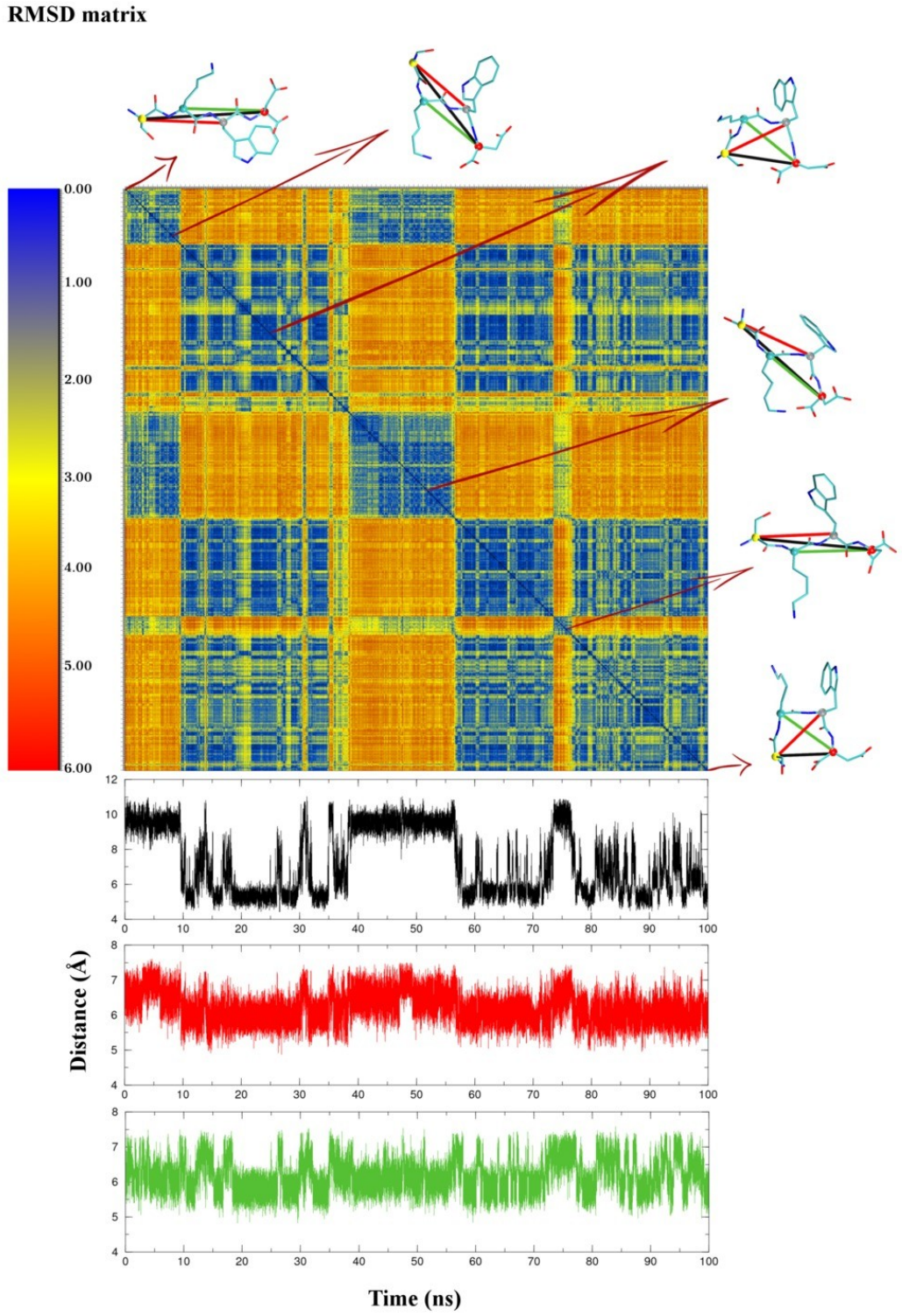
Supplementary Material

On the foldability of tryptophan-containing tetra- and pentapeptides: an exhaustive molecular dynamics study

Panagiota S. Georgoulia & Nicholas M. Glykos*

Department of Molecular Biology and Genetics, Democritus University of Thrace, University campus, 68100 Alexandroupolis, Greece, Tel +30-25510-30620, Fax +30-25510-30620, http://utopia.duth.gr/~glykos/ , glykos@mbg.duth.gr

