Computationally designing antimicrobial peptides for *Acinetobacter Baumannii*

Ayan Majumder,^{†,‡} Malay Ranjan Biswal,^{†,‡} and Meher K. Prakash^{*,†}

*Theoretical Science Unit, Jawaharlal Nehru Centre for Advanced Scientific Research,
Jakkur, Bengaluru, India.*

E-mail: meher@jncasr.ac.in

Abstract

Acinetobacter Baumannii, which is mostly contracted in hospital stays, has been developing resistance to all available antibiotics, including the last line of drugs, such as carbapenem. Because of its quick adaptation there is an immediate need to design new antibiotics, possibly antimicrobial peptides (AMPs) to which bacteria do not develop resistance easily. Our threefold goal was to curate the available activity of AMPs on the same strain of *A. Baumannii*, build a neural network model for predicting their activity and use it to rationally pre-screen for lead generation from the thousands of naturally occurring AMPs. By curating and analyzing the recent activity data from 81 AMPs on ATCC 19606 strain, we develop a quantitative AMP activity prediction model. We selected three other models with comparable performance against a test set with known activities. With the goal of inspiring further studies on AMP drug candidates and their rational shortlisting, we made activity predictions for the entire

*To whom correspondence should be addressed

[†]Theoretical Science Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bengaluru, India.

[‡]Contributed equally to this work

database of AMPs using all the models. To handle the uncertainty of training with a small data set, highlighted peptides which had consistent results from all models.

Acinetobacter Baumannii (*A. Baumannii*)¹ mainly implicated in hospital infections and is responsible for 80% of the infections caused by *Acinetobacter* species. *A. Baumannii* can also be found on normal human skin, but it generally does not pose threat towards a healthy person,²⁻⁵ besides the not so frequent skin and soft tissue infection, infection in the surgical site, urinary tract infection etc.^{6,7} In the past 30 years, *A. Baumannii* has evolved into a multidrug resistant (MDR)⁸⁻¹⁰ opportunistic pathogen that selectively infects seriously ill patients in intensive care unit (ICU), trauma or burn patients.^{2-4,11} The presence of intrinsic efflux pump and high rates of genetic adaptation, contributes to adaptation against the antibiotics.¹²⁻¹⁴ Besides, it also possesses several beta-lactamase genes which offer resistance against beta-lactam antibiotics.^{15,16} *A. Baumannii* has also been developing resistance against carbapenem¹⁷ which had been one of the last line of drugs against it. Combination therapies such as of colistin, polymixin B, and tigecycline are used to treat MDR strains, but these are complex compared to a single drug when it comes to quantification of the effect and the validation of their safety.¹⁸⁻²⁰ Due to the growing concern about MDR, there is a surge for new types of antimicrobial agents.

Antimicrobial peptides (AMP) are a fundamental part of the innate defense system and are reportedly present in organisms from bacteria and fungus to humans.^{21,22} Although several modes of AMP activity, including DNA damage,²³ RNA damage²⁴ and targeting a regulatory enzyme²⁵ or proteins²⁶ have been proposed, it is generally believed that the positively charged AMPs act by disrupting the bacterial membrane.²⁷⁻²⁹ Because of this fundamental difference in the mechanism compared to the traditional drugs, it is believed that the bacteria do not develop resistance easily against AMPs.²¹ The low toxicity of AMPs towards human cells and their low tendency to result in resistant strains makes them an ideal rational choice as the next generation antimicrobial agents,³⁰⁻³² possibly eventually becoming an effective drug for *A. Baumannii*.

Quantitative Structure and Activity Relationship (QSAR)³³ is an approach in computer aided rational drug design, which uses biophysical or biochemical parameters of the molecules to develop a quantitative relation with the measured activities. Once validated, the computational model can be used for predicting the activities of the possible drug candidates and for pre-screening them. Recent studies have developed QSAR relation using 29 small molecule drug candidates which act on the oxphos metabolic path of *A. Baumannii*.³⁴ As noted above, since with AMP based drugs, bacteria are less likely to develop resistance to, we focus on QSAR for AMPs against *A. Baumannii*.

Several experimental groups have independently evaluated the activity of AMP against *A. Baumannii*. The goal of the present work is threefold. We curated these experimental result against a single, well studied, target ATCC 19606 strain. We developed a computational model using neural networks to rationally predict the activity from the biochemical attributes of the AMP. Since *A. Baumannii* is a growing threat, while realizing the potential limitations of training on 81 peptides, we also predict the activity of all AMPs from the AMP database on the *A. Baumannii*.

The data on the activity of AMPs on *A. Baumannii* is scattered in literature. We curated the data mainly with the goal of developing a quantitative model, and hence restricted the focus to the most commonly studied strain. The activity data of AMPs on *A. Baumannii* ATCC 19606 strain was curated from different sources³⁵⁻⁴⁷ and is presented in **Supplementary table 1**. The curated AMP data set had the activity of 81 AMPs with length ranging from 10 to 43 amino acids and charges in the range +1 to +12. Of these, for 69 AMPs the MIC was available (referred to as exact MIC data), and for the remaining 12, only the lower bound of MIC (referred to as the lower bound data).

