Doctoral Dissertation

Studies on Functional regulation by Biomaterial interaction using an integrated *in silico-in vitro* approach

Radhika BIYANI

Supervisor: Masahiro TAKAGI

Graduate School of Advanced Science and Technology,
Japan Advanced Institute of Science and Technology, Japan
[Materials Science]

September 2022

ABSTRACT

Studies on Functional regulation by Biomaterial interaction using an integrated in silico-in vitro approach

The knowledge of regulating the function of proteins or peptides directed by the biomaterial-interaction based on the structural diversity and conformational dynamics by integrating both *in vitro* laboratory-based experiments and *in silico* computational-based molecular docking experiments has been studied.

The thesis is organized into 4 major chapters:

The first chapter is based on the interaction between ***protein*** and ***small molecule***. It is titled *"Identification and evaluation of small molecule-based natural compounds for inhibition of HIV-1 Reverse Transcriptase activity"*. This chapter talks about the use of small-molecule based natural-compounds that are derived from plants, as biomaterials that interact and regulate the function of reverse transcriptase (RT) enzyme which is responsible for causing HIV-1 (Human immunodeficiency virus-1) disease. For this purpose, few Indian plants (not native but abundantly grown as weeds in India) have been chosen and screened for identification of inhibitory activity to RT enzyme from HIV Type-1 using different Reverse Transcription Assays. The results suggest that water-extracts of leaves of *Argemone mexicana* plant strongly inhibited the DNA polymerase activity of HIV-1 RT, indicating they contain organic compounds that inhibit the enzyme activity. This study thus leads to the understanding of how small molecules based natural compounds as biomaterials can interact with the functional proteins and inhibit their activity.

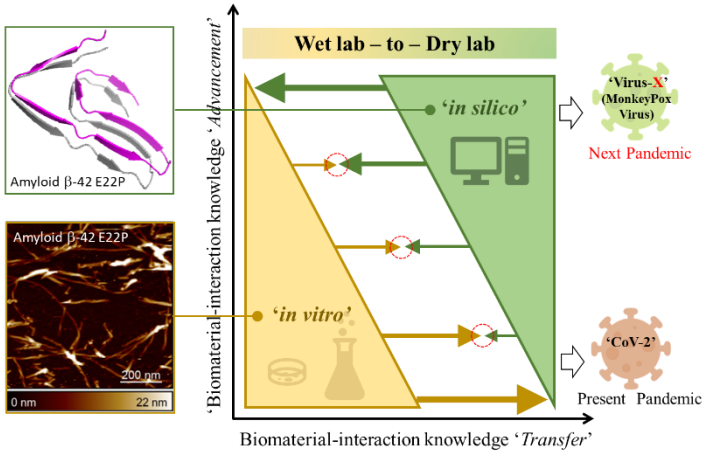
The second chapter is based on the interaction between ***protein*** and ***nucleic acid***. It is titled *"Co-operative interaction between RNA polymerase and Recombinase for Reverse Transcriptase-based improved isothermal nucleic acid amplification assay"*. This chapter talks about the bio-interaction of different functional proteins that can co-assist and interact with each other and help to develop an improved diagnostic assay for infectious diseases. By understanding the biomolecular interaction of RNA polymerase enzyme and Recombinase enzyme, a new method has been developed, which is termed as RICCA (RNA isothermal co-assisted coupled amplification). We integrated the essentials of both types of isothermal amplification methods, i.e. RNA specific amplification and RPA (recombinase polymerase amplification) into a simple format of 'sample-in and answer-out' with a primary focus on the detection of low copy numbers of viral RNA directly from COVID-19 saliva samples without the need for any laboratory handling or sample preprocessing. We report the development of a completely homogeneous, isothermal, highly sensitive, and ultrarapid method for detecting virus RNA target sequences for the on-site (low resource settings) molecular diagnosis of COVID-19 and other infectious diseases. We further plan to advance this study by utilizing the *in silico* approach for optimization of primer designing that can be further utilized for the application of RICCA for new variants.

The third chapter is based on the interaction between ***peptide*** and ***peptide***. It is titled *"Understanding biophysical properties of oligomerization and fibril formation of amyloid beta 42 conformers"*. This chapter deals with the diversity of molecular conformation of small peptide, amyloid beta 42 (Aβ42), which is caused by a single amino-acid substitution in the peptide chain. This causes one conformer to be more toxic, i.e. E22P-Aβ42 (mutation at 22nd position of amino acid chain that changes Glutamic acid (E) to proline(P)) and the other one to be less-toxic i.e.E22- Aβ42 (wild-type). To understand how a single amino acid change in the sequence can lead to several folds increased toxicity and increased aggregation of the peptide leading to increased risk to disease, macromolecular analysis was made using SDS-PAGE followed by making a microscopic analysis of the aggregation dynamics of both the conformers using atomic force microscopy (AFM) studies. Our results disclose the formation of amorphous aggregates in E22P-Aβ42 that are stem-based network-like structures while formation of mature fibrils in E22-Aβ42 that are sphere-based flexible structures. A relative comparison is made between the biophysical properties of E22P-Aβ42 and E22-Aβ42 that reveals high stiffness, lesser periodicity, and higher rigidity in E22P-Aβ42. *In silico* studies were performed by molecular docking that revealed atomic scale details like number of beta sheets, and number of residues in beta sheets involved, and the dihedral angle between beta sheets involved in the formation of E22-Aβ42 and E22P-Aβ42. We propose a systematic model of fibril formation that helps in understanding the molecular basis of conformational

transitions in the Aβ42 species. These findings will have significant implications to our understanding of the structural basis of toxicity caused by conformational diversity in Aβ42 species.

The fourth chapter is based on the interaction between ***peptide*** and ***nucleic acid***. It is titled *"Identification and evaluation of aptamer-based synthetic compounds for regulation of toxic conformer of amyloid beta 42"*. Based on the understanding of Chapter 3, this chapter was designed to identify aptamer-based synthetic compounds as biomaterials for specific recognition of both the conformers of Aβ42 using an advanced screening method. We identified novel 70-nt DNA aptamer sequences using in vitro competitive selection method SELCOS. Utilizing this advanced screening approach helps in identification of unique aptamer sequences that can specifically recognize their respective targets and not the competitive target. Our results revealed that the selected aptamer candidates (Apt-W1 for E22-Aβ42 and Apt-T2 for E22P-Aβ42) show high binding affinity to their respective targets. Further, their binding was confirmed by electrochemical evaluation, SDS PAGE and qPCR analysis. The *in silico* studies support the supposition that the number of beta sheets decrease on interaction of aptamer with Aβ42 indicating the inhibitory function of selected aptamer towards aggregation of Aβ42. We believe this work could be a promising research tool for further studies about toxicity caused by E22P-Aβ42, and thereafter regulate the toxicity.

From these studies in chapter 1-4, we can thus draw a general understanding of how integrating the *in vitro* and *in silico* studies can help in a better understanding of the biological systems. We therefore give a 'dry-lab concept', as shown in below figure to develop a platform that can 'integrate' the power of *in vitro* and *in silico* systems to regulate the function of proteins with the help of structure and biomaterial-interaction. This integration can help to gradually reduce the dependency of *in vitro* system on the *in silico* system and develop a



fully functional in silico platform for the future.

The DRY-Lab concept

KEYWORDS: biomaterial interaction, *in silico-in vitro* approach, aptamer, amyloid beta, reverse transcriptase

TABLE OF CONTENTS

GENERAL INTRODUCTION

Understanding interactions between biomaterials

We often think of proteins as nutrients in the food we eat or the main component of muscles or the building blocks of life, but proteins are also the microscopic molecules inside of cells that perform diverse, and the most vital jobs required for the structure, function, and regulation of the body's tissues and organs. Proteins are heteropolymers of amino acids, with the amino acids linked in a chain by peptide bonds and for a protein to be functional, its polypeptide chain needs to be 'properly' folded into its native conformation. Nearly all the information on how to reach the native fold is present in the amino acid sequence itself, in the nature of the side chains and the order in which they are linked. To quote Anfinsen from his Nobel laureate lecture in 1972 on the folding of proteins: *"The native conformation is determined by the totality of interatomic interactions and hence by the amino acid sequence, in a given environment."* Therefore, the intramolecular and intermolecular interaction of/between proteins play a crucial role in the behavior and function of proteins and can be categorized into three major classes: (i) proteins and small molecules interaction, e.g., allosteric binding between hemoglobin and oxygen leads to the formation of oxyhemoglobin that can transport oxygen to different parts of the body; (ii) protein and nucleic acids interaction, e.g., the process of transcription involves RNA polymerase and DNA that can synthesis RNA copies; (iii) proteins and proteins interactions, e.g., the process of signal transduction where the different proteins like kinase and phosphatase interact to regulate the process.

Due to their central role in biological function, protein-interactions are vital to control the mechanisms leading to healthy and diseased states in living beings. Many diseases are often caused by mutations or changes affecting the binding interface or leading to biochemically dysfunctional allosteric changes in proteins. These changes because of protein interactions can cause protein malfunctions like over expression or absence of protein expression, misfolding of proteins, aggregation of proteins, loss of function of proteins, over-production,

or less-production of proteins/ functional proteins, etc. For e.g., most of the neurodegenerative disorders like Alzheimer's disease, Parkinson's disease, Huntington's disease, Creutzfeldt-Jakob disease, cystic fibrosis, Gaucher's disease and many others are caused by simply misfolding of proteins due to slight changes in their interactions. Therefore, protein interaction networks can elucidate the molecular basis of disease, which in turn can inform methods for prevention, diagnosis, and treatment.

In this stream, 'Biomaterial Science' is setting new standards of emerging technologies for various diseases and their related research. Biomaterials science, is a new, developing branch of science that encompasses new concepts in biomaterials design, studies into the interaction of biomaterials with the body, and the use of biomaterials to answer fundamental biological questions. Exploring the potential of biomaterials and their interaction with the biological system has particularly become important after the COVID-19 pandemic because biomaterials can offer a range of possibilities to develop disease models, protective, diagnostic, therapeutic and monitoring measures.

Biomaterial, 'the smart molecule', is defined as any material that can be either synthetic or natural in origin, and that has been engineered to take a form, that can alone or as a part of complex system, be used to direct, by control of interactions with components of the biological systems, to augment any (protein) function of the body, in order to maintain or improve the quality of life of the individual. When a biomaterial is placed in the human body, different types of molecular interactions take place at the biomaterial surface. These different types of interactions are affected by the surface properties of the biomaterial such as surface chemistry, wettability, and surface roughness. It is at first required that any kind of biomaterial must be 'biocompatible' with the biological system and in case of natural biomaterials, they inherently show biomimetism because of their natural presence. Natural or naturally derived biomaterials have been demonstrated to show several advantages over synthetic biomaterials like biocompatibility, biodegradability, and remodeling. Therefore, the

use of natural or naturally derived biomaterials are the preferred choices for diseases and related research.

The development of biomaterials has shifted from the development of simple, biocompatible, and naturally derived biomaterials to advanced biomaterials with functions to mimic biological systems because of their limited diversity. To develop these advanced materials, biomaterials are often functionalized as ligands for cell attachment or protein binding. One such smart class of next-generation synthetic biomaterials can be 'biomolecules as biomaterials'. Using biomolecules like small single-stranded DNA as biomaterials can have mixed properties of both natural, and synthetic biomaterials which can function as ligands and can specifically and selectively recognize their respective target protein and most importantly, regulate their function.

Significance of biomaterials interaction-based protein function regulation

Before understanding the exact molecular mechanism of protein function regulation based on biomaterial interactions, it is critically important to study and understand the dynamic behavior of proteins which is governed by the existing space of structural diversity and conformational dynamics of proteins and/or biomaterials. To investigate this, a detailed knowledge of the molecular conformations for the structure- and function-based knowledge is pre-requisites not only at macroscopic and microscopic levels but also at atomic levels.

One possible solution to attain this to utilize the power of computers. With the advent of digitalization, the computers and the integrated application of IoT have turned into a basic necessity in our lives and society. In the present COVID-19 pandemic, we have realized the power of computers towards a transformation from physical world to virtual world. The history of computing is not only evolution, but also revolution and we are moving towards the Artificial Intelligence era. Based on this, we can say that the current need is advancement in this field because we need to prepare ourselves for the future. When we talk about the scientific world, recent advancements in the computer-based technologies have made it

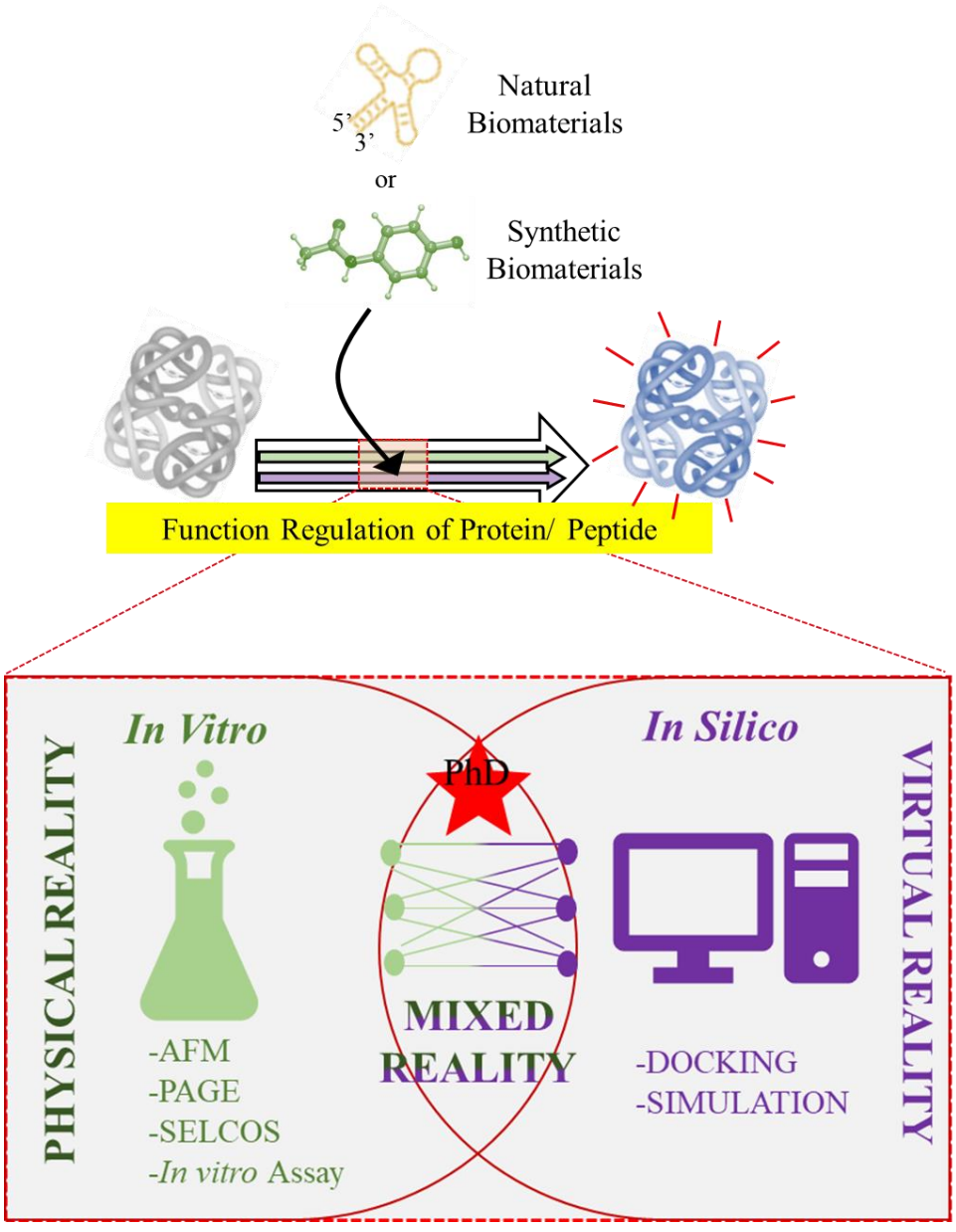
possible to predict the atomic interactions that take place between biomolecules. It has now become feasible to create models that can simulate the molecule behavior based on our physical understandings. The conformations, dynamics and properties of biomolecules can now be predicted theoretically. However, all this is possible along with the experimental approaches and not independently. In other words, the *in silico* experiments today are merely used to understand and/or validate the information that is generated by *in vitro* studies because in silico system is dependent on the *in vitro* system. At present, all the major understandings of the biological functions are derived from the laboratory based 'wet' experiments. All the wet-information (information acquired by in-vitro methods) is then input into the 'dry' system (computer-based) and we get some information that adds on to our understanding.

To prepare a fully functional platform for *in silico* system, it is required to integrate the *in silico* system with the *in vitro* system. Development of a platform that can co-assist and couple the *in silico* power of data-driven modelling technologies and the primary molecular information generated by *in vitro* technologies can speed up the progress in this field and in future, shift the research focus on *in silico* platform towards re-creating the new information with relaxed dependency (gradually) on *in vitro* system.

This study:

To address the above-mentioned issues, this doctoral study specially designed to explore the knowledge of regulating the function of proteins or peptides directed by the biomaterial-interaction based on the structural diversity and conformational dynamics using both *in vitro* laboratory-based experiments and *in silico* computational-based molecular docking experiments. A schematic drawing is shown in **Box 1**. Three important key words form the basis of this thesis include ***Biomaterial Interaction, Functional Regulation*** and ***Integrated in silico-in vitro approach*.**

Box 1: Schematics of 'Study on Structure and Biomaterial interaction-based Functional Regulation of Proteins and Peptides using an integrated in silico-in vitro approach'



GUIDE TO THESIS

My research work in the doctoral course has led me down to organize my doctoral dissertation into 5 chapters. All the chapters are products of individual (but related) studies and made significant enough to include all the information that is needed to read them independently.

Chapter 1 is based on the interaction between ***protein*** and ***small molecule***. It is titled *"Identification and evaluation of small molecule-based natural compounds for inhibition of HIV-1 Reverse Transcriptase activity"*. This chapter talks about the use of small-molecule based natural-compounds as biomaterials that interacts and regulate the function of reverse transcriptase (RT) enzyme which is responsible for causing HIV-1 (Human immunodeficiency virus-1) disease. The concept of 'Ayurveda', which is a commonly practiced Indian traditional natural system of medication has been utilized in this chapter as an alternative to the existing HIV-1 disease therapies. With the idea of making available a natural medication for treating diseases like HIV-AIDS, few Indian plants (not native but abundantly grown as weeds in India) were chosen and screened for identification of inhibitory activity to RT enzyme from HIV Type-1 using different Reverse Transcription Assays by the fluorescent dye PicoGreen and [^{3}H]-dTTP. Our results suggest that water-extracts of leaves of *Argemone mexicana* strongly inhibited the DNA polymerase activity of HIV-1 RT, indicating they contain substances that inhibit the enzyme activity. Neither thermal treatment at 100°C nor proteinase K treatment of the extracts abolished the inhibitory activity, suggesting the inhibitory substance is an organic compound rather than a protein. This study thus leads to the understanding of how small molecules based natural compounds as biomaterials can interact with the functional proteins and inhibit their activity. We further plan to advance this study by *in silico* approach to understand the molecular mechanism of inhibition that can further help in advancing towards drug development.

The results produced in this chapter has been published in Journal of Biological Macromolecules (Publisher: Elsevier) as Radhika Biyani et al. "Inhibition of HIV-1 Reverse Transcriptase Activity by the Extracts of Indian Plants." Journal of Biological Macromolecules 20, no. 1 (2020): 17-22.

Chapter 2 is based on the interaction between ***protein*** and ***nucleic acid***. It is titled *"Co-operative interaction between RNA polymerase and Recombinase for Reverse Transcriptase-based improved isothermal nucleic acid amplification assay"*. This chapter talks about the bio-interaction of different functional proteins that can co-assist and interact with each other and help to develop an improved diagnostic assay for infectious diseases. By understanding the biomolecular interaction of RNA polymerase enzyme and Recombinase enzyme, a new method has been developed, which is termed as RICCA (RNA isothermal co-assisted coupled amplification). In this study, we aimed to advance the bioengineering of the earlier studies performed on RNA specific amplification and RPA (recombinase polymerase amplification), and then integrate the essentials of both types of isothermal amplification methods into a simple format of 'sample-in and answer-out' with a primary focus on the detection of low copy numbers of viral RNA directly from COVID-19 saliva samples without the need for any laboratory handling or sample preprocessing. In this regard, we report the development of a completely homogeneous, isothermal, highly sensitive, and ultrarapid method for detecting virus RNA target sequences for the on-site (low resource settings) molecular diagnosis of COVID-19 and other infectious diseases. We further plan to advance this study by utilizing the *in silico* approach for optimization of primer designing that can be further utilized for the application of RICCA for new variants.

The results produced in this chapter has been published in Scientific Reports (Publisher: Nature Portfolio) as Radhika Biyani et al. "Development of robust isothermal RNA amplification assay for lab-free testing of RNA viruses." Scientific Reports 2021: 11, 1, 1-13.

This chapter also contributed to make patent applications as follows: Japanese patent app. 2020-093786. International Patent app. PCT/JP2021/018799

Chapter 3 is based on the interaction between ***peptide*** and ***peptide***. It is titled *"Understanding biophysical properties of oligomerization and fibril formation of amyloid beta 42 conformers"*. This chapter deals with the diversity of molecular conformation of small peptide, amyloid beta 42 (Aβ42), which is caused by a single amino-acid substitution in the peptide chain. This causes one conformer to be more toxic, i.e. E22P- Aβ42 (mutation at 22nd position of amino acid chain that changes Glutamic acid (E) to proline(P)) and the other one to be less-toxic i.e. E22- Aβ42 (wild-type). This chapter was designed to make a deeper

understanding of how a single amino acid change in the sequence can lead to several folds increased toxicity and increased aggregation of the peptide leading to increased risk to disease. Macromolecular analysis was made using SDS-PAGE followed by making a deeper analysis of the aggregation dynamics of both the conformers using atomic force microscopy (AFM) studies. Our results disclose the formation of amorphous aggregates in E22P-Aβ42 that are stem-based network-like structures while formation of mature fibrils take place in the less-toxic conformer of Aβ42, E22-Aβ42 that are sphere-based flexible structures. A relative comparison is made between the biophysical properties of E22P-Aβ42 and E22-Aβ42 that reveals high stiffness, lesser periodicity, and higher rigidity in E22P- Aβ42. *In silico* studies were performed by molecular docking that revealed atomic scale details like number of beta sheets, and number of residues in beta sheets involved in the formation of E22-Aβ42 and E22P-Aβ42. A characteristic difference in the dihedral angle is also observed by molecular docking that forms the basis of aggregation in both the variants. These findings will have significant implications to our understanding of the structural basis of toxicity caused by conformational diversity in Aβ42 species. In the end, we propose a systematic model of fibril formation that helps in understanding the molecular basis of conformational transitions in the Aβ42 species.

Chapter 4, is based on the interaction between ***peptide*** and ***nucleic acid***. It is titled *"Identification and evaluation of aptamer-based synthetic compounds for regulation of toxic conformer of amyloid beta 42"*. Based on the understanding of Chapter 3, this chapter was designed to identify aptamer-based synthetic compounds as biomaterials for specific recognition of both the conformers of Aβ42 using an advanced screening method. Forces that govern the recognition, binding and specific interaction force aptamer-target protein were studied.

In this chapter, firstly, we identified a novel 70-nt DNA aptamer sequences using in vitro competitive selection method SELCOS, in which E22P-Aβ42 is taken as the positive target and E22-Aβ42 is the competitive target protein. Utilizing this advanced screening approach helps in identification of unique aptamer sequences that can specifically recognize their respective targets and not the competitive target. Our results revealed that the selected aptamer candidates show high binding affinity to their respective targets. Further, their binding was confirmed by electrochemical evaluation, SDS PAGE and qPCR analysis. The in

silico studies support the supposition that the number of beta sheets decrease on interaction of aptamer with Aβ42 indicating the inhibitory function of selected aptamer towards aggregation of Aβ42. We believe this work could be a promising research tool for further studies about toxicity caused by E22P-Aβ42.

Chapter 5, is titled *"General conclusion: A novel 'dry-lab' concept for biomaterial interaction-based protein function regulation"*. With the current trend being to utilize the in-silico experiment-based results to merely validate and/or understand the information generated by in vitro-based studies, this chapter highlights the idea of how *in silico* studies can be incorporated and integrated into/along with the *in vitro* studies. On the basis of the insights made in the succeeding chapters (especially Chapter 3), it is evident that using the information generated by both the *in silico* studies with *in vitro* studies can lead to a better understanding of the system. Therefore, this chapter gives a concept to develop a platform which can co-assist and couple the *in silico* power of data-driven modelling technologies and the primary molecular information generated by *in vitro* technologies that can speed up the progress in this field. The idea is to shift the research focus on *in silico* platform towards re-creating the new information with relaxed dependency *(gradually*) on in vitro system. This chapter is the result of the understandings made in this thesis and can thus serve as the general outcome of this thesis.

LIST OF FIGURES AND TABLES

Figures:

16. Direct saliva-to-RICCA-to-LF assay for lab-free testing of SARS-CoV-2 virus. **(Fig 2.9)**

17. Electrophoretic analysis to check sensitivity and specificity of direct saliva-to RICCA assay. **(Fig 2.10)**

18. Comparative analysis of real-time amplification of the RT-PCR and RICCA reaction. RT-PCR and RICCA were performed with (20 pg template for CoV-229E RNA) or without template. To determine the reaction efficiency (rapidity), the Ct values are plotted against the reaction time. **(Fig 2.11)**

19. Healthy Brain (Left) vs Alzheimer's Brain (Right) **(Fig 3.1)**

20. Sequence information of variants E22-Aβ42 and E22P-Aβ42 showing mutation at 22nd position where amino acid glutamic acid changes to proline that leads to higher aggregation and toxicity. **(Fig 3.2)**

21. A schematic illustration of the structure of the toxic and non-toxic conformer of Aβ42. The turn at 22-23 position is the 'toxic-turn', which draws Y10 and M35 closer to accelerate the radical transition and contributes to the hydrophobic core at C-terminus. **(Fig 3.3)**

22. Aβ42 monomers can form higher order assemblies ranging from low molecular weight oligomers, including dimers, trimers, tetramers, and pentamers, to mid-range molecular weight oligomers including hexamers, nonamers and dodecamers to protofibrils and fibrils. **(Fig 3.4)**

23. Time dependent aggregation of E22P-Aβ42 and E22-Aβ42 on Silver-stained Tricine SDS-PAGE analysis confirms rapid aggregation in E22P-Aβ42. (a) 16% T and (b) 10% T along with 16% T gel. **(Fig 3.5)**

24. (a) FM-AFM images of oligomerization and fibril formation in E22-Aβ42 and E22P-Aβ42 (b) E22-Aβ42 at 0 hours when incubated on the mica surface for 5 hours (c) 72 hour incubation of E22-Aβ42 and E22P-Aβ42 (6 μm scale) showing mature fibril formation in E22-Aβ42 and amorphous aggregation (i) along with mature fibrils (ii) in E22P-Aβ42 **(Fig 3.6)**

25. High resolution imaging of E22-Aβ42 and E22P-Aβ42 fibrils using FM-AFM **(Fig 3.7)**

26. Different types of fibrils and nature of coils formed in E22P- Aβ42.Type-1 fibril showing on overlapping fibril that coils over the base fibril. Type-2 and type-3 showing different nature of coiling including loose and compact coiling. Type-4

Tables:

PUBLICATIONS AND CONFERENCES

Awards:

1. Recipient of 'People's Choice Award' at The Hiraku 3MT Competition 2020.
2. Best Poster Presentation Award at 7th AIST International Imaging Conference (2020).

Publications:

1. Biyani, R., Sharma, K., Kojima, K., Biyani, M., Sharma, V., Kumawat, T., ... & Biyani, M. (2021). Development of robust isothermal RNA amplification assay for lab-free testing of RNA viruses. Scientific Reports, 11(1), 1-13.
2. Ishizuka, K., Tsutsumi, Y., Baba, M., Biyani, R., Meng, C. W., Biyani, M., ... & Yasukawa, K. (2020). Inhibition of HIV-1 Reverse Transcriptase Activity by the Extracts of Indian Plants. Journal of Biological Macromolecules, 20(1), 17-22.
3. Biyani, R., Hirata, K., Okhmula, K., Yurtsever, A., Maezono, R., Takagi, M., Fukuma, T & Biyani, M. "Biophysical properties of fibril structure of toxic conformer of amyloid-β42: characterization by atomic force microscopy and molecular docking" (to be submitted in *ACS Nano*)
4. Okhmula, K., Biyani, R., Hongo, K., Biyani, M., Takagi, M., Maezono, R. "Structural and Molecular Mechanism of Amyloid β42 Oligomerization: A Computational Study" (to be submitted in *Journal of Physical Chemistry B*)
5. Hashimoto, Y., Biyani, R., Takamura, Y., Biyani, M. "Simple detection of Amyloid-beta using peptide aptamer and DEPSOR based hybrid biosensor for Alzheimer's disease" (In preparation)
6. Biyani, R.,Okhmula, K., Hongo, K., Maezono, R., Biyani, M. & Takagi, M. "In-vitro and in-silico based development of DNA Aptamers for Aβ42 conformers" (In preparation)

Academic Meetings:

1. Biyani, R., Hirata, K., Fukuma,T., Biyani, M., Takagi, M. (2021). Imaging of the selective interaction between DNA aptamer and toxic conformer of β-amyloid 42 by Frequency Modulation Atomic Force Microscopy. (International) Oral Presentation at the 16th India-Japan Bilateral Conference, India

2. Biyani, R., Biyani, M., Takagi, M. (2021). Development of selective DNA aptamer for Amyloid beta 42 variants using SELCOS. (Domestic) Oral presentation at the 15th Symposium on Biorelevant Chemistry, organized by Division of Biofunctional Chemistry and Division of Biotechnology, The Chemical Society of Japan.

3. Biyani, R., Biyani, M., Takagi, M. (2020). Virtual reality for Virus-X. (International) Oral presentation at the BICON-2020 Virtual Conference, organized by Biyani Group of Colleges, India.

4. Biyani, R. et al. (2020). Identification of the Inhibitory Activity of HIV-1 Reverse Transcriptase in the Extracts of Indian Plants. (International) Poster presentation at the 7th AIST International Imaging Conference, Japan.

ACKNOWLEDGEMENT

I bow in reverence to the almighty for blessing me with power, patience, and confidence, which has helped me in completing the present work. I am hereby presenting my Doctoral Dissertation as a part of the fulfilment of my Doctor of Philosophy in Material Science.

The joy and satisfaction that come along with the successful completion of any work would be in vain unless I mention the names of the people who made it possible, whose constant guidance and encouragement helped me persevere against all odds.

First of all, I take this opportunity to express a deep sense of heartfelt gratitude to my supervisor, **Prof. Masahiro Takagi**, School of Materials Science, JAIST, for providing me a platform to envision science in more fundamental approach. I am thankful for his cordial support, valuable information, and unstinted guidance, which helped me complete my study through various stages. I am very much indebted for his encouraging words from the very first day I entered JAIST for my Master Course. His wide knowledge and logical way of thinking have been of great value to me. It was he who believed that I could early finish my doctor course in a period of two years. The perpetual guidance for quality work given by over the last four years shall carry me a long way in the journey of life on which I am about to embark.

It gives me immense pleasure to thank my second supervisor **Prof. Yuzuru Takamura**, School of Materials Science, JAIST, for keeping complete faith in me and allowing me to use his well-equipped laboratory facilities during the entire course of my research. His exemplary guidance, monitoring, and constant encouragement throughout have been of great assistance to me.

I would also like to express my gratitude to my supervisor for minor research **Prof. Kenzo Fujimoto**, School of Materials Science, JAIST for keeping belief in me. I am obliged to have had your support and encouragement.

Next, I owe a special thanks to **Prof. Manish Biyani**, Research Professor, JAIST, who is also my uncle, for introducing me to the scientific world and providing a start to my scientific career. I grew up seeing his devotion to science, and that planted the seed of contribution towards science in me right from my childhood days. Thank you for making me what I am

today. It wouldn't have been possible without you by my side. I sincerely thank you for thoroughly directing my theme and for keeping my spirit up! Along with this, I wish to acknowledge Prof. Biyani's start-up **BioSeeds Corporation** for all the experimental and laboratory support. I would also like to thank Dr. Kirti Sharma, BioSeeds Corporation who has been like an elder sister to me, and for her camaraderie throughout, especially for all the hours we have shared and spent together in the laboratory.

I also take this opportunity to convey my sincere and honourable thanks to **Prof. Kiyoshi Yasukawa**, Graduate School of Agriculture, Kyoto University for kindly accepting and hosting a part of my research and for giving me this opportunity to work under his able guidance at Kyoto University. I am obliged for the valuable information and relentless support provided by him and all the members of Yasukawa laboratory (especially Ms. Keiko Ishizuka, Mr. Yuto Tsutsumi, and Ms. Misato Baba) in their respective fields.

I sincerely thank **Prof. Takeshi Fukuma** and the team, WPI-Nano Life Science Institute (NanoLSI), Kanazawa University in supporting and hosting AFM experiments for my research. I must thank **Dr. Kaito Hirata** for not only teaching but assisting me in the AFM experiments. I am truly obliged, and all your contribution is kindly acknowledged.

I would also like to thank **Prof. Ryo Maezono** and the team, School of Information Science, JAIST for generously helping me with the *in-silico* experiments and for their constructive suggestions and discussions throughout. I wish to recall my thankfulness to **Mr. Kenji Okumura** for his warm support.

I would also like to extend my heartiest thanks to the members of Takagi Laboratory, Senior Lecturer **Dr. Naofumi Shimokawa**, Secretary **Ms. Chihono Odaka**, Researchers Ms. Mariko Nagata, and Aiko Masuda for providing me with a friendly research environment and motivating me to furnish my research before time. I wish to extend all my love and regards to all my lab mates, Dr. Guo Jingyu and Mr. Yusuke Nakatani for being a big support system and the best colleagues. I thank you all for assisting me whenever I needed you and for helping me to improve my Japanese language skills.

There are many names outside my research life that have provided me with the personal foundation that has been indispensable to shaping my professional life.

First and foremost, I dedicate this work in honor of my family, especially my late grandfather whose dream of seeing me as a doctor is about to come true. I thank my parents and my family for their constant encouragement and silent support without which this would never be possible. I am extremely grateful for your love and care. No amount of words can express my heartfelt appreciation for all the encouragement, inspiration, and motivation that you all have filled in me always and throughout.

I express my heartfelt gratitude and all due respect to **Prof. Madhu Biyani**, Kanazawa University, who is also my aunt for not only the technical discussions and useful guidance that she gave throughout my research period, but also for all the love and motherly care she showered me with along with providing me a lively home-like environment in Japan.

Most importantly I would like to give a special thanks from the bottom of my heart to my little sister and all my siblings for their unconditional love that kept me cheerful. With all my warmth and affection, I also remember the unfailing support of all my dear friends who were always emotionally present for me.

I also wish to acknowledge the **Ministry of Education, Science and Culture of the Japanese Government** for awarding me the MEXT Scholarship and helping me to devote my full time to research.

And finally, I would like to thank all those who rendered me direct or indirect help in completing my study.

Radhika Biyani

CHAPTER -1

Identification and evaluation of small molecule-based natural compounds for inhibition of HIV-1 Reverse Transcriptase activity

(Interaction between protein and small molecule)

1.1 INTRODUCTION

1.1.1 Human Immunodeficiency Virus (HIV) and Reverse Transcriptase (RT) Enzyme

Reverse transcriptase (RT) [EC 2.7.7.49] is an enzyme used to generate complementary DNA (cDNA) from an RNA template by the process of reverse transcription. RT is one of the key players in the mechanism of infection in Human Immunodeficiency Virus type-1 (HIV-1) disease. RT from HIV-1 possesses RNA- and DNA-dependent DNA polymerase activities and RNase H activities that plays a major role in the disease progression [1]. RT is thus an appropriate target for inhibiting HIV-1 replication.

RT is a heterodimer consisting of two distinct subunits, but related chains: 66-kDa p66 subunit and a 51-kDa p51 subunit. p66 is the larger of two subunits. The former includes the fingers, palm, thumb, and connection subdomains and the RNase H domain, while the latter has all four subdomains except the RNase H domain. The fingers, palm, and thumb subdomains are involved in the DNA polymerase activity, and the RNase H domain is in the RNase H activity. These two domains are in general responsible for the binding of the DNA:RNA substrate [2]. The figure below **(Fig 1.1)** shows the crystallographic structure of HIV-1 RT where the two subunits p51 and p66 are colored and the active sites of polymerase and nuclease are highlighted.

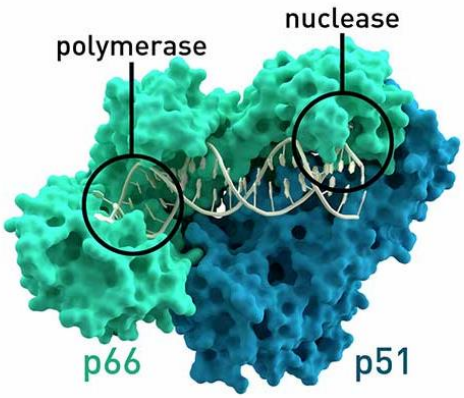
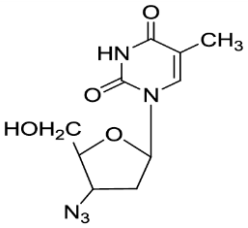


Figure 1.1 : Structure of HIV-1 RT (Source: Wikipedia)

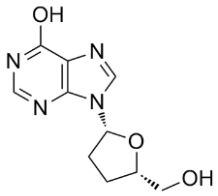
So far, seven nucleosides and five non-nucleoside HIV-1 RT inhibitors have been approved for utilization in HIV-1 therapy. However, these inhibitors block the DNA Polymerase activity and not the RNase H activity. Considering that the emergence of RT inhibitor-resistant HIV-1 RT inhibitors that can block the RNase H activity are

important subjects of study. Various natural products such as flavonoids and alkaloids have been reported to inhibit HIV-1 RT activity. However, being not so specific in nature, it is important to screen different chemicals for the same.

Nucleoside analog RT inhibitors

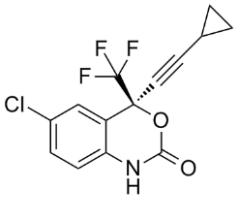


Zidovudine

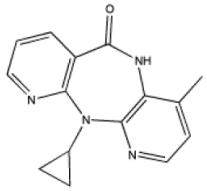


Didanosine

Non-Nucleoside analog RT inhibitors



Efavirenz



Nevirapine

In the current HIV-1 therapy, nucleoside/nonnucleoside RT inhibitors, protease inhibitors, and integrase inhibitors are used, which is highly expensive. In addition, the long-term use of such inhibitors has caused the emergence of multi-drug resistant HIV-1 strains, which is a major issue. Therefore, the development of a novel and cheap HIV-1 RT inhibitor is anticipated.

1.1.2 *Brasenia schreberi*: An important drug source for HIV inhibition

B. schreberi, is a kind of perennial water plant that grows naturally in shallow freshwater ponds and swamps **(Fig.1.2)**. It is a member of the Cabombaceae (water-shield) family. More than 90 percent of this hydrophyte is made up of water. *B. schreberi*, is widely cultivated in Japan where it is commonly known as Junsai. It has long been a part of the Japanese food culture. Akita in Japan has a sustainable cultivation of Junsai.

Figure 1.2: *B.schreberi* plant (Source: Wikipedia)

Junsai, being edible has several medicinal uses. Recently, Yasukawa et al. reported that ethanol and water extracts of *B. schreberi* are not only strongly inhibiting the DNA polymerase activity of HIV-1 RT (Human immunodeficiency virus type 1-reverse transcriptase) to incorporate dTTP, but also the activity of RNase H to hydrolyze the RNA strand of RNA-DNA hybrid [3]. This strongly supports the fact that Junsai contains some substances that can inhibit the HIV-1 replication by inhibiting either the DNA polymerase activity and/or RNase H activity of HIV-1 RT.

1.1.3 Ayurveda: The Indian holistic natural healing system

Ayurveda is the 5000-year-old renowned system of medicine in India. The focus of Ayurveda is to integrate and balance the body, mind, and spirit, rather than focusing on individual symptoms. Ayurveda is mainly based on herbs, plants, flowers, fruits, vegetables, and all vegetation that grows around us in plenty. More than 77 percent of Indian households use Ayurveda products [4,5].

