Res Chem Intermed (2015) 41:7517–7540
DOI 10.1007/s11164-014-1841-0

Synthesis and chemoinformatics analysis of *N*-aryl-β-alanine derivatives

Kazimieras Anusevicius · Vytautas Mickevicius · Maryna Stasevych · Victor Zvarych · Olena Komarovska-Porokhnyavets · Volodymyr Novikov · Olga Tarasova · Tatyana Gloriozova · Vladimir Poroikov

Received: 7 April 2014 / Accepted: 1 October 2014 / Published online: 11 October 2014

Abstract Carbohydrazides of *N*-substituted β-amino acids exhibit a variety of different biological activities including antibacterial, antiviral, fungicidal, anti-helminthic, anticancer, antiinflammatory, etc. New potentially biologically active *N*-(4-iodophenyl)-β-alanine derivatives, *N*-(4-iodophenyl)-*N*-carboxyethyl-β-alanine derivatives, and their cyclization products were designed and synthesized. To determine the most propitious directions for further investigation of the obtained compounds, we tried to appraise their biological activity in silico using the ChemSpider and chemical structure lookup service (CSLS), chemical similarity assessment (Integrity and SuperPred), and machine learning methods [prediction of activity spectra for substances (PASS)]. No useful hints on potential biological activity of the obtained novel compounds were delivered by ChemSpider, CSLS, Integrity or SuperPred. In contrast, PASS predicted some biological activities that could be verified experimentally. Neither antibacterial nor antifungal activity was predicted for the compounds under study despite these actions being known for compounds from this chemical class. Evaluation of antibacterial (*Escherichia coli* B-906, *Staphylococcus aureus* 209-P, and *Mycobacterium luteum* B-91) and antifungal (*Candida tenuis* VKM Y-70 and *Aspergillus niger* F-1119) activities in vitro did not reveal any significant antimicrobial action, which corresponds to the

Electronic supplementary material The online version of this article (doi:10.1007/s11164-014-1841-0) contains supplementary material, which is available to authorized users.

K. Anusevicius · V. Mickevicius
Kaunas University of Technology, Radvilėnų Pl. 19, 50254 Kaunas, Lithuania

M. Stasevych (✉) · V. Zvarych · O. Komarovska-Porokhnyavets · V. Novikov
Lviv Polytechnic National University, Bandera Str. 12, 79013 Lviv, Ukraine
e-mail: maryna_s@inbox.ru

O. Tarasova · T. Gloriozova · V. Poroikov
Orekhovich Institute of Biomedical Chemistry, Pogodinskaya Str. 10/8, 119121 Moscow, Russia

computational prediction. Advantages and shortcomings of chemical similarity and machine learning techniques in computational assessment of biological activities are discussed. Based on the obtained results, we conclude that academic organic chemistry studies could provide a significant impact on drug discovery due to the novelty and diversity of the designed and synthesized compounds; however, practical utilization of this potential is narrowed by the limited facilities for assaying biological activities.

Keywords *N*-Aryl-β-alanines · Organic synthesis · Antibacterial and antifungal activity · Biological potential · Computational predictions · Drug discovery in academia

Introduction

β-Amino acids are promising intermediates in organic, bioorganic, medical, and peptide chemistry. Many compounds with structural fragments of β-amino acids—peptides, coenzymes, antibiotics, and alkaloids—are found in nature. A group of very important β-amino acid derivatives are the β-lactams, which are present in antibiotics, human leukocyte elastase inhibitors, and cholesterol uptake inhibitors. The anticancer compound taxol also has a structural component of β-alanine. An extensive description of the synthesis and application of β-alanines is presented in [1]. Likewise, *N*-substituted β-amino acids and their derivatives are important and have been widely studied for many different purposes, exhibiting biological activity, e.g., antibacterial [2] and antifungal [3, 4], as well as growth stimulation of agricultural crops [5, 6], antagonism of lysophosphatidic acid (EDG-2 receptor) [7], etc. *N*-Substituted β-amino acids are also excellent synthons for synthesis of azetidine [8–10], pyridinone [11–13], dihydrouracil [12, 14, 15], and quinolinone [16–18] heterocyclic systems. Modification of β-amino acids as potential substances for organic synthesis is presently a prevailing trend. One of the ways to modify *N*-substituted β-amino acids is to use carbohydrazides of *N*-substituted β-amino acids in synthesis of heterocyclic compounds [19–21]. On the other hand, many azole derivative compounds have shown a variety of different biological activities, including antibacterial and antifungal activity [22].

The aim of this work is to synthesis and evaluate in silico new potentially biologically active azoles from *N*-(4-iodophenyl)-β-alanine hydrazide, to establish their biological potential.

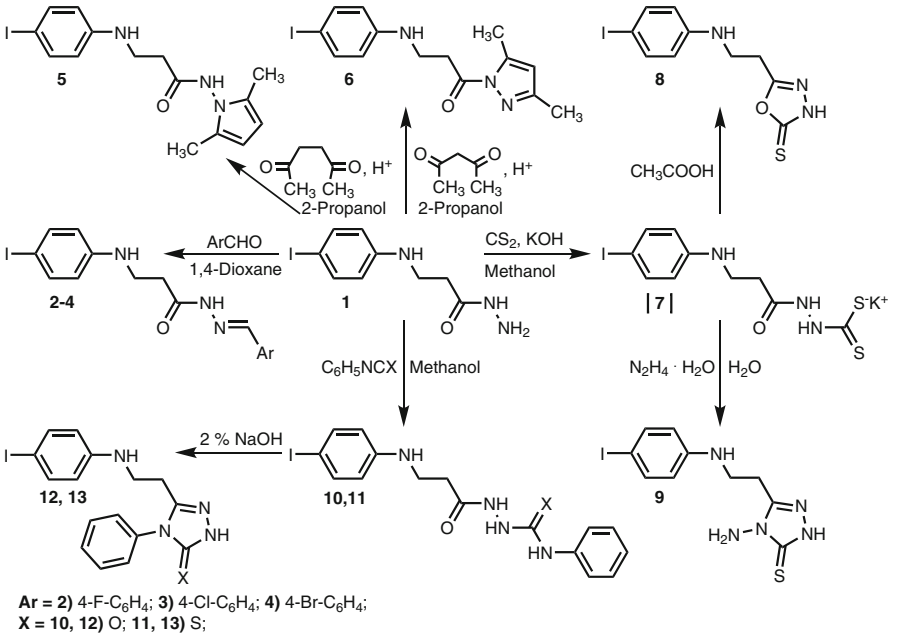
Nowadays, computer-aided methods are widely used in drug discovery, particularly for lead finding and optimization [23]. Target-based methods are mostly applied to find compounds inhibiting a particular endogenous macromolecule, whose three-dimensional structure is determined by X-ray analysis or nuclear magnetic resonance (NMR) [24]. Thus, such methods are focused on a particular target and are less applicable for evaluation of the biological potential of drug-like organic compounds. In contrast to this, ligand-based approaches provide information about the general biological activity profiles of drug-like compounds [25],

either based on chemical similarity assessment or on machine learning methods. Thus, for the purpose of our study, we selected ligand-based methods, in particular ChemSpider, chemical structure lookup service (CSLS), Integrity, SuperPred, and PASS (see below for description of "Computational studies").

Results and discussion

Chemistry

In continuation of our interest in the chemistry of *N*-substituted β-amino acids, *N*-(4-iodophenyl)-β-alanine hydrazide **1** was synthesized by reaction of *N*-(4-iodophenyl)-β-alanine and hydrazine hydrate in toluene under reflux. The hydrazones **2–4** were synthesized in good yields (89–97 %) by reacting *N*-(4-iodophenyl)-β-alanine hydrazide **1** with aromatic aldehydes in methanol at reflux (Scheme 1). The structure of the hydrazones **2–4** was established mainly on the basis of ^{1}H and ^{13}C NMR spectra. Theoretically, this compound, possessing amide and azomethine groups, can exist as an inseparable mixture of four isomers. The amide group determined the splitting of resonances in the ^{1}H and ^{13}C NMR spectra due to the restricted rotation around the amide bond. The ^{1}H and ^{13}C NMR spectral data allowed us to conclude that the geometrical isomers of the azomethine group were not observed and hydrazones **2–4c** in dimethyl sulfoxide (DMSO)-d_6 solution existed as a mixture of *E*/*Z* isomers, where the *Z* isomer predominated due to hindered rotation around the CO–NH bond.



Scheme 1 Synthesis of new *N*-(4-iodophenyl)-β-alanine derivatives **2–13**

Azole derivatives can be obtained by interaction of carbohydrazide and diketones. Dimethylpyrrole derivative **5** was prepared by refluxing a mixture of the corresponding carbohydrazide **1**, 2,5-hexanedione, 2-propanol, and a catalytic amount of glacial acetic acid (Scheme 1). 3,5-Dimethylpyrazole derivative **6** was prepared in good yield by refluxing a mixture of carbohydrazide **1**, 2,4-pentanedione, 2-propanol, and a catalytic amount of hydrochloric acid (Scheme 1). The resonances at 2.18 and 2.47 ppm (CH_3 groups), and 6.18 ppm (C=CH=C group) in the ^{1}H NMR spectra confirmed the formation of pyrazole rings. The resonances at 13.4 and 14.0 ppm of CH_3 groups in the ^{13}C NMR spectra of molecule **6** and the signals at 111.1 ppm, which were attributed to the CH, proved the presence of a five-membered heterocycle. The resonances observed at 143.1 and 148.1 ppm were ascribed to the carbons of the N–C and N=C groups of the heterocycle.

In this work, 1,3,4-oxadiazole **8** was prepared by refluxing carbohydrazide **1**, carbon disulfide, and potassium hydroxide in 2-propanol, followed by dissolution of the intermediate potassium dithiocarbazate **7** in water and treatment of the obtained solution with acetic acid to pH 6. The presence of broad singlets at 14.20 ppm of NH proton in ^{1}H NMR spectra and two resonances at 162.5 and 177.7 ppm, attributed to the N=C and C=S groups, respectively, in ^{13}C NMR spectra confirmed the formation of a five-membered oxadiazole ring in compound **8** (Scheme 1).