Overall, the comprehensive collection of the data on AMP activity allowed an analysis relative the various biophysical parameters (P_i) which are commonly used for developing quantitative relation with activities: (1) charge, which draws the AMPs selectively to anionic membrane,

(2) length, and how it is commensurate with the membrane thickness reflecting the activity^{48,49} (3) molecular weight, which gives an idea of the bulkiness and membrane penetration efficiency (4) hydrophobic moment (μ_H),⁵⁰ which quantifies the amphipathic characters required to form pores in the membrane, (5) aliphatic index,⁵¹ which indicates the aliphatic content of the peptide, (6) grand average of hydropathy (GRAVY) based on Kyte-Doolittle hydropathicity scale,⁵² (7) *in vivo* aggregation propensity, calculated by using a web-based software AGGRESCAN⁵³ and (8) *in vitro* (β -sheet) aggregation propensity, calculated by using TANGO software (with ionic strength 0.02 M, pH 7.0 and temperature 298 K).⁵⁴ The *in vitro* aggregation, before interaction with the membrane can at times stop proteolytic degradation⁵⁵ by the bacteria but in many other cases reduce the drug potency.^{56,57} Further, the aggregation propensity affects the Barrel-Stave⁵⁸ and carpet mechanisms⁴⁸ of action differently. The distribution of the eight parameters for all the curated AMPs are given in **Supplementary Figure 1** and their individual relation with MIC in **Figure 1**, which shows that each of the parameters individually is not sufficient to describe the activity.

Figure 1: MIC versus different parameters. The AMP used in the analysis is shown in Supplementary Table 1.

Artificial neural network (ANN) method is used to obtain the relationship between the various above-mentioned parameters and the peptide activity (MIC values). Since the available

data is limited, we did not discard the lower bound data in the quantitative model. We used 63 of the exact MIC values and 3 of the lower bound values, which were treated as the MIC for this analysis, for training (59 data points) and validation (7 data points). 6 data points with exact MIC were used for a quantitative test and the 9 remaining lower bound data points for an independent qualitative test. The activity of the AMPs was predicted by ANN model with an open module for machine learning called Scikit-learn⁵⁹ in Python. For the activation function, logistic function was used and low memory BFGS optimization algorithm was used as a solver. Two independent neural network calculations, one with a hidden layer of 6 neurons, another with 8 neurons were implemented for the quantitative activity prediction. 2500 trial runs in each case were made by taking 50 different random initializations for the input biases and 50 random choices for the training and validation sets. We screened the results of these 2500 trials with $R^2_{training} > 0.8$ and $R^2_{validation} > 0.6$, and identified 47 models with 6 neurons and 56 models with 8 neurons, to perform the test.

We created two independent test sets, one in which a quantitative MIC comparison was made and another quantitative one in which the calculated MIC was checked if it was more than the experimental lower bound. The selected models should simultaneously satisfy two conditions that at least 6 predictions in lower bound data set were correct and $R^2_{test} > 0.6$ for the exact data. 4 models satisfied these conditions, with $R^2_{test} > 0.85$ and they were selected. The best among these models (referred to as Model-1), had good predictions ($R^2_{training} = 0.954$, $R^2_{validation} = 0.877$ and $R^2_{test} = 0.874$). The experimental MIC for the quantitative data set versus MIC values predicted from Model-1 is shown in **Figure 2**. Results obtained from other three models are given in **Supplementary figure 2**.

It is important to know, which are the parameters (P_i) that are most responsible for the activity on *A. Baumannii*. In the combined training and validation set used for accepting the models, we replaced (P_i) with its average $\langle P_i \rangle$ and measure the difference $\Delta R^2_{P_i} = R^2_{training+validation} - R^2_{training+validation, \langle P_i \rangle}$. $\Delta R^2_{P_i}$ is treated as reflecting the importance of

Figure 2: Comparison of the experimental and calculated MIC ($\mu\text{g/mL}$) of curated AMPs on *A. Baumannii* obtained from Model-1. Training (purple circles), validation (orange squares) and test (green diamonds) sets are shown. The data used in the analysis is shown in Supplementary Table 1.

the parameter. The results obtained from Model-1 are given in **Figure 3** and those from the other three models are given in **Supplementary Figure 3**. From our calculations, we found out that the *in vitro* aggregation propensity in the most important parameter in Model-1.

Figure 3: The relative importance of the different parameters in Model-1 is shown. *In vitro* aggregation propensity influences the outcomes of the predictions the most.

Considering the health threat *A. Baumannii* is posing, and the potential of AMPs for antibiotic-resistance-free activity, we propose a rational basis for the design of AMPs for *A. Baumannii*. Our models were used to predict the MIC values of the 2338 AMPs ob-

tained from AMP database^{60–62} (<https://aps.unmc.edu>). In order to reduce the risk of a poorly trained ANN model with limited data, we filtered the predictions from all models for a consistent prediction that is within $\Delta\text{MIC} \leq 5\mu\text{g/ml}$ (**Supplementary table 3**).

To our knowledge ours is the only QSAR study for predicting AMP activity against *A. Baumannii*. The present work is different from the only the only other QSAR in two different ways, using AMPs in stead of small molecules for a better tolerance to antibiotic resistance and a slightly larger set (81 AMPs compared to 29 small molecules). Using the ANN models we developed, we could make quantitative predictions for the entire database of naturally occurring AMPs. We hope that our work will inspire the further studies quantifying the activity of AMPs on *A. Baumannii*, some of which may follow the activity predictions and others that differ offer an opportunity to retrain the ANN models.

