Figure S1.

Alvinellacin purification and molecular identification. Material eluting at 60% acetonitrile (ACN) upon solid phase extraction was loaded onto a C18 column (250×4 mm, Vydac). Elution was performed with a linear gradient of acetonitrile in acidified water (dotted line), and absorbance was monitored at 225 nm. Each individually collected fraction was tested for its antimicrobial activity (white bar) and its immunoreactivity to the alvinellacin Ab by DIA (grey bar). Fractions containing antimicrobially active alvinellacin were further purified by two additional RP-HPLC purification steps. Asterisk shows the active final fraction containing alvinellacin.

doi:10.1371/journal.pone.0095737.s001

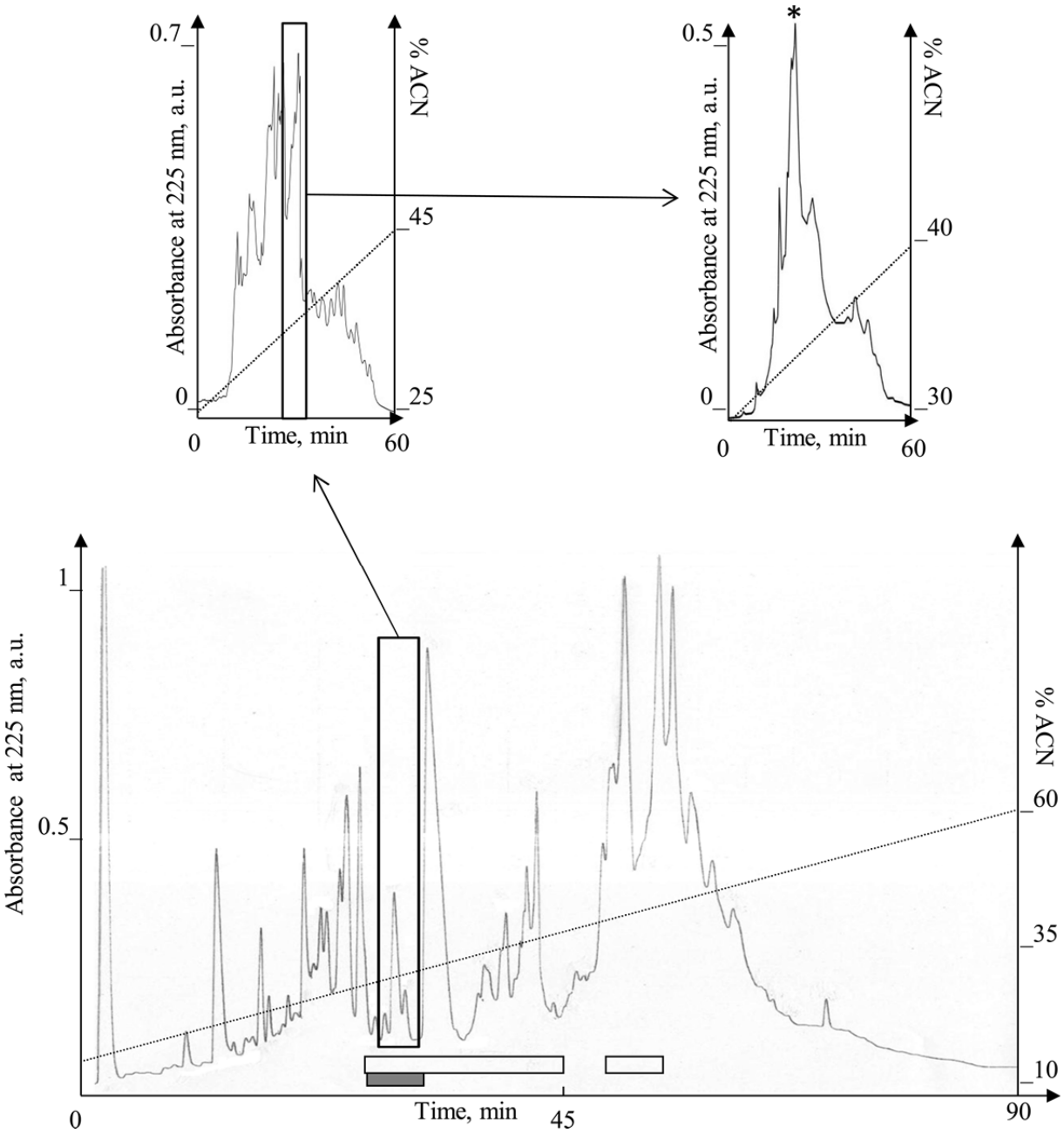

Figure S2.

