

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Therapeutic Development of Two Pathogenic Amyloids

**Permalink**

<https://escholarship.org/uc/item/0m817304>

**Author**

Zhao, Anni

**Publication Date**

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Therapeutic Development of Two Pathogenic Amyloids

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Biochemistry and Molecular Biology

By

Anni Zhao

2013

©Copyright by

Anni Zhao

2013

## ABSTRACT OF THE DISSERTATION

Therapeutic Development of Two Pathogenic Amyloids

by

Anni Zhao

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2013

Professor David S. Eisenberg, Chair

Amyloids are structured aggregates formed by misfolded proteins. Research has shown 25+ diseases associated with amyloid deposits *in vivo*. No cure has yet been developed for the majority of these conditions and the development of disease-modifying drugs has been disappointing. In the past year, several pharmaceutical companies have simultaneously withdrawn drug trials because of the ineffective clinical results. However, these treatments have been developed based on the molecular understanding of amyloid proteins from a few decades earlier. The amyloid field has made promising advancements in the past several years that offer potential for new therapeutic strategies. We now know that a subset of amyloidosis is caused by the deposition of the mature fibrils. Others are caused by the low molecular weight oligomers formed along the fibrillation pathway. Atomic views of these amyloid species were made available with the first steric zipper structure published in 2005 revealing the fibril spine and a

cylindrin structure published in 2012, revealing the core of a toxic oligomer. These structures provide opportunities for new therapeutic strategies against amyloid diseases.

In the course of my PhD work, I have investigated two amyloid protein systems. The first project focused on amyotrophic lateral sclerosis, a progressive motor neuron disease caused by oligomer formation of the amyloid protein superoxide dismutase I. I identified a region in the protein that is particularly important for the cytotoxicity, and solved its crystal structure. This work revealed a potential drug target for amyotrophic lateral sclerosis, a painful disease with no available treatment option. The second project focuses on an extracellular peptide whose fibril formation drastically enhances HIV infection. I developed inhibitors that prevented fibril formation and reduced HIV infection rate in a functional assay. Importantly, these inhibitors were designed based on the steric zipper structure, common to many amyloid fibrils. As such, this approach has the potential to be adapted to the other 25+ amyloid systems.

The dissertation of Anni Zhao is approved.

James U. Bowie

Irvin S.Y. Chen

Feng Guo

Todd O. Yeates

David S. Eisenberg, Committee Chair

University of California, Los Angeles

2013

## DEDICATION

This work is dedicated to my parents, particularly my mother who has been my role model for pursuing a successful career as a woman; and to my husband, Howard.

## TABLE OF CONTENTS

<b>Introduction and Dissertation Layout</b> .....	1
<b>Chapter 1: A Therapeutic Target for Lou Gehrig's Disease</b> .....	8
<b>Chapter 2: Structure based L-peptide inhibitors of SEVI fibrils</b> .....	53
<b>Chapter 3: Structural-Based Design of Non-Natural Amino Acid Inhibitors of Amyloid Fibril Formation</b> .....	82
<b>Chapter 4: Small Molecules that Disaggregate Preformed Fibrils</b> .....	122

## LIST OF FIGURES

<b>Figure 0.1</b> The cylindrin structure of an amyloid oligomer.....	1
<b>Figure 0.2</b> Crystal structure of a fibril .....	4
<b>Figure 1.1</b> Crucial role of glycine in cylindrin formation .....	29
<b>Figure 1.2</b> Mapping of SOD1 disease-related mutations on SOD1 .....	30
<b>Figure 1.3</b> Recombinant SOD1 mutant production .....	31
<b>Figure 1.4</b> Both SOD1 <sup>G93A</sup> and SOD1 <sup>G33V/G93A</sup> form abundant fibrils .....	32
<b>Figure 1.5</b> G33V abolishes the toxicity of SOD1 <sup>G93A</sup> in PC12 neuronal cells. ....	33
<b>Figure 1.6</b> G33V abolishes the toxicity of SOD1 <sup>G93A</sup> in motor neurons .....	34
<b>Figure 1.7</b> Crystal structure of the K11L open barrel .....	35
<b>Figure 1.8</b> Prefibrillar K11L specie is toxic to mammalian cells.....	36
<b>Figure 1.9</b> Concentration dependent toxicity .....	37
<b>Figure 1.10</b> Native MS spectrum of K11L and K11LG6V.....	38
<b>Figure 1.11</b> Segments P11L and K11L are both toxic to mammalian cells.....	39
<b>Figure 1.12</b> Glycine to valine mutation abolishes the toxicity of K11L .....	40
<b>Figure 1.13</b> Fibril diffraction .....	41
<b>Figure 1.14</b> G33V abolishes the toxicity of K11L in motor neurons .....	42
<b>Figure 1.15</b> Model for SOD1 aggregate using the K11L structure as a spine .....	43



<b>Figure 1.16</b> Toxicities of monomer and amyloid preparations .....	44
<b>Figure 1.17</b> G33V does not abolish the toxicity of WT SOD1 in motor neurons s .....	46
<b>Figure 1.18</b> Valine mutation creates steric clash .....	47
<b>Figure 2.1</b> High propensity segment prediction .....	68
<b>Figure 2.2</b> Fibrillation assay of proline mutant .....	69
<b>Figure 2.3</b> Crystal structure of KLIMY .....	70
<b>Figure 2.4</b> Crystal structure of GGVLVN .....	71
<b>Figure 2.5</b> Fibril diffraction and structural model .....	72
<b>Figure 2.6</b> 248PAP286 expression and purification .....	73
<b>Figure 2.7</b> SEVI fibrillation kinetics .....	74
<b>Figure 2.8</b> L-peptide inhibitors against seeded and unseed fibrillation.....	75
<b>Figure 2.9</b> L-peptide inhibitor does not inhibit A $\beta$ fibrillation .....	77
<b>Figure 2.10</b> L-peptide inhibitors prevent SEVI-mediated infection .....	78
<b>Figure 3.1</b> Design and characterization of peptide inhibitors of amyloid fibril formation .....	84
<b>Figure 3.2</b> Designed D-peptide delays tau K12 fibril formation in a sequence specific manner .....	84
<b>Figure 3.3</b> Mechanism of interaction .....	85
<b>Figure 3.4</b> Designed non-natural peptide inhibits 248PAP286 fibril formation.....	85
<b>Figure 3.5</b> D-TLKIVW design model .....	91
<b>Figure 3.6</b> Other designed D-amino acid tau inhibitor models .....	93
<b>Figure 3.7</b> Steric repulsion between d-TLKIVW and VQIVYK .....	95
<b>Figure 3.8</b> Designed d-peptides delay VQIVYK and tau K12 fibril formation .....	96
<b>Figure 3.9</b> D-TLKIVW delays fibril formation in a ratio-dependent manner .....	98
<b>Figure 3.10</b> D-TLKIVW decreases the fibril propagation rate of VQIVYK and the tau constructs K12 and K19 .....	100
<b>Figure 3.11</b> Scrambled inhibitors demonstrate sequence specificity .....	102
<b>Figure 3.12</b> L-TLKIVW diastereomer has little effect on tau fibril formation .....	104

<b>Figure 3.13</b> D-TLKIVW is not a general inhibitor of fibril formation .....	106
<b>Figure 3.14</b> Quantification of nanogold-labeled D-TLKIVW bound to K19 fibrils .....	107
<b>Figure 3.15</b> D-TLKIVW interacts with VQIVYK and K19 fibrils .....	108
<b>Figure 3.16</b> The 500 MHz <sup>1</sup> H spectra of D-TLKIVW and D-LKTWIV .....	109
<b>Figure 3.17</b> D-TLKIVW does not interact with soluble VQIVYK .....	110
<b>Figure 3.18</b> D-TAKIVW has little effect on tau fibril formation .....	111
<b>Figure 3.19</b> Binding orientations of designed inhibitor of 248PAP286 fibril formation .....	113
<b>Figure 3.20</b> The chemical structure of the non-natural peptide WW61 .....	115
<b>Figure 3.21</b> Inhibitory effect of the designed peptide WW61 on seeded 248PAP286 fibril formation .....	116
<b>Figure 3.22</b> Inhibitory effect of the designed peptide WW61 on unseeded 248PAP286 fibril formation .....	117
<b>Figure 3.23</b> The non-natural peptide WW61 shows a ratio-dependent inhibition effect on seeded 248PAP286 .....	118
<b>Figure 3.24</b> Negative-stain electron micrographs show the effect of the inhibitor WW61 on 248PAP286 fibril formation .....	119
<b>Figure 3.25</b> No inhibitory effect of the control peptides GIHKQK and PYKLWN on seeded 248PAP286 fibril formation .....	120
<b>Figure 3.26</b> Control HIV infectivity assays .....	121
<b>Figure 4.1</b> Inhibitors are not active in seminal fluid .....	135
<b>Figure 4.2</b> EGCG is not active in dissolving SEVI fibrils .....	136
<b>Figure 4.3</b> ThT binds strongly to SEVI fibrils .....	137
<b>Figure 4.4</b> Preliminary experiments with various SEVI concentrations .....	138
<b>Figure 4.5</b> SEVI only control .....	139

## LIST OF TABLES

<b>Table 1.1</b> K11L data collection and refinement statistics .....	57
<b>Table 2.1</b> KLIMY data collection and refinement statistics .....	79
<b>Table 3.1</b> Comparison of D-peptides and controls .....	122
<b>Table 3.2</b> GGVLVN data collection and refinement statistics .....	123
<b>Table 4.1</b> List of small molecules that facilitate dissociation of SEVI fibrils .....	140
<b>Table 4.2</b> List of small molecules that stabilize SEVI fibrils .....	143

## Acknowledgements

I would like to begin by acknowledging my advisor and mentor, David Eisenberg. I have learned so much for David for the past six years, both as a scientist and as a person. I want to thank David for his patience, generosity, guidance and continuous support.

I want to thank my other committee members James Bowie, Irvin Chen, Feng Guo and Todd Yeates. I want to thank James Bowie for introducing me to structural biology through my first graduate course at UCLA. I want to thank Irvin Chen for encouraging me and helping me start with my HIV project. I want to thank Feng Guo for his advice when I was stuck on producing my protein target. I want to thank Todd Yeates for making the complex theories of crystallography comprehensible.

Additionally, I would like to thank the dedicated scientists that I have worked with. Cong Liu and Dan Li told me most of the experimental techniques and spent many hours helping me develop my project. Lin Jiang and I have worked together for most of my graduate program, developing inhibitors and building structural models. Mike Sawaya taught me solving crystal structure and Elon Hartman, a good friend and a source of support, taught me everything I know about NMR. I want to also thank Alice Soragni, who helped me with most of the cell work; James Stroud, who taught me EM; Art Laganowsky, who taught me cloning techniques; Lisa Johnson, who taught me peptide synthesis; Magdalena Ivanova who helped me express and purify SOD1; Meytal Landau, Luki Goldshimdt and Marcin Apostol who guided me through the first year in lab; and Rebecca Nelson, Zhefeng Guo, and Minglei Zhao with whom I had many scientific discussions. I want to thank Dan Anderson and Cindy Chau for keeping the lab in order; Duilio Cascio for data collection at the synchrotron trips; Alex Lisker for helping with

computer problems; the staff in the Chemistry and Biochemistry Department and the Molecular Biology Institute for their help during the past six years.

I have also been very fortunate to collaborate with excellent scientists. My HIV project is a collaborative effort with Lin Jiang, Jiyong Park, Onofrio Zirafi, Jan Munch, Robert Damoiseaux, Robert Peterson and Elon Hartman. My SOD1 project is a collaborative effort with Smriti Sangwan, Rebecca Nelson, Cong Liu, Katrina Adam and Ben Novitch. Additionally, I want to thank the Bill and Melinda Gates Foundation, the UCLA-CalTech Joint Center for Translational Medicine, the Edward W. Pauley Foundation and Dr. Audree Fowler for financial support.

I came to the United States from China in high school, speaking very little English. I want to thank the great teachers who have taught me English and American culture at Ithaca High School. I have also made great friends in the Eisenberg lab. I want to thank Weixia Yu, for feeding me with her delicious home cooking; Alice Soragni, Angie Soriaga, Smriti Sangwan, Lorena Saelices Gomez and Alex Bajaj for keeping life fun.

I want to thank Howard for making me a happier and better person. Finally, I want like to thank my parents Lu and Kejian for all the sacrifices they have made for me. To my parents, thank you for always standing behind me with your unconditional love and support.

## VITA

2004 -2007	Research Assistant Department of Pharmaceutical Sciences University of Wisconsin, Madison
2007	Bachelor of Science, Biochemistry University of Wisconsin, Madison
2008 - 2009	Teaching Assistant Department of Chemistry and Biochemistry University of California, Los Angeles
2009	Candidate of Doctor of Philosophy Biochemistry and Molecular Biology University of California, Los Angeles
2009 – 2013	Graduate Student Researcher Department of Chemistry and Biochemistry University of California, Los Angeles

### Publications

1. Zhao A et al. (2011) Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature*. 2011 Jun 15;475(7354):96-100.
2. Massey-Gendel E et al. (2009) Genetic selection system for improving recombinant membrane protein expression in *E. coli*. *Protein Science*. Feb;18(2):372-83.

3. Xiong MP et al. (2007) Poly(aspartate-g-PEI800), a polyethylenimine analogue of low toxicity and high transfection efficiency for gene delivery. *Biomaterials*. Nov;28(32):4889-900.
4. Bae Y et al. (2007) Mixed polymeric micelles for combination cancer chemotherapy through the concurrent delivery of multiple chemotherapeutic agents. *Journal of Control Release*. Oct 8;122(3):324-30.
5. Forrest ML et al. (2006) Lipophilic prodrugs of Hsp90 inhibitor geldanamycin for nanoencapsulation in poly(ethylene glycol)-b-poly(epsilon-caprolactone) micelles. *Journal of Control Release*. Nov 28;116(2).

## Introduction and Dissertation Layout

Amyloidosis is a class of 25+ diseases include Alzheimer's disease, Parkinson's disease, Amyotrophic Lateral Sclerosis, Type II Diabetes, Mad Cow Disease and many more (Westermarck *et al.*, 2005). Recently, P53 aggregation has been shown to cause the development of cancer and an amyloid fibril, described in my thesis work, has been found to enhance HIV transmission (Ano Bom *et al.*, 2012; Munch *et al.*, 2007). All of these conditions share the hallmark of *in vivo* fibril deposition and the fibrils formed by different amyloid proteins share common structural features. They have an elongated, unbranched morphology under electron microscopy and can all be positively stained by histological dyes Congo Red and Thioflavin T (Westermarck *et al.*, 2002). X-ray diffraction revealed that a cross- $\beta$  conformation is present in the core structure of these fibrils (Westermarck *et al.*, 2002).

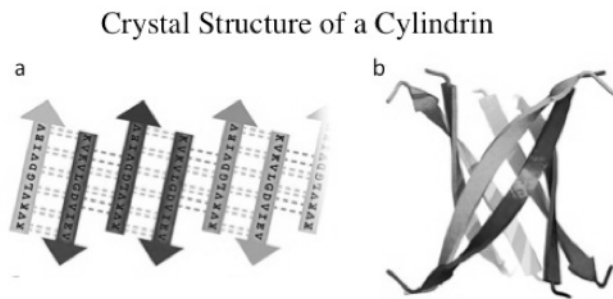
### Oligomers as the toxic species

For a number of amyloidosis, particularly the neurodegenerative diseases, the cytotoxic entities are not the mature fibrils observed *in vivo*. Rather, the disease phenotypes are caused by amyloid oligomers. These oligomers are soluble, lower molecular weight species formed along the fibrillation pathway (Eisenberg and Jucker, 2012). A number of studies have support the role of oligomers in diseases. For example, in both Alzheimer's and ALS patients, the development of the disease symptoms have been observed before the fibrils could be detected. Soluble oligomers formed *in vitro* have shown more cytotoxicity than mature fibrils (Johnston *et al.*, 2000; Karch *et al.*, 2009; Wang *et al.*, 2003; Wang *et al.*, 2002). More over, small molecules that stabilize the fibrils, shifting the equilibrium from oligomers to mature fibrils have shown to reduce cytotoxicity (Bieschke *et al.*, 2011). Collectively these results support that in many of the



neurodegenerative diseases, the low molecular weight, soluble oligomers are the toxic entity (Broersen *et al.*, 2010; Caughey *et al.* 2003).

In 2012, Arthur Laganowsky *et al.* published the first atomic structure of a hexamer by an amyloid-forming segment from the amyloid protein  $\alpha$ -B-Crystalline (Figure 1). The structure is a tightly packed barrel with six anti-parallel  $\beta$ -strands. Each strand has a strong and a weak



**Figure 1. The cylindrin structure of an amyloid oligomer**

interface due to the different number of hydrogen bonds to the neighboring strands. The interior of the cylindrin is formed entirely by hydrophobic residues. We termed this structure cylindrin.

Laganowsky *et al.* further demonstrated that the oligomer formed by the segment resembles properties of the oligomers formed by full-length proteins. It is toxic to mammalian cells and can be recognized by A11, a conformational antibody selectively binds to toxic oligomers (Kayed *et al.* 2003)

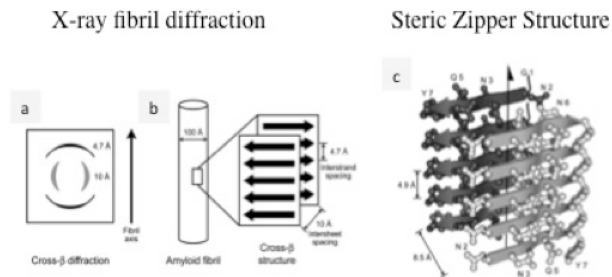
Chapter 1 describes my work on a cellular protein, superoxide dismutase I (SOD1). Mutations on SOD1 lead to motor neuron degeneration and the development of familial amyotrophic lateral sclerosis (fALS). Previous work has shown that this cytotoxicity is likely to be caused by oligomers of  $\Delta$ SOD1 (Karch *et al.*, 2009). This chapter is a manuscript in preparation on the identification of a region in  $\Delta$ SOD1 that is crucial for its toxicity. According to the ALS Database, 163 disease-related mutations have been identified across this protein of 153 residues. The longest conserved segment is a seven-residue region near the N-terminus (30-KVWGSIK-36). Building on the cylindrin structure, Rebecca Nelson, Luki Goldschmidt and Mike Sawaya has developed an algorithm to predict regions in other amyloid proteins that have

high propensity to form such structures. When applying this Rosetta-Profiling algorithm, 30-KVWGSIK-36 resides in an 11-residue segment which showed high propensity to form a cylindrin-like structure. Thus, we hypothesize that this region is important for  $\Delta$ SOD1 toxicity by enabling the toxic oligomer formation.

To test if this region is important for SOD1 toxicity, we made a single G33V mutation, engineered to disrupt the formation of a cylindrin like oligomer. We found that this point mutation abolishes the toxicities of both full length SOD1 and the high propensity segment in isolation, while maintaining their ability to form abundant fibrils. These results suggest that the mutation selectively knocks out oligomer formation and that cylindrin-like oligomers may be responsible for SOD1 toxicity. We further solved the crystal structure of this 11-residue segment assembled into an open barrel like structure. Upon closer look, the glycine to valine mutation would create major steric clashes, prohibiting such structure from forming. This supports our hypothesis that this N-terminal region is important for SOD1 toxicity for its role in enabling the toxic oligomer formation. Upon further validation *in vivo*, this region could be an important therapeutic target for fALS.

### **Fibrils as the toxic species**

In a subset of amyloidosis, including many systematic disorders, fibrils are the toxic species that causes the disease phenotypes (Eisenberg and Jucker, 2012). Inhibition of fibril formation thus offers a viable therapeutic strategy. In 2005, Rebecca Nelson *et al.* obtained the first atomic detail of a cross- $\beta$  fibril structure formed by an amyloidogenic peptide from the yeast prion protein Sup35 (Figure 2) (Nelson *et al.*, 2005). Building on this structure, our lab further developed a computational approach to predict fibril-forming regions in other amyloid proteins



**Figure 2. Crystal structure of a fibril**

(Thompson *et al.* 2006). This work laid a foundation for my thesis project, developing structure-based amyloid inhibitors.

Chapters 2 through 4 discuss my work on the amyloid fibril termed Semen-Derived Enhancer of Viral Infection (SEVI).

This fibril is formed by the aggregation of a peptide segment generated from prostatic acidic phosphatase, an extracellular protein that is present in semen in milligram quantities. The amyloid fibrillar form of this peptide drastically enhances the HIV infection rate (Munch *et al.*, 2007).

In Chapter 2, I discussed the structure study of the SEVI fibrils and my work on developing inhibitors against SEVI fibril formation. I have included crystal structures of two regions within the SEVI peptide revealing the potential target for inhibition. In collaboration with Lin Jiang, we have predicted a SEVI fibril model incorporating the crystal structure and biochemical data. Based on this structure, we developed inhibitors that are designed to cap the growing end of the fibrils and prevent the formation of mature fibrils. This chapter includes five L-peptide inhibitor candidates from three rounds of computational design and experimental testing. In our later work, we incorporated non-natural amino acid to enhance the activities of the inhibitors. The non-natural inhibitor is discussed in Chapter 3.

Chapter 3 is a reprint of the publication Structure-based Design of Non-natural Amino-acid Inhibitors of Amyloid Fibril Formation. Sievers SA\*, Karanicolas J\*, Chang HW\*, Zhao A\*, Jiang L\*, Zirafi O, Stevens JT, Münch J, Baker D, Eisenberg D. Nature. 2011 Jun 15;475(7354):96-100. We have applied this approach to two amyloid proteins. Stuart Sievers,

John Karanicolas and Howard Change developed D-peptide inhibitors against Tau fibril formation, a condition lead to Alzheimer's and Parkinson's diseases. Lin Jiang and myself developed inhibitors with non-natural amino acid derivatives against SEVI fibril formation. Different from chapter 2, this chapter is a more in-depth discussion and characterization specifically on the SEVI inhibitor WW61, which showed the most promising inhibition activity.

Chapter 4 reveals the unpublished work on the search for small molecules that are able to disaggregate pre-existing fibrils. After the initial work on the SEVI inhibitors, we have begun to learn that the SEVI fibrils are likely to have formed in the seminal fluid before ejaculation. Infectivity assay from our collaborator's laboratory showed that "fresh" semen samples had already exhibited the HIV enhancing properties (Kim *et al.*, 2010). As fibrils are tightly packed, stable species, *de novo* design of fibril disaggregators is extremely challenging. Instead, we adapted a high-throughput screening approach. We developed a fluorescent-based assay and screened 2,000 natural compounds for their ability to disaggregate pre-formed fibrils. From this first round of screening, we have obtained a subset of compounds with various activities. While the activities need to be optimized to have therapeutic values, this set of inhibitors may offer a starting point for further design of small molecule inhibitors.

During my graduate study, I have investigated two amyloid systems. In the SEVI system, I have developed two strategies for inhibitor development. In the SOD1 system, I have identified an important site that is potentially responsible for the toxicity of this protein. The results of these projects provide valuable information for further therapeutic development.

## References

Ano Bom AP, Rangel LP, Costa DC, de Oliveira GA, Sanches D, Braga CA, Gava LM, Ramos CH, Cepeda AO, Stumbo AC, De Moura Gallo CV, Cordeiro Y, Silva JL. (2012)

Mutant p53 aggregates into prion-like amyloid oligomers and fibrils: implications for cancer. *J Biol Chem.* 287(33):28152-62.

Bieschke J, Herbst M, Wiglenda T, Friedrich RP, Boeddrich A, Schiele F, Kleckers D, Lopez del Amo JM, Grüning BA, Wang Q, Schmidt MR, Lurz R, Anwyl R, Schnoegl S, Fändrich M, Frank RF, Reif B, Günther S, Walsh DM, Wanker EE. (2011). Small-molecule conversion of toxic oligomers to nontoxic  $\beta$ -sheet-rich amyloid fibrils.

*Nat Chem Biol.* 2011 Nov 20;8(1):93-101.

Broersen, K., Rousseau, F., and Schymkowitz, J. (2010). The culprit behind amyloid beta peptide related neurotoxicity in Alzheimer's disease: oligomer size or conformation? *Alzheimer's research & therapy* 2, 12.

Caughey, B., and Lansbury, P.T. (2003). Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annual review of neuroscience* 26, 267-298.

Eisenberg D, Jucker M. (2012) The amyloid state of proteins in human diseases. *Cell.* 148(6):1188-203.

Karch, C.M., Prudencio, M., Winkler, D.D., Hart, P.J., and Borchelt, D.R. (2009). Role of mutant SOD1 disulfide oxidation and aggregation in the pathogenesis of familial ALS. *Proceedings of the National Academy of Sciences of the United States of America* 106, 7774-7779.

Laganowsky A, Liu C, Sawaya MR, Whitelegge JP, Park J, Zhao M, Pensalfini A, Soriaga AB, Landau M, Teng PK, Cascio D, Glabe C, Eisenberg D. (2012). Atomic view of a toxic amyloid small oligomer. *Science.* 335(6073):1228-31

Münch J, Rucker E, Ständker L, Adermann K, Goffinet C, Schindler M, Wildum S, Chinnadurai R, Rajan D, Specht A, Giménez-Gallego G, Sánchez PC, Fowler DM, Koulov A, Kelly JW, Mothes W, Grivel JC, Margolis L, Keppler OT, Forssmann WG, Kirchhoff F. (2007) Semen-derived amyloid fibrils drastically enhance HIV infection. *Cell.* 131(6):1059-71.

Nelson R, Sawaya MR, Balbirnie M, Madsen A, Riek C, Grothe G, Eisenberg D. (2005) Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature.* 435, 773-778.

Nelson R, Eisenberg D. (2006) Structural models of amyloid-like fibrils. *Adv Protein Chem.* 73:235-82.

Nelson R, Sawaya MR, Balbirnie M, Madsen AØ, Riekel C, Grothe R, Eisenberg D. (2005). Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* 435, 773–778

Kayed R, Head E, Thompson JL, McIntire TM, Milton SC, Cotman CW, Glabe CG. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*. Apr 18;300(5618):486-9.

Thompson MJ, Sievers SA, Karanicolas J, Ivanova MI, Baker D, Eisenberg D. (2006). The 3D profile method for identifying fibril-forming segments of proteins. *Proc. Natl Acad. Sci. USA* 103, 4074–4078

Westermarck, P., Benson, M. D., Buxbaum, J. N., Cohen, A. S., Frangione, B., Ikeda, S., Masters, C. L., Merlini, G., Saraiva, M. J., and Sipe, J. D. (2005). Amyloid: toward terminology clarification. Report from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid* 12, 1-4.

Wetzel, R. (2002). Ideas of order for amyloid fibril structure. *Structure (Camb.)* 10, 1031-6.

Johnston, J.A., Dalton, M.J., Gurney, M.E., and Kopito, R.R. (2000). Formation of high molecular weight complexes of mutant Cu, Zn-superoxide dismutase in a mouse model for familial amyotrophic lateral sclerosis. Proceedings of the National Academy of Sciences of the United States of America 97, 12571-12576.

Karch, C.M., Prudencio, M., Winkler, D.D., Hart, P.J., and Borchelt, D.R. (2009). Role of mutant SOD1 disulfide oxidation and aggregation in the pathogenesis of familial ALS. Proceedings of the National Academy of Sciences of the United States of America 106, 7774-7779.

Kim, K.A., Yolamanova, M., Zirafi, O., Roan, N.R., Staendker, L., Forssmann, W.G., Burgener, A., Dejucq-Rainsford, N., Hahn, B.H., Shaw, G.M., *et al.* (2010). Semen-mediated enhancement of HIV infection is donor-dependent and correlates with the levels of SEVI. *Retrovirology* 7, 55.

Wang, J., Slunt, H., Gonzales, V., Fromholt, D., Coonfield, M., Copeland, N.G., Jenkins, N.A., and Borchelt, D.R. (2003). Copper-binding-site-null SOD1 causes ALS in transgenic mice: aggregates of non-native SOD1 delineate a common feature. *Human molecular genetics* 12, 2753-2764.

Wang, J., Xu, G., and Borchelt, D.R. (2002). High molecular weight complexes of mutant superoxide dismutase 1: age-dependent and tissue-specific accumulation. *Neurobiology of disease* 9, 139-148.

# Chapter 1

This Chapter is a draft of a manuscript tentatively titled, “The critical role of Gly 33 in SOD1<sup>G93A</sup> oligomerization and toxicity”

## Abstract

Familial amyotrophic lateral sclerosis (fALS) is a neurodegenerative disease often associated with the conversion of mutant superoxide dismutase I (SOD1) to amyloid form. Over 160 such mutations have been identified throughout this 153-residue protein with no apparent pattern explaining toxic gain of function. Here, we report a segment in SOD1 (28-PVKVWGSIKGL-38) whose integrity is essential for the toxicity of a disease mutant SOD1<sup>G93A</sup>. Previous work showed that the pathogenic entities in fALS are the soluble  $\Delta$ SOD1 oligomers. Using a recently published structure of a toxic amyloid oligomer as a 3D profile, we detected that residues 28 to 38 in SOD1 has high propensity to form a toxic oligomer. Additionally, this segment has not found to be mutated in any patient presenting with SOD1-linked ALS, suggesting that it may be required for the formation of toxic species. We found that a single mutation within this segment (G33V) is able to abolish the toxicity of both the segment itself and full length SOD1<sup>G93A</sup> for mammalian cells including motor neurons. We further determined the crystal structure of the segment elucidating the critical role of Gly 33 in enabling the formation of toxic oligomers.

## Introduction

Amyotrophic lateral sclerosis (ALS) leads to the degeneration of motor neurons in the central nerve system. ALS patients develop progressive paralysis of muscles leading to difficulties swallowing and breathing and few survive beyond 3 – 5 years of disease onset. ALS can be sporadic or familial. Familial ALS (fALS) accounts for ~20% of ALS of cases, of which ~10% are linked to mutations in the gene encoding the superoxide dismutase I (SOD1) protein (Chattopadhyay and Valentine, 2009; Wong et al., 1995). Over 160 SOD1 mutations have been identified in fALS patients (Chattopadhyay and Valentine, 2009). A subset of the mutations appears to destabilize the native protein structure leading to the aggregation of mutant SOD1 ( $\Delta$ SOD1) into amyloid deposits (Valentine et al., 2005). However, most of the mutations are distributed throughout the SOD1 protein with no obvious pattern for explaining toxic gain of function.

The development of therapeutics for fALS has been challenging without understanding of the toxic entity at the atomic level. Mice models suggest that the development of the disease follows a gain-of-function mechanism. SOD1<sup>G93A</sup> expressing mice, but not SOD1 null mice or mice expressing WT SOD1, develop ALS-like symptoms. This suggests the development of fALS is linked to  $\Delta$ SOD1 forming amyloid species (Mead et al., 2011; Reaume et al., 1996). Additionally, disease mice develop ALS-like symptoms before the accumulation of large fibrillar aggregates, suggesting the toxic species in fALS is a prefibrillar structure formed by  $\Delta$ SOD1 (Johnston et al., 2000; Karch et al., 2009; Wang et al., 2003; Wang et al., 2002).

A range of studies suggests that the toxic entities in neurodegenerative diseases are low molecular weight, soluble oligomers, rather than mature fibers. Although the atomic details of the oligomers are not clear, a number of studies have support the roles of these species in



diseases. For example, in both Alzheimer's and ALS patients (Karch et al., 2009; Lesne et al., 2006; Tomic et al., 2009), the development of the disease symptoms have been observed before the fibrils could be detected. Soluble oligomers formed *in vitro* have shown more cytotoxicity than mature fibrils. Moreover, small molecules that stabilize the fibrils, shifting the equilibrium from oligomers to mature fibrils have shown to reduce cytotoxicity (Haass and Selkoe, 2007; Lesne et al., 2006; Ono et al., 2009; Shankar et al., 2008). Collectively these results support that in many of the neurodegenerative diseases, the low molecular weight, soluble oligomers are the toxic entities.

Amyloid oligomers have been found present in an ensemble of sizes and shapes (Eisenberg and Jucker, 2012). These species tend to be transient and present at low concentration. Furthermore, they are in equilibrium with the monomers, fibrils and with each other (Chiti and Dobson, 2006, 2009; Eisenberg and Jucker, 2012), complicating biochemical isolation and characterization. Bypassing these difficulties, Laganowsky *et al.* used a computation method to identify an amyloid-forming segment (K11V) from the amyloid protein  $\alpha$ B-crystallin (Laganowsky et al., 2012). This segment forms abundant fibrils at elevated temperature upon vigorous shaking. However, under non-agitated conditions in physiological conditions, the segment assembles into a hexameric oligomer. Laganowsky *et al.* further showed that this oligomer exhibits properties of oligomers formed by full-length proteins: toxicity to mammalian cells and recognition by A11, a conformational antibody selectively binds to amyloid oligomers (Kay et al., 2003). Laganowsky *et al.* determined the crystal structure of K11V oligomer, the first view at the atomic level of an intermediate amyloid species. It shows a tightly packed barrel formed by six anti-parallel  $\beta$ -strands, distinct from previous structures in the Protein Data Bank. Each  $\beta$ -strand of the cylinder forms a strong and a weak interface due to the different number of

hydrogen bonding to its two neighboring strands. The interior of the cylindrin is formed entirely by hydrophobic residues (Figure 1).

Here, I report a segment around Gly 33 in the SOD1 protein that is necessary for the toxicity of the fALS disease mutant SOD1<sup>G93A</sup>. According to the ALS database ([http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene\\_id=SOD1](http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SOD1)), 163 disease-related fALS mutations have been identified throughout this protein of 153 residues. We observe that the longest conserved segment is a seven-residue stretch (30-KVWGSIK-36) (Figure 2) that lies within the cylindrin-compatible segment identified by our 3D Profile algorithm, 28-PVKVWGSIKGL-38, abbreviated P11L. We hypothesize that Gly 33 may be an important residue for  $\Delta$ SOD1 toxicity because it enables the formation of a cylindrin-like structure. To test if  $\Delta$ SOD1 toxicity is caused by the formation of a cylindrin-like structure, we engineered a glycine to valine mutation in the background of the disease-related SOD1 mutant G93A. This single G33V mutation abolishes the toxicity of SOD1<sup>G93A</sup> in both PC12 neuronal cells and ES-derived motor neurons. We determined the crystal structure of the 11-residue segment revealing the potential role of Gly 33 in enabling SOD1 oligomerization.

## **Results**

### **Prediction of cylindrin forming segments in superoxide dismutase I.**

Numerous studies have shown that amyloid oligomers are a mixture of various sizes (Broersen et al., 2010; Caughey and Lansbury, 2003; Fandrich, 2012; Lashuel et al., 2002; Last et al., 2011; Orte et al., 2008; Sakono and Zako, 2010). Despite this polymorphic nature, many toxic oligomers, including the K11V cylindrin bind to the conformational antibody A11, indicating that they share common structural features (Glabe 2006). The cylindrin structure has

enabled us to identify segments in other amyloid proteins that could form toxic oligomers (Glabe, 2006; Kaye et al., 2003; Tomic et al., 2009). This strategy has shown to be successful with predicting fibril structures (Goldschmidt et al., 2010; Sawaya et al., 2007; Thompson et al., 2006). Using this method, Rebecca Nelson, Mike Sawaya and Luki Goldschmidt identified 11-residue sequence segments from other amyloid proteins with cylindrin-forming compatibility. Since the K11V cylindrin is so far the only atomic structure of an amyloid oligomer, we developed a computational approach based on this structure to explore 11-residue segments in other amyloid proteins that can form cylindrin or cylindrin-like structures. We have identified 26 putative segments in 6 amyloid proteins. Examining the sequence pattern of these high propensity segments, we found that many share a characteristic glycine at the central position. Revisiting the cylindrin structure, we hypothesized that this central glycine is critical for formation of a cylindrin-like oligomer structure (Figure 1).

Most cylindrin-compatible sequence segments, including the K11V segment of alpha-B-crystallin, contain a glycine residue at the central position, residue 6. Revisiting the K11V structure, we noticed that if one were to unroll the barrel, the interior of the barrel sheet contains a knob-and-hole motif (figure 1). The “knobs” are formed by Val 4 and Val 8, and the “holes” are created by Gly 6. In the interior of the barrel, the “knobs” fit into the “holes” and any residue that is larger than glycine is likely to disrupt this tight packing and preclude the formation of a cylindrin barrel. This suggests that a glycine at position 6 of an 11-residue cylindrin-compatible segment may be necessary for the formation of toxic oligomers.

**Residue 30 – 36 is the longest stretch of conserved sequence in the collection of ALS-linked SOD1 mutant sequences.**

In the SOD1 protein, our application of the 3D Profile method resulted in two overlapping segments clustered near the N-terminus of the protein. Searching through the ALS database ([http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene\\_id=SOD1](http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SOD1)), we discovered that these segments are centered on 30-KVWGSIK-36, a highly conserved segment in all SOD1 disease-related mutants. It is in fact the longest stretch of conserved sequence identified so far (Figure 2). In our 3D Profile prediction, segment 29-PVKVWGSIKGL-39, with a conserved glycine at the central position, has a high propensity to form a cylindrin-like oligomer structure. Combining these two observations, we hypothesize that SOD1 toxicity is directly related to a structure that is formed by the 11-residue segment. To test this hypothesis, we designed glycine mutation, based on the predicted structural model of P11L cylindrin. The G33V mutation does not break  $\beta$ -strands or manipulate charge; it is simply a slightly bulkier residue that would disrupt the tight packing observed in cylindrin. We made recombinant proteins SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup>. Native SOD1<sup>G93A</sup> protein appears to have WT like properties and SOD1<sup>G93A</sup> mice model is one of the few well-studied fALS disease models (Valentine et al., 2005).

### **Glycine 33 is not crucial for the SOD1 expression or of a native-like structure of SOD1<sup>G93A</sup>.**

An alternative explanation for the high conservation of Gly 33 in disease mutants is that this residue is crucial for the cellular production of the SOD1 protein. The native structure of the SOD1 contains an eight-stranded  $\beta$ -barrel and that residue 30-37 spans an entire  $\beta$  strand ( $\beta$ 3) in the native structure (Valentine et al., 2005). In our work, we have adapted the E. coli expression system to produce recombinant proteins. We do not observe significant difference in the expression levels of SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup>. Additionally, both constructs form stable dimers that were purified by HPLC – SEC and confirmed with dynamic light scattering

experiments (Figure 3). The glycine to valine mutation did not have major disruptive effect on protein production. In fact, the two constructs behaved very similarly during the expression and purification process, suggesting that Gly 33 is not crucial for the cellular production of the SOD1 protein.

### **SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> both form abundant amyloid fibrils.**

We expressed recombinant proteins SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup>, both of which form abundant amyloid fibrils after 24 hours of incubation at 37°C with continuous shaking (Figure 4). The fibrillation assay was carried out at partially denaturing conditions described by published protocols (Chattopadhyay and Valentine, 2009). At the end of the fibrillation assay, we pelleted the fibrils for 30 minutes at 13,000 rpm and tested the protein concentration in the pellet. We found no major difference in the fibril concentration between the two preparations (Figure 4d). EM images showed that both constructs formed abundant fibrils.

### **SOD1<sup>G93A</sup> is toxic to PC12 cells and SOD1<sup>G33V/G93A</sup> rescues the toxicity.**

We next compared the toxicity of the SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> amyloid preparations in mammalian cells. In these amyloid preparations, the various amyloid species such as monomers, oligomers and fibrils are present in equilibrium. We added these amyloid preparations to the cell medium and evaluated toxicity using the MTT reduction assay. SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> preparations were added to cells and incubated for 24 hours at 37°C. We found that while both SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> form abundant amyloid fibrils, only SOD1<sup>G93A</sup> is toxic, reducing the cell viability to roughly 65% of the control. SOD1<sup>G33V/G93A</sup> amyloid preparations, while also exhibiting abundant fibrils showed little toxicity both at 4 μM and at 8

$\mu\text{M}$  (Figure 5). All the protein concentrations discussed in this manuscript refer to the corresponding monomer concentrations.

### **SOD1<sup>G93A</sup> is toxic to motor neurons and SOD1<sup>G33V/G93A</sup> rescues the toxicity.**

Since fALS is a motor neuron disease, we next tested the effect of G33V mutation with embryonic stem cell (ES) derived motor neuron cultures. These motor neurons carry a GFP reporter gene, and toxicities can be directly observed from the morphology of the cells fluorescent microscopy. We added the amyloid preparations into the cell medium and compared the toxicity of SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> after a 24 hour incubation. At 12  $\mu\text{M}$  protein concentration, SOD1<sup>G93A</sup> showed a clear toxic effect with the absence of axons and shrinking cell bodies. In contrast, at the same concentration, SOD1<sup>G33V/G93A</sup> appeared to be non-toxic to motor neurons in our assays with cell morphology essentially identical to that of the control (Figure 6). Thus, from both the MTT reduction assay and the direct observation of cell morphology, we find that a single mutation G33V, while not disrupting fibril formation, is able to rescue the toxicity of SOD1<sup>G93A</sup>. This result is consistent with previous studies that suggest SOD1 toxicities are likely to be caused by low molecular weight oligomers rather than the mature fibrils (Citation).

### **The K11L crystal structure shows an open barrel with out-of-registered $\beta$ strands.**

To further test our hypothesis that SOD1 toxicity is directly related to a structure that is formed by the 11-residue segment, we determined crystal structure of this segment and tested the ability of the segment itself to form toxic structures. To aid crystallization, we replaced the initial proline of the 11-residue segment by lysine, termed K11L. This K11L peptide was chemically synthesized, crystallized in hanging drops and found to diffract to 2.1Å resolution. The structure

was determined by SIRAS with phases provided by KI soaking. The asymmetric unit of the structure is a continuous chain of  $\beta$  strands in an open cylindrin-like barrel that executes a corkscrew-like helix with  $8_1$  symmetry (Figure 7). The strands are anti-parallel and are out-of-register by two residues. Each strand forms a strong interface to one of its two neighbors in the corkscrew with eight inter-chain hydrogen bonds and a weak interface to its other neighbor with six inter-chain hydrogen bonds. Additionally, the valine residues form a hydrophobic inner groove, while the Lys 3 residues align on the outer surface, forming a stripe of positive charges on the open barrel. This structure share similar patterns with both the cylindrin and the out-of-register fibril structure (Laganowsky et al., 2012; Liu et al., 2012).

