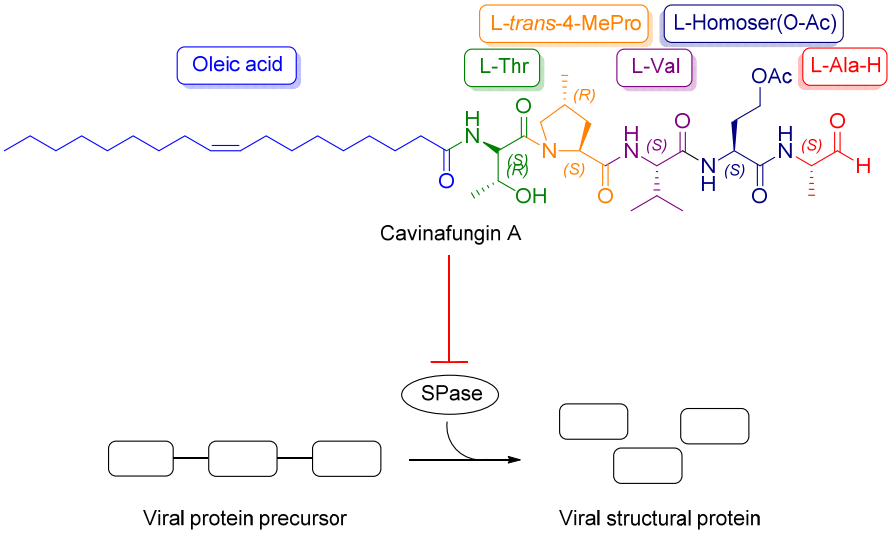
Synthesis and Structural Modifications of Antiviral Natural Product Cavinafungin A



Ting Fung Lam

University of Southern California

Loker Hydrocarbon Research Institute

Qualifying Examination Proposal 1

2017

ABSTRACT

Cavinafungin A, a natural product isolated from the fungus *Colispora cavincola*, was recently discovered to potently and selectively inhibit Zika and dengue viral reproduction in human host cells. It binds and inhibits host signal peptidase SPC18, which processes viral structural proteins for viral reproduction. In pursuance of inventing urgently needed therapies against these viruses, I hereby propose an *in vitro* method to study the inhibitory kinetics of cavinafungin A, as well as structural analogs that will render the natural product as a suitable drug candidate (Figure 1). Hence, the experimental results may move us a step closer to developing a new antiviral drug to treat both Zika fever and dengue fever.

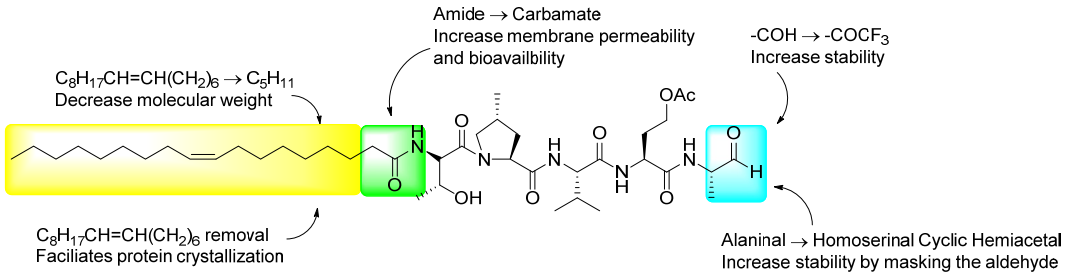


Figure 1. Summary of structural modifications of CavA and their expected effects.

INTRODUCTION

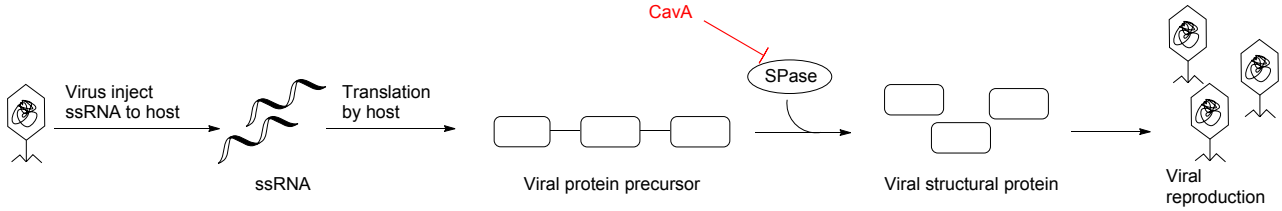
In 2015, a major global outbreak of Zika virus spread rapidly throughout more than 33 countries, infecting a large number of people, such that the World Health Organization declared a public health emergency of international concern.[1] This highly infectious flavivirus is transmitted mainly through mosquitos and is associated with newborn birth defects. Due to the lack of US Food and Drug Administration (US FDA) approved treatments against Zika virus, potential therapies against this virus are being investigated.[1] Similarly, dengue fever, caused by another flavivirus, results in ~1 million infections each year predominantly in Southeast Asia.[2] Flavivirus infects host cells by injecting RNA and reproduces with the aid of translational machinery of the host (Figure 2a). Both viruses have caught attention among governments and the healthcare industry to develop measures to control the spread of these viruses.

In 2017, an anti-viral natural product named cavinagunfin A (CavA, **1**) was discovered to potently inhibit production of Zika and dengue viruses in human cells by Estoppey et al.[3] Cavinafungin A inhibits virus replication in host human cells through the actions of binding and inhibiting protein SPC18 (aka SEC11A, UniProtKB: P67812), an essential host endoplasmic

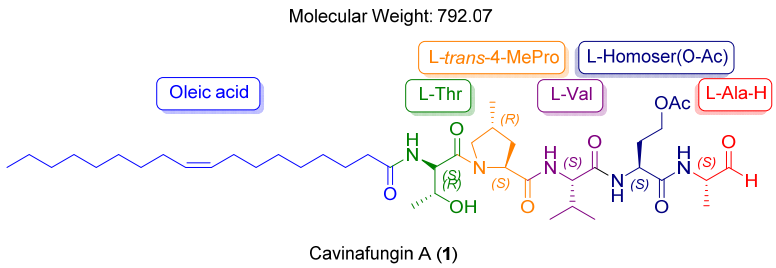
reticulum (ER) signal peptidase (SPase) subunit which processes viral structural proteins to facilitate their formation.[3-5] It is also the first small molecule inhibitor of human SPase reported to date.[3] Ortiz-Lopez et al. isolated CavA as a linear lipopeptide in 2015 from *Colispora cavincola*, with the amino acid sequence of Oleate-Thr-(trans-4-MePro)-Val-HomoSer(O-Ac)-Ala-H (Figure 2b).[6] CavA functions by reversibly trapping Ser56 of SPC18 via its aldehyde moiety, forming a hemiacetal reversibly (Figure 2c).[3]

Although cavinafungin A efficiently blocks the reproduction of Zika and dengue viruses in human cells, its molecular structure has to be modified in order to improve its chemical and pharmacological properties for antiviral therapies. For instance, the peptidyl aldehyde group is highly reactive and may react with other biomolecules prior to reaching the target, so it will not be suitable as a drug candidate.[7-8] Furthermore, the binding kinetics data are yet to be reported. Crystal structure of the protein target, SPC18, as well as the inhibitor-bound protein remain to be solved. Natural products are remarkable sources of lead compounds for drug discovery due to their structural diversities. In fact, it is estimated that 55% of approved drugs are natural products or their derivatives.[9] Thus, CavA has the potential to serve as a lead compound for developing drug candidates for antiviral therapies.Taken together, an interdisciplinary medicinal chemistry and pharmacology approach to further investigate the inhibitory activity of CavA towards its target SPase, as well as optimizing the lead compound is necessary to design potential therapies of treating flaviviral diseases.

