Electronic Supplementary Information

Machine Learning Approach-based Array Sensor for Rapidly Predicting Mechanisms of Action for Antibacterial Compounds

Zhijun Li,[‡, a] Kun Jin,[‡, a] Hong Chen,[‡, b, c] Liyuan Zhang,[*, d, e] Guitao Zhang,[a] Yizhou Jiang,[a] Haixia Zou,[a] Wentao Wang,[a] Guangpei Qi,[a] Xiangmeng Qu[*,a]

‡ These authors contributed equally to this work.

[a]Key Laboratory of Sensing Technology and Biomedical Instruments of Guangdong Province and School of Biomedical Engineering, Sun Yat-Sen University, Shenzhen, 518107, China
[b]Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, Xiamen 361005, China
[c]Jiujiang Research Institute of Xiamen University, Jiujiang 332000, China
[d]Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[e]School of Petroleum Engineering, State Key Laboratory of Heavy Oil Processing, China University of Petroleum (East China), Qingdao, 266580, China

*Corresponding author:
Xiangmeng Qu, PhD
quxm5@mail.sysu.edu.cn
Liyuan Zhang, PhD
liyuanzhang@seas.harvard.edu

Table S1. The set of feature variables.

Confusion matrix	Characteristic variables
I	A20-GO
II	A20-GO+T20-GO
III	A20-GO+T20-GO+C20-GO
IV	A20-GO+T20-GO+C20-GO+A20-MoS_2
V	A20-GO+T20-GO+C20-GO+A20-MoS_2+T20-MoS_2
VI	A20-GO+T20-GO+C20-GO+A20-MoS_2+T20-MoS_2+C20-MoS_2
VII	A20-GO+T20-GO+C20-GO+A20-MoS_2+T20-MoS_2+C20-MoS_2+A20-WS_2
VIII	A20-GO+T20-GO+C20-GO+A20-MoS_2+T20-MoS_2+C20-MoS_2+A20-WS_2+T20-WS_2
IX	A20-GO+T20-GO+C20-GO+A20-MoS_2+T20-MoS_2+C20-MoS_2+A20-WS_2+T20-WS_2+A20-WS_2

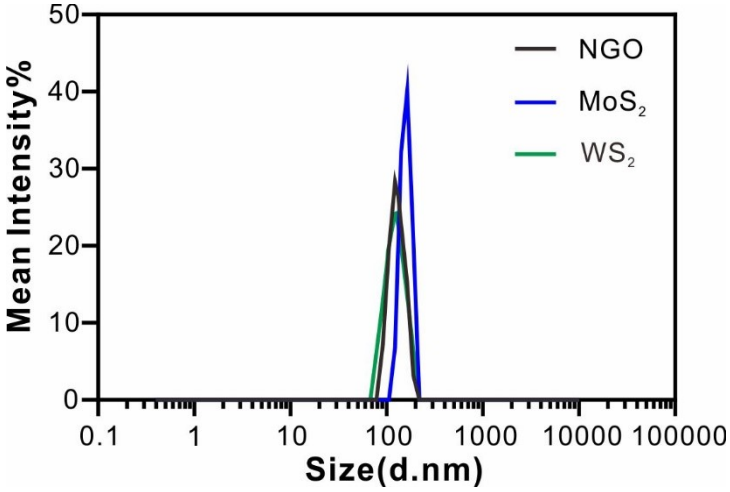


Fig. S1 DLS characterization of NGO, MoS2 and WS2, respectively.

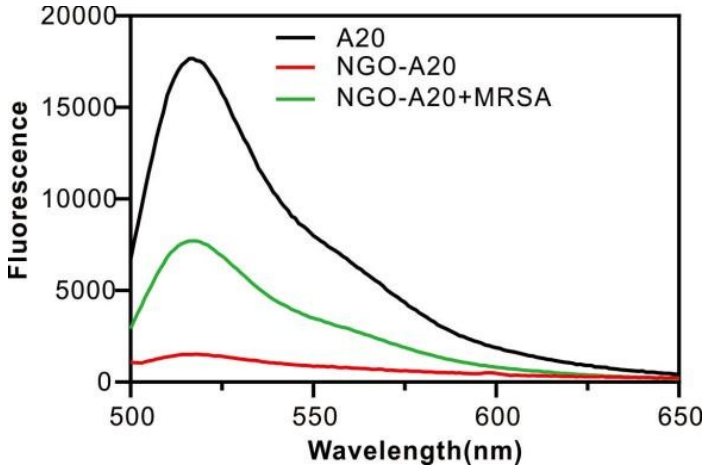


Fig. S2 Fluorescence signal of free ssDNA reporter (10 nM, A20, labeled as black curve, A20) binding with NGO conjugates (40 μg/mL, labeled as red curve, NGO-A20), and in the presence of MRSA (6×108 CFU/mL, green curve, labeled as NGO-A20+MRSA).

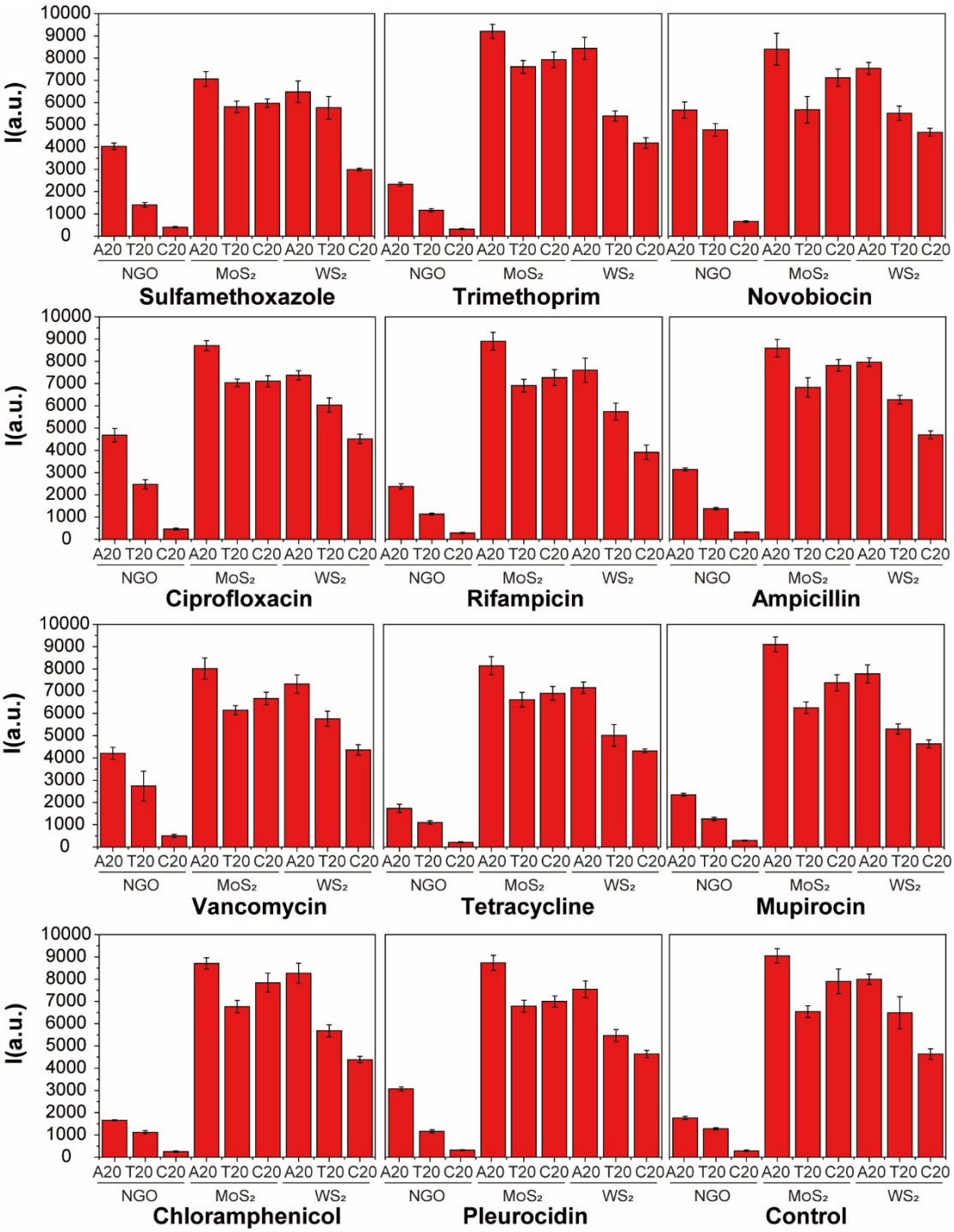


Fig. S3 Fluorescence intensity of the relative (I (a.u)) of antibacterial compounds after adding bacterial solution in the training cohort.

