European Journal of Medicinal Chemistry 270 (2024) 116362

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: www.elsevier.com/locate/ejmech

Research paper

An update on antibacterial AlkylGuanidino Ureas: Design of new derivatives, synergism with colistin and data analysis of the whole library☆

Claudia Ardino [a], Filomena Sannio [b], Giulio Poli [c], Salvatore Galati [c], Elena Dreassi [a], Lorenzo Botta [d,e], Jean-Denis Docquier [b,d], Ilaria D'Agostino [a,c,*]

[a] Department of Biotechnology, Chemistry and Pharmacy, University of Siena, via Aldo Moro 2, I-53100, Siena, Italy
[b] Department of Medical Biotechnologies, University of Siena, Viale Mario Bracci 16, I-53100, Siena, Italy
[c] Department of Pharmacy, University of Pisa, via Bonanno Pisano 6, I-56126, Pisa, Italy
[d] Lead Discovery Siena s.r.l., Via Vittorio Alfieri 31, I-53019, Castelnuovo Berardenga, Italy
[e] Department of Ecological and Biological Sciences, University of Tuscia, Largo dell'Università snc, I-01100, Viterbo, Italy

ARTICLE INFO

Keywords:
Alkylguanidine
Urea
Antibacterial agents
Guanidines
Cheminformatics
Machine-learning
Molecular dynamics
Membranes
Synergism
Colistin

ABSTRACT

Antimicrobial resistance (AMR) represents one of the most challenging global Public Health issues, with an alarmingly increasing rate of attributable mortality. This scenario highlights the urgent need for innovative medicinal strategies showing activity on resistant isolates (especially, carbapenem-resistant Gram-negative bacteria, methicillin-resistant *S. aureus*, and vancomycin-resistant enterococci) yielding new approaches for the treatment of bacterial infections. We previously reported AlkylGuanidino Ureas (AGUs) with broad-spectrum antibacterial activity and a putative membrane-based mechanism of action. Herein, new tetra- and mono-guanidino derivatives were designed and synthesized to expand the structure-activity relationships (SARs) and, thereby, tested on the same panel of Gram-positive and Gram-negative bacteria. The membrane-active mechanism of selected compounds was then investigated through molecular dynamics (MD) on simulated bacterial membranes. In the end, the newly synthesized series, along with the whole library of compounds (more than 70) developed in the last decade, was tested in combination with subinhibitory concentrations of the last resort antibiotic colistin to assess putative synergistic or additive effects. Moreover, all the AGUs were subjected to cheminformatic and machine learning analyses to gain a deeper knowledge of the key features required for bioactivity.

1. Introduction

Morbidity and mortality related to infections from drug-resistant pathogens have worryingly increased during the last decades [1]. As reported by the World Health Organization (WHO), none of the current antibacterial drug candidates could sufficiently address the lack of efficient treatments for infectious diseases [2]. Several pharmaceutical and biotech companies have deserted the field due to the low-profit margin [3], contributing to further weakening the already crumbling antibiotic research and development infrastructure. Although controversial, it seems that the COVID-19 outbreak refocused the attention of the scientific community on antimicrobial resistance (AMR), described as a "silent pandemic" [4]. The large number of secondary bacterial infections in both homecare and hospitalized COVID-19 patients became a significant concern [5,6] and the consequent increased use of antibiotics [7] is expected to involve a selective pressure on bacteria to develop resistant phenotypes [8]. Lately, the threatening emergency of resistant Gram-negative pathogens led to a renewed interest in colistin, known as the last resort antibiotic [9], representing one of the few therapeutic options against infections from multi-drug resistant (MDR) *Acinetobacter baumannii, Pseudomonas aeruginosa*, and *Klebsiella pneumoniae* [10]. However, resistance to polymyxins is not rare, being colistin a yet outdated drug also widely misused in animal husbandry and then put aside in lieu of more selective antibiotics.

☆ DEDICATION: The paper is dedicated to the memory of Prof. Maurizio Botta (1950–2019) who first conceived and guided this research project and then continued to inspire its latest progress in the field. Meeting him and discussing with him MedChem research matters have represented a great privilege for all of us. The smile for a new milestone achieved will always be the best picture of him in our mind.

* Corresponding author. Department of Pharmacy, University of Pisa, via Bonanno Pisano 6, I-56126, Pisa, Italy.
E-mail address: Ilaria.dagostino@unipi.it (I. D'Agostino).

https://doi.org/10.1016/j.ejmech.2024.116362
Received 3 February 2024; Received in revised form 21 March 2024; Accepted 25 March 2024
Available online 28 March 2024
0223-5234/© 2025 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Therefore, there is a rising interest in discovering new chemical classes exhibiting novel mechanisms of action, showing either direct activity or being used as an adjuvant in combination with older or obsolete antibiotics, whose activity might be restored against drug-resistant bacteria.

Recently, we reported the chemical class of AlkylGuanidino Ureas (AGUs) being endowed with notable antibacterial properties [11–13]. Among them, compounds **1** and **2** (Fig. 1) emerged as the best derivatives in terms of antibacterial activity, with minimal inhibitory concentrations (MICs) ranging from 0.5 to 16 μg/mL (Table 1).

Preliminary studies suggested bacterial membranes as the putative target of AGUs and a salt-bridge-like interaction with the negatively charged bacterial membranes is likely to be established [14] without provoking cell lysis, as assessed by a validated protocol on simulated membranes both in vitro and *in silico* [14].

The encouraging results on previously reported AGUs led us to focus our synthetic efforts first on designing derivatives of compound **1** obtained by changing the length of the methylene spacer between urea and guanidino functions and their *N*-substitutions. Then, a second generation of AGUs based on the molecular scaffold of simplified hit compound **2** was prepared. Antibacterial susceptibility on a panel of eight bacterial strains, including Gram-positive and Gram-negative microorganisms, highlighted some compounds with a potent antibacterial profile and gave relevant information on the chemical features modulating the biological activity, providing robust structure-activity relationship (SAR) data.

Overall, more than 70 antibacterial AGUs were prepared and a large dataset of bioactivity was collected, always addressing the following structural modification for further optimization processes. To fully assess the potential of such compounds, we investigated their effect on bacterial membranes and, thereby, their antibacterial effect when in combination with colistin. Moreover, thanks to the considerable amount of biological data generated so far for the whole library [11–13,15], we implemented a machine-learning protocol to correlate molecular features of the compounds with their antibacterial activity. This approach proved to help highlight the molecular features yielding broad-versus narrow-spectrum antibacterials, with the latter benefiting from more controlled side effects of such therapies.

2. Results and discussion

2.1. Design, synthesis, biological activity, and interaction with bacterial membranes of new AGU derivatives

2.1.1. Design and synthesis

Derivatives **1**, **3**–**10** (Table 1) were designed to further explore the impact of the aliphatic chain length between the urea and the guanidino moieties on the antibacterial activity. Symmetric compounds with the same spacer length on the urea *N,N*-substitution (3–6) and asymmetric ones (**7**–**10**) were afforded through two synthetic approaches. The symmetric ones were obtained by coupling bisguanilated carbamoyl derivatives [14] (Scheme S1 and S2 in *Supporting Information*), whereas the latter ones required homologs of key urea intermediate **11** [15], namely **12** and **13** (Scheme 1), accessible through the time-effective and high-yielding synthetic approach shown in Scheme S3.

Reagents and conditions: (i) TFA, dry CH_2Cl_2, r.t., 0.5 h, N_2; (ii) *N,N′*-DiBoc-*1H*-pyrazolecarboxamidine (for cpd **14**, **16**, and **18**) or cyclopropylmethyl derivative (for cpd **15**, **17** and **19**), DIPEA, THF, r.t., 16–24 h; (iii) TFA, CH_2Cl_2, sealed flask, r.t., 5–7 h; (iv) Hydrogenation procedure *A* or *B*; (v) appropriate **GA**, DIPEA, THF, r.t., 16–24 h (for cpd **22**–**26**); *N*-Succinimidyl *N*-methylcarbamate, dry DIPEA, dry CH_2Cl_2, r.t., 5 h, N_2 (for cpd **27**); or *N*-methylthiourea, dry DIPEA, dry CH_2Cl_2, r.t., 24 h, N_2 (for cpd **28**).

Orthogonally protected ureas **11**–**13,** bearing azido and trityl (Tr) groups, made chemical diversity rapidly accessible within the series. Indeed, several derivatives were furnished by subsequent selective deprotections and late-stage functionalizations [15]. Hence, **11**–**13** were detritylated and guanylated with the appropriate di-Boc-protected Guanylating Agent (GA, Scheme S4 in *Supporting Information*), such as the commercially available *N,N′*-DiBoc-1*H*-pyrazolecarboxamidine or the cyclopropylmethyl substituted one, [11,12,14,16,17] providing the bisguanylated derivatives **14**–**19**. Then, *tert*-butoxycarbonyl (Boc) groups of intermediates **14** and **15** were cleaved providing final compounds **20** and **21**. Furthermore, compounds **14** and **15** were again used, along with intermediated **16**–**19**, as starting points for additional functionalizations. Indeed, azido moieties of **14**–**19** were either reduced and subsequently guanylated, yielding compounds **22**–**26,** or reacted with *N*-succinimidyl *N*-methylcarbamate or *N*-methylisothiocyanate to afford *N*-methylurea (**27**) and *N*-methylthiourea (**28**), respectively. Boc groups were then hydrolyzed affording final compounds **7**–**10** and **29**–**31**.

Also, derivatives **32**–**37**, belonging to second-generation AGUs characterized by a different synthetic path, were obtained as previously reported [15] (Scheme S5 in *Supporting Information*) in order to gain preliminary SAR around the alkyl spacer length and the guanidino substituents.

2.1.2. In vitro antibacterial activity

All final compounds **3**–**10, 29a-m, 30**–**37** and intermediates **11**, **20**, and **21** were tested on a panel of eight representative Gram-positive and Gram-negative bacterial strains, as reported in Table 1.

Biological results for symmetric derivatives **3** and **4**, characterized by spacers composed of 7 and 9 carbon atoms respectively, seem to confirm that eight is the optimal length of the methylene spacer, as already emerged from previous SAR analysis on the library [15]. However, asymmetric derivatives **7**–**10**, endowed with different alkyl spacers also displayed interesting biological profiles. Also, compounds **5** and **6** (Table 1) showed overall comparable MICs to those of parent compound **1** on Gram-positive bacteria. Among analogues bearing two 6-methylenes spacers (**5**–**8**), some differences can be observed: derivative **5** slightly improves the antibacterial activity against *S. aureus*, whereas a large reduction in activity against Gram-negative strains was observed for derivatives **7** and **8** (Table 1). Moreover, although compound **10** showed worsened MIC values, its structural isomer **9** displayed an enhanced antibacterial profile, especially against Gram-positive strains, suggesting 6 methylenes and a terminal unsubstituted guanidine could be optimal to reach a good antibacterial potency. Interestingly, although derivatives **9** and **10** are structural isomers and their predicted physicochemical properties (data not shown) are comparable, a discrepancy in antibacterial potency emerged. Thus, we could assume that the distance between the two guanidines, thus two positive charges, plays a

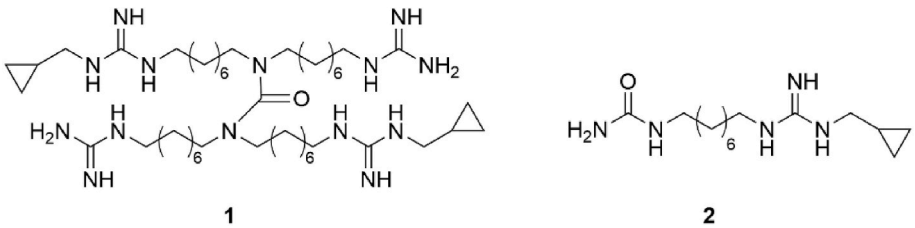
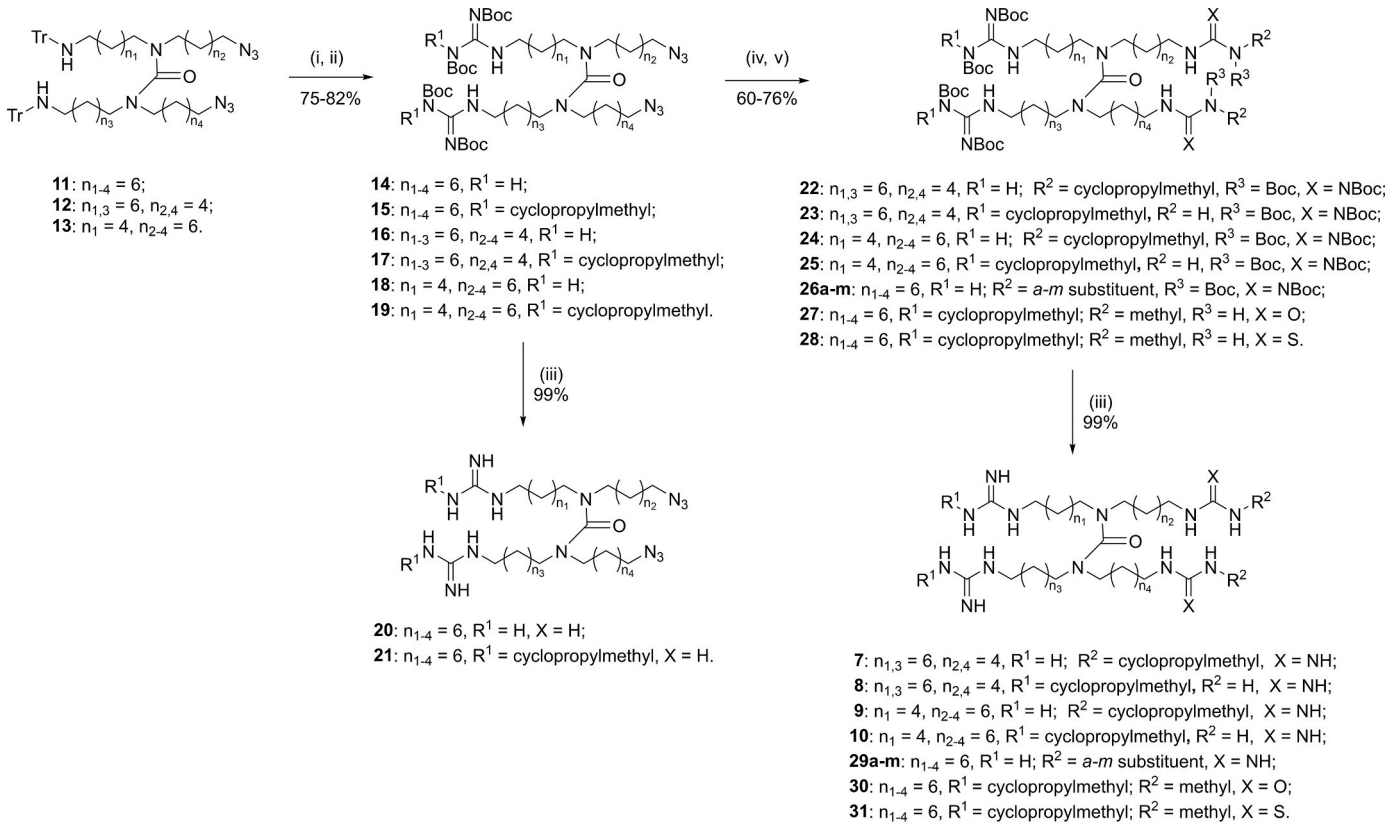


Fig. 1. Structure of compounds **1** and **2**.