This makes the study of plants for HIV-inhibition more promising to give a natural medication for treating diseases like HIV- AIDS as well [6,7]. In this study, we have chosen and screened below mentioned 5 plants for identification of inhibitory activity to reverse transcriptase enzyme from HIV Type-1. All these plants are not native but are abundantly grown as weeds in India.

a) Parthenium hysterophorus

Kingdom	**Plantae**
Order	**Asterales**
Family	**Asteraceae**
Genus	**Parthenium**
Species	***P. hysterophorus***

P. hysterophorus (commonly known as Gajar Ghas in India) is an annual, highly prolific, and pernicious weed that grows almost everywhere on earth. It is utilized as an ethnomedicine to treat different infectious and chronic diseases. In addition, the

extracts of *P. hysterophorus* have been found to be useful in treating anemia, fever, healing of wounds, heart troubles and sores. Some studies also report the possibility of anti- HIV activity [8,9].

b) Argemone mexicana

Kingdom **Plantae**
Order **Ranunculales**
Family **Papaveraceae**
Genus **Argemone**
Species ***A.mexicana***

A. mexicana is a species of poppy. Despite being poisonous in nature, all parts of this plant are used in traditional medicine to treat chronic diseases of skin, gastrointestinal tract, and respiratory tract. In addition, it exhibits allelopathic, larvicidal, nematocidal, and antimicrobial potential along with chemo sterilant and wound healing activities. It occurs in Ethiopia, Mexico, United States, and many warm places of India. It is also a very commonly used traditional medicine in Taiwan. It is used for treatment of fever, pain, diarrhea, and cancer. Several alkaloids and fatty acids have been isolated from this plant species [10,11].

c) Solanum xanthocarpum

Kingdom **Plantae**
Order **Solanaceae**
Family **Solanaceae**
Genus **Solanum**
Species ***S. xanthocarpum***

S. xanthocarpum is a species of herb. It is a medicinal plant used mostly in India. Some part of the plant, especially fruit, is poisonous. *S. xanthocarpum* is an annual herb which grows as a wild plant in many parts of India. In Indian local dialect it is

known as Kantakari or Bhatkatiya. Fruits and leaves are considered as a valuable herbal product for traditional healers in treatment of many common diseases in many parts of India. Phytoconstituents present in *S. xanthocarpum* are used as anti-fertility, anti-inflammatory, anti-allergic agents and as potential fungicide [12].

d) Calotropis procera

Kingdom	**Plantae**
Order	**Gentianales**
Family	**Apocynaceae**
Genus	**Calotropis**
Species	***C. procera***

C. procera is a species of flowering plant. The milky sap contains a complex mix of chemicals, some of which are steroidal heart poisons. *C. procera* is an important medicinal plant whose leaves and roots have multiple uses. The plant is a rich source of many bioactive compounds which are of medicinal importance. *C. procera* has a bitter, healing, laxative and anthelmintic properties and it relieves strangury, cures ulcers and acts as an expectorant. The leaves are used to relieve stomach pain, flower is a tonic, appetizer, stomachic, cures piles, asthma and wounds. In Indian traditional medicine, the latex of this plant has been used for the treatment of skin diseases, rheumatism, and aches. It has been reported to exhibit potent anti-inflammatory, analgesic, and weak antipyretic activities when administered orally [13].

e) Thevetia peruviana

Kingdom	**Plantae**
Order	**Gentianales**
Family	**Apocynaceae**
Genus	**Cascabela**
Species	***C. thevetia***

T.peruviana is an evergreen shrub or small tree. The plant is widely used in folk medicine. In India, it has become a household remedy for several diseases. It grows widely throughout tropical and subtropical regions. Its generic name is yellow oleander. Despite its toxicity, oleander has been used as an abortifacient, to treat congestive heart failure, malaria, leprosy, indigestion, ringworm, venereal disease and even as a suicide instrument [14]

OBJECTIVE: Identification and evaluation of small molecule-based natural compounds for inhibition of HIV-1 Reverse Transcriptase activity

1.2 MATERIALS AND METHODS

1.2.1 Materials

$p(dT)_{15}$ was purchased from Fasmac (Tokyo, Japan). [methyl-^{3}H]dTTP (1.52 TB q/mmol) and poly(rA) were from GE Healthcare (Buckinghamshire, UK). Recombinant HIV-1 RT was expressed in Escherichia coli and purified from the cells as described previously by Konishi et al. (2013) [15]. The RT concentration was determined using Protein Assay CBB Solution (Nacalai Tesque, Kyoto, Japan) with bovine serum albumin (Nacalai Tesque) as standard.

1.2.2 Extraction of phytochemicals from crude plant parts

a) Fresh leaves, stem, fruits, and flowers of different plants were collected from botanical garden of Science Faculty Campus, University of Allahabad and Company Garden, Allahabad (Table 1).

b) All the plant parts were shade-dried and freeze-dried, crushed and ground into fine powder with electronic mixture. The powdered material was sequentially extracted using different solvents: hexane (HX), ethyl acetate (EA), methanol (ME) and water (AQ) at high temperature (HT) (temperature equivalent to the boiling point of the respective solvents) using Soxhlet apparatus **(Table 1.1)**. The extracts were also prepared in these solvents at room temperature (RT) using shaker. The ratio of plant material and the solvent was 1:10. All the extracts prepared were concentrated/dried using rotavapor.

Table 1.1. Details of plants used in experiment

Sample No.	Plant name	Part of plant	Solvent
1	***P. hysterophorus***	**Leaf**	Aqueous (RT)
2			Ethyl acetate (RT)
3			Hexane (HT)
4			Ethyl acetate (HT)

5		**Flower**	Ethyl acetate (RT)
6			Methanol (HT)
7			Ethyl acetate (HT)
8	***A. mexicana***	**Leaf**	Methanol (RT)
9			Ethyl acetate (HT)
10			Methanol (HT)
11			Aqueous (HT)
12		**Fruit**	Ethyl acetate (HT)
13			Ethyl acetate (RT)
14	***S. xanthocarpum***	**Leaf**	Hexane (HT)
15		**Fruit**	Ethyl acetate (HT)
16			Ethyl acetate (RT)
17			Aqueous (RT)
18	***C. procera***	**Flower**	Aqueous (HT)
19			Ethyl acetate (HT)
20	***T. peruviana***	**Fruit**	Aqueous (HT)

1.2.3 Reverse Transcription Assay using the fluorescent dye PicoGreen

The EnzChek Reverse Transcriptase Assay is a convenient, efficient, and an inexpensive assay for measuring the reverse transcriptase activity. The major role to this method is of the PicoGreen dsDNA quantitation reagent, which preferentially detects dsDNA or RNA-DNA heteroduplexes over single-stranded nucleic acids or free nucleotides. In the assay, reverse transcriptase activity in a biological sample generates long RNA-DNA heteroduplexes from a mixture of a long poly(A) template, an oligo-dT primer and dTTP. The RNA-DNA heteroduplexes formed are then detected by the PicoGreen reagent. In less than an hour, samples can be read in a fluorometer or microplate reader with filter sets appropriate for fluorescein (FITC).

The assay is sensitive, detecting as little as 0.02 units of HIV-1 reverse transcriptase, and has about a 50-fold linear range. Its simplicity also makes it useful for automated high-throughput screening of reverse transcriptase inhibitors.

EnzChek Reverse Transcriptase Assay Kit (Thermo Fisher Scientific, Waltham, MA) was used for the experiments.

a) Sample Preparation for assay: The twenty plants samples were dissolved into 1ml 10% respective solvents from which they were extracted to get a saturated solution. (Sample 20 was incubated at 100°C for 15 mins to dissolve it in water) Samples were then filtered using a 5µm filter to get a homogenous extract and stored at 4°C until use.

b) Preparation of Poly (rA)-p$(dT)_{16}$: 5µl Poly (A) ribonucleotide template and 5µl of oligo d$(T)_{16}$ primer were mixed together in a nuclease-free microfuge tube. The mixture was then incubated at room temperature for one hour to allow the primer to anneal to the template. The template/primer solution was then diluted (200-fold) into polymerization buffer (60 mM Tris-HCl buffer (pH 8.1), 60 mM KCl, 8 mM MgCl2, 13 mM DTT, 100 µM dTTP). (10µL of template/primer solution into 2.0mL of polymerization buffer). For each sample, 96 µl of this mixture (T/P PB) was added into vial and incubated on 37°C until use.

c) Preparation of Samples of Reverse Transcriptase: Recombinant HIV-1 RT was expressed in *Escherichia coli* and purified from the cells. The RT concentration was determined using Protein Assay CBB with bovine serum albumin as standard. HIV- Reverse Transcriptase Enzyme (10.5µM) was diluted to a concentration of 125nM in PDGT Buffer. 10µl RT was added to 40 µl sample volume to make RT concentration 25nM. 2 µl 200mM EDTA was put in each reaction well of the microplate. (For n samples, 4n wells are required)

d) Run the reaction: **At 0 min**: 20µl T/P PB was added into the wells and 5µl 25nM RT-Sample mixture was also added to the wells. Reaction just starts and stops because EDTA was already present in the wells.

At 2.5, 5, 7.5 mins: 24µl 25nM RT-Sample mixture was added to the vial having 96 µl T/P PB and was mixed by pipetting. aliquots (25µl) from the reaction mixture were taken out at 2.5, 5 and 7.5 mins and added to the wells having EDTA to stop the reaction. After the reaction, microplate was wrapped with foil and incubated on ice for 30 mins.

e) Staining with PicoGreen Dye: PicoGreen Dye was diluted with 1X TE (10 mM Tris-HCl buffer (pH 7.5), 1 mM EDTA).

f) 173µl PicoGreen Solution was added into each well (25 µl), mixed well and the microplate was incubated at room temperature for 10 mins after covering with aluminum foil (Buffer in the ratio 1:350).

g) Quantitate the reverse transcriptase activity: The fluorescence at 523nm was measured using EnSight (PerkinElmer,Waltham,MA) with the excitation wavelength of 502nM. Relative activity was measured as:

$$\text{Relative activity (\%)} = \frac{\text{Slope of sample}}{\text{Slope of water}} \text{ X } 100$$

1.2.4 Reverse Transcription Assay using the [^{3}H]-dTTP

a) Poly(rA)-p(dT)15 was prepared by annealing p(dT)15 and poly(rA) (around 350 base).

b) Reverse transcription reaction (50 µl) was carried out in 25 mM Tris-HCl buffer (pH 8.3), 50 mM KCl, 2 mM DTT, 5 mM MgCl2, 25 µM poly(rA)-p(dT)15 (this concentration is expressed based on p(dT)15), 0.2 mM [^{3}H]dTTP (1.85 Bq/pmol) and 5nM RT at 37°C.

c) An aliquot (20µl) was taken from the reaction mixture at 2.5 and 5 min and immediately spotted onto the glass filter (GF/C 2.5 cm; Whatman, Middlesex, UK).

d) Unincorporated [^{3}H]dTTP was removed by three washes of chilled 5%(w/v) trichloroacetic acid (TCA) for 10 min each, followed by one wash of chilled 95% ethanol.

e) The radioactivity retained on the dried filters was counted in 2.5 ml of Ecoscint H. The initial reaction rate was estimated from the time-course for incorporation of [^{3}H]dTTP. One unit is defined as the amount which incorporates 1 nmol of dTTP into poly(rA)-p(dT)15 in 10 min. Relative activity as measured as:

$$\text{Relative activity (\%)} = \frac{\text{Activity with extracts}}{\text{Activity without extracts}}$$

1.2.5 Liquid-Liquid extraction of Sample 1 to identify the nature of inhibitor

a) Sample 1 was dissolved up to saturation in 2ml water and then filtered using a 5µm filter to get a homogenous extract. Two aliquots of 300µl each were taken out.

b) Equal amount of 10% Ethyl Acetate and 10% Benzene were added to them respectively in two different vials. After vortexing for 5 mins, the mixture was centrifuged for 10mins at 15000 rpm at 1°C.

c) Water (about 350 µl) and solvent (about 250 µl) layers were separated gently using pipette and put to dry overnight.

d) 10% Ethyl Acetate was added to each vial keeping the same initial volume as before drying.

e) Reverse Transcription Assay was carried out using the EnzChek Reverse Transcriptase Assay Kit.

1.3 RESULT AND DISCUSSION

1.3.1 Optimizing the RT concentration in the presence & absence of *Brasenia schreberi*

It has been reported that ethanol- and water-extracts of Brasenia schreberi (Junsai) exhibit strong inhibitory activities for the DNA polymerase activity of HIV-1 RT. The RT concentration that has been used in all the experiments was optimized at 5nM, 10nM and 40nM concentrations of RT in the presence and absence of Junsai (control plant). 5mg/ml Junsai was used for experiments. 5nM is considerably the most suitable condition for measurement as can be seen from **Fig. 1.3**.

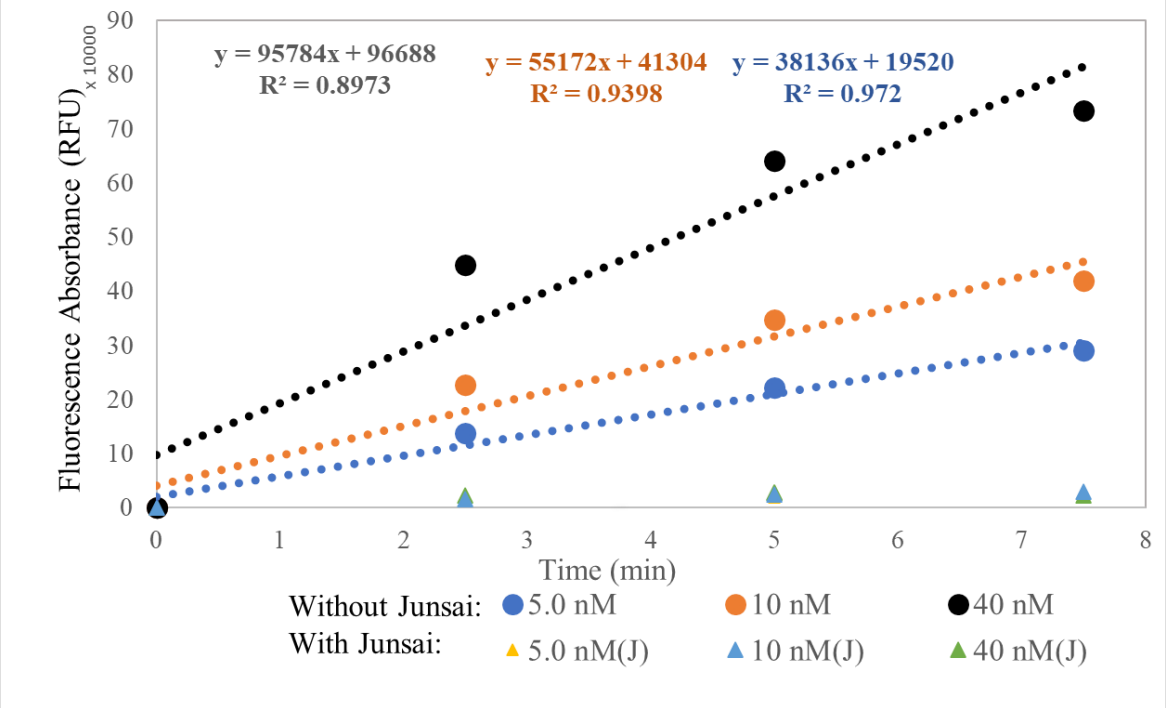


Figure 1.3: Optimized RT concentration for Enz Chek Assay

1.3.2 Checking the influence of different solvents on RT activity for Enz Chek Assay

The influence of different solvents (DMSO, Benzene and Water) at different concentrations was checked using the EnzChek Reverse Transcriptase Assay. This was done to check if the presence of solvent influences the inhibition of RT Activity. DMSO cannot be used since it inhibits the enzyme activity. Benzene is insoluble and possibly causes the denaturation of RT and thus cannot be used. 50% and 30% water

is insoluble to RT, thus it was concluded that all the samples can be prepared with 10% solvents in water **(Table 1.2)**.

Solvent	Extract of plant samples						
	Hexane	**Ethyl Acetate**			**Methanol**		
	10%	50%	30%	10%	50%	30%	10%
DMSO	X	X	X	X	X	X	X
Benzene	X	X	X	X	X	X	X
Water	O	X	X	O	X	X	O

Table 1.2: Influence of different solvents on RT activity

1.3.3 Screening of HIV-1 RT inhibitory activity

The inhibitory activities of 20 extracts for HIV-1 RT were measured and compared by [^{3}H]-dTTP assay and fluorescent dye PicoGreen assay **(Fig.1.4)**. The reaction was carried out at 37°C with HIV- 1 RT at 5 nM and extracts at 8% v/v. The reaction without extracts was also carried out as a control. Relative activity was defined as the ratio of the reaction rate with extracts to that without it.

Water extracts of the leaf of *P. hysterophorus* (Plant No. 1) and water extracts of the leaf of *A.mexicana* (Plant No. 11) inhibited HIV-1 RT activity by more than 90%, while ethyl acetate, hexane, and methanol extracts of the same plants inhibited HIV-1 RT activity by 17-58%. The difference in inhibitory activity between water and other extracts was probably due to the different inhibitory constituents which were extracted by water and not other solvents. Extracts of the leaf, fruit, and flower of *S. xanthocarpum*, *C. procera*, and *T. peruviana* inhibited HIV-1 RT activity by 10-80%. For subsequent analysis, the water extracts from leaf of *P. hysterophorus* (Plant No. 1) and *A. mexicana* (Plant No.11) were selected because of their high inhibitory activity.

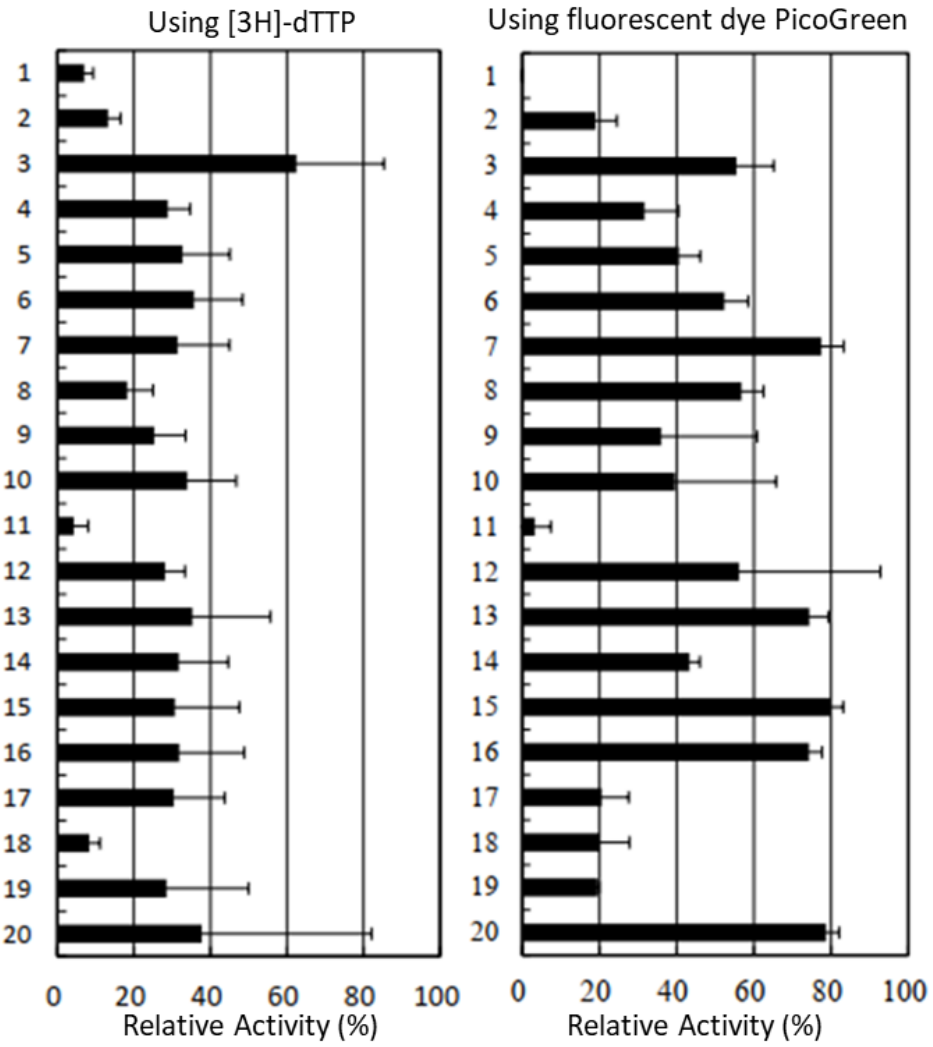
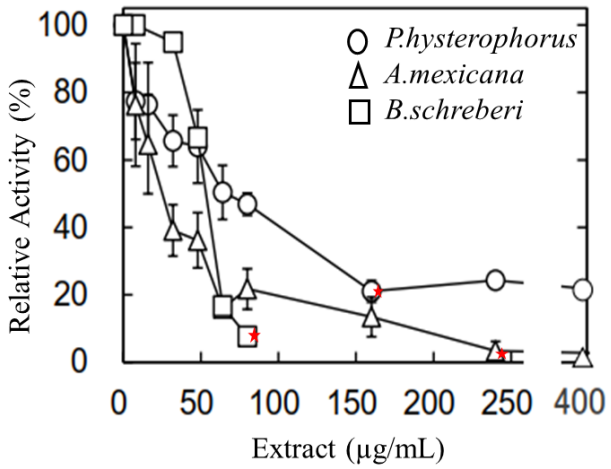


Figure 1.4: Inhibition of HIV-1 RT activity by 20 extracts

1.3.4 HIV-1 RT inhibitory activity of water extracts from the leaf of *A. mexicana*

The effect of increasing concentration of water extracts from the leaf of *P. hysterophorus* and *A. mexicana* on HIV-1 RT activity were examined **(Fig. 1.5)**. *Brasenia schreberi* (Junsai) was also examined as a comparison. The inhibition of HIV-1 RT by water extracts of the leaf of *P. hysterophorus* is shown in a circle, A. mexicana. Is shown in a triangle and water extracts of B. schreberi in a square. The results of the assay using [^{3}H]-dTTP are shown. Relative activity was defined as the ratio of the reaction rate with extracts to that without it. One of the representative data of the average of triplicate determinations with SD value is shown.

The reaction rate decreased with increasing concentration of extracts. The relative activity reached 20% at 160 µg/ml for *P. hysterophorus*, 0% at 240 µg/ml for *A. mexicana*, and 0% at 80 µg/ml for *B. schreberi*. The concentrations giving 50% effective inhibition (IC50 values) were estimated as 60.4 µg/ml for *P. hysterophorus*, 23.7 µg/ml for *A. mexicana*, and 50.2 µg/ml for *B. schreberi*. These results indicate that the inhibitory effect of *A. mexicana* was higher than that of *P. hysterophorus*. Considering that the IC50 values for the DNA polymerase activity of HIV-1 RT of the various plant extracts used for traditional medicine were in the range of 2 to 100 µg/ml, the IC50 values of *A. mexicana* and *P.hysterophorus* were high enough [16].



Plant	Concentration at 50% inhibition (IC_{50})
P.hysterophorus	60.4 ug/ml
A.mexicana	23.7 ug/ml
B.Schreberi (Control)	50.2 ug/ml

Figure 1.5: Inhibition of HIV-1 RT by water extracts of the leaf of P. hysterophorus, A. mexicana. and B. schreberi

1.3.5 Effects of thermal treatment and proteinase treatment on the inhibitory activity

In order to see if the inhibitory substance of the leaf of *A. mexicana* is an organic compound or a protein, the inhibitory activity of the extracts that had received the thermal treatment at 100°C for 10 min or proteinase K treatment was examined **(Fig. 1.6)**. The results of the assay using [^{3}H]-dTTP is shown. Relative activity was defined as the ratio of the reaction rate with extracts to that without it. The concentrations of the extracts of *A.mexicana* and B. schreberi were 160 and 400 µg/ml, respectively, in the thermal treatment at 100°C for 5 min, 160 and 400 µg/ml, respectively, in the proteinase K treatment (50 µg/ml in 20 mM Tris-HCl buffer (pH 7.5) at 37°C for 60 min), and 64 and 160 µg/ml, respectively, in the RT reaction. The concentrations of

the solution received were 0.8 mg/mL for A. mexicana and 2 mg/mL for B. Schreberi. To adjust the concentration at RT reaction to be 8% v/v, the concentrations for thermal and protein K treatments were determined, and therefore are different. One of the representative data of the average of duplicate determinations are shown. Neither thermal treatment nor proteinase K treatment of the extracts abolished the inhibitory activity. The same results were obtained for water extracts of *B. schreberi*. These results suggest that the inhibitory substance in the leaf of *A. mexicana* is an organic compound rather than a protein.

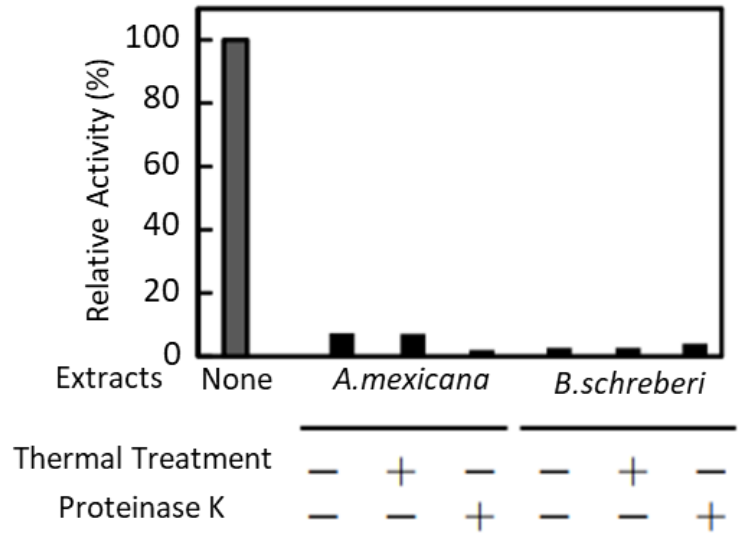


Figure 1.6: Effects of thermal treatment and proteinase treatment on the inhibitory activity.

1.3.6 Advancing studies by *in-silico* approach

After confirming the nature of novel inhibitor compound as an organic hydrophilic compound, the next step is the structure identification of the compound. *In-silico* studies can then be utilized to study the molecular interaction between inhibitor compound and target RT at atomic levels with the help of molecular docking. Similar studies have been performed for the molecular docking prediction for the molecular mechanism of inhibition, followed by comparison of the conformational and steric parameters [17]. Thus, understanding the molecular mechanism of inhibition can help in advancing the drug development and clinical testing stages, that can help in speeding up the progress in the therapeutic world.

1.4 CONCLUSIONS

Interaction between small molecules and proteins has been studied in this chapter. HIV-1 RT (protein in this study) activities of the extracts (small molecule in this study) from 5 different Indian plants namely *Parthenium hysterophorus, Argemone mexicana, Solanum xanthocarpum, Calotropis procera*, and *Thevetia peruviana* was evaluated with the Enz Chek assay and the Tritium assay.

The concentration of RT was optimized to 5 nM and samples were prepared in 10% solvents (Hexane, Ethyl Acetate, Methanol) with water. In Enz Chek assay, water extract of the leaves of *P.hysterophorus* and *A.mexicana* showed strong inhibition in common. In Tritium assay, water extract of the leaf of *A.mexicana* has the strongest inhibitory activity. Neither thermal treatment at 100ºC nor proteinase K treatment of the extracts abolished the inhibitory activity, suggesting that the inhibitory substance is an organic hydrophilic compound rather than a protein. It was observed that the *A.mexicana* shows a higher inhibition activity than the currently reported highest HIV-1 inhibitory activity by Junsai creating a lot of potential to study more plants.

The purification and identification of the HIV-1 RT inhibitory compound from *A.mexicana* followed by studying the molecular mechanism with the help of *in-silico* studies can be performed in future to advance the study. This can be followed by in-vivo studies to elucidate and compare the function of inhibitor compound inside the cell, as reported in case of some RT inhibitors [18,19].

Various enzyme inhibitors have been so far identified in natural products. In the case with HIV-1 RT, 5,6,7-trihydroxyflavone (baicalein) [20] and catechins containing a galloyl moiety such as (-)-epicatechin 3-gallate (ECG) [21] and (-)-epigallocatechin 3-gallate (EGCG) [22] were reported to inhibit DNA polymerase activity of HIV-1 RT.

However, India has numerous plant resources, and our results suggest that there might be unidentified Indian plants which possess strong HIV-1 RT inhibitory activity. Combined with Ayurveda technique, such plants might be developed as a novel inexpensive pharmaceutical for HIV-1 therapy.

REFERENCES

1. Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. A., Hughes, S. H., & Arnold, E. (2009). Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. Journal of molecular biology, 385(3), 693-713.
2. Jäger, J., Smerdon, S. J., Wang, J., Boisvert, D. C., & Steitz, T. A. (1994). Comparison of three different crystal forms shows HIV-1 reverse transcriptase displays an internal swivel motion. Structure, 2(9), 869-876.
3. Hisayoshi, T., Shinomura, M., Yokokawa, K., Kuze, I., Konishi, A., Kawaji, K., ... & Sakuda, S. (2015). Inhibition of the DNA polymerase and RNase H activities of HIV-1 reverse transcriptase and HIV-1 replication by Brasenia schreberi (Junsai) and Petasites japonicus (Fuki) components. *Journal of natural medicines*, *69*(3), 432-440.
4. https://theselfimprovementblog.com/self-improvement/self-improvement-tips/ayurvedic-medicine-benefits/
5. https://economictimes.indiatimes.com/industry/healthcare/biotech/healthcare/77-percent-indian-households-use-ayurvedic-products-pwc-report/articleshow/66773295.cms?from=mdr
6. Hisayoshi, T., Shinomura, M., Konishi, A., Tanaka, J., Shimoda, H., Hata, K., ... & Yasukawa, K. (2014). Inhibition of HIV-1 reverse transcriptase activity by Brasenia schreberi (Junsai) components. Journal of Biological Macromolecules, 14(1), 59-65.
7. Hisayoshi, T., Shinomura, M., Yokokawa, K., Kuze, I., Konishi, A., Kawaji, K., ... & Sakuda, S. (2015). Inhibition of the DNA polymerase and RNase H activities of HIV-1 reverse transcriptase and HIV-1 replication by Brasenia schreberi (Junsai) and Petasites japonicus (Fuki) components. Journal of natural medicines, 69(3), 432-440.
8. Kumar, S., Chashoo, G., Saxena, A. K., & Pandey, A. K. (2013). Parthenium hysterophorus: aprobable source of anticancer, antioxidant and anti-HIV agents. BioMed Research International, 2013.

9. Kushwaha, V. B., & Maurya, S. (2012). Biological utilities of Parthenium hysterophorus. Journal of Applied and Natural Science, 4(1), 137-143.

10. Srivastava, N., Chauhan, A. S., & Sharma, B. (2012). Isolation and characterization of some phytochemicals from Indian traditional plants. Biotechnology research international, 2012.

11. Joel. H., Elizondo-Lu´evano, Roc´ıo CastroR´ıos, Eduardo S´anchez-Garc´ıa, Magda E.Hern´andez-Garc´ıa, Javier VargasVillarreal, Osvelia, E. Rodr´ıguez-Luis, and Abelardo Ch´avez-Montes. (2018) Can. J. Physiol. Pharmacol. 1-8

12. Kumar, S., & Pandey, A. K. (2014). Medicinal attributes of *Solanum xanthocarpum* fruit consumed by several tribal communities as food: an in vitro antioxidant, anticancer and anti HIV perspective. BMC complementary and alternative medicine, 14, 112.

13. Mohanraj, R., Rakshit, J., & Nobre, M. (2010). Anti HIV-1 and antimicrobial activity of the leaf extracts of *Calotropis procera*. International Journal of Green Pharmacy (IJGP), 4(4).

14. Tewtrakul, S., Nakamura, N., Hattori, M., Fujiwara, T., & Supavita, T. (2002). Flavanone and flavonol glycosides from the leaves of *Thevetia peruviana* and their HIV-1 reverse transcriptase and HIV-1 integrase inhibitory activities. Chemical and Pharmaceutical Bulletin, 50(5), 630-635.

15. Konishi, A., Shinomura, M., & Yasukawa, K. (2013). Enzymatic characterization of human immunodeficiency virus type 1 reverse transcriptase for use in cDNA synthesis. Applied biochemistry and biotechnology, 169(1), 77-87.

16. Baba, M., Kakue, R., Leucht, C., Rasor, P., Walch, H., Ladiges, D., & Yasukawa, K. (2017). Further increase in thermostability of Moloney murine leukemia virus reverse transcriptase by mutational combination. Protein Engineering, Design and Selection, 30(8), 551-557.

17. da SM Forezi, L., Ribeiro, M. M., Marttorelli, A., Abrantes, J. L., Rodrigues, C. R., Castro, H. C., ... & de Souza, M. C. B. (2020). Design, synthesis, in vitro and in silico

studies of novel 4-oxoquinoline ribonucleoside derivatives as HIV-1 reverse transcriptase inhibitors. European Journal of Medicinal Chemistry, 194, 112255.

18. Abram, M. E., Tsiang, M., White, K. L., Callebaut, C., & Miller, M. D. (2015). A cell-based strategy to assess intrinsic inhibition efficiencies of HIV-1 reverse transcriptase inhibitors. Antimicrobial agents and chemotherapy, 59(2), 838-848.

19. Hukezalie, K. R., Thumati, N. R., Coˆte, H. C., & Wong, J. M. (2012). In vitro and ex vivo inhibition of human telomerase by anti-HIV nucleoside reverse transcriptase inhibitors (NRTIs) but not by non-NRTIs. PLoS One, 7(11), e47505.

20. Ono, K., Nakane, H., Fukushima, M., Chermann, J. C., & Barré-Sinoussi, F. (1989). Inhibition of reverse transcriptase activity by a flavonoid compound, 5, 6, 7-trihydroxyflavone. Biochemical and biophysical research communications, 160(3), 982-987.

21. Nakane, H., & Ono, K. (1990). Differential inhibitory effects of some catechin derivatives on the activities of human immunodeficiency virus reverse transcriptase and cellular deoxyribonucleic and ribonucleic acid polymerases. Biochemistry, 29(11), 2841-2845.

22. Li, S., Hattori, T., & Kodama, E. N. (2011). Epigallocatechin gallate inhibits the HIV reverse transcription step. Antiviral Chemistry and Chemotherapy, 21(6), 239-243.

CHAPTER -2

Co-operative interaction between RNA polymerase and Recombinase for Reverse Transcriptase-based improved isothermal nucleic acid amplification assay

(Interaction between protein and nucleic acid)

2.1 INTRODUCTION

With the emergence of new infectious virus strains that result in outbreaks, the identification and isolation of infected suspects (symptomatic or asymptomatic) constitute the most effective method for preventing human-to-human disease transmission and pandemics [1,2]. In this stream, a simple yet reliable, rapid and field-deployable molecular diagnostic tool for the detection of nucleic acid sequences of infectious viruses is vital for the massive screening of the population and has been greatly needed worldwide during the current global wave of coronavirus disease 2019 (COVID-19). Hundreds of molecular tests have been rapidly introduced but still face an arduous journey to reach mass usage, due mainly to the shortcomings in the existing laboratory-based testing paradigm for RNA-based viral diagnostics [3]. First, nucleic acid isolation from clinical samples, which is routinely performed in a laboratory for PCR testing, a current gold-standard diagnostic test, constitutes a major bottleneck. This bottleneck becomes more challenging for RNA-based viruses because rapid RNA degradation contributes to poor clinical sensitivity [4]. Second, the requirements of bulky expensive instrumentation, such as thermal cyclers with fluorometry, hinder the application of current PCR testing in the field.

To obtain PCR-similar molecular testing outside a centralized laboratory, various isothermal (single-temperature) nucleic acid amplification methods have been devised and are in a continuous race to achieve a performance similar to that of PCR tests. The major candidates include recombinase polymerase amplification (RPA) [5-9], nucleic acid sequence-based amplification (NASBA) or RNA-specific amplification [10-12], and loop-mediated isothermal amplification (LAMP) [13]. The advantage of isothermal amplification methods over PCR is that they do not require bulky instrumentation, enable rapid (10-60 min) amplification of nucleic acids at a constant temperature (e.g., 37-65°C) and therefore improve throughput in situations in which large numbers of clinical samples must be processed and facilitate point-of-care diagnosis. Recent advances in saliva-based isothermal amplification-based assays, such as RT-LAMP or RT-RPA, have encouraged the testing of SARS-CoV-2-containing clinical samples [14-17]. Although encouraging, isothermal amplification methods have also been shown to produce more false-positive (i.e., lower test specificity due to nonspecific amplification) results than PCR testing [18].

Specifically, false-positive signals in RPA, which are more likely to occur due to nontarget-triggered amplification reactions or primer-dependent artifacts, are often exacerbated at low reaction temperatures and are magnified with increases in the reaction time [19]. To circumvent this issue, the application of probe chemistry [20], portable gel electrophoresis [21], and CRISPR-based technology [22,23] has been demonstrated to improve the performance of RPA. Noticeably, the problem of nonspecific amplification in RPA generally occurs when the viral loads in the tested samples are low [24]. Additionally, if the target is RNA, the amplification methods depend on the efficiency of the initial molecular step, reverse transcription, which can greatly affect post DNA-specific amplification efficiency at low concentrations of the RNA template [25] and has also been reported to have an effect of apparent false-negative results (from 17% to as high as 48%) due to the role of the SARS-CoV-2 viral load dynamics [26]. Therefore, to maintain the integrity of RNA in the reaction sample, a method to support reverse transcription and effectively amplify a low viral RNA load is crucial.

RNA-specific amplification specifically amplifies a target RNA sequence at a temperature near 40°C with reverse transcriptase (RT) and RNA polymerase [10-12]. In RNA-specific amplification, RT synthesizes promoter-bearing double-stranded (ds) DNA with the help of its RNase H activity. RNA polymerase continues to induce in vitro transcription to produce copies of RNA fragments that are subsequently recycled as RNA templates for the synthesis of promoter-bearing dsDNA. However, the formation of promoter-bearing dsDNA has a possibility of failure if the initial low target RNA template copy number in clinical samples is further removed due to ready degradation by the action of host nucleases and/or metal ions (metal-ion-based RNA cleavage). Noticeably, the half-life of exogenous viral RNA in host saliva is approximately 30 sec [27]. During this failure event, enzymes and primers are consumed in vain; thus, it takes time for trace amounts of template RNA to enter the amplification cycle due to such failure. To circumvent this issue, we proposed to power the RNA cycle by entering the DNA cycle using RPA, a DNA-specific amplification method that can amplify a target DNA sequence at a constant temperature similar to that used in the RNA-specific amplification method using a recombinase enzyme, a single-stranded DNA-binding protein (SSB), and a strand-displacing polymerase. As of now, we have not come across

any study that has reported the co-assisted amplification of RNA and DNA cycles using an isothermal amplification method. In this study, we aimed to advance the bioengineering of our earlier studies on RNA-specific amplification [11,12] and RPA [8,9,21] and then integrate the essentials of both types of isothermal amplification methods into a simple format of 'sample-in and answer-out' with a primary focus on the detection of low copy numbers of viral RNA directly from COVID-19 saliva samples without the need for any laboratory handling or sample preprocessing.

In this regard, we report the development of a completely homogeneous, isothermal, highly sensitive, and ultrarapid method for detecting virus RNA target sequences for the on-site (low resource settings) molecular diagnosis of COVID-19 and other infectious diseases. We named this method the RICCA (RNA Isothermal Co-assisted and Coupled Amplification) reaction. The current manual prototype concurrently measures samples within 15-30 min with a sensitivity as low as a few copies per microliter of living bacterial cells (*E. coli*) and heat-inactivated viruses (SARS-CoV-2 and HCoV-229E).

OBJECTIVE: Co-operative interaction between RNA polymerase and Recombinase for Reverse Transcriptase-based improved isothermal nucleic acid amplification assay

2.2 MATERIALS AND METHODS

2.2.1 Preparation of standard SARS-CoV-2 or CoV-229E RNA templates

The sequence of the SARS-CoV-2 N gene was identified from GenBank (accession number Mn908947), and the 134-nt single-stranded DNA fragment of the nucleocapsid protein gene, corresponding to the RNA sequence at position 28707-28840, was synthesized in vitro. A T7 RNA polymerase promoter sequence was added to the 5' terminal end, and the double-stranded DNA construct was synthesized by overlap extension PCR using DNA oligomers (**Table 1**) purchased from a commercial supplier (Eurofins Genomics, Tokyo, Japan). The RNA transcript was synthesized using a cell-free transcription kit (RiboMAX, Promega, Madison, WI, USA) and purified (RNeasy, Qiagen, Tokyo, Japan). The concentration of purified RNA was determined spectrophotometrically (Nanodrop ND-2000, Thermo Fisher Scientific, Waltham, MA), and standard templates ranging from 10^9 to 1 RNA copies/µl were serially diluted and freshly prepared in Tris-EDTA (TE) buffer.

