Development of aptamer probes against two wheat α-gliadin peptides containing DQ2.5 restricted celiac epitopes glia-α1a, glia-α1b, glia-α2 and glia-α3

A thesis

Submitted by

PAPORI BURAGOHAIN

For the award of the degree

of

Doctor of Philosophy

DEPARTMENT OF BIOSCIENCES & BIOENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI 781039, ASSAM, INDIA

NOVEMBER 2019

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
Department of Biosciences and Bioengineering
Guwahati-781039

DECLARATION

I do hereby declare that the matter embodied in this thesis entitled ***"Development of aptamer probes against two wheat α-gliadin peptides containing DQ2.5 restricted celiac epitopes glia-α1a, glia-α1b, glia-α2 and glia-α3"*** is the result of investigations carried out by me under the supervision of **Prof. Utpal Bora** and **Prof. Ajaikumar B. Kunnumakkara** and is submitted to the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Assam, India for the award of degree of *Doctor of Philosophy in Biosciences and Bioengineering*. This work has not been submitted elsewhere for any degree or diploma in any institute or university to the best of my knowledge and belief.

In keeping with the general practice of reporting scientific observations, due acknowledgments have been made wherever the work described is based on the findings of other investigators and copyright permissions have been taken from respective publishers.

Papori Buragohain 22.11.19

November 2019

Papori Buragohain
(Roll No. 136106006)

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
Department of Biosciences and Bioengineering
Guwahati-781039

CERTIFICATE

This is to certify that the matter embodied in this thesis entitled ***"Development of aptamer probes against two wheat α-gliadin peptides containing DQ2.5 restricted celiac epitopes glia-α1a, glia-α1b, glia-α2 and glia-α3"*** is the result of research work carried out by **M/s. Papori Buragohain** (Roll No. 136106006) under my supervision, in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati-781039, Assam, India for the award of degree of Doctor of Philosophy in Biosciences and Bioengineering.

The work embodied in this thesis has not been submitted elsewhere for the award of any degree.

22/11/19

Prof. Utpal Bora
(Main Supervisor)

Prof. Ajaikumar B. Kunnumakkara
(Co-supervisor)

November 2019

Dedicated to my life's coach,

My beloved Maa, Papa, Abhijit,

And to all who has been a part of this journey……

I owe it all to you!

ACKNOWLEDGMENT

Writing this thesis has been very exciting and extremely fulfilling for me giving me an opportunity to reveal everyone about the span of more than half a decade that I spent in IIT Guwahati and Biosciences and Bioengineering Lab (BERL) which was nothing less than wonderful experiences. Few years ago when I started as a PhD student, it turned out to be life's experience taught by small molecules called "Aptamers". Words cannot really summarize the days spent in the lab with my fellow lab mates and friends, the joy when something worked for the first time, with hope for good results and the sadness and tiredness of each failed attempt. In this journey many people have contributed in different ways, personally and professionally. I cannot find a better way of thanking all, than to dedicate this thesis as a humble gesture of gratitude.

Thanks to all my well-wishers, friends and all those close to my heart, for tolerating my agonies and stupidity with your love and encouragement. Your endless love and support was the only thrust in helping me to reach this stage of life. I am glad to take this opportunity to thank you all as a token of love from my side.

I would like to start with my mentors Prof. Utpal Bora and Dr. Ajaikumar B. Kunnumakkara and convey my deepest sense of gratitude for all your guidance, suggestions, moral support and encouragement throughout these years. I would like to thank Prof. Bora for introducing me to this amazing field of aptamers. It has been an absolute pleasure to be a part of your dynamic BERl group. Your out-of-the-box ideas always drove us to think in a different way. I must acknowledge the freedom you always allowed to explore on my own and at the same time guiding me through the right path. Thank you Sir for all your valuable guidance and support for all these years.

Next I would like to thank my doctoral committee chairman Dr. Nitin Chaudhary and members Prof. Gautam Barua and Dr. Lalit M. Pandey for continuously evaluating my work, giving suggestions for improvement and encouraging me in times of low energy. Thank you all for your valuable advice and help me refine this work.

I would also extend my heartfelt thanks to the present and past Heads of the department of Biosciences and Bioengineering for providing the essential facilities and a favorable academic and research environment. I would also thank my departmental staff members, Nurul Da, Niranjan Da, Dipankar Da, Prarthana Ma'am, Binoy Da, Dipangkar Da, Pankaj Da, and all others for the immeasurable help in the time of need. I would also thank all the faculty members, and the entire family of the Dept. of Biosciences and Bioengineering for your help and assistance in some way or the other. I would also like to thank the Centre for Central Instruments Facility, IIT Guwahati, Dept. of Chemistry, IIT Guwahati and Centre for the Environment, IIT Guwahati for providing access to various instrumentation facilities without which the work would not have been possible. I would also acknowledge IIT Guwahati for providing all the in-campus facilities and a healthy environment for all these years. My heartfelt thanks to the whole IIT Guwahati fraternity for your help in some way or the other.

I am sincerely thankful to Ministry of Human Resource Development, Govt. of India, for the financial support in the form of fellowship and Dept. of Biotechnology, Govt. of India, for research grants to my mentor.

Influence of some persons remain incomparable to anything in your life, which have always encouraged me, inspired me and deep rooted the foundations of fulfilling me dreams. My school teacher Dwipayan Das Gupta and my undergraduate teacher Dr. Mostaque A Laskar finds special mention for their contribution in teaching me the lessons for leading a meaningful life and always inspiring and guiding me through the right way when I was lost. I am also indebted to my grad school teacher and mentor; rather I would say my

brother Dr. Debojit Borah for the guidance and support during my Masters and inculcating in me the interest for pursuing research. I owe you a lot.

My heartfelt thanks to the whole BERL family, past and present, to be always with me whenever I needed you the most. I am sincerely thankful to the past lab members Dr. Arghya, Dr. Suradip, Dr. Manav, Dr. Manoj, Dr. Disco, Dr. Sunita, Dr. Deepika, Dr. Sambhavi, Swagata Ba, Krishna, Jyotismita Bow, Jyotirekha, Deepjyoti, Shamim and Dr. Birru. My sincere thanks to all my present lab mates Jonjyoti, Hasna, Vimal, Kabiraj, Dharitri, Biju, Adhiraj, Manash, Tinka, Pulak, Dr. Jintu, Nayan and Dr. Pragya Ma'am for providing the best and enjoyable environment both inside and outside the lab. Birru you are missed a lot. You guys were always the supporting bricks when I was about to fall and pulled me back into track through your love and encouragement. Thank you all for being a beautiful family. You guys are really hard workers. Wish you all the best of luck for your life ahead. You guys always rock.

I would also express my sincere thanks to all the members of the Cancer Biology Lab and Neurospora Group for helping me in times of need and sharing this journey with me. Thank you Dr. Nandu, Dibakar, Devivasha, Amrita, Kishore, Kishan, Harsha, Darshana, Abhishek , Ajit and all of you for your unconditional help and support. Thanks a lot.

My life outside the lab cannot be complete without my friends and extended family, Pallavi, Pushpita, Sumitha, Garima Di, Dooradarshi, Sou, and Papi. We have always shared the best moments in my life. My life would have been incomplete without you. You guys were the only support system throughout my PhD. Your love and endless support always soothed me when I was in hyper states. I have experienced all the shades of love in our friendship. You guys are my lifeline. You deserve the best in life. Wish you all the best. Always be with me like this forever, because undoubtedly I always need you all near to my heart.

Finally, it's all yours my super rocking family. The anchor and the real coaches of my life. Papa, Ma, my better half Abhijit, my sister Pori and the entire family of my in-laws, thank you all for your endless love, support and unwavering faith in me. Without you I would not have been the person I am today. You all are the biggest of all my strength. Ma and Papa I cannot thank you enough for the opportunities that you have given me and always been there for me whatever be the reason. I am fortunate to have parents like you. You have always given me and my sister all the best things in our life. Abhijit, thanks is only a six letter word for the infinite love and support you have showered on me. Love you to the infinity and beyond. I am always grateful to the love of my life, my beloved grandmother (Aita) and grandfather (Koka), for they have always wished to see me graduate. They would have been the happiest and proudest of all. I couldn't explain how lucky I am to have inherited some of your values in life through my parents. I love you the most my love, Aita and Koka. Really, I am missing you both.

Last but not the least I thank that power who have always helped me overcome all my problems in life. I don't know from where it comes, but it will surely, always come, whenever I need. Thank you for putting the smile back always, in its place.

"The greatest religion is to be true to your own nature and have faith in yourself."- Swami Vivekananda.

Papori
April 2019

CONTENTS

Chapter 3

3. Selection and characterization of specific DNA aptamers against the α-gliadin peptide PFPQPQLPYPQPQLPY

Synopsis

Celiac disease or a gluten sensitive enteropathy is a chronic inflammatory syndrome of the small intestine caused by the ingestion of gluten present in cereals like wheat, barley, rye and oats in susceptible individuals with positive HLA-DQ2/8 haplotypes. This HLA association plays the central role in $CD4^+$ T cell-mediated celiac disease pathogenesis. The major toxicity factors implicated in celiac disease is mainly attributed to the immunodominant epitopes and peptides present in prolamins of wheat. The only current available treatment for celiac disease is strict avoidance of gluten and its food products. This necessitates the need for strict monitoring of gluten levels in gluten containing foodstuffs and gluten-free products, ensuring food and health safety for celiac sufferers. Currently, antibody based immunoassays are considered as the gold standard for gluten monitoring. But antibody based methods suffer from some inherent disadvantages mostly attributed to high cost of production of antibodies, stability issues and batch to batch variations. Therefore, substitute to conventional antibodies could be addressed by next generation high affinity molecules called aptamers. Nucleic acid aptamers are highly stable and cheap alternatives of antibodies with no batch to batch variations and comparable binding affinity and specificity. Nucleic acid aptamers are short, oligonucleotide sequences that can be generated towards a range of target molecules (Gopinath, 2006; Ilgu & Nilsen-Hamilton, 2016; Wu, Chen, Wu & Zhao, 2015). Moreover aptamers are chemically synthesized with the added advantage of easy scale-up in short time and ease of chemical modification to further fabricate

them into cheap biosensing device. Hence, they can be utilized as a sensitive and cost effective detection platform for gluten assessment.

The present investigation is centered on the development of DNA aptamer probes against two native non-deamidated peptides PFPQPQLPYPQPQLPY (P16) and PQPFRPQQPYPQSQPQY (P17) of wheat α-gliadin containing four DQ2.5 restricted epitopes glia-α1a, glia-α1b, glia-α2 and glia-α3 implicated in celiac disease. For selecting specific aptamers against the two α-gliadin peptides P16 and P17, peptides were synthesized manually in the laboratory by Fmoc solid-phase peptide synthesis (SPPS), purified by reverse-phase high-performance liquid chromatography (RP-HPLC) and confirmed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF). Circular Dichroism spectroscopic analysis was also done to study the secondary structures of the peptides. *In vitro* selection of DNA aptamers against the peptides P16 and P17 was done using Systematic Evolution of Ligands by EXponential enrichment (SELEX). The final enriched pool from the last round of SELEX was cloned into TA cloning vector and transformed. Positive clones were selected by blue/white screening and confirmed by PCR using specific primers. The positive clones were sequenced by Sanger's method to identify the potential aptamer candidates. The sequencing results revealed two highly enriched candidates AG1 and AG2 against the peptide P16, while four enriched candidates AG5, AG75, AG110 and AG122 against the peptide P17. Once the aptamer sequences were obtained their secondary structure was predicted by M-fold web server and QGRS mapper softwares. Mfold studies revealed the lowest δG values for AG1, AG75 and AG122 with values -11.87 kcal/mol, -14.76 kcal/mol and -11.76 kcal/mol, respectively. QGRS mapper predicted the capability of forming G-

quadruplex structures for the aptamers AG1, AG75 and AG122. These features of the aptamers show their capability to form stable structures as compared to other screened candidates. The *in silico* structural studies of the enriched aptamers were confirmed by circular dichroism (CD) studies. CD spectra analyses of the aptamers revealed the correctly hybridized structures of the aptamers as deduced by Mfold.

On the basis of this, preliminary binding analyses between the aptamers and their respective targets were conducted by CD. CD binding studies revealed an increase in ellipticity of the CD spectra of the aptamers AG1, AG2 and AG75 upon target binding. This change in ellipticity is due to the induced CD (ICD) spectra of the aptamer-target complex confirming the binding of aptamers with their target peptides. The aptamer AG122 did not show any marked change in ellipticity of its CD spectra upon target binding. Further binding analysis by isothermal titration calorimetry (ITC) confirmed the high binding affinity of the aptamers AG1 towards peptide P16 and AG75 and AG122 towards P17 peptide. ITC studies for the aptamer AG2 didn't show significant binding affinity and hence not considered further in the study. Also, a specificity study was conducted for the aptamer AG1 with the related peptide P17 which did not show any significant heat change in the ITC isotherm, and hence it was considered highly specific towards its target P16. The value of the dissociation constants calculated from the ITC isotherms were 1.54 µM for AG1, 3.7 µM for AG75 and 10.9 µM for AG122. Finally, the binding affinity study based on electrophoretic mobility shift assay (EMSA) were also done for the aptamers AG1 and AG75, which further confirmed the successful formation of the binding complex between the respective aptamers and their target peptides.

Hence, there was a successful generation of potential and sensitive aptamer candidates against the targeted wheat α-gliadin peptides. Altogether two potential anti-gliadin DNA aptamer candidates were generated; AG1 aptamer specific against the gliadin peptide P16 and another aptamer AG75 against the gliadin peptide P17. Thus, these novel anti-gliadin aptamers with high binding affinity towards their targets are potential candidates as an alternative biorecognition probe for the development of highly sensitive detection of gluten.

List of Figures

List of Tables

Introduction

In recent years autoimmune diseases have become important medical problems worldwide. Autoimmunity arises because of the hypersensitivity of the immune system resulting in the production of antibodies against our own body, failing to discriminate between self and non-self-entities (Patris, Vandeput & Kauffmann, 2016). The celiac (or coeliac) disease is one such disorders of the immune system having autoimmune phenomenon. The disease was first reported in the 1st and 2nd centuries AD by a Greek physician Aretaeus. However, Samuel Gee gave the first modern scientific description of the disease in 1888 (Losowsky, 2008). Celiac disease or gluten sensitivity is a chronic inflammatory syndrome of the small intestine caused by the ingestion of gluten and other prolamins present in related cereals like barley, rye and oats in individuals with a genetic predisposition towards HLA and/or non-HLA genes. This is a multifactorial disease having systemic manifestations with gastrointestinal or non-gastrointestinal symptoms or no symptoms in affecting individuals of all age groups (Parzanese *et al.*, 2017; Tye-Din, Galipeau & Agardh, 2018). This gluten sensitivity is characterized by destruction of the small intestinal mucosa and an inability to digest prolamin containing foods and cereals. The only available cure for this disease is to avoid gluten and adhering to a lifelong strict gluten-free diet (Scherf & Poms, 2016).

Currently, celiac disease is considered to be a global problem with an increasing prevalence of 0.7 - 2% worldwide including the combined estimates of both the diagnosed and undiagnosed population (Rewers, 2005). In Europe and the USA, the statistical probability estimates that celiac disease affects around 0.5% -

1.26% of the total population (Green & Jabri, 2003). Amongst the affected population, only 10%–15% or fewer have been diagnosed and treated (Kagnoff, 2007). Therefore, it is considered as one of the most under-diagnosed diseases and is said to have iceberg type epidemiological characteristics (Fasano & Catassi, 2001). According to a very recent study conducted in the Indian population, celiac disease affects 1% of the north-Indian population, which is attributed mainly to the presence of wheat consuming population (Ramakrishna *et al.*, 2016). This data suggests that India is also having a major population of celiac sufferers and hence need to be taken care off. Thus, gluten detection has become very important in the current scenario for ensuring gluten-free assessment in the food industries and assuring health safety for all. Current conventions given by Codex Alimentarius (Commission for International Food Standards founded by the Food and Agriculture Organization and World Health Organization) allow gluten-free labelling of manufactured foods only when the content of gluten in it does not exceed 20 ppm (or 20 mg gluten/kg) (García *et al.*, 2005). This parameter has set a minimum standard to be maintained for the detection of trace amounts of gluten guaranteeing the safety of gluten sensitive patients. Existing standard methods of gluten detection include mostly ELISA (Enzyme-linked immunosorbent assay) based techniques that employ antibodies raised against the toxic and immunogenic fragments of prolamins which are then used in food screening assays as capture or recognition probe against gluten (Mitea *et al.*, 2008). However, ELISA based immunoassays suffer from many drawbacks, like the antibodies used in the immunoassays are proteins and sensitive to degradation under reducing agents and disaggregating chemicals (like β-mercaptoethanol and guanidine chloride) which are employed in the gluten

extraction process, making it difficult to recognize gluten in both processed and un-processed gluten-containing foods. Also, the production of antibodies is relatively expensive, time-consuming, requiring the sacrifice of living animals, and suffers from batch to batch variation. A novel intervention called “chemical antibodies” or nucleic-acid aptamers, being a type of non-protein receptor offer added advantage over protein antibodies in terms of stability, reliability and off-course sensitivity.

