



**HAL**  
open science

# Overlap of On-demand Ultra-large Combinatorial Spaces with On-the-shelf Drug-like Libraries

Mariana Perebyinis, Didier Rognan

► **To cite this version:**

Mariana Perebyinis, Didier Rognan. Overlap of On-demand Ultra-large Combinatorial Spaces with On-the-shelf Drug-like Libraries. *Molecular Informatics*, In press, 10.1002/minf.202200163 . hal-03830333

**HAL Id: hal-03830333**

**<https://hal.science/hal-03830333v1>**

Submitted on 26 Oct 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.

DOI: 10.1002/minf.200((full DOI will be filled in by the editorial staff))

# Overlap of on-demand ultra-large combinatorial spaces with on-the-shelf drug-like libraries.

Mariana Perebyinis,<sup>[a]</sup> and Didier Rognan\*<sup>[a]</sup>

**Abstract:** On-demand combinatorial spaces are shifting paradigms in early drug discovery, by considerably increasing the searchable chemical space to several billions of compounds while securing their synthetic accessibility. We here systematically compared the on-the-shelf available drug-like chemical space (9 million compounds) to three on-demand ultra-large (ODUL) combinatorial fragment spaces (REAL, CHEMriya, GalaXi) covering 32 billion of readily accessible molecules. Surprisingly, only one space (REAL)

intersects almost entirely the currently available drug-like space, suggesting that it is the only ODUL widely suitable for in-stock hit expansion. Of course, expanding a preliminary ODUL hit in the same chemical space is the best possible strategy to rapidly generate structure-activity relationships. All three spaces remain well suited to early hit finding initiatives since they all provide numerous unique scaffolds that are not described by on-the shelf collections.

**Keywords:** drug discovery, compound libraries, chemical space, hit finding, hit expansion

The pressure to identify chemically novel and patentable drug candidates has led the pharmaceutical industry, at the turn of the century, to massively invest in miniaturized high-throughput screening facilities and assemble legacy collections of millions of drug-like compounds.<sup>[1]</sup> In parallel, chemical suppliers have quickly offered commercially available collections of comparable sizes,<sup>[2]</sup> notably available to the academia for virtual screening and follow-up cherry picking.<sup>[3]</sup>

Despite the 10-fold increase in the size of such screening decks within 20 years, the drug-like chemical space available in-stock for screening (ca. 15 million unique compounds)<sup>[4]</sup> is just a drop in the ocean of synthesizable drug-like compounds ranging from  $10^{20}$  to  $10^{63}$  according to various estimates.<sup>[5]</sup> As a first attempt to increase drug-like space by several orders of magnitude, Reymond et al. enumerated, from simple molecular graphs, all possible organic drug-like molecules with an upper limit of atom numbers and elements, leading to the GDB-17 database of 166 billion molecules.<sup>[6]</sup> Such giant databases, although useful to discover novel bioactive compounds,<sup>[7]</sup> have three main drawbacks: (i) corresponding compounds are not physically available, (ii) their synthesis remain problematic as so far as no retrosynthetic schemes are provided, (iii) their practical usage requires enormous computing resources for just storage and virtual screening. To address these issues, Pfizer pioneered the design of an organic-reaction driven chemical space featuring  $10^{18}$  virtual compounds.<sup>[8]</sup> Instead of enumerating all possible molecules, the corresponding chemical space is combinatorially encoded by both the building blocks and multicomponent organic reactions necessary to synthesize them,<sup>[9]</sup> thereby enabling a high synthesis success rate (ca. 80%)<sup>[9-10]</sup> and multibillion sizes. Both proprietary and commercial fragment spaces have accordingly been described peaking at an impressive number of  $10^{26}$  tangible compounds for the GSK XXL space.<sup>[11]</sup> Mining such gigantic spaces requires dedicated

