Interim Progress Report for CDFA Agreement Number 12-0130-SA.

Title of Project:
Building a next generation chimeric antimicrobial protein to provide rootstock-mediated resistance to Pierce's Disease in grapevines

Principal investigator (PI)
Abhaya M. Dandekar, Plant Sciences Department; UC Davis, Davis, CA 95616
Phone: 530-752-7784; fax: 530-752-8502; Email: amdandekar@ucdavis.edu

Co-Principal Investigator (Co-PI)

Cooperators
Hossein Gouran, Plant Sciences Department; UC Davis, Davis, CA 95616
Phone: 530-752-5325; fax: 530-752-8502; Email: hgouran@ucdavis.edu

My Phu, Plant Sciences Department; UC Davis, Davis, CA 95616
Phone: 530-752-5325; fax: 530-752-8502; Email: mlphu@ucdavis.edu

Sandeep Chakraborty. Biological Sciences Department, Tata Institute of Fundamental Research, Mumbai India 400005. Phone: 011-91-2222-78-2606, Fax: 011-91-2222-80-4610; Email: sanchak@gmail.com

B.J. Rao, Biological Sciences Department, Tata Institute of Fundamental Research, Mumbai India 400005. Phone: 011-91-2222-78-2606, Fax: 011-91-2222-80-4610; Email: bjrao@tifr.res.in

Ana M. Ibáñez, Plant Sciences Department; UC Davis, Davis, CA 95616
Phone: 530-752-5325; fax: 530-752-8502; Email: amibanez@ucdavis.edu

Time period:
5 months (10/01/2013 to 02/28/2014)

Introduction
Xylella fastidiosa (Xf), the causative agent of Pierce's Disease, has a complex lifestyle requiring colonization of plant and insect. Its growth and developmental stages include virulence responses that stimulate its movement *in planta* and its ability to cause disease in grapevines (Chatterjee et al., 2008). Resistance to this pathogen must be multifaceted to block the pathogen at different stages of its complex lifestyle. A key issue is the reservoir of bacterial inoculum already present in California that poses an immediate threat in the presence of a significant insect vector like the GWSS. Chemical pesticides are now used to suppress the GWSS population, which is effective but does not reduce this inoculum reservoir. Resistance mechanisms capable of degrading the reservoir could prevent further spread of the disease. It is critical to know whether any resistance mechanism under consideration can clear *Xf* and if so, by what mechanism. The resistance mechanism must limit spread and movement of the bacterium *in planta* and block transmission of the disease by insect vectors. We have previously shown that *Xf* exposed to xylem fluid from resistant lines expressing NE-CB shows significant mortality. Our group has successfully designed and tested a NE-CB chimeric protein that specifically targets *Xf* in plant xylem (the site of infection), rapidly clears the pathogen, and blocks infection (Dandekar et al., 2009, 2012; Kunkel et al., 2007). The protein contains two separate domains. A surface binding domain recognizes outer membrane proteins; we have previously shown that it recognizes and cleaves mopB, a major *Xf* outer membrane protein (Dandekar et al., 2012). This surface recognition domain is encoded by a synthetic gene derived from the human innate defense protein neutrophil elastase (NE) (Dandekar et al., 2012; Kunkel

et al., 2007). The second, CB domain is a clearance domain, connected with a flexible linker to the C-terminal of NE. This domain is a synthetic gene that encodes an antimicrobial peptide, cecropin B (CB), that specifically lyses Gram-negative bacteria like *Xf* (Andrès and Dimarcq, 2007). The two domains work in tandem to recognize and lyse *Xf*. Our current hypothesis for the mode of action is that NE binds to the surface of *Xf* via its mopB outer membrane protein, bringing the cecropin B peptide close to the bacterial surface where it can lyse and kill the pathogen. Transgenic expression of this protein in tobacco and grape has provided phenotypic evidence of bacterial clearance and biochemical evidence of mopB degradation by NE (Dandekar et al., 2012). A major concern is that the presence of a protein of human origin in grapevines is potentially controversial to groups opposed to GMOs. Therefore, substituting NE and CB proteins derived from plants, ideally from grapevine, would be less controversial.