11: n_{1-4} = 6;
12: $n_{1,3}$ = 6, $n_{2,4}$ = 4;
13: n_1 = 4, n_{2-4} = 6.

14: n_{1-4} = 6, R^1 = H;
15: n_{1-4} = 6, R^1 = cyclopropylmethyl;
16: n_{1-3} = 6, n_{2-4} = 4, R^1 = H;
17: n_{1-3} = 6, $n_{2,4}$ = 4, R^1 = cyclopropylmethyl;
18: n_1 = 4, n_{2-4} = 6, R^1 = H;
19: n_1 = 4, n_{2-4} = 6, R^1 = cyclopropylmethyl.

22: $n_{1,3}$ = 6, $n_{2,4}$ = 4, R^1 = H; R^2 = cyclopropylmethyl, R^3 = Boc, X = NBoc;
23: $n_{1,3}$ = 6, $n_{2,4}$ = 4, R^1 = cyclopropylmethyl, R^2 = H, R^3 = Boc, X = NBoc;
24: n_1 = 4, n_{2-4} = 6, R^1 = H; R^2 = cyclopropylmethyl, R^3 = Boc, X = NBoc;
25: n_1 = 4, n_{2-4} = 6, R^1 = cyclopropylmethyl, R^2 = H, R^3 = Boc, X = NBoc;
26a-m: n_{1-4} = 6, R^1 = H; R^2 = *a-m* substituent, R^3 = Boc, X = NBoc;
27: n_{1-4} = 6, R^1 = cyclopropylmethyl; R^2 = methyl, R^3 = H, X = O;
28: n_{1-4} = 6, R^1 = cyclopropylmethyl; R^2 = methyl, R^3 = H, X = S.

20: n_{1-4} = 6, R^1 = H, X = H;
21: n_{1-4} = 6, R^1 = cyclopropylmethyl, X = H.

7: $n_{1,3}$ = 6, $n_{2,4}$ = 4, R^1 = H; R^2 = cyclopropylmethyl, X = NH;
8: $n_{1,3}$ = 6, $n_{2,4}$ = 4, R^1 = cyclopropylmethyl, R^2 = H, X = NH;
9: n_1 = 4, n_{2-4} = 6, R^1 = H; R^2 = cyclopropylmethyl, X = NH;
10: n_1 = 4, n_{2-4} = 6, R^1 = cyclopropylmethyl, R^2 = H, X = NH;
29a-m: n_{1-4} = 6, R^1 = H; R^2 = *a-m* substituent, X = NH;
30: n_{1-4} = 6, R^1 = cyclopropylmethyl; R^2 = methyl, X = O;
31: n_{1-4} = 6, R^1 = cyclopropylmethyl; R^2 = methyl, X = S.

R^1 = H;
R^2 = **a**: 1-propyl; **b**: 2-propyl; **c**: 1-butyl; **d**: 2-methyl-1-propyl; **e**: but-2-en-1-yl; **f**:1-cyclopentyl; **g** (*E*)-geranyl; **h**: 2-methoxyethyl; **i**: 2-(morpholin-4-yl)ethyl; **l**: 2-thiomethoxyethyl; **m**: 2-(1-adamantyl)ethyl.

Scheme 1. Synthesis of AGUs **7**–**10**, **20**, **21**, and **29**–**31**.

crucial role in the bioactivity, as already observed [15]. This could allow guanidinium moieties to access specific sites of the membranes and, thus, to properly reach and interact with the negatively charged phospholipids heads. In fact, the typical distribution of bacterial phospholipids in the inner and outer leaflets of the bilayer is known to affect properties like membrane potential, permeability, shape, and surface charge [18]. It is indeed reported that in Gram-positive strains, phosphatidylinositol and phosphatidylethanolamine are preferentially distributed in the inner leaflet, whereas phosphatidylglycerol is more represented in the outer one, with cardiolipin being found in both of them [19]. Loss or alteration of natural lipid asymmetry might be translated into relevant biological consequences for the bacterium and several antimicrobial peptides and other membrane-active agents might induce a flip-flop of phospholipids within the membranes, scrambling their asymmetry [18].

Further investigation was performed by exploring the *N*-substituents on the guanidino moieties, which were selected according to their contribution to the total steric hindrance, lipophilicity, and electronic distribution of the whole structure. The cyclopropylmethyl group of compound **1** was spatially deconstructed or simplified to furnish linear and branched substituents, affording 1-propyl (**29a**), 2-propyl (**29b**), 1-butyl (**29c**), 2-methyl-1-propyl (**29d**) and but-2-en-1yl (**29e**) guanidino derivatives. These modifications should provide both a slight reduction in lipophilicity and a mild change in steric hindrance [20], allowing us to understand how the physicochemical properties affect the biological activity. While **29b** showed an enhanced broad-spectrum antibacterial activity, derivatives **29a** and **29c-e** displayed a biological profile comparable to **1** (Table 1). Interestingly, *N*-cyclopentyl guanidines (**29f**) improved the antibacterial activity (Table 1), probably due to their capability to establish a large number of hydrophobic interactions [20].

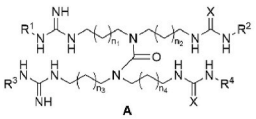
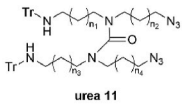
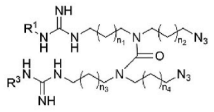
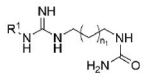
Then, we selected the geranyl chain as a representative unsaturated substituent due to its well-known antimicrobial activity [21–23]. Geranyl derivative **29g** was synthesized and tested, resulting in a broad-spectrum antibacterial activity compared to its parent compound **1** (Table 1).

Finally, guanidino substituents with different degrees of polarity were explored, by preparing polar (**29h**) and **29i**) and non-polar (**29l** and **29m**) representatives (Table 1). The presence of polar substituents was generally associated with an improvement of the broad-spectrum activity, compared to compound **1**. In fact, methoxyethyl derivative **29h**, benefiting from a big H-bonding capability, showed an improved antibacterial profile. Also, morpholinoethyl derivative **29i**, endowed with an increased number of positive charges and conceived to be more selective towards Gram-positive and Gram-negative [24], exerted a promising activity, especially against Gram-positive strains. However, while no significant information could be gained from data on the thiomethoxyethyl compound **29l**, the presence of bulky and hindered substituents, such as the adamantanoethyl group (**29m**), was not tolerated.

Furthermore, two of the four guanidines on traditional AGUs were substituted with the uncharged *N*-methylurea and *N*-methylthiourea functions in derivatives **30** and **31** respectively. As expected, a decrease in antibacterial activity was noticed (Table 1). The biological evaluation of the intermediates **11**, **20**, and **21** highlights a different behavior: while urea **11** resulted to be overall inactive, both of the bisguanylated **20** and **21** displayed moderate activity against Gram-positive strains (Table 1).

A focused library of second-generation AGUs was prepared by performing chemical derivatizations on the molecular scaffold of compound **2** in terms of the spacer length and the guanidine moiety

Table 1
MICs of AGUs derivatives and control antibiotics (colistin, vancomycin, and daptomycin) on representative bacterial strains.

Compound	*General Structure*	 A	 urea 11	Minimal Inhibitory Concentration [μg/mL][a]							
		 B	 C	Bacterial strains							
		n_1	R^1	n_2	R^2	n_3	R^3	n_4	R^4	X	Bsu	Efa	Sau	Spy	Eco	Kpn	Aba	Pae
1[b]	**A**	6		6	H	6		6	H	NH	2	2	2	1	4	2	8	16
3	**A**	5		5	H	5		5	H	NH	2	8	4	2	8	8	32	32
4	**A**	7		7	H	7		7	H	NH	4	4	2	4	8	16	8	16
5	**A**	4		4	H	6		6	H	NH	2	4	1	0.5	8	4	16	16
6	**A**	5		5	H	4		4	H	NH	2	4	4	1	16	8	32	32
7	**A**	6		4	H	6		4	H	NH	1	4	1	0.25	2	2	32	32
8	**A**	4		6	H	4		6	H	NH	1	8	1	0.25	4	4	64	128
9	**A**	6		6	H	6		4	H	NH	0.5	1	0.5	0.25	1	1	16	16
10	**A**	6		6	H	4		6	H	NH	4	4	4	1	8	4	64	128
29a	**A**	6		6	H	6		6	H	NH	1	1	1	0.5	0.5	1	8	8
29b	**A**	6		6	H	6		6	H	NH	0.5	0.5	0.5	0.25	0.5	1	8	8
29c	**A**	6		6	H	6		6	H	NH	1	2	2	1	2	2	4	8
29d	**A**	6		6	H	6		6	H	NH	2	4	4	2	4	4	8	16
29e	**A**	6		6	H	6		6	H	NH	2	2	2	1	2	2	8	16
29f	**A**	6		6	H	6		6	H	NH	1	1	1	0.25	1	1	4	8
29g	**A**	6		6	H	6		6	H	NH	2	2	2	1	2	2	4	8
29h	**A**	6		6	H	6		6	H	NH	0.5	1	0.5	0.5	1	2	8	16
29i	**A**	6		6	H	6		6	H	NH	1	2	1	1	4	4	32	64
29l	**A**	6		6	H	6		6	H	NH	1	2	1	0.5	2	4	16	32
29m	**A**	6		6	H	6		6	H	NH	16	16	16	8	32	32	64	128
30	**A**	6	H	6	Me	6	H	6	Me	O	4	8	4	4	16	16	64	128
31	**A**	6	H	6	Me	6	H	6	Me	s	4	4	4	2	8	16	128	64

(*continued on next page*)

Table 1 (*continued*)

Compound	*General Structure*									Minimal Inhibitory Concentration [μg/mL][a]							
										Bacterial strains							
		n_1	R^1	n_2	R^2	n_3	R^3	n_4	R^4	X	Bsu	Efa	Sau	Spy	Eco	Kpn	Aba	Pae
Urea 11		6		6		6		6			n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
20	**B**	6	H			6	H				8	16	8	2	32	32	64	n.a.
21	**B**	6				6					32	8	8	4	16	16	64	128
2[b]	**C**	6									1	1	1	0.5	2	1	16	16
32	**C**	4									128	128	128	128	n.a.	n.a.	n.a.	n.a.
33	**C**	6	H								64	n.a.	64	16	n.a.	n.a.	n.a.	n.a.
34	**C**	6	Me								128	128	64	32	n.a.	n.a.	n.a.	n.a.
35	**C**	6	Et								64	128	64	32	n.a.	n.a.	n.a.	n.a.
36	**C**	6									64	128	64	32	n.a.	n.a.	n.a.	n.a.
37	**C**	6									64	64	32	16	128	128	n.a.	n.a.
Ref. cpd		**VAN**									0.5	1	1	0.5	–	–	–	–
		COL									–	–	–	–	0.5	0.5	1	0.5
		DAP									1	1	0.5	0.12	–	–	–	–

[a] MICs are expressed as the average values calculated from experiments performed at least in triplicate. Colistin (COL), Vancomycin (VAN), and Daptomycin (DAP) antibiotics were used as controls in these assays. Tested organisms were *Bacillus subtilis* ATCC 6633 (Bsu); *Enterococcus faecalis* ATCC 19433 (Efa); *Streptococcus pyogenes* ATCC 12344 (Spy); *Staphylococcus aureus* ATCC 25923 (Sau); *Escherichia coli* CCUG^T (Eco); *Klebsiella pneumoniae* ATCC 13833 (Kpn); *Acinetobacter baumannii* ATCC 17978 (Aba); *Pseudomonas aeruginosa* ATCC 27853 (Pae). n.a.: not active at the highest tested concentration (256 μg/mL); -: not determined.

[b] Data for compounds **1** and **2** were already reported [14,15].

substitutions. Shortening the spacer length from 8 to 6 methylene (**32**), resulted in a complete loss of biological activity (Table 1). Different guanidino tails were chosen: starting from the unsubstituted guanidine moiety of **33**, derivatives substituted with methyl (**34**), ethyl (**35**), isopropyl (**36**), and crotyl (**37**) groups were also provided. Such modifications were overall detrimental to the activity, especially on Gram-negative microorganisms (**33**–**37**, Table 1). These findings seem to suggest a specific target for compound **2**, besides the interaction with bacterial membranes. Hence, further studies are required to elucidate the mode of action of compound **2** to proceed with a rational design of new derivatives.