Figure S1

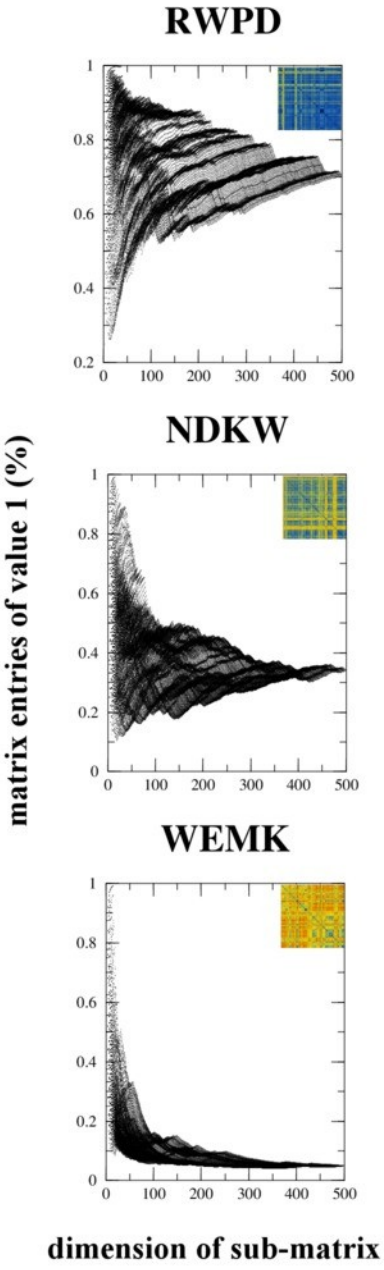


RMSD matrix, evolution of interatomic distances and snapshot-structures of a representative

tetrapeptide, SKWD, calculated from a 100-ns long trajectory (3rd pass in Table 1):

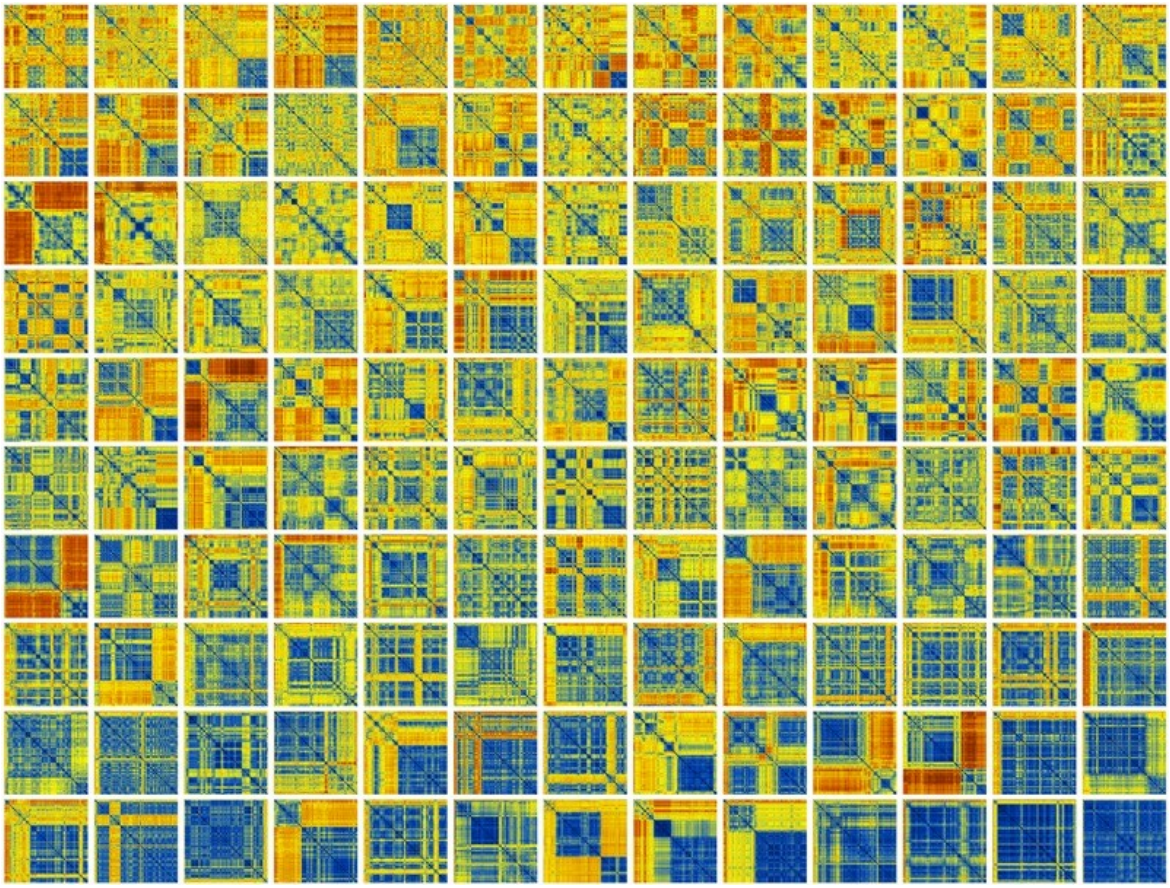
The upper panel shows a color representation of the square matrix of the RMSDs between all possible structures from successive frames of the trajectory (using all-heavy atoms). The origin (t=0) is at the top left-hand corner and the linear color scale ranges from dark blue (0Å), through yellow (~3Å), to dark red (6Å) as shown in the ruler on the left. Snapshot-structures taken directly from the trajectory are shown in licorice with Cα atoms in cpk representation. The interatomic vectors are indicated using black for the distance between 1-4 Cα atoms, red for the distance between 1-3 Cα atoms and green for the distance between 2-4 Cα atoms. The lower panel shows the evolution of the three interatomic distances between all three pairs of Cα atoms using the same colors as the ones used in the upper panel.

