

УДК 577.182.0
©Коллектив авторов

ОБЩИЕ ЧЕРТЫ АНТИБАКТЕРИАЛЬНЫХ СОЕДИНЕНИЙ: АНАЛИЗ ВЫБОРКИ 10⁴ ВЕЩЕСТВ

М.С. Веселов¹, П.В. Сергеев^{2,3}, И.А. Остерман², Д.А. Скворцов²,
А.Я. Головина³, Е.С. Андреянова³, И.Г. Лаптев², Ф.И. Плетнев²,
С.А. Евфратов², Е.И. Марусич¹, С.В. Леонов¹, Я.А. Иваненков¹,
А.А. Богданов^{2,3}, О.А. Донцова^{2,3}*

¹Московский физико-технический институт (государственный университет),
141700, Московская область, г. Долгопрудный, Институтский переулок, д. 9

²Московский государственный университет имени М.В. Ломоносова,
Химический факультет, 119991, Москва, Ленинские горы, 1, стр. 3

³Московский государственный университет имени М.В. Ломоносова,
Научно-исследовательский институт физико-химической биологии им. А.Н. Белозерского,
119992, Москва, Ленинские горы, 1, стр. 40; тел.: (495)9395418;
эл.почта: petya@genebee.msu.ru

Антибактериальные соединения – один из важнейших классов лекарственных препаратов. При высокопроизводительном поиске органических соединений могут быть выявлены потенциальные антибиотики, действующие на самые разные мишени в бактериальной клетке. Мы использовали выборку 9820 органических веществ, скомпонованную на основе максимизации химического разнообразия для высокопроизводительного скрининга антибактериальной активности. В результате роботизированного скрининга было обнаружено 103 соединения, проявивших антибактериальную активность в отношении модельной грамотрицательной бактерии, *Escherichia coli*. Для исходной выборки и обнаруженных активных молекул был рассчитан ряд молекулярных дескрипторов и выявлены основные различия в свойствах этих двух групп молекул. Для наглядного сравнения протестированных молекул и известных антибактериальных соединений была построена самоорганизующаяся карта Кохонена. Анализ результатов нелинейного картирования показал, что обнаруженные соединения, проявившие антимикробную активность, отличаются от хемотипов уже известных антибактериальных молекул и представляют интерес для дальнейшего исследования.

Ключевые слова: антибиотик, бактерия, скрининг, химические свойства

DOI: 10.18097/PBMC20156106785

ВВЕДЕНИЕ

Распространение устойчивых к антибиотикам возбудителей бактериальных заболеваний наносит значительный ущерб системе здравоохранения, приводит к увеличению трудопотерь и смертности среди населения [1-4]. Поиск новых антибиотиков ведётся как среди продуктов жизнедеятельности организмов, в первую очередь почвенных [5], так и среди химических соединений, синтезированных в лаборатории. Для первой категории, основную проблему представляет повторное обнаружение ранее идентифицированных антибиотиков [6]. Поскольку химическая формула активного соединения, производимого вновь обнаруженным штаммом

почвенного микроорганизма, заранее неизвестна, исследователи очень часто встречаются изоляты бактерий или грибов, продуцирующих антибиотики, обнаруженные ранее. При скрининге химически синтезированных соединений, формула каждого из них известна заранее, что облегчает задачу исследователей. Однако, органические соединения, созданные в лаборатории и не являющиеся аналогами известных ранее антибактериальных препаратов, крайне редко обнаруживают антибактериальную активность [7]. Для обнаружения принципиально новых классов антибактериальных соединений необходим непредвзятый поиск среди больших массивов химических соединений с очень низкой вероятностью обнаружения антибактериальных

* - адресат для переписки

веществ. Для того, чтобы облегчить такой поиск, необходимо компоновать соединения для скрининга используя набор описательных характеристик-дескрипторов, позволяющих повысить вероятность обнаружения антибактериальных веществ. Наиболее известным среди них является правило пяти, предложенное Кристофером Липински [8]. Позднее оказалось, что это правило не вполне подходит для поиска антибиотиков [5].

В нашей работе мы использовали библиотеку химических соединений, собранную по принципу максимизации химического разнообразия, для непредвзятого поиска веществ, обладающих антибактериальными свойствами и выявления их общих свойств.