MS spectrum of native alvinellacin. Analysis of purified alvinellacin by MALDI TOF-MS shows a m/z value of 2,600.35 MH+ which perfectly matches the theoretical mass of the peptide including two disulfide bonds.

doi:10.1371/journal.pone.0095737.s002

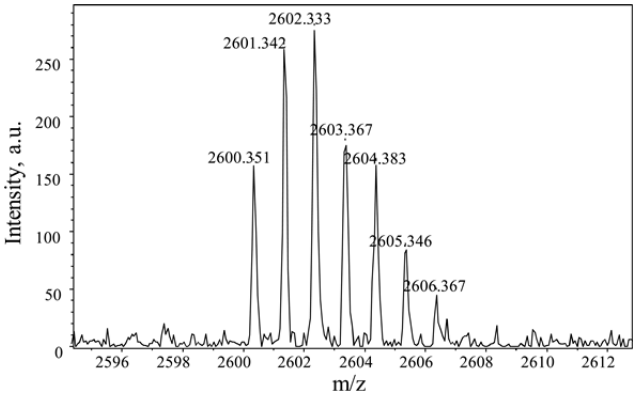

Figure S3.

Sequence alignments of the precursors of alvinellacin, capitellacin, and two arenicin isoforms.

doi:10.1371/journal.pone.0095737.s003

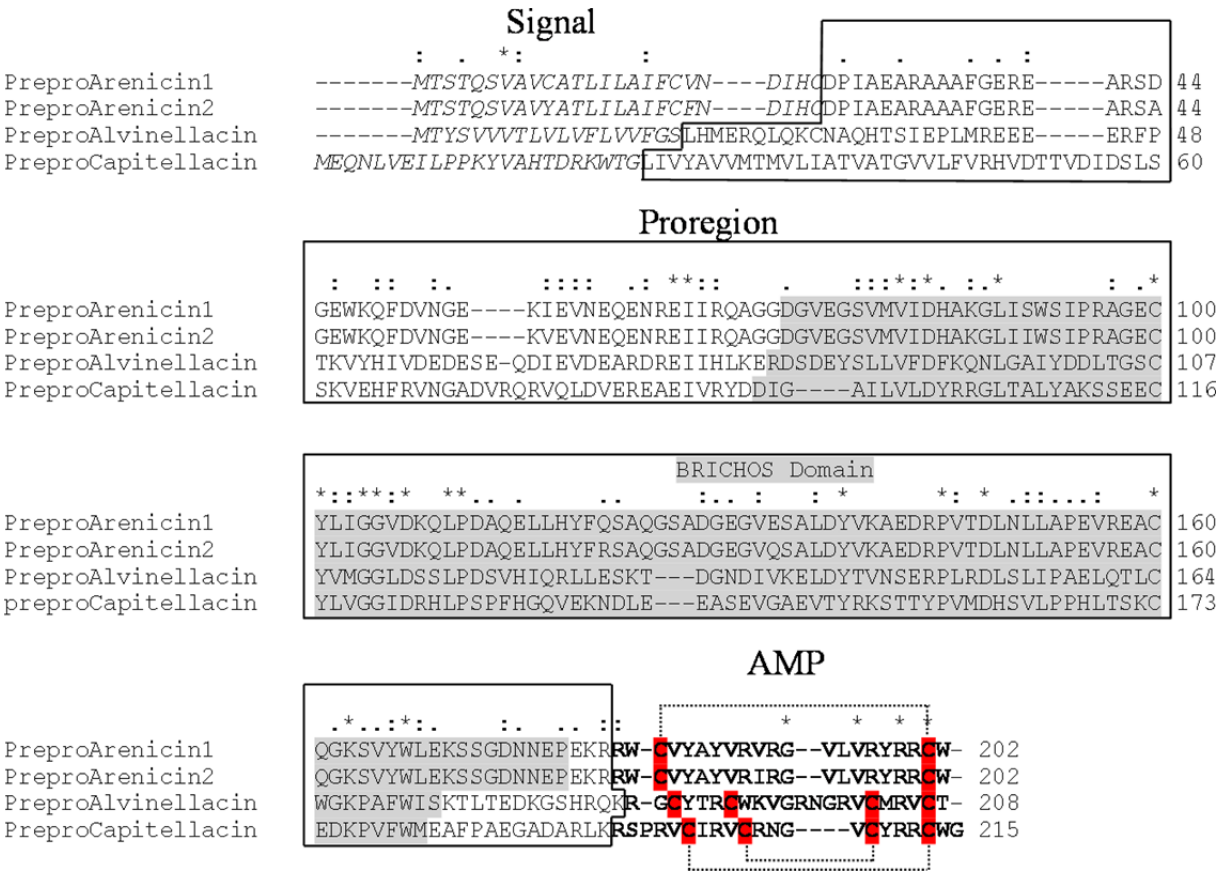

Figure S4.