Figure S2



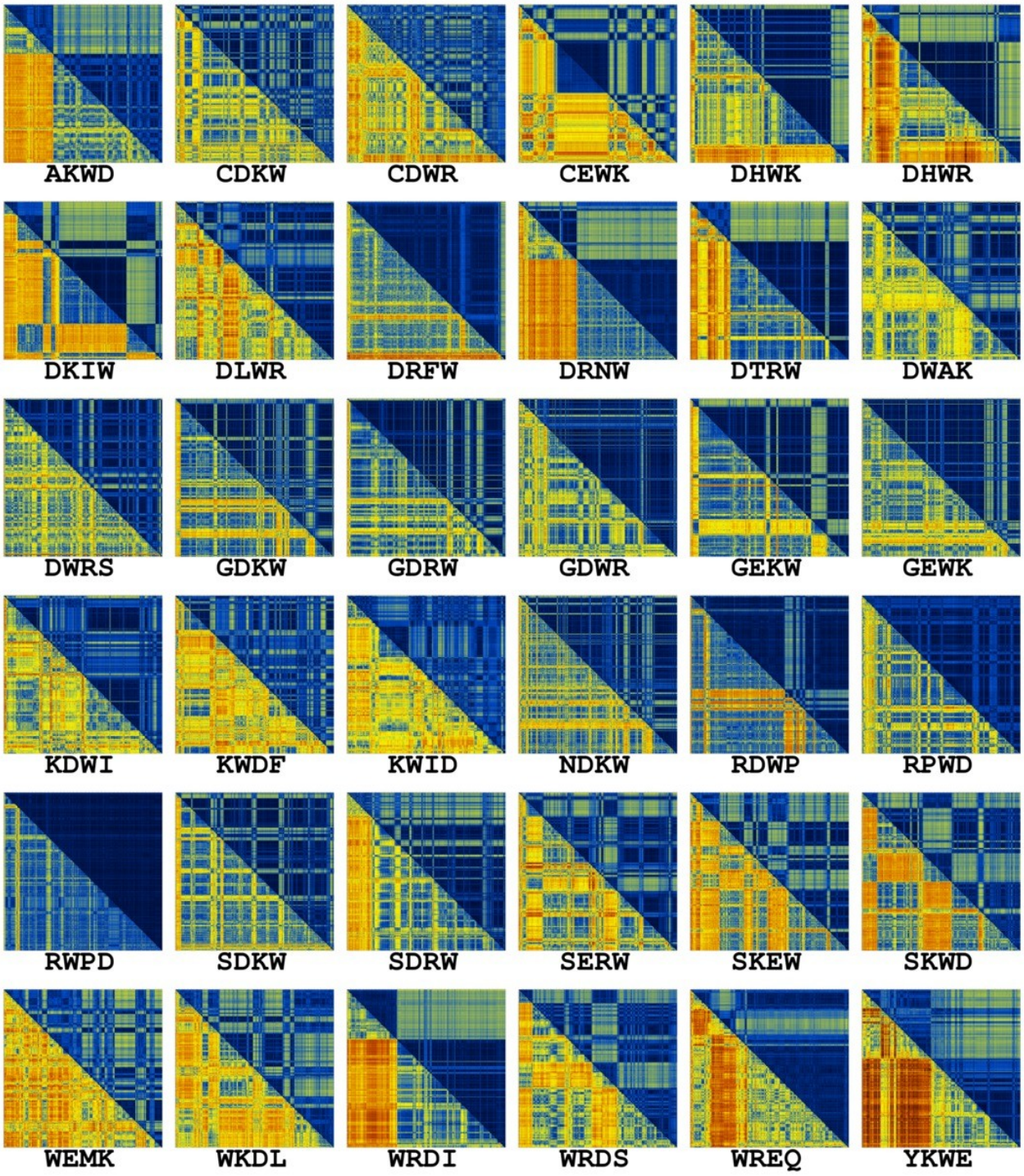
Graphical interpretation of the TF2 function: Three representative peptides with the highest (RWPD, TF2 score:6438), median (RWPD, TF2 score:1120) and lowest (RWPD, TF2 score:40) TF2 score from the 3rd pass (Table 1) are shown. In all three panels the inset shows the RMSD matrix to be analyzed. The main graphs show the distribution of the percent of matrix entries of value 1 (RMSD ≤ 2.0Å) (y-axis) with the dimension of the symmetric submatrix (x-axis) (step 4). The product of median and mode of the distribution is used to rank the peptides for their foldability.