descriptors and algorithms<sup>[11]</sup> for 2D similarity searches,<sup>[12]</sup> space comparisons<sup>[13]</sup> and quite-recently for maximum common substructure (MCS) searches.<sup>[14]</sup> Interestingly, on-demand combinatorial spaces weakly overlap (usually between 0.2 and 2%)<sup>[13]</sup> and thereby constitutes orthogonal chemical spaces for either preliminary hit finding<sup>[15]</sup> or minute-scale identification of analogues from in-stock compounds (SAR by Space)<sup>[16]</sup> that can be readily synthesized and tested. The democratization of fragment spaces leads to the question of their most suitable usage in early drug discovery. Shall they be reserved to hit expansion or directly used for hit identification? Up to now, initial hit finding relying on virtual screens has been mostly restricted to commercial screening libraries of 'on-the-shelf' compounds. The chemical overlap between on-demand combinatorial spaces and commercial drug-like screening decks is however unknown to date. The only reported attempt to decipher this overlap just concerned 100 representative drugs.<sup>[12b]</sup>

We therefore systematically studied the overlap between a representative library of nine million commercially available in-stock drug-like compounds and three ultra-large combinatorial spaces totalling 30 billion on-demand compounds, to assess whether such fragment spaces contain known drug-like scaffolds for immediate hit finding or expansion strategies. To achieve this goal, we first updated the in-house Bioinfo-DB database, a repository of in-stock drug-like screening compounds successfully used in the laboratory for virtual screening

---

[a] M. Perebyinis, D. Rognan  
Laboratoire d'Innovation Thérapeutique, UMR7200  
CNRS-Université de Strasbourg  
74 route du Rhin, F-67400 Illkirch  
\*e-mail: rognan@unistra.fr  
phone: +33 3 68 85 42 35  
fax: +33 3 68 85 43 10



Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/minf>. ((DOI will be filled in by the editorial staff)).

## Drug-likeness of ultra-large combinatorial spaces

campaigns.<sup>[17]</sup> In its current version, it relies on 25 trustable suppliers (Table S1) selected for their ability to provide requested compounds within 3-4 weeks with an acceptable price-quality ratio (typically 5 mg of >90% pure powder for 50 €). An automated workflow was design to read, clean and process supplier files, notably removing duplicates (while keeping track of all possible suppliers), ionising at physiological pH and filtering for drug-likeness (Figure S1). Along the process, the number of available compounds decreased from 21 to 9.1 million (Table S2), a number inferior to that reported in ZINC20,<sup>[4]</sup> but better representative to our viewpoint of the currently purchasable drug-like (CPDL) space. To limit the number of queries, the CPDL space was next restricted to their 2.4 million unique Bemis-Murcko (BM) frameworks.<sup>[18]</sup> Compounds sharing the same BM framework are likely to share the same synthetic route. Pairwise comparison of CPDL BM frameworks with on-demand ultra-large fragment spaces (ODULs), also organized according their most tangible synthetic access, makes therefore some sense from a medicinal chemistry perspective.

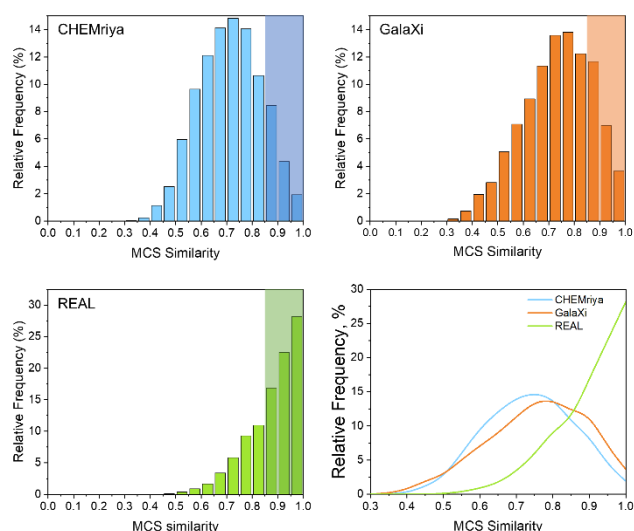
In the current work, three ODULs, encoded as fragment spaces were selected: Enamine REAL space,<sup>[19]</sup> OTAVA CHEMriya<sup>[20]</sup> and WuXi GalaXi.<sup>[21]</sup> REAL space is the first ODUL to have been reported in 2007 by the Ukrainian supplier Enamine.<sup>[22]</sup> It currently comprises 29 billion compounds obtained via more than 170 well-validated parallel synthesis protocols applied to over 112 000 qualified reagents and building blocks. The average synthesis success of REAL compounds is ca. 80% and most of them are obtained in one-pot synthesis (one or two steps) in less than 8 weeks.<sup>[9]</sup> Due to its high practical availability, REAL space (or subsets of it) has been frequently screened in silico the last 3 years, with remarkable successes in terms of chemical novelty and potency of primary hits.<sup>[15]</sup>

