



Identification of the sortase enzyme and its substrates in *Renibacterium salmoninarum*:  
Solving problems with bioinformatics.

Samuel N. Crane

Undergraduate, University of Washington

Educational Appointee, Microbiology  
Northwest Fisheries Science Center, NMFS, NOAA



## Sortase and Its Substrates

- Attaches particular proteins to cell wall
- These cell wall proteins are associated with virulence
- Inhibition of sortase negatively impacts pathogenicity
- **Sortase and its substrates are targets for novel therapeutics.**

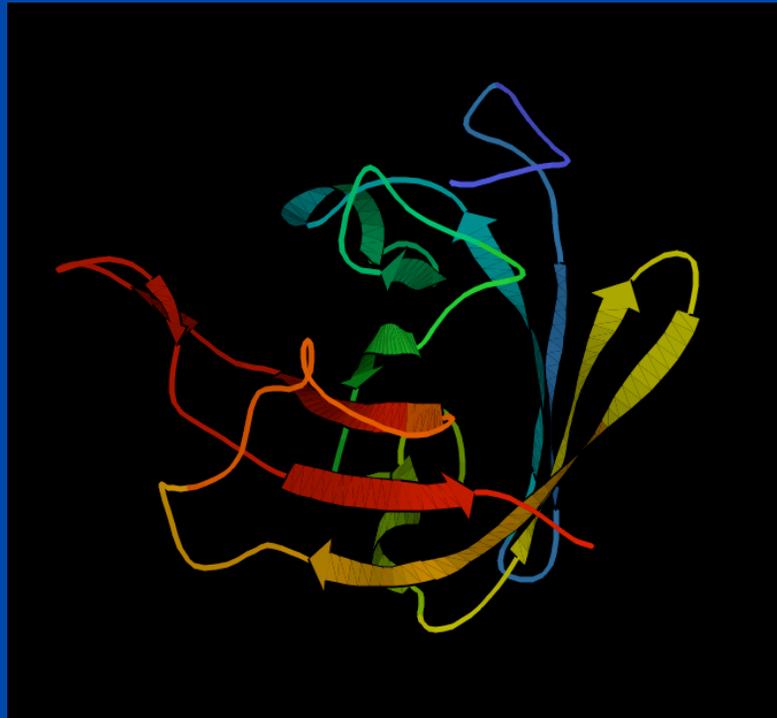
Is there at least one sortase gene in *Renibacterium salmoninarum*?  
If so, how many substrates are there in the genome?  
What are their functions?



- System characteristics
- Finding Sortase
- Finding Substrates
- Next Steps

# The Sortase Enzyme

- A cysteine peptidase
- Bound to the plasma membrane
- Binds select proteins to cell wall via transpeptidation



Ilangovan *et al* (2001) *PNAS* **98**: 6056-6061

# Sortase Substrates

## Sortase substrate

- Signal peptide sequence (SEC secretion pathway)
- Cleavage site
  - e.g. LPXTG
- C-terminal transmembrane region
- Charged tail