Table 2.1: Primers used in this study

	Name	Base position[a]	Sequences (5' to 3')
SARS-CoV-2 RNA	Anti-sense strand RNA-specific amplification- 134 nts fragment (28,707-28,840)		
	N_Sarbeco_F-	28706-28724	CACATTGGCA CCCGCAATC
	T7-N_Sarbeco_R-	28814-28833	*AATTCTAATACGACTCACTATAGGGAGA*GAGGAACGAGAAGAGGCTTG
	Sense strand RNA-specific amplification- 134 nts fragment (28,707-28,840)		
	T7-N_Sarbeco_F+	28706-28724	*AATTCTAATACGACTCACTATAGGGAGA*CACATTGGCACCCGCAATC
	N_Sarbeco_R+	28814-28833	GAGGAACGAGAAGAGGCTTG
	RICCA reaction-134 nts fragment (28,707-28,840)		
	DIG-N_Sarbeco_F-	28706-28724	[Digoxigenin]CACATTGGCACCCGCAATC
	BIO-N_Sarbeco_R-	28814-28833	[Biotin]GAGGAACGAGAAGAGGCTTG
	Anti-sense strand RNA-specific amplification- 164 nts fragment (28,195-28,358)		
	2019-nCoV_N1-CDC_F	28195-28214	GTTGTTCGTTCTATGAAGAC
	T7-2019-nCoV_N1-CDC_R	28335-28358	*AATTCTAATACGACTCACTATAGGGAGA*TCTGGTTACTGCCAGTTGAATCTG
	Lateral flow assay probes for 164 nts fragment (28,195-28,358)		
	5DIG-CDC-P1	28201-28230	[Digoxigenin]CGTTCTATGAAGACTTTTTAGAGTATCATG
	3-BIOTIN-CDC-P1	28231-28260	AAATCTACTTTAGATTTTGTTGTGCTTGCA[Biotin]
CoV-229E RNA	Anti-sense strand RNA-specific amplification		
	229E_F	25151-25175	TTTTCCGACGTGCTCGAACTTTTTG
	T7-229E_R	25261-25285	*AATTCTAATACGACTCACTATAGGGAGA*CGCTCAACAAGGTCACAGTAATGCC
	RICCA reaction		
	DIG-229E_F	25151-25175	[Digoxigenin]TTTTCCGACGTGCTCGAACTTTTTG
	BIO-229E_R	25261-25285	[Biotin]CGCTCAACAAGGTCACAGTAATGCC
E. Coli 16S rRNA	RNA-specific amplification		
	ECA619_F	0618-0600	AGCCCGGGGATTTCACATC
	T7-ECA415_R	0415-0435	*AATTCTAATACGACTCACTATAGGGAGA*GGCCTTCGGGTTGTAAAGTAC
	RICCA reaction		
	ECA415_F	0415-0442	[Digoxigenin]GGCCTTCGGGTTGTAAAGTACTTTCAGC
	ECA619_R	0618-0593	[Biotin]AGCCCGGGGATTTCACATCTGACTTA

The italicized sequence is the T7 polymerase-binding sequence. [a]The base position corresponds to that described in the sequence deposited in GenBank.

2.2.2 Preparation of standard *Escherichia coli* rRNA template

The sequence of *Escherichia coli* (*E. coli*) 16S rRNA was identified from GenBank (accession number J01859.1), and the 204-nt single stranded DNA fragment of 16S rRNA, corresponding to the RNA sequence at position 415-618, was synthesized in vitro. The test strain (NBRC 3301) of nonpathogenic *E. coli* was obtained commercially (NITE Biological Resource Center, Tokyo, Japan) and grown in L broth at 37°C for 24 h. The cell counts (CFUs per mL) in serially diluted bacterial culture were determined using a JuLI-smart fluorescent cell analyzer and conventional hemocytometer. For isothermal amplification reactions, cells were harvested from 1 mL of culture by 5 min of centrifugation at 15,000 x *g*, and RNA extraction was

performed by disrupting the bacterial cells first by sonication and then by treatment with NP40 reagent (Surfact-Amps NP40, Thermo Fisher Scientific). The solution was then centrifuged at 15,000 x *g* for 5 min to sediment cell debris. The supernatant was used directly as a template for isothermal amplification reactions or was stored at -20°C until use.

2.2.3 RNA-specific amplification assay

The reaction mixture was prepared by mixing 5.0 µl of substrate stock solution [17.6 µl of nuclease-free water (NFW), 5.6 µl of 1 M $MgCl_2$, 19.8 µl of 1.9 M Tris HCl (pH 8.6), 0.6 µl of 40 U/µl RNase Inhibitor, 3.3 µl of 100 mM DTT, 41.3 µl of 2.0 mM dNTP (each), 17.1 µl of 58 mM NTP (each) and 4.8 µl of 250 mM inosine triphosphate], 5.0 µl of primer stock solution [32.5 µl of NFW, 21.5 µl of 2 M KCl, 6.6 µl of 50 µM sense forward primer, 6.6 µl of 50 µM promoter-bearing antisense reverse primer and 42.9 µl of dimethyl sulfoxide] and 2.5 µl of RNA template in a PCR tube and incubated at 65°C for 5 min (optional) and 41°C for 5 min. The reaction was started by adding 2.5 µl of enzyme stock solution [4.0 µl of 10 mg/mL Bovine Serum Albumin, 4.4 µl of 20,000 unit/mL AMV Reverse Transcriptase (Life Sciences Inc, Petersburg, FL), 11.1 µl of 60% sorbitol and 35.5 µl of T7 RNA Polymerase of 50,000 unit/mL (Toyobo)] and continued at 41°C for 10 to 30 min. The amplified products were analyzed using 1-inch 6% denatured polyacrylamide gel [24] and SYBR Gold staining.

2.2.4 RNA-specific amplification assay with DNA-specific amplification

The reaction was performed using a TwistAmp Basic kit (TwistDx, Cambridge, MA, USA). Briefly, the reaction mixture for RNA-specific amplification (5.0 µl of substrate stock solution, 5.0 µl of primer stock solution, and 2.5 µl of standard RNA or sample) was mixed with 5 µl of primer mix (10 µM digoxin-modified sense/forward primer and 10 µM biotin-modified antisense/reverse primer) and 29.5 µl of primer-free rehydration buffer. The resultant 49.5 µl of mixture and 2.5 µl of enzyme stock solution were added to the dried enzyme pellet supplied by the TwistAmp Basic kit. The reaction was initiated by adding 2.5 µl of 280 mM magnesium acetate and immediately placed at 41°C for 10 min. The amplified

products were analyzed using a 1-inch 6% denatured polyacrylamide gel [24] and SYBR Gold staining.

2.2.5 Gel electrophoretic analysis

A 1-inch electrophoretic device developed in-house was used for gel-based detection [24]. Ten-well microcassettes with 6% T polyacrylamide gel were prepared using 40(w/v)%-acrylamide/bis (19:1) (Nacalai Tesque, Kyoto, Japan), 5X TBE, 10% ammonium persulfate (Fujifilm Wako Pure Chemicals) and tetramethyl ethylenediamine (Fujifilm Wako Pure Chemicals, Osaka, Japan) and used as a 1-inch (2.5 × 2.5 cm) precast polyacrylamide gel to analyze the RNA-specific amplification or RICCA reaction products. Paper pads (1.0 × 2.0 cm) were soaked in 2 mL of 1X TBE buffer and used as a source of running buffer. One microliter of each sample was loaded in the micro gel wells with 0.5 µL of 6X gel loading dye. A 50-bp DNA ladder was also loaded in one well to confirm the size of the obtained DNA fragments and comigrated with the samples at 100 V for 5-7 min. A 300-µl volume of 10X SYBR gold nucleic acid gel stain (Thermo Fisher Scientific) was poured on top of the gel, and the bands were visualized using the blue LED flashlight installed in the palm-sized electrophoretic device. Images were acquired using a standard camera-equipped cellular phone.

2.2.6 Real-time analysis

The real-time RICCA or PCR assay was performed with a Thermal Cycler Dice Real Time System (Thermo Fisher Scientific). The real-time RICCA reaction was prepared in a 50-µl reaction volume as described in an earlier section with the addition of 800× SYBR™ Green I dye (Invitrogen). For real-time PCR, the first reverse transcription reaction was performed using MM4 reverse transcriptase [32]. The composition of the 20-µl reaction volume was as follows: 2 µl of 10 nM MM4 enzyme, 1 µL of 1 mM dNTP mixture, 2 µL of 10× RT buffer, 0.5 µL of RNase inhibitor (20 U), 1 µL of RNA template, and 3.5 µL of nuclease-free water. The reaction program was 42°C for 30 min followed by incubation at 65°C for 5 min. Real-time PCR was conducted using the TB Green™ Premix Ex Taq™ II Kit (Tli RnaseH Plus, Takara Bio, Kusatsu, Japan) according to the manufacturer's instructions. The composition of the

25-µl reaction volume was as follows: 12.5 µl of 2×□SYBR Premix Ex Taq II (Tli RNaseH Plus), 1 µL of forward primer (10 µM), 1 µL of reverse primer (10 µM), 2 µL of template (reverse transcription product), and 8.5 µL of nuclease-free water. The following thermal program was used: reverse transcription at 55°C for 15 min followed by 90°C for 30 sec and 35 cycles of amplification (30 sec at 94°C, 30 sec at 60°C, and 30 sec at 72°C).

2.2.7 Lateral flow immunochromatography assay

The lateral flow dipstick (LFD) was provided by Asahi Kasei Pharma, Japan. The amplification products from the RICCA reaction were added to the sample pad of the LFD strip and allowed to migrate by capillary action, and during process, the biotinylated RICCA product hybridized with a digoxin-labeled DNA probe and complexed with a gold-labeled anti-digoxin antibody. This hybridization product was trapped by a biotin ligand and bound to a lateral flow test strip to form an immune complex bound at the test line (T). The experimental procedure and determination method were as follows: 5 µl of RICCA product was added to 95 µl of NFW, the mixture was added to the sample mat on the test paper strip, and the result was read after 5-10 min.

2.2.8 Direct saliva-to-RICCA-to-LF assay

Heat-inactivated SARS-CoV-2 was obtained from ATCC (ATCC VR-1986HK; American Type Culture Collection, Manassas, VA, USA). The concentration of virus in lot #70037781 was estimated by the preinactivation titer (6.45 x 10^6 TCID50/mL), and the RNA copy number was determined by ddPCR (4.2 x 10^5 genome copies/µl). Prior to the RICCA reaction, the stock aliquot vial was centrifuged and serially diluted in phosphate-buffered saline (PBS) with replicates of 42,000, 4,200, 420, 42, and 4.2 copies/µl. Two microliters of diluted heat-inactivated SARS-CoV-2 virus was spiked with and without 3 µl of human saliva, which was freshly collected using a SalivaBio Oral Swab (SOS, Salimetrics LLC, USA). The 5-µl SARS-CoV-2 spiked samples were mixed with 0.5 µl of Proteinase K (Takara) in 0.2-mL nuclease-free and thin-walled microcentrifuge tubes by simple tapping or vortexing followed by heating in a portable heat block (Mini-Block Bath, MD-Mini,

AXEL, Japan) at 95°C for 5 min. To the lid of the tube, 10 µl of substrate stock solution, 10 µl of primer stock solution and 5 µl of enzyme stock solution were added, and the vial was then spun down the vial and incubated at 41°C for 15-30 min. Of the 30-µl reaction sample, 28 µl was utilized for the lateral flow (LF) assay, and 2 µl of the sample was used for gel electrophoresis analysis. To prevent possible contamination during liquid handling in the LF assay, we introduced a closed system by designing a device to continuously perform the RICCA reaction and LF assay without opening the RICCA tube. To this end, a prototype is fabricated using a 3D printer that can hold RICCA tubes and LF assay strips. In this design, after a tube is inserted and mechanically forced down by closing the device, a crack forms in the bottom of the tube, and the reaction solution flows out and moves up on the LF strip by capillary action. For the LF assay, 30 µl of the RICCA product was mixed with RNA detection reagents (1 µl of each 20 µM LF assay probe (Table 1), 15 µl of 1 M KCl, and 53 µl of nuclease-free water) in a 1.5-mL microcentrifuge tube, and the mixture was then incubated at 60°C for 2 min. An LF strip was inserted in the tube, and the test line was observed within 3 min.

2.3 RESULTS

2.3.1 RICCA concept

A workflow of the RICCA 'sample-in and result-out' approach using a lateral flow (LF) device is shown in **Fig. 2.1A**. First, a saliva sample is acquired from the subject using a saliva sample collection tube (e.g., SalivaBio Oral Swab, Salimetrics, LLC.). Second, the direct lysis of virus cells is performed by heating the saliva sample with RNA stabilizing reagent (protease) and/or mixing the saliva sample with detergent (e.g., 1% NP40 or 0.01% Triton X-100) containing buffer (phosphate-buffered saline, pH 7.4). Third, the lysed sample and the RICCA reagents were mixed to initiate a one-pot isothermal nucleic acid amplification reaction near 37°C using body heat or portable heat block. Finally, the results are visualized by the naked eye using a hand-held DNA LF device.

As shown in **Fig. 2.1B**, a key concept is the dual-function application of RNA and DNA cycles in the RICCA reaction, including a modified and advanced version of our previously developed RNA-specific amplification method [11,12], in which the formation of promoter-bearing dsDNA is essential for the template RNA to enter the RNA-specific amplification cycle. This step, however, is the rate-limiting step if the copy number of the target template RNA in the reaction is very low.

Furthermore, the possibility of failure in the formation of promoter-bearing dsDNA becomes severe due to the ready degradation of template RNA by the presence of exogenous nucleases in clinical samples such as saliva, which limits the RNA template that can enter the amplification cycle. Therefore, an alternative route for the formation of promoter-bearing dsDNA is essential and is provided in RICCA by the inclusion of RPA, which provides a bypass route and assists RNA amplification, similar to RPA by supplying T7 promoter-bearing dsDNA.

Therefore, in RICCA, coupled RNA-specific and DNA-specific amplification reactions co-assist each other by providing a favorable environment (i.e., specific template) to each other and synergistically improve the efficiency of both amplifications, i.e., RNA-specific amplification provides a DNA template to RPA, and RPA provides sufficient template for RNA-specific amplification. In the RICCA

protocol, the clinical sample of RNA viruses is added directly into a reaction tube containing all necessary components for the isothermal amplification of RNA-specific amplification and DNA-specific amplification.

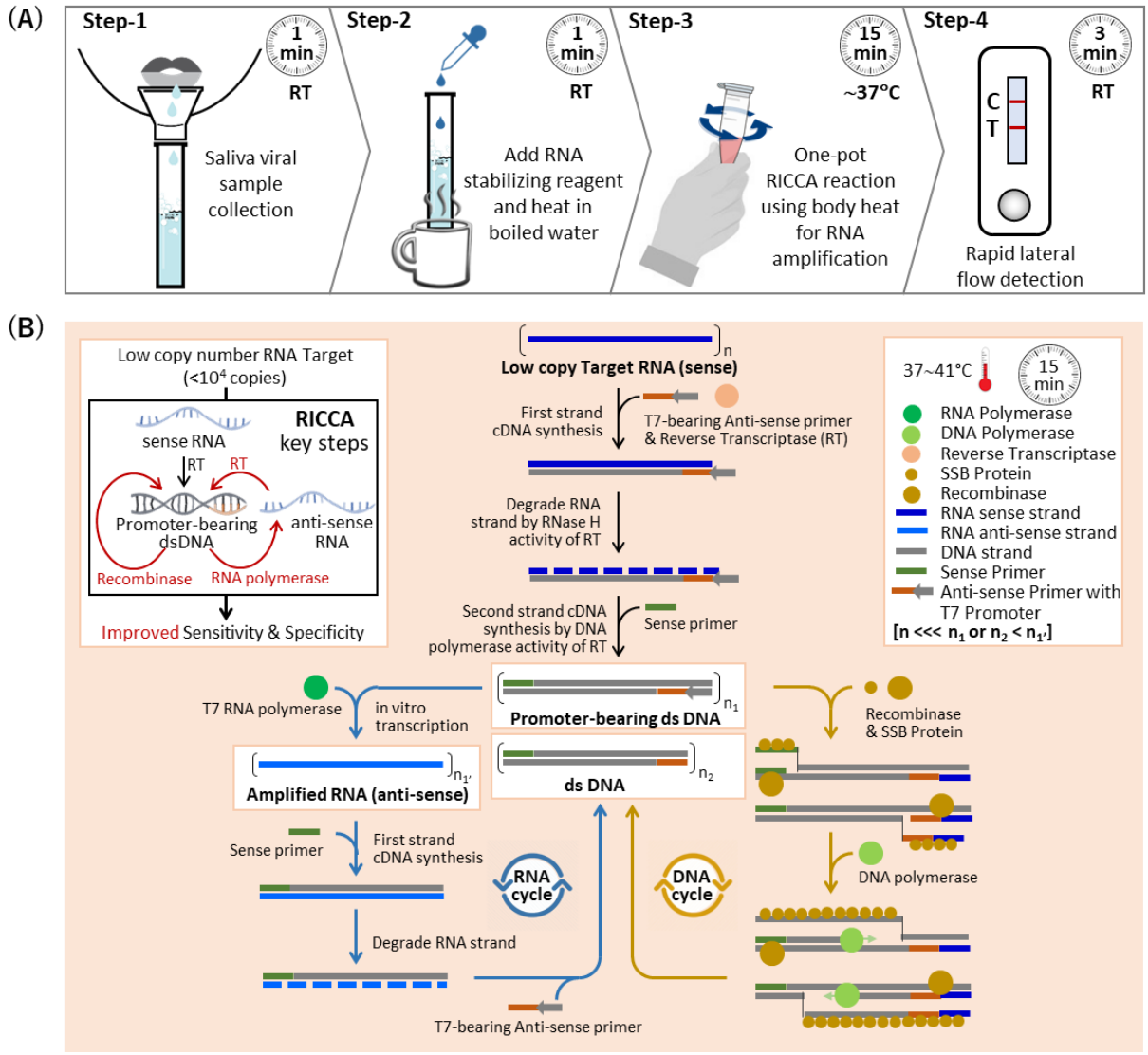


Figure 2.1: RICCA (Rapid Isothermal Co-Assisted Coupled Amplification of RNA and DNA) test concept and mechanism. (A) A working flowchart for the rapid, reliable and field-deployable molecular testing of RNA virus with a sample-to-answer period of less than 20 min using a lateral flow device-powered RICCA test. The four simple operation steps are as follows: a 1-min protocol for easy and painless saliva sampling is used for the collection of biological virus samples (step-1) and for the release of viral RNA (step-2). Next, the lysed solution is directly used for one-pot coupled isothermal RNA and DNA amplification reactions at a temperature near 37°C for 15-30 min (step-3). Qualitative results are visualized by the naked eye using a DNA lateral flow immunochromato graphic assay in 3 min (step 4). **(B)** Schematics of the mechanism and the key steps involved in the RICCA assay to overcome the inevitable challenging features of conventional isothermal amplification methods.

2.3.2 Demonstration of the RICCA concept for one-pot ultrasensitive RNA detection directly from cells

First, we demonstrated the field-deployable application of a one-pot RICCA reaction for 'sample-in and answer-out' testing. For this purpose, an initial assessment of the RICCA platform was performed by detecting RNA directly from nonclinical bacterial cells (i.e., a nonpathogenic strain of living *E. coli*) in the presence of human saliva (**Fig. 2.2A**). We designed sequence-specific primers to target a 204-base (or "nt") RNA fragment of the 16S rRNA gene from the *E. coli* bacterial genome. **Fig. 2.2B** shows the nucleotide sequence of this fragment. To amplify the antisense RNA, the T7 promoter sequence needed for RNA-specific amplification was added to the reverse primer. We then performed RNA-specific amplification using human saliva spiked with *E. coli* cells as templates in the range between $7.3X10^8$ and 8.6 CFU per µl. As shown in **Fig. 2.2C**, RICCA with RNA cycle (i.e., RNA-specific amplification only) produced an amplicon of 204 bases only in the reaction with the highest template concentration ($7.3X10^8$ CFU/µl). However, RICCA with an RNA plus DNA cycle produced amplicons of both RNA and DNA in all reactions, including the lowest template concentration (8.6 CFU/µl). Considering the presence of approximately 20,000 ribosomes in a single cell of *E. coli*, the detection limit was estimated to be less than 10^5 copies of template RNA in the presence of saliva. This one-pot reaction confirmed the utility of RICCA for field application with ultrahigh sensitivity and within 10 min of the start of the reaction. Importantly, no false-positive (nonspecific) signal was obtained with the RICCA reaction with the negative control sample (no template).

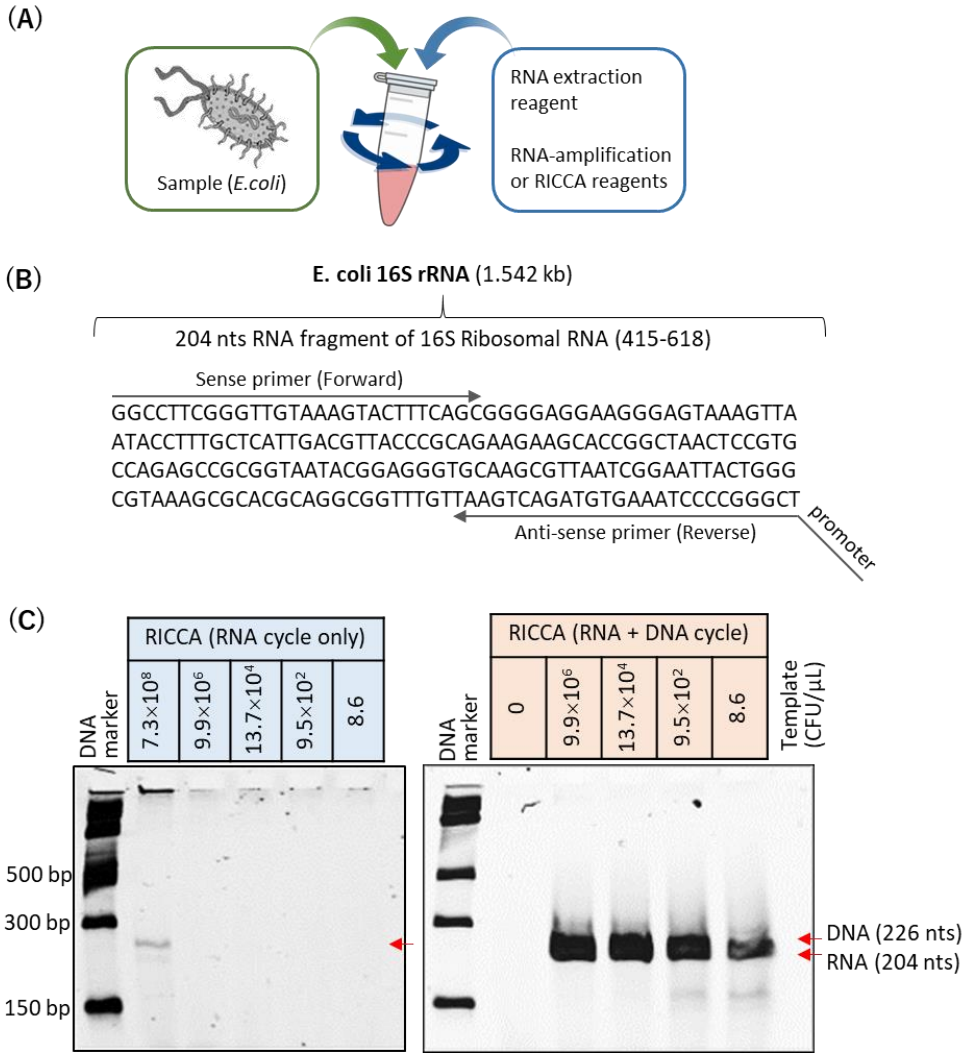


Figure 2.2: Demonstration of one-pot RICCA for field application. (A) Schematic view of the one-pot reaction protocol for RNA extraction to RNA/DNA amplification. **(B)** Target sequence of the *E. coli* genome for amplification. The sequence to which primers bind is indicated by arrows. **(C)** Electrophoretic analysis for comparatively evaluating the detection limits of RNA-specific amplification without (left) and with (right) DNA-specific amplification. *E. coli* cells at levels ranging from 10^9 to 10^1 were mixed with saliva (5% v/v) and used directly as the template for RNA extraction followed by RNA or DNA amplification. The desired amplified product (204-nt RNA or 226-bp DNA) is indicated by red arrows.

2.3.3 RNA-specific amplification assay (RICCA with RNA cycle only) for SARS-CoV-2

To investigate the application of RICCA for the detection of COVID-19, we started by exploring the effect on the efficiency of RNA-specific amplification for amplifying sense or antisense RNA from sense RNA as a template. For this purpose, we compared the sensitivity and specificity of RNA-specific amplification for amplifying the sense or antisense RNA of SARS-CoV-2. We designed sequence-specific primers to target a 134-bp fragment of the nucleocapsid phosphoprotein gene from the SARS-CoV-2 viral genome. As shown in **Fig. 2.3A**, the nucleotide sequence of this fragment. Standard RNA (406 bases) was prepared by in vitro transcription using cloned DNA. For the amplification of antisense or sense RNA, the T7 promoter sequence needed for RNA-specific amplification was added to the reverse or forward primer, respectively. We then performed RNA-specific amplification using 20 ng ($0.87X10^{11}$ copies) of SARS-CoV-2 RNA. As shown in **Fig. 2.3B**, a clear amplicon of 134 bases was observed in the amplification of antisense RNA. The RNA-specific amplicon was further confirmed by postreaction treatment of the amplified product with DNase or RNase. However, this was not the case with the amplification of sense RNA, and a very faint band was observed with the gel electrophoretic analysis. A possible explanation of this failure is demonstrated in **Fig. 2.4**. Based on these results, we chose the primer set to amplify antisense RNA for the RICCA reaction. Noticeably, the appearance of a band of template RNA (406 base) was visible with sense RNA amplification, which indicated the inefficiency of the template for full utilization in the reaction. We then examined the limit of detection (LOD) of RNA-specific amplification of antisense RNA using serial dilutions of 20 ng ($0.87X10^{11}$ copies) to 20 fg ($0.87X10^{5}$) of SARS-CoV-2 RNA standard template spiked into a 15-µl reaction. As shown in **Fig. 2.3C**, RNA-specific amplification produced visible amplicons of 134 bases at decreasing intensities with decreasing template copies. This result demonstrated the lowest LOD of 20 pg ($0.87X10^{8}$) within 10 min of the start of the amplification reaction.

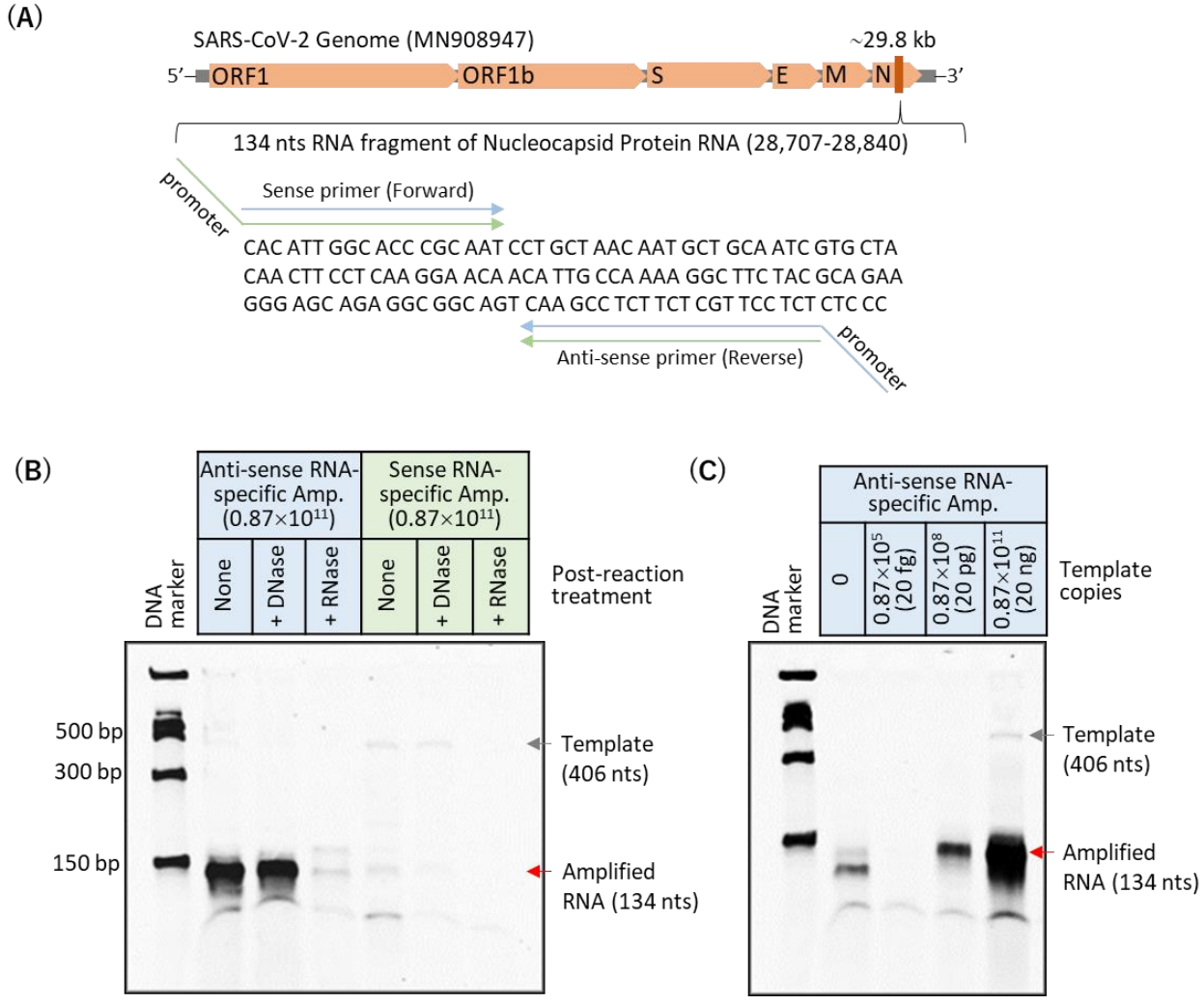


Figure 2.3: Evaluation of the specificity of RNA-specific amplification of sense or antisense strands of SARS-CoV-2 RNA. (A) Structure of the SARS-CoV-2 genome and a target sequence for amplification. The sequence to which primers bind is indicated by arrows. **(B)** Electrophoretic analysis of RNA-specific amplification of antisense or sense strands in 10 min. The desired amplified product (134 nt RNA) is shown by the red arrow and clearly observed with the amplification of the antisense strand of RNA. RNA-specific amplicons were confirmed by postreaction treatment with RNase-free DNase or RNase (RNase ONE and RNase H). **(C)** Electrophoretic analysis for evaluating the detection limit of RNA-specific amplification using SARS-CoV-2 RNA template at a level ranging from 10^{11}-10^{5} and a zero copy number with a reaction incubation time of 10 min.

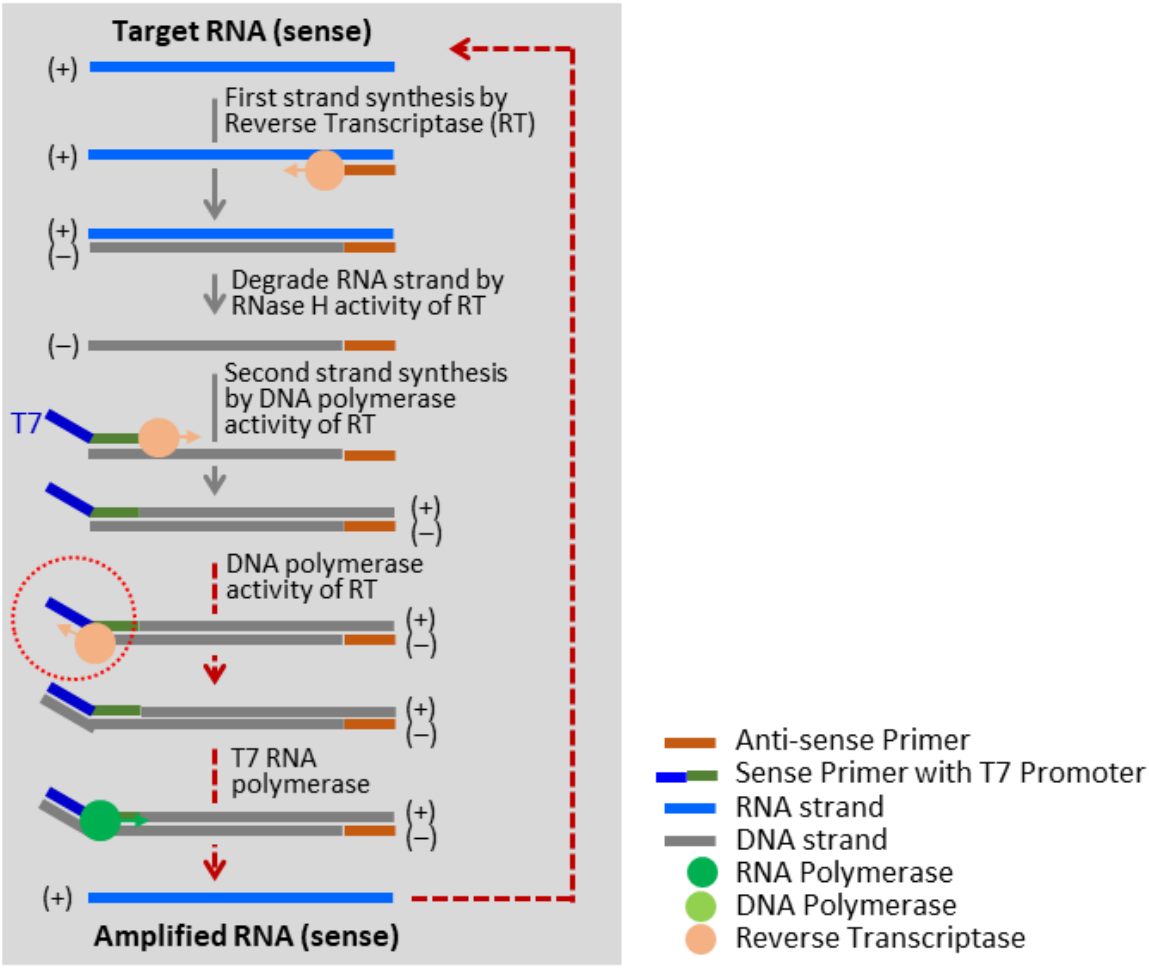


Figure 2.4: Schematics of the mechanism and the key steps involved in RNA-specific amplification. The formation of promoter-bearing dsDNA which is essential for the template RNA to enter into the RNA-specific amplification cycle is synthesized inefficiently while targeting the synthesis of sense RNA strand from sense RNA template (shown by red arrows).

2.3.4 Optimization and sensitivity evaluation of RNA cycle in RICCA for SARS-CoV-2

To improve the efficiency of the antisense RNA-specific amplification reaction, we optimized (i) the temperature, (ii) the reaction time, (iii) the selectivity for the targeted template, and (iv) the inhibitory effect of saliva on the reaction efficiency. To test the effect of temperature, we performed RNA-specific amplification at 37°C, 39°C, 41°C and 43°C for 10 and 20 min each. As shown in **Fig. 2.5A,** amplification occurred at all temperatures, and 41°C appeared to be optimal for the specific target of the SARS-CoV-2 genome. Furthermore, increasing the time from 10 to 20 min increased the yield of amplicons many-fold. Therefore, we further examined the effect of longer reaction times on higher yields. As shown in **Fig. 2.5B**, a higher increase in the yield of the specific amplicon was consistently observed after 30 min of reaction. However, nonspecific amplicons of higher and lower molecular weights were observed only with the 30-min reaction if the reaction did not include template. This result is consistent with the fact that nontarget-triggered amplification due to a lower reaction temperature in isothermal amplification is inevitable and magnified with increases in the reaction time. To further investigate this issue, we examined the efficiency of RNA-specific amplification to amplify extremely diluted template copies ranging from 200 fg ($0.87X10^6$ copies per reaction) to 200 ag (870 copies per reaction or 29 copies/µl) versus longer reaction time. Our results suggested that the yield of a specific amplicon with a lower copy number of template increased with increases in the reaction time but also resulted in nonspecific amplicons; thus, a longer incubation time for the RNA-specific amplification reaction should be avoided, particularly if the sample has a low viral load. Additionally, we noticed that the success of reaction using lower copy number of template (29 copies/µl) was highly depending on the freshly prepared synthetic RNA template and contamination-free experiments since the lower number of smaller size of RNA template can be readily degraded and removed from the system. Subsequently, to examine the primer selectivity for the targeted template of SARS-CoV-2 RNA in RNA-specific amplification, we performed the reaction with a template from a different source (non-SARS-CoV-2). As shown in **Fig. 2.5C**, no

specific amplicons were observed when the SARS-CoV-2 primers were used to amplify non-SARS-CoV-2 template copies in the range between 20 pg and 20 fg. In contrast and as expected, the analysis of the yield of specific amplicons increased with increases in the template copies of SARS-CoV-2 RNA revealed that the signals of the desired amplicon (134 nt) could be clearly observed using as little as 20 fg (1,740 copies/μl) within 10 min of the start of the reaction.

Then, given the bottleneck issues in the handling of nasopharyngeal swabs and with the current growing interest in testing COVID-19 using saliva, we examined whether the presence of digestive enzymes such as RNases in human saliva can readily degrade intact SARS-CoV-2 RNA. Therefore, to adapt RNA-specific amplification to direct saliva samples, we tested the efficiency of RNA-specific amplification in the presence of human saliva. As shown in **Fig. 2.5D**, the results indicated that the presence of human saliva in the reaction exerts pronounced inhibitory effects on the RNA amplification process. The inclusion of more than 10% v/v saliva to amplify 20 ng of template RNA induced almost 100% inhibition of the reaction, and minimal inhibition (below 66%) was observed at concentrations below 5% saliva in the RNA-specific amplification reaction. To further improve the sensitivity by utilizing higher concentrations of saliva, we followed the Saliva-Direct approach [28], which was recently reported to bypass the need for RNA extraction (isolation and purification). Interestingly, we were able to perform successful RNA-specific amplification reactions with the inclusion of a maximum of 10% human saliva using our modified protocol.

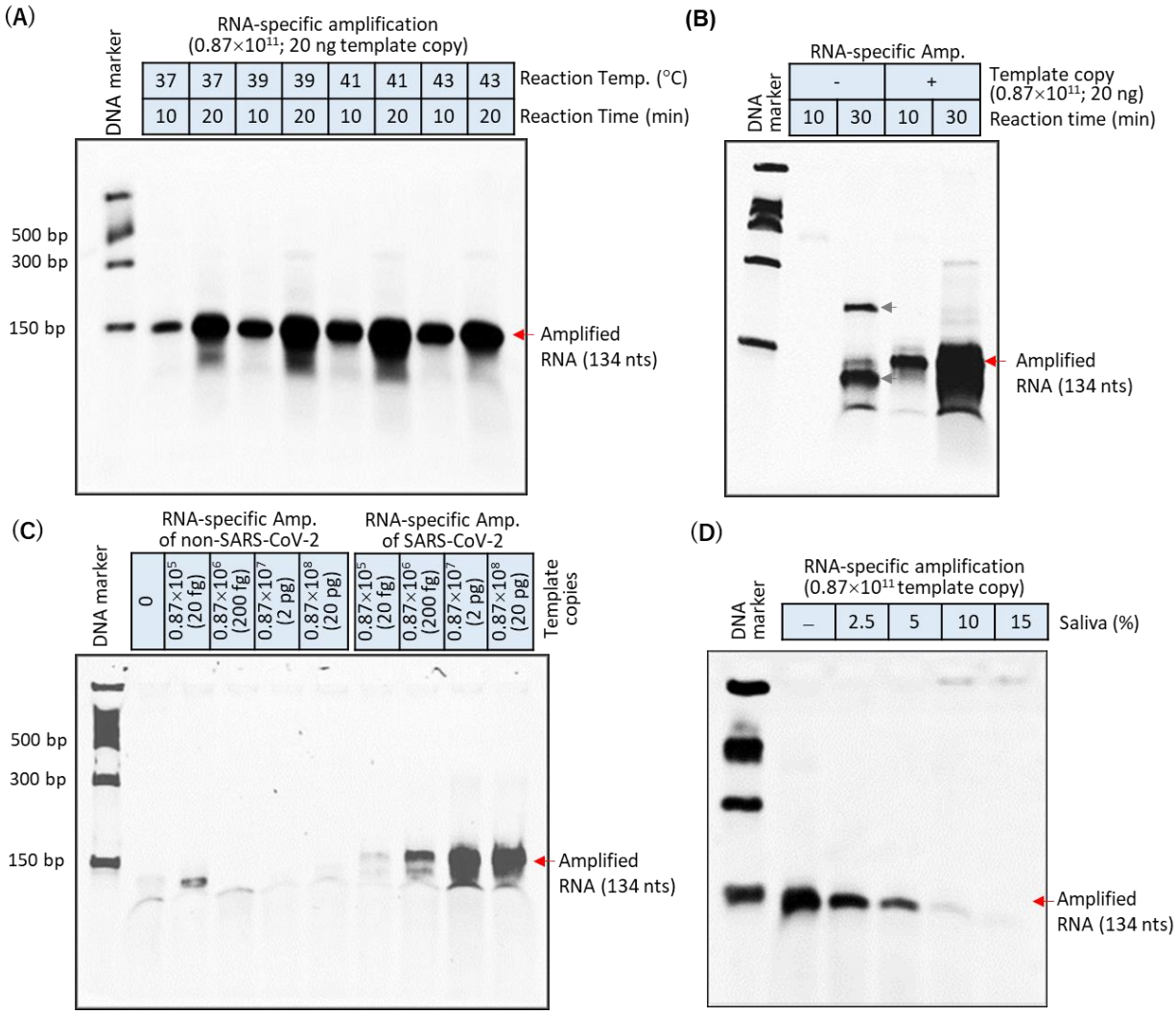


Figure2.5: Optimization of antisense strand RNA amplification. (A) The effect of temperature in the range of 37-43°C on the reaction efficiency of RNA amplification was checked at 10 and 20 min. **(B)** The specificity of RNA amplification was evaluated with high (10^{11}) or no template copy numbers and shorter (10 min) or longer (30 min) reaction times. **(C)** Evaluation of the selectivity of RNA amplification for the detection of the targeted template (SARS-CoV-2) over a longer template (non-SARS-CoV-2) at a copy number range of 10^5-10^8 with a 10-min reaction incubation time. **(D)** The inhibition effect of saliva on the reaction efficiency of RNA amplification was tested in the range of 2.5 to 15% saliva in the reaction. Red and gray arrows indicate true-positive (134 nt) and false-positive amplified products, respectively.