References

- (1) Nemec, A.; Musilek, M.; Maixnerova, M.; De Baere, T.; van der Reijden, T. J.; Vanechoutte, M.; Dijkshoorn, L. *International journal of systematic and evolutionary microbiology* **2009**, *59*, 118–124.
- (2) Towner, K. *Journal of Hospital Infection* **2009**, *73*, 355–363.
- (3) Struelens, M.; Carlier, E.; Maes, N.; Serruys, E.; Quint, W. G.; Van Belkum, A. *Journal of Hospital infection* **1993**, *25*, 15–32.
- (4) Davis, K. A.; Moran, K. A.; McAllister, C. K.; Gray, P. J. *Emerging infectious diseases* **2005**, *11*, 1218.
- (5) Young, L. S.; Sabel, A. L.; Price, C. S. *Infection Control & Hospital Epidemiology* **2007**, *28*, 1247–1254.
- (6) Sievert, D. M.; Ricks, P.; Edwards, J. R.; Schneider, A.; Patel, J.; Srinivasan, A.;

- Kallen, A.; Limbago, B.; Fridkin, S. *Infection Control & Hospital Epidemiology* **2013**, *34*, 1–14.
- (7) Weiner, L. M.; Webb, A. K.; Limbago, B.; Dudeck, M. A.; Patel, J.; Kallen, A. J.; Edwards, J. R.; Sievert, D. M. *infection control & hospital epidemiology* **2016**, *37*, 1288–1301.
- (8) Gordon, N. C.; Wareham, D. W. *International journal of antimicrobial agents* **2010**, *35*, 219–226.
- (9) Falagas, M. E.; Karageorgopoulos, D. E. *Clinical infectious diseases* **2008**, *46*, 1121–1122.
- (10) Pachón, J.; Vila, J. *Current opinion in investigational drugs (London, England: 2000)* **2009**, *10*, 150–156.
- (11) Öncül, O.; Keskin, Ö.; Acar, H. V.; Küçükardalı, Y.; Evrenkaya, R.; Atasoyu, E. M.; Top, C.; Nalbant, S.; Özkan, S.; Emekdaş, G.; Çavuşlu, Ş.; Us, M. H.; Pahsa, A.; Gökben, M. *Journal of Hospital Infection* **2002**, *51*, 47–51.
- (12) Davies, J.; Davies, D. *Microbiology and molecular biology reviews* **2010**, *74*, 417–433.
- (13) Tan, S. Y.-Y.; Chua, S. L.; Liu, Y.; Høiby, N.; Andersen, L. P.; Givskov, M.; Song, Z.; Yang, L. *Genome biology and evolution* **2013**, *5*, 807–818.
- (14) Hood, M. I.; Mortensen, B. L.; Moore, J. L.; Zhang, Y.; Kehl-Fie, T. E.; Sugitani, N.; Chazin, W. J.; Caprioli, R. M.; Skaar, E. P. *PLoS pathogens* **2012**, *8*, e1003068.
- (15) Fournier, P.-E.; Vallenet, D.; Barbe, V.; Audic, S.; Ogata, H.; Poirel, L.; Richet, H.; Robert, C.; Mangenot, S.; Abergel, C.; Nordmann, P.; Weissenbach, J.; Raoult, D.; Claverie, J.-M. *PLoS genetics* **2006**, *2*, 1–11.
- (16) Sahl, J. W.; Gillece, J. D.; Schupp, J. M.; Waddell, V. G.; Driebe, E. M.; Engelthaler, D. M.; Keim, P. *PloS one* **2013**, *8*, e54287.

- (17) Nemec, A.; Krizova, L.; Maixnerova, M.; van der Reijden, T. J.; Deschaght, P.; Passet, V.; Vaneechoutte, M.; Brisse, S.; Dijkshoorn, L. *Research in microbiology* **2011**, *162*, 393–404.
- (18) Fan, B.; Guan, J.; Wang, X.; Cong, Y. *PloS one* **2016**, *11*, e0157757.
- (19) Kim, W.-Y.; Moon, J.-Y.; Huh, J. W.; Choi, S.-H.; Lim, C.-M.; Koh, Y.; Chong, Y. P.; Hong, S.-B. *PLoS One* **2016**, *11*, e0150642.
- (20) Rao, G. G.; Ly, N. S.; Diep, J.; Forrest, A.; Bulitta, J. B.; Holden, P. N.; Nation, R. L.; Li, J.; Tsuji, B. T. *International journal of antimicrobial agents* **2016**, *48*, 331–336.
- (21) Zasloff, M. *nature* **2002**, *415*, 389.
- (22) Hancock, R. E.; Sahl, H.-G. *Nature biotechnology* **2006**, *24*, 1551.
- (23) Ghosh, A.; Kar, R. K.; Jana, J.; Saha, A.; Jana, B.; Krishnamoorthy, J.; Kumar, D.; Ghosh, S.; Chatterjee, S.; Bhunia, A. *ChemMedChem* **2014**, *9*, 2052–2058.
- (24) Hao, G.; Shi, Y.-H.; Tang, Y.-L.; Le, G.-W. *Journal of microbiology* **2013**, *51*, 200–206.
- (25) Dangel, A.; Ackermann, N.; Abdel-Hadi, O.; Maier, R.; Önder, K.; Francois, P.; Müller, C. W.; Pané-Farré, J.; Engelmann, S.; Schrenzel, J.; Jürgen, H.; Christian, L. *The FASEB Journal* **2013**, *27*, 4476–4488.
- (26) Nguyen, L. T.; Haney, E. F.; Vogel, H. J. *Trends in biotechnology* **2011**, *29*, 464–472.
- (27) Yeaman, M. R.; Yount, N. Y. *Pharmacological reviews* **2003**, *55*, 27–55.
- (28) Brogden, K. A. *Nature reviews microbiology* **2005**, *3*, 238.
- (29) Lohner, K.; Blondelle, S. *Combinatorial chemistry & high throughput screening* **2005**, *8*, 241–256.

