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 (03/01/2014to 07/31/2014)

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.

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) package (http://www.sanchak.com/clasp/). Details of the protocol and workflow used to make this selection have been described (Chakraborty et al., 2013). The P14a from *Solanum lycopersicum* (tomato) was the protein of 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. We chose, from several matching 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).

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

To test the efficacy of *Vs*P14a in the clearing of *Xf,* we codon-optimized and chemically synthesized VsP14a after adding a 3xFlag purification tag (Sigma Aldrich) and cloned into an an expression vector pEAQ-HT in addition into a binary vector for transient protein expression in tobacco (*N. benthamiana;* Sainsbury et al., 2009). Total protein was extracted using an apoplastic wash method and tested the ability inhibit growth of *E.coli* and *Xylella*. After four hours, the *E.coli* with the extract from the plant expressing the empty vector showed growth, while growth of those exposed to VsP14 was completely inhibited. VsP14a 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. These results are encouraging and indicate that we have found the desired replacement for NE.

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 1) 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 1). 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.

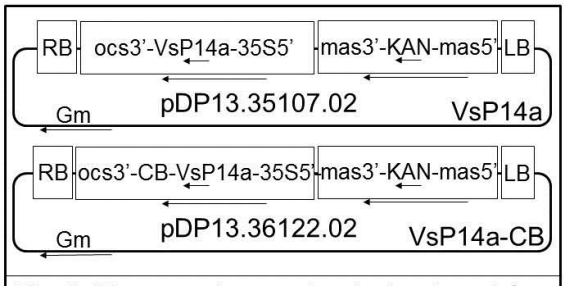

Fig. 1: Binary vector constructs developed for *in planta* efficacy testing of VsP14a based CAP proteins in tobacco and grapevine plants

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), a similar approach 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.

Initially we 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 the 4 Lys residues, Lys10, Lys11, Lys16 and Lys29 as the input motif from CB (PDBid:2IGR), allowing Lys to matched by Lys, Arg or His. As indicated in our last report this strategy revealed several good candidate proteins. We are now focusing on VsHAT52 a 52 amino acid segment of the plasma membrane H^+ ATPase, whose structure matches very well with the CB. However, we have reexamined our approach and focused instead on the alpha helical structure itself in order to generate a greater diversity of candidates that could serve as replacements for the CB lytic peptide domain of our CAP previously described (Dandekar et al., 2012). In order to understand better the functionality of the alpha helical domains of CB we have developed two new computational tools PAGAL and SCALPEL to better predict the antimicrobial activities in

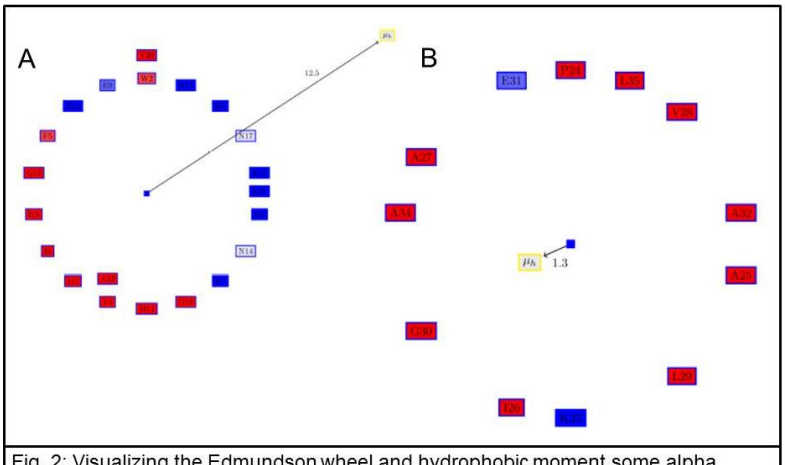

Fig. 2: Visualizing the Edmundson wheel and hydrophobic moment some alpha helices: The color coding is as follows: all hydrophobic residues are colored red, while hydrophilic residues are colored blue - dark blue for positively charged residues, medium blue for negatively charged residues and light blue for amides. (A) N-Terminal helix of CB. (B) C-Terminal helix of CB.

portions of existing proteins. PAGAL (**P**roperties **a**nd corresponding **g**raphics of **al**pha helical structures in proteins) in an open source software that implements previously known and established methods of evaluating the properties of alpha helical structures, providing very useful information of the amphipathicity, hydrophobicity and charge moments within these structures. We have successfully used PAGAL to search 4000 plant proteins in the PDB database to create a database that contains a listing of the properties of each and every alpha helical structure present within these proteins. A key feature of lytic peptides is the distribution of hydrophobic and charged residues on the surface of the protein. For proteins like cecropin that have an alpha helical structure we used PAGAL to evaluate CB. CB contains two α helixes (AHs), joined by a short stretch of random coil. Fig 2A and 2B shows the Edmundson wheel and hydrophobic moment of the two AHs. It can be seen that the N-Terminal AH has a large hydrophobic moment, as well as a specific positive charge distribution. This can also be seen in a Pymol rendering of the peptide surface (Fig 3). The Pymol script for this rendering is automatically generated by PAGAL. On the other hand, the C-Terminal AH of CB does not have either of the above two properties.