Aptamers are a unique class of chemically synthesized recognition elements that can be selected and developed against various targets (small molecules, peptides, whole cells, micro-organisms etc.). They are single-stranded nucleic acids (DNA or RNA) or peptides that adopt unique three-dimensional structure upon binding to their targets (Sun & Zu, 2015). They can be the ideal choice for specific detection of gluten due to various reasons: a) cost-effective, b) affinity comparable to monoclonal antibodies with no batch to batch variation, c) flexibility for a wide variety of chemical modifications for use in diverse analytical platforms (Nimjee, Rusconi & Sullenger, 2005). In the context of gluten detection, aptamers come as a novel class of molecules offering potential properties which can be explored extensively.

Considering the sensitivity of aptamers and the need for a cost-effective alternative for antibodies the present investigation is focused on the development of specific DNA aptamer probes against two native, non-deamidated wheat alpha-gliadin peptides using SELEX (Systematic Evolution of Ligands by Exponential enrichment) strategy. This investigation also characterizes and elucidates the interaction between the target and the aptamer.

Based on the above idea, the thesis is divided into the following chapters.

Chapter 1: Review of Literature

The objective of this chapter is to give a systematic literature review on the present status and progress of gluten detection, keeping in view the up-to-date advances in the quantification methods and the challenges faced in its detection. This chapter briefs about gluten and its toxicity factors, describing in detail about celiac disease, its pathogenesis and epidemiology. The chapter also focuses on nucleic acid aptamers, its characteristic features and SELEX (Systematic Evolution of Ligands by Exponential enrichment) technology with special mention in its advancements. This section also highlights the potential advantages of using aptamers as an alternative recognition element in analytical assays. Finally this chapter briefs about the advances in aptamer-based detection of gluten.

Chapter 2: Synthesis and characterization of two wheat α-gliadin peptides PFPQPQLPYPQPQLPY and PQPFRPQQPYPQSQPQY containing celiac disease epitopes

This chapter describes the synthesis, purification and characterization of two α-gliadin peptides of Wheat-containing epitopes which are implicated in celiac disease. These immunodominant epitopes within the gluten protein have the ability to provoke an inflammatory enteric reaction resulting in an abnormal immune

response in genetically predisposed persons. In the present work, two native peptides of wheat α-gliadin: (i) PFPQPQLPYPQPQLPY (P16) (residues 61-76), and (ii) PQPFRPQQPYPQSQPQY (P17) (residues 86-102) were selected as targets for gluten detection, constituting the DQ2.5 restricted toxic and immunogenic epitopes glia-α1a (glia-α9), glia-α1b (glia-αIII), glia-α2 and glia-α3 (glia-α20). These peptides were manually synthesized in the laboratory by Fmoc solid phase peptide synthesis (SPPS) using Wang resin. The purification of the synthesized peptides were done by reverse-phase high-performance liquid chromatography (RP-HPLC) and mass characterization done by matrix-assisted laser/desorption ionization (MALDI) mass spectrometry using MALDI TOF/TOF mass analyzer. Circular dichroism spectra were also recorded to ascertain the integrity and secondary structural conformations in the peptides.

Chapter 3: Selection and characterization of specific DNA aptamers against the α-gliadin peptide PFPQPQLPYPQPQLPY

This chapter describes the *in vitro* selection of specific DNA aptamers against the α-gliadin peptide PFPQPQLPYPQPQLPY (P16) (residues 61-76) containing three overlapping DQ2.5 restricted toxic epitopes glia-α1a, glia-α1b, and glia-α2 through magnetic bead-based SELEX process. A total of 16 iterative rounds of SELEX were conducted to enrich the aptamer pool. The enriched aptamers from the last round of SELEX was cloned and sequenced. Sequencing results revealed two highly enriched sequences AG1 and AG2. The sequences were synthesized and considered for further analysis. *In silico* structure predictions of the enriched aptamer

candidates were done my M-fold and QGRS mapper software to revealing diverse structures of the aptamers. *In silico* structure predictions were further validated by circular dichroism studies. Binding affinity studies were done by Isothermal Titration Calorimetry (ITC), circular dichroism studies and electrophoretic mobility shift assay (EMSA). The dissociation constant of the candidate aptamer was (K_d value) calculated from ITC studies.

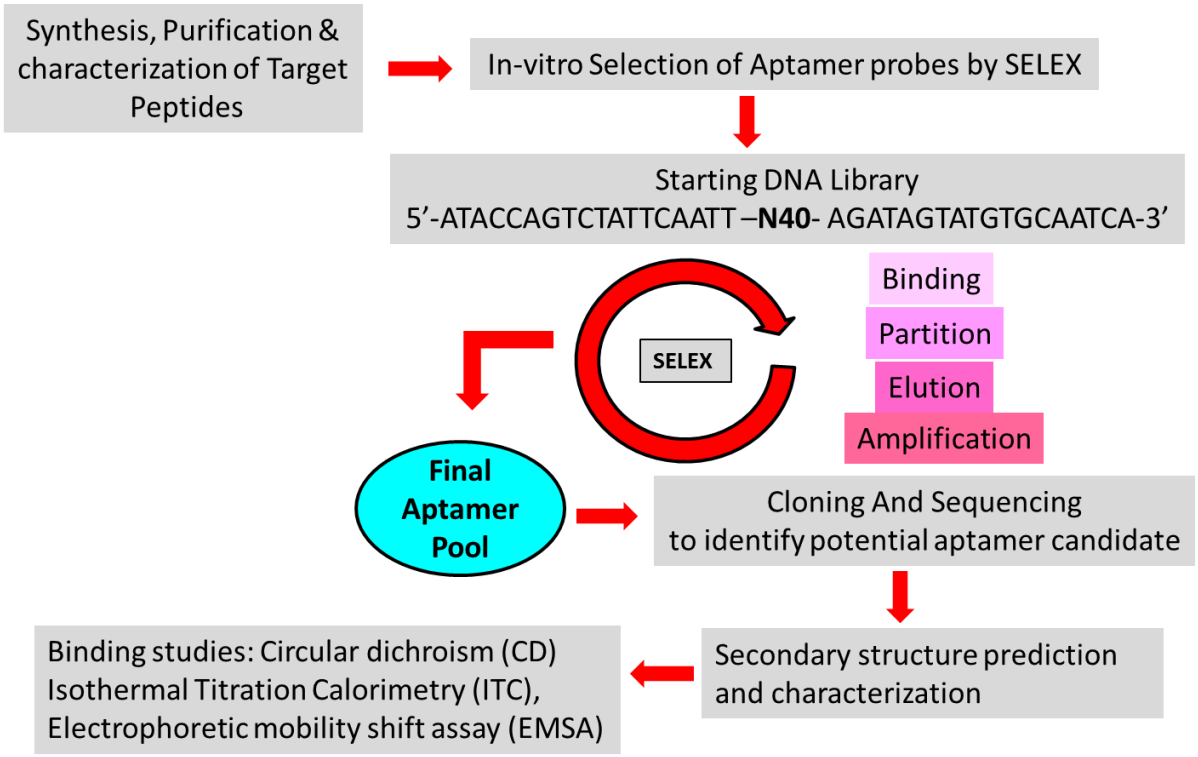
Chapter 4: **Selection and characterization of DNA aptamers against the α-gliadin peptide PQPFRPQQPYPQSQPQY**

This chapter describes the *in vitro* selection of DNA aptamers against another α-gliadin peptide, PQPFRPQQPYPQSQPQY (P17) (residues 86-102) with an immunogenic epitope DQ2.5-glia-α3 (glia-α20). A total of 14 iterative cycles of SELEX were conducted to fully enrich the aptamer pool. Four enriched aptamer sequences AG5, AG75, AG110 and AG122 were obtained after cloning and sequencing candidates from the last round of selection. *In silico* secondary structure analyses of the enriched aptamer candidates were done by M-fold and QGRS mapper softwares and further confirmed by circular dichroism studies. Binding affinity studies were done by Isothermal Titration Calorimetry (ITC) studies, circular dichroism studies and electrophoretic mobility shift assay (EMSA) studies. Subsequently, the dissociation constant (K_d value) was calculated from ITC studies.

Chapter 5: Summary and Future Prospects

This chapter discusses the summary and future direction of the study. Overall, the present study has led to the successful development of one specific high affinity aptamer probe against the wheat α-gliadin peptide PFPQPQLPYPQPQLPY (P16) and another aptamer against the peptide PQPFRPQQPYPQSQPQY (P17) containing celiac disease epitopes for highly sensitive detection of gluten. This investigation entails a detailed study from the generation of aptamers, their characterization and the diverse binding patterns involved with their respective target peptides. Binding studies revealed strong binding affinity of developed aptamers towards its target, confirmed by their dissociation constants. This study validates the practical feasibility of using these anti-gliadin aptamers in the gluten-free assessment of cereals and foodstuffs ensuring safety for celiac patients. Thus, the present investigation has the scope for further studies in understanding aptamer-gluten interactions and can help in the fabrication of highly sensitive and low-cost aptamer-based biosensing platforms for the detection of gluten.

THE WORK OUTLINE



1

Review of Literature

1.1 Celiac disease: Epidemiology, Aetiology and Pathogenesis

In the current scenario, Wheat is one of the most dominant food crops in temperate countries. It is considered as an important food source for humans and is grown in a major land area. Statistics of the FAO (Food and Agriculture Organization) state that the annual global production of wheat in 2013 was 715 million tonnes (Ozuna *et al.*, 2015b) (https://faostat3.fao.org/2013). Globally several billion individuals depend on wheat cereal and its products, making it the largest primary commodity. At the same time, there is much speculation about the ever increasing incidences of gluten intolerance and gluten-related clinical disorders with the increase in consumption of gluten-containing food grains (wheat, barley, rye and oats) and its foodstuffs. Gluten is the basic constituent of diet within the western population, but then with the shift in food habits in the eastern population towards western culture witness an increasing occurrence of gluten-related conditions among the eastern population as well. In this situation, celiac disease, and other non-celiac disorders have emerged as a worldwide phenomenon. Amongst these celiac disease accounts for ~1% (1 in every 100 individuals) of the worldwide human population with a large number of them remaining undiagnosed and untreated (Elli *et al.*, 2015; Sapone *et al.*, 2011). Celiac disease toxicity arising from gluten is attributed to the presence of toxic and immunogenic epitopes, which are T-cell stimulatory peptides with 9-mer amino acid core region derived from wheat prolamins, HMW-glutenins, LMW-glutenins, hordeins in barley, secalins in rye and avenins in oats (Sollid, Qiao, Anderson, Gianfrani & Koning, 2012). Fig 1.1 illustrates the classification of cereals based on gluten-related toxicity.

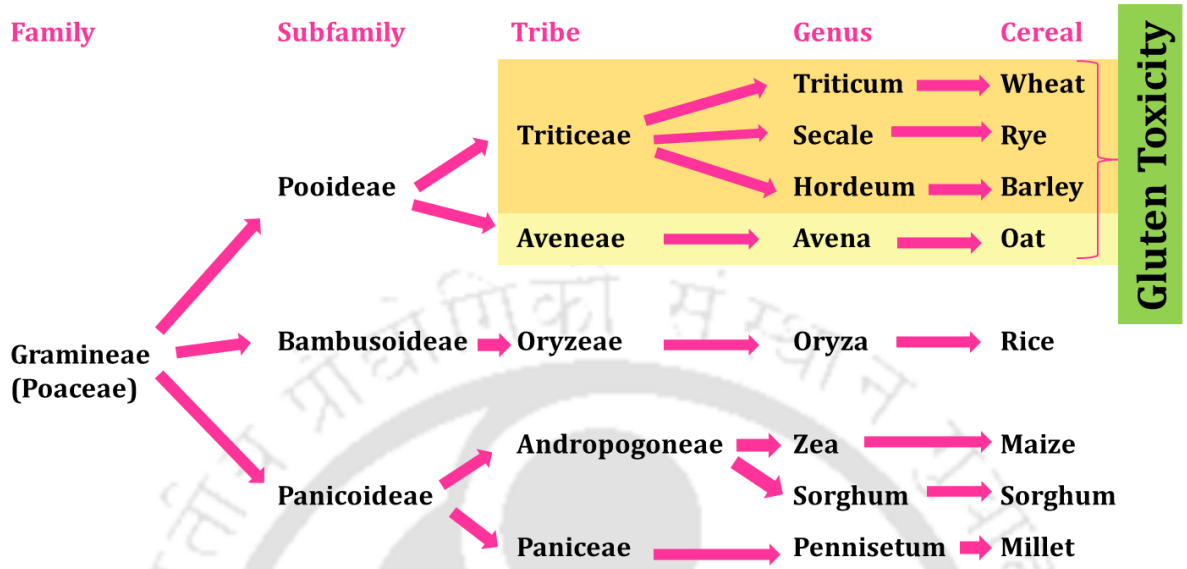


Fig. 1.1 Taxonomic classification of cereals, illustrating the cereals containing celiac disease toxic epitopes (modified from Simón *et al.*, 2017)

1.1.1 Epidemiology

Till a few years back celiac disease was believed to be a disease of the western population with Europe being historically determined as a place with a high prevalence of celiac sufferers. But until recently, global population studies have estimated that the disease is not only common to people of European or western origin but also common in other parts of the world, including the Middle East, North West and East Africa, Southern America and South Asia (Cataldo & Montalto, 2007). Epidemiological studies based on serological tests in general population showed a prevalence of 0.5%-1% in the Western countries (Green and Jabri, 2003; Rubio–Tapia *et al.*, 2009; Tack *et al.*, 2010; Scanlon and Murray, 2011). In the United States, the prevalence of celiac disease is seen to be increasing suddenly from

1:133 in patients not at risk to 1:56 in the case of symptomatic patients (McAllister, Williams & Clarke, 2018). Globally, the prevalence of this disease is increasing at an alarming rate. A recent global study based on systematic review and meta-analysis found the disease prevalence to be reported from across the globe having a prevalence of 1.4% based on serological test results and 0.7% based on biopsy tests (Singh *et al.*, 2018). Celiac disease is now known to affect people of every age group with higher prevalence in women vs men and higher prevalence among children than adults (Singh *et al.*, 2018). It is considered as the most common food intolerances in the world (Simón *et al.*, 2017). Also, global population-based studies in both general and risk group individuals of the developing countries showed a higher prevalence rate similar to those in the Western population. Celiac disease incidence is seen to have a direct relationship between the wheat-consuming population having increased frequency of HLA-DQ haplotypes and the disease prevalence (J. Y. Kang, Kang, Green, Gwee & Ho, 2013; Lionetti, Gatti, Pulvirenti & Catassi, 2015). This celiac risk profile among the population is also seen to be influenced by the geographic region and ethnicity. Like in Europe, the disease is more prevalent in Finland and Sweden than in Germany and Italy. Also in the USA the disease is more prevalent among the Caucasians than that among the African Americans (Lebwohl, Sanders & Green, 2017).

