Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2023

Supplementary Information

VITAS, a sensitive in vivo selection assay to discover enzymes producing antiviral natural products

Aws Fahd Alharbi, Hayun Kim, Dhirish Chumroo, Yuxuan Ji, Mohammed Hakil, Kourosh H. Ebrahimi*

Institute of pharmaceutical Science, King's College London, London, UK

Kourosh.ebrahimi@kcl.ac.uk

SI Methods

Chemical and reagents. All chemicals were reagent grade and were obtained from Merck or Fisher Scientific.

Creating the expression constructs for SAND enzymes. The expression construct for TtSAND is obtained as described previously[1,2]. The expression constructs for human SAND (hSAND), and PpSAND was prepared as explained for TtSAND. The construct for hSAND lacks the N-terminus hydrophobic domain (67 amino acids, residues number 2-68 highlighted in green). The amino acid sequence of each protein is described below:

>tr|M1WKF0|M1WKF0_PSEP2 Predicted Fe-S oxidoreductases OS=Pseudodesulfovibrio piezophilus (strain DSM 21447 / JCM 15486 / C1TLV30) OX=1322246 GN=BN4_12251 PE=4 SV=1
MSTVETLATDRFIPAINLFVTKHCNMRCRFCFGSCKMRSPLSSQDQDGVFVDVIRQCHQQ
GISKITFVGGEPLLYPKLKLLIRLAHDLGITTCVVSNGALLTKEWLREVSGMLDWIGISI
DSLSVDTNWSIGRISNGVPMSKLVYEQLVDWVHDYGMRLKINTTVCRWNHHEDMSSFYRD
TNPHRIKMFQALTIDGVNDEESTKFSVSDEQFTHYVERHLRQGIKAVAEASNDMVGSYLM
VSPDGCFFDNTHGSYRLSRPISRVGFSSAIKDISVNHTKFMDRGGMYRW

>sp|Q8WXG1|RSAD2_HUMAN Radical S-adenosyl methionine domain-containing protein 2 OS=Homo sapiens OX=9606 GN=RSAD2 PE=1 SV=1
MWVLTPAAFAGKLLSVFRQPLSSLWRSLVPLFCWLRATFWLLATKRRKQQLVLRGPDETK
EEEEDPPLPTTPTSVNYHFTRQCNYKCGFCFHTAKTSFVLPLEEAKRGLLLLKEAGMEKI
NFSGGEPFLQDRGEYLGKLVRFCKVELRLPSVSIVSNGSLIRERWFQNYGEYLDILAISC
DSFDEEVNVLIGRGQGKKNHVENLQKLRRWCRDYRVAFKINSVINRFNVEEDMTEQIKAL
NPVRWKVFQCLLIEGENCGEDALREAERFVIGDEEFERFLERHKEVSCLVPESNQKMKDS
YLILDEYMRFLNCRKGRKDPSKSILDVGVEEAIKFSGFDEKMFLKRGGKYIWSKADLKLD
W

The *E. coli* codon-optimised gene for the expression of each protein was obtained from GeneArt (ThermoFisher Scientific). The N-terminal hydrophobic residues were removed from the synthetic gene of human SAND (or RSAD2). The removed peptide sequence is: WVLTPAAFAGKLLSVFRQPLSSLWRSLVPLFCWLRATFWLLATKRRKQQLVLRGPDETKEE EEDPPL). Each gene was subcloned into pBAD/His C vector using KpnI and EcoRI restriction sites as described for TtSAND[1]. The presence of the correct insert was confirmed using agarose gel electrophoresis.

Creating the expression construct for toxin LdrD. The E. coli codon-optimised gene for the expression of toxin LdrD was obtained from GeneArt (ThermoFisher Scientific). The amino acid sequence of the LdrD protein is shown below.

>sp|Q6BF25|LDRD_ECOLI Small toxic polypeptide LdrD OS=Escherichia coli (strain K12) OX=83333 GN=ldrD PE=1 SV=1
MTFAELGMAFWHDLAAPVIAGILASMIVNWLNKRK

The LdrD gene was subcloned into pET28a vector using NdeI and EcoRI restriction sites. The presence of the correct insert was confirmed using agarose gel electrophoresis.

