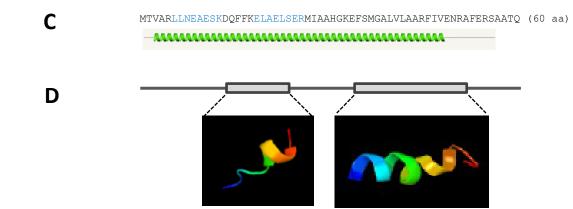


B 4078856 cgagcaccag gaggaaatcc atgacagtag cacgattgct aaatgaagcc gaaagtaagg atcagttctt caaagagctg gcggaactat cggagcgcat gatcgccgcg cacggaaaag agttctcgat gggcgccttg gtgctggcgg cgaggttcat cgtcgaaaat agggcgttcg agcgaagcgc agcaacacag tga 4079058



S2 Fig. Blr0566_ISGA gene and protein. A) The blr0566_ISGA gene is weakly expressed at the mRNA level and is probably co-transcribed with upstream genes. Shown are cDNA reads and relevant annotated features (annotated genes, TSSs, promoter and terminators) at the blr0566_ISGA locus [15]. Only the plus strand is shown. The differential RNA-seq (dRNA-seq) was described previously [15]. The analyzed RNA was isolated from exponentially growing, free-living cells (F) in liquid cultures and from nodules (N). RNA samples were treated (+) or not treated (–) with terminal exonuclease (TEX), which degrades 5'-monophosphorylated (processed) transcripts. The scale of each library is indicated (Reads). B) Gene sequence with upstream region. The genomic coordinates are indicated, the start and stop codons are in red and the putative Shine-Dalgarno sequence is underlined. C) Protein sequence with proposed secondary structure. Blue, peptides detected by proteomics [15]. The peptides were detected only in nodules. D) Proposed three-dimensional structure of indicated protein parts . Phyre² was used for secondary and tertiary structure modeling [43].