Figure S3



RMSD matrices of the 130 tetrapeptides (2nd pass in Table 1) calculated from the 30ns-long trajectories: The order from left to right and top to bottom follows the increase in the score received by the scoring function TF2. Each image corresponds to an individual peptide and is a color representation of the square matrix of the RMSDs between all possible structures from successive frames of the trajectory. For each matrix the origin (t=0) is at the top left-hand corner. All matrices were calculated using all heavy atoms (~40 atoms) and the linear color scale is kept the same, ranging from dark blue (0Å), through yellow (~3Å), to dark red (6Å).

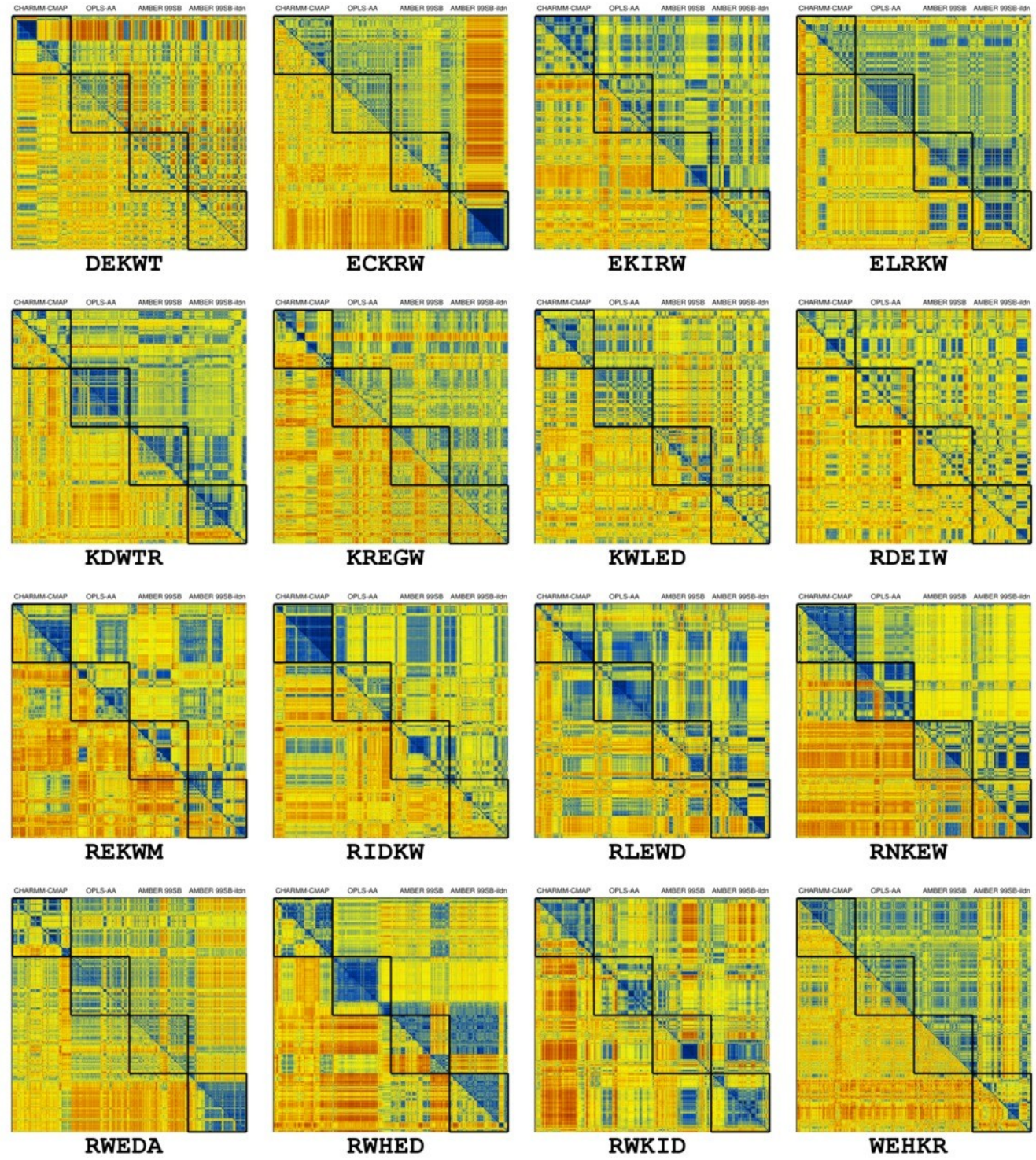
Figure S4



RMSD matrices of the 36 tetrapeptides (3rd pass in Table 1) calculated from the 100ns-long trajectories:

Each image corresponds to an individual peptide and is a color representation of the square matrix of the RMSDs between all possible structures from successive frames of the trajectory. For each matrix the origin (t=0) is at the top left-hand corner. All matrices were calculated using only backbone atoms (above the diagonal) and all heavy atoms (below the diagonal). The linear color scale is kept the same, ranging from dark blue (0Å), through yellow (~3.1Å), to dark red (6.2Å).