Preparation of samples and VITAS assay. To perform the VITAS assay, first competent E. coli BL21-AI cells were transformed with the pBAD/His C plasmid carrying the gene of a SAND, or pET28a plasmid carrying the gene of LdrD, or both SAND and LdrD plasmids. After transformation, cells were spread on LB-Agar plates containing 100 µg/ml ampicillin (transformed with pBAD/His C plasmid only), 80 µg/ml kanamycin (transformed with pET28a plasmid only), and 100 µg/ml ampicillin and 80 µg/ml kanamycin (transformed with pBAD/His C & pET28a plasmids). The concentration of arabinose in the LB-Agar was varied between 0-0.04%. After spreading cells on the plates, they were incubated at 37 °C overnight. Single colonies were taken from the plate and inoculated in 2 mL LB media containing relevant antibiotics. The LB media was added to reusable glass tubes with phenolic screw caps (Fisher Scientific). The concentration of arabinose and IPTG in the LB media was varied. After inoculation, the media was incubated in a shaker (200 rpm) at 37 °C. Subsequently, optical density (OD) was measured at 600 nm. Measurements were repeated at least 3 times with different biologically independent samples.

Protein expression and purification. Each protein was overexpressed in TOP10 E. coli cells. A single colony of cells transformed with a pBAD/His C plasmid containing the SAND gene was inoculated in 25 ml of LB media containing 100 µg/ml ampicillin. The flask was incubated in a shaker (200 rpm) at 37 °C overnight (about 16 hours). The next day, the preculture was added to 250 ml TB media. The culture was incubated in a shaker (200 rpm) at 37 °C until OD at 600 nm reached 0.5-08. At this point, SAND's expression was induced by adding a 0.04% (mg/ml) final concentration of arabinose and the culture was further incubated in the shaker for another 7-8 hours. Subsequently, cells were collected by centrifugation at 5000 rpm for 20 min, and cells were frozen in a -80 freezer. Cells were resuspended in 5 ml of lysis buffer (50 mM Tris, 300 mM NaCl, pH 8.0 containing 2% triton, 0.01 mg/ml PMSF, 0.05 mg/ml DNase, and 1 mM DTT). The lysis buffer was supplemented with 1x cell lytic B buffer (ThermoFisher). Cells were lysed using a probe sonicator (Branson Sonifier). The lysis of *E. coli* cells was performed on ice. The conditions of sonication were: On cycle 30 seconds, off cycle 15 second, number of on-cycle 10, %amplitude 30%. After sonication, the lysed cells were centrifuged for 40 min at 5000 rpm to remove cell debris and collect the cell-free extract containing soluble proteins. The cell-free extract was immediately transferred to an anaerobic glovebox (Belle Technology) with an O_2 concentration of less than 5-8 ppm. cOmpleteTM His-tag purification resin (Roche) and gravity chromatography were used to purify each protein. 0.5 ml of resin was used per 250 ml of TB culture. The resin was equilibrated using equilibration buffer (10 times the volume of the reissue): 50 mM Tris, 300 mM NaCl, pH 7.5, 2 mM Imidazole, 0.2% triton, 1 mM DTT. Then the cell-free extract was added to the resin and incubated in the glovebox for 1 hr with stirring. Next, the cell-free extract containing the resin was added to a gravity column (ThermoFisher). The flow-through was collected. The resin was washed with the Wash buffer (10-time the volume of the resin): 50 mM Tris, 300 mM NaCl, pH 7.5, 10 mM Imidazole, 0.2% Triton, 1 mM DTT. After washing, the protein of interest was eluted using elution buffer (2-time the volume of resin): 50 mM Tris, 300 mM NaCl, pH 7.5, 200 mM Imidazole, 0.2% Triton, 1 mM DTT. To remove imidazole PD10 desalting column was used. The PD10 column (Cytiva) was equilibrated with 50 mM MOPS, 100 mM NaCl, pH 7.0**.** The purified protein was added to the column and eluted using 50 mM MOPS, 100 mM NaCl, pH 7.0 according to the manufacturer's protocol. The overexpression and purification of each protein were confirmed using SDS-PAGE gel electrophoresis. The protein concentration was measured using nanodrop by measuring absorbance at 280 nm or using BCA assay. The

purified protein was flash-frozen in liquid nitrogen and kept in a -80 freezer until further analysis.