Interestingly, a recent serological study from India, showed the presence of celiac autoantibodies which are highly prevalent among the North Indian population (1.23%) as compared to the North Eastern (0.87%) and South Indian (0.10%) population (Ramakrishna *et al.*, 2016). The prevalence of celiac disease in the North

Indian populations (mainly children), was recognized since 1960s with estimates of about 0.3% - 1% of cases reported from all over India. This suggested the presence of symptomatic celiac disease in the North Indian states with a prevalence of 1% (Ramakrishna *et al.*, 2016). This prevalence in India is seen to be based on dietary preference and ethnicity with the highest number from the North India followed by the North Eastern India and least prevalent in Southern India. It was also observed that the population from these three regions in India had similar predisposition through HLA-DQ factors but due to change of food habits, the disease showed a different trend in its prevalence. However, mostly celiac disease can be asymptomatic, with patients not diagnosed or presenting atypical symptoms. A screening study in North India recognized positive tissue transglutaminase antibodies in 1.4 % of the population and confirmed celiac disease only in 1 %, suggesting around 5 to 8 million population in India might have undiagnosed and silent celiac disease (Freeman, Chopra, Clandinin & Thomson, 2011; Murch, 2016). Therefore, the epidemiology of CD is said to have iceberg type of characteristic, with more number of undiagnosed or silent cases of the disease than the diagnosed ones (Fasano & Catassi, 2001; Ramakrishna *et al.*, 2016). Fig.1.2 illustrates the iceberg characteristic of typical celiac disease prevalence.

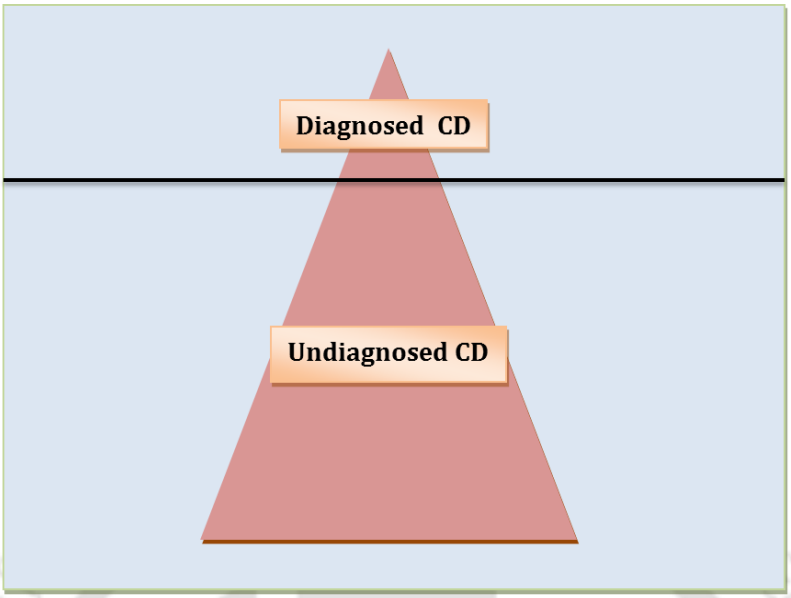


Fig. 1.2 Iceberg Model of celiac disease (modified from Fasano and Catassi, 2001)

1.1.2 Aetiology and Pathogenesis

Celiac disease or celiac sprue is a gluten mediated enteropathy which was earlier assumed to be a pediatric disease is now better agreed to be a multifaceted condition of the immune system having systemic manifestations (Jason Allan Tye-Din, Galipeau & Agardh, 2018). Celiac disease represents a jigsaw of the immune system that is attributed to the genetic factors containing the HLA haplotypes, HLA-DQ2 and/or 8 (Lionetti *et al.*, 2014). HLA system is a gene complex encoding the MHC (major histocompatibility complex) class II molecules that presents foreign antigen-derived peptides triggering a T-cell mediated adaptive immune reaction. These HLA genotypes along with other non-HLA genetic variations are said to precipitate in coeliac disease (Tack *et al.*, 2010). The interaction of gluten-peptides towards HLA-DQ haplotypes is more strongly assisted by the process of deamidation in specific Gln (glutamine) to Glu (glutamic) acid residues by tissue

transglutaminase, which renders the peptides a negative charge making it to bind more strongly with the HLA-DQ molecules. But it is less clear why only a small proportion of the individuals with high-risk genetic factors and exposure to gluten actually develop the disease (Shamir *et al.*, 2014). This fact makes it clear that apart from the genetic factors, other factors that disturb the immune regulatory mechanism are also crucial which cannot be ignored in the context of disease development and its increasing prevalence (Meresse, Malamut & Cerf-Bensussan, 2012). These factors may include a range of environmental elements like rotavirus infection, milk-feeding patterns during the first year of life and other host-microbiome interactions (Parzanese *et al.*, 2017).

Celiac disease pathogenesis involves both the innate and adaptive immune responses towards dietary gluten. The innate mechanism targets the epithelium directly and the adaptive branch involves $CD4^+$ T-cells by secretion of autoantibodies (Slot, Van Der Fels-Klerx, Bremer & Hamer, 2016). The primary mechanism of pathogenesis is mainly a series of events comprising of the following stages: i) direct innate immune response of the intestinal epithelium against the ingested toxic gluten peptides, ii) deamidation (Q to E modification) of the peptides by tissue transglutaminase enzyme, iii) modified peptides binding to HLA-DQ2/8 with subsequent stimulation of CD4+ T-cells, iv) release of cytokines and interferon-γ (INF-γ) causing mucosal damage (Sollid & Jabri, 2005). Fig. 1.3 illustrates the immune mechanisms involved in celiac disease pathogenesis. This establishes the dominant role of HLA-DQ2/8 in celiac disease pathogenesis. The innate immune system mediated response, which is not yet well understood, is attributed to a gliadin

peptide (amino acid 31-43) which triggers the release of interleukins (IL-15) that effects the expression of NKG2D receptor on intraepithelial lymphocytes (IEL) and MICA on the enterocytes causing damage to the epithelial cells, a hallmark of celiac disease (Koning, Gilissen & Wijmenga, 2005; Parzanese *et al.*, 2017; Shamir *et al.*, 2014).

The only safe and currently available treatment for celiac disease is totally avoiding gluten-containing diet and sticking to a lifelong strict gluten-free food, which is proved to be helpful in celiac sufferers, helping in the healing of the intestinal mucosa aiding in absorption and reversal of the disease (Abadie, Sollid, Barreiro & Jabri, 2011; Arentz-Hansen *et al.*, 2000; Jason Allan Tye-Din *et al.*, 2018).

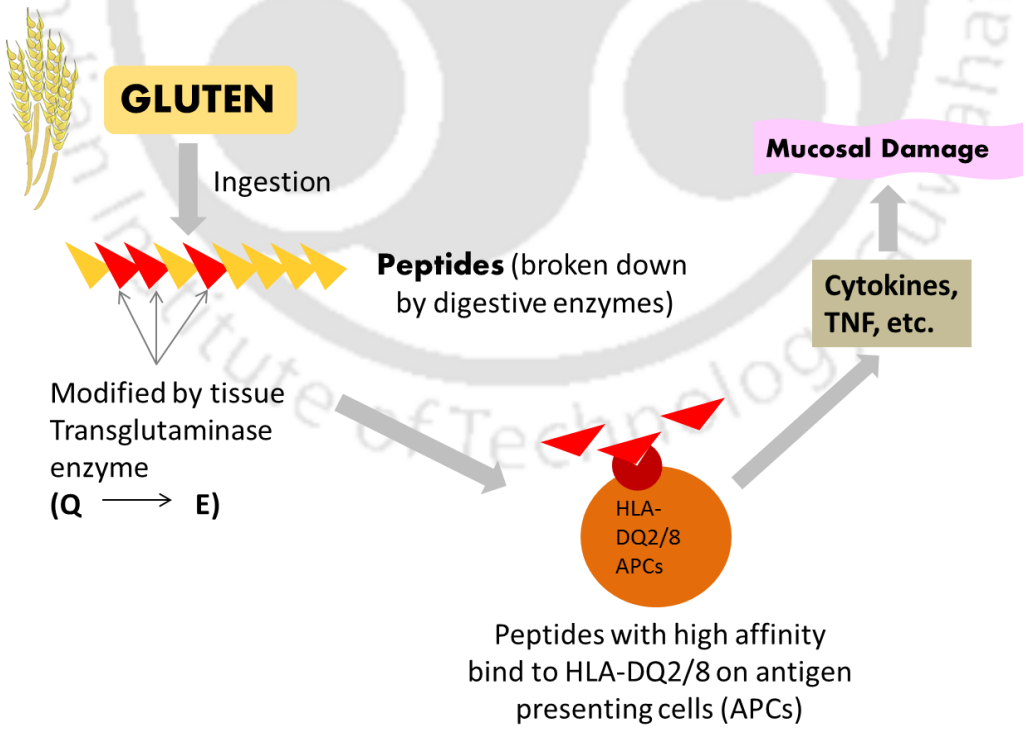


Fig. 1.3 Diagrammatic illustration of the adaptive immune mechanisms involved in celiac disease pathogenesis (adapted from Horvath & Mehta, 2000)

1.2 Gluten: the main causative agent of celiac disease

Gluten is the Latin word meaning “glue”. It gives bread dough its fluffy and gluey texture. It is present in extensive range of products ranging from packaged foods to medicines and supplements. Gluten was probably introduced into our diet thousands of years back particularly within a human population. But the time period seems to be short for the species to adapt towards the processing of this complex protein mixture, consequently developing intolerance towards the undigested fragments produced in the intestine resulting in the destruction of the villi and causing celiac disease (Miranda-Castro, de-los-Santos-Álvarez, Miranda-Ordieres & Lobo-Castañón, 2016). With the increasing prevalence of celiac disease, the incidence of other gluten-related conditions like wheat allergy and non-celiac gluten sensitivity (NCGS) are also increasing with the increase in gluten consumers all over the world (Czaja-Bulsa, 2015; Lebenthal, Shteyer & Branski, 2008; Murch, 2016). This scenario necessitates a need for sensitive gluten detection platform ensuring health security among celiac patients.

According to Osborne fractions (1907) cereal proteins are categorized into four main groups: albumins (water-soluble) , globulins (salt-soluble), prolamins (60-70% alcohol-soluble) and glutelins (dilute acid or base soluble) based on their extraction and solubility (Osborne, 1907). Fig 1.4 depicts the classical Osborne classification of wheat proteins. Gluten is an umbrella term for the storage proteins: prolamins and glutelins present in wheat (gliadins and glutenins), rye (secalins) and

barley (hordeins). Wheat Gluten is the complex mixture of gliadins and glutenins in a ratio of 1:1 which accounts for 80-85% of cereal proteins (Barak, Mudgil & Khatkar, 2015; Simón *et al.*, 2017). The viscoelastic qualities of wheat dough are said to depend on the ratio of these two fractions. Gliadin maintains the extensibility and cohesive properties whereas glutenin on the other side decides the elasticity and strength. Gluten proteins have common structural characteristics with unique amino acid composition mostly rich in proline and glutamine with less number of charged amino-acids. Hence, the term prolamin and glutelin are used. The presence of proline is advantageous for a storage protein because of its ability to form β-turns which is more efficient in packaging proteins in a lesser space than α-turns. This compact and tight structure in these proteins hinders its complete hydrolysis by peptidases forming undigested immunogenic and toxic peptides in the small intestine which is the major-trigger implicated in celiac disease pathogenesis (Kagnoff, 2007; Simón *et al.*, 2017).

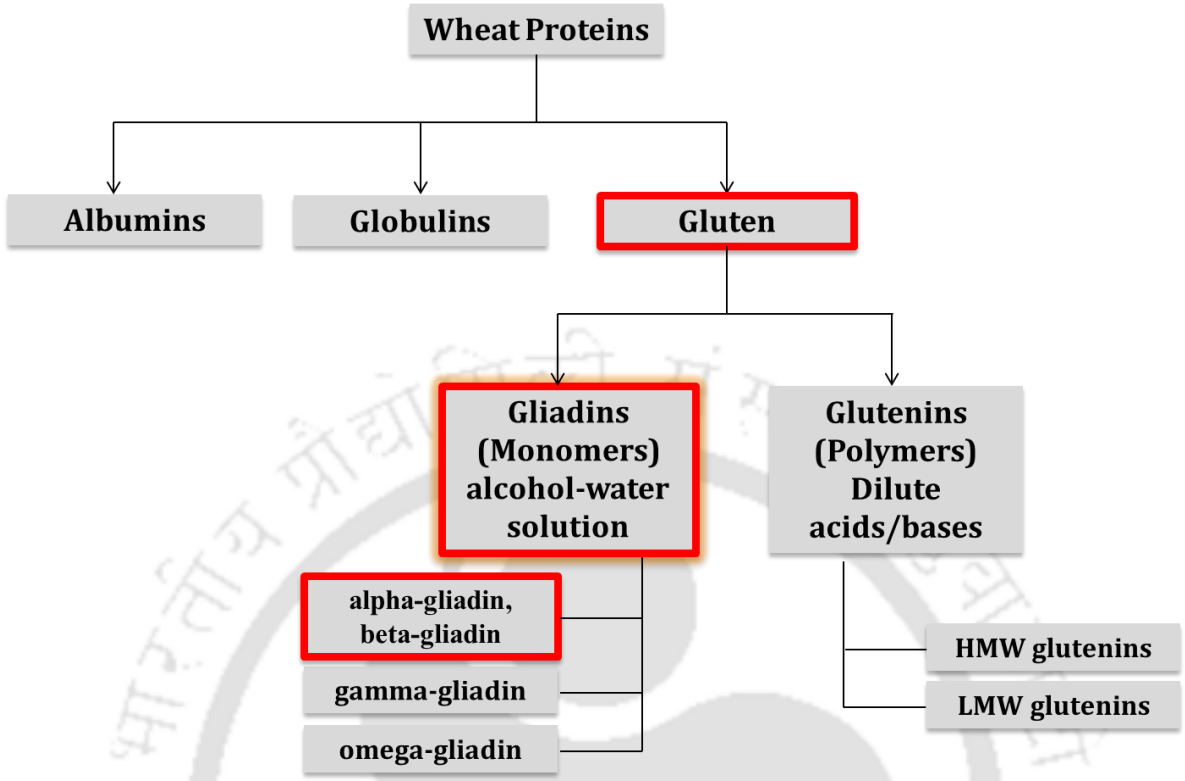


Fig. 1.4 Osborne classification of wheat cereal proteins (adapted from Simón *et al.*, 2017)

1.2.1 Gliadins: Structure and toxicity factors

The major gluten storing seed proteins in wheat are gliadins which account for half (40-50%) of the total wheat seed proteins with an important role in nutritional and flour processing quality (Urade, Sato & Sugiyama, 2018). Its composition differs in different varieties depending upon the type of cereals. Gliadins of wheat and the homologs present in rye, barley and oats are the main targets of gluten detection. Gliadins are a heterogeneous mixture of polypeptides divided into four groups depending on their electrophoretic mobility (fastest to slowest: α > β > γ > ω) in Acid-Polyacrylamide gel electrophoresis (A-PAGE) (Balakireva & Zamyatnin, 2016; G

Chirdo & Arranz, 2015; Shewry, Tatham, Forde, Kreis & Miflin, 1986). Gliadins are mainly monomers consisting of many polypeptide chains linked by hydrogen bonds and hydrophobic interactions plus intramolecular disulfide bonds (α/β- and γ-gliadin) or no di-sulfide bonds (ω-gliadin) with molecular weights ranging from 30-78 kDa (Qi, *et al.*, 2006). Based on structural homology, α and β-gliadins were grouped together as α-type gliadins. Gliadins are grouped as S-rich (α/β-type and γ-type gliadins) and S-poor (ω-type gliadins) based on the presence or absence of cysteine residues (Barak *et al.*, 2015; Shewry *et al.*, 1986). All the types of gliadins have distinctive secondary structures with α/β-type and γ-type gliadins containing both α-helices (30-35%) and β-sheets (10-20%) allowing them to be stabilized by hydrogen bonds and hydrophobic interactions plus di-sulfide linkages forming additional compact structures. While ω-gliadins form randomly coiled β-turns only without forming overall any compact structure (Barak *et al.*, 2015).