# Substrate Families

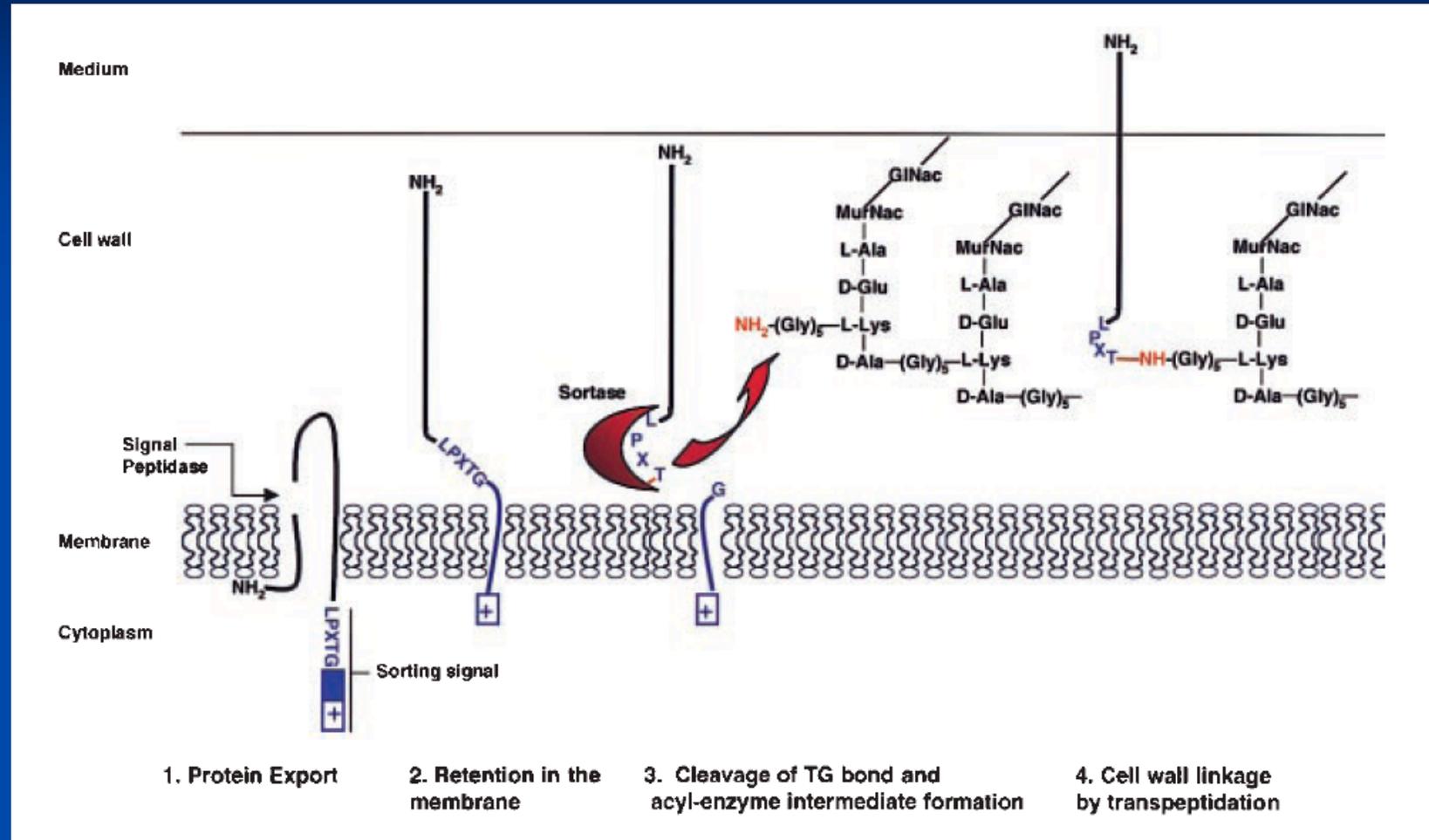
Differences in cleavage pattern exist

Following Comfort and Clubb 2004\*:

- **SrtA:** LPxTG-[EDS]
- **SrtB:** NP-[QK]-T-[NGS]-[DA]
- **Family 3:** [LI]-PxTGG
- **Family 4:** LPxTA-[TS]
- **Family 5:** LAxTG

\*Comfort and Clubb. (2004). *Infection and Immunity*. 72(5): 2710-2722.

# The Sortase Reaction



# Sortase



# Finding Sortase: Pairwise Comparisons

Search *R. sal.* genome for homologs to known sortase genes.

Organism	No. Sortase Genes	No. <i>R sal</i> Homologs
<i>Streptomyces coelicolor</i>	7	1
<i>Streptomyces avermitilis</i>	4	1
<i>Corynebacterium diphtheriae</i>	6	1
<i>Corynebacterium efficiens</i>	5	1
<i>Corynebacterium glutamicum</i>	1	1

# Finding Sortase: Hidden Markov Models

A HMM is a statistical model of a multiple sequence alignment.