#### **K11L segment in isolation is toxic to cells.**

We found that the K11L segment showed significant toxicity to mammalian cells in a concentration dependent manner (Figure 8a and Figure 9). The K11L peptide was synthesized and dissolved in Tris buffer and incubation at 37°C without agitation. Interestingly, we found that a K11L sample prepared at a different condition, overnight agitation at 37 °C, forms abundant amyloid fibrils after. However, these fibril preparations showed toxicity only at 400  $\mu$ M (Figure 8a). To examine the species present in the non-agitated preparations, we conducted native mass spectroscopy. The spectrum indicates the presence oligomers with a range of sizes from dimers to dodecamers (Figure 10). We have also observed spherical species under EM, but not fibrillar species (Figure 8c). We further verified that the wildtype sequence P11L also showed toxicity to mammalian cells (Figure 11). Thus, the observed toxicity was not an artifact of the proline to lysine mutation. This segment insolation is toxic to mammalian cells.

### **The glycine to valine mutation abolishes K11L toxicity.**

In contrast to the wild type K11L segment, the segment carries the valine mutation (K11L<sup>G6V</sup>) forms a gel-like consistency immediately upon dissolving. Electron micrographs show abundant fibril formation immediately upon dissolving in buffer solution (Figure 12). X-ray fibril diffraction on the freshly dissolved K11L<sup>G6V</sup> sample showed that the most abundant species observed under electron microscopy has the characteristic cross- $\beta$  diffraction pattern observed from other amyloid fibrils (Figure 13). The SOD1 construct K11L<sup>G6V</sup> designed to prevent cylindrin formation is not toxic to cells whether as freshly dissolved or aged (Figure 10, Figure 12a). We then compared the toxicities of K11L and K11L<sup>G6V</sup> in ES derived motor neurons using MTT reduction assay. The K11L peptide is toxic to motor neurons and K11L<sup>G6V</sup> abolishes toxicity (Figure 14). These results with the high propensity peptide segment are consistent with that of the full-length SOD1 proteins. Both of these studies showed that Gly 33 is important for the toxicity of SOD1<sup>G93A</sup>.

### **Effect of G33V on WT SOD1**

We have also tested the effect of G33V on wildtype SOD1 protein. We have found that both SOD1 and SOD1<sup>G33V</sup> are able to form amyloid fibrils under partially destabilizing conditions with the presence of TCEP and EDTA. While the G33V abolishes the toxicity of WT SOD1 in PC12 cells (Figure 16), it does not have apparent effect on motor neurons (Figure 17). Both WT SOD1 and SOD1<sup>G33V</sup> are toxic to motor neurons when added extracellularly. Additionally, toxicity experiments on freshly dissolved SOD1 proteins indicated that both WT and G93A already appeared to be toxic to cells. Although after 24 hours of agitation, both preparations became more toxic. Interestingly, G33V mutation had no effect in the toxicity of



freshly dissolved WT protein, but had a mild rescue effect on G93A construct (Figure 16). This seemed to suggest that although both the WT and G93A constructs are toxic to HeLa cells, the toxicity mechanisms may be different.

## **Discussion**

### **Gly 33 is important for the formation of a cylindrin-like oligomer**

Our Rosetta-profiling algorithm predicted that an 11-residue segment 28-PVKVWGSIKGL-38 in SOD1 has high propensity to form a cylindrin like oligomer (Figure 1). In the cylindrical oligomers, the central glycine (Gly 33) appears to be a crucial residue. Any amino acid with a larger side chain would create steric hindrance that would disrupt the tight packing of the predicted oligomer structure (Figure 1b). This region also covers the only  $\beta$  strand that has not found to be mutated (Figure 2). All other strands have multiple mutations, two of which are mutated at all positions. Even though the specific role of the individual mutation is not clear, these disease mutants are all capable of forming a structure that is cytotoxic suggesting that this region is important for the formation of toxic  $\Delta$ SOD1 oligomers.

### **Extracellular SOD1 aggregates are toxic to mammalian cells and G33V abolishes this toxicity**

To assess the role of Gly 33 in  $\Delta$ SOD1 toxicity, we engineered a glycine to valine mutation to prevent the formation of a cylindrin-like oligomer. We tested the effect of G33V mutation on the toxicities of four constructs, full-length proteins SOD1<sup>G93A</sup>, SOD1<sup>G33V/G93A</sup> and segments in isolation K11L and K11L<sup>G6V</sup>. Although SOD1 is an intracellular protein selectively toxic to motor neurons(Chattopadhyay and Valentine, 2009), so far there has not been well-

established cell-line for toxicity studies with intracellular expressed SOD1. As such, we tested cytotoxicity of these amyloid preparations by adding the protein to the cell medium. Similar assays have been conducted with many amyloid proteins (Laganowsky et al., 2012). Additionally, recent work suggested that many amyloid proteins (Kfoury et al., 2012; Prusiner, 2012), including SOD1 (Grad et al., 2011), have prion like properties that the toxic species are capable of propagating from cell to cell. Here, we have shown that G33V mutation does not disrupt fibril formation that all four constructs form abundant amyloid fibrils. However, in our assay, only SOD1<sup>G93A</sup> and K11L amyloid preparations are toxic to mammalian cells, including motor neurons, while SOD1<sup>G33V/G93A</sup> and K11L<sup>G6V</sup> abolishes the toxicities. These results suggest that Gly 33 is important for observed toxicity of SOD1<sup>G93A</sup>.

### **Pre-fibrillar species are more toxic than fibrils**

Our study of K11V in isolation shows that the fibrillar species is not responsible for the observed cytotoxicity. The agitated K11L sample, which forms abundant fibrils is much less toxic than the non-agitated K11L sample where spherical species was observed. Additionally, K11L<sup>G6V</sup>, which forms abundant fibrils immediately upon dissolving, abolishes the toxicity. This is consistent with the previous studies that the proteins that have higher propensity to aggregate are less toxic (Finder et al., 2010). Moreover, we found that while K11L<sup>G6V</sup> is non-toxic at 800  $\mu$ M, when testing toxicity with a serial dilution of the K11L<sup>G6V</sup>, we begin to observe toxicity at a much lower concentration of 80  $\mu$ M. We speculate that this concentration is below a critical concentration that shifts the equilibrium to oligomer formation. Together, these results suggest that the oligomeric species, rather than the mature fibril, is the toxic entity and Gly 33 is necessary for toxicity because it enables oligomer formation.

## **Crystal structure of K11L**

To understand the effect of the G33V at an atomic level, we determined the crystal structure of K11L. This segment forms an open-barrel with out-of-registered  $\beta$  strands, resembling a deformed cylindrin. In the interior of the open-barrel, Val 2 - Gly 6 -Val 4 from adjacent strands form an apolar stripe. The small size of Gly 6 enables Val 2 and Val 4 to form van der waals interactions. When the out-of-register beta sheet rolls up into a barrel, Val 2 and Val 4 pack tightly against Gly 6. The G6V mutation creates significant steric hindrance with valines lining up in the interior of the barrel, precluding the formation of the open barrel (Figure 18). The packing in the interior of the structure is consistent with our observation that K11L<sup>G6V</sup> forms fibrils significantly faster than K11L. We hypothesize that the glycine to valine mutation abolishes cytotoxicity by preventing the formation of the toxic oligomer, which contains a barrel like structure.

## **The relationship of 3D structure to toxicity**

This study offers a molecular explanation for the relationship of Gly residues to amyloid toxicity. A number of studies have shown a relationship between the glycine residues and toxicity of amyloid proteins. For example, in the amyloid protein IAPP, patients carry the S20G mutation develop early onset, more severe form of the noninsulin-dependent diabetes mellitus (Sakagashira et al., 2000). Meanwhile, the NNFGAIL segment in the wild type IAPP sequence was believed to be important for the development of the disease (Jaikaran and Clark, 2001). The relationship between the glycine residues and toxicity has also been demonstrated with Amyloid- $\beta$  proteins. Studies have shown that the C-terminal glycines in the A $\beta$  sequences are important for the cytotoxicity. A G37L mutation in A $\beta$  significantly reduced A $\beta$  toxicity in various toxicity

assays, while not disrupting fibril formation (Fonte et al., 2011; Harmeier et al., 2009; Kanski et al., 2002). This observation is very similar to the SOD1 system that we have studied here.

Although the proposed mechanism was different, these authors have also come to the conclusion that a particular glycine residue is important for the formation of toxic A $\beta$  oligomers.

### **Out-of-register amyloid species are more toxic than in-registered fibrils**

While the open barrel appears to differ from the cylindrin structure at first glance, the two structures share structural similarities, namely antiparallel  $\beta$  strands in an out-of-registered arrangement with a strong and a weak interface. It is noteworthy that in-registered amyloid fibers appear to be more common than out-of-register amyloid species, and that many studies (Chiti and Dobson, 2006; Haass and Selkoe, 2007; Sakono and Zako, 2010) have suggested that minor species are responsible for toxicity. This link of a minor species to toxicity is supported by our work, particularly the study with the target segment in isolation, where the mature fibril preparations are non-toxic to both Hela cells and motor neurons. Previous work has demonstrated that there may be an alternative pathway that leads to the antiparallel out-of-register species including out-of-register oligomers and out-of-register fibrils. Solid state NMR has shown that A $\beta$  mutant protein can associate into anti-parallel arrangement (Qiang et al., 2012; Tycko et al., 2009). Our lab has obtained crystal structures of an amyloidogenic segment from  $\beta$ 2-microglobulin that formed both parallel in-registered fibrils and anti-parallel out-of-registered fibrils. The out-of-register fibrils are more toxic than the in-registered fibrils formed by the same segments (Liu et al., 2012).

**K11V mutation is likely to disrupt not just the cylindrin formation, but also other toxic oligomers with structural similarities.**

The G33V mutation is likely to disrupt multiple oligomers that contain a barrel-like structure. We know that the valine mutation would disrupt the formation of cylindrin and the open barrel structure reported in this work. Amyloid oligomers are polymorphic in nature and observed toxicity is likely to be caused by more than one species. Since the cylindrin and the open barrel are the only two prefibrillar structures reported so far, it is difficult to assess the specific effect of the this mutation on other toxic structures. However, the observation that this single point mutation abolishes the toxicity of the SOD1<sup>G93A</sup> strongly suggests that the mutation is likely to disrupt more than one toxic species. Examination of the structures of K11L and cylindrin suggests that the valine mutation could prevent formation of multiple barrel like structures. When an oligomer structure is formed by identical segments from each protein subunit arranged in an out-of-registered fashion, the same residues align in the interior of the barrel. In the P11L sequence (PVKVGSIKGL), the interior of the barrel is likely to have interactions formed between valine and isoleucine. Small residues, particularly glycine, permit adjacent residues to form a tightly packed barrel. Thus, the glycine to valine mutation is likely to disrupt not just the cylindrin formation, but also other toxic oligomers with structural similarities. This is consistent with our native MS analysis on the K11L and K11L<sup>G6V</sup> preparations. We observed that the valine mutation led to an overall reduction of oligomer abundance. Our study has identified a site in SOD1<sup>G93A</sup> that is important for the formation of toxic species.

## **Full length SOD1 oligomer model resembles a lipid binding protein**

Because our crystal structure of K11L is of only a segment, we modeled an SOD1 oligomer with K11L as the spine. This model of the full length protein appears to resemble the structure of a lipid-binding protein. To build the oligomer model around the K11L structure with no steric clash, it was necessary to slightly unravel the SOD1 monomer, exposing the N-terminal region of the protein (Figure 15). In the resulting model of the full-length SOD1 oligomer a hydrophobic groove formed by valines in the K11L structure is maintained in the inner surface. The K11V oligomer bears a striking resemblance to the lipid-binding moieties of the protein bactericidal/permeability-increasing protein (BPI). This suggests a potential mechanism for the toxicity of SOD1 oligomers.

## **Materials and Method**

### **Expression and purification of SOD1 variant constructs**

All SOD1 constructs are expressed in an E coli system. The mutations discussed in this paper are based on a SOD1 background strain that carries a C6A, C111S double mutation to simplify the purification. SOD1 gene is inserted into the pET22b vector from Novagen with NcoI and Sall restriction enzyme digestion site. This background construct was obtained from Professor Joan Valentine's lab at UCLA. The plasmid is transformed into BL21(DE3)Gold expression cell line. For expression, we inoculate 10 ml LB + Amp (50mg/mL) from frozen stock and grow overnight. 10ml of starting culture is added to 2-L flask of 1L LB + Amp and grow for 3 hours at 37 °C to  $OD_{600} = 0.6$ . IPTG was then added to 1mM and  $Zn^{2+}$  was added to 0.05 mM and grow for additional 3 hours. Bacteria pellet was collected by centrifuging 8,000 xg for 10 minute at 4 °C.

To release protein from the periplasm, we resuspend the cell pellet thoroughly in 30 ml of 30 mM Tris-HCl pH 8, 20% sucrose and stir slowly at room temperature for 10 min. Collect the cells by centrifugation at  $10,000 \times g$  at  $4^{\circ}\text{C}$  for 10 min. Remove all of the supernatant and discard. Then thoroughly resuspend the pellet in 30 ml of ice-cold 5 mM  $\text{MgSO}_4$  and stir the cell suspension slowly for 10 min on ice. During this step, the periplasmic proteins are released into the buffer. Centrifuge at  $4^{\circ}\text{C}$  for 10 min at  $10,000 \times g$  to pellet the shocked cells. Keep the supernatant that contains the proteins. We then remove contaminant proteins with  $\text{AmSO}_4$  precipitation. We take the supernatant from the last step of double osmotic shock in to a beaker that is put on ice with a stir bar and slowly add  $\text{AmSO}_4$  to 326g/L and let the sample stir for additional 30min on ice. Centrifuge at  $10,000 \times g$  for 15min and collect supernatant which contains AS-SOD1, then filter with 0.45um filter. We then remove  $\text{AmSO}_4$  with a Phenyl sepharose column. Buffer A contains 2 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.15 M NaCl, 0.05 M  $\text{KH}_2\text{PO}_4$ , pH 7.0, adjusted with 1 M KOH. Buffer B contains 0.15 M NaCl, 0.05 M  $\text{KH}_2\text{PO}_4$ , pH 7.0, adjusted with 1 M KOH. Load protein from  $\text{AmSO}_4$  precipitation step to 1x5ml hydrophobic column; Wash with buffer A to baseline and run linear gradient from 100%A to 100%B with 50ml volume. Peak contains SOD should come out before 50%. There is also a small peak of contaminant protein elute near 100% buffer.

## **SEC-HPLC**

This sample is concentrated and loaded onto a prep-sized SEC-HPLC column. The peak containing AS-SOD elutes at around 60 minute. Since SOD1 has low A280 Absorbance and a contaminant protein elutes immediately before SOD1, to obtain pure protein, either low small amount of sample, or collect the peak, concentrate and reload onto the SEC-HPLC column.

Proteins were filtered through a 0.22  $\mu\text{m}$  Centrex MF filter (Whatman, Florham Park, NJ). Filtered preparations were injected on a 21.5 mm x 60 cm Tosohaas G3000SW column (Tosoh Bioscience, King of Prussia, PA) equilibrated in SEC buffer (25mM sodium phosphate, 100mM sodium sulfate pH 6.5) at a flow rate of 3 mL/min. Absorbance at 220nm and 280nm were recorded using a Waters 2487 dual  $\lambda$  absorbance detector (Waters, Milford, MA). Protein standards were monitored by absorbance at 280nm.

### **SOD1 demetallation**

The SOD1 discussed in this study are demetallated proteins. The sample purified with SEC-HPLC are dialyzed in four buffers at 4 °C. This is a modified version of a protocol developed by Professor John Hart's group. Protein preparations is put in a 10K dialysis cassette and exchanged against buffer 1 (100mM Na Acetate pH 3.8, 10mM EDTA pH 8.0) twice, 12 hours each. Then Exchange against buffer 2 (100mM Na Acetate pH 3.8, 100mM NaCl) for 4 hours, then in Buffer 3 (100mM Na Acetate pH 5.5, 100mM NaCl) for 4 hours, and lastly in Buffer 4 (20mM K phosphate pH7.0) twice for 6-10 hours each time. To avoid metal contamination, all stock solutions (except for EDTA) have been treated with Chelex 100 resin (Bio-Rad) to remove trace metals and all beakers have been rinsed with metal-free water.

### **Fibrillation assay**

Stock protein after metal removal have been concentrated with a 10K concentrator to 800  $\mu\text{M}$  (monomer) stock concentration and aliquot to 50  $\mu\text{l}$  and stored at -80 °C. Fibrillation assay was carried out in 1.5ml eppendorf tubes under acidic conditions with 80  $\mu\text{M}$  SOD1, 50 mM Sodium Acetate, 1mM EDTA and 10 mM TCEP, pH 3.5. Preparations were agitated at 900 rpm



at 37 °C for 24 hours and amyloid aggregates can be observed as white precipitant. These sample were then transferred to a 10K Slide-A-Lyzer MINI Dialysis Devices and exchanged against 2L of buffer with 100mM Tris and 50mM NaCl, pH 7.5 for 8 hours at 4 °C. Preparations were than collected for toxicity assay.

### **PC12 and HeLa Cell Culture and Viability Assay**

Cell viability was investigated using a CellTiter 96 aqueous non-radioactive cell proliferation assay kit (MTT) (Promega cat. #G4100). HeLa and PC12 cells were used to assess the toxic effect of full length SOD1 and 11-residue peptide segment. HeLa cells were cultured in DMEM medium with 10% fetal bovine serum and PC-12 cells were cultured in ATCC-formulated RPMI 1640 medium (ATCC; cat.# 30-2001) with 10% heat-inactivated horse serum and 5% fetal bovine serum. Cells were maintained at 37 °C in 5% CO<sub>2</sub>. For all toxicity experiments, 96-well plates (Costar cat. # 3596) were used. HeLa and PC-12 cells were plated at 10,000 cells per well. Cells were cultured for 20h at 37 °C in 5% CO<sub>2</sub> prior to addition of peptide preparations. 10 µl of sample was added to each well containing 90 µL medium, and allowed to incubate for 24h prior to the addition of 15 µl Dye solution (Promega. cat. #G4102) into each well, followed by incubation for 4h at 37°C in 5% CO<sub>2</sub>. After incubation, 100 µl solubilization Solution/Stop Mix (Promega cat. #G4101) was added to each well. After 12h incubation at room temperature, the absorbance was measured at 570nm. Background absorbance was recorded at 700nm. Each of the experiments was repeated 3 times with 4 replicates per sample per concentration. The concentrations of the amyloid preparations were based on the corresponding monomeric protein/peptides. The results were normalized using the buffer-treated cell as 100% viability and 0.2% SDS-treated as 0% viability.

## **Differentiation of mESCs to motor neurons**

Hb9::eGFP mESCs were maintained and differentiated into motor neurons as previously described (Wichterle and Peljto 2008). Briefly, mESCs were first plated on gelatin to remove MEFs prior to differentiation, and then plated in 60mm bacterial petri dishes in core motor neuron medium (DMEM/F12, Neurobasal, 10% Knockout Serum Replacement, Pen-Strep, Glutamax) to induce embryoid body (EB) formation. Two days later, N2 supplement (1x), Retinoic Acid (1uM; Sigma), and SAG (1uM; Calbiochem) were added to the EBs. Media was changed every two days. All reagents are from Invitrogen, unless otherwise noted.

## **Plating Hb9::GFP mESC-derived motor neurons for toxicity assay**

After five days post Retinoic Acid and SAG addition, Hb9::GFP EBs were dissociated using ice-cold 0.25% Trypsin-EDTA for 6 minutes at room temperature, followed by trituration in L-15 medium (Hyclone). Dissociated motor neurons were plated in core motor neuron medium with GDNF, BDNF, CNTF (all 10ng/ml; Peprotech) on 96-well plates that were previously coated with Poly-O-Ornithine (0.01%; Sigma) and Laminin (5ug/ml; BD Biosciences). Approximately 80,000 cells were plated per well. The next day, recombinant proteins were added to the cells. Fluorescent images were taken after 24 hours of incubation.

## **K11L structure determination**

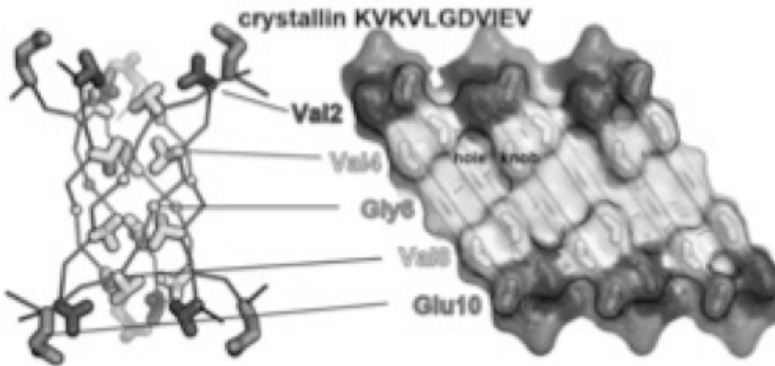
Crystals of K11L were grown in hanging drop VDX plates (Hampton Research, Aliso, Viejo, CA). Protein is dissolved as 50mg/ml in 50mM Tris-base. Reservoir solution contains 0.2M Na Citrate pH 5, 13% PEG 6000. Thin needle like crystals appear after 2-3 days at ambient temperature. For heavy metal soaking, weigh 0.008g KI and dissolve KI in 100 uL of

reservoir solution as 0.5M KI solution. Mix 6.5 microliters KI solution with 3.5 microliters of mother liquor solution with 10% glycerol stock for cryoprotection. Soak crystal 30 seconds in the cryo-KI solution and flash freeze in liquid nitrogen.

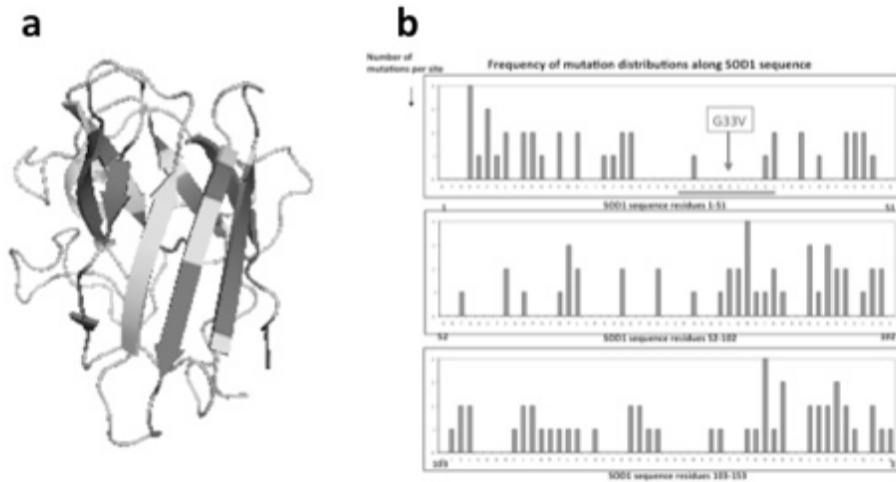
All data were collected at Advanced Photon Source (Chicago, IL) beam lines 24-ID-C and 24-ID-E. Single crystals were mounted with CrystalCap HT Cryoloops (Hampton Research, Aliso Viejo, CA). All data were processed using DENZO and SCALEPACK or XDS. K11L phase was obtained by heavy atom soaking and structure was solved by SIRAS, followed by model building using COOT. All model refinement was done using REFMAC

### **Native Nanoelectrospray Mass Spectrometry**

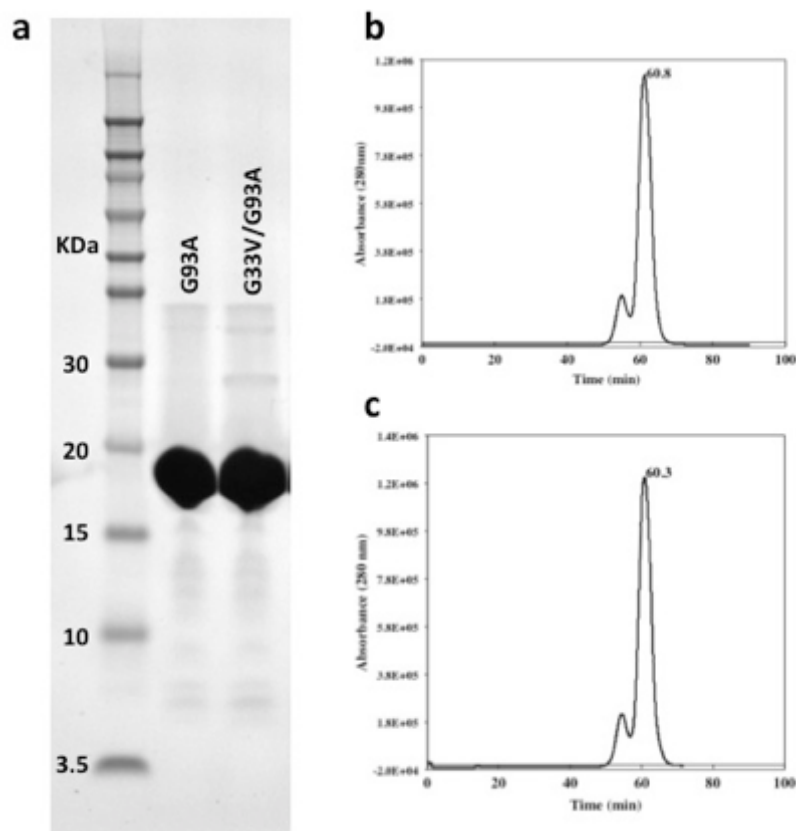
K11L and K11LG6V were analyzed by direct nanospray injection (for review see (62)). Fractions were individually loaded into a 2- $\mu\text{m}$  internal diameter externally coated nanospray emitter (ES380, Thermo) and desorbed by adjusting the spray voltage to maintain an ion current between 0.1 and 0.2  $\mu\text{A}$ . A hybrid linear ion-trap/FTICR mass spectrometer was used for the analysis (7T, LTQ FT Ultra, Thermo Scientific, Bremen, Germany). Individual charge states of multiply protonated complex ions were selected for isolation and collisional activation in the linear ion trap followed by detection of the resulting product ions in the FTICR cell. Xtract software (Thermo Scientific, Bremen, Germany) was used to compute monoisotopic mass from the measured isotopomer profile.



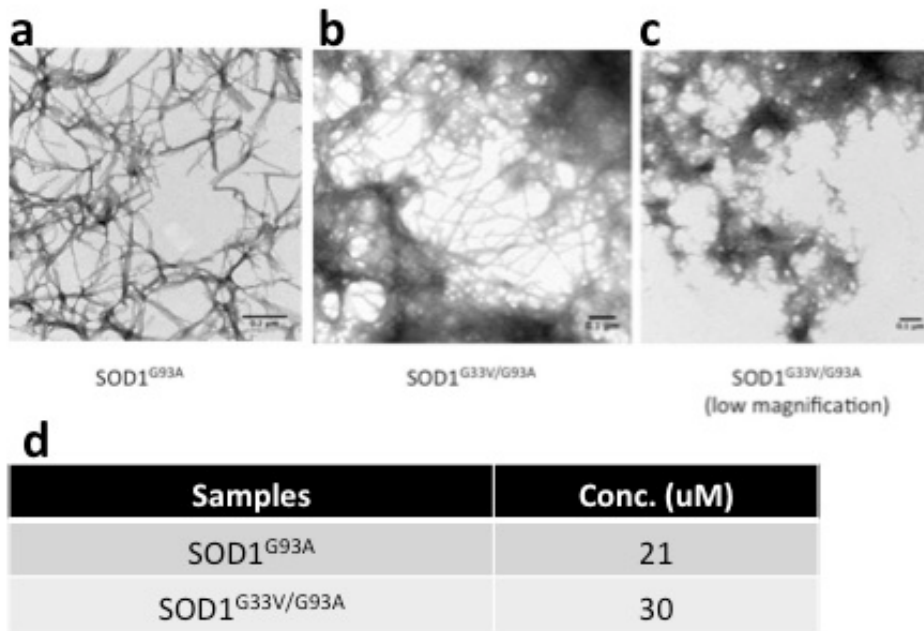
**Figure 1. Crucial role of glycine in cylindrin formation.** Structural model of the cylindrin unrolled into a flat out-of-register  $\beta$ -sheet, revealing the interior of the cylindrin and the “knobs-and-holes” motif. Val 6 and Val 8 form the “knob” as labeled and Gly 6 creates the “holes”. The small size of the glycine is crucial in enabling the formation of the cylindrin structure.



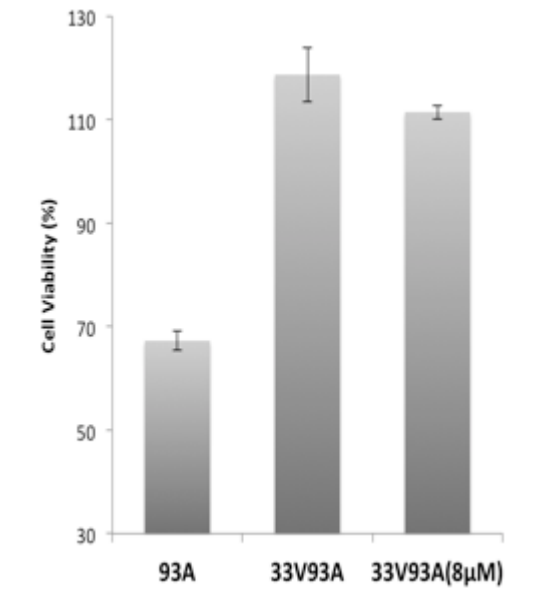
**Figure 2. Mapping of disease-related mutations on SOD1.** **a**, Amino acids positions that have found to be mutated in fALS patients are colored dark grey ([http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene\\_id=SOD1](http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SOD1)). Except for  $\beta$ 3, all other  $\beta$ -strands have multiple mutations and two of the strands are mutated at all positions. **b**, Plotting the number of mutations per amino acid position shows that the mutations are distributed through the SOD1 sequence with no apparent pattern. We identified that the seven residues centered around Gly 33, making  $\beta$ 3 is the longest stretch of sequence that is present in all the disease related mutants.



**Figure 3. Recombinant SOD1 mutant production.** **a**, SDS-PAGE of recombinant SOD1<sup>G93A</sup> and SOD<sup>G33V/G93A</sup> eluting as monomers with expected molecular weight. **b**, SEC-HPLC chromatogram at 220 nm, showing SOD1<sup>G93A</sup> eluting as a dimer at 60.8 minute. **c**, SEC-HPLC for SOD1<sup>G33V/G93A</sup> showed similar chromatogram. A peak containing a contaminant protein eluted immediately before the SOD1 construct. SOD1 fractions, for both constructs, were only collected after the A220 nm absorption reaches its maximum values. The resulting preparations contain fairly pure SOD1 proteins (**a**).

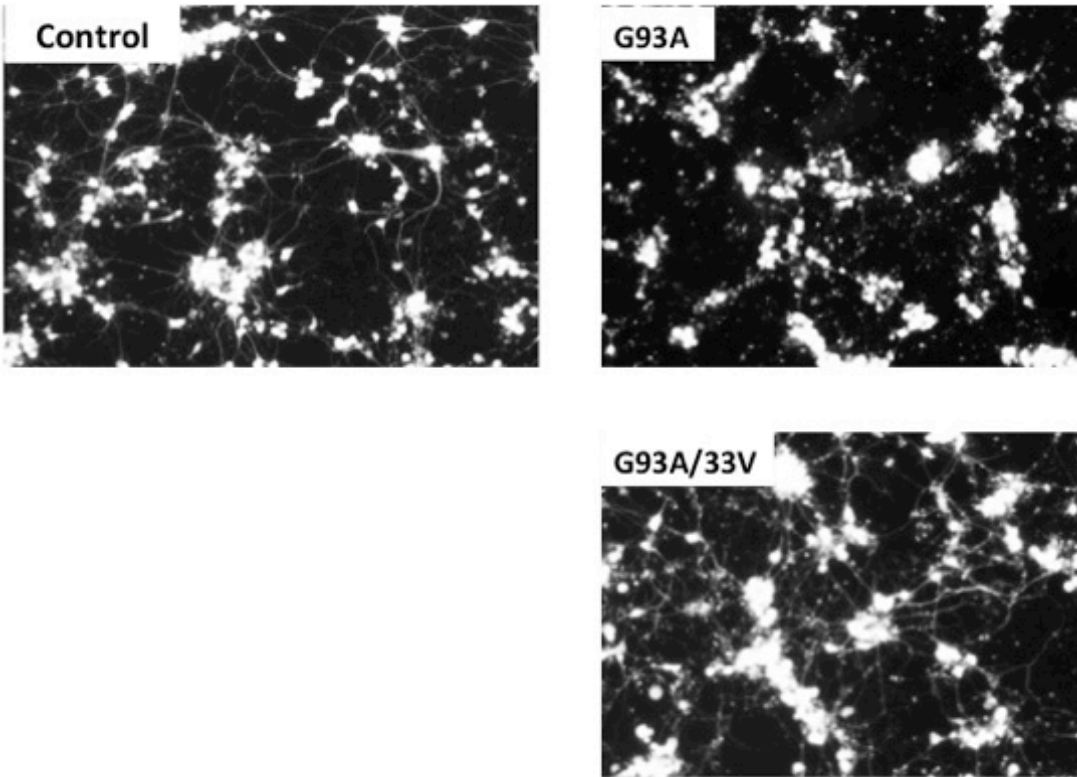


**Figure 4. Both SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> form abundant fibrils.** EM images of SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> fibrils. 40  $\mu$ M of monomeric proteins were dissolved in partially denatured conditions and incubated at 37 °C for 24 hours with shaking and fibril preparations were further dialyzed in a physiologically compatible buffer. Abundant fibrils can be observed under negative staining electron microscopy for the post-dialysis preparations SOD1<sup>G93A</sup> (a) and SOD1<sup>G33V/G93A</sup> (b, c). d, Fibril preparations were pelleted by centrifuging at 13,000 xg for 30 min and proteins concentrations in the pellets was determined. The protein concentration present in the pellet is comparable between the two preparations. G33V mutation does not disrupt SOD1<sup>G93A</sup> fibril formation and that both constructs are capable of forming abundant fibrils.

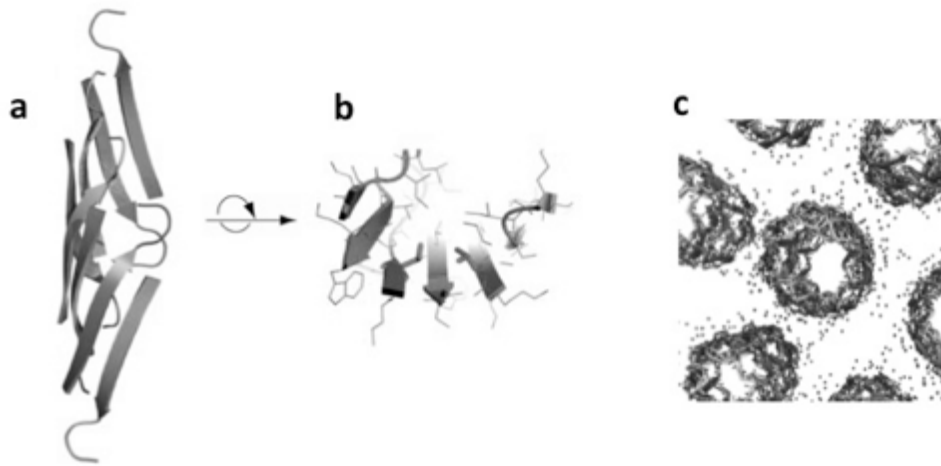


**Figure 5. G33V abolishes the toxicity of SOD1<sup>G93A</sup> in PC12 neuronal cells.** Post-dialysis amyloid preparations containing abundant fibrils were added to PC12 cells. Cells were plated as 10,000 cells per well. SOD1<sup>G93A</sup> was added as 4 µM final concentration. SOD1<sup>G33V/G93A</sup> was added as both 4 µM and 8 µM final concentrations. Three replicates were done with each condition. MTT reagents were added after 24 hours of incubation at 37 °C. SDS was used for 0% cell viability and toxicity data was normalized against a negative control with dialysis buffer. SOD1<sup>G93A</sup> was toxic at 4 µM and reduced cell viability to ~65% while G33V mutation abolishes the toxicity at both 4 µM and 8 µM. Standard deviations are showing here as error bars.

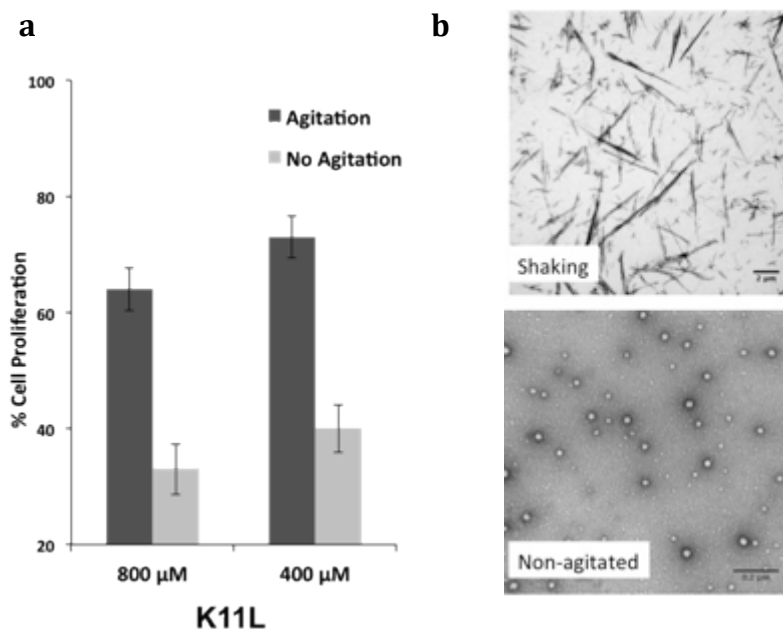




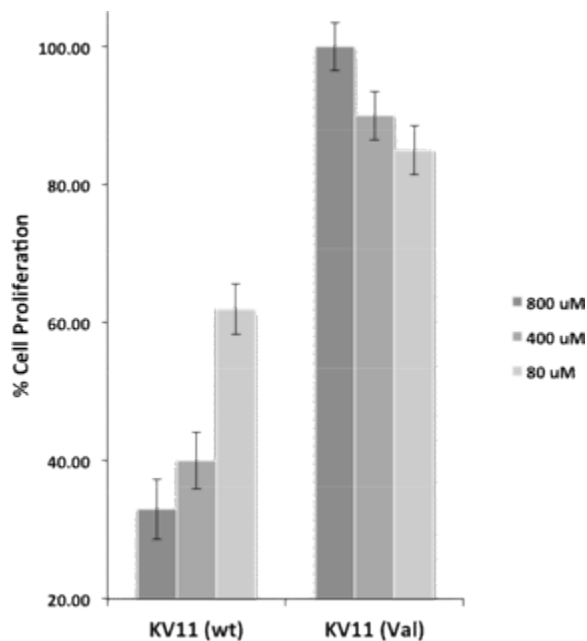
**Figure 6. G33V abolishes the toxicity of SOD1<sup>G93A</sup> in motor neurons.** Post-dialysis SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> amyloid preparations were added to ES-derived motor neurons six day post differentiation. Motor neurons were plated as 80,000 cells per well. Amyloid preparations were added as 12  $\mu$ M final concentration. Dialysis buffer was added to control preparations. Fluorescent images were taken 24 hours after cells were incubated with the protein preparations. Buffer treated cells showed healthy neurons with abundant axon networks. In SOD1<sup>G93A</sup> treated cells, a clear toxic effect can be observed. The morphology of the neurons was significantly different with shrinking cell bodies and majority of the axons missing. In contrary, SOD1<sup>G33V/G93A</sup> treated cells showed normal morphology that is similar to the control preparations.



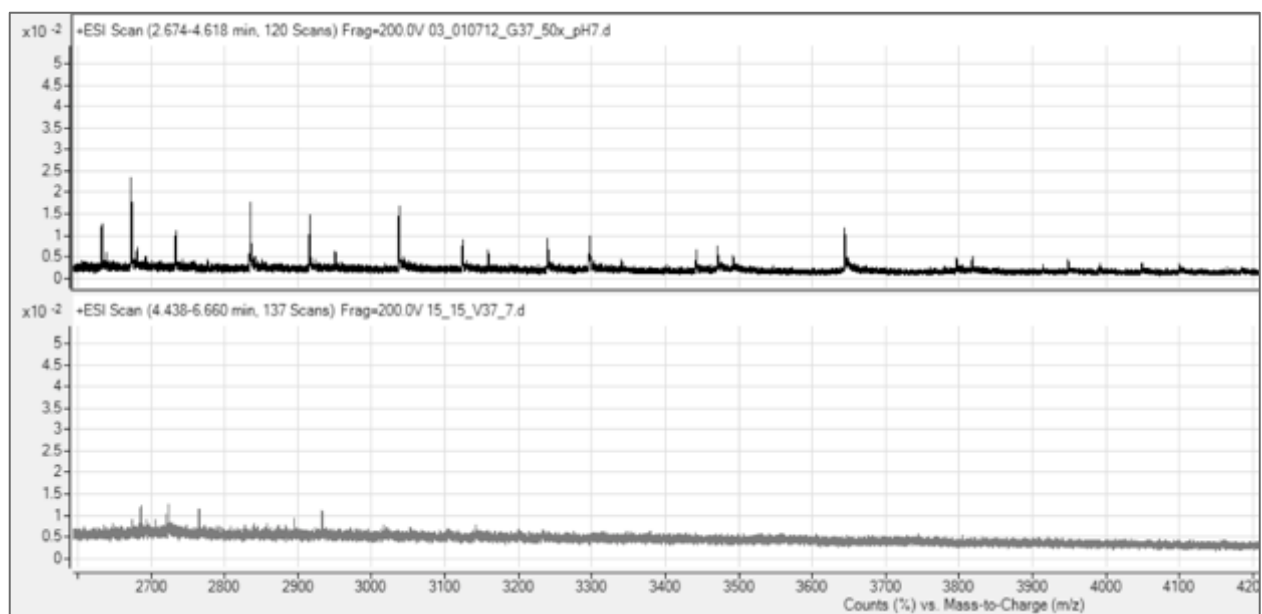
**Figure 7. Crystal structure of the K11L open barrel.** Crystal structure of K11L reveals an asymmetric unit containing eight antiparallel  $\beta$ -strands. These strands are arranged in an out-of-register fashion forming an open-barrel (a). Top down view (b) shows the interior of the structure is composed of valine residues and the exterior of the structure is composed of tryptophan and lysine residues. Strong van der waal interactions are formed with valine in the interior of the barrel. Val 2 and Val 4 pack tightly against the glycine in the adjacent strand. The interior of the open barrel hydrophobic, and excludes water (c).