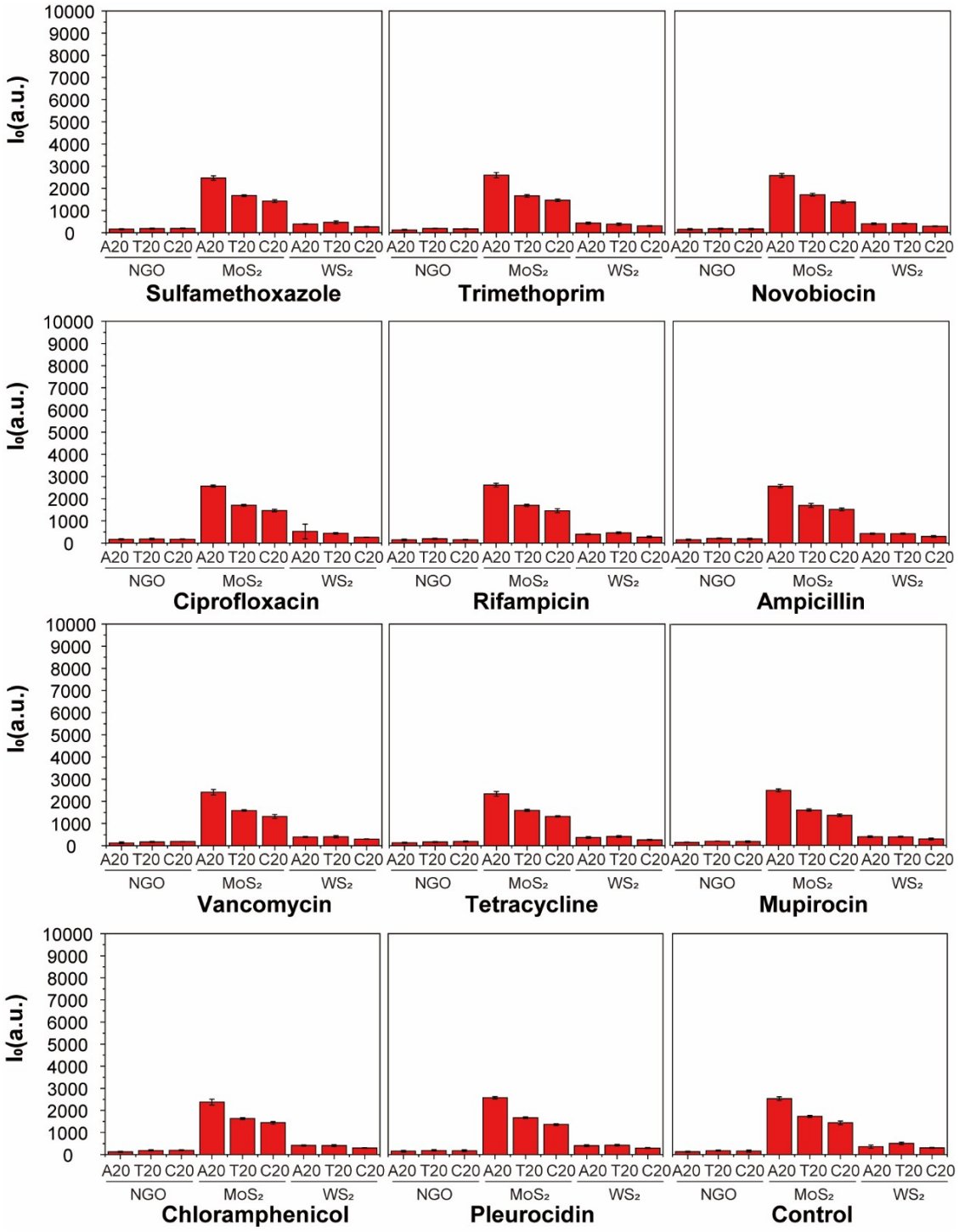


Fig. S4 Fluorescence intensity of the initial (I_0 (a.u)) of sensing elements.

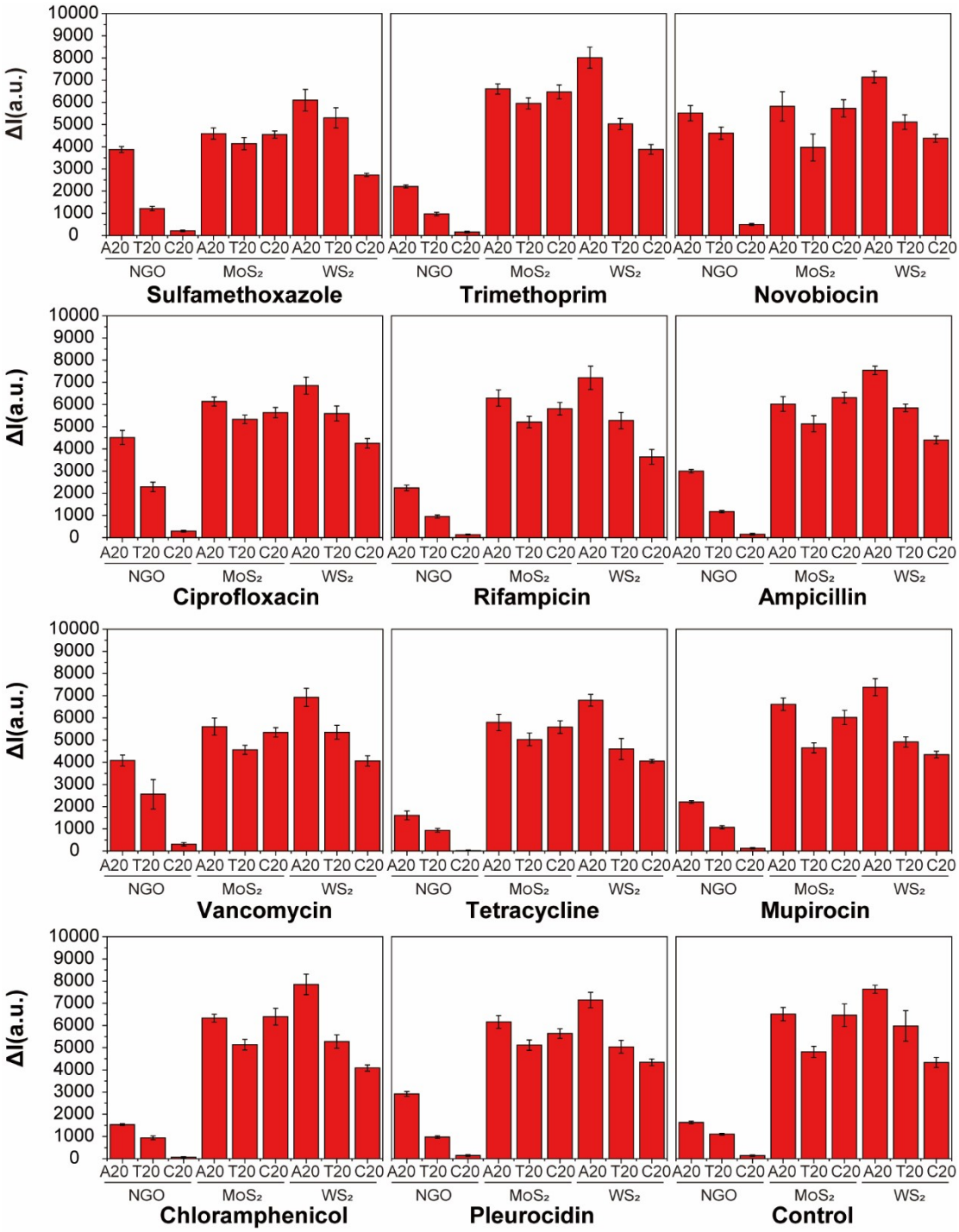


Fig. S5 The absolute fluorescence response (ΔI (a.u) =I-I_0) of antibacterial compounds after adding bacterial solution in the training cohort.

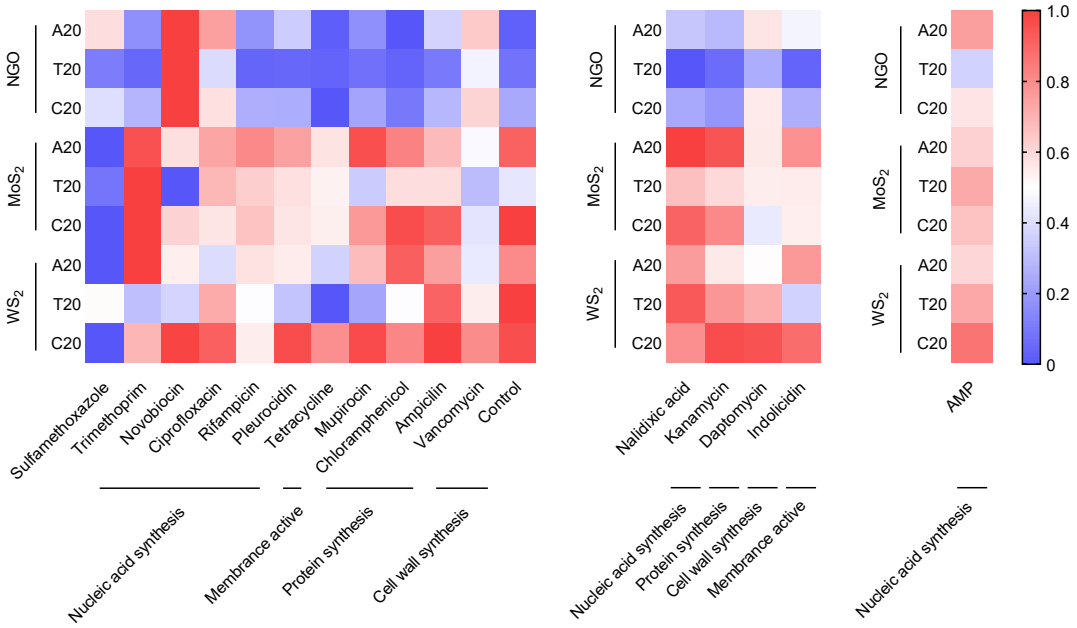


Fig. S6 A relative heat map for each sensing element to different antibacterial compounds. Each sensor element sensing data of different antibacterial compounds are normalized separately.