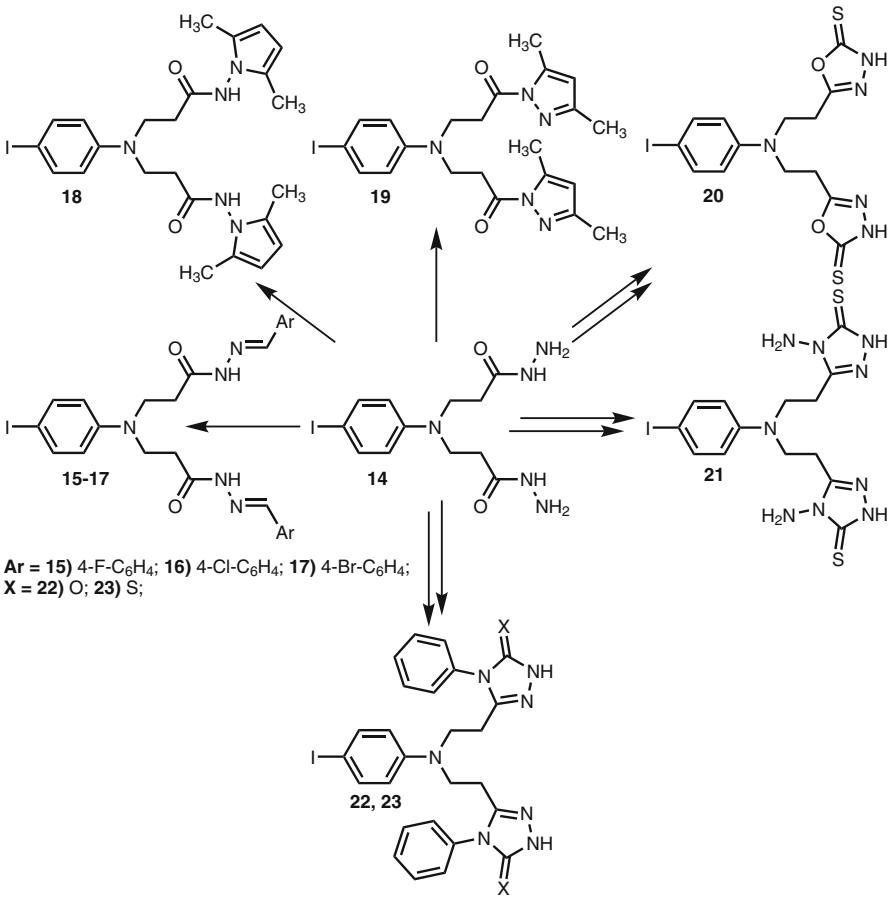
The 4-amino-1,2,4-triazole derivative **9** was prepared by heating intermediate dithiocarbazate **7** with hydrazine in water solution under reflux, and then the reaction mixture was acidified with acetic acid to pH 6. Compound **9** was also prepared by heating 1,3,4-oxadiazole **8** with hydrazine in 2-propanol solution under reflux, and then the reaction mixture was acidified with acetic acid to pH 6, but the reaction yield was low (24 %). The characteristic resonances at 13.51 ppm (NH) in the ^{1}H NMR spectrum and at 150.3 ppm ($CH_2\underline{C}$=) and 165.8 ppm (C=S) proved the formation of a 1,2,4-triazole ring in compound **9**. The presence of a NH_2 group was confirmed by the resonance multiplet at about 5.00–5.25 ppm in the ^{1}H NMR spectrum.

The starting compounds for the synthesis of phenyl-substituted triazoles (Scheme 1)—*N*-phenylhydrazinecarboxamide **10** and *N*-phenylhydrazinecarbothioamide **11**—were synthesized from hydrazides **1** with phenyliso- and phenylisothiocyanates in methanol. Triazoles **12**, **13** were obtained by refluxing compounds **10**, **11** in aqueous 2 % NaOH for 3 h with subsequent acidification of the reaction mixture with acetic acid. The sharp singlet at 11.72 ppm, assigned to the proton of the NH group, in the ^{1}H NMR spectrum, and resonances at 145.0, 154.3 ppm were assigned to C=N, C=O carbon atoms in the ^{13}C NMR spectrum, proving the formation of a five-membered heterocycle in compound **12**. The resonances at 150.3 ppm (C=N) and at 167.5 ppm (C=S) in the ^{13}C NMR spectrum proved the formation of a thioanalogous five-membered heterocycle in compound **13**.

Synthesis of compounds **15**–**23** (Scheme 2) was carried out in earlier work [26].

Computational studies

To identify the most promising directions for experimental research into the synthesized compounds, we decided to use an in silico approach. There exist several



Scheme 2 Synthesis of new bis-*N*-(4-iodophenyl)-β-alanine derivatives **15**–**23**

tools that may provide estimation of probable biological activity for drug-like compounds based on assessment of structural similarity or machine learning methods. All these tools require the structural formulae as input information, which should be presented as MOL files (for a single compound) or as an SDF file (for the set including all 20 compounds). We prepared such files using the ISIS Draw and ISIS Base programs [27].

ChemSpider

First, we tried to determine whether the synthesized compounds corresponded to known chemicals available via ChemSpider [28]. ChemSpider is a freely available chemical database providing structure search access to over 29 million structures from hundreds of data sources. No identical structure was found for any of the 20 synthesized compounds (see, e.g., the example search for compound **5** in Fig. 1).

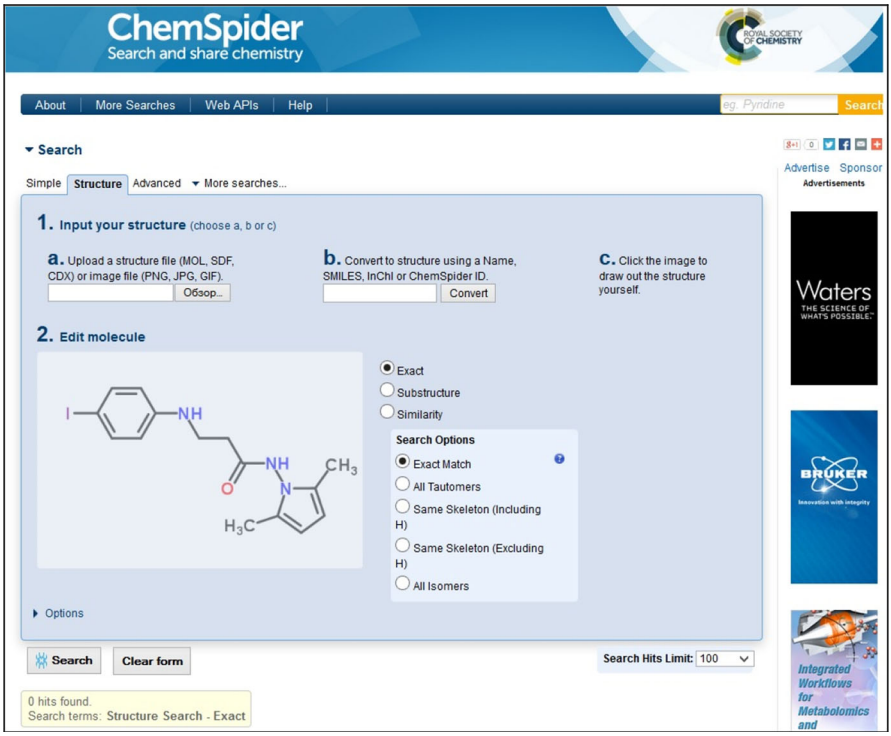


Fig. 1 Example search for compound **5** using ChemSpider

Chemical structure lookup service

Second, we tried to find the synthesized compounds among the known chemicals in different chemical databases using the CSLS [29]. The CSLS search engine provides the retrieval of 74 million indexed structures (46 million unique structures) from over 100 chemical databases (PubChem, NIST, EPA, etc.). As one may see from Fig. 2, no identical structure was found for any of the 20 synthesized compounds as well. Therefore, all synthesized compounds do not have any equivalent structure among the 46 million unique structures in the different databases available through CSLS.

Therefore, structure searches using both ChemSpider and CSLS did not find any identical structures among the dozens of millions of structures in over 100 different databases. This result provides additional evidence regarding the relative novelty of the structures synthesized in this work.

Integrity

Then, we tried to find similar structures within the Integrity database (Thomson Reuters) [30]. Integrity is a commercially available Internet portal providing information about drugs and pharmacological substances that are currently under different stages of research and development as potential human medicines. It

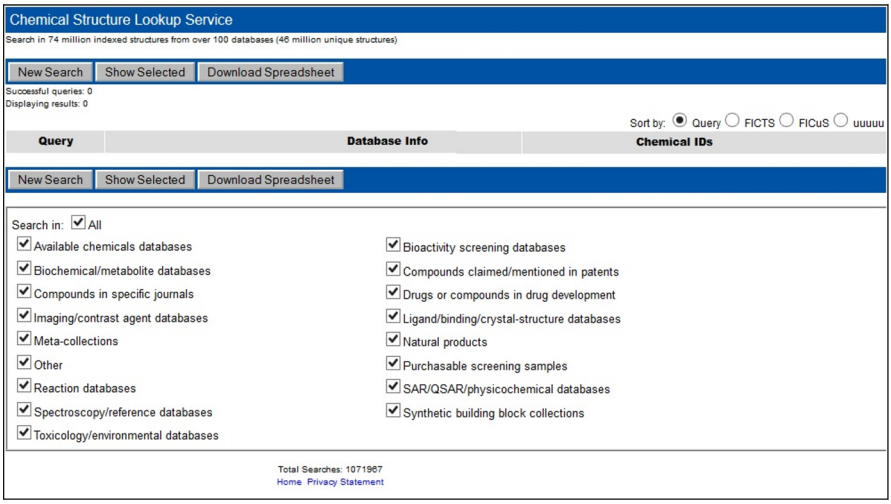


Fig. 2 Results of search for 20 structures using CSLS service

contains information about ca. 330,000 chemical structures. In addition to the exact structure search, which gave no results for any of the 20 synthesized structures, in Integrity it is possible to perform a search based on structure similarity. If some structures that have known biological activity and that are similar to the synthesized compounds could be identified, it is reasonable to suggest that our 20 compounds may have the same biological activities.