The major toxicity towards celiac disease is mostly attributed to wheat gliadins depending on their amino acid sequence and molecular properties. For this, many epitopes and peptide regions are proposed on the basis of their immunogenicity and toxicity. Despite the number of immunogenic peptides, a few of them from the gliadin fraction of Wheat appear to be highly toxic. Thus, gliadin fraction is considered as the principal toxic component of wheat gluten in celiac sufferers (Shan *et al.*, 2005; Simón *et al.*, 2017). The peptides derived from α/β and ω-gliadins cause immune reactions in adults and peptides from low–molecular weight (LMW) glutenins and γ-gliadins cause toxicity mostly in children and sometimes in adults (Simón *et*

al., 2017). The following paragraphs describe the different types of wheat gliadins based on their molecular and structural characteristics.

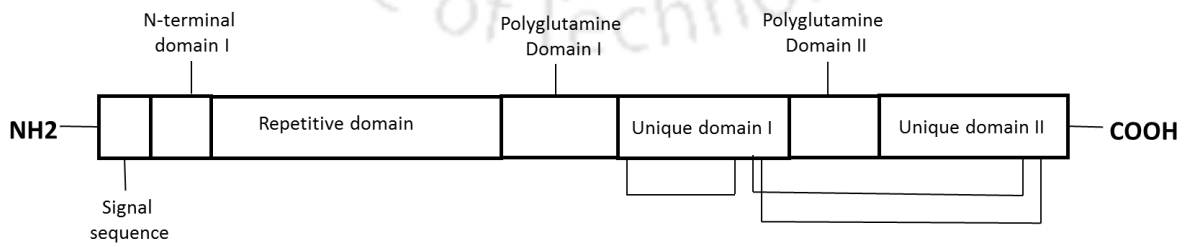
The primary structure of α-type gliadin proteins consist of 250-300 amino acid residues composed of 5-14 residues in N-terminal region followed by a N-terminal repetitive domain consisting of around 134 residues mostly rich in proline and glutamine. There is a small N-terminal signal sequence of ~20 amino acids which is removed from the mature protein. The repetitive domain is followed by the first polyglutamine stretch, the unique domain I, second polyglutamine stretch and the unique domain II (Qi *et al.*, 2006). The main stretch of variability lies between the two polyglutamine domains with the most immunodominant 33-residue peptide located in the N-terminal repetitive domain. The C-terminal non-repetitive domain is about 150 amino acid residues. These C-terminal region unique regions are the place of 6 cysteine residues forming three intramolecular disulfide linkages (Barak *et al.*, 2015; Ozuna *et al.*, 2015a; Wieser, 2007). Fig. 1.5, A represents the typical primary structure of α-type gliadins.

The γ-type gliadins contain a 20 amino-acid signal peptide followed by 12 residue non-repetitive N-terminal region, a 72-161 residue highly variable repetitive region, again followed by a non-repetitive domain containing mainly cysteine, a Gln rich stretch and finally the C-terminal non-repetitive region with two conserved cysteine residues (Qi *et al.*, 2006). The α/β and γ-type gliadins contain a higher number of cysteine than ω-type with the possibility for the earlier two to interact more with the glutenin fraction forming intermolecular disulfide linkage. The α- and γ-

gliadin repetitive region is rich in β-reverse turns forming an stretched structure while the non-repetitive region forms α-helix (Barak *et al.*, 2015; Wieser, 2007). Fig. 1.5, B represents the typical primary structure of γ-gliadins.

The distribution of different forms of gliadins depends largely on the wheat variety (genotype) but it can be generalized that α and γ-type gliadins account for a larger percentage of gliadin compared to ω-gliadins. The ω-type gliadins are categorized in the molecular weight range of 44-80 kDa with the presence of 80% of the proline, glutamine and phenylalanine with no sulfur-containing amino acids (Hsia & Anderson, 2001; Wieser, 2007). Hence unable to form disulfide bonds and lack a compact structure. Also, the ω-type gliadins are more polar than the α- and γ-type with lesser surface hydrophobicity. The structure of ω-gliadins is simple with a 19 amino-acid residue signal peptide followed by 10-11 amino-acid N-terminal region, a repetitive domain covering approximately 90% of the protein and finally a short C-terminal region (Qi *et al.*, 2006). Fig 1.5, C represents the typical primary structure of ω-gliadins.

A



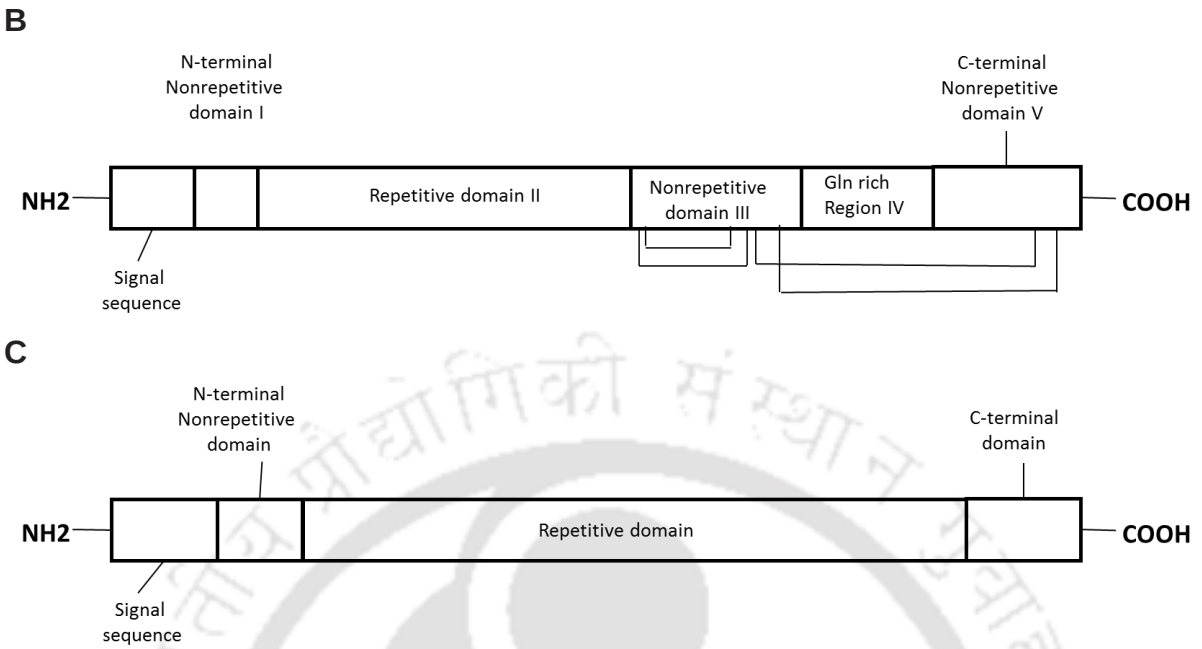


Fig. 1.5 Graphical representation of the molecular structure of (A) α/β-type gliadin, (B) γ-type gliadin, and (C) ω-type gliadins (adapted from Qi *et al.*, 2006)

Amongst all three types of gliadins, peptides corresponding to α-gliadin fraction have been widely studied for their toxicity towards celiac disease. From the available studies of several groups it is well-known that α-gliadin fraction contributes the most immunodominant and toxic epitopes with proline-rich repetitive domains in gliadins mostly responsible for holding these toxic epitopes (Anderson, Degano, Godkin, Jewell & Hill, 2000; Bergseng, Sidney, Sette & Sollid, 2008; Ciccocioppo, Di Sabatino & Corazza, 2005; Fraser *et al.*, 2003; Koning *et al.*, 2005; Molberg *et al.*, 2005; Sollid *et al.*, 2012; Vader *et al.*, 2002). Studies conducted by few groups have listed these immunodominant and stimulatory peptide regions in α-gliadin comprising of glutamine and proline-rich epitopes which are the 9-residue epitopic core region binding to HLA-DQ 2 and HLA-DQ 8 on antigen presenting cells (APCs) which has been confirmed both *in vivo* and *in vitro* T-cell stimulatory immune reactions against

them (Arentz-Hansen *et al.*, 2000; Arentz–Hansen *et al.*, 2002; Qiao *et al.*, 2005; Sjöström *et al.*, 1998; Tye-Din *et al.*, 2010; Vader *et al.*, 2003). Table 1.1 lists the HLA-DQ 2.5 restricted celiac disease relevant gliadin epitopes. The underlined glutamic acid (E) residues are modified by tissue transglutaminase.

Table 1.1 List of T-cell stimulating gliadin peptides relevant in celiac disease (modified from Sollid *et al.*, 2012)

Protein	HLA-DQ 2.5 molecule	Epitopes	References
α-Gliadin	glia-α1a	PFPQPE<u>E</u>LPY	(Arentz-Hansen *et al.*, 2000)
α-Gliadin	glia-α1b	PYPQPE<u>E</u>LPY	(Arentz-Hansen *et al.*, 2002)
α- Gliadin	glia-α2	PQPE<u>E</u>LPYPQ	(Arentz-Hansen *et al.,* 2000)
α- Gliadin	glia-α3	FRPE<u>E</u>QPYPQ	(Vader *et al.,* 2003)
γ-Gliadin	glia-γ1	PQQSFPE<u>E</u>QQ	(Sjostrom *et al.,* 1998)
γ-Gliadin	glia-γ2	IQPE<u>E</u>QPAQL	(Vader *et al.,* 2003)
γ-Gliadin	glia-γ3	QQPE<u>E</u>QPYPQ	(Arentz-Hansen *et al.,* 2002)
γ-Gliadin	glia-γ4a	SQPE<u>E</u>QE<u>E</u>FPQ	(Arentz-Hansen *et al.,* 2002)
γ-Gliadin	glia-γ4b	PQPE<u>E</u>QE<u>E</u>FPQ	(Qiao *et al.,* 2005)
γ-Gliadin	glia-γ4c	QQPE<u>E</u>QPFPQ	(Arentz-Hansen *et al,.* 2002)
γ-Gliadin	glia-γ4d	PQPE<u>E</u>QPFCQ	(Qiao *et al.,* 2005)
γ-Gliadin	glia-γ5	QQPFPE<u>E</u>QPQ	(Arentz-Hansen *et al.,* 2002)
ω-Gliadin	glia-ω1	PFPQPE<u>E</u>QPF	(Tye-Din *et al.,* 2010)
ω-Gliadin	glia-ω2	PQPE<u>E</u>QPFPW	(Tye-Din *et al.,* 2010)

1.3 Techniques for detection of Gluten

Gluten detection is a matter of concern in the current scenario because of its relation to celiac disease and other gluten sensitivities. In the past years, there was an increase in the prevalence of celiac disease cases form over the world with consumers preferring gluten-free products safer. Moreover, gluten-free diet is the only treatment available for celiac sufferers to prevent the disease. So there is an obvious market demand for gluten-free products. But adherence to a strict gluten-free diet is difficult because apart from food sources, gluten is also found in a range of products in a modified form including drugs, supplements, cosmetics etc. (Scherf & Poms, 2016). So, the most important question arises that the tag mark "gluten-free" written in the products is how much really gluten-free because those products can be cross-contaminated with gluten during manufacture. So such claims should be regulated with a sensitive detection method. Moreover, Codex Alimentarius an international body concerned to develop coordinated international food standards has defined a gluten-free product to be gluten-free only when the amount of gluten in it doesn't exceed 20mg/kg in total, and gluten-low when the content of gluten is between 20-100 mg/kg in total (Miranda-Castro *et al.*, 2016; Scherf & Poms, 2016). These regulations further necessitate the need for a standard analytical gluten detection platform. Complying with the legal regulations, the quantitative determination of gluten is currently done by ELISA (enzyme-linked immunosorbent assay) based immunological methods using antibodies, which is the currently accepted method of choice for gluten detection. But, many alternative methods like

proteomics, molecular techniques, mass spectrometry, genomics and other novel methods like aptamers are also being explored.

1.3.1 Immunological techniques

1.3.1.1 ELISA (Enzyme-Linked Immunosorbent Assay)

Immunological methods are based on antibodies (immunoglobulins) which are raised against a particular antigen. In this case, it can be different prolamin fractions of cereal proteins or specific toxic and immunogenic peptides found in them. Production of antibodies requires immunization of animals like mice, rabbits or goats with the specific antigen and later sacrificing the animal for extracting the developed antibody. Both polyclonal (pAb) and monoclonal antibodies (mAb) have been screened against known immunogenic and toxic epitopes of gluten (Chirdo, Añón & Fossati, 1995; Ellis, Rosen-Bronson, O'Reilly & Ciclitira, 1998; Freedman, Galfre, Gal, Ellis & Ciclitira, 1987; Friis, 1988; McKillop, Gosling, Stevens & Fottrell, 1985; Redondo *et al.*, 2005; Troncone *et al.*, 1986). Two principles of ELISAs namely sandwich and competitive ELISA is the recommended method of choice. In the sandwich format the antigen is sandwiched between the capture and the detection antibody and so the antigen requires at least two antigenic epitopes for binding to both the immunoglobulins. Therefore, this format of ELISA suffers from the disadvantage that it can only be used to detect gluten in large intact proteins or peptides with two epitopic regions and making it unsuitable for hydrolyzed gluten proteins which may have only one epitopic region (Scherf & Poms, 2016). The

competitive ELISA format uses only one antibody which requires only one antigenic epitope helpful in the quantification of small peptides and intact proteins along with hydrolyzed gluten fractions. In the competitive format, the standard antigen is immobilized on the surface and there is a competition between the sample antigen and the standard antigen for the antigen binding site in the antibody. In this format, the intensity of the chromogenic reaction is inversely proportional to the amount of antigen in the sample (Simón *et al.*, 2017).

ELISA kits are mostly developed against the wheat gliadin or its sub-fractions (α, γ and ω), but due to structural homology among the prolamins, they also recognize other prolamins like hordeins and secalins to the same extent. This results in inaccurate measurement of signal towards the presence of a particular antigen. Out of the several antibodies, the majority of the ELISA test kits are based on monoclonal antibodies developed by Skerritt & Hill (1990) (401.21 also called as Skerritt mAb) against the heat stable ω-gliadins; R5 mAb developed against pentapeptide QQPFP by Valdés, García, Llorente & Méndez (2003); the G12 and A1 mAb against the 33-residue α-gliadin peptide by Morón *et al.* (2008); and α20 mAb developed by (Mitea *et al.*, 2008).

In the 1990s, Skerritt and Hill developed a gluten detection sandwich ELISA method against the heat stable ω-gliadins which recognizes the epitopes PQPQPFPQE and PQQPPFPEE (Skerritt & Hill, 1990). The 401.21 mAb (Skerritt mAb) also recognizes the corresponding prolamins from rye and barley along with the high molecular weight (HMW) glutenin subunits and presumably low molecular

weight (LMW) glutenins (Haraszi, Chassaigne, Maquet & Ulberth, 2011). This method was approved by the AOACI (Association of Official Analytical Chemists International) and was used as the standard method for many years. The main disadvantages of this method is that it suffers from low sensitivity (160mg gluten/kg), also it reacts weakly to hordein proteins leading to an underestimation of the total hordein content in barley. Also due to the large variability among the ω-gliadins in different cereal species, this method leads to miscalculation of the total gliadins in the immunological result (Mena & Sousa, 2015; Slot *et al.*, 2016; Wieser, 2008).

The R5 mAb involves both the sandwich and competitive ELISA formats developed against a ω-secalin extract from rye recognizing the epitope QQPFP and its related sequences LQPFP, QQQFP, QLPFP (Méndez, Vela, Immer & Janssen, 2005; Sorell *et al.*, 1998; Valdés *et al.*, 2003). The sandwich R5 ELISA is very sensitive and specific towards gluten proteins and is used for the determination of the low concentration of antigens. It can recognize the gliadins, hordeins and secalins to the same extent but not able to recognize oat avenins (Simón *et al.*, 2017). The R5 competitive format was developed to estimate the small peptides in hydrolyzed gluten products and was recognized as the AACCI (American Association of Cereal Chemists International) approved method 38-55.01. The sandwich R5 ELISA method validated by collaborative studies was also adopted as AACCI approved method 38-50.01 (Koehler *et al.*, 2013; Mena & Sousa, 2015; Méndez *et al.*, 2005).