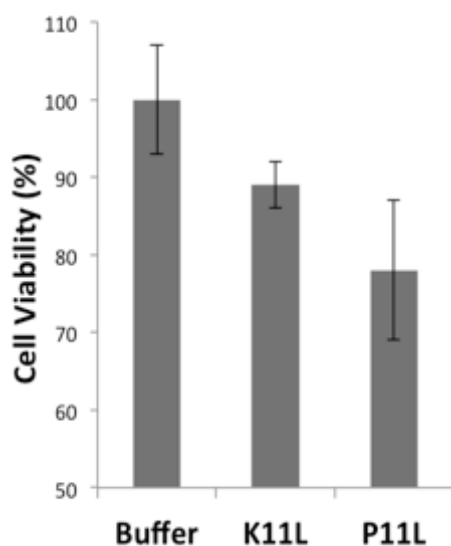
**Figure 8. Prefibrillar K11L specie is toxic to mammalian cells.** Synthetic K11L peptide was dissolved in 50 mM Tris-base as 8 mM and incubated at 37 °C either with or without shaking for over night and then added to HeLa cells. MTT reagents were added after 24 hours of incubation at 37 °C. SDS was used for 0% cell viability and toxicity data was normalized against a negative control. K11L non-agitated preparations were significantly more toxic than preparations formed under agitated conditions. EM images showed abundant fibrils are present in the agitated sample and spherical species is present in the non-agitated, toxic sample.



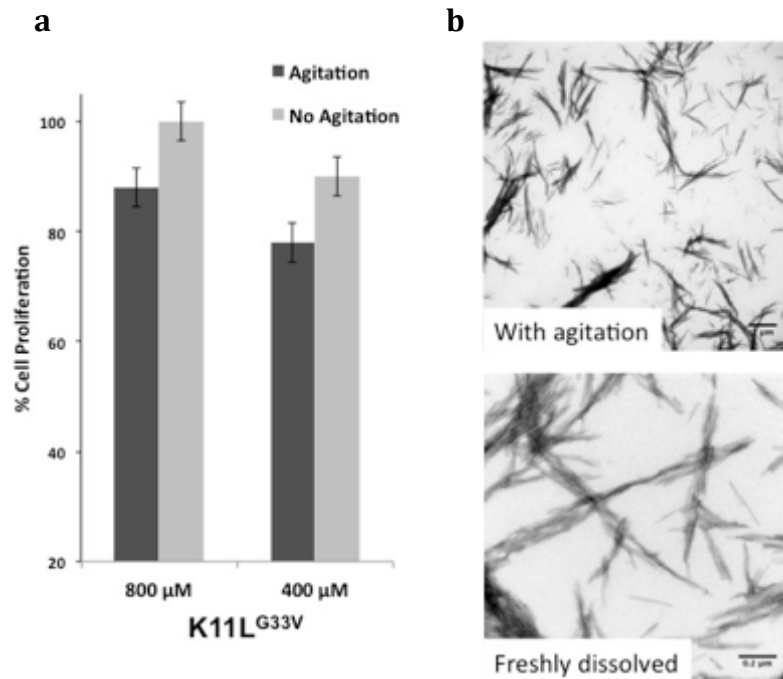
**Figure 9. Concentration dependent toxicity.** Synthetic K11L and K11LG6V peptide were dissolved in 50 mM Tris-base as 8 mM and incubated at 37 °C without shaking for over night. Preparations were diluted and added to HeLa cells at indicated concentrations. MTT reagents were added after 24 hours of incubation at 37 °C. SDS was used for 0% cell viability and toxicity data was normalized against a negative control. K11L is toxic to cell in a concentration dependent manner. K11LG6V rescues the toxicity of K11L also in a concentration dependent manner, but followed an opposite trend. At 800  $\mu$ M, K11LG6V completely abolished the toxicity. However, at a lower concentration, 80  $\mu$ M, the peptide became mildly toxic. Since K11LG6V form fibrils immediately upon dissolving and that the fibrils are not the toxic species, we speculate that the low concentration destabilized the K11LG6V fibril and favored the release of toxic oligomers.



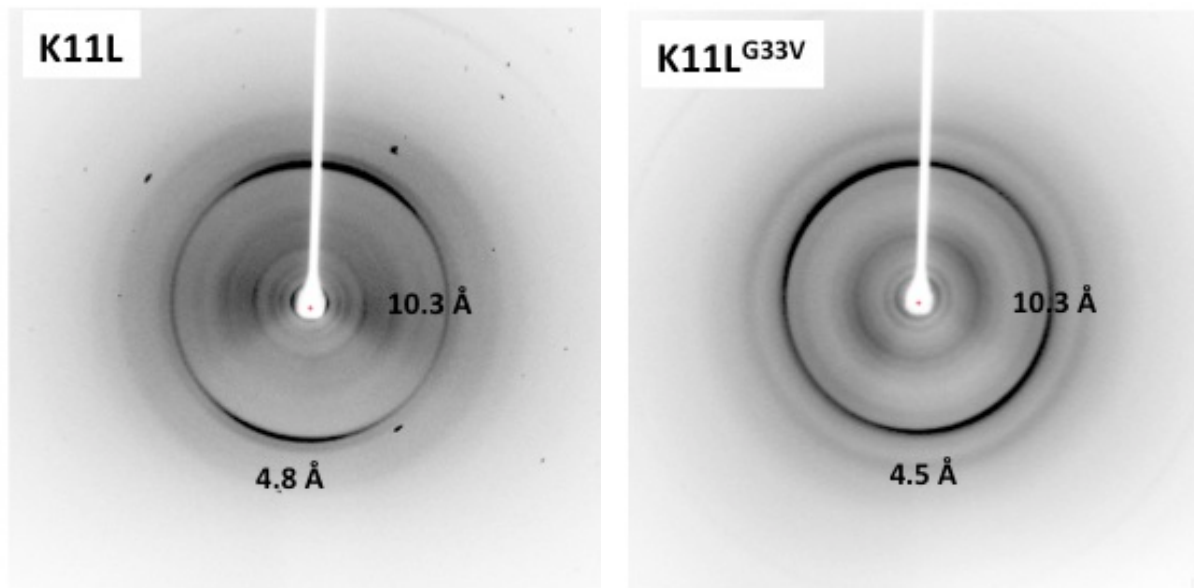
**Figure 10. Native MS spectrum of K11L (top) and K11LG6V (Bottom).** Synthetic peptides were dissolved in Tris-base and incubated at 37 °C for over night under non-agitated condition. Oligomers of various sizes were present in K11L preparations. Small amount of oligomers were still present in K11LG6V, but are present in a much lower amount.



**Figure 11. Segments P11L and K11L are both toxic to mammalian cells.** Chemically synthesized peptides were incubated at 37 °C in 50mM Tris-base over night. HeLa cells were plated as 10,000 cells per well. Peptide preparations were added as 250  $\mu$ M final concentration. Tris-base was added as control. Three replicates were done with each condition. MTT reagents were added after 24 hours of incubation at 37 °C. SDS was used for 0% cell viability and toxicity data was normalized against a negative control. The native segment P11L (PVKVVWGSIKGL) was toxic to HeLa cells. K11L (KVKVVWGSIKGL), constructed to aid crystallization also showed toxicity but with a milder effect.

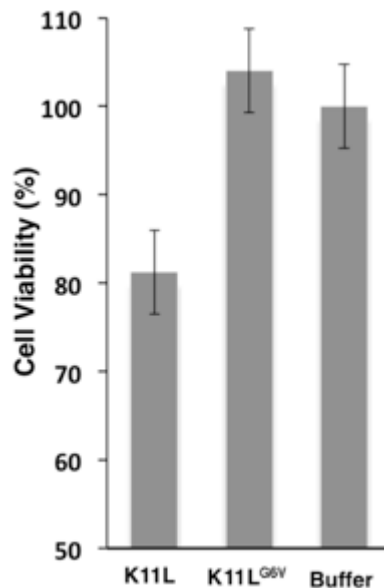


**Figure 12. Glycine to valine mutation abolishes the toxicity of K11L.** Synthetic K11LG6V (KVKVWVSIKGL) peptide was chemically synthesized and dissolved in 50 mM Tris-base as 8 mM and incubated at 37 °C either with or without shaking for over night and then added to HeLa cells. MTT reagents were added after 24 hours of incubation at 37 °C. SDS was used for 0% cell viability and toxicity data was normalized against a negative control. K11LG6V formed a gel like consistency immediately after dissolving. EM image showed abundant fibrils were present both in the freshly dissolved sample and the agitated sample. Regardless of the preparation methods, these preparations are either nontoxic or are only mildly toxic, all of which are significantly less toxic than the K11L preparations.

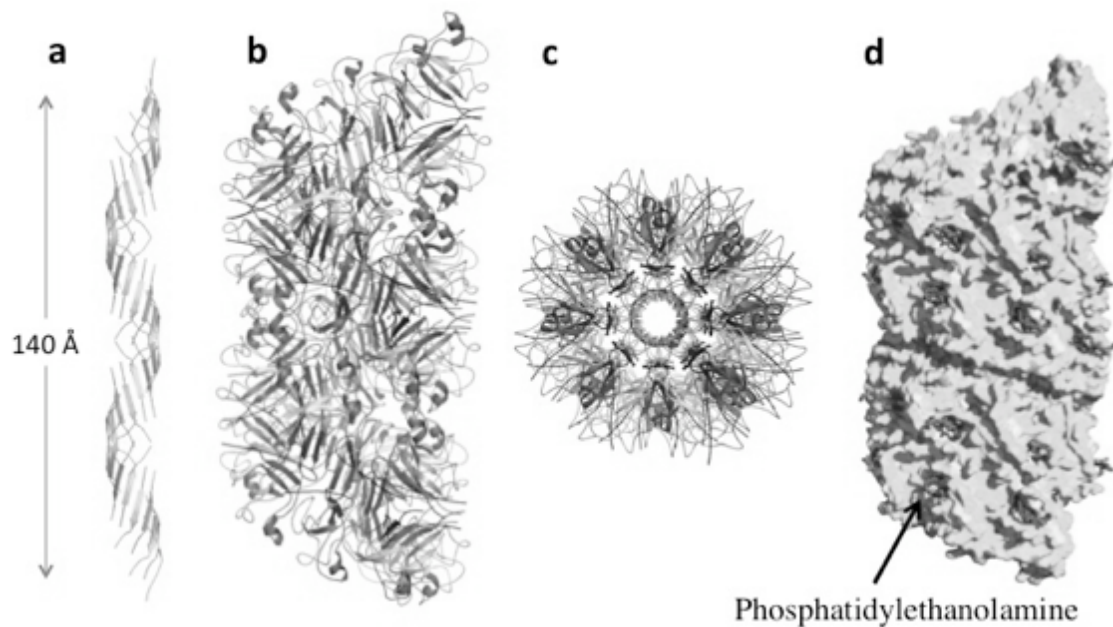


**Figure 13. Fibril diffraction.** K11L and K11LG6V agitated preparations were pelleted and aligned at room temperature. X-ray fibril diffraction for both preparations showed the cross- $\beta$  diffraction pattern with strong diffractions at  $\sim 4.8$  Å and  $\sim 10$  Å along the meridional and the equatorial axis respectively. This data suggest that both preparations contain typical amyloid fibrils.

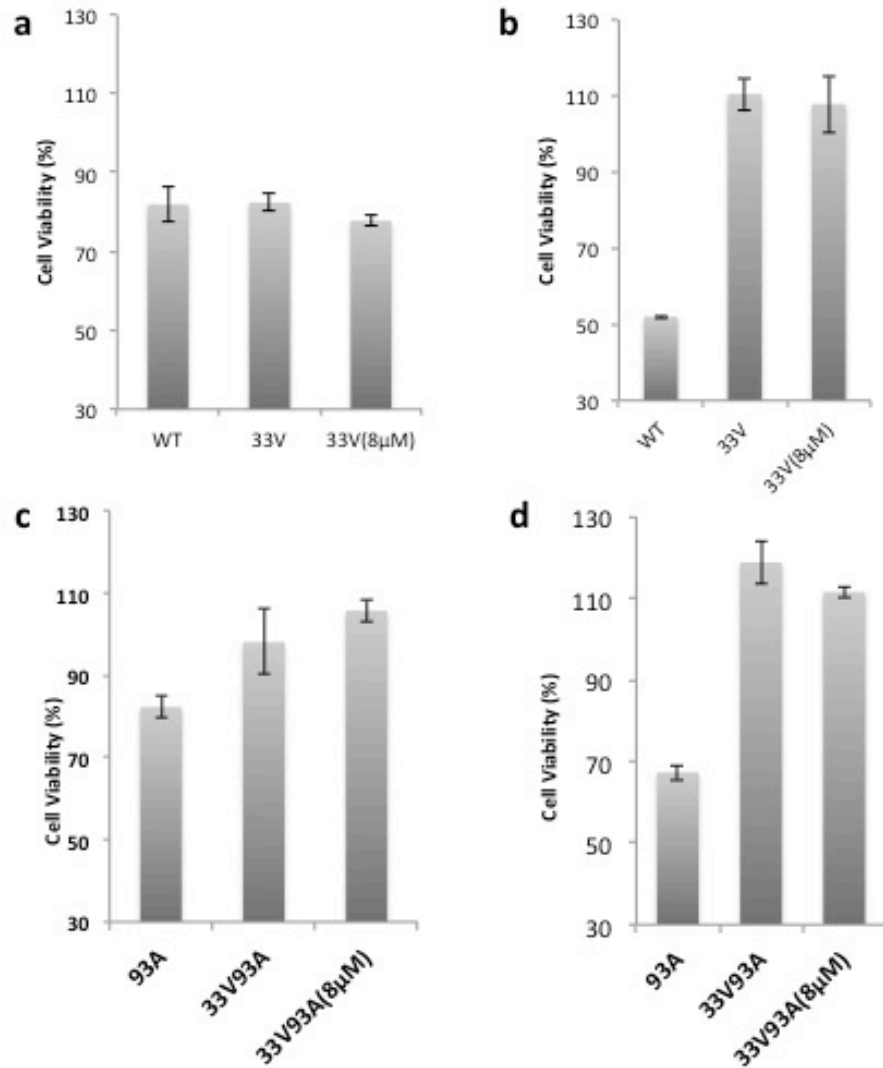




**Figure 14. G33V abolishes the toxicity of K11L in motor neurons.** Post-dialysis K11L and K11LG6V amyloid preparations were added to ES-derived motor neurons six day post differentiation. Motor neurons were plated as 80,000 cells per well. Amyloid preparations were added as 250  $\mu$ M final concentration. Dialysis buffer was added to control preparations. MTT reagents were added after 24 hours of incubation at 37 °C. SDS was used for 0% cell viability and toxicity data was normalized against a negative control. The segment in isolation was less toxic than full-length protein due to the absence of the rest of the amino acids that could facilitate the development of toxicity. Additionally, this segment appeared to be more toxic to HeLa cells than to motor neurons. Nonetheless, it was clear that the glycine to valine mutation rescues toxicity of the K11L segment.

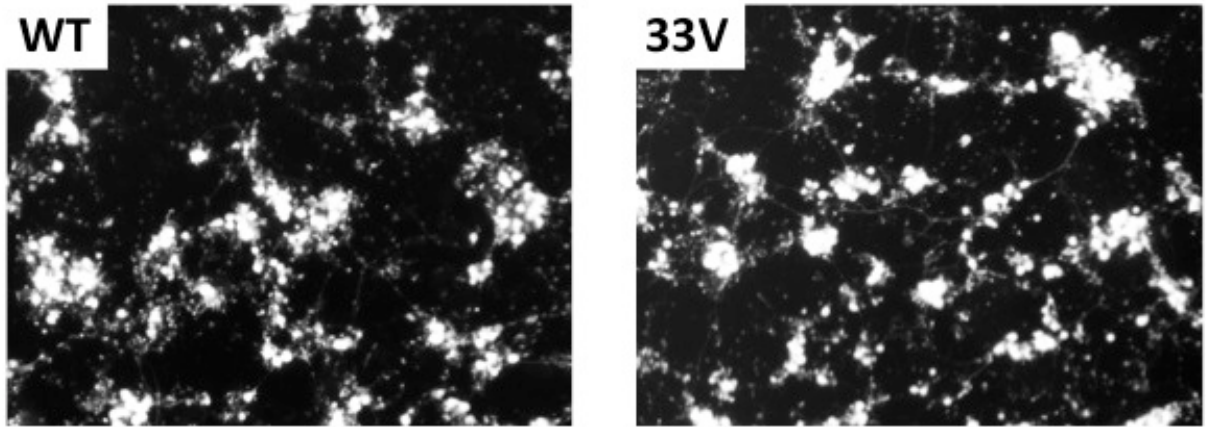


**Figure 15. Model for SOD1 aggregate using the K11L structure as a spine.** Using the structure of K11L (a), we were able to model how full length SOD1 might assemble into such structure. Monomeric protein used here contained a partially denatured structure. Both side view (b) and top view (d) of the modeled structure is showing here. A hydrophobic groove that was present in the K11L structure was maintained after the full-length protein was modeled. The structure of the groove resembles that of a lipid binding protein, the human bactericidal/permeability-increasing protein (BPI). As such, we modeled the phosphatidylethanolamine molecule into the predicted lipid-binding sites as indicated (d). Co-crystal structures are needed to validate this model. Unfortunately, we have not been able to obtain diffraction-quality crystals from detergent screens.

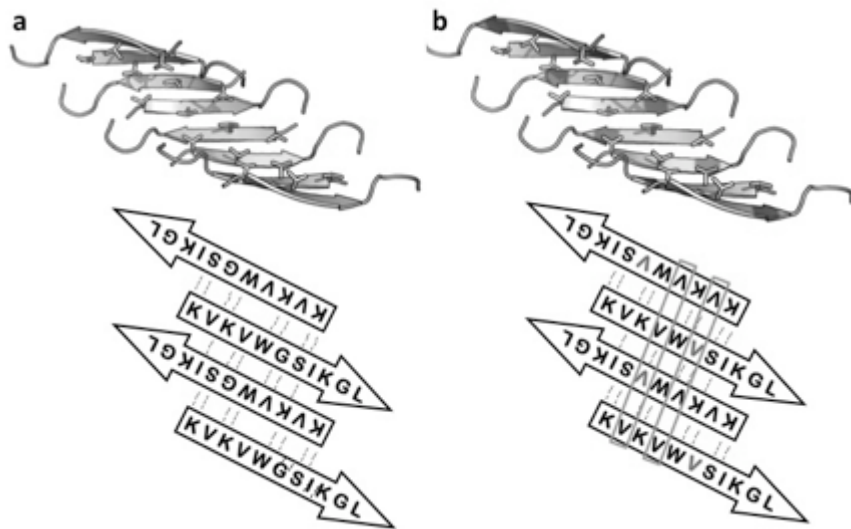


**Figure 16 Toxicities of monomer and amyloid preparations.** For freshly dissolved preparations (**a, c**), proteins SOD1, SOD1<sup>G33V</sup>, SOD1<sup>G93A</sup> and SOD1<sup>G33V/G93A</sup> were dissolved in dialysis buffer as describe in previous sections and added to HeLa cells as either 4 μM or 8 μM as indicated. For amyloid preparations (**b, d**), recombinant proteins were dissolved in a partially denaturing condition and incubated at 37 °C with continuous shaking for 24 hours. The amyloid preparations were then dialyzed and added to HeLa cells as either 4 μM or 8 μM as indicated. MTT reagents were added after 24 hours of incubation at 37 °C. SDS was used for 0% cell

viability and toxicity data was normalized against a negative control. Even in freshly dissolved preparations (**a, c**), both WT and G93A appeared to already be toxic to cells. However, both preparations become more toxic after 24 hours of incubation with shaking (**b, d**). G33V was able to abolish the toxicities of both WT and G93A in the amyloid preparations (**b, d**). Interestingly, G33V mutation had no effect in the toxicity of freshly dissolved WT protein, but had a mild rescue effect on G93A construct. This seemed to suggest that the WT and G93A constructs might have different effects on HeLa cells, both of which led to the observed toxicity detected by MTT assays.



**Figure 17. G33V does not abolish the toxicity of WT SOD1 in motor neurons.** Post-dialysis SOD1 and SOD1<sup>G33V</sup> amyloid preparations were added to ES-derived motor neurons six day post differentiation. Motor neurons were plated as 80,000 cells per well. Amyloid preparations were added as 12  $\mu$ M final concentration. Dialysis buffer was added to control preparations. Fluorescent images were taken 24 hours after cells were incubated with the protein preparations. Control preparations show healthy neurons with well-connected axons. Both SOD1 and SOD1<sup>G33V</sup> treated cells showed ill-looking cell morphology majority of the axons missing. No appearance differences could be observed between the two preparations.



**Figure 18. Valine mutation creates steric clash.** In the K11L structure (a), The interior of the open barrel was made with interactions between Val 2 and Val 4. Gly 6 was positioned on a strand that was in between the Val2 and Val 4, enabling the valine interactions. In the structural model of the engineered mutant (b), where glycine is replaced by valine, rolls of valine would line up in the interior creating major steric hindrance. Formation of such barrel structure would not be possible with K11L<sup>G6V</sup> peptide.

**Table 1. K11L data collection and refinement statistics**

<b>Data collection</b>	
Beam line	APS 24-ID-E
Resolution Å	2.1
Total unique reflections	11489
Total reflections observed	43500
Space group	P212121
Rsym	9.5% (41.8%)
I/σ	11.14 (3.6)
Completeness	95.8 %
Wavelength Å	0.9791
Unit Cell dimensions a b c (Å) alpha beta gamma (°)	33.09 44.39 71.43 90.000 90.0 90.0
<b>Refinement</b>	
Resolution Å	2.1- 19.3
Reflections for refinement (after merging Friedel Pairs)	6249
Rfree/Rwork (%)	.25/.22
Molecules in the Asymmetric unit	8
Solvent content (%)	54.6
Matthews coefficient	2.71
Total water molecules	52
Total iodide Atoms	2
Glycerol Molecule	1
Rmsd bond length (Å)	0.010
Rmsd angles (°)	0.88
Ramachandran plot:	
Allowed	100
Generous	0
Disallowed	0

## Reference

Broersen, K., Rousseau, F., and Schymkowitz, J. (2010). The culprit behind amyloid beta peptide related neurotoxicity in Alzheimer's disease: oligomer size or conformation? *Alzheimer's research & therapy* 2, 12.

Caughey, B., and Lansbury, P.T. (2003). Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annual review of neuroscience* 26, 267-298.

Chattopadhyay, M., and Valentine, J.S. (2009). Aggregation of copper-zinc superoxide dismutase in familial and sporadic ALS. *Antioxidants & redox signaling* 11, 1603-1614.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. *Annual review of biochemistry* 75, 333-366.

Chiti, F., and Dobson, C.M. (2009). Amyloid formation by globular proteins under native conditions. *Nature chemical biology* 5, 15-22.

Eisenberg, D., and Jucker, M. (2012). The amyloid state of proteins in human diseases. *Cell* 148, 1188-1203.

Fandrich, M. (2012). Oligomeric intermediates in amyloid formation: structure determination and mechanisms of toxicity. *Journal of molecular biology* 421, 427-440.

Finder, V.H., Vodopivec, I., Nitsch, R.M., and Glockshuber, R. (2010). The recombinant amyloid-beta peptide A $\beta$ 1-42 aggregates faster and is more neurotoxic than synthetic A $\beta$ 1-42. *Journal of molecular biology* 396, 9-18.

Fonte, V., Dostal, V., Roberts, C.M., Gonzales, P., Lacor, P., Magrane, J., Dingwell, N., Fan, E.Y., Silverman, M.A., Stein, G.H., *et al.* (2011). A glycine zipper motif mediates the formation of toxic beta-amyloid oligomers in vitro and in vivo. *Molecular neurodegeneration* 6, 61.

Glabe, C. (2006). *Biomedicine*. Avoiding collateral damage in Alzheimer's disease treatment. *Science* 314, 602-603.

Goldschmidt, L., Teng, P.K., Riek, R., and Eisenberg, D. (2010). Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proceedings of the National Academy of Sciences of the United States of America* 107, 3487-3492.

Grad, L.I., Guest, W.C., Yanai, A., Pokrishevsky, E., O'Neill, M.A., Gibbs, E., Semenchenko, V., Yousefi, M., Wishart, D.S., Plotkin, S.S., *et al.* (2011). Intermolecular transmission of superoxide dismutase 1 misfolding in living cells. *Proceedings of the National Academy of Sciences of the United States of America* 108, 16398-16403.



Haass, C., and Selkoe, D.J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nature reviews Molecular cell biology* 8, 101-112.

Harmeier, A., Wozny, C., Rost, B.R., Munter, L.M., Hua, H., Georgiev, O., Beyermann, M., Hildebrand, P.W., Weise, C., Schaffner, W., *et al.* (2009). Role of amyloid-beta glycine 33 in oligomerization, toxicity, and neuronal plasticity. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29, 7582-7590.

Jaikaran, E.T., and Clark, A. (2001). Islet amyloid and type 2 diabetes: from molecular misfolding to islet pathophysiology. *Biochimica et biophysica acta* 1537, 179-203.

Johnston, J.A., Dalton, M.J., Gurney, M.E., and Kopito, R.R. (2000). Formation of high molecular weight complexes of mutant Cu, Zn-superoxide dismutase in a mouse model for familial amyotrophic lateral sclerosis. *Proceedings of the National Academy of Sciences of the United States of America* 97, 12571-12576.

Kanski, J., Varadarajan, S., Aksenova, M., and Butterfield, D.A. (2002). Role of glycine-33 and methionine-35 in Alzheimer's amyloid beta-peptide 1-42-associated oxidative stress and neurotoxicity. *Biochimica et biophysica acta* 1586, 190-198.

Karch, C.M., Prudencio, M., Winkler, D.D., Hart, P.J., and Borchelt, D.R. (2009). Role of mutant SOD1 disulfide oxidation and aggregation in the pathogenesis of familial ALS. *Proceedings of the National Academy of Sciences of the United States of America* 106, 7774-7779.

Kayed, R., Head, E., Thompson, J.L., McIntire, T.M., Milton, S.C., Cotman, C.W., and Glabe, C.G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* 300, 486-489.

Kfoury, N., Holmes, B.B., Jiang, H., Holtzman, D.M., and Diamond, M.I. (2012). Trans-cellular propagation of Tau aggregation by fibrillar species. *The Journal of biological chemistry* 287, 19440-19451.

Laganowsky, A., Liu, C., Sawaya, M.R., Whitelegge, J.P., Park, J., Zhao, M., Pensalfini, A., Soriaga, A.B., Landau, M., Teng, P.K., *et al.* (2012). Atomic view of a toxic amyloid small oligomer. *Science* 335, 1228-1231.

Lashuel, H.A., Hartley, D., Petre, B.M., Walz, T., and Lansbury, P.T., Jr. (2002). Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature* 418, 291.

Last, N.B., Rhoades, E., and Miranker, A.D. (2011). Islet amyloid polypeptide demonstrates a persistent capacity to disrupt membrane integrity. *Proceedings of the National Academy of Sciences of the United States of America* 108, 9460-9465.

Lesne, S., Koh, M.T., Kotilinek, L., Kaye, R., Glabe, C.G., Yang, A., Gallagher, M., and Ashe, K.H. (2006). A specific amyloid-beta protein assembly in the brain impairs memory. *Nature* *440*, 352-357.

Liu, C., Zhao, M., Jiang, L., Cheng, P.N., Park, J., Sawaya, M.R., Pensalfini, A., Gou, D., Berk, A.J., Glabe, C.G., *et al.* (2012). Out-of-register beta-sheets suggest a pathway to toxic amyloid aggregates. *Proceedings of the National Academy of Sciences of the United States of America* *109*, 20913-20918.

Mead, R.J., Bennett, E.J., Kennerley, A.J., Sharp, P., Sunyach, C., Kasher, P., Berwick, J., Pettmann, B., Battaglia, G., Azzouz, M., *et al.* (2011). Optimised and rapid pre-clinical screening in the SOD1(G93A) transgenic mouse model of amyotrophic lateral sclerosis (ALS). *PloS one* *6*, e23244.

Ono, K., Condron, M.M., and Teplow, D.B. (2009). Structure-neurotoxicity relationships of amyloid beta-protein oligomers. *Proceedings of the National Academy of Sciences of the United States of America* *106*, 14745-14750.

Orte, A., Birkett, N.R., Clarke, R.W., Devlin, G.L., Dobson, C.M., and Klenerman, D. (2008). Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proceedings of the National Academy of Sciences of the United States of America* *105*, 14424-14429.

Prusiner, S.B. (2012). Cell biology. A unifying role for prions in neurodegenerative diseases. *Science* *336*, 1511-1513.

Qiang, W., Yau, W.M., Luo, Y., Mattson, M.P., and Tycko, R. (2012). Antiparallel beta-sheet architecture in Iowa-mutant beta-amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America* *109*, 4443-4448.

Reaume, A.G., Elliott, J.L., Hoffman, E.K., Kowall, N.W., Ferrante, R.J., Siwek, D.F., Wilcox, H.M., Flood, D.G., Beal, M.F., Brown, R.H., Jr., *et al.* (1996). Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury. *Nature genetics* *13*, 43-47.

Sakagashira, S., Hiddinga, H.J., Tateishi, K., Sanke, T., Hanabusa, T., Nanjo, K., and Eberhardt, N.L. (2000). S20G mutant amylin exhibits increased in vitro amyloidogenicity and increased intracellular cytotoxicity compared to wild-type amylin. *The American journal of pathology* *157*, 2101-2109.

Sakono, M., and Zako, T. (2010). Amyloid oligomers: formation and toxicity of A $\beta$  oligomers. *The FEBS journal* *277*, 1348-1358.

Sawaya, M.R., Sambashivan, S., Nelson, R., Ivanova, M.I., Sievers, S.A., Apostol, M.I.,

Thompson, M.J., Balbirnie, M., Wiltzius, J.J., McFarlane, H.T., *et al.* (2007). Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* *447*, 453-457.

Shankar, G.M., Li, S., Mehta, T.H., Garcia-Munoz, A., Shepardson, N.E., Smith, I., Brett, F.M., Farrell, M.A., Rowan, M.J., Lemere, C.A., *et al.* (2008). Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nature medicine* *14*, 837-842.

Thompson, M.J., Sievers, S.A., Karanicolas, J., Ivanova, M.I., Baker, D., and Eisenberg, D. (2006). The 3D profile method for identifying fibril-forming segments of proteins. *Proceedings of the National Academy of Sciences of the United States of America* *103*, 4074-4078.

Tomic, J.L., Pensalfini, A., Head, E., and Glabe, C.G. (2009). Soluble fibrillar oligomer levels are elevated in Alzheimer's disease brain and correlate with cognitive dysfunction. *Neurobiology of disease* *35*, 352-358.

Tycko, R., Sciarretta, K.L., Orgel, J.P., and Meredith, S.C. (2009). Evidence for novel beta-sheet structures in Iowa mutant beta-amyloid fibrils. *Biochemistry* *48*, 6072-6084.

Valentine, J.S., Doucette, P.A., and Zittin Potter, S. (2005). Copper-zinc superoxide dismutase and amyotrophic lateral sclerosis. *Annual review of biochemistry* *74*, 563-593.

Wang, J., Slunt, H., Gonzales, V., Fromholt, D., Coonfield, M., Copeland, N.G., Jenkins, N.A., and Borchelt, D.R. (2003). Copper-binding-site-null SOD1 causes ALS in transgenic mice: aggregates of non-native SOD1 delineate a common feature. *Human molecular genetics* *12*, 2753-2764.

Wang, J., Xu, G., and Borchelt, D.R. (2002). High molecular weight complexes of mutant superoxide dismutase 1: age-dependent and tissue-specific accumulation. *Neurobiology of disease* *9*, 139-148.

Wong, P.C., Pardo, C.A., Borchelt, D.R., Lee, M.K., Copeland, N.G., Jenkins, N.A., Sisodia, S.S., Cleveland, D.W., and Price, D.L. (1995). An adverse property of a familial ALS-linked SOD1 mutation causes motor neuron disease characterized by vacuolar degeneration of mitochondria. *Neuron* *14*, 1105-1116.

## Chapter 2

Structure based L-peptide inhibitors of SEVI fibrils.

### Abstract

248PAP286 is a proteolytic peptide naturally present in semen. It forms abundant amyloid fibrils that drastically enhance HIV infection. The fibrillar form of the peptide is termed Semen-derived Enhancer of Viral Infection (SEVI). This enhancement is only observed with the amyloid fibrils, monomeric peptide is completely inactive. Thus preventing the fibril formation could potentially reduce the rate of HIV transmission. I obtained crystal structures of two regions in 248PAP286, GGVLVN and KLIMY, that have high propensity to form the core structure of SEVI fibrils. Mutation studies confirmed that both regions are important for SEVI formation. Bases on these structures, we designed L-peptide inhibitors. These inhibitors are designed to interact with GGVLVN and KLIMY and cap the growing end of the fibril to inhibit fibril growth. From three rounds of design and testing, we obtained 5 active L-peptide inhibitors that prevent both seeded and unseeded SEVI formation. These inhibitors also counteract SEVI's activity in enhancing HIV infection in a cell-based assay.

## Introduction

Amyloids are insoluble aggregates formed by misfolded cellular proteins. These proteins contain diverse sequence compositions, native structures and functions. Upon environmental triggers that are not entirely understood, they can all enter the amyloid state, forming aggregates that share a similar structure. These amyloid deposits can be positively stained by histological dyes such as Thioflavin T and Congo Red. They form elongated, un-branched fibrillar structure under electron microscope (Eisenberg and Jucker, 2012). Additionally, these fibrillar structures formed by different amyloid proteins all share a similar cross-beta diffraction pattern, suggesting that similar structural architecture is present in these fibrils (Sipe and Cohen, 2000). The spine of the fibril is made of in-registered  $\beta$ -strands stacked along the fiber axis, forming paired sheets that are parallel to each other. The kinetics of fibril formation is described as a nucleation dependent process. The nucleation phase is considered the rate-limiting step and produces a lag time before aggregation can be detected by histological dyes.

25+ diseases have been linked to amyloid formation. An amyloid protein discussed in this work, has been found to enhance HIV transmission. HIV pandemic is a serious concern with millions of new infections occurring each year. Among these cases, more than 80% are acquired through sexual intercourse. Since the development of effective HIV vaccines is still challenging, new strategies for prevention are of urgent need. Pre- and post -exposure prophylaxis (PrEP and PEP respectively) has been a growing interest. On-going clinical trials of PrEP and PEP entirely involve the use of antiretroviral drugs. Here, we tested an innovative strategy of host-directed, pre- and potentially post-exposure prophylaxis for HIV prevention. A host-directed method is likely to bypass the rapid emergence of a drug-resistant virus. In addition, our strategy does not

involve manipulation of the host immune system. This offers a potential unconventional, safe and effective way of preventing the transmission of HIV.

In 2007, our collaborator Jan Munch's laboratory identified and characterized a peptide fragment in semen that can drastically enhanced the HIV infection rate. This peptide fragment is generated from the prostatic acidic phosphatase (PAP), a protein that is present in semen in milligram quantities. With it present, one to three virions is sufficient for HIV infectivity (Dimitrov et al., 1993; Munch et al., 2007). Surprisingly, the enhancement is not merely due to the presence of this peptide segment (248PAP286). Munch et al. showed that newly synthesized peptides are completely inactive. HIV infectivity is only enhanced when 248PAP286 forms structured aggregates, amyloid fibrils. These PAP fibrils are termed Semen-derived Enhancer of Viral Infection (SEVI) (Munch et al., 2007). This study proposed a new role for amyloid fibrils in enhancing HIV infectivity that can lead to a new opportunity to prevent HIV transmission.

Amyloid fibrils have been studied for many years due to their association with neurodegenerative diseases. It is known that despite the difference in amyloid proteins' primary sequences, the fibrils (including SEVI) formed all share similar characteristics, such as cross-beta diffraction pattern and dye binding affinity (Nelson et al., 2005; Sawaya et al., 2007; Sunde and Blake, 1997). The development of structure-based inhibitors that exhibit high specificities is made possible by our understanding of the fibril atomic structure. In 2005, Rebecca Nelson *et al.* solved the first crystal structure of the cross- $\beta$  fibril spine (Nelson et al., 2005). This is a seven-residue peptide segment from Sup35, a yeast prion protein that forms amyloid-like fibrils. Because amyloid fibrils formed by various proteins share similar core structure, we used this structure from Sup35 as a template and developed computational algorithms that predict segments in other amyloid proteins that have high propensity to form the spine of the fibrils. The

effectiveness of this algorithm has been validated by several studies (Sawaya et al., 2007; Thompson et al., 2006; Wang et al., 2008; Wiltzius et al., 2008). The spines of the fibrils exhibit extensive hydrogen bonding and are important for the stability of the fibrils. These regions represent important targets for inhibition of fibril formation. Using this prediction algorithm, we have detected six segments in 248PAP286 that have high propensity to form amyloid fibrils. Initial computational modeling of the fibril interface suggests two of these six segments are especially plausible. The importance of these regions has been further confirmed with mutation studies. I determine the atomic structure of the two high propensity segments. These structures were served as target templates for inhibitor design.

The most desirable inhibitor should bind to fibril forming segments of 248PAP286 with high affinity and specificity. A design algorithm has developed by our lab aiming to establish a well-packing interface between the fibril capping D-peptide and steric zipper core of amyloid protein. One inhibitor is designed to cap one beta-sheet at each end of the fibril, thus blocking the elongation of fibril at early stage of the process. This design method has shown success *in vitro* with tau, a classic amyloid protein associated with Alzheimer's disease. Since a substantial amount of experimental results indicate amyloid fibrils share the same general core structure, our algorithm, with optimization, can be used to design inhibitors for all amyloid fibrils. This same approach was used in designing SEVI fibril blockers. After three rounds of design and testing, we have identified a set of active L-peptide inhibitors. In collaboration with Jan Munch, top *in vitro* candidates were tested with HIV in semen. In this study, we have obtained effective SEVI aggregation inhibitors that counteracted SEVI mediated enhancement of HIV infection.

## Results

### High propensity segment prediction

248PAP286 is a 39-residue peptide that forms amyloid fibrils both in vivo and in vitro. Our prediction algorithm predicted three regions in this peptide to have high propensity to form the steric zipper core structure of the fibril spine, residues 248 - 253, 260 – 268 and 279 – 286 (Figure 1a). Previous work has shown that truncation at the N-terminus does disrupt fibril formation. In contrary, deleting of the C-terminal residues abolishes fibril formation of synthetic peptides. Additionally, I generated a valine to proline mutation in region 260-268, which also abolishes 248PAP286's ability to form amyloid fibrils (Figure 2). Therefore, we focus our study on region 260 – 268 and region 279 – 286. We have chemically synthesized peptide segments 260-GGVLVN-265 and 282-KLIMY-286. We showed that both segments in isolation are capable of forming abundant fibrils as seen under electron microscopy (Figure 1b, c).

### Crystal structures of GGVLVN and KLIMY

We obtained crystal structure of the amyloidogenic segments GGVLVN and KLIMY. Both crystals were obtained using hanging drop vapor diffusion method. Both GGVLVN and KLIMY form steric zipper structures containing mating  $\beta$ -sheet tightly packed together with a dry interface. KLIMY crystal diffracted to 1.7Å and structure was solved by molecular replacement (Table 1). KLIMY structure contains anti-parallel  $\beta$  strands forming the mating sheets. Four possible steric zipper interfaces have been identified (Figure 3). GGVLVN formed class 1 steric zipper as defined by Sawaya et al. (ref). This structure is composed of parallel, in registered  $\beta$  strands with only one type of steric zipper interface present (Figure 4). The details of GGVLVN structure and refinement statistics are reported in chapter 3.



## **SEVI fibril model**

GGVLVN and KLIMY both form steric zipper structures and are capable of forming fibrils in isolation. When predicting a putative SEVI fibril model for the development of inhibitors, we hypothesized that GGVLVN forms the fibril spine for following reasons. Large number of studies has shown that the most mature fibrils are made of parallel in-registered  $\beta$  strands. Other steric zipper structures with similar structural motif to GGVLVN have been obtained from regions that are shown to be critical for other amyloid fibril formation, including Tau, Sup35 and Amyloid  $\beta$ . Fibril diffraction on SEVI showed classic cross- $\beta$  diffraction pattern suggesting that SEVI fibrils share similar core structure as the other well studied amyloid proteins (Figure 5a). Therefore, we took the GGVLVN crystal structure as the spine of SEVI fibril, and we speculate that KLIMY region is important for fibril stability by forming hetero zipper interaction with GGVLVN. We then modeled the rest of the peptide using Rosetta Energy minimization method and top-scoring model is shown in figure 5b.

## **Fibrillation of recombinant 248PAP286**

To further characterize SEVI fibrils, we developed a method to express and purify recombinant 248PAP286 peptide in an *E. coli* expression system. We have found that 248PAP286 has high tendency to be degraded by bacterial protease. After testing a dozen or so constructs, we found that MBP fusion is the only construct that gave the expression of intact 248PAP286 peptide (Figure 6) . The peptide is cleaved from MBP using a TEV protease with no extra residue left from the cleavage site (Figure 6). We have found that this recombinant peptide form abundant fibrils after incubation at 37 °C with continuous shaking overnight. The lag time of this fibril formation at 5mg/ml is around 3 hours  $\pm$  1hour. When small amount (5% v/v) of

preformed SEVI fibrils was added to the fibrillation assay, the lag time is reduced to less than 1 hour (Figure 7). The seeding effect is characteristic of amyloid fibrillation kinetics. We used this recombinant peptide for the study of the activities of our designed inhibitors.

### **Inhibitors of SEVI formation**

The GGVLVN and KLIMY structures gave us templates for designing structure-based inhibitors. We went through three rounds of optimization and testing. Fibrillation kinetics with the addition of active inhibitors is shown in figure 8. In the last round, we have incorporated non-natural amino acid from which we obtained the inhibitor WW61. Details of the non-natural inhibitor design and characterization of WW61 is discussed in chapter 3. In this chapter, we only report L-peptide inhibitors with promising activities.

To test the activity of the inhibitors we adapt a fluorescent-based assay with Thioflavin T dye. ThT only binds to fibrils, but not monomers or oligomers. Therefore, we evaluate the rate of fibril formation by monitor the lag time. A longer lag time should be expected with the addition of active inhibitors. As detailed in chapter 3, these inhibitors are designed to bind to the growing end of the amyloid fibrils that prevent elongation and therefore the formation of mature fibrils. In this study, we tested the inhibitors against both unseeded and seeded SEVI fibrillation.

Designed inhibitors should show inhibitory activities in both assays. Using these criteria, we have identified 5 active L-peptide inhibitors, WHRLYI, WHKLYI, HYRLYM, HYKWW and YTWKW. The results showed in this chapter are with inhibitors added as 2-fold molar excess. We have tested these L-peptide inhibitors as equal-molar or less, and observed little activity. Interestingly, YTWKW is an inhibitor designed to block the interaction of KLIMY. It is more active against unseeded fibrillation, and much less active against seeded fibrillation. This is

consistent with our hypothesis that GGVLVN forms the main interaction at the core structure of SEVI fibril. Furthermore, we tested one of the active inhibitor HYRLYN against fibrillation of a different amyloid protein A $\beta$ . HYRLYN is completely inactive against A $\beta$  fibril formation, supporting our design that the inhibitors are selective to GGVLVN or KLIMY structures.