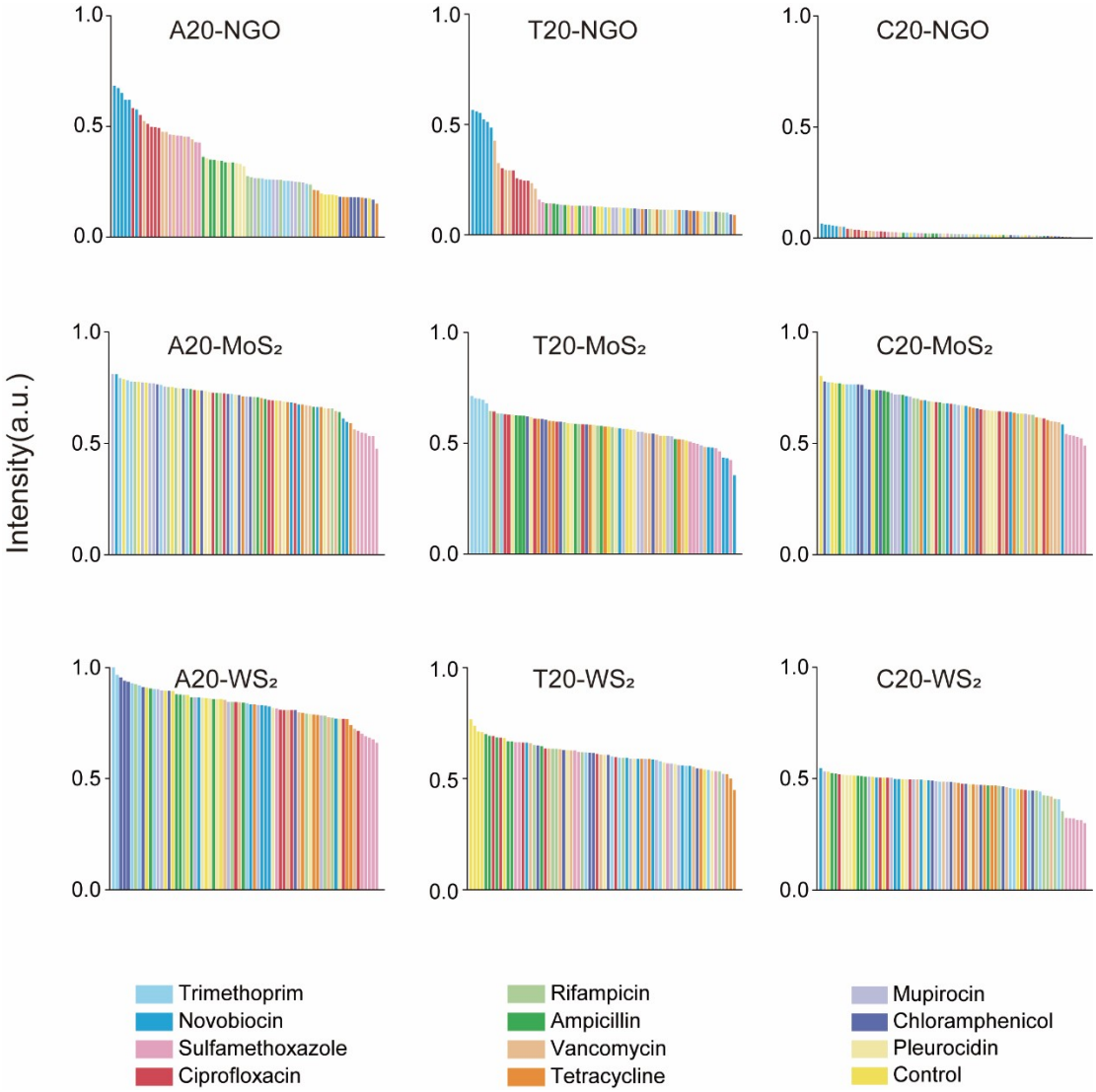


Fig. S7 Fluorescence response levels of all nine sensing elements (A20-NGO, T20-NGO, C20-NGO, A20-MoS_2, T20-MoS_2, C20-MoS_2, A20-WS_2, T20-WS_2, and C20-WS_2) of reference antibacterial compounds in the training cohort. The signal level of the antibacterial compounds sorted from high to low.

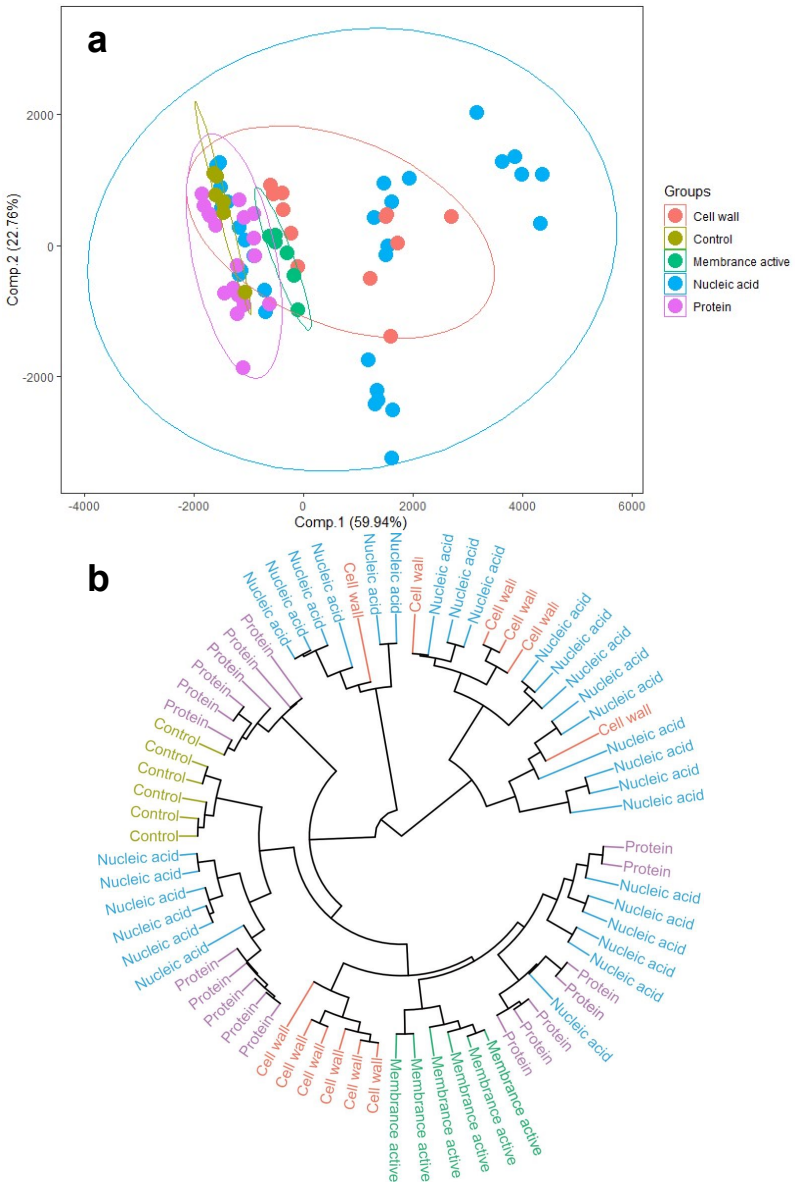


Fig. S8 Determination MOA using the sensor array combine different algorithms. a) PCA algorithm, b) HCA algorithm.

