MatAlign vs. SSAP

Double dynamic programming approach for protein structure alignment was pioneered by the SSAP method [TO89]. The basic idea of double dynamic programming is to apply multiple *low-level* dynamic programming procedures in a number of *local* setups, and consolidate these low-level results into the *global* alignment by a single *high-level* dynamic programming procedure.

Although MatAlign also employs double dynamic programming, its approach is **substantially different** from that of SSAP in that:

1. For low-level dynamic programming:
   - SSAP superimposes (rotates and translates) two *local coordinate frames* of two residues, each from one protein, and align their *views* defined by the *direction*, *orientation*, *sequence distance*, and *spatial distance* components.
   - MatAlign does not require any rotation and translation. It aligns the *distance profiles* of two residues, each from one protein, defined by the corresponding rows in their respective distance matrices. For speed optimization, *reduced rows (reduced distance profiles)* can be used.
2. For high-level dynamic programming:
   - SSAP accumulates (sums up) the scores of all corresponding low-level dynamic programming matrix cells that involve in their respective low-level alignments in order to calculate the score of a particular cell in the high-level dynamic programming matrix. Mathematically, for two proteins *A* and *B*:

     $S_{high}[i, j] = \sum_{a \in A, b \in B} S_{low}^{a,b}[i, j]$ for any $i \in A$ and $j \in B$. ($S_{low}^{a,b}[i, j] = 0$ if it is not in the alignment path of $S_{low}^{a,b}$ matrix.)

   - MatAlign uses the corresponding final alignment score (stored in the bottom-right cell of the corresponding low-level alignment matrix) as the score of a particular cell in the high-level dynamic programming matrix. Mathematically:

     $S_{high}[i, j] = S_{low}^{i,j}[|A_i|, |B_j|]$ for any $i \in A$ and $j \in B$, where $|A_i|$ and $|B_j|$ are the lengths of i^{th} and j^{th} rows in *A*'s and *B*'s distance matrices respectively.

     This enables MatAlign to use *reduced rows* for speed optimization, since the dimensions of the low-level dynamic programming matrices and that of the high-level matrix need not be the same. (SSAP's summing-up approach cannot be adapted for such an optimization, since the high and low-level matrices requires the same dimension.)

Reference

[TO89] W. R. Taylor and C. A. Orengo, "Protein structure alignment", *Journal of Molecular Biology*, **208**:1–22, 1989.