### **SEVI inhibitors reduces HIV infection rate in a functional assay**

SEVI was identified to increase HIV infection. Therefore, we tested these inhibitors in a functional assay. In these assays, we incubate 248PAP286 peptide *in vitro* with and without inhibitors and these samples were added into an HIV infectivity assay. We have shown that the inhibitors that prevented SEVI aggregation also abolished SEVI mediated HIV infection. The infectivity with the presence of inhibitors is comparable to that of monomeric 248PAP286 peptide, consistent with our design rationale that these inhibitors prevent the formation of HIV enhancing material.

### **Discussion**

284PAP286 is an extracellular peptide that is naturally present in semen. The amyloid form of this peptide drastically enhances HIV infection. This discovery offered a relationship between HIV infection and other amyloid diseases, such as Alzheimer's disease, Parkinson's disease, and bovine spongiform encephalopathy. A study published in 2002 by Wojtowicz et. al. has suggested that  $\beta$ -amyloid fibrils which cause neurodegeneration in Alzheimer's disease also enhance HIV infection, albeit with a lower efficiency (Wojtowicz et al., 2002). Both studies propose a new role for amyloid fibrils in enhancing HIV infectivity that can lead to a new opportunity to prevent HIV transmission.

The amino acid composition of 248PAP284 peptide includes a number of positively charged residues at physiological pH. Previous work on amyloid fibrils suggest that individual peptide unite arrange in a parallel in-registered fashion. Thus, upon fibril formation, positive residues will line up along the side of the fibrils and provide local high concentration of positive charges. These cations are suggested to facilitate HIV fusion to the target cells by reduce the repulsion force between the negatively charged membranes. A previous study showed that a mutant peptide with all the lysine and arginine residues mutated to alanine still forms amyloid fibrils (Roan et al., 2011; Roan et al., 2009). However, the fibrils formed by the mutant peptide no longer show HIV enhancing activity. Therefore, inhibitors that prevent this fibril formation could potentially reduce the HIV infection rate.

In this study, we have identified two regions in the 248PAP284 that are important for SEVI formation and obtained crystal structures of these segments. GGVLVN formed a parallel in-registered steric zipper structure and KLIMY formed an antiparallel steric zipper structure. Based on previous knowledge about amyloid fibrils, we know that the core structures in majority of these fibrils are made with parallel  $\beta$  strands. Therefore, we believe GGVLVN and KLIMY stabilize SEVI fibrils in different ways. Since GGVLVN forms the most common motif observed in amyloid fibril spines, we modeled this region to form the core structure and KLIMY stabilize the fibrils by secondary interaction with GGVLVN.

We then designed L-peptide inhibitors based on the structure of GGVLVN and KLIMY. We identified 5 inhibitors that prevent both seeded and unseeded SEVI formation. Additionally, this inhibitor prevented SEVI mediated HIV infection. While testing these inhibitors, we have noticed that these peptide inhibitors are only active at two fold molar excess or more. On average, at this concentration, the lag time can be delayed by a factor of two. At five fold molar excess;

the inhibition can last for at least 2 days. Very little activity was observed at equal molar or sub molar ratio of inhibitor to 248PAP286. It is worth noting that this ratio is the molar ratio of inhibitor to monomeric peptide. Since the inhibitors are designed to only bind to the growing end of the fibrils, and each fibril is made with hundreds of peptides, the effective molar ratio between inhibitor and SEVI fibrils is significantly lower. The specific binding affinity is difficult to calculate due to the lack of understanding of the number of fibril ends present in the solution.

We have also noticed that the lag time for SEVI formation varies significantly among replicates. The variation is generally larger with different batches of recombinant proteins. Additionally, the storage at -20 °C for prolonged period of time also leads to a change in fibrillation kinetics. Since the nucleation of fibrillation is a stochastic event, these variations in lag times are commonly observed. For these reasons, we have always included a 248PAP286 only control when testing these inhibitors. Additionally, adding preformed seeds can reduce the variations in lag time among replicates.

Even though the active inhibitors consistently show inhibitory activities, we do observe different levels of inhibitor activities with different batches of 248PAP286 peptides. Due to these variations, quantification of the inhibitor activity is challenging. Additionally, in this study, we tested the activity of these inhibitors with synthetic 248PAP286 peptides. Further experiments need to be done to evaluate the activities of the inhibitors in the presence of seminal fluid.

Our work showed that structured based inhibitor design is viable strategy for preventing amyloid fibril formation. Due to the common structure present in majority of the amyloid fibrils, this strategy can be used as a general technology platform that is useful for developing inhibitors for other amyloid proteins. This approach is particularly useful for systematic amyloidosis where 1) there is a clear link between fibril deposit and the development of the diseases and 2) the fibril

deposit is outside of the brain since these peptide inhibitors cannot cross the blood brain barrier. Furthermore, stability of the inhibitors needs to be evaluated in the content of body fluid with the presence of many proteases. Our lab has developed computation method that incorporates non-natural residues to these peptide inhibitors as discussed in chapter 3.

## **Materials and Methods**

### **248PAP286 peptide expression and purification**

248PAP286 expression and purification have been detailed in chapter 1. Briefly, synthetic DNA (Invitrogen, Carlsbad, CA) has been PCR amplified with 5' SacI and 3' XhoI restriction sites. It is then cloned into a custom vector, a chimera constructed from the NdeI and XhoI digestion products pET15b (Novagen, Gibbstown, NJ), and the maltose binding protein (MBP) gene from pMALC2X (New England Biolabs, Ipswich, MA), resulting in an N-terminal His-tag MBP fusion vector. SEVI mutant was made with site directed mutagenesis with QuikChange Site-Directed mutagenesis kit from Stratagene. MBP-248PAP286 fusion proteins are expressed in BL21-DE3(Gold) cells (Agilent Technologies, Santa Clara, CA) in LB miller media supplied with 100 µg/mL ampicillin (Fisher Scientific, Pittsburgh, PA). Protein expression is induced with 0.5mM IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) once culture reached an OD600 ~0.6. Strong expression level is observed after 2-3 hours of induction at 34 °C.

All buffers used in this purification need to be pre-chilled, since the 248PAP286 peptide is largely disordered and degradation-prone. His-MBP- 248PAP286 fusion protein is first purified with a HisTrap-HP column, where the target protein elutes at 300 mM Imidazole. The elute sample is then dialyzed in a Slide-A-Lyzer 10,000 MWCO dialysis cassette (Pierce, Thermo Fisher Scientific, Rockford, IL), against buffer 50 mM Tris buffer pH 7.5 and 50 mM

NaCl at 4 °C for at least 6 hours. The dialyzed sample was then pooled and 1/500 volume of TEV protease stock was added. TEV cleavage reaction was incubated again at 4 °C for overnight where close to 100% cleavage rate can be achieved. TEV protease reaction was then re-loaded onto the HisTrap-HP column. In this purification, buffer A is 20mM Tris pH7.5, 500mM NaCl and buffer B is 20mM Tris pH7.5, 500mM NaCl, 500mM Imidazole. Load protein sample and completely wash of the contaminants with buffer A and then elute 248PAP286 with 10% buffer B. 248PAP286 has two histadines which enables the peptide to bind to Ni-column with some affinity. This peak is then further purified using reverse phase high performance liquid chromatography (RP-HPLC) on a 2.2 x 25 cm Vydac 214TP101522 column equilibrated in buffer RA (0.1% trifluoroacetic acid (TFA)/water) and eluted over a linear gradient from 0% to 100% buffer RB (Acetonitrile/0.1% TFA) in 40 minutes at a flow rate of 9 mL/min. Absorbance at 220nm and 280nm were recorded using a Waters 2487 dual  $\lambda$  absorbance detector (Waters, Milford, MA). Peak fractions containing peptide were assessed for purity by a MALDI-TOF mass spectrometry (Voyager-DE-STR, Applied Biosystems, Carlsbad, CA). Pooled fractions were frozen in liquid nitrogen and lyophilized.

### **Prepare SEVI peptide stock**

Lyophilized SEVI peptide stock was dissolved in HFIP as 10mg / ml to remove any preformed fibril seeds or small aggregates. Aliquot the sample as 1ml stocks, each contains 10mg peptide. Leave the sample in a chemical hood for HFIP to evaporate over night. Then remove the HFIP completely with a SpeedVac concentrator. Samples were stored at -80 °C.

### **Prepare seed stock**

Each aliquot of the SEVI sample is dissolved as 5mg/ml SEVI in PBS and was agitated at 37 °C for 2days to form mature fibrils. Fibril seeds were prepared by sonication at 25% amplitude, 1 second on; 1second off for total of 20 second. Seed solution was stored at 4 °C.

### **Inhibition Assay Set up**

Inhibitors were chemically synthesized from CS Bio Co. 248PAP286 was made recombinant from E coli. 248PAP286 monomer final concentration is 0.66mM (3mg/ml) and inhibitor final concentration is 1.32mM. In seeded experiment, seed final concentration is 3% (v/v). Inhibitor peptides were dissolved in PBS as 5X working stock and were filtered with 0.45 µm filters. 248PAP286 were dissolved as 1.25X working stock and filtered with 0.45 µm filters. Add inhibitor to 248PAP286 working stock and rotate mix SEVI and inhibitor solutions for 10 min. Add ThT solution to final 10 µM. Pipet mix the seed solution and then add to 248PAP286 and inhibitor mix to 3% (v/v). Pipet this mixture to a 96-well Nunc optical clear bottom black plate and sealed with Corning pressure seal tape. 3 replicates were used for each assay condition. In control samples, PBS buffer was used instead of inhibitor peptide. Kinetic assay was carried out with a Varioskan plate reader (Thermo Fisher Scientific Inc.). Fluorescent signal was measured at excitation wavelength of 440 nm and emission wavelength 482 nm at 15 minute intervals. The assay plate was incubated in the plate reader at 37 °C with continuous shaking at 960 rpm with a diameter of 1mm.



## **Electron Microscopy**

5  $\mu\text{L}$  of sample was applied to glow discharged 400 mesh carbon-coated formvar films on copper grids for three minutes. The grid was washed 3 times with 0.22  $\mu\text{m}$  filtered water and stained with 1% uranyl acetate for one minute. Grid were examined with JEM 1200-EX is a transmission electron microscope with wide angle (top mount) BioScan 600W 1x1K digital camera (Gatan). Images were processed with Digital Micrograph acquisition software (Gatan).

## **KLIMY structure determination**

KLIMY peptide was chemically synthesized from CS Bio Co. Dry peptide was dissolved as 10 mg/ml in water. Reservoir solution contains 1.0M Ammonium Sulfate, 0.1M HEPES pH7.0. Crystals appear at ambient temperature after 1 day. Single crystal was mount onto a glass needle and store at room temperature. X-ray data was collected at the Advanced Photon Source (Chicago, IL) beam line 24-ID-E. Data were processed using DENZO (52) and XDS (53). Phase information was obtained by molecular replacement using PHASER. Model refinement was done using REFMAC. Statistics pertaining to structure determination are provided in table 1.

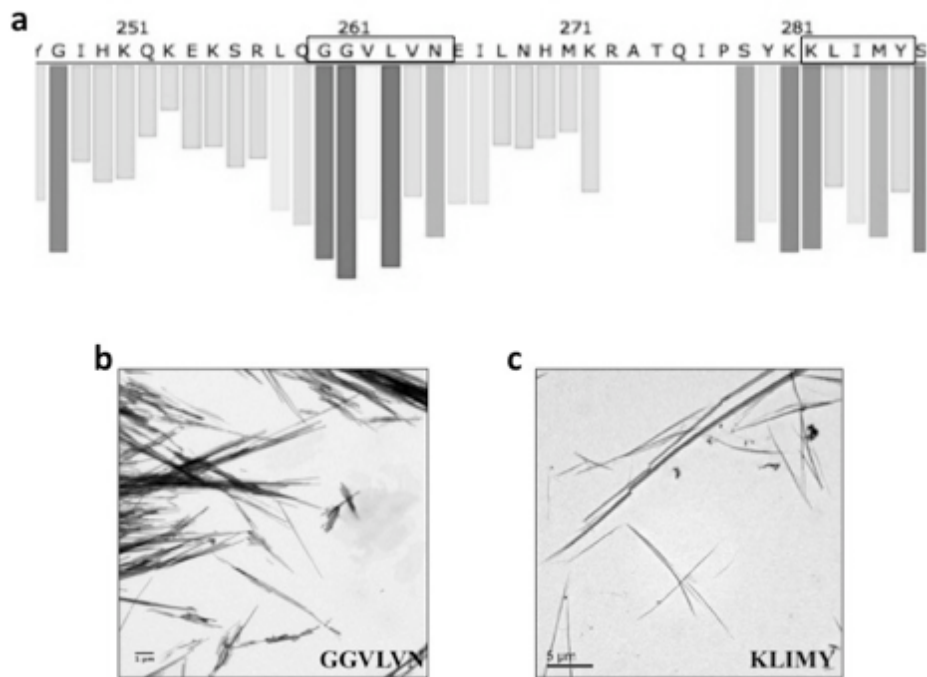
## **Fiber diffraction**

SEVI fibril samples were concentrated by centrifugation at 13,000  $\times g$  for 30 min followed by washing with water two times to remove the salts in PBS buffer. Then, the pellets were resuspended into 10  $\mu\text{L}$  water. The 5  $\mu\text{L}$  of the suspension was pipetted in between two closely spaced, fire-polished glass rods. The samples are dried at room temperature overnight. During this process, SEVI fibrils are aligned in between the glass rod. The fibril samples on the glass rod were mounted on the in-house X-ray machine using a Rigaku FR-D X-ray generator

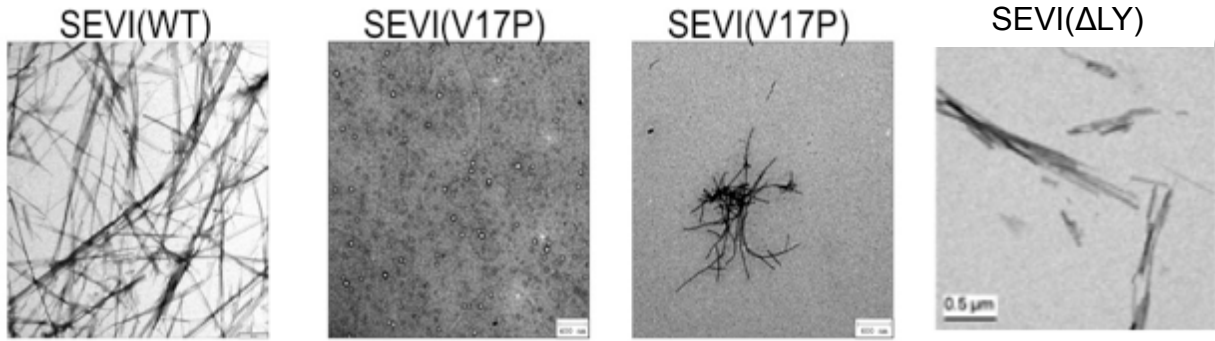
equipped with an RIGAKU HTC Imaging plate detector. The diffraction data were collected with 5° rotation oscillation for 5-min exposures.

### **HIV infectivity assay**

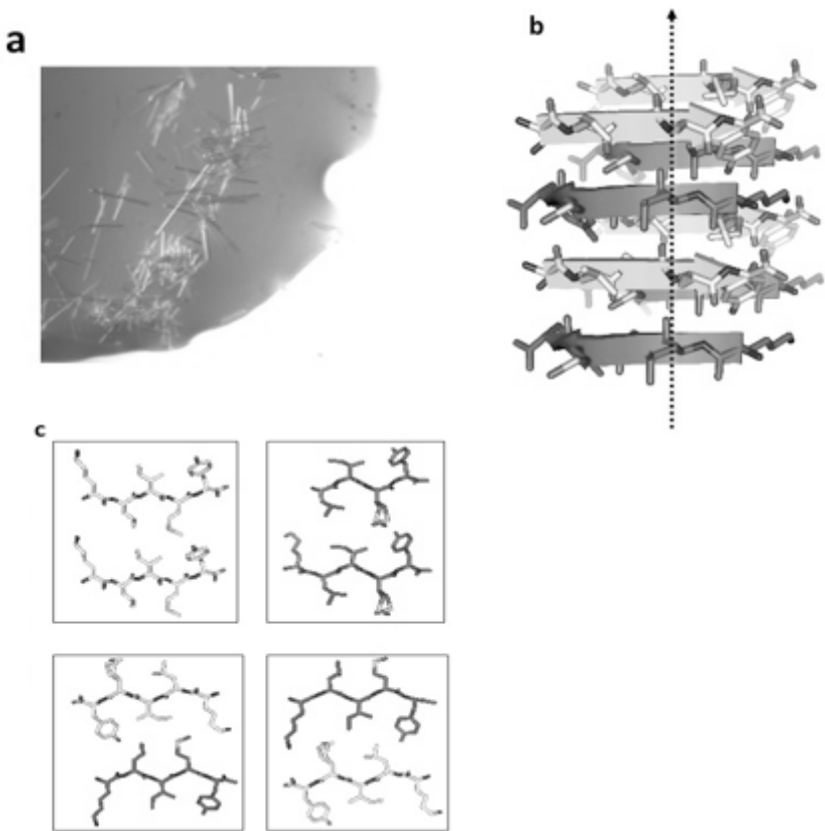
The CCR5 tropic molecular HIV-1 clone NL4\_3/92TH014-2 (Papkalla et al., 2002) was generated by transient transfection of 293T cells with proviral DNA. Supernatants were collected 48 hrs later and p24 concentrations determined by ELISA. TZM-bl reporter cells encoding a lacZ gene under the control of the viral LTR promoter were obtained through the NIH AIDS Research and Reference Reagent Program and were provided by Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc.) (Platt et al., 1998). 40  $\mu$ l HIV-1 containing 0.1 ng p24 antigen was incubated with 40  $\mu$ l dilutions of mixtures of 248PAP286 and inhibitory peptides that were either freshly prepared (t=0) or that had been agitated for 23 hrs (t=23). Peptide concentrations and experimental conditions during agitation have been described above. Thereafter, 20  $\mu$ l of virus/248PAP286/inhibitor mixtures were used to infect 180  $\mu$ l TZM-bl cells seeded the day before ( $1 \times 10^5$ /well). 2 days later infection rates were determined by quantifying  $\beta$ -galactosidase activities in cellular lysates using the Gal-Screen assay as recommended by the manufacturer (Applied Biosystem, T1027). Luminescence was recorded on an Orion microplate luminometer.



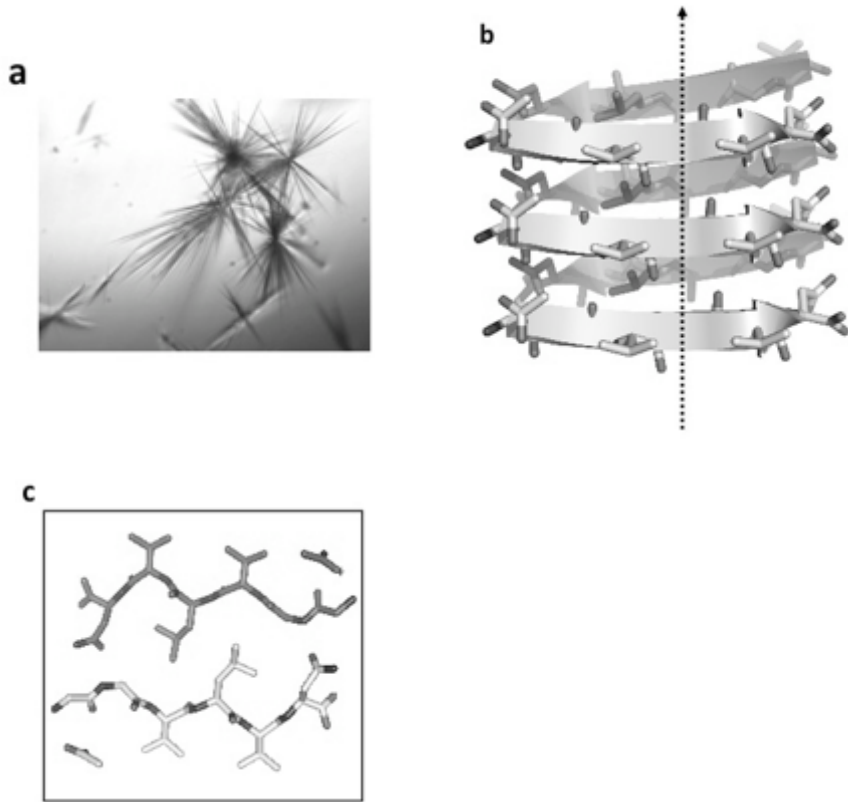
**Figure 1. High propensity segment prediction.** Rosetta-profiling algorithm generated three regions in 248PAP286 that have high propensity to form the core structure of SEVI fibrils (a). Truncation study has shown that the N-terminal residues are not required for SEVI formation. Thus, we chemically synthesized the other two high propensity segments GGVLN and KLIMY. Electro micrographs show that both segments form abundant fibrils as 10 mg/ml in water (b, c).



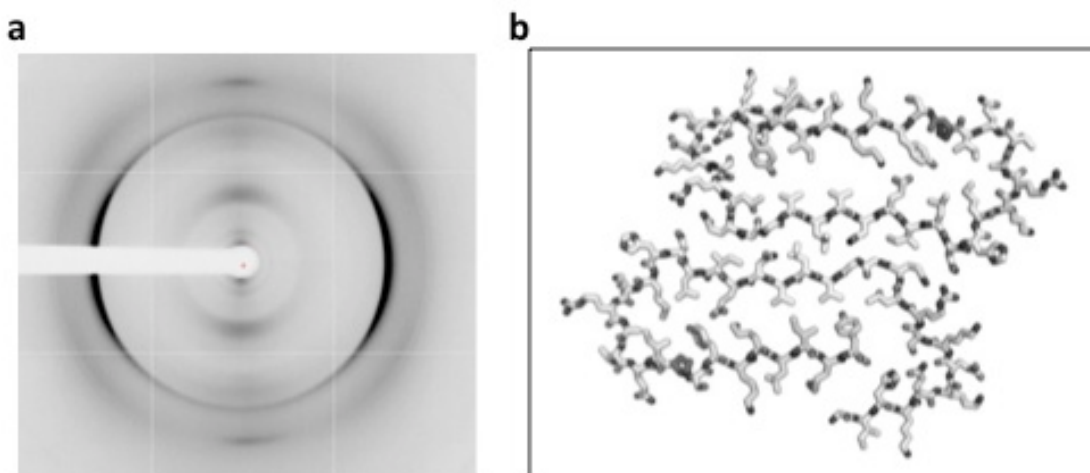
**Figure 2. Fibrillation assay of proline mutant.** To confirm the role of the high propensity segments, I made a proline mutation in the GGVLVN region of the 248PAP286 peptide. Recombinant peptides were dissolved in PBS at 10mg/ml, and incubated at 37 °C with shaking for over night. WT 248PAP286 formed abundant fibrils and V17P mutation disrupted the fibril formation. In the 248PAP286<sup>V17P</sup> sample, most abundant species under EM is spherical aggregates with sparse fibrils. Munch et al constructed a truncation mutant at the C-terminus where LY was deleted from KLIMY region. This construct also significantly disrupted fibrils formation. These mutagenesis studies confirmed our prediction results that GGVLVN and KLIMY is important for fibril formation.



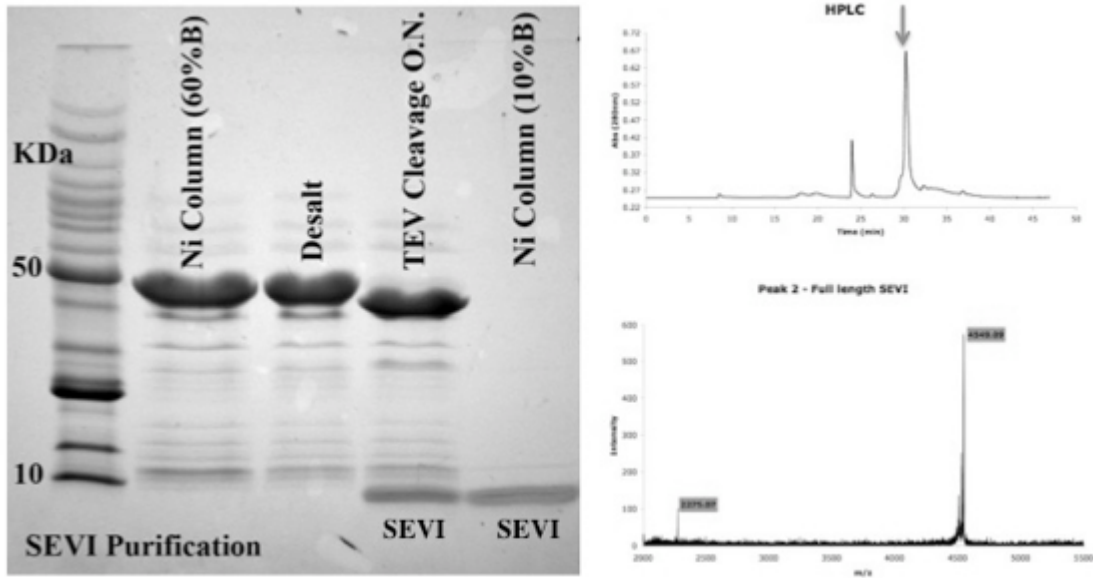
**Figure 3. Crystal structure of KLIMY.** KLIMY form thick needle like crystals (a). Side view of two mating sheet represent the structure of a putative fibril spine (b). Arrow indicates the fibril axis. Each  $\beta$ -sheet is made of anti-parallel  $\beta$ -strands and KLIMY forms dry interface. Four possible steric zipper interfaces are shown in c.



**Figure 4. Crystal structure of GGVLVN.** GGVLV form thin needle crystals (**a**). Side view of two mating sheet represent the structure of a putative fibril spine (**b**). Arrow indicates the fibril axis. Each  $\beta$ -sheet is made of parallel  $\beta$ -strands and the steric zipper interface is shown in **c**.

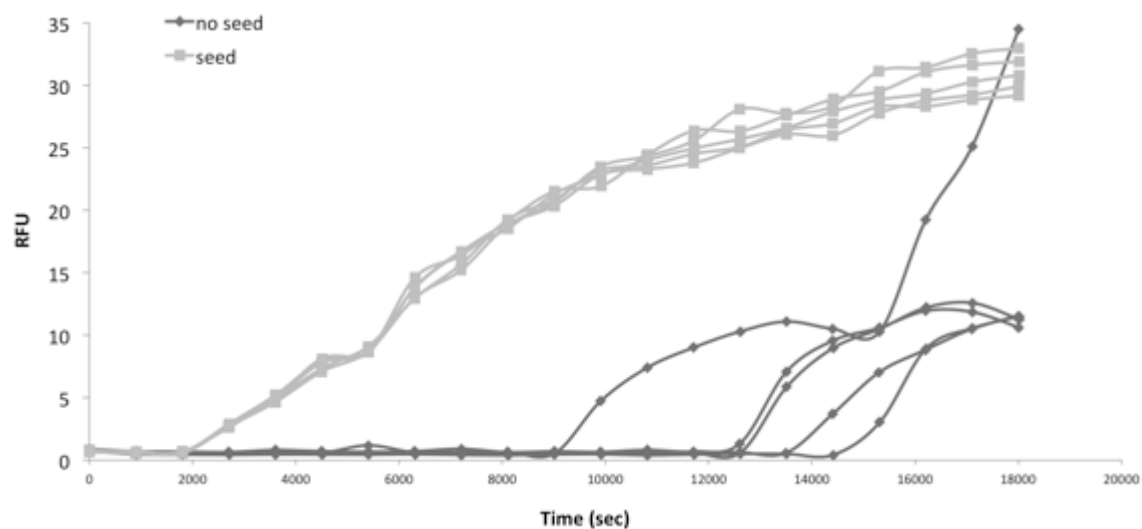


**Figure 5. Fibril diffraction and structural model.** SEVI fibrils were aligned at room temperature. **(a)** X-ray fibril diffraction indicates that SEVI forms the typical cross- $\beta$  confirmation as observed with other amyloid fibrils. We did not observe a diffraction at  $\sim 9.6$  Å that would indicate an anti-parallel fibril structure. **(b)** We used GGVLVN as the core of the SEVI fibrils and modeled the rest of the peptide with KLIMY forming hetero interactions with GGVLVN to stabilize the fibril structure. N-terminal residues are not engaged in major interactions. Computation model was refined with Rosetta-energy minimization method.



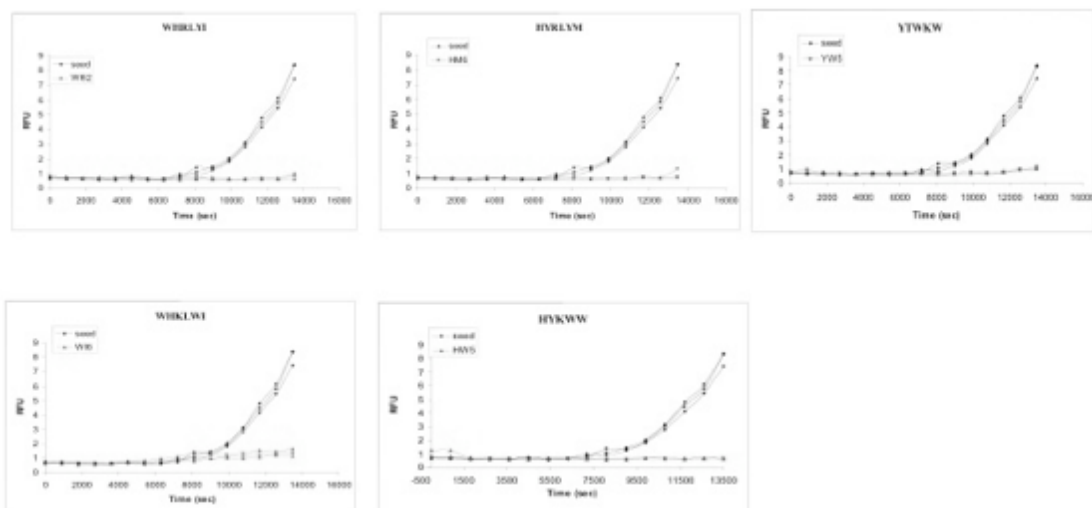
**Figure 6. 248PAP286 expression and purification.** 248PAP286 peptide was expressed as a MBP fusion in *E. coli*. 248PAP286 was first purified with Ni-NTA column and then purified with RP-HPLC. MS spectrum confirm the HPLC elute peak contains pure 284PAP286 peptide.



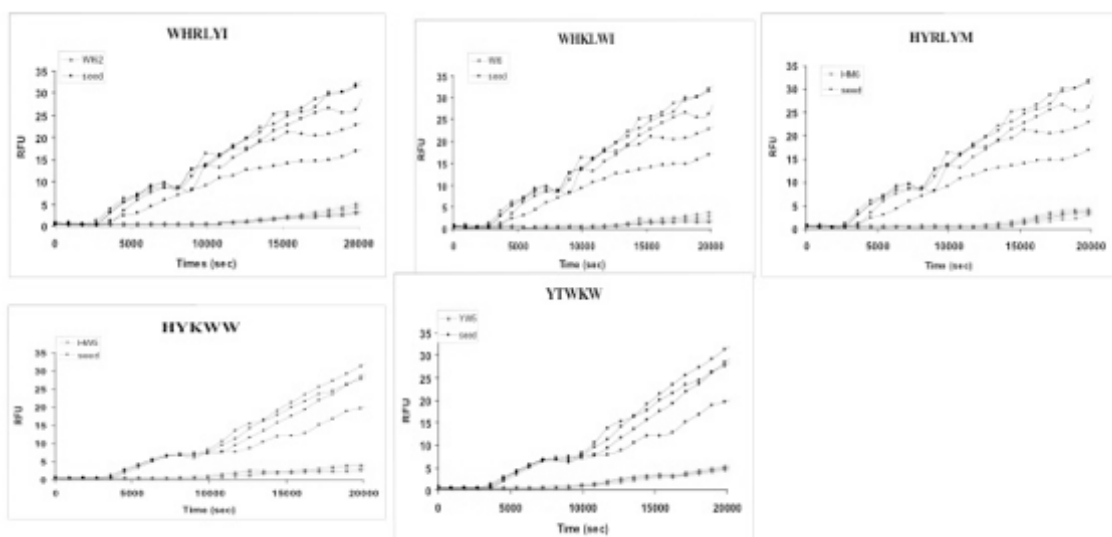


**Figure 7. SEVI fibrillation kinetics.** 5 mg/ml of 248PAP286 peptides were dissolved in PBS either with (light grey) or without seed (dark grey). Both samples were incubated at 37 °C with continuous shaking. Unseeded sample showed nucleation dependent kinetics with large variation in lag time (3 hour  $\pm$  1 hour). Seeded sample form fibrils within the first hour with very little variation in lag times. 5% (v/v) seed was used in this assay.

**a**

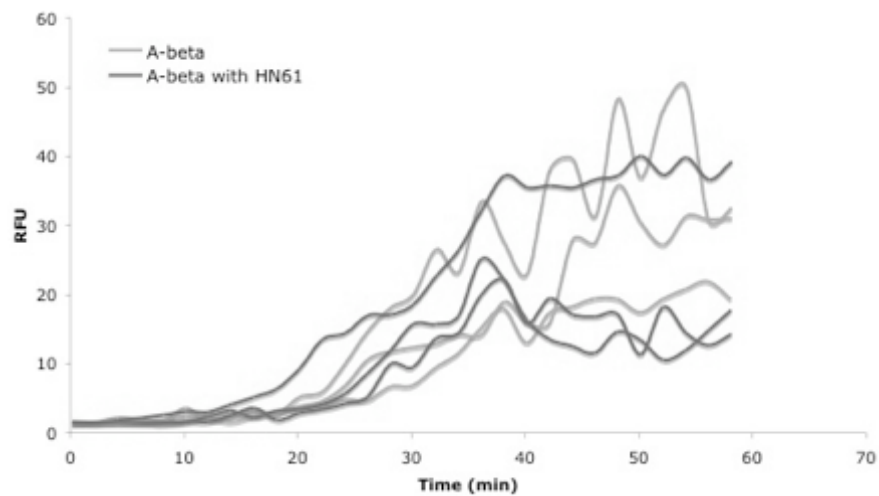


**b**

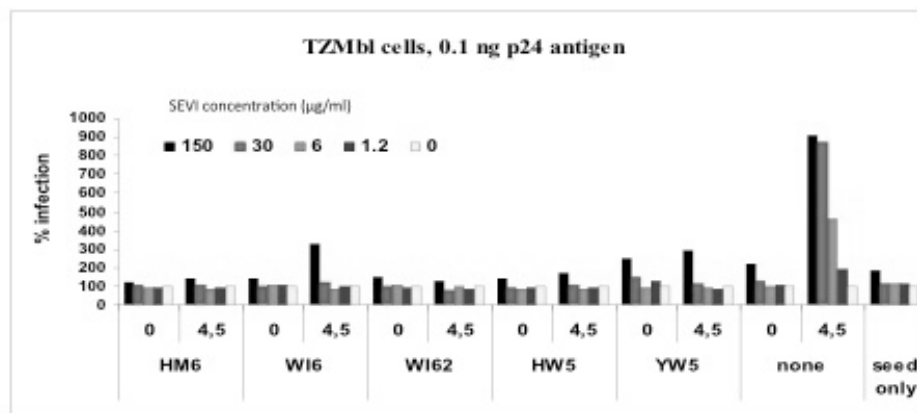
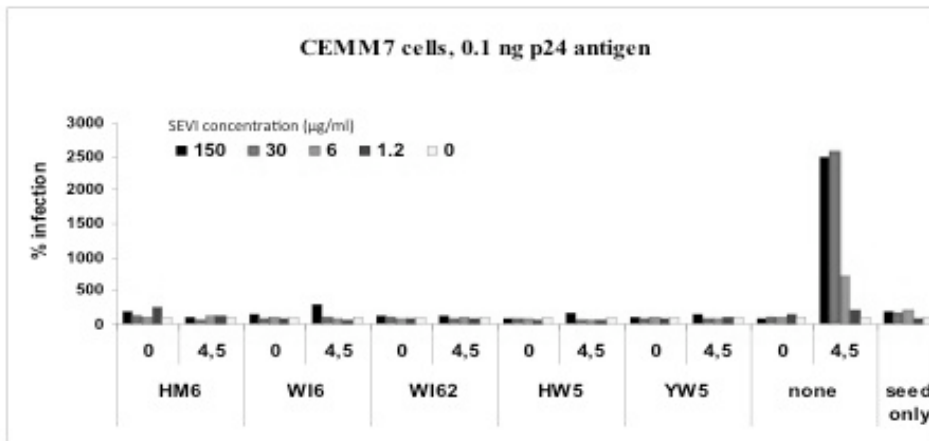


**Figure 8. L-peptide inhibitors against seeded and unseeded fibrillation.** 2-fold molar excess of inhibitor peptides were added to either unseeded (**a**) or seeded (**b**) SEVI fibrillation assays. This figure showed five inhibitors that are able to delay both seeded and unseeded fibrillation. However,

these inhibitors were not able to prevent fibril formation, particularly under seeded conditions, fibril formation is observed near the end of the assay. YTWKW is designed based on KLIMY structure and all other active inhibitors are designed based on GGVLVN structure.



**Figure 9. L-peptide inhibitor does not inhibit A $\beta$  fibrillation.** We tested one of the active L-peptide inhibitors at two fold molar excess against A $\beta$ . The inhibitor designed based on SEVI structure is completely inactive against A $\beta$ , suggesting that the inhibition is not due to non-specific interactions.



**Figure 10. L-peptide inhibitors prevent SEVI-mediated infection.** 40  $\mu$ l HIV-1 containing 0.1 ng p24 antigen was incubated with 40  $\mu$ l dilutions of mixtures of 248PAP286 and inhibitory peptides that were either freshly prepared (t=0) or that had been agitated for 23 hrs (t=23) without seed. 20  $\mu$ l of virus/248PAP286/inhibitor mixtures were used to infect TZM-bl cells and CEMM7 cells seeded the day before. 2 days later infection rates were determined by quantifying  $\beta$ -galactosidase activities in cellular lysates using the Gal-Screen assay. All five active inhibitors significantly reduced rate of SEVI mediated infection in both cell lines.

**Table 1. KLIMY data collection and refinement statistics**

<b>Data Collection</b>	<b>KLIMY</b>
Space group	P1
Resolution	1.7Å
Unit cell dimensions	9.6 11.8 19.3 93.3 90.8 103.1
Measured reflections	2349
Unique reflections	849
Overall completeness	93.80%
Last shell completeness	96.60%
Overall Rsym	0.146
Last shell Rsym	0.581
Overall I/sigma	6.0
Last shell I/sigma	1.8
<b>Refinement</b>	
Rwork	0.248
Rfree	0.261
rmsd bond length	0.004
rmsd bond angle	0.7
Average B factors	12.4

## Reference

- Dimitrov, D.S., Willey, R.L., Sato, H., Chang, L.J., Blumenthal, R., and Martin, M.A. (1993). Quantitation of human immunodeficiency virus type 1 infection kinetics. *Journal of virology* 67, 2182-2190.
- Eisenberg, D., and Jucker, M. (2012). The amyloid state of proteins in human diseases. *Cell* 148, 1188-1203.
- Munch, J., Rucker, E., Standker, L., Adermann, K., Goffinet, C., Schindler, M., Wildum, S., Chinnadurai, R., Rajan, D., Specht, A., *et al.* (2007). Semen-derived amyloid fibrils drastically enhance HIV infection. *Cell* 131, 1059-1071.
- Nelson, R., Sawaya, M.R., Balbirnie, M., Madsen, A.O., Riekel, C., Grothe, R., and Eisenberg, D. (2005). Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435, 773-778.
- Roan, N.R., Muller, J.A., Liu, H., Chu, S., Arnold, F., Sturzel, C.M., Walther, P., Dong, M., Witkowska, H.E., Kirchhoff, F., *et al.* (2011). Peptides released by physiological cleavage of semen coagulum proteins form amyloids that enhance HIV infection. *Cell host & microbe* 10, 541-550.
- Roan, N.R., Munch, J., Arhel, N., Mothes, W., Neidleman, J., Kobayashi, A., Smith-McCune, K., Kirchhoff, F., and Greene, W.C. (2009). The cationic properties of SEVI underlie its ability to enhance human immunodeficiency virus infection. *Journal of virology* 83, 73-80.
- Sawaya, M.R., Sambashivan, S., Nelson, R., Ivanova, M.I., Sievers, S.A., Apostol, M.I., Thompson, M.J., Balbirnie, M., Wiltzius, J.J., McFarlane, H.T., *et al.* (2007). Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* 447, 453-457.
- Sipe, J.D., and Cohen, A.S. (2000). Review: history of the amyloid fibril. *Journal of structural biology* 130, 88-98.
- Sunde, M., and Blake, C. (1997). The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Advances in protein chemistry* 50, 123-159.
- Thompson, M.J., Sievers, S.A., Karanicolas, J., Ivanova, M.I., Baker, D., and Eisenberg, D. (2006). The 3D profile method for identifying fibril-forming segments of proteins. *Proceedings of the National Academy of Sciences of the United States of America* 103, 4074-4078.
- Wang, L., Maji, S.K., Sawaya, M.R., Eisenberg, D., and Riek, R. (2008). Bacterial inclusion bodies contain amyloid-like structure. *PLoS biology* 6, e195.
- Wiltzius, J.J., Sievers, S.A., Sawaya, M.R., Cascio, D., Popov, D., Riekel, C., and Eisenberg, D. (2008). Atomic structure of the cross-beta spine of islet amyloid polypeptide (amylin). *Protein science : a publication of the Protein Society* 17, 1467-1474.

Wojtowicz, W.M., Farzan, M., Joyal, J.L., Carter, K., Babcock, G.J., Israel, D.I., Sodroski, J., and Mirzabekov, T. (2002). Stimulation of enveloped virus infection by beta-amyloid fibrils. *The Journal of biological chemistry* 277, 35019-35024.



## Chapter 3

This Chapter is a reprint of:

Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation

Sievers SA\*, Karanicolas J\*, Chang HW\*, Zhao A\*, Jiang L\*, Zirafi O, Stevens JT, Münch J, Baker D, Eisenberg D. *Nature*. 2011 Jun 15;475(7354):96-100

\*These authors contributed equally to this work.

