Development and validation of methodology for designing and analysis of virtual combinatorial libraries based on defined reaction pathways

Rafał Kurczab[a], Paweł Zajdel[b], Maciej Pawłowski[b], Andrzej J. Bojarski[a]

[a]Department of Medicinal Chemistry, Institute of Pharmacology Polish Academy of Sciences, 12 Smętna Street, 31-343 Kraków

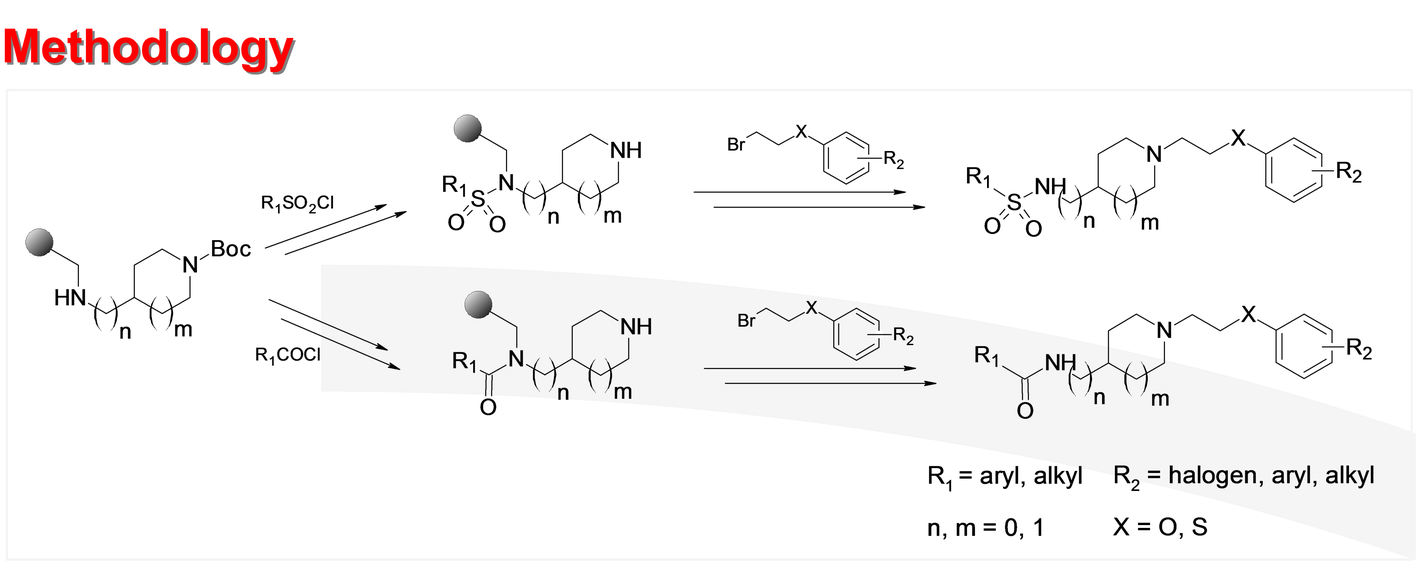
[b]Department of Medicinal Chemistry, Jagiellonian University Medical College, 9 Medyczna Street, 30-688 Kraków

Introduction

New synthetic methods in combinatorial chemistry, such as parallel solid-phase synthesis, enable preparation of large number of compounds for drug development screening campaigns in a very fast and efficient way. Large libraries of compounds are usually synthesized as combinations of different building blocks (BB's) [1,2]. However, the number of compounds that can be synthesized using elaborated synthesis protocol, and tested for biological activity (even in HTS) is often limited by the project's budget. In that case, methodology enabling prioritization and helping in decision making of which compound, from the available synthetic space, should be obtained first, would be very useful.

Herein we present a new approach to generating and ranking of the virtual combinatorial library based on defined chemical reactions. For this prupose the different chem– and bioinformatic methods and tools were used and combined in multistep protocol.