*It is a statement about the probability of finding particular residues at particular positions within the alignment.*

Search the *R. sal.* genome with a HMM built from known sortases

Used a published sortase model:

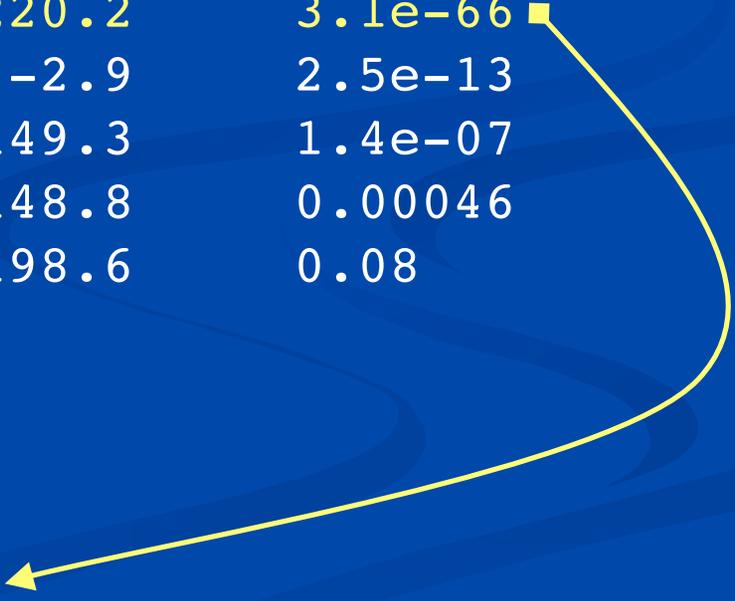
- [www.doe-mbi.ucla.edu/Services/Sortase/](http://www.doe-mbi.ucla.edu/Services/Sortase/)
- Comfort and Clubb. (2004). *Infection and Immunity*. **72**(5): 2710-2722.

# Finding Sortase: Hidden Markov Models

Scores for sequence family classification:

Model	Score	E-value
-----	-----	-----
family_5_sortases	220.2	3.1e-66
family_4_sortases	-2.9	2.5e-13
family_3_sortases	-149.3	1.4e-07
sortase_A_sortases	-148.8	0.00046
sortase_B_sortases	-198.6	0.08

LAXTG





# The *R. sal* Sortase

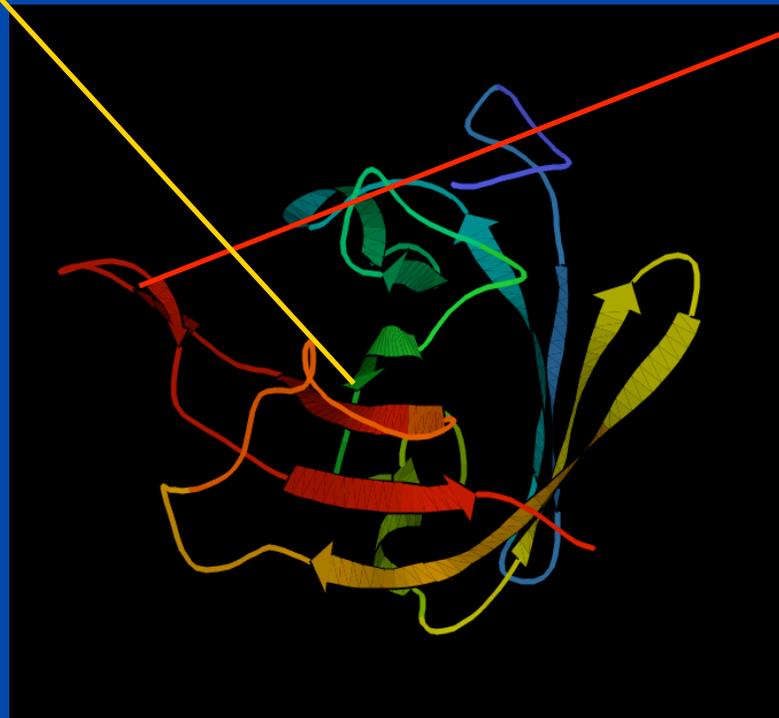
ERGO ID	RRSA01227
Location	1542346 – 1543194
Length	283 aa, 849 bp
Mol. Weight	30.7 kDa
GC Content	58.3% (+2.03%)