### Abstract

This chapter discusses our work on developing structure-based amyloid inhibitors against fibril formation. We have applied this approach simultaneously to two amyloid model proteins. Lin Jiang and myself developed inhibitors with non-natural amino acid derivatives against SEVI fibril formation. Stuart Sievers, John Karanicolas and Howard Chang developed D-peptide inhibitors against Tau fibril formation. The SEVI inhibitor described here WW61 is optimized based on a L-peptide inhibitor discussed in chapter 3. We incorporated two non-natural amino acid residues to enhance the inhibition activity. At two fold molar excess, this inhibitor efficiently delayed fibrillation for two days. In a functional assay, we showed that WW61 was able to block the formation of viral-enhancing material and prevent the SEVI mediated enhancement of HIV infection.

## Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation

Stuart A. Sievers<sup>1\*</sup>, John Karanikolas<sup>2,3\*</sup>, Howard W. Chang<sup>1\*</sup>, Anni Zhao<sup>1\*</sup>, Lin Jiang<sup>1\*</sup>, Onofrio Ziraffi<sup>4</sup>, Jason T. Stevens<sup>5</sup>, Jan Münch<sup>1</sup>, David Baker<sup>2</sup> & David Eisenberg<sup>1</sup>

Many globular and natively disordered proteins can convert into amyloid fibrils. These fibrils are associated with numerous pathologies<sup>1</sup> as well as with normal cellular functions<sup>2,3</sup>, and frequently form during protein denaturation<sup>4,5</sup>. Inhibitors of pathological amyloid fibril formation could be useful in the development of therapeutics, provided that the inhibitors were specific enough to avoid interfering with normal processes. Here we show that computer-aided, structure-based design can yield highly specific peptide inhibitors of amyloid formation. Using known atomic structures of segments of amyloid fibrils as templates, we have designed and characterized an all-D-amino-acid inhibitor of the fibril formation of the tau protein associated with Alzheimer's disease, and a non-natural L-amino-acid inhibitor of an amyloid fibril that enhances sexual transmission of human immunodeficiency virus. Our results indicate that peptides from structure-based designs can disrupt the fibril formation of full-length proteins, including those, such as tau protein, that lack fully ordered native structures. Because the inhibiting peptides have been designed on structures of dual- $\beta$ -sheet 'steric zippers', the successful inhibition of amyloid fibril formation strengthens the hypothesis that amyloid spines contain steric zippers.

The finding that dozens of pathologies, including Alzheimer's disease, are associated with amyloid fibrils has stimulated research on fibril inhibition. One approach uses the self-associating property of proteins that form fibrils to poison fibril formation with short peptide segments<sup>6–11</sup>. A second approach is based on screening for molecules that can disrupt fibril formation<sup>12,13</sup>. Here we take a third approach to fibril inhibition: structure-based design of non-natural peptides targeted to block the ends of fibrils. With advanced sampling techniques and by minimizing an appropriate energy function, we identify novel candidate inhibitors computationally from a large peptide space that interact favourably with our template structure. This approach has been made possible by the determination of several dozen fibril-like atomic structures of segments from amyloid-forming proteins<sup>14–16</sup>.

These structures reveal a common motif called a steric zipper, in which a pair of  $\beta$ -sheets is held together by the interdigitation of their side chains<sup>17</sup>. Using as templates the steric-zipper structures formed by segments of two pathological proteins, we have designed inhibitors that cap fibril ends. As we show, the inhibitors greatly slow the fibril formation of the parent proteins of the segments, offering a route to designed chemical interventions and supporting the hypothesis that steric zippers are the principal structural elements of these fibrils.

One of the two fibril-like steric zippers that we have chosen as a target for inhibitor design is the hexapeptide VQIVYK, residues 306–311 of the tau protein, which forms intracellular amyloid fibrils in Alzheimer's disease<sup>18</sup>. This segment has been shown to be important for fibril formation of the full-length protein and itself forms fibrils

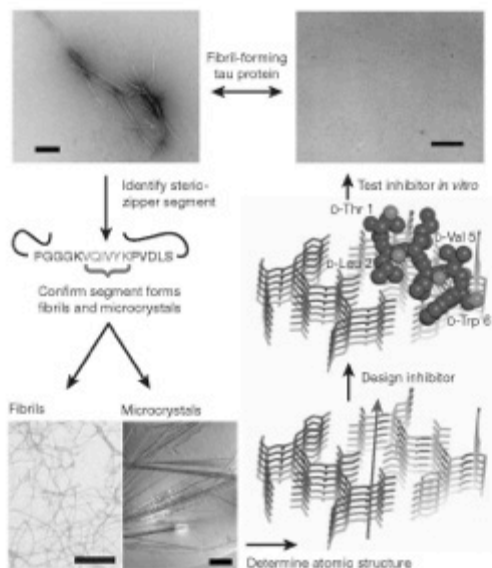
with biophysical properties similar to full-length tau fibrils<sup>19,20</sup>. Our second template for inhibitor design, identified by the '3D profile' algorithm<sup>21,22</sup>, is the steric-zipper structure of the peptide segment GGVLVN from the amyloid fibril formed by <sup>244</sup>PAP<sup>236</sup>, a proteolytic fragment containing residues 248–286 of prostatic acid phosphatase, a protein abundant in semen. <sup>244</sup>PAP<sup>236</sup> fibrils, also known as semen-derived enhancer of virus infection (SEVI), enhance human immunodeficiency virus (HIV) infection by orders of magnitude in cell culture studies, whereas the monomeric peptide is inactive<sup>23</sup>.

Our computational approach to designing non-natural peptides that inhibit fibril formation is summarized in Fig. 1 for the VQIVYK segment of tau protein; the same general strategy is used for the GGVLVN segment of <sup>244</sup>PAP<sup>236</sup>. In both systems, we design a tight interface between the inhibiting peptide and the end of the steric zipper to block additional segments from joining the fibril. By sampling L or D amino acids, or commercially available non-natural amino acids, we can design candidate inhibitors with side chains that maximize hydrogen bonding and hydrophobic interactions across the interface.

We propose that the steric-zipper structures of the VQIVYK and GGVLVN segments represent the spines of the fibrils formed by the parent proteins containing these segments. Supporting our hypothesis are our results that D-amino-acid inhibitors designed on the VQIVYK steric-zipper template inhibit fibril formation not only of the VQIVYK segment, but also of two tau constructs, K12 and K19<sup>21,24</sup> (Fig. 2a). Similarly, the peptide composed of non-natural amino acids designed on the GGVLVN template inhibits the fibril formation of <sup>244</sup>PAP<sup>236</sup> and greatly inhibits the HIV infectivity of human cells in culture.

To design a D-amino-acid hexapeptide sequence that interacts favourably with the VQIVYK steric zipper<sup>17</sup>, and prevents further addition of tau molecules to the fibril, we used the Rosetta software<sup>25</sup>. This led to the identification of four D-amino-acid peptides: D-TLKIVW, D-TWKLVL, D-DYYFEF and D-YVIER, in which the prefix signifies that all  $\alpha$ -carbon atoms are in the D configuration (Fig. 2b, c, Supplementary Figs 1 and 2 and Supplementary Table 1). In the D-TLKIVW design model (Fig. 2b, c and Supplementary Fig. 1), the inhibitor packs tightly across the top of the VQIVYK steric-zipper structure, maintaining all main-chain hydrogen bonds. The side-chain hydrogen bonding between layers of stacked Gln 307 residues is replaced in the designed interface by an interaction with D-Lys 3. Several hydrophobic interactions between D-TLKIVW and the two VQIVYK  $\beta$ -strands contribute to the favourable binding energy (Supplementary Table 1). In the design, the D-peptide blocks the addition of another layer of VQIVYK, both above the D-peptide and across on the mating  $\beta$ -sheet (Supplementary Fig. 3). D-Leu 2 of the designed inhibitor prevents the addition of a VQIVYK molecule above it through a steric clash with Ile 308 of VQIVYK and on the mating sheet through a clash with Val 306 and Ile 308 (Supplementary Fig. 3). These steric clashes involving D-Leu 2 are intended to block fibril growth.

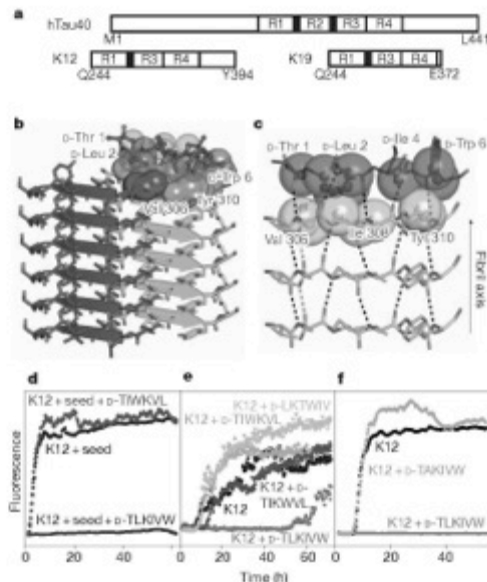
<sup>1</sup>Departments of Biological Chemistry and Chemistry and Biochemistry, Howard Hughes Medical Institute, UCLA, Box 951970, Los Angeles, California 90095-1570, USA, <sup>2</sup>Department of Biochemistry and Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, USA, <sup>3</sup>Center for Bioinformatics and Department of Molecular Biosciences, University of Kansas, 1220 Sornoykide Avenue, Lawrence, Kansas 66045-7534, USA, <sup>4</sup>Institute of Molecular Virology, University Hospital Ulm, Meyerhofstrasse 1, 89081 Ulm, Germany. \*These authors contributed equally to this work.



**Figure 1 | Design and characterization of peptide inhibitors of amyloid fibril formation.** Tau constructs form fibrils *in vitro*<sup>24</sup> (top left scale bar, 200 nm). The VQIVYK segment in isolation forms fibrils and microcrystals (bottom left fibril scale bar, 200 nm; microcrystal scale bar, 100  $\mu$ m). The atomic structure of the fibril-like VQIVYK segment reveals a characteristic sterio-zipper motif<sup>25</sup> comprising a pair of interacting  $\beta$ -sheets (purple and grey) running along the fibril axis (grey arrow) (bottom right). We designed a D-amino-acid peptide to bind to the end of the sterio-zipper template and prevent fibril elongation (middle right). The D-peptide (red) is designed to satisfy hydrogen bonds and make favourable non-polar interactions with the molecule below, while preventing the addition of other molecules above and on the opposite  $\beta$ -sheet. As shown *in vitro*, the designed D-peptide prevents the formation of fibrils when incubated with tau K19 (upper right; scale bar, 200 nm).

We used fluorescence spectroscopy and electron microscopy to assess whether the designed D-peptides inhibit the fibril formation of the tau segment VQIVYK and of the tau constructs K12 and K19. Of our designed inhibitors, D-TLKIVW is the most effective (Supplementary Fig. 4). Electron microscopy, performed after three days, verified that incubation with equimolar D-TLKIVW prevents K19 fibril formation, which would otherwise have occurred within the elapsed time (Fig. 1, upper right). D-TLKIVW delays fibril formation of VQIVYK, K12 and K19 even when present in sub-equimolar concentration (Supplementary Fig. 5). A fivefold molar excess of D-TLKIVW delays K12 fibril formation for more than two weeks in some experimental replicates (Supplementary Fig. 5c, d). In tenfold molar excess, D-TLKIVW prevents the fibril formation of K12 for more than 60 hours in the presence of preformed K12 fibril seeds, suggesting that the peptide interacts with fibrils (Fig. 2d). Also, kinetic analysis shows that the fibril elongation rate decreases in the presence of increasing concentrations of inhibitor peptide (Supplementary Fig. 6). The large increase in lag time in unseeded reactions may be due to interactions with small aggregates formed during the process of fibril formation.

To investigate the specificity of the designed inhibitor, we tested scrambled sequence variants of D-TLKIVW that have poor (that is, high) calculated energies and unfavourable packing (Supplementary Table 1). The scrambled peptides D-TIKWVL, D-TIWKVL and D-LKTWIV have little inhibitory effect when present at an equimolar



**Figure 2 | Designed D-peptide delays tau K12 fibril formation in a sequence-specific manner.** a, Tau construct composition<sup>26</sup>. The longest human tau isoform found in the central nervous system, hTau40 (Uniprot ID, P10636-8), contains four microtubule-binding repeats, R1 to R4, whereas K12 and K19 lack R2. The black bars at the amino terminal of R2 and R3 represent the fibrillogenetic segments VQIVYK and VQIVYK, respectively. b, The inhibitor D-TLKIVW (red) is designed to interact with atoms on both  $\beta$ -strands of the VQIVYK sterio-zipper (grey) primarily through hydrophobic packing and hydrogen-bonding interactions. c, The inhibitor interacts with the VQIVYK  $\beta$ -strand below. The transparent spheres show where the two molecules interact favourably. Black and red dashes indicate main-chain and side-chain hydrogen bonds, respectively. Stereo views of b and c are shown in Supplementary Fig. 1. d, The seeded fibril formation of 50  $\mu$ M K12 in the presence and absence of a tenfold molar excess of peptide was monitored by Thioflavin S fluorescence. In the presence of the scrambled peptide D-TIWKVL (dark green) and alone (black), seeded K12 fibril formation occurs with almost no lag time. However, D-TLKIVW prevents fibril formation for days (maroon). e, At equimolar concentrations, D-TLKIVW (red) inhibits the fibril formation of 50  $\mu$ M K12. D-TIKWVL (blue), with only three residues scrambled, shows weak inhibition. However, no inhibition is observed for either D-TIWKVL (green) or D-LKTWIV (cyan). f, The replacement of D-Leu 2, designed to clash with VQIVYK on the opposite sheet, with D-Ala eliminates the inhibition of fibril formation.

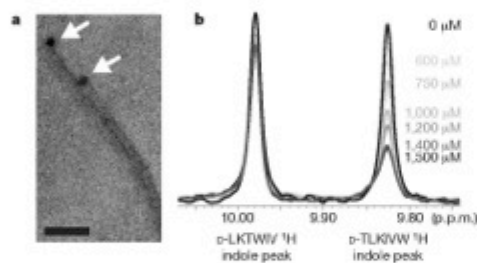
ratio with VQIVYK, K12 and K19 (Fig. 2e and Supplementary Fig. 7), showing that the inhibition is sequence specific. Also, the diastereomer, L-TLKIVW, is less effective than D-TLKIVW (Supplementary Fig. 8). As a further test of the specificity of our design, we confirmed that D-TLKIVW is unable to block the fibril formation of amyloid- $\beta$ , which also is associated with Alzheimer's disease (Supplementary Fig. 9). This suggests that the D-peptide inhibitor is not general to amyloid systems, but is specific to the VQIVYK interface in tau protein. Such specificity is essential for designed inhibitors if they are not to interfere with proteins that natively function in an amyloid state<sup>1</sup>.

To confirm that the designed D-peptide inhibits in accordance with the design model (Fig. 2b, c and Supplementary Fig. 1), we performed several additional tests. First we visualized the position of the inhibitor D-TLKIVW relative to fibrils of the tau construct K19 using electron microscopy. We covalently linked Monomaleimido Nanogold particles

both to the inhibitor and, separately, to a scrambled hexapeptide, D-LKTWIV. We used a blind counting assay and found that, relative to Nanogold alone, D-TLKIVW shows a significant binding preference for the end of fibrils, in contrast to the scrambled control peptide, D-LKTWIV (Fig. 3a and Supplementary Fig. 10).

As a further test of the model, we used NMR to characterize the binding affinity of D-TLKIVW for tau fibrils. The  $^1\text{H}$  NMR spectra for D-TLKIVW were collected in the presence of increasing concentrations of VQIVYK or K19 fibrils. Because neither K19 nor VQIVYK contains tryptophan, we were able to monitor the  $^1\text{H}$  resonance of the indole proton of the tryptophan in our inhibitor. When bound to a fibril, the inhibitor, D-TLKIVW, is removed from the soluble phase and the  $^1\text{H}$  resonance is diminished<sup>36</sup> (Fig. 3b and Supplementary Fig. 11). As a control, we also measured spectra for the non-inhibiting peptide D-LKTWIV present with D-TLKIVW in the same reaction mixture. As shown in Fig. 3b, the presence of VQIVYK fibrils at a given concentration reduces the D-TLKIVW indole resonance much more than it does the D-LKTWIV indole resonance. Spectra of the two peptides are shown in Supplementary Fig. 12. By monitoring the D-TLKIVW indole resonance over a range of VQIVYK fibril concentrations, we estimate the apparent dissociation constant of the interaction between D-TLKIVW and VQIVYK fibrils to be  $\sim 2\ \mu\text{M}$  (Supplementary Fig. 11a and Methods). This value corresponds to a standard free binding energy of  $-7.4\ \text{kcal mol}^{-1}$ , with  $-2.5\ \text{kcal mol}^{-1}$  from non-polar interactions and  $-4.9\ \text{kcal mol}^{-1}$  from six hydrogen bonds (Methods). Repeating the NMR binding experiment with K19 fibrils yields a similar trend (Supplementary Fig. 11b). To determine whether D-TLKIVW has affinity for soluble VQIVYK, we measured  $^1\text{H}$  NMR spectra of D-TLKIVW and D-LKTWIV in the presence of increasing amounts of soluble VQIVYK. Only a slight change in the respective chemical shifts of the indole proton peaks of D-TLKIVW and D-LKTWIV is observed, even at a 70-fold molar excess of VQIVYK (Supplementary Fig. 13). This, together with the ability of the peptide to prevent seeded fibril formation, suggests that D-TLKIVW does not interact with monomers but rather with a structured, fibril-like species.

As another test of our design model, we replaced the D-Leu residue with D-Ala in D-TLKIVW. Our structural model suggests that D-Leu 2 of D-TLKIVW is important for preventing tau fibril formation because of its favourable interaction with the Ile residue of the VQIVYK molecule below and with Ile and the first Val of VQIVYK across the steric zipper (Fig. 2b, c and Supplementary Fig. 1). The D-Ala replacement

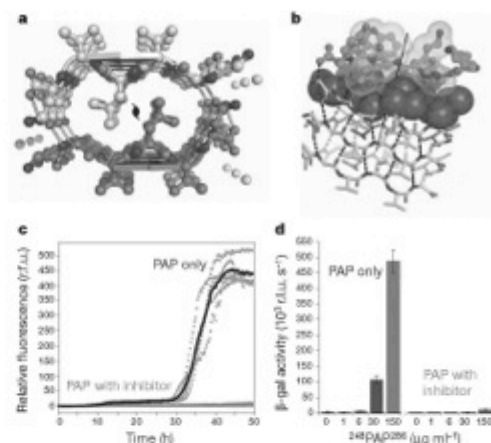


**Figure 3 | Mechanism of interaction.** **a**, Nanogold covalently bound to D-TLKIVW localizes at the ends (arrows) of two tau K19 fibrils. Scale bar, 50 nm. **b**, The inhibitor D-TLKIVW binds to fibrils with an estimated affinity constant in the low micromolar range, as shown by the indole proton region of the 500-MHz  $^1\text{H}$  NMR spectra of D-TLKIVW (9.83 p.p.m.) and D-LKTWIV (9.98 p.p.m.) in the presence of increasing concentrations of VQIVYK fibrils. The resonance of the D-TLKIVW indole proton is reduced in the presence of increasing concentrations of VQIVYK fibrils, whereas the indole proton signal for the scrambled control peptide D-LKTWIV is only slightly affected. Fibril solutions contained 0–1,500  $\mu\text{M}$  VQIVYK monomers, as indicated.

eliminates these interactions and, furthermore, removes a steric clash that would occur were another VQIVYK molecule placed across from the inhibitor (Supplementary Fig. 3 and Supplementary Table 1). When the D-Ala variant is incubated with VQIVYK and the tau constructs, it has no inhibitory effect on fibril formation (Fig. 2f and Supplementary Fig. 14). This confirms that D-Leu 2 is critical for the efficacy of D-TLKIVW, consistent with our model.

In summary, although our electron microscopy, NMR and D-Ala replacement results support a model in which the designed peptide D-TLKIVW binds to the ends of tau fibrils, they do not constitute proof that the inhibitors bind exactly as anticipated in the designs (Supplementary Fig. 15).

To expand on our design methodology, we computationally designed an inhibitor of  $^{248}\text{PAP}^{256}$  fibril formation containing non-natural L-amino acids (Fig. 4b and Supplementary Fig. 16), using the GGVLVN structure as a template (Fig. 4a and Supplementary Table 2). This peptide, Trp-His-Lys-chAla-Trp-hydroxyTic (WW61), contains an Ala derivative,  $\beta$ -cyclohexyl-L-alanine (chAla) and a Tyr/Pro derivative, 7-hydroxy-(S)-1,2,3,4-tetrahydroquinoline-3-carboxylic acid (hydroxyTic), both of which increase contact area with the GGVLVN template. The non-natural chAla forms hydrophobic interactions with the Leu residue in the steric-zipper interface, and hydroxyTic supports the favourable placement of chAla through hydrophobic packing (Fig. 4b and Supplementary Fig. 16b).



**Figure 4 | Designed non-natural peptide inhibits  $^{248}\text{PAP}^{256}$  fibril formation.** **a**, The view down the fibril axis of the crystal structure of the GGVLVN steric zipper reveals two mating  $\beta$ -sheets with parallel, in-register  $\beta$ -strands (hydrogen bonds, green dashed lines; water molecules, yellow spheres). **b**, View roughly perpendicular to a fibril of three layers, with the atoms of the side chains of the top layer shown as purple spheres. On top is a designed non-natural peptide inhibitor, Trp-His-Lys-chAla-Trp-hydroxyTic (blue; see Supplementary Fig. 16). **c**, The inhibitor blocks  $^{248}\text{PAP}^{256}$  fibril formation, as shown by monitoring Thioflavin T fluorescence. With a two-fold molar excess of the inhibitor (pale red), the fluorescence remains low over the course of the experiment for all five replicates, unlike in the absence of inhibitor (grey). Mean fluorescence values are shown as solid red and black lines with and without the inhibitor, respectively. r.f.u., relative fluorescence units. **d**, HIV infection rates were determined by monitoring  $\beta$ -galactosidase ( $\beta$ -gal) activity. Agitated  $^{248}\text{PAP}^{256}$  alone efficiently increases viral infection, whereas  $^{248}\text{PAP}^{256}$  mixtures incubated with inhibitor were unable to enhance HIV infection. Peptide concentrations during virion treatment are indicated on the x axis. Error bars show the s.d. of three measurements per sample. r.l.u., relative light units.

Moreover, we propose that the bulky side chains and steric constraints of hydroxyTic provide hindrance to further fibril growth.

This designed peptide, WW61, effectively delays both seeded and unseeded fibril formation of  $^{248}\text{PAP}^{286}$  *in vitro* (Fig. 4c and Supplementary Figs 17 and 18). In the presence of a twofold molar excess of this inhibitor, seeded fibril formation is efficiently blocked for more than two days (Fig. 4c). Furthermore, we see that increasing the concentration of this inhibitor extends the fibril formation lag time (Supplementary Fig. 19). These inhibition assay results were further confirmed by electron microscopy (Supplementary Fig. 20). As a control for specificity, we tested the effect of GHKQK, from the amino terminus of  $^{248}\text{PAP}^{286}$ , and PYKLWN, a peptide with the same charge as WW61. Neither peptide affected fibril formation kinetics, indicating that the inhibitory activity of the designed peptide is sequence specific (Supplementary Fig. 21).

Because  $^{248}\text{PAP}^{286}$  fibrils (SEVI) have been shown to enhance HIV infection<sup>22</sup>, using a functional assay we investigated whether WW61 is able to prevent this enhancement. In this experiment, we treated HIV particles with  $^{248}\text{PAP}^{286}$  solutions that had been agitated for 20 hours (to allow fibril formation) in the presence or absence of WW61, and infected TZM-bl indicator cells. As has been previously observed, SEVI efficiently enhanced HIV infection<sup>22</sup>. However,  $^{248}\text{PAP}^{286}$  incubated with the designed inhibitor prevented HIV infection (Fig. 4d).

We performed several control experiments to verify that the lack of infectivity observed in the assay is indeed due to the inhibition of SEVI formation. First we confirmed that in the absence of SEVI the designed inhibitor WW61 does not affect HIV infectivity (Supplementary Fig. 22a). We also found that the control peptides GHKQK and PYKLWN, which do not inhibit  $^{248}\text{PAP}^{286}$  fibril formation, fail to decrease HIV infectivity (Supplementary Fig. 22b). Additionally, we observed that WW61 has no inhibitory effect on polylysine-mediated HIV infectivity<sup>22</sup>, further ruling out a non-specific electrostatic interaction mechanism (Supplementary Fig. 22a). Together, these results demonstrate that a peptide capable of preventing  $^{248}\text{PAP}^{286}$  fibril formation also inhibits the generation of virus-enhancing material.

Structure-based design of inhibitors of amyloid fibril formation has been challenging in the absence of detailed information about the atomic-level interactions that form the fibril spine. So far, one of the most successful structure-based approaches to preventing fibril formation has been to stabilize the native tetrameric structure of transthyretin<sup>23</sup>. That approach is well suited to the prevention of fibril formation of proteins with known native structures, but other proteins involved in amyloid-related diseases, such as tau protein, amyloid- $\beta$  and  $^{248}\text{PAP}^{286}$ , lack fully ordered native structures<sup>24</sup>. Our structure-based approach makes it possible to design inhibitors independent of native structure. Instead, the templates are atomic-level structures of short, fibril-forming segments<sup>24,25</sup>. By using these fibril-like templates, and adopting computational methods successful in designing novel proteins and protein-protein interfaces<sup>26,27</sup>, we have created specific inhibitors of proteins that normally form fibrils. These results support the hypothesis that the steric zipper is a principal feature of tau-related and SEVI fibrils, and suggest that, with current computational methods and steric-zipper structures, we have the tools to design specific inhibitors to prevent the formation of other amyloid fibrils.

## METHODS SUMMARY

We used crystal structures of hexapeptide segments of VQIVYK and GGVLVN as templates to design peptide inhibitors using the Rosetta software<sup>28</sup>. Briefly, this algorithm searches possible side-chain conformations (called rotamers) of all amino acids in a peptide  $\beta$ -strand backbone stacked onto the fibril end of both segment structures. The Rosetta software is extended to sample the approximate side-chain conformation of non-natural D and L amino acids by adapting side-chain torsion angles from those in their natural counterparts. The lowest energy set of side-chain rotamers is identified by combinatorial optimization of a potential consisting of a term for the Lennard-Jones potential, an orientation-dependent hydrogen-bond potential term, an implicit solvation term and a structure-derived side-chain and backbone torsional potential term.

Full Methods and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 6 December 2010; accepted 21 April 2011.

Published online 15 June 2011.

1. Westermark, P. et al. A primer of amyloid nomenclature. *Amyloid* **14**, 179–183 (2007).
2. Maj, S. K. et al. Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **325**, 328–332 (2009).
3. Fowler, D. M., Koulou, A. V., Salch, W. E. & Kelly, J. W. Functional amyloid – from bacteria to humans. *Trends Biochem. Sci.* **32**, 217–224 (2007).
4. Astbury, W. T. & Dickinson, S. The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochem. J.* **29**, 2351–2360 (1935).
5. Calamia, M., Chit, F. & Dobson, C. M. Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophys. J.* **89**, 4201–4210 (2005).
6. Tjernberg, L. O. et al. Arrest of  $\beta$ -amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.* **271**, 8545–8548 (1996).
7. Firdes, M. A. Peptide inhibitors of  $\beta$ -amyloid aggregation. *Curr. Top. Med. Chem.* **2**, 417–423 (2002).
8. Sciametta, K. L., Gordon, D. J. & Meredith, S. C. Peptide-based inhibitors of amyloid assembly. *Methods Enzymol.* **413**, 273–312 (2006).
9. Soto, C., Kindy, M. S., Baumann, M. & Frangione, B. Inhibition of Alzheimer's amyloidosis by peptides that prevent  $\beta$ -sheet conformation. *Biochem. Biophys. Res. Commun.* **226**, 672–680 (1996).
10. Kekkonen, N., Stott, K., Amey, H., Mason, J. M. & Doig, A. J. N-methylated peptide inhibitors of amyloid aggregation and toxicity. Optimization of the inhibitor structure. *Biochemistry* **45**, 9906–9918 (2006).
11. Sato, T. et al. Inhibitors of amyloid toxicity based on  $\beta$ -sheet packing of A $\beta$ 40 and A $\beta$ 42. *Biochemistry* **45**, 5503–5516 (2006).
12. Larbig, G., Pichardt, M., Lloyd, D. G., Schmidt, B. & Mandelkow, E. Screening for inhibitors of tau protein aggregation into Alzheimer paired helical filaments: a ligand based approach results in successful scaffold hopping. *Curr. Alzheimer Res.* **4**, 315–323 (2007).
13. Wesschan, K. et al. Selection of D-amino-acid peptides that bind to Alzheimer's disease amyloid peptide A $\beta$ <sub>1–42</sub> by mirror image phage display. *ChemBioChem* **4**, 748–753 (2003).
14. Nelson, R. et al. Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
15. Sawaya, M. R. et al. Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
16. Wittzus, J. J. et al. Molecular mechanisms for protein-encoded inheritance. *Nature Struct. Mol. Biol.* **16**, 973–978 (2009).
17. Selkoe, D. J. Alzheimer's disease: genes, proteins, and therapy. *Physiol. Rev.* **81**, 741–766 (2001).
18. Gour, W. J. et al. The formation of straight and twisted filaments from short tau peptides. *J. Biol. Chem.* **279**, 26868–26875 (2004).
19. von Bergen, M. et al. Assembly of a protein into Alzheimer paired helical filaments depends on a local sequence motif ( $^{100}\text{VQIVYK}^{111}$ ) forming  $\beta$ -structure. *Proc. Natl Acad. Sci. USA* **97**, 5129–5134 (2000).
20. Goldschmidt, L., Teng, P. K., Riek, R. & Eisenberg, D. Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proc. Natl Acad. Sci. USA* **107**, 3487–3492 (2010).
21. Thompson, M. J. et al. The 3D profile method for identifying fibril-forming segments of proteins. *Proc. Natl Acad. Sci. USA* **103**, 4074–4078 (2006).
22. Münch, J. et al. Semen-derived amyloid fibrils drastically enhance HIV infection. *Cell* **131**, 1059–1071 (2007).
23. Friedhoff, P., von Bergen, M., Mandelkow, E. M., Davies, P. & Mandelkow, E. A nucleated assembly mechanism of Alzheimer paired helical filaments. *Proc. Natl Acad. Sci. USA* **95**, 15712–15717 (1998).
24. Wille, H., Drewes, G., Biemart, J., Mandelkow, E. M. & Mandelkow, E. Alzheimer-like paired helical filaments and antiparallel dimers formed from microtubule-associated protein tau *in vitro*. *J. Cell Biol.* **118**, 573–584 (1992).
25. Kuhlman, B. et al. Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364–1368 (2003).
26. Chen, Z., Krause, G. & Rief, B. Structure and orientation of peptide inhibitors bound to  $\beta$ -amyloid fibrils. *J. Mol. Biol.* **354**, 760–776 (2005).
27. Roan, N. R. et al. The cationic properties of SEVI underlie its ability to enhance human immunodeficiency virus infection. *J. Virol.* **83**, 73–80 (2009).
28. Petraas, H. M., Klabunde, T., Sacchettini, J. & Kelly, J. W. Structure-based design of N-phenyl piperazine transthyretin amyloid fibril inhibitors. *J. Am. Chem. Soc.* **122**, 2178–2192 (2000).
29. Schweers, O., Schonbrunn-Hanebeck, E., Marx, A. & Mandelkow, E. Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for  $\beta$ -structure. *J. Biol. Chem.* **269**, 24290–24297 (1994).
30. Fleishman, S. J. et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816–821 (2011).

Supplementary information is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. I. Ivanova, J. Corn, T. Korferme, D. Anderson, M. S. Sawaya, M. Phillips, S. Sambasivan, J. Park, M. Landau, A. Laginowsky, Q. Zhang, R. Cutts, F. Guo, T. Yeates, J. Nowick, J. Zheng and M. J. Thompson for discussions; the HHMI, NIH, NSF, Gates Foundation and Joint Center for Translational Medicine for support; R. Peterson for help with NMR experiments; E. Mandelkow for providing tau constructs; R. Riek for providing amyloid- $\beta$ ; and J. Stroud for amyloid- $\beta$  preparation.

**RESEARCH LETTER**

Support came from the Damon Runyon Cancer Research Foundation (J.K.), the Ruth L. Kirschstein National Research Service Award (H.W.C.), the programme for junior professors by the Ministry of Science, Baden-Württemberg (J.M.), and a UCLA-XBIRT bioinformatics traineeship (S.A.S.).

**Author Contributions** S.A.S., J.K., D.B., J.M. and D.E. designed the project. J.K. and S.A.S. created the design protocol. J.K. designed the  $\alpha$ -peptides. L.J. expanded the design methodology and designed the non-natural amino-acid peptides. S.A.S., H.W.C. and A.Z. performed the fluorescence experiments and electron microscopy, and analysed kinetic data. A.Z. determined the structure of GGVLVN. O.Z. performed the HIV

infectivity experiments. J.T.S. determined the tau fibril elongation rates. S.A.S. performed the NMR experiments. S.A.S., J.K. and D.E. wrote the manuscript and coordinated contributions by other authors.

**Author Information** Atomic coordinates and structure factors for the reported GGVLVN structure have been deposited in the Protein Data Bank with accession code 3PPD. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.E. ([david@mbi.ucla.edu](mailto:david@mbi.ucla.edu)).

## METHODS

**Computational design.** Computational designs were carried out using the Rosetta software<sup>37</sup> (<http://www.rosettacommons.org>). This algorithm involves building side-chain rotamers of all amino acids onto a fixed protein backbone. The lowest energy set of side-chain rotamers is then identified as those which minimize an energy function containing a Lennard-Jones potential, an orientation-dependent hydrogen-bond potential, a solvation term, amino-acid-dependent reference energies and a statistical torsional potential that depends on the backbone and side-chain dihedral angles.

**D-amino-acid tau inhibitors.** The crystal structure of VQIVYK (ref. 15; Protein Data Bank ID, 2ON9) was used as a starting scaffold for computational design. To take full advantage of the statistical nature of the rotamer library and some terms in the Rosetta energy function, the stereochemistry of the fibril scaffold was inverted so that design would take place using L amino acids. An extended L-peptide was aligned with the N, C and O backbone atoms of the D-fibril scaffold. This L-peptide was subsequently redesigned, keeping all atoms of the D-fibril fixed. The stereochemistry of the final design model was then inverted, yielding a D-peptide designed to cap an L-fibril. We inspected the finished models to confirm that inversion of the stereochemistry at the Thr and Ile C $\beta$  atoms did not make the design energetically unfavourable. Energetic consequences of incorporating a D inhibitor peptide in the middle of an L fibril were subsequently evaluated to ensure that fibril propagation could not continue after association of an inhibitor. Calculations of the area buried and shape complementarity were performed with AREAMOL<sup>38</sup> and SC<sup>39</sup>, respectively.

**L-peptide <sup>240</sup>PAP<sup>350</sup> inhibitors.** The crystal structure of GGVLNV (PDB ID, 3PPD) was used as a template for the following design procedure. An extended L-peptide was aligned according to crystal symmetry. Small, random perturbations of the L-peptide were performed to optimize the rigid-body arrangement between the fibril template and the peptide inhibitor. Full sequence optimization of the inhibitor was performed using the Rosetta software package, allowing residues directly contacting the inhibitor to repack; other scaffold residues remained fixed. Because the design calculations use a discrete rotamer representation of the side chains, we next performed simultaneous quasi-Newtonian optimization of the inhibitor rigid-body orientation, the side-chain torsion angles and, in some cases, the backbone torsion angles using the full-atom Rosetta energy function. This optimization was essential to the subsequent assessment of the inhibition of the design. Several iterative runs of small perturbations in inhibitor placement, interface design and refinement were performed to improve hydrogen-bonding and packing interactions. The designs that ranked highest on the basis of the total binding energy between the inhibitor and the fibril scaffold and the interfacial shape complementarity<sup>35</sup> were subsequently synthesized and tested.

For each initial active L-peptide design, the non-natural L amino acids were incorporated using a growth strategy. Non-natural amino acids, structurally similar to those of initial active designs, were selected on the basis of their solubility, side-chain shape and commercial availability. Side-chain conformations were approximately sampled by adopting side-chain torsion angles from those in their natural counterparts. Sequence optimization of the inhibitor was performed and the optimal set of rotamers identified using Monte Carlo simulated annealing with the full-atom energy function described above. The resulting designs were ranked on the basis of the total binding energy between the inhibitor and the fibril scaffold.

**Tau construct expression and purification.** pNG2 expression vectors (derived from pET-3b<sup>33</sup>) containing either the K12 or K19 gene were provided by E. Mandelkow<sup>24</sup>. Expression in BL21(DE3) *Escherichia coli*<sup>25</sup> was induced with 1 mM isopropyl thioallostoside when the absorbance  $A_{600nm}$  was between 0.8 and 1.0, and cells were collected after 3–4 h. K12 and K19 were purified on the basis of previously described methods<sup>25</sup>. Cells were pelleted for 20 min at 4,700g and resuspended in 20 mM MES, pH 6.8, 1 mM EDTA, 0.2 mM MgCl<sub>2</sub>, 5 mM DTT, 1 mM PMSF and a protease inhibitor cocktail. The cells were sonicated for 2.5 min and, following addition of NaCl to bring cell lysate to 0.5 M NaCl, the lysate was boiled for 20 min. The lysate was sedimented at 30,000g for 20 min and dialyzed twice against 20 mM MES, pH 6.8, 50 mM NaCl, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, 2 mM DTT and 0.1 mM PMSF at 4 °C. The dialysate was pelleted for 20 min at 30,000g and filtered before cation exchange chromatography on an AKTA Explorer (GE Pharmacia) with a HighTrap HP SP 5-ml column (GE Healthcare). The sample was eluted with a linear gradient of up to 60% buffer B (20 mM MES, pH 6.8, 1 M NaCl, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, 2 mM DTT and 0.1 mM PMSF). Size exclusion chromatography was optionally performed with a Superdex 75 10/300 GL column (GE Healthcare) in PBS buffer (137 mM NaCl, 3 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) with 1 mM DTT on the AKTA Explorer depending on preparation purity as assessed by SDS polyacrylamide gel electrophoresis.

**Tau construct inhibition assays.** Fibril formation assays were performed on the basis of previously published protocols<sup>25, 36</sup>. Reaction mixtures (150  $\mu$ l) containing

50  $\mu$ M tau K12 or K19, as determined by the Micro BCA Protein Assay Kit (Pierce), were incubated in 250 mM sodium phosphate buffer, pH 7.4, with 1 mM DTT, 12.5  $\mu$ M heparin (average molecular mass, 6,000 Da; Sigma) and 10  $\mu$ M Thioflavin S (ThS; MP Bio). Inhibitor peptides (CS Bio, Celtek Biosciences) were dissolved in 250 mM phosphate buffer, pH 7.4, to 0.5 mM and added at specified molar ratios. Reactions were split into a minimum of three replicates in black, 96-well, optically clear plates (Nunc), sealed with Corning pressure-sensitive sealing tape and monitored using either a Varioskan plate reader (Thermo Scientific), for K12, or a SpectraMax M5, for K19. The fluorescence signal was measured every 15 min with excitation and emission wavelengths of 440 and 510 nm, respectively, at 37 °C, with continuous shaking at 900 r.p.m. with a diameter of 1 mm for K12, and with quiescent incubation with shaking 2 s before each reading for K19. Plots showing the fluorescence trace of the replicate with median lag time for each sample were created using R<sup>40</sup>. Plots of lag time depict the mean time value at which each replicate crossed an arbitrary fluorescence value above noise background (values were selected per experiment and applied to all samples). Error bars represent the standard deviation of the replicate lag times for each sample.

**Seeded K12 fibril formation assays.** Seeds were produced by incubating 50  $\mu$ M K12 as above, but without ThS present, and were added at 0.25% (v/v). Peptide stock concentrations were 0.75 mM and were added at a final concentration of tenfold molar excess relative to soluble K12. Reaction mixtures were otherwise prepared and monitored as above.

**VQIVYK inhibition assays.** The VQIVYK fibril formation assay was modified from a previously published protocol<sup>25</sup>. Buffers and plates were kept on ice to delay VQIVYK fibril formation while the reaction mixtures were prepared. Replicate solutions of 180  $\mu$ l of 25 mM MOPS, pH 7.2, 100  $\mu$ M ThS and inhibitor peptides were added to black, clear-bottomed, 96-well Nunc plates with 1/8-inch PTFE beads (Orange Products). Acetylated and amidated VQIVYK (GenScript) was dissolved in H<sub>2</sub>O to 1.3 mM and filtered through a Millipore Microcon 100-kDa filter device at 14,000g for 5 min at 4 °C to remove large aggregates (final concentration, ~1 mM). Filtered VQIVYK (20  $\mu$ l) was added to each reaction well. ThS fluorescence was monitored at room temperature every 2 min using a SpectraMax M5 fluorometer with 2 s of mixing before each reading.

**Amyloid- $\beta$  fibril formation assay.** Lyophilized amyloid- $\beta$ (1–42) was diluted to 0.2 mg ml<sup>-1</sup> in 50 mM NH<sub>4</sub>OH and filtered with a 0.2- $\mu$ m filter. The reaction mixture contained a final concentration of 11.5  $\mu$ M amyloid- $\beta$ (1–42), 10  $\mu$ M Thioflavin T (ThT), 23 mM NH<sub>4</sub>OH in 100 mM bicine, pH 9.1, and 11.5  $\mu$ M D-TLKIVW in reactions with peptide present. Reactions were split into four replicates and the ThT fluorescence signal was measured every minute (excitation wavelength, 440 nm; emission wavelength, 510 nm), at 37 °C, with continuous shaking at 960 r.p.m. with a 1-mm diameter in a Varioskan fluorometer.

**Electron microscopy.** Sample (5  $\mu$ l) was applied to glow-discharged, 400-mesh carbon-coated, formvar films on copper grids (Ted Pella) for 3 min. Grids were rinsed twice with distilled water and stained with 1% uranyl acetate for 90 s. Grids were examined in a Hitachi H-7000 transmission electron microscope at 75 keV or a JEOL JEM1200-EX operating at 80 keV.

**Tau fibril formation kinetic analysis.** The nucleation ( $k_1$ ) and propagation ( $k_2$ ) rates were determined by fitting the form of the Finke-Watzky two-step mechanism<sup>41</sup>. Plateau values were determined and the remaining parameters were fitted using the 'leasqr' nonlinear least-squares regression function (<http://fly.ist.cnr.it/pub/software/octave/leasqr/>) through the OCTAVE software package (<http://www.gnu.org/software/octave/>).

