

# ChemGPS-NP – tuned for navigation in biologically relevant chemical space

*Josefin Larsson,<sup>†</sup> Johan Gottfries,<sup>‡</sup> Sorel Muresan,<sup>§</sup> and Anders Backlund<sup>\*,†</sup>*

Division of Pharmacognosy, Department of Medicinal Chemistry, BMC, Uppsala University, Box 574, S-751 23 Uppsala, Sweden, Department of Medicinal Chemistry, AstraZeneca R&D Mölndal, S-431 83 Mölndal, Sweden, and GDECS Computational Chemistry, AstraZeneca R&D Mölndal, S-431 83 Mölndal, Sweden.

---

\* To whom correspondence should be addressed. Tel: +46-18-4714498. Fax: +46-18-509101.

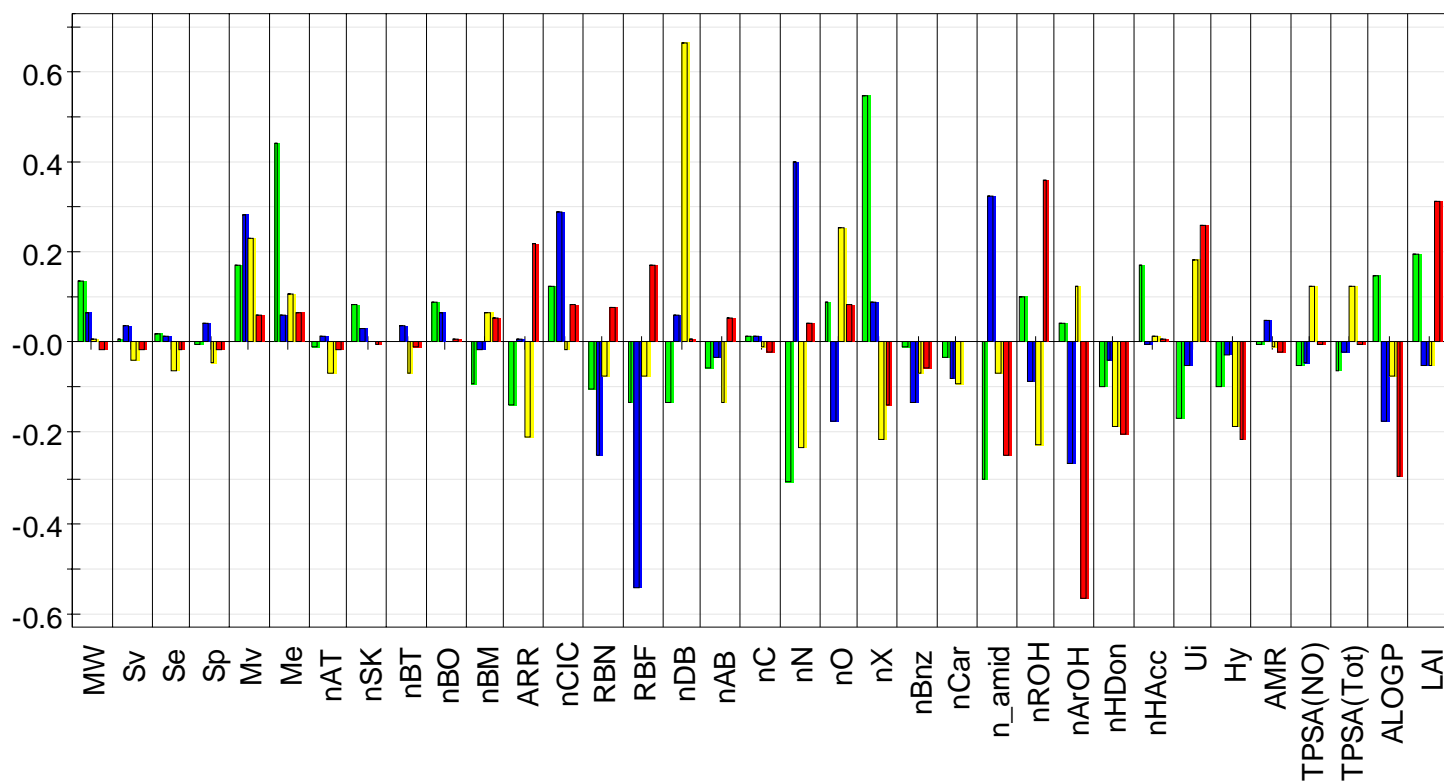
Email: anders.backlund@fkog.uu.se.

<sup>†</sup> Uppsala University.

<sup>‡</sup> Department of Medicinal Chemistry, AstraZeneca R&D Mölndal.

<sup>§</sup> GDECS Computational Chemistry, AstraZeneca R&D Mölndal.

## Supporting information



**Supplementary Figure S1. Interpretation of PC5-PC8.** Loading for each of the 35 descriptors in PC5 (green), PC6 (blue), PC7 (yellow), and PC8 (red). The Y-axis denotes loading and the X-axis gives the molecular descriptors.