>RRSA01227 - *Renibacterium salmoninarum* (IG-152)

```
GTGGAATTGGGCACTTCCAGAGGCAGCCAGGAAGCCAATATTGGCGTGGATGAATTGTTTCCGGTGCCC  
TGAAACGGGACCGGTCTGATCGCGGACGGCCTAGGCGATCCAGGCCGCGAAAGGGTGTTTTTCGCACCAT  
CGTGCAAGTCTTTGGTGAGTTACTTATTACGCTCGGTGTGATCTTAATGCTCTTCGTCCGGCTGGGAATTG  
TGGTGGACCAACATTCAAGTCCGACCAAACTCAGCAACAGGCTGTTCAAGCAGTTTGGCGAGAATTCAAAG  
GCCCCGTTACGCCGAGGCTTCCGCGCAACCAACTATGGCGACCCGGTGGTAACCAAGGCGCCCGATGC  
GGCCGGTGAAACTTTTCGGCTTGGCATAACATCCCTCGATTTGGTGCCGATTACAAACCGCGGCCCTTGGTG  
CAAGGTACCGCGCAGCGGAGCTGGACACCCTAGGACTAGGTCAACACTACACCTCTACGGCGATGCCGG  
GAGCAGTGGGCAATTTCCGCGTTGCTGGTCACCGGCAACGCACGGTGGCGGTGCTCGATGCCATTCACGC  
TTTGGTCCCGGGGACAAGATCTATGTTCAAGACCAAGACGGCTACTACACATACGTCTTCCGGAACAGC  
GAGATTGTGCTGCCGACGCAAACCAGCGTTCTGGCTCCGGTGCCACCCAGAGCTCTGCTCAACCAACTG  
ACCGTTATTTGACGTTGACCAGCTGCAATCCGCGGTTTGGCGTCGCTGAGCGTTTCATCGCATAACCGGT  
GCTGGAATCTTGGCAACCAGCTAGCGCAGGACCGCCAGCAGAGATCGCGCAGCAAGTTCAAGCCGCAGCA  
GGACAGGGATAG
```

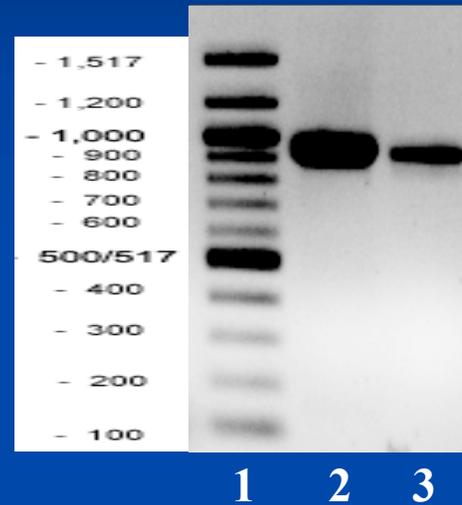
# The *R. sal* Sortase

Pfam HMM    engnavlaGHRvfGfgltlygtlFrnLdklkkGDkiyvetknetfavYkvysilevVePsdvdvLddvkdgkdgskkggaelTLiTCtPFFi  
                   gn+++aGHR            +t+g++    + +l    GDkiyv t+++++ Y            e+V P++++vL    v+   ++   +   +++lTL+ C P  
 RRS01227    AVGNFAVAGHR-----QTHGAVLDAIHALVPGDKIYVQTQDGYIT-YVFRNS-EIVLPTQTSVLAPVP-TQSSAQPTDRYLTLTSCNP--R