**Preparation of peptide-gold conjugates.** Peptide-Nanogold conjugates were prepared as described earlier for similarly sized peptides<sup>42</sup>. Briefly, 60 nmol of the peptides CGGG-(D)-TLKIVW and CGGG-(D)-LKTWV (CS Bio) were dissolved in 110  $\mu$ l of phosphate-buffered saline (20 mM, pH 6.5, 0.15 M NaCl), added to 6 nmol of Monomaleimido Nanogold (Nanoprobes), dissolved in 200  $\mu$ l H<sub>2</sub>O and incubated for 1 h at room temperature (22 °C) with constant rotation. Peptide-Nanogold conjugates were separated from excess unbound peptides by membrane centrifugation (Microcon-10 system, Amicon) using a molecular mass cut-off of 10 kDa. Peptide-Nanogold conjugates were then diluted into phosphate-buffered saline, aliquoted and stored at -20 °C for no longer than one month.

**Preparation of K19 fibrils.** K19 fibrils were generated by incubating 100  $\mu$ M soluble K19 with 25  $\mu$ M 6-kDa heparin overnight at 37 °C in phosphate buffer (50 mM, pH 7.4). K19 fibrils were sonicated for 15 s, using a microtip set to 35% amplitude. Residual heparin and small oligomers were removed by centrifuging the mixture through a 100-kDa Microcon concentrator for 10 min at 14,000g, washing the retentate with phosphate buffer and repeating three times; the retentate was restored to its original volume with phosphate buffer. These short fibril segments were stored at 4 °C for no longer than one week. For NMR studies, fibril samples were similarly prepared, but were washed in H<sub>2</sub>O and concentrated to 2 mM K19 (by monomer).

**Preparation of samples for Nanogold binding experiments.** Nanogold conjugated inhibitor (or control) (10 nM) was incubated with 1.67  $\mu\text{M}$  K19 fibrils (by monomer) in MOPS buffer (25 mM, pH 7.2) for 1 h. We applied 5  $\mu\text{l}$  of it to a glow-discharged, 400-mesh carbon-stabilized copper grid (Ted Pella) for 3 min. The grids were washed twice with  $\text{H}_2\text{O}$  and 10  $\mu\text{l}$  of the Goldenhance reagent was applied for 10 s. The grids were washed five times with  $\text{H}_2\text{O}$  and negatively stained with 2% uranyl acetate.

**Quantification and localization of Nanogold binding.** For each sample, 75 Nanogold particles  $\leq 15$  nm in diameter were counted and classified as bound or unbound. The 15-nm cut-off was chosen to exclude unbound, but adjacent, particles enlarged by Goldenhance that only apparently bind fibrils. To establish the localization of the binding observed, individual Nanogold particles bound to fibrils were categorized as bound to the fibril end or side. In both of these experiments, sample identities were concealed from the microscopist to ensure unbiased counting. Grids were examined with a JEOL JEM1200-EX and images were recorded using DIGITALMICROGRAPH (Gatan).

**Statistical analysis of Nanogold binding.** We compared counts of Nanogold-conjugated peptides and unconjugated Nanogold bound to fibrils or localizing to fibril ends. Twenty-one unconjugated Nanogold particles out of 75 counted bound to fibrils. We modelled Nanogold particles bound to fibrils using a binomial distribution with parameters  $n = 75$  (sample size: number of observations) and  $P = 0.28$  (probability of success). In a separate experiment, 22 unconjugated Nanogold particles bound to fibrils that localized to fibril ends, following a binomial distribution with  $n = 105$  and  $P = 0.21$ .

Because the number of counts is fairly large, we assumed a normal distribution and used a standard Z-test to compare the number of bound Nanogold-peptide conjugates with the expected distribution based on the number of bound, unconjugated Nanogold particles. We used an analogous analysis to determine the significance of localization to fibril ends.

The numbers of Nanogold-d-TLKIVW conjugates bound to fibrils ( $n_{\text{bound}} = 43$ ,  $n = 75$ ) and bound Nanogold-d-TLKIVW conjugates localizing to the end of fibrils ( $n_{\text{end}} = 49$ ,  $n = 86$ ) were significantly different from the corresponding numbers for Nanogold alone, whereas the number of Nanogold-d-LKTWIV conjugates bound ( $n_{\text{bound}} = 15$ ,  $n = 75$ ) or the number localized to fibril ends ( $n_{\text{end}} = 17$ ,  $n = 100$ ) did not differ significantly from the corresponding numbers for Nanogold alone.

**VQIVYK preparation for binding studies.** Acetylated and amidated VQIVYK peptide (Genscript) was dissolved to 1 mM in 25 mM MOPS, pH 7.2, and incubated at room temperature for at least 24 h. Fibrils were washed with  $\text{H}_2\text{O}$ , concentrated using an Amicon ultracentrifugal filter with a 3-kDa molecular mass cut-off and resuspended in  $\text{H}_2\text{O}$  to a final concentration (by monomer) of 4 mM. Soluble VQIVYK was prepared by dissolving VQIVYK peptide (CS Bio) with free amino and carboxy termini in  $\text{H}_2\text{O}$ .

**$^1\text{H}$  NMR sample preparation and measurements.** NMR samples were prepared with 5%  $\text{D}_2\text{O}$  and 10 mM NaOAc, pH 5.0. D-peptides were added from 1 mM stocks in  $\text{H}_2\text{O}$  to a final concentration of 100  $\mu\text{M}$ . Soluble and fibrillar VQIVYK and tau protein were added at indicated concentrations to make a final volume of 550  $\mu\text{l}$ .  $^1\text{H}$  NMR spectra measured at 500 MHz were collected on a Bruker DRX500 at 283 K.  $\text{H}_2\text{O}$  resonance was suppressed through presaturation. Spectra were processed with XWINNMR 3.6.

**Binding constant estimations.** NMR data were analysed to estimate a binding constant for the interaction between d-TLKIVW and VQIVYK fibrils. At about 1,000  $\mu\text{M}$  VQIVYK (concentration as monomer), 50% of d-TLKIVW is bound (Supplementary Fig. 11). The steric-zipper model suggests that there are two monomers per 4.7 Å (0.47 nm) layer in a fibril<sup>14</sup>—such that the number of monomers per fibril is given by [fibril length (nm)]  $\times$  [2 monomers per 0.47 nm]—and we estimate the fibril concentration using the monomer concentration:  $[\text{VQIVYK}_{\text{bound}}] = [\text{VQIVYK}_{\text{monomer}}]/[\text{monomers per fibril}]$ . If we assume one binding site and estimate from electron microscopy an average length of ~140 nm per fibril, then there are about 600 monomers per fibril, and the apparent dissociation constant is about 2  $\mu\text{M}$ .

**Hydrogen-bonding energy calculation.** We used AREAIMOL<sup>31</sup> to calculate the non-polar and polar areas buried by the interaction between d-TLKIVW with the VQIVYK steric zipper (Fig. 2b, c and Supplementary Fig. 1). We calculate buried areas of 201, 24 and 102 Å<sup>2</sup> for carbon, nitrogen and oxygen atoms, respectively. Using the atomic solvation parameters of ref. 43, we estimate that the free energy of transferring the inhibitor from a non-polar phase to an aqueous phase,  $\Delta G_{\text{desolvation}}$  is approximately 2.5 kcal mol<sup>-1</sup>. On the basis of an apparent dissociation constant of 2  $\mu\text{M}$ , we estimate the total free energy change of bringing the inhibitor into contact with the VQIVYK steric-zipper template,  $\Delta G_{\text{binding}}$  to be 7.4 kcal mol<sup>-1</sup>. From the interaction model (Fig. 2c and Supplementary Fig. 1), we maintain six hydrogen bonds between d-TLKIVW and VQIVYK, and estimate the free energy change per hydrogen bond to be  $(\Delta G_{\text{binding}} - \Delta G_{\text{desolvation}})/6$ , or  $-0.8$  kcal mol<sup>-1</sup>.

**GGVLVN crystallization and structure determination.** The GGVLVN peptide was dissolved in 10 mM Tris, pH 9, at 1.8 mg ml<sup>-1</sup> and crystallized in 10% (w/v) PEG-8000, 0.1 M MES, pH 6.0, and 0.2 M Zn(OAc)<sub>2</sub>. X-ray diffraction data was collected at APS beamline 24-ID-E. Phases were determined by molecular replacement using an idealized  $\beta$ -strand in PHASER<sup>32</sup>. Crystallographic refinement was performed using REFMAC<sup>33</sup>. Model building was performed with COOT<sup>34</sup> and illustrated with PYMOL<sup>35</sup>.

**$^{249}\text{PAP}^{249}$  fibril formation and inhibition.** Fmoc- $\beta$ -cyclohexyl-L-alanine and Fmoc-7-hydroxy-(S)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid were purchased from AnaSpec and the inhibitor peptide Trp-His-Lys-chAla-Tryp-hydroxyTic (WW61) was synthesized by Celtek Biosciences.  $^{249}\text{PAP}^{249}$  and WW61 were dissolved as  $\times 1.25$  and  $\times 5$  stocks in PBS, respectively, and filtered with a 0.1- $\mu\text{m}$  filter.  $^{249}\text{PAP}^{249}$  was diluted with PBS to 0.66 mM and ThT was added to 10  $\mu\text{M}$  final concentration. Samples were optionally mixed with 1.32 mM WW61 and vortexed. Five replicates of 150  $\mu\text{l}$  were immediately dispensed into a 96-well plate. In dose-response experiments, WW61 final concentrations were 0.33, 0.66 and 1.32 mM. Plates were continuously agitated at 960 r.p.m. at 37 °C, and ThT fluorescence readings were recorded (excitation wavelength, 440 nm; emission wavelength, 482 nm) at 15-min intervals with a Varioskan Flash fluorometer. Lag time was determined when fluorescence crossed an arbitrary value (3 r.f.u.) above background.

**Effect of WW61 on fibril-mediated enhancement of HIV-1 infection.** The CCR5 tropic molecular HIV-1 clone NL4\_3/92TH014-2<sup>36</sup> was generated by transient transfection of 293T cells with proviral DNA. Supernatants were collected 48 h later and p24 concentrations determined by ELISA. T2M-bl reporter cells encoding a lucZ gene under the control of the viral LTR promoter were obtained through the NIH AIDS Research and Reference Reagent Program and provided by Dr John C. Kappes, Dr Xiaoyun Wu and Tranzyme<sup>37</sup>. HIV-1 (40  $\mu\text{l}$ ) containing 0.1 ng of p24 antigen was incubated with 40- $\mu\text{l}$  dilutions of mixtures of  $^{249}\text{PAP}^{249}$  and inhibitory peptide, WW61, that was either freshly prepared or had been agitated for 23 h. Peptide concentrations and experimental conditions during agitation were similar to those described above. Thereafter, 20  $\mu\text{l}$  of the mixtures were used to infect 180  $\mu\text{l}$  of T2M-bl cells seeded the day before (10<sup>6</sup> per well). Two days later, infection rates were determined by quantifying  $\beta$ -galactosidase activities in cellular lysates using the Gal-Screen assay (Applied Biosystems, T1027). Luminescence was recorded on an Orion microplate luminometer as relative light units per second.

**Effect of WW61 on polylysine-mediated enhancement of HIV-1 infection.** Polylysine (Sigma Aldrich) (50  $\mu\text{l}$ ) was mixed with an equal volume of WW61. Thereafter, 35- $\mu\text{l}$  fivefold dilutions of the polylysine-WW61 mixture or polylysine alone were incubated with the same volume of virus and incubated for 5 min at room temperature. Polylysine-WW61 concentrations were 100, 20, 4, 0.8, 0.16, 0.032, 0.064 and 0  $\mu\text{g ml}^{-1}$  during pre-incubation with virus stocks. Thereafter, 20  $\mu\text{l}$  of each mixture was added to 180  $\mu\text{l}$  of T2M-bl cells. The infection rate was determined two days later as described above.

**Effect of WW61, GIHKQK and PYKLWN on HIV-1 infection.** Each peptide (40  $\mu\text{l}$ ) was incubated with an equal volume of virus containing 1 ng of p24 antigen for 5 min at room temperature. Peptide concentrations were 150, 30, 6, 1.2 and 0  $\mu\text{g ml}^{-1}$  during pre-incubation with virus stocks. Thereafter, 20  $\mu\text{l}$  of each mixture was added separately to 180  $\mu\text{l}$  of T2M-bl cells (tenfold dilution) and the infection rate was determined as above.

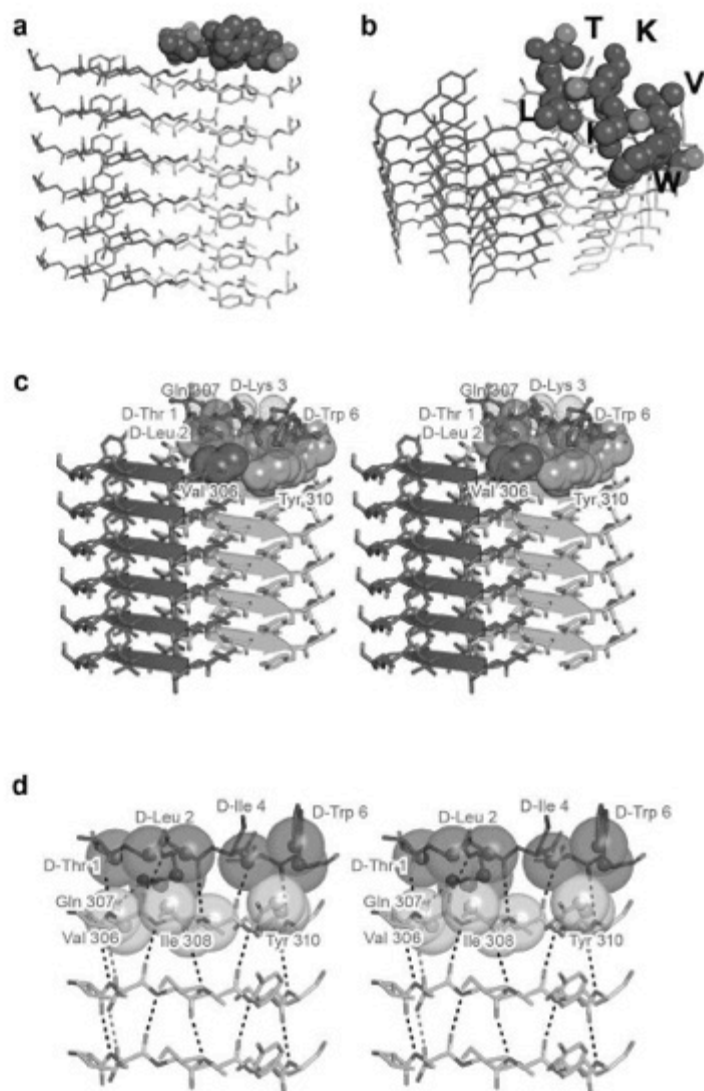
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Lawrence, M. C. & Colman, P. M. Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950 (1993).
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89 (1990).
- Biernat, J. et al. The switch of tau protein to an Alzheimer-like state includes the phosphorylation of two serine-proline motifs upstream of the microtubule binding region. *EMBO J.* **11**, 1593–1597 (1992).
- Barghorn, S., Biernat, J. & Mandelkow, E. Purification of recombinant tau protein and preparation of Alzheimer-paired helical filaments in vitro. *Methods Mol. Biol.* **299**, 35–51 (2005).
- Friedhoff, P., Schneider, A., Mandelkow, E. M. & Mandelkow, E. Rapid assembly of Alzheimer-like paired helical filaments from microtubule-associated protein tau monitored by fluorescence in solution. *Biochemistry* **37**, 10223–10230 (1998).
- Pérez, M., Valpuesta, J. M., Medina, M., Montejó de Garçon, E. & Avila, J. Polymerization of tau into filaments in the presence of heparin: the minimal sequence required for tau-tau interaction. *J. Neurochem.* **67**, 1183–1190 (1996).
- Schweers, O., Mandelkow, E. M., Biernat, J. & Mandelkow, E. Oxidation of cysteine-322 in the repeat domain of microtubule-associated protein tau controls the in vitro assembly of paired helical filaments. *Proc. Natl Acad. Sci. USA* **92**, 8463–8467 (1995).



39. R Development Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna) (<http://www.r-project.org>) (2008).
40. Rojas-Quijano, F. A., Morrow, D., Wise, B. M., Brancia, F. L. & Goss, W. J. Prediction of nucleating sequences from amyloidogenic propensities of tau-related peptides. *Biochemistry* **45**, 4638–4652 (2006).
41. Morris, A. M., Watzky, M. A., Agar, J. N. & Finkbe, R. G. Fitting neurological protein aggregation kinetic data via a 2-step, minimal "Ockham's razor" model: the Finkbe-Watzky mechanism of nucleation followed by autocatalytic surface growth. *Biochemistry* **47**, 2413–2427 (2008).
42. Schmidt, K., Segond von Banchet, G. & Heppelmann, B. Labelling of peptides with 1.4-nm gold particles to demonstrate their binding sites in the rat spinal cord. *J. Neurosci. Methods* **87**, 195–200 (1999).
43. Eisenberg, D., Wesson, M. & Yamashita, M. Interpretation of protein folding and binding with atomic solvation parameters. *Chem. Scr.* **29A**, 217–221 (1989).
44. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
45. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
46. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
47. DeLano, W. L. PyMOL, Molecular Viewer (<http://www.pymol.org>) (2002).
48. Papkalis, A., Munch, J., Otto, C. & Kirchhoff, F. Nef enhances human immunodeficiency virus type 1 infectivity and replication independently of viral coreceptor tropism. *J. Virol.* **76**, 8455–8459 (2002).
49. Platt, E. J., Wehrly, K., Kuhlmann, S. E., Chesbro, B. & Kabat, D. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophage-tropic isolates of human immunodeficiency virus type 1. *J. Virol.* **72**, 2855–2864 (1998).

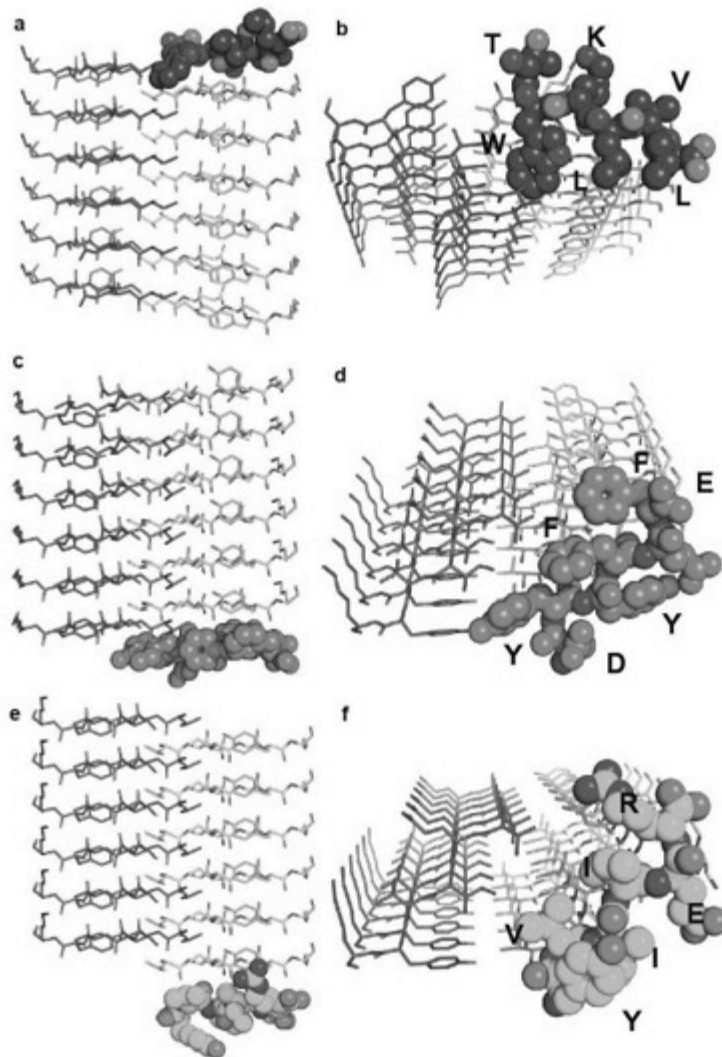
# SUPPLEMENTARY INFORMATION

doi:10.1058/nature10154



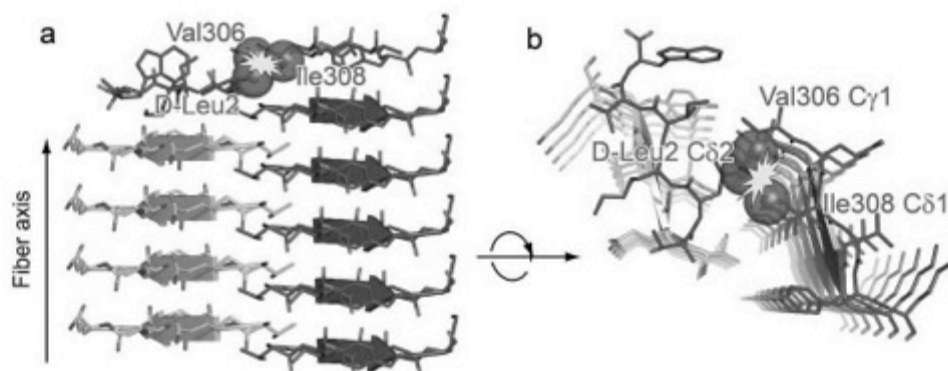
**Supplementary Figure 1. D-TLKIVW design model.** The modeled interaction of D-TLKIVW with VQIVYK (pdb id: 2ON9) is shown. **a**, The side view and top-down view, **b**,

of the designed inhibitor, D-TLKIVW (red), interacting with the VQIVVK steric zipper structure in purple and grey. **c**, The wall-eyed stereo view of the modeled D-TLKIVW segment on the VQIVVK steric-zipper structure shows complementarity between D-TLKIVW (red ball-and-stick representation) and both  $\beta$ -strands of VQIVVK. The D-TLKIVW molecule (dot-sphere representation) fits well on the surface of the pair of  $\beta$ -sheets from VQIVVK molecules (shown as spheres). **d**, In this stereo view, the purple  $\beta$ -sheet of VQIVVK is removed to show the interactions between D-TLKIVW and the grey VQIVVK  $\beta$ -sheet. The side chains of D-Leu2, D-Ile4, and D-Trp6 of D-TLKIVW make favorable interactions with the grey sheet from the VQIVVK molecule. Transparent spheres show interacting atoms in the model. D-Leu2 and D-Ile4 also interact with the N-terminal Val306 and Ile308 of the mating VQIVVK molecule. D-Lys3 of D-TLKIVW replaces the side chain hydrogen bonding interaction of stacked Gln in the VQIVVK structure (red dashes). Black dashes show the main chain hydrogen bonds between layers of peptide.

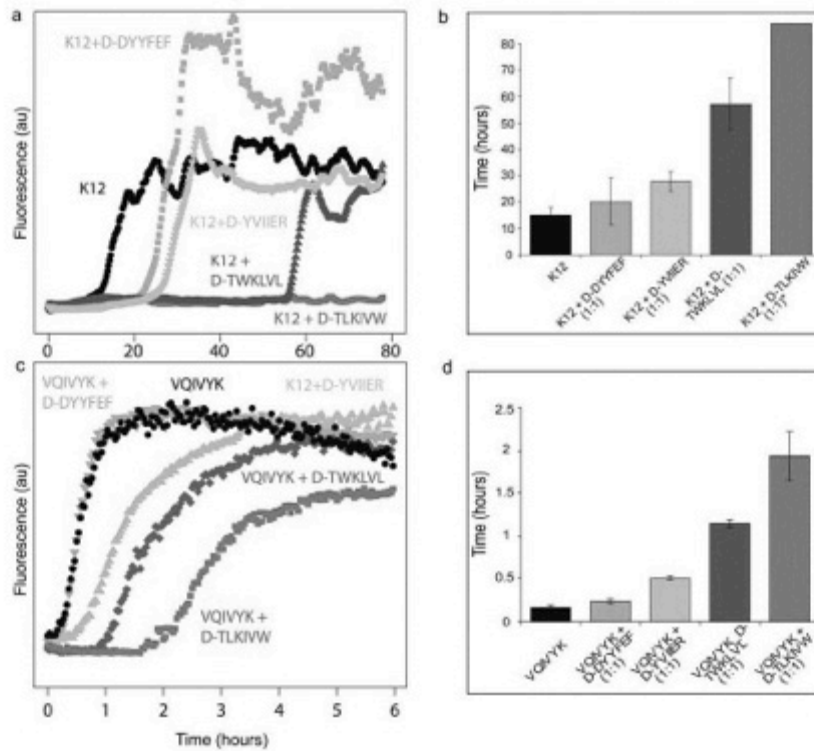


**Supplementary Figure 2. Other designed D-amino acid tau inhibitor models.** Besides D-TLKIVW, three additional D-amino acid peptides were predicted. The side view **a**, and top-down view **b**, of D-TWKLVL (in blue) show how the D-TWKLVL model is

predicted to interact with VQIVYK (purple and grey  $\beta$ -sheets). The bulky D-Trp2 residue would prevent the addition of another VQIVYK molecule on top of the purple  $\beta$ -sheet. **c** and **d**, D-DYYFEF was designed to interact with both faces of the hydrophobic VQIVYK steric zipper bottom, with the inner facing tyrosine and phenylalanine residues interacting with the  $\beta$ -sheet across the steric zipper interface. **e** and **f**, The bottom inhibitor D-YVIIER was designed to interact with the bottom grey  $\beta$ -sheet of VQIVYK, testing the hypothesis that an arginine residue would pack favorably against Tyr311 of VQIVYK.



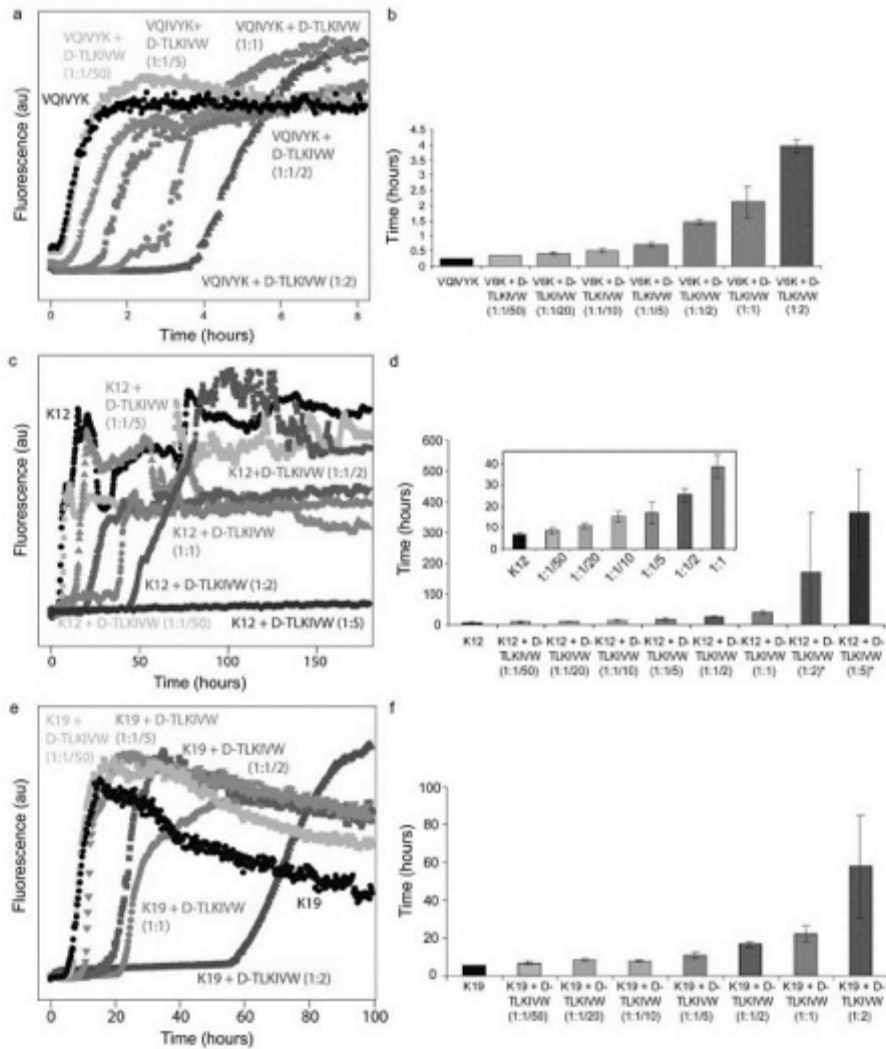
**Supplementary Figure 3. Steric repulsion between D-TLKIWW and VQIVYK molecules disrupts fibril growth.** The modeled interaction between D-TLKIWW and VQIVYK suggests that growth cannot continue on the  $\beta$ -sheet across from the sheet capped by D-TLKIWW. **a**, The view perpendicular to the fibril axis shows a steric clash between adding an additional VQIVYK segment on the purple sheet opposite the  $\beta$ -strand capped by the inhibitor molecule. **b**, Rotating the view so that the fibril axis is perpendicular to the plane of the figure shows that the C $\delta$ 2 of D-Leu2 of D-TLKIWW provides steric interference, preventing the addition of a VQIVYK molecule across from it, specifically via a clash with the Ile308 C $\delta$ 1 and Val306 C $\gamma$ 1 atoms.



**Supplementary Figure 4. Designed D-peptides delay VQIVYK and tau K12 fibril formation.** **a**, The fibril formation of 50  $\mu$ M K12 was monitored using ThS fluorescence. Traces are shown for the replicate with median lag time. The tau K12 construct forms fibrils after about 15 hours (black). Incubating tau with equimolar quantities of each D-peptide resulted in no inhibition (D-DYYFEF, green), moderate delay (D-YVIER, cyan), a lag of more than 50 hours (D-TWKLVL, blue), or inhibition for several days (D-TLKIWW, red). K12 fluorescence traces were smoothed to reduce signal noise. **b**, Lag time plots show the variation in the time for fibril formation to begin (error bars represent 1 s.d.). While D-DYYFEF does not show clear inhibition, the other 3 inhibitors clearly

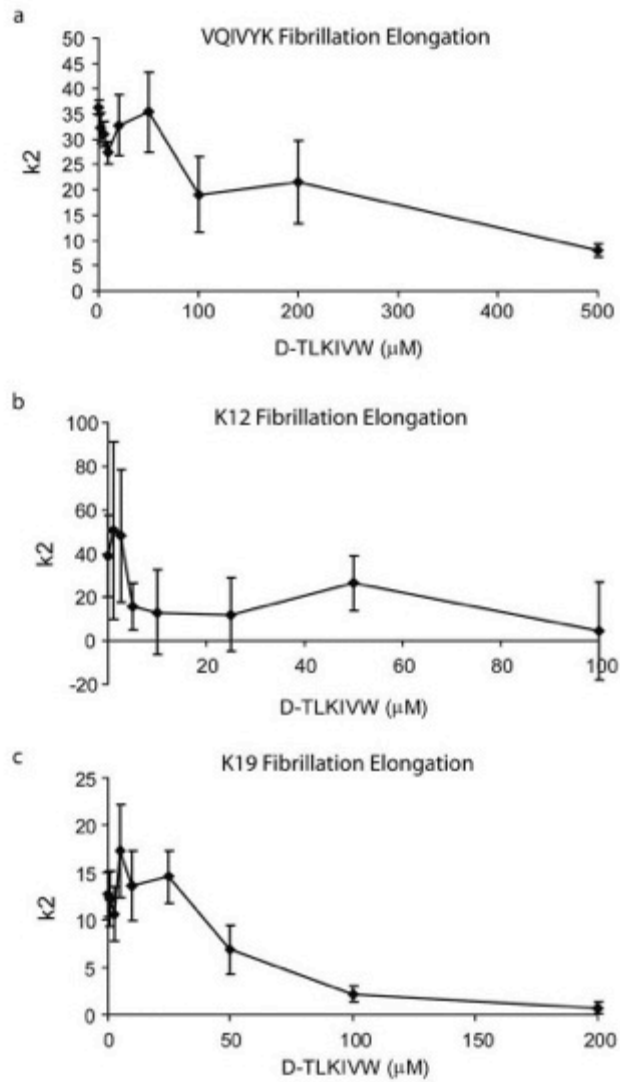
delay fibril formation, with D-TLKIVW delaying fibril formation for more than 100 hours in this experiment. The asterisk indicates that not all replicates had begun fibril formation at the end of the experiment and mean lag time could not be accurately determined. **c**, The fibril formation of 100  $\mu$ M VQIVYK begins in about 15 minutes alone or in the presence of equimolar D-DYYFEF, but is delayed by equimolar D-YVIIEE, D-TWKLVL, and D-TLKIVW. **d**, Lag time plots indicate little variation among the 6 replicates in this experiment and a significant delay due to D-TLKIVW.





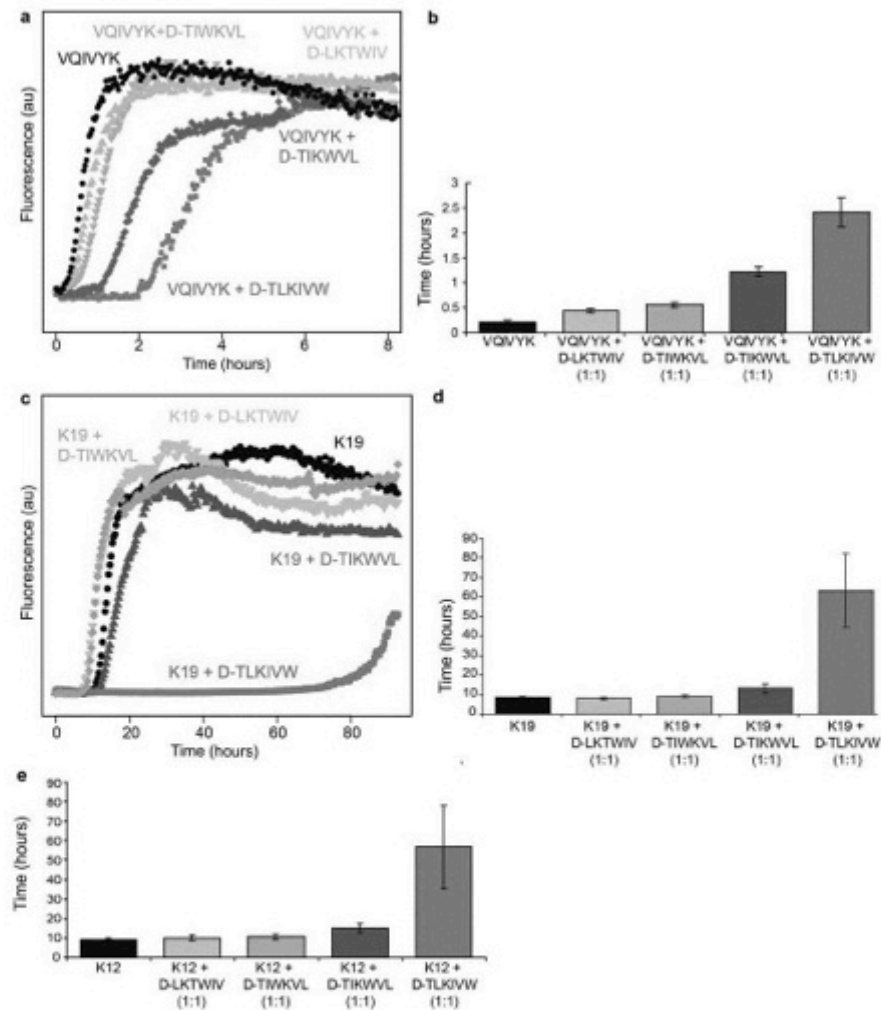
**Supplementary Figure 5. D-TLKIWW delays fibril formation in a ratio-dependent manner.** Monitoring ThS fluorescence as a function of time yields the fluorescence curves of **a**, 100  $\mu$ M VQIVYK; **c**, 50  $\mu$ M K12; and **e**, 50  $\mu$ M K19. All samples are delayed

longer by higher concentrations of the D-TLKIVW peptide. The fluorescence of the replicate with the median lag time is shown. Modest inhibitory effects can be seen at sub-stoichiometric ratios for all samples. A delay of several days can be seen when excess D-TLKIVW is present. Lag times for each sample demonstrate the effect of increasing inhibitor concentrations and the variation among replicates for **b**, VQIVYK; **d**, K12; and **f**, K19. Error bars represent the standard deviation of the lag time for four (VQIVYK) and 8 (K12 and K19) replicates at each ratio in this experiment. In panel **d**, the inset expands the time axis to show the differences between the peptide ratios at lower concentrations of the D-peptide. Asterisks indicate that not all replicates had begun to form fibrils at the conclusion of the experiment and a mean lag time and standard deviation could not be determined accurately.



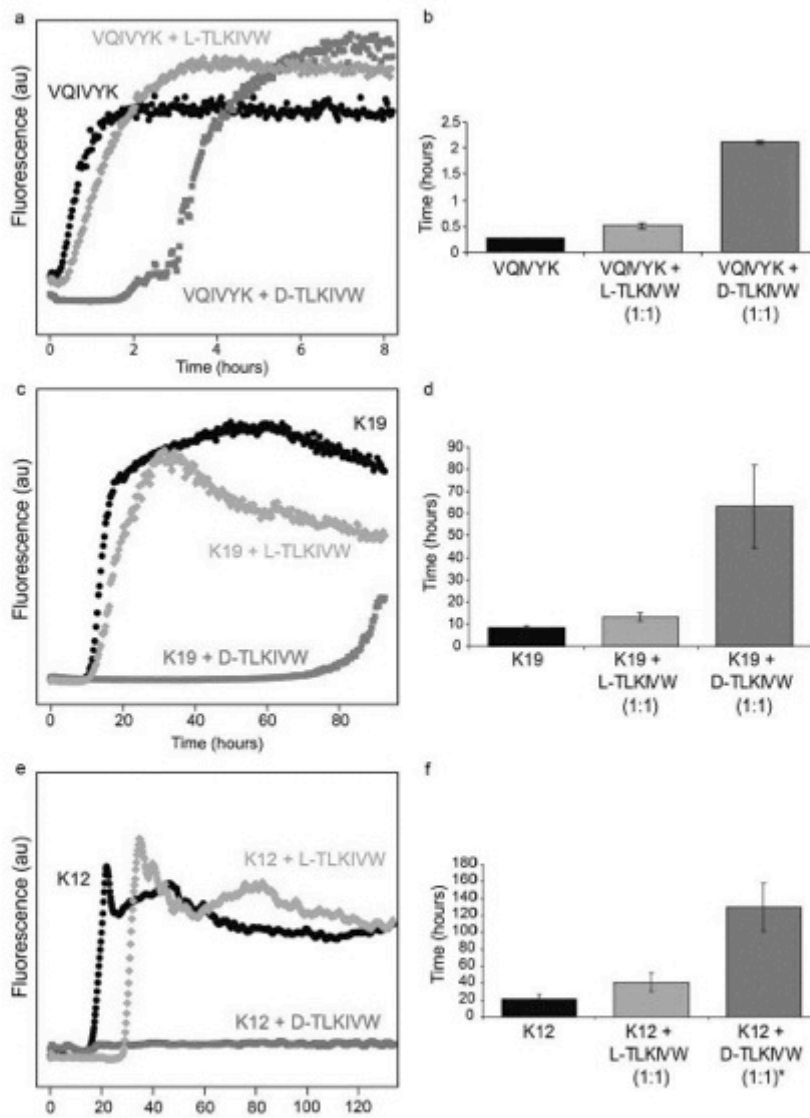
**Supplementary Figure 6. D-TLKIVW decreases the fibril propagation rate of VQIVYK and the tau constructs K12 and K19. Plots of  $k_2$  (proportional to the fibril**

elongation rate) as a function of increasing D-TLKIWW concentrations in fibril formation experiments of **a**, 100  $\mu\text{M}$  VQIVYK; **b**, 50  $\mu\text{M}$  K12; and **c**, 50  $\mu\text{M}$  K19 show that at higher concentrations of D-TLKIWW the rate of fibril growth after initiation is slower. The plots shown are determined from the data shown in Supplementary Fig. 5 with 4 replicates for each sample in the VQIVYK assay and 8 replicates for each assay with the tau constructs.



**Supplementary Figure 7. Scrambled inhibitors demonstrate sequence specificity in the design model.** Monitoring ThS fluorescence shows that in the presence of equimolar peptide the fibril formation of **a**, 100  $\mu$ M VQIVYK; **c**, 50  $\mu$ M K19; and 50  $\mu$ M K12 (Fig. 2e) is delayed significantly by D-TLKIWW and only slightly by D-TIKWVL (with

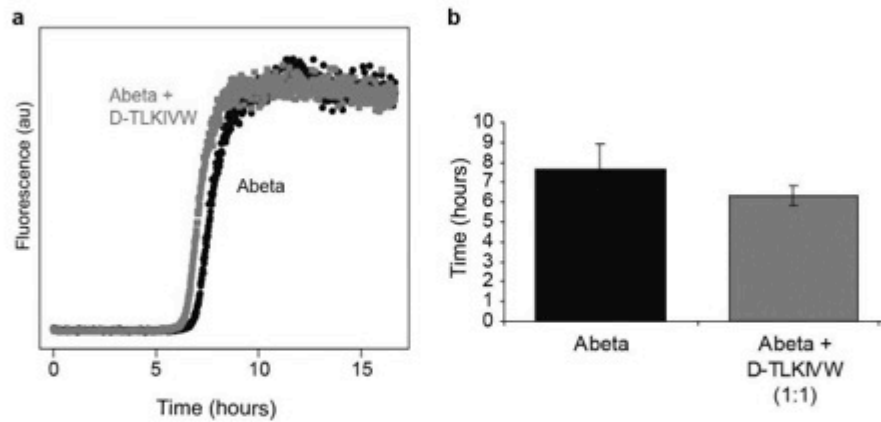
3 residues on 1 side of the  $\beta$ -strand permuted), while D-TIWKVL and D-LKTMIV have no effect on tau fibril formation. The traces of replicates with the median lag time are shown in **a** and **c**. **b**, **d**, and **e**, Lag time plots for VQIVYK, K19, and K12 fibril formation show the variation of replicates in each experiment. The error bars represent the standard deviations of the lag times for the 6, 4, and 3 replicates in the VQIVYK, K19, and K12 experiments, respectively.