UV-visible spectroscopy analysis of [FeS] cluster. Samples were prepared in an anaerobic glovebox. Immediately after the protein purification, an aliquot of the protein was added to a 1 ml Quartz cuvette. The sample was sealed tightly and then moved outside the glovebox. UV-visible absorbance spectrum was immediately recorded using a LAMBDA 365 UV-Vis spectrophotometer (PerkinElmer). The spectrophotometer was blanked using the elution buffer before the measurements of the samples. The spectrometer lamp was changed at 400 nm leading to a small drift observed in the absorbance spectra.

SI figures

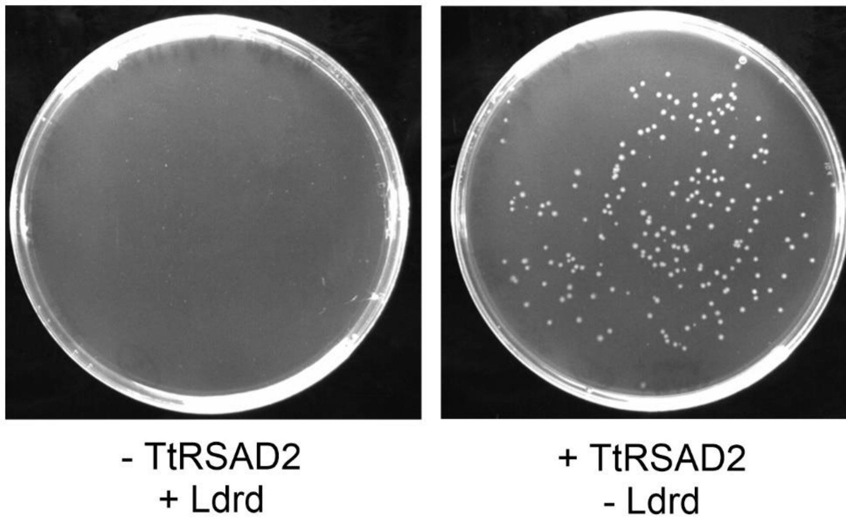


Fig. S1. Expression of LdrD in *E. coli* BL21 cells effectively blocks growth. (a) The *E. coli* BL21 cells were transformed with the plasmid expressing LdrD or TtSAND. After transformation, cells were spread on agar plates containing 0.04% arabinose and 1 mM IPTG and kanamycin (+LdrD) or ampicillin (+TtSAND). The plates were incubated overnight at 37 °C. Measurements were repeated three times.