2.1.3. Molecular dynamics on simulated bacterial membranes

Molecular modeling studies were performed on selected AGUs to evaluate their potential interaction with bacterial membranes. For this aim, we employed a molecular dynamics (MD) protocol previously validated as a reliable and efficacious tool for this class of compounds [15]. According to its biological profile, compound **29h** exerted the best broad-spectrum antibacterial activity among the newly synthesized derivatives bearing guanidine substituents with higher polarity compared to the parent compound **1**. To shed light on its behavior towards bacterial membranes, it was selected for MD simulation studies. In particular, compound **29h** was complexed with two different phospholipid bilayer systems solvated by a layer of explicit water molecules above each membrane leaflet. The first system, representing Gram-positive bacterial membranes, included only palmitoyloleoyl phosphatidylglycerol (POPG) molecules within the bilayer, while the second one was composed of a mixture of palmitoyloleoyl phosphatidylethanolamine (POPE) and POPG in a 6:4 ratio to represent Gram-negative bacterial membranes. Both ligand-membrane-water systems were subjected to a 600 ns MD simulation protocol aimed at evaluating the interactions between ligands and phospholipid membranes (see Materials and Methods for details). As previously predicted for the reference compound **1** [13,15], compound **29h** succeeds in establishing H-bonds and salt-bridge interactions with the polar heads of the POPG lipids early from the beginning of the MD. After about 50 ns the ligand assumes a carpet-like model of membrane interaction. As the simulation proceeds, the compound rapidly sinks among the membrane phospholipids, with its central ureidic moiety far below the level of the phosphate groups of the upper membrane leaflet. During the remaining ns of MD, **29h** averagely keeps this type of interaction with the membrane, maintaining most of its structure embedded within the upper leaflet, protruding the polar extremities of its branches below the level of the phosphate groups, thus also transiently altering the integrity of the membrane surface (Fig. 2A).

The distribution of the electron density of **29h** during the MD, confirms the substantial protrusion of the ligand within the POPG membrane. As shown in Fig. 2C, the density peak associated with the phosphate groups of the membrane is at about 19.5 Å from the bilayer center (see also Fig. S1 in *Supporting Information*), while the density peak related to the ligand is only about 15.4 Å away from it (thus about 4 Å below the phosphate groups). Moreover, the plot shows that only a small portion of the ligand density resides outside the membrane. Finally, the analysis of the ligand-membrane interactions showed that **29h** is able to form 10 H-bonds with 5 different phospholipids that are maintained for more than 200 ns of simulation. These results suggested that **29h** is able

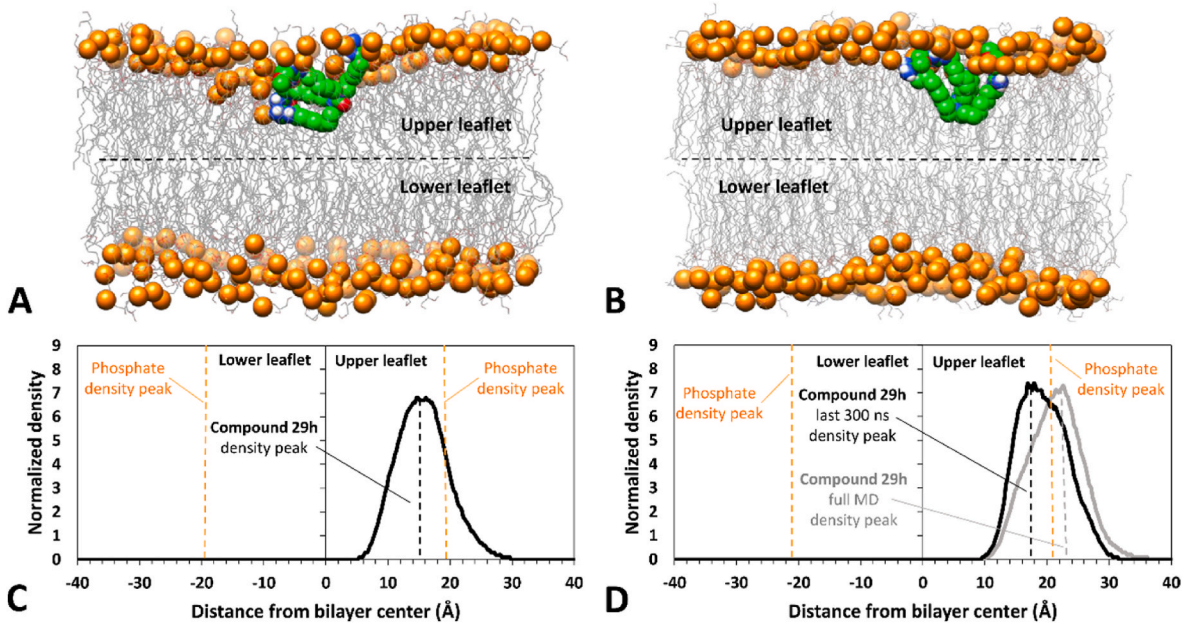


Fig. 2. Representative snapshots extracted after about 300 ns from the MD simulation with (**A**) pure POPG and (**B**) mixed POPE/POPG membrane in presence of compound **29h**. The ligand is shown as green spheres, while the phospholipids are represented as gray wires, and their phosphorous atoms are highlighted as orange spheres. **C-D**) Electron density profile of **29h**, with the corresponding peaks highlighted by dashed lines, obtained from the MD with the pure POPG and mixed POPE/POPG membrane, respectively. The density peaks of the lipid phosphate groups are indicated with orange dashed lines.

to strongly interact with the lipid bilayer and be retained among the membrane phospholipids. The results of the MD simulation performed with the mixed POPE/POPG membrane system showed a slightly different outcome, which seems to be consistent with the less pronounced activity of **29h** against Gram-negative bacteria compared to Gram-positive ones. In fact, although the ligand is able to rapidly interact with the membrane and form stable H-bonds with different phospholipids of both types (showing 7 H-bonds with 4 lipids maintained for more than 200 ns of MD), **29h** maintains a carpet-like model of interaction for almost the first half of the MD. After that time, its central and more lipophilic moiety protrudes towards the bilayer center and stays below the level of the phosphate groups (Fig. 2B), similarly to what observed with the pure POPG membrane. In fact, although the density peak of the ligand calculated for the whole MD simulation (around 22.5 Å) is slightly above the peak associated with the phosphate groups (around 21.0 Å, see also Fig. S2 in *Supporting Information*), in the second half of the MD the ligand is predominantly placed within the membrane, showing a density peak around 17.5 Å (Fig. 2D). These results suggest that compound **29h** is equally able to efficiently interact with the mixed POPE/POPG membrane, although it requires more time to reach the same level of internalization achieved within the pure POPG one. This aspect may be connected to the different lipid composition of the two membrane systems, which significantly affect the compactness of the phospholipid bilayer. In fact, the average area per lipid observed in the pure POPG membranes during the simulations was above 62 $Å^2$, whereas for the mixed POPE/POPG membranes a value of average area per lipid below 56 $Å^2$ was calculated. These results, which are in line with experimental and *in silico* data reported in literature [25], suggest that the presence of POPE lipids in the mixed POPE/POPG membranes increases the compactness of the phospholipid bilayer, thus making it more resistant to the internalization of the compound.

Aiming at broadening our understanding of their mode of interactions with membranes, compound **29m**, bearing bulky and strongly lipophilic substituents on the guanidine branches, proved also to be appealing for additional MD studies, along with compound **30**, endowed with neutral ureidic groups instead of the usual charged guanidines. Compound **29m** showed a similar mode of interaction both with the pure POPG and the mixed POPG/POPE membranes. Although it is able to rapidly form multiple long-lasting H-bond interactions with the phospholipids, the ligand does not protrude inside the upper membrane leaflet. Differently from what noticed for **29h**, **29m** only succeeds in inserting its two terminal hydrophobic groups within the membrane, while the rest of the molecule remains mainly outside the bilayer (Fig. 3A and B).

In fact, the density peak of the ligand is about 4 Å above that of the upper leaflet phosphate groups of both membrane bilayers (Fig. 3C and D). On top of that, in the simulation with the mixed POPG/POPE membrane, it is even possible to identify a small but distinct density peak around 16 Å, associated with the adamantyl groups (Fig. 3D), which represent the only moieties of **29m** protruded within the membrane (Fig. 3B). These results suggest that, due to their very high lipophilicity, the terminal adamantyl groups of **29m** tend to be trapped within the membrane to prevent their exposure to the solvent, which hampers the ligand from being properly internalized within the membrane. Nevertheless, even the internalization of the bulky adamantyl group alone might still affect the ordering, packing, and fluidity of the lipidic composition of the membrane, having an impact on the overall functionality of the membrane. For this reason, we cannot exclude that the poor antibacterial profile of **29m** results from a combination of additional factors, still related to the presence of the adamantyl groups, such as the hampered penetration through the thick layer of teichoic acids and peptidoglycan of the cell wall. On the other hand, the two ureidic moieties of **30** seem to seriously prevent the ligand from interacting with phospholipids. The MD simulation performed on the POPG membrane showed that only the guanidines of **30** are able to interact with the membrane, while the rest of the molecule remains fully solvent-exposed (Fig. S3A in *Supporting Information*). In fact, the density peak of **30** (28.5 Å) is about 9 Å above that of the phosphate groups of the upper membrane leaflet (Fig. S3C). Moreover, the ligand remains anchored to the membrane for just about 320 ns of the whole simulation, without forming any long-lasting H-bond and fluctuating within the bulk solvent for almost half of the MD. In fact, the presence of a second small ligand density peak about 35 Å away from the lower membrane leaflet demonstrates that **30** occasionally crosses the edges of the periodic system

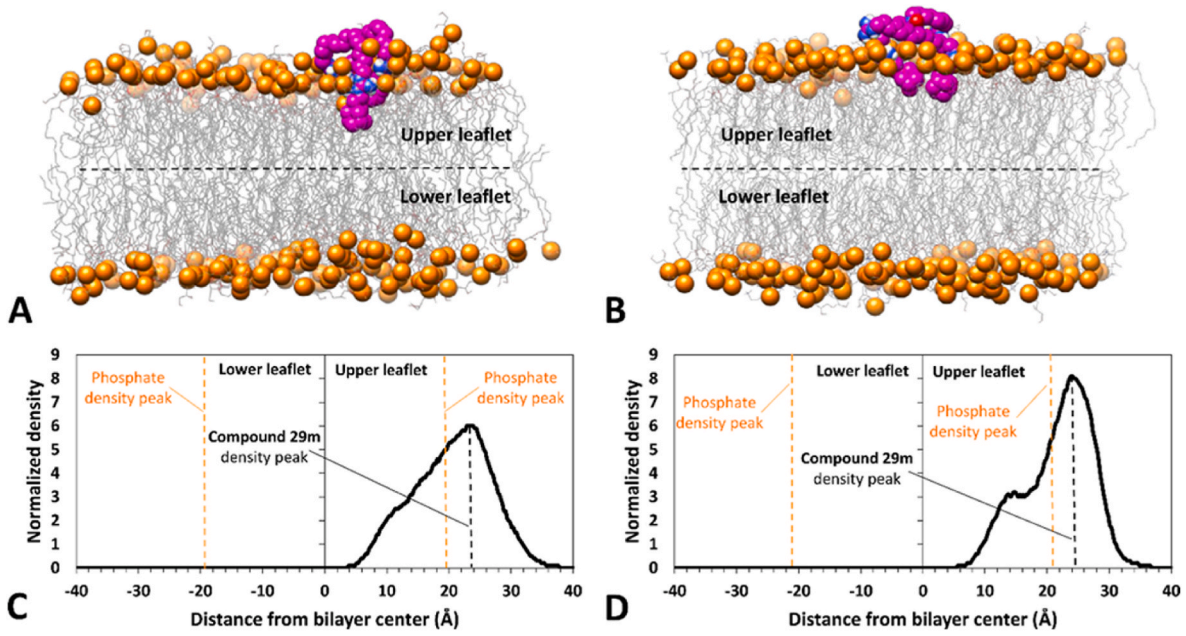


Fig. 3. Representative snapshots extracted after about 300 ns from the MD simulation with (**A**) pure POPG and (**B**) mixed POPE/POPG membrane, in presence of compound **29m**. The ligand is shown as purple spheres, while the phospholipids are represented as gray wires, and their phosphorous atoms are highlighted as orange spheres. **C-D**) Electron density profile of **29m**, with the corresponding peaks highlighted by dashed lines, obtained for the MD with the pure POPG and mixed POPE/POPG membrane, respectively. The density peaks of the lipid phosphate groups are indicated with orange dashed lines.

box, appearing at the other side of the membrane together with water molecules. Finally, the outcome of the MD simulation on the POPE/POPG membrane suggested that the ligand is less likely to interact with this kind of simulated membrane. Compound **30** does not even succeed in assuming a carpet-like model of interaction, as already highlighted with previously analyzed parent compounds. In fact, compound **30** engages with phospholipids for just a short time, while remaining totally immersed in the bulk solvent for almost the whole simulation (Fig. S3B). This is confirmed by the density plot of the ligand during the MD, which is characterized by almost two equally high peaks around 35 Å away from both sides of the bilayer center, indicating that **30** freely transits across the system box edges, remaining on average at the center of the solvent layers. These results may explain the low activity of the ligand against Gram-positive and, in particular, Gram-negative bacteria.

2.2. *The whole AGUs library: synergism with colistin and chemoinformatics*

2.2.1. *Evaluation of synergism with colistin*

Clinically relevant bacteria species are faster evolving towards ultra-resistant phenotypes, narrowing the therapeutic spectrum of colistin [26]. In this frame, *A. baumannii* still represents a critical priority to tackle, as it has rapidly evolved over decades becoming resistant to carbapenems, third-generation cephalosporins, and even colistin [27].