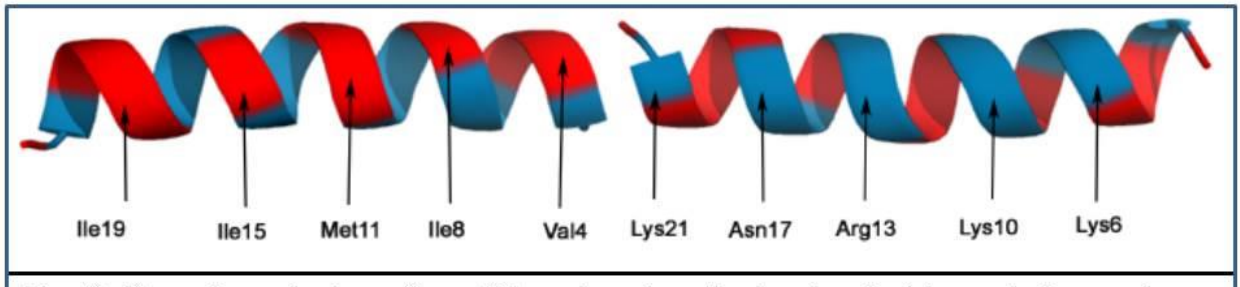

Fig. 3: Pymol rendering of peptides showing the hydrophobic and charged surfaces for the N-terminal helix of CB. All hydrophobic residues are colored red, while hydrophilic residues are colored blue.

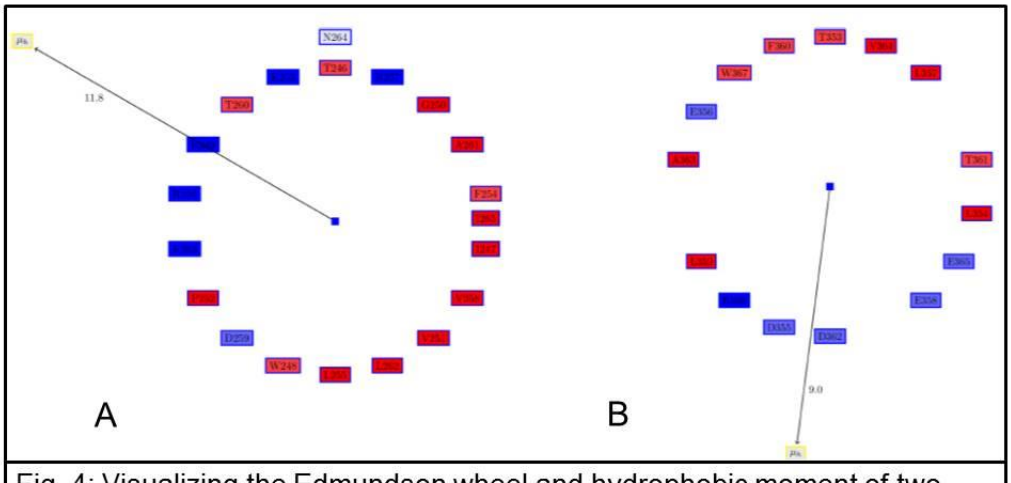

Fig. 4: Visualizing the Edmundson wheel and hydrophobic moment of two Vitis alpha helices: The color coding is as follows: all hydrophobic residues are colored red, while hydrophilic residues are colored blue - dark blue for positively charged residues, medium blue for negatively charged residues and light blue for amides. (A) The 20 aa residue PPC20 with positively charged residues. (B) 15 aa residue ISS15 with negatively charged residues.