Intact protein MS spectrum of alvinellacin measured by nanoESI-Orbitrap MS. (A) Full range MS survey spectrum. (B) Zoom-in of the [M+5H]5+ charge state species in a. A small species (indicated as asterisk) found next to the major component was identified as the methionine oxidation product of alvinellacin. The experimentally determined monoisotopic MW of alvinellacin was 2,599.2221 Da. (C) Display of theoretical MW (2,599.2067 Da) of alvinellacin and its isotope distribution at charge state 5. The results indicated that all four cysteines are involved in the formation of disulfide bonds.

doi:10.1371/journal.pone.0095737.s004

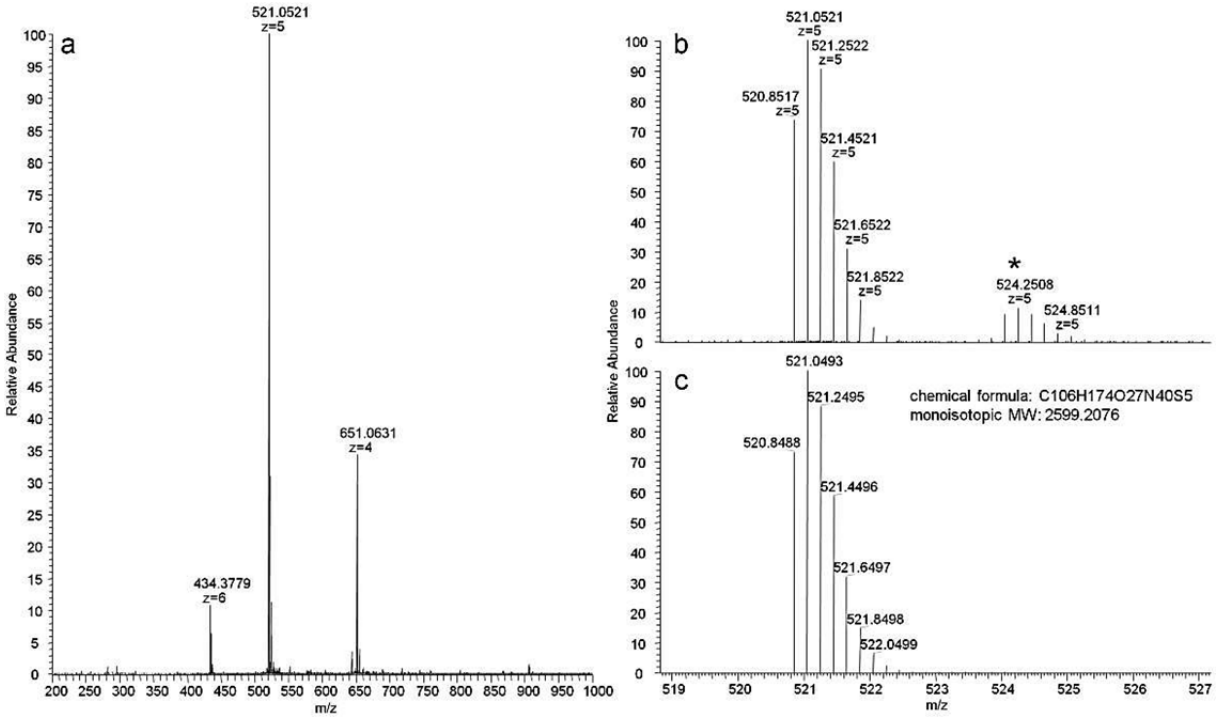

Figure S5.