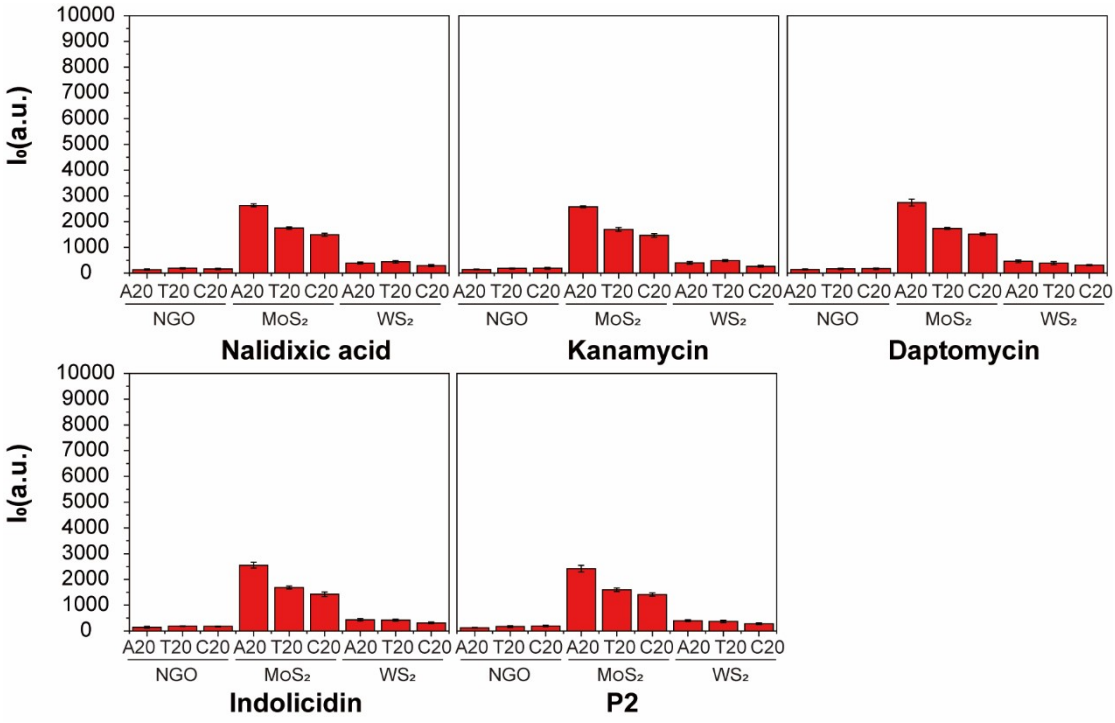


Fig. S9 Fluorescence intensity of the initial (I_0 (a.u.)) of antibacterial compounds before adding bacterial solution in the validation cohort.

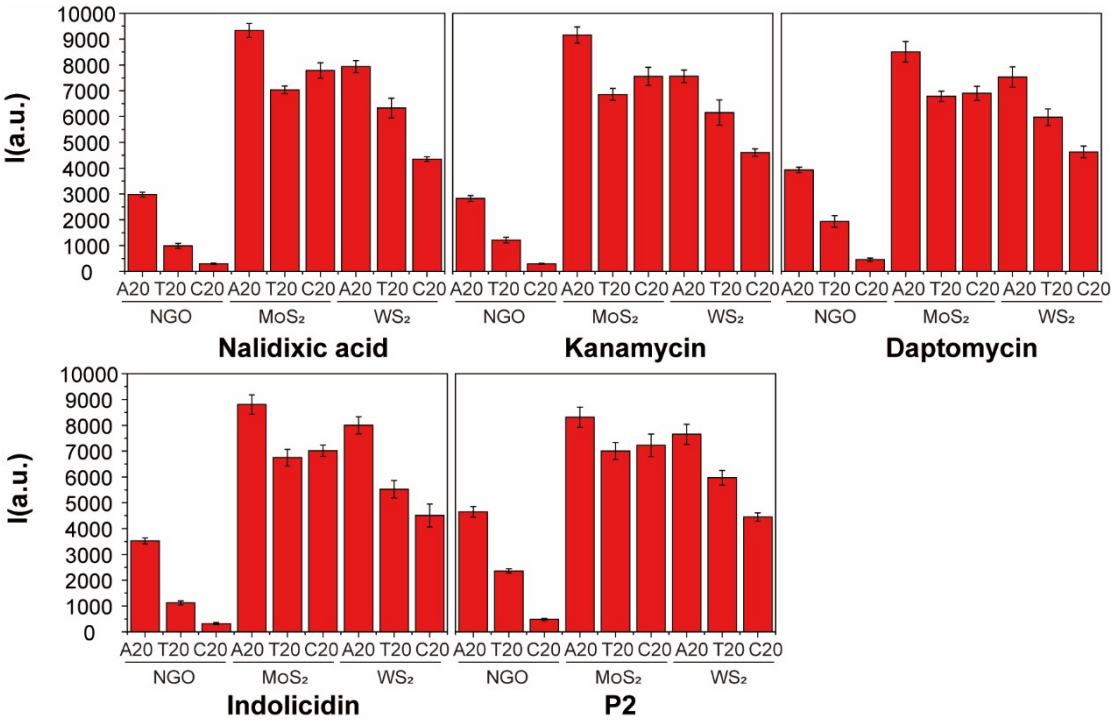


Fig. S10 Fluorescence intensity of the relative (I (a.u)) of antibacterial compounds after adding bacterial solution in the validation cohort.

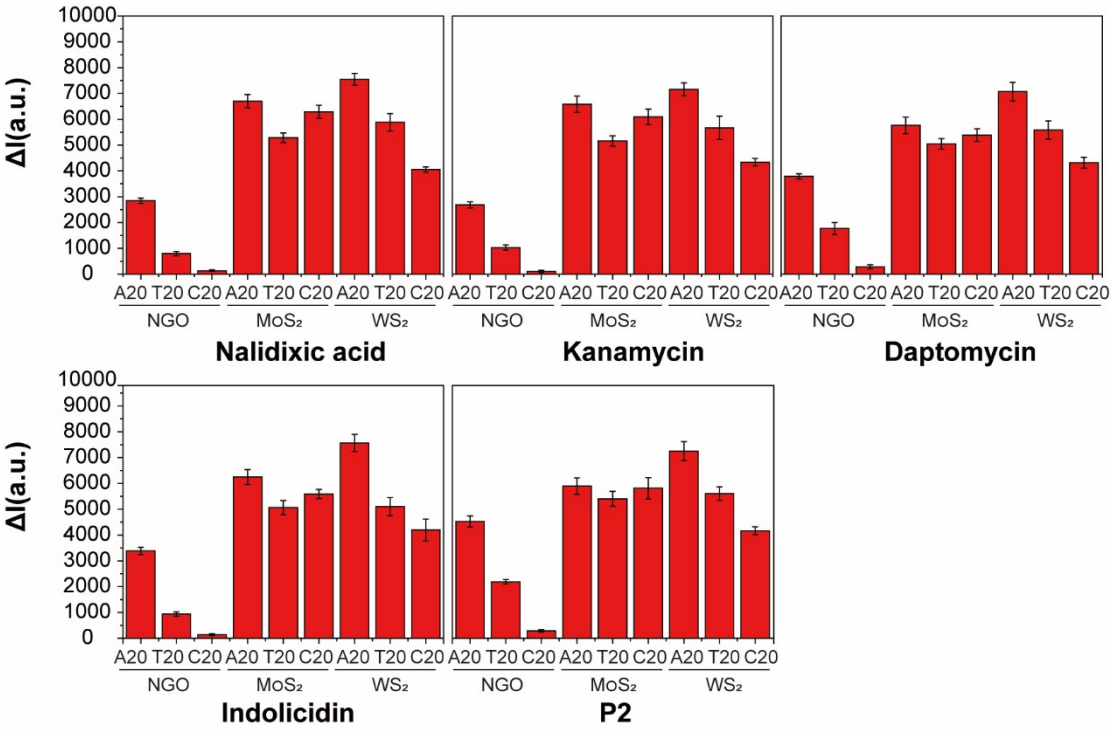


Fig. S11 The absolute fluorescence response (ΔI (a.u) =I-I_0) of antibacterial compounds after adding bacterial solution in the validation cohort.

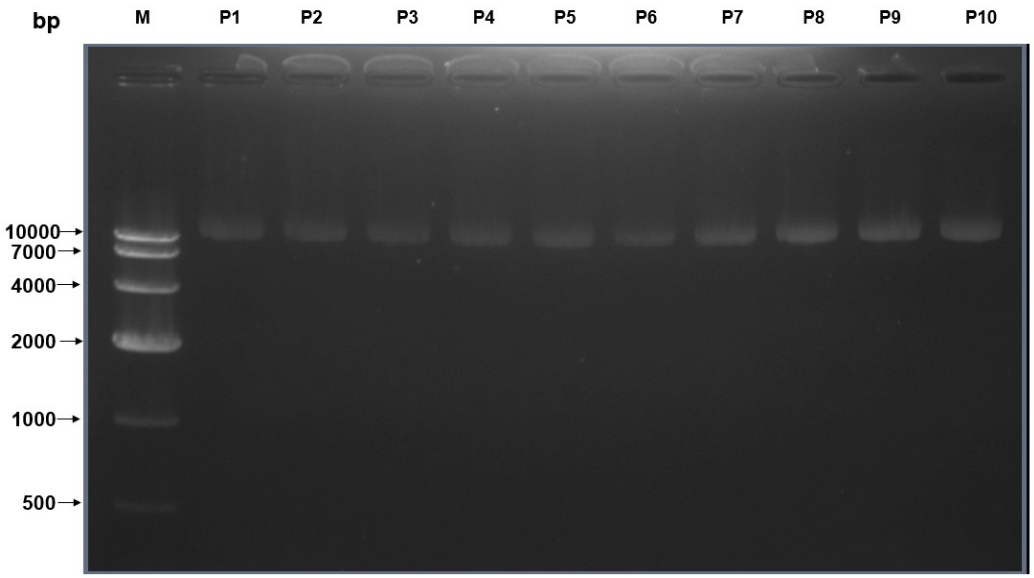


Fig. S12 Electrophoresis of phage single stranded DNA in 1% agarose.