МЕТОДИКА

Биологическое тестирование

Для тестирования была использована библиотека 9820 химических соединений "Разнообразие", предоставленная ЗАО "Исследовательский Институт Химического разнообразия". Основа компоновки библиотеки – максимизация различий между веществами. Библиотека была представлена в виде 103 96-луночных планшетов, каждая лунка содержала 1 мг индивидуального вещества. К веществам добавляли по 200 мкл ДМСО, растворяли вещества с помощью десятикратного пипетирования и отбирали по 1,5 мкл раствора на тестирование антимикробной активности. Раствор тестируемых веществ наносили на сплошной газон *Escherichia coli* штамма JW5503 (Δ tolC) на поверхности агаризованной среды LB. Пипетирование проводили с помощью автоматизированной станции Janus Extended ("Perkin Elmer", США) Чашки с газоном бактерий и образцами тестируемых веществ инкубировали сутки при 37°C, после чего анализировали визуально по наличию зон ингибирования, свободных от бактерий.

Соединения и базы данных

Выборка из 500 известных антибактериальных соединений, активных против *E. coli*, была составлена на основе базы данных Thomson Integrity Database [9]. В процессе набора были исключены соединения с молекулярным весом >900 Да, комплексы металлов, и удалены соли. Для построения карты Кохонена также был составлен набор из 500 соединений, набранных случайным образом из 9717 молекул, у которых не была обнаружена антибактериальная активность. Процедура кластеризации соединений проводилась в программе ChemoSoft [10] со следующими параметрами: минимальное количество соединений в кластере – 5, минимальный коэффициент подобия Танимото – 0,5.

Молекулярные дескрипторы

Расчёт молекулярных дескрипторов осуществлялся с помощью специализированных компьютерных

программ: MOE [11], Chemosoft и SmartMining [12]. Для каждого соединения было рассчитано более 200 уникальных дескрипторов, включая топологические, структурные, электронные, термодинамические и трёхмерные. Для снижения размерности пространства входных переменных был проведен анализ главных компонент в программном пакете Orange [13]. В качестве входных признаков для построения карты Кохонена были использованы первые 10 главных компонент, описывающих более 90% вариативности системы. Отбор дескрипторов, вносящих наибольший вклад в разделение различных групп соединений, осуществляли при помощи t-статистики Стьюдента.

Построение карты Кохонена

Построение карты Кохонена осуществляли с помощью программного комплекса SmartMining со следующими основными параметрами: размер карты 10×10, общее количество итераций – 1000, начальный радиус обучения – 5, начальная скорость обучения – 0,3, мера подобия между объектами в многомерном пространстве признаков – Евклидова, активационная функция нейронов – нормальная функция Гаусса.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для обнаружения принципиально новых антибактериальных соединений, не являющихся производными ранее изученных антибиотиков, необходимо проводить непредвзятый высокопроизводительный скрининг больших библиотек органических веществ. Ингибиторы каждой из молекулярных мишеней в бактериальной клетке, конечно, имеют свои особенности. Более того, ингибиторы разных мишеней могут сильно отличаться по свойствам друг от друга. Невозможно представить себе, что на основе общих свойств, таких, как молекулярный вес, заряд или гидрофобность можно было бы достоверно предсказать, что какое-либо новое соединение, непохожее на известные ингибиторы, непременно будет хорошим антибактериальным препаратом. Тем не менее, поскольку высокопроизводительный скрининг библиотек десятков тысяч соединений в любом случае предполагает очень низкую долю активных соединений среди них, важным усовершенствованием может служить даже небольшое повышение вероятности обнаружения активных соединений.

Для оптимизации подбора библиотек органических соединений с целью поиска среди них антибактериальных мы провели пилотный эксперимент по проверке антибактериальной активности среди 9820 органических соединений, не являющихся аналогами ранее известных антибактериальных препаратов и имеющих максимально различающееся химическое строение. Библиотека этих соединений была использована для роботизированного скрининга активности по подавлению роста *E. coli* на агаризованной среде.

В данном эксперименте мы не задавались целью искать мишень соединений, обнаруживших антибактериальную активность. Мы хотели лишь выявить сходные черты строения антибактериальных веществ для оптимизации подбора соединений в библиотеки для дальнейших поисков потенциальных антибиотиков.