2.3.5 Evaluation of the demonstration of the RICCA concept for SARS-CoV-2

To overcome the challenge of low sensitivity in RNA-specific amplification if the template copy number is excessively low (below picograms), we performed a comparative analysis to demonstrate the advantage of the DNA cycle in the RICCA reaction. We sequentially performed reactions by excluding and including a DNA-specific amplification step in the RICCA assay to amplify low template copy numbers ranging from 20 ng ($0.87X10^{11}$) to 20 fg ($0.87X10^{5}$). A clear difference between RICCA with RNA cycle only and RICCA with RNA plus DNA cycle could be observed by electrophoretic analysis (**Fig. 2.6A**). The desired amplicon (134 base) in the electrophoretic analysis was observed using RICCA with an RNA cycle only with high template copies (20 ng); however, for all cases, RICCA with an RNA plus DNA cycle could produce visible amplicons (134-base RNA amplicon and 156-base DNA amplicon) at intensities that decreased with decreases in the template copies. This result demonstrated the lowest LOD for RICCA, and 20 fg (87,000) was amplified within 10 min of the 50 µl-scale reaction start, which corresponds to 1,740 template copies per microliter. The efficiency of RICCA for equally producing both DNA-specific and RNA-specific amplicons was further confirmed by postreaction treatment of the amplified product with DNase or RNase, respectively (**Fig. 2.6B**). For further evaluation, the amplicon from the RICCA reaction was examined by DNA sequencing and confirmed as in **Fig. 2.7**.

Next, to check the efficiency of RICCA in the presence of saliva, we performed a comparative LOD evaluation for RICCA with RNA cycle only and RICCA with RNA plus DNA cycle in the absence or presence of saliva. For this purpose and to ensure the absence of traces of carryover template contamination, we freshly synthesized a new RNA template for another species of human coronavirus, i.e., CoV-229E, which is one of the viruses responsible for the common cold. As shown in **Fig. 2.6C**, electrophoretic analysis confirmed that RNA-specific amplification was not successful only in the case of reaction with saliva; however, including DNA-specific amplification with RNA-specific amplification in RICCA could produce the desired amplified products (135-base RNA and 129-base, 163-base DNA) with no difference between the presence and absence of saliva. This

finding surely demonstrates that a co-assisted and coupled approach involving RNA and DNA cycles in RICCA can amplify lower copies of RNA template even in the inhibitory presence of oral saliva at levels as high as 10% v/v in reaction. To further demonstrate and evaluate the co-assisting nature of RNA-specific and DNA-specific amplification, RICCA was performed in the presence and/or absence of both primer sets for both RNA-specific and DNA-specific amplification reactions. As shown in **Fig. 2.6D**, RICCA was not successful in the absence of any of the individual primer sets, which further confirms that promoter-bearing second-strand DNA is essential to seed the RICCA reaction. In addition to sensitivity, the specificity of the RICCA test is also an important indicator because the proportion of coinfections is highly expected in COVID-19. Therefore, to examine and ensure the absence of cross-reactivity, primer sets for SARS-CoV-2 and CoV-229E were used for the RNA-specific amplification of CoV-229E and SARS-CoV-2, respectively. The results confirmed the absence of cross-reactivity between SARS-CoV-2 and CoV-229E and indicated the robustness of the RICCA assay, as demonstrated by the very high specificity of the designed primers and template listed in **Table 2.1 and Figure 2.8**.

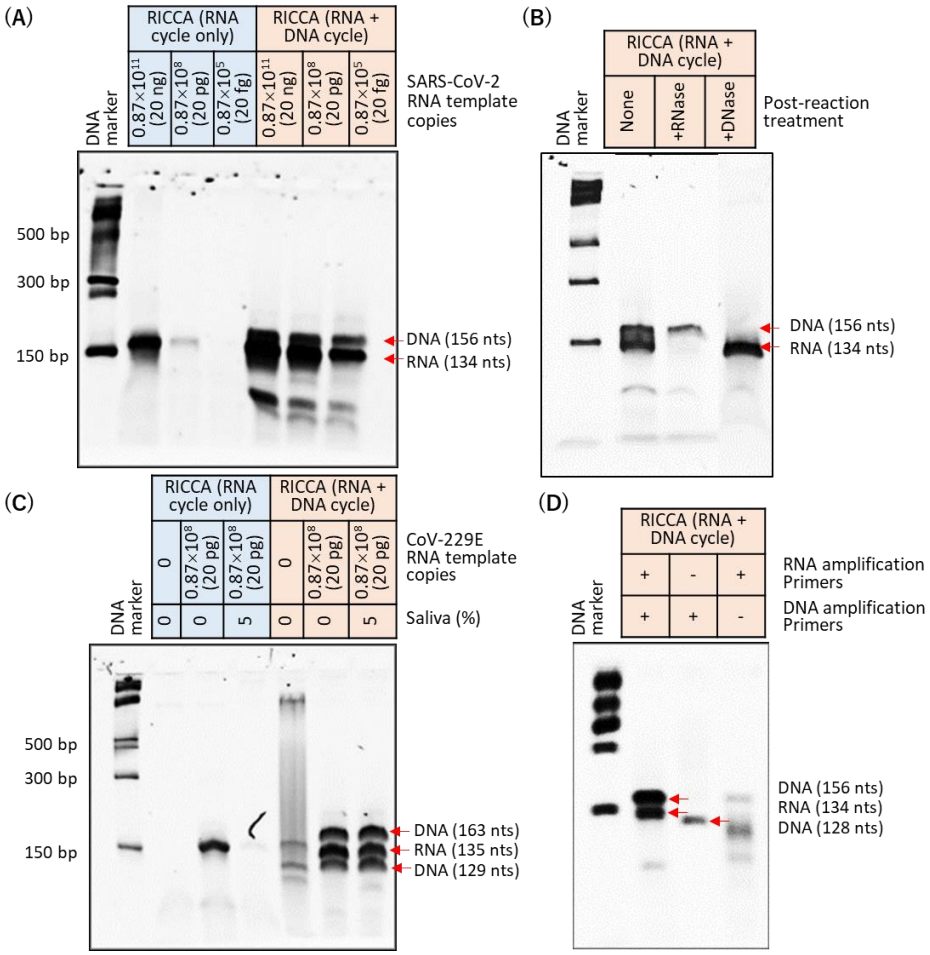


Figure 2.6: Proof-of-concept demonstration of RICCA. (A) Electrophoretic analysis to comparatively evaluate the detection limit of RNA-specific amplification only versus RNA-specific amplification with DNA-specific amplification to amplify low template copy numbers ranging from 20 ng to 20 ag. The desired amplified products (134-nt RNA and 156-bp DNA) are shown by red arrows in the gel electrophoretic analysis. **(B)** RNA- or DNA-specific amplicons were confirmed by treatment with DNase or RNase, respectively. **(C)** Electrophoretic analysis to compare the detection limit of RNA-specific amplification only versus RNA-specific amplification with DNA-specific amplification in the absence or presence of saliva using a freshly synthesized template for CoV-229E RNA. The desired amplified product (135-nt RNA or 163-bp DNA) is indicated by red arrows. **(D)** Electrophoretic analysis to evaluate the co-assisted efficiency of RNA- and/or DNA-specific amplification during the RICCA reaction in the absence and/or presence of primers for RNA and/or DNA-specific amplifications.

```
Template source   CACATTGGCACCCGCAATCCTGCTAACAATGCTGCAATCGTGCTACAACTTCCTCAAGGA
RICCA amplicon    ----------------------------NNNGCAATCGTGCTACAACTTCCTCAAGGA
                                               . .**************************

Template source   ACAACATTGCCAAAAGGCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCT
RICCA amplicon    ACAACATTGCCAAAAGGCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCT
                  ************************************************************

Template source   CGTTCCTCTCTCCC----------------------
RICCA amplicon    CGTTCCTCTCTCCCTATAGTGAGTCGTATTAGAATTA
                  **************
```

Figure 2.7: The sequence alignment of the SARS-CoV-2 RICCA amplicon with the actual source of template SARS-CoV-2. The RICCA amplicon was amplified using 20 ag template copy number.

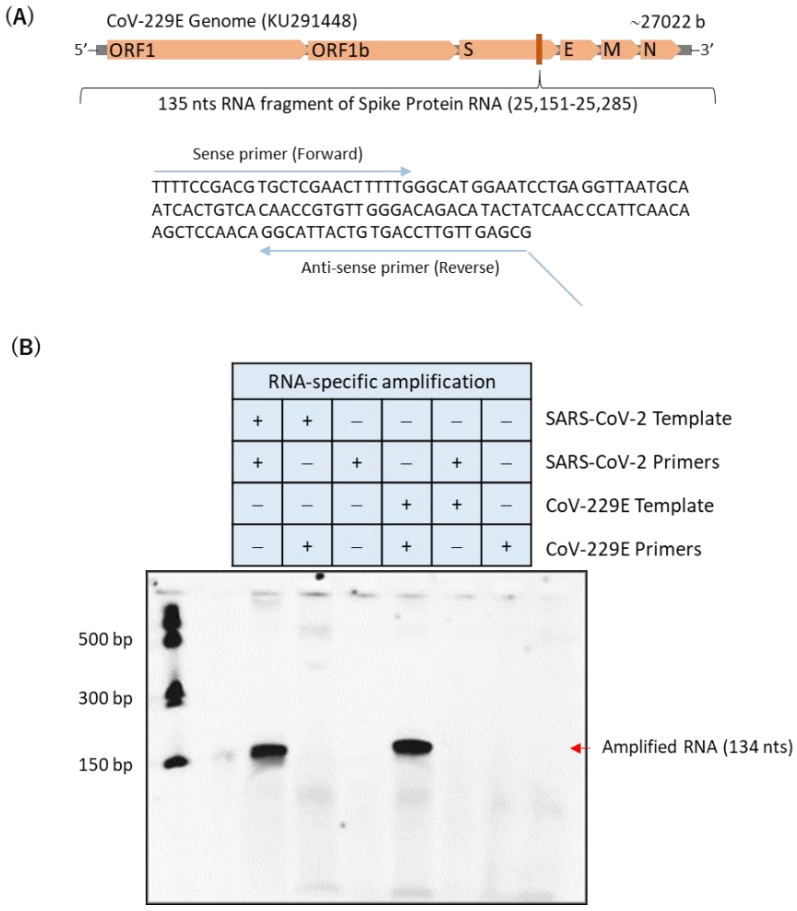


Figure 2.8: Electrophoretic analysis to check cross-reactivity in RNA-specific amplification using template and primer sets of SARS-CoV-2 and/or CoV-229E. (A) Structure of CoV-229E genome and a target sequence for amplification. The sequence to which primers bind are indicated by arrows. (B) Electrophoretic analysis of RNA-specific amplification of anti-sense amplicon in 10 minutes.

2.3.6 Development of a direct saliva-to-RICCA-to-LF assay for lab-free testing of SARS-CoV-2 virus

To demonstrate the RICCA approach for lab-free diagnosis, we followed the saliva-direct approach [28], which was recently reported to bypass the need for an RNA extraction (isolation and purification) step, and a commercial lateral flow (LF) assay that allows rapid and accurate read-out by the naked eye. For the saliva-to-RICCA reaction, we designed a one-pot assay using sequential combinations of multiple components in a one-pot reaction approach.

As shown in **Fig. 2.9A**, we used a 0.5-mL screw-cap Eppendorf tube with a total of three screw caps consisting of the necessary amounts of RNA stabilizing reagent, RNA amplification reagent and RNA detection reagent to run a single test reaction. For long-term storage, co-dried reagents can be stored in the cap using a glass fiber matrix [29]. For saliva collection, a simple method with a SalivaBio Oral Swab (SOS, Salimetrics LLC, USA) was used [30]. Ten percent human saliva spiked with heat-inactivated SARS-CoV-2 virus was first mixed with RNA stabilizing reagent (cap #1) and heated for 2-5 min at 80-95°C. Heating at 95°C demonstrated the remarkable robustness of the procedure. The saliva samples were then mixed thoroughly with RNA amplification reagent (cap #2) by sequentially replacing the caps followed by incubating for 15-30 min at 41°C. Finally, cap #3 was used to introduce the RNA detection reagent into the reaction mixture.

For the LF assay, we designed a pair of digoxigenin-labeled DNA probes and a biotin-labeled DNA probe to enable hybridization with RNA amplicons at the 5'-terminal end and 3'-terminal end, respectively.

For high sensitivity, we opted to use an LF assay with a streptavidin patterned line and gold nanoparticle-labeled anti-digoxigenin antibodies for visualization. Subsequently, to demonstrate the RICCA approach for contamination-free (cross-contamination between specimens or synthetically derived PCR amplicon contamination) diagnosis [31], we aimed to newly design the so-called closed system for the saliva-to-RICCA-to-LF assay, which can avoid aerosol contamination during liquid handling in the LF assay.

As shown in **Fig. 2.9B**, the RICCA vial and LF strip are placed inside the LF device, and the device is closed by tightening the upper lid. As a result, the bottom side of the RICCA vial approaches an inbuilt blade at the bottom of the LF device and cracks, which allows the RICCA reaction mixture to come out and move up on the LF strip by capillary action. As a result, a positive test reaction showed the appearance of a band (test line) within 3 min for. As shown in **Fig. 2.9A'**, RNA amplicons yielded a clear band in the RICCA-to-LF assay with ultrahigh sensitivity, and a concentration as low as 2.8 copies/µl of SARS-CoV-2 virus could be detected. We then demonstrated the application of a closed system using an LF device and tested the sensitivity in the presence of 10% human saliva (**Fig. 2.9B')**. Interestingly, we found that the result from the direct saliva-to-RICCA-to-LF assay is fortunate and important and exhibits high sensitivity with levels as low as 2.8 copies/µl of SARS-CoV-2 virus (true-positive control reaction) and high specificity (true-negative control reaction). We further confirmed the results by analyzing the RNA amplicons with a size of 164 nt by gel electrophoresis (**Fig. 2.10**). These results demonstrate the remarkable simplicity and robustness of the RICCA assay for the lab-free and rapid testing of COVID-19.

2.3.7 Early phase clinical trials

Total 47 (25 positive, 21 Negative and 1 Unknown) clinical samples were received from SMS Medical College, Jaipur, India and were tested with RICCA-RNA amplification kit and compared with the results from qRT-PCR. A 87% positive agreement and a 81% negative agreement were obtained. Low agreement with negative samples may be due to difference in results analysis (quantitative of RT-PCR and qualitative (end product) of our system). A 100% positive agreement and a 100% negative agreement were obtained when the samples were compared with in-house RT-PCR method using PAGE analysis. **Table 2.2** shows the test results of SARS-CoV-2 eluted RNA samples from SMS Medical College.

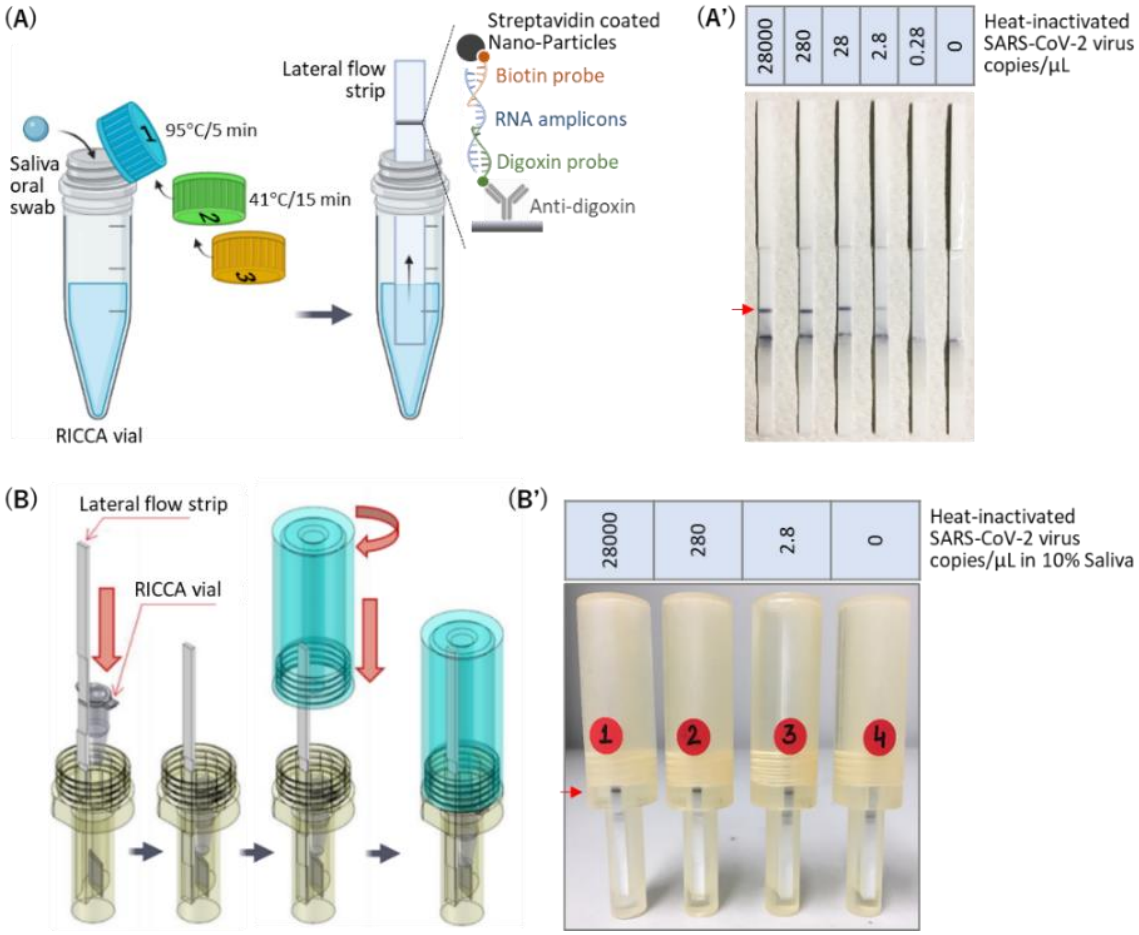


Figure 2.9: Direct saliva-to-RICCA-to-LF assay for the lab-free testing of SARS-CoV-2 virus. Schematic view of the assay (A) and the closed system (B). Detection of the heat-inactivated SARS-CoV-2 virus in the absence (A') and presence (B') of 10% human saliva using a lateral flow assay. The true amplified products are indicated by red arrows.

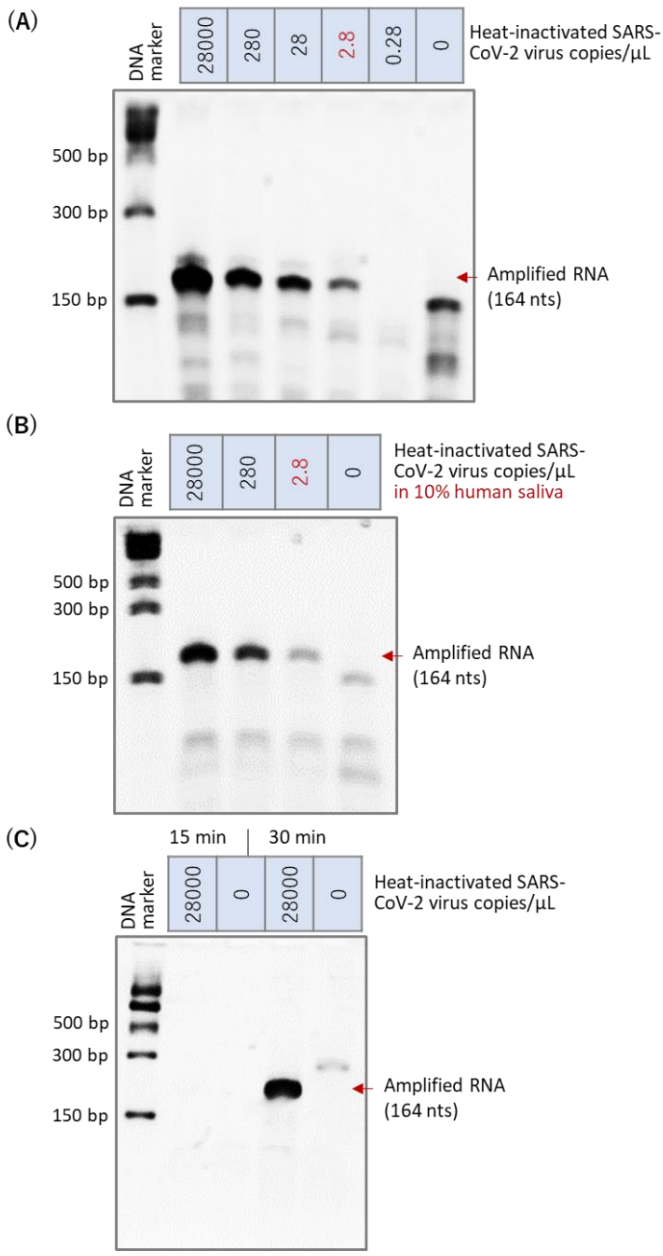


Figure 2.10: Electrophoretic analysis to check sensitivity and specificity of direct saliva-to-RICCA assay. Electrophoretic analysis to comparatively evaluate the detection limits of RNA-specific amplification of SARS-CoV-2 virus in absence (A) and presence (B) of 10% human saliva in 30 minutes of reaction time. (C) The effect of reaction time on the reaction efficiency of RNA amplification with and without SARS-CoV-2 virus template was checked at 15 and 30 min.

Table 2.2: Clinical trials test results of SARS-CoV-2 eluted RNA samples from SMS Medical College.

S.No.	Sample			Results				
	Date	Code	Ct-value	Ours (RNA Amp.)	RT-PCR (qPCR)	Match	RT-PCR (PAGE)	Match
1	29-Jan	NA	NA	Pos.	Pos.	o	-	-
2	29-Jan	NA	NA	Pos.	Pos.	o	-	-
3	29-Jan	NA	NA	Neg.	Neg.	o	-	-
4	29-Jan	NA	NA	Neg.	Neg.	o	-	-
5	09-Feb	NA	NA	Pos.	Pos.	o	-	-
6	09-Feb	NA	NA	Neg.	Neg.	o	-	-
7	09-Feb	NA	NA	Pos.	Pos.	o	-	-
8	09-Feb	NA	NA	Neg.	Neg.	o	-	-
9	25-Feb	142	14-22	Pos.	Pos.	o	-	-
10	25-Feb	143	14-20	Pos.	Pos.	o	-	-
11	25-Feb	144	NA	Neg.	Neg.	o	-	-
12	25-Feb	505	22	Pos.	Pos.	o	-	-
13	25-Feb	506	20-22	Pos.	Pos.	o	-	-
14	25-Feb	507	20	Pos.	Pos.	o	-	-
15	8-Mar	64	32	Neg.	Pos.	×	-	-
16	8-Mar	65	24-26	Neg.	Pos.	×	-	-
17	8-Mar	41	33	Neg.	Pos.	×	-	-
18	8-Mar	125	NA	Neg.	Neg.	o	-	-
19	8-Mar	75	20	Pos.	Pos.	o	-	-
20	8-Mar	63	15	Pos.	Pos.	o	-	-
21	8-Mar	117	NA	Neg.	Neg.	o	-	-
22	8-Mar	32	18-20	Pos.	Pos.	o	-	-
23	8-Mar	82	32	No elute	Pos.	-	-	-
24	8-Mar	42	32	Pos.	Pos.	o	-	-
25	8-Mar	2	31	Pos.	Pos.	o	-	-
26	8-Mar	50	28	No elute	Pos.	-	-	-
27	8-Mar	123	NA	Neg.	Neg.	o	-	-
28	8-Mar	124	NA	Neg.	Neg.	o	-	-
29	8-Mar	118	NA	Neg.	Neg.	o	-	-
30	27-Mar	95	24-28	Pos.	Pos.	o	-	-
31	27-Mar	97	27-29	Pos.	Pos.	o	-	-
32	27-Mar	128	26-32	Pos.	Pos.	o	-	-
33	27-Mar	137	19-25	Pos.	Pos.	o	Pos.	o
34	27-Mar	139	19-25	Pos.	Pos.	o	-	-
35	27-Mar	1	NA	Neg.	Neg.	o	-	-
36	27-Mar	2	NA	Pos.	Neg.	×	Pos.	o
37	27-Mar	3	NA	Pos.	Neg.	×	-	-
38	30-Mar	06	NA	Neg.	Unknown	o	-	-
39	30-Mar	01	NA	Neg.	Neg.	o	-	-
40	30-Mar	02	NA	Neg.	Neg.	o	-	-
41	30-Mar	03	NA	Neg.	Neg.	o	-	-
42	30-Mar	04	NA	Neg.	Neg.	o	-	-
43	30-Mar	05	NA	Pos.	Neg.	×	-	-
44	30-Mar	5377	14-19	Pos.	Pos.	o	-	-
45	11-May	CMHO	NA	Pos.	Neg.	×	Pos.	o
46	11-May	Dr.-1	NA	Neg.	Neg.	o	Neg.	o
47	11-May	Dr.-2	NA	Neg.	Neg.	o	Neg.	o

2.3.8 Advancing studies by *in-silico* approach

One of the major challenges in the developed system is in the process of primer designing. In conventional PCR-based system, primer designing is comparatively easy because of the high-temperature being used that can easily unfold the DNA to make the interaction between primer binding site and primers very easily possible. However, in isothermal amplification-based system (41°C in our system), the primer binding sites on the template (genomic RNA) are not easily exposed for the primers to come and interact. This process is governed by enzymes that regulate the process of primer binding to the template. This makes primer designing, a very crucial, yet challenging part of RICCA.

Along with criteria for primer designing in conventional methods-based systems that includes homology with other sequences, primer length, GC content, 5' and 3' end composition, possibility of dimer formation and T_m of forward/reverse primers, additional factors responsible for primer designing in RICCA includes SSB protein gp32 and loading factors Uvsy and Uvsx.

The in-silico approach is being recently used for primer designing [33] and thus can be utilized and applied for primer designing in RICCA. RICCA can be performed for any kind of target with just one additional step of primer designing, and this can make RICCA applicable for different viruses.

2.4 CONCLUSIONS

Interaction between nucleic acid and proteins has been studied in this chapter. We introduced and demonstrated a robust isothermal amplification assay, termed the RICCA, that enables a simple one-pot 'sample-in and answer-out' format for the ultrarapid detection of low copy numbers of RNA virus directly from saliva, with a total assay time within 30 min and with a sensitivity as low as few copies of SARS-CoV-2 virus per microliter. The RICCA assay requires only a heat block and a tube prefilled with a mixture containing the necessary reagents (substrate, primer, and enzyme) to simultaneously perform RNA-specific and DNA-specific amplifications. We estimated that the material cost for in-house reagents for RICCA is very affordable (RNA-specific amplification reagents cost USD 2.6 per reaction, and DNA-specific amplification reagents cost USD 1.3 per reaction); thus, RICCA has the potential to decentralize COVID-19 molecular diagnosis to realize a 'lab-free, lab-quality' mega-testing platform. To evaluate the rapidity of RICCA, a real-time analysis was used to assess the performance of real-time RICCA in a reference with conventional real-time RT-PCR approach (**Fig. 2.11**). A cycle threshold (Ct) value for RICCA was significantly faster than Ct value for RT-PCR. These results indicated that the RICCA reaction speed was excellent for rapid diagnosis. Higher fluorescence for RICCA control than RICCA positive after longer times can be explained as false positive signals because of non-specific amplification in the absence of template, which is a drawback of isothermal amplifications methods. Thus, this system is designed for short reaction times. Also, it can be observed that RICCA positive curve starts as the rising part of the curve indicating early start of the reaction. This indicates high specificity of the system that can start amplification as soon as the template is added to the reaction.

In overall conclusion, we have devised a novel approach RICCA using an improved isothermal RNA amplification method. In order to show the potential of RICCA to contribute to COVID-19 diagnosis, early phase clinical trials have already been performed at hospitals in India and Japan and trials using a large number of clinical samples are now in the planning. Following this, we hypothesized that a robotic and mobile platform equipped with an automatic dispenser unit (a desk-top prototype is

under development) for RICCA could be applied to taking COVID-19 diagnostics directly to consumers by mitigating the logistical burden of sample transport and could provide large-scale population testing in remote or resource-limited settings. More importantly, this platform could facilitate the expansion of COVID-19 testing to other infectious viral or bacterial strains with only an additional step of primer design.

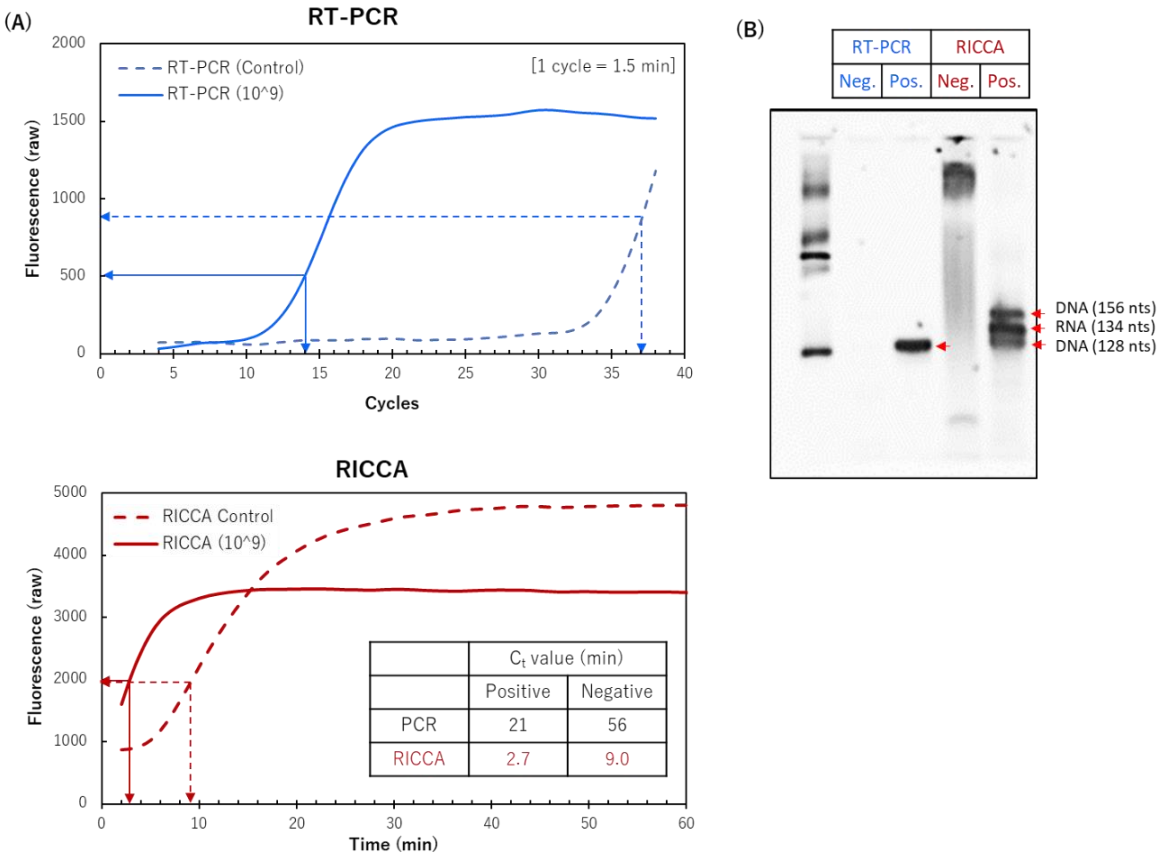


Figure 2.11: Comparative analysis of real-time amplification of the RT-PCR and RICCA reaction. RT-PCR and RICCA were performed with (20 pg template for CoV-229E RNA) or without template. To determine the reaction efficiency (rapidity), the Ct values are plotted against the reaction time. (A) A Ct value for RICCA was observed at a threshold time near 3 min, which was significantly 7-times faster than RT-PCR (threshold time near 21 min). (B) The end products for all the reactions were also confirmed by gel electrophoretic analysis.

REFERENCES

1. Ferretti, L. *et al.* Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, (2020)
2. Salathé, M. *et al.* COVID-19 epidemic in Switzerland: on the importance of testing, contact tracing and isolation. *Swiss Med. Wkly* **150**, w20225 (2020)
3. Vandenberg, O., Martiny, D., Rochas, O., van Belkum, A. & Kozlakidis, Z. Considerations for diagnostic COVID-19 tests. *Nat. Rev. Microbiol.* **19**, 171–183 (2021)
4. Relova, D. *et al.* Impact of RNA Degradation on Viral Diagnosis: An Understated but Essential Step for the Successful Establishment of a Diagnosis Network. *Vet. Sci.* **5**, 19 (2018)
5. Piepenburg, O., Williams, C. H., Stemple, D. L. & Armes, N. A. DNA detection using recombination proteins. *PLoS Biol.* **4**, e204 (2006)
6. Lobato, I. M. & O'Sullivan, C. K. Recombinase polymerase amplification: Basics, applications and recent advances. *Trends Analyt. Chem.* **98**, 19–35 (2018)
7. Li, J., Macdonald, J. & von Stetten, F. Review: a comprehensive summary of a decade development of the recombinase polymerase amplification. *Analyst* **144**, 31–67 (2018)
8. Kojima, K. *et al.* Solvent engineering studies on recombinase polymerase amplification. *J. Biosci. Bioeng.* **131**, 219–224 (2021)
9. Juma, K. M. *et al.* Optimization of reaction condition of recombinase polymerase amplification to detect SARS-CoV-2 DNA and RNA using a statistical method. *Biochem. Biophys. Res. Commun.* **567**, 195-200 (2021)
10. Kievits, T. *et al.* NASBA^{TM} isothermal enzymatic in vitro nucleic acid amplification optimized for the diagnosis of HIV-1 infection. *J. Virol. Methods* **35**, 273–286 (1991)
11. Ishiguro, T. *et al.* Intercalation activating fluorescence DNA probe and its application to homogeneous quantification of a target sequence by isothermal sequence amplification in a closed vessel. *Anal. Biochem.* **314**, 77–86 (2003)
12. Yasukawa, K., Agata, N. & Inouye, K. Detection of *cesA* mRNA from *Bacillus cereus* by RNA-specific amplification. *Enzyme Microb. Technol.* **46**, 391–396 (2010)

13. Notomi, T. *et al.* Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res.* **28**, e63–e63 (2000)

14. Lalli, M. A. *et al.* Rapid and extraction-free detection of SARS-CoV-2 from saliva by colorimetric reverse-transcription loop-mediated‒isothermal amplification. *Clin. Chem.* **67**, 415–424 (2020).

15. Anahtar, M. N. *et al.* Clinical assessment and validation of a rapid and sensitive SARS-CoV-2 test using reverse transcription loop-mediated isothermal amplification without the‒need for RNA extraction. *Open Forum Infect. Dis.* **8**, ofaa631 (2020)

16. Yang, Q. *et al.* Saliva TwoStep for rapid detection of asymptomatic SARS-CoV-2 carriers. *Elife* **10**, e65113 (2021)

17. Xia, S., Chen, X. Single-copy sensitive, field-deployable, and simultaneous dual-gene detection of SARS-CoV-2 RNA via modified RT–RPA. *Cell Discovery* **6**, 1–4 (2020)

18. Zou, Y., Mason, M. G., Botella, J. R. Evaluation and improvement of isothermal amplification methods for point-of-need plant disease diagnostics. *PloS One* **15**, e0235216 (2020)

19. Li, J., Macdonald, J. & von Stetten, F. Correction: Review: a comprehensive summary of a decade development of the recombinase polymerase amplification. *Analyst* **145**, 1950–1960 (2020)

20. Powell, M. et al. New Fpg probe chemistry for direct detection of recombinase polymerase amplification on lateral flow strips. *Anal. Biochem.*, **543**, 108-115 (2018)

21. Rathore, H., Biyani, R., Kato, H., Takamura, Y. & Biyani, M. Palm-size and one-inch gel electrophoretic device for reliable and field-applicable analysis of recombinase polymerase amplification. *Anal. Methods* **11**, 4969–4976 (2019)

22. Huang, Z. *et al.* Ultra-sensitive and high-throughput CRISPR-p owered COVID-19 diagnosis. *Biosens. Bioelectron.* **164**, 112316 (2020)

23. Qian, J. *et al.* An enhanced isothermal amplification assay for viral detection. *Nat. Commun.* **11**, 5920 (2020)

24. Subsoontorn, P., Lohitnavy, M. & Kongkaew, C. The diagnostic accuracy of isothermal nucleic acid point-of-care tests for human coronaviruses: A systematic review and meta-analysis. *Sci. Rep.* **10**, 22349 (2020)

25. Chandler, D. P., Wagnon, C. A. & Bolton, H., Jr. Reverse transcriptase (RT) inhibition of PCR at low concentrations of template and its implications for quantitative RT-PCR. *Appl. Environ. Microbiol.* **64**, 669–677 (1998)
26. Jarvis, K.F., Kelley, J.B. Temporal dynamics of viral load and false negative rate influence the levels of testing necessary to combat COVID-19 spread. *Sci Rep* **11**, 9221 (2021)
27. Park, N. J., Li, Y., Yu, T., Brinkman, B. M. N. & Wong, D. T. Characterization of RNA in saliva. *Clin. Chem.* **52**, 988–994 (2006)
28. Vogels C. B. F. *et al.* SalivaDirect: A simplified and flexible platform to enhance SARS-CoV-2 testing capacity. *Med.* **2**, 263-280.e6 (2021)
29. Kumar S. et al. Long-term dry storage of enzyme-based reagents for isothermal nucleic acid amplification in a porous matrix for use in point-of-care diagnostic devices. *Analyst* **145**, 6875-6886 (2020)
30. Sobczak L., Gorynski K. Evaluation of swabs from 15 commercially available oral fluid sample collection devices for the analysis of commonly abused substances: doping agents and drugs of abuse. *Analyst* **145**, 7279-7288 (2020)
31. Huggett J.F. *et al.* Cautionary note on contamination of reagents used for molecular detection of SARS-CoV-2. *Clin. Chem.* **66**, 1369-1372 (2020)
32. Yasukawa, K., Mizuno, M., Konishi, A. & Inouye, K. Increase in thermal stability of Moloney murine leukaemia virus reverse transcriptase by site-directed mutagenesis. *J. Biotechnol.* **150**, 299–306 (2010)
33. Higgins, M., Ravenhall, M., Ward, D., Phelan, J., Ibrahim, A., Forrest, M. S., ... & Campino, S. (2019). PrimedRPA: primer design for recombinase polymerase amplification assays. Bioinformatics, 35(4), 682-684.

CHAPTER -3

Understanding biophysical properties of oligomerization and fibril formation in amyloid beta 42 conformers

(Interaction between peptide and peptide)

3.1 INTRODUCTION

3.1.1 The role and challenges of Amyloid βeta in Alzheimer's Disease

Alzheimer's Disease:

Alzheimer's disease (AD) is the most common type of brain disorder, causing more than 60-70% cases of dementia [1]. It affects many million people worldwide, and this number is escalating considerably. The current Alzheimer's statistics report that the projected number of people in the world population that will be suffering from AD will cross the 100 million mark by the year 2050 [2]. The current challenges in AD lie not only in therapy (no cure has been reported till date that can slow down the progression of the disease) but diagnosis as well [3]. The diagnosis of dementia is comparatively easy but to determine AD at an early stage is still a clinical challenge. Dementia due to AD is the last stage of the disease and the neuropathological changes in brain start accumulating about 10-15 years before the signs of dementia [4,5]. Early diagnosis of AD is thus particularly important for prognosis and treatment of the disease [6].

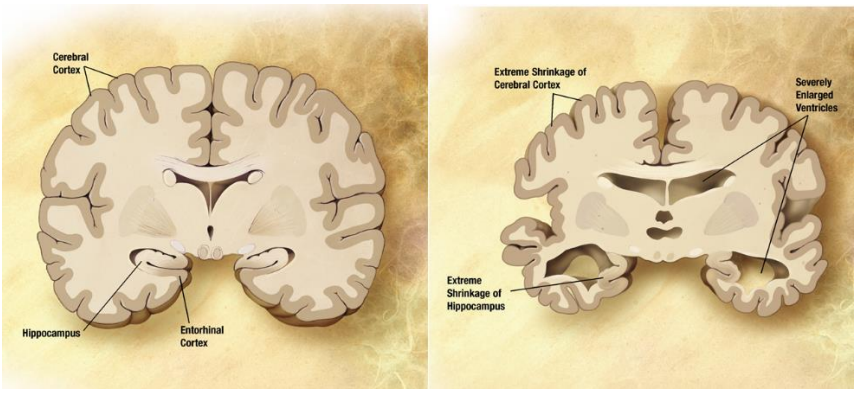


Fig 3.1: Healthy Brain (Left) vs Alzheimer's Brain (Right)
{Source: https://en.wikipedia.org/wiki/Alzheimer%27s_disease}

Amyloid Beta Peptide:

The amyloid hypothesis proposes that Amyloid Beta (Aβ) is the main cause of the disease and that the misfolding of this extracellular Aβ protein leads to its accumulation triggering cognitive decline [7,8]. Accumulated Aβ peptide is the main

component of the senile plaque that causes neurotoxicity and is the clear indication of AD [9]. Thus, Aβ is the main cause of neuronal death in the brains of patients with AD.