We then developed the second program SCALPEL (**S**earch **c**haracteristic **alp**ha h**el**ical peptides in the PDB database and locate it in the genome) to search for alpha helical structures of a particular type. We searched for the smallest peptide with a large hydrophobic moment and a high proportion of positively charged residues on the hydrophilic side. Once we obtain a particular hit that has the right properties that we are interested in investigating we then use BLAST to find the corresponding protein from grapevine. In this case we identified a 20 aa region of phosphoenolpyruvate carboxylase, a key enzyme in the C4-photosynthetic carbon

cycle form grapevine that we call PPC20 (Fig 4A). We also searched for the smallest peptide with a large hydrophobic moment and a high proportion of negatively charged residues on the hydrophilic side. In this case we identified a 15 aa region of Isoprene Synthase form grapevine that we call ISS15 (Fig 4B). It can be visualized very clearly in Fig 4 that very large hydrophobic moment for both of these proteins. In Fig 5A one can the PYMOL rendering of the protein surface that shows the highly hydrophobic surface of PPC20 and in Fig 5B one can clearly see the positively charged surface with the exception of the single Asp that is the only negative residue in an positive surface. In Fig 5C one can the PYMOL rendering of the protein surface that shows the highly hydrophobic surface of ISS15 and in Fig 5D one can clearly see the negatively charged surface.

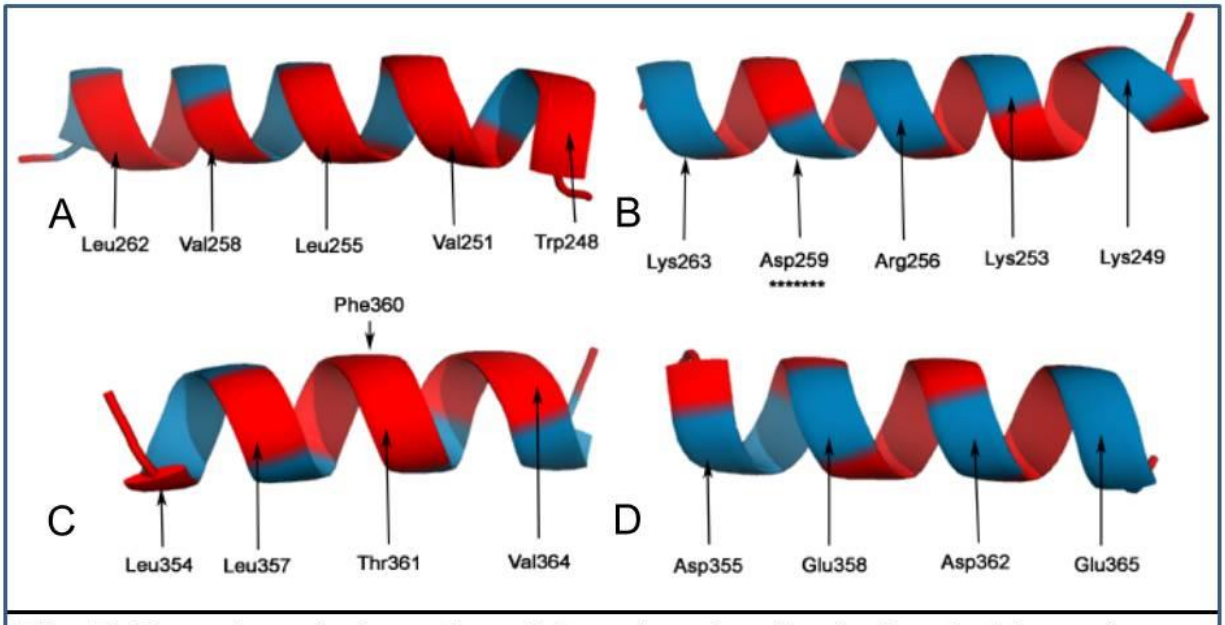

Fig. 5: Pymol rendering of peptides showing the hydrophobic and charged surfaces for the two potential lytic peptides extracted from grapevine genes. All hydrophobic residues are colored red, while hydrophilic residues are colored blue. (A) hydrophobic surface for PPC20 (B) charged surface for PPC20 (C) hydrophobic surface for ISS15 (D) charged surface for ISS15.

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

Using these two programs PAGAL and SCALPEL we have focused on our phase 1 search which is currently underway where we are examining all small proteins and will compare them to the N-terminal 21 aa domain of CB that we call CBNT21 (+ve charge). We determined using these two programs that this domain to be the most active of the two alpha helical domains of CB. We would predict would be the most active in terms of lytic properties. We identified three small proteins in citrus PPC20 (+ve charge very similar to CBNT21 in its structural properties); CsHAT22 (+ve charge smaller version of the 52 aa CsHAT that is currently being tested in plants), ISS15 (-ve charge, very small protein). We also will test CATH12 smallest defensin (12 aa) to identify how small a peptide we can use successfully against our pathogens of interest. All 5 proteins have been synthesized and are currently being tested for their lytic activity against *Xylella* to determine their range of clearance of this pathogen.