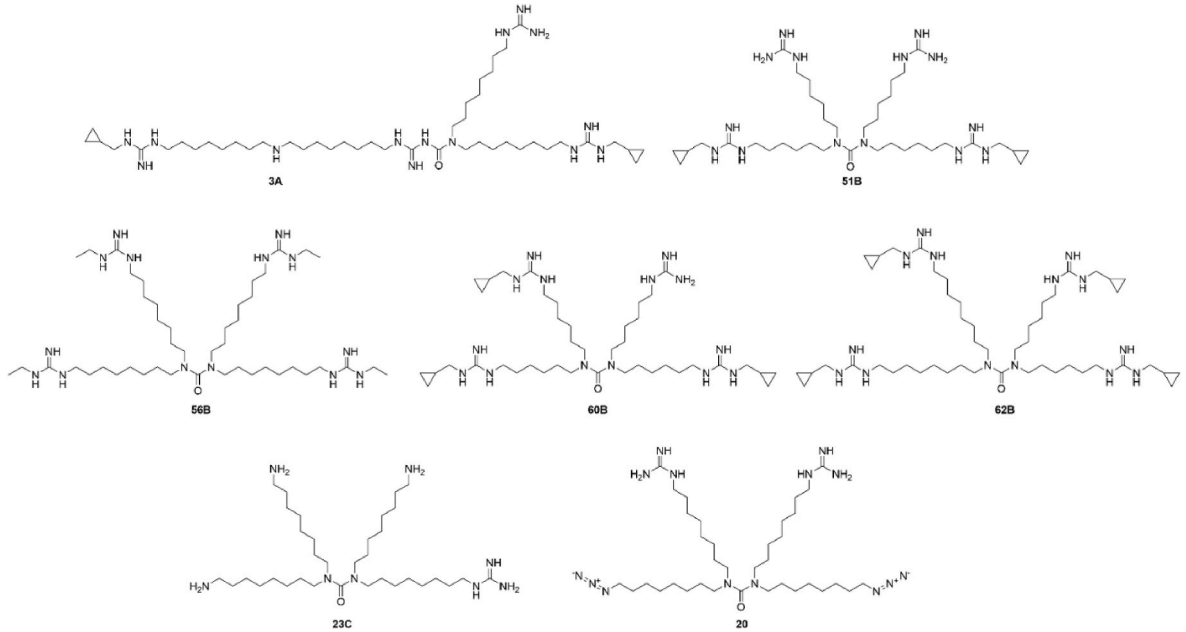


Fig. 4. Structures of representative AGU compounds involved in the synergism screening with colistin.

On top of that, clinical isolates of *A. baumannii* proved to be unsusceptible to more recent combination and synergistic therapies [28].

Our study throughout the years in the field of antibacterial AGUs was indeed aimed to provide new valuable options to treat such pathogens, proposing an emerging class of antibacterials endowed with very promising activity. However, only good or moderate antibacterial activity, if any, was achieved towards *A. baumannii* or the also challenging *P. aeruginosa* and *K. pneumoniae* strains, even for the most promising hits of the series (Table S1 in *Supporting Information*).

Hence, the whole AGUs library was tested in combination with a fixed subinhibitory concentration of colistin (0.25 x MIC). Colistin not only targets bacterial membranes but might behave as well as an antibiotic enhancer [29], representing sometimes a winning strategy to broaden the activity spectrum of selected antibiotics towards more challenging and resistant pathogens. Synergism was determined for the whole library towards two different *A. baumannii* strains, the reference ATCC 17978 and the clinical isolate N50, while some selected compounds were also tested against *E.coli*, *K. pneumoniae,* and *P. aeruginosa* (Table S2 in *Supporting Information*).

The extensive screening campaign for the identification of compounds with synergistic activities led to remarkable observations, especially on compounds reported in Fig. 4, and related data are reported in Table 2.

Direct antibacterial activity was expressed as minimal inhibitory concentration (MIC) and synergistic activity with subinhibitory concentrations of colistin (COL) was expressed as minimal antibacterial concentration (MAC). nd: not determined. For the complete data set see Table S2 in *Supporting Information.*

On one hand, the activity of some potent compounds was only merely affected by the presence of colistin, while some other inactive or poorly active compounds showed instead a variable degree of potentiation when co-tested with colistin. Interestingly, **3A** (Fig. 4) [11] characterized by a narrow spectrum and being active only on Gram-positive strains (Table S1), showed an impressive potentiation (up to 512-fold) of its activity in the presence of colistin with all tested Gram-negative strains, including nonfermenters and *A. baumannii* N50 (Table 2).

Compound **51B** (Fig. 4) [15], a close analogue of **3A**, also gained significant activity in presence of colistin, especially on *P. aeruginosa* and *A. baumannii* but not on the MDR-resistant isolate, indicating potential cross-resistance mechanisms impacting the synergistic activity (Table 2). However, it is even more puzzling that another close analogue, compound **56B** (Fig. 4) [15], showed the highest synergistic activity on that colistin-resistant isolate, while exerting a moderate direct-acting activity (MIC 64 μg/mL, Table 2). This also suggests that modification of the outer membrane composition, and therefore its chemical-physical properties, could be circumvented by subtle chemical modifications of the AGU backbone to obtain compounds tailored to interact specifically with a certain type of bacterial membrane. This could also be one of the factors determining the activity profile of some compounds of the AGU library, such as **60B** or **62B** (Fig. 4 and Table 2) [15], which showed a much better profile on Enterobacteriaceae rather than on non-fermenters.

The potentiation effect was also species-dependent for **23C** [15] and **20** (Fig. 4 and Table 2). In fact, colistin significantly enhanced the activity of **23C** (Fig. 4 and Table 2) on *P. aeruginosa* but only moderate effects were observed on *K. pneumoniae* and both the strains of *A. baumannii.*

Overall, these data also support a potential interest of AGUs as antibiotic adjuvants, and in particular with membrane-acting antibiotics. This would allow us to obtain potent antibacterial activity at lower concentrations of both antibacterials and could be potentially relevant for the last-resort polymyxins, which are characterized by suboptimal selectivity. This study also underlines the potential role of AGUs as resistance breakers, i.e. able to restore the in vitro activity of colistin on colistin-resistant extensively-drug resistant clinical isolates.

2.2.2. *Molecular descriptors and machine learning analyses of the AGU library*

As of today, more than 70 molecules, belonging to the AGU class have been reported. Intensive synthetic efforts have been made to improve the preparation and assess the bioactivity of the compounds, along with investigating their putative mechanisms of action. However, due to the considerable amount of data collected for the library, a comprehensive SAR analysis is not as trivial to perform as it might seem. In fact, the risk of excluding from consideration the biodiversity and the morphological/phenotypical differences within the tested bacterial strains is estimated to be high. Thus, the whole library of compounds (Table S1 in *Supporting Information*) was submitted first to an investigatory study of the molecular descriptors distribution shared within different groups of compounds, that were defined by their various range of antibacterial activity. Moreover, the library was also submitted to a machine learning-based approach to gain insights into the chemical features required for antibacterial activity.

The chemical space of the compounds was examined by analyzing the distribution of values of 123 different molecular descriptors calculated for the studied analogues, in relation to their activity against either Gram-positive or Gram-negative bacteria, which were considered separately. To simplify the analysis, the compounds were grouped into three different classes based on their antibacterial activity: highly active, partially active, and inactive (see Materials and Methods for details). With respect to the activity against Gram-positive bacteria two descriptors, number of H-bond donors (NumHDonors) and topological polar surface area (TPSA), showed an interesting distribution of values. From the distribution of the NumHDonors values it is clear that compounds inactive against Gram-positive bacteria exhibit a low number of H-bond donors compared to both partially and highly active compounds

Table 2
Antibacterial activity of representative AGU compounds.

		3A	51B	56B	60B	62B	23C	20
E.coli CCUG^T	**MIC** (μg/mL)	>256	16	2	8	4	32	32
	MAC (0.12 μg/mL COL)	0.5	1	nd	8	nd	0.5	32
	Potentiation fold (MIC/MAC)	>512	16	nd	1	nd	64	1
K. pneumoniae ATCC 13833	**MIC** (μg/mL)	>256	32	4	16	4	64	32
	MAC (0.12 μg/mL COL)	1	16	nd	nd	nd	32	2
	Potentiation fold (MIC/MAC)	>256	2	nd	nd	nd	2	16
P. aeruginosa ATCC 27853	**MIC** (μg/mL)	>256	256	64	128	128	128	>256
	MAC (0.25 μg/mL COL)	2	8	nd	nd	128	8	16
	Potentiation fold (MIC/MAC)	>128	32	nd	nd	1	16	>16
A. baumannii ATCC 17978	**MIC** (μg/mL)	>64	128	16	64	64	64	64
	MAC (0.25 μg/mL COL)	0.125	1	0.5	4	4	8	64
	Potentiation fold (MIC/MAC)	>256	128	32	16	16	8	1
A. baumannii N50	**MIC** (μg/mL)	>64	>256	64	nd	64	32	32
	MAC (2.0 μg/mL COL)	0.5	>256	0.25	nd	2	4	32
	Potentiation fold (MIC/MAC)	>128	1	256	nd	32	8	1

(Fig. 5A). In particular, while most of the inactive compounds present an average of 5 H-bond donors, most of the highly active compounds show between 10 and 15 H-bond donors. Similarly, the distribution of values of TPSA (Fig. 5B) demonstrates that inactive compounds present a relatively low topological polar surface area, with a peak around 100 Å^2, whereas the majority of highly active compounds present a TPSA between 200 and 300 Å^2. A third descriptor that showed a distribution of values comparable to that observed for NumHDonors and TPSA was Net Charge. In particular, highly active compounds present a peak of Net Charge around 4, while inactive compounds showed generally lower Net Charge values, with a peak around 1 (Fig. S4). This is consistent with the fact that, in general, AGUs bearing a low number of charged branches, such as compounds **32**–**37**, are endowed with very weak activities against Gram-positive bacteria. These results suggest that compounds active against Gram-positive strains require a number of H-bond donor groups between 10 and 15, a TPSA around 250 Å^2, and a net charge around 4. Furthermore, the presence of charged moieties and H-bond donor groups certainly contributes to expand the polar surface area of the compounds and may justify the similar statistical distributions observed for the three descriptors among the compounds (Fig. 5 and Fig. S4 *in Supporting Information*).

On the contrary, the analysis of molecular descriptors taking into consideration the activities of the compounds against Gram-negative bacteria showed a generally homogeneous distribution of descriptors among inactive and active compounds, which did not suggest any potential SAR.

In addition to this analysis, an application of a machine learning strategy was performed to provide a more comprehensive insight concerning the relationship between molecular descriptors and the activity of the compounds library. Two different machine learning models employing a Random Forest algorithm, which considered the activity of the compounds against Gram-positive and Gram-negative bacteria, respectively, were developed. The compounds were considered as grouped into the same three classes used for the previous analysis: highly active, partially active, and inactive (see Experimental section for details). Then, to simplify the training of the models, for each of the two datasets only the two most represented compound classes were selected. Specifically, for Gram-positives the model was trained with highly active and inactive compounds, while for Gram-negatives the partially active and inactive classes were considered. Therefore, the Gram-positive model was trained with 64 compounds and the Gram-negative model was trained with a dataset of 72 instances. The performance of the models was evaluated by 10-fold cross-validation, in which the data set was divided recurrently by random sampling into 70% training compounds and 30% test compounds. The performance of the model for Gram-positive classification was found to have an average Matthew's Correlation Coefficient (MCC) value of 0.53, while for Gram-negative models the MCC was found to have an average value of 0.54. These results demonstrated that both models were able to perform more than 75% of correct predictions, thus showing a satisfying performance. Both models were then subjected to a protocol aimed at deriving the impact of the molecular descriptors on the final predictions. In particular, we applied the SHapley Additive exPlanations (SHAP) approach to estimate the importance of molecular descriptors for model predictions. The tree-based SHAP approach was applied to rationalize the classification of the compounds based on their activity against Gram-positive and Gram-negative bacteria (see Materials and Methods for details). The importance of the features was derived and ranked according to their magnitude. Specifically, a SHAP value greater than 0 corresponds to a positive impact of the feature on the prediction of active compounds, while a negative value indicates that the feature affects the predictions addressing them towards the inactive class. The analysis of the top 10 features with a positive impact on Gram-positive classification showed that, among several molecular descriptors that can be less easily correlated with chemical structure, it is possible to find the presence of NumHDonors, which has a significant SHAP value. In particular, as indicated by the SHAP ranking diagram in Fig. 6A, a high value of NumHDonors is favorable for the prediction of active compounds. This result is in agreement with the molecular descriptor distribution analysis (Fig. 5A), which confirms that an appropriate and high value of the chemical groups acting as H-bond donors is a key attribute for identifying compounds active against Gram-positive bacteria.

Interestingly, for what regards the classification of compounds based on their activity against Gram-negative bacteria (Fig. 6B), the NHOH descriptor was found to be the third one in terms of importance for predicting the activity. This feature takes into account the number of NH and OH groups within the structure of the compound. In more detail, the SHAP analysis suggests that a low number of NH and OH groups is important for classifying compounds as active against Gram-negative strains. Since such groups can generally act as H-bond donors, these results may indicate an opposite trend with respect to that observed for the classification of compounds based on the activity against Gram-positive strains. Overall, these results could represent a first step to help define SARs aimed at identifying key features to optimize antimicrobial compounds against both Gram-positive and Gram-negative strains.

2.3. Toxicity and PAINS prediction

VenomPred tool [30] was employed to assess *in silico* cytotoxicity for the whole set of compounds, resulting in being safe for carcinogenicity, hepatotoxicity, mutagenicity, and estrogenicity [31]. Also, the search for pan-assay interference compounds (PAINS) resulted in no alerts for filters A, B, and C [32,33].

