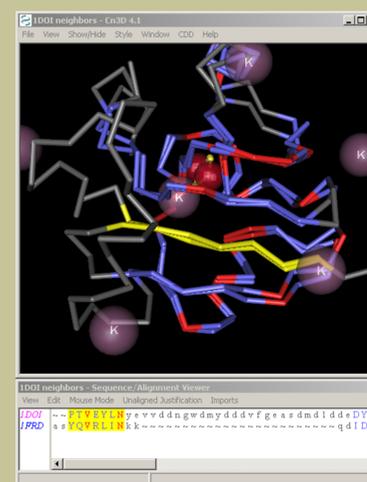


Figure 1. VAST alignment of 1DOI and 1FRD. Aligned residues are shown in red/blue, with red indicating identical residue types. VAST alignments consist of continuous sequence blocks. In this case, the block highlighted in yellow corresponds to a single beta strand.

Alignment Editing



Figure 2. Before and after views of an edited alignment, highlighting many sequence alignment mistakes made obvious by structure alignment. The same residues are shown in yellow in both alignments.

Cn3D's Import Window

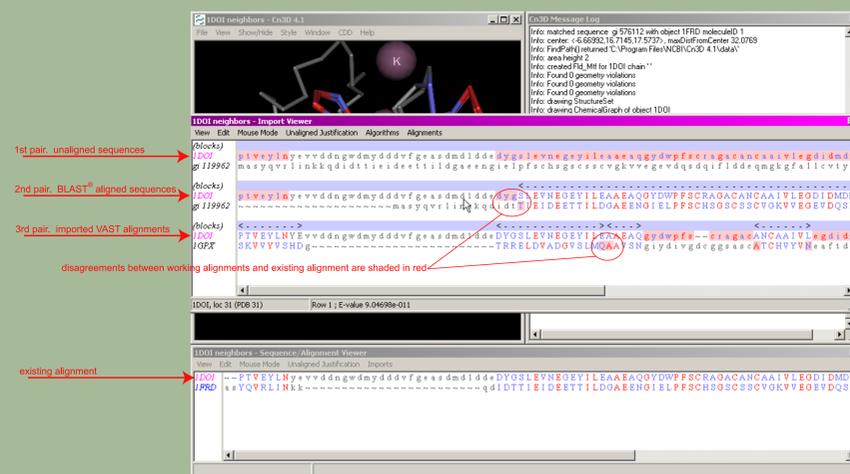


Figure 3. Importing new sequences into an existing alignment. A sequence is shown as originally imported, with no alignment (first pair in the import window), and then aligned to 1DOI with BLAST (second pair). A structure, 1GPX, is shown imported with its VAST alignment to the master structure 1DOI (third pair).

Alignment Annotation

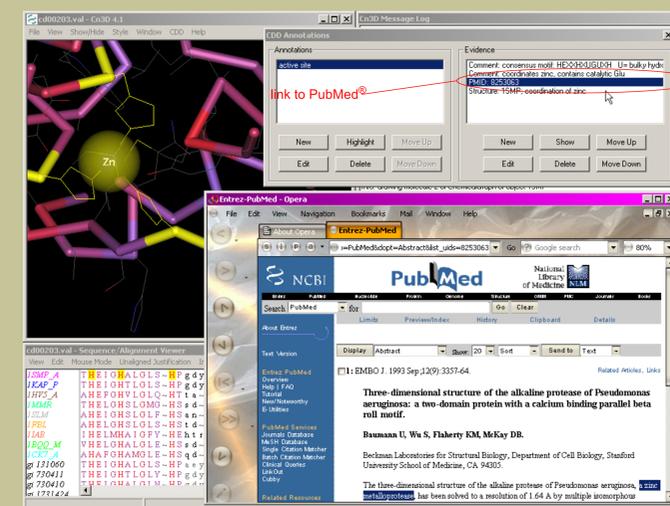


Figure 4. Annotating alignments in Cn3D. Shown is an annotation of important active site residues which are highlighted in yellow, along with curated links to the primary literature.