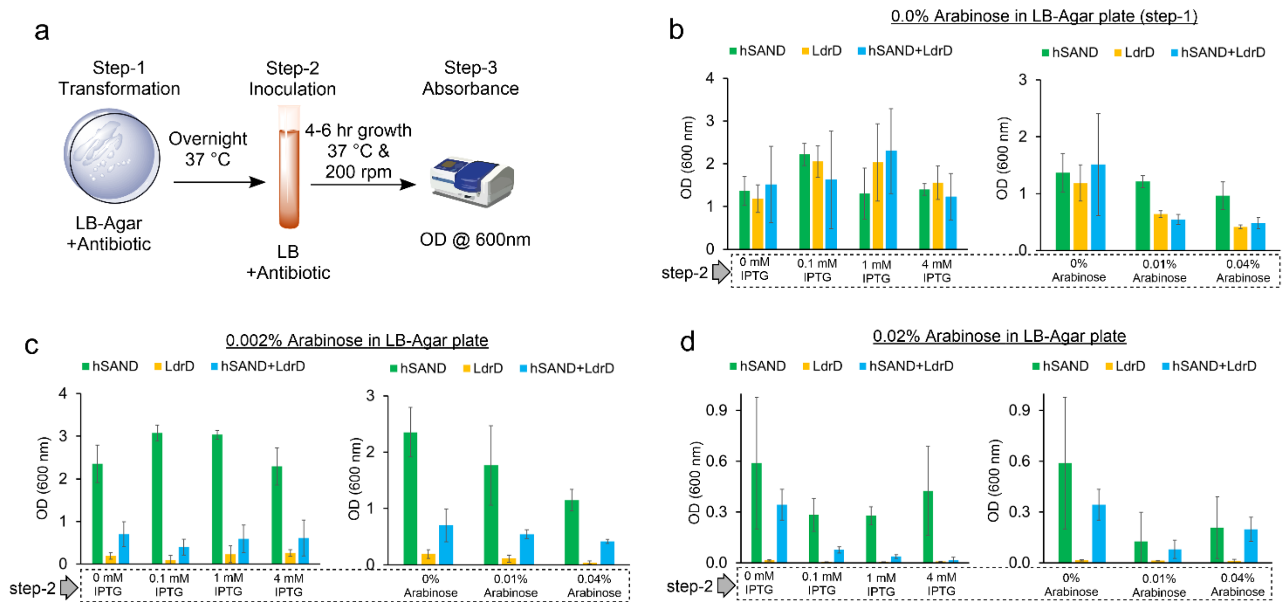


Fig. S2. The VITAS assay was optimised by varying the concentration of inducing agents. (a) A scheme representing the three steps used to test the VITAS assay. Arabinose was added in step-1, and arabinose or IPTG were added in step-2. (b) No arabinose in Step-1 and 0-4 mM IPTG or 0-0.4% arabinose in Step-2; (c) 0.002% arabinose in Step-1 and 0-4 mM IPTG or 0-0.4% arabinose in Step-2; and (d) 0.02% arabinose in Step-1 and 0-4 mM IPTG or 0-0.4% arabinose in Step-2. The optical density (OD) was measured at 600 nm after 5 hours of incubation. Data are averages of measurements using three biologically independent samples ± standard deviations.