List of objectives

The goal of this project is to redesign our existing therapeutic NE-CB CAP, replacing the human NE and insect CB domains with plant/grapevine orthologs, and to validate the efficacy of the new CAP components in providing resistance to Pierce's Disease in grapevines. We are now following the goals, objectives and activities as proposed in the revised proposal submitted last year (2013) and approved for 2 years.

Goal: Redesign the NE-CB chimeric antimicrobial with a plant elastase and plant-derived lytic domain and test its efficacy to combat Pierce's Disease in transgenic tobacco and grapevines.

Objective 1: Redesign the chimeric antimicrobial protein by substituting a plant counterpart (plant elastase or PE) for the human neutrophil elastase (NE) component and demonstrate its efficacy for bacterial clearance.

Activity 1: Identify a suitable plant elastase candidate that is comparable to human neutrophil elastase in active site structure using the CLASP computational tool.

Activity 2: Construct vectors and test *in planta*-produced protein for efficacy in killing *Xf* in culture.

Activity 3: *In planta* efficacy testing: construct binary vectors (PE-CB), transform grapevine and tobacco, and test transgenic tobacco and grapevine for clearance of *Xf* and resistance to Pierce's Disease symptoms.

Objective 2: Redesign the chimeric antimicrobial protein by substituting a plant/grapevine counterpart (plant lytic domain or PLD) for the insect-derived Cecropin B (CB) component in the lytic domain and demonstrate its efficacy for bacterial clearance.

Activity 4: Identify a suitable PLD candidate that is comparable to insect-derived Cecropin B in primary and secondary structure using CLASP and other computational tools.

Activity 5: Test synthetic PLD protein for efficacy in killing *Xf* in culture.

Activity 6: *In planta* testing of the efficacy of grape-derived CAP components using transient expression of *Vs*P14-LPD and *Vs*P14m-LPD.

Objective 3: Construct and test a fully plant-derived CAP (PE-PLD) and test its ability to confer resistance to Pierce's Disease in grapevine rootstocks.

Activity 7: Construct a PE-PLD binary vector, transform grapevine and tobacco, and evaluate *Xf* resistance and Pierce's Disease development.

Description of activities conducted to accomplish each objective, and summary of accomplishments and results for each objective

Objective 1: Redesign the chimeric antimicrobial protein by substituting a plant counterpart (plant elastase or PE) for the human neutrophil elastase (NE) component and demonstrate its efficacy for bacterial clearance.

Since the CAP components work synergistically we will replace them one at a time maintaining the other original component. In this time period we focused our efforts in replacing the human neutrophil elastase (NE) with a plant/grapevine version of NE (PE). In order to do this we need to discover an appropriate protein in plants and possibly more appropriately in grapevine that has the same activity as NE.

Activity 1: Identify a suitable plant elastase candidate that is comparable to human neutrophil elastase in active site structure using the CLASP computational tool.

This activity has been successfully accomplished. A plant PE candidate protein was identified using the CLASP (Chakraborty et al., 2011) and PROMISE packages (http://www.sanchak.com/clasp/). The potential differences between reactive atoms in the active site of the target proteins were calculated using the Adaptive Posson-Boltzmann Solver (APBS) and PDB2PQR packages (Baker et al., 2001 and Dolinsky et al., 2004). The NE structure (1B0F) was chosen from Protein Data Bank (PDB). The elastase scaffold residues Ser195, His57, Asp102, Ser214, and Gly193 were chosen based on their known interaction with the elastase enzymatic function. The best plant candidate based on this analysis was a pathogenesis related (PR) protein P14a from tomato (1CFE). Details of the protocol and workflow used to make this selection have been recently described (Chakraborty et al., 2013). The P14a from *Solanum lycopersicum* (tomato) is a promising choice since we found similar, highly conserved proteins from other plant species, including grapevine. Additionally, the PR superfamily is widely distributed in animals, plants, and fungi. Some of these related to CRISP have demonstrated, substrate-specific protease activity (Milne et al., 2003). Furthermore, a striking structural homology was observed between P14a and a protein found in snake venom that has been demonstrated to be an elastase (Bernick and Simpson 1976). The structure of the 135 amino acid mature P14a protein from tomato was also determined using NMR (Fernandez et al., 1997). Based on these considerations, we used the tomato P14a to find an appropriate P14a in *Vitis*. We chose, from several matching these criteria, the P14a from *Vitis shuttleworthii* (Vs), as this species is resistant to PD (Walker, personal communication). More recently, 21 different PR-1 genes from grapevine, including those from Vs, were characterized and shown to confer resistance to bacterial disease (Li et al., 2011). Further analysis of the predicted active site of *Vs*P14a using CLASP indicated that removal of amino