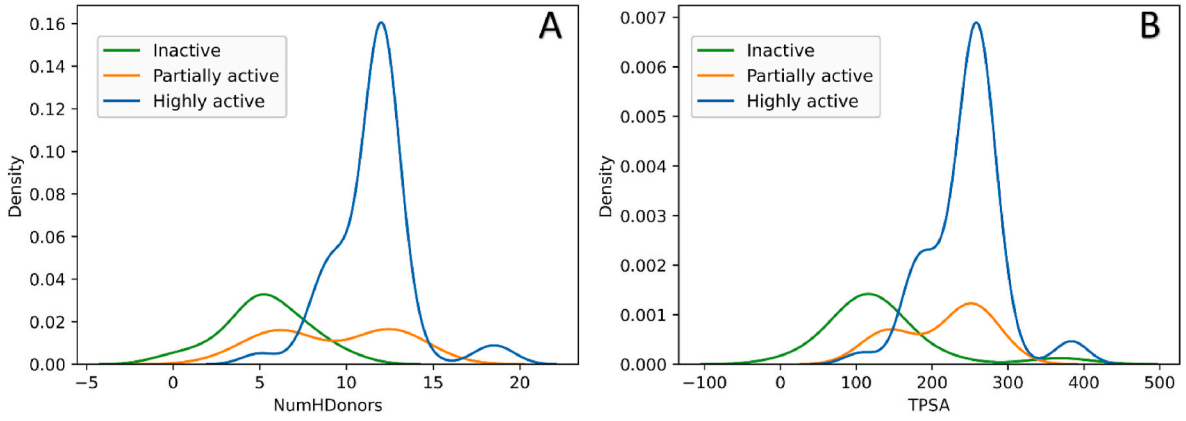


Fig. 5. Statistical distribution of molecular descriptors. The distributions for NumHDonors (A) and TPSA (B) according to Gram-positive classification are shown.

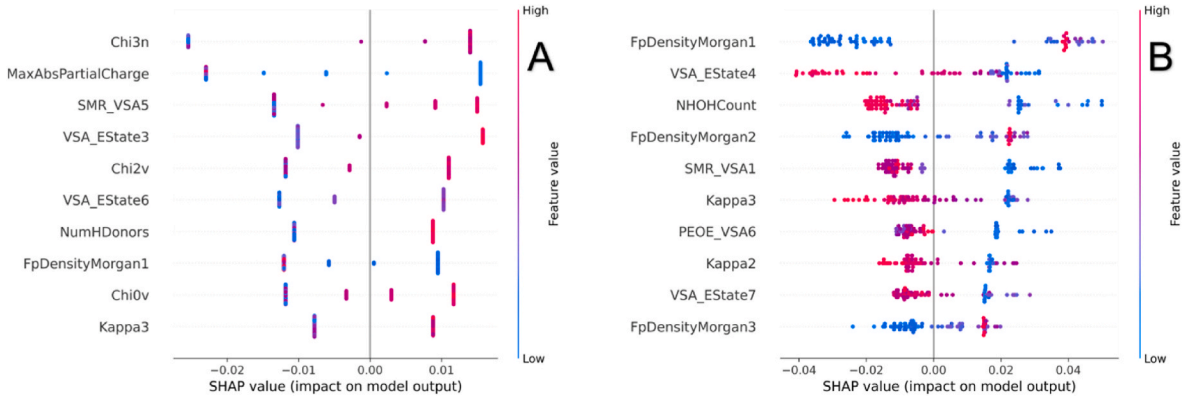


Fig. 6. Interpretation of model-based compound activity class prediction. The extracted features were ranked and the top 10 are represented for Gram-positive (A) and Gram-negative (B) classification.

3. Conclusion

In this work, we presented the development of a new AGU series of 30 derivatives including tetra- and mono-guanylated ureas as perspective antibacterial agents. MIC assessment enabled the extension of robust SAR data, leading to relevant assumptions, *e.g.,* at least one 6-methylenes spacer bearing an unsubstituted guanidine is required to boost the antibacterial activity, as suggested by biological results for **9.** The removal (**11**) or the decrease in the number (**20** and **21**, Table 1) of guanidino functions causes a reduction or a complete loss of antibacterial activity. The same trend was detected when guanidines were replaced with isosteric urea (**30**) or thiourea (**31**, Table 1) functions. Branched (**29b**) or cyclic (**29f**, Table 1) guanidino substituents improve the antibacterial profile, especially against Gram-positive strains. In addition, the activity on tested Gram-positive bacteria was promising with linear or cyclic polar guanidino substituents of **29h** and **29i** respectively, that are more tolerated over small or bulky non-polar groups of **29l** and **29m**, which showed a significant decrease in activity. Six derivatives belonging to a second-generation AGU series were synthesized to further explore the newly developed scaffold of compound **2**. The lack of activity of derivatives **32**–**37** (Table 1) suggests that compound **2** could exhibit a unique and specific mode of action. MD studies performed on representative derivatives (**29h**, **29m**, and **30**) suggested that both the presence of bulky lipophilic groups on the guanidine branches and the replacement of the guanidine moieties with non-charged substituents may hamper the proper interaction of the compounds within the membranes and their insertion into the bilayer. Furthermore, the whole library of AGUs, including the herein presented 30 new compounds and 44 previously reported derivatives [11,12,14,15], were also investigated as potential enhancers of colistin antibacterial activity. Our results showed that some AGUs, particularly compound **56B**, exerted a potent synergistic activity with colistin on colistin-resistant *A. baumannii* clinical isolate, underlining the potential applicability of these derivatives as resistance breakers. Finally, cheminformatics based on molecular features distribution and machine-learning analyses performed on the entire library of AGUs highlighted that the number of H-bond donor groups within the compounds can represent a discriminatory characteristic for their activity against Gram-positive or negative bacteria. Also, a high number of H-bond donor groups and a high net charge seem to be correlated with antibacterial activity towards Gram-positive strains.

In conclusion, our study provides novel insights into the AGU chemical class as promising antibacterial agents and their mechanism of action, with MD simulations providing a rationale for their interaction with membranes. Those findings outline the potential of those compounds to be further optimized, alone or in combination as adjuvants, to address the global issue of antibiotic resistance.

4. Material and methods

4.1. Chemistry

General Chemistry. All commercially available chemicals and solvents were used as purchased. CH_2Cl_2 was dried over calcium hydride and Tetrahydrofuran (THF) was dried over sodium and benzophenone before use. Degassed dry CH_2Cl_2 was prepared from freshly dried CH_2Cl_2 via three or more freeze-pump-thaw cycles. Dry *N,N*-diisopropylethylamine (DIPEA) was purchased from Merck. Dry triethylamine (TEA) was dried over KOH and distilled under a nitrogen atmosphere. Anhydrous reactions were performed into flame-dried glassware after three cycles of vacuum/dry nitrogen and were run under a positive pressure of dry nitrogen. Chromatographic separations were performed on columns packed with silica gel (230–400 mesh, for flash technique). TLCs were visualized under UV light and stained with ninhydrin, bromocresol green, or basic permanganate stains. ^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker Avance 400 or 600 spectrometers, at 400/600 and 100/150 MHz, respectively. Spectra are reported in parts per million (δ scale) and internally referenced to the $CDCl_3$ and CD_3OD signal, respectively at δ 7.26 and 3.31 ppm. Chemical shifts for carbon are reported in parts per million (δ scale) and referenced to the carbon resonances of the solvent ($CDCl_3$ at δ 77.00 and CD_3OD at δ 49.00 ppm). Data are shown as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, qi = quintet, m = multiplet and/or multiplet resonances, bs = broad singlet), coupling constants (*J*) in Hertz (Hz), and integration. Mass spectra (LCMS ESI) were acquired using an Agilent 1100 LC-MSD VL system (G1946C) by direct injection with a 0.4 mL/min flow rate using a binary solvent system of 95/5 MeOH/H_2O. UV detection was monitored at 221 or 254 nm. Mass spectra were acquired in positive or negative mode scanning over the mass range 100–1500 *m/z*, using a variable fragmentor voltage of 0–70 V.

Determination of purity. The purity of final products (≥95%) was assessed by HPLC-UV-MS analysis as reported [15].

4.1.1. Characterization of final compounds 3–11, 20–21, 30-37

Procedure previously reported by using the appropriate Boc-protected urea.

1,3-bis(7-carbamimidamidoheptyl)-1,3-bis({7-[*N*'-(cyclopropylmethyl)carbamimidamido] heptyl})urea trifluoroacetate salt (3): Starting material: **40**. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 0.31 (m, 4H); 0.62 (m, 4H); 1.10 (m, 2H); 1.40 (m, 24H); 1.56 (m, 8H); 1.63 (m, 8H); 3.09 (d, 4H, J = 7.2 Hz); 3.16–3.24 (m, 16H). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 2.5; 9.5; 26.0; 26.5; 27.5; 28.3; 28.8; 28.9; 29.2; 41.0; 41.1; 44.2, 46.9; 47.1; 48.1; 151.0; 155.6; 161.0; 165.4. LCMS *m/z* (ES+) = 198.2 $[M + 4H]^{4+}$, 264.0 $[M + 3H]^{3+}$, 395.5 $[M + 2H]^{2+}$. Yield: 99%.

1,3-bis(9-carbamimidamidononyl)-3-{9-[*N*'-(cyclopropylmethyl)carbamimidamido]nonyl}-1-{8-[*N*'-(cyclopropylmethyl)carbamimidamido]octyl}urea trifluoroacetate salt (4): Starting material: **41**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 0.28 (m, 4H); 0.61 (m, 4H); 1.12 (m, 2H); 1.41 (m, 24H); 1.55 (m, 8H); 1.65 (m, 8H); 3.12 (d, 4H, J = 7.2 Hz); 3.23 (m, 16H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 2.3; 9.6; 26.0; 26.6; 27.5; 28.3; 28.8; 29.0; 29.3; 41.0; 41.3; 44.2, 47.0; 47.2; 48.1; 151.0; 155.7; 161.1; 165.4. LCMS m/z (ES+) = 223.3 $[M + 4H]^{4+}$, 297.4 $[M + 3H]^{3+}$, 444.7 $[M + 2H]^{2+}$. Yield: 99%.

3-(6-carbamimidamidohexyl)-1-(8-carbamimidamidooctyl)-3-{6-[*N*'-(cyclopropylmethyl)carbamimidamido]hexyl}-1-{8-[*N*'-(cyclopropylmethyl)carbamimidamido]octyl}urea trifluoroacetate salt (5): Starting material: **42**. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 0.25 (m, 4H); 0.56 (m, 4H); 1.04 (m, 2H); 1.30 (m, 24H); 1.50–1.56 (m, 16H); 2.88 (d, 4H, J = 7.6 Hz); 3.03–3.18 (m, 16H) ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 2.4; 9.5; 26.2; 28.4; 28.9; 39.3; 41.0; 45.8; 52.4; 155.9; 157.2; 165.8. LCMS m/z (ES+) = 263.6 $[M + 3H]^{3+}$; 395.1 $[M + 2H]^{2+}$, 789.0 $[M + H]^+$. Colorless oil. Yield: 99%.

1-(7-carbamimidamidoheptyl)-3-(6-carbamimidamidohexyl)-1-{7-[*N*'-(cyclopropylmethyl)carbamimidamido]heptyl}-3-{6-[*N*'-(cyclopropylmethyl)carbamimidamido]hexyl}urea trifluoroacetate salt (6): Starting material: **43**. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 0.32 (m, 4H); 0.62 (m, 4H); 1.10 (m, 2H); 1.41 (m, 20H); 1.56–1.63 (m, 16H); 3.10 (d, 4H, J = 7.2 Hz); 3.18–3.24 (m, 16H). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 2.4; 9.5; 26.3; 28.3; 28.8; 28.9; 39.1; 41.0; 44.9; 47.2; 52.5; 155.9; 157.2; 161.2; 165.4. LCMS m/z (ES+) = 254.6 $[M + 3H]^{3+}$; 381.3 $[M + 2H]^{2+}$; 761.7 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(6-carbamimidamidohexyl)-1,3-bis({8-[*N*'-(cyclopropylmethyl)carbamimidamido]octyl})urea trifluoroacetate salt (7): Starting material: **22**. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 0.27 (m, 4H); 0.58 (m, 4H); 1.06 (m, 2H); 1.30 (m, 26H); 1.53-1.51-1.55 (m, 8H); 1.54–1.60 (overlapped, m, 8H); 3.06 (d, 4H, J = 7.2 Hz); 3.16 (m, 16H). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 2.4; 9.5; 26.0; 26.2; 26.5; 27.4; 27.5; 28.4; 28.5; 28.8; 28.9; 41.0; 45.8; 156.0; 157.2; 161.6; 165.8. LCMS m/z (ES+) = 264.0 $[M + 3H]^{3+}$; 395.5 $[M + 2H]^{2+}$; 790.1 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis({6-[*N*'-(cyclopropylmethyl)carbamimidamido]hexyl})urea trifluoroacetate salt (8): Starting material: **23**. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 0.27 (m, 4H); 0.58 (q, 4H, J = 5.2 Hz); 1.06 (m, 2H); 1.31 (m, 26H); 1.52–1.55 (m, 8H); 1.54–1.59 (overlapped, m, 8H); 3.06 (d, 4H, J = 6.8 Hz); 3.16 (m, 16H). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 2.4; 9.5; 26.0; 26.2; 26.5; 27.4; 27.5; 28.4; 28.5; 28.8; 28.9; 41.0; 45.8; 156.0; 157.2; 161.6; 165.8. LCMS m/z (ES+) = 264.0 $[M + 3H]^{3+}$; 395.5 $[M + 2H]^{2+}$; 790.1 $[M + H]^+$. Colorless oil. Yield: 99%.