- (30) Oyston, P.; Fox, M.; Richards, S.; Clark, G. *Journal of medical microbiology* **2009**, *58*, 977–987.
- (31) Poole, K. *Current opinion in investigational drugs (London, England: 2000)* **2003**, *4*, 128–139.
- (32) Guaní-Guerra, E.; Santos-Mendoza, T.; Lugo-Reyes, S. O.; Terán, L. M. *Clinical Immunology* **2010**, *135*, 1–11.
- (33) Jenssen, H.; Fjell, C. D.; Cherkasov, A.; Hancock, R. E. *Journal of peptide science: an official publication of the European Peptide Society* **2008**, *14*, 110–114.
- (34) Sayiner, H. S.; Abdalrahm, A. A.; Basaran, M. A.; Kovalishyn, V.; Kandemirli, F. *Medicinal Chemistry* **2018**, *14*, 253–268.
- (35) Vila-Farres, X.; De La Maria, C. G.; López-Rojas, R.; Pachón, J.; Giralt, E.; Vila, J. *Clinical Microbiology and Infection* **2012**, *18*, 383–387.
- (36) Peng, S.-Y.; You, R.-I.; Lai, M.-J.; Lin, N.-T.; Chen, L.-K.; Chang, K.-C. *Scientific reports* **2017**, *7*, 11477.
- (37) Jayamani, E.; Rajamuthiah, R.; Larkins-Ford, J.; Fuchs, B. B.; Conery, A. L.; Vilcin-skas, A.; Ausubel, F. M.; Mylonakis, E. *Antimicrobial agents and chemotherapy* **2015**, AAC–04198.
- (38) Jaśkiewicz, M.; Neubauer, D.; Kazor, K.; Bartoszewska, S.; Kamysz, W. *Probiotics and antimicrobial proteins* **2018**, 1–8.
- (39) Mohamed, M. F.; Brezden, A.; Mohammad, H.; Chmielewski, J.; Seleem, M. N. *Scientific Reports* **2017**, *7*, 6953.
- (40) Andrä, J.; Monreal, D.; de Tejada, G. M.; Olak, C.; Brezesinski, G.; Gomez, S. S.; Goldmann, T.; Bartels, R.; Brandenburg, K.; Moriyon, I. *Journal of Biological Chemistry* **2007**,

- (41) Kohn, E. M.; Shirley, D. J.; Arotsky, L.; Picciano, A. M.; Ridgway, Z.; Urban, M. W.; Carone, B. R.; Caputo, G. A. *Molecules* **2018**, *23*, 329.
- (42) Li, D.; Yang, Y.; Tian, Z.; Lv, J.; Sun, F.; Wang, Q.; Liu, Y.; Xia, P. *Oncotarget* **2017**, *8*, 55958.
- (43) Saugar, J. M.; Rodríguez-Hernández, M. J.; Beatriz, G.; Pachón-Ibañez, M. E.; Fernández-Reyes, M.; Andreu, D.; Pachón, J.; Rivas, L. *Antimicrobial agents and chemotherapy* **2006**, *50*, 1251–1256.
- (44) Almaaytah, A.; YaâËŽu Ajingi, A. A.; Tarazi, S.; AlsharâËŽi, N.; Al-Balas, Q. *Infection and drug resistance* **2017**, *10*, 1.
- (45) Thomsen, T. T. Peptide Antibiotics for ESKAPE Pathogens: Past, Present and Future Perspectives of Antimicrobial Peptides for the Treatment of Serious Gram-Negative and Gram-Positive Infections. Ph.D. thesis, Department of Biology, Faculty of Science, University of Copenhagen, 2016.
- (46) Papo, N.; Oren, Z.; Pag, U.; Sahl, H.-G.; Shai, Y. *Journal of Biological Chemistry* **2002**, *277*, 33913–33921.
- (47) Almaaytah, A.; Qaoud, M. T.; Abualhaijaa, A.; Al-Balas, Q.; Alzoubi, K. H. *Infection and drug resistance* **2018**, *11*, 835.
- (48) Shai, Y. *Peptide Science: Original Research on Biomolecules* **2002**, *66*, 236–248.
- (49) Deslouches, B.; Phadke, S. M.; Lazarevic, V.; Cascio, M.; Islam, K.; Montelaro, R. C.; Mietzner, T. A. *Antimicrobial agents and chemotherapy* **2005**, *49*, 316–322.
- (50) Gautier, R.; Douguet, D.; Antonny, B.; Drin, G. *Bioinformatics* **2008**, *24*, 2101–2102.
- (51) Ikai, A. *The Journal of Biochemistry* **1980**, *88*, 1895–1898.
- (52) Kyte, J.; Doolittle, R. F. *Journal of molecular biology* **1982**, *157*, 105–132.