The amyloid fibrils of the senile plaques found in AD patients consists mainly of the 40-mer or 42-mer amyloid β proteins which are known as Aβ40 and Aβ42, respectively [10]. The only difference between Aβ42 and Aβ40 is the two additional C-terminal residues on Aβ42. Aβ40 and Aβ42 are produced from Amyloid Precursor Protein (APP) which is cleaved by beta secretase and gamma secretase to yield the Aβ [11]. Nevertheless, it is suspected that Aβ42 form of Aβ is more critical in the pathogenesis of AD [12]. Although Aβ42 makes up for about less than 10% of total Aβ in normal people, it is more abundant in AD brain tissue than in age-matched controls. Aβ42 aggregates at much faster rates than Aβ40 does and thus has higher neurotoxicity. Several reasons have been reported to be responsible for the toxicity caused by the aggregation of Aβ42. Few such reasons include oligomer hypothesis, size and stability of Aβ42 species [13] and 'toxic-turn' at position 22 and 23 [14]. Conformational diversity of Aβ42 species is also said to have a potent role in causing toxicity [15]. Hence, Aβ42 is the primary and the key component of senile plaque deposits [16].

3.1.2 Toxic Conformer of Aβ42

Many studies have reported a toxic conformer of Aβ42 with a turn at Glu22 and Asp23 position in the 42 amino acid chain which has been described and differentiated with the wild-type Aβ42 which has a turn at Gly25 and Ser26 position [17]. The mutation is called E22P, which means E (Glutamic Acid) at 22nd position changes to P (Proline) in toxic conformer of Aβ42. Conformation of the toxic conformer of Aβ42 is such that it induces close interaction between Tyr10 and Met35 for formation of the S-oxidized radical cation of Met35 and a C-terminal hydrophobic core for stabilization of radicals for long-lasting oxidative stress [18]. This toxic conformer exhibited a strong neurotoxicity and synaptotoxicity and thereby making it more prone to risk [19]. Fig 3

Mutant Type/ Toxic Conformer (E22**P**)

1'-DAEFRHDSGYEVHHQKLVFFA**P**DVGSNKGAIIGLMVGGVVIA-42'

Proline

Wild Type/ Less-toxic Conformer (E22)

1'-DAEFRHDSGYEVHHQKLVFFA**E**DVGSNKGAIIGLMVGGVVIA-42'

Glutamic Acid

Figure 3.2: Sequence information of variants E22-Aβ42 and E22P-Aβ42 showing mutation at 22nd position where amino acid glutamic acid changes to proline that leads to higher aggregation and toxicity.

In simple terms, the proportion of Aβ42 toxic conformer, in which residues Gln15~Ala21 and Val24~Ile32 are involved in forming an intermolecular parallel β-sheet of Aβ42 aggregates, may contribute to AD pathology with a much greater risk than the wild-type Aβ42 does. Moreover, the C-terminal hydrophobic core derived from another turn at Gly38 and Val39, in addition to the intramolecular anti-parallel β-sheet (Met35~Gly37 and Val40~Ala42), accelerates the aggregation (oligomerization) of Aβ42. The aggregation and toxicity of E22P-Aβ42 with a turn at positions 22 and 23 are said to be stronger than those of wild-type Aβ42, whereas only E22P-Aβ42 with a turn breaking Val residue was found in Aβ40 mutants. These results well supported the existence of turn formation at positions 22 and 23.

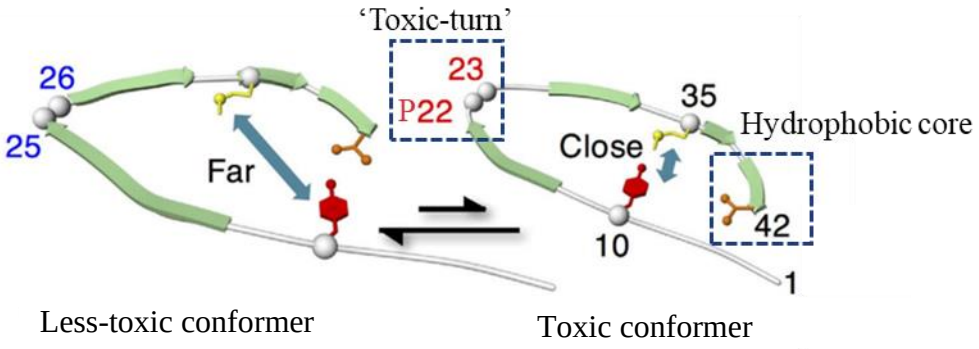


Figure 3.3: A schematic illustration of the structure of the toxic and non-toxic conformer of Aβ42. The turn at 22-23 position is the 'toxic-turn', which draws Y10 and M35 closer to accelerate the radical transition and contributes to the hydrophobic core at C-terminus. *{Source: Nature2017;7:1181 }*

3.1.3 Aβ42 assembly

Aβ monomers aggregate into various types of associations, including oligomers, protofibrils and amyloid fibrils. Aβ monomers first aggregate into different forms of oligomers, which can then form regular fibrils and then the fibrils further assemble to amyloid plaques. Among the various Aβ species, structured oligomers are proposed to be more toxic than any other form [20].

There are several statements that report about Aβ peptides being directly involved in the pathogenesis of AD and that their aggregative capacity is closely related to neurotoxicity [21]. However, the precise mechanism of their neurotoxic effects through aggregation is yet unclear.

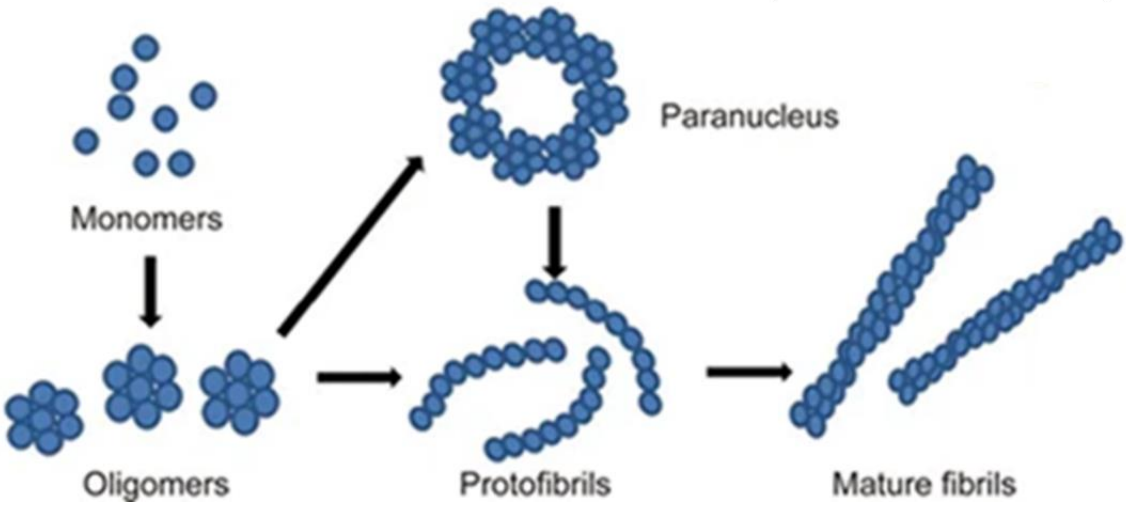


Figure 3.4: Aβ42 monomers can form higher order assemblies ranging from low molecular weight oligomers, including dimers, trimers, tetramers, and pentamers, to mid-range molecular weight oligomers including hexamers, nonamers and dodecamers to protofibrils and fibrils. *{Source: Nature, APS. 2017; 38.9}*

Aβ42 is cleaved from the amyloid precursor protein (APP) by β- and γ- secretases and aggregates into the intermolecular β-sheets from Aβ monomers into amyloid fibrils via oligomers/ protofibrils. The evidence suggests that soluble oligomeric forms of Aβ42 are more important for the pathogenesis of AD as compared to Aβ42 fibrils

because of their potent neurotoxicity [22-24]. Differential toxicity of oligomers is caused because of several reasons that include highly unstable disordered structures of oligomers as compared to stable organized molecules in case of fibrils, smaller size of oligomers that allow easy diffusion in tissues, exposed hydrophobic surfaces in beta sheets and more open active ends present in oligomers allow improved interaction with cellular targets [25]. Some evidence suggests that the toxic conformer of Aβ42 exists mainly in soluble oligomeric forms and not in senile plaques [26-27].

Key objectives of this chapter:

The toxic conformer of Aβ42 was reported 20 years ago, however only few reports were made available on the precise molecular structure and aggregation mechanism of Aβ42 before 2015 [28]. This was mainly because of the diversified conformation diversity and highly aggregative and unstable character of this small peptide molecule that it became difficult to perform NMR analysis or X-ray crystallography [29,30]. Although some groups have also reported the supramolecular structure of various Aβ42 species under controlled conditions [31], however it is still challenging to make a understanding of the conformational diversity of the various Aβ42 conformers.

The purpose of this study is to understand and differentiate the biophysical properties of the toxic conformer of Aβ42, E22P-Aβ42 with the non-toxic conformer of Aβ42, E22-Aβ42. It is interesting to understand the mechanism that leads to a 200-fold increased toxicity by a mutation in single amino acid.

OBJECTIVE: Understanding biophysical properties of oligomerization and fibril formation in amyloid beta 42 conformers

3.2 MATERIALS AND METHODS

3.2.1 Pretreatment of Aβ42 species

Synthetic E22-Aβ42 was purchased from Anaspec, Japan and E22P-Aβ42 was custom synthesized from Anygen, Korea. The pretreatment of Aβ42 conformers was done according to the previous literatures [32]. To prevent the rapid aggregation and to obtain a homogenous solution of monomeric Aβ42 in unstructured conformation, it was dissolved in 100% ice-cold HFIP (1,1,1,3,3,3-hexafluoro-2-propanol) at 1 mM concentration to remove any pre-existing aggregates, if any. After incubation at room temperature for 30 min, the solution was sonicated for 1 min and aliquoted, then left open to dry in hood overnight. To remove any further traces of HFIP, Aβ42 was vacuum dried under freezing conditions for 90 min. The resultant film of Aβ42 was stored at −80 °C until use.

For the oligomerization and fibril formation of E22-Aβ42 and E22P-Aβ42, the dried film was diluted in DMSO upto 200uM and slightly vortexed. The solution was then diluted 10 times with 10mM HCl upto final concentration of 20uM. After vortexing, the sample was sonicated for few seconds followed by incubation at 37°C for different time intervals in protein low binding tubes. Incubation in acidic environment (~ pH 2) accelerated the fibril formation as compared to incubation in neutral environment (~pH 7.4).

3.2.2 Tricine SDS-PAGE

Aβ42 film was dissolved in 0.1% NH_4OH, followed by 10-fold dilution upto 25μM with phosphate buffered saline (PBS) buffer (137 mM NaCl, 10 mM phosphate, 2.7 mM KCl; pH 7.4). The solutions were then incubated at 37°C for 0/1/6/24 hrs respectively for aggregation to take place. Tricine SDS Page was carried out using previously reported literatures [33]. A 16% running gel, 10% spacer gel, and a 4% stacking gel in gradient were prepared following a previously established protocol. Electrophoresis was carried out at 30 V for approximately 15 min until the sample completely entered the stacking gel. Then, the voltage was increased to 110 V, and the electrophoresis was run at this voltage until it was finished. Precision Plus protein

standard Protein marker was used in all the experiments. Silver staining was carried out to visualize the results.

3.2.3 Atomic Force Microscopy imaging of Aβ42 species

The FM-AFM (Frequency Modulation Atomic Force Microscopy) imaging was performed to characterize the morphologies and aggregation behavior of E22-Aβ42 and E22P-Aβ42. A modified mica was used for the fixation of Aβ42 oligomers or fibrils. For mica modification, 50µl Poly-l-lysine (PLL) solution was dropped on freshly cleaved mica. After incubation for 3 min, the mica was rinsed several times with MilliQ water. 100µl of diluted Aβ42 samples in water with a final concentration of 2 µM were then dropped on the mica surface and incubated for 10 min. The sample solution was then exchanged with imaging buffer of 100 µl PBS solution. The FM-AFM imaging was performed using a home-built AFM system with a low noise cantilever deflection sensor [34,35] and a high stability photothermal cantilever excitation system [36,37] at room temperature in a liquid environment. To control this AFM scanning and oscillation, a commercially available AFM controller (Nanonis OC4 and RC4, SPECS) was used. Commercially available cantilevers (160AC, Opus) with a nominal spring constant of 26 N/m and a nominal resonance frequency of 120 kHz in liquid were used. The obtained AFM images were rendered using Gwyddion software.

3.2.4 Nanomechanical mapping of Aβ42 species

Elasticity measurement and thereafter, the Young's Modulus values were obtained using the Bruker BioScope AFM system. PLL-modified mica was used for nanomechanical mapping of Aβ42 fibrils. Several images were obtained at different locations across the mica surface to ensure a high degree of reproducibility of the obtained fibril structures. Firstly, an overview topographic image was acquired to locate individual fibrils, followed by imaging at a relatively small scan area to perform the Young's modulus mapping. The imaging was performed in PeakForce tapping mode in liquid environments at room temperature. The topographic images and the Young's modulus maps were recorded over a scan area of 500 nm x 500 nm2 (256 x 256 pixels). A force setpoint of 1.2 nN was applied for all scans to prevent damage to the fibril. Commercially available cantilevers (240AC, OPUS) with a

nominal spring constant of 2 N/m and a nominal resonance frequency of 25 kHz in a liquid were used. AFM image rendering and data processing were performed using the NanoScope Analysis software (version 1.9, Bruker, Billerica, MA, USA). The Sneddon model was fitted on the approach curves to extract the Young's modulus. The fitting was limited within the range of 10-90% of maximum force to reduce the risk of substrate effect as much as possible.

3.2.5 Molecular Docking of Aβ42 species

An initial three-dimensional monomer structure of E22-Aβ42 was extracted from E22-Aβ42 fibril crystal (Accession number 5-OQV in the protein data bank (PDB)) [38]. The corresponding structure of E22P-Aβ42 monomer was constructed using the homology modelling software MODELLER [39] from its E22/E22P residue sequences and the three-dimensional structure of E22-Aβ42 monomer (as a related homologous protein).

Their dimer structures were predicted by the high ambiguity driven docking (HADDOCK) simulations on the HADDOCK webserver 2.4.[40,41].In order for the HADDOCK simulations to explore the stable conformation efficiently, the Ambiguous interaction restraint (AIR) was used to restrict the conformational space, where distances between an active site and an active/passive site are limited by adding the AIR energy term to the docking score; the active site is defined as a residue that forms binding interfaces and the passive site is defined as a solvent-accessible residue around the active sites [42]. These sites can be usually given by NMR titration experiments; however, if unavailable, a data-driving site prediction program CPORT46 can be also used to predict active/passive sites. In the present study, CPORT was applied to obtain active/passive sites of our E22P/E22-Aβ42 monomers. All 42 residues were predicted as active sites and then used for an input of the HADDOCK simulation.

The present HADDOCK simulation follows the four procedures: (i) randomization of orientations and rigid-body energy minimization (EM), (ii) simulated annealing in torsion angle space (TAD-SA), (iii) short-time molecular dynamics (MD) with waters as explicit solvent molecules, and (iv) clustering and ranking of MD-relaxed structures. (i) Initially, 1000 complex conformers treated as rigid bodies were

randomly generated and energetically optimized. Then 200 conformations with higher docking scores were selected, where the docking score is given from the interaction potential energy based on the weighted force field and desolvation entropy term derived buried surface. Furthermore, HADDCOK automatically removed conformers that never form the binding interface. (ii) TAD-SA was applied to the remaining conformers in order to obtain the conformers with optimized torsion angles. In order to consider structural flexibility in TAD-SD, the five residues (20,21,22,23,24) around the loop region was specified as fully flexible residue. Because the residues around loop in Aβ42 would be more flexible and it seems affect the predicted conformer complex. (iii) The resultant conformers were partially relaxed by the MD simulation. Note that the above mentioned five residues were treated as fully flexible. (iv) The final conformers were clustered into 2 and 5 clusters for E22/E22P dimers respectively using the fractional common contacts [43] at the binding interface. Each cluster is ranked by averaging the docking scores over the conformers within the cluster. To support out stacking hypothesis, we will refer to theoretical structures of E22P/E22-Aβ42 dimers selected from top 1 conformer in top 1 cluster.

3.3 RESULTS

3.3.1 Structure prediction of E22-Aβ42 and E22P-Aβ42 by molecular docking

Three-dimensional structure of both the variants, E22-Aβ42 and E22P-Aβ42 was predicted using molecular docking. **Fig.3.5** shows the sequence and monomer structure of toxic conformer, E22P-Aβ42 (a) and less toxic conformer, E22-Aβ42 (b) where mutation at 22nd position in E22-Aβ42 causes amino acid E (glutamic acid) to change into amino acid P (proline). B1 to B6 indicates the number of beta sheet structures. A total number of 4 beta sheets (B2, B3, B4, B6) are involved in E22-Aβ42 whereas 2 additional beta sheets (B1 and B5) are involved in E22P-Aβ42 which indicates more intra-chain interaction of beta sheets. The number of residues involved in the beta sheet formation are also varied in both the cases. Around 28 amino acids out of 42 (67%) are engaged in the formation of beta sheets in E22P-Aβ42 whereas around 16 amino acids out of 42 (38%) are engaged in E22-Aβ42. Length of beta sheets in E22P-Aβ42 includes 4 residues in B1 (Phe4 to Asp7), 6 residues in B2 (Tyr10 to Gln15), 5 residues in B3 (Leu17 to Ala21), 6 residues in B4 (Ser26 to Ile31), 4 residues in B5 (Leu34 to Gly37) and 3 residues in B6 (Val39 to Ile41), respectively. Length of beta sheets in E22-Aβ42 includes 3 residues in B2 (His13 to Gln15), 4 residues in B3 (Val18 to Ala21), 6 residues in B4 (Ser26 to Ile31), and 3 residues in B6 (Val39 to Ile41), respectively. Higher number of residues in beta sheets can cause more intra-molecular interactions between the amino acids of Aβ42. This intra-molecular interaction of beta sheets can cause the building blocks of both the conformers to be significantly different. It has been studied that the beta sheet content can influence the properties of amyloid aggregation [44] . The beta sheets B1 and B5, which are present only in E22P-Aβ42 can cause its structure to be more compact because of the inter-molecular interaction and thus appear more rod-like and less flexible. However, in E22-Aβ42, the structure can appear more flexible and less compact because of the absence of specific interaction in the residues towards the N'-terminus causing this end to be flexible.

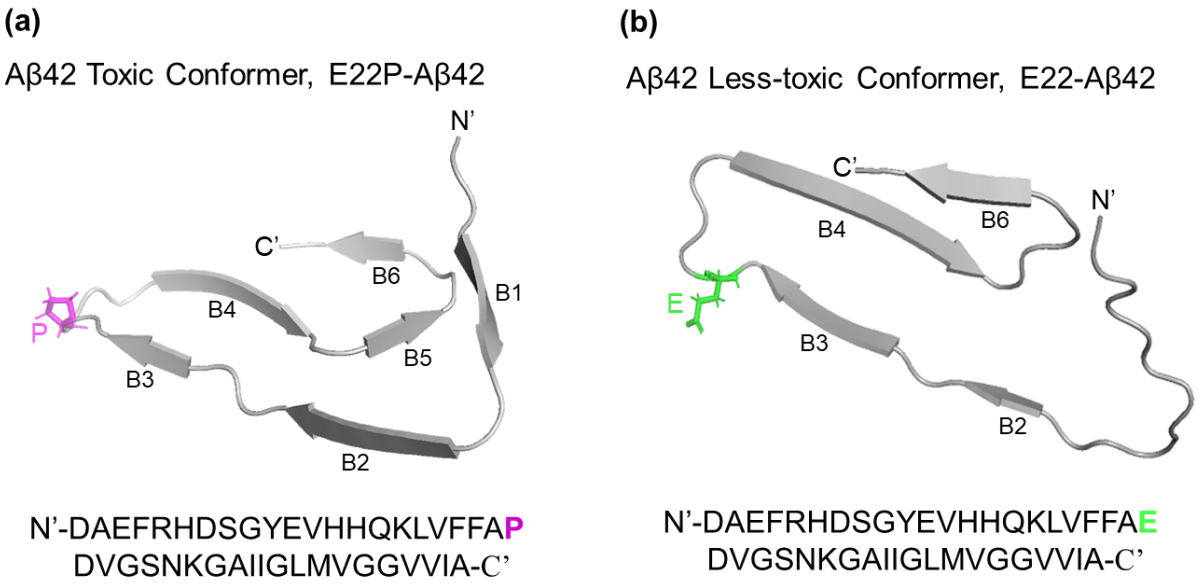


Figure 3.5: Monomer structure prediction using molecular docking and sequence information of variants (a) E22P-Aβ42 and (b) E22-Aβ42 showing mutation at 22nd position where amino acid proline (P) is replaced by glutamic acid (E) that causes higher toxicity in the former. B1-B6 indicates the number of beta sheets involved in the structure formation. 6 beta sheets are involved in E22P-Aβ42, and 4 beta sheets are involved in E22-Aβ42.

3.3.2 Macroscopic analysis of time dependent aggregation of E22-Aβ42 and E22P-Aβ42 by SDS-PAGE

Macroscopic analysis of time dependent aggregation state of E22-Aβ42 and E22P-Aβ42 was performed using silver-stained Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Tricine-SDS-PAGE). Freshly dissolved Aβ42 samples and Aβ42 samples after 4-, 8-, 24-, 72-hours incubation were examined. The 16%T resolving gel in SDS-PAGE revealed three discrete bands that exhibited a range of species from 5-20 kDa (**Fig. 3.6a**) and the 10%T spacer gel revealed 2 major bands that exhibited low molecular weight (LMW) aggregates and high molecular weight (HMW) aggregates species above 100 kDA **(Fig 3.6b)**. As previously reported [45-47], the intensity of all bands appeared to be higher with longer incubation time. Very interestingly, the lowest molecular weight band at 5 kDA in **Fig 3.6**, which is likely to represents monomer, shows the most prominent increasing intensity with increasing incubation time. This observation led to the conclusion that a possible dimer structure constituted the basic building blocks for Aβ42 aggregation which is

much faster in E22P-Aβ42. This SDS-PAGE result confirms that E22P-Aβ42 has a stronger and higher ability to aggregate as compared to E22-Aβ42.

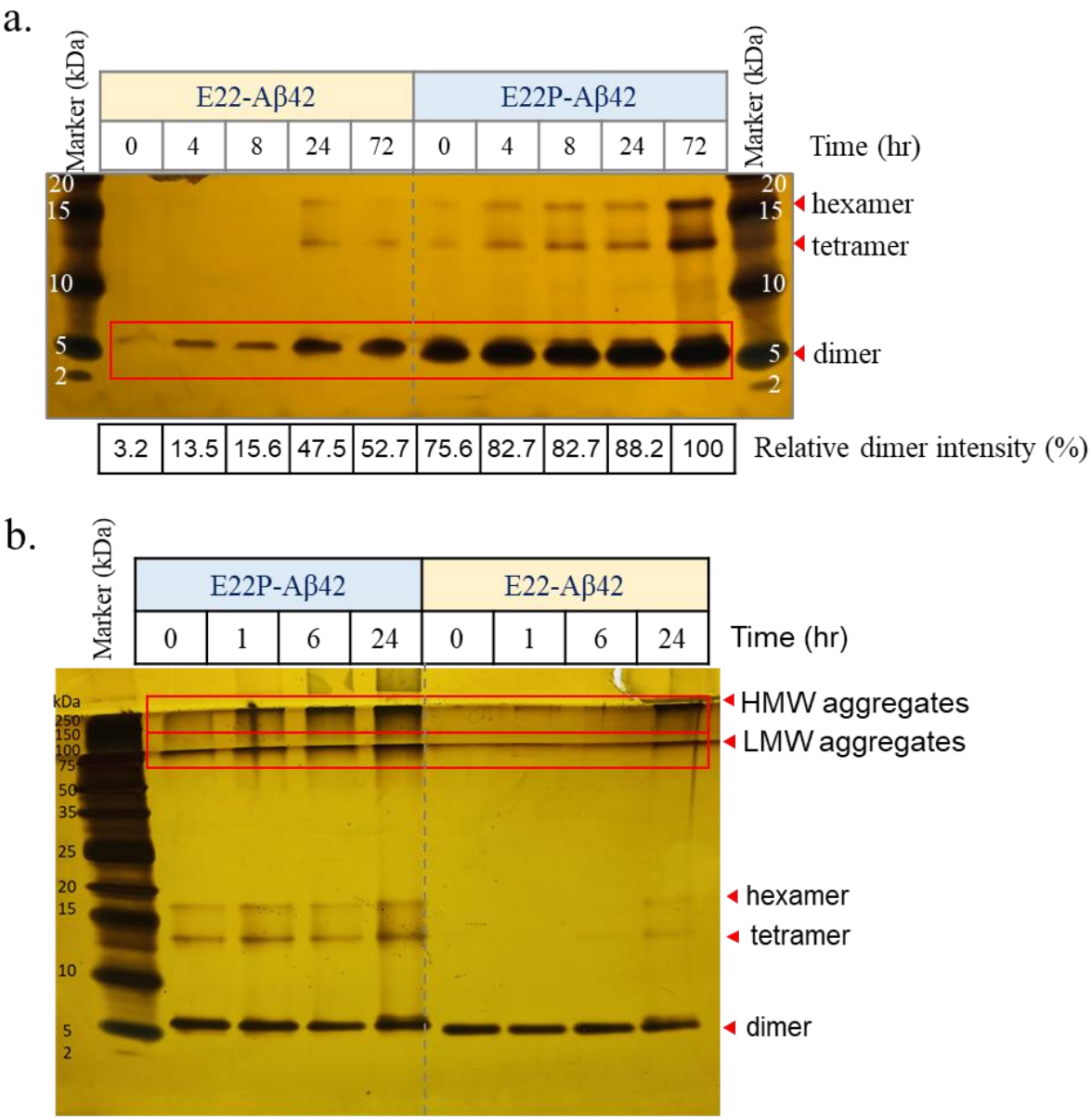


Figure 3.6: Time dependent aggregation of E22P-Aβ42 and E22-Aβ42 on Silver-stained Tricine SDS-PAGE analysis confirms rapid aggregation in E22P-Aβ42. (a) 16% T and (b) 10% T along with 16% T gel.

3.3.3 Microscopic analysis of time dependent aggregation of E22-Aβ42 and E22P-Aβ42 by FM-AFM

Given the potential artifacts of SDS in SDS-PAGE analysis, it is critical to further study this interesting observation and characterize the assembling nature of Aβ42 aggregation. For this reason, it is important to consider a method that are not dependent on SDS. Atomic force microscopy (AFM) offers the opportunity to draw a detailed microscopic analysis of the structural differences in the oligomerization process of both the conformers of Aβ42. Here we utilized frequency modulation atomic force microscopy (FM-AFM). E22-Aβ42 and E22P-Aβ42 were allowed to aggregate for different times at 37°C, and AFM images were obtained at 24- and 72-hours incubation respectively. **Fig.3.7a s**hows the AFM images of fibril formation in E22-Aβ42 and E22P-Aβ42. **Fig.3.7b** shows different oligomeric species along with fibrillar structures when AB42 was left on mica for 5 hours. Oligomeric structures can be seen at 24 hours in both cases, however in E22P-Aβ42 it appears that the oligomerization has already scaled up, and we can see comparatively bigger oligomeric structures. The higher distribution in E22P-Aβ42 indicates rapid oligomerization. The inset images show high-resolution imaging of the oligomeric structures (50 nm scale) where the interaction of 3 nucleating units is visible in E22P-Aβ42, while 3 discrete nucleating units are seen in E22- Aβ42. At 72 hours, long and mature fibrils can be seen in E22-Aβ42 whereas amorphous aggregation along with mature fibrils can be seen in E22P-Aβ42. **(Fig 3.7c)**. Secondary branched and amorphously distributed structures were observed in E22P-Aβ42 along with the mature fibrils as compared to E22-Aβ42 which has few, long and linear mature fibrils. The structural difference caused by a single amino acid mutation in E22-Aβ42 and E22P-Aβ42 causes E22P-Aβ42 to aggregate more rapidly and amorphously than E22-Aβ42 does.

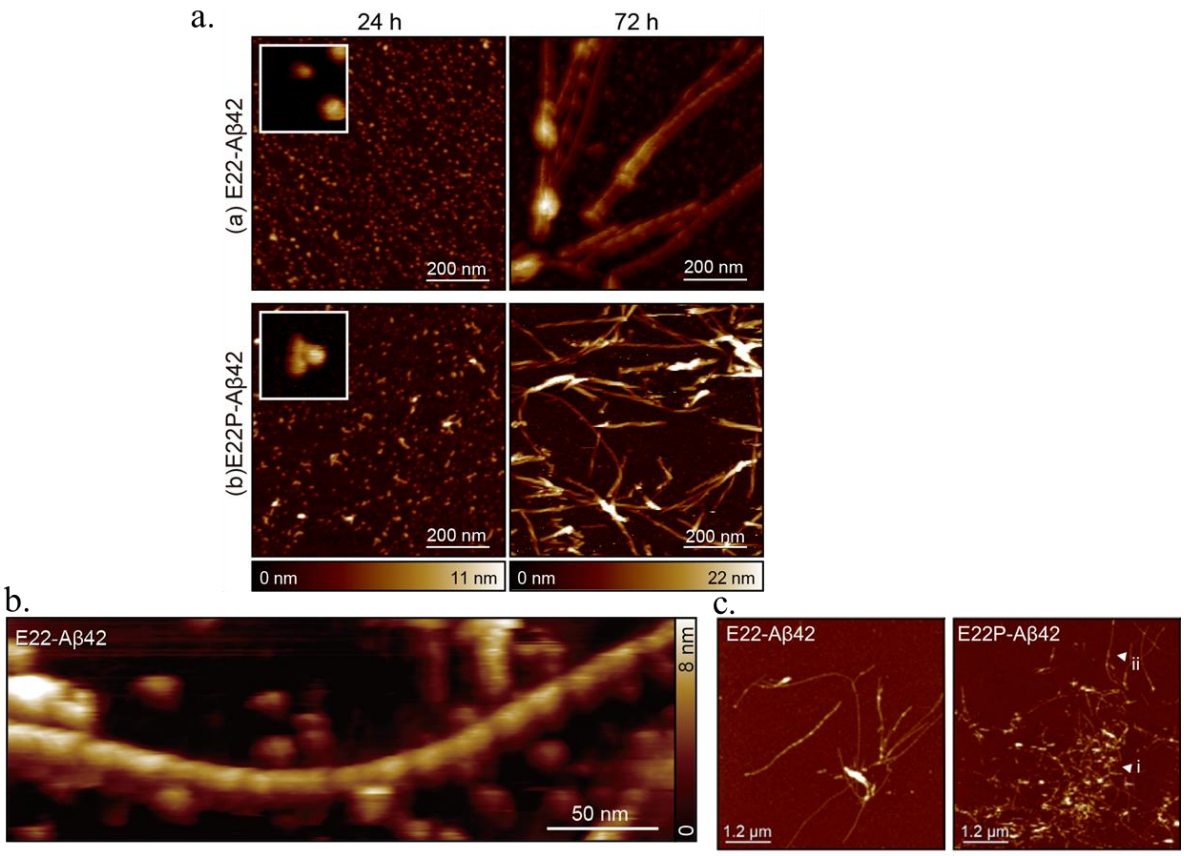


Figure 3.7: (a) FM-AFM images of oligomerization and fibril formation in E22-Aβ42 and E22P-Aβ42 after incubation at 37°C for 24 and 72 hrs respectively show the formation of mature fibrils from spherical oligomeric structures in E22-Aβ42 and formation of amorphous aggregation from fibril-like oligomeric structures in E22P-Aβ42. The inset images show high resolution imaging of E22-Aβ42 and E22P-Aβ42 oligomers (50nm scale). (b) E22-Aβ42 at 0 hours when incubated on the mica surface for 5 hours resulted in different oligomeric species in the imaging. During the oligomerization process, a mixture of different oligomeric species are present in the solution that includes small oligomers, large oligomers and fibril that confirms building blocks coming together for fibril formation. (c) 72 hour incubation of E22-Aβ42 and E22P-Aβ42 (6 µm scale) showing mature fibril formation in E22-Aβ42 and amorphous aggregation (i) along with mature fibrils (ii) in E22P-Aβ42

3.3.4 Characterization of biophysical properties of E22-Aβ42 and E22P-Aβ42 fibrils by high-resolution imaging

High resolution imaging of E22-Aβ42 and E22P-Aβ42 fibrils was performed to understand the nature of coiling in the fibrils. **Fig.3.7** indicates that the fibrils are more tightly coiled in the case of E22P-Aβ42 unlike the loose coiling in E22-Aβ42. This can be explained with the help of periodicity, which is the measure of the length between two turns that is comparable in E22-Aβ42 and E22P-Aβ42 for different turns as shown by a cyan colored dotted line (A-B in E22-Aβ42 and C-D in E22P-Aβ42).This indicates the varied nature of coiling in E22P-Aβ42 with loose coiling at some points that gradually become rigid coiling in the later stages of coiling. The height of fibrils at different points in E22-Aβ42 fibrils is varied with an average of 19.0±4.9 nm (points i, ii, iii) while in E22P-Aβ42 is relatively constant with an average of 7.9±0.96 nm (points iv, v vi). It can be inferred that the high height of fibril in E22-Aβ42 is because of the flexible monomeric structure of E22-Aβ42 caused by the absence of specific interaction between beta sheets in the residues towards the N'-terminus region as shown in **Fig.3.5**, and the compact-monomeric structure of E22P-Aβ42 caused by the involvement of 2 additional beta sheets is resulting in an overall decreased height of the fibril. It can also be inferred from **Fig.3.8** that the number of fibrils intertwined with each other can be different for E22-Aβ42 and E22P-Aβ42. The loose and flexible structure of E22-Aβ42 can allow more number of fibrils to coil with each other to form 1 fibril that can also result in increased overall height as compared to E22P-Aβ42.

Fig.3.9 shows different types of fibrils and the nature of coils formed in E22P- Aβ42. Type-1 fibril shows a fibril overlapping and coiling around the base fibril causing it to become more compact as indicated by the arrow. Type-2 and type-3 fibrils show the nature of coiling in the fibril. The type-2 fibril is coiled loosely over its base fibril leading to larger periodicity whereas in type-3, the fibril is coiled tightly over its base fibril leading to decreased periodicity and thus appearing as a rigid structure. Type-4 fibrils show how multiple mature fibrils can also intertwine with each other because of the self-assembly of the fibrillar structure. Different types of fibrils, including loosely coiled and tightly coiled can interact with each other to give an aggregated appearance.

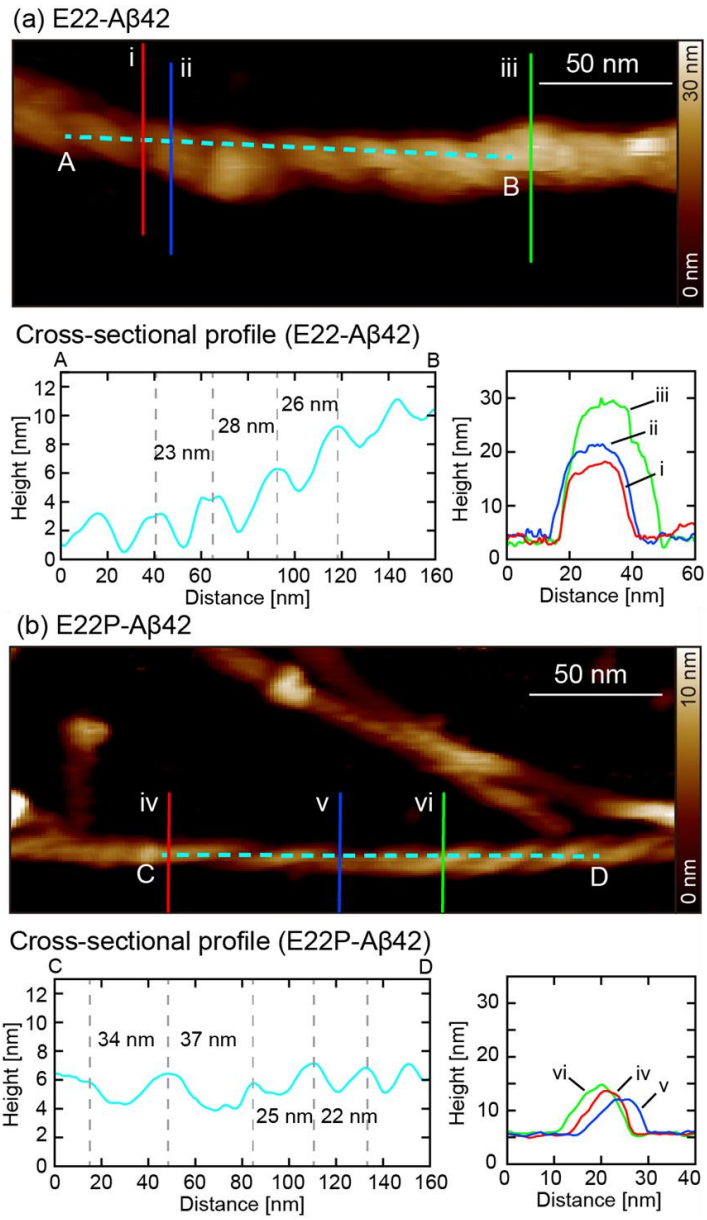


Figure 3.8: High resolution imaging of E22-Aβ42 and E22P-Aβ42 fibrils using FM-AFM. Unlike (a) E22-Aβ42 which appears as a non-uniform assembly, the small width of (b) E22P-Aβ42 appears to be more constrained structures with a uniform assembly. In cross-sectional profile of height vs distance, periodicity, and height for E22-Aβ42 and E22P-Aβ42 respectively are shown. The periodicity of E22-Aβ42 fibril is relatively constant as compared to the variable periodicity of E22P-Aβ42 (Points A-B and C-D) and the height of E22-Aβ42 is greater and variable as compared to the lesser and constant in E22P-Aβ42 (Points i-vi).

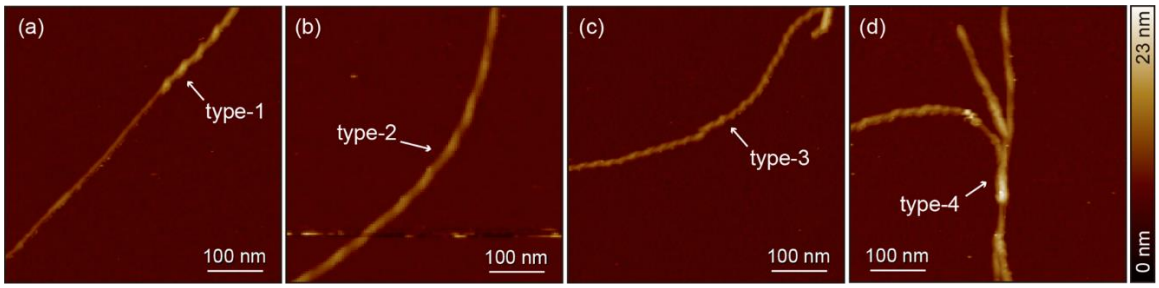


Figure 3.9: Different types of fibrils and nature of coils formed in E22P-Aβ42. Type-1 fibril showing on overlapping fibril that coils over the base fibril. Type-2 and type-3 showing different nature of coiling including loose and compact coiling. Type-4 fibril showing multiple mature fibrils coiled around each other interacting with different mature fibrils giving an aggregated appearance.

3.3.5 Young Modulas Mapping of E22-Aβ42 and E22P-Aβ42 fibrils for elasticity measurement

The mechanical properties of E22-Aβ42 and E22P-Aβ42 fibrils, i.e. their stiffness/elasticity, can play an essential role in the biological processes [48], and thus help us to understand and determine the level and effects of toxicity caused by E22P-Aβ42. To characterize the mechanical properties of the fibril based on their stiffness and elasticity, the Young's Modulus mapping was investigated through the use of a contact mechanics model provided by Sneddon et al. [49] for fibrils of E22-Aβ42 and E22P-Aβ42 at 48-hour incubation. A representative example of young's modulus map acquired for both E22-Aβ42 and E22P-Aβ42 is shown in **Figure 3.10**. This shows the topography (a) and the corresponding young's modulus (b) map of E22-Aβ42 (top) and E22P-Aβ42 (bottom) on PLL-modified mica substrate with setpoint 1.2nN and a 240 AC cantilever. The area occupied by the fibril appears darker and is characterized by a lower value of the young's modulus than the supporting substrate. **Fig 3.10c** shows the graphical representation of young's modulus vs distance for 4 different points taken on the fibril. Disturbances/variations in the signals of E22-Aβ42 confirm the non-uniform rough assembly, unlike uniform signals in E22P-Aβ42, which indicate smooth surface. It is because of the uneven assembly that at some points the young's modulus value is high and at some points is low, thus creating a varied signal.

An example of 20 representative curves (gray) obtained for each fibril is shown in **Fig. 3.10d**.The average of 20 curves is averaged (red) and young's modulus is then

extracted by fitting Sneddon Model curve (green) to each force curve. The histogram shown in **Fig. 3.10e** indicates the distribution of the young's modulus values compiled from the 18 fibrils in E22-Aβ42 and 24 fibrils in E22P-Aβ42. Gaussian fitting to the histogram of the calculated Young's modulus values revealed an average value of 0.59±0.44 GPa in E22-Aβ42 fibril and 0.73±0.45 GPa in E22P-Aβ42. This is the relative difference between both the fibrils that implies that E22P-Aβ42 has a greater young's modulus and indicates greater stiffness than E22-Aβ42. The topographic images of E22-Aβ42 and E22P-Aβ42 fibrils also indicate the stiffness of fibrils based on the number of kinks (**Figure 3.11**).