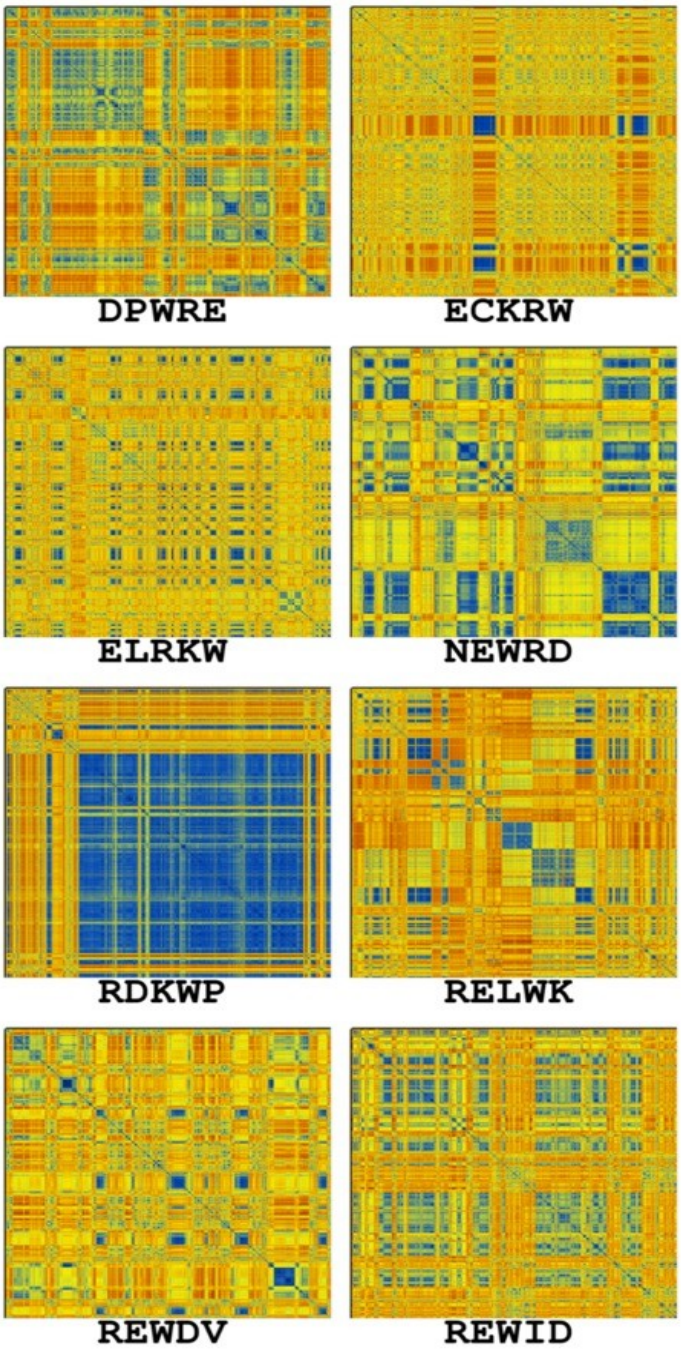
Figure S5



Inter-RMSD matrices for the 16 pentapeptides and the 4 different force fields (complementary to Figure 3):

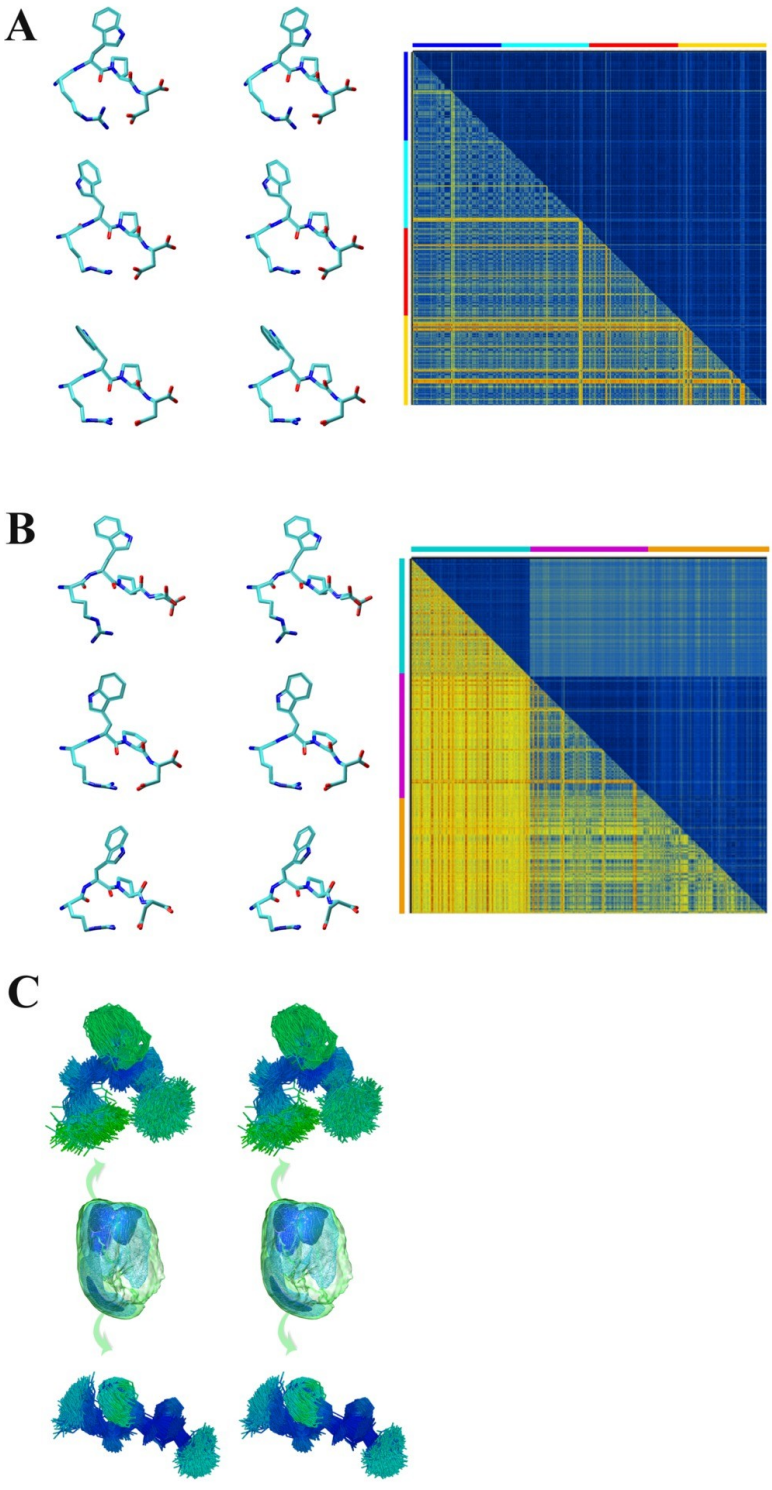
Each image corresponds to an individual peptide and is a color representation of the square matrix of the (inter) RMSDs between all possible structures from successive frames of an artificial trajectory (480ns total simulation time per peptide) produced by concatenating four independent trajectories produced by four force fields (3rd pass in Table 1). The limits of each trajectory are indicated with square black boxes, with the corresponding force field, CHARMM-CMAP, OPLS-AA, AMBER99SB and AMBER99SB-ildn, noted on the top. All matrices were calculated using only backbone atoms (above the diagonal) and all heavy atoms (below the diagonal). The linear color scale is kept the same, ranging from dark blue (0Å), through yellow (~3.9Å), to dark red (7.8Å).