- (53) Conchillo-Solé, O.; de Groot, N. S.; Avilés, F. X.; Vendrell, J.; Daura, X.; Ventura, S. *BMC bioinformatics* **2007**, *8*, 65.
- (54) Fernandez-Escamilla, A.-M.; Rousseau, F.; Schymkowitz, J.; Serrano, L. *Nature biotechnology* **2004**, *22*, 1302.
- (55) Oren, Z.; Lerman, J. C.; Gudmundsson, G. H.; Agerberth, B.; Shai, Y. *Biochemical Journal* **1999**, *341*, 501.
- (56) Kustanovich, I.; Shalev, D. E.; Mikhlin, M.; Gaidukov, L.; Mor, A. *Journal of Biological Chemistry* **2002**, *277*, 16941–16951.
- (57) Feder, R.; Dagan, A.; Mor, A. *Journal of Biological Chemistry* **2000**, *275*, 4230–4238.
- (58) Ehrenstein, G.; Lecar, H. *Quarterly reviews of biophysics* **1977**, *10*, 1–34.
- (59) Pedregosa, F. et al. *Journal of machine learning research* **2011**, *12*, 2825–2830.
- (60) Wang, Z.; Wang, G. *Nucleic acids research* **2004**, *32*, D590–D592.
- (61) Wang, G.; Li, X.; Wang, Z. *Nucleic acids research* **2008**, *37*, D933–D937.
- (62) Wang, G.; Li, X.; Wang, Z. *Nucleic acids research* **2015**, *44*, D1087–D1093.

Computationally designing antimicrobial peptides for *Acinetobacter Baumannii*

Supplementary Information

Ayan Majumder,^{†,‡} Malay Ranjan Biswal,^{†,‡} and Meher K. Prakash^{*,†}

*Theoretical Science Unit, Jawaharlal Nehru Centre for Advanced Scientific Research,
Jakkur, Bengaluru, India.*

E-mail: meher@jncasr.ac.in

*To whom correspondence should be addressed

[†]Theoretical Science Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bengaluru, India.

[‡]Contributed equally to this work

Supplementary Table 1

Supplementary Table 1: Details of all the 81 antimicrobial peptides used in the present work.

Sl. No	Name	Sequence	MIC(μ g/ml)	References
1	Bactenecin	RLCRIVVIRVCR	64	1
2	Cecropin A	KWKLFKKIEKVGQNIRDGIKAGPAVAVVGQATQIAK	32	1
3	Cecropin B	KWKIFKKIEKVGGRNIRNGIAGPAVAVLGEAKAL	32	1
4	Cecropin P1	SWLSKTAKKLENSAKKRISSEGIATIAIQGGPR	1.6	1
5	Histatin 8	KFHEKHHSHRGY	32	1
6	HNP-1	ACYCRIPACIAGERRYGTCTIYQGRWAFCC	50	1
7	HNP-2	CYCRIPACIAGERRYGTCTIYQGRWAFCC	50	1
8	Indolicidin	ILPWKPWWPWR	8	1
9	Magainin I	GIGKFLHSAGKFGKAFVGEIMKS	64	1
10	Magainin II	GIGKFLHSAGKFGKAFVGEIMNS	256	1
11	Mastoparan	INLKALAALAKKIL	4	1
12	Melittin	GIGAVLKVLTTGLPALISWIKRKRQQ	4	1
13	β -Defensin	DHYNVSSGGQCLYSACPIFTKIQTGTCYRGKAKCKK	256	1
14	LysAB2 P0	NPEKALEPLIAIQIAIKGMLNGWFTGVGFRRKR	237.66	2
15	LysAB2 P1	EKALEKLIATQKAIKGMNGWFTGVGFRRKR	28.48	2
16	LysAB2 P2	EKALEKLIATQKAIKGMLAGWFTGVGARRKR	55.06	2
17	LysAB2 P3	NPEKALEKLIATQKAIKGMNGWFTGVGFRRKR	30.12	2
18	LL-37	LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES	4	3
19	Aurein 1.2	GLFDIICKIAESF -NH ₂	16	3
20	CAMEL	KWKLFKKIGAVLKVL -NH ₂	2	3
21	Citropin 1.1	GLFDVIKKVASVIGGL -NH ₂	16	3
22	Omiganan	ILRWPWWPWRK -NH ₂	32	3
23	r-Omiganan	KRRWPWWPWRLI -NH ₂	16	3

