Supplementary Material for

Factors Stabilizing β-Sheets in Protein Structures
from a Quantum-Chemical Perspective

Martin Culka and Lubomír Rulišek*

1 Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences,
Flemingovo náměstí 2, 166 10, Praha 6, Czech Republic

* rulisek@uochb.cas.cz
Search for Global Minimum

For the calculation of strain energy along the protein fold on overlapping tripeptide basis, we had to find a global minimum for each tripeptide in water solution. We generated 200 conformers for each tripeptide using molecular dynamics (MD) at 1000 K, as described in the main text, Methods section. Here, we compare the global minima found using this methodology and much more robust but computationally very demanding bias-exchange metadynamics approach used in our previous study.\(^1\) We generated 200 conformers for the VIV tripeptide studied previously using the 1000K MD approach and compare the energetic minimum of this set with the minimum of the much larger set of 2522 unique VIV conformers derived from metadynamics published as SI of the reference 1. Using the QM(DFT-D3)/COSMO-RS methodology as described in reference 1 for free energy evaluation, we found out that the energetic minima of the two conformer sets lie 0.008 kcal.mol\(^{-1}\) from each other and are structurally almost indistinguishable. When the simplified QM(DFT-D3)/COSMO-RS methodology (see below) is used for the evaluation of energies within both conformer sets, the global minima lie 0.6 kcal.mol\(^{-1}\) and have slightly more different structure.

Comparison of QM(DFT-D3)/COSMO-RS Methodologies

We recalculated energies for both sets of VIV conformers described above using simplified QM(DFT-D3)/COSMO-RS methodology, which does not account for thermal vibrations and calculates electronic energies with a smaller DZVP-DFT basis set. See Methods section in the main text for a complete description. Note that the DZVP-DFT basis set has been, however, shown to perform very well in benchmarks.\(^2\) We compare the way how both approaches (original from ref. 1 and simplified from this paper) rank the conformers. In other words, we calculate mean absolute error (MAE) of relative energies within each conformer set. For the large metadynamics set, we
get MAE=1.43 kcal.mol\(^{-1}\). In the case the 1000K MD conformer set, the MAE is 0.77 kcal.mol\(^{-1}\).

Note that although those errors are not negligible, we rely on error cancelation when evaluating profiles of the strain energy rather than directly interpreting absolute values.

References