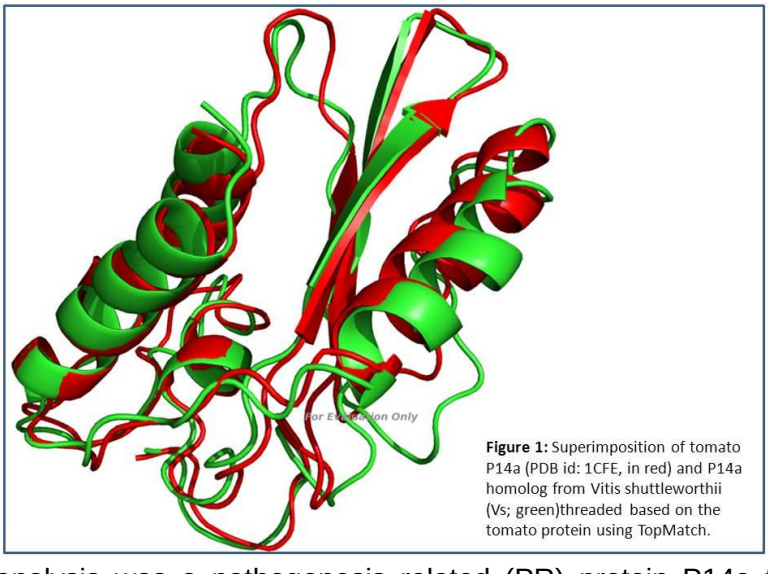

Figure 1: Superimposition of tomato P14a (PDB id: 1CFE, in red) and P14a homolog from Vitis shuttleworthii (Vs; green)threaded based on the tomato protein using TopMatch.

acids 74 and 77 had the potential to improve elastase/protease activity. The sequence of VsP14a without the two amino acids is designated VsP14m.

Activity 2: Construct vectors and test *in planta*-produced protein for efficacy in killing *Xf* in culture.

To test the efficacy of *Vs*P14a and *Vs*P14m proteins in the clearing of *Xf,* we codon-optimized and chemically synthesized the two versions of VsP14a after adding a 3xFlag purification tag (Sigma Aldrich). VsP14a-3xFlag and VsP14m-3xFlag genes were cloned into pEAQ-HT, a binary vector specific for transient protein expression system in the tobacco species *N. benthamiana* (Sainsbury et al., 2009) (Fig 2A). The constructed binary vectors were transformed into a super-virulent *Agrobacterium* strain (EHA105). As a negative control, an empty vector (pEAQ-HT) was also transformed into the same strain of *Agrobacterium*. Tobacco leaves were harvested from greenhouse-grown plants and vacuum-infiltrated with *Agrobacterium* suspensions containing *Vs*P14a and the empty vector. Agro-infiltration conditions were optimized as described earlier (Huang et al., 2009). The infiltrated leaves were harvested six days post-infiltration. Total protein was extracted using an apoplastic wash method to extract secreted proteins present in the leaf apoplast. Using Anti-Flag M2 antibody (Sigma), we detected *Vs*P14a on the first run. We tested the ability of the crude leaf protein extract from plants transformed with both constructs to inhibit growth of *E.coli* (Fig 2C) and *Xylella* (Fig 2D). After four hours, the *E.coli* with the extract from the plant expressing the empty vector showed growth, while growth of those exposed to VsP14A was completely inhibited (Fig 2C). Proteins VsP14a and VsP14m also inhibited the growth of *Xylella* while protein obtained from the empty vector did not. By 50 minutes, 50% of the population was killed, but mortality reached a plateau at ~ 60% mortality (Fig 2D). These results are encouraging and indicate that we may have found the desired replacement for NE. Now it must be tested *in planta* for protection against *Xf* infection. We also developed antibodies against VsP14a and used these to detect the protein in the plants using Western blots. Proteins isolated from Thompson Seedless (TS) and *Vitis shuttleworthii* infected and not infected with *Xylella* showed that the P14a in *V. shuttleworthii* was expressed in both infected and uninfected tissues. In TS, it was only expressed in infected tissues and un-infected tissues expressed no protein corresponding to P14a (Fig 2B).