Similarity searching finds molecules that are similar to a query structure. The calculation applies the Tanimoto coefficient (TC), which has two arguments: (1) the fingerprint (a bit string that contains structural information on the molecule) of the query structure, and (2) the fingerprint of the molecule in the Integrity database. The TC is calculated by the following formula:

$$\mathrm{TC} = N_{\mathrm{AB}}/(N_{\mathrm{A}} + N_{\mathrm{B}} - N_{\mathrm{AB}}),$$

where N_{A} and N_{B} are the number of bits set for molecule A and B, respectively, and N_{AB} is the number of bits common to the two molecules.

The similarity threshold specifies a lower limit for the TC. If a TC is greater than this threshold, the query structure and the given structure from the database are considered to be similar. In our study, we used 70 % similarity as a lower threshold, which can still provide a reasonable basis for the inductive conclusion that such structures would have similar biological activities.

Figure 3 presents an example query for compound **12**.

The result of the similarity search for compound **12** is given in Fig. 4.

As one can see from Fig. 4, the only compound with greater than 70 % similarity to the structure of **12** is the compound with ID # 443457 from the Integrity database. This compound was studied by the Birla Institute of Technology and Science as an agent for treatment of neuropathic pain and as a potential antiepileptic drug.

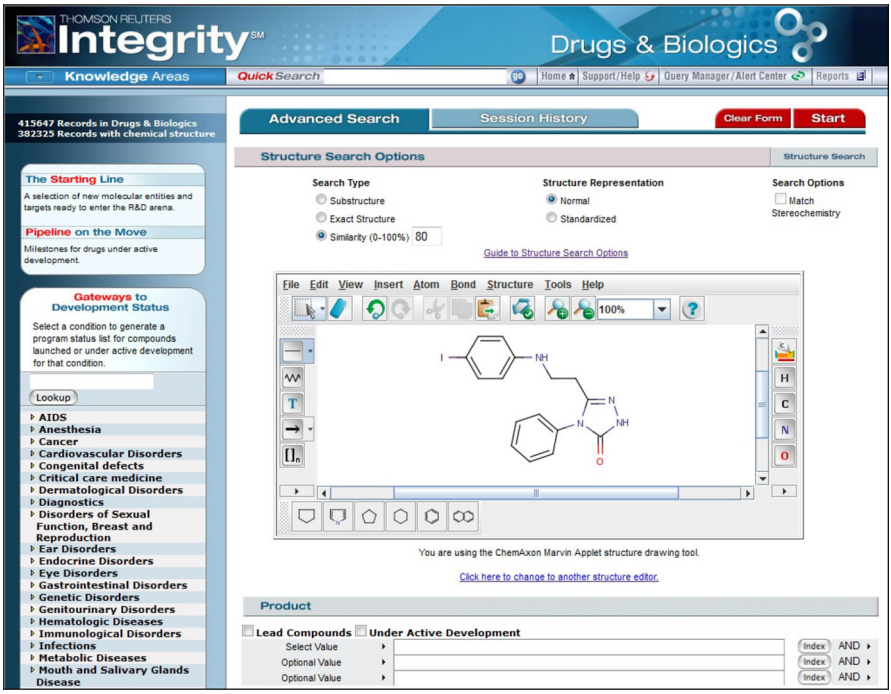


Fig. 3 Example similarity search using the structure of compound **12** as a query (threshold TC > 70 %)

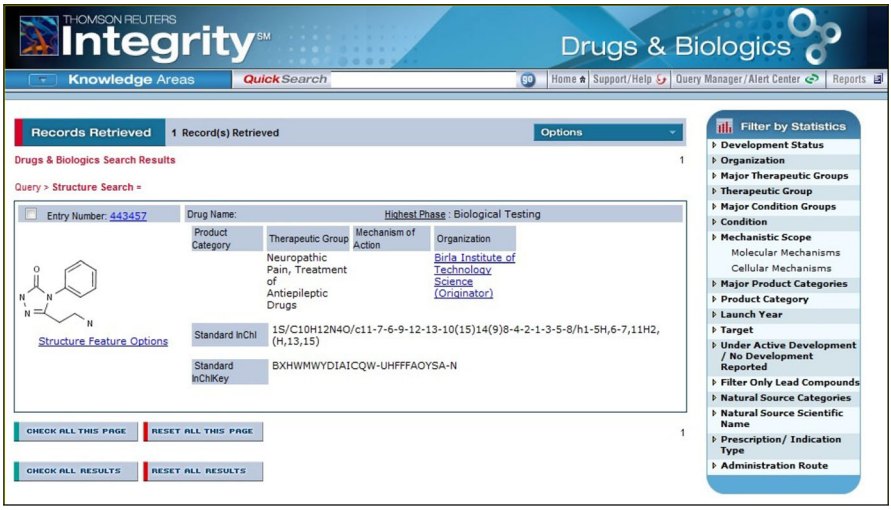
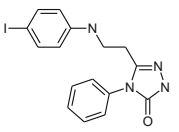
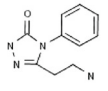
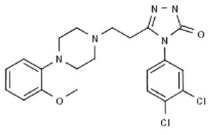
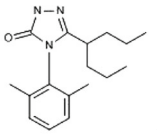
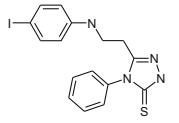
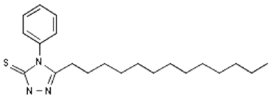
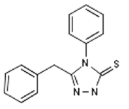
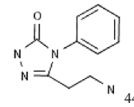
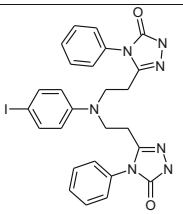
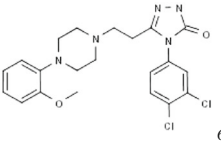
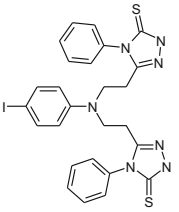
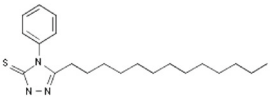
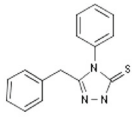


Fig. 4 Result of similarity search using the structure of compound **12** as a query (threshold TC > 70 %)

For 16 of the 20 synthesized compounds, no similar structures were found using the threshold of TS > 70 %. The results for the other four compounds are given in Table 1.

Table 1 Results of similarity search for compounds **12**, **13**, **22**, and **23** (threshold TC > 70 %)

Query structure	Similar structures (TS>70%)	Biological activities known for similar structures
 **12**	 443457	Neuropathic Pain, Treatment of Antiepileptic Drugs
	 663821	Anxiolytics Antidepressants 5-HT1A Receptor Agonists Signal Transduction Modulators
	 796958	Neuropathic Pain, Treatment of Signal Transduction Modulators TNF-alpha Production Inhibitors
 **13**	 218806	Lipoprotein Disorders, Treatment of Cholesteryl Ester Transfer Protein (CETP) Inhibitors
	 326397	Stroke, Treatment of Angina pectoris, Treatment of Neutral Sphingomyelinase (N-SMase) Inhibitors
	 443457	Neuropathic Pain, Treatment of Antiepileptic Drugs
 **22**	 663821	Anxiolytics Antidepressants 5-HT1A Receptor Agonists Signal Transduction Modulators
 **23**	 218806	Lipoprotein Disorders, Treatment of Cholesteryl Ester Transfer Protein (CETP) Inhibitors
	 326397	Stroke, Treatment of Angina pectoris, Treatment of Neutral Sphingomyelinase (N-SMase) Inhibitors

From the results presented in Table 1, it is clear that, for these 4 of the20 synthesized compounds, only five similar structures were found in the Integrity database when using the TC > 70 % threshold. One may suggest that biological activities known for these five structures, which are presented in Table 1, could be found in the query structures. However, it is necessary to bear in mind the result obtained by Yvonne Martin and coworkers [31] that there is only a 30 % chance that structures similar at the TS > 85 % threshold will exhibit similar biological activities.

SuperPred

SuperPred is a freely available web tool that provides estimation of biological activity based on chemical similarity [32, 33]. This tool calculates the TC values for the similarity between two compounds using their structural fingerprints as generated by the Chemistry Development Kit [34]. Biological activities are described here in terms of the Anatomical Therapeutic Chemical (ATC) classification and using a list of human proteins that are considered as pharmacological targets. Estimation of chemical similarity for each of the synthesized 20 compounds using SuperPred did not give TC > 70 % for any pairs of compounds (see data presented in Electronic Supplementary Material 1). The highest TC value, calculated for the similarity between compound **14** and 3-[acetyl(3-amino-2,4,6-triiodophenyl)amino]-2-methylpropanoic acid, was about 65 %; however, no biological activity was predicted in this case. Bearing in mind the above-mentioned poor accuracy of biological activity prediction based on chemical similarity [31], one may conclude that SuperPred does not provide any useful hints about the biological activity of the 20 synthesized compounds.

