Guowei Qi, Michail D. Vrettas, Carmen Biancaniello, Maximo Sanz-Hernandez, Conor T. Cafolla, John W. R. Morgan, Yifei Wang, Alfonso De Simone and David J. Wales

1) Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK
2) Department of Pharmacy, University of Naples Federico II, 80131 Naples, Italy
3) Department of Life Sciences, Imperial College Longon, South Kensington, London SW7 2AZ, UK
* Corresponding authors: dw34@cam.ac.uk or alfonso.desimone@unina.it

INTRODUCTION

Biomolecular simulations are very useful methods to overcome the inherent limitations of experimental structural elucidation techniques in the study of dynamics of flexible proteins. Previous works have shown that the performance of these computational approaches can be improved through the definition of hybrid restraint potentials, where molecular mechanics force fields are combined with experimental data. Using NMR chemical shifts (CSs), we developed a hybrid restraint potential based on NapShift, a machine learning model trained to predict chemical shifts from protein sequence and structure. The functioning of our method was evaluated in basin-hopping (BH) global optimization, energy landscape (EL) sampling and molecular dynamics (MD) simulations with promising results.

METHODS

We developed NapShift, a program in python based on simple feed-forward, fully-connected, artificial neural networks (ANNs) to predict CS values for each backbone atom (N, C, Ca, Cβ, H and Ha) using protein sequence and dihedral angles of tripeptides as inputs.

RESULTS

In comparison with other CS predicting methods, NapShift shows improved accuracy on 250 PDB files used as the test set.

The performance of the NapShift hybrid restraint potential was tested in BH global optimization, EL analysis and in MD simulations. The following results come from one of different systems explored: tryptophan 1 (trpzip), a 12-residue peptide that folds into a β-hairpin with specific packing of tryptophan sidechains. Using trpzip distorted conformations as the starting point, the NapShift CS restraints implementation was able to lead to better structure prediction via BH and increased local stability within MD simulations.

CONCLUSIONS

• Napshift is an accurate CS prediction method.
• NapShift ANN functions are fully differentiable with respect to atomic coordinates, allowing to use this method for structural refinement.
• The implementation of NapShift hybrid potential into a molecular mechanics force field improves the efficiency of BH global optimization, EL analysis and MD simulations.

REFERENCES