a



b



c

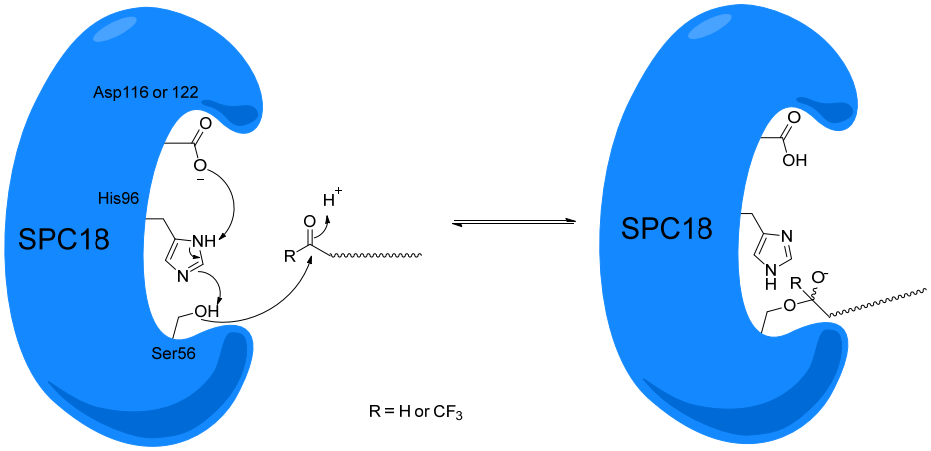
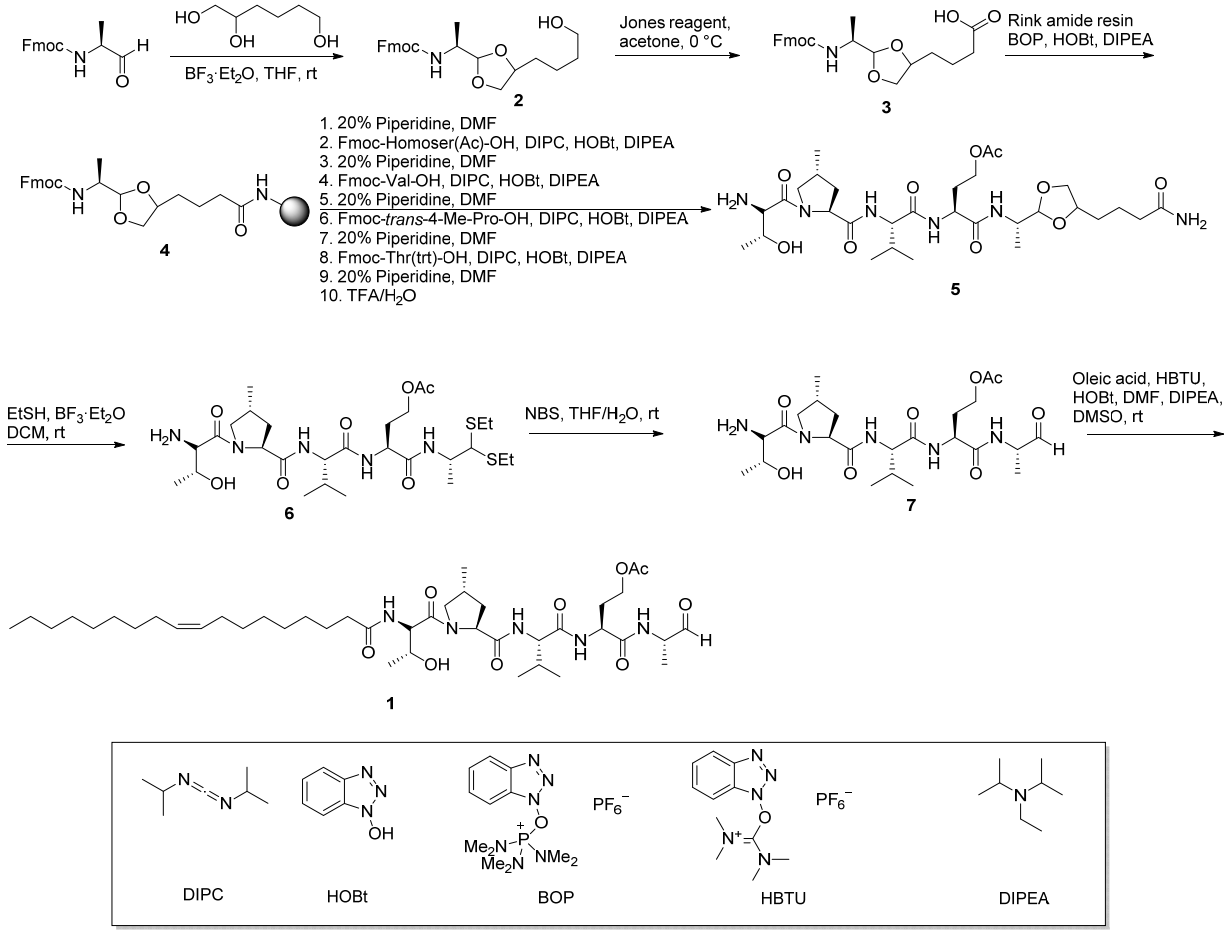


Figure 2. a) Mechanism of SPase-dependent viral reproduction; b) Structure of cavinafungin A (**1**) with each component labeled; c) Enzymatic mechanisms of SPC18's SHD catalytic triad;

STRATEGY

I. Total Synthesis of Cavinafungin A (1)



Scheme 1.

The total synthesis of CavA (**1**) will begin with a recently reported solid phase peptide synthesis strategy for synthesizing peptide aldehydes by Konno et al., which works with peptides containing nucleophilic residues.[10] Beginning with commerically available Fmoc-L-alaninal (CAS #: 146803-41-0) and the linker 1,2,6-Hexanetriol, compound **2** is afforded by $BF_3 \cdot Et_2O$ catalyzed acetal formation reaction (Scheme 1). The free alcohol group is subsequently oxidized by Jones reagent, and coupled to Rink amide resin by BOP/HOBt reagents to yield **4**. Standard peptide chain elongation with Fmoc-protected amino acids followed by TFA cleavage of the resin affords **5**. Fmoc-Homoser(O-Ac)-OH is afforded by standard Fmoc protection of homoserine followed by acetylation with acetic anhydride, whereas Fmoc-*trans*-4-MePro-OH can be made by standard Fmoc protection of commercially available *trans*-4-methyl-L-proline (CAS#:23009-50-9, Carbosynth Limited). Acetal **5** is treated with ethanethiol under $BF_3 \cdot Et_2O$ to generate thioacetal **6**, which is subjected to *N*-bromosuccinimide treatment to afford peptidyl aldehyde **7**.[11-12] Finally, **7** will be coupled with oleic acid by treatment with HBTU with the presence of alcohol and aldehyde.[13-14]

II. Protein expression and *In vitro* enzymatic activity assays

SPC18 is a serine hydrolase possessing conserved catalytic Ser-His-Asp residues and functions by a classical SHD catalytic triad.[15-17] This protein is the catalytic subunit of a signal peptidase complex consisting 5 different proteins (SPC18, SPC21, SPC22/23, SPC12 and SPC25) in human,[18] and importantly, the protein subunit SPC22/23, also known as SPCS3, (UniProtKB: P61009) is necessary and sufficient for activating signal peptidase activity of SPC18 for their yeast and chicken oviduct homologs.[4, 16, 19] Both SPC18 and SPCS3 are membrane proteins, in which their *N*-terminal domains of approximately 30 residues span through the cytoplasm and ER membrane, while the catalytic domains are located in the ER lumen.[16] SPase recognizes specific consensus signal sequences in their substrates to cleave the protein at specific positions. Typically, the -1 and -3 positions upstream of the cleavage sites are small, uncharged residues.[16]

Although various *in vitro* experiments have been performed using purified yeast or hen homologs of SPC18 and SPCS3, the human proteins were never expressed from their cloned cDNAs.[18] More importantly, the crystal structures of SPC18 itself, and the SPC18/SPCS3 complex, with or without substrate binding have not been solved.[18] Thus, in order to synthesize enough quantities of proteins to perform *in vitro* experiments, protein expression from cloned

cDNAs should be attempted. In order to obtain water-soluble SPC18 and SPCS3, truncated versions of these proteins will be expressed in *E. coli*. Hence, SPC18 (Δ2-36) and SPCS3 (Δ2-32) will be co-expressed for enzymatic assay experiments, such that only the ER lumen catalytic domains would be expressed.[20] Note that a truncated bacterial SPase I (Δ2-75) was shown to be catalytically active, and it was able to be crystallized.[16] Consequently, the inhibitory activities of CavA towards the SPC18/SPCS3 complex co-expressed from their cDNAs will be assayed *in vitro* because this permits investigation of the structures of inhibitor-bound proteins as well as the enzymatic kinetics. The enzyme substrate will be pro-OmpA-nuclease A (pONA), a synthetic protein which was found to be an active substrate of chicken oviduct SPase.[21] These kinetics data such as inhibitor binding affinities (K_i) and thermodynamic parameters of inhibitor binding will offer clues for structural optimization of the compounds.[22]