Sl. No	Name	Sequence	MIC(μ g/ml)	References
24	Pexiganan	GIGKFLKKAKKFGKAFVKILKK -NH ₂	2	3
25	Temporin A	FLPLIGRVLSGIL -NH ₂	128	3
26	rr	WLRRIKAWLRR	24.86	4
27	rr2	WIRRIKKWIRRVHK	3.95	4
28	rr3	WLRRIKAWLRRKRK	31.46	4
29	rr4	WLRRIKAWLRRIKA	3.73	4
30	CLP-19	CRKPTFRRLKWKIKFKFKC	80.36	5
31	C18G	ALWKKLLKLLKSAAKLG -NH ₂	3.87	6
32	C18G-Arg	ALWRLLRLRLRSARRLG -NH ₂	8.48	6
33	NK-2	KILRGVCKKIMRTFLRRISKDILTGGK -NH ₂	2	7
34	NK27	KILRGVSKKIMRTFLRRISKDILTGGK -NH ₂	2	7
35	N17	KILRGVSKKIMRTFLRR -NH ₂	8	7
36	C20	KKIMRTFLRRISKDILTGGK -NH ₂	8	7
37	C20-DK	KKIMRTFLRRISKILTGGK -NH ₂	8	7
38	NK23a	KISKKIMRTFLRRISKDILTGGK -NH ₂	4	7
39	NK23b	KILRGVSKKIMRRISKDILTGGK -NH ₂	8	7
40	NK22b	KILGVSKKIMRRISKDILTGGK -NH ₂	16	7
41	NK23c	KILRGVSKKIMRTFLRRILTGGK -NH ₂	4	7
42	NK19a	KISKKIMRTFLRRILTGGK -NH ₂	4	7
43	NK19b	KILRGVSKKIMRRILTGGK -NH ₂	2	7
44	NK19b-KR	RILRGVSRIMRRILTGR -NH ₂	8	7
45	NK13	KISKKIMRTFLRR -NH ₂	256	7
46	CA(1-8)M(1-18)	KWKLFFKKIGIGAVLKVLTTGLPALIS -NH ₂	4	8
47	CA(1-7)M(2-9)	KWKLFFKKIGAVLKVL -NH ₂	2	8
48	CA(1-7)M(5-9)	KWKLKKVVKVL -NH ₂	4	8
49	Pepcon	FLFSLIPSAIGGLISAFK	37.62	9

Sl. No	Name	Sequence	MIC(μ g/ml)	References
50	BP100	KKLFKKILKYL -NH ₂	8.5	10
51	RW-BP10	RRLFRRLRLWL -NH ₂	8.5	10
52	BP202	KRLFRKILKYL -NH ₂	4.5	10
53	BP203	KKLFKKILRYL -NH ₂	8.5	10
54	H4	KFKKLFKKLSPVIGKEFKRIVERIKRFLR	36.34	11
55	BR001	KWKLFKKIEKVGQNIRDGIKAGPAVAVVGATQIAK -NH ₂	10	12
56	BR002	GWLKKIGKKIERVGGHTRDATIQGLGIAQAANVAATAR -NH ₂	10	12
57	BR003	GGLKKLGKKLEGAGKRVFNAAEKALPVVAGAKALRK	5	12
58	BR005	RGFRKHFNKLKKVKHTISETAHVAKDTAVIAGSGAAVVAAT -NH ₂	20	12
59	BR029	KWKIFKKIEKAGRNIRDGIKAGPAVSVVGEAATIIYKTG	20	12
60	BR031	GWLDFGKRIERVGGHTRDATIQIGVAQAANVAATVRG	20	12
61	BR032	GWLKKIGKKIERVGGHTRDATIQVLGVAQAANVGPATARG	20	12
62	BR033	GWLKKIGKKIERVGGHTRDATIQTIGVAQAANVAATLKG	20	12
63	BR034	GWLKKFGKKIERVGGHTRDATIQIGVAQAANVAATLKG	20	12
64	BR035	GWLKKIGKKIERVGGHTRDASIQAIGIAQAANVAATARG	10	12
65	BR036	GWLKKIGKKIERVGGHTRDATIQVLGVAQAANVAATARG	20	12
66	BR037	GLVKKIGKKIERVGGHTRDASIQAIGIAQAANVAATARG	20	12
67	Amphipathic-1L	LKLLKKLLKLLKLL -NH ₂	1.8	13
68	Scrambled-4L	KLKLLKLLKLLKLL -NH ₂	3.61	13
69	Segregated-5L	KKKLLLLLLLLKKK -NH ₂	90.22	13
70	Buforin I	AGRKGQGGKVRKAKTRSSRAGLQFPVGRVHLLRKGNV	>256	1
71	Histatin 5	DSHAKRHHGYKRFHEKHHSHRGY	>256	1
72	rr1	WKRRIKIWKIR	>438.06	4
73	C18G-His	ALWHLLHLLHSAHHLG -NH ₂	>31.93	6
74	I10	KKIMRTFLRR -NH ₂	>128	7
75	NK15	KILRGVSKRILTGGK -NH ₂	>128	7

Sl. No	Name	Sequence	MIC(μ g/ml)	References
76	NK14	KILGVSKRILTGKK -NH ₂	>256	7
77	NK11	KISKRILTGKK -NH ₂	>256	7
78	NK10	ISKRILTGKK -NH ₂	>256	7
79	BR030	KWKFYKKIERVGGNIRDGIIKAGPAVQVVGQPPRYIKENRFYS	>20	12
80	BR043	AGFRKRFNKLVKVKVHTIKETANVSKDVAIVAGSGVAVGAAMG	>20	12
81	BR044	GFRKRFNKLVKVKVHTIKETANVSKDVAIVAGSGVAVGAAMG	>20	12

Supplementary Table 2

Supplementary Table 2: Details of the 12 lower-bound AMP data, these 9 were used for a qualitative test. Before accepting a model, at least 6 of these 9 results were verified to be satisfying the condition and then a quantitative test condition is used with the test data set from the exact data.