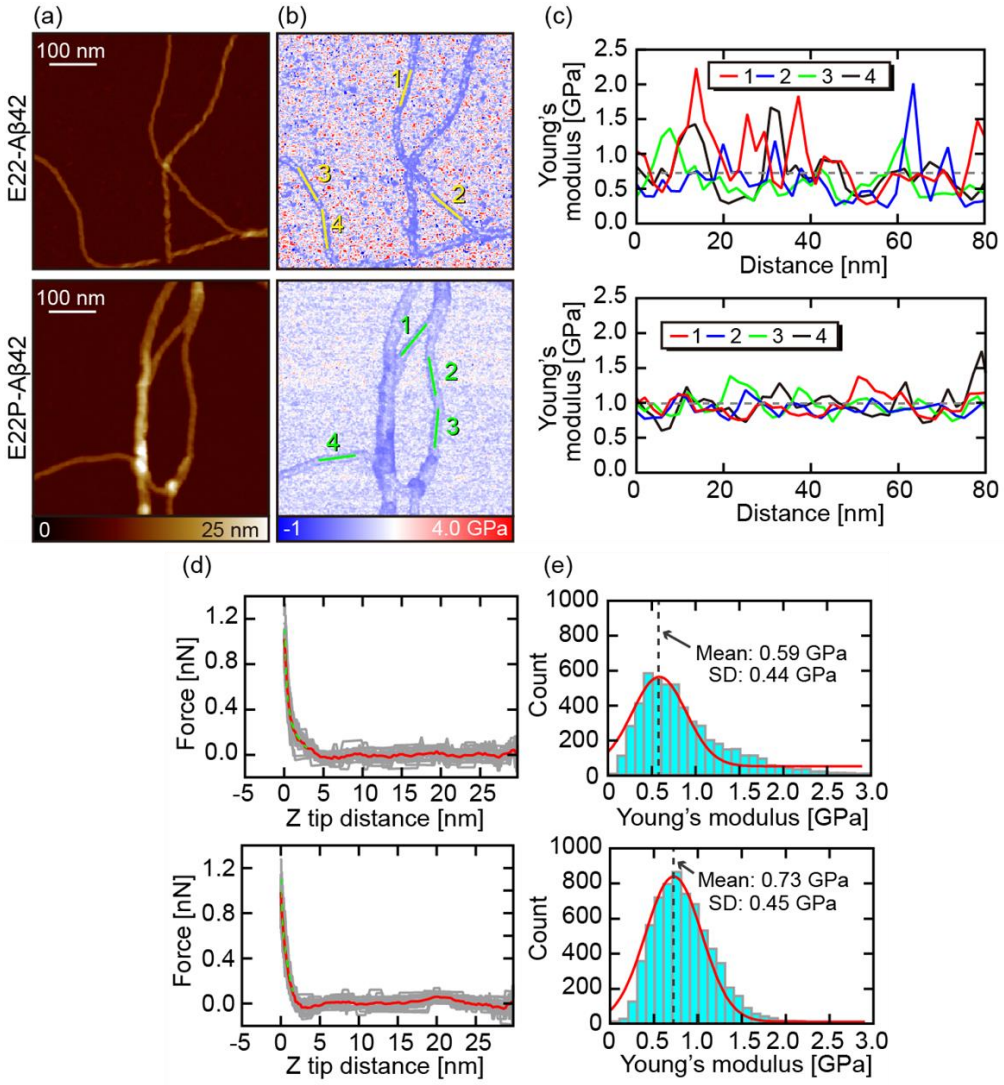


Figure 3.10: Mapping the mechanical properties of E22-Aβ42 (top) and E22P-Aβ42 (bottom) fibrils were recorded with setpoint 1.2 nN and a 240AC cantilever. Representative topography (a) and corresponding young's modulus map (b) of E22-Aβ42 and E22P-Aβ42 fibrils. These images include 500 x 500 nm^2 (256 × 256 pixels). The area occupied by the Aβ42 appears blue and is characterized by a lower value of young's modulus. (c) Young's Modulus vs distance graphs from the position marked with 1,2,3 and 4 in the elasticity map. (d) A typical force–distance (F(z)) curve (gray for 20 force curves randomly obtained from each fibril and red for average) with the fitted Sneddon model curve (green) obtained. (e) Histogram showing the distribution of young's modulus values measured for E22-Aβ42 and E22P-Aβ42 fibrils. The solid red line represents the fitted Gaussian function. The mean young's modulus value of the histogram was 0.53±0.44 GPa for E22-Aβ42 and 0.73±0.45 GPa for E22P-Aβ42 fibril.

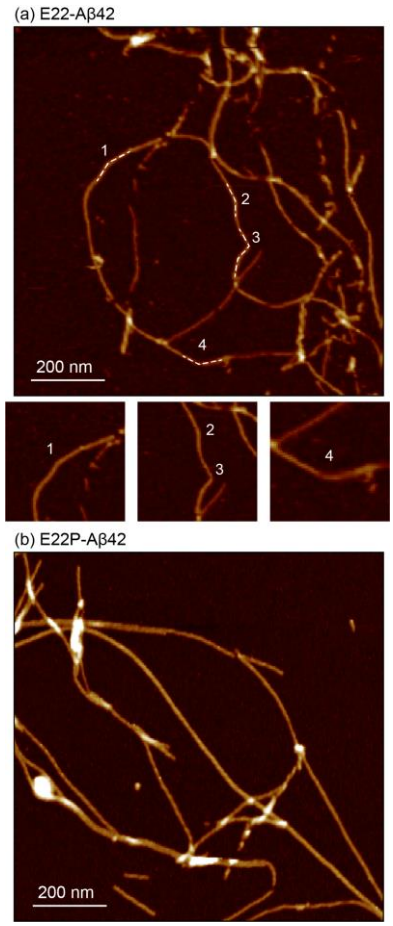


Figure 3.11: The topographic images of E22-Aβ42 and E22P-Aβ42 fibrils also indicating the stiffness of fibrils based on the kinks. Dotted lines in E22-Aβ42 shows kinks (bends) that can be caused due to flexible nature of the fibril. Stiff and rigid fibrils tend to break and not form kinks as shown in E22P-Aβ42.

3.3.6 Dihedral Angle Calculation between beta sheets using molecular docking

E22P-Aβ42 monomer was predicted by homology modelling from the residues of E22/E22P residue sequences. Thereafter, dimer structures were predicted using the monomer structures. Out of the several clusters that were prepared for the dimer structure in both relaxed and non-relaxed way, the top-structure was selected on the basis of its binding energies in the relaxed form. Score of selected structure for E22-Aβ42 is -276.4±2.4 (cluster size- 160; i-RMSD/Å- 2.2±2.0) and E22P-Aβ42 is -272.5±1.3 (cluster size- 113; i-RMSD/Å- 2.8±1.6). Studying the difference at nucleus formation level in E22-Aβ42 and E22P-Aβ42 can be crucial to understand the mechanism and cause of toxicity caused by E22P-Aβ42. For this, the difference in dihedral angle between two monomers of Aβ42 was calculated from 4 coordinates in both beta sheets B2 and B3 of E22-Aβ42 and E22P-Aβ42: N-terminus N atom of β-sheet of chain1, C-terminus C atom of β-sheet of chain1, N-terminus N atom of β-sheet of chain2 and C-terminus C atom of β-sheet of chain2.The N atom of N-terminus and the C atom of C-terminus of the β-sheet region were then connected as one vector and the torsion angle between them is calculated using the Praxeolitic formula. **Figure 3.12a** shows the calculated dihedral angle in E22-Aβ42 and **Figure 3.12b** shows the calculated dihedral angle in E22P-Aβ42. The beta sheets B2 and B3 of the dimers of E22-Aβ42 (E22-B2 and E22-B3) are aligned with an angle of 18° and 10°, respectively.

The beta sheets B2 and B3 of the dimers of E22P-Aβ42 (E22P-B2 and E22P-B3) are aligned with an angle of 37° and 43°, respectively. Because of the greater torsion angle in E22P-Aβ42, it can form a more strained and rigid structure unlike E22-Aβ42, which seems to form a more relaxed structure because of the relative smaller rotational angle. This can also be explained in the terms of rate of aggregation, which will be higher in E22P-Aβ42 because of the inter-molecular interaction of beta sheets in E22P-Aβ42.

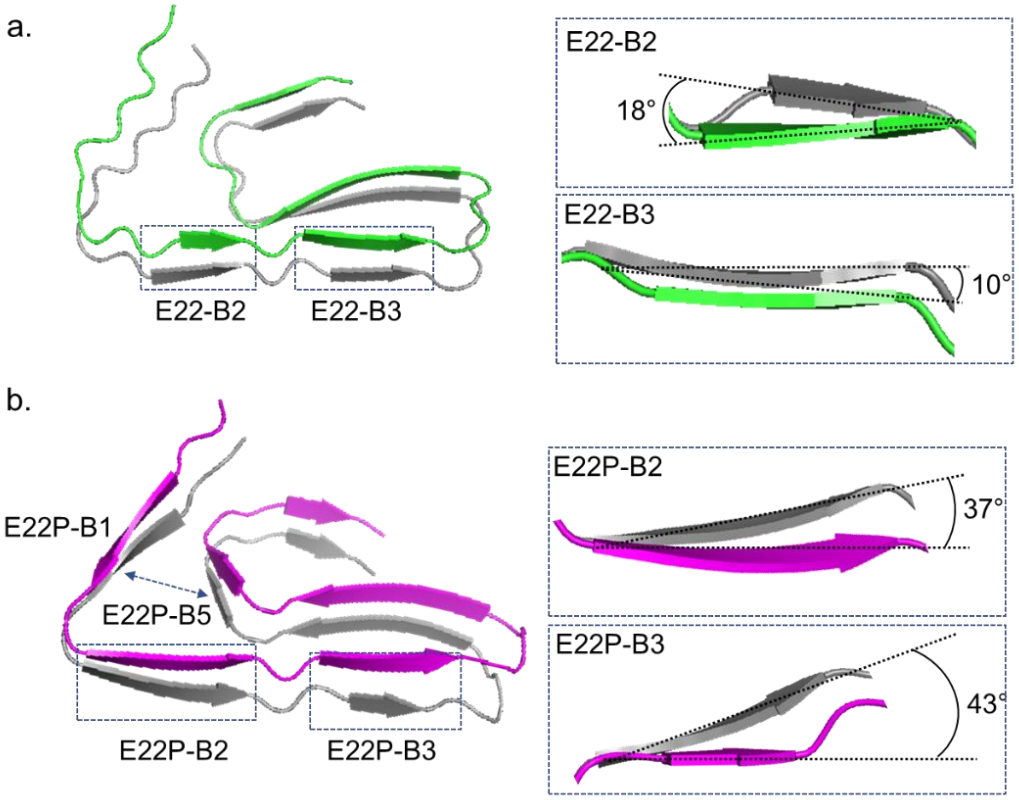


Figure 3.12: Dihedral Angle measurement between the dimers of beta sheets B2 and B3 of (a) E22 -Aβ42 and (b) E22P-Aβ42. Inset shows enlarged image of dihedral angle as 18° in E22-B2, 10° in E22-B3, 37° in E22P-B2 and 43° in E22P-B3.

3.4 CONCLUSIONS

Interaction between peptide and peptide has been studied in this chapter. In order to understand the role played by Aβ42 aggregation in the pathogenesis of Alzheimer's Disease and so thus be able to develop effective strategies for its treatment or diagnosis, it is necessary to have a clear understanding of aggregation at the molecular level. It is also important to understand how one amino acid substitution in E22P-Aβ42 can cause it to follow a different aggregation pathway, thereby making it more toxic than its wild type variant, E22-Aβ42. Previous studies have shown that amyloid beta peptide can self-assemble from monomeric structures to oligomeric followed by fibrillar structures, however different aggregation pathways for E22-Aβ42 and E22P-Aβ42 does not depend only on the amino acid sequence. Although the precise mechanism of aggregation *in-vivo* is still ambiguous to some extent, in this study we have captured the aggregation pathway of E22-Aβ42 and E22P-Aβ42 and propose a systematic structural model for the process of fibril formation in E22-Aβ42 and E22P-Aβ42 with the help of integrated *in-vitro* and *in-silico* analysis in this study.

As shown in **Figure 3.13**, our results indicate that the formation of additional intramolecular antiparallel β-sheet structure in monomer unit of E22P-Aβ42 (B1 and B5, as shown by red color in **Fig.3.13b**) than E22-Aβ42 caused a fluctuation in dihedral angle restraints (twist angle) between β sheets layers of two molecules of E22P-Aβ42 during dimerization, which offers as an important characteristic to determine the molecular basis of amorphous aggregation in E22P-Aβ42. As shown in **Fig.3.12b**, E22P-Aβ42 exhibits the high dihedral angle (~43° in B3 and ~37° in B2 β sheets layers), whereas the dihedral angle in E22-Aβ42 is comparatively lower (~10° in B3 and ~18° in B2 β sheets layers). This observation clearly suggests that the interacting two monomer units in E22P-Aβ42 can possess the higher twisting structures and favors spiral rod-like oligomerization, whereas the E22-Aβ42 favors sphere-like oligomerization. As shown in **Fig.3.9**, the AFM images reveals that the fibrils in E22P-Aβ42 have a multistranded helical shape with twisted ribbon-like structures, which can lead to the formation of abundant branching and so thus the formation of amorphous aggregation assembly unlike linear mature fibril assembly in E22-Aβ42. In case of E22-Aβ42, the short dihedral angle and absence of additional beta sheets allow more uniform and ordered interaction between monomer units and linear oligomerization can take place resulting into linear mature fibrils. As

shown in **Fig.3.13**, we describe a molecular model for E22P-Aβ42 and E22-Aβ42 fibrils that is based on molecular docking and AFM experimental data. It is observed from the AFM results that E22-Aβ42 forms sphere-like oligomers while E22P-Aβ42 forms spiral rod-like oligomers. This forms the basis of lego-type assembly in E22-Aβ42 and spiral-type assembly in E22P-Aβ42. The oligomerization pathway follows a sigmoidal growth curve where the rate of oligomerization is initially slow during the nucleation phase and then speeds up after the nuclear formation takes place during the elongation phase. In E22-Aβ42, the process of mature fibril formation from the nuclear formation stage is slower, while it is very rapid for E22P-Aβ42 to oligomerize from nucleus to amorphous aggregation. These findings provide an explanation of the characteristic properties of the amorphous aggregation nature of E22P-Aβ42.

Our study to understand the increased toxicity of E22P-Aβ42 can also be supported with the reference of oligomer-driven impairment of ubiquitin-dependent proteasome function [50]. First, additional beta sheets can cause increased and stronger interaction in E22P-Aβ42, thus making it resistant to degradation by proteasome. Also, higher stiffness (as observed by young's modulus calculation in **Fig.3.10**) can make the E22P-Aβ42 resistant to degradation by proteasome due to obvious rigidity reasons. Resistance to degradation can lead to accumulation of the peptide causing increased endoplasmic reticulum stress which makes E22P-Aβ42 more toxic as compared to E22-Aβ42. We anticipate that the finding reported in this study can guide to design the future experiments for our deeper understanding on the structural basis of toxicity caused by conformational diversity in amyloid-β42.

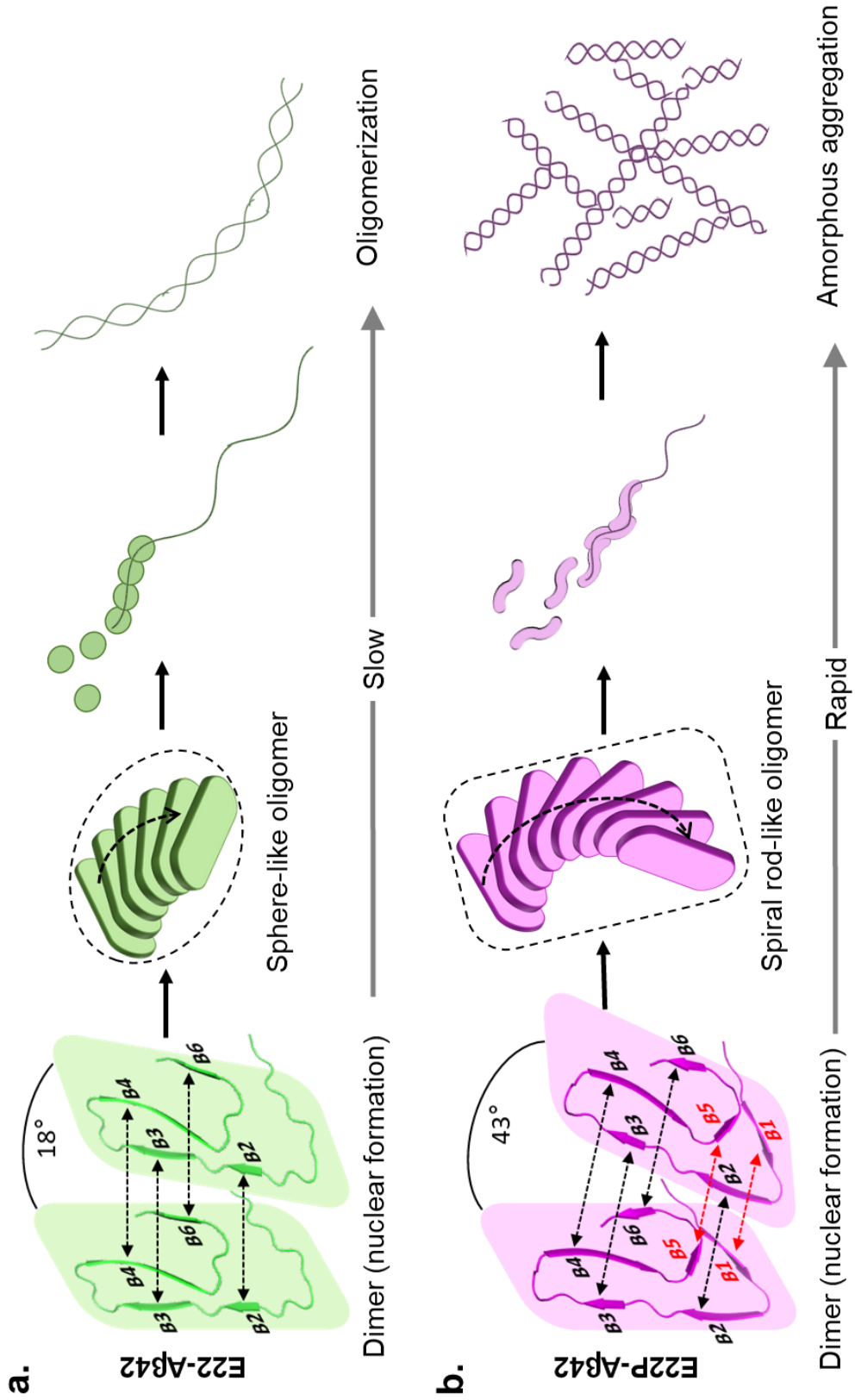


Figure 3.13: A systematic structural modeling of aggregation pathway in E22-Aβ42 and E22P-Aβ42.

In summary, the measurement and understanding of biophysical characteristics of different conformers of Aβ42 could thus give insight into the conformational dynamics of less toxic conformer E22-Aβ42 and toxic conformer E22P-Aβ42.The exact molecular mechanisms underlying the Aβ aggregation and its role in the process of neurodegeneration are still unknown. The progress in this field is challenging because of the varied conformation diversity and highly aggregative and unstable characteristics of Aβ42 peptide. This study presents new understandings about the molecular basis of fibril formation in E22-Aβ42 and E22P-Aβ42 that leads to increased toxicity in E22P-Aβ42 **(Summarized in Table 3.1)**. Molecular docking revealed that number of beta sheets involved in the formation of E22P-Aβ42 is 6 while in E22-Aβ42 is 4. Also, the residues involved in beta sheet formation of E22P-Aβ42 are 29% more than E22-Aβ42. The high-resolution imaging using FM-AFM revealed the amorphous nature of aggregation in E22P-Aβ42 while formation of mature fibrils take place in E22-Aβ42. Also, the macroscopic and microscopic analysis confirmed that rate of aggregation is higher in E22P-Aβ42. A relative comparison is made between the biophysical properties of E22-Aβ42 and E22P-Aβ42 that revealed stretched and branched fibrillar structure with varied periodicity in E22P-Aβ42 and relaxed and linear fibrillar structure with constant periodicity in E22-Aβ42. The Young's Modulus of E22P-Aβ42 fibrils was found to be 0.73 GPa and E22-Aβ42 was 0.59 GPa that indicates high stiffness and rigidity of E22P-Aβ42. We finally propose a molecular model of mechanism of aggregation in E22-Aβ42 and E22P-Aβ42 that causes increased toxicity in E22P-Aβ42 based on the understandings derived from this study. These findings will have significant implications to our understanding of the structural basis of toxicity caused by conformational diversity in Aβ42 species.

Table 3.1: Relative comparative analysis of biophysical properties of E22-Aβ42 and E22P-Aβ42

	E22P-Aβ42	**E22-Aβ42**
Number of beta sheets involved	6	4
Residues involved in beta sheets formation	28/42 aa (67%)	16/42 aa (38%)
- First beta sheet (B1)	4 (F04-D07)	-
- Second beta sheet (B2)	6 (Y10-Q15)	3 (H13-Q15)
- Third beta sheet (B3)	5 (L17-A21)	4 (V18-A21)
- Fourth beta sheet (B4)	6 (S26-I31)	6 (S26-I31)
- Fifth beta sheet (B5)	4 (L34-G37)	-
- Sixth beta sheet (B6)	3 (V39-I41)	3 (V39-I41)
Dihedral Angle		
- B2	37°	18°
- B3	43°	10°
Oligomeric Structure	Rod-like	Spherical
Nature of aggregation	Amorphous	Mature Fibrils
Aggregation Rate	Faster	Slower
Periodicity	Varied	Less-varied
Rigidity & Stiffness (GPa)	0.73 ± 0.45	0.59 ± 0.44
Coiling	Stretched	Relaxed
Fibrillar Structure	Branched	Unbranched
Kinks	Few	Many

REFERENCES

1. Holtzman, D. M., Morris, J. C., & Goate, A. M. (2011). Alzheimer's disease: the challenge of the second century. *Science translational medicine*, *3*(77), 77sr1-77sr1.
2. Alzheimer's Association. (2010). 2010 Alzheimer's disease facts and figures. *Alzheimer's & dementia*, *6*(2), 158-194.
3. Redulla, R. (2020). Reminiscence Therapy for Dementia. *Issues in Mental Health Nursing*, *41*(3), 265-266.
4. Beason-Held, L. L., Goh, J. O., An, Y., Kraut, M. A., O'Brien, R. J., Ferrucci, L., & Resnick, S. M. (2013). Changes in brain function occur years before the onset of cognitive impairment. *Journal of Neuroscience*, *33*(46), 18008-18014.
5. Murphy, M. P., & LeVine III, H. (2010). Alzheimer's disease and the amyloid-β peptide. *Journal of Alzheimer's disease*, *19*(1), 311-323.
6. Chan, H. N., Xu, D., Ho, S. L., Wong, M. S., & Li, H. W. (2017). Ultra-sensitive detection of protein biomarkers for diagnosis of Alzheimer's disease. *Chemical science*, *8*(5), 4012-4018.
7. Musiek, E. S., & Holtzman, D. M. (2015). Three dimensions of the amyloid hypothesis: time, space and'wingmen'. *Nature neuroscience*, *18*(6), 800-806.
8. Chen, G. F., Xu, T. H., Yan, Y., Zhou, Y. R., Jiang, Y., Melcher, K., & Xu, H. E. (2017). Amyloid beta: structure, biology and structure-based therapeutic development. *Acta Pharmacologica Sinica*, *38*(9), 1205-1235.
9. Finder, V. H., & Glockshuber, R. (2007). Amyloid-β aggregation. *Neurodegenerative Diseases*, *4*(1), 13-27.
10. Gu, L., & Guo, Z. (2013). Alzheimer's Aβ42 and Aβ40 peptides form interlaced amyloid fibrils. *Journal of neurochemistry*, *126*(3), 305-311.
11. Seigel, G., Gerber, H., Koch, P., Bruestle, O., Fraering, P. C., & Rajendran, L. (2017). The Alzheimer's disease γ-secretase generates higher 42: 40 ratios for β-amyloid than for p3 peptides. *Cell reports*, *19*(10), 1967-1976.
12. Findeis, M. A. (2007). The role of amyloid β peptide 42 in Alzheimer's disease. *Pharmacology & therapeutics*, *116*(2), 266-286.
13. Verma, M., Vats, A., & Taneja, V. (2015). Toxic species in amyloid disorders: Oligomers or mature fibrils. Annals of Indian Academy of Neurology, 18(2), 138.

14. Murakami, K., Masuda, Y., Shirasawa, T., Shimizu, T., & Irie, K. (2010). The turn formation at positions 22 and 23 in the 42-mer amyloid β peptide: The emerging role in the pathogenesis of Alzheimer's disease. Geriatrics & gerontology international, 10, S169-S179.
15. Nguyen, P. T., Zottig, X., Sebastiao, M., Arnold, A. A., Marcotte, I., & Bourgault, S. (2021). Identification of transmissible proteotoxic oligomer-like fibrils that expand conformational diversity of amyloid assemblies. Communications Biology, 4(1), 1-14.
16. Citron, M. (2004). Strategies for disease modification in Alzheimer's disease. Nature Reviews Neuroscience, 5(9), 677-685.
17. Murakami, K., Masuda, Y., Shirasawa, T., Shimizu, T., & Irie, K. (2010). The turn formation at positions 22 and 23 in the 42-mer amyloid β peptide: The emerging role in the pathogenesis of Alzheimer's disease. *Geriatrics & gerontology international*, *10*, S169-S179.
18. Murakami, K., Tokuda, M., Suzuki, T., Irie, Y., Hanaki, M., Izuo, N., & Tokuda, T. (2016). Monoclonal antibody with conformational specificity for a toxic conformer of amyloid β42 and its application toward the Alzheimer's disease diagnosis. Scientific reports, 6(1), 1-12.
19. Izuo, N., Kasahara, C., Murakami, K., Kume, T., Maeda, M., Irie, K., ... & Shimizu, T. (2017). A Toxic Conformer of Aβ42 with a Turn at 22–23 is a Novel Therapeutic Target for Alzheimer's Disease. *Scientific Reports*, *7*(1), 1-13.
20. Heo, C., Chang, K. A., Choi, H. S., Kim, H. S., Kim, S., Liew, H., ... & Suh, Y. H. (2007). Effects of the monomeric, oligomeric, and fibrillar Aβ42 peptides on the proliferation and differentiation of adult neural stem cells from subventricular zone. Journal of neurochemistry, 102(2), 493-500.
21. Parihar , M. S., & Hemnani, T. (2004). Alzheimer's disease pathogenesis and therapeutic interventions. *Journal of Clinical Neuroscience*, *11*(5), 456-467.Murakami, K., Tokuda, M., Suzuki, T., Irie, Y., Hanaki, M., Izuo, N., ... & Tokuda, T. (2016).
22. Ono K. Alzheimer's disease as oligomeropathy. Neurochem Int. 2018;119:57–70
23. Walsh DM, Klyubin I, Fadeeva JV, et al. Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation in vivo. Nature. 2002;416:535–539.

24. Benilova I, Karran E, De Strooper B. The toxic Aβ oligomer and Alzheimer's disease: an emperor in need of clothes. Nat Neurosci. 2012;15:349–357
25. Verma, M., Vats, A., & Taneja, V. (2015). Toxic species in amyloid disorders: Oligomers or mature fibrils. Annals of Indian Academy of Neurology, 18(2), 138.
26. Izuo, N., Kasahara, C., Murakami, K., Kume, T., Maeda, M., Irie, K., ... & Shimizu, T. (2017). A toxic conformer of Aβ42 with a turn at 22–23 is a novel therapeutic target for Alzheimer's disease. Scientific reports, 7(1), 1-13.
27. Tomiyama, T., Nagata, T., Shimada, H., Teraoka, R., Fukushima, A., Kanemitsu, H., ... & Mori, H. (2008). A new amyloid β variant favoring oligomerization in Alzheimer's-type dementia. Annals of neurology, 63(3), 377-387.
28. Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F., & Tycko, R. (2002). A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. Proceedings of the National Academy of Sciences, 99(26), 16742-16747.
29. Colvin, M. T., Silvers, R., Ni, Q. Z., Can, T. V., Sergeyev, I., Rosay, M., ... & Griffin, R. G. (2016). Atomic resolution structure of monomorphic Aβ42 amyloid fibrils. Journal of the American Chemical Society, 138(30), 9663-9674.
30. Xiao, Y., Ma, B., McElheny, D., Parthasarathy, S., Long, F., Hoshi, M., ... & Ishii, Y. (2015). Aβ (1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. Nature structural & molecular biology, 22(6), 499-505.
31. Wälti, M. A., Ravotti, F., Arai, H., Glabe, C. G., Wall, J. S., Böckmann, A., ... & Riek, R. (2016). Atomic-resolution structure of a disease-relevant Aβ (1–42) amyloid fibril. Proceedings of the National Academy of Sciences, 113(34), E4976-E4984.
32. Stine, W. B., Jungbauer, L., Yu, C., & LaDu, M. J. (2010). Preparing synthetic Aβ in different aggregation states. In Alzheimer's Disease and Frontotemporal Dementia (pp. 13-32). Humana Press, Totowa, NJ.
33. Schägger, H. (2006). Tricine–sds-page. Nature protocols, 1(1), 16-22.
34. Fukuma, T., Kimura, M., Kobayashi, K., Matsushige, K., & Yamada, H. (2005). Development of low noise cantilever deflection sensor for multienvironment frequency-modulation atomic force microscopy. Review of Scientific Instruments, 76(5), 053704.

35. Fukuma, T., & Jarvis, S. P. (2006). Development of liquid-environment frequency modulation atomic force microscope with low noise deflection sensor for cantilevers of various dimensions. Review of Scientific Instruments, 77(4), 043701.
36. Fukuma, T., Onishi, K., Kobayashi, N., Matsuki, A., & Asakawa, H. (2012). Atomic-resolution imaging in liquid by frequency modulation atomic force microscopy using small cantilevers with megahertz-order resonance frequencies. Nanotechnology, 23(13), 135706.
37. Fukuma, T. (2009). Wideband low-noise optical beam deflection sensor with photothermal excitation for liquid-environment atomic force microscopy. Review of Scientific Instruments, 80(2), 023707.
38. Gremer, G., Schölzel, D., Schenk, C., Reinartz, E., Labahn, J., Ravelli, R., Tusche, M., IGLESIAS, C., Hoyer, W., & Schröder, G. (2017). Fibril structure of amyloid-β (1-42) by cyro-electron microscopy. Science 358(6359), 116-119.
39. 42. Sali, A., & Brundell, T. (1993). Comparative protein modeling by satisfaction of spatial restraints. Journal of Molecular Biology, 234(3), 779-815.
40. Zundert, G., Rodrigues, J., Trellet, M., Schmitz, C., Kastritis, P., Karaca, E., Melquiond, A., Dijik, M., Vries, S., & Bonvin, A. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes, 428(4), 720-725.
41. Honorato, R., Koukos, P., García, B., Tsaregotodtsev, A., Verlato, M., Giachetti, A., Rosato, A., & Bonvin, A. (2021). Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem. Frontiers in Molecular Biosciences, 8, 1-7.
42. Dominguez, C., Boelens, R., & Bonvin, A. (2012). HADDOCK: A Protein-Protein Docking Based on Biochemical or Biophysical Information. Journal of American Chemical Society, 125(7), 1731-1737.
43. Rodrigues, J., Trellet, M., Christophe, S., Kastritis, P., Karaca, E., Melquiond, A., & Bonvin, A. (2012). CPORT: A Consensus Interface Predictor and Its Preformance in Prediction-Driven Docking with HADDOCK. PROS ONE, 6(3), e17695.
44. Ruggeri, F. S., Adamcik, J., Jeong, J. S., Lashuel, H. A., Mezzenga, R., & Dietler, G. (2015). Influence of the β-sheet content on the mechanical properties of aggregates during amyloid fibrillization. Angewandte Chemie, 127(8), 2492-2496.

45. Cerf, E., Sarroukh, R., Tamamizu-Kato, S., Breydo, L., Derclaye, S., Dufrêne, Y. F.,& Raussens, V. (2009). Antiparallel β-sheet: a signature structure of the oligomeric amyloid β-peptide. Biochemical Journal, 421(3), 415-423.
46. Stine, W. B., Jungbauer, L., Yu, C., & LaDu, M. J. (2010). Preparing synthetic Aβ in different aggregation states. In Alzheimer's Disease and Frontotemporal Dementia (pp. 13-32). Humana Press, Totowa, NJ.
47. Chang, Y. J., & Chen, Y. R. (2014). The coexistence of an equal amount of Alzheimer's amyloid- β 40 and 42 forms structurally stable and toxic oligomers through a distinct pathway. The FEBS journal, 281(11), 2674-2687.
48. Adamcik, J., Berquand, A., & Mezzenga, R. (2011). Single-step direct measurement of amyloid fibrils stiffness by peak force quantitative nanomechanical atomic force microscopy. Applied Physics Letters, 98(19), 193701.
49. I. N. Sneddon, Int. J. Eng. Sci., 1965, 3, 47–57
50. Nandi, D., Tahiliani, P., Kumar, A., & Chandu, D. (2006). The ubiquitin-proteasome system. Journal of biosciences, 31(1), 137-155.

CHAPTER -4

Identification and evaluation of aptamer-based synthetic compounds for regulation of toxic conformer of amyloid beta 42

(Interaction between nucleic acid and peptide)

4.1 INTRODUCTION

4.1.1 Challenges with Aβ42 in Alzheimer's Disease

Alzheimer's Disease (AD) is well known to be characterized by a progressive deterioration of the episodic memory. As the disease progresses onto a later stage, symptoms that include aphasia, apraxia, executive disorders, psychiatric symptoms etc. start to develop [1]. This, however, is the story of the later stages of AD. Early onset of dementia is defined to start 10-15 years before the critical symptoms start to appear. The early symptoms are identical to that of normal ageing and thus are often misunderstood. Since it is not being able to be diagnosed at an early stage, no specific therapy that can cause a delay in the AD progression have been developed. The social and economic burden of Alzheimer's disease is therefore high. [2].

At present, the neuropathologic test of the brain is the gold standard for diagnosing AD [3]. The plaques and tangles in the brain tissues are the hallmarks of the disease. This complex process makes the need to use specific biomarkers important that can help in rapid and early diagnosis of AD. Aβ42, which is the main causative agent, and the key biomarker for the pathogenesis of AD can thus be particularly important from diagnostic point of view.

Antibodies are an ideal choice to produce biomarkers. However, antibody-based tools rely exclusively on antibody production for diagnosis and for this reason they are the major cause of diagnostic delays [4]. A few antibodies have been reported until now, however the success rates are still low because the small size of Aβ42 (~5kDa protein) and large size of antibodies makes it very difficult to create a suitable environment for their functionality. To work with a small target molecule, a small biomarker is required that fits well with Aβ42.

4.1.2 Aptamers as a new class of modern affinity reagents

Aptamers are single-stranded oligonucleotide or peptide molecules that bind to a specific target molecule and have emerged as smart class of biomolecules in medicine and biology [5]. Aptamers are selected for high-specificity and strong-affinity binding

to molecular targets by selecting them from a large random sequence pool [6]. The term ‘aptamer’ is derived from the Latin word “altus”, which means ‘to fit’ and the Greek word “meros”, which means ‘particles’ and is referred to biomolecules (nucleic acids and peptides) which can form complex shapes and thus may act as scaffolds for molecular interactions (Fig 1.4) [7.8]. Aptamers can bind selectively to a target protein and, in general, can be raised against any kind of target which makes them a better choice (as compared to antibodies) for several applications. Aptamers are more efficient than antibodies, owing to the ease of modifying active groups, adding useful small compounds (e.g., fluorescent dyes, biotin), renaturing, and designing structural changes [9]. The binding ability and specificity of aptamers are comparable to (or, in certain cases, superior to) those of antibodies. When compared, the general size of an antibody ranges from about 150-170kDa, while an aptamer can be as small as 12-25 kDa (~ 30-80 nucleotides). Thus, aptamers have the potential for application as molecular recognition elements in research and diagnostics [10]. Aptamers are also widely being used to understand the protein behavior and protein function analysis [11].

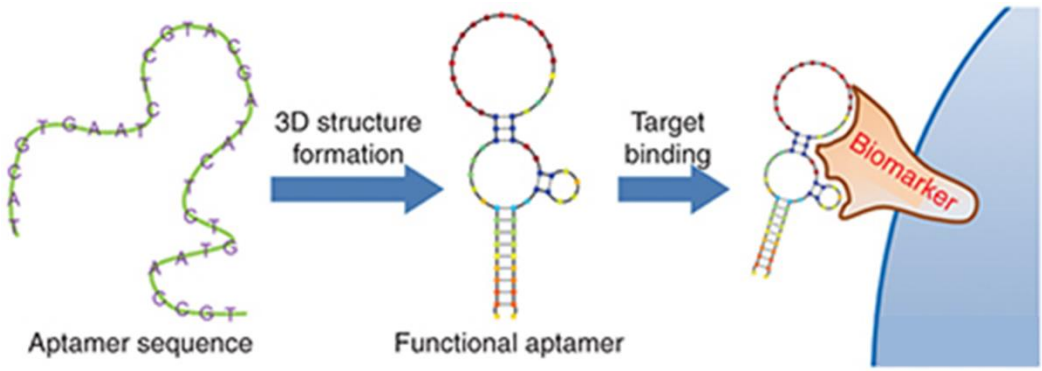


Fig 4.1: Schematic diagram of aptamer binding to its target molecule *{source: Mol Ther Nucleic Acids. 2014; 3(8): e182}*

4.1.3 Potential of aptamers to target different conformers of Aβ42

The focused challenge highlighted in the above section was the need of a small molecule that fits well with small sized biomarker (Aβ42 in our case). Aptamers can be an ideal choice for the same.

Recent studies which have reported a toxic conformer of Aβ42 (E22P-Aβ42) ,where amino acid E (Glutamic Acid) at 22nd position of the 42 amino acid long peptide (as in wild-type Aβ42, i.e. E22- Aβ42) changes to P (Proline) has been extensively studied in Chapter-3 [12]. E22P-Aβ42 exhibits a strong neurotoxicity and synaptotoxicity and thereby making it more prone to risk [13,14]. Several studies that confirm the rapid oligomerization and toxic behavior of E22P-Aβ42 have been reported [15,16].It is also reported that E22P-Aβ42 does not forms senile plaques in AD patients but are present in soluble oligomeric forms [17]. However, it is still a topic of interest for researchers to understand and regulate the molecular mechanism and toxic behavior of E22P-Aβ42 at structural levels.

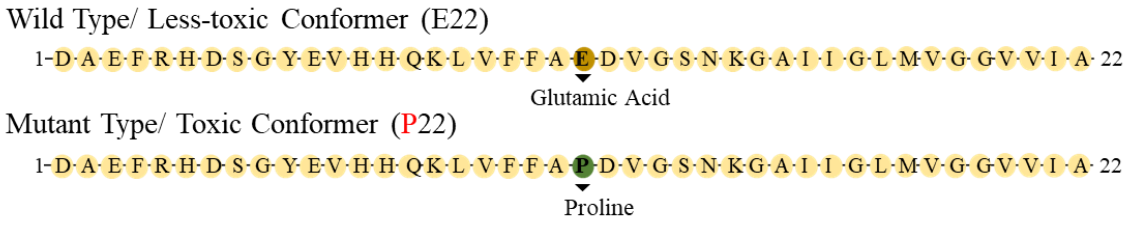


Figure 4.2: Targets E22-Aβ42 and E22P-Aβ42: Sequence information of variants E22-Aβ42 and E22P-Aβ42 showing mutation at 22nd position where amino acid glutamic acid changes to proline that leads to higher aggregation and toxicity.

Considering the recent finding which highlights the toxic conformer of Aβ42, it can be very effective to characterize aptamer that can identify the toxic, short-lived metastable structure of Aβ42. Furthermore, the study can be very useful if we are able to differentiate between the toxic and non-toxic conformers of Aβ42.

4.1.4 SELCOS: The competitive non-SELEX approach for screening aptamers.

Systematic Evolution of Ligands by Exponential Enrichment (SELEX) has been extensively utilized to screen aptamer molecules since long back [18]. Many aptamer-based studies have been performed in the field of Aβ research. Several aptamers have been selected for Aβ which includes DNA Aptamer targeting Aβ40 oligomers [19], RNA Aptamer targeting Aβ40 fibrils and oligomers [20], RNA Aptamer targeting other amyloid fibrils [21], DNA Aptamer targeting α-synuclein that also recognizes Aβ40 oligomers [22,23], Peptide Aptamer targeting Aβ42 fibrils [24], RNA Aptamer targeting toxic type Aβ42 [25] etc.

Aptamers in all these cases are screened by the SELEX method, which is typically carried out for a single target protein in solution and by performing repeated rounds of selection and enrichment using PCR Amplification. A new competitive non-SELEX method (SELCOS) has been reported for the selection of highly selective aptamer candidates for a target protein in the presence of a highly similar competitive target molecule [26]. The principle used is that in the presence of highly competitive molecule, enriched screening of aptamers can take place, and thus we can precisely distinguish between highly similar targets without multiple PCR amplification steps. Specific binding can be improved than non-specific binding during competitive selection.