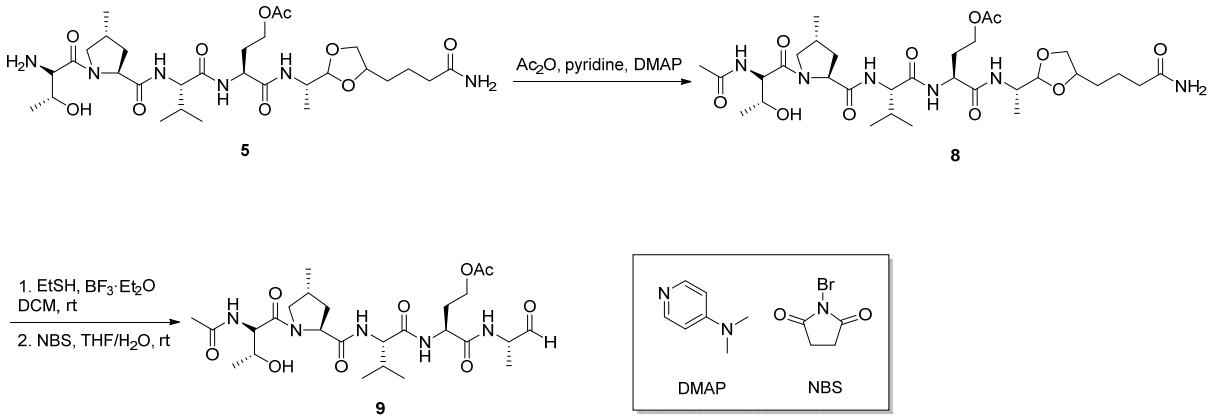
III. *In vivo* activity assays

In order to test the inhibitory activities of CavA analogs *in vivo*, signal peptide cleavage activities by SPase in human 293T cells will be assayed as previously reported by Estoppey et al.[3] Briefly, HEK293 cells will be transfected with plasmids expressing viral precursor protein substrates, C_{14}-prM-E and 2K-NS4B-V5, which become viral structural proteins after enzymatic cleavage. The cleavage activities will be analyzed with Western blot using anti-prM, anti-E, and anti-NS4B antibodies. Activities of SPase towards host endogenous substrates will be assayed beginning with transfecting HeLa cells and INE-1E cells with plasmids expressing their endogenous substrates pre-provasopressin and pre-proinsulin respectively. Then they will be labeled with radioactive [^{35}S]-Met and [^{35}S]-Cys to distinguish against the existing proteins from the experimental enzymatic cleavage products, and analyzed by Western blot combined with autoradiography.[3].

IV. Fatty chain length reduction

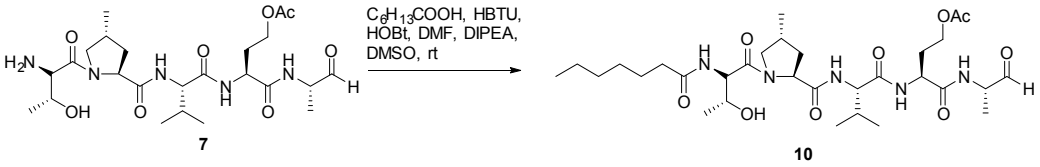
The fatty chain of CavA was proposed to function as an anchor to the ER membrane when the compound binds the target ER membrane protein SPC18.[3, 23] For *in vitro* experiments involving no membranes or membrane domains in the proteins, removal of the carbon chain is proposed by capping the free amine with an acetyl group (Scheme 2).[10] Compound **5** will be subjected to acetylation of the free amine by Ac_2O, and subsequently the acetal will be removed in a similar fashion to yield aldehyde **9**. Importantly, crystallization of the inhibitor bound

protein can be difficult in the presence of a flexible carbon chain. Thus, **9** can be utilized for *in vitro* experiments as well as crystallographic analysis.



Scheme 2.

The molecular weight of **1** is greater than average compared to a majority of approved drugs (Figure 2b), so it should be attempted to be reduced in order to improve its drug-like properties. Buzder-Lantos et al. designed a similar linear lipopeptide for inhibiting *E. coli* SPase, and reported that the fatty chain was essential for inhibitory activities against the full-length SPase protein.[23-24] The length of the fatty chains of the synthetic *E. coli* SPase inhibitors is 10 carbons, so the chain length of **1** may be reduced without affecting activity. Furthermore, in order to estimate their absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of the compounds of interest, *in silico* calculations were implemented with Schrödinger® QikProp™ program (Table 1).[25-26] In particular, the predicted parameters of **1** imply that it is physically too large and non-polar, so it may be improved by reducing its size as well as removing non-polar groups (Figure 3). Hence, a panel of analogs of **1** will be synthesized with various carbon chain lengths. For the purpose of illustrating the latter part of this proposal, the carbon chain length is set to be 7 carbons tentatively, after analyzing the *in silico* calculation results as well as comparing with other drug candidates with fatty chains such as MGS2010, a prodrug of the glutamate receptor agonist MGS0039.[27-28] Hence, **10** can be synthesized from **7** by coupling with heptanoic acid (Scheme 3).



Scheme 3.

```
      QikProp v5.1 (rel 11) - Properties Predictions for
            1

Primary Metabolites & Reactive FGs:
> Metabolism likely: allylic H -> alcohol
> Metabolism likely: allylic H -> alcohol
> Metabolism likely: alpha, beta dehydrogenation at carbonyl
> Metabolism likely: alpha, beta dehydrogenation at carbonyl
> Metabolism likely: secondary alcohol -> ketone
> Metabolism likely: alpha, beta dehydrogenation at carbonyl
> Metabolism likely: alpha, beta dehydrogenation at carbonyl
> Metabolism likely: alpha, beta dehydrogenation at carbonyl
> Reactive FG: unhindered ester
> Metabolism likely: alpha, beta dehydrogenation at carbonyl

Principal Descriptors:                                (Range 95% of Drugs)
    Solute         Molecular Weight       =   792.067  ( 130.0 / 725.0)*
    Solute         Dipole Moment (D)      =     5.767  (   1.0 /  12.5)
    Solute         Total        SASA      =  1406.800  ( 300.0 /1000.0)*
    Solute         Hydrophobic  SASA      =  1155.340  (   0.0 / 750.0)*
    Solute         Hydrophilic  SASA      =   237.527  (   7.0 / 330.0)
    Solute         Carbon Pi    SASA      =    13.933  (   0.0 / 450.0)
    Solute         Weakly Polar SASA      =     0.000  (   0.0 / 175.0)
    Solute         Molecular Volume (A^3)=  2681.822  ( 500.0 /2000.0)*
    Solute         vdW Polar SA (PSA)     =   238.895  (   7.0 / 200.0)*
    Solute         No. of Rotatable Bonds=    30.000  (   0.0 /  15.0)*
    Solute as Donor -    Hydrogen Bonds =     1.000  (   0.0 /   6.0)
    Solute as Acceptor - Hydrogen Bonds =    14.700  (   2.0 /  20.0)
    Solute Globularity   (Sphere = 1)    =     0.664  (  0.75 /  0.95)*
    Solute Ionization Potential (eV)     =     9.346  (   7.9 /  10.5)
    Solute Electron Affinity    (eV)     =    -0.101  (  -0.9 /   1.7)
Predictions for Properties:
  QP Polarizability (Angstroms^3)        =    82.043M (  13.0 /  70.0)*
  QP log P  for      hexadecane/gas      =    27.046M (   4.0 /  18.0)*
  QP log P  for      octanol/gas         =    36.272M (   8.0 /  35.0)*
  QP log P  for      water/gas           =    20.578M (   4.0 /  45.0)
  QP log P  for      octanol/water       =     5.040  (  -2.0 /   6.5)
  QP log S  for   aqueous solubility     =    -8.059M (  -6.5 /   0.5)*
  QP log S - conformation independent    =    -6.161M (  -6.5 /   0.5)
  QP log K hsa Serum Protein Binding     =    -0.143M (  -1.5 /   1.5)
  QP log BB for      brain/blood         =    -4.841M (  -3.0 /   1.2)*
  No. of Primary Metabolites             =         9  (   1.0 /   8.0)*
  Predicted CNS Activity (-- to ++)      =        --
  HERG K+ Channel Blockage: log IC50     =    -0.170M (concern below -5)
  Apparent Caco-2 Permeability (nm/sec)  =        11M (<25 poor, >500 great)
  Apparent MDCK   Permeability (nm/sec)  =        21M (<25 poor, >500 great)
  QP log Kp for skin permeability        =    -2.968M (Kp in cm/hr)
  Jm, max transdermal transport rate     =     0.000M (micrograms/cm^2-hr)
  Lipinski Rule of 5 Violations          =         3  (maximum is 4)
  Jorgensen Rule of 3 Violations         =         3  (maximum is 3)
  % Human Oral Absorption in GI (+-20%)  =        36M (<25% is poor)
  Qual. Model for Human Oral Absorption  =       lowM (>80% is high)

     A * indicates a violation of the 95% range. # stars = 13
    An M indicates MW is outside training range.
```