GalaXi space<sup>[21]</sup> was reported first in October 2019, as a research collaboration between WuXi AppTech, a Chinese CRDMO (Contract Research, Development and Manufacturing Organization), and the German software editor BioSolveIT. GalaXi space comprises 1.8 billion virtual compounds developed using various building block couplings that utilize more than 20 different types of reactions and no more than 2 synthetic steps. Most compounds were defined using novel WuXi scaffolds combined with in-stock building blocks.

The first release of the CHEMriya space<sup>[20]</sup> was reported in June 2021 by the Canadian supplier Otava Chemicals, as a research collaboration with BioSolveIT, and contains 12 billion accessible on-demand molecules, based on 30 000 building blocks and 44 in-house reactions. Several multi-component and ring-closure reactions are claimed by the supplier to provide a vast chemical diversity to this space, with an average synthesis time of 4-6 weeks. Neither GalaXi nor CHEMriya Spaces, have been publicly reported yet to yield bioactive hit compounds.

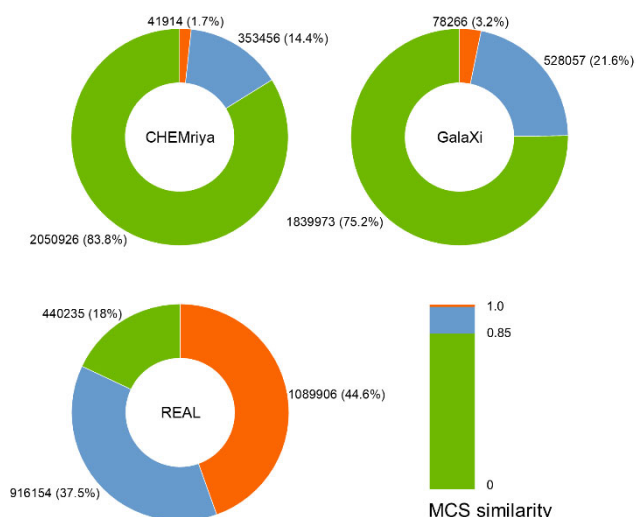
All three spaces are available at BioSolveIT<sup>[23]</sup> and encoded as fragment spaces connectable by specific organic reactions. Neither the fragments (building blocks) nor the organic reactions to combine them, are freely available. Due to their huge size, current cheminformatics software and computing facilities are unable to fully enumerate the three corresponding spaces. The exact BM scaffold composition of the three ODULs was therefore unknown to us.

Chemical overlap between CPDL and ODUL spaces was addressed at the MCS level, thanks to the recently described SpaceMACS algorithm,<sup>[14]</sup> that enables to directly interrogate a fragment combinatorial space and estimate MCS similarity to any SMILES string. In the current study, only the most similar ODUL compound to each CPDL BM scaffold (highest MCS similarity) was retained. The distribution of the 2.4 million similarity values is strikingly different across the three ODULs (Figure 1).