Ilangovan *et al* (2001) *PNAS* **98**: 6056–6061

# Sortase Expression

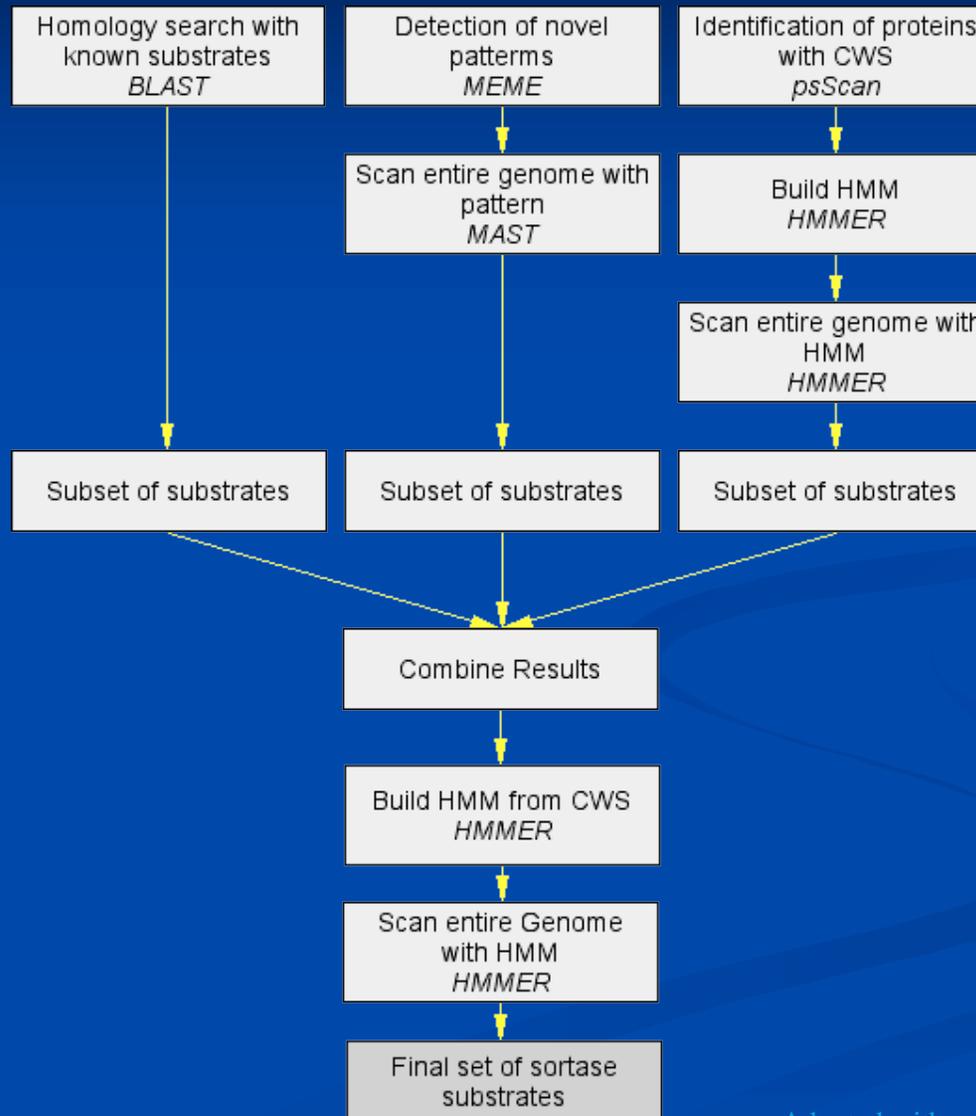


1. 100bp marker
2. Sortase PCR product
3. Sortase RT-PCR product

# Substrate



# Finding Substrates: Parallel Approaches



# Finding Substrates: psScan

## Manual cell wall sorting signal verification

- Cleavage motif in last 60 aa
- 2–3 transmembrane regions
- Charged tail
- Signal peptide