1-(6-carbamimidamidohexyl)-3-(8-carbamimidamidooctyl)-1,3-bis({8-[*N*'-(cyclopropylmethyl)carbamimidamido]octyl})urea trifluoroacetate salt (9): Starting material: **24**. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 0.28 (m, 4H); 0.58 (m, 4H); 1.06 (m, 2H); 1.32 (m, 28H); 1.52–1.56 (m, 8H); 1.55–1.59 (overlapped, m, 8H); 3.06 (d, 4H, J = 6.8 Hz); 3.16 (m, 16H). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 2.4; 9.5; 26.0; 26.2; 26.5; 27.5; 28.3; 28.4; 28.8; 28.9; 40.9; 41.0; 41.1; 45.8; 156.0; 157.2; 161.3; 165.8. LCMS m/z (ES+) = 273.3 $[M + 3H]^{3+}$; 409.4 $[M + 2H]^{2+}$; 817.7 $[M + H]^+$; 839.7 $[M + Na]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1-{6-[*N*'-(cyclopropylmethyl)carbamimidamido]hexyl}-3-{8-[*N*'-(cyclopropylmethyl)carbamimidamido]octyl}urea trifluoroacetate salt (10): Starting material: **25**. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 0.28 (m, 4H); 0.59 (m, 4H); 1.07 (m, 2H); 1.36 (m, 28H); 1.53 (m, 8H); 1.59 (m, 8H); 3.06 (d, 2H, J = 6.8 Hz); 3.13–3.21 (m, 16H).^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 2.4; 9.5; 26.0; 26.2; 26.5; 27.5; 28.3; 28.4; 28.8; 28.9; 40.9; 41.0; 41.1; 45.8; 156.0; 157.2; 161.3; 165.8. LCMS m/z (ES+) = 205.1 $[M + 4H]^{4+}$;273.2 $[M + 3H]^{3+}$; 409.3 $[M + 2H]^{2+}$, 817.7 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-azidooctyl)-1,3-bis(8-carbamimidamidooctyl)urea trifluoroacetate salt (20): Starting material: **14**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.35 (m, 32H); 1.51–1.54 (m, 8H); 1.55–1.61 (m, 8H); 3.16 (m, 12H); 3.26 (m, 4H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 26.2; 26.3; 26.4; 26.5; 27.4; 27.5; 28.4; 28.7; 28.8; 28.9; 29.0; 40.9; 50.9; 156.0; 157.6; 165.9. LCMS m/z (ES+) = 235.9 $[M + 3H]^{3+}$; 353.4 $[M + 2H]^{2+}$; 705.8 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-azidooctyl)-1,3-bis({8-[*N*'-(cyclopropylmethyl)carbamimidamido]octyl})urea trifluoroacetate salt (21): Starting material: **15**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 0.27 (q, J = 4.8 Hz, 4H); 0.59 (q, J = 4.8 Hz, 4H); 1.08 (m, 2H); 1.33 (m, 32H); 1.52 (t, J = 6.6 Hz, 8H); 1.59 (t, J = 6.5 Hz, 8H); 3.05 (d, J = 7.8 Hz, 4H); 3.14 (m, 12H); 3.27 (t, J = 8 Hz, 4H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 3.1; 8.9; 26.7; 27.1; 27.3, 27.4; 29.3; 29.5; 29.8; 31.6; 41.0; 45.3; 47.6; 159.3; 163.6. LCMS m/z (ES+) = 407.3 $[M + 2H]^{2+}$; 813.9 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis[8-(*N*'-propylcarbamimidamido)octyl]urea trifluoroacetate salt (29a): Starting material: **26a**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 0.98 (t, 6H, J = 7.2 Hz); 1.36 (m, 32H); 1.50–1.54 (m, 8H); 1.57–1.64 (m, 12H); 3.13–3.20 (m, 20H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 21.8; 26.2; 26.5; 27.5; 28.4; 28.8; 41.0; 41.0; 42.7; 156.0; 157.2; 165.8. LCMS m/z (ES+) = 274.0 $[M + 3H]^{3+}$; 411.6 $[M + 2H]^{2+}$; 822.2 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis({8-[*N*'-(propan-2-yl)carbamimidamido]octyl})urea trifluoroacetate salt (29b): Starting material: **26b**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.23 (d, 12H, J = 6.4 Hz); 1.30 (m, 32H); 11.50–1.55 (m, 8H); 1.54–1.57 (overlapped, m, 8H); 3.16 (m, 16H); 3.73 (m, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 21.2; 26.1; 26.2; 26.5; 27.5; 28.3; 28.4; 28.8; 28.9; 40.9; 41.0; 43.4; 46.8; 155.1; 157.4; 165.8. LCMS m/z (ES+) = 274.7 $[M + 3H]^{3+}$; 411.6 $[M + 2H]^{2+}$, 822.1 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis[8-(*N*'-butylcarbamimidamido)octyl]-1,3-bis(8-carbamimidamidooctyl)urea trifluoroacetate salt (29c): Starting material: **26c**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 0.97 (t, 6H, J = 7.2 Hz), 1.36–1.44 (m, 32H), 1.49–1.61 (24H), 3.13–3.21 (m, 20H).^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 22.8; 25.0; 26.2; 26.5; 26.8; 27.5; 28.4; 28.8; 29.0; 41.0; 43.0; 156.0; 157.2; 165.8. LCMS m/z (ES+) = 283.8 $[M + 3H]^{3+}$; 425.2 $[M + 2H]^{2+}$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis({8-[*N*'-(2-methylpropyl)carbamimidamido]octyl})urea trifluoroacetate salt (29d): Starting material: **26d**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): ^{1}H NMR (400 MHz, MeOD) δ = 0.97 (d, 12H, J = 6.4 Hz); 1.30–1.37 (m, 32H); 1.51–1.55 (m, 8H); 1.56–1.61 (m, 8H); 1.86 (m, 2H); 3.01 (d, 4H, J = 6.8 Hz); 3.13–3.21 (m, 16H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 18.5; 18.8; 26.2; 26.5; 27.5; 28.0; 28.4; 28.8; 41.0; 156.2; 157.3; 165.7. LCMS m/z (ES+) = 283.8 $[M + 3H]^{3+}$; 425.2 $[M + 2H]^{2+}$. Colorless oil. Yield: 99%.

1,3-bis(8-{*N*'-[(2E)-but-2-en-1-yl]carbamimidamido}octyl)-1,3-bis(8-carbamimidamidooctyl)urea trifluoroacetate salt (29e): Starting material: **26e**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.36 (m, 32H); 1.53 (m, 8H); 1.59 (m, 8H); 1.73 (d, 6H, J = 6.4 Hz); 3.13–3.20 (m, 16H); 3.76 (d, 4H, J = 5.6 Hz); 5.46–5.53 (m, 2H); 5.71–5.79 (m, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 16.3; 26.2; 26.6; 27.5; 28.4; 28.4; 28.9; 29.0; 41.0; 41.1; 42.5; 124.9; 128.6; 155.8; 157.5; 165.8. LCMS m/z (ES+) = 282.6 $[M + 3H]^{3+}$; 423.3 $[M + 2H]^{2+}$; 845.7 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis[8-(*N*'-cyclopentylcarbamimidamido)octyl]urea trifluoroacetate salt (29f): Starting material: **26f**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.30 (m, 32H); 1.52–1.60 (m, 16H); 1.65 (m, 8H); 1.76 (m, 4H); 1.97–2.03 (m, 4H); 3.17 (m, 16H); 3.86 (m, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 22.9; 26.2; 26.5; 27.5; 28.4; 28.5; 28.8; 31.9; 40.9; 45.8; 155.6; 157.2; 163.3; 165.8. LCMS m/z (ES+) = 292.0 $[M + 3H]^{3+}$; 437.6 $[M + 2H]^{2+}$, 873.2 $[M + H]^+$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis(8-{*N*'-[(2E)-3,7-

dimethylocta-2,6-dien-1-yl] carbamimidamido}octyl)urea trifluoroacetate salt (29g): Starting material: **26g**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.36 (m, 32H); 1.48–1.64 (m, 8H); 1.56–1.62 (overlapped s, m, 20H); 1.71 (s, 6H); 1.84 (m, 4H); 2.10 (m, 4H); 3.13–3.20 (m, 16H); 3.83 (d, 4H, $J = 6.8$ Hz); 5.08 (m, 2H); 5.28 (m, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 24.6; 26.2; 26.5; 27.5; 28.3; 28.4; 29.0; 38.9; 41.0; 118.5; 134.4; 137.5; 140.3; 156.1; 157.2; 161.3; 165.8. LCMS *m/z* (ES+) = 337.3 $[M + 3H]^{3+}$; 505.6 $[M + 2H]^{2+}$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis({8-[*N*'-(2-methoxyethyl)carbamimidamido] octyl})urea trifluoroacetate salt (29h): Starting material: **26h**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.35 (m, 32H); 1.50 (m, 8H); 1.54 (m, 8H); 3.16 (m, 16H); 3.36–3.39 (m, 4H); 3.39 (overlapped, s, 6H); 3.52 (t, 4H, $J = 5.2$ Hz). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 26.2; 26.6; 27.5; 28.3; 28.4; 28.8; 28.9; 40.9; 41.2; 41.4; 57.7; 78.7; 156.5; 1572; 165.8. LCMS *m/z* (ES+) = 285.4 $[M + 3H]^{3+}$; 427.5 $[M + 2H]^{2+}$; 876.2 $[M + Na]^{+}$. Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis(8-{*N*'-[2-(morpholin-4-yl)ethyl] carbamimidamido}octyl)urea trifluoroacetate salt (29i): Starting material: **26i**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.33 (m, 32H); 1.52 (t, 8H, $J = 7.3$ Hz); 1.58 (m, 8H); 3.13–3.17 (m, 12H); 3.15–3.21 (overlapped, m, 8H); 3.20–3.25 (overlapped, m, 8H); 3.63 (m, 4H); 3.92 (m, 8H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 27.1; 27.3; 29.3; 29.5; 41.1; 41.8; 43.7; 53.2; 54.9; 66.8; 157.2; 163.6. LCMS *m/z* (ES+) = 241.8 $[M + 4H]^{4+}$; 322.1 $[M + 3H]^{3+}$; 482.6 $[M + 2H]^{2+}$; 964.2 $[M + H]^{+}$Colorless oil. Yield: 99%.

1,3-bis(8-carbamimidamidooctyl)-1,3-bis(8-{*N*'-[2-(methylsulfanyl)ethyl] carbamimidamido}octyl)urea trifluoroacetate salt (29l): Starting material: **26l**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.33 (m, 32H); 1.48–1.54 (m, 8H); 1.53–1.59 (overlapped, m, 8H); 2.11 (s, 6H); 2.68 (m, 4H); 3.16 (m, 20H); 3.87 (m, 4H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 15.6; 27.1; 27.3; 27.4; 29.3; 29.5; 30.3; 41.0; 41.1; 44.8; 47.6; 157.3; 158.2; 164.0. LCMS *m/z* (ES+) = 222.3 $[M + 4H]^{4+}$; 296.1 $[M + 3H]^{3+}$; 443.4 $[M + 2H]^{2+}$; 885.7 $[M + H]^{+}$Colorless oil. Yield: 99%.

1,3-bis(8-{*N*'-[2-(adamantan-1-yl)ethyl] carbamimidamido} octyl)-1,3-bis(8-carbamimidamidooctyl)urea trifluoroacetate salt (29m): Starting material: **29m**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.33 (m, 48H); 1.54 (m, 24H); 1.68 (m, 8H); 1.76 (m, 8H); 3.16 (m, 20H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 27.1; 27.3; 27.9; 28.9; 29.3; 29.5; 36.1; 37.2; 37.4; 41.1; 42.2; 47.6; 157.3; 158.0; 162.1 LCMS *m/z* (ES+) = 266.6 $[M + 4H]^{4+}$; 354.7 $[M + 3H]^{3+}$; 531.5 $[M + 2H]^{2+}$ Colorless oil. Yield: 99%.

1,3-bis({8-[*N*'-(cyclopropylmethyl)carbamimidamido] octyl})-1,3-bis({8-[(methylcarbamoyl)amino] octyl})urea trifluoroacetate salt (30): Starting material: **27**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 0.29 (m, 4H); 0.59 (m, 4H); 1.07 (m, 2H); 1.33 (m, 32H); 1.53 (m, 12H); 1.60 (m, 4H); 2.69 (s, 6H); 3.07 (d, 2H, $J = 8.0$ Hz); 3.13–3.21 (m, 16H).^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 2.4; 9.5; 25.5; 26.2; 26.4; 27.5; 28.5; 28.9; 29.0; 29.9; 39.6; 41.1; 45.8; 157.2; 161.3; 165.1. LCMS *m/z* (ES+) = 205.1 $[M + 4H]^{4+}$;273.2 $[M + 3H]^{3+}$; 409.3 $[M + 2H]^{2+}$, 817.7 $[M + H]^{+}$. Colorless oil. Yield: 99%.

1,3-bis({8-[*N*'-(cyclopropylmethyl)carbamimidamido] octyl})-1,3-bis({8-[(methylcarbamothioyl)amino] octyl})urea trifluoroacetate salt (31): Starting material:**28**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 0.29 (m, 4H); 0.59 (m, 4H); 1.07 (m, 2H); 1.36 (m, 32H); 1.53 (m, 8H); 1.60 (m, 8H); 2.93 (s, 6H); 3.07 (d, 2H, $J = 7.2$ Hz); 3.15–3.21 (m, 16H).^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 2.5; 9.6; 25.5; 26.2; 26.4; 27.5; 28.5; 28.9; 29.0; 29.9; 31.5; 41.1; 45.8; 46.8; 157.2; 165.7; 182.8. LCMS *m/z* (ES+) = 454.5 $[M + 2H]^{2+}$, 907.9 $[M + H]^{+}$. Colorless oil. Yield: 99%.

{6-[*N*'-(cyclopropylmethyl)carbamimidamido] hexyl}urea trifluoroacetate salt (32): Starting material **61**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 0.27 (q, $J = 4.8$ Hz, 2H); 0.59 (q, J = , 2H); 1.06 (m, 1H); 1.31 (m, 4H); 1.50 (m, 2H); 1.60 (m, 2H); 3.05 (d, $J = 7.8$ Hz, 2H); 3.10 (t, $J = 7.2$ Hz, 2H); 3.19 (t, $J = 7.2$ Hz, 2H). ^{13}C NMR (150 Hz, CD_3OD) δ (ppm): 3.6, 10.2, 26.2, 26.6, 27.5, 28.4, 29.0, 40.3, 41.1, 46.8, 159.3, 162.3. LCMS *m/z* (ES+) = 256.1 $[M + H]^{+}$. Colorless oil. Yield: 99%.