Time-course analysis of the proteolytic cleavage of alvinellacin. The products of alvinellacin digestion were analyzed by nanoESI-Orbitrap MS. (A) Peptide MS survey spectra of alvinellacin digested with Lys-C at 35°C (overnight). (B) Subsequent digestion of the Lys-C-digest with trypsin after 30 min; (C) after 2 h; (D) after 18 h at 37°C. The identities of the peptides are summarized in Table S2.

doi:10.1371/journal.pone.0095737.s005

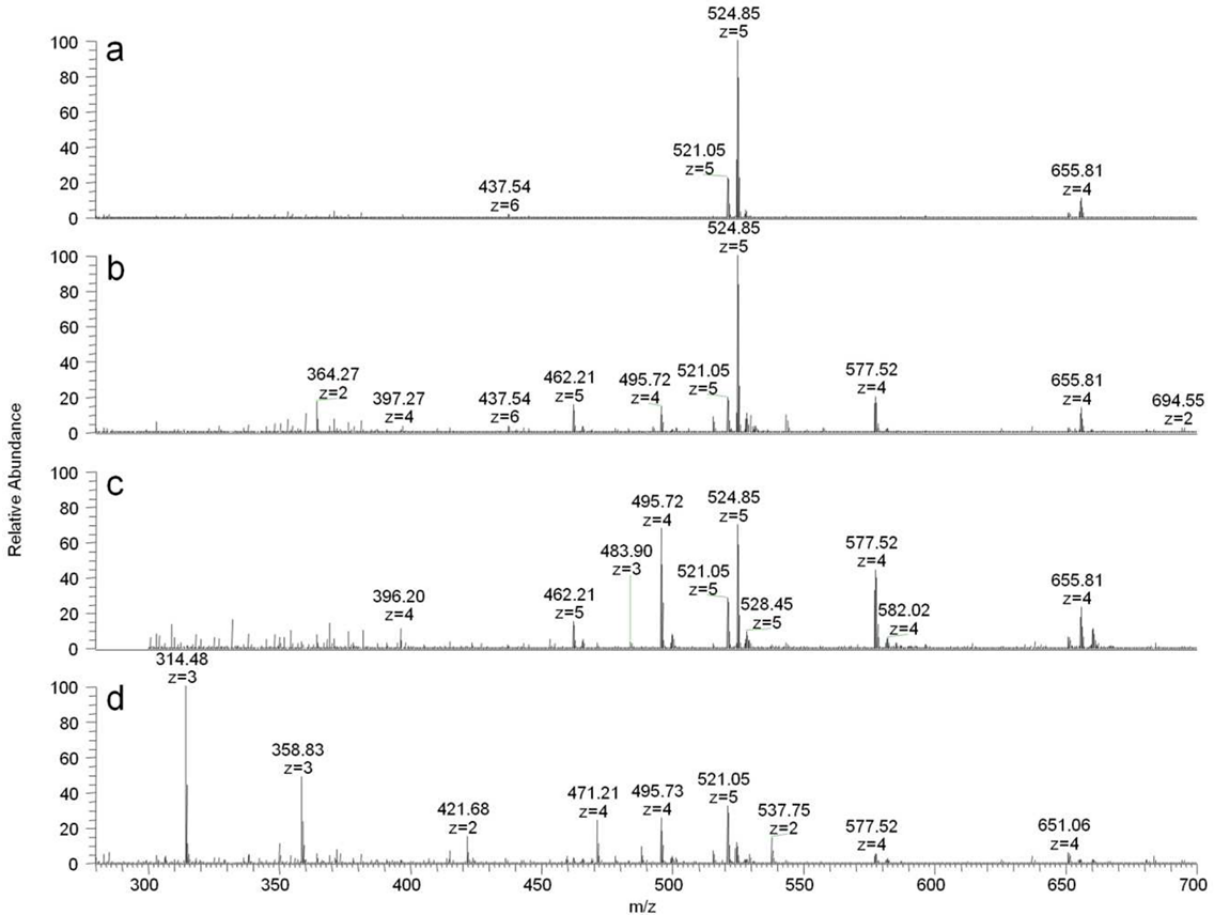

Figure S6.

Alvinellacin and capitellacin gene structures. (A) As opposed to CDS (648 bp), the alvinellacin gene is rather long (1949 bp from the initial methionine to the stop codon) with a 5 introns/6 exons structure and a first large intron of 442 bp. Introns are all inserted in phase 0 with the exception of the last one in phase 1. (B) Alignment of the translated regions of the alvinellacin and capitellacin genes. The intron splicing positions (triangles) are nearly conserved. The BRICHOS domains are shaded and the AMP sequences are in bold type.