Из 9820 соединений, прошедших тестирование, антибактериальная активность была выявлена у 103 соединений. Таким образом, вероятность обнаружения антибактериальной активности в использованной выборке примерно 1%. Группа для анализа общих свойств была составлена из 103 активных молекул. В этом анализе исходную выборку 9820 соединений использовали в качестве контроля. Поскольку количества соединений в выборках отличались на два порядка величины, для удобства визуализации результатов анализа мы нормировали число соединений, обладающих определенным свойством, на общее число соединений выборки. Таким образом, получали вероятность того, что соединение из выборки обладает именно таким свойством.

Оказалось, что распределение антимикробных соединений по молекулярному весу практически не отличалось от распределения исходной выборки (рис. 1А). Единственное отличие было в том, что распределение соединений, проявивших антимикробную активность, было более узким, чем распределение исходной выборки. Наибольшее число активных веществ имело молекулярную массу примерно 350 Да. Наибольшее количество антимикробных соединений имело 2-3 ароматических кольца (рис. 1Б), больше чем в среднем по исходной выборке. Соответственно, меньше была средняя "разветвленность", Загреб-индекс активных молекул (рис. 1В). Полярная поверхность антимикробных соединений была в среднем меньше (рис. 1Г), чем в исходной выборке. Меньше было число акцепторов водородной связи (рис. 1Д) при примерно равном числе доноров (рис. 1Е). В активных соединениях можно было чаще всего встретить 1-2 акцептора водородной связи, в то время как в исходной выборке чаще были представлены соединения с 2-3 акцепторами водородной связи (рис. 1Д). Также небольшим, в среднем 4, было количество атомов N и O (рис. 1Ж), в противоположность общей выборке, содержащей в среднем 6 N и O атомов на соединение. При сходном распределении молекулярных масс, уменьшение полярной поверхности, акцепторов водородных связей и полярных атомов N и O с неизбежностью привело к большей средней липофильности антимикробных соединений по сравнению с соединениями исходной выборки (рис. 1З). Анализ особенностей антибактериальных соединений по другим дескрипторам, таким, как площадь доступной поверхности, момент инерции, количество связей со свободным вращением и Ван-дер-Ваальсовый объем не выявили существенных различий (данные не показаны).

Исходная выборка 9820 соединений была кластеризована в программном пакете ChemoSoft с помощью метода, описанного в [14], на основе структурного подобия Танимото. В результате 6737 соединений распределились по 1088 кластерам, а остальные 3083 были идентифицированы в качестве соединений, не принадлежащих ни одному кластеру. При анализе состава кластеров выяснилось, что из 103 активных соединений 63 попали в 48 кластеров, а остальные 40 – в группу некластеризованных структур. При этом в большинстве случаев кластер содержал одно либо два активных соединения, остальные же (от 4 до 7 соединений в кластере) принадлежали к группе неактивных. Это позволяет предположить, что различия в заместителях соединений в рамках одного кластера (хемотипа) сильно влияют на антибактериальную активность.

Помимо выявленных различий в свойствах исходной выборки протестированных молекул и обнаруженных антимикробных соединений, также представляет интерес сравнение с соединениями, обладающими известной антибактериальной активностью. Стоит отметить, что исследование отдельных характеристик, представленных молекулярными дескрипторами, не позволяет в полной мере охарактеризовать расположение различных групп молекул относительно друг друга в общем химическом пространстве. Для решения подобных задач в настоящее время широко применяются методы нелинейного картирования [15], которые позволяют получать двумерные или трёхмерные отображения многомерных пространств свойств исследуемых выборок соединений. Как один из наиболее наглядных и удобных методов картирования, был выбран метод самоорганизующихся карт Кохонена [16], реализованный в программе SmartMining. Соединения, находящиеся в ячейках (нейронах), которые расположены близко на карте Кохонена, обладают близкими свойствами в исследуемом многомерном пространстве.

Общая выборка для построения карты Кохонена включала 103 обнаруженных в данной работе антимикробных соединений (группа 1), 500 соединений с известной антибактериальной активностью относительно *E. coli* (группа 2) и 500 случайно выбранных из 9717 молекул, для которых не была обнаружена антибактериальная активность (группа 3). Области преимущественной локализации трех выделенных групп молекул (рис. 2) были поочередно идентифицированы на общей карте Кохонена. В положениях соединений из этих групп наблюдаются заметные различия, и для каждой группы можно выделить области повышенной заселённости.

