

Supporting Information

Ceriotti et al. 10.1073/pnas.1108486108

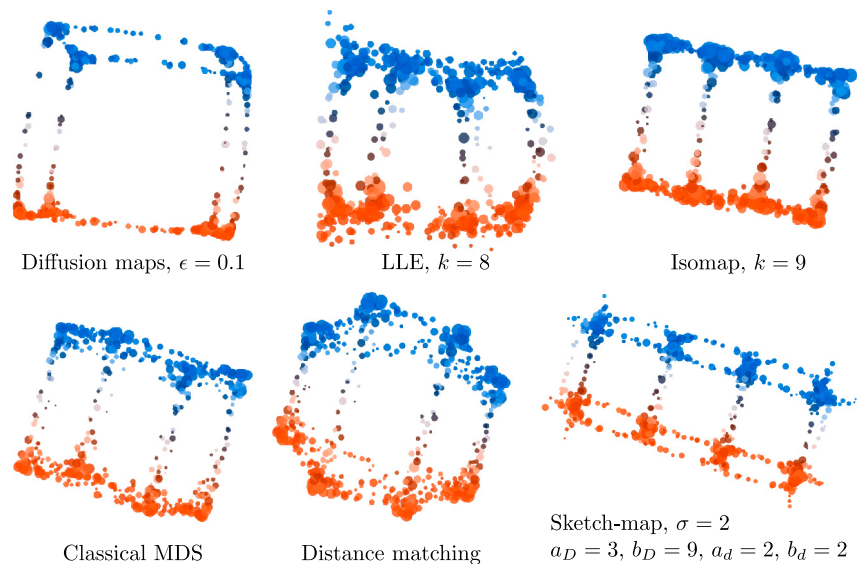


Fig. S1. 2D projections of the landmark points selected from the dataset depicted in Fig. 2 of the main text. As for sketch-map distances were calculated in a way that takes the periodicity of the space into account. Projections obtained using a number of commonly used manifold learning algorithms are shown along with the parameters used (for all the algorithms a number of different parameters were tried and similar results were obtained). Projected points are colored, using the key shown in Fig. 1C of the text, in accordance with the value of one of the three underlying variables. This figure makes clear that sketch-map gives a far better representation of the topology of the 3D distribution than any of the other algorithms.

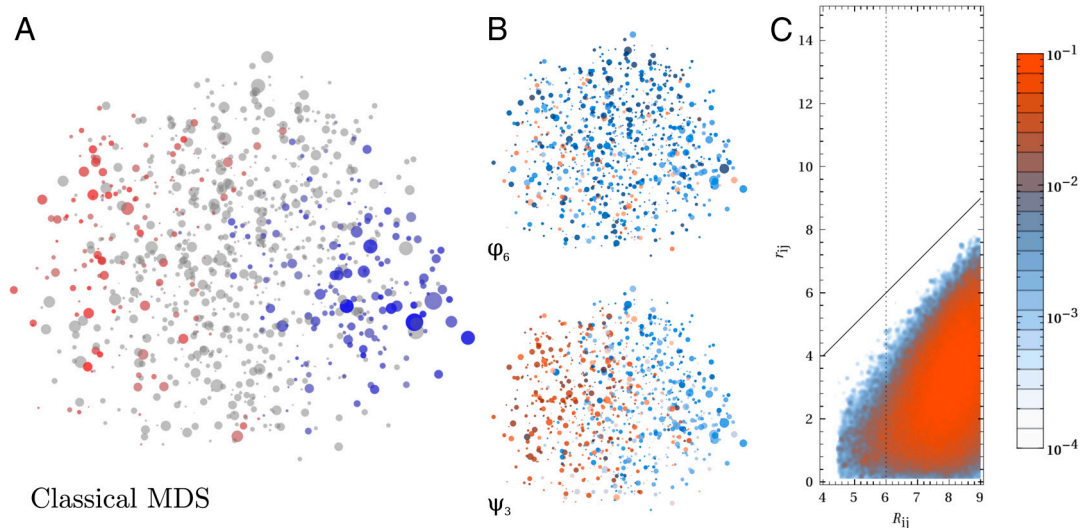


Fig. S2. Diagnostic information on the embedding of the ala12 landmark points using classical multidimensional scaling (MDS). Distances here are calculated using the differences between torsional angles in the landmark configurations taking into account the periodicity of the space. **A** should be compared with Fig. 4 of the main text as here points are colored according to the protein secondary structure. **B** should be compared with Fig. 5C of the main text as here points are colored according to the value of a particular dihedral angle. Lastly, **C** shows the joint probability distribution of the distances between two frames in high dimension (R_{ij}) and the distance between the corresponding low-dimensional projections (r_{ij}).