PASS

Since our attempts to identify the same or similar structures with known biological activities among dozens of millions of chemical substances from several Internet resources did not help in determination of the most promising directions for testing the biological activity of our 20 synthesized compounds, we decided to predict their biological activity using the PASS computer program [35, 36]. The prediction of activity spectra for substances (PASS) estimates the probabilities of belonging to the classes of "actives" (Pa) and "inactives" (Pi) for over 6,000 biological activities on the basis of analysis of structure–activity relationships for a training set including more than 300,000 biologically active compounds. The machine learning method implemented in PASS is based on a Bayesian algorithm [36] and multilevel neighborhoods of atoms (MNA) descriptors [37]. The average accuracy of PASS prediction estimated by a leave-one-out cross-validation procedure for the whole training set and all predictable biological activities is about 95 %; similar values were obtained in 20-fold cross-validation. The freely available web resource PASS Online [35] has been available since 2000 [38–40], and more than 50 papers with independent confirmation of the prediction results for various biological activities and different chemical classes have been published [35, 40].

Table 2 Statistics of number of compounds with activities predicted with Pa > 50 %

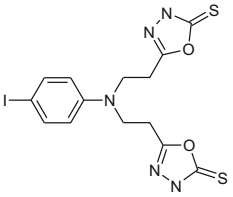
No.	Number of compounds	Biological activity
1	20	Complement factor 1s inhibitor
2	8	DNA ligase 1 inhibitor
3	8	Protein disulfide-isomerase inhibitor
4	6	Epstein–Barr nuclear antigen 1 inhibitor
5	6	NACHT, LRR, and PYD domains-containing protein 3 inhibitor
6	5	PfA-M1 aminopeptidase inhibitor
7	5	Eukaryotic translation initiation factor 4H inhibitor
8	4	Potassium sparing diuretic
9	4	Peptide agonist
10	3	Taurine dehydrogenase inhibitor
11	3	Ubiquinol-cytochrome-c reductase inhibitor
12	2	Nonstructural protein 1 (influenza A) inhibitor
13	2	Transferase inhibitor
14	2	Analgesic
15	2	Ethanolamine oxidase inhibitor
16	2	Glutamine-phenylpyruvate transaminase inhibitor
17	2	Maillard reaction inhibitor
18	2	Penicillin binding protein antagonist
19	2	Penicillin-binding protein ampH inhibitor
20	2	G-protein-coupled receptor kinase inhibitor
21	2	Beta-adrenergic receptor kinase inhibitor
22	2	Polyadenylate-binding protein 1 inhibitor
23	2	Amine dehydrogenase inhibitor
24	2	Beta-Lysine 5,6-aminomutase inhibitor
25	2	Centromere associated protein inhibitor
26	2	Transcription factor STAT1 inhibitor
27	2	Botulin neurotoxin A light chain inhibitor
28	1	Manganese peroxidase inhibitor
29	1	Trimethylamine dehydrogenase inhibitor
30	1	HIV rev inhibitor
31	1	Phosphatidylserine decarboxylase inhibitor
32	1	Fusarinine-C ornithinesterase inhibitor
33	1	Transferase stimulant
34	1	Plectin inhibitor
35	1	NADPH-cytochrome-c2 reductase inhibitor
36	1	Phosphopantothenoylcysteine decarboxylase inhibitor
37	1	Aspartate-phenylpyruvate transaminase inhibitor
38	0	Spermidine dehydrogenase inhibitor
39	0	Neuropeptide Y2 antagonist
40	0	Glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase inhibitor

We performed PASS Online predictions of the pharmacotherapeutic effects and mechanisms of action for the 20 synthesized compounds, and then analyzed the obtained results using the PharmaExpert computer program [41]. A report on the prediction results for all the compounds as generated by PharmaExpert is presented in Electronic Supplementary Material 2. Statistics on the number of compounds for which a particular activity is predicted with Pa > 50 % for the top 40 activities are given in Table 2.

As one can see from the data presented in Table 2, complement factor 1s inhibiting activity is predicted for all 20 compounds; DNA ligase 1 inhibiting and protein disulfide-isomerase inhibiting activities are predicted for 8 of the 20 compounds; Epstein–Barr nuclear antigen 1 inhibiting and NACHT, LRR, and PYD domains-containing protein 3 inhibiting activities are predicted for 6 of the 20 compounds; PfA-M1 aminopeptidase inhibiting and eukaryotic translation initiation factor 4H inhibiting activities are predicted for 5 of the 20 compounds; etc. These activities can be considered as typical for the chemical series under consideration because they are predicted for an essential fraction of the whole set of compounds [42]. In contrast, transcription factor STAT1 inhibiting, botulin neurotoxin A light chain inhibiting, manganese peroxidase inhibiting, trimethylamine dehydrogenase inhibiting, HIV rev inhibiting, etc. activities are predicted just for one or two derivatives, and they can be considered as minor for this chemical series [39]. Spermidine dehydrogenase inhibiting, neuropeptide Y2 antagonistic, and glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase inhibiting activities are predicted with probabilities Pa < 50 %, and the probability to find the activity in experiments is rather small.

It is necessary to stress that, at the threshold Pa > 50 %, neither antibacterial nor antifungal activity is predicted by PASS, which corresponds to the negative results of experimental testing of our 20 compounds on *Escherichia coli* B-906, *Mycobacterium luteum* B-917, *Candida tenuis* VKM Y-70, and *Aspergillus niger* F-1119 (see below).

These PASS predictions provide hints about priorities on biological activities which should be tested for each synthesized derivative, to increase the chances of finding biological activity in experiments. Let us consider, for example, the prediction results obtained with Pa > 50 % for compound **20**:



> <ACTIVITY_PREDICTION>
33 Substructure descriptors; 0 new.
5 Possible activities.
0.680 0.009 Complement factor 1s inhibitor
0.573 0.005 Polyadenylate-binding protein 1 inhibitor
0.569 0.006 Transcription factor STAT1 inhibitor
0.540 0.029 Analgesic
0.556 0.085 Protein disulfide-isomerase inhibitor

Activity that is predicted with the highest probability, so-called focal activity, should be tested first. For compound **20** the focal activity is as a complement factor 1s inhibitor. Then, one should test the activities that have the second, third, etc. ranks in the predicted activity spectra: polyadenylate-binding protein 1 inhibitor, transcription factor STAT1 inhibitor, analgesic, and protein disulfide-isomerase inhibitor.

There are only four different focal activities for all 20 compounds: complement factor 1s inhibitor (compounds **5**–**13**, **18**–**23**), DNA ligase 1 inhibitor (compounds **2**–**4**, **15**–**17**), beta-lysine 5,6-aminomutase inhibitor (compound **1**), and taurine dehydrogenase inhibitor (compound **14**). It is clear that the majority of the focal activities (e.g., complement factor 1s inhibitor) correspond to typical activities in this chemical series; however, some of the focal activities (e.g., beta-lysine 5,6-aminomutase inhibitor) correspond to minor activities in this chemical series.

Analysis of the statistics of activities predicted at Pa $>$ 50 % shows that even 20 *N*-(4-iodophenyl)-β-alanine derivatives, *N*-(4-iodophenyl)-*N*-carboxyethyl-β-alanine derivatives, and their cyclization products demonstrate wide pharmacological potential. Let us consider this potential in more detail.

A complement, as an essential component of the immune system, is of substantial relevance for destruction of invading microorganisms and for maintaining tissue homeostasis, including protection against autoimmune diseases [43]. Complement activation is tightly controlled by its regulators to prevent complement-mediated self-tissue injury [44]. It has been well established that excessive complement activation contributes to the pathogenesis of many autoimmune diseases, including rheumatoid arthritis [45, 46], type I diabetes [47], myasthenia gravis [48], glomerulonephritis [49], as well as transplanted graft rejection [50]. Thus, to focus on complement inhibition appears to be a logical approach to stop the process of inflammatory disorders. Compounds with predictive complement factor 1s inhibiting activity include all 20 synthesized compounds.

A hallmark of many neurodegenerative diseases, including Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, and Huntington's disease, is accumulation of misfolded proteins within neurons, leading to cellular dysfunction and cell death. Comprehensive understanding of how the protein disulfide-isomerase cell death pathway is controlled may aid in discovering therapeutics to treat protein misfolding diseases [51].

DNA reparation (DNA ligase I) inhibitors have potential to concomitantly increase the killing of cancer cells and reduce damage to normal tissues and cells if either the damaging agent or the inhibitor can be selectively delivered to cancer cells. A problem with the screening of random chemical libraries for DNA ligase inhibitors is that many of the hits are likely to be nonspecific inhibitors that either bind to the DNA substrate or are nucleotide analogs that inhibit a large number of adenosine triphosphate (ATP)-dependent enzymes [52].

PfA-M1 aminopeptidase localizes differently in trophozoites and schizonts, suggesting two different roles in critical steps of the intraerythrocytic life of the parasite [53]. Proteases that are expressed during the erythocytic stage of *Plasmodium falciparum* are newly explored drug targets for treatment of malaria

[54]. Some of our compounds are predicted to be potential inhibitors of PfA-M1 aminopeptidase.

Substances for which peptide agonist activity was predicted are probably resistant to proteolytic attack, and may have prolonged activity as peptide-based drugs used to treat corneal ulceration, emphysema, rheumatoid arthritis, peptic ulcers, hypertension, and multiple sclerosis.

Ubiquinol cytochrome C oxidoreductase (bc_1 complex) is an essential component of the electron transport systems critical for respiration in prokaryotes and eukaryotes, and photosynthesis in some bacteria [55, 56]. In fact, in many species, ubiquinol-cytochrome C oxidoreductase activity was found to be sensitive to quinone analogs such as myxothiazol, stigmatellin, antimycin, hydroxyquinoline *N*-oxide (HQNO), and diuron [57]. A variety of inhibitors of the bc_1 complex have been important in elucidating the Q cycle hypothesis and, together with recent structural and genetic experiments, have also been useful in mapping the features of the Qo and Qi domains and their relationship to the reaction mechanism [58]. The inhibitors fall into two main categories: those that bind at the Qi site (class II), such as the antibiotic antimycin, and those that bind to the Qo site, such as strobilurin A (class I) [58]. The obtained information [59] may be useful for understanding the function of the bc_1 complex, and for design of novel antiinfective agents for use as human remedies and in agricultural applications.