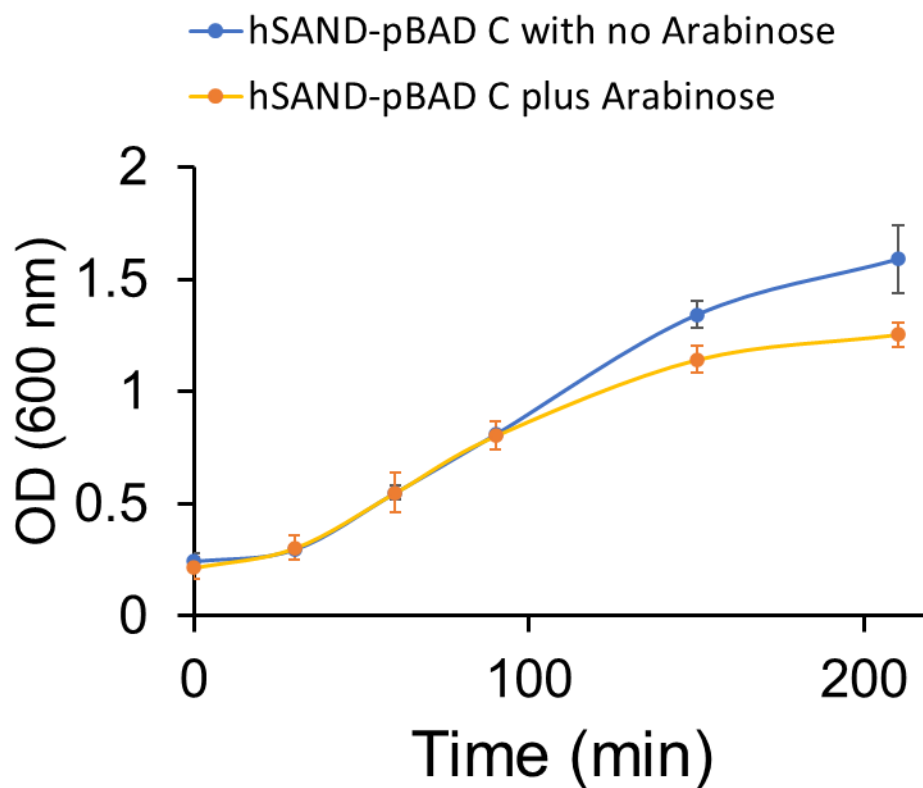


Fig. S3. Expression of hSAND slightly inhibits growth of *E. coli* cells. Top10 cells transformed with the construct expressing hSAND were treated with no arabinose or 0.02% arabinose (final concentration). Arabinose was added when OD reached circa 0.5. After addition of arabinose and the expression of hSAND the growth of *E. coli* cells slowed down.