**Figure 1.** Distribution of the pairwise MCS similarity between 2.4 million CPDL Bemis-Murcko scaffolds and the most similar ODUL compound. Frequencies apply to CPDL scaffolds. Shaded sections indicate similarity values higher than 0.85.

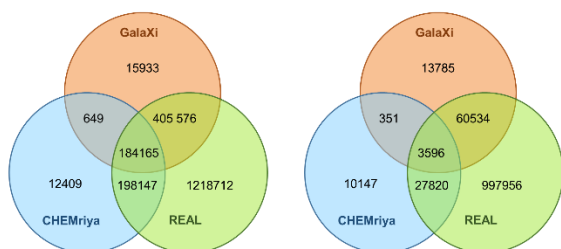
CHEMriya and GalaXi hits follow an overall normal distribution with a mean around MCS similarity values of 0.70 and 0.75, respectively. REAL hits are clearly shifted towards significantly higher similarity values with a half-normal distribution peaking at the maximal value of 1.0 (Figure 1). Preliminary trials prompted us to define an MCS similarity value of 0.85 to discriminate similar from dissimilar scaffolds. We acknowledge that this threshold is purely subjective, it has however the merit to match medicinal chemistry intuitions, notably with respect to bioisosterism rules (Figure S2). ODUL hits were thus classified in three categories according to their MCS similarity to CPDL scaffolds: (i) dissimilar (MCS similarity < 0.85), (ii) similar ( $0.85 \leq$  MCS similarity < 1), (iii) identical (MCS similarity = 1; Figure 2). According to this classification, only a low percentage of CPDL scaffolds have similar (or identical) counterparts in CHEMriya and GalaXi (16.2 and 24.8%, respectively; Figure 2). Conversely, 82% of pairwise comparisons to REAL compounds exceed the MCS similarity threshold of 0.85 with 44.6% of CPDL scaffolds being strictly retrieved in REAL space (Figure 2). Although we were expecting higher similarity values for the REAL space that positions itself as the largest publicly available drug-like space, such striking differences with the other two ODULs were unexpected.



**Figure 2.** Donut plot of the pairwise MCS similarity between 2.4 million CPDL Bemis-Murcko scaffolds and the most similar compounds from three ODULs (CHEMriya, GalaXi and REAL). Statistics (numbers, percentages) apply to CPDL scaffolds.

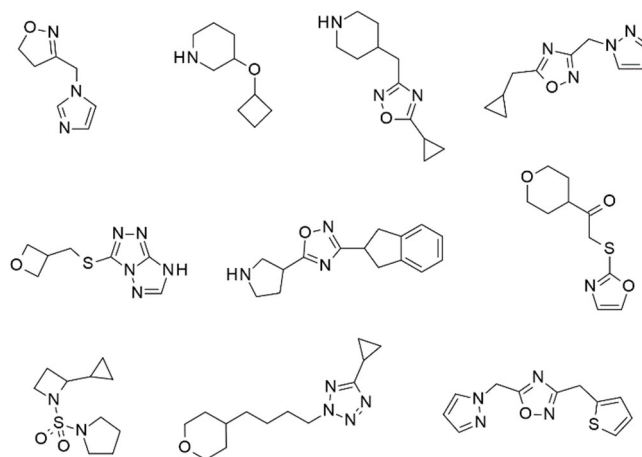
The above analysis might have been biased by the predominance of Enamine in contributing to the Bioinfo-DB CPDL space (almost 20% of the final number drug-like compounds, Table S2). Indeed, 32% of Bioinfo-DB scaffolds (811 276 out of 2 453 944) stem from Enamine compounds out of which 640 280 (25%) are unique to Enamine and not found in other Bioinfo-DB suppliers. Since it is a pretty significant number, we checked whether the removal of Enamine and Otava as suppliers in Bioinfo-DB would dramatically change the proportion of Bioinfo-DB scaffolds retrieved in the three ODULs (GalaXi needed not to be considered because the WuXi supplier does not contribute yet to Bioinfo-DB CPDL space). The distributions of MCS similarities between the new “unbiased” Bioinfo-DB and the three ODULs remain almost identical to that originally observed (Figure S3). We can therefore safely conclude that our initial analysis has not been biased by the Enamine predominance in Bioinfo-DB compounds and that only Enamine REAL space tightly overlaps with Bioinfo-DB space.