Figure S6



RMSD matrices for 8 pentapeptides (4th pass in Table 1) and the AMBER99SB-ildn force field: Each matrix is calculated for 1μs-long trajectory and using all heavy atoms (~40 atoms). The linear color scale is kept the same, ranging from dark blue (0Å), through yellow (~3.7Å), to dark red (7.5Å).

Figure S7



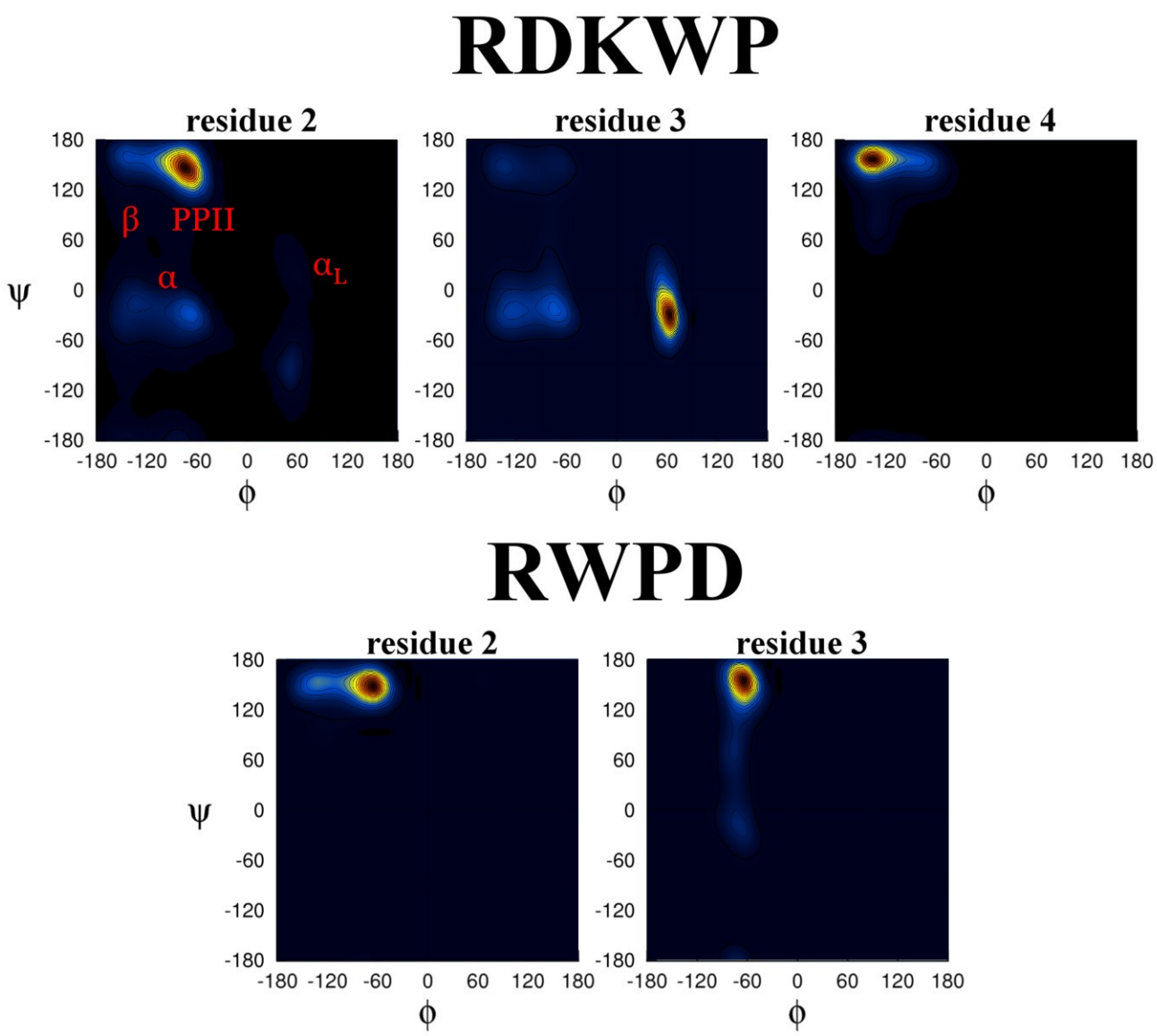
The tetrapeptide RWPD :

