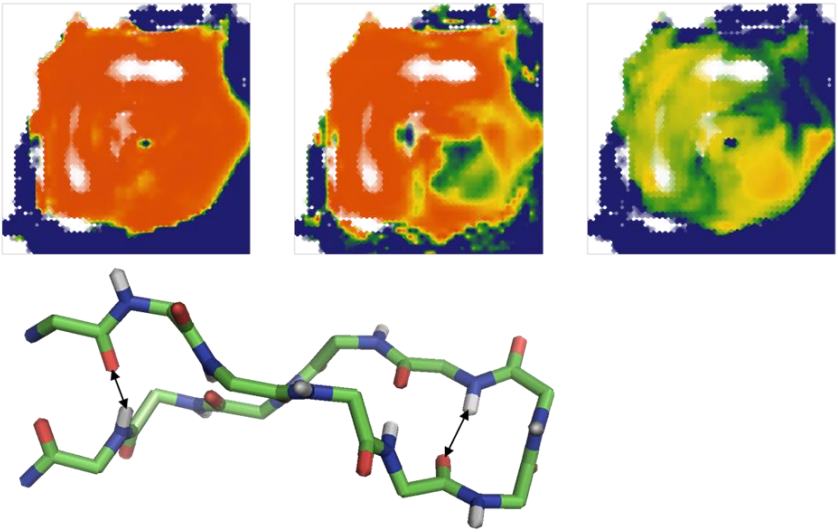
D.Horvath[1]
I. Baskin[2]
G. Marcou[1]
A.Varnek[1]

GENERATIVE TOPOGRAPHIC MAPPING OF CONFORMATIONAL SPACE

[1] Laboratoire de Chémoinformatique, UMR 7140 CNRS-Univ. Strasbourg, 1 rue Blaise Pascal, Strasbourg 67000, France
[2] Moscow State University, Russia
dhorvath@unistra.fr

Herein, Generative Topographic Mapping (GTM) [1] was challenged to produce planar projections of the high-dimensional conformational space of complex molecules (the 1LE1 peptide). GTM is a probability-based mapping strategy, and its capacity to support property prediction models serves to objectively assess map quality (in terms of regression statistics). The properties to predict were total, non-bonded and contact energies, surface area and fingerprint darkness. Map building and selection was controlled by a previously introduced evolutionary strategy allowed to choose the best-suited conformational descriptors, options including classical terms and novel atom-centric autocorrellograms. The latter condensate interatomic distance patterns into descriptors of rather low dimensionality, yet precise enough to differentiate between close favorable contacts and atom clashes. A subset of 20K conformers of the 1LE1 peptide, randomly selected from a pool of 2M geometries (generated by the S4MPLE tool [2]) was employed for map building and cross-validation of property regression models.

The GTM build-up challenge reached robust three-fold cross-validated determination coefficients of Q^2=0.7…0.8, for all modeled properties. Mapping of the full 2M conformer set produced intuitive and information-rich property landscapes. Functional and folding subspaces appear as well-separated zones, even though RMSD with respect to the PDB structure was never used as a selection criterion of the maps.



1. Kireeva N. et al. *Mol. Inf.*, 2012, **31**: 301-312.
2. Hoffer L. et al. *Molecules (Basel, Switzerland)*,2015, **20**: 8997-9028.