We next analysed how many ODUL hits (MCSs found similar/identical to CPDL scaffolds) overlap in chemical space by counting the number of Bioinfo-DB scaffolds uniquely found in each ODUL, and those shared by two or all three ultra-large spaces (Figure 3).



**Figure 3.** Overlap in drug-like BM scaffolds found similar (A, MCS similarity > 0.85) or identical (MCS similarity = 1) to three ODUL spaces.

A Venn diagram analysis clearly show that CHEMriya and GalaXi hits significantly overlap with REAL hits (Figure 3). Considering an MCS similarity above 0.85, most of CHEMriya and GalaXi hits are indeed shared with REAL. The number of unique hits in CHEMriya and GalaXi spaces (12 409 and 15 933) is much lower than that observed with REAL that achieves to gather an impressive number of unique MCS hits (1 218 172, 60%) not found in the two other spaces. We could identify 12 284 BM scaffolds with a MCS similarity of 1.0 to REAL space but lower than 0.5 for both CHEMriya and GalaXi. According to our hit-likeness filter (Table S3), 12 076 of them (98%) are hit-like and worth being considered for hit finding initiatives. We could not detect any obvious bias in their molecular properties (Table S4), a randomly selected samples being shown in Figure 4.



**Figure 4.** Randomly-chosen CPDL scaffolds found in REAL space (MCS similarity = 1) but not in CHEMriya or GalaXi (MCS similarity ≤ 0.50).

Still considering a MCS similarity threshold of 0.85, only 7.5% of CPDL BM scaffolds (184 165, Figure 3A) are shared by the three spaces. Out of this limited space, only 0.14 % (3 596, Figure 3B) are strictly shared, confirming previous studies highlighting that on-demand combinatorial spaces weakly overlap.<sup>[12b, 13]</sup>

From the current study, it appears that only REAL space overlaps the chemical space currently described by in-stock drug-like compounds. This observation does not disqualify the two other combinatorial spaces (CHEMriya, GalaXi), but suggests different usages of the three ODULs. On the one hand, REAL is ideally suited to hit finding and remains clearly the combinatorial space to consider first for CPDL hit expansion because the likelihood that it contains close analogues of a preliminary drug-like hit is by far the highest. Success stories from such SAR by space have already been reported<sup>[16]</sup> and will be numerous in a very near future. On the other hand, the three ultra-large spaces are well suited to find hits in chemical spaces that are orthogonal, but not necessarily better, to the CPDL space; notably when no valuable hits have been identified from stock compounds. The question why CHEMriya and GalaXi scaffolds are not on-shelf is a difficult question. It is already known that REAL, CHEMriya and GalaXi spaces do not overlap.<sup>[13]</sup> Moreover, many on-the-shelf screening decks have been historically

### Drug-likeness of ultra-large combinatorial spaces

built from scaffolds which were exhaustively exemplified, therefore defining a restricted space. It seems that REAL just reproduces it and the other two ODULs not. We are not aware whether it was the intention of OTAVA and WuXI not to overlap it. We hypothesize that the combined use of proprietary building blocks and space-specific organic reactions (e.g. multi-component and ring-closure reactions) really defines orthogonal screening spaces. Of course, expanding a preliminary ODUL hit in the same chemical space is the best possible strategy to rapidly generate structure-activity relationships.