Как видно из рисунка 2, довольно сильное разделение наблюдается между известными антибактериальными соединениями (рис. 2А) и соединениями, протестированными в нашей работе (рис. 2Б,В), что говорит о значительном различии в свойствах этих групп молекул.

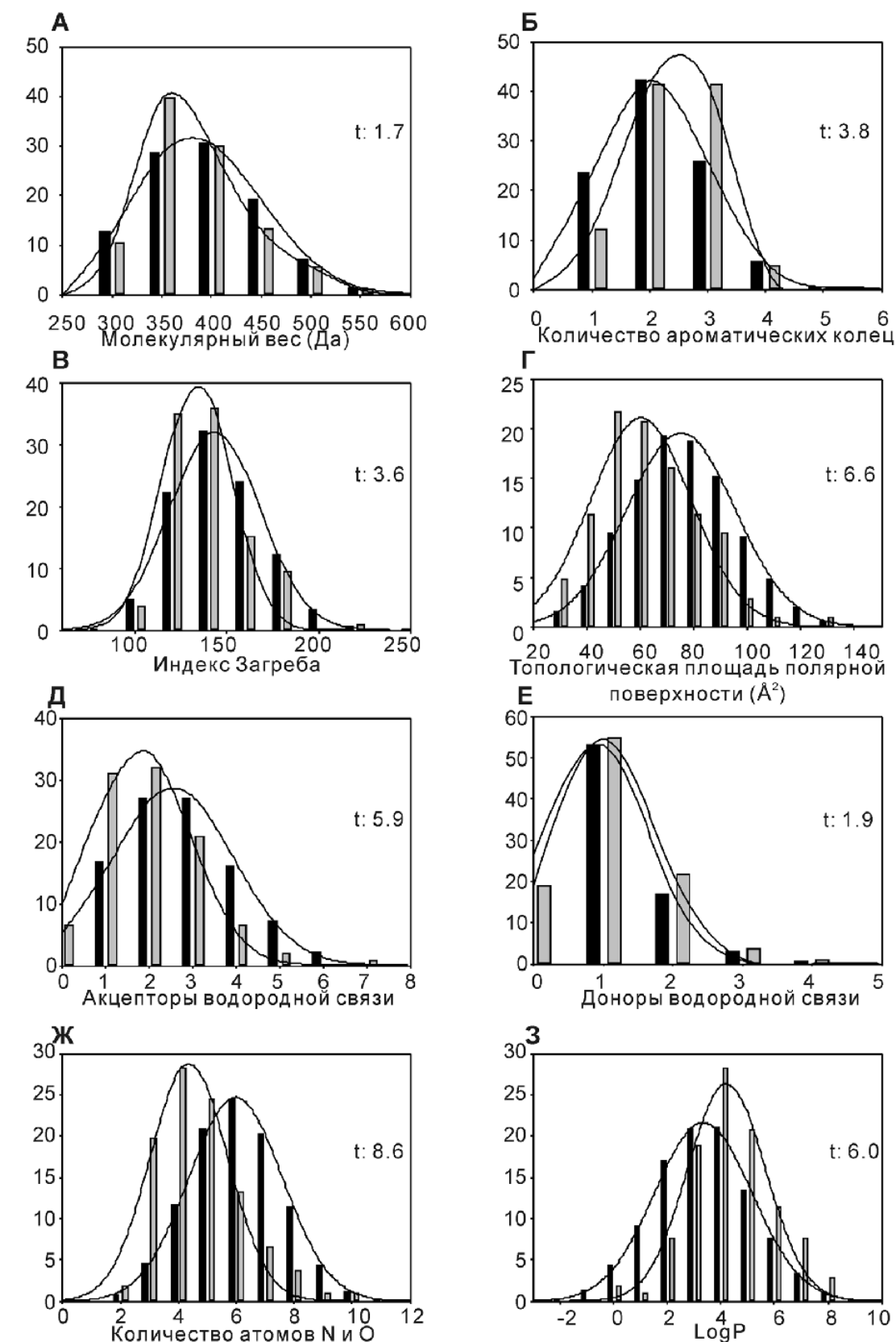


Рисунок 1. Свойства антимикробных соединений. Представлено распределение доли соединений по соответствующим свойствам-дескрипторам: молекулярной массе (А), числу ароматических колец (Б), Загреб-индексу разветвленности (В), площади полярной поверхности (Г), числу акцепторов (Д) и доноров (Е) водородной связи, числу полярных атомов N и O (Ж) и липофильности (З). Чёрными столбиками показаны доли соединений исходной выборки, серыми - соединений, проявивших антимикробную активность.

