## Supporting information for Generating ampicillin-level antimicrobial peptides with activity-aware generative adversarial networks

Andrejs Tucs,<sup>†</sup> Duy Phuoc Tran,<sup>‡</sup> Akiko Yumoto,<sup>¶</sup> Yoshihiro Ito,<sup>§,¶</sup> Takanori Uzawa,<sup>§,¶</sup> and Koji Tsuda<sup>\*,†,||,⊥</sup>

†Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwa-no-ha, Kashiwa, Chiba 277-8561, Japan

‡School of Life Sciences and Technology, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8550, Japan

¶Emergent Bioengineering Materials Research Team, RIKEN Center for Emergent Matter Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

§Nano Medical Engineering Laboratory, RIKEN Cluster for Pioneering Research, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

||RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

⊥Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Tsukuba, Ibaraki 305-0047, Japan

E-mail: tsuda@k.u-tokyo.ac.jp

Phone: +81 (4)-7136-3983. Fax: +81 (4)-7136-3975

Table S1: Activity predictor scores of generated peptides for different  $\lambda$  values

| $\lambda = 0$       | $\lambda = 0.5$   | $\lambda = 1$     |
|---------------------|-------------------|-------------------|
| $0.949 {\pm} 0.008$ | $0.952{\pm}0.012$ | $0.950{\pm}0.012$ |



Figure S1: The purities of the peptides were quantified using a high-performance liquid chromatograph. We found over 80% purities for AMP2 and AMP9NC, over 70% for AMP8NC, 60% for AMP1 and over 90% purity for other peptides including AMP4. In this analysis, a mixture of solvent A (0.1% TFA in water) and B (0.1% TFA in acetonitrile) was used as the mobile phase; and a linear gradient from 0% to 100% B for 60 min at a flow rate of 1.0 mL/min was applied.



Figure S2: A typical feature of 310-helical structure was observed for most peptides including AMP4. The spectra were obtained using a circular dichroism spectropolarimeter (J1500, JASCO). We used 1cm pathlength cuvette for 0.1mg/ml AMP2, AMP3 and AMP4 and 0.1cm pathlength cuvette for the other peptides (1mg/ml).



Figure S3: Activity predictor performance. (a) Receiver operating characteristic curve (ROC); (b) Precision-recall curve for activity predictor (PR). For ROC the corresponding area under curve AUC=0.86, whereas for PR AUC=0.89.