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

We have successfully the construction of binary vector for the expression of VsP14a-PLD, in order to test the efficacy of the 52 aa H+ATPase sequence described above (Fig 6). This needs to be incorporated in to an *Agrobacterium* host to create a function transformation system for both transient expression in tobacco and for the transformation of grapevines (Activity 7).

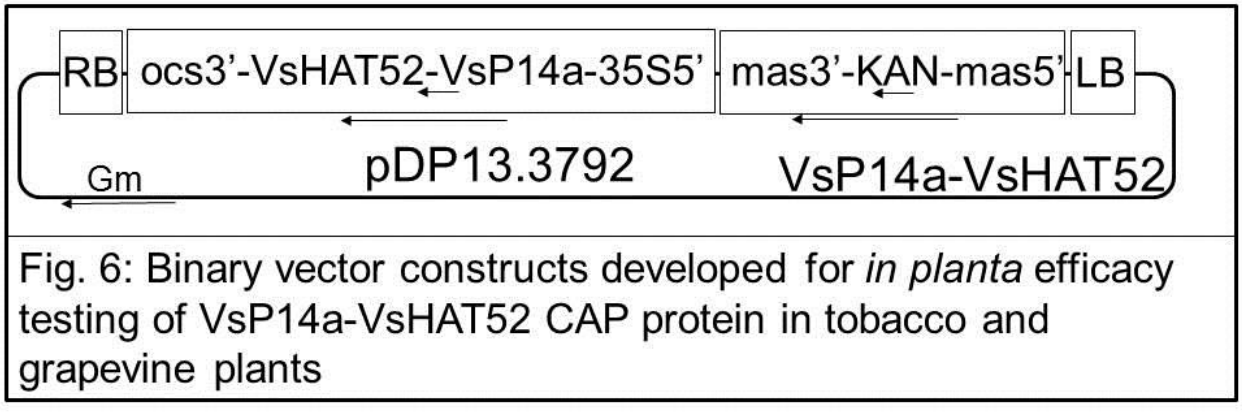

Fig. 6: Binary vector constructs developed for *in planta* efficacy testing of VsP14a-VsHAT52 CAP protein in tobacco and grapevine plants

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.

Dandekar, A.M. A.M. Ibanez, H. Gouran, M. Phu, B.J. Rao and S Chakraborty. 2012. Building the next generation chimeric antimicrobial protein to provide rootstock-mediated resistance to Pierce's Disease in grapevines. Pierce's Disease Research Reports, December 2012. pp. 89-93.

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.

Dandekar, A.M., H. Gouran, S Chakraborty, M. Phu, B.J. Rao and A.M. Ibanez,. 2013. Building the next generation chimeric antimicrobial protein to provide rootstock-mediated resistance to Pierce's Disease in grapevines. Proceedings of Pierce's Disease Research Symposium held December 16-18, 2013 at the Hiatt Regency Hotel, Sacramento, California. pp. 89-94.

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 of 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

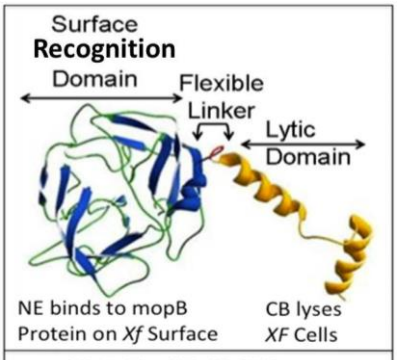

Fig 7: The NE-CB CAP

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 grapevines and confirm whether it functions as effectively as the human protein. This addresses concerns about expressing a human protein in grapevines.

Layperson summary of project accomplishments

We have successfully investigated the replacement of the surface recognition domain (SRD) of our NE-CB chimeric antimicrobial protein (Fig. 7) 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. We have enhanced the computational tools and will discover additional candidates in the grapevine genome that can be used to replace CB component of CAP.

Status of funds.

All of the funds for the second year have been expended. We have been funded for 3 years we will have begun spending for year 3 starting on July 1, 2014.

Summary and status of intellectual property associated with the project

We have made any intellectual property disclosures it UC Case No. 2014-776-1. For "Computational methods developed to identify alpha helical peptides that contain anti-microbial properties encoded in any specified genome. This invention provides the source code, and an example flow for choosing alpha helical anti-microbial peptides from plant genomes".

Literature cited

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.

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.