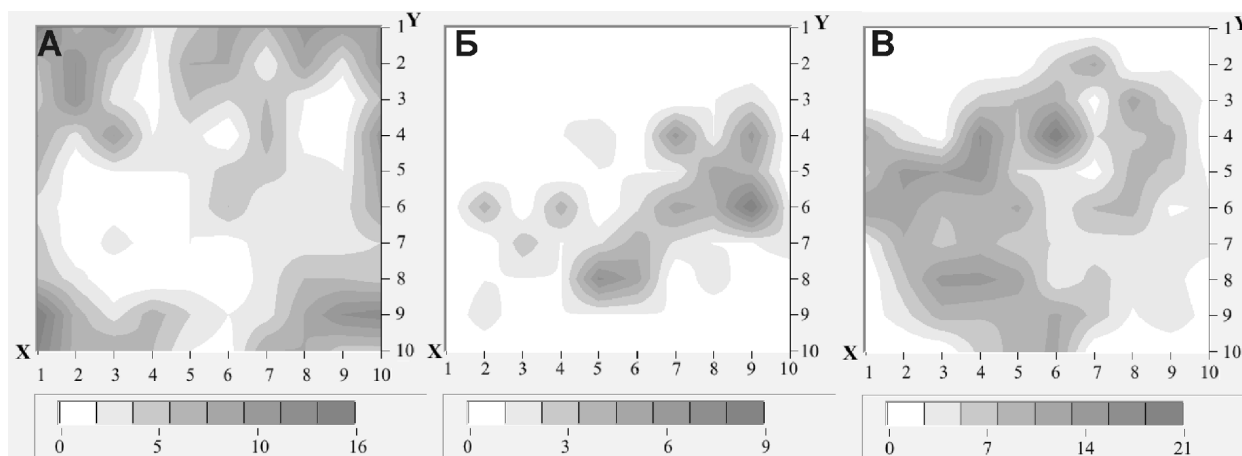


Рисунок 2. Области распределения репрезентативных групп соединений на карте Кохонена: известные соединения, обладающие антибактериальной активностью (А), обнаруженные антибактериальные хиты (Б), протестированные соединения, не обладающие антибактериальной активностью (В). По осям X и Y обозначены координаты узлов (нейронов) сети, градиентом серого цвета представлена плотность соединений в узлах.

Таблица. Разделение групп соединений на карте Кохонена.

Группа*	1	2	3	Не классифицировано	Всего
1	77/75%	24/23%	24/23%	0/0%	103/100%
2	103/21%	366/73%	23/5%	8/2%	500/100%
3	57/11%	70/14%	365/73%	8/2%	500/100%

Примечание. * 1 - обнаруженные в работе соединения-хиты, 2 - протестированные неактивные соединения, 3 - молекулы с известной антибактериальной активностью.

Таким образом, протестированные соединения расположены в отдельной области химического пространства относительно известных антибактериальных соединений.

Разделение между группой активных (рис. 2Б) и неактивных (рис. 2В) среди протестированных соединений наблюдается в меньшей степени, но тоже значительно. Обнаруженные антимикробные соединения занимают отдельную область и локализованы преимущественно в правой части карты, хотя отдельные соединения охватывают довольно значительную область на карте. Это говорит о высокой степени разнообразия в свойствах молекул в данной группе.

Каждая точка (нейрон) на карте классифицировалась как ячейка, принадлежащая определенной группе соединений, если в ней преобладают соответствующие молекулы (в процентном содержании). Средний процент разделения групп соединений из общей выборки на карте составил 73,6% (таблица). Существенное количество неактивных соединений, классифицированных как активные (21%) говорит о том, что незначительное изменение в свойствах или структуре молекулы может привести к потере антибактериальной активности.