**Supplementary Table 1.** The datasets used to develop ChemGPS-NP v.1

Database, and number of compounds	Reference
Dictionary of Natural Products ChemGPS	167,169 551 Chapman & Hall/CRC Press, <a href="http://www.chemnetbase.com">www.chemnetbase.com</a> Oprea, T.I. & Gottfries, J. <i>J Comb Chem</i> <b>3</b> , 157-166 (2001)
Amino acids	98 Doctoral thesis, Sandberg, M. (1998) <a href="http://www.chem.umu.se">www.chem.umu.se</a>
ASINEX platinum collection	128,140 <a href="http://www.asinex.com">www.asinex.com</a>
Bioactives	4,635 <a href="http://chembank.broad.harvard.edu/data_downloads/">chembank.broad.harvard.edu/data_downloads/</a>
Bionet Screening database	43,410 <a href="http://www.keyorganics.ltd.uk/scdownloads.htm">www.keyorganics.ltd.uk/scdownloads.htm</a>
Cytotoxic NPs	31 Lee, K.-H. <i>Med Res Rev</i> <b>19</b> , 569-596 (1999)
Cytotoxic NPs	144 Kim J. & Park, E.J. <i>Curr Med Chem Anti-canc Agents</i> <b>2</b> , 485-537 (2002)
Flavors	338 <a href="http://www.flavornet.org">www.flavornet.org</a>
G&J Chemicals Screening	1,620 <a href="http://www.keyorganics.ltd.uk/scdownloads.htm">www.keyorganics.ltd.uk/scdownloads.htm</a>
Guppy NPs	158 <a href="http://www.mosmedchemlabs.com">www.mosmedchemlabs.com</a>
Hitfinder	16,000 <a href="http://www.maybridge.com">www.maybridge.com</a>
IBS2005n	42,349 <a href="http://www.ibscreen.com">www.ibscreen.com</a>
IBS_Natural_update	6,457 <a href="http://www.ibscreen.com">www.ibscreen.com</a>
Ligand_3D_9	7,494 <a href="http://www.timtec.net">www.timtec.net</a>
Maybridge	61 <a href="http://www.maybridge.com">www.maybridge.com</a>
MB_build_blocks_eur	12,466 <a href="http://www.maybridge.com">www.maybridge.com</a>
MB_HitFinder_V4	16,000 <a href="http://www.maybridge.com">www.maybridge.com</a>
MB_Screening_Collection	58,855 <a href="http://www.maybridge.com">www.maybridge.com</a>
Mdpi46	514 <a href="http://www.mdpi.org">www.mdpi.org</a>
Merlin Screening	2,602 <a href="http://www.keyorganics.ltd.uk/scdownloads.htm">www.keyorganics.ltd.uk/scdownloads.htm</a>
NATPROD	683 <a href="http://www.msdiscovery.com">www.msdiscovery.com</a> (private communication)
NatProd Collection	720 <a href="http://www.msdiscovery.com/download.html">www.msdiscovery.com/download.html</a>
Natural_552	552 <a href="http://www.chemdiv.com">www.chemdiv.com</a> (private communication)
Natural cyclooxygenase inhibitors	248 Larsson, J., Gottfries, J., Bohlin, L. & Backlund, A. <i>J Nat Prod</i> <b>68</b> , 985-991 (2005)
NATURAL_PROD_EXTENDED	3,342 <a href="http://www.timtec.net">www.timtec.net</a>
NATx_A_what's new	1,014 <a href="http://www.ac-discovery.com">www.ac-discovery.com</a>
NATx_A	7,838 <a href="http://www.ac-discovery.com">www.ac-discovery.com</a>
NATx_HD	633 <a href="http://www.ac-discovery.com">www.ac-discovery.com</a>
NC_15933	15,933 <a href="http://www.chemdiv.com">www.chemdiv.com</a>
NCI	250,251 <a href="http://cactus.nci.nih.gov/ncidb2/download.html">cactus.nci.nih.gov/ncidb2/download.html</a>
NPs fkog	979 Master thesis, Baumgarten, A., div. of Pharmacognosy, Uppsala University (2005)
Pesticides	1,657 <a href="http://www.bcp.org/epesticidemanual/main.html">www.bcp.org/epesticidemanual/main.html</a>
Pharma	46,707 <a href="http://www.nanosyn.com">www.nanosyn.com</a>
Pheromones	511 <a href="http://www.pherolist.slu.se/pherolist.php">www.pherolist.slu.se/pherolist.php</a>
Poly-all	2,339 <a href="http://www.acbblocks.com">www.acbblocks.com</a>
ScreeningCollection	59,493 <a href="http://www.maybridge.com">www.maybridge.com</a>
Semi_natural_6049	6,049 <a href="http://www.chemdiv.com">www.chemdiv.com</a> (private communication)
Sesquiterpene lactones	1,389 Master thesis, Ottosson, H., div. of Pharmacognosy, Uppsala University (2005)
SHARK	103,208 <a href="http://www.mosmedchemlabs.com">www.mosmedchemlabs.com</a>
Specs_Aug2005	604 <a href="http://www.specs.net">www.specs.net</a>
Specs_NP_1mg_April	381 <a href="http://www.specs.net">www.specs.net</a>
Specs_NP_1mg_Aug	606 <a href="http://www.specs.net">www.specs.net</a>
Specs_SMILES	400 <a href="http://www.specs.net">www.specs.net</a>
STARFISH	73,675 <a href="http://www.mosmedchemlabs.com">www.mosmedchemlabs.com</a>
TDS001	10,460 <a href="http://www.timtec.net">www.timtec.net</a>