Name	Expt. MIC	Model-1	Model-2	Model-3	Model-4
Histatin 5	>256.000000	32.097823	31.692110	22.089586	23.036572
C18G-His	>31.927500	239.101949	99.836883	67.044164	45.931436
I10	>128.000000	395.654143	284.503188	436.228521	407.402294
NK15	>128.000000	6.742059	11.875128	8.123219	13.202811
NK14	>256.000000	30.195290	11.933689	8.123217	13.202904
NK10	>256.000000	267.735769	239.177329	433.381930	276.532836
BR030	>20.000000	18.974162	275.746286	379.442897	20.284478
BR043	>20.000000	73.831330	23.773969	19.715427	20.283674
BR044	>20.000000	131.697049	23.772869	19.715427	20.283674

Supplementary Figure 1: Histogram of all the parameters. The data used in the analysis is shown in Supplementary Table 1.

Supplementary Figure 2: Comparison of the experimental and calculated MIC ($\mu\text{g/mL}$) of curated AMPs on *A. Baumannii* obtained from (a) Model-2, (b) Model-3, (c) Model-4. Training (purple circles), validation (orange squares) and test (green diamonds) sets are shown. The data used in the analysis is shown in Supplementary Table 1.

Supplementary Figure 3: The relative importance of the different parameters is shown for (a) Model-2, (b) Model-3, (c) Model-4.

Supplementary Table 3

Supplementary Table 3: Using the 4 different models, we predicted the activity of 2338 naturally occurring AMPs documented in the AMP database. Of these the AMPs which had consistent predictions from any two models ($\Delta\text{MIC} \leq 5 \mu\text{g/ml}$) were selected and presented in this table.

Name	Sequence	Length	Model-1	Model-2	Model-3	Model-4
AP01465	VNWKKVLGKIIKVAK	15	12.11	11.98	8.14	12.78
AP02058	GLWNTIKEAGKKFAINVLDKIRCGIAGGCKT	31	19.24	20.63	20.54	20.28
AP01834	GILDTFKNMALNAKSAGVSVLNALSCKLSKTC	33	23.23	20.65	19.72	23.41
AP00113	GLLSGLKKVGVKHAVKNVAVSLMDSLCKISGDC	33	18.97	23.21	19.72	20.28
AP00822	GIFNVFKGALKTAGKHVAGSLNQLKCKVSGEC	33	19.00	21.49	19.72	20.32
AP01927	GLFNVFKVGVKNVLKNVAGSLMDNLCKVSGEC	33	18.97	22.93	19.72	20.28
AP00092	SLFSLIKAGAKFLGKNLLKQGAQYAACKVSKEC	33	18.97	23.18	19.72	20.31
AP00680	GLFGRLRDSLQGGGQKILEKAERIWCXIKDIFR	33	18.98	23.77	19.72	18.92
AP00085	SLFSLIKAGAKFLGKNLLKQGACYAACKASKQC	33	18.97	23.74	19.72	20.28
AP00228	LTCEIDRSLCLLHCRLLKGYLRAYCSQQKVCRCVQ	34	18.97	23.22	19.72	20.32
AP02212	GWLKKIGKKIERVGGHTRGLGIAQIAANVAATAR	34	18.97	22.91	19.72	20.86
AP00368	GLFRRLRDSIRRGQQKILEKARRIGERIKDIFRG	34	18.97	23.52	19.72	19.51
AP00892	PDITKLNKIKLTATCKVISKGASMCKVLFDKKKQE	36	18.97	23.64	19.72	19.43
AP00577	GLFTLIKGAALIGKTVAKEAGKTGLELMACKITNQC	37	18.97	23.63	19.72	20.28
AP00733	LLGDFFRKAREKIGEEFKRIVQRIKDFLRNLVPTES	37	18.95	19.21	19.72	20.25
AP01253	GLMSLFKGVLTAGKHIFKNVGGSLLDQACKITGEC	37	18.97	23.24	20.48	20.22
AP01254	GLMSLFRGVLTAGKHIFKNVGGSLLDQACKITGEC	37	18.97	23.04	20.41	20.17
AP01420	GIFSLIKGAAQLIGKTVAKEAGKTGLELMACKVTKQC	37	18.97	23.67	19.72	20.28
AP01705	GIFSLIKGAAKLITKTVAKEAGKTGLELMACKVTNQC	37	18.97	23.72	19.72	20.28
AP01716	GVFSFLTKGAKLLGSTLLKMAGKAGAEHLACKATNQC	37	19.02	22.06	19.72	20.28
AP01870	GVFTLIKATQLIGKTLGKEVGKTGLELMACKITKQC	37	19.09	23.51	21.64	20.28