В результате работы выяснилось, что вероятность обнаружения антимикробной активности среди органических соединений разнообразного строения, без предварительного отбора по сходству с известными антибиотиками, составляет около 1%. “Типичное” антимикробное соединение имеет несколько ароматических колец, образованных по большей части атомами углерода, достаточно гидрофобно и сравнительно невелико. При помощи построенной карты Кохонена для иллюстрации результатов скрининга было обнаружено, что протестированные соединения обладают свойствами, значительно отличающимися от известных соединений, обладающих активностью против *E. coli*, а соединения-хиты локализованы в отдельной области химического пространства. Таким образом, можно предположить, что обнаруженные хиты принадлежат к новым классам малых молекул, обладающих антибактериальной активностью, и представляют интерес для дальнейшего исследования.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации (Соглашение номер 14.607.21.0086, уникальный идентификатор проекта RFMEFI60714X0086) и ЗАО “ИИХР”.

ЛИТЕРАТУРА

1. Fowler T., Walker D., Davies S.C. (2014) Ann. N-Y Acad. Sci., **1323**, 1-10.
2. Andersson D.I., Hughes D. (2014) Nat. Rev. Microbiol., **12**, 465-478.
3. WHO. *Antimicrobial resistance: global report on surveillance* (2014) Available from: http://www.who.int/drugresistance/documents/AMR_report_Web_slide_set.pdf?ua=1.
4. Дебабов Д.В. (2012) Биотехнология, **4**, 7-17.
5. Lewis K. (2013) Nat. Rev. Drug Discov., **12**, 371-387.
6. Ito T., Masubuchi M. (2014) J Antibiot (Tokyo), **67**, 353-360.
7. Payne D.J., Gwynn M.N., Holmes D.J., Pompliano D.L. (2007) Nat. Rev. Drug Discov., **6**, 29-40.
8. Lipinski C.A. (2004) Drug Discov. Today Technol., **1**, 337-341.
9. Thomson Integrity Database. Available from: <https://integrity.thomson-pharma.com>
10. Chemosoft Software, V3. Official site: <http://chemdiv.com>
11. Molecular Operatin Environment (MOE). Official site: <https://www.chemcomp.com>
12. Balakin K.V., Ekins S. (2009) *Pharmaceutical Data Mining: Approaches and Applications for Drug Discovery*. Wiley, USA.
13. Orange Data mining. Official site: <http://orange.biolab.si>
14. Butina D. (1999) J. Chem. Inf. Comput. Sci., **39**(4), 747-750.
15. Ivanenkov Y.A., Bovina E.V., Balakin K.V. (2009) Russ. Chem. Rev., **78**(5), 465-483
16. Kohonen T. (1989) *Self-Organization and Associative Memory*, 3rd ed. Springer-Verlag, Berlin.

Поступила: 15. 06. 2015.

COMMON FEATURES OF ANTIBACTERIAL COMPOUNDS:
AN ANALYSIS OF 10⁴ COMPOUNDS LIBRARY

M.S. Veselov¹, P.V. Sergiev², I.A. Osterman², D.A. Skvortsov², A.Ya. Golovina², E.S. Andreyanova²,
I.G. Laptev², P.I. Pletnev², S.A. Evfratov², E.I. Marusich¹, S.V. Leonov¹, Y.A. Ivanenkov¹,
A.A. Bogdanov², O.A. Dontsova²

¹Moscow Institute of Physics and Technology (State University), Moscow region, Russia, 117303

²Department of Chemistry and A.N. Belozersky Institute of Physico-Chemical Biology,
Lomonosov Moscow State University,
Moscow, Russia, 119992; tel.: (495)9395418; e-mail: petya@genebee.msu.ru

Antibacterial compounds are one of the essential classes of clinically important drugs. High throughput screening allowed revealing potential antibiotics active towards any molecular target in bacterial cell. We used a library of 9820 organic compounds with highly diversified structures to screen for antibacterial activity. As the result of automated screening, 103 compounds were found to possess antibacterial activity against *Escherichia coli*. The properties of these compounds were compared with those of initial library. Non-linear Kohonen mapping was used to analyze the differences between non-active molecules from initial library, identified antibacterial hits and compounds with reported antibacterial activity. It was found that identified antibacterial compounds are located in the separated area of chemical space. It can be therefore suggested that these molecules belong to novel classes of antibacterial compounds and could be studied further.

Key words: antibiotic, bacteria, screening, molecular descriptors