A limitation of the current study is that large combinatorial spaces cannot be fully enumerated. Therefore, if one knows which ODUL scaffolds are common to the CPDL space, we unfortunately could not determine the number and chemical structures of the unique ODUL scaffolds that do not intersect CPDL space, and how many compounds are described by these unique scaffolds. The latter are difficult to analyse in terms of standard drug-likeness properties. Two-dimensional similarity search in large spaces is now possible using either feature trees,<sup>[12b]</sup> topological fingerprints<sup>[12a]</sup> or maximum common substructures,<sup>[14]</sup> virtual screening tools (shape matching, pharmacophore search, docking) able to mine large combinatorial spaces are still at their infancy<sup>[15a]</sup> and need to be developed<sup>[11, 24]</sup> to fully leverage the potential of large combinatorial spaces in preliminary hit finding.

## Computational Methods

### On-the-shelf drug-like libraries (Bioinfo-DB)

The Bioinfo-DB database (v.22.1) was set-up using a fully automated Pipeline Pilot (Dassault Systèmes, Paris, France) protocol in five steps (Figure S2): (i) reading sd files from 25 suppliers and defining unique commercial catalogue identifiers and vendor's name, (ii) removing compounds with a bad topology (e.g. pentavalent carbon atoms), more than two undefined stereocenters, and stripping counter-ions, (iii) detecting duplicates after generating canonical SMILES strings and keeping track of the main supplier (first-read supplier) and back-ups, (iv) filtering for drug-likeness thanks to in-house rules (Table S3) implemented in OpenEye Filter v.2021.2.0. (OpenEye Scientific Software, Santa Fe, NM 83507; U.S.A) and subsequent ionization at pH = 7.4. Noteworthy, potential duplicates arising from the ionization step were searched again and removed, if necessary, (v) structure standardization and annotation by computed molecular descriptors (molecular weight, polar surface area, hydrogen bond donor and acceptor counts, number of rotatable bonds and rings, aqueous solubility, logP, violations of Lipinski's rule-of-five) were estimated. The final number of compounds in Bioinfo-DB v22.1 is 9 306 375, all 2D structures being saved in single sd file.

Bemis-Murcko scaffolds<sup>[18]</sup> were computed for the full Bioinfo-DB set using RDKit v. 2020.09.1.0 (<https://www.rdkit.org/>), yielding 2 453 944 unique scaffolds saved as SMILES strings.

### On-demand ultra-large combinatorial spaces (ODULs)

Three ODULs in binary fragment space format (fsf) were kindly provided by M. Gastreich (BiosolveIT GmbH, Sant-Augustin, Germany): REAL (Enamine, 2021-04 version, 19

billion compounds), CHEMriya (Otava, 2021-04 version, 11 billion compounds) and GalaXi (WuXi, 2021-11 version, 2.1 billion compounds).

### MCS similarity searches

Maximum common substructure searches were computed with SpaceMACS<sup>[14]</sup> using each Bioinfo-DB scaffold as a SMILES query and the three ODUL as searching spaces. To fasten pairwise MCS similarity searches, 246 query jobs of 10 000 SMILES strings each were submitted using a single thread (GalaXi), 16 threads (CHEMriya) and 32 threads (REAL). The most similar hit, according to pairwise MCS similarity,<sup>[14]</sup> was kept for each ODUL. Average computing times for each job ranged from 1-2 h (CHEMriya), 4-5 hours (GalaXi) to 7-11 h (REAL).

## Acknowledgements

We sincerely thank M. Gastreich (BioSolveIT, Sankt Augustin, Germany) for kindly providing commercial fragment spaces in SpaceMACS readable format, and M. Rarey (Center for Bioinformatics, Hamburg, Germany) for making a binary version of SpaceMACS available to us. The Calculation Center of the IN2P3 (CNRS, Villeurbanne, France) is acknowledged for allocation of computing time and excellent support. We last thank OpenEye Scientific Software for the generous allocation of academic licenses.