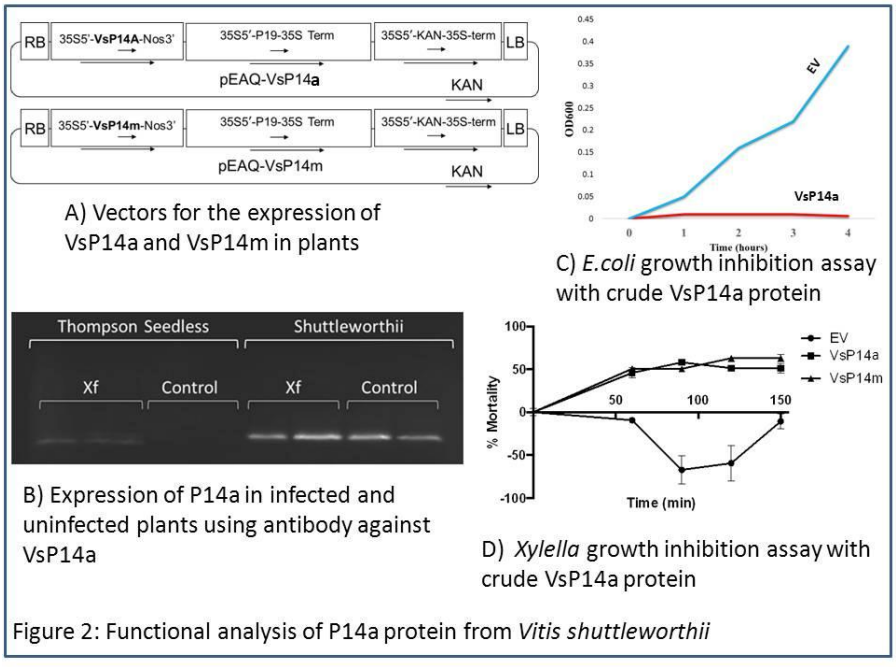

Figure 2: Functional analysis of P14a protein from *Vitis shuttleworthii*

Activity 3: *In planta* efficacy testing: construct binary vectors (PE-CB), transform grapevine and tobacco, and test transgenic tobacco and grapevine for clearance of *Xf* and resistance to Pierce's Disease symptoms.

We have completed the construction of two binary vectors, one for the expression of VsP14a by itself (Fig 3) and the other that links the VsP14a sequence to CB. This recreates the original CAP protein but with the SRD domain being swapped with the P14a protein (Fig 3). In the first construct the P14a coding sequence has been fused to a signal peptide from the Ramy3D protein which is cleaved upon expression in the plant. Next to the signal peptide cleavage site and at the N-terminal of P14a is a 3XFLAG sequence to improve the immune detection of the P14a protein. In the second construct the coding sequence is fused to the CB sequence via the 4 amino acid flexible linker sequence used in the original CAP design (Dandekar et al., 2012). Both coding regions have a TMV omega sequence in the 5' non-coding region to improve translation efficiency. The regulatory sequence in both constructs is a double CaMV35S promoter sequence. The binary vectors have been introduced into a disarmed *Agrobacterium* strain in order to reconstitute a functional plant transformation system. Both vectors are being used to transform grapevine rootstock and SR1 tobacco to evaluate the efficacy of these two constructs.

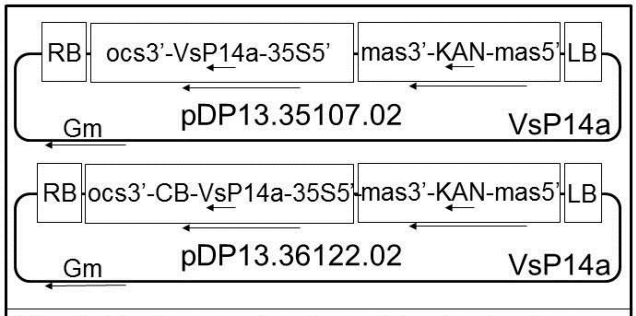

Fig. 3: Vector construct used for *in planta* efficacy testing of VsP14a in tobacco and grapevine plants

Objective 2: Redesign the chimeric antimicrobial protein by substituting a plant/grapevine counterpart (plant lytic domain or PLD) for the insect-derived Cecropin B (CB) component in the lytic domain and demonstrate its efficacy for bacterial clearance.