doi:10.1371/journal.pone.0095737.s006

A

```
ATG ACG TAT TCT GTA GTT GTG ACG CTG GTC CTA GTG TTT CTT GTC GTC TTC GGT AGT CTG CAT ATG GAA
 M   T   Y   S   V   V   V   T   L   V   L   V   F   L   V   V   F   G   S   L   H   M   E
CGG CAG CTG CAG AAA TGC AAC GCG CAG gtaagattaccttgccgtgtttgttacgccaacagattggtgctgtcgccataat
 R   Q   L   Q   K   C   N   A   Q
tataatcatcatggaatttaaaatacattcagccttcactttgccgtcgatgtttttccaaacgaagccagcatttcgaaaaccacctga
atagtcatcttgaacacaataatcactatcagtttcgcaagaattgtcctaaatataatagctggccacctcaacatcactgccaacatca
agtatttgatcacatgaaatgaggcgcagttgacttgaagtttatctatgagtctttgtgctacagtttcacggaacgtctcgtgaatcta
gcgatcccacaggtggaatgcaatgtctttacttgattgattatctgcgttaagcaccaacaaataaatatcaactcttctatttgatgtc
acattcttctctttgtttatcag CAT ACT TCG ATT GAA CCC CTG ATG CGT GAA GAA GAG GAG CGC TTT CCT ACA
                        H   T   S   I   E   P   L   M   R   E   E   E   E   R   F   P   T
AAG gtaaatacctgactggtgtctattattactgcaacgtgacgcaagcaaacgattccaatactggatgaatggatcgaccaatcgatt
 K
gatcgatctggcaaacgtttcag GTT TAT CAC ATT GTG GAC GAT GAT GAA ACC GAA CAA GAC ATC GAA GTA GAC
                        V   Y   H   I   V   D   D   D   E   T   E   Q   D   I   E   V   D
CAA GCA CGT GAC CGG GAG ATA ATC CAT TTG AAG GAG CGC GAT AGT GAT GAA TAT TCA TTA CCT GTC TTC
 Q   A   R   D   R   E   I   I   H   L   K   E   R   D   S   D   E   Y   S   L   P   V   F
GAT TTC AAG CAG gtatcatcttctacttacatttttactacgtaatggaaatgcaaacaattctcaccaagtttaatcgaagcctaa
 D   F   K   Q
atgtctaacatcgaagaatgaggtatggtttcatatagttgcaccaccctatggctaaactaaactgcagtgagacgatattgataaaaag
aagcattttaaaagagaaatactggtgaccaccagaataggtctgcgatgttttatttcttttctggctgaattgttttggtgcag AAT
                                                                                        N
CTC GGA GCC ATT TAC GAC GAT CTT ACC GGA TCG TGT TAC GTC ATG GGT GGC CTT GAC AGT AGT TTG CCA
 L   G   A   I   Y   D   D   L   T   G   S   C   Y   V   M   G   G   L   D   S   S   L   P
GAC AGT GTA CAT ATA CAG CGA TTG CTT GAA AAC AAG gtaaacatttgctagactgctgaaccaaatagcaataggtagc
 D   S   V   H   I   Q   R   L   L   E   N   K
taatacggcaggtaatacttcagttgggccacattaaatttgcgggctcagattccgtctgcagaaaggtggttagcgtggctattttcat
ggcatatttggaagttaaagaataatggtcgaacgttctggaaacatttgaccagttttgatcttatctag ACT GAT GGC AAT GAC
                                                                          T   D   G   N   D
ATC GTG AAG GAA CTC GAT TAC ACC GTC AAC TCT GAA CGT CCA CTG AGA GAT CTG AGC CTG ATT CCA GCC
 I   V   K   E   L   D   Y   T   V   N   S   E   R   P   L   R   D   L   S   L   I   P   A
GAG CTC CAG ACG TTG TGT TGG GGA AAG CCT GTC TTC TGG ATC AGT AAG ACT CTA ACC GAA GAC AAA GGT
 E   L   Q   T   L   C   W   G   K   P   V   F   W   I   S   K   T   L   T   E   D   K   G
TCT C gtaagtagctacaaattaacgtcggaattgggaacgaaatcattcctaatgtatacgatgattgcatatcgtaccggtggcgtct
 S
tctgttatctgtatgcgatacgggctggttttaagcgtagaattatgtagcttgtaagttgcagacgtgggcttcgcgtacatgtgtgggt
gtgtaggtgtgggattattcgatggaaatagcattgtgttgtcttcgtttcag AT CGT CAG AAG AGA GGT TGT TAC ACA CGT
                                                      H   R   Q   K   R   G   C   Y   T   R
TGT TGG AAA GTT GGT AGG AAC GGA CGT GTT TGT ATG CGT GTT TGT ACA TAA ctcacctgcttcatttcactgag
 C   W   K   V   G   R   N   G   R   V   C   M   R   V   C   T  STOP
aaaa
```

B