Figure 3. Predicted ADMET properties of CavA (**1**) by QikProp™.[25]

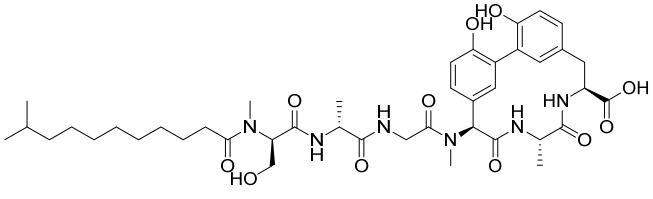
molecule	mol_MW	#rotor	SASA	volume	QPlogPo/w	QPPCaco	PercentHumanOralAbsorption	SAfluorine	RuleOfFive	RuleOfThree
1	792.067	30	1406.8	2681.82	5.04	11.249	36.395	0	3	3
10	639.788	20	1052.94	2036.24	1.007	11.604	25.982	0	2	2
17	707.786	20	1140.96	2158.75	2.388	37.218	43.126	107.501	2	2
18	641.76	19	1035.45	2008.72	1.522	20.434	33.393	0	2	1
22	709.759	19	1119.67	2130.82	2.772	36.851	45.297	101.433	2	1
26	671.787	19	1072.58	2064.22	0.844	25.283	18.117	0	3	0

Table 1. Selected ADMET parameters of compounds of interest predicted by QikProp™. Mol_MW: Molecular weight; #rotor: Number of rotatable bonds; SASA: Total solvent accessible area (Å^2); volume: Total solvent accessible volume (Å^3); QPlogPo/w: Octabol/water partition coefficient; QPPCaco: Caco-2 cell permeability (nm/sec); PercentHumanOralAbsorption: Human oral absorption in 0 – 100% scale; SAFluorine: Solvent accessible area of fluorine atoms (Å^2); RuleofFive: Number of Lipinski's rule of five violations; RuleOfThree: Number of Jorgensen's rule of three violations.[25]

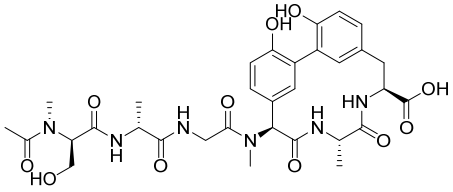
V. *In silico* docking simulations

In order to estimate binding energies and structures for the compounds of interest, and to exclude potentially poorly binding compounds, *in silico* docking experiments with Schrödinger® Glide™ were performed with CavA and its derivatives using reported crystal structure of *E coli.* SPase I bound to a natural product antibiotic arylomycin A2 (**11**) (PDB: 1T7D), whose structure is similar to CavA.[26, 29-30] Due to the flexibility of the fatty chain of **11**, electron density could not be observed, so the observed structure of the inhibitor of the crystal structure becomes **12**.[29] Since the crystal structure of SPC18 has not been solved, the closest SPase homolog of *E. coli* structure with a similar ligand bound is selected for docking calculations. Docking constraints were imposed such that the reactive aldehyde group aligns near the active site Ser90 of SPase I, and the N-terminal of **9** aligns near Pro83 in a similar fashion for **12**.[29] Docking simulation with **12** resulted in a docking score of -6.90 kcal/mol, which approximates the ligand binding free energy (Table 2). Using the same calculation settings, the docking score for **9** was -5.70 kcal/mol. Their predicted structures show that the inhibitor binds between two β-strands of the protein (Figure 4b). Unsurprisingly, **12** scores slightly better than **9**, but the result indicates that *E. coli* SPase I still serves as a reasonable approximation to SPC18. In addition, the distances between the nucleophilic oxygen atom of Ser90 to the electrophilic carbonyl carbon of the ligands are also estimated (Table 2). Hence, the results provide a rough estimation of their binding affinities to SPC18.

a



Arylomycin A2 (**11**)



12

b

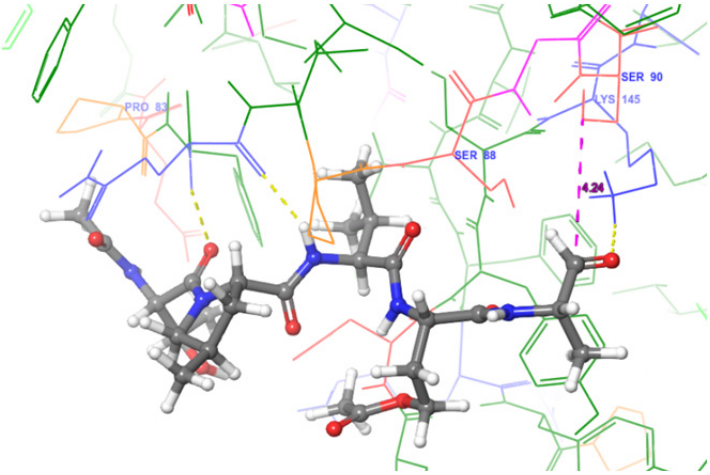


Figure 4. a) Structures of arylomycin A2 (**11**) and **12**; b) Predicted inhibitor-bound structures of active site of SPase I with **9**.

Compound	Score (kcal/mol)	Ser90 to C=O Distance (Å)
12 (control)	-6.90	3.99
9	-5.70	4.24
16	-5.24	4.92
19	-5.59	4.07
20	-7.42	4.19
23	-6.76	5.35
24	-5.76	5.03
27	-5.14	4.17
28	-4.84	4.53

Table 2. Docking simulation scores and predicted distances between Ser90 to the carbonyl carbon of the compounds of interest using Schrödinger® Glide™.[26, 30]

VI. Replacing aldehyde functionality with trifluoromethyl ketone

As mentioned earlier, aldehydes are generally not amenable for drug candidates due to the high reactivity, so the aldehyde of CavA will be substituted to a trifluoromethyl ketone functionality. Incorporating fluorine atoms to drugs and drug candidates has recently attracted a lot of attention due to their high electronegativity, which may improve drug physical and chemical properties.[31] In particular, trifluoromethyl ketones (TFMK) are known to behave as reversible competitive inhibitors to serine proteases by forming a hemiacetal upon binding (Figure 2b). Due to the strong inductive effect of fluorine atoms, the carbonyl carbon is very electron deficient, and allows nucleophilic attack by the serine hydroxyl group of the enzyme.[31-33] TFMK is also featured in the serine hydrolase inhibitor drug candidate GR148672X by GlaxoSmithKline (Figure 5).[34] Moreover, peptidyl TMFKs were observed to be more potent than their aldehyde counterparts, while retaining their specificities towards the enzyme targets, as

reported by Abeles et al. for a similar chymotrypsin inhibitor (Figure 5).[33] The synthesis of compound **16** begins with protection of the free alcohol of **9** with silyl ether,[35] follow by addition of CF_3 to aldehyde through Ruppert-Prakash reagent ($TMSCF_3$) with cesium fluoride as the initiator, in the presence of amides and ester forming trifluoromethyl alcohol **14** (Scheme 4).[36-37] Standard Dess-Martin periodinane oxidation followed by TBAF deprotection affords **16** as the TFMK.[31, 36] Similarly, compound **17** can be synthesized from **10** in a similar fashion (Scheme 5).