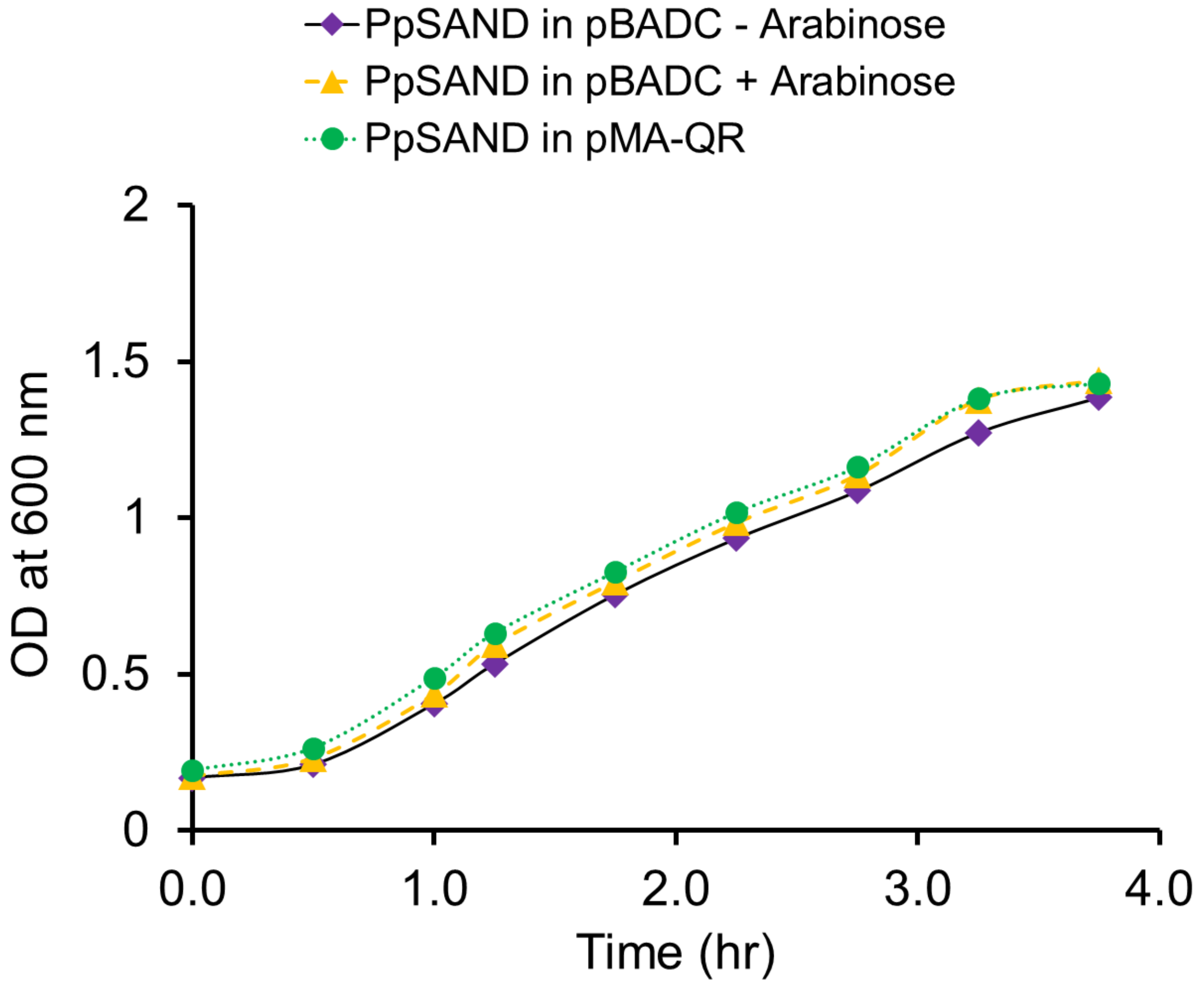


Fig. S4. Expression of PpSAND in E. coli does not cause any toxicity. The E. coli Top10 cells were transformed with the pMA-QR plasmid containing the codon-optimised gene of PpSAND (green) or with the pBAD/His C plasmid containing the codon-optimised gene of PpSAND (purple and orange). Expression of PpSAND was induced (orange) at 1:15 hr by the addition of 0.004% arabinose. The pMA-QR plasmid does not contain any promotor for the protein expression. Data are averages of three measurements ± standard deviation.

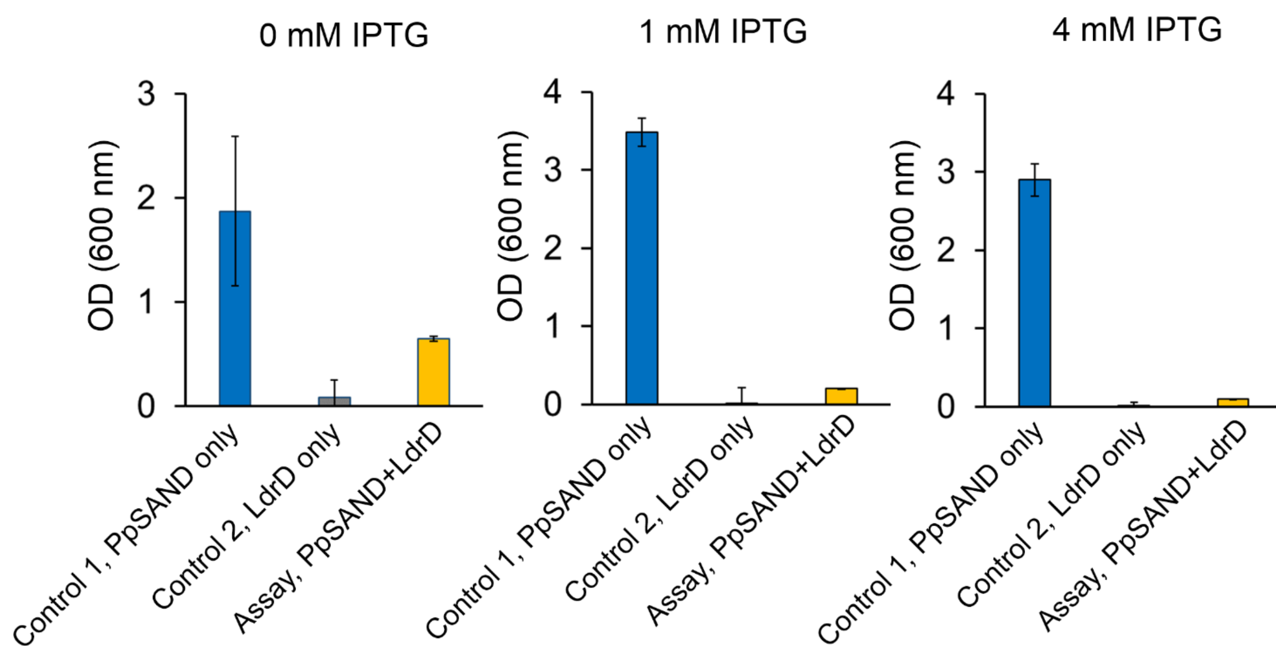


Fig. S5. The expression of LdrD and the response of the VITAS assay can be modulated by varing the concentration of IPTG. The assay was tested using PpSAND with 0.02% arabinose in the LB-agar plate and three different concentrations of IPTG in the LB media. Data are average of three biologically independent measurements ± standard deviation.