The G12 and A1 monoclonal antibodies were raised against the 33-mer gluten peptide from α-gliadin which is considered the most toxic peptide and the principal immunodominant factor responsible for gluten toxicity (Morón *et al.*, 2008). G12 was used as the detection antibody recognizing the epitope QPQLPY while the A1 mAb was used as the capture antibody reacting with the epitope QLPYPQP (Scherf & Poms, 2016; Slot *et al.*, 2016). These antibodies offer high selectivity towards the 33-mer in wheat and similar peptides in rye and barley with lower detection capacity for oat avenins. Both competitive and sandwich ELISA formats are available. The G12 sandwich format is in compatibility with reducing extraction solvents offering a limit of quantification (LOQ) of 4mg gluten/kg (Mena & Sousa, 2015; Scherf & Poms, 2016). A collaborative study approved the G12 sandwich assay as AACCI approved method 38-52.01 (Don, Halbmayr-Jech, Rogers & Koehler, 2014).

Other immunological ELISA methods include the PN3 antibody raised against the 19-mer synthetic α-gliadin peptide LGQQQPFPPQQPYPQPQPF which detects the epitope QQQPFP. This mAb detects gliadins, hordeins, secalins, avenins and low molecular weight (LMW) glutenins but it does not react with High molecular weight (HMW) glutenins (Haraszi *et al.*, 2011; Simón *et al.*, 2017). Another monoclonal antibody Glia-α20 was raised against the most immunogenic α-gliadin peptide PFRPQQPYPQP with the heptapeptide RPQQPYP as the minimal T-cell recognition epitope (Mitea *et al.*, 2008; Scherf & Poms, 2016). This ELISA method is reported to detect other homologous epitopes in barley, rye and oats (Slot *et al.*,

2016). Table 1.2 represents a few of the important commercially available ELISA based kits for detection of gluten with a comparison of their analytical efficiency.

Table 1.2 Commercially available ELISA and LFD (lateral flow device) based gluten detection kits (modified from Miranda-Castro *et al.*, 2016)

Antibody	Commercial name	Manufacturer	Design format	Limit of detection (LOD) in ppm
Skerritt	Biokits	Neogen	Sandwich	1
	Gluten Check	BioCheck	Sandwich	5
	EZ Gluten	ELISA technologies	LFD	10
R5	TransiaPlateProlamins	Bio Control Systems	Sandwich	3
	Veratox	Neogen	Sandwich	5
	RidaScreenGliadincompetitive	R-Biopharm	Competitive	5
	Rida Quick		LFD	5
	Ingezim gluten	Ingenasa	Sandwich	3
G12	Gluten Tox Sandwich	Biomedal	Sandwich	0.6
	Gluten Tox Competitive		Competitive	3
	Gluten Tox ELISA Sticks		LFD	3
	Agra Quant	RomerLab	Sandwich	2
	Agra-Strip		LFD	5

1.3.1.2 Other Immunological detection techniques

Lateral flow devices (LFDs) and dipstick are a type of immuno-chromatographic method that is used as a rapid qualitative or semi-quantitative indication of the presence or absence of the target. This format employs an antibody covered zone and a nanoparticle conjugated secondary antibody on the surface of a strip. When the liquid sample (containing gluten) is applied to the strip, the antibody conjugate and the gluten sample interacts and flows through the strip and comes in contact with the fixed antibody covered zone where color is formed indicating the presence of gluten in the sample. The LFDs formats are available based on G12, R5 and Skerritt monoclonal antibodies (Table 1.2) and other polyclonal antibody formats (Mena & Sousa, 2015; Scherf & Poms, 2016)

Immunosensors are analytical devices that use a biological element plus a physicochemical transducer. There are some biosensors using Resonance Enhanced Absorption (REA) or Surface Plasmon Resonance (SPR) and the Luminex xMAP Technology (Slot *et al.*, 2016). An immunosensor was raised against the immunodominant gliadin peptide which detects alpha-gliadins with a limit of detection (LOD) of 5.5 µg/L. Another type of biosensor developed was based on the adsorption of anti-gliadin Fab fragments on gold surfaces (Mena & Sousa, 2015). Apart from these, some multiplex immunoassays using multi-analyte profiling platform are also being developed combining various food allergens and gluten in a single stage providing simultaneous detection of multiple targets (C. Y. Cho, Nowatzke, Oliver & Garber, 2015; Slot *et al.*, 2016).

1.3.2 Non-immunological techniques

1.3.2.1 Proteomic-based approaches

The use of proteomic-based approaches in food analysis has become a significant tool for the characterization and quantification of gluten proteins and peptides. Due to the structural similarity and homology between amino acid sequences among different prolamins (e.g. gliadins, hordeins or secalins) the exact identification of celiac disease epitopes become difficult along with other shortcomings of the available methodologies. Proteomics complement the immunological techniques helping in further validation of the obtained results. In cereal protein analysis proteomics plays the role of a key technology which has high sensitivity for separation, identification, characterization and finally quantification. Proteomic-based approaches include mass spectrometry (MS) techniques like MALDI-TOF (matrix-assisted laser desorption ionization time-of-flight) and LC-MS (liquid-chromatography mass spectrometry) which quantitate different proteins and peptides based on the relation of mass/charge. One of the most important contributions of proteomics in gluten-related celiac toxicity is the identification of the immunogenic HLA-DQ epitopes involved in celiac disease (Alves, D'Almeida & Ferreira, 2017). MALDI-TOF MS was the first non-immunological technique to be explored to quantitate intact gluten proteins involved in celiac disease from a large quantity of wheat, barley, rye and oat cultivars which were later optimized for gluten analysis (Camafeita, Alfonso, Mothes & Méndez, 1997; Mena & Sousa, 2015). Another study by Hermando *et al*., 2003 shows the presence of a typical mass

fingerprint ranging from molecular weight 30-45 kDa corresponding to gliadin components in gluten-free foods like rice and maize using MALDI-TOF MS which also validated the R5 ELISA results (Hernando, Valdes & Méndez, 2003). But the lowest limit of detection of gliadin by this method was 100 ppm which is insufficient in terms of sensitivity regulations for gluten-free foods. The most important disadvantage of MALDI-TOF MS came due to the complexity of the protein mixture and loss of resolution after heat modification and hydrolysis of gliadins leading to ambiguous identification of gluten. So, MALDI-TOF MS is suitable only for the semi-quantitative determination of gluten proteins (Scherf & Poms, 2016). These limitations were overcome by tandem mass spectrometry detection technique LC-MS/MS. This technique combines primary fractionation by HPLC (high-performance liquid chromatography) followed by multi-enzymatic digestion of the protein resulting in different peptide and protein fractions whose mass measurements were done by high resolution MS or MS/MS and finally matching the results with the database (Alves *et al.*, 2017).

Despite the sensitivity of the technique, the analysis of gluten by LC-MS/MS requires expensive equipment and expertise which is available only in a few research laboratories. Also, this technique is used in complementation to ELISA for validation purpose. Currently, there is no LC-MS/MS capable enough to distinguish all the peptides of the gluten proteins, specifically glutenins (Mena & Sousa, 2015; Scherf & Poms, 2016). But still, it holds to be the most promising non-immunological method for precise quantification of trace amounts of gluten proteins in both native and hydrolyzed forms.

1.3.2.2 Genomic-based approaches

Genomic approaches are indirect methods of gluten detection targeting a marker (DNA or RNA) indicative of the presence of gluten rather than the gluten itself. Therefore, these methods mandatorily require an exact correlation between DNA and gluten concentrations, because legal regulations have established the limit of detection in terms of gluten concentration for quantification purposes. Still, these methods have gained importance and popularity because of the stability of DNA molecules. Similar to protein-based approaches, successful implementation of this method also depends on its sensitivity and specificity for analysis of gluten. Genomic approaches rely on the amplification of specific DNA fragments of the offending component by polymerase chain reaction (PCR) using specific oligonucleotide primers offering selectivity and amplification of the particular ingredient of interest. Then these PCR products can be analyzed by electrophoresis or southern blotting or DNA sequencing. Qualitative PCR method suffers from many disadvantages, therefore, quantitative-PCR or real time-PCR (Q-PCR/RT-PCR) is the preferred one over conventional PCR for gluten quantification in cereals and foodstuffs.

Q-PCR uses expensive laboratory equipment and chemicals but offers rapid and accurate quantification of DNA in real time (Haraszi et al., 2011). Sandberg et al. 2003 have developed a Q-PCR method for the estimation of oat products contaminated by gluten using melting curve analysis that gave good correlation with the Skerritt ELISA (Sandberg, Lundberg, Ferm & Malmheden Yman, 2003). Another Q-PCR method developed by Mujico *et al.* (2011) has shown to detect wheat DNA in

gluten-free foods. This method employs a fluorescent dye SYBR green I or TaqMan probe for rapid quantification. This assay is a highly specific and sensitive platform presenting a quantification limit of 20 picogram (pg) of DNA per mg of food. Comparing this method to commercially available R5 ELISA revealed that with the exception of some highly processed and hydrolyzed food products, the rest of the foods with prolamin levels above the R5 ELISA detection limit (1.5 mg/kg) gave positive results with this Q-PCR platform (Mujico, Lombardía, Mena, Méndez & Albar, 2011). However, these DNA based approaches are difficult to be implemented in case of processed and hydrolyzed food products because of large amount of DNA damage (Scherf & Poms, 2016).

New genomic approaches like next-generation sequencing (NGS) which doesn't directly quantify gluten but can be used to quantify between immunogenic and non-immunogenic celiac epitopes. Salenjitn and his team employed a dedicated 454 RNA-amplicon sequencing platform to determine the three major α-gliadin toxic epitopes and their variants in different durum wheat cultivars (Salentijn *et al.*, 2013). Another group of researchers studied and analyzed the entire α-gliadin fragment containing toxic celiac epitopes in diploid, tetraploid and hexaploid wheat varieties using NGS and Sanger sequencing. The study identified six types of α-gliadins (type 1-6) with only a single type containing all the immunogenic epitopes and peptides, and the other five types are devoid of the immunogenic epitopes (Ozuna *et al.*, 2015).

1.4 Main challenges in gluten detection

Several analytical methods of gluten detection are developed but it is clear that only a very few of the available methods are able to successfully detect gluten present in various foodstuffs and food matrices. Despite major efforts and technologies, there are a few important challenges faced in gluten detection that has to be overcome for sensitive and specific detection of gluten. At the same time, the method should be cost-effective so that it can be fabricated easily in any format. The successful quantification of gluten lies in the following main factors which include, firstly the efficient and complete extraction of gluten from both raw and processed food matrices plus its compatibility with the type of detection method, secondly the need of a standard reference protein and finally the accuracy given by sensitivity and specificity of the method (Miranda-Castro *et al.*, 2016).

The efficient and complete extraction of gluten from food matrices faces a big hurdle because of the complex nature and physicochemical properties of the gluten proteins. The solubility of the gluten proteins is largely hindered by their heterogeneous surface properties, inter and intra-chain disulfide bonds and other modifications like high-temperature treatment, hydrolysis, transamidation, deamidation, etc. which occurs during food processing. Some of these modifications like hydrolysis is useful in decreasing or abolishing the toxicity of gluten proteins and increasing the solubility, but fragmentation occurring as a result of hydrolysis complicates and makes the analysis of gluten more challenging (Mena & Sousa, 2015; Scherf & Poms, 2016). Likewise, deamidation can decrease the affinity of

antibodies towards the gluten proteins and peptides making immunoassays difficult for correct estimation of gluten. Also, the heat treatment of food products lead to the formation of protein aggregates which may further add difficulties to the extraction process (Mena & Sousa, 2015). The most frequently used extraction method uses aqueous alcohol (60% ethanol or 50% propanol) which helps in extracting mainly the prolamin fraction from raw gluten containing foods but not capable of extracting gluten out of processed foods because of protein aggregation. To solve these problems different extraction methods utilizing reducing agents like beta-mercaptoethanol are being used to break protein aggregations but they also interfere in the immunoassays making the detection problematic. So some other extraction solutions like the UPEX (universal prolamin and glutelin extraction solution) which uses Tris (2-carboxyethyl)-phosphine (TCEP) capable of disrupting disulfide linkages and N-Lauroyl sarcosine capable of disaggregating proteins proved helpful in extracting total gluten from various food matrices of processed and unprocessed gluten-containing foods (Simón *et al.*, 2017).

The other significant problem is the lack of a standard reference protein to be used for calibrating the methods for gluten detection. Till date, the well-characterized reference material is the gliadin standard called PWG-gliadin from the Prolamin Working Group (PWG) on Prolamin Analysis and Toxicity which is a highly purified form of gliadin, containing a mixture of 28 European wheat cultivars (Van Eckert et al., 2006). PWG gliadins were extracted with 60% ethanol post removal of the non-gluten proteins (albumins and globulins) by extraction using 0.4 M NaCl solution. The gliadin extracts were then concentrated, desalted by ultrafiltration, freeze-dried, and

homogenized, after which the left over material is called as PWG-gliadin (Wieser, 2008). But the standard can only be used as a reference against the prolamin fraction of gluten limiting its usage to other fractions like glutenin. However this gliadin standard is the only best-characterized reference protein and utilized as the correct reference material in many studies for validation purposes (Amaya-González, de-los-Santos-Álvarez, Miranda-Ordieres & Lobo-Castañón, 2014; Amaya-González, López-López, *et al.*, 2015; Miranda-Castro *et al.*, 2016; Svigelj *et al.*, 2018; Wieser, 2008).

The third challenge which needs to be addressed is the requirement for a sensitive and cost-effective method for the successful detection of gluten. Currently, the method of choice is the ELISA based commercial kits using the Skerritt, R5 and G12 antibodies which are available in two different assay formats, sandwich and competitive ELISA formats (Miranda-Castro *et al.*, 2016). However, there are few significant inadequacies of these techniques as mentioned in the earlier sections. Also the characterizations of different wheat, rye, barley and oat varieties together with the surplus diversity of gluten-containing food products in the market, antibody based detection and screening is not a feasible idea, as it will be very expensive. Also, detection with antibodies suffers from reproducibility issues as sometimes degradation of antibodies during transport and storage reduces its applicability. Also the specificity of monoclonal antibodies suffers from batch to batch variability challenging the uniformity of the detection technique. In this respect, aptamers are an outstanding alternative capable of replacing antibody-based technologies with their high affinity, specificity and selectivity. Moreover with the growing need to

develop a rapid and inexpensive method for gluten detection, aptamer technology can provide an ideal alternative towards designing robust bio-sensing platform.

1.5 Nucleic acid Aptamers and SELEX

Nucleic acid molecules have long been known to fold into characteristic three-dimensional structures (naturally occurring ribozymes) that can catalyze reactions and control gene expression (Doudna & Cech, 2002; Weigand & Suess, 2009). But it was only with the advent of the advanced technologies like solid-phase synthesis and polymerase chain reaction (PCR) which is capable of generating large population of degenerate oligonucleotides and amplify them, the synthetic functional nucleic acid molecules called “aptamers” have arrived (Dunn, Jimenez, & Chaput, 2017). Aptamers are a group of promising bio-recognition molecules capable of binding to various targets ranging from small molecules, peptides and proteins with high affinity and specificity (Amaya-González, López-López, *et al.*, 2015; E. J. Cho, Lee & Ellington, 2009; Vasilescu & Marty, 2016). The word aptamer is derived from the Latin word “aptus” meaning to fit (Jayasena, 1999). Nucleic acid aptamers are short, single-stranded DNA or RNA molecules capable of spontaneously folding into specific three-dimensional structures characterized by stems, loops, bulges, hairpins, triplexes, and/or quadruplexes (Gopinath, 2006). These 3-dimensional assemblies make them bind with a wide variety of targets driven by structure compatibility and non-covalent interactions (Regina Stoltenburg, Reinemann & Strehlitz, 2007). With respect to binding affinities, aptamers resemble and, in many applications, are used as an alternative to antibodies. Aptamers show greater chemical and thermal

stability, can easily be developed artificially *in vitro* in a simple microfuge tube through an evolutionary technique called SELEX (Systematic Evolution of Ligands by Exponential enrichment) (Irvine, Tuerk & Gold, 1991; Regina Stoltenburg *et al.*, 2007; Tuerk & Gold, 1990). SELEX generate aptamers with high binding affinities ranging from micromolar (µM) to picomolar (pM) range (depending upon the type of targets) virtually towards any type of target, which is very difficult when raising antibodies (Keefe, Pai & Ellington, 2010).