Methodology



Defined synthetic pathway

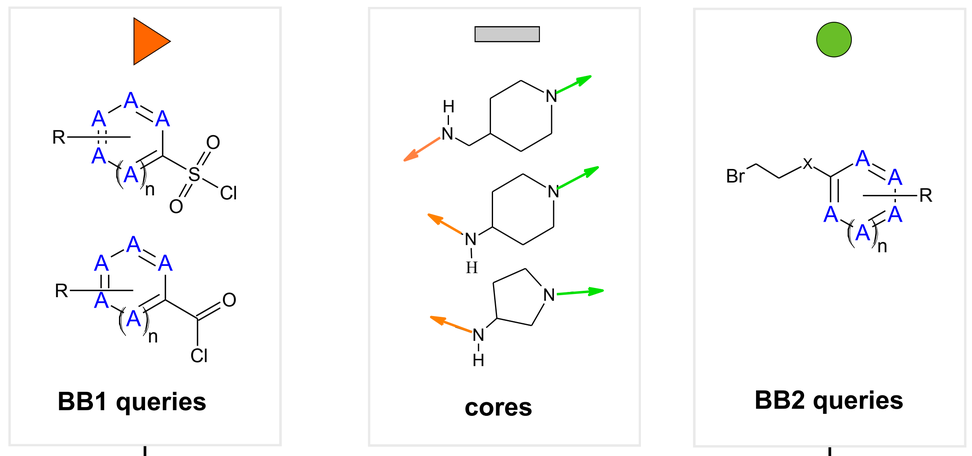
The elaborated solid-phase methodology employed attachment of the primary core amines to the BAL-MBHA -PS resin, treatment of the resin-bound Boc-protected amines with sulfonyl or acyl chlorides (BB1), subsequent Boc removal, and final alkylation of the secondary amines with halogen derivatives (BB2) [3].

Definition of automatic molecular query

Two sets of queries were defined, first (BB1 queries) describing molecular pattern for different substituted aromatic and heteroaromatic sulfonyl and acyl chlorides, second for aryloxyalkyl- or arylthioalkyl- halides (BB2 queries).

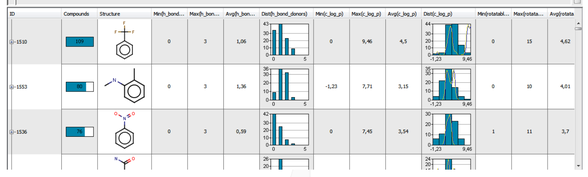
Building Blocks Database Searching

The database (26 commercial Vendors adapted) containing over 8.5 M 'stock available' building blocks and intermaediates was screened using substructure searching algorithm and defined queries.



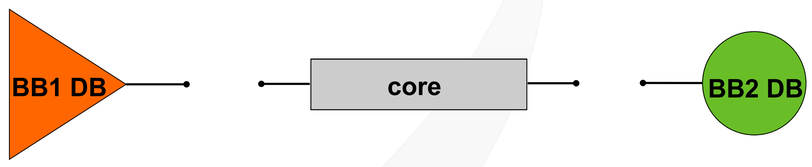
Buidling Blocks Clustering

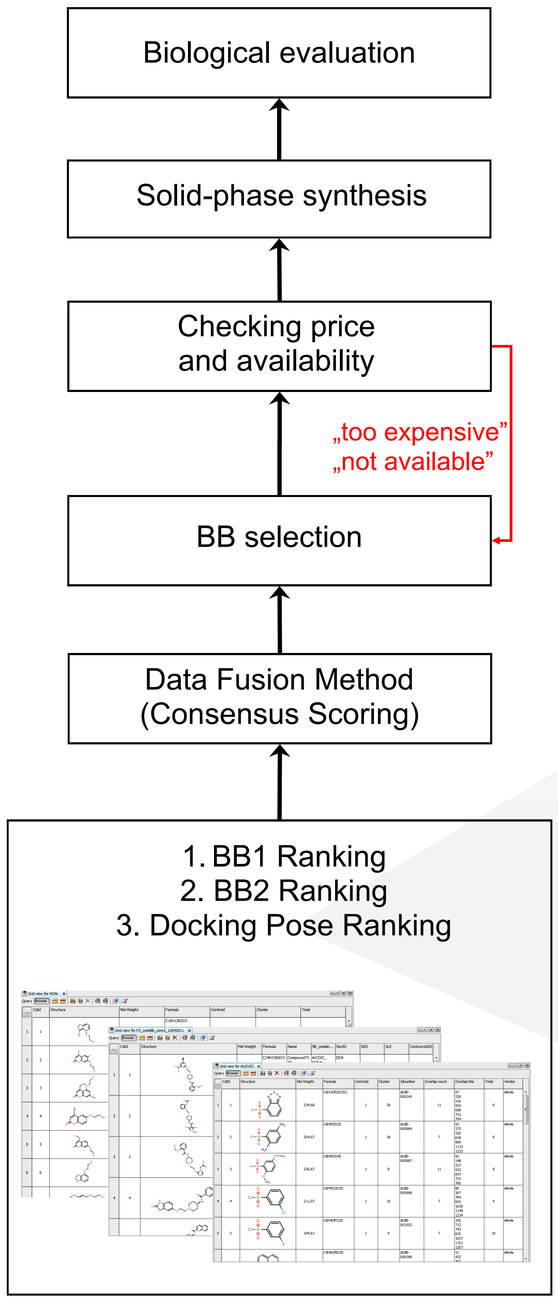
Selected subsets (BB1 and BB2) were clustered hierarchically using Chemical Hashed Fingerprint and Tanimoto metric. For each cluster the proportional number of building blocks to its size were chosen to the next phase.