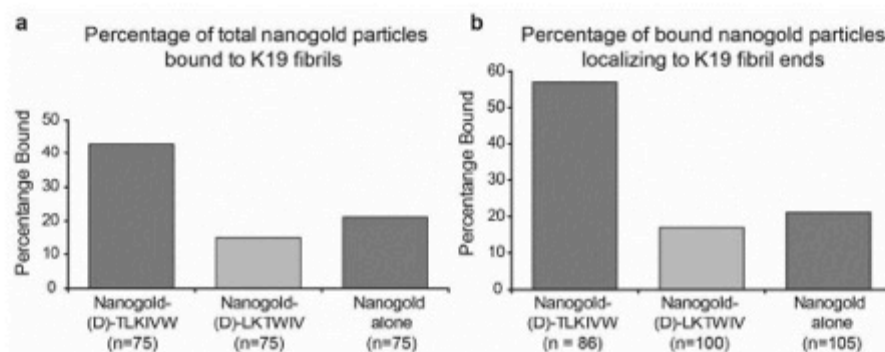
**Supplementary Figure 8. L-TLKIVW diastereomer has little effect on tau fibril formation.** The fibril formation of the tau peptide and constructs alone is shown in

black. Comparing inhibition by equimolar D-TLKIVW to its diastereomer shows that D-TLKIVW (red) delays fibril formation much longer than L-TLKIVW (green). Fibril formation plots showing the ThS fluorescence traces of the sample with median lag time demonstrate that the designed D-amino acid peptide is much more effective than its L-amino acid diastereomer. **a** and **b**, show the fibril formation plot and lag time plots for 100  $\mu$ M VQIVYK in the presence of either D- or L-TLKIVW, with 4 replicates per sample. **c** and **d**, show the fluorescence and lag time plots for 50  $\mu$ M K19 fibril formation with 4 replicates. **e**, The fibril formation plots for 50  $\mu$ M K12 with the 2 peptides are shown. Each K12 fluorescence trace was smoothed to reduce signal noise. **f**, The lag time plot for K12 with each diastereomer shows that the L-peptide has a minor effect on K12 fibril formation compared to the much larger effect of the D-amino acid stereoisomer. The asterisk for K12 plus equimolar D-TLKIVW indicates that 7 of 8 replicates had not begun to form fibrils after 140 hours in the experiment and the mean and standard deviation could not be accurately calculated, with the mean being underestimated.

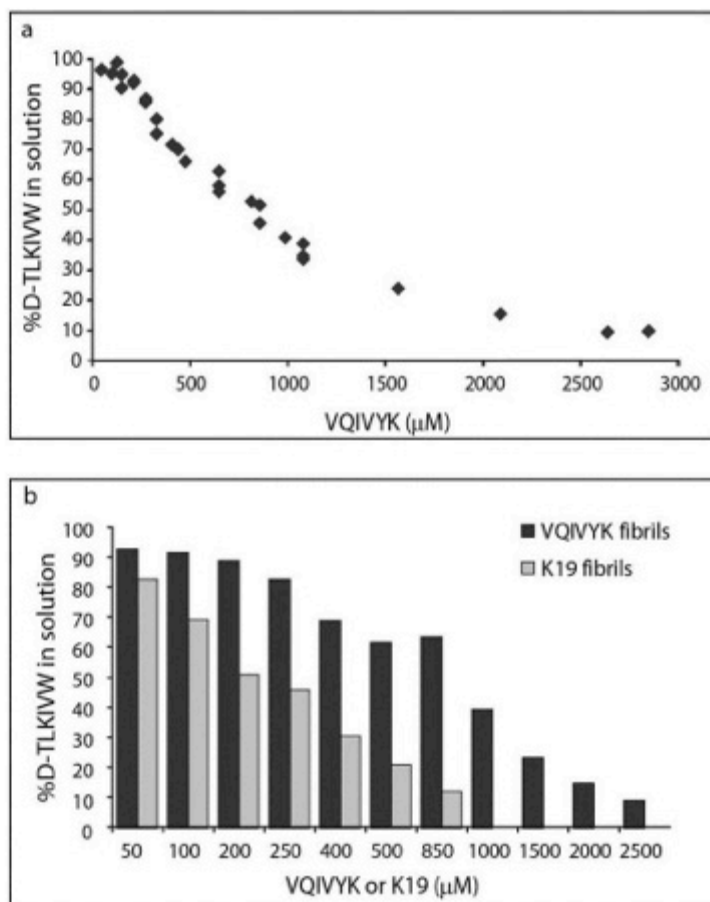




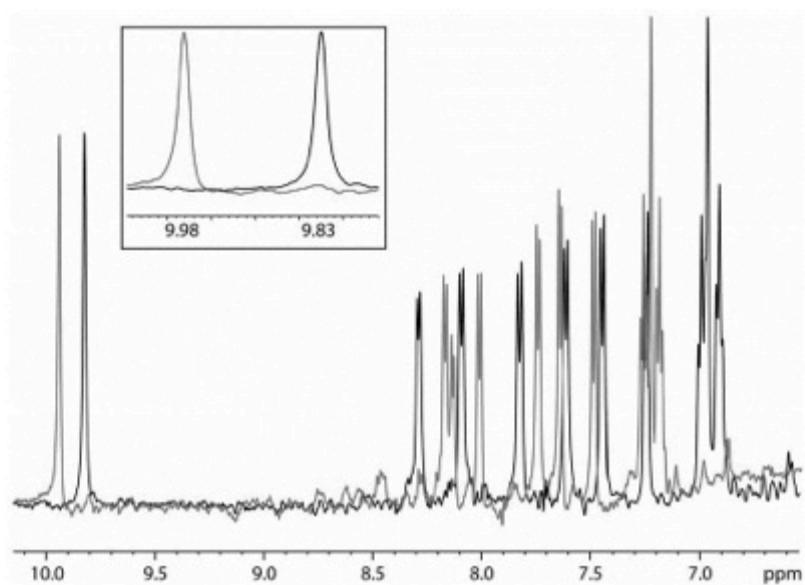
**Supplementary Figure 9. D-TLKIVW is not a general inhibitor of fibril formation. a,** As monitored by ThT fluorescence, fibril formation of 11.5  $\mu$ M amyloid beta peptide (black) is not delayed by the addition of equimolar D-TLKIVW (red). The trace of the replicate with median lag time is displayed for both samples. **b,** The lag time plot indicates that D-TLKIVW does not have a significant effect on amyloid beta fibril formation. Error bars represent the standard deviation of the lag times of the four replicates per sample.



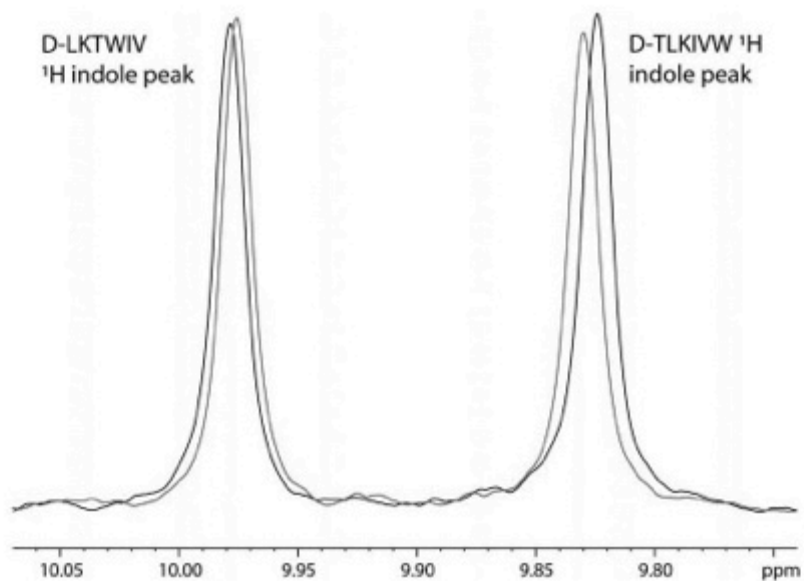
**Supplementary Figure 10. Quantification of nanogold-labeled D-TLKIVW bound to K19 fibrils.** **a**, This plot shows the percentage of nanogold particles bound to fibrils compared to those unbound. Nanogold-bound D-TLKIVW molecules have a preference to interact with K19 fibrils compared to nanogold-bound D-LKTWIV, and nanogold alone. **b**, This plot shows the percentage of nanogold particles bound to K19 fibrils that localize to fibril ends. Of nanogold particles found bound to K19 fibrils, nanogold-bound D-TLKIVW molecules have a preference for K19 fibril ends when compared to nanogold-bound D-LKTWIV, and unbound nanogold.



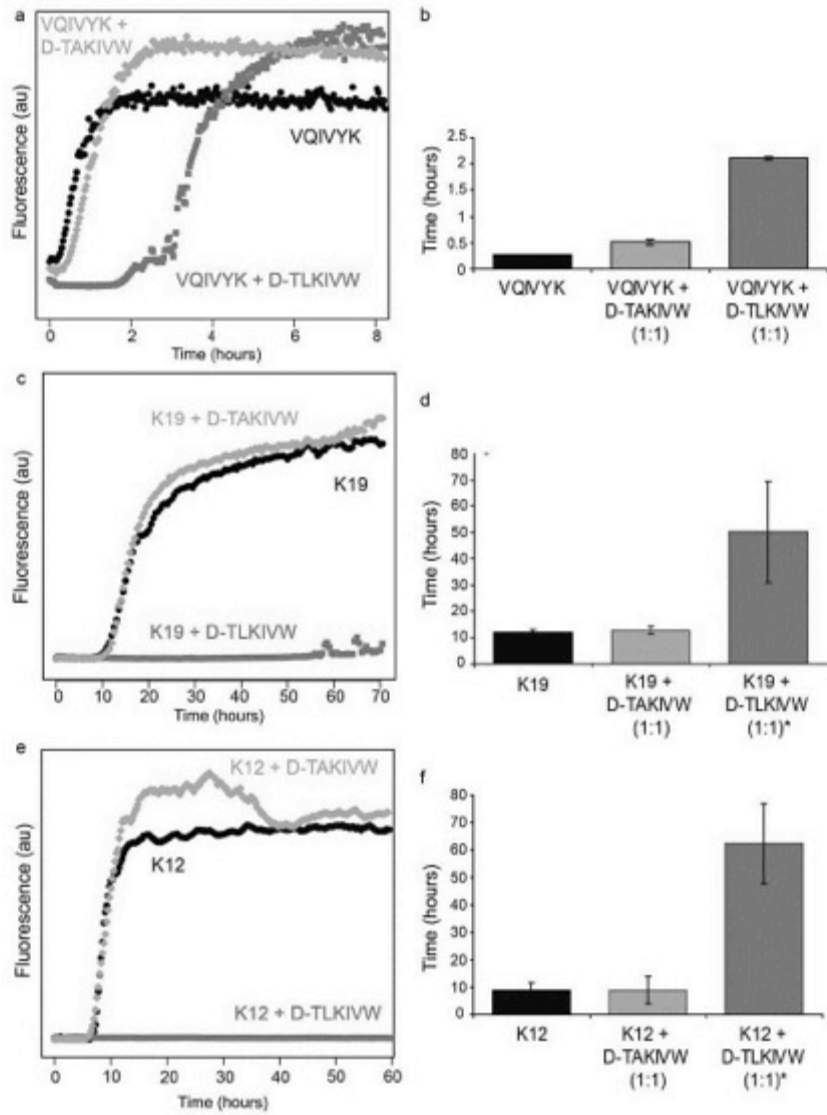
**Supplementary Figure 11. D-TLKIWW interacts with VQIVYK and K19 fibrils.** **a**, The plot shows the percentage of D-TLKIWW no longer in solution as a function of increasing concentrations of fibrillar VQIVYK. **b**, The plot compares the percentage of D-TLKIWW no longer in solution in the presence of VQIVYK fibrils (dark blue) and K19 fibrils (light blue). The concentrations shown are for VQIVYK or K19 monomers in each fibril sample, since fibril concentration can only be estimated.



**Supplementary Figure 12. The 500 MHz <sup>1</sup>H spectra of D-TLKIVW and D-LKTWIV.** The 500 MHz <sup>1</sup>H NMR spectra of D-TLKIVW and D-LKTWIV are shown in black and red, respectively. The inset shows the Trp indole region of both peptides, with the peak of D-TLKIVW around 9.83 ppm and D-LKTWIV around 9.98 ppm.



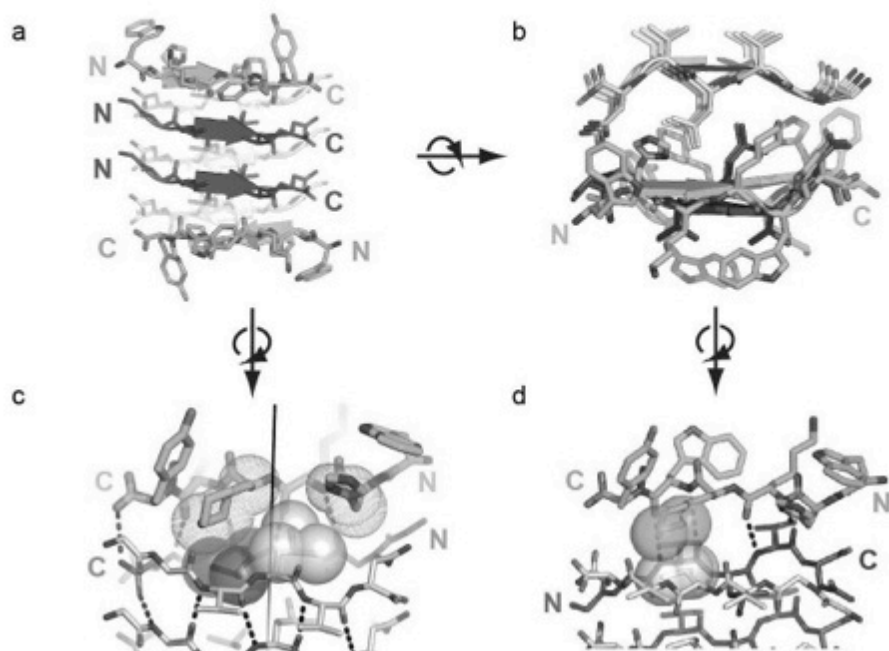
**Supplementary Figure 13. D-TLKIWW does not interact with soluble VQIVYK.** The indole region of the 500 MHz <sup>1</sup>H NMR spectra of 100  $\mu$ M D-TLKIWW and D-LKTWIV peptides in the absence (black) presence of 7 mM soluble VQIVYK (red) is shown. Small differences in the indole peaks for both D-TLKIWW and D-LKTWIV when 70-fold excess VQIVYK is present indicate that neither peptide interacts strongly with soluble VQIVYK.



**Supplementary Figure 14. D-TAKIWW has little effect on tau fibril formation.**

Equimolar D-TLKIWW (red) delays fibril formation while D-TAKIWW (green) has no effect.

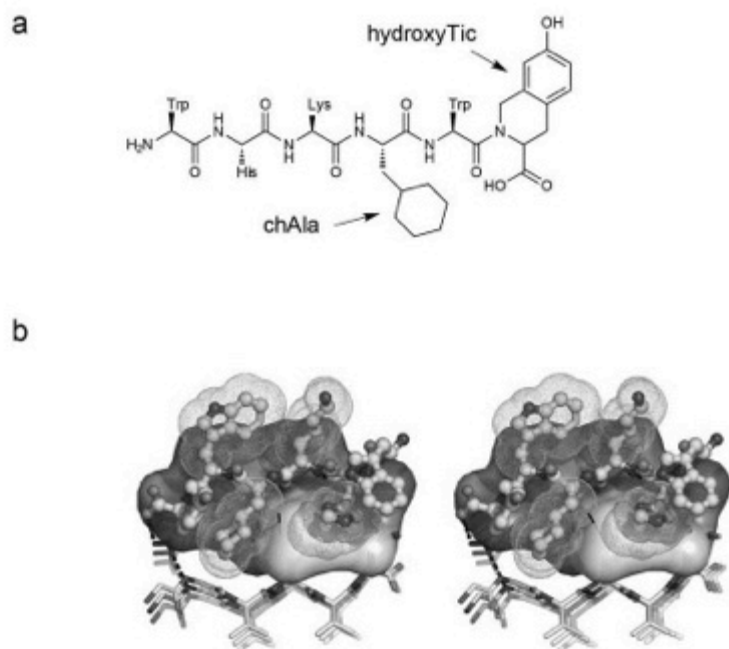
Fibril formation plots show the ThS fluorescence trace of the sample with median lag time and indicate that the substitution of D-Leu2 with D-Ala in the designed inhibitor eliminates inhibition activity. **a** and **b**, show the fibril formation plot and lag time plots for 100  $\mu$ M VQIVYK (black) in the presence of either D-TLKIVW (red) or D-TAKIVW (green), with 4 replicates per sample. **c** and **d**, show the plots for K19 fibril formation with 8 replicates per sample. **e** and **f**, show the plots for K12 fibril formation with 12 replicates per sample. Each K12 fluorescence trace was smoothed to reduce signal noise. Asterisks indicate that not all replicates had begun fibril formation by the end of the experiment, preventing the accurate determination of the mean and standard deviation in these cases. Error bars represent the standard deviations of the replicates for each sample.



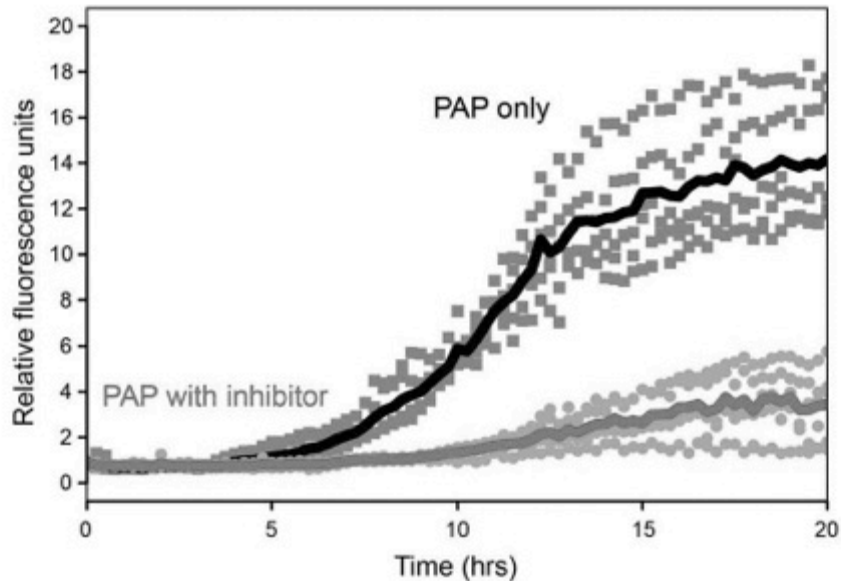
**Supplementary Figure 15. Binding orientations of designed inhibitor of  $^{248}\text{PAP}^{296}$  fibril formation.** The peptide inhibitor, WW61, designed using the GGVLVN template from PAP is shown. **a** and **b** show the side view and top-down views of the designed inhibitor interactions, respectively. Two possible binding modes are proposed: 1) the peptide inhibitor (cyan) aligns on the top of the fibril parallel to one of the  $\beta$ -sheets; 2) the peptide inhibitor (cyan) aligns on the bottom of the fibril in an anti-parallel orientation relative to one of the fibril's  $\beta$ -sheets. The N- and C- termini of the  $\beta$ -strands of the fibril and inhibitor are labeled. **c**, The side view of the parallel binding mode (1). This binding mode is from the original design with favorable calculated binding energy. The side chains of chAla and His (cyan dotted spheres) make favorable interactions with the



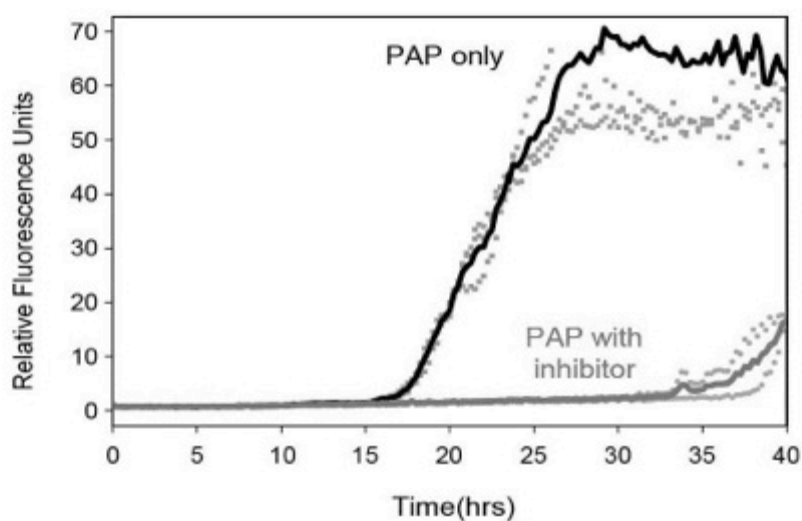
core residues Leu (purple and grey spheres) from the GGVLVN molecule. Only the contact atoms are shown in sphere representation to highlight their apolar interactions. The black dashed lines represent hydrogen bonds. 1 backbone-side-chain and 4 backbone-backbone hydrogen bonds between the inhibitor and fibril end are shown. **d**, The side view of the anti-parallel binding mode (2). This binding mode is not used in our original design model, and it has less favorable interactions, as calculated. The anti-parallel binding mode shares a similar binding interface to the parallel model with the same contacting residue pairs and atomic interactions. Although our original design is based on the parallel mode with more favorable calculated binding energy, we cannot exclude the possibility of the peptide inhibitor binding to the fibril end in some other way, such as in an anti-parallel orientation.



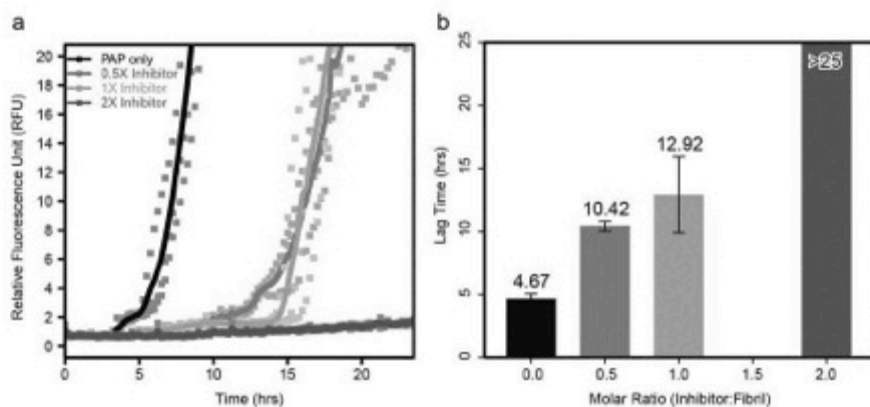
**Supplementary Figure 16. The chemical structure of the non-natural peptide Trp-His-Lys-chAla-Trp-hydroxyTic (WW61).** **a**, ChAla represents  $\beta$ -cyclohexyl-L-alanine and hydroxyTic is 7-hydroxy-(*S*)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid. The N- and C-termini of the peptide are free. **b**, The wall-eyed stereo image shows how the designed inhibitor peptide (cyan ball-and-stick) is predicted to interact with the peptide segment GGVLN. The surface representation of the fibrillar GGVLN highlights the tight interactions between the non-natural chAla from the inhibitor and the core residue Leu from the fibrillar GGVLN template. The 4 inner residues of the peptide inhibitor are shown in dot-sphere representation to emphasize their shape complementarity to the outer residues of the inhibitor, and to the fibrillar GGVLN structure.



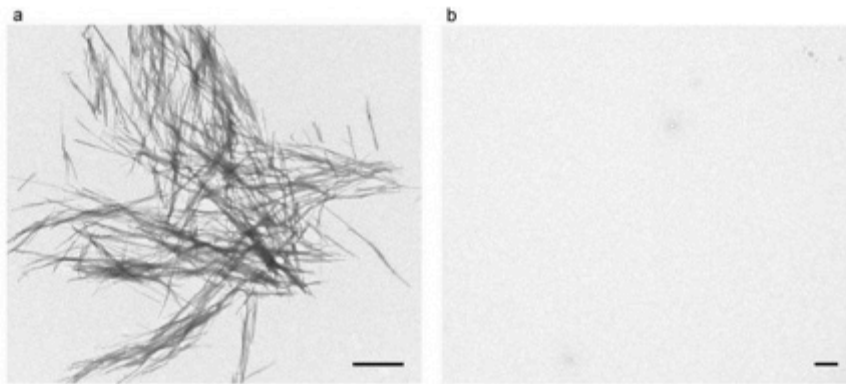
**Supplementary Figure 17. Inhibitory effect of the designed peptide WW61 on seeded  $^{248}\text{PAP}^{206}$  fibril formation.**  $^{248}\text{PAP}^{206}$  fibril formation was monitored by observing ThT fluorescence over time for samples with 5 replicates. In the presence of preformed  $^{248}\text{PAP}^{206}$  fibrils, there is an increase in fluorescence in  $^{248}\text{PAP}^{206}$  alone after about 5 hours and fibrils are observed by electron microscopy. In the presence of 2-fold excess inhibitor, there is a small rise in signal, but fibrils were not observed by electron microscopy (Supplementary Fig. 20). The signal remains low for more than 50 hours (see Fig. 4c). Solid red and black lines show mean fluorescence values with and without inhibitor, respectively.



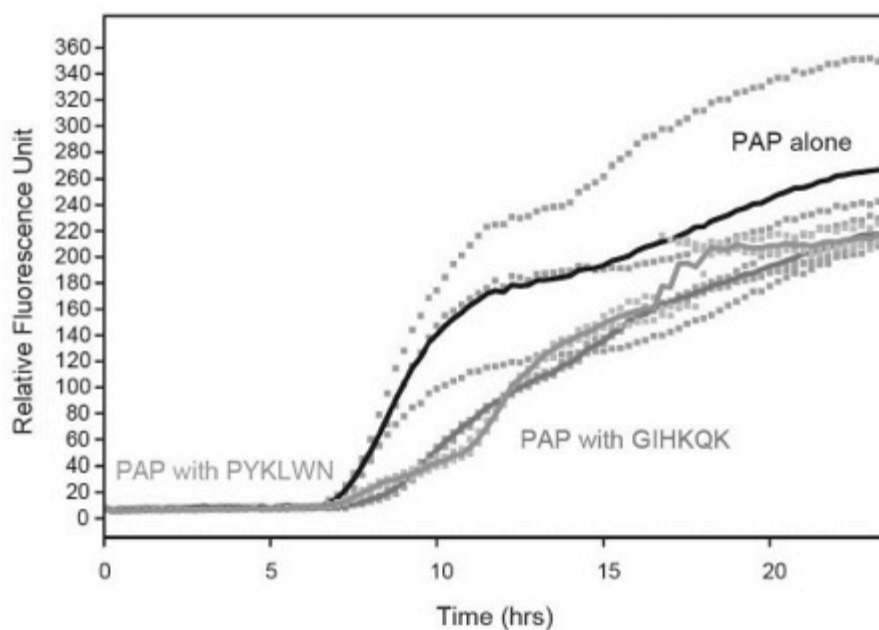
**Supplementary Figure 18. Inhibitory effect of the designed peptide, WW61, on unseeded  $^{248}\text{PAP}^{286}$  fibril formation.**  $^{248}\text{PAP}^{286}$  fibril formation was monitored by observing ThT fluorescence over time for samples in triplicate. The lag time for 0.66 mM  $^{248}\text{PAP}^{286}$  alone (grey dots) was  $17 \pm 0.2$  hours (s.d.), based on an arbitrary fluorescence value cutoff of 3. When a 2-fold molar excess of the inhibitor peptide was present, the lag time of  $^{248}\text{PAP}^{286}$  fibril formation (pale red dots) was extended to  $35 \pm 3$  hours (s.d.). Solid red and black lines show mean fluorescence values with and without inhibitor, respectively.



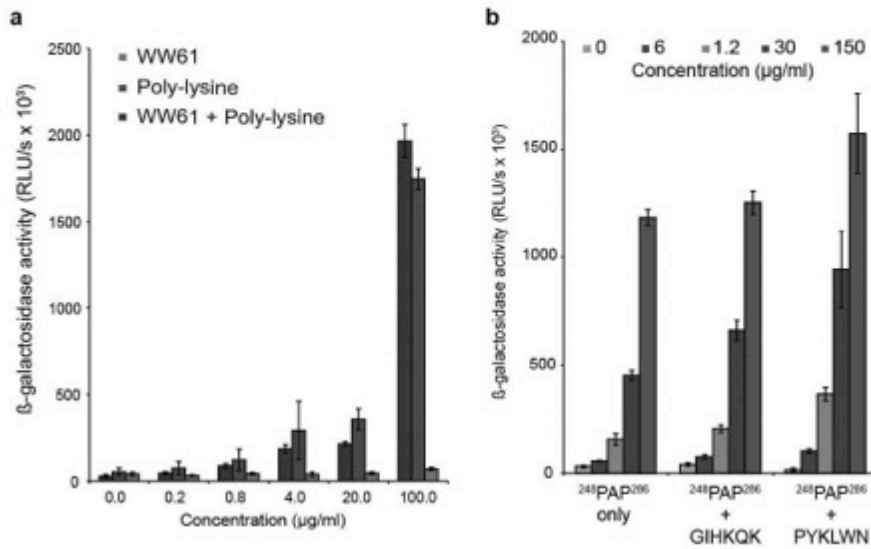
**Supplementary Figure 19.** The non-natural peptide WW61 shows a ratio-dependent inhibition effect on seeded  $^{248}\text{PAP}^{286}$  fibril formation. **a**,  $^{248}\text{PAP}^{286}$  fibril formation was monitored by observing ThT fluorescence over time when varying the molar ratio of peptide inhibitor and  $^{248}\text{PAP}^{286}$ . Three replicates were plotted for each molar ratio (PAP only, 1:2, 1:1, 1:0.5). The solid lines show the average value of the fluorescence traces. **b**, The plot of molar ratio vs. lag time of  $^{248}\text{PAP}^{286}$  fibril formation. The lag time for  $^{248}\text{PAP}^{286}$  fibril formation is determined by the starting time whose ThT fluorescence is greater than an arbitrary fluorescence value of 2. The error bar indicates the standard deviation of the lag times for the three replicates of each sample.



**Supplementary Figure 20.** Negative-stain electron micrographs show the effect of the inhibitor WW61 on  $^{248}\text{PAP}^{286}$  fibril formation. **a**, The electron micrograph shows fibrils of  $^{248}\text{PAP}^{286}$ . **b**, Fibril formation of  $^{248}\text{PAP}^{286}$  is inhibited by a 2-fold excess of the non-natural amino acid inhibitor, WW61. Samples were prepared following 20 hours of incubation with  $^{248}\text{PAP}^{286}$  seeds (scale bars,  $1\mu\text{m}$ ).



**Supplementary Figure 21. No inhibitory effect of the control peptides GIHKQK and PYKLWN on seeded  $^{248}\text{PAP}^{286}$  fibril formation.** GIHKQK, a peptide segment from the N-terminal region of  $^{248}\text{PAP}^{286}$ , not involved in SEVI formation, and PYKLWN, an unrelated peptide with the same charge as the designed inhibitor WW61, were selected as negative controls. ThT fluorescence was used to monitor the fibril formation of 0.66 mM  $^{248}\text{PAP}^{286}$  over time in the absence (grey dots) and presence of the 2-fold molar excess of the control peptides, GIHKQK (red dots) and PYKLWN (green dots). The solid lines show the average value of the fluorescence traces.



**Supplementary Figure 22. Control HIV infectivity assays.** HIV infection rates were determined by measuring reporter enzyme activity in cellular lysates. The mean values derived from triplicate infections are shown. Error bars represent one standard deviation. RLU/s; relative light units per second. **a**, WW61 alone does not affect HIV-1 infection and does not affect poly-lysine mediated infectivity enhancement. **b**, Control peptides GIHKQK and PYKLWN do not affect HIV-1 infection rates.



**Supplementary Table 1.** Comparison of D-peptides and controls.

<b>Sequence</b>	<b><math>\Delta</math>Binding energy</b>	<b>Area buried (Å<sup>2</sup>)</b>	<b>Shape complementarity</b>	<b>Inhibitory activity</b>
<b>Designed inhibitors</b>				
D-TLKIVW	-1.7	327	0.78	++++
D-TWKLVL	0.3	342	0.81	+++
D-YVIIER	8.8	261	0.80	++
D-DYYFEF	1.2	358	0.78	+
<b>Negative controls</b>				
D-TAKIVW	3.0	313	0.79	-
L-TLKIVW	3.4	386	0.72	+
D-TIKWWL	9.9	313	0.72	+
D-TIWKVL	11.7	316	0.65	-
D-LKTWIV	12.2	350	0.72	-

Activity is ranked from highest activity (++++) to no activity (-). Each peptide model was minimized using RosettaDesign allowing both the template and inhibitor to move freely. The change in binding energy shown is in reference to an equivalent L-VQIVYK peptide in the same position. Values for area buried and shape complementarity were determined from the final models.

**Supplementary Table 2. Data collection and refinement statistics (Molecular replacement)**

GGVLVN	
<b>Data collection</b>	
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	4.83, 17.7, 40.3
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 90.00, 90.00
Resolution (Å)	1.50 (1.55-1.50)*
<i>R</i> <sub>sym</sub>	0.149 (0.256)
<i>I</i> / <i>σ</i> <i>I</i>	8.9 (4.5)
Completeness (%)	86.4 (74.0)
Redundancy	4.9 (2.6)
<b>Refinement</b>	
Resolution (Å)	1.5
No. reflections	548
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.182/0.219
No. atoms	
Protein	39
Ligand/ion	5
Water	1
B-factors	
Protein	8.7
Ligand/ion	17.5
Water	13.8
R.m.s deviations	
Bond lengths (Å)	0.003
Bond angles (°)	1.291

Datasets from two crystals were used in GGVLVN structural determination.

\*Highest resolution shell is shown in parenthesis.

## Chapter 4

### Small Molecules that Dissolve Preformed Fibrils

#### Abstract

*In vivo* deposition of amyloid fibrils is a hallmark in the pathogenesis of amyloid diseases. In some of these diseases, fibrils are the toxic species. Inhibitors that dissolve preformed fibrils provide an important therapeutic option. In other diseases, oligomers are the toxic species and fibrils serve as a reservoir of toxic oligomers. In these conditions, stabilization of the fibrils may provide viable therapeutic strategies. Most current studies have identified inhibitors that prevent fibrillation but have little to no effect against preformed fibrils. Here, I used SEVI as a model system and screened a library of 2,000 small molecules. In this assay condition, SEVI fibrils is partially destabilized and slowly self-depolymerize over ~18 hours. Using this assay, I identified 9 compounds that facilitate the dissociation of fibrils and 10 compounds that appear to stabilize SEVI fibrils.

## Introduction

Semen-derived enhancer of virus infection (SEVI) is a proteolytic fragment from human prostatic acidic phosphatase (248PAP286). It is an extracellular peptide that is naturally present in semen at a concentration of ~35 µg/ml. Upon releasing from the parent protein, the peptide fragments spontaneously assemble into amyloid fibrils, termed SEVI (Munch et al., 2007). 248PAP286 carries a number of residues that are positively charged under physiological pH. Similarly to other amyloid fibrils, SEVI has a classic cross- $\beta$  fibril diffraction pattern indicating that a parallel in-registered arrangement in the fibril ultrastructure. In the fibril structure model, the positive residues from individual peptide unit line up to form a local high concentration of cationic residues. This facilitates the fusion of virus and target cells, both of which have negatively charged membranes. As a result, SEVI fibrils enhance HIV infection for more than 100,000-fold (Munch et al., 2007; Roan et al., 2009). This enhancement is only observed with the fibrillar form of the peptide. Monomeric species is completely inactive (Munch et al., 2007). Thus, SEVI fibrils present a potential target for reducing the incidence of HIV transmission.

We have previously designed structure-based peptide inhibitors that efficiently disrupted the monomeric peptide from assemble into mature SEVI fibrils (Sievers et al., 2011). Recently, Cu(II) and Zn(II) have also been shown to have inhibitory activities against SEVI formation (Sheftic et al., 2012). In both studies, the inhibitors were only active when mixed with monomeric 248PAP286. However, in a follow up study on SEVI in 2010, infectivity assay conducted with “fresh” semen samples showed significant enhancement of HIV infection (Kim et al., 2010). This indicates that SEVI fibrils are likely to have already formed before ejaculation. Furthermore, additional seminal peptides have been identified to form amyloid fibrils and enhance HIV infection. These include a peptide released from a region different from

248PAP286 (Arnold et al., 2012), as well as peptides released from semen coagulum protein (Roan et al., 2011). Therefore, it is possible that the enhancement of infection is a collective effect of multiple amyloid species. These recent developments on semen-derived amyloid fibrils suggest that viable treatment options should not only prevent fibril formation, but also be able to disaggregate preformed fibrils. Furthermore, because of the multiple amyloid fibrils present in semen, a broad-spectrum inhibitor is highly desirable.

Majority of the small molecule inhibitors discovered thus far have activity only on preventing fibril formation. Most of them are shown to remodel the fibrillation pathway and trap the amyloid proteins into non-toxic oligomeric aggregates that are not cytotoxic (Ehrnhoefer et al., 2008; Klunk et al., 1998; Kroth et al., 2012; Ladiwala et al., 2011). Furthermore, in many of these studies, fluorescent-based assay was used to evaluate the relative abundance of the fibrils. Certain small molecules have quenching effects, misleading the data interpretation (Kroth et al., 2012). EGCG is the only compound so far that has been shown to dissolve preformed fibrils when added as equal-molar concentration (Bieschke et al., 2010). This disaggregation effect of EGCG has been shown with  $\alpha$ -synuclein and amyloid- $\beta$  fibrils (Bieschke et al., 2010). However, EGCG was not able to effectively dissolve preformed SEVI fibrils. Rational-design of small molecule inhibitors has been challenging for amyloid fibrils. The surface of the fibril is a flat structure, lacking a well-defined drug pocket that is typically preferred for *de novo* design. X-ray crystallography has shown that several of the small molecules can slide along the spine of the fibrils (Landau et al., 2011). Additionally, amyloid fibrils are metastable structures with tightly packed interfaces. Therefore, the inhibitor must have very high affinity to be able to disrupt the interactions present in the preformed fibrils.

Here, we took a high-throughput screening approach to identify compounds that can

potentially dissolve fibrils. We screened a small molecule library of 2,000 compounds that were identified from natural sources. We monitored the abundance of the fibrils over 16 hours upon adding compounds. We identified a set of 9 compounds that are capable of destabilizing SEVI fibrils. Additionally, we have also identified 10 compounds that stabilize SEVI fibrils. For neurodegenerative diseases such as Alzheimer's disease, numerous studies have shown that the pathogenic species is the amyloid oligomers and fibrils are non-toxic byproducts (Haass and Selkoe, 2007). Since oligomers and fibrils co-exist in equilibrium, fibril-stabilizing compounds could shift this equilibrium towards fibrils. This approach could reduce the abundance of oligomers and thus be useful for treating this class of diseases.

## **Results**

In my previous work, I have developed inhibitors that prevent fibril formation. In these assays, inhibitor peptides were mixed with monomeric 248PAP286 peptides and prevent the association of these peptides to form elongated fibrils. In a functional assay where the inhibitors and 248PAP286 peptides are added to cell medium, SEVI mediated HIV infection was inhibited (Sievers et al., 2011). However, we later learned that these inhibitors are not active if seminal fluid instead of purified 248PAP286 peptides were used in this assay (Figure 1). Previous work has shown that “fresh” semen sample already has HIV enhancing material, suggesting that SEVI fibrils are formed before ejaculation (Kim et al., 2010). Our inhibitors designed to prevent elongation of the fibrils are not active against preformed fibrils.

I have also tested EGCG's ability to disaggregate SEVI fibrils. EGCG is a polyphenol that has been well studied *in vitro* as an amyloid inhibitor. This molecule has shown to remodel the fibrillation pathway towards non-toxic amyloid oligomers. Its activity has been demonstrated

in amyloid  $\beta$  and  $\alpha$ -synuclein proteins (Bieschke et al., 2010). In this system, equal molar of EGCG was added. The fibril samples disaggregate and spherical oligomers can be observed under electron microscope. Here, I have shown that at 5x and 10x molar excess, EGCG was not able to dissolve preformed SEVI fibril even after 24 hours of incubation at ambient temperature (Figure 2).

To identify compounds that can potentially dissolve preformed SEVI fibrils, we adapted a high throughput screening approach. Amyloid fibrils can be recognized by the fluorescent dye Thioflavin T. At excitation wavelength 440 nm and emission wavelength 480 nm, fluorescence can only be observed with fibrils, but not monomeric peptides. Therefore, we use ThT based assay as readout for the detection of the presence of SEVI fibrils. Since high throughput assays require substantial amount of SEVI peptides that is challenging to be obtained using the recombinant protein production methods, we first investigated the lowest concentration that would allow us to differentiate monomers from fibrils using ThT fluorescent assays. We found that ThT dye binds strongly to SEVI fibrils and significant fluorescent signals were observed even at 0.1 mg/ml peptide concentrations (Figure 3a). Additionally, since small molecules in the compound library are dissolved in DMSO, we also tested the effect of DMSO on the fibril stability and the fluorescent readings. We observed that 2% final DMSO concentration had no effect on the fluorescent readout (Figure 3b).

The high throughput assay design is as following. Preformed SEVI fibrils were diluted to 0.1mg/ml in PBS with 10  $\mu$ M ThT. This solution was pipet into glass bottom 384-well plates at 10  $\mu$ l per well. 100 nl of small molecules were added as final concentration of 10  $\mu$ M and 1% final DMSO concentration. Fluorescent readings were taken every two hours for 16-48 hours. We have found that without the presence of small molecules, SEVI fibril self-dissociate at sub-

mg/ml concentration (figure 4). At the target concentration of 0.1mg/ml, majority of the SEVI fibrils dissociate into monomeric peptide after roughly 30 hours of incubation at ambient temperature. Similar depolymerization kinetics has been observed with heat-induced dissociation of insulin fibrils and  $\beta$ 2-microglobulin fibrils (Kardos et al., 2011; Shamma et al., 2011).

To examine the effect of small molecule compounds, we plotted the depolymerization curve over 16 hour of incubation for each of the 2,000 compounds. The control well with 0.1 mg/ml SEVI showed the depolymerization curve that have seen in our preliminary assay (Figure 5). I then looked through each of the plot and identified 16 compounds that had effects on the depolymerization kinetics of SEVI fibrils. These compounds can be categorized as two broad types. 9 out of the 19 compounds facilitate the dissociation of SEVI fibrils (Table 1). In control well, significant dissociation of fibrils can be observed after 12 hours, while with these compounds, significant dissociation was observed between 6-8 hours. 10 out of the 19 compounds showed an opposite effect (Table 2). They appear to stabilize SEVI fibrils over the 16 hours of incubation. The chemical structures and the kinetic curves of these compounds are listed in table 1 and table 2.