1.5.1 SELEX (Systematic Evolution of Ligands by Exponential enrichment)

For more than two decades since the discovery of SELEX method by three independent research groups of Joyce (La Jolla), Szostak (Boston) and Gold (Boulder) developed the *in vitro* selection method for generating catalytic or ligand-binding RNA molecules from RNA libraries and revolutionized the way we look at nucleic acid ligands (Ellington & Szostak, 1990; K. Kang & Lee, 2012; Robertson & Joyce, 1990; Tuerk & Gold, 1990). SELEX is an enrichment technique involving iterative cycles of *in vitro* selection where a chemically synthesized, random, combinatorial, single-stranded nucleic-acid library of defined length and variability (~10^{13}-10^{15} different sequence motifs) is challenged to bind its target under defined conditions and parameters (Klug & Famulok, 1994). The SELEX library is designed in such a manner that the variability of the SELEX library is determined by the length of the randomized region (varies mostly from 20-90 nucleotides in length). For example, if the length of the random region is 40 nucleotides then the randomness of the given library will be represented by 4^{40} possible combinations of

sequences which is equal to ~10^{24} different individual sequences. Each nucleic acid sequence in the library consists of a central random region and flanked by two primer binding sequences on both the sides available for PCR. Fig. 1.6 depicts a typical structure of an aptamer sequence.

Fig. 1.6 Diagrammatic illustration of a nucleic acid aptamer sequence showing a central random region flanked by two primer binding sites on 5' and 3' ends

In general, the SELEX process involves repeated rounds of binding, separation and amplification steps allowing the starting nucleic acid library to undergo selection with its cognate molecule. In the first binding step, the random aptamer library is incubated with the target, followed by separation of the non-binding molecules by washing steps and then eluting nucleic acids with higher binding affinity to the target of interest. After elution, the strong binders are amplified by PCR or RT-PCR. This amplified pool is then used for the next round of selection. This cycle of binding, partitioning and amplification continues till the highest affinity binders are obtained. Also during the repeated cycles, the selection pressure is increased (for e.g. by increasing washing steps, decreasing binding time, etc.) to generate aptamer molecules with remarkable binding affinity and specificity towards its target. Typically SELEX is carried out for 5-20 rounds of selection. At the last round of selection when the aptamer pool is enriched with strong binders, they are

cloned and sequenced for further analysis and characterization. Fig. 1.7 illustrates a typical SELEX method.

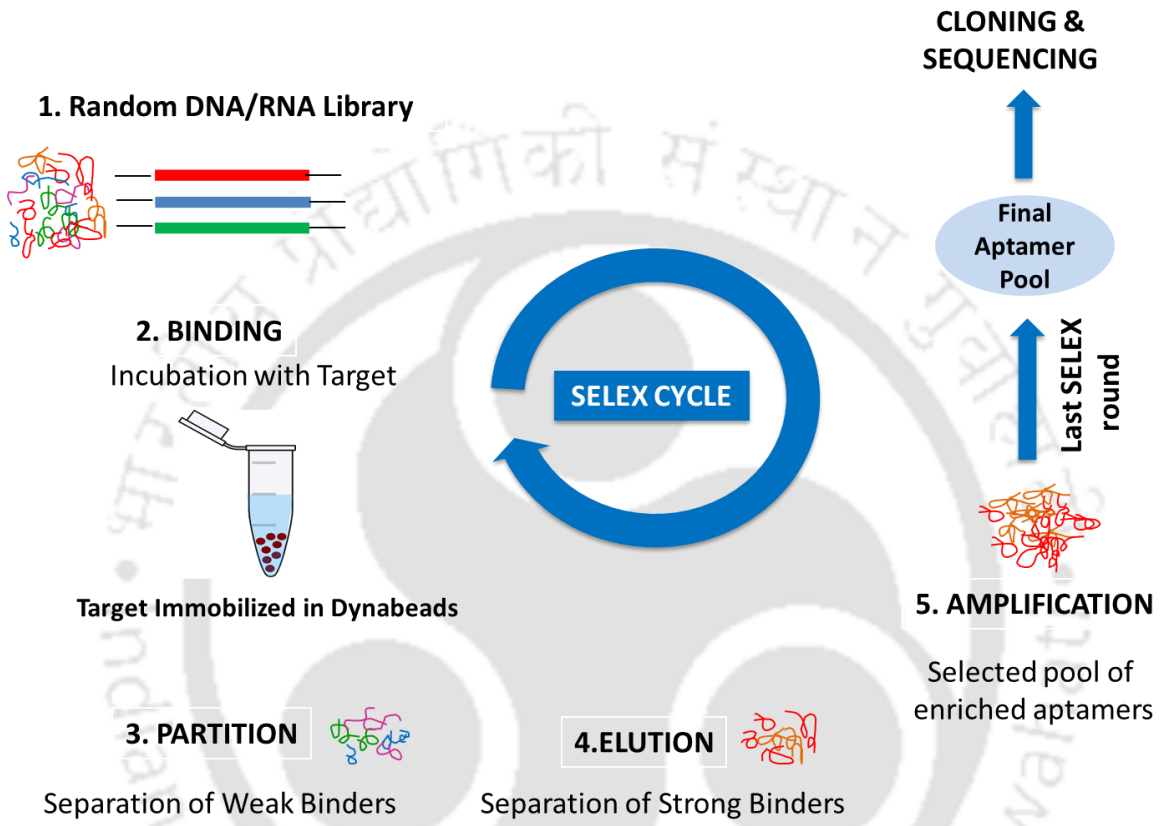


Fig. 1.7 Diagrammatic illustration of different steps involved in SELEX cycle

Ever since the discovery of the SELEX process in 1990, it has taken a giant leap in improving the technology with useful modifications, depending upon the target, desired features of the aptamers and the required applications. In fact, there is no any so-called standardized SELEX protocol for a particular target, rather a SELEX method has to be adapted depending upon the requirement. Since then innumerable variations of the SELEX process are being developed to enhance the properties of the aptamers and increase their efficiency so that they can be applied

in different fields like therapeutics, diagnostics, drug delivery, environmental monitoring and sensing applications. One of the classic examples of a successful aptamer in the market is Macugen, an anti-angiogenic drug for age-related macular degeneration (AMD) with US FDA (Food and Drug Administration) approval and available to patients (Maberley, 2005; Ruckman *et al.*, 1998). Another popular successful aptamer-based diagnostic kit is OTA sense developed by Neoventures Biotechnology Inc. for the detection of Ochratoxin A (OTA), a toxin produced by Aspergillus and Penicillium species of fungi (Penner, 2012). Table 1.3 describes different modifications of the SELEX method since the time of its first development.

Table 1.3 A list of modifications of the SELEX process (modified from Darmostuk *et al.* 2015; Stoltenburg *et al.*, 2007)

SELEX type (reference)	Description
Blended SELEX (Smith, Kirschenheuter, Charlton, Guidot & Repine, 1995)	Offering additional properties to aptamers beyond binding by addition of non-nucleic acid molecules, by covalently modifying the oligonucleotide library.
Photo SELEX (Golden, Collins, Willis & Koch, 2000; Jensen, Atkinson, Willis, Koch & Gold, 1995)	Aptamers bear light-sensitive molecules which on UV excitation bind to targets through covalent binding and increase the efficiency of aptamer binding to targets.

Chimeric SELEX (Burke and Willis, 1998)	Two different pools of oligonucleotide library used to obtain chimeric aptamers with more than one feature. Each of the parent libraries is first selected towards different targets and then the selected aptamers fused together to obtain the chimera.
Multi-Stage SELEX (Wu & Curran, 1999)	A special type of chimeric SELEX. In this method, after the fusion of the preselected aptamers, a reselection to the target is done.
Deconvolution SELEX (Morris, Jensen, Julin, Weil & Gold, 1998)	Used to generate aptamers for complex targets, and then separate the relevant aptamers binding to target structures from the non-binding oligonucleotides. This method is utilized in the identification of certain proteins of a complex target.
Speigelmer Technology (Klußmann, Nolte, Bald, Erdmann & Fürste, 1996)	Aptamers are first selected towards the enantiomeric form of the target. Only the aptamer candidates with L-nucleotides are synthesized to bind with the native form of the target.
Crossover SELEX (Hicke *et al.,* 2001)	Parallel selection of aptamers against both protein target and the cell bearing the protein target is carried out. Aptamers obtained from the cell-SELEX are reselected against the protein target to obtain high-affinity aptamers towards the target.
Toggle SELEX (van Eckert *et al.*, 2006)	Utilized to select aptamers with cross-reactivity towards two homologous targets by switching between targets in alternate rounds.

Indirect SELEX (Kawakami, Imanaka, Yokota & Sugimoto, 2000)	Generation of metal-ion dependent aptamers that bind to their targets only in the presence of a metal-ion.
Genomic SELEX (Gold *et al.*, 1997; Shimada, Fujita, Maeda & Ishihama, 2005; Singer, Shtatland, Brown & Gold, 1997)	cDNA library of an organism's genome is utilized as the SELEX library to bind the organism's own proteins or metabolites as target identifying significant interactions.
In-vivo SELEX (Coulter, Landree & Cooper, 1997)	In this method a transient transfection is done in an iterative mode in cultured vertebrate cells.
Tissue SELEX (Li *et al.*, 2009)	Aptamers are selected towards a complex tissue target such as collections of cells in diseased tissues useful in clinical specimen.
TECS SELEX (Ohuchi, Ohtsu & Nakamura, 2006)	A type of cell SELEX, where targets expressed on the cell surface is directly used to isolate aptamers.
FluMag SELEX (R Stoltenburg, Reinemann & Strehlitz, 2005)	Oligonucleotide library after the first round of selection is amplified with fluorescein modified primers offering an ease of quantifying the enrichment process with fluorescence based methods.
Capture SELEX (R Stoltenburg, Nikolaus & Strehlitz, 2012)	It's a type of FluMag SELEX where the oligonucleotide library itself is immobilized on magnetic beads instead of the target.

Capillary Electrophoresis SELEX (Mendonsa & Bowser, 2004)	This SELEX utilizes capillary electrophoresis for separation of bound and unbound aptamers. This method helps in the reduction of SELEX rounds.
On-chip selection (Asai, Nishimura, Aita & Takahashi, 2004)	Rapid and automated selection using microarray technology in combination with point mutations of selected sequences using a genetic algorithm
Truncation SELEX (Pagratis, Gold, Shtatland & Javornik, 2001)	Several variations of this method is available where the fixed regions of the oligonucleotide library are truncated or eliminated reducing the chances of the fixed sequences to participate in the selection process.
Expression Cassette SELEX (Martell, Nevins & Sullenger, 2002)	Used to develop vectors bearing aptamer sequences with the advantage of expression of the functional aptamers in mammalian cells.
Single microbead SELEX (Tok & Fischer, 2008)	This method utilizes a single microbead bound with the target for aptamer selection.
Sol-gel SELEX (Park *et al.*, 2009)	A specially fabricated micro-fluidic device with sol-gel arrays is used in target immobilization. This method is useful in the selection of multiple aptamers at the same time.
GO SELEX (WooáKim & BockáGu, 2012)	This method utilizes Graphene Oxide for adsorption of the target-oligonucleotide complex with ease of separation for the bound and unbound

	aptamers without immobilization of the target.

1.6 Advantages of aptamers over antibodies

Several advantages of aptamers over antibodies make them more preferred bio-recognition molecules as compared to antibodies. Aptamers behave in a way similar to antibodies in terms of binding to its target by folding into a three-dimensional structure in sequence dependent manner driven by non-covalent interaction (Gopinath, 2006). Aptamers bind to their targets with dissociation constants (Kd) in picomolar (pM) to nanomolar (nM) range for various protein molecules and in micromolar (µM) range for smaller target molecules comparable to their antibody counterparts (E. J. Cho *et al.*, 2009). Aptamer binding is highly specific interaction discriminating targets that share similar structural domains. Beyond these advantages, aptamers have presented extraordinary potential in analytical based applications because they can be easily modified or conjugated via linkers, reporter moieties or other functional groups (E. J. Cho *et al.*, 2009). Specifically, aptamers offer immense advantages to be used as a bio-sensing platform. One of the most important advantages is the regeneration capacity of the aptamer-based biosensors as compared to immunosensors which has a poor capability to regenerate their surface (Jayasena, 1999). As discussed, Table 1.4 elucidates a few significant and relevant advantages of aptamers over antibodies.

Table 1.4 Properties of aptamers versus antibodies (Keefe *et al.*, 2010; Nimjee, Rusconi & Sullenger, 2005; Nimjee, White, Becker & Sullenger, 2017; Rubio, Homs & O'Sullivan, 2004)

Antibodies	Aptamers
Selection is done *in vivo* inside a biological system, therefore difficult to raise against some targets like toxins or non-immunogenic targets.	Selection of aptamers is done *in vitro* with extreme accuracy and reproducibility, offering selection capability to different targets.
Antibodies have limited shelf life, sensitive to temperature and undergo irreversible denaturation.	Aptamers are robust, stable for longer period, can be transported in ambient conditions and easily regenerated if denatured.
Production of antibodies requires the sacrifice of living animals.	Aptamers are artificially synthesized *in vitro* without sacrificing animals.
Production of antibodies is difficult to scale up.	Production of aptamers can be effortlessly scaled up.
Immunogenic	Non- immunogenic
Production and screening is costly.	Cost-effective
Limited modification and conjugation resulting in reduced activity.	Easy for modification and conjugation enhancing its applicability.
Problem with isolation of cross-reactive compounds.	Cross-reactive molecules can be eliminated using toggle strategy.

1.7 Developments in aptamer technology for detection of gluten

In the recent years few research groups have explored the field of aptamers for detection of gluten and its toxic and immunogenic peptides. All targets are not that easy for aptamer generation and many of them are prone to failure. Gluten is one such target that is very hydrophobic and hence doesn't go that easy with nucleic acids that are hydrophilic in nature. Moreover, the gluten peptides are also negatively charged at the pH used for the aptamer binding buffer. Despite these challenges, few groups have developed aptamers against gliadin and gluten peptides. One group has selected a DNA aptamer against gliadin, called G33 aptamer through the development of a competitive real-time apta-PCR format for the detection of gluten (Pinto *et al.*, 2014). Yet another group from Spain has developed aptamers Gli1 and Gli4 against the 33-residue peptide from α2-gliadin with powerful detection of gluten (Amaya-González *et al.*, 2014). Gli4 aptamer showed the highest affinity towards its target and was used as a recognition receptor for an electrochemical based enzyme linked aptamer-assay using magnetic particles. This assay was sensitive in detecting hordeins, secalins and avenins, with no cross-reactivity to other gluten free crops (corn, soy, rice). This aptamer based assay was claimed to be more sensitive than the reference immunoassay for the same target (Amaya-González *et al.*, 2014). The same group conducted research with the Gli1 aptamer to illustrate detection of gluten in different food matrices including the hydrolyzed ones and the results revealed good correlation with R5 and G12 immunoassays (Amaya-González, de-los-Santos-Álvarez, Miranda-Ordieres & Lobo-Castañón, 2015).