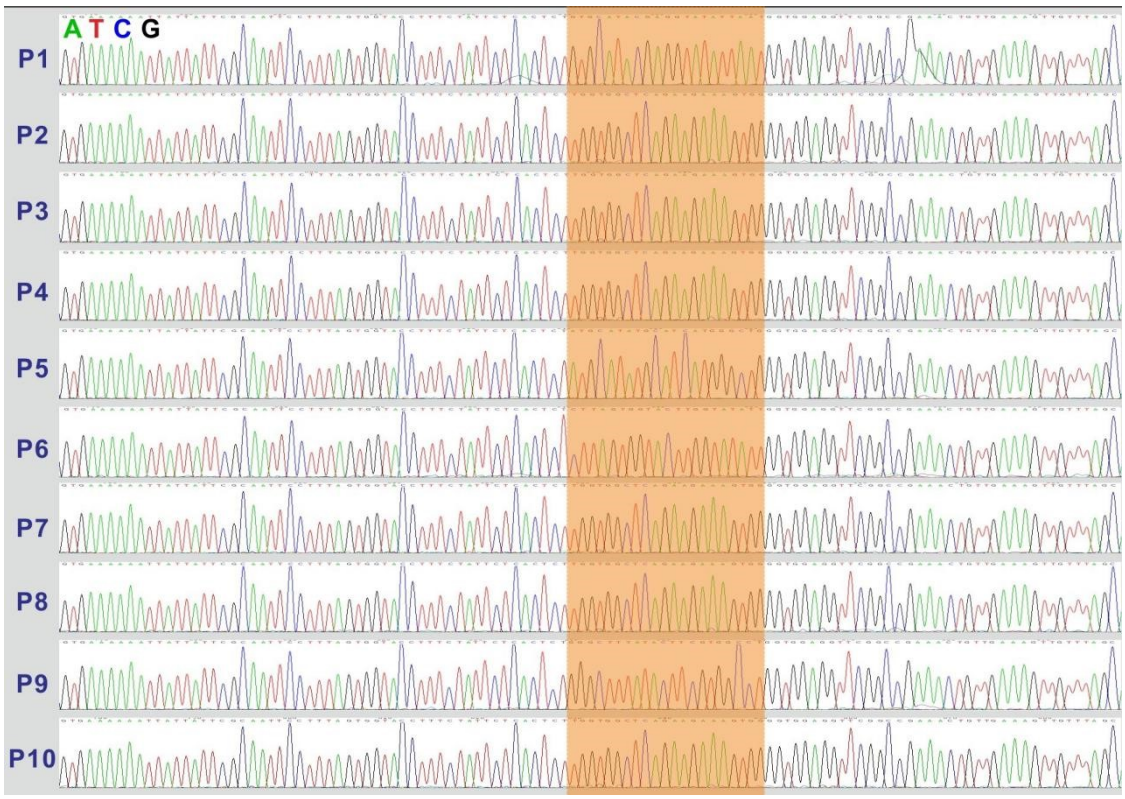


Fig. S13 The DNA sequencing results of P1-P10.

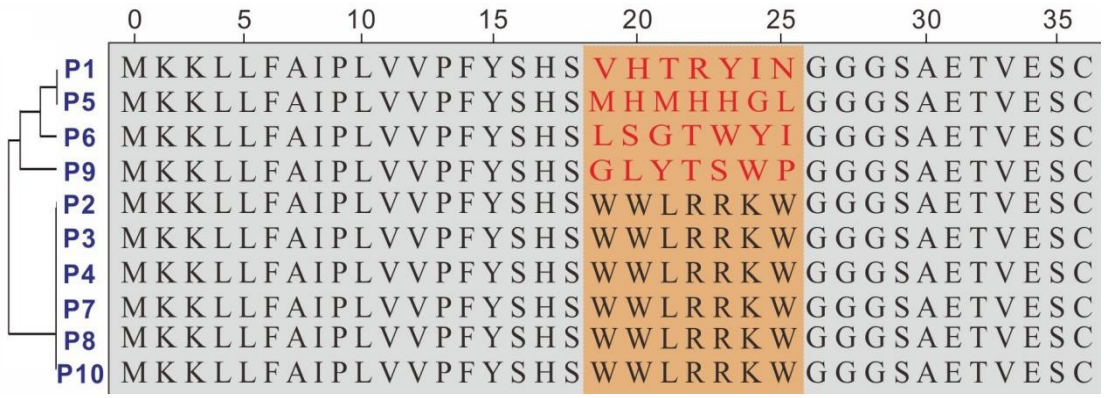


Fig. S14 The comparison results of amino acid sequence. The base sequence obtained from cloning P1-P10 was translated into amino acid sequence.

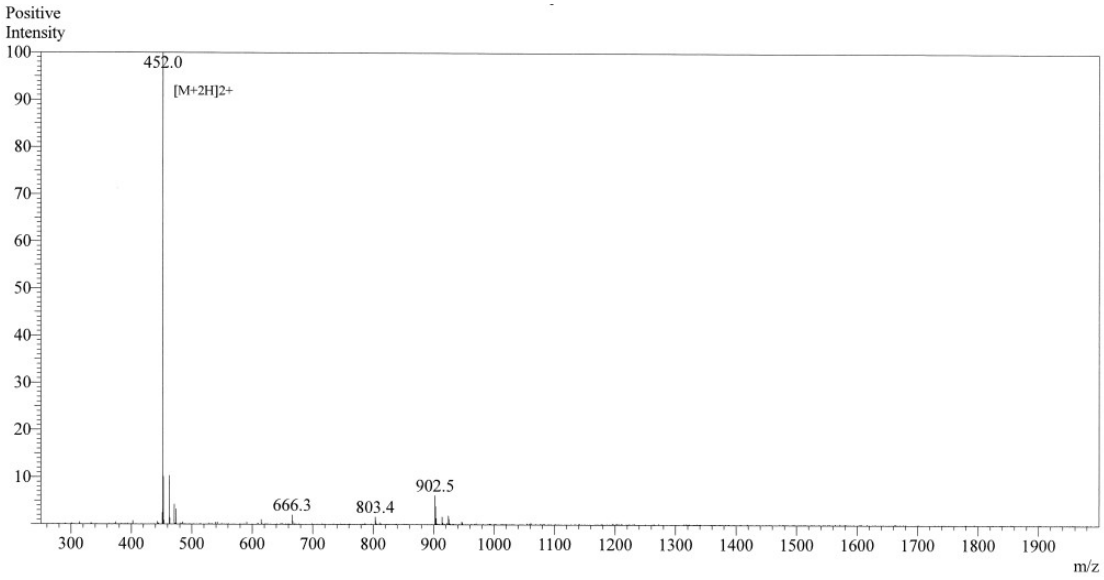


Fig. S15 The Mass spectrometric detection of P1.

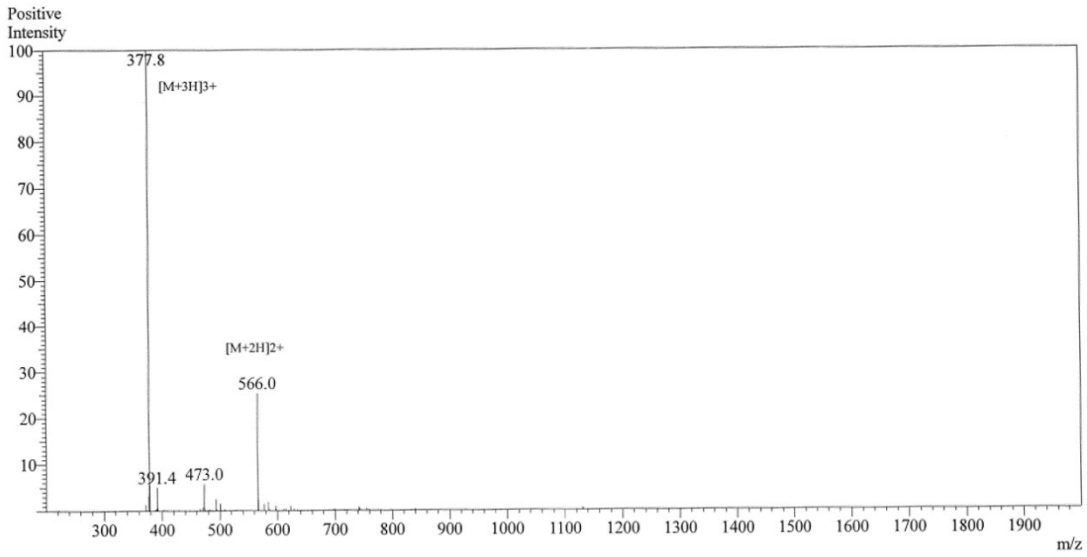


Fig. S16 The Mass spectrometric detection of P2.