Comparison of the obtained Pa data for our 20 compounds shows that the Pa values for monosubstituted derivatives exceed those for disubstituted derivatives; also, wider predicted biological activity spectra are observed in monosubstituted derivatives.

Neither antibacterial nor antifungal activities were predicted by PASS for compounds under study with $Pa > 0.5$. However, since such activities are known for compounds from this chemical class, we decided to study their antibacterial and antifungal activities in vitro. The experimental results should shed light on the antimicrobial action of our compounds, and also act as a negative control for computer-aided predictions.

Biology

Compounds **1**–**6**, **8**, **9**, and **12**–**23** were selected to identify the influence of azole fragments in the case of mono- and disubstituted derivatives of *N*-(4-iodophenyl)-β-alanine (substances **10** and **11** were not chosen for antimicrobial study because they were used only as intermediate compounds in synthesis of triazoles **12**, **13**). These substances were evaluated for their antibacterial and antifungal activity against strains of *E. coli* B-906, *Staphylococcus aureus* 209-P, *M. luteum* B-917 (as a nonpathogenic test bacterial culture representative of the genus *Mycobacterium*), *C. tenuis* VKM Y-70, and *A. niger* F-1119 by the diffusion technique [60] and serial dilution technique [61] (determination of minimal inhibition concentration, MIC) to correlate the in silico investigation with experimental data for the new potentially biologically active compounds based on *N*-(4-iodophenyl)-β-alanine hydrazides. It was established that all tested compounds showed no antibacterial or fungicidal action at concentrations of 0.5 and 0.1 % (diffusion technique) against strains

E. coli, *S. aureus*, *M. luteum*, *C. tenuis*, and *A. niger*. No influence of the studied compounds on the growth of microorganisms was observed at concentrations of 3.9–500 mkg/ml (serial dilution technique). However, in the same experiment, antimicrobial activity was found for the reference antibacterial and antifungal drugs (vancomycin and nystatin at concentration of 0.5 and 0.1 %, respectively).

Conclusions

In silico studies of synthesized new potentially biologically active *N*-(4-iodophenyl)-β-alanine derivatives, *N*-(4-iodophenyl)-*N*-carboxyethyl-β-alanine derivatives, and their cyclization products allowed their biological potential to be established. Application of several freely available and commercial tools (ChemSpider and CSLS) showed the relative novelty of the synthesized derivatives; however, similarity searches (Integrity and SuperPred) did not provide many fruitful ideas regarding which kinds of biological activity of the synthesized compounds should be tested with high priority. This finding corresponds to earlier conclusions that there is a rather small chance of revealing the same activity in pairs of compounds with TC exceeding 85 % [30].

In contrast to the similarity assessment, machine learning methods implemented in the PASS program predicted probability to be active Pa > 50 % for 37 different biological activities, including typical activities for this chemical series (e.g., complement factor 1s inhibitor) and minor activities for this chemical series (e.g., taurine dehydrogenase inhibitor). To achieve the maximum efficiency in discovery of new biologically active compounds, it is necessary to study the biological activity of each compound starting from those predicted with the highest probability (so-called focal activity), continuing in descending order of probability to be active (Pa).

However, the predicted activities fall into very different pharmacological categories and may have various mechanisms of action. Therefore, it is rather difficult for an academic team to synthesize new chemical compounds with great pharmacological potential according to computational predictions to check their diverse biological activities. Based on the current study, we can only conclude that antibacterial and antifungal activities were not predicted by PASS with Pa > 50 %, and that these activities were not found in in vitro assays.

The essential role of academic studies in drug design and discovery is well established [62–65]. Being more flexible in their research topics, people working in academia often generate new ideas that may be translated into practice through further research and development. However, the past experience of many organic chemists shows that the main bottleneck in this process is the limited facilities for assaying biological activities. Creating new academic drug R & D centers, as observed during recent years [62], may significantly improve this situation. Collaboration through multidisciplinary teams and wide application of computational methods represent the current mainstream of new drug discovery projects [66].

Experimental

Chemistry

^{1}H and ^{13}C NMR spectra were recorded on a Varian Unity Inova (300, 75 MHz) spectrometer operating in Fourier-transform mode, using DMSO-d_6 as solvent and tetramethylsilane (TMS) as internal reference (chemical shifts in δ, ppm). Infrared (IR) spectra (v, cm^{-1}) were recorded on a PerkinElmer Spectrum BX FT–IR spectrometer using KBr tablets. Elemental analyses were performed with a CE-440 elemental analyzer. Melting points were determined with an automatic APA1 melting point apparatus and are uncorrected. Thin-layer chromatography (TLC) was performed with Merck silica gel 60 F_{254} (Kieselgel 60 F_{254}) silica gel plates.

General procedure for synthesis of hydrazones **2–4**

A mixture of hydrazide **1** (0.92 g, 3 mmol) and 4 mmol of the corresponding aromatic aldehyde in 30 ml 1,4-dioxane was heated under reflux for 3 h. The reaction mixture was cooled down, the precipitate filtered, then washed with 2-propanol and crystallized from methanol to give hydrazones **2–4**.

*N′-[(Z/E)-(4-Fluorophenyl)methylidene]-3-[(4-iodophenyl)amino]propanehydrazide (**2**)*

White solid (1.16 g, 94 %), m.p. 169–170 °C (methanol). IR, v: 1,231 (ArNH), 1,599 (N=C), 1,660 (CO), 3,173 (CONH), 3,354 (CH_2NH). ^{1}H NMR (300 MHz, DMSO-d_6): δ = 2.43–2.50 (0.4 (2H), m, CH_2CO), 2.88 (0.6 (2H), t, J = 7.0 Hz, CH_2CO), 3.26–3.36 (2H, m, CH_2NH), 5.92 (1H, t, J = 5.8 Hz, CH_2NH), 6.42–7.78 (8H, m, H_{ar}), 7.98 (0.6 (1H), s, N=CH), 8.15 (0.4 (1H), s, N=CH), 11.36, 11.41 (1H, 2s, CONH). ^{13}C NMR (75.4 MHz, DMSO-d_6): δ = 31.7, 33.8 (CH_2CO), 38.2, 38.8 (CH_2NH), 75.8, 75.9, 114.5, 114.6, 115.6, 115.9, 128.6, 128.7, 129.0, 129.1, 130.7, 130.7, 130.9, 137.0, 141.6, 144.8, 148.1, 148.2, 161.1, 164.4 (C-ar), 167.0, 172.8 (CONH). Found, %: C 46.53, H 3.54, N 10.38. Calculated, %: C 46.73, H 3.68, N 10.22. $C_{16}H_{15}FIN_3O$.

*N′-[(Z/E)-(4-Chlorophenyl)methylidene]-3-[(4-iodophenyl)amino]propanehydrazide (**3**)*

White solid (1.24 g, 97 %), m.p. 182–183 °C (methanol). IR, v: 1,264 (ArNH), 1,596 (N=C), 1,664 (CO), 3,168 (CONH), 3,348 (CH_2NH). ^{1}H NMR (300 MHz, DMSO-d_6): δ = 2.43–2.50 (0.4 (2H), m, CH_2CO), 2.89 (0.6(2H), t, J = 7.0 Hz, CH_2CO), 3.31 (2H, dd, ^{1}J = 6.7 Hz, ^{2}J = 12.7 Hz, CH_2NH), 5.92 (1H, t, J = 5.4 Hz, CH_2NH), 6.41–7.73 (8H, m, H_{ar}), 7.98 (0.6 (1H), s, N=CH), 8.14 (0.4 (1H), s, N=CH), 11.41, 11.48 (1H, 2s, CONH). ^{13}C NMR (75.4 MHz, DMSO-d_6): δ = 31.6, 33.8 (CH_2CO), 38.2, 38.7 (CH_2NH), 75.8, 114.6, 114.6, 128.2, 128.5, 128.8, 133.1, 133.2, 134.0, 134.3, 137.1, 141.4, 144.6, 148.1, 148.2 (C-ar), 167.1,

172.9 (CONH). Found, %: C 44.81, H 3.83, N 9.98. Calculated, %: C 44.93, H 3.54, N 9.83. $C_{16}H_{15}ClIN_3O$.

*N′-[(Z/E)-(4-Bromophenyl)methylidene]-3-[(4-iodophenyl)amino]propanehydrazide (**4**)*

White solid (1.36 g, 96 %), m.p. 181–182 °C (methanol). IR, *v*: 1,288 (ArNH), 1,598 (N=C), 1,659 (CO), 3,167 (CONH), 3,334 (CH_2NH). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.42$–2.50 (0.4 (2H), m, CH_2CO), 2.88 (0.6 (2H), t, $J = 7.0$ Hz, CH_2CO), 3.31 (2H, q, $J = 6.8$ Hz, CH_2NH), 5.92 (1H, t, $J = 5.8$ Hz, CH_2NH), 6.41–7.65 (8H, m, H_{ar}), 7.95 (0.6 (1H), s, N=CH), 8.12 (0.4 (1H), s, N=CH), 11.41, 11.47 (1H, 2s, CONH). ^{13}C NMR (75.4 MHz, DMSO-d_6): 31.6, 33.8 (CH_2CO), 38.2, 38.7 (CH_2NH), 75.8, 114.6, 114.6, 122.8, 128.4, 128.7, 131.7, 133.4, 133.5, 137.1, 141.5, 144.6, 148.1, 148.2 (C-ar), 167.1, 172.1 (CONH). Found, %: C 40.64, H 3.42, N 9.05. Calculated, %: C 40.70, H 3.20, N 8.90. $C_{16}H_{15}BrIN_3O$.

*N-(2,5-Dimethyl-1H-pyrrol-1-yl)-3-[(4-iodophenyl)amino]propanamide (**5**)*

A mixture of hydrazide **1** (0.92 g, 3 mmol), 2,5-hexanedione (0.57 g, 5 mmol), 2-propanol (10 ml), and acetic acid (0.5 ml) was heated under reflux for 6 h, then cooled, the precipitate was filtered off, washed with 2-propanol and crystallized from 2-propanol to give dimethylpyrrole derivative **5**.