Name	Sequence	Length	Model-1	Model-2	Model-3	Model-4
AP01873	GVFTLIKATQLIGKTLGKELGKTGLEIMACKITKQC	37	23.56	23.62	19.92	20.28
AP00736	RLGDILQKAREKIEGGLKLVQKIKDFGKFPARTES	37	16.98	20.55	19.72	19.82
AP01867	GLFTLIKGAALKIGKTTAKEAGKTGKLEMACKITNQC	37	18.98	20.06	19.72	20.28
AP01930	GIFSLIKTAAKFVGKNNLKQAGKAGVEHLACKANNQC	37	18.97	18.55	19.72	20.29
AP01159	KWKIFKKIEHMQNIRDGLIKAGPAVQVVGQAATIIYKG	38	18.97	23.27	19.72	20.28
AP00839	KVPVGAIKKGGKAIKTGLGVGAAGTAHEVYSHIRNRH	38	18.97	22.99	19.72	20.26
AP00135	GWLKIGKKIERVQGHTRDATIQTIGVAQQAANVAATLK	39	18.97	19.68	19.72	20.28
AP00231	GWLKIGKKIERVQGHTRDATIQVIGVAQQAANVAATAR	39	18.97	18.51	19.72	20.28
AP00232	GWLRKIGKKIERVQGHTRDATIQVLGIAQQAANVAATAR	39	18.97	20.26	19.72	20.28
AP01563	GWLRKLGGKIERIGQHTRDASIQLGIAQQAANVAATAR	39	18.97	21.46	19.72	20.28
AP02031	KPFKKLEKVGRNIRDGIIKAGPAVAVIGQATSIARPTGK	39	18.97	23.42	19.72	20.28
AP00266	GKREKCLRRNGFCAFLKCPTLSVISGTCSRFQVCKTLLG	40	18.95	23.77	19.72	20.28
AP00301	GILDTIKSIASKVWNSKTVQDLKRKGINWVANKLGVSPQAA	41	18.97	22.70	19.73	20.28
AP01796	ASFPWSCPSLSGVCRKVCLPTELFPGPLGCGKGLCGVSHFL	42	19.19	24.13	22.36	20.28
AP01323	LPVNEAQRQVGGYCGLRICNFPSRFLGLCTRNPCCSRVWV	42	18.97	23.79	19.72	19.39

References

- (1) Vila-Farres, X.; De La Maria, C. G.; López-Rojas, R.; Pachón, J.; Giralt, E.; Vila, J. *Clinical Microbiology and Infection* **2012**, *18*, 383–387.
- (2) Peng, S.-Y.; You, R.-I.; Lai, M.-J.; Lin, N.-T.; Chen, L.-K.; Chang, K.-C. *Scientific reports* **2017**, *7*, 11477.
- (3) Jaśkiewicz, M.; Neubauer, D.; Kazor, K.; Bartoszewska, S.; Kamysz, W. *Probiotics and antimicrobial proteins* **2018**, 1–8.
- (4) Mohamed, M. F.; Brezden, A.; Mohammad, H.; Chmielewski, J.; Seleem, M. N. *Scientific Reports* **2017**, *7*, 6953.

- (5) Li, D.; Yang, Y.; Tian, Z.; Lv, J.; Sun, F.; Wang, Q.; Liu, Y.; Xia, P. *Oncotarget* **2017**, *8*, 55958.
- (6) Kohn, E. M.; Shirley, D. J.; Arotsky, L.; Picciano, A. M.; Ridgway, Z.; Urban, M. W.; Carone, B. R.; Caputo, G. A. *Molecules* **2018**, *23*, 329.
- (7) Andrä, J.; Monreal, D.; de Tejada, G. M.; Olak, C.; Brezesinski, G.; Gomez, S. S.; Goldmann, T.; Bartels, R.; Brandenburg, K.; Moriyon, I. *Journal of Biological Chemistry* **2007**,
- (8) Saugar, J. M.; Rodríguez-Hernández, M. J.; Beatriz, G.; Pachón-Ibañez, M. E.; Fernández-Reyes, M.; Andreu, D.; Pachón, J.; Rivas, L. *Antimicrobial agents and chemotherapy* **2006**, *50*, 1251–1256.
- (9) Almaaytah, A.; Yau Ajingi, A. A.; Tarazi, S.; Alshari, N.; Al-Balas, Q. *Infection and drug resistance* **2017**, *10*, 1.
- (10) Thomsen, T. Peptide Antibiotics for ESKAPE Pathogens: Past, Present and Future Perspectives of Antimicrobial Peptides for the Treatment of Serious Gram-Negative and Gram-Positive Infections. Ph.D. thesis, 2016.
- (11) Almaaytah, A.; Qaoud, M. T.; Abualhaijaa, A.; Al-Balas, Q.; Alzoubi, K. H. *Infection and drug resistance* **2018**, *11*, 835.
- (12) Jayamani, E.; Rajamuthiah, R.; Larkins-Ford, J.; Fuchs, B. B.; Conery, A. L.; Vilcin-skas, A.; Ausubel, F. M.; Mylonakis, E. *Antimicrobial agents and chemotherapy* **2015**, AAC-04198.
- (13) Papo, N.; Oren, Z.; Pag, U.; Sahl, H.-G.; Shai, Y. *Journal of Biological Chemistry* **2002**, *277*, 33913–33921.