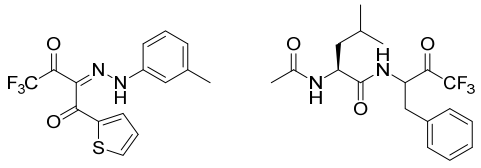
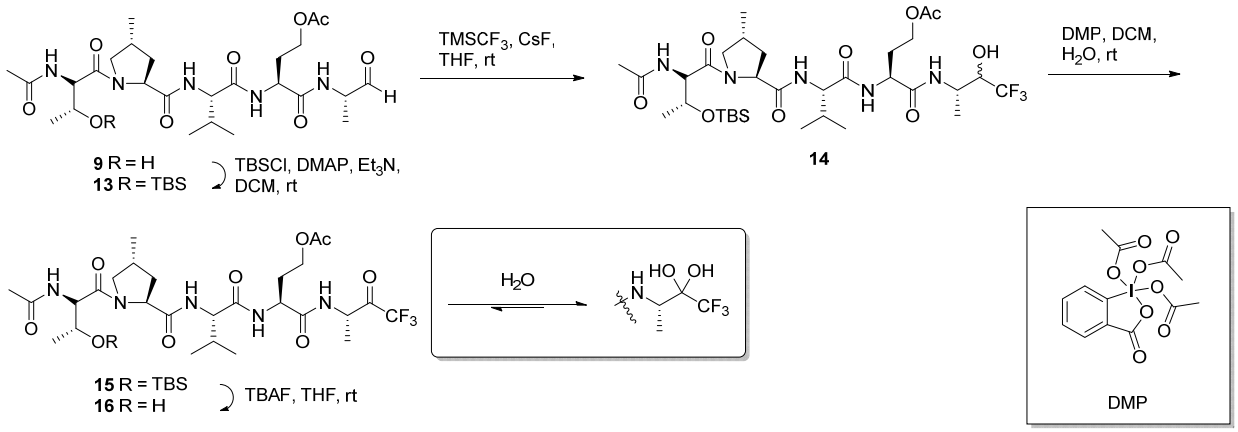
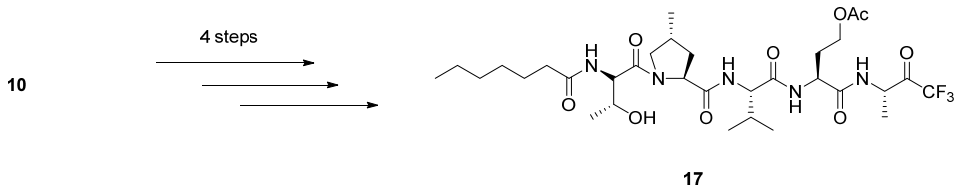


Figure 5. Examples of known enzyme inhibitors featuring TMFK moiety.[33-34]



Scheme 4.



Scheme 5.

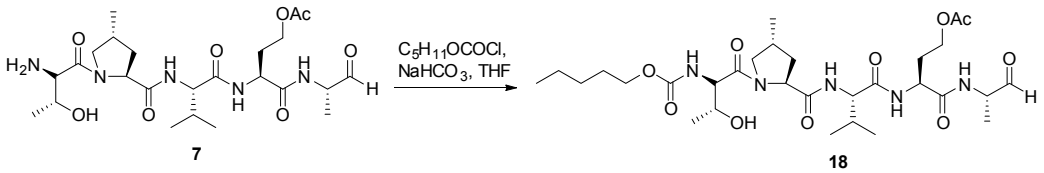
Addition of fluorine atoms also allows binding and kinetics studies by ^{19}F nuclear magnetic resonance (NMR) both *in vivo* and *in vitro*.[38] Liang et al. observed that new fluorine signals appeared downfield of the unbound inhibitor upon binding to the enzyme.[32] Hence, by using NMR kinetics protocols, the intensities of both bound and unbound signals can be recorded

over time, and the enzyme-inhibitor kinetics can be calculated.[38] In addition, they observed a majority of the inhibitor to exist as hydrated form when it is dissolved in water, which may also apply to **16** (Scheme 4). Consequently, the enzymes are expected to react with the inhibitors in the ketone form, as an intramolecular like process to afford strong binding activities (Figure 2b).[32-33]

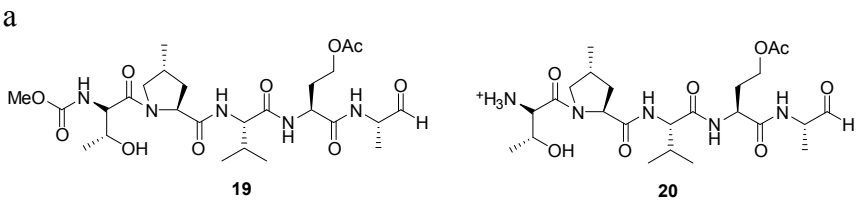
VII. Prodrug approach with carbamate functionality

Because the putative function of the carbon chain of CavA is to anchor the membrane, CavA may be modified such that the chain will be cleaved upon binding to its target protein, which may increase its potency by tighter binding. Consequently, I propose to replace the amide linkage of the fatty chain to a carbamate to render it as a prodrug.[39] Prodrugs are inactive or less active molecules that undergo chemical transformation *in vivo* after ingestion to generate the active drug molecule, and they are promiscuously featured in approved drugs such as Capecitabine, an anti-cancer drug.[27, 39] Furthermore, prodrugs are designed by coupling non-polar groups in order to increase membrane permeability.[27] Carbamates are usually more stable than an ester but more labile than an amide, so they are susceptible to esterase hydrolysis, which makes them preferable for prodrug design. In addition, they are also known to increase the bioavailability of many other drugs.[28, 39]

The synthesis of **18** is accomplished by simply coupling the free amine of peptide **7** with an alkyl chloroformate in the presence of a base while not tampering the aldehyde (Scheme 6).[40-41] In order to keep the length of the molecule consistent, the carbon chain length of **18** is reduced to 5 carbons. Due to the chemical stability of the carbamate functionality, **18** is anticipated to pass through the cell and ER membranes and binds its target, before being metabolized to release its active warhead.[39] Moreover, the carbamate should be exposed to the solvent, because it links the membrane-binding fatty chain and the protein-binding amino acids of the molecule, as illustrated by *in silico* simulation (Figure 6b). Therefore, water molecules can attack the carbamate of **18** while it is bound to its target. Subsequently, a free amine is generated as an ammonium ion under physiological pH. *In silico* docking simulation results show that the ammonium ion of **20** is capable of making additional cation-anion interactions with the amino acid residues of the protein target (Figure 6c), so the drug is presumed to be more potent by binding more tightly to the target (Table 2).



Scheme 6.



