Supplementary Figure S12: Multiple sequence alignment of the predicted Type II thioesterase (orf-a) along with the known TEII of Surfactin-Synthetase (SrfAD) and Rifamycin (RifR).

The alignment highlights the conserved active site triad using red triangles in all the three sequences, S82-D192-H217 in orf-a; S94-D200-H228 in RifR [1]; S86-D189-H216 in SrfAD [2]. The predicted orf-a protein harbours a Pfam thioesterase domain (PF00975, HMMER hit, positions 15-236 with E-value=4.2e-7) which belongs to the alpha-beta hydrolase clan (CL0028). The same sequence region also gives a structural HHPrep hit to 2K2Q_B (Surfactin synthetase thioesterase subunit) with an E-value of 4.8e-24 as well as to 3FLA_A (RifR alpha-beta hydrolase thioesterase) with an E-value of 2.7e-17.

Reference List