[8-(carbamoylamino)octyl] urea trifluoroacetate salt (33): Starting material **62**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.36 (m, 8H); 1.48 (m, 2H); 1.58 (m, 2H); 3.08 (t, $J = 6.5$ Hz, 2H); 3.16 (t, $J = 6.6$ Hz, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 27.2, 27.3, 29.3, 29.5, 40.3, 41.1, 157.3, 163.3. LCMS *m/z* (ES+) = 230.2 $[M + H]^{+}$. Colorless oil. Yield: 99%.

[8-(*N*'-methylcarbamimidamido)octyl] urea trifluoroacetate salt (34): Starting material **63**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.36 (m, 8H); 1.47 (m, 2H); 1.58 (m, 2H); 2.84 (s, 3H); 3.05 (d, $J = 7.2$ Hz, 2H); 3.16 (t, $J = 7.2$ Hz, 2H); 3.19 (t, $J = 7.2$ Hz, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 27.2, 27.3, 29.3, 29.5, 38.7, 40.3, 41.1, 157.3, 163.3. LCMS *m/z* (ES+) = 244.4 $[M + H]^{+}$. Colorless oil. Yield: 99%.

[8-(*N*'-ethylcarbamimidamido)octyl] urea trifluoroacetate salt (35): Starting material **64**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.21 (t, $J = 7.2$ Hz, 3H); 1.36 (m, 8H); 1.48 (m, 2H); 1.59 (m, 2H); 3.06 (t, $J = 7.2$ Hz, 2H); 3.17 (t, $J = 7.2$ Hz, 2H); 3.23 (q, $J = 7.2$ Hz, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 14.6, 27.2, 27.3, 29.3, 29.5, 39.2, 40.3, 41.1, 157.3, 163.3. LCMS *m/z* (ES+) = 258.3 $[M + H]^{+}$. Colorless oil. Yield: 99%.

{8-[*N*'-(propan-2-yl)carbamimidamido] octyl}urea trifluoroacetate salt (36): Starting material **65**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.22 (d, $J = 6.6$ Hz, 6H); 1.36 (m, 8H); 1.48 (m, 2H); 1.59 (m, 2H); 3.08 (t, $J = 7.2$ Hz, 2H); 3.18 (t, $J = 7.2$ Hz, 2H); 3.72 (m, 1H). ^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 22.8, 27.2, 27.3, 29.3, 29.5, 40.3, 41.1, 47.1, 157.3, 163.3. LCMS *m/z* (ES+) = 72.1 $[M + H]^{+}$. Colorless oil. Yield: 99%.

(8-{*N*'-[(2E)-but-2-en-1-yl] carbamimidamido}octyl)urea trifluoroacetate salt (37): Starting material **66**. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm): 1.37 (m, 8H); 1.47 (m, 2H); 1.58 (m, 2H); 1.72 (d, $J = 5.8$ Hz, 2H); 3.08 (t, $J = 7.2$ Hz, 2H); 3.17 (t, $J = 7.2$ Hz, 2H); 3.74 (m, 2H); 5.49 (m, 1H); 5.73 (m, 1H).^{13}C NMR (150 MHz, CD_3OD) δ (ppm): 15.5, 27.2, 27.3, 29.3, 29.5, 40.3, 41.1, 44.7, 128.1, 128.4, 157.3, 163.3. LCMS *m/z* (ES+) = 84.2 $[M + H]^{+}$. Colorless oil. Yield: 99%.

4.2. In vitro antibacterial activity testing

Bacterial strains, including representatives of both Gram-positive and Gram-negative bacteria, were obtained from the ATCC or CCUG culture collections, except for *A. baumannii* N50, a clinical isolate showing carbapenem- and colistin-resistance isolate in Italy [34]. Compounds were resuspended in DMSO at a final concentration of 50 or 100 mg/mL and subsequently diluted in the culture medium. MIC and MBC values of the compounds were determined using the microdilution broth method using Mueller-Hinton II broth (Becton Dickinson, Le Pont de Claix, France) as recommended by the Clinical Laboratory Standards Institute (CLSI) and following standard procedures [26,35]. Bacterial inoculum was 5×10^5 CFU/mL and MIC was determined after 18–24 h of incubation at 35 ± 2 °C. Synergistic activity (as MAC) was determined as the MIC of the compound in the presence of a subinhibitory concentration ($0.5 \times$ MIC) of colistin (fixed concentration of 0.12 μg/mL for *E. coli* CCUG^T and *K. pneumoniae* ATCC 13833, 0.25 μg/mL for *A. baumannii* ATCC 17978 and *P. aeruginosa* ATCC 27853, or 2 μg/mL for the resistant *A. baumannii* N50 clinical isolate).

4.3. Molecular dynamics

The ligands (**29h**, **29m**, and **30**) were built with Maestro and then subjected to energy minimization performed with Macromodel until a convergence value of 0.05 kcal/Å•mol, by employing the conjugate gradient (CG) algorithm, MMFFs force field, and a distance-dependent dielectric constant of 1.0. The CHARMM-GUI Membrane Builder tool [36] was used for the generation of two different phospholipid bilayers with dimensions 80 × 80 Å, solvated with a 30 Å water layer on both

sides: a first one constituted by pure POPG molecules and a second one composed of a mixture of POPE/POPG 6:4 molecules. Six different systems were then created by complexing each compound with both solvated lipid bilayers. MD simulations were performed on the six systems using AMBER, version 20 [37]. General Amber force field (GAFF) parameters were used for the ligands, whose partial charges were assigned using the Antechamber suite of AMBER 20, based on the AM1-BCC method. Lipid17 force field was used for the parametrization of the phospholipids, while TIP3P explicit solvent model was used for the solvent. Sodium ions were added to achieve system neutrality, which is necessary to use Particle Mesh Ewald (PME) electrostatics in the MD studies. Nevertheless, the old Amberff99 parameters were used for sodium ions in order to prevent them from producing a condensation effect on the lipid bilayers due to strong interactions with the negatively charged lipid head groups [38]. The systems were energy minimized in two steps, both consisting of 5000 cycles of steepest descent (SD) followed by CG algorithm. In the first step, only the minimization of the solvent was performed, since a harmonic potential of 100 kcal/mol•Å^2 was applied on all solute atoms. In the second step, no position restraint was applied and the whole system was thus energy minimized. The minimized systems were used as input structures for the MD simulations, which were run as previously performed [15], following a protocol adapted from previous MD studies of membrane-containing systems [39,40].

All MD studies were run using PME electrostatics, with a cutoff of 10 Å for the non-bonded interactions and periodic boundary conditions. SHAKE algorithm was used to constrain all bonds involving hydrogen atoms and a time step of 2.0 fs was thus used for the simulation. Initially, the temperature of the systems was gradually raised from 0 to 300 K through a brief constant-volume MD simulation where a position restraint of 5 kcal/mol·Å^2 was applied to the phospholipid molecules. The systems were then relaxed through a 500 ps constant-pressure MD simulation in which the harmonic potential applied to the lipid molecules was gradually removed and Langevin thermostat was used to maintain the temperature at 300 K. Finally, 600 ns constant-pressure MD simulation production was performed using the Monte Carlo barostat with anisotropic pressure scaling. All analyses performed on the MD trajectories obtained as a result of the MD simulations were carried out using the Cpptraj program implemented in AMBER 20 [40].

4.4. Cheminformatics analysis

4.4.1. Data sets

The data used to develop the machine learning models included the whole AGUs library that we recently synthesized. For each of the 74 compounds of the series, we considered the corresponding biological activity, expressed as MIC value, against ten different bacterial strains. Specifically, 4 Gram-negative (*E. coli* CCUG^T, *K. pneumoniae* ATCC 13833, *A. baumannii* ATCC 17978, *P. aeruginosa* ATCC 27853) and 4 Gram-positive (*B. subtilis* ATCC 6633, *E. faecalis* ATCC 19433, *S. aureus* ATCC 25923, *S. pyogenes* ATCC 12344) strains were considered. The compounds were labeled into three classes based on their biological activity against Gram-positive and Gram-negative bacteria. Specifically, compounds with at least three out of five MIC values against Gram-positive bacteria $\leq$4 μg/mL were labeled as highly active, compounds with at least three out of five MIC values between 4 and 32 μg/mL were considered as partially active, whereas compounds that did not meet any of the two criteria were reported as inactive. The same classification was performed in terms of activity against Gram-negative bacteria. Based on this classification, the final dataset was composed of 46 highly active, 13 weakly active, and 15 inactive compounds for Gram-positive bacteria, whereas 6 highly active, 29 weakly active, and 39 inactive compounds for Gram-negatives.

4.4.2. Analysis of molecular properties

The compounds contained in the datasets were represented through 123 molecular descriptors. Molecular descriptors are experimentally measured or theoretically derived properties of a molecule. More specifically, they are quantitative representations of physical, chemical, or topological characteristics of molecules that describe the molecular structure from different aspects. In this work, we employed 123 RDKit [41] descriptors calculated with the built-in deepchem functions [42]. To perform a SAR analysis with respect to activities against Gram-positive and Gram-negative bacteria, two different machine learning models based on the Random Forest algorithm were generated. Random Forest is a machine learning model that uses several decision trees trained with the data of the training set. Each tree provides a single prediction, while the final output of the models corresponds to the most voted prediction by the individual trees. The models were subjected to a 10-fold cross-validation, and for each fold, the model was trained with 70% of the corresponding datasets, while the remaining 30% of the data was used as the test used to assess the model's performance. MCC was used to evaluate the model performance [43].

4.4.3. Feature importance

The impact of molecular descriptors on the models predictions was computed with the SHAP Python library [44]. SHAP concept allows us to estimate the importance of an individual molecular descriptor in the outputs provided by the machine learning model. Values of feature contributions present different magnitudes and different signs according to the impact on the model's predictions and the class to which they are referred [45]. Specifically, a positive sign identifies a contribution to the prediction of activity, while a negative sign results in a contribution to the prediction of inactivity. Since in this work we employed models based on decision trees, a recently developed algorithm for the exact calculation of SHAP values was applied [46].

4.5. Toxicity and PAINS predictions

SMILES of the whole set of compounds were obtained by Marvin 17.21.0, Chemaxon [47]. SMILES were then imported in MolBook UNIPI [31] and *in silico* toxicity evaluation was performed using VenomPred software [30]. *In silico* PAINS assessment was performed using the web tools SwissADME [32] and False Positive Remover [33], as well as MolBook UNIPI software.

CRediT authorship contribution statement

Claudia Ardino: Writing - original draft, Writing - Review & Editing, Validation, Methodology, Investigation, Formal Analysis, Data curation, Conceptualization.

Filomena Sannio: Formal analysis, Data curation.

Giulio Poli: Writing - original draft, Writing - Review & Editing, Software, Formal analysis, Data curation.

Salvatore Galati: Writing - original draft, Writing - Review & Editing, Software, Formal analysis, Data curation.

Elena Dreassi: Resources, Data curation.

Lorenzo Botta: Resources, Data curation.

Jean-Denis Docquier: Writing - original draft, Writing - Review & Editing, Validation, Methodology, Supervision, Formal analysis, Data curation, Conceptualization.

Ilaria D'Agostino: Writing - original draft, Writing - Review & Editing, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The authors declare the work is partially supported by Lead Discovery Siena s.r.l. They declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We thank Lead Discovery Siena s.r.l. for partial financial support. We also thank Prof. Elena Petricci for the discussion on the topic. This work was supported (to J.D.D.) in part by the Italian MUR (Ministero dell'Università e Ricerca) in the frame of the PNRR PE-13 ("Piano Nazionale di Ripresa e Resilienza, Partenariato Esteso 13, Malattie infettive emergenti") INF-ACT project (One Health Basic and Translational Research Actions addressing Unmet Needs on Emerging Infectious Diseases).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmech.2024.116362.