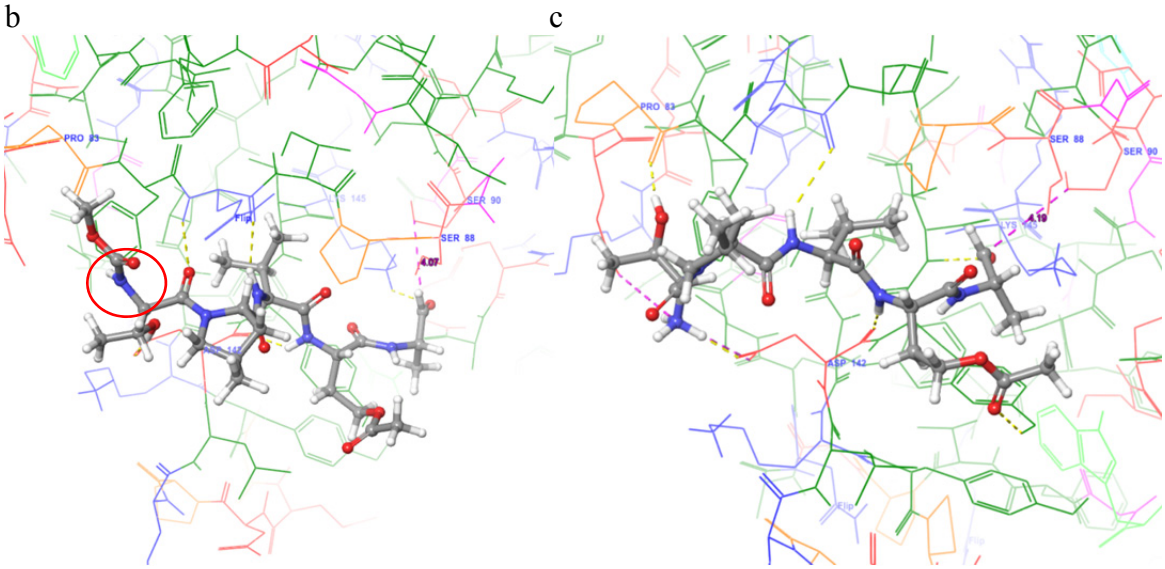
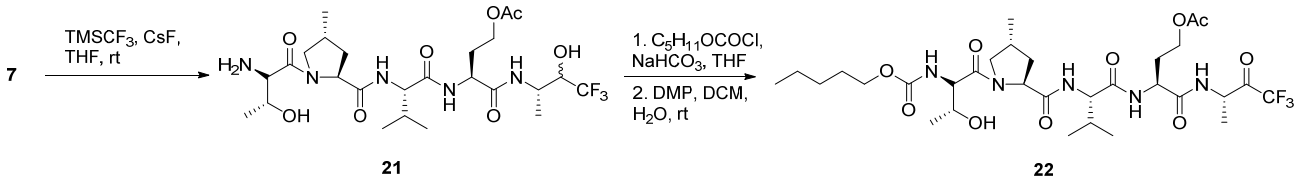


Figure 6. a) Structures of **19** and **20** used for docking simulation. b,c) Docking simulations of **19** and **20** bound to SPase I.

VIII. Carbamate and TMFK combination

Next, the carbamate and TMFK functionalities are combined to generate compound **22**. The synthesis of **22** begins with trimethylfluorination of intermediate **7** to yield **21** (Scheme 7). Subsequently, **21** is subjected to carbamate formation followed by DMP oxidation to yield compound **22** as previously mentioned.[40] This functionality combination was previously observed to have high oral activities by researchers at AstraZeneca for a similar peptide inhibitor of elastase, so it is amenable for the optimization of CavA.[39-40]



Scheme 7.

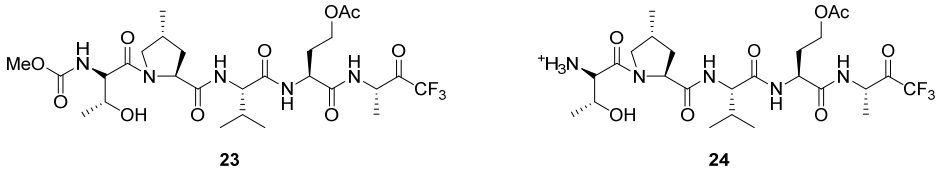
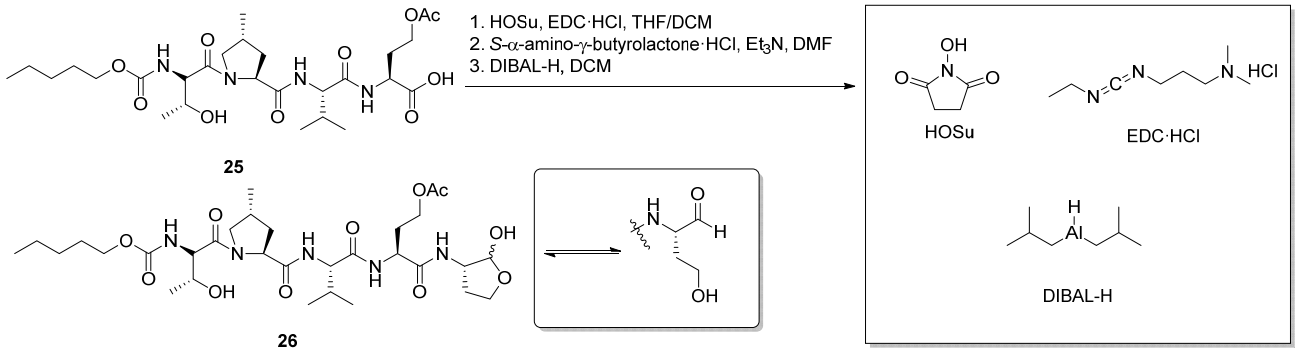


Figure 7. Structures of **23** and **24** used for docking simulations.

IX. Hemiacetal as a prodrug for aldehyde

Because replacing the reactive aldehyde group with TMFK alters its reactivity, an alternative prodrug approach that masks the aldehyde functionality may be desirable. Hence, the C-terminal alaninal residue will be substituted by a cyclic hemiacetal, which is expected to open up to form the active homoserine aldehyde before binding to the enzyme.[8, 42] Compound **26** will be synthesized beginning with peptide **25**, which can be synthesized by standard solid-phase peptide synthesis protocols. Subsequently, activation of the carboxylic acid, followed by coupling with the amino lactone and DIBAL reduction affords compound **26** as previously reported (Scheme 8).[8] This idea was successfully applied for a similar peptidyl aldehyde functioning as a calpain inhibitor, and it significantly increased the metabolic stability of the compound as reported by Nakamura et al. Nonetheless, the activity of the compound was reduced presumably due to the amendment of the active residue.[42] In addition, docking simulation results in lower scores for compounds **27** and **28** (Table 2). Taken together, replacing the aldehyde with a hemiacetal may serve as an alternative strategy towards adopting TMFK functionality.



Scheme 8.

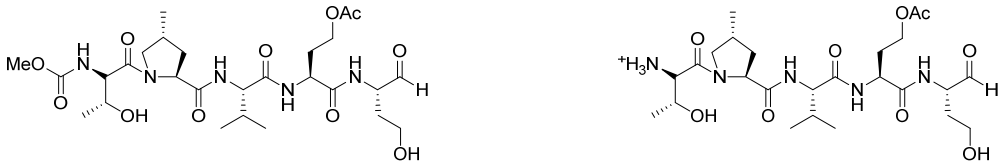


Figure 8. Structures of **27** and **28** used for docking simulations.

X. Activity based probes

A variety of tool compounds can be synthesized by adding an alkyne functionality onto the fatty chains of the proposed compounds, followed by attaching affinity or fluorescence tags by click chemistry. For instance, activity-based probes **29** and **30** (Figure 9) can be easily designed by incorporating fluorophosphonate as the irreversible warhead and either the fluorophore rhodamine (**29**) or biotin (**30**) attached by copper-catalyzed azido-alkyne cycloaddition (CuAAC).[43-44] Because serine hydrolases are promiscuous in human cells, the compounds might bind to other undesired target enzymes. Thus, probe **29** can be utilized to measure the amount of bound protein as well as identifying any non-specific protein binding partners through in-gel fluorescence visualizations. Moreover, by using standard biotin-streptavidin affinity chromatography protocols, probe **30** can be used for enriching any proteins that binds to the molecule of interest.[44] Note that similar activity-based probes had been successfully utilized for studying azadipeptide nitrile inhibitors for cysteine proteases cruzain and rhodesain.[44-45]

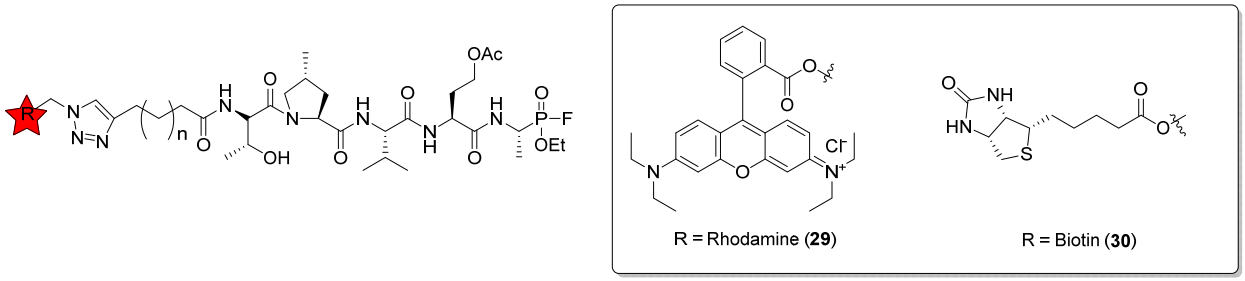


Figure 9. Structures of activity-based probes **29** and **30**.