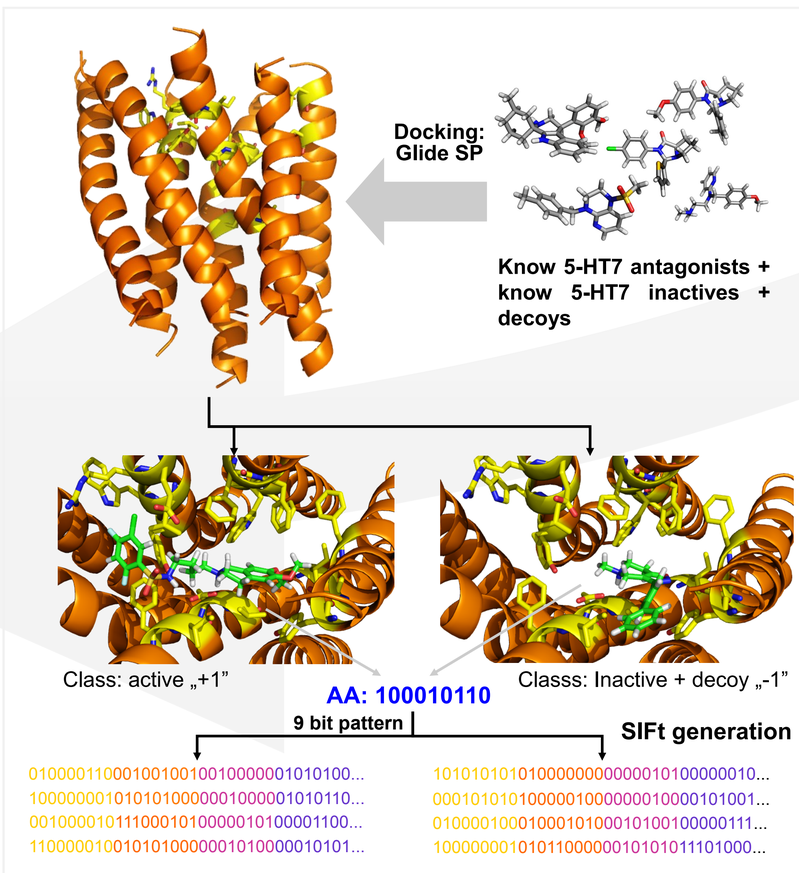


Generation of Virtual Combinatorial Library

Finally, selected building blocks databases BB1 and BB2 were iteratively combined with each core. In this way, all possible combinations, i.e. virtual compounds, were produced.







Multistep Virtual Screening Protocol

Physiochemical Filter (Lipinski Rule of 5, Veber rules, pKa)

ADMET Filter (reactive functional groups, aqueous solubility, gut-blood barrier, blood-brain partition coefficient)

3D Pharmacophore Models (a linear combination of 6 different pharmacophore models of 5-HT7 antagonists)

Docking Protocol (cross-docking to 6 different conformations of 5-HT7 homology models with ASP3.32 spatial constrain)

SVM Prediction Model (the Structural Interaction Fingerprints generated for docked know actives and inactives were used to build SVM (support vector machine) predictive models. They were further used to rank the combinatorial compounds docked to 6 receptor conformations).

References

[1] Kubinyi, H. Combinatorial and computational approaches in structure-based drug design, *Curr. Opin. Drug Disc. Dev.*, **1998**, *1*, 16.

[2] Boehm, M.; Wu, T. Y.; Claussen, H.; Lemmen, C. Similarity Searching and Scaffold Hopping in Synthetically Accessible Combinatorial Chemistry Spaces, *J. Med. Chem.*, **2008**, *51*, 2468.

[3] Zajdel, P.; Grychowska, K.; Pawłowski, M. Solid-phase synthesis of sulfonamide derivatives of differently substituted alkylamines in the search for CNS receptor ligands. III Konwersatorium Chemii Medycznej, P-55, Lublin 2010.

Acknowledgments

This study was partly supported by a grant PNRF-103-AI-1/07 from Norway through the Norwegian Financial Mechanism and MNiSW Grant No. N N405 671540.

Polish-Norwegian Research Fund
www.cns-platform.eu
norway grants