## References

- [1] J. R. Broach, J. Thorner, *Nature* **1996**, *384*, 14-16.
- [2] J. J. Irwin, B. K. Shoichet, *J. Chem. Inf. Model.* **2005**, *45*, 177-182.
- [3] K. H. Bleicher, H. J. Bohm, K. Muller, A. I. Alanine, *Nat. Rev. Drug. Discov.* **2003**, *2*, 369-378.
- [4] J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, R. A. Sayle, *J. Chem. Inf. Model.* **2020**, *60*, 6065-6073.
- [5] a) R. S. Bohacek, C. McMartin, W. C. Guida, *Med Res Rev* **1996**, *16*, 3-50; b) P. Ertl, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 374-380; c) P. G. Polishchuk, T. I. Madzhidov, A. Varnek, *J. Comput. Aided Mol. Des.* **2013**, *27*, 675-679.
- [6] L. Ruddigkeit, R. van Deursen, L. C. Blum, J. L. Reymond, *J. Chem. Inf. Model.* **2012**, *52*, 2864-2875.
- [7] K. Meier, J. Arus-Pous, J. L. Reymond, *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 2074-2077.
- [8] Q. Hu, Z. Peng, S. C. Sutton, J. Na, J. Kostrowicki, B. Yang, T. Thacher, X. Kong, S. Mattaparti, J. Z. Zhou, J. Gonzalez, M. Ramirez-Weinhouse, A. Kuki, *ACS Comb. Sci.* **2012**, *14*, 579-589.
- [9] O. O. Grygorenko, D. S. Radchenko, I. Dziuba, A. Chuprina, K. E. Gubina, Y. S. Moroz, *iScience* **2020**, *23*, 101681.
- [10] T. Hoffmann, M. Gastreich, *Drug Discov. Today* **2019**, *24*, 1148-1156.
- [11] W. A. Warr, M. C. Nicklaus, C. A. Nicolaou, M. Rarey, *J. Chem. Inf. Model.* **2022**, *62*, 2021-2034.
- [12] a) L. Bellmann, P. Penner, M. Rarey, *J Chem Inf Model* **2021**, *61*, 238-251; b) U. Lessel, C. Lemmen, *ACS Med. Chem. Lett.* **2019**, *10*, 1504-1510.
- [13] L. Bellmann, P. Penner, M. Gastreich, M. Rarey, *J. Chem. Inf. Model.* **2022**, *62*, 553-566.
- [14] R. Schmidt, R. Klein, M. Rarey, *J. Chem. Inf. Model.* **2022**, *62*, 2133-2150.
- [15] a) A. A. Sadybekov, A. V. Sadybekov, Y. Liu, C. Iliopoulos-Tsoutsouvas, X. P. Huang, J. Pickett, B. Houser, N. Patel, N. K. Tran, F. Tong, N. Zvonok, M. K. Jain, O. Savych, D.