## **Discussion**

Amyloid diseases can be caused by either mature fibrils or soluble oligomers. In many diseases that involve mature fibrils, the disease symptoms only appear after the fibrils have already been formed. Therefore inhibitors that can disaggregate preformed fibrils have important therapeutic potentials. In other diseases where oligomers are the toxic species, mature fibrils are considered a benign reservoir for toxic oligomers. In this class of diseases, it maybe beneficial to identify compounds that can stabilize the fibrils, shifting the equilibrium to mature fibrils and



reduce the population of toxic oligomers in the solution.

The search for small molecule inhibitors has not been very successful. Most of these compounds are metal chelators (Bush, 2003), dyes (Klunk et al., 1998) and polyphenolic natural products (Bieschke et al., 2010; Ehrnhoefer et al., 2008; Ladiwala et al., 2010; Ramassamy, 2006). The arguably most effective compound so far is EGCG a polyphenol that has been shown to eliminate A $\beta$  toxicity by converting the mature fibrils to non-toxic oligomers. However, the concentration required for EGCG to have a significant effect is fairly high. Additionally, the strategies for rational-design of such inhibitors are limited. The most common approach is to design inhibitors that tightly bind to monomeric proteins, sequester it from solution and preventing it from forming fibrils. However, many amyloid proteins are disordered in solution, including 248PAP286, prohibiting the possibility of structure-based design. Targeting fibrils are also extremely challenging. Despite our understanding of the atomic structure, fibrils are very stable species and it is difficult to imagine a design approach that could disaggregate preformed fibrils. Currently, high-throughput screening is perhaps the approach that has the most potential.

Here, I identified 9 inhibitors that facilitate the disaggregation of SEVI fibrils. In this assay, fibrils were added at 20  $\mu$ M corresponding monomer concentration and the inhibitor are active at 10  $\mu$ M, half of the amyloid peptide concentration. To my best knowledge, this is the lowest molar ratio of inhibitor to amyloid peptide that has been reported for dissolving preformed fibrils. Because SEVI fibrils are not stable at 0.1mg/ml concentration, it is possible that these molecules only facilitate dissociation of fibrils and are only active when the fibrils are present in a partial denaturing conditions. Therefore, it is difficult to conclude the effectiveness of these compounds from these preliminary results. Additionally, it would also be interesting to see if these compounds have cross reactivity for other amyloid proteins. We have learned that

SEVI fibrils only represent one of the many fibrils that potentially present in semen. Effective inhibitors should have broad-spectrum activities.

In this work, due to the limited peptide material, we had to use a fibril concentration that is lower than what we typically use in other *in vitro* assay. Because we observed self-depolymerization of the fibrils at this concentration, it allowed us to identify a set of small molecules that can stabilize fibrils. This class of compounds can potentially be important for treating amyloid diseases where oligomers are the toxic species. In fact a few studies have tested the approach, where compounds that stabilize amyloid  $\beta$  fibrils reduces the A $\beta$ 's toxicity in cell-based assay (Bieschke et al., 2012). It would be interesting to see if the compounds identified from our screening show similar activities.

## **Materials and Methods**

### **PAP248-286 peptide expression and purification**

PAP248-286 expression and purification have been detailed in chapter 1. Briefly, synthetic DNA (Invitrogen, Carlsbad, CA) has been PCR amplified with 5' SacI and 3' XhoI restriction sites. It is then cloned into a custom vector, a chimera constructed from the NdeI and XhoI digestion products pET15b (Novagen, Gibbstown, NJ), and the maltose binding protein (MBP) gene from pMALC2X (New England Biolabs, Ipswich, MA), resulting in an N-terminal His-tag MBP fusion vector. MBP-PAP248-286 fusion proteins are expressed in BL21-DE3(Gold) cells (Agilent Technologies, Santa Clara, CA) in LB miller media supplied with 100  $\mu$ g/mL ampicillin (Fisher Scientific, Pittsburgh, PA). Protein expression is induced with 0.5mM IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) once culture reached an OD600  $\sim$ 0.6. Strong

expression level is observed after 2-3 hours of induction at 34 °C.

All buffers used in this purification need to be pre-chilled, since the PAP248-286 peptide is largely disordered and degradation-prone. His-MBP- PAP248-286 fusion protein is first purified with a HisTrap-HP column, where the target protein elutes at 300 mM Imidazole. The elute sample is then dialyzed in a Slide-A-Lyzer 10,000 MWCO dialysis cassette (Pierce, Thermo Fisher Scientific, Rockford, IL), against buffer 50 mM Tris buffer pH 7.5 and 50 mM NaCl at 4 °C for at least 6 hours. The dialyzed sample was then pooled and 1/500 volume of TEV protease stock was added. TEV cleavage reaction was incubated again at 4 °C for overnight where close to 100% cleavage rate can be achieved. TEV protease reaction was then re-loaded onto the HisTrap-HP column. In this purification, buffer A is 20mM Tris pH7.5, 500mM NaCl and buffer B is 20mM Tris pH7.5, 500mM NaCl, 500mM Imidazole. Load protein sample and completely wash of the contaminants with buffer A and then elute PAP248-286 with 10% buffer B. PAP248-286 has two histadines which enables the peptide to bind to Ni-column with some affinity. This peak is then further purified using reverse phase high performance liquid chromatography (RP-HPLC) on a 2.2 x 25 cm Vydac 214TP101522 column equilibrated in buffer RA (0.1% trifluoroacetic acid (TFA)/water) and eluted over a linear gradient from 0% to 100% buffer RB (Acetonitrile/0.1% TFA) in 40 minutes at a flow rate of 9 mL/min. Absorbance at 220nm and 280nm were recorded using a Waters 2487 dual  $\lambda$  absorbance detector (Waters, Milford, MA). Peak fractions containing peptide were assessed for purity by a MALDI-TOF mass spectrometry (Voyager-DE-STR, Applied Biosystems, Carlsbad, CA). Pooled fractions were frozen in liquid nitrogen and lyophilized. Dried peptide powders were stored in desiccant jars at -20 °C.

## **EGCG Experiments**

SEVI fibrils were formed as 5 mg/ml in PBS with continuous shaking at 37°C for 24 hours. Fibril samples were then diluted with PBS to 1 mg/ml. Samples were added as 200 µl to a black, 96-well, optically clear plates (Nunc) (Thermo fisher Scientific, Waltham, MA). EGCG (Sigma Aldrich, St. Louis, MO) was dissolved in water at 20X molar excess and then added to SEVI fibrils as 5 times excess and 10X excess final concentrations. Thioflavin T (Sigma Aldrich, St. Louis, MO) stock solution is at 1mM in water and was added to the samples as 10 µM final concentration. Plate was sealed with Corning pressure-sensitive sealing tape. Fluorescent readings were taken with SpectraMax M5 microplate reader (Molecular Devices, Sunnyvale, CA) with excitation wavelength 440nm and emission wavelength 480nm at t=0 hours and t=24 hours. Reaction mixture was stored at ambient temperature during the 24 hour incubation.

## **Fibrillation Assay**

Dried peptide powders were dissolved as 5mg/ml in Phosphate Buffer Saline (Sigma Aldrich, St. Louis, MO) buffer in a sealed eppendorf tube and agitated at 37 °C, 900 rpm for 2 days (Torrey Pines Scientific, Carlsbad, CA). These samples were then diluted to 1mg/ml with PBS, continue the agitation for additional 24 hours. Fibril samples were stored at 4°C. For monomeric samples, lyophilized peptide were dissolved in PBS immediately before fluorescent reading was taken.

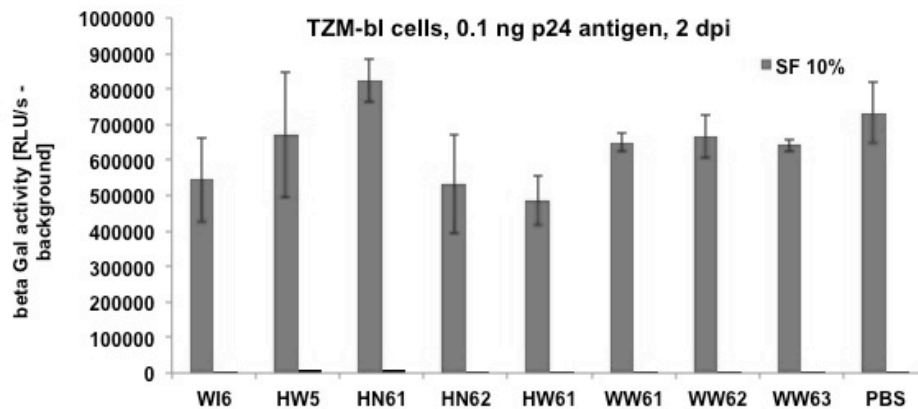
## High Throughput Assay

Thioflavin T (Sigma Aldrich, St. Louis, MO) stock solution is at 1mM (100x) in water. Glass-bottom small volume 384 well plates were ordered from Greiner Bio-one (Greiner Bio-one, Monroe, NC). Fibril samples were further diluted to 0.1 mg/ml with PBS shortly before samples were added. ThT stock solution was added as 100X dilution to the fibril samples and the monomer samples. Samples were mixed and incubated for 15 minutes for ThT to bind. 10  $\mu$ l of fibrils or monomers with ThT were then pipette into the 384 well plates. The plates were sealed and spanned at 3,000 rpm for 3 minute for sample to settle to the bottom of the plate. Test wells received compounds with use of a Biomek FX (Beckman Coulter, Brea, CA) with a 500 nl custom pin tool (V&P Scientific, San Diego, CA) at a target concentration of 10 $\mu$ M with a maximal DMSO concentration of 1%. DMSO was used as the solvent for compound addition. DMSO was also added to the positive and negative wells to the same final concentration. Fluorescence readings were taken with SpectraMax microplate reader (Molecular Devices, Sunnyvale, CA) with excitation wavelength 440nm and emission wavelength 480nm. Fluorescent readings were taken every hour for 16-48 hours.

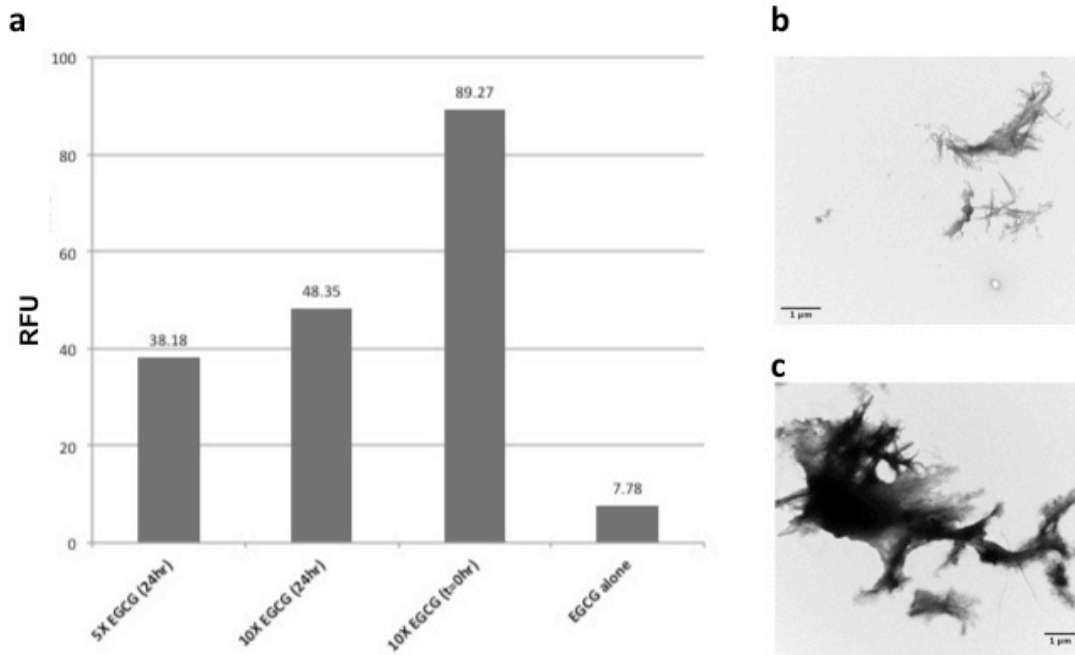
Chemical libraries were prepared by the Molecular Shared Screening Resource core facility at University of California, Los Angeles. Approximately 2,000 small molecules from multiple libraries were screened, including the Enzo Life Sciences bioactive compound library (bioactive lipids, endocannabinoids, ion channel ligands, kinase and phosphatase inhibitors, orphan receptor ligands, ~500 compounds), the Prestwick library (>1000 FDA-approved compounds), the Microsource Spectrum Collection (~2000 compounds), the ChemBridge DiverSet (~30,000 compounds), and others.

## Inhibitors are not active in seminal fluid

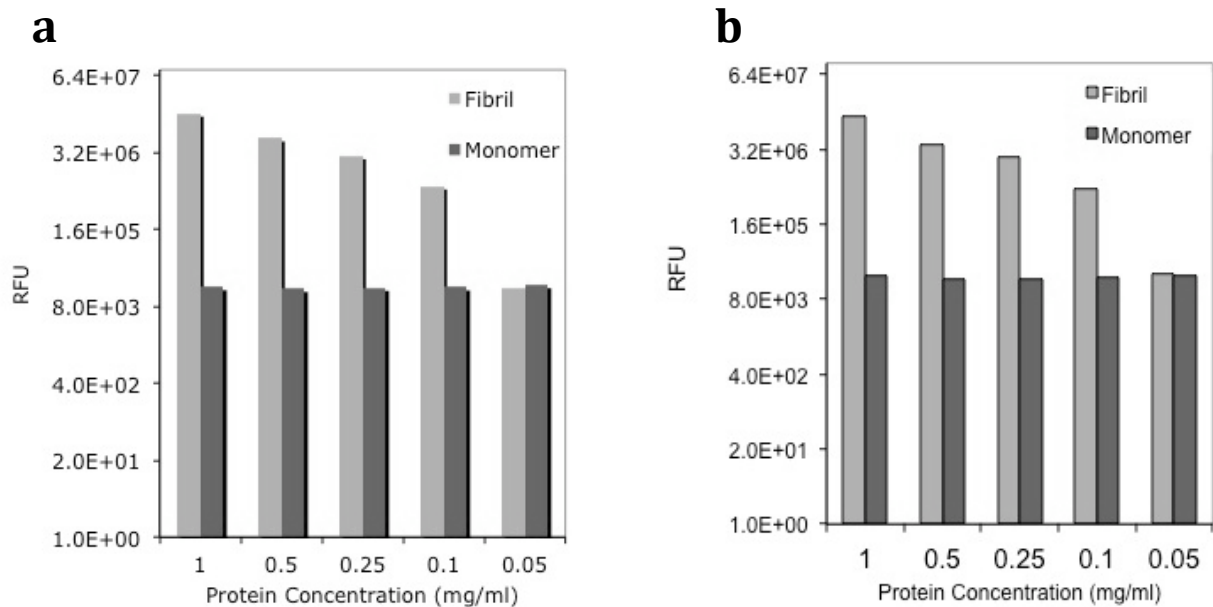
(Work done by Jan Munch, University of Ulm, Germany)



**Figure 1. Inhibitors are not active in seminal fluid.** 40  $\mu$ l HIV-1 containing 0.1 ng p24 antigen was incubated with 10% SF and inhibitory peptides that had been agitated for 23 hrs ( $t=23$ ). Thereafter, 20  $\mu$ l of virus/SF/inhibitor mixtures were used to infect 180  $\mu$ l TZM-bl cells seeded the day before ( $1 \times 10^5$ /well). 2 days later infection rates were determined by quantifying  $\beta$ -galactosidase activities in cellular lysates using the Gal-Screen assay. Peptide inhibitors have only mild effect in reducing HIV infection.

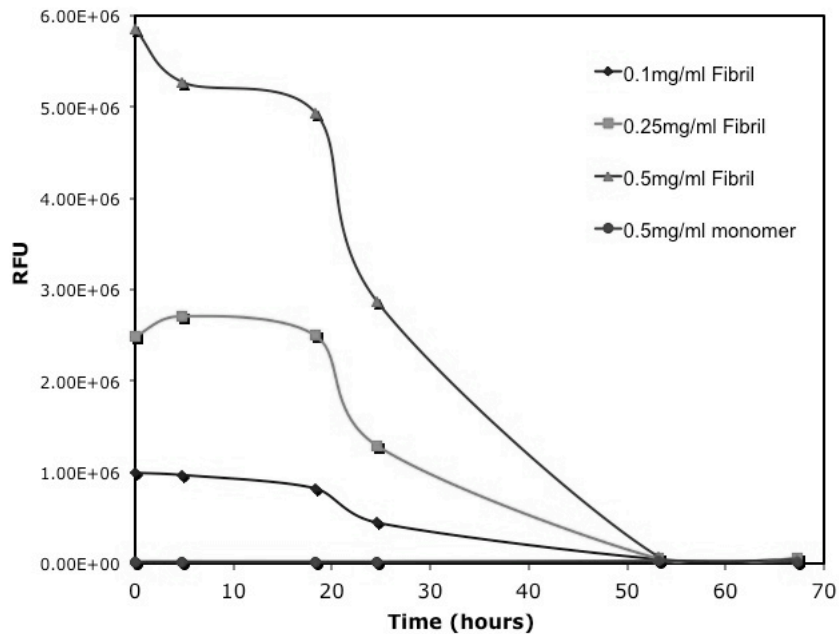


**Figure 2. EGCG is not effective in dissolving SEVI fibrils.** SEVI fibrils were formed as 5 mg/ml in PBS with continuous shaking at 37°C for 24 hours. Fibril samples were then diluted with PBS to 1 mg/ml. EGCG was dissolved in water and added to fibril samples as 5 times excess and 10X excess final concentrations. Fluorescent readings were taken after immediately adding EGCG and after 24 hour of incubation at room temperature. Fluorescent reading suggest that EGCG was able to reduce fluorescent reading both at 5X and 10X after 24 hour of incubation (**a**). However, under EM, fibrils were observed with SEVI only (**b**) and SEVI with 10x EGCG (**c**). The observed reduction in fluorescent readings could be due to EGCG disrupting the binding of ThT rather than disaggregating fibrils.

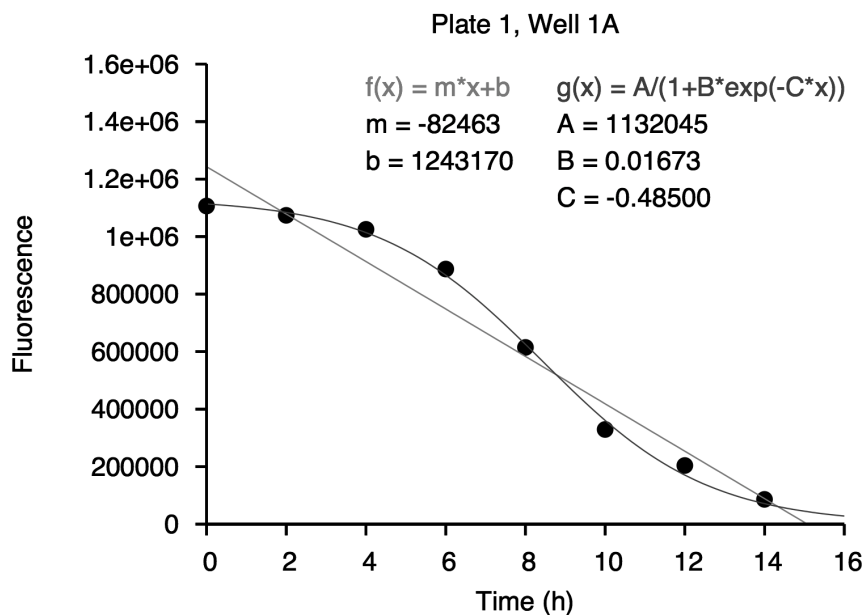


**Figure 3. ThT binds strongly to SEVI fibrils.** SEVI fibrils and monomers were added to 96-well clear bottom plate at the indicated concentrations. 10  $\mu$ M ThT was added and reading were taken after 10 minutes incubation. SEVI fibrils have strong affinity for ThT dye (**a**). Significant difference can be observed event at 0.1 mg/ml. Adding 5% DMSO does not effect ThT binding (**b**).



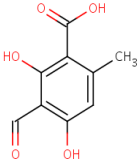
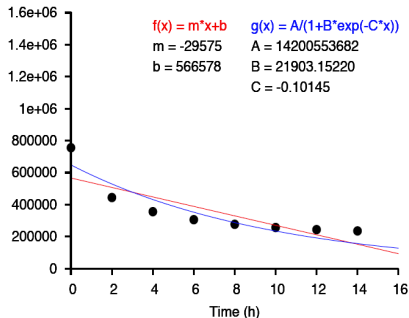
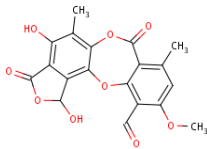
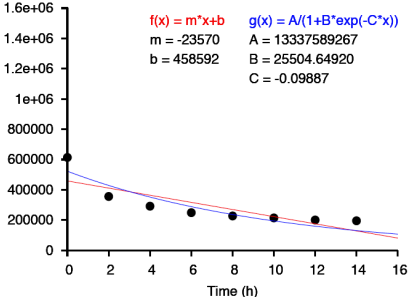
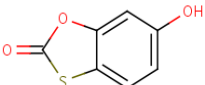
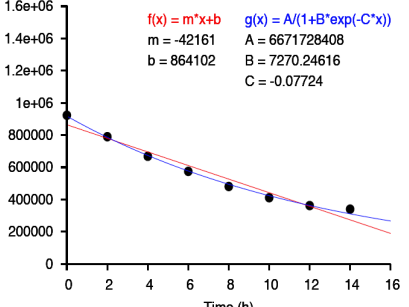
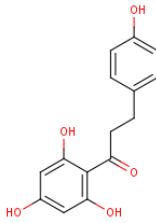
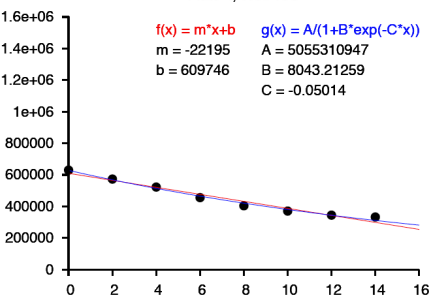


**Figure 4. Preliminary experiments with various SEVI concentrations.** 10  $\mu$ l of either SEVI fibrils or 248PAP826 monomers were added to 384-well glass bottom plate with 1% DMSO. Fluorescent readings were taken at 5-10 hour intervals for 52 hours. At the sub-mg/ml concentrations, SEVI fibrils appear to be unstable and gradually dissociate into monomers. Majority of the fibrils depolymerize after 24 hours incubation at room temperature.



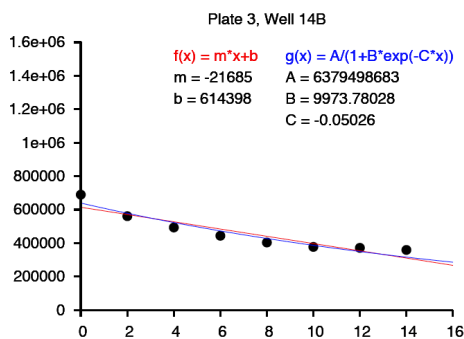
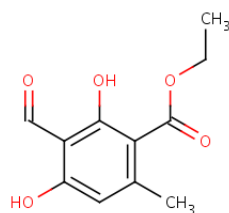
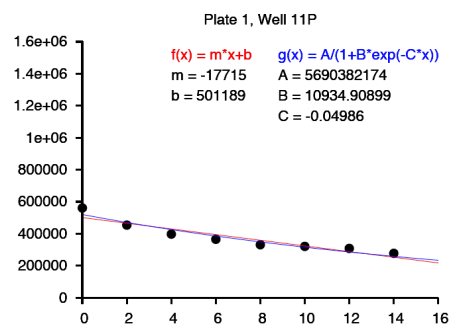
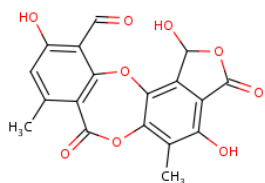
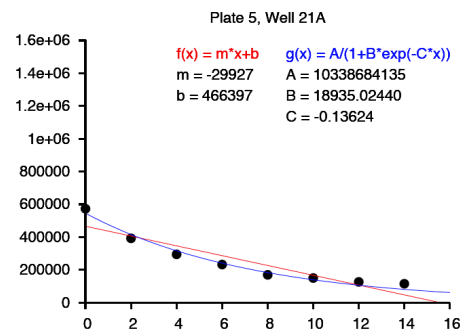
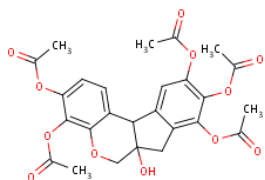
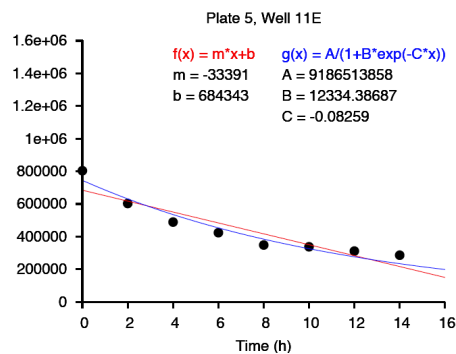
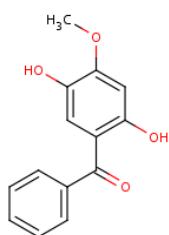
**Figure 5. SEVI only control.** 10  $\mu$ L of 0.1 mg/ml SEVI fibrils was added to 384-well glass bottom plate and incubated at room temperature for 16 hours. Fluorescent readings were taken every two hours. SEVI fibrils gradually depolymerize. The rate of depolymerization here is slightly faster than in the preliminary assays with majority of the fibrils are dissociated to monomers after 16 hours incubation, although the shapes of the best fitting curves were similar.

**Table 1. Small molecules that facilitates SEVI disaggregation**

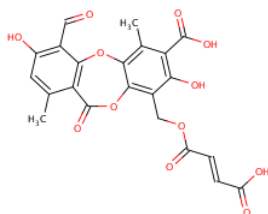
Destabilizing Compounds	Activities
 <p>Chemical structure of 2,4,6-trihydroxyacetophenone: <chem>CC(=O)c1c(O)c(O)c(O)c1</chem></p>	<p>Plate 4, Well 18O</p>  <p> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -29575</math>    <math>A = 14200553682</math>  <math>b = 566578</math>    <math>B = 21903.15220</math>  <math>C = -0.10145</math> </p>
 <p>Chemical structure of a complex polyphenolic compound: <chem>CC1=C(C(=O)OC2=C(O)C(=O)OC2O)C(O)C(=O)C1</chem></p>	<p>Plate 4, Well 11N</p>  <p> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -23570</math>    <math>A = 13337589267</math>  <math>b = 458592</math>    <math>B = 25504.64920</math>  <math>C = -0.09887</math> </p>
 <p>Chemical structure of a benzothiazole derivative: <chem>Oc1ccc2c(c1)sc(=O)o2</chem></p>	<p>Plate 4, Well 11M</p>  <p> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -42161</math>    <math>A = 6671728408</math>  <math>b = 864102</math>    <math>B = 7270.24616</math>  <math>C = -0.07724</math> </p>
 <p>Chemical structure of a polyphenolic compound with a hydroxyethyl side chain: <chem>Oc1ccc(O)cc1C(=O)CC(O)c2c(O)c(O)c(O)c2</chem></p>	<p>Plate 4, Well 13O</p>  <p> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -22195</math>    <math>A = 5055310947</math>  <math>b = 609746</math>    <math>B = 8043.21259</math>  <math>C = -0.05014</math> </p>

## Destabilizing Compounds

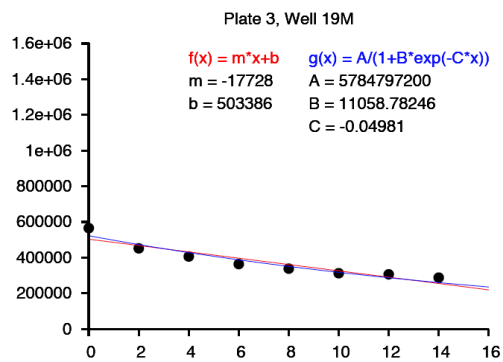
## Activities



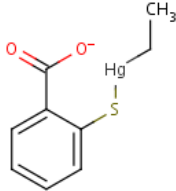
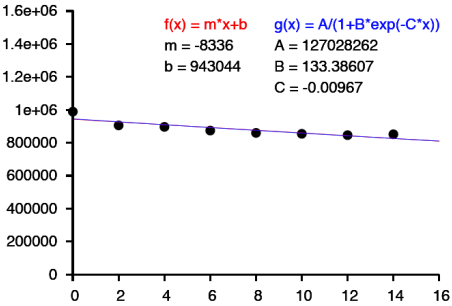
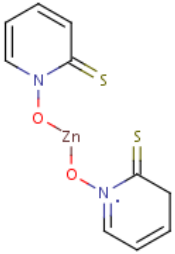
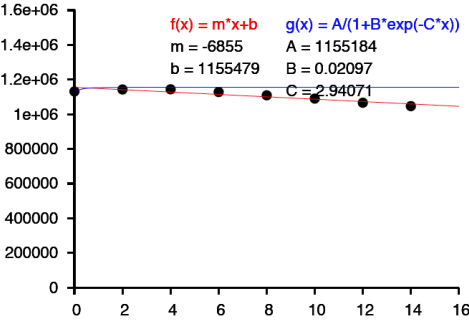
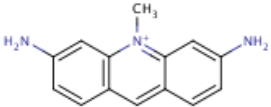
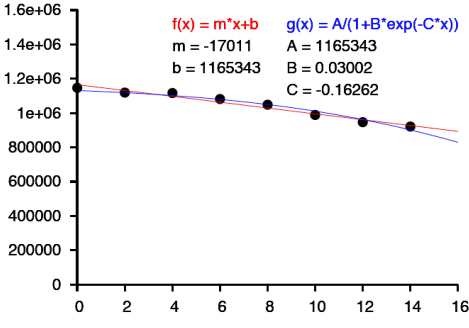
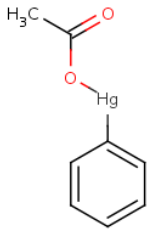
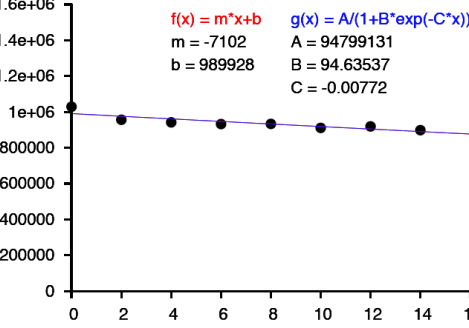
## Destabilizing Compounds



## Activities



**Table 2. Small molecules that stabilize SEVI fibrils**

Stabilizing Compounds	Activities
<p data-bbox="284 506 321 531">Na<sup>+</sup></p> 	<p data-bbox="889 363 1003 384">Plate 2, Well 6D</p>  <p data-bbox="857 401 1133 478"> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -8336</math>    <math>A = 127028262</math>  <math>b = 943044</math>    <math>B = 133.38607</math>  <math>C = -0.00967</math> </p>
	<p data-bbox="898 741 1011 762">Plate 2, Well 8A</p>  <p data-bbox="857 779 1149 856"> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -6855</math>    <math>A = 1155184</math>  <math>b = 1155479</math>    <math>B = 0.02097</math>  <math>C = 2.94071</math> </p>
<p data-bbox="256 1276 289 1302">Cl<sup>-</sup></p> 	<p data-bbox="898 1140 1011 1161">Plate 2, Well 18H</p>  <p data-bbox="857 1178 1149 1255"> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -17011</math>    <math>A = 1165343</math>  <math>b = 1165343</math>    <math>B = 0.03002</math>  <math>C = -0.16262</math> </p>
	<p data-bbox="898 1518 1011 1539">Plate 3, Well 20K</p>  <p data-bbox="857 1556 1149 1633"> <math>f(x) = m \cdot x + b</math>    <math>g(x) = \frac{A}{1+B \cdot \exp(-C \cdot x)}</math>  <math>m = -7102</math>    <math>A = 94799131</math>  <math>b = 989928</math>    <math>B = 94.63537</math>  <math>C = -0.00772</math> </p>

## Stabilizing Compounds

## Activities

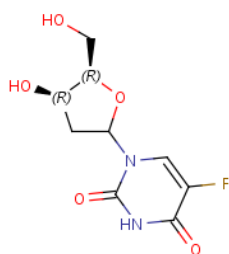


Plate 3, Well 22N

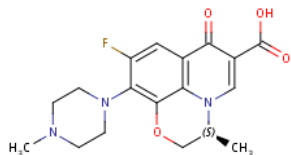
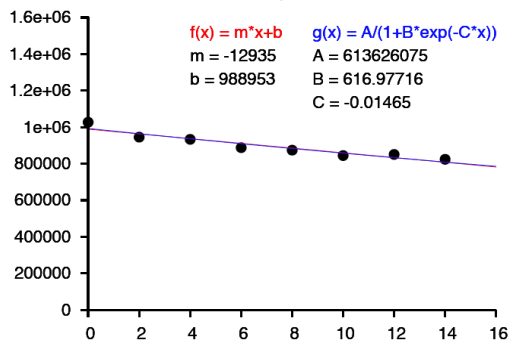


Plate 4, Well 11G

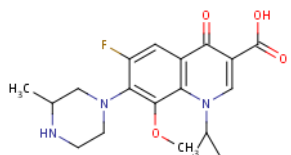
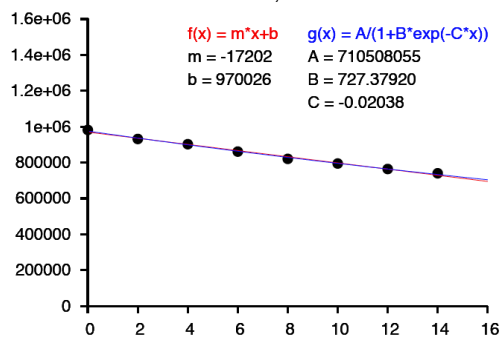


Plate 4, Well 19G

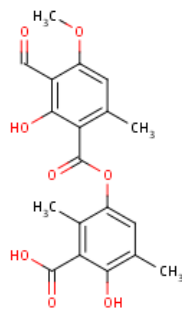
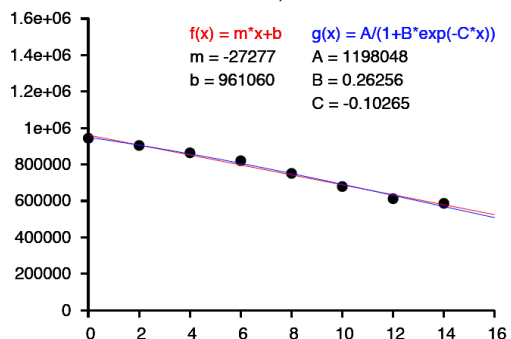
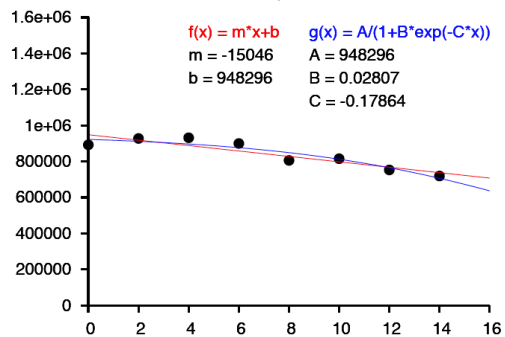
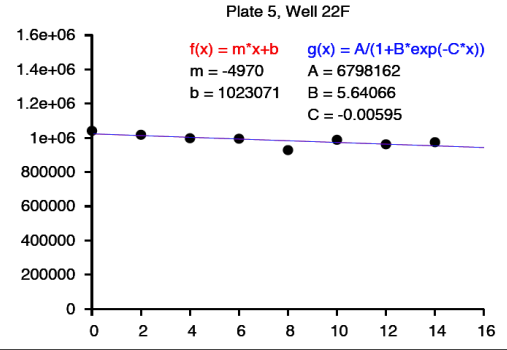
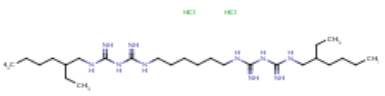
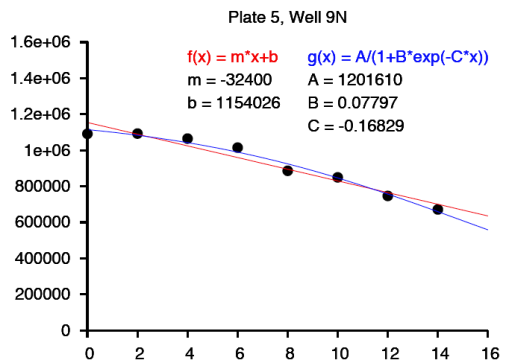
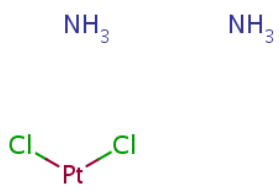


Plate 5, Well 15A



**Stabilizing Compounds**

**Activities**





## References

Arnold, F., Schnell, J., Zirafi, O., Sturzel, C., Meier, C., Weil, T., Standker, L., Forssmann, W.G., Roan, N.R., Greene, W.C., *et al.* (2012). Naturally occurring fragments from two distinct regions of the prostatic acid phosphatase form amyloidogenic enhancers of HIV infection. *Journal of virology* 86, 1244-1249.

Bieschke, J., Herbst, M., Wiglenda, T., Friedrich, R.P., Boeddrich, A., Schiele, F., Kleckers, D., Lopez del Amo, J.M., Gruning, B.A., Wang, Q., *et al.* (2012). Small-molecule conversion of toxic oligomers to nontoxic beta-sheet-rich amyloid fibrils. *Nature chemical biology* 8, 93-101.

Bieschke, J., Russ, J., Friedrich, R.P., Ehrnhoefer, D.E., Wobst, H., Neugebauer, K., and Wanker, E.E. (2010). EGCG remodels mature alpha-synuclein and amyloid-beta fibrils and reduces cellular toxicity. *Proceedings of the National Academy of Sciences of the United States of America* 107, 7710-7715.

Bush, A.I. (2003). Copper, zinc, and the metallobiology of Alzheimer disease. *Alzheimer disease and associated disorders* 17, 147-150.

Ehrnhoefer, D.E., Bieschke, J., Boeddrich, A., Herbst, M., Masino, L., Lurz, R., Engemann, S., Pastore, A., and Wanker, E.E. (2008). EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nature structural & molecular biology* 15, 558-566.

Haass, C., and Selkoe, D.J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nature reviews Molecular cell biology* 8, 101-112.

Kardos, J., Micsonai, A., Pal-Gabor, H., Petrik, E., Graf, L., Kovacs, J., Lee, Y.H., Naiki, H., and Goto, Y. (2011). Reversible heat-induced dissociation of beta2-microglobulin amyloid fibrils. *Biochemistry* 50, 3211-3220.

Kim, K.A., Yolamanova, M., Zirafi, O., Roan, N.R., Staendker, L., Forssmann, W.G., Burgener, A., Dejuq-Rainsford, N., Hahn, B.H., Shaw, G.M., *et al.* (2010). Semen-mediated enhancement of HIV infection is donor-dependent and correlates with the levels of SEVI. *Retrovirology* 7, 55.

Klunk, W.E., Debnath, M.L., Koros, A.M., and Pettegrew, J.W. (1998). Chrysamine-G, a lipophilic analogue of Congo red, inhibits A beta-induced toxicity in PC12 cells. *Life sciences* 63, 1807-1814.

Kroth, H., Ansaloni, A., Varisco, Y., Jan, A., Sreenivasachary, N., Rezaei-Ghaleh, N., Giriens, V., Lohmann, S., Lopez-Deber, M.P., Adolfsson, O., *et al.* (2012). Discovery and structure activity relationship of small molecule inhibitors of toxic beta-amyloid-42 fibril formation. *The Journal of biological chemistry* 287, 34786-34800.

Ladiwala, A.R., Dordick, J.S., and Tessier, P.M. (2011). Aromatic small molecules remodel toxic soluble oligomers of amyloid beta through three independent pathways. *The Journal of biological chemistry* 286, 3209-3218.

Ladiwala, A.R., Lin, J.C., Bale, S.S., Marcelino-Cruz, A.M., Bhattacharya, M., Dordick, J.S., and Tessier, P.M. (2010). Resveratrol selectively remodels soluble oligomers and fibrils of amyloid Abeta into off-pathway conformers. *The Journal of biological chemistry* 285, 24228-24237.

Landau, M., Sawaya, M.R., Faull, K.F., Laganowsky, A., Jiang, L., Sievers, S.A., Liu, J., Barrio, J.R., and Eisenberg, D. (2011). Towards a pharmacophore for amyloid. *PLoS biology* 9, e1001080.

Munch, J., Rucker, E., Standker, L., Adermann, K., Goffinet, C., Schindler, M., Wildum, S., Chinnadurai, R., Rajan, D., Specht, A., *et al.* (2007). Semen-derived amyloid fibrils drastically enhance HIV infection. *Cell* 131, 1059-1071.

Ramassamy, C. (2006). Emerging role of polyphenolic compounds in the treatment of neurodegenerative diseases: a review of their intracellular targets. *European journal of pharmacology* 545, 51-64.

Roan, N.R., Muller, J.A., Liu, H., Chu, S., Arnold, F., Sturzel, C.M., Walther, P., Dong, M., Witkowska, H.E., Kirchhoff, F., *et al.* (2011). Peptides released by physiological cleavage of semen coagulum proteins form amyloids that enhance HIV infection. *Cell host & microbe* 10, 541-550.

Roan, N.R., Munch, J., Arhel, N., Mothes, W., Neidleman, J., Kobayashi, A., Smith-McCune, K., Kirchhoff, F., and Greene, W.C. (2009). The cationic properties of SEVI underlie its ability to enhance human immunodeficiency virus infection. *Journal of virology* 83, 73-80.

Shammas, S.L., Knowles, T.P., Baldwin, A.J., Macphee, C.E., Welland, M.E., Dobson, C.M., and Devlin, G.L. (2011). Perturbation of the stability of amyloid fibrils through alteration of electrostatic interactions. *Biophysical journal* 100, 2783-2791.

Sheftic, S.R., Snell, J.M., Jha, S., and Alexandrescu, A.T. (2012). Inhibition of semen-derived enhancer of virus infection (SEVI) fibrillogenesis by zinc and copper. *European biophysics journal : EBJ* 41, 695-704.

Sievers, S.A., Karanicolas, J., Chang, H.W., Zhao, A., Jiang, L., Zirafi, O., Stevens, J.T., Munch, J., Baker, D., and Eisenberg, D. (2011). Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* 475, 96-100.