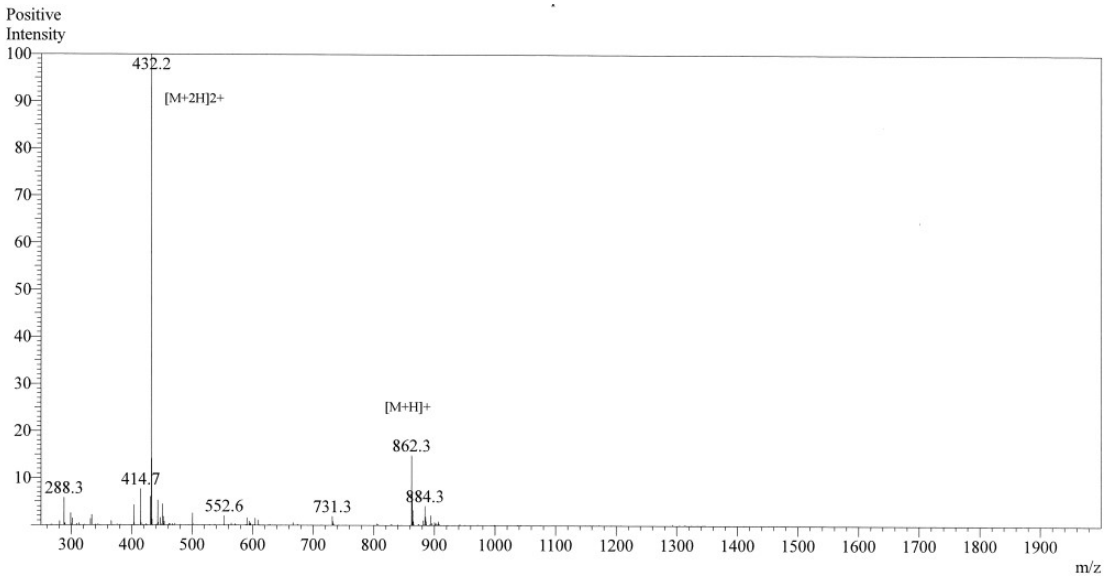


Fig. S17 The Mass spectrometric detection of P5.

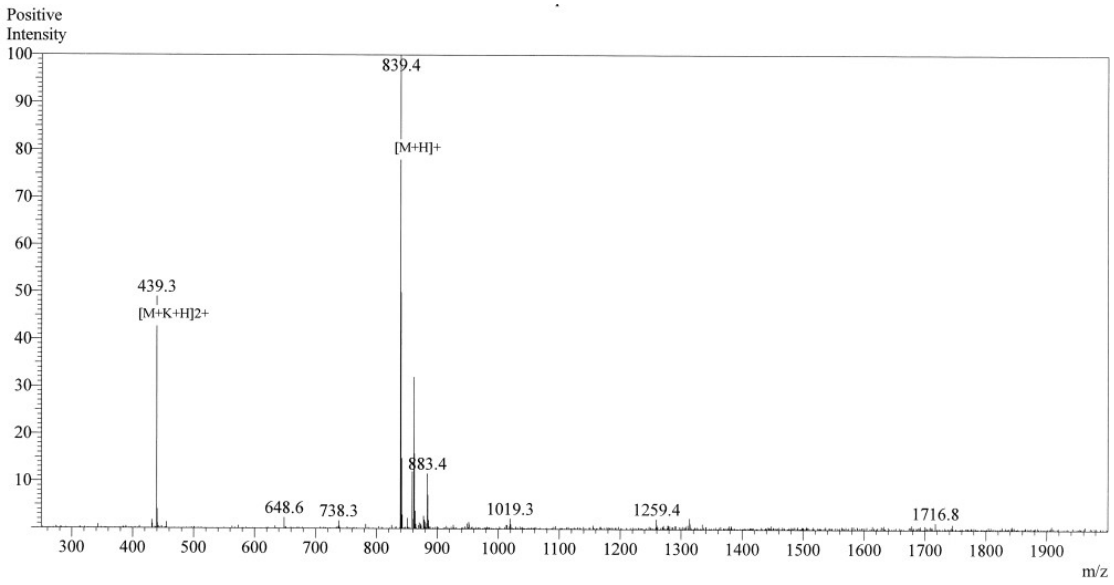


Fig. S18 The Mass spectrometric detection of P6.

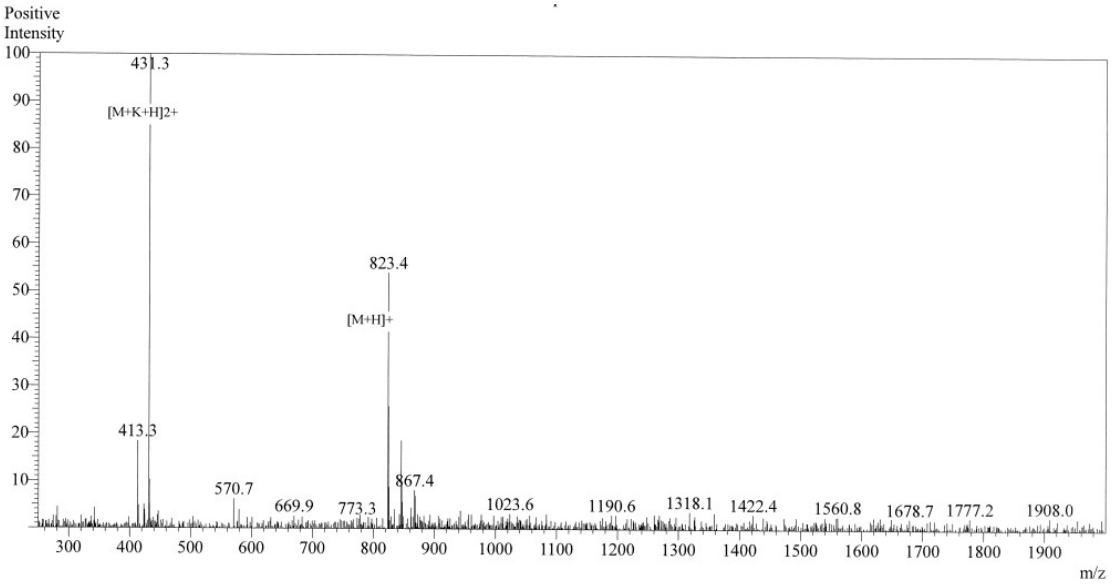
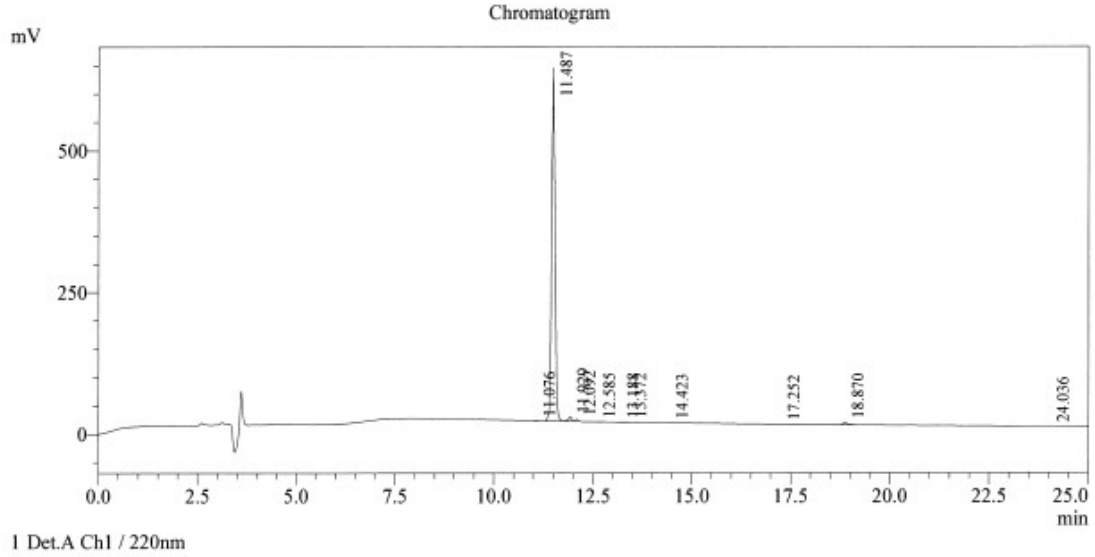


Fig. S19 The Mass spectrometric detection of P9.