- S. Radchenko, S. P. Nikas, N. A. Petasis, Y. S. Moroz, B. L. Roth, A. Makriyannis, V. Katritch, *Nature* **2022**, *601*, 452-459; b) J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O'Meara, T. Che, E. Alga, K. Tolmachova, A. A. Tolmachev, B. K. Shoichet, B. L. Roth, J. J. Irwin, *Nature* **2019**, *566*, 224-229; c) R. M. Stein, H. J. Kang, J. D. McCorvy, G. C. Glatfelter, A. J. Jones, T. Che, S. Slocum, X. P. Huang, O. Savych, Y. S. Moroz, B. Stauch, L. C. Johansson, V. Cherezov, T. Kenakin, J. J. Irwin, B. K. Shoichet, B. L. Roth, M. L. Dubocovich, *Nature* **2020**, *579*, 609-614; d) C. Gorgulla, A. Boeszoermenyi, Z. F. Wang, P. D. Fischer, P. W. Coote, K. M. Padmanabha Das, Y. S. Malets, D. S. Radchenko, Y. S. Moroz, D. A. Scott, K. Fackeldey, M. Hoffmann, I. Iavniuk, G. Wagner, H. Arthanari, *Nature* **2020**, *580*, 663-668; e) A. Alon, J. Lyu, J. M. Braz, T. A. Tummino, V. Craik, M. J. O'Meara, C. M. Webb, D. S. Radchenko, Y. S. Moroz, X. P. Huang, Y. Liu, B. L. Roth, J. J. Irwin, A. I. Basbaum, B. K. Shoichet, A. C. Kruse, *Nature* **2021**, *600*, 759-764; f) A. A. Sadybekov, R. L. Brouillette, E. Marin, A. V. Sadybekov, A. Luginina, A. Gusach, A. Mishin, E. Besserer-Offroy, J. M. Longpre, V. Borshchevskiy, V. Cherezov, P. Sarret, V. Katritch, *Biomolecules* **2020**, *10*.
- [16] F. M. Klingler, M. Gastreich, O. O. Grygorenko, O. Savych, P. Borysko, A. Griniukova, K. E. Gubina, C. Lemmen, Y. S. Moroz, *Molecules* **2019**, *24*.
- [17] a) E. Kellenberger, J. Y. Springael, M. Parmentier, M. Hachet-Haas, J. L. Galzi, D. Rognan, *J. Med. Chem.* **2007**, *50*, 1294-1303; b) M. Cazorla, J. Premont, A. Mann, N. Girard, C. Kellendonk, D. Rognan, *J. Clin. Invest.* **2011**, *121*, 1846-1857; c) N. Weill, C. Valencia, S. Gioria, P. Villa, M. Hibert, D. Rognan, *Mol. Inform.* **2011**, *30*, 521-526; d) C. de Graaf, C. Rein, D. Piwnica, F. Giordanetto, D. Rognan, *ChemMedChem* **2011**, *6*, 2159-2169; e) .A. Ayrolles-Torro, T. Imberdis, J. Torrent, K. Toupet, I. V. Baskakov, G. Poncet-Montange, C. Gregoire, F. Roquet-Baneres, S. Lehmann, D. Rognan, M. Pugnieri, J. M. Verdier, V. Perrier, *J. Neurosci.* **2011**, *31*, 14882-14892; f) J. R. Pallandre, C. Borg, D. Rognan, T. Boibessot, V. Luzet, S. Yesylevskyy, C. Ramseyer, M. Pudlo, *Eur. J. Med. Chem.* **2015**, *103*, 163-174; g) S. A. Jacques, I. Kuhn, O. Koniev, F. Schuber, F. E. Lund, A. Wagner, H. Muller-Steffner, E. Kellenberger, *J. Med. Chem.* **2015**, *58*, 3582-3592; h) C. Rivat, C. Sar, I. Mechaly, J. P. Leyris, L. Diouloufet, C. Sonrier, Y. Philipson, O. Lucas, S. Mallie, A. Jouvenel, A. Tassou, H. Haton, S. Venteo, J. P. Pin, E. Trinquet, F. Charrier-Savournin, A. Mezghrani, W. Joly, J. Mion, M. Schmitt, A. Pattyn, F. Marmigere, P. Sokoloff, P. Carroll, D. Rognan, J. Valmier, *Nat. Commun.* **2018**, *9*, 1042.
- [18] G. W. Bemis, M. A. Murcko, *J. Med. Chem.* **1996**, *39*, 2887-2893.
- [19] <https://enamine.net/compound-collections/real-compounds/real-space-navigator> (accessed 2022-07-06)
- [20] <https://otavachemicals.com/products/chemriya> (accessed 2022-07-06)
- [21] <https://labnetwork.com/frontend-app/p/#!/library/virtual> (accessed 2022-07-06)
- [22] A. Shivanyuk, S. V. Ryabukhin, A. V. Bogolubsky, A. Tolmachev, *Chem. Today* **2007**, *25*, 58-59.
- [23] [https://www.biosolveit.de/infiniSee/#chemical\\_spaces](https://www.biosolveit.de/infiniSee/#chemical_spaces) (accessed 08-14-2022)
- [24] L. Howes, *C&EN* **2022**, *100*.

Received: ((will be filled in by the editorial staff))

Accepted: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))