XI. Live virus assays

The inhibitory effects of the proposed compounds against Zika and dengue viruses will be measured using live viruses. A549 cells will be incubated with multiple dengue virus strains such as Dengue virus type 1 (ATCC® VR-1254™) or Zika virus (ATCC® VR-84™) as well as with the compounds of interest.[3] The surviving quantities of dengue and Zika viruses will be measured by labeling with Dylight 488® conjugated flavivirus group antigen 4G2 antibody (Novus®). Alternatively, a recently reported recombinant luciferase Zika virus can be utilized, such that it can be quantified by measuring luciferase activity.[46] Due to the lack of equipment meeting biosafety level 2 standards for highly contagious viruses, the experiments will be conducted in Department of Molecular Microbiology & Immunology at Keck School of Medicine.

CONCLUSION

Due to the wide-spreading of Zika and dengue viruses around the globe, new antiviral therapies are urgently needed to control the spread of these viruses. CavA has the potential to serve as a lead compound for drug discovery against flaviviruses, by inhibiting the viral protein processing activity of *H. sapiens* signal peptidase SPC18. The proposed strategies above may take us a step forward in developing a new medicine to treat Zika fever and dengue fever patients. In addition, the experimental results may allow a better understanding of how endogenous enzymes assist viral reproduction, and opens a new pathway of thwarting these viruses.

REFERENCES

1. Petersen, L. R.; Jamieson, D. J.; Powers, A. M.; Honein, M. A., Zika Virus. *New England Journal of Medicine* **2016,** *374* (16), 1552-1563.

2. Guzman, M. G.; Halstead, S. B.; Artsob, H.; Buchy, P.; Farrar, J.; Gubler, D. J.; Hunsperger, E.; Kroeger, A.; Margolis, H. S.; Martinez, E.; Nathan, M. B.; Pelegrino, J. L.; Simmons, C.; Yoksan, S.; Peeling, R. W., Dengue: a continuing global threat. *Nat Rev Micro* **2010,** *8* (12 Suppl), S7-16.
3. Estoppey, D.; Lee, C. M.; Janoschke, M.; Lee, B. H.; Wan, K. F.; Dong, H.; Mathys, P.; Filipuzzi, I.; Schuhmann, T.; Riedl, R.; Aust, T.; Galuba, O.; McAllister, G.; Russ, C.; Spiess, M.; Bouwmeester, T.; Bonamy, G. M. C.; Hoepfner, D., The Natural Product Cavinafungin Selectively Interferes with Zika and Dengue Virus Replication by Inhibition of the Host Signal Peptidase. *Cell Reports* **2017,** *19* (3), 451-460.
4. Fang, H.; Mullins, C.; Green, N., In Addition to SEC11, a Newly Identified Gene,SPC3, Is Essential for Signal Peptidase Activity in the Yeast Endoplasmic Reticulum. *Journal of Biological Chemistry* **1997,** *272* (20), 13152-13158.
5. Meyer, H.-A.; Hartmann, E., The Yeast SPC22/23 Homolog Spc3p Is Essential for Signal Peptidase Activity. *Journal of Biological Chemistry* **1997,** *272* (20), 13159-13164.
6. Ortíz-López, F. J.; Monteiro, M. C.; González-Menéndez, V.; Tormo, J. R.; Genilloud, O.; Bills, G. F.; Vicente, F.; Zhang, C.; Roemer, T.; Singh, S. B.; Reyes, F., Cyclic Colisporifungin and Linear Cavinafungins, Antifungal Lipopeptides Isolated from Colispora cavincola. *Journal of Natural Products* **2015,** *78* (3), 468-475.
7. Jurczak, J.; Golebiowski, A., Optically active N-protected .alpha.-amino aldehydes in organic synthesis. *Chemical Reviews* **1989,** *89* (1), 149-164.
8. Nakamura, M.; Yamaguchi, M.; Sakai, O.; Inoue, J., Exploration of cornea permeable calpain inhibitors as anticataract agents. *Bioorganic & Medicinal Chemistry* **2003,** *11* (7), 1371-1379.
9. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products* **2016,** *79* (3), 629-661.
10. Konno, H.; Sema, Y.; Tokairin, Y., Synthetic study of peptide aldehyde via acetal/thioacetal transformation: application for Lys/Ser-containing peptides. *Tetrahedron* **2015,** *71* (21), 3433-3438.
11. Kinoshita, M.; Taniguchi, M.; Morioka, M.; Takami, H.; Mizusawa, Y., Synthetic Studies of Amphotericin B. III. An Enantiospecific Synthesis of the C-1–C-19 Segment of the Amphotericin B Aglycon. *Bulletin of the Chemical Society of Japan* **1988,** *61* (6), 2147-2156.
12. Lallana, E.; Freire, F.; Seco, J. M.; Quiñoá, E.; Riguera, R., The 1H NMR Method for the Determination of the Absolute Configuration of 1,2,3-prim,sec,sec-Triols. *Organic Letters* **2006,** *8* (20), 4449-4452.
13. Kottani, R.; Valiulin, R. A.; Kutateladze, A. G., Direct screening of solution phase combinatorial libraries encoded with externally sensitized photolabile tags. *Proceedings of the National Academy of Sciences* **2006,** *103* (38), 13917-13921.
14. Flinn, N. S.; Quibell, M.; Turnell, W. G.; Monk, T. P.; Ramjee, M. K., Novel Chemoselective Linkage Chemistry toward Controlled Loading of Ligands to Proteins through In Situ Real-Time Quantification of Conjugate Formation. *Bioconjugate Chemistry* **2004,** *15* (5), 1010-1020.
15. Tjalsma, H.; Stöver, A. G.; Driks, A.; Venema, G.; Bron, S.; van Dijl, J. M., Conserved Serine and Histidine Residues Are Critical for Activity of the ER-type Signal Peptidase SipW of Bacillus subtilis. *Journal of Biological Chemistry* **2000,** *275* (33), 25102-25108.
16. Paetzel, M.; Karla, A.; Strynadka, N. C. J.; Dalbey, R. E., Signal Peptidases. *Chemical Reviews* **2002,** *102* (12), 4549-4580.