The goal of this objective is to identify a CB like protein in plants to replace the lytic domain of CAP described earlier (Dandekar et al., 2012). Unlike the search for the HNE where the focus was the congruence of the active site (spatial and electrostatic) an approach which cannot be used for CB as it does not have an enzymatic function. The approach for looking for a CB like in protein in plants focused instead on the highly structured nature of CB during the search process.

Activity 4: Identify a suitable PLD candidate that is comparable to insect-derived Cecropin B in primary and secondary structure using CLASP and other computational tools.

We have used an approach similar to that described above in Activity 1 to identify a replacement component for CB. However, instead of comparing the reactive atoms as was done for the HNE matching algorithm, in this case we match for the Cα atoms of specific residues based on the overall shape of CB. Thus, we adopted a slightly different computational flow in this case. First, we did a keyword search, 'plants' in http://www.pdb.org/ that yielded about 1000 proteins. Each pdb was expanded on basis of each chain. For example, PDB X.pdb with chains A and B resulted in PDB files XA.pdb and XB.pdb. This list was filtered based on a 80% similarity redundancy. Then

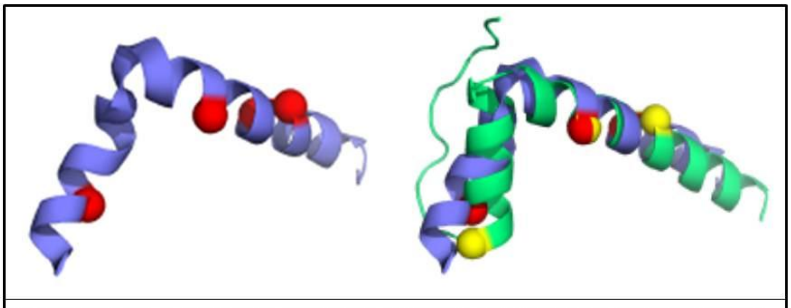

Fig 4: Cecropin B structure: (A) native structure (PDBid: 2IGR) showing chosen motif (Lys10, Lys11, Lys16 and Lys29). (B) Superimposition of Cecropin B (PDBid: 2IGR, in blue) with Plasma membrane H+ ATPase (PDBid:2O98, in green) showing structural similarity

proteins larger than 60 aa in length were ignored. APBS was run on each pdb (the ones that failed electrostatic analyses, a few, were excluded). We were able to extract 200 proteins which we then analyzed further using CB motif. The CB structure is characterized by two helices (Fig 4A). We chose specific residues from these two helices such that the residues were polar, and had stereochemical matches. For example, Lys can be replacing by either Arg or His, and thus constitutes a good candidate. We chose, Lys10, Lys11, Lys16 and Lys29 as the input motif from CB (PDBid:2IGR; Fig 4A), allowing Lys to matched by Lys, Arg or His.

Table 1: Best matches in set of approximately 200 proteins when queried using the motif (Lys10, Lys11, Lys16 and Lys29) from CB (PDBid:2IGR)

PDB	**Chain**	**Length**	**Description**
1GPT	A	47	GAMMA-1-H THIONIN
2KGQ	A	53	Brazzein
1N4N	A	47	floral defensin-like protein 1
2O98	P	52	Plasma membrane H+ ATPase
1MR4	A	47	Nicotiana alata plant defensin 1 (NaD1)
1KRL	A	44	MONELLIN, CHAIN A
2DS2	A	33	Sweet protein mabinlin-2 chain A
2HUG	A	57	Signal recognition particle 43 kDa protein, c

Our analysis has identified good candidates (Table 1). It is noteworthy that several defense-like proteins feature in the list, allowing us to speculate that another peptide might be a good replacement for CB in the chimera. We have evaluated eight protein candidates. We are now focusing on a 52 amino acid segment of the plasma membrane H^+ ATPase, whose structure matches very well with the CB structure (Table 2; Fig 4B).

Table 2: Potential and spatial congruence of `duplicate' scaffolds to the active residues in CB (PDBid:1EH5) and Plasma membrane H+ ATPase (PDBid: 2O98, Chain P) : D = Pairwise distance in Angstroms. PD = Pairwise potential difference.