White solid (0.80 g, 70 %), m.p. 136–137 °C (2-propanol). IR, *v*: 1,262 (ArNH), 1,660 (CO), 3,292 (CONH), 3,345 (CH_2NH). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 1.95$, 2.00 (6H, 2s, CH_3), 2.52 (2H, t, $J = 6.6$ Hz, CH_2CO), 3.30–3.39 (2H, m, CH_2NH), 5.61 (0.9 (4H), s, CH–CH), 5.70 (0.1 (4H), s, CH–CH), 5.95 (1H, t, $J = 5.9$ Hz, CH_2NH), 6.34 (0.2 (2H), d, $J = 8.9$ Hz, $H_{ar\text{-}2.6}$), 6.47 (0.8 (2H), d, $J = 8.9$ Hz, $H_{ar\text{-}2.6}$), 7.29 (0.2 (2H), d, $J = 8.8$ Hz, $H_{ar\text{-}3.5}$), 7.35 (0.8 (2H), d, $J = 8.8$ Hz, $H_{ar\text{-}3.5}$), 10.17 (0.1 (2H), 2s, NH), 10.60 (0.9 (2H), 2s, NH). ^{13}C NMR (75.4 MHz, DMSO-d_6): $\delta = 10.9$ ($2CH_3$), 32.8 (CH_2CO), 38.8 (CH_2NH), 76.1 (C-4), 102.8 (CH–CH), 114.7 (C-2.6), 126.7 (N–CCH_3), 137.1 (C-3.5), 148.0 (C-1), 170.1 (CO). Found, %: C 47.15, H 4.72, N 10.80. Calculated, %: C 47.01, H 4.73, N 10.96. $C_{15}H_{18}IN_3O$.

*1-(3,5-Dimethyl-1H-pyrazol-1-yl)-3-[(4-iodophenyl)amino]propan-1-one (**6**)*

A mixture of hydrazide **1** (0.92 g, 3 mmol), 2,4-pentanedione (0.5 g, 5 mmol), 2-propanol (10 ml), and conc. hydrochloric acid (0.5 ml) was heated under reflux for 7 h, then cooled, the precipitate was filtered off and washed with water. Compound **6** was crystallized from a mixture of 2-propanol and water (2:1).

White solid (0.84 g, 76 %), m.p. 93–94 °C (2-propanol). IR, *v*: 1,257 (ArNH), 1,311, 1,324, 1,361, 1,378 (CH_3), 1,594 (N=C), 1,728 (CO), 3,323 (CH_2NH). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.18$ (3H, s, CH_3), 2.47 (3H, s, CH_3), 3.15–3.45 (4H, m, $NHCH_2CH_2$), 5.98 (1H, t, $J = 4.6$ Hz, CH_2NH), 6.18 (1H, s, CH), 6.46 (2H, d, $J = 8.5$ Hz, $H_{ar\text{-}2.6}$), 7.35 (2H, d, $J = 8.5$ Hz, $H_{ar\text{-}3.5}$). ^{13}C NMR (75.4 MHz, DMSO-d_6): $\delta = 13.4$ (CH=CCH_3), 14.0 (CH-CCH_3), 34.5 (CH_2CO), 37.9

(CH_2NH), 76.0 (C-4), 111.1 (CH=C), 114.6 (C-2.6), 137.1 (C-3.5), 143.1 (CH=CCH$_3$), 148.1 (C-1), 151.3 (CH-CCH$_3$), 171.7 (CO). Found, %: C 45.38, H 4.39, N 11.23. Calculated, %: C 45.54, H 4.37, N 11.38. $C_{14}H_{16}IN_3O$.

*5-{2-[(4-Iodophenyl)amino]ethyl}-1,3,4-oxadiazole-2(3H)-thione (**8**)*

A solution of hydrazide **1** (1.53 g, 5 mmol), potassium hydroxide (0.84 g, 15 mmol), carbon disulfide (0.76 g, 10 mmol), and methanol (20 ml) was heated under reflux for 24 h, then volatile fractions were evaporated under reduced pressure. The residue was dissolved in water (30 ml), and the solution was acidified with acetic acid to pH 6. The formed residue was filtered off, washed with water, and dried. Compound **8** was crystallized from a mixture of 2-propanol and water (2:1).

Light-yellow solid (1.44 g, 83 %), m.p. 111–112 °C (2-propanol). IR, *v*: 1,172 (C=S), 1,261 (ArNH), 1,592 (N=C), 3,101 (NNH), 3,415 (CH_2NH). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.93$ (2H, t, $J = 6.6$ Hz; CH_2C=), 3.30–3.40 (2H, m, CH_2NH), 6.02 (1H, br. s. CH_2NH), 6.44 (2H, d, $J = 8.8$ Hz; $H_{ar\text{-}2.6}$), 7.34 (2H, d, $J = 8.8$ Hz; $H_{ar\text{-}3.5}$), 14.20 (1H, br. s, NNH). ^{13}C NMR (75.4 MHz, DMSO-d_6): $\delta = 25.0$ (CH_2C=), 38.8 (CH_2NH), 76.5 (C-4), 114.7 (C-2.6), 137.1 (C-3.5), 147.7 (C-1), 162.5 (OC=N), 177.7 (C=S). Found, %: C 34.23, H 3.99, N 12.16. Calculated, %: C 34.60, H 2.90, N 12.10. $C_{10}H_{10}IN_3OS$.

*4-Amino-5-{2-[(4-iodophenyl)amino]ethyl}-2,4-dihydro-3H-1,2,4-triazole-3-thione (**9**)*

A mixture of hydrazide **1** (1.53 g, 5 mmol), potassium hydroxide (0.85 g, 15 mmol), carbon disulfide (0.76 g, 10 mol), and methanol (20 ml) was heated under reflux for 24 h, then volatile fractions were evaporated under reduced pressure; the residue was dissolved in water (5 ml), and hydrazine hydrate (1.00 g, 20 mmol) was added. The mixture was heated under reflux for 30 h, diluted with 10 ml of water, cooled, and acidified with acetic acid to pH 6. The formed residue was filtered off, washed with water, and compound **9** was crystallized from 2-propanol.

White solid (1.07 g, 59 %), m.p. 137–138 °C (2-propanol). IR, *v*: 1,182 (C=S), 1,238 (ArNH), 1,590 (N=C), 3,171 (NNH), 3,271 (NNH_2), 3,395 (CH_2NH). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.86$ (2H, t, $J = 7.1$ Hz, CH_2C=), 3.36 (2H, t, $J = 7.1$ Hz, CH_2NH), 5.00–6.25 (3H, br. s, NH_2, CH_2NH), 6.46 (2H, d, $J = 8.5$ Hz, $H_{ar\text{-}2.6}$), 7.34 (2H, d, $J = 8.4$ Hz, $H_{ar\text{-}3.5}$), 13.51 (1H, s, NNH) ppm. ^{13}C NMR (75.4 MHz, [D_6]DMSO): $\delta = 24.3$ (CH_2C=), 39.3 (CH_2NH), 76.3 (C-4), 114.7 (C-2.6), 137.1 (C-3.5), 147.8 (C-1), 150.3 (CH_2C=), 165.8 (C=S). Found, %: C 33.45, H 3.48, N 19.18. Calculated, %: C 33.25, H 3.35, N 19.39. $C_{10}H_{12}IN_5S$.

*2-{3-[(4-Iodophenyl)amino]propanoyl}-N-phenylhydrazinecarboxamide (**10**)*

A mixture of hydrazide **1** (1.53 g, 5 mmol), phenyl isocyanate (0.71 g, 6 mmol), and 30 ml methanol was heated under reflux for 2 h. The reaction mixture was

cooled, the precipitate was filtered off, then washed with methanol, and compound **10** was crystallized from methanol.

White solid (0.92 g, 57 %), m.p. 157–158 °C (methanol). IR, *v*: 1,230 (ArNH), 1,604–1,641 (C=O), 3,313–3,374 (NH). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.42$ (2H, t, $J = 6.9$ Hz, CH_2CO), 3.25 (2H, dd, $^{1}J = 6.8$ Hz, $^{2}J = 13.0$ Hz, $\underline{CH_2}NH$), 5.86 (1H, t, $J = 5.6$ Hz, $CH_2\underline{NH}$), 6.40–7.50 (9H, m, H_{ar}), 8.04, 8.37, 8.71, 8.90, 9.01, 9.74 (3H, 6s, CONH). ^{13}C NMR (75.4 MHz, DMSO-d_6): $\delta = 32.9$ ($\underline{CH_2}CO$), 39.7 (CH_2N), 76.1, 114.7, 118.4, 121.8, 137.1, 139.5, 148.1 (C-ar), 155.3 (NHCONH), 170.6 ($CH_2\underline{CO}NH$). Found, %: C 45.19, H 3.92, N 13.28. Calculated, %: C 45.30, H 4.04, N 13.21. $C_{16}H_{17}IN_4O_2$.

*2-{3-[(4-Iodophenyl)amino]propanoyl}-N-phenylhydrazinecarbothioamide (**11**)*

A mixture of hydrazide **1** (1.53 g, 5 mmol), phenyl isothiocyanate (0.81 g, 6 mmol), and 50 ml methanol was heated under reflux for 3 h. The reaction mixture was cooled, the precipitate filtered off, then washed with methanol, and compound **11** was crystallized from methanol.