17. VanValkenburgh, C.; Chen, X.; Mullins, C.; Fang, H.; Green, N., The Catalytic Mechanism of Endoplasmic Reticulum Signal Peptidase Appears to Be Distinct from Most Eubacterial Signal Peptidases. *Journal of Biological Chemistry* **1999,** *274* (17), 11519-11525.
18. Walker, S. J.; Lively, M. O., Chapter 778 - Signal Peptidase (Eukaryote) A2 - Rawlings, Neil D. In *Handbook of Proteolytic Enzymes*, Salvesen, G., Ed. Academic Press: 2013; pp 3512-3517.
19. Baker, R. K.; Lively, M. O., Purification and characterization of hen oviduct microsomal signal peptidase. *Biochemistry* **1987,** *26* (26), 8561-8567.
20. Shelness, G. S.; Lin, L.; Nicchitta, C. V., Membrane topology and biogenesis of eukaryotic signal peptidase. *Journal of Biological Chemistry* **1993,** *268* (7), 5201-5208.
21. Chatterjee, S.; Suciu, D.; Dalbey, R. E.; Kahn, P. C.; Inouye, M., Determination ofKm andkcat for Signal Peptidase I Using a Full Length Secretory Precursor, pro-OmpA-nuclease A. *Journal of Molecular Biology* **1995,** *245* (4), 311-314.
22. Steadman, V. A.; Pettit, S. B.; Poullennec, K. G.; Lazarides, L.; Keats, A. J.; Dean, D. K.; Stanway, S. J.; Austin, C. A.; Sanvoisin, J. A.; Watt, G. M.; Fliri, H. G.; Liclican, A. C.; Jin, D.; Wong, M. H.; Leavitt, S. A.; Lee, Y.-J.; Tian, Y.; Frey, C. R.; Appleby, T. C.; Schmitz, U.; Jansa, P.; Mackman, R. L.; Schultz, B. E., Discovery of Potent Cyclophilin Inhibitors Based on the Structural Simplification of Sanglifehrin A. *Journal of Medicinal Chemistry* **2017,** *60* (3), 1000-1017.
23. Buzder-Lantos, P.; Bockstael, K.; Anné, J.; Herdewijn, P., Substrate based peptide aldehyde inhibits bacterial type I signal peptidase. *Bioorganic & Medicinal Chemistry Letters* **2009,** *19* (10), 2880-2883.
24. Zhong, W.; Benkovic, S. J., Development of an Internally Quenched Fluorescent Substrate forEscherichia coliLeader Peptidase. *Analytical Biochemistry* **1998,** *255* (1), 66-73.
25. QikProp. *Schrödinger, LLC, New York, NY* **2017**.
26. Maestro. *Schrödinger, LLC, New York, NY* **2017**.
27. Rautio, J.; Kumpulainen, H.; Heimbach, T.; Oliyai, R.; Oh, D.; Jarvinen, T.; Savolainen, J., Prodrugs: design and clinical applications. *Nat Rev Drug Discov* **2008,** *7* (3), 255-270.
28. Nakamura, M.; Kawakita, Y.; Yasuhara, A.; Fukasawa, Y.; Yoshida, K.; Sakagami, K.; Nakazato, A., In vitro and in vivo evaluation of the metabolism and bioavailability of ester prodrugs of mgs0039 (3-(3,4-dichlorobenzyloxy)-2-amino-6-fluorobicyclo[3.1.0]hexane-2,6-dicarboxylic Acid), a potent metabotropic glutamate receptor antagonist. *Drug Metabolism and Disposition* **2006,** *34* (3), 369-374.
29. Paetzel, M.; Goodall, J. J.; Kania, M.; Dalbey, R. E.; Page, M. G. P., Crystallographic and Biophysical Analysis of a Bacterial Signal Peptidase in Complex with a Lipopeptide-based Inhibitor. *Journal of Biological Chemistry* **2004,** *279* (29), 30781-30790.
30. Glide. *Schrödinger, LLC, New York, NY* **2017**.
31. Kelly, C. B.; Mercadante, M. A.; Leadbeater, N. E., Trifluoromethyl ketones: properties, preparation, and application. *Chemical Communications* **2013,** *49* (95), 11133-11148.
32. Liang, T. C.; Abeles, R. H., Complex of .alpha.-chymotrypsin and N-acetyl-L-leucyl-L-phenylalanyl trifluoromethyl ketone: structural studies with NMR spectroscopy. *Biochemistry* **1987,** *26* (24), 7603-7608.
33. Imperiali, B.; Abeles, R. H., Inhibition of serine proteases by peptidyl fluoromethyl ketones. *Biochemistry* **1986,** *25* (13), 3760-3767.
34. Bachovchin, D. A.; Cravatt, B. F., The pharmacological landscape and therapeutic potential of serine hydrolases. *Nat Rev Drug Discov* **2012,** *11* (1), 52-68.

35. Leon, R.; Jawalekar, A.; Redert, T.; Gaunt, M. J., Catalytic enantioselective assembly of complex molecules containing embedded quaternary stereogenic centres from simple anisidine derivatives. *Chemical Science* **2011,** *2* (8), 1487-1490.
36. Walter, M. W.; Felici, A.; Galleni, M.; Soto, R. P.; Adlington, R. M.; Baldwin, J. E.; Frère, J.-M.; Gololobov, M.; Schofield, C. J., Trifluoromethyl alcohol and ketone inhibitors of metallo-β-lactamases. *Bioorganic & Medicinal Chemistry Letters* **1996,** *6* (20), 2455-2458.
37. Inglis, S. R.; Zervosen, A.; Woon, E. C. Y.; Gerards, T.; Teller, N.; Fischer, D. S.; Luxen, A.; Schofield, C. J., Synthesis and Evaluation of 3-(Dihydroxyboryl)benzoic Acids as d,d-Carboxypeptidase R39 Inhibitors. *Journal of Medicinal Chemistry* **2009,** *52* (19), 6097-6106.
38. Marsh, E. N. G.; Suzuki, Y., Using 19F NMR to Probe Biological Interactions of Proteins and Peptides. *ACS Chemical Biology* **2014,** *9* (6), 1242-1250.
39. Ghosh, A. K.; Brindisi, M., Organic Carbamates in Drug Design and Medicinal Chemistry. *Journal of Medicinal Chemistry* **2015,** *58* (7), 2895-2940.
40. Veale, C. A.; Bernstein, P. R.; Bohnert, C. M.; Brown, F. J.; Bryant, C.; Damewood, J. R.; Earley, R.; Feeney, S. W.; Edwards, P. D.; Gomes, B.; Hulsizer, J. M.; Kosmider, B. J.; Krell, R. D.; Moore, G.; Salcedo, T. W.; Shaw, A.; Silberstein, D. S.; Steelman, G. B.; Stein, M.; Strimpler, A.; Thomas, R. M.; Vacek, E. P.; Williams, J. C.; Wolanin, D. J.; Woolson, S., Orally Active Trifluoromethyl Ketone Inhibitors of Human Leukocyte Elastase. *Journal of Medicinal Chemistry* **1997,** *40* (20), 3173-3181.
41. Tomich, P. K.; Bohanon, M. J.; Turner, S. R.; Strohbach, J. W.; Thaisrivongs, S.; Thomas, R. C.; Romines, K. R.; Yang, C. P.; Aristoff, P. A.; Skulnick, H. I., 4-hydroxy-benzopyran-2-ones and 4-hydroxy-cycloalkyl[b]pyran-2-ones useful to treat retroviral infections. Google Patents: 1997.
42. Shirasaki, Y.; Nakamura, M.; Yamaguchi, M.; Miyashita, H.; Sakai, O.; Inoue, J., Exploration of Orally Available Calpain Inhibitors 2: Peptidyl Hemiacetal Derivatives. *Journal of Medicinal Chemistry* **2006,** *49* (13), 3926-3932.
43. Chen, B.; Ge, S.-S.; Zhao, Y.-C.; Chen, C.; Yang, S., Activity-based protein profiling: an efficient approach to study serine hydrolases and their inhibitors in mammals and microbes. *RSC Advances* **2016,** *6* (114), 113327-113343.
44. Sanman, L. E.; Bogyo, M., Activity-Based Profiling of Proteases. *Annual Review of Biochemistry* **2014,** *83* (1), 249-273.
45. Yang, P. Y.; Wang, M.; Li, L.; Wu, H.; He, C. Y.; Yao, S. Q., Design, Synthesis and Biological Evaluation of Potent Azadipeptide Nitrile Inhibitors and Activity-Based Probes as Promising Anti-Trypanosoma brucei Agents. *Chemistry - A European Journal* **2012,** *18* (21), 6528-6541.
46. Shan, C.; Xie, X.; Muruato, Antonio E.; Rossi, Shannan L.; Roundy, Christopher M.; Azar, Sasha R.; Yang, Y.; Tesh, Robert B.; Bourne, N.; Barrett, Alan D.; Vasilakis, N.; Weaver, Scott C.; Shi, P.-Y., An Infectious cDNA Clone of Zika Virus to Study Viral Virulence, Mosquito Transmission, and Antiviral Inhibitors. *Cell Host & Microbe* **2016,** *19* (6), 891-900.