PDB	**Active site atoms (a,b,c,d)**		**ab**	**ac**	**ad**	**bc**	**bd**	**cd**
2IGR Chain A	Lys10,Lys11,Lys16,Lys29, Cecropin B	D	3.8	9.9	22.1	8.7	18.9	15.3
		PD	8.0	-0.5	84.2	-8.4	76.2	84.7
2O98 Chain P	Arg919,Arg920,Arg925,Lys943, Plasma membrane H+ ATPase	D	3.8	9.6	23.8	8.4	21.0	17.2
		PD	-20.3	29.3	56.8	49.6	77.0	27.4

Activity 5: Test synthetic PLD protein for efficacy in killing *Xf* in culture.

Since the 52 aa H+ATPase is large so we cannot have it cheaply synthesized as a protein. We are developing additional computational tools to better evaluate individual alpha helical domains so we can identify smaller proteins that can be more cheaply synthesized for testing. We have begun developing these tools, expanding activity 4.

Activity 6: Conduct *in planta* efficacy testing of the grape-derived CAP components using transient expression of *Vs*P14-LPD and *Vs*P14m-LPD.

We have begun the construction of binary vectors for the expression of VsP14a-PLD, in order to test the efficacy of 52 aa H+ATPase sequence described above.

Objective 3: Construct and test a fully plant-derived CAP (PE-PLD) and test its ability to confer resistance to Pierce's Disease in grapevine rootstocks.

Activity 7: Construct a PE-PLD binary vector, transform grapevine and tobacco, and evaluate *Xf* resistance and Pierce's Disease development.

Publications produced and pending, and presentations made that relate to the funded project.

Chakraborty S., R. Minda, L. Salaye, A.M. Dandekar, S.K. Bhattacharjee and B.J. Rao. 2013. Promiscuity-based enzyme selection for rational directed evolution experiments. In, J. Samuelson (ed.), "Enzyme Engineering: Methods and Protocols". Pub: Springer New York. Methods in Molecular Biology. 978: 74-78.

Research relevance statement, indicating how this research will contribute towards finding solutions to Pierce's disease in California.

Xylella fastidiosa (Xf), the causative agent of Pierce's Disease, has a complex lifestyle requiring colonization of plant and insect. Its growth and developmental stages include virulence responses that stimulate its movement *in planta* and its ability to cause disease in grapevines (Chatterjee et al., 2008). Thus, any control or resistance measure must by necessity be multifaceted to block this pathogen at different stages in its complex lifestyle. A key issue for the industry is the reservoir of bacterial inoculum already present in California that poses an immediate threat in the presence of a significant insect vector like the GWSS. Chemical pesticides are now used to suppress the GWSS population, which is effective but does not reduce this reservoir of bacterial inoculum. Resistance mechanisms must be directed to degrade this inoculum and prevent the further disease spread. It is critical to know whether any resistance mechanism under consideration can clear *Xf* and if so, by what mechanism. The resistance mechanism must limit spread and movement of the bacterium *in planta* and block transmission of the disease by insect vectors. We have previously shown that *Xf* exposed to xylem fluid from resistant lines expressing NE-CB shows significant mortality (Dandekar et al., 2012). Transgenic grapevine lines expressing the NE-CB chimeric protein show very reduced or no PD symptoms: less xylem blockage and leaf scorching. Currently, some of these lines are bring field tested in two locations, one with natural PD pressure due to an abundant supply of Glassy Winged Sharp Shooters (GWSS) and one without GWSS but field inoculated with *Xf* bacteria. In this proposal, we seek to swap the human NE domain with an equivalent protein from a plant source and confirm whether it functions as effectively as the human protein. This addresses concerns about expressing a human protein in grapevines.