In panel A we show results from the simulations in the four temperatures (4th pass in Table 1) and the CHARMM22 force field. The RMSD matrix is calculated for the concatenated trajectory (their order follows the raise in temperature) and using backbone atoms (above the diagonal, maximum rmsd 3.1Å) and all heavy atoms (below the diagonal, maximum rmsd 3.1Å). On the left we show in stereodiagram the representative structures observed during all simulations, each corresponding to a distinct cluster of structures, as obtained from principal component analysis in Cartesian space and using all heavy atoms. From top to bottom the structures have occupancies of 10-17%, 17-54%, 12%, 1.3% and 2.9%, of the total simulation time, depending the temperature of the simulation.

In panel B we show results from the simulations using three force fields (5th pass in Table 1). The RMSD matrix is calculated for the concatenated trajectory (AMBER99SB is indicated with cyan, CHARMM27-CMAP with magenta and OPLS-AA with orange) and using backbone atoms (above the diagonal, maximum rmsd 3.1Å) and all heavy atoms (below the diagonal, maximum rmsd 3.1Å). On the right we show in stereodiagram the representative structure of the prominent cluster observed during each simulation, as obtained from principal component analysis in Cartesian space and using all heavy atoms. From top to bottom, the AMBER, CHARMM and OPLS structures have occupancies of 40%, 63% and 53% of the total simulation time (all rest observed clusters have occupancies less than 6%).

In panel C we show the three-dimensional folding landscape from the trajectory obtained at 320K with the AMBER99SB force field. The middle is a wall-eyed stereodiagram of the projection of the trajectory on the space defined by the top three principal components as obtained from a principal component analysis in Cartesian space using all-heavy atoms. Three isosurfaces are drawn at the mean density (white transparent isosurface), 1σ above mean (pink wireframe) and 6σ above mean (magenta solid surface) of the map distribution. The mauve picks correspond to the two cluster of structures (with 40% and 6% occupancy of the total simulation time) observed during the trajectory. For each cluster we show (in wall-eyed stereodiagram) 500 snapshot-structures, obtained directly from the trajectory. The structures are colored according to their atomic rms fluctuation from the calculated (for the cluster) average structure using a color gradient from blue to red (maximum rmsf, 5.55Å).

Figure S8



Ramachandran plots: The diagrams on the top row show the Ramachandran plots for the 2μs-long folding trajectory (Table 1) for the RDKWP peptide (Figure 4) and the AMBER99SB-ildn force field [only residues 2-4 (inclusive) have been used for the calculation]. The bottom row corresponds to the Ramachandran plots for the 1μs-long folding trajectory (Table 1) for the RWPD peptide (Figure S7) and the AMBER99SB force field [only residues 2-3 (inclusive) have been used for the calculation]. Contour lines are drawn at the same (arbitrary) interval. The α, β, polyproline (PPII) and α_L Ramachandran regions are indicated in the first panel.