White solid (2.05 g, 93 %), m.p. 155–156 °C (methanol). IR, *v*: 1,172 (C=S), 1,229 (ArNH), 1,682 (C=O), 3,064, 3,133, 3,272 (NH). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.43$–2.50 (2H, m, CH_2CO), 3.27 (2H, dd, $^{1}J = 6.7$ Hz, $^{2}J = 12.6$ Hz, CH_2NH), 5.85 (1H, t, $J = 5.4$ Hz, $CH_2\underline{NH}$), 6.40–7.55 (9H, m, H_{ar}), 9.57 (2H, s, $\underline{NH}CS\underline{NH}$), 9.98 (1H, s, CONH). ^{13}C NMR (75.4 MHz, DMSO-d_6): $\delta = 33.0$ ($\underline{CH_2}CO$), 38.6 (CH_2NH), 76.2, 114.7, 125.0, 125.8, 128.0, 137.1, 139.0, 148.1 (C-ar), 170.4 (C=O), 180.8 (C=S). Found, %: C 43.49, H 4.03, N 12.58. Calculated, %: C 43.65, H 3.89, N 12.72. $C_{16}H_{17}IN_4OS$.

*5-{2-[(4-Iodophenyl)amino]ethyl}-4-phenyl-2,4-dihydro-3H-1,2,4-triazole-3-one (**12**)*

The semicarbazide **10** (0.85 g, 2 mmol) and 20 ml of aqueous 2 % sodium hydroxide solution were heated under reflux for 4 h, cooled, and acidified with acetic acid to pH 6. The formed residue was filtered off, washed with water, and dried. Compound **12** was crystallized from 2-propanol.

Light-blue solid (0.71 g, 87 %), m.p. 182–183 °C (2-propanol). IR, *v*: 1,591 (C=N), 1,706 (C=O), 3,175 (NH), 3,345 ($CH_2\underline{NH}$). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.60$ (2H, t, $J = 7.3$ Hz, CH_2C=), 3.15 (2H, dd, $^{1}J = 7.1$ Hz, $^{2}J = 13.6$ Hz, $\underline{CH_2}NH$), 5.88 (1H, t, $J = 6.0$ Hz; $CH_2\underline{NH}$), 6.18–7.57 (9H, m, H_{ar}), 11.72 (1H, s, NNH). ^{13}C NMR (75,4 MHz, [D_6]DMSO): $\delta = 25.5$ ($\underline{CH_2}C$=), 39.0 (CH_2NH), 76.1, 114.4, 127.5, 128.6, 129.4, 132.8, 137.0, 147.7 (C-ar), 145.0 (C=N), 154.3 (C=O). Found, %: C 47.49, H 3.84, N 13.95. Calculated, %: C 47.31, H 3.72, N 13.79. $C_{16}H_{15}IN_4O$.

*5-{2-[(4-Iodophenyl)amino]ethyl}-4-phenyl-2,4-dihydro-3H-1,2,4-triazole-3-thione (**13**)*

The thiosemicarbazide **11** (0.88 g, 2 mmol) and 20 ml of aqueous 2 % sodium hydroxide solution were heated under reflux for 3 h, cooled down, and acidified

with acetic acid to pH 6. The formed residue was filtered off, washed with water, and dried. Compound **13** was crystallized from 2-propanol.

White solid (0.80 g, 95 %), m.p. 156–157 °C (2-propanol). IR, *v*: 1,339 (C=S), 1,572 (C=N), 3,102 (NH), 3,393 ($CH_2\underline{NH}$). ^{1}H NMR (300 MHz, DMSO-d_6): $\delta = 2.63$ (2H, t, $J = 7.2$ Hz, CH_2C=), 3.18 (2H, t, $J = 7.0$ Hz, $\underline{CH_2}NH$), 5.50–6.0 (1H, br. s, $CH_2\underline{NH}$), 6.15–7.60 (9H, m, H_{ar}), 13.77 (1H, s, $\underline{N}NH$). ^{13}C NMR (75.4 MHz, DMSO-d_6): $\delta = 25.0$ ($\underline{CH_2}C$=), 39.2 (CH_2NH), 76.2, 114.4, 128.3, 129.4, 129.4, 133.6, 137.0, 147.5 (C-ar), 150.3 (C=N), 167.5 (C=S). Found, %: C 45.65, H 3.69, N 13.26. Calculated, %: C 45.51, H 3.58, N 13.27. $C_{16}H_{15}IN_4S$.

Antibacterial and antifungal activity testing in vitro

Diffusion technique

The antimicrobial activity of the compounds was evaluated by diffusion in agar on solid nutrient medium (nutrient agar—for bacteria, wort agar—for fungi). Disks (5 mm diameter) were soaked in 0.02 mg/ml solutions of compounds in DMSO. Disks were put on an exponentially growing plated culture. The microbial loading was 10^9 cells/1 ml. The plates were then incubated for bacteria for 24 h at 35 °C and for fungi for 48–72 h at 28–30 °C. The results were recorded by measuring the zones surrounding the disk. Control disk contained vancomycin (for bacteria) or nystatin (for fungi) as standard.

Serial dilution technique

Testing was performed in flat-bottomed 96-well tissue culture plates. The tested compounds were dissolved in DMSO; the concentration range was 500–1.9 μg/ml. Inoculum of bacteria and fungi was inoculated in nutrient medium (nutrient meat extract—for bacteria, wort—for fungi). The duration of incubation was 24–72 h at 37 °C for bacteria and 30 °C for fungi. The results were estimated according to the presence or absence of growth of microorganisms.

Computational studies using PASS

Prediction of the biological activity spectra for the synthesized compounds was performed using the PASS computer program [35, 36]. The latest PASS version (2012.11.22) predicts 6,400 kinds of biological activity based on analysis of a training set including information about ~330,000 drugs, drug candidates, lead compounds, etc. The average accuracy of prediction estimated in the leave-one-out cross-validation procedure for the whole training set is about 95 %. The PASS Online version is freely available for the scientific community via a website. Based on the PASS predictions, new pharmaceutical agents from various chemical classes with different kinds of biological activity have been discovered [35, 36]. PASS allows simultaneous prediction of the interaction of chemicals with many biological targets [67]. Based on the prediction results using the PharmaExpert computer program, a substance can be selected from the pleiotropic effects. In this way,

antihypertensive [68] and antiinflammatory [69] pharmacological agents with dual mechanisms of action have been identified, and nootropic properties of antihypertensive drugs unrelated to their antihypertensive activity were discovered [70]. It has also been shown to predict the biological activity of the components of herbal medicines [71].

PASS input information is presented as MOL or SDF files with the structural formulae of the compounds under study; the PASS output is presented as a list of probable activities with two estimated probabilities: Pa—the probability to be "active," and Pi—the probability to be "inactive."

The chemical descriptors used in the PASS analysis are called MNA [37]. They are automatically generated on the basis of the MOL file of a molecule. The list of unique MNA descriptors currently consists of more than 75,800 different items. New ones are added to this list if found in a novel compound, refreshing the training set. MNA descriptors are effectively applied in structure–activity relationship (SAR), quantitative SAR (QSAR), and similarity analyses. They can also be used as keys or fingerprints to cluster libraries of chemical compounds, select representative subsets from chemical databases, etc.

The list of predicted activities for a certain chemical compound is called the biological activity spectrum. It includes both the main and side pharmacological effects (e.g., antihypertensive, hepatoprotective, sedative, etc.), molecular mechanisms of action (5-hydroxytryptamine agonist, acetylcholinesterase inhibitor, adenosine uptake inhibitor, etc.), specific toxicities (carcinogenic, hallucinogenic, hepatotoxic, etc.), antitargets (ATPase inhibitor, CYP3A4 inhibitor, HERG channel blocker, etc.), terms associated with drug metabolism (CYP1A substrate, CYP1A1 human substrate, CYP3A4 substrate, etc.), terms associated with the transport of drugs (P-glycoprotein substrate, P-glycoprotein inhibitor, P-glycoprotein inductor, etc.), and terms associated with gene expression (APOA1 expression enhancer, ErbB-2 expression inhibitor, etc.).

The result of the prediction is presented as a list of activities with the appropriate Pa, which estimates the probability for the compound to be active or inactive for each type of biological activity. Its values vary from 0.000 to 1.000. The higher the Pa value, the lower the predicted probability of obtaining false positives in biological testing; For example, if one selects for testing only compounds for which a particular activity is predicted with $Pa > 0.9$, the expected probability of finding inactive compounds in the selected set is very low; however, about 90 % of active compounds are missed. If one lowers the Pa threshold to 0.8, the probability of finding inactive compounds is still low, but (only) about 70 % of active compounds are missed; etc. PASS uses the criterion $Pa = Pi$ as the default threshold; i.e., all compounds with $Pa > Pi$ are declared as being active.

Another aspect of the predictions is the compounds' novelty. If one limits oneself to high Pa values, one may find close analogs of known biologically active substances among the tested compounds; For example, for $Pa > 0.7$, the chance of finding the activity in an experiment is high, but some of the compounds may be close analogs of known pharmaceutical agents. If one chooses $0.5 < Pa < 0.7$, the chances of obtaining activity in the experiment are lower, but the compounds may be less similar to known pharmaceutical agents. For $Pa < 0.5$, the chances of

obtaining activity in the experiment are even lower, but if it is found, the compound might happen to be a new chemical entity.

The prediction results were obtained for all compounds using the PASS version which predicts 4,169 types of biological activity including 84 pharmacological effects, 3,965 molecular mechanisms, 110 metabolism-related actions, 2 gene expression regulation, 4 transporter-related actions, 4 side effects, and toxicity.

If PASS predictions are analyzed for a set of compounds from a certain chemical series, the predicted activities may be defined as typical (predicted for the majority of derivatives from this series) or minor (predicted only for some compounds) [42].

Therefore, the computer program PASS can be effectively used for the following purposes [66]: finding compounds with required properties and without undesirable side effects, revealing new effects and mechanisms of action of known substances from corporate and personal databases, selecting the most propitious compounds from a set of available samples for high-throughput screening, and determining the more relevant screens for a particular compound.

Acknowledgments This work was performed with partial financial support of the State Fund for Fundamental Research of Ukraine (agreement F53/97-2013) and the Russian Foundation for Basic Research (projects 13-04-90425, 12-07-00597).