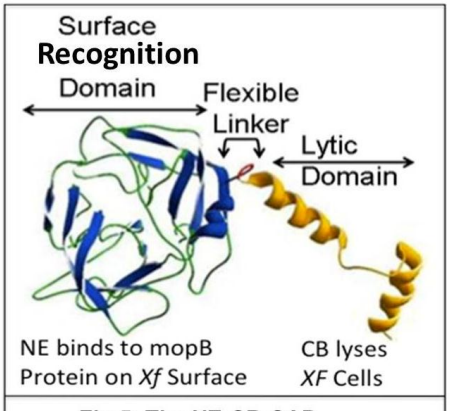

Fig 5: The NE-CB CAP

Layperson summary of project accomplishments

We have successfully investigated the replacement of the surface recognition domain (SRD) of our NE-CB chimeric antimicrobial protein with a protein from grapevine. Using the recently described computational tool (CLASP, Chakraborty et al., 2011) we identified the VsP14a protein from grapevine *Vitis shuttleworthii* based on its structural and conformational similarity with the active site of the human neutrophil elastase (NE domain). In order to verify the biological activities of this protein we have constructed a gene cassette to express and produce this VsP14A protein in plants. This step has also been accomplished and we have been able to make this protein in tobacco leaf extracts and successfully test this protein to evaluate its ability to inhibit the growth of *E.coli* a gamma-proteobacteria and Xylella. We found that addition of the VsP14a protein to the media inhibits the growth of E.coli. Results with Xylella show that it is able to clear Xylella cells when viewed under a microscope. We have identified a grapevine candidate to replace the CB component using computational tools.

Status of funds.
All of the funds for the first year have been expended. We have been funded for years 2 and 3 for which we have begun spending for year 2 starting on July 1, 2013.

Summary and status of intellectual property associated with the project
We have not made any intellectual property disclosures to date.

Literature cited

Baker, N.A., D. Sept, S. Joseph, M.J. Holst and J.A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA 98:10037–10041.

Bernick, J.J., and W. Simpson.1976 Distribution of elastase-like enzyme activity among snake venoms. Comp Biochem Physiol B 54: 51–54

Chakraborty, S., R. Minda, L. Salaya, S.K. Bhattacharjee and B.J. Rao. 2011. Active site detection by spatial conformity and electrostatic analysis-unravelling a proteolytic function in shrimp alkaline phosphatase. PLoS ONE 6(12): e28470. Doi:10.1371/journal.pone.0028470.

Chakraborty, S., and B.J. Rao. 2012. A measure of the promiscuity of proteins and characteristics of residues in the vicinity of the catalytic site that regulates promiscuity. PLoS ONE 7(2): e32011.

Chakraborty S., R. Minda, L. Salaye, A.M. Dandekar, S.K. Bhattacharjee and B.J. Rao. 2013. Promiscuity-based enzyme selection for rational directed evolution experiments. In, J. Samuelson (ed.), "Enzyme Engineering: Methods and Protocols". Pub: Springer New York. Methods in Molecular Biology. 978: (In Press).

Chatterjee, S., R.P.P. Almeida, and S.E. Lindow. 2008. Living in two worlds: the plant and insect lifestyles of *Xyllela fastidiosa*. Annual Review of Phytopathology 46:243-271.

Dandekar, A.M., H. Gouran, A.M. Ibanez, S.L. Uratsu, C.B. Agüero, S.McFarland, Y. Borhani, P.A. Fieldstein, G.E. Bruening, R. Nascimento, L. Goulart, P.E. Pardington, A. Choudhary, M. Norvell, R. Civerolo and Goutam Gupta. 2012. An engineered innate immune defense protects grapevines from Pierce's Disease. Proc. Nat. Acad. Sci. USA 109(10): 3721-3725.

Dolinsky, T.J., J.E. Nielsen, J.A. McCammon and N.A. Baker. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res 32: W665–W667

Fernandez, C., T. Szyperski, T. Bruyere, P. Ramage, E. Mosinger and K. Wuthrich. 1997. NMR solution structure of the pathogenesis-related protein P14a. J. Mol. Biol. 266: 576-593.

Huang, T-K., M.A. Plesha, B.W. Falk, A.M. Dandekar and K.A. McDonald. 2009. Bioreactor strategies for improving production yield and functionality of a recombinant human protein in transgenic tobacco cell cultures. Biotechnol. Bioeng. 102: 508-520.

Li, Z.T., S.A Dhekney and D.J. Gray. 2011. PR-1 gene family of grapevine: a uniquely duplicated PR-1 gene from Vitis interspecific hybrid confers high level resistance to bacterial disease in transgenic tobacco. Plant Cel Rep. 30:1-11.

Milne, T.J., G. Abbenante, J.D. Tyndall, J. Halliday and R.J. Lewis. 2003. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. J Biol Chem 278: 31105-31110.

Sainsbury, F., E.C. Thuenemann, G.P. Lomonossoff. 2009. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol J. 7: 682-693.