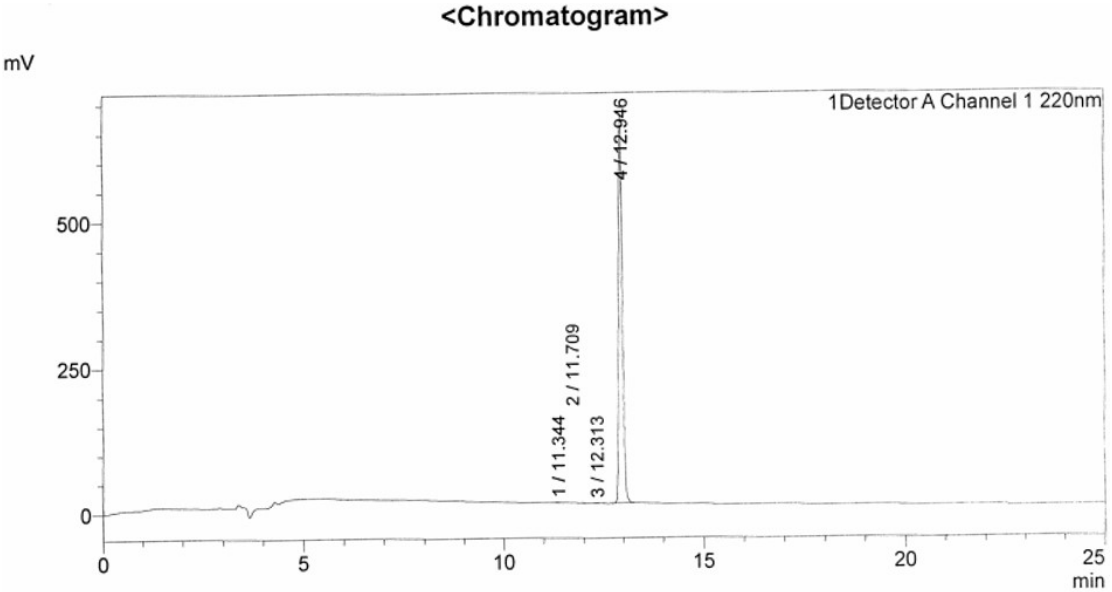


Peak Table

Detector A Ch1 220nm

Peak#	Ret. Time	Area	Height	Area %
1	11.076	2436	473	0.066
2	11.487	3609093	622168	97.407
3	11.929	37646	6995	1.016
4	12.092	14297	3131	0.386
5	12.585	2713	310	0.073
6	13.188	1102	233	0.030
7	13.372	3202	591	0.086
8	14.423	1979	350	0.053
9	17.252	1829	297	0.049
10	18.870	29792	4070	0.804
11	24.036	1085	115	0.029
Total		3705174	638734	100.000

Fig. S20 The purity of P1 (HPLC).

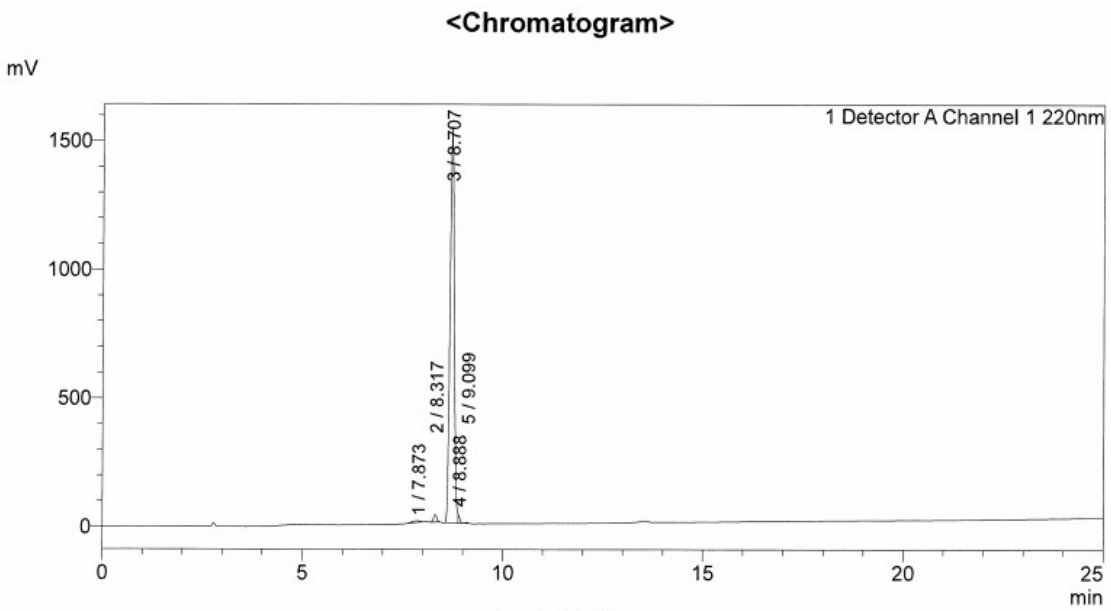


<Peak Table>

Detector A Channel 1 220nm

Peak#	Ret. Time	Area	Height	Area%
1	11.344	3036	243	0.072
2	11.709	3417	128	0.081
3	12.313	6997	957	0.166
4	12.946	4196480	665042	99.681
Total		4209930	666370	100.000

Fig. S21 The purity of P2 (HPLC).

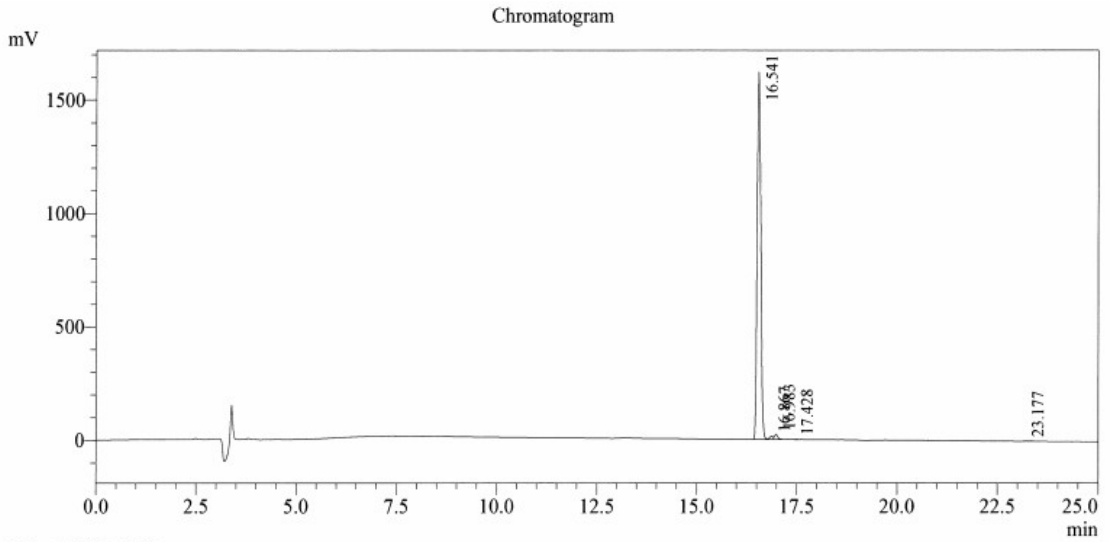


<Peak Table>

Detector A Channel 1 220nm

Peak#	Ret. Time	Area	Height	Area%
1	7.873	101497	5933	0.877
2	8.317	163873	30448	1.415
3	8.707	11205164	1540174	96.787
4	8.888	83614	34675	0.722
5	9.099	23035	4725	0.199
Total		11577183	1615955	100.000

Fig. S22 The purity of P5 (HPLC).

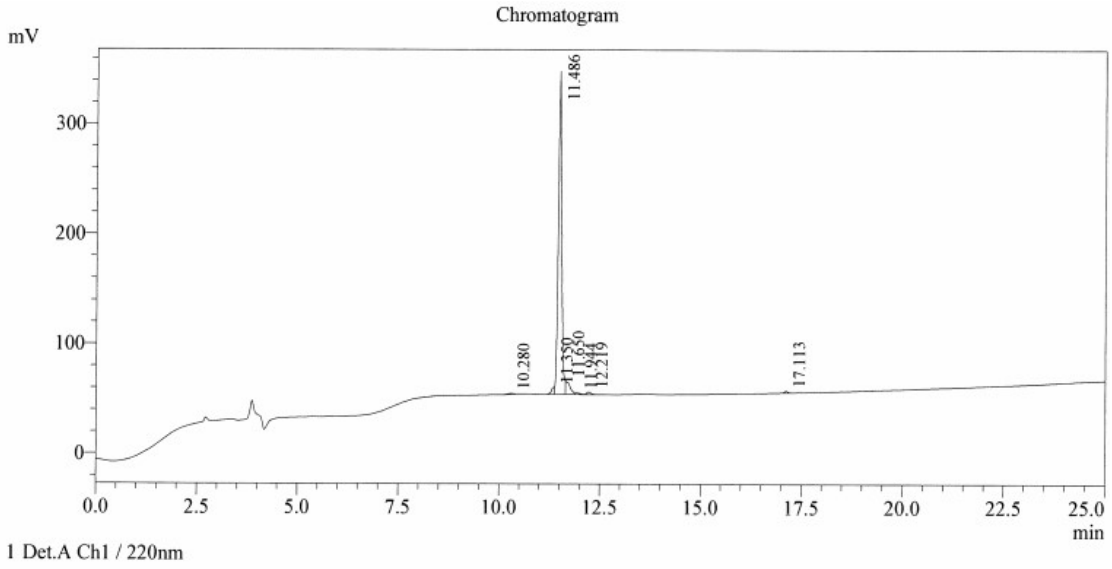


Peak Table

Detector A Ch1 220nm

Peak#	Ret. Time	Area	Height	Area %
1	16.541	9911140	1619894	97.622
2	16.867	81531	14201	0.803
3	16.983	143987	24105	1.418
4	17.428	9679	1682	0.095
5	23.177	6217	136	0.061
Total		10152554	1660017	100.000

Fig. S23 The purity of P6 (HPLC).



