

Poster presentation

Open Access

Visualisation and exploitation of the chemical space covered by patents

C Tyrchan* and S Muresan

Address: AstraZeneca R&D, DECS Global Compound Sciences, SC2, Pepparedsleden 1, 43183 Mölndal, Sweden

* Corresponding author

from 4th German Conference on Chemoinformatics
Goslar, Germany. 9–11 November 2008

Published: 5 June 2009

Chemistry Central Journal 2009, **3**(Suppl 1):P12 doi:10.1186/1752-153X-3-S1-P12

This abstract is available from: <http://www.journal.chemistrycentral.com/content/3/S1/P12>

© 2009 Tyrchan and Muresan; licensee BioMed Central Ltd.

Understanding the chemical space covered by relevant patents is an important step in early stage medicinal chemistry projects. The claimed chemical space is usually defined by generic Markush structures which are exemplified by a number of synthesized compounds with associated biological profiles.

Extracting chemical entities (manual, semi-automatic or fully automatic) from patents is not a trivial task. Once the chemical space of a patent (prophetic and/or exemplified compounds) is extracted, computational methods can be applied for data exploitation and visualisation. Here, we present two chemoinformatic tools for structural space and structural activity relationship (SAR) analysis.

To visualise the chemical space of a patent we have used ChemGPS [1] combined with Spotfire [2]. ChemGPS maps the absolute position of compounds in a PCA-derived multidimensional space defined by a set of 423 reference compounds and 72 physicochemical descriptors. To exploit the SAR information from patents we have developed a protocol using SciTegic's PipelinePilot [3]. The compounds (structures and associated biological data) are exported from IBEX [4].

To exemplify our approaches, we have extracted from IBEX four GPCR patents WO2006023462, WO2007122156, US20060019998 and US20070208005. The SAR analysis of a patent takes, at most, a couple of minutes as does the calculation of the ChemGPS map. The output of the PipelinePilot protocol can be imported in other software for further analysis.

The results clearly show the benefit of such automated patent analysis methods when combined with appropriate data sources. They can help medicinal chemists to understand and quickly overview the SAR content of a patent for an efficient decision making process.

References

1. Oprea TI, Gottfries J: **Chemography: The Art of Navigating in Chemical Space.** *J Comb Chem* 2001, **3**:157-166. [<http://spotfire.tibco.com/Products/Default.aspx>].
2. [<http://accelrys.com/products/scitegic>].
3. [<http://www.gvkbio.com/informatics.html>].
- 4.