Following this, we utilized SELCOS to screen DNA aptamer for E22P-Aβ42 in the presence of its highly similar counterpart, E22-Aβ42 to make a deeper understanding about the differences in both the target peptides.

This chapter is thus focused on identifying novel DNA aptamer molecules as modern affinity reagents that can selectively and specifically identify E22-Aβ42 and E22P-Aβ42.

OBJECTIVE: Identification and evaluation of aptamer-based synthetic compounds for regulation of toxic conformer of amyloid beta 42

4.2 MATERIALS AND METHODS

4.2.1 Materials

All DNA strands including DNA Library and primer sequences used for selection and evaluation were ordered from Eurofins (Tokyo, Japan). Biotinylated E22P-Aβ42 was custom synthesized from Scrum Inc. (Tokyo, Japan) and AnyGen. Co., Ltd, Korea. Biotinylated E22-Aβ42 was ordered from Anaspec, Japan). Dynabeads MyOne streptavidin magnetic beads T1 were purchased from ThermoFisher Scientific, Japan. Gold colloids (40nm) were purchased from BBI Solutions, Japan. Disposable electrochemical screen-printed (DEP) chip was obtained from BioSeeds Co. Ltd., Japan. TB Green® Fast qPCR Mix and HotStart Ex Taq ® Hot Start Version were purchased from TaKaRa, Japan. Precision Plus protein standard Protein marker was purchased from BioRad and PCR Marker was purchased from Promega. Buffers and solutions used were prepared in laboratory. Salts and chemicals used preparations were from Wako, Dojindo, Nacalai tesque and Sigma.

4.2.2 Preparation of Aβ42 species

The pretreatment of Aβ42 was done according to the previous literatures [27]. To prevent the rapid aggregation and to maintain the monomer state of Aβ42 (both E22-Aβ42 and E22P-Aβ42), it was dissolved in ice-cold HFIP (1,1,1,3,3,3-hexafluoro-2-propanol) at 1 mM concentration. After incubation at room temperature for 30 min, the solution was sonicated for 1 min and aliquoted, then left open to dry in hood overnight. To remove any further traces of HFIP, Aβ42 was vacuum dried under freezing conditions for 90 min. The resultant film of Aβ42 was stored at −80 °C until use.

4.2.3 In vitro selection of DNA Aptamers by SELCOS

Library Design and Primers: The DNA library used for selection was made up of a random 30-nucleotide region flanked by a 20-nucleotide primer region on both sides, specifically,5'-
AGCAGCACAGAGGTCAGATG(N30)CCTATGCGTGCTACCGTGAA-3'

(Library Size: 10^{18} molecules). For PCR amplification, the forward primer 5'-AGCAGCACAGAGGTCAGATG-3' and the biotinylated reverse primer 5'-TTCACGGTAGCAGCGATAGG-3' were used.

Immobilization of target on beads: 530pmol biotinylated E22P-Aβ42 monomers were immobilized onto 20µl streptavidin-coated magnetic beads (1µm) by incubating the magnetic beads with E22P-Aβ42 in PBS (pH 7.4) for 30 min at room temperature using gentle rotation. Magnetic beads were then washed with 0.1% BSA in phosphate buffer saline (PBS) twice and resuspended in PBS. The protocol stated by the manufacturer was followed. For counter selection, biotinylated E22-Aβ42 monomers were immobilized onto the streptavidin-coated magnetic beads in a similar way.

Negative Selection: To renature oligonucleotides into thermodynamically stable secondary and tertiary structures, the library was heated to 95°C for 5 min and immediately cooled to 4°C and placed for 15 min on ice, followed by incubation at room temperature for 5 min. To remove any sequences from the library that binds non-specifically to the magnetic beads, the library was incubated with the bare streptavidin coated magnetic beads on which target will be immobilized.1.5 nmol library was incubated with beads for 15 mins at room temperature. The flow-through was separated using magnet and used as the library for selection process.

Selection: Target E22P- Aβ42 immobilized on magnetic beads and free competitive target, E22-Aβ42 were incubated together with 10pmol of the ssDNA library in the presence of the binding buffer (20 mM Tri-HCl buffer (pH 7.4), 100 mM NaCl, 5 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 0.01% Tween 20) for 30 min at room temperature with gentle rotation. The reaction volume was kept 1 ml. The supernatant was then removed with the help of a magnetic separator by washing twice with washing buffer (20 mM Tri-HCl buffer (pH 7.4), 200 mM NaCl, 5 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 0.01% Tween 20). In each wash, the sample solution was mixed gently, and the supernatant that consisted of unbound sequences and free competitive target with its binders was removed carefully. The same procedure was repeated 3 more times with a successive addition of 50 pmol, 250 pmol, and 1000 pmol of the ssDNA library, keeping the same incubation time and washing frequency constant to 30 min

and two times respectively. The selection pressure was increased by changing the NaCl concentration from 200 mM to 400mM, 800mM and 1600 mM, respectively, in each successive step. Since the competitive target, E22- Aβ42 was free, 500 pmol of it was also added in each successive step. Finally, the selected aptamer DNA pools, which were bounded on magnetic beads, were recovered by heat treatment. The aptamer pool bound beads were resuspended into 50 µl Binding Buffer and heated to 95°C for 5 min (with gentle tapping each minute) followed by immediate removal of the supernatant with the help of magnetic separator. Elution was done twice for complete recovery of selected pool.

Table 4.1: Buffer Composition for SELCOS

SELCOS Buffers	Composition	
Binding Buffer (BB)	20mM Tris-HCl 5mM KCl 2mM $MgCl_2$ 1mM $CaCl_2$ 0.01% Tween 20	100 mM NaCl
Washing Buffer 1 (WB-1)		200 mM NaCl
Washing Buffer 2 (WB-2)		400 mM NaCl
Washing Buffer 3 (WB-3)		800 mM NaCl
Washing Buffer 4 (WB-4)		1600 mM NaCl

Counter Selection: The selected aptamer-pool for E22P-Aβ42 which was eluted from beads was then briefly incubated for 30 min with bead-immobilized competitive target, i.e., 500 pmol biotinylated E22-Aβ42 immobilized on streptavidin-coated magnetic beads. This was done to remove any non-specific aptamer candidates, if present. The flow through was collected and this counter selection round was repeated for 5 continuous times to enrich the selected pool. Reaction volume was 1 ml and 500 µl flow through was taken for each counter round. The enriched-selected pool for E22P-Aβ42 and selected pool for E22-Aβ42 (counter-target beads from counter round) by SELCOS were then amplified by PCR (initial incubation at 98 °C for 2 min, followed by 30 cycles of 98 °C for 10 s, 59 °C for 5 s, and 72 °C for 10 s, and finally,

72 °C for 4 min). Gel electrophoresis was used to monitor the successful amplifications using an 8% polyacrylamide gel with 8 M urea at a temperature of 55 °C.

Table 4.2: Primer Sequences for SELCOS

Primer	Sequence	Label
Forward Primer	5′-AGCAGCACAGAGGTCAGATG-3′	-
Reverse Primer	5′-TTCACGGTAGCAGCGATAGG-3′	5′ Biotin-labelled

4.2.4 Electrochemical Analysis of selected pool

Electrochemical measurements using DEP-chip were performed to confirm the binding of selected pool to their respective targets. DEP-Chip, which is screen printed electrode works on the principle of 3-electrode system for the electrochemical analysis, with a 3mm carbon-based working electrode, counter electrode and Ag/AgCl reference electrode. For the assay, 27 pmol E22-Aβ42, E22P-Aβ42 and 1% BSA (Negative Control) were separately immobilized to per µl of streptavidin magnetic beads. A 10 µl sample made of gold colloids conjugated to thiolated ssDNA of selected pool of E22-Aβ42 was incubated with bead immobilized positive target E22-Aβ42, competitive target E22P-Aβ42 and control target BSA. Likewise, 10µl sample made of gold colloids conjugated to thiolated ssDNA of selected pool of E22P-Aβ42 was incubated with bead immobilized positive target E22P-Aβ42, competitive target E22-Aβ42 and control target BSA. After incubation for 30 mins at room temperature under gentle rotation conditions in the presence of PBS buffer, the bound gold colloids on target immobilized magnetic beads were collected using magnet. The magnetic beads were then rinsed with washing buffer that was used for selection (20 mM Tri-HCl buffer (pH 7.4), 200 mM NaCl, 5 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 0.01% Tween 20). The pool bound magnetic beads were then diluted 10 folds and resuspended in the binding buffer(20 mM Tri-HCl buffer (pH 7.4), 100 mM NaCl, 5 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 0.01% Tween 20). 2µl sample was dropped and dried at 37°C on the working electrode of DEP chip, and thereafter connected to the electrochemical analyzer system. 25 µl of 0.5M HCl was dispensed onto the DEP-

chip to electrooxidate the gold colloids at a constant potential of +1.25 V for 40 s, immediately followed by DPV detection from +0.8 to 0 V, with a step amplitude of 4 mV, a pulse amplitude of 50 mV, and a pulse period of 0.2 s with a scan rate of 20 mV/s.

4.2.5 Sequencing and analysis

The selected pools for E22-Aβ42 and E22P-Aβ42 were sequenced by Next-Generation Sequencing. Further, for filtering, clustering, and identifying aptamer candidates for further evaluation, web-based tools Clustal Omega and UNAFold web server for multiple sequence alignment and secondary structure analysis, respectively were used.

4.2.6 Quantitative Aptamer-Target Binding Analysis by qPCR

66pmol Aβ42 was first bound on 2.5ul streptavidin magnetic beads followed by incubation with 10pM heat-cooled aptamers, in PBS for 15 mins. Aptamers bound to the target coated on beads were then immediately separated with the help of magnet and the flow through was discarded. Bound aptamers were washed with 400mM NaCl and 0.01% Tween 20 in PBS to remove off the non-specific binders, if any. The quantification of aptamers bound to the bead immobilized Aβ42 was carried out by qPCR analysis (Mx3000P, Stratagene) using SYBR Green I after eluting the aptamers from the beads by heating at 95°C for 3 mins followed by separation with magnet. 25µL qPCR reaction was set up containing 5 µL template ,12.5µL of TB Green® Fast qPCR Mix, 1µL each of 10µM forward and reverse primers, 0.5µL of RoX II reference dye and the rest of the volume was made up with NFW. The reactions were carried out in triplicates. The thermal cycling procedure started with the initial denaturation for 30 s at 95°C and was followed by 50 rounds of amplification. Each round of operation included denaturation at 95°C for 10 s, annealing at 59°C for 5 s, extension at 72°C for 10 s and final extension at 72°C for 4 min. Poly A sequence was used a negative control.

4.3 RESULTS

4.3.1 Competitive Screening of DNA aptamers for E22-Aβ42 and E22P-Aβ42

We performed in vitro selection of DNA aptamers that can specifically recognize these highly similar conformers of Aβ42 using competitive non-SELEX approach. Competitive non-SELEX utilizes the concept of enrichment in the equilibrium state of the interaction between the targets and the pool of ligands. Presence of a competitor molecule can enhance the specificity of the selected candidate. **Fig 4.3** shows the schematics of the selection process used in this study to obtain specific DNA aptamers for E22-Aβ42 and E22P-Aβ42.

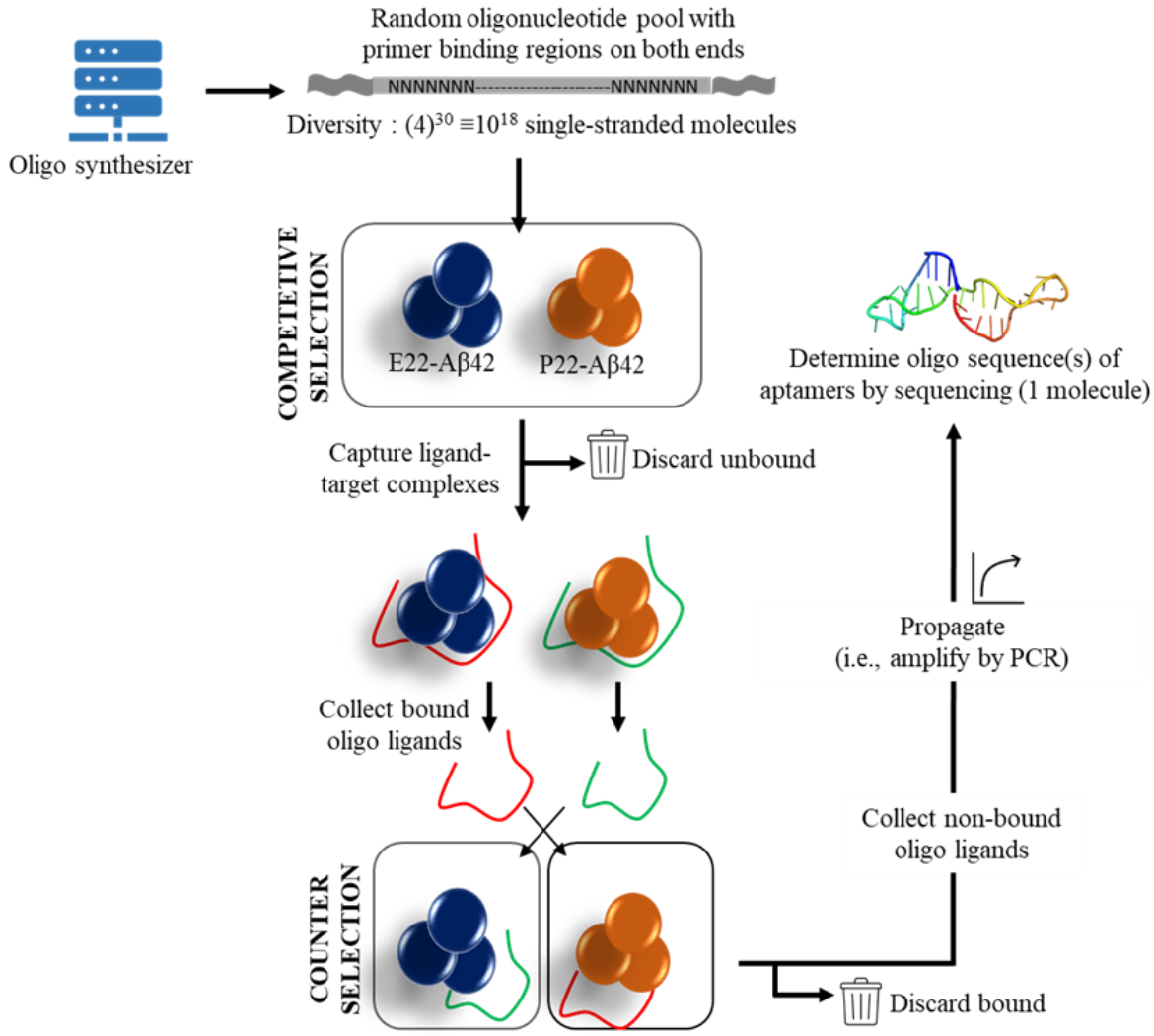


Figure 4.3: Schematics of selection of aptamers for E22P-Aβ42 using competitive Non-SELEX (SELCOS) method.

4.3.2 Evaluation of selected pool by gel electrophoresis

Competitive non-SELEX utilizes the concept of enrichment in the equilibrium state of the interaction between the targets and the pool of ligands. Presence of a competitor molecule can enhance the specificity of the selected candidate. Considering this, *in vitro* selection of DNA aptamers for E22P- Aβ42 was carried out in the presence of its competitive target, i.e., E22-Aβ42. 5-rounds of counter selection were also carried out to remove any non-specific candidates and confirmed by PCR and gel electrophoresis.

Polyacrylamide gel electrophoresis (PAGE), which is a common conventional approach to separate DNA based on size, was used to evaluate the selection. The enriched-selected pool for E22P-Aβ42 (flow through from counter selection round) and selected pool for E22-Aβ42 (counter-target beads from counter round) were then amplified using the following conditions: 20 cycles at initial incubation at 98 °C for 2 min, followed by 98 °C for 10 s, 59 °C for 5 s, and 72 °C for 10 s, and finally a 4 min extension at 72 °C.

The successful amplifications were analyzed by PAGE using 8% polyacrylamide gel with 8 M urea at a temperature of 55°C. The appearance of clear bands, both before and after counter selection round confirms the selection and enrichment. **Figure 4.4** (left panel) shows the amplified selected pool after the 5 counter selection rounds.

However, the appearance of band in **Figure 4.4** (right panel) reveals that the selected pool that binds to toxic Aβ42 still binds to the E22-Aβ42 even after 5 rounds of counter selection.

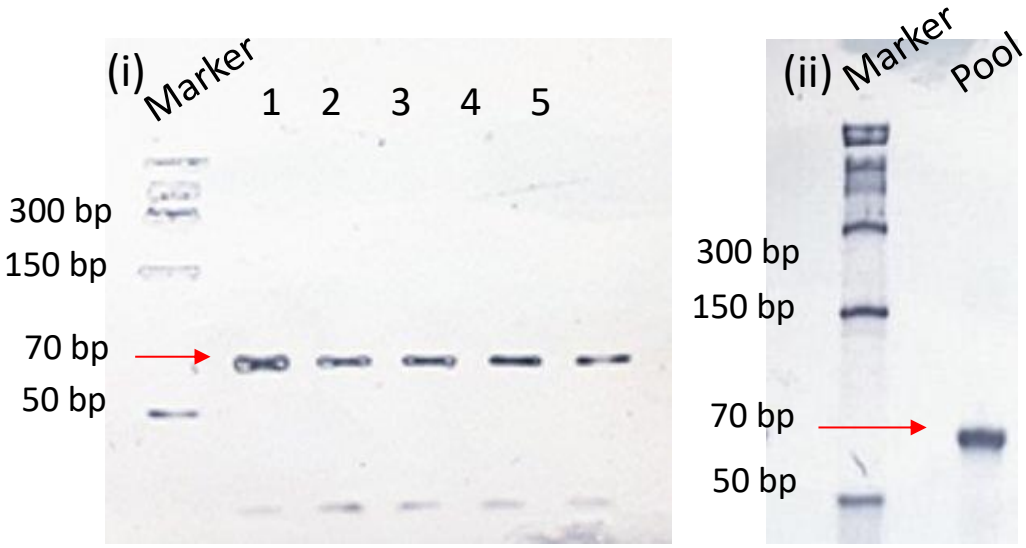


Figure 4.4: (i) PCR amplified pool after counter selection rounds using non-competitive SELEX. Lanes 1-5 are enriched selected pool after counter rounds 1-5. #5 is the selected pool for E22P-Aβ42. (ii) PCR amplified pool after eluting the sequences from counter-target beads after 5th counter-selection round. This is the selected pool for E22-Aβ42.

4.3.3 Analysis of selected pool enrichment by temperature gradient gel electrophoresis

Temperature Gradient Gel Electrophoresis (TGGE), is an easy and sensitive well-established method to monitor the diversity of oligonucleotides. It is based on the melting behavior of double stranded DNA in a temperature gradient which is sequence-enrichment dependent. As the dsDNA is electrophoresed through the denaturing temperature gradient, the dsDNA starts to melt. The diversity can be readily observed by the melting profile from native double-stranded DNA to denatured single-stranded DNA using TGGE analysis.

The TGGE analysis of the initial random pool (library before selection) and the selected pool (after 5th round of counter selection) was carried out using 8% polyacrylamide gel with 6 M urea at a temperature gradient from 15°C to 55°C. The complicated transition pattern in the TGGE profile of initial random pool is drastically changed to a linear transition pattern in the TGGE profile of selected pool (**Figure**

4.5). This result supported the successful enriched selection of DNA aptamer against E22P-Aβ42.

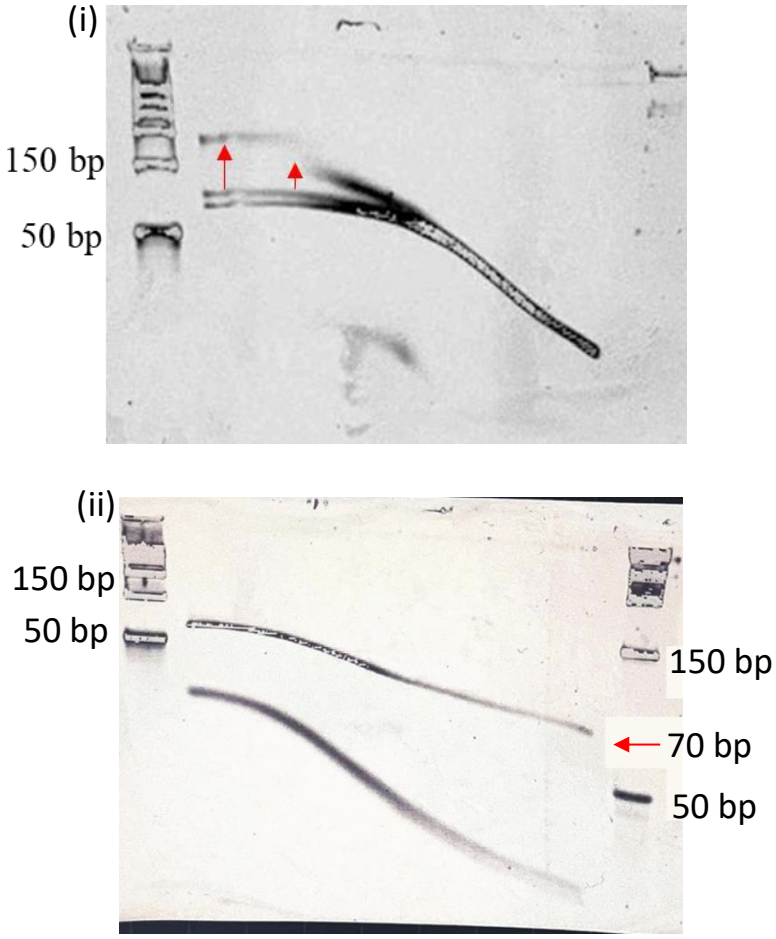


Figure 4.5: Diversity analysis of initial library and selected library by TGGE.
(i) TGGE profile of initial library (ii) TGGE profile of selected pool

4.3.4 Electrochemical binding evaluation for E22-Aβ42 and E22P-Aβ42 pool

The binding affinity of the selected pool of E22-Aβ42 for E22-Aβ42 and selected pool for E22P-Aβ42 for E22P-Aβ42 were evaluated by differential pulse voltammetry (DPV) using the electrochemical sensor DEPSOR. The DEPSOR system has been recently developed as a rapid and sensitive method for evaluating the affinity between target proteins and selected aptamers. A DPV assay was performed followed by the interaction between the bead bound protein and selected pool-conjugated gold

nanoparticles. **Figure 4.6** shows the schematics of the assay. A clear signal peak can be detected sensitively when gold nanoparticles bound to bead-attached target protein are detected onto the working electrode of DEPSOR. **Figure 4.7a** shows the obtained voltammograms and corresponding bar graph that represents the relative binding affinity between the selected ssDNA pool for E22-Aβ42 and E22-Aβ42 .The current peak heights suggest that the ssDNA pool obtained for E22-Aβ42 showed 3.85-fold higher affinity than the competitor E22P-Aβ42 protein and 3.3-fold higher affinity than the nontargeted BSA protein. Similarly, in **Figure 4.7b** the current peak heights suggest that the ssDNA pool obtained for E22P-Aβ42 showed 5.91-fold higher affinity than the competitor E22-Aβ42 protein and 7.72-fold higher affinity than the nontargeted BSA protein. The signals in each case include average of 3 independent readings. These results confirmed the successful competitive selection of aptamer molecules for E22-Aβ42 and E22P-Aβ42 with high specificity.

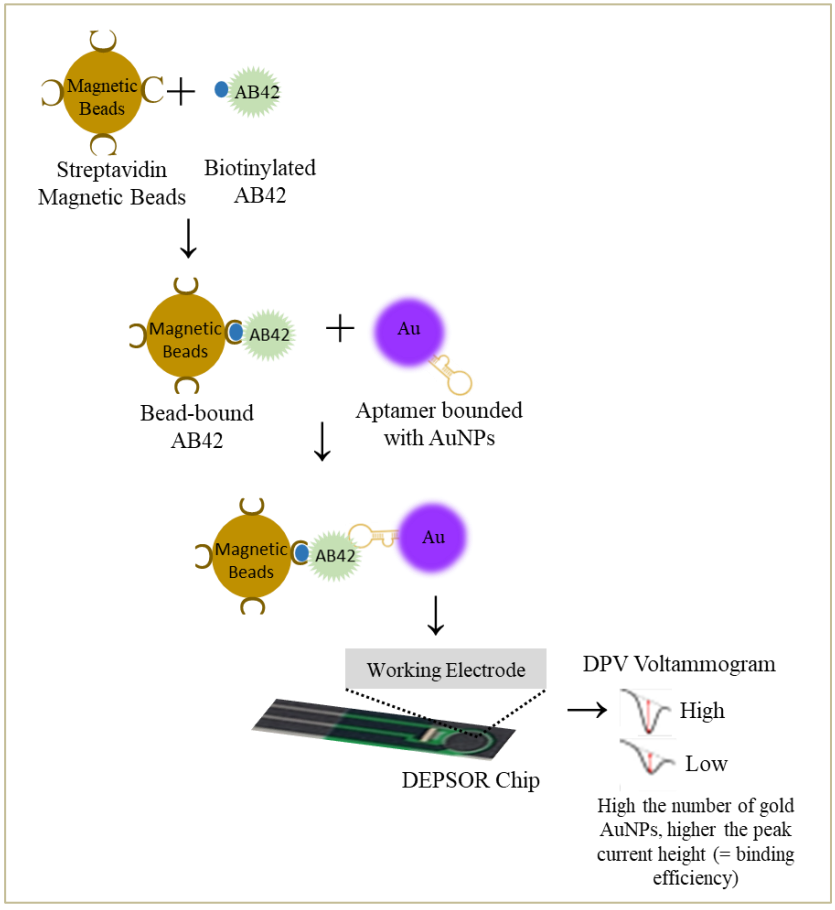


Figure 4.6: Schematics of the assay for electrochemical evaluation of the selected pool for E22-Aβ42 and E22P-Aβ42

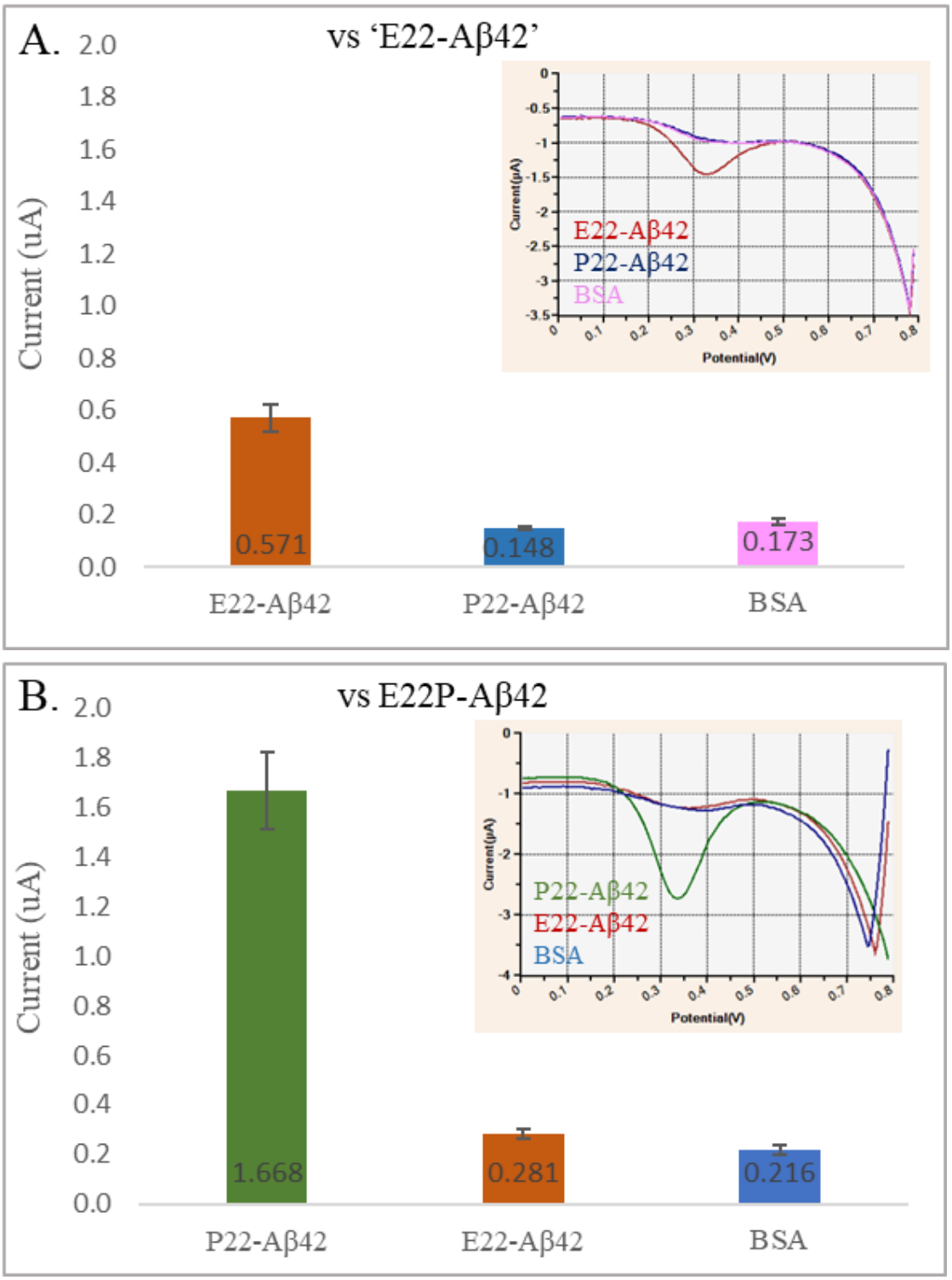


Figure 4.7: The selected pool for (A) E22-Aβ42 and (B) E22P-Aβ42 was measured against their respective targets, their competitive targets and BSA respectively. The DPV curves and corresponding graphs are shown.

4.3.5 Identification of aptamer sequences by next-generation sequencing

The selected ssDNA pools for E22-Aβ42 and E22P-Aβ42 were analyzed by next-generation sequencing (NGS). In case of E22-Aβ42, a total of 430K sequences were read, out of which 333K sequences were fit for analysis. Out of these, 100 sequences were selected with maximum number of repeats. In the case of E22P-Aβ42, a total of 387K sequences were read, out of which 300K sequences were fit for analysis. Out of these, again 100 sequences were selected with maximum number of repeats. Next, from the top 100 repeated sequences in both E22-Aβ42 and E22P-Aβ42, a total of 36 and 33 sequences were specific to E22-Aβ42 and E22P-Aβ42 respectively **(Figure 4.8A)**. None of the selected sequences for E22-Aβ42 were present in the sequences of E22P-Aβ42 and vice-versa. These selected 36 and 33 sequences were aligned in a phylogenetic tree to cluster the sequences according to their sequence homology and sequence distance by using ClustalW2 software **(Figure 4.8B)**. Also, the secondary structures of selected 36 and 33 sequences were predicted by mfold software. Thereafter, based on the secondary structure pattern consisting of loops and bulges, the aptamers were grouped into one of the following groups:

A) Acute-bulge shape hairpin stem/loop;

B) Right-bulge shape hairpin stem/loop;

C) Obtuse-bulge shape hairpin stem/loop;

D) Linear shape hairpin stem/loop and

E) Complex shape hairpin stem/loop (**Figure 4.9)**.

Further, 6 representative aptamer candidates for E22-Aβ42 and 2 for E22P-Aβ42 were selected based on their binding energies and dominant secondary structures from the groups as shown in **Figure 4.10**. The sequence details and properties of these 6+2 aptamer candidates are listed in **Table 4.3.** These aptamer sequences were then chemically synthesized for further evaluation.

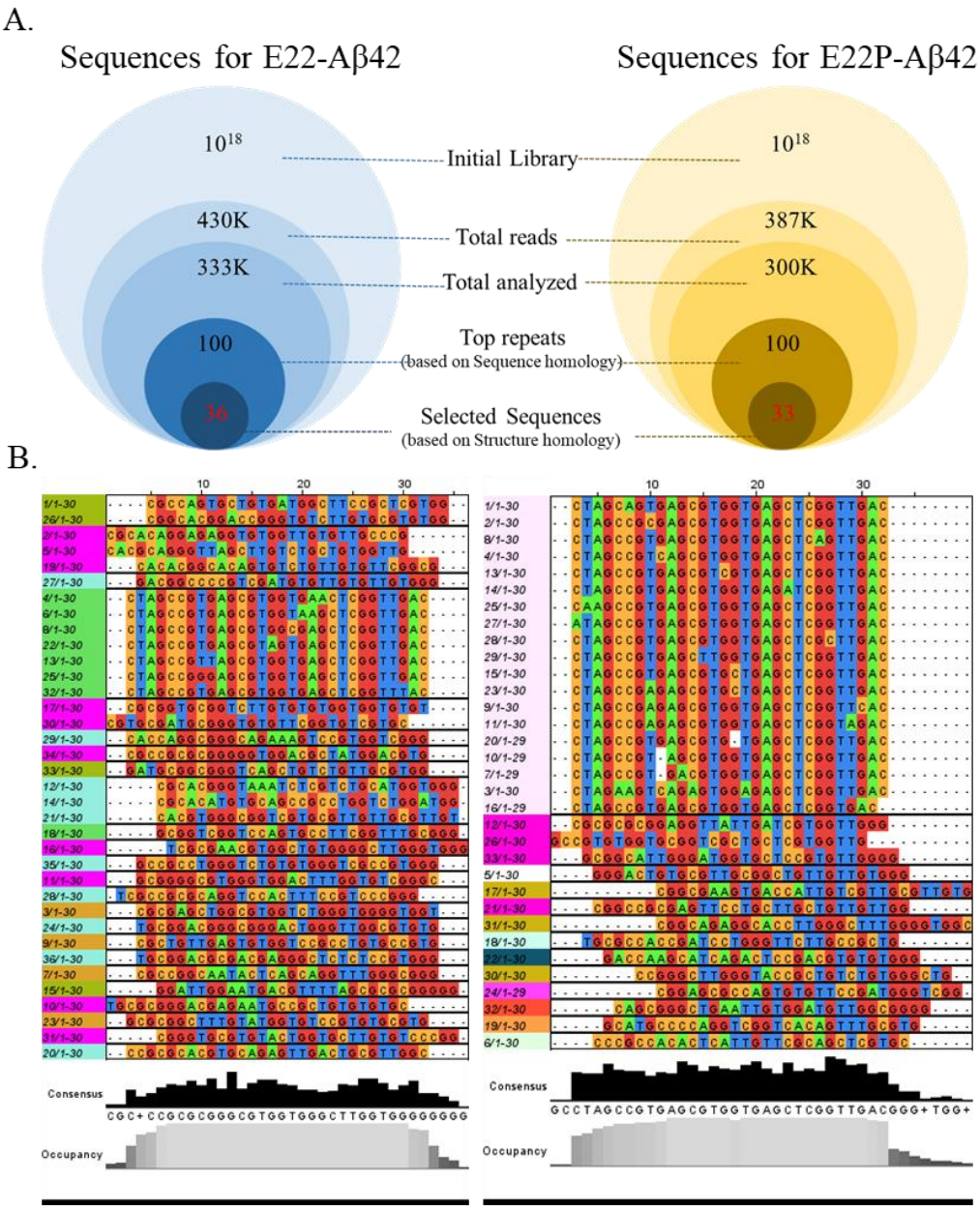


Figure 4.8: NGS Sequencing Analysis of the selected pool for E22-Aβ42 and E22P-Aβ42: (A) A venn diagram showing the hierarchy of selected sequences starting from 10^{18} molecules during the selection process to a final of 36 and 33 sequences in E22-Aβ42 and E22P-Aβ42 respectively after the next generation sequencing. (B) The selected 36 and 33 sequences were aligned in a phylogenetic tree to cluster the sequences according to their sequence homology and sequence distance by using ClustalW2 software.

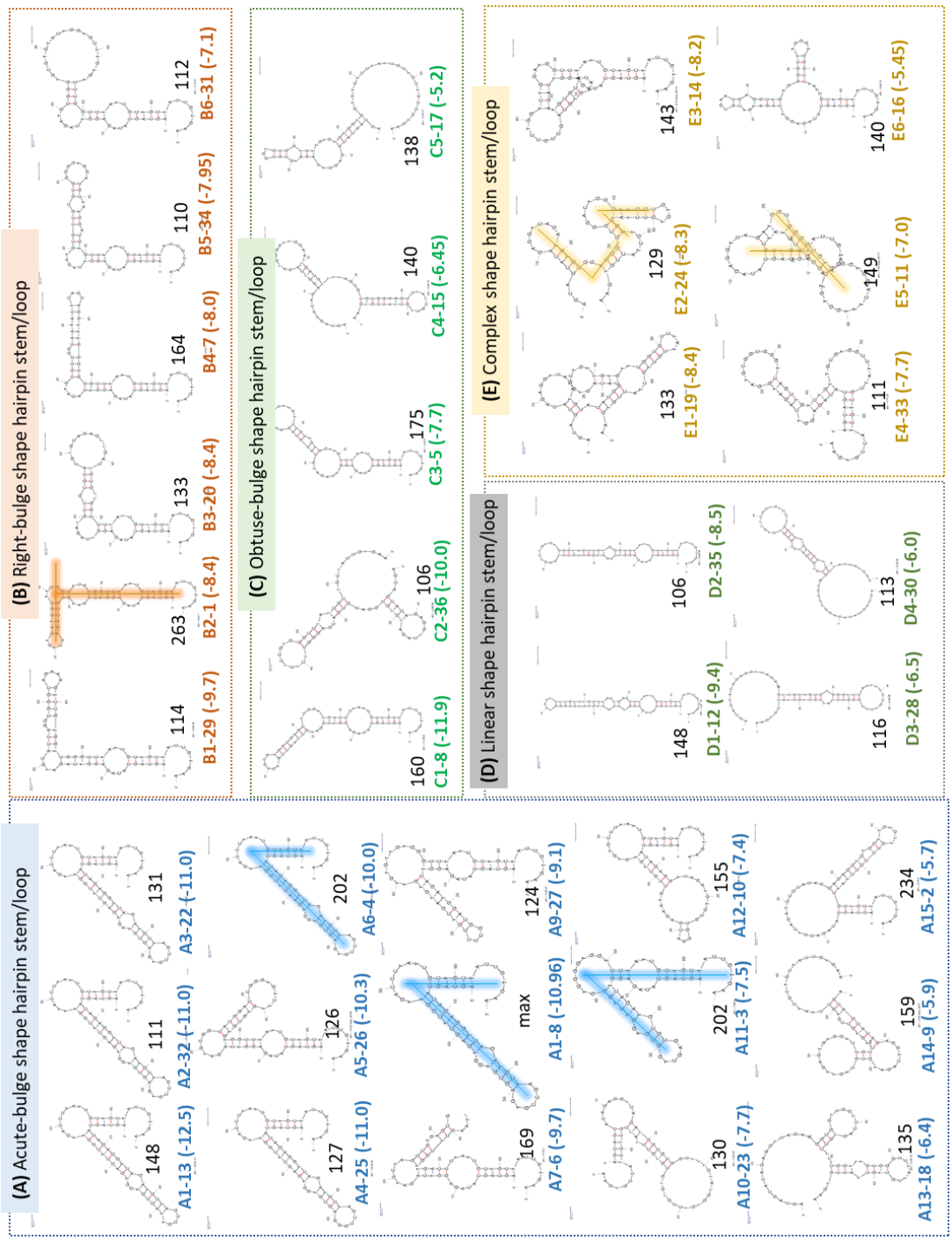


Figure 4.9 (A): Five secondary structure groups of selected aptamer candidates for E22- Aβ42

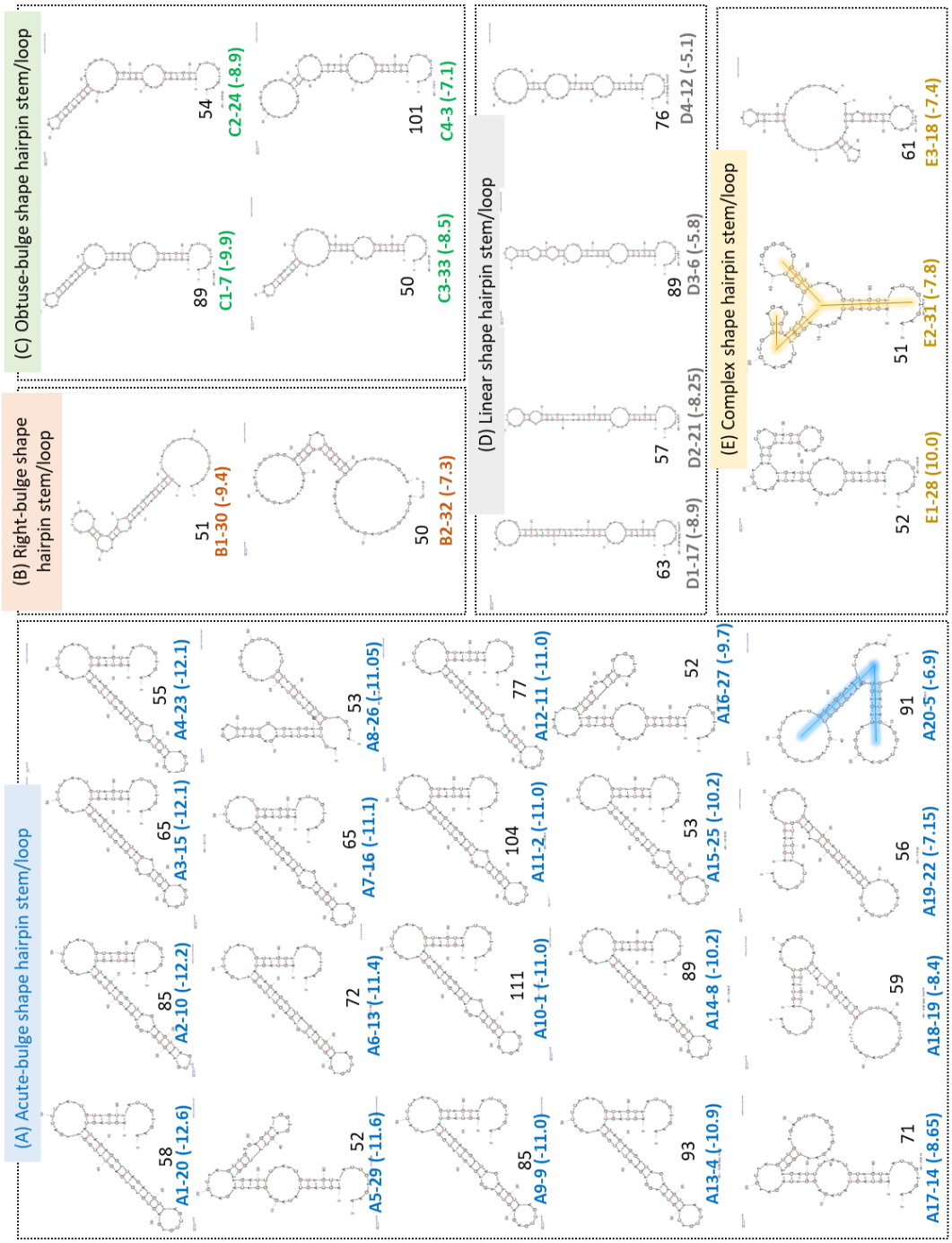


Figure 4.9 (B): Five secondary structure groups of selected aptamer candidates for E22P- Aβ42

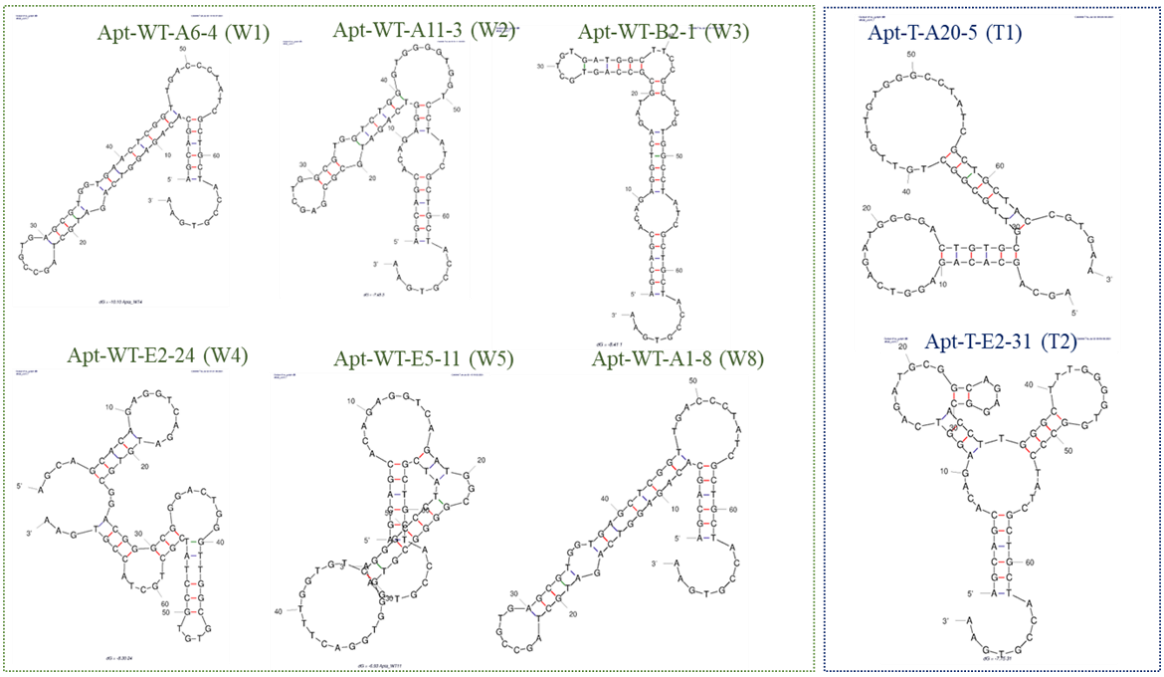


Figure 4.10: 2-D secondary structure of representative aptamer candidates for E22-Aβ42 (6) and E22P-Aβ42 (2) predicted by mfold. The parenthesis shows the aptamer code.