Peak Table

Detector A Ch1 220nm

Peak#	Ret. Time	Area	Height	Area %
1	10.280	7973	819	0.418
2	11.350	30697	6159	1.611
3	11.486	1737838	293955	91.219
4	11.650	91590	12158	4.808
5	11.944	10766	1463	0.565
6	12.219	16407	2355	0.861
7	17.113	9847	1693	0.517
Total		1905117	318602	100.000

Fig. S24 The purity of P9 (HPLC).

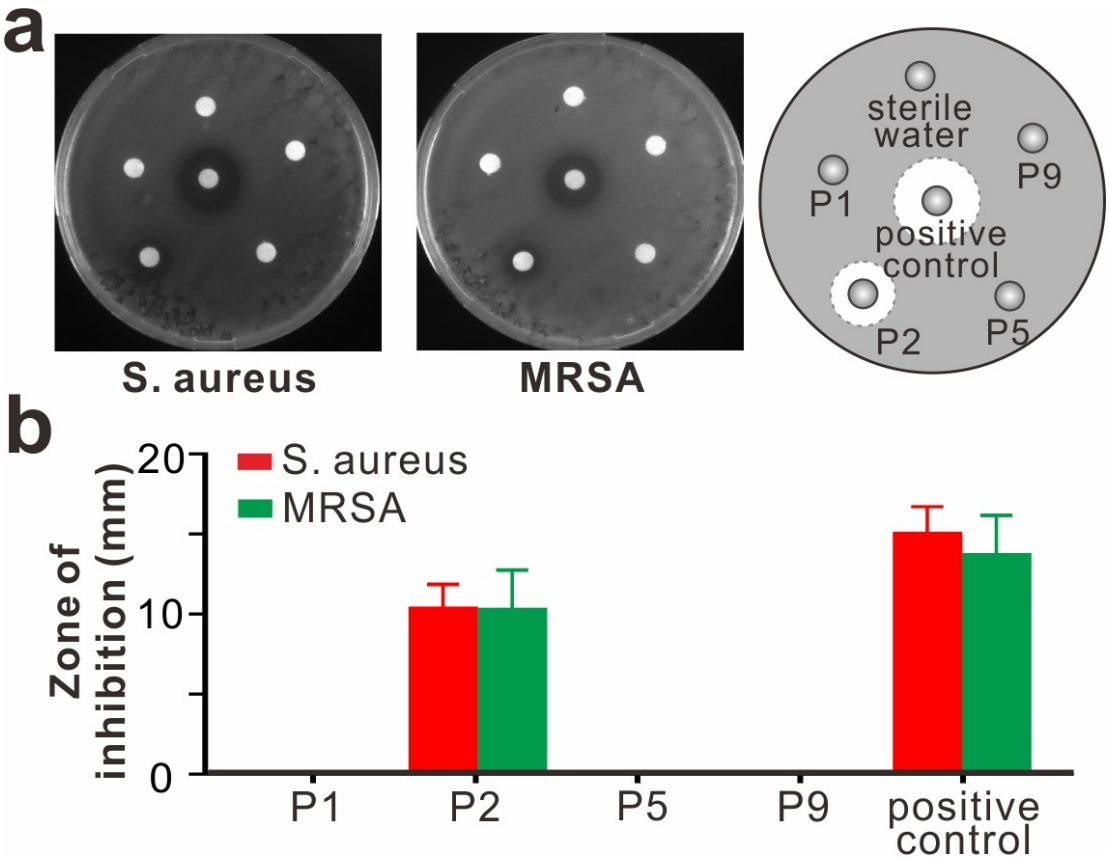


Fig. S25 Antimicrobial susceptibility test. Positive control: ampicillin, 0.1 mg/mL. Negative control: sterile water.

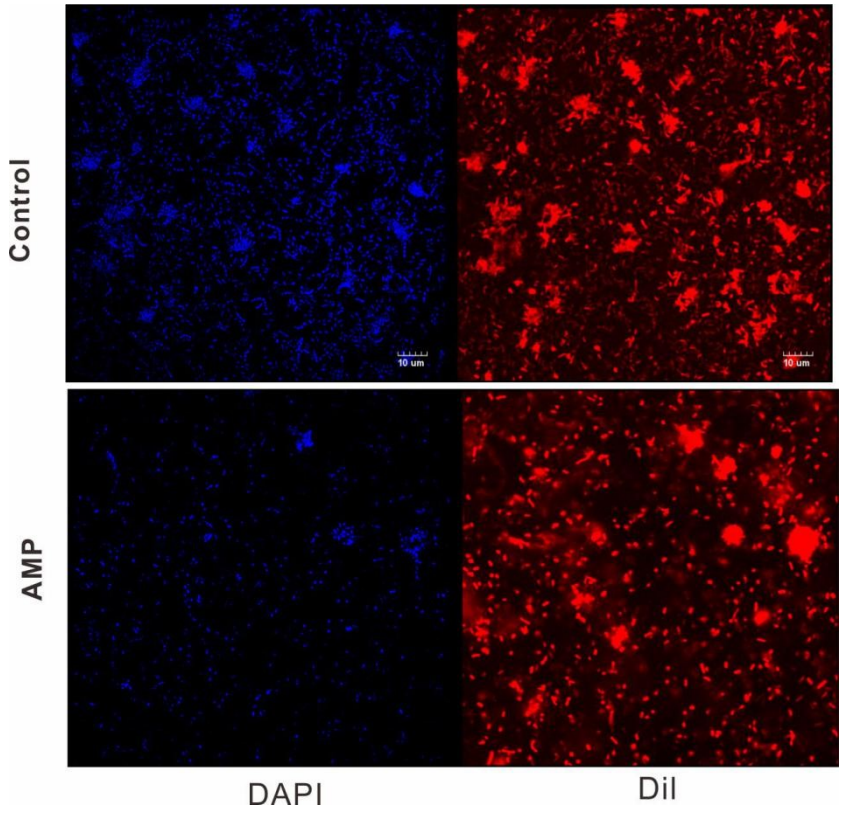


Fig. S26 Fluorescence images of cells stained with DAPI (a DNA-specific fluorescent probe) and Dil (cytomembrane staining), untreated and treated with AMP for (Scale bar 10 um).

Table S2. Summary of antibacterial compounds.

Antibiotic name	MIC (ug/mL)	Target
	Nucleic acid synthesis	
Sulfamethoxazole	64	Nucleotide biosynthesis and DNA replication
Trimethoprim	64	Dihydrofolate reductase
Novobiocin	0.25	DNA gyrase B
Ciprofloxacin	0.25	DNA gyrase A
Nalidixic acid	64	DNA gyrase A
Rifampicin	0.03	DNA-dependent RNA polymerase
	Protein synthesis	
Tetracycline	0.25	30S ribosome
Mupirocin	0.5	Isoleucyl t-RNA synthetase
Chloramphenicol	8	50S ribosome
Kanamycin	128	30S ribosome
	Cell wall synthesis	
Ampicillin	1	Penicillin-binding proteins
Daptomycin	32	Cell wall synthesis
Vancomycin	0.5	Binds D-Ala-D-Ala terminal
	Membrane active	
Pleurocidin	64	Cell membrane
Indolicidin	16	Cell membrane
	Unknown	
P2	64	

Table S3. Summary of peptides.

Peptide	Sequence
Pleurocidin	GWGSFFKKAAHVGKHVGKAALTHYL
Indolicidin	ILPWKWPWWPWRR
Small molecule AMP (P2)	WWLRRKW

Table S4. The performances of NNGA.

Accuracy	Cohen Kappa	F1-score	AUC
100%	100%	100%	100%

Table S5. The results of base sequence translation into amino acid sequence.

Peptide	Sequence	Probability
P1	VHTRYIN	1/10
P2	WWLRRKW	6/10
P5	MHMHHGL	1/10
P6	LSGTWYI	1/10
P9	GLYTSWP	1/10