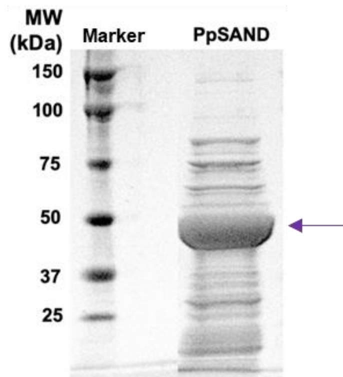
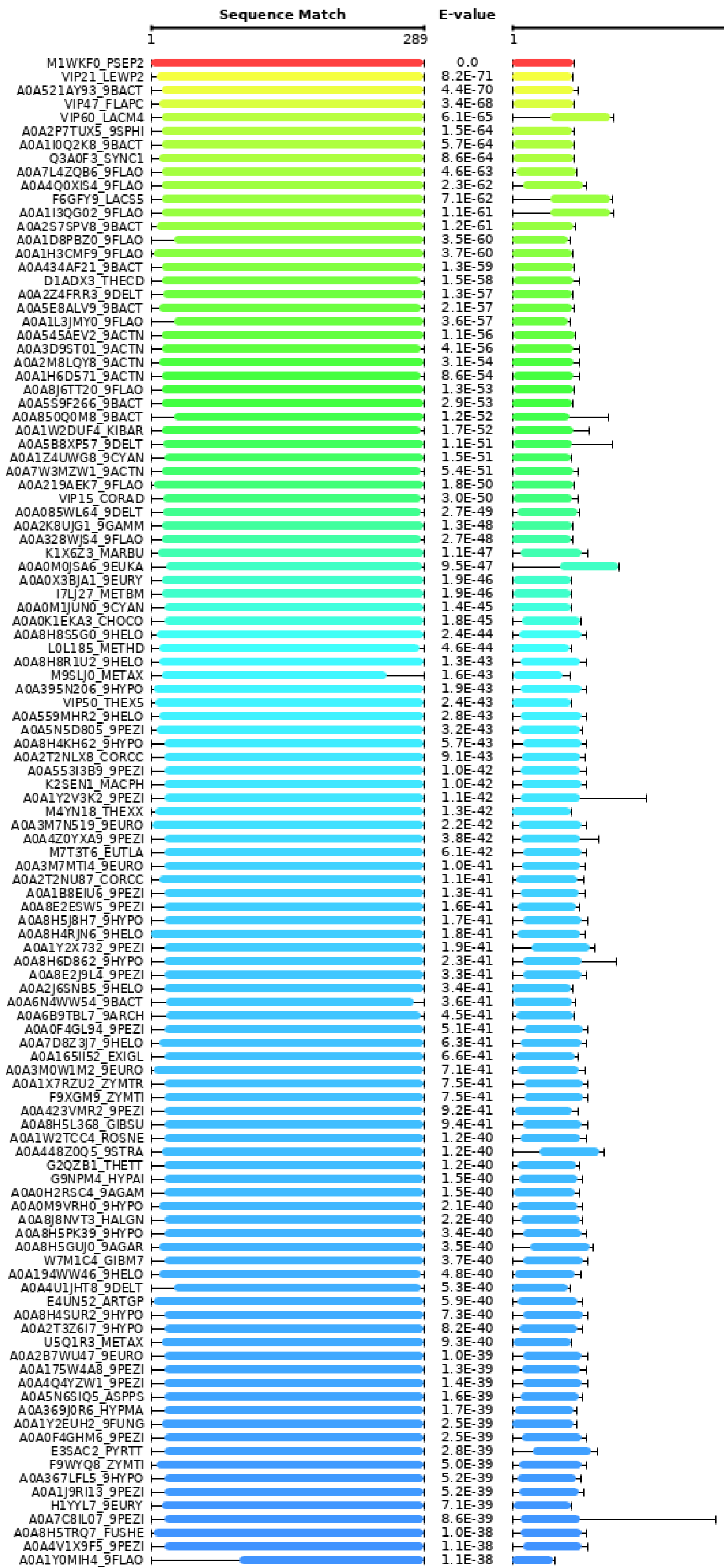
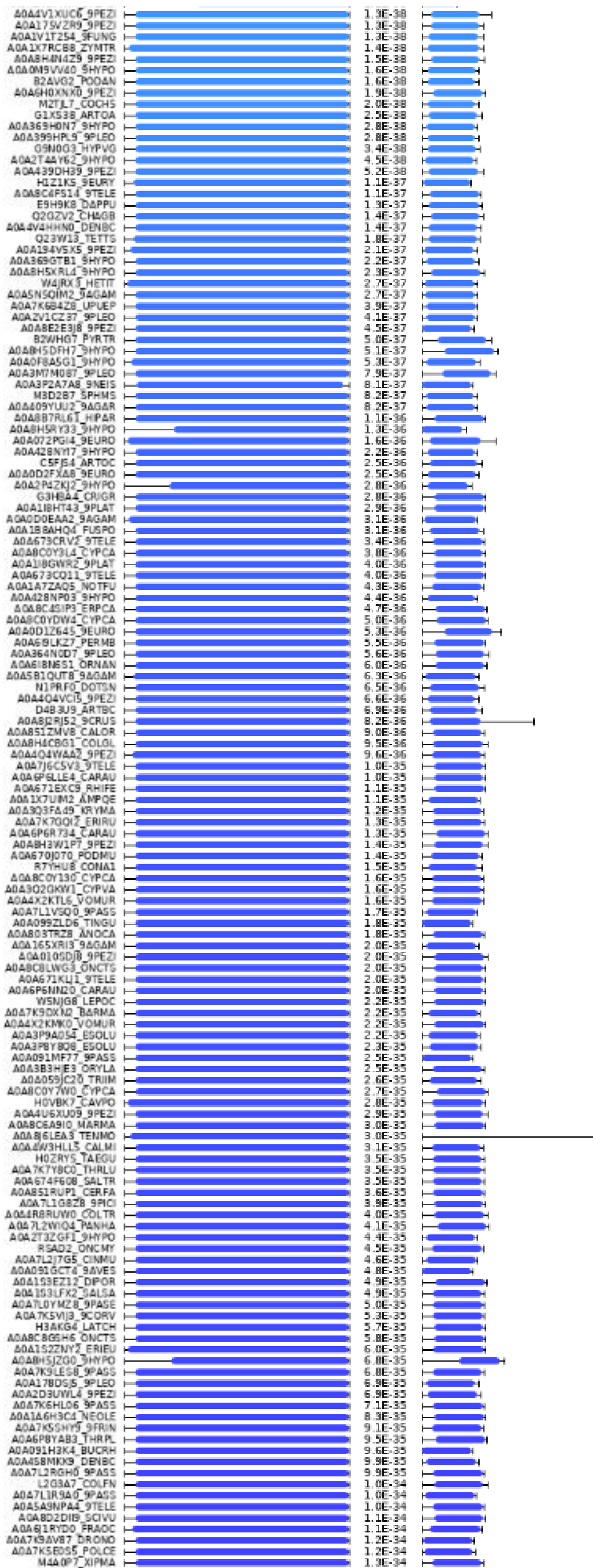


Fig. S6. SDS-PAGE analysis of purified PpSAND. Pre-casted polyacrylamide gels (Any kD^{TM}) (Bio-Rad Laboratories) were used and the protein marker was Protein^{TM} All Blue Standard 10-250kDa (Bio-Rad Laboratories).





E-value 0.0 8.2E-72 3.265E-53 6.515E-44 1.3E-34

European Bioinformatics Institute 2006-2020. EBI is an Outstation of the European Molecular Biology Laboratory.

Fig. S7. Amino acid blast analysis suggests that several uncharacterised homologues of PpSAND exist in different organisms.

References

[1] K. H. Ebrahimi, S. B. Carr, J. McCullagh, J. Wickens, N. H. Rees, J. Cantley, F. Armstrong, *FEBS Lett.* **2017**, *591*, 2394–2405.

[2] K. H. Ebrahimi, J. Rowbotham, J. McCullagh, W. S. James, *ChemBioChem* **2020**, *21*, 1605–1612.