References

[1] World Health Organization (WHO), 2019 Antibacterial Agents in Clinical Development: an Analysis of the Antibacterial Clinical Development Pipeline, 2019.
[2] 2020 Antibacterial Agents in Clinical and Preclinical Development, 2020.
[3] E. Mossialos, C. Morel, S. Edwards, J. Berenson, M. Gemmill-Toyama, D. Brogan, Policies and Incentives for Promoting Innovation in Antibiotic Research, World Health Organization for the European Observatory on Health Systems and Policies, Geneva, 2010.
[4] A.R. Mahoney, M.M. Safaee, W.M. Wuest, A.L. Furst, The silent pandemic: emergent antibiotic resistances following the global response to SARS-CoV-2, iScience 24 (2021) 102304, https://doi.org/10.1016/j.isci.2021.102304.
[5] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (2020) 497–506, https://doi.org/10.1016/S0140-6736(20)30183-5.
[6] M. Ripa, L. Galli, A. Poli, C. Oltolini, V. Spagnuolo, A. Mastrangelo, C. Muccini, G. Monti, G. De Luca, G. Landoni, L. Dagna, M. Clementi, P. Rovere Querini, F. Ciceri, M. Tresoldi, A. Lazzarin, A. Zangrillo, P. Scarpellini, A. Castagna, A. Andolina, M.B. Redaelli, G. Bigai, A. Bigoloni, G. Borio, S. Bossolasco, E. Bruzzesi, M.G. Calabrò, S. Calvisi, C. Campochiaro, D. Canetti, V. Canti, J. Castellani, B. Castiglioni, G. Cavalli, L. Cavallo, M. Cernuschi, M. Chiurlo, M. Cilla, E. Cinel, P. Cinque, C. Conte, V. Da Prat, A. Danise, R. De Lorenzo, A. Dell'Acqua, R. Dell'Acqua, E. Della Torre, L. Della Torre, G. Di Terlizzi, I. Dumea, F. Farolfi, M. Ferrante, C. Frangi, L. Fumagalli, G. Gallina, B. Germinario, N. Gianotti, M. Guffanti, H. Hasson, F. Lalla, M. Lanzillotta, R. Li Voti, E. Messina, C. Molinari, E. Moizo, M. Montagna, G. Morsica, S. Nozza, M. Pascali, A. Patrizi, M. Pieri, A. Poloniato, D. Prestifilippo, G. Ramirez, M. Ranzenigo, J. Sapienza, F. Seghi, G. Tambussi, C. Tassan Din, S. Turi, C. Uberti-Foppa, C. Vinci, Secondary infections in patients hospitalized with COVID-19: incidence and predictive factors, Clin. Microbiol. Infection 27 (2021) 451–457, https://doi.org/10.1016/j.cmi.2020.10.021.
[7] P. Nori, K. Cowman, V. Chen, R. Bartash, W. Szymczak, T. Madaline, C. Punjabi Katiyar, R. Jain, M. Aldrich, G. Weston, P. Gialanella, M. Corpuz, I. Gendlina, Y. Guo, Bacterial and fungal coinfections in COVID-19 patients hospitalized during the New York City pandemic surge, Infect. Control Hosp. Epidemiol. 42 (2021) 84–88, https://doi.org/10.1017/ice.2020.368.
[8] A.A. Karbelkar, A.L. Furst, Electrochemical diagnostics for bacterial infectious diseases, ACS Infect. Dis. 6 (2020) 1567–1571, https://doi.org/10.1021/acsinfecdis.0c00342.
[9] R.L. Nation, J. Li, Colistin in the 21st century, Curr. Opin. Infect. Dis. 22 (2009) 535–543, https://doi.org/10.1097/QCO.0B013E328332E672.
[10] A. Corona, D. Cattaneo, Dosing colistin properly: let's save "our last resort old drug!,", Clin. Infect. Dis. 65 (2017) https://doi.org/10.1093/CID/CIX388, 870–870.
[11] C. Zamperini, G. MacCari, D. Deodato, C. Pasero, I. D'Agostino, F. Orofino, F. De Luca, E. Dreassi, J.D. Docquier, M. Botta, Identification, synthesis and biological activity of alkyl-guanidine oligomers as potent antibacterial agents, Sci. Rep. 7 (2017) 1–11, https://doi.org/10.1038/s41598-017-08749-6.
[12] M. Botta, G. Maccari, S. Sanfilippo, F. De Luca, J.-D. Docquier, D. Deodato, Linear guanidine derivatives, methods of preparation and uses thereof, WO2016/055644 A1, Int. Patent Appl. (2016 April 14).
[13] C. Ardino, F. Sannio, C. Pasero, L. Botta, E. Dreassi, J.D. Docquier, I. D'Agostino, The impact of counterions in biological activity: case study of antibacterial alkylguanidino ureas, Mol. Divers. 29 (2022) 1–11, https://doi.org/10.1007/S11030-022-10505-6.
[14] C. Pasero, I. D'Agostino, F. De Luca, C. Zamperini, D. Deodato, G.I. Truglio, F. Sannio, R. Del Prete, T. Ferraro, D. Visaggio, A. Mancini, M.B. Guglielmi, P. Visca, J.D. Docquier, M. Botta, Alkyl-guanidine compounds as potent broad-spectrum antibacterial agents: chemical library extension and biological characterization, J. Med. Chem. 61 (2018) 9162–9176, https://doi.org/10.1021/acs.jmedchem.8b00619.
[15] I. D'Agostino, C. Ardino, G. Poli, F. Sannio, M. Lucidi, F. Poggialini, D. Visaggio, E. Rango, S. Filippi, E. Petricci, P. Visca, L. Botta, J.-D.D. Docquier, E. Dreassi, Antibacterial alkylguanidino ureas: molecular simplification approach, searching for membrane-based MoA, Eur. J. Med. Chem. 231 (2022) 114158, https://doi.org/10.1016/J.EJMECH.2022.114158.
[16] L.J.I. Balestri, I. D'Agostino, E. Rango, C. Vagaggini, R. Marchitiello, M. Mariotti, A. Casian, D. Deodato, G.I. Truglio, F. Orofino, M. Sanguinetti, F. Bugli, L. Botta, E. Dreassi, Focused library of phenyl-fused macrocyclic amidinoureas as antifungal agents, Mol. Divers. 26 (2022) 3399–3409, https://doi.org/10.1007/S11030-022-10388-7.
[17] L.J.I. Balestri, C.I. Trivisani, F. Orofino, D. Fiorucci, G.I. Truglio, I. D'Agostino, F. Poggialini, L. Botta, J.-D. Docquier, E. Dreassi, Discovery and optimization of a novel macrocyclic amidinourea series active as acidic mammalian chitinase inhibitors, ACS Med. Chem. Lett. (2023), https://doi.org/10.1021/ACSMEDCHEMLETT.2C00472.
[18] D. Marquardt, B. Geier, G. Pabst, Asymmetric lipid membranes: towards more realistic model systems, Membranes 5 (2015) 180–196, https://doi.org/10.3390/MEMBRANES5020180, 5 (2015) 180–196.
[19] M.R.J. Salton, Structural and functional asymmetry of bacterial membranes, Membranes and Transport (1982) 277–282, https://doi.org/10.1007/978-1-4684-4082-9_34.
[20] J.G. Topliss, Utilization of operational schemes for analog synthesis in drug designt, J. Med. Chem. 15 (1972) 1006–1011, https://doi.org/10.1021/JM00280A002.
[21] G. Asghari, M. Jalali, E. Sadoughi, Antimicrobial activity and chemical composition of essential oil from the seeds of artemisia aucheri boiss, Jundishapur J. Nat. Pharm. Prod. 7 (2012) 11, https://doi.org/10.5812/jjnpp.3530.
[22] R. Bonikowski, P.ₙwitakowska, M.S. Sienkiewicz, M. Zakłos-Szyda, Selected compounds structurally related to acyclic sesquiterpenoids and their antibacterial and cytotoxic activity, Molecules 20 (2015) 11272, https://doi.org/10.3390/MOLECULES200611272.
[23] K. Sugamoto, Y.I. Matsusita, K. Matsui, C. Kurogi, T. Matsui, Synthesis and antibacterial activity of chalcones bearing prenyl or geranyl groups from Angelica keiskei, Tetrahedron 67 (2011) 5346–5359, https://doi.org/10.1016/J.TET.2011.04.104.
[24] C. Ghosh, J. Haldar, Membrane-active small molecules: designs inspired by antimicrobial peptides, ChemMedChem 10 (2015) 1606–1624, https://doi.org/10.1002/CMDC.201500299.
[25] G. Shahane, W. Ding, M. Palaiokostas, M. Orsi, Physical properties of model biological lipid bilayers: insights from all-atom molecular dynamics simulations, J. Mol. Model. 25 (2019), https://doi.org/10.1007/s00894-019-3964-0.
[26] F. Sannio, A. Brizzi, R. Del Prete, M. Avigliano, T. Simone, C. Pagli, T. Ferraro, F. De Luca, M. Paolino, F. Corelli, C. Mugnaini, J.D. Docquier, Optimization of pyrazole compounds as antibiotic adjuvants active against colistin- and carbapenem-resistant acinetobacter baumannii, Antibiotics 11 (2022) 1832, https://doi.org/10.3390/ANTIBIOTICS11121832, 11 (2022) 1832.
[27] WHO. https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed. (Accessed 19 October 2023).
[28] M. Zalacain, C. Lozano, A. Llanos, N. Sprynski, T. Valmont, C. de Piano, D. Davies, S. Leiris, C. Sable, A. Ledoux, I. Morrissey, M. Lemonnier, M. Everett, Novel specific metallo-β-lactamase inhibitor ANT2681 restores meropenem activity to clinically effective levels against NDM-positive enterobacterales, Antimicrob. Agents Chemother. 65 (2021), https://doi.org/10.1128/AAC.00203-21.
[29] R.J. Melander, C. Melander, The challenge of overcoming antibiotic resistance: an adjuvant approach? ACS Infect. Dis. 3 (2017) 559–563, https://doi.org/10.1021/ACSINFECDIS.7B00071.
[30] S. Galati, M. Di Stefano, E. Martinelli, M. Macchia, A. Martinelli, G. Poli, T. Tuccinardi, VenomPred: a machine learning based platform for molecular toxicity predictions, Int. J. Mol. Sci. 23 (2022) 2105, https://doi.org/10.3390/IJMS23042105.
[31] S. Galati, M. Di Stefano, M. Macchia, G. Poli, T. Tuccinardi, MolBook UNIPI—Create, manage, analyze, and share your chemical data for free, J. Chem. Inf. Model. 63 (2023) 3977–3982, https://doi.org/10.1021/ACS.JCIM.3C00278.
[32] A. Daina, O. Michielin, V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci. Rep. 7 (1) (2017) 1–13, https://doi.org/10.1038/srep42717, 7 (2017).
[33] J.B. Baell, G.A. Holloway, New substructure filters for removal of Pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays, J. Med. Chem. 53 (2010) 2719–2740, https://doi.org/10.1021/jm901137j.
[34] M.M. D'Andrea, T. Giani, S. D'Arezzo, A. Capone, N. Petrosillo, P. Visca, F. Luzzaro, G.M. Rossolini, Characterization of pABVA01, a plasmid encoding the OXA-24 carbapenemase from Italian isolates of acinetobacter baumannii, Antimicrob. Agents Chemother. 53 (2009) 3528, https://doi.org/10.1128/AAC.00178-09.
[35] Clinical and Laboratory Standards Institute, Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria that Grow Aerobically. Approved Standard, tenth

ed., CLSI Document M07-A10; Clinical and Laboratory Standards Institute, Wayne, PA, USA, 2015.
[36] Y. Li, J. Liu, J.C. Gumbart, Preparing membrane proteins for simulation using CHARMM-GUI, Methods Mol. Biol. 2302 (2021) 237–251, https://doi.org/10.1007/978-1-0716-1394-8_13.
[37] Case Ross, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, et al., Amber (2021).
[38] Å.A. Skjevik, B.D. Madej, C.J. Dickson, C. Lin, K. Teigen, R.C. Walker, I.R. Gould, Simulation of lipid bilayer self-assembly using all-atom lipid force fields, Phys. Chem. Chem. Phys. 18 (2016) 10573–10584, https://doi.org/10.1039/C5CP07379K.
[39] B.M. Bizzarri, L. Botta, D. Aversa, N.B. Mercuri, G. Poli, A. Barbieri, N. Berretta, R. Saladino, L-DOPA-quinone mediated recovery from GIRK channel firing inhibition in dopaminergic neurons, ACS Med. Chem. Lett. 10 (2019) 431–436, https://doi.org/10.1021/acsmedchemlett.8b00477.
[40] F. De Logu, S. Li Puma, L. Landini, T. Tuccinardi, G. Poli, D. Preti, G. De Siena, R. Patacchini, M.G. Tsagareli, P. Geppetti, R. Nassini, The acyl-glucuronide metabolite of ibuprofen has analgesic and anti-inflammatory effects via the TRPA1 channel, Pharmacol. Res. 142 (2019) 127–139, https://doi.org/10.1016/j.phrs.2019.02.019.
[41] RDKit: open-source cheminformatics. https://www.rdkit.org.
[42] B. Ramsundar, P. Eastman, P. Walters, V. Pande, O' Reilly), Deep Learning for the Life Sciences, 2019.
[43] M. Di Stefano, S. Galati, G. Ortore, I. Caligiuri, F. Rizzolio, C. Ceni, S. Bertini, G. Bononi, C. Granchi, M. Macchia, G. Poli, T. Tuccinardi, Machine learning-based virtual screening for the identification of Cdk5 inhibitors, Int. J. Mol. Sci. 23 (2022) 10653, https://doi.org/10.3390/IJMS231810653.
[44] S.M. Lundberg, P.G. Allen, S.-I. Lee, A unified approach to interpreting model predictions, part of, Advan. Neural Inform. Process. Syst. 30 (NIPS 2017) (2017).
[45] R. Rodríguez-Pérez, J. Bajorath, Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions, J. Comput. Aided Mol. Des. 34 (2020) 1013–1026, https://doi.org/10.1007/S10822-020-00314-0.
[46] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.I. Lee, From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. 2 (2020), https://doi.org/10.1038/s42256-019-0138-9.
[47] ChemAxon, https://www.chemaxon.com.

Abbreviations used

AMR: Antimicrobial Resistance
AGU: Alkylguanidinourea
ATCC: American Type Culture Collection
Boc: tert-Butyloxycarbonyl
CCUG: Culture Collection of University of Gothenburg
CFU: colony-forming unit
CLSI: Clinical Laboratory Standards Institute
COL: colistin
DAP: Daptomycin
DIPEA: N,N-diisopropylethylamine
DMF: N,N-dimethylformamide
DMSO: N,N-dimethyl sulfoxide
GA: Guanylating Agent
GAFF: General Amber Force Field
MAC: Minimal Antibacterial Concentration
MBC: minimal bactericidal concentration
MCC: Matthew's Correlation Coefficient
MD: Molecular Dynamics
MDR: Multi-Drug Resistance
MAC: minimal antibacterial activity
MIC: minimal inhibition concentration
NumHDonors: Number of H-bond Donors
PAINS: pan-assay interference compounds
PE: Petroleum Ether
PME: Particle Mesh Ewald
POPE: Palmitoyloleoyl phosphatidylethanolamine
POPG: Palmitoyloleoyl phosphatidylglycerol
SAR: Structure Activity Relationship
SD: Steepest Descent
SHAP: Shapley Additive exPlanations
TEA: N,N,N-triethylamine
TFA: trifluoroacetic acid
THF: tetrahydrofuran
TPSA: Topological Surface Area
Tr: trityl
VAN: vancomycin
WHO: World Health Organization