2

Synthesis and characterization of two wheat α-gliadin peptides PFPQPQLPYPQPQLPY and PQPFRPQQPYPQSQPQY containing celiac disease epitopes

2.1 Introduction and overview

Gluten is recorded as the first protein to be studied by scientists as early as 1700s. Jacopo Beccari, a professor of chemistry, at the University of Bologna, isolated and studied gluten in 1728 for the first time (Urade, Sato & Sugiyama, 2018). Gliadins are an important sub-fraction of gluten having importance in food science studies. But despite their long history, it was difficult to determine their structures as they aggregate in aqueous solution. So, they were extracted in 60-70% alcohol or dilute acidic solution by Osborne fractionation method for studies. Gliadins are a large family of proteins with related amino acid sequences, commonly classified as α/β-, γ- and ω-gliadins based on their amino acid composition and molecular weight (Barak, Mudgil & Khatkar, 2015). Structural studies of the gliadins reveal that they possess globular structure with α-type having a dense globular form structure whereas γ-type and ω-type possess extended rod-type arrangement (Barak *et al.*, 2015). The N-terminal repeat domains of α- and γ-gliadins have dominant polyproline-II structure and β-reverse turns while the C-terminal region have α-helix-rich secondary structure stabilized by intramolecular disulfide linkages. On the other hand, ω-gliadins also possess polyproline and β-turn conformation, but without any α-helix secondary structure. In α-gliadins the amino terminal domain retains the most characteristic immunogenic fragment related to celiac disease toxicity. Many studies have reported the presence of a 33-residue toxic fragment (57-89) in the amino-terminal region of α-gliadins presenting the most strong adaptive immune response in celiac patients (Balakireva & Zamyatnin, 2016; Herrera *et al.*,

2015). This 33-residue fragment comprises of six overlying copies of three types of DQ2.5 restricted celiac epitopes glia-α1a, glia-α1b and glia-α2. Another DQ2.5 restricted epitope, glia-α3 showing a little overlap with the 33-residue peptide is also associated with celiac disease. These specific amino acid sequences within the protein have the ability to provoke an autoimmune enteric response destroying the villi of the small intestine resulting in malabsorption and systemic manifestations which are attributed in celiac disease. With the advances in the studies of celiac disease causing epitopes, it is now well understood that the celiac disease precipitating factors are grouped into two class based on their strength and mechanism of action of causing the disease. One group of peptides which are considered the "toxic group" is responsible for the mucosal damage when administered in-vivo. While the other group called the "immunogenic group" is responsible for specifically stimulating HLA-DQ containing T-cells *in vitro*, which were isolated from peripheral blood of celiac patients (Balakireva & Zamyatnin, 2016). All the different types of gluten proteins have their own distinct type of immunogenicity conferred by their own sets of toxic and/or immunogenic peptides or epitopes.

This chapter describes the synthesis, purification and characterization of two α-gliadin peptides of wheat containing few toxic and immunogenic epitopes from the N-terminal repetitive domain of the protein. Since our aim is to develop specific DNA aptamers as probes against celiac disease epitopes for sensitive detection of gluten, two peptide stretches from the α-gliadin protein fraction has been selected as the targets for the present study. The selected peptide sequences are as follows:

- Sequence 1: PFPQPQLPYPQPQLPY (amino acid 61-76) (P16) containing three DQ2.5 restricted overlapping toxic epitopes glia-α1a (PFPQPQLPY), glia-α1b (PYPQPQLPY) and glia-α2 (PQPQLPYPQ) (H. Arentz-Hansen *et al.*, 2000; H. Arentz-Hansen *et al.*, 2002), and
- Sequence 2: PQPFRPQQPYPQSQPQY (amino acid 86-102) (P17), containing the DQ2.5 restricted immunogenic epitope glia-α3 (FRPQQPYPQ) (Vader *et al.*, 2003).

The position of the two peptides containing the respective epitopic region in the α-gliadin protein (accession number AJ133612) is depicted diagrammatically in Fig. 2.1. Also the characteristic properties of the peptide sequences P16 and P17 are elaborated in Table 2.1. The selected peptide sequences were synthesized manually in the laboratory following Fmoc-chemistry and their subsequent purification was done by reverse-phase high-performance liquid chromatography (HPLC). The intact mass characterization was done in a MALDI-TOF/TOF mass analyzer to confirm the presence of correct peptide species. Circular Dichroism spectroscopy was also done to ascertain the secondary structural properties of the peptides.

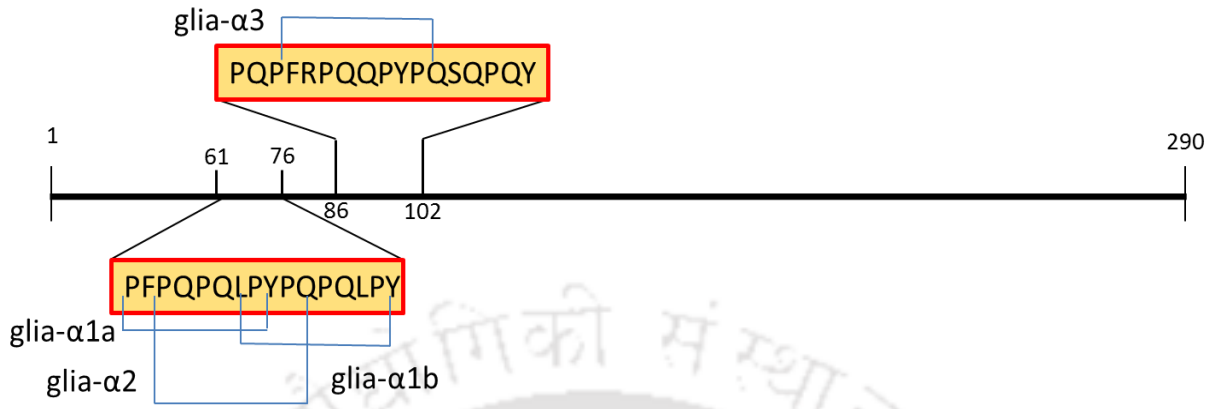


Fig. 2.1 Location of the peptides PFPQPQLPYPQPQLPY (P16) and PQPFRPQQPYPQSQPQY (P17) in the α-gliadin protein sequence showing the respective DQ2.5 restricted celiac disease epitopes glia-α1a, glia-α-1b, glia-α2 and glia-α3

Table 2.1 Properties of the Peptides P16 and P17

Peptide Sequence	No. of residues	Molecular weight (g/mol)	Extinction coefficient ($M^{-1}cm^{-1}$)	Isoelectric point (pI)	Net charge at pH 7
PFPQPQLPYPQPQLPY (amino acid 61-76) (P16)	16	1910.17	2560	3.74	0
PQPFRPQQPYPQSQPQY (amino acid 86-102) (P17)	17	2086.26	2560	9.91	+1

2.2 Materials and Methods

2.2.1 Chemicals and reagents

Wang resin and Nα-(9-Fluorenylmethoxycarbonyl)-amino acids (Fmoc-AA) were purchased from Advanced ChemTech, Louisville, Kentucky, USA. Analytical grade peptide synthesis chemicals, dichloromethane (DCM), N,N-dimethylformamide (DMF), tri-isopropylsilane (TIS), piperidine, and HPLC-grade solvents, trifluoroacetic acid (TFA) and acetonitrile (CH_3CN) were purchased from Merck, India. All other analytical grade chemicals, 1-hydroxybenzotriazole (HOBt), N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl) uranium hexafluorophosphate (HBTU), 4-(N,N dimethylamino)-pyridine (DMAP), N,N-diisopropylethylamine (DIPEA), N,N'-diisopropylcarbodimide (DIC), p-chloranil, acetic anhydride and α-Cyano-4-hydroxycinnamic acid (4-HCCA) MALDI-matrix was procured from Sigma-Aldrich Chemicals Pvt. Ltd., India.

2.2.2 Methods

2.2.2.1 Synthesis of the target α-gliadin peptides

The two peptides PFPQPQLPYPQPQLPY (P16) and PQPFRPQQPYPQSQPQY (P17) were manually synthesized in-house by solid phase peptide synthesis following standard Fmoc-chemistry. Wang resin with resin loading capacity of 1.2 mmol/g was used as the initial solid-support for synthesizing

peptide sequences. Wang resin 200 mg was taken in the reaction vessel and swelled in a solution of DCM/DMF in the ratio 9:1 (v/v) at a concentration of 10-15 mg/ml for 2-3 hours and again in DMF for 30 minutes. For the attachment of the first amino acid, in a round bottom flask, 5-equivalency (relative to resin loading capacity) of the first amino-acid was dissolved in DMF and an equal equivalent of HOBT added to it. Just enough of DMF was added to the amino-acid mixture to completely dissolve the constituents. In another flask 0.3 equivalent of DMAP was dissolved in a little amount of DMF. Then 5 equivalents (relative to resin capacity) of DIC added into the amino-acid mix followed by the addition of dissolved DMAP solution. The activation of the amino acid was done for 8-10 minutes at 0°C prior to its addition to the pre-swelled resin. Nitrogen gas was sparged into the reaction vessel to create an inert environment to aid in the stability of the coupling reagents. The reaction was kept at room temperature with agitation till 2-3 hours. The coupling was repeated once more if required followed by N-terminal capping of the unreacted hydroxyl groups on the resin using 2 equivalents each of acetic anhydride and DIPEA, then kept in agitation for another 30 minutes at room temperature. The resin was washed with DMF (3x times) followed by 1:1 DMF/DCM wash and three times with DCM again. Measurement of the substitution level was monitored by weighing out a small amount of the dried resin by measuring the amount of Fmoc released using spectrophotometry. After coupling, the Fmoc deprotection was accomplished with 20% piperidine in DMF for 20-30 minutes at room temperature followed by filtering the solution and washing the resin with DMF. Coupling of the subsequent amino acids was also done manually by using 5 equivalents of Fmoc-amino acids (5 ml/g of resin), 5 equivalents HBTU (1 M solution) and 10 equivalents of DIPEA in DMF.

Equal equivalency of HOBT was added to the reaction mixture to further decrease the racemisation. The reaction was allowed to continue for 1 to 1.5 hours under nitrogen blanket and washed with DMF (3-4 times) followed by deprotection using 20% piperidine in DMF. The cycle of coupling, deprotection and washing was repeated until the desired peptide sequence was assembled.

2.2.2.2 Chloranil test

Qualitative measurement of coupling and deprotection steps was done by Chloranil test using reagent A (2% acetaldehyde in DMF) and reagent B (2% p-chloranil in DMF). A little amount of resin, at the tip of a micro-tip was taken out from the reaction vessel in a microfuge tube and one drop each of reagents A and B was added and kept for 5 minutes to observe the color change.

2.2.2.3 Peptide cleavage from the resin

After assembling the full length of the peptide sequence, it was cleaved from the resin beads using a freshly prepared cleavage cocktail comprising of trifluoroacetic acid, water and tri-isopropylsilane in the ratio of 95/2.5/2.5 and cleaved for 2-3 hours. Following the cleavage, the peptide solution was filtered through a small plug of glass wool and the resin was rinsed with small portions of TFA. The recovered filtrate is then precipitated using 10x ice-cold diethyl-ether and kept overnight at 4ºC for maximum recovery. The crude precipitate was washed 3-4 times more with diethyl ether to remove any residual scavengers and collected by

centrifugation. Peptides were then dissolved in 1:1 mixture of acetonitrile/water and lyophilized.

2.2.2.4 HPLC Purification of the synthesized peptides

The crude peptides were purified on Agilent Technologies, USA, 1260 Infinity HPLC instrument, in Guwahati Biotech Park on a semi-preparative reverse-phase C18 column using a linear gradient of water and acetonitrile (10-100%) in the presence of 0.1% trifluoroacetic acid. The purified peptides were subsequently confirmed for their purity by analytical reverse-phase HPLC (Agilent Technologies, USA, 1260 Infinity HPLC, in Guwahati Biotech Park). The mobile phases used were, solvent A: 10% acetonitrile and 0.1% TFA in water and solvent B: 0.1% TFA in acetonitrile with temperature of the column maintained at 25°C. The purified peptides were then freeze-dried and stored at -20°C until use. Stock solution of the lyophilized peptides were prepared in deionized water and concentration was estimated using the molar-absorption coefficient of 1280 $M^{-1}cm^{-1}$ at 280nm for the tyrosine amino acid containing P16 and P17 peptides in a UV-spectrophotometer.

2.2.2.5 Characterization of the peptides by mass spectrometry analysis

Characterization of the purified peptide sequences was done using matrix-assisted laser desorption/ionization mass spectrometry on a Bruker, Autoflex Speed MALDI-TOF/TOF mass analyser. The mass spectra were recorded in a positive ion and reflectron mode. 1µl saturated solution of α-Cyano-4-hydroxycinnamic acid

matrix (15mg/ml) dissolved in a mixture of acetonitrile and 0.1% trifluoroacetic acid in the ratio 40:60 was mixed with 2 ul of the 20 µM peptide solution and deposited into the MALDI sample spotting plate. The peptide samples were analyzed by the in-built software associated with the instrument.

2.2.2.6 Secondary structure analysis by circular dichroism spectroscopy

Circular dichroism spectra of the peptides were recoded using a Jasco J-1500 spectropolarimeter in a 1 mm path length quartz-cuvette and 1 nm bandwidth. The spectrum was recorded in the wavelength range of 190-250 nm at a scan-rate of 100nm/min, wavelength interval of 0.5 nm, time constant of 2 seconds and an average of 4 scans. The spectra were corrected for baseline with water contribution subtracted from the spectra and smoothed using the Savitzky-Golay algorithm (Savitzky & Golay, 1964). A concentration of 20µM for both the peptides P16 and P17 were diluted in deionized water from their respective stock solutions (kept at -20°C) and their spectra were recorded at 25°C maintaining the PMT voltage (HT voltage) <500 volts. The recorded spectra were then analyzed by the freely available circular dichroism spectra analysis soft-wares, CDPro software suite (http://lamar.colostate.edu/~sreeram/CDPro) and Dichroweb using the algorithms CONTINLL, CDSSTR SELCON3 and K2D (Sreerama & Woody, 2000; Whitmore & Wallace, 2004).

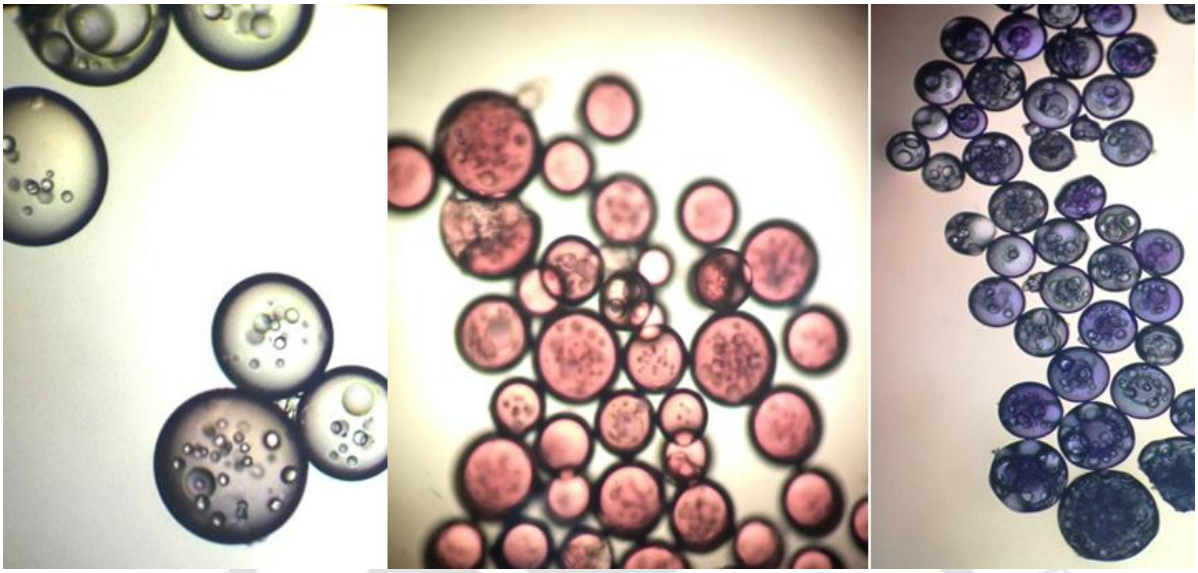
2.3 Results

2.3.1 Synthesis of the target α-gliadin peptides

Peptide synthesis involves alternating cycles of coupling and deprotection. When the desired length of the peptide is assembled, it is cleaved from the solid support. The coupling procedure includes activation of the free carboxyl group of amino acid using uronium salts of benzotriazole under alkaline environment and deprotection involves the removal of the amine protecting Fmoc group. The two peptides were successfully synthesized and monitored in every step by chloranil test.

2.3.2 Chloranil Test

The qualitative assessment of the coupling and deprotection steps was monitored by Chloranil test. This test stain the free amines on the resin either turning them dark red in case of primary amines or dark blue for secondary amines indicating a positive test; and colorless to yellowish beads indicates a negative test (Fig 2.2, A). For every round of successful coupling of amino acids, chloranil test gives a negative result confirming the completion of coupling step. While for every successful round of deprotection, the test gave a positive result showing dark red or dark blue beads (Fig 2.2, B).



