HAL
open science

SenSaaS: Shape-based Alignment by Registration of Colored Point-based Surfaces

Dominique Douguet, Frédéric Payan, Lucas Grandmougin

To cite this version:

Dominique Douguet, Frédéric Payan, Lucas Grandmougin. SenSaaS: Shape-based Alignment by Registration of Colored Point-based Surfaces. 22th GGMM (Group of Graphism and Molecular Modeling) and the 10th SFCi (French Society of Chemoinformatics), Sep 2021, Lille, France. hal-03519840

HAL Id: hal-03519840

https://hal.science/hal-03519840v1

Submitted on 10 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

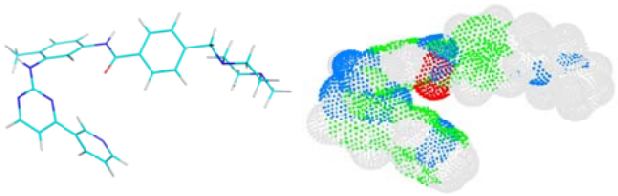
SENSAAS: Shape-based Alignment by Registration of Colored Point-based Surfaces

Dominique Douguet[1], Frédéric Payan[2], Lucas Grandmougin[2]
[1] Institut de Pharmacologie Moléculaire et Cellulaire CNRS UMR 7275, Valbonne, France
[2] Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis CNRS UMR7271, Sophia Antipolis, France

Introduction

3D point clouds or 3D meshes are data structures used in many fields well known by large audience (robotics, 3D reconstruction, games, autonomous navigation...) to model surfaces or volumes. Such 3D representations can also be relevant in chemistry to describe molecules although they were not the most used so far[1].

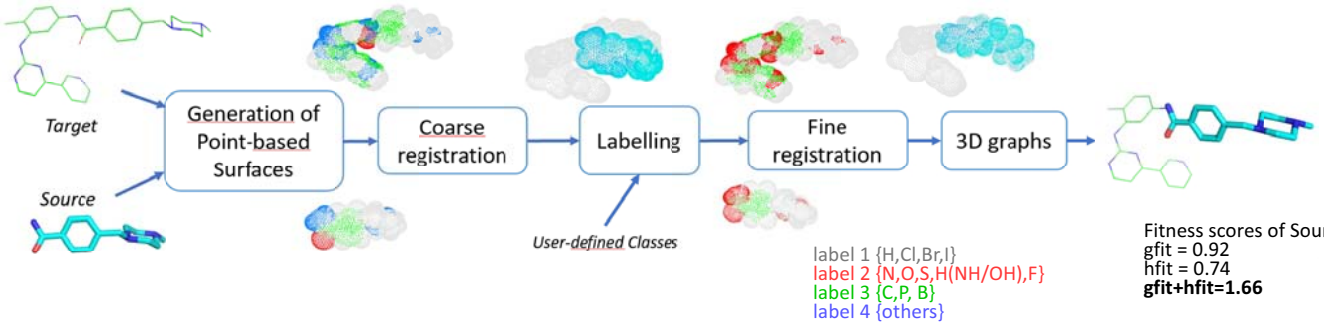
SENSAAS (SENsitive Surface As A Shape) is a shape-based alignment software using 3D point-based representation of the van der Waals surface[2]. SENSAAS is an original tool that combines methods dedicated to 3D registration[3], initially developed for the fusion of 3D point clouds collected by devices such as depth cameras or LiDAR scanners.



3D graph 3D point cloud

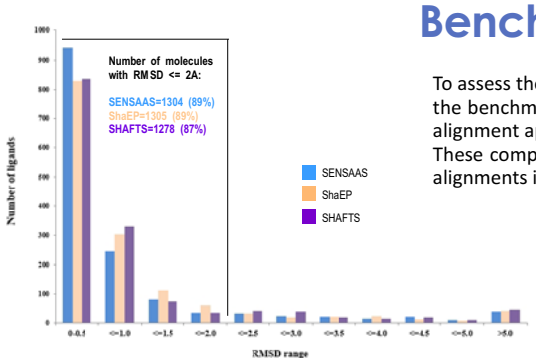
Registration of 3D Point Clouds

Let us consider two molecules that we want to superimpose in order to find similarities. SENSAAS gives a transformation matrix as output, leading to the 3D alignment of the Source molecule on the Target molecule (which is kept fixed). The general idea is to align surface representations of the two molecules, to then compute a similarity score.