## Tripartite pattern search

L-P-X-[TA]-[GNSD]-X(1,11)-[VIFAGTSMWLCNRK](14,20)-X(6)-[RK](1,5)

5 hits

1 true positive

L-A-X-[TA]-[GNSD]-X(1,11)-[VIFAGTSMWLCNRK](14,20)-X(6)-[RK](1,5)

19 matches

8 true positives

## Cleavage motif search

L-[PA]-X-[TA]-[GNSD]

698 matches

9 true positives



# Finding Substrates: MEME/MAST

[meme.sdsc.edu](http://meme.sdsc.edu)

MEME discovers short highly conserved regions (putative motifs)

MAST searches for matches to MEME motifs

*No new substrates!!*

Best matching motif:

N-[HS]-GDSGGPL-[IV]-[IV]-[GN]-[GN]-K-[IL]-VGVCS-[TG]

A serine protease pattern

# Finding Substrates: Pairwise Comparisons

Search *R. sal.* genome for homologs to known substrates.

*Corynebacterium diphtheriae*: 16 substrates

*R. salmoninarum* homologs: 11 total hits, 6 unique

Result: *No true positives!!*

Although, partial aminopeptidase C (EC 3.4.22.40)

Currently searching with more homologs.

# The *R. sal.* Substrates

ORF ID	Function / Annotation	Pattern	Length (aa)	Pattern Position	
				Start	End
RRSA00417	Peptidase (EC 3.4.-.-)	LAnTG	986	946	979
RRSA00636	Putative hydrolase (EC 3.-.-.-)	LAaTG	890	858	889
RRSA01045	Hypothetical protein	LAaTG	118	84	117
RRSA01248	Putative partial cell-surface adhesin	LAnTG	637	602	635
RRSA01787	Acid phosphatase (EC 3.1.3.2)	LAaTG	387	352	385
RRSA02499	Partial sialidase (EC 3.2.1.18)	LAeTG	676	642	674
RRSA02971	Hypothetical membrane protein	LAiTG	110	50	91
RRSA00949	Hypothetical protein	LPvAG	142	97	128
RRSA02546	Cell surface adhesin (C-terminus)	LPdTG	78	43	77



## Next Steps

### This Project:

- Inspect partial substrates
- Screen additional substrate homologs

### Further Investigations:

- Fish immunization to assess protective nature
- Development of anti-infectives



## Next Steps

### Inhibition of sortase reaction by anti-infectives

- *Novel approach to drug therapy*
- *Reduced virulence*
- *Does not affect cell viability*
- *No selection pressure for resistance.*

## Next Steps

Small molecule inhibitors:

**Methyl methane thiosulfonate**

Phenyl vinyl sulfones

Phenylsulfonyl 1-propene

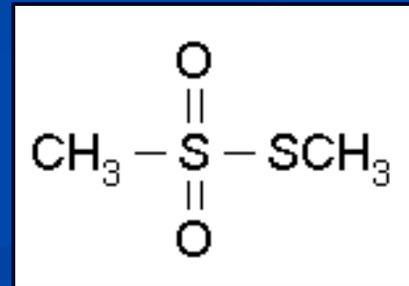
Methyl vinyl sulfone

Ethyl vinyl sulfone

Divenyl sulfone

Phenylsulfonyl ethylene

Organic mercurials





# Thank you

Dr. Mark Strom

*Project co-PI and my supervisor*

Donald Chen

*My fellow undergraduate and conspirator*

Dr. Sudheesh Ponneraserry

*For the PCR work*

The NWFSC Microbiology lab

*For always listening to me rattle on about genomics*

The UW Carlson Center for Experiential Learning

*For the great opportunity!*