References

1. E. Juaristi, A. Soloshonok, *Enantioselective synthesis of β-aminoacids* (Wiley, New York, 2005)
2. K. Anusevičius, I. Jonuškienė, V. Mickevičius, Monatsh. Chem. **144**, 1883 (2013)
3. S. Pelosi Jr., US Pat. 1976/3980689, Dec.14 (1976)
4. T. Aboul-Fadl, A.R. Khallil, Arzneim. Forsch. **53**, 526 (2003)
5. V. Mickevicius, A. Voskienė, I. Jonuskienė et al., Molecules **18**, 15000 (2013)
6. I. Jonuskiene, S. Kuusiene, V. Mickevicius, Chem. Technol. **49**, 53 (2008)
7. H. Habashita, M. Terakado, S. Nakade, T. Seko, US Patent 2005/0256160 A1. 17 Nov 2005
8. S. Dekeukeleire, M. D'hooghe, K.W. Turnroos, N. De Kimpe, J. Org. Chem. **75**, 5934 (2010)
9. I. Ojima, S. Inaba, M. Nagai, Synthesis **7**, 545 (1981)
10. S. Kanwar, S.D. Sharma, ChemInform **37**, 11 (2006)
11. V. Mickevicius, R. Vaickelioniene, G. Mikulskiene, N. Sewald, Chem. Heterocycl. Compd. **6**, 874 (2005)
12. K. Anusevičius, I. Jonuskienė, V. Mickevicius, Monatsh. Chem. **144**, 1883 (2013)
13. V.K. Tiwari, N. Tewary, D. Katiyjar, R.P. Tripathi, Monatsh. Chem. **138**, 1297 (2007)
14. R. Vaickelioniene, V. Mickevicius, G. Mikulskiene, Molecules **10**, 407 (2005)
15. R. Patino-Molina, I. Cubero-Lajo, M.J. Perez de Vega, M.T. Garcia Lopez, R. Gonzalez-Muniz, Tetrahedron Lett. **48**, 3613 (2007)
16. J.R. Merchant, D.S. Clothia, J. Chem. Soc. Perkin Trans. **1**, 932 (1972)
17. O. Prakash, D. Kumar, R.K. Saini, S.P. Singh, Synth. Commun. **15**, 2167 (1994)
18. V. Novikovaite, Z.J. Beresnevicius, G. Mikulskiene, Chem. Heterocycl. Compd. **43**(1), 113 (2007)
19. I. Tumosienė, K. Jonuskienė, Z. Kantminienė, J. Beresnevicius, Monatsh. Chem. **143**, 1441 (2012)
20. K. Anusevicius, V. Mickevicius, G. Mikulskiene, Chemija **21**, 127 (2010)
21. A.S. Anees, A. Amit, N. Sidqui, M. Amit, Indian J. Chem. Sect. B **44**, 838 (2005)
22. H. Bektaş, N. Karaali, D. Şahin, A. Demirbaş, Ş.A. Karaoglu, N. Demirbaş, Molecules **15**, 2427 (2010)
23. W.L. Jorgensen, Science **303**, 1813 (2004)
24. L.W. Tari, Structure-based drug discovery. Methods Mol. Biol. **841**, 1–385 (2012)
25. K. Heikamp, J. Bajorath, Chem. Biol. Drug Des. **81**, 33 (2013)
26. K. Anusevicius, R. Vaickelioniene, V. Mickevicius, G. Mikulskiene, J. Heterocycl. Chem. **50**, 309 (2013)

27. http://www.accelrys.com. Accessed 2014
28. http://www.chemspider.com. Accessed 2014
29. http://www.cactus.nci.nih.gov/cgi-bin/lookup/search. Accessed 2014
30. http://www.integrity.thomson-pharma.com. Accessed 2014
31. Y.C. Martin, J.L. Kofron, L.M. Traphagen, J. Med. Chem. **45**, 4350 (2002)
32. http://www.bioinformatics.charite.de/superpred. Accessed 2014
33. M. Dunkel, S. Günther, J. Ahmed, B. Wittig, R. Preissner, Nucleic Acids Res. **36**, W55 (2008)
34. C. Steinbeck, C. Hoppe, S. Kuhn, M. Floris, R. Guha, E.L. Willighagen, Curr. Pharm. Des. **12**, 2111 (2006)
35. http://www.way2drug.com/passonline. Accessed 2014
36. D.A. Filimonov, V.V. Poroikov, in *Chemoinformatics approaches to virtual screening*, ed. by A. Varnek, A. Tropsha (RSC Publishing, Cambridge, 2008), pp. 182–216
37. D. Filimonov, V. Poroikov, Yu. Borodina, T. Gloriozova, J. Chem. Inf. Comput. Sci. **39**, 666 (1999)
38. A. Lagunin, A. Stepanchikova, D. Filimonov, V. Poroikov, Bioinformatics **16**, 747 (2000)
39. A. Sadym, A. Lagunin, D. Filimonov, V. Poroikov, SAR QSAR Environ. Res. **14**, 339 (2003)
40. A. Geronikaki, D. Druzhilovsky, A. Zakharov, V. Poroikov, SAR QSAR Environ. Res. **19**, 27 (2008)
41. V. Poroikov, A. Lagunin, D. Filimonov, in *QSAR and molecular modelling in rational design of bioactive molecules*, ed. by E. Aki Sener, I. Yalcin (CADD & D Society, Ankara, 2005), pp. 514–515
42. A. Geronikaki, V. Poroikov, D. Hadjipavlou-Litina, D. Filimonov, A. Lagunin, R. Mgonzo, Quant. Struct. Activ. Relationsh. **18**, 16 (1999)
43. M.J. Walport, N. Engl, J. Med. **344**, 1058 (2001)
44. M.J. Walport, N. Engl, J. Med. **344**, 1140 (2001)
45. W.C. Song, M.R. Sarrias, J.D. Lambris, Immunopharmacology **49**, 187 (2000)
46. E. Munthe, H.M. Hoyeraal, S.S. Froland, O.J. Mellbye, E. Kess, J.B. Natvig, Rheumatology **6**, 43 (1975)
47. S. Ruddy, K.F. Austen, Ann. N.Y. Acad. Sci. **256**, 96 (1975)
48. O. Radillo, A. Nocera, A. Leprini, S. Barocci, T.E. Mollnes, M. Pocecco, M. Pausa, U. Valente, C. Betterle, F. Tedesco, Clin. Immunol. Immunopathol. **79**, 217 (1996)
49. S.J. Piddlesden, S. Jiang, J.L. Levin, A. Vincent, B.P. Morgan, J. Neuroimmunol. **71**, 173 (1996)
50. L.H. Perrin, P.H. Lambert, P.A. Miescher, J. Clin. Invest. **56**, 165 (1975)
51. M. Chrupcala, S. Pomer, G. Staehler, R. Waldherr, C. Kirschfink, Transpl. Int. **7**, 650 (1994)
52. B.G. Hoffstrom, A. Kaplan, R. Letso, R. Schmid, G.J. Turmel, D.C. Lo, B.R. Stockwell, Nat. Chem. Biol. **6**, 900 (2010)
53. S. Madhusudan, I.D. Hickson, Trends Mol. Med. **11**, 503 (2005)
54. M. Allary, J. Schrevel, I. Florent, Parasitology **125**, 1 (2002)
55. M. Flipo, T. Beghyn, V. Leroux, I. Florent, B.P. Deprez, R.F. Deprez-Poulain, J. Med. Chem. **50**, 1322 (2007)
56. R.B. Gennis, B. Barquera, B. Hacker, D.S.R. Van, S. Arnaud, A.R. Crofts, E. Davidson, K.A. Gray, F. Daldal, J. Bioenerg. Biomembr. **25**, 195 (1993)
57. E. Darrouzet, M. Valkova-Valchanova, T. Ohnishi, F. Daldal, J. Bioenerg. Biomembr. **31**, 275 (1999)
58. S. Ouchane, I. Agalidis, C. Astier, J. Bacteriol. **184**, 3815 (2002)
59. S.O. Pember, L.C. Fleck, W.K. Moberg, M.P. Walker, Arch. Biochem. Biophys. **435**, 280 (2005)
60. Performance Standards for Antimicrobial Disk Susceptibility Tests; Approved Standard-11th edn, Clinical and Laboratory Standards Institute, Wayne, Pennsylvania, (2012)
61. Method for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard-9th edn, Clinical and Laboratory Standards Institute, Wayne, Pennsylvania, (2012)
62. J. Frearson, P. Wyatt, Expert Opin. Drug Discov. **5**, 909 (2010)
63. W. Rhodes, J. Transl. Med. **10**(Suppl 2), A42 (2012)
64. D.M. Huryn, L.O. Resnick, P. Wipf, J. Med. Chem. **56**, 7161 (2013)
65. S. Ekins, B.A. Bunin, Methods Mol. Biol. **993**, 139 (2013)
66. L. Esser, B. Quinn, Y. Li, M. Zhang, M. Elberry, L. Yu, C.A. Yu, D. Xia, J. Mol. Biol. **341**, 281 (2004)
67. V.V. Poroikov, D.A. Filimonov, J. Comput. Aid. Molec. Des. **11**, 819 (2002)
68. A.A. Lagunin, O.A. Gomazkov, D.A. Filimonov, T.A. Gureeva, E.A. Dilakyan, E.V. Kugaevskaya, YuE Elisseeva, N.I. Solovyeva, V.V. Poroikov, J. Med. Chem. **46**, 3326 (2003)

69. A.A. Geronikaki, A.A. Lagunin, D.I. Hadjipavlou-Litina, PhT Eleftheriou, D.A. Filimonov, V.V. Poroikov, I. Alam, A.K. Saxena, J. Med. Chem. **51**, 1601 (2008)
70. S.A. Kryzhanovskii, R.M. Salimov, A.A. Lagunin, D.A. Filimonov, T.A. Gloriozova, V.V. Poroikov, Pharm. Chem. J. **45**, 605 (2012)
71. A. Lagunin, D.A. Filimonov, V.V. Poroikov, Curr. Pharm. Des. **16**, 1703 (2010)