Table 4.3: Properties of DNA Aptamers selected for E22-Aβ42 and E22P-Aβ42 by SELCOS

S No.	Target	Aptamer Name[a]	Aptamer Code	Aptamer Sequence (5'-3')[c]	Number of repeats	ΔG (kcal/mol)[d]
1	E22-Aβ42	WT-A6-4	W1	5'PBS-CTAGCCGTGAGCGTGGTGAACTCGGTTGAC-3'PBS	202	-10
2		WT-A11-3	W2	5'PBS-CCGCGAGCTGGCGTGGTCTGGGTGGGGTGGT-3'PBS	202	-7.5
3		WT-B2-1	W3	5'PBS-CCGCCAGTGCTGTGATGGCTTCCGCTCGTGG-3'PBS	263	-8.4
4		WT-E2-24	W4	5'PBS-TGCGGACGGGCGGGACTGGGTTGGCGTGTG-3'PBS	129	-8.3
5		WT-E5-11	W5	5'PBS-GCGGGGCGTGGGTGGACTTTGGTGTCGGGC-3'PBS	149	-7
6		WT-A1-8	W8	5'PBS-CCTAGCCGTGAGCGTGGTGAGCTCGGTTGAC-3'PBS	124939	-10.96
1	P22-Aβ42	T-A20-5	T1	5'PBS-GGGACTGTGCGTTGCGGCTGTTGTTGTGGG-3'PBS	91	-6.9
2		T-E2-31	T2	5'PBS-CGGCAGAGGCACCTTGGGCTTTGGGGTGGC-3'PBS	51	-7.8

[a]Nomenclature for NGS analyzed aptamers were systematically done based on the secondary stem-loop structure
[b]Aptamer code is systematically assigned according to the secondary structure analysis-based clustering.
[c]5'-PBS and 3'-PBS are primer binding sequences (AGCAGCACAGAGGTCAGATG and CCTATCGCTGCTACCGTGAA, respectively)
[d]ΔG (kcal/mol) is free energy value and calculated from Mfold online tool

4.3.6 Binding evaluation of aptamer candidates by qPCR

The selected aptamer candidates were evaluated for their binding affinity using the qPCR analysis. **Figure 4.11** shows the schematics of the binding evaluation assay. The magnetic bead-bound Aβ42 was incubated with the aptamers to allow the binding to take place and were separated after 15 mins with the help of a magnet. Aptamers bound to the magnetic bead bound Aβ42 were then washed and eluted by temperature method. Washing was done to remove any non-specific binders. Quantification of the aptamer elute was done using qPCR.

Initially, binding affinity of Apt-W1 to E22-Aβ42 was evaluated in order to estimate and optimize the experimental conditions. **Figure 4.12** shows the amplification plot of the binding interaction of Apt-W1(true aptamer) to E22-Aβ42. PolyA sequence was used as control sequence (false aptamer).Positive and negative controls were also carried out. Only Apt-W1 was amplified as positive control and no template was used in the negative control reaction. The CT values obtained in this case were, ~16 for the positive control, ~22 for Apt-W1 with E22-Aβ42, ~28 for Poly-A aptamer with E22-Aβ42 and ~36 for negative control. A clear difference was observed in the CT values of the 4 cases. The CT value in the case of negative control can be explained as the carryover contamination.

Next, the 6 aptamer candidates Apt-W1, Apt-W2, Apt-W3, Apt-W4, Apt-W5 and Apt-W8 were evaluated for their binding affinity to E22-Aβ42 (**Figure 4.13a**). The binding order was found to be: W1> W3> W4> W8> W2> W5 and Poly A aptamer showed no binding. This confirms that all the aptamer candidates are binders of E22-Aβ42, with W1 showing the maximum binding affinity, making it a good candidate for further studies. Similarly, the 2 aptamer candidates Apt-T1 and Apt-T2 were evaluated for their binding affinity to E22P-Aβ42. Apt-T2 showed higher binding affinity as compared to Apt-T1(**Figure 4.13b**).

Next, to check the specificity of Apt-W1 and Apt-T2 to E22P-Aβ42 and E22-Aβ42 respectively, cross binding evaluation was performed. Apt-W1 and Apt-T2 were both evaluated against both E22-Aβ42 and E22P-Aβ42. As it can be seen in **Figure 4.14**, for E22-Aβ42, Apt-W1 binds strongly as compared to Apt-T2, and similarly for E22P-Aβ42, Apt-T2 binds strongly as compared to Apt-W1. This gives a preliminary idea that even though the aptamers can bind to both the conformers, but the selected aptamers are specific to some extent to their respective targets.

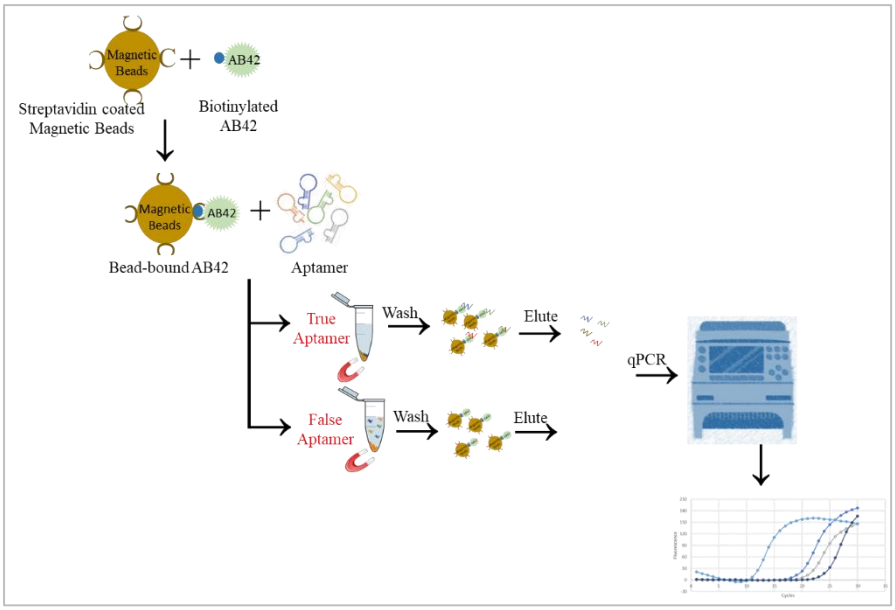


Figure 4.11: Schematics of qPCR binding evaluation assay

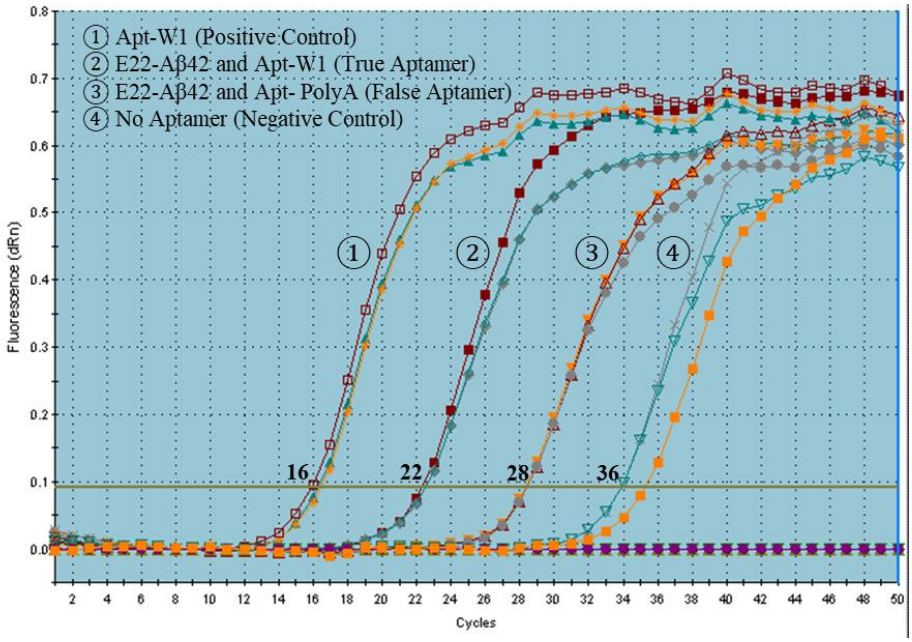
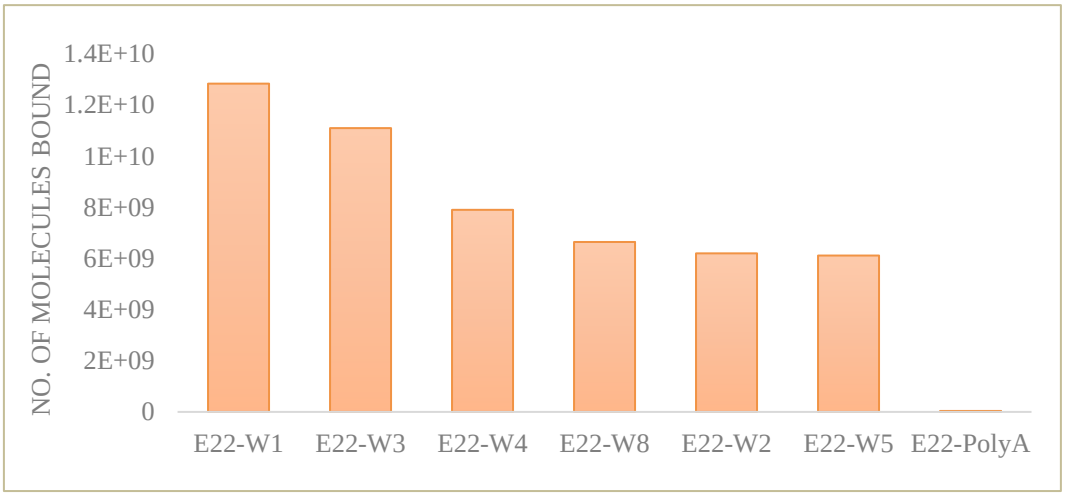


Figure 4.12 : Amplification plots of the binding interaction of Aptamer W1 with E22-Aβ42. PolyA sequence has been used as control sequence. CT values are mentioned near threshold fluorescence line.

(a)



(b)

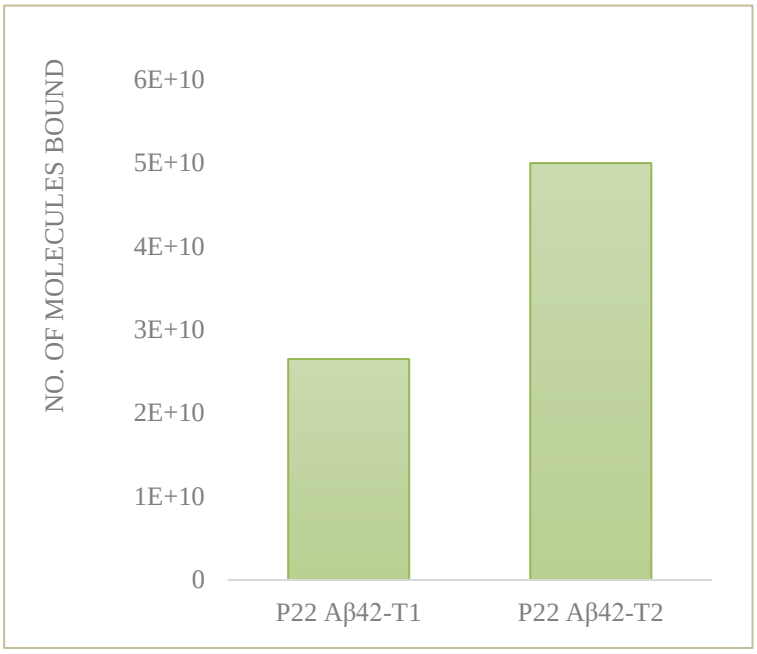


Figure 4.13 : qPCR Analysis of the affinity binding of selected representative aptamers (a) Apt-W1, Apt-W2, Apt-W3, Apt-W4, Apt-W5 and Apt-W8 to E22-Aβ42 and (b) Apt-T1 and Apt-T2 to E22P-Aβ42. Corresponding bar graphs of the number of molecules bound calculated from the CT values are shown.

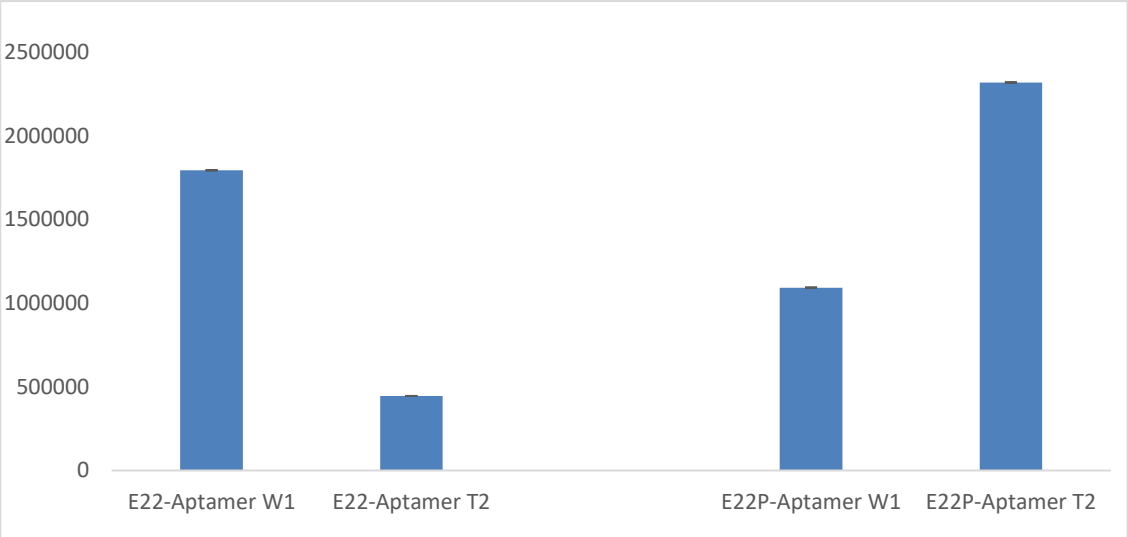


Figure 4.14 : qPCR Analysis of the cross-binding evaluation of Apt-W1 to E22-Aβ42 and E22P-Aβ42 and Apt-T2 to E22-Aβ42 and E22P-Aβ42. Corresponding bar graphs of the number of molecules bound calculated from the CT values are shown.

4.3.7 Docking prediction for binding of aptamers to Aβ42

Three dimensional strucures of Aptamers-W1 and T2 were predicted using in-silico studies as shown in **Figure 4.15**. Docking prediction was then done to evaluate the binding of Aptamer-W1 with E22-Aβ42 and Aptamer-T2 with E22P-Aβ42 as shown in **Figure 4.16**. After the interaction of Aptamer-W1 and E22-Aβ42, it was observed that only 1 out of 4 beta sheet is remaining in E22-Aβ42. Likewise in case of interaction of Aptamer-T2 and E22P-Aβ42, it was observed that only 3 out of 6 beta sheets are remaining in E22P-Aβ42. This indicates that the aptamer on interaction with Aβ42 can inhibit the beta sheet formation, and thus prevent the aggregation of Aβ42.

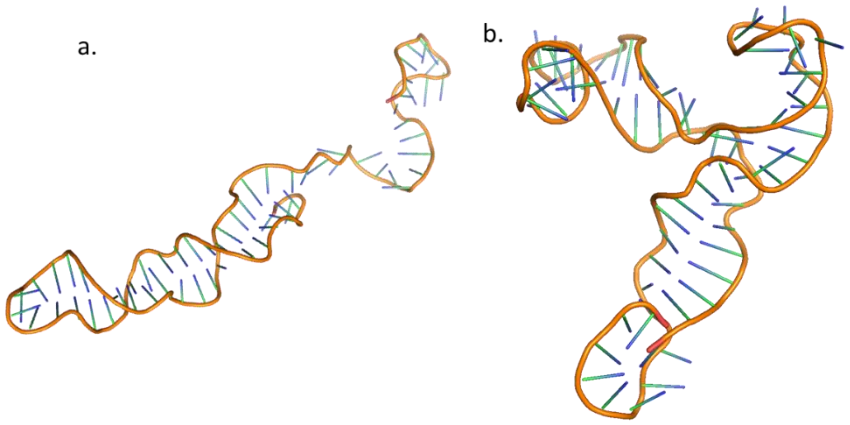


Figure 4.15: 3D-structure prediction of (a) Aptamer W1 and (b) Aptamer T2

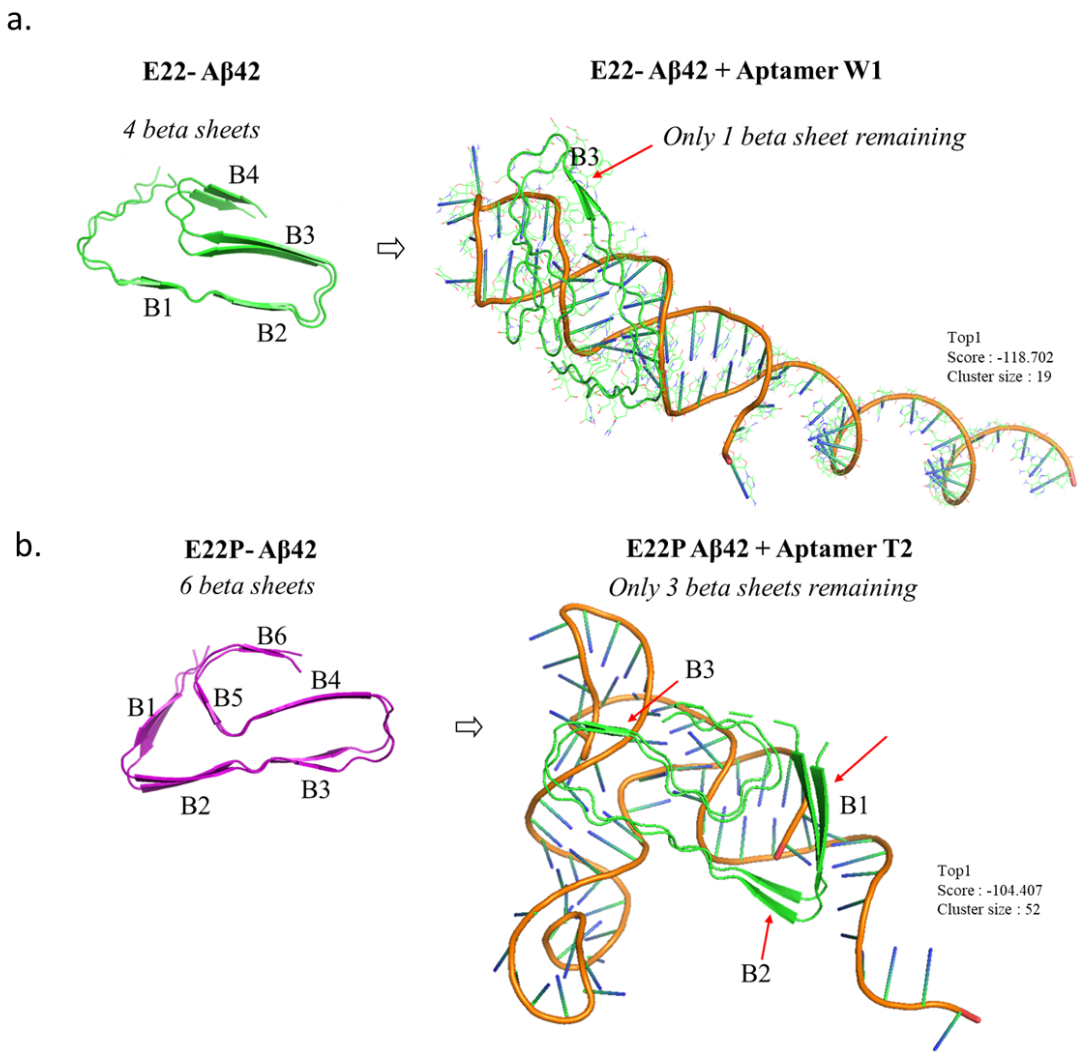


Figure 4.16: Docking prediction of binding of (a) Apt-W1 to E22-Aβ42 , and (b) Apt-T2 to E22P-Aβ42

4.4 CONCLUSIONS

Interaction between nucleic acid and peptide has been studied in this chapter. We identified novel 70-nt DNA aptamer sequences for E22-Aβ42 and E22P-Aβ42 using an in vitro competitive selection method SELCOS. The PAGE results showed a clear band after each round of counter selection clearly indicates the selection of a DNA aptamer for E22P-Aβ42. The TGGE results show a complicated transition pattern in the profile of initial random pool which is drastically changed to a linear transition pattern in the profile of selected pool. This confirms the enrichment of library during successive selection rounds.

Apt-W1 is a potential binder for E22-Aβ42 and Apt-T2 is a potential binder for E22P-Aβ42. We confirmed the binding evaluation of selected aptamers using electrochemical approach and qPCR analysis. Furthermore, preliminary molecular docking indicates that Apt-W1 and Apt-T2 can reduce the number of beta sheets on interaction with E22-Aβ42 and E22P-Aβ42, respectively indicating the selected aptamers can inhibit the beta sheet aggregation of Aβ42 .

We thus believe that our work can promote the application of aptamer in the field of Aβ research and help to make a clear understanding of the molecular mechanism of neurotoxicity caused by E22P-Aβ42.

REFERENCES:

1. Tarawneh, R., & Holtzman, D. M. (2012). The clinical problem of symptomatic Alzheimer disease and mild cognitive impairment. *Cold Spring Harbor perspectives in medicine*, *2*(5), a006148.

2. Castro, D. M., Dillon, C., Machnicki, G., & Allegri, R. F. (2010). The economic cost of Alzheimer's disease: Family or public-health burden?. *Dementia & Neuropsychologia*, *4*(4), 262-267.

3. Beach, T. G., Monsell, S. E., Phillips, L. E., & Kukull, W. (2012). Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005–2010. *Journal of neuropathology and experimental neurology*, *71*(4), 266-273.

4. Molefe, P. F., Masamba, P., Oyinloye, B. E., Mbatha, L. S., Meyer, M., & Kappo, A. P. (2018). Molecular application of aptamers in the diagnosis and treatment of cancer and communicable diseases. *Pharmaceuticals*, *11*(4), 93.

5. Zhou, J., & Rossi, J. (2017). Aptamers as targeted therapeutics: current potential and challenges. *Nature reviews Drug discovery*, *16*(3), 181-202.

6. Zhang, Y., Lai, B. S., & Juhas, M. (2019). Recent advances in aptamer discovery and applications. *Molecules*, *24*(5), 941.

7. Ku, T. H., Zhang, T., Luo, H., Yen, T. M., Chen, P. W., Han, Y., & Lo, Y. H. (2015). Nucleic acid aptamers: An emerging tool for biotechnology and biomedical sensing. *Sensors*, *15*(7), 16281-16313.

8. Keefe, A. D., Pai, S., & Ellington, A. (2010). Aptamers as therapeutics. *Nature reviews Drug discovery*, *9*(7), 537-550.

9. Liu, J., Cao, Z., & Lu, Y. (2009). Functional nucleic acid sensors. *Chemical reviews*, *109*(5), 1948-1998.

10. McConnell, E. M., Holahan, M. R., & DeRosa, M. C. (2014). Aptamers as promising molecular recognition elements for diagnostics and therapeutics in the central nervous system. *Nucleic acid therapeutics*, *24*(6), 388-404.

11. Tombelli, S., Minunni, M., & Mascini, M. (2005). Analytical applications of aptamers. Biosensors and Bioelectronics, 20(12), 2424-2434.

12. Murakami, K., Masuda, Y., Shirasawa, T., Shimizu, T., & Irie, K. (2010). The turn formation at positions 22 and 23 in the 42-mer amyloid β peptide: The emerging role in the pathogenesis of Alzheimer's disease. Geriatrics & gerontology international, 10, S169-S179.

13. Morimoto, A., Irie, K., Murakami, K., Ohigashi, H., Shindo, M., Nagao, M., ... & Shirasawa, T. (2002). Aggregation and neurotoxicity of mutant amyloid β (Aβ) peptides with proline replacement: importance of turn formation at positions 22 and 23. Biochemical and biophysical research communications, 295(2), 306-311.

14. Irie, K., Murakami, K., Masuda, Y., Morimoto, A., Ohigashi, H., Hara, H., ... & Shirasawa, T. (2007). The toxic conformation of the 42-residue amyloid beta peptide and its relevance to oxidative stress in Alzheimer's disease. Mini Reviews in Medicinal Chemistry, 7(10), 1001-1008.

15. Morimoto, A., Irie, K., Murakami, K., Masuda, Y., Ohigashi, H., Nagao, M., ... & Shirasawa, T. (2004). Analysis of the secondary structure of β-amyloid (Aβ42) fibrils by systematic proline replacement. Journal of Biological Chemistry, 279(50), 52781-52788.

16. Izuo, N., Kasahara, C., Murakami, K., Kume, T., Maeda, M., Irie, K., ... & Shimizu, T. (2017). A toxic conformer of Aβ42 with a turn at 22–23 is a novel therapeutic target for Alzheimer's disease. Scientific reports, 7(1), 1-13.

17. Tomiyama, T., Nagata, T., Shimada, H., Teraoka, R., Fukushima, A., Kanemitsu, H., ... & Mori, H. (2008). A new amyloid β variant favoring oligomerization in Alzheimer's-type dementia. Annals of neurology, 63(3), 377-387.

18. Yüce, M., Kurt, H., Hussain, B., & Budak, H. (2018). Systematic evolution of ligands by exponential enrichment for aptamer selection. In *Biomedical Applications of Functionalized Nanomaterials* (pp. 211-243). Elsevier.

19. Keefe, A. D., Pai, S., & Ellington, A. (2010). Aptamers as therapeutics. Nature reviews Drug discovery, 9(7), 537-550.

20. Takahashi, T., Tada, K., & Mihara, H. (2009). RNA aptamers selected against amyloid β-peptide (Aβ) inhibit the aggregation of Aβ. *Molecular Biosystems*, *5*(9), 986-991.

21. Rahimi, F. (2018). Aptamers Selected for Recognizing Amyloid β-Protein—A Case for Cautious Optimism. *International journal of molecular sciences*, *19*(3), 668.

22. Tsukakoshi, K., Abe, K., Sode, K., & Ikebukuro, K. (2012). Selection of DNA aptamers that recognize α-synuclein oligomers using a competitive screening method. Analytical chemistry, 84(13), 5542-5547.

23. Tsukakoshi, K., Harada, R., Sode, K., & Ikebukuro, K. (2010). Screening of DNA aptamer which binds to α-synuclein. Biotechnology letters, 32(5), 643-648.

24. Ghimire Gautam, S., Komatsu, M., & Nishigaki, K. (2015). Strong Inhibition of Beta-Amyloid Peptide Aggregation Realized by Two-Steps Evolved Peptides. *Chemical biology & drug design*, *85*(3), 356-368.

25. Murakami, K., Obata, Y., Sekikawa, A., Ueda, H., Izuo, N., Awano, T., ... & Irie, K. (2020). An RNA aptamer with potent affinity for a toxic dimer of amyloid β42 has potential utility for histochemical studies of Alzheimer's disease. Journal of Biological Chemistry, 295(15), 4870-4880.

26. Kushwaha, A., Takamura, Y., Nishigaki, K., & Biyani, M. (2019). Competitive non-SELEX for the selective and rapid enrichment of DNA aptamers and its use in electrochemical aptasensor. *Scientific reports*, *9*(1), 1-11.

27. Stine, W. B., Jungbauer, L., Yu, C., & LaDu, M. J. (2010). Preparing synthetic Aβ in different aggregation states. In Alzheimer's Disease and Frontotemporal Dementia (pp. 13-32). Humana Press, Totowa, NJ.

CHAPTER -5

General Conclusion: A novel ‘dry-lab’ concept for biomaterial interaction- based protein function regulation

5.1 UNDERSTANDINGS FROM THE THESIS

The work done in this thesis is dedicated to investigate the regulation of protein/peptide function with the help of biomaterials on the basis of their structure and interaction to proteins/peptides. The idea to achieve this was using an integration of the *in silico* and *in vitro* approach to make a better and a broader understanding than what individually *in silico* approach or *in vitro* approach could make. Considering today's time, making a complete understanding on the basis of 'only' computer-based theoretical methods is not very reliable and thus the *in silico* methods are used to assist and confirm the information that we obtain by *in vitro* methods. Completely depending on 'only' in vitro methods also has challenges that include consumption of a whole lot of time to obtain information and gain understanding. Also, the study at atomic levels requires the use of complicated methods like NMR which is not very easy in case of all biomolecules depending on their size and stability issues.

To address these situations, we first worked on the identification of small molecules based natural compounds that can work as biomaterials and regulate the function of protein in the first chapter. In a nutshell, we screened various extracts of 5 different Indian plants to evaluate their inhibitory activity for HIV-1 reverse transcriptase (RT). We found that water-leaf extracts of the leaves of Argemone mexicana strongly inhibits the DNA polymerase activity of HIV-RT which indicates that they contain substances that can regulate the enzyme (functional protein) function. Further findings revealed that the inhibitory substance is not a protein but an organic compound. We can thus conclude how biomaterial interaction (natural small-molecule compounds in this case) can help in the regulation of protein function.

Next, we worked on the idea of bio-interaction of different functional proteins that can interact and co-assist each other to develop an improved diagnostic assay in the second chapter. In a nutshell, we developed a new method RICCA (RNA isothermal co-assisted coupled amplification), which is a completely homogenous, isothermal, highly sensitive and rapid method for detecting virus RNA for on-site diagnosis of infectious diseases like COVID-19. Here we thus conclude how the biomolecular interaction of different functional proteins (RNA polymerase and Recombinase in this case) can co-assist each other to improve nucleic acid based amplification.

In the third chapter, we worked on the diversity in molecular conformations of small-sized proteins (peptides) to understand the biophysical properties of their behavior. In a nutshell, we worked on two different conformers of Aβ42 which have only one amino acid substitution difference that is responsible for making one conformer several folds more toxic than the other (More toxic - E22P, and less toxic – E22). In order to understand what causes this change in their behavior and the aggregation dynamics of the peptide, the molecular basis of the conformational transitions in the Aβ42 species was studied. Initially *in vitro* macroscopic studies based on gel electrophoresis was done followed by microscopic studies using atomic force microscopy in order to understand the difference in aggregation behavior of the conformers. We also performed *in silico* studies that revealed atomic-level information that was not possible/ rather very difficult to perform using the in vitro experiments because of the small size of the peptide. Based on the findings of both in vitro and in silico methods, we finally propose a systematic model of fibril formation in both the conformers of Aβ42 by a mixed knowledge from both the approaches. Here we conclude, how in vitro and in silico independently give certain piece of information, however we can advance our knowledge and level of understanding by co-studying both the systems together.

In the fourth chapter, we worked on the screening of aptamer-based synthetic compounds as biomaterials for specific recognition of different conformers of Aβ42 that are studied in the third chapter. In a nutshell, DNA aptamers were screened for E22-Aβ42 and E22P- Aβ42 using a new competitive screening approach, SELCOS that can help in specific recognition of highly similar targets. Using several in-vitro assays we confirmed the binding of selected aptamers to their respective targets. The in-silico experiments are still in progress that will reveal the molecular basis of aptamer-peptide interaction. Thereafter, again the information acquired by both in vitro methods and in silico methods will be co- studied to make a combined understanding.

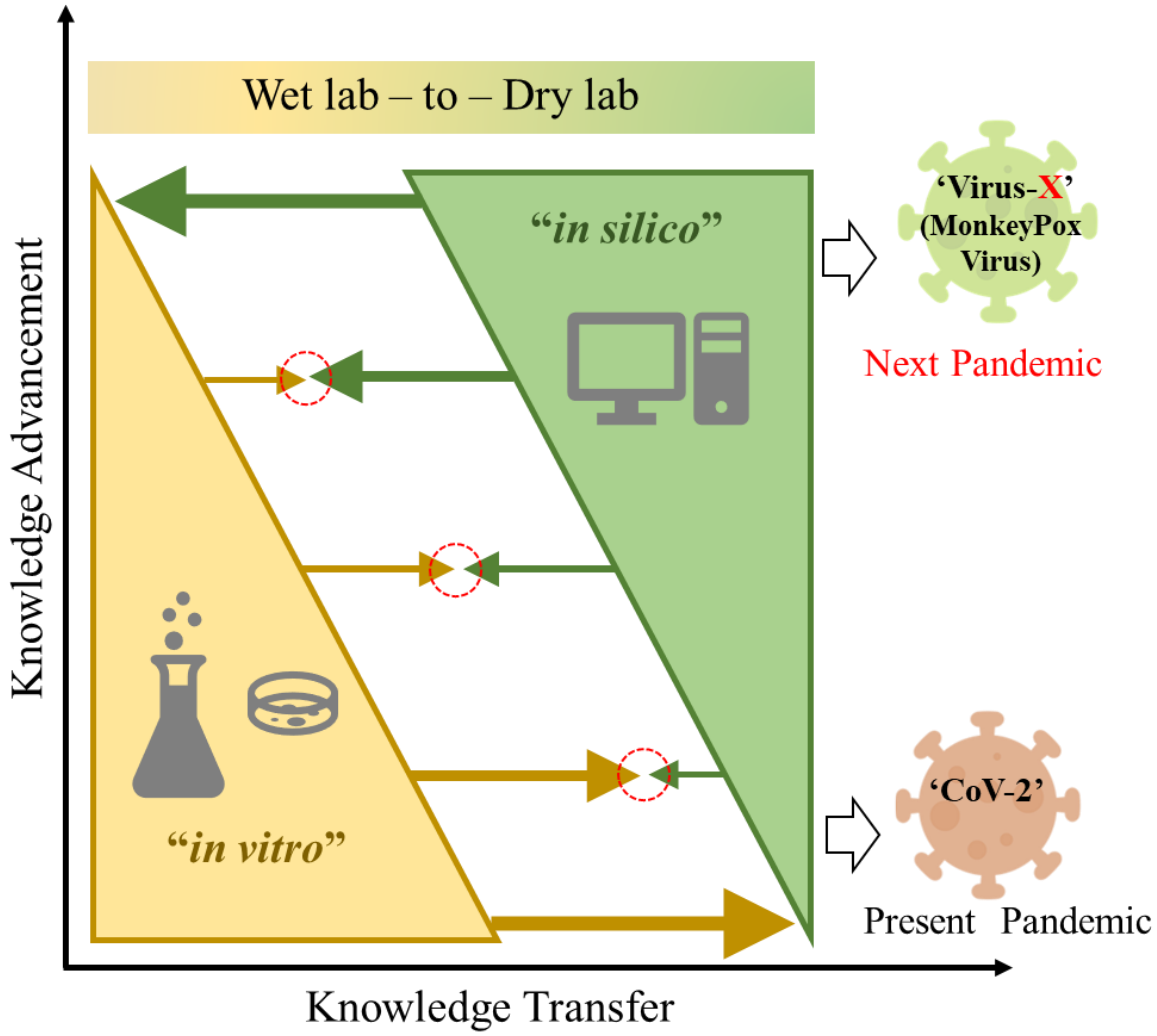
5.2 FUTURE PROSPECTS: THE 'DRY-LAB' CONCEPT

From this thesis, we can thus draw a general understanding of how integrating the in vitro and in silico studies can help in a better understanding of the biological systems. In the present times, we are more inclined towards the 'wet-lab', i.e. in vitro experiments for relying on the information that we acquire and 'dry-lab', i.e. in silico experiments are playing a supporting role. Gradually, we are shifting our focus towards making understandings from the in silico experiments. However, still the ratio of knowledge acquired by *in vitro* methods is higher as compared to *in silico* methods. 'Co-studying' and 'integrating' both the approaches can help us to develop a smart system for the future where (gradually) we can rely completely on the information gathered by *in silico* methods.

The basic idea is when everything around us is getting virtual in the present COVID-era, why can't laboratories get virtual? Considering, CoV-2 virus only, it took more than a year for the researchers to come up with a vaccine solution that eventually led the infection to become endemic and then a pandemic. Completely relying on the wet-lab is time consuming and it always will be, but utilizing the power of 'Artificial intelligence', we can gradually convert this time consuming process from years to months to days.

For instance, in my study, co-studying the conformers of Aβ42 using the *in vitro* and *in silico* approach helped gain an insight into the atomic levels and well as macromolecular levels which wouldn't have been possible by either one of them individually. To reach a stage where 'dry-lab' can be used to deliver the complete information individually, it is very important to first integrate both the systems that can gradually lead to knowledge advancement. If this can be achieved, we can prepare ourselves before we are hit by any other pandemic 'Virus-X'.

Box 2. 'The Dry-LAB' Concept'



This thesis gives a concept to develop a platform that can 'integrate' the power of in vitro and in silico systems to regulate the function of proteins with the help of structure and biomaterial-interaction. This integration can help to gradually reduce the dependency of in vitro system on the in silico system and develop a fully functional in silico platform for the future.