The last step results into a final transformation matrix (rotation + translation), that is applied to the Source molecule to get the final alignment of 3D graphs. In parallel, fitness scores similar to Tversky coefficients are proposed to evaluate the embedding of a point cloud into another one.

Benchmarking AZ data set



To assess the efficiency of SENSAAS, we tested its ability to reproduce the superimposition of X-ray structures of the benchmarking AstraZeneca (AZ) data set and, compared its results with those generated by the two shape-alignment approaches ShaEP[4] and SHAFTS[5].
These comparisons showed that SENSAAS is as accurate as reference methods but, it generates more accurate alignments in the first precision interval (RMSD <= 0.5 Å).

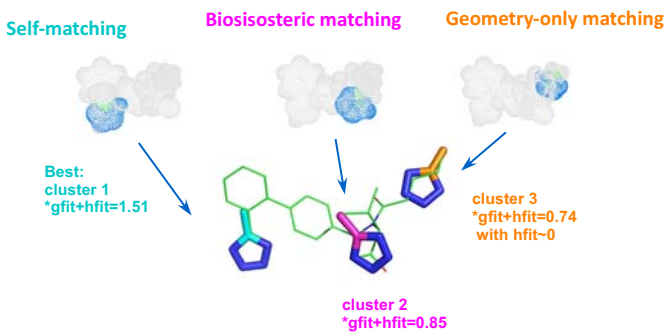
		Reproduced ligands (RMSD ≤ 2.0 Å)			Mean RMSD		
Category	Number of ligands	SENSAAS	ShaEP	SHAFTS	SENSAAS	ShaEP	SHAFTS
easy	185	**183 (98.9%)**	178 (96.2%)	181 (97.8%)	**0.37**	0.5	0.46
moderate	991	**927 (93.5%)**	923 (93.1%)	922 (93%)	**0.8**	0.88	0.86
hard	190	148 (77.9%)	**151 (79.4%)**	144 (75.8%)	**1.65**	1.8	1.71
unfeasible	99	46 (46.4%)	**53 (53.5%)**	31 (31.3%)	2.69	**2.58**	3.47
Total =	**1465 (100%)**	1304 (89%)	**1305 (89%)**	1278 (87%)	**0.97**	1.06	1.09
average run time / pairwise alignment		5.7s	0.51s	**0.46s**			

Submatching properties

We also showed that SENSAAS is able to identify local similarities.

SENSAAS possesses sub-matching properties which allow to align substructures or small fragments on a large molecule.

In the following example, three distinct alignments of a tetrazole fragment on the drug Valsartan are displayed:



* score (gfit+hfit) ranges between 0 (not aligned) and 2 (perfect matching)

Conclusion

SENSAAS provides a relevant contribution to shape-based alignment methods, especially in the field of lead optimization where scaffold hopping and bioisosteric replacement properties of a method are out of importance to identify promising compounds.

- The code and documentation are available on GitHub[6].

- Tutorials for installation and use are available on YouTube[7].

- A demo is available at https://chemoinfo.ipmc.cnrs.fr/SENSAAS

References:
1 Kumar, A. and Zhang, K. Y. J., A, *Front Chem* **2018**, 6, 315.
2 Douguet D. and Payan F., *Mol Inform.* **2020**, 8, 2000081.
3 Q.-Y. Zhou, J. Park, V. Koltun, arXiv 2018, 1801.09847. http://www.open3d.org
4 M. J. Vainio, J. S. Puranen, M. S. Johnson, *J. Chem. Inf. Model.* **2009**, 49, 492–502.
5 X. Liu, H. Jiang, H. Li, *J. Chem. Inf. Model.* **2011**, 51, 2372–2385.
6 https://github.com/SENSAAS/sensaas
7 https://www.youtube.com/channel/UC3cjM1j8cQ-95ev0DNxRMOA