A **B** **C**

Fig. 2.2 Microscopic view of the resin beads post staining with chloranil reagent indicating, (A) colorless to yellowish beads for negative test; (B) dark red beads in the presence of primary amines and (C) dark blue beads in the presence of secondary amines indicating a positive test

2.3.3 Peptide cleavage from the resin

The assembled sequence of the peptide is cleaved from the solid support by using a cleavage cocktail containing trifluoroacetic acid (TFA). Wang resin is acid labile and therefore can be easily cleaved with TFA.

2.3.4 HPLC purification of the synthesized peptides

A gradient of 10-100% solvent B after 10 minutes in a total of 60 minutes run time was set for P16 and 50 minutes run time for P17 with UV absorbance recorded at 214 nm, 254 nm and 280 nm. HPLC chromatograms of the peptides revealed major characteristic peaks at 21.933 minutes for P16 (Fig 2.3, A) and 19.90 minutes for P17 (Fig 2.3, B). The purity of the synthesized peptides were >90%, with P16 having 95% purity and P17 with 93% purity as ascertained by reversed-phase analytical HPLC. Percentage of purity for each of the peptide species was calculated by integrating the area under the peak for the respective peptides. The percentage of purity from the chromatograms of the respective peptides was determined by the equation: percentage of peptide purity= [(Peak area under UV detection for the desired peptide)/ (Peak area under UV detection for all other peaks including the desired peptide peak)]*100.

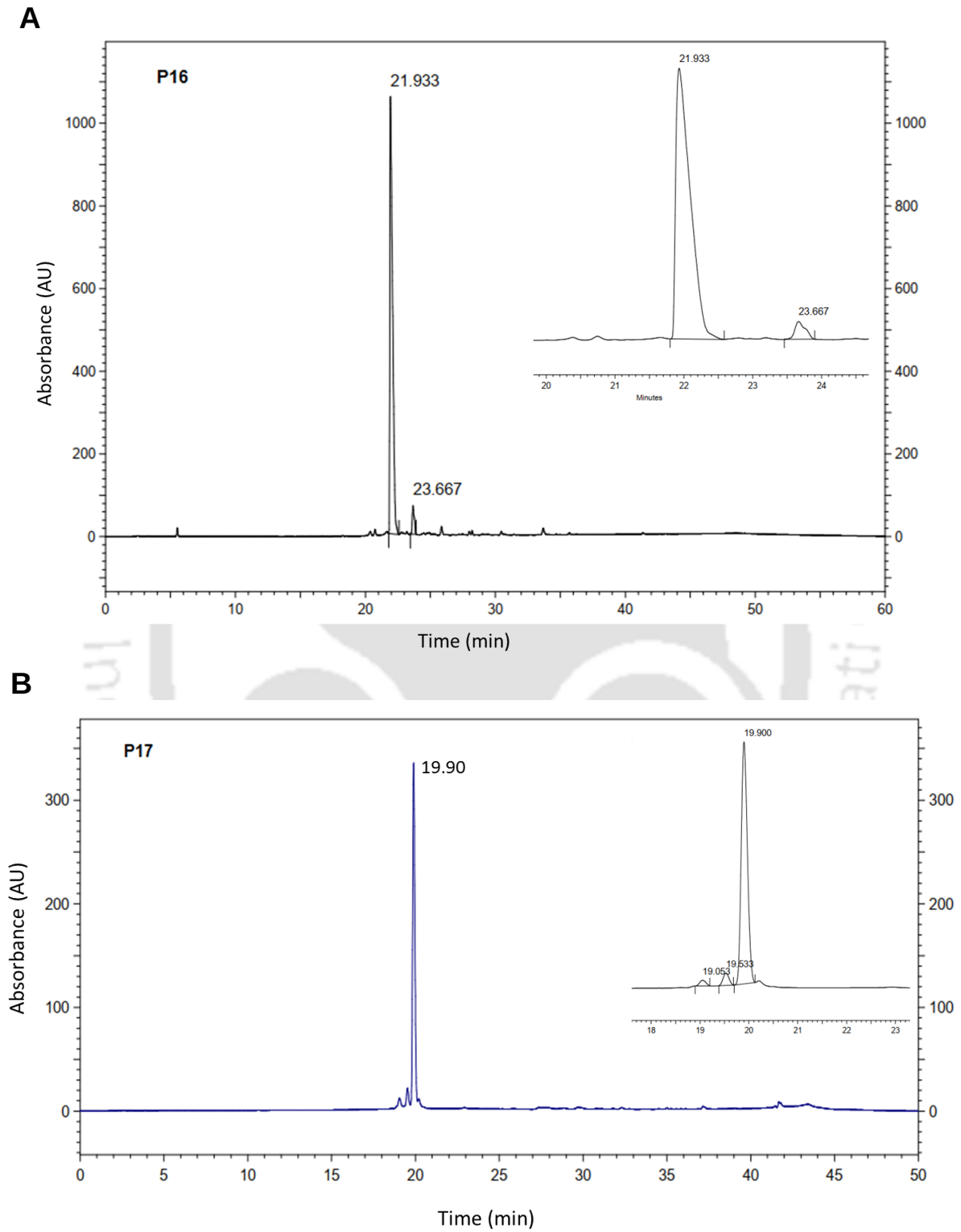


Fig. 2.3 Reverse-phase HPLC chromatograms of the the purified peptides, (A) P16 showing a peak at retention time 21.933 minutes; and (B) P17 shows a peak at retention time 19.90 minutes

2.3.5 Characterization of the peptides by mass spectrometry analysis

The mass characterization of the peptides by mass spectrometry analysis was done to reveal their correct intact molecular weight (in mass/charge) for the respective peptides. Matrix-associated laser desorption/ionization leads to the ionization of the sample-matrix mixture generating ions which are separated depending on their m/z ratio through the time-of-flight analyzer. A spectral profile of these ions were obtained and analyzed by the software, giving a mass representation of the species under investigation. The mass profile of the peptide P16 (Fig. 2.4, A) and P17 (Fig 2.4, B) reveals the m/z corresponding to their respective adducts. The calculated monoisotopic masses for P16 were 1932.372 Da and 1954.369 Da corresponding to the sodium adduct (M+Na and M+2Na-H) and for P17 were 2087.853 Da and 2109.867 Da corresponding to the hydrogen (M+H) and the sodium (M+Na) adducts respectively. The mass spectrum with m/z corresponding to their respective adducts did not signify any impurity in the peptide and confirmed the presence of the pure peptide species as desired, with the adduct mass added to their correct intact molecular weight.

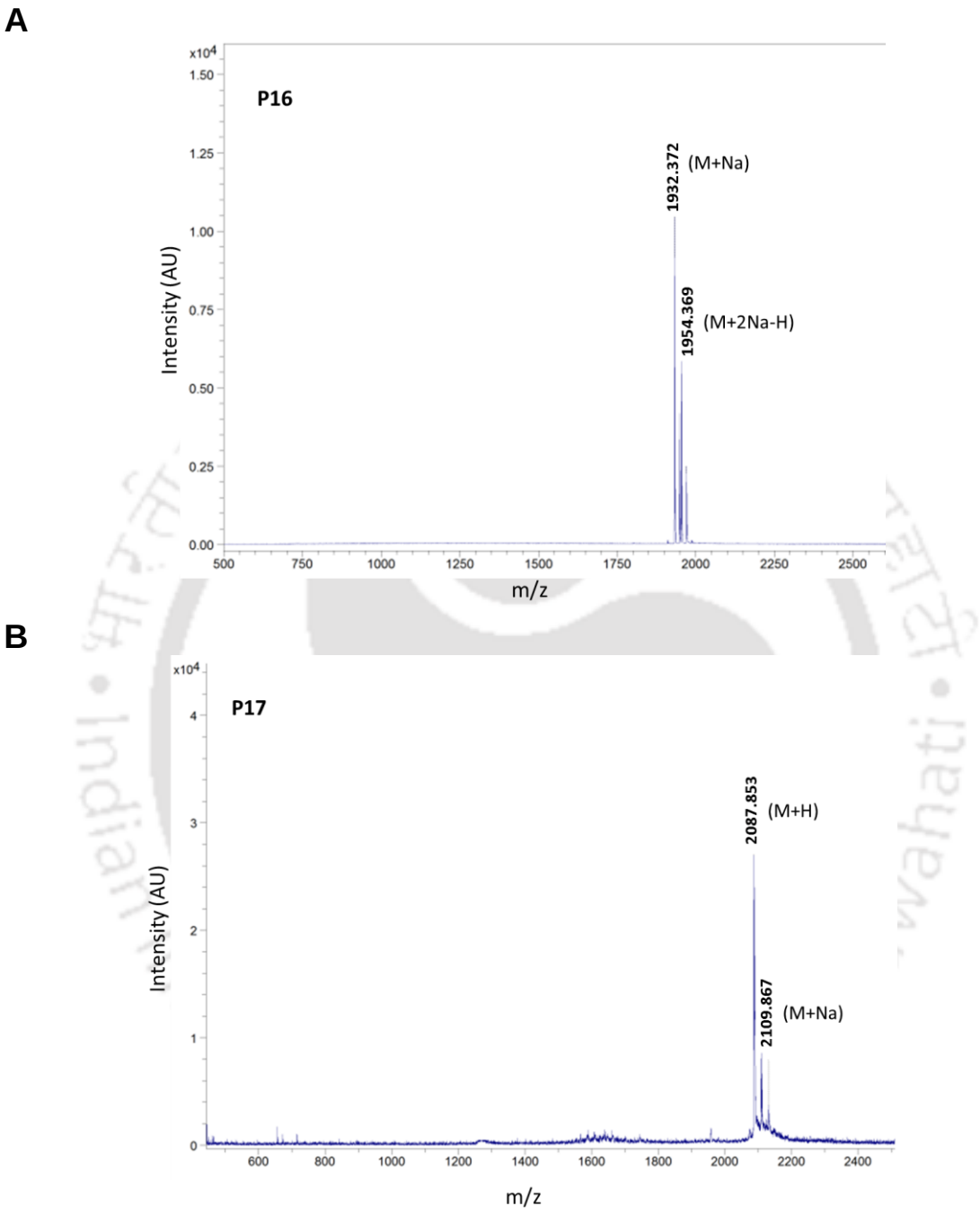
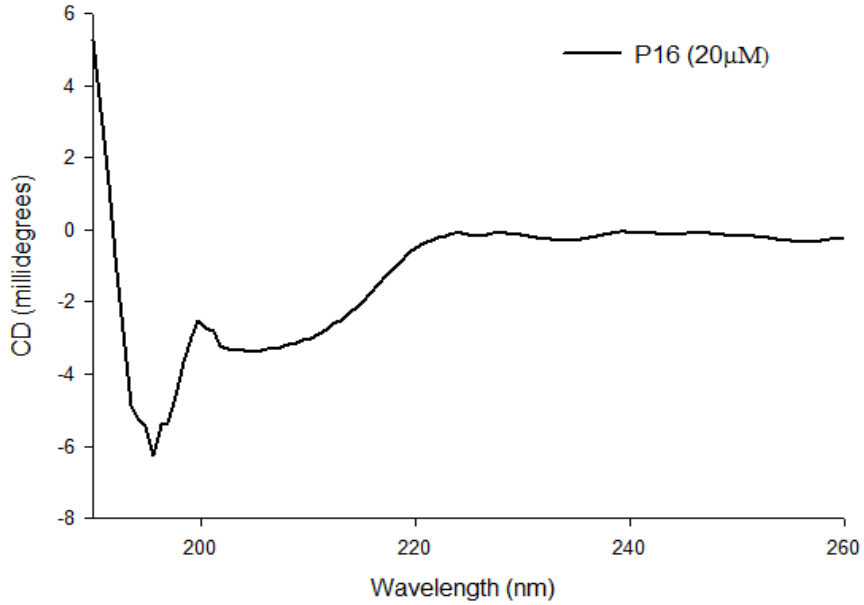


Fig. 2.4 MALDI-TOF mass spectra of the purified peptides, (A) The calculated monoisotopic masses for P16 are 1932.372 Da (M+Na) and 1954.369 Da (M+2Na-H) corresponding to the sodium adducts; and (B) for P17 are 2087.853 Da and 2109.867 Da corresponding to the hydrogen (M+H) and sodium adducts (M+Na) respectively (M denotes the mass of the parent compund)

2.3.6 Secondary structure analysis by circular dichroism spectroscopy

Circular dichroism studies were done to ascertain the secondary structural conformation of the peptides. The spectra for both the peptides were recorded in the far-UV wavelength range of 190 nm to 260 nm at 25°C. The CD spectra for both the peptide P16 (Fig 2.5, A) and P17 (Fig 2.5, B) shows a strong negative minima from ~196-200 nm revealing the presence of mostly random structure. The CD spectra of the peptides were analyzed by two protein secondary structure analysis softwares, CDPro software suite and Dichroweb using the CD spectroscopic data. The CDPro software suite and Dichroweb analyse the secondary structural components of the peptides using popular algorithms like SELCON3, CONTINLL, CDSSTR and K2D comparing the experimental data with the calculated results. The softwares compare the calculated CD spectrum from each of these programs with the experimental spectra and also calculate the RMSD between the two spectrums. Hence, the best calculated result was selected based on the lowest RMSD (root mean square deviation) and NRMSD (normalised root mean square deviation) values. The estimated amount of secondary structures as calculated from the softwares had an average of ~42% of random structure, ~23% of turns and lesser amount of beta structure without any helix for the peptide P16. Similarly P17 consists of 44% of random structure, ~24% of turns, far lesser beta structure and insignificant amount of helix structure. Table 2.2 shows the values of different secondary structural components as calculated from these softwares with the lowest NRMSD values.

A



B

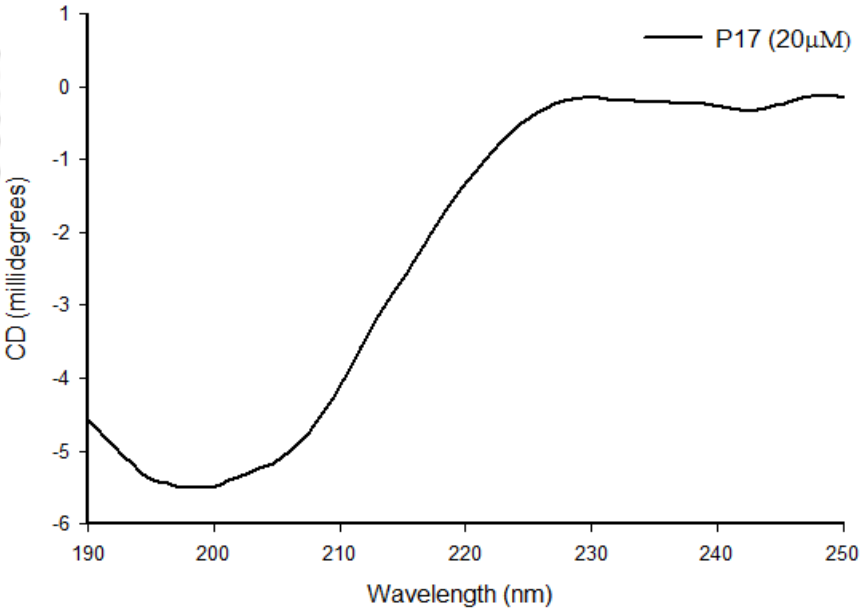


Fig. 2.5 Far-UV circular dichroism spectra of both the peptides, (A) P16 and (B) P17 showing negative ellipticity in the range ~196-200 nm

Table 2.2 Calculated secondary structural fractions of peptides P16 and P17 by circular dichroism analysis softwares CDPro software suite and Dichroweb

Peptide	Analysis Software	Secondary Structural components						NRMSD
		Helix (1)	**Helix (2)**	**Strand (1)**	**Strand (2)**	**Turns**	**Random**	
P16	**CDPro** CONTINLL	0.00	0.04	0.08	0.12	**0.25**	**0.39**	0.075
	Dichroweb CDSSTR	0.00	0.05	0.18	0.10	**0.20**	**0.46**	0.023
P17	**CDPro** CONTINLL	0.00	0.06	0.11	0.09	**0.28**	**0.44**	0.040
	Dichroweb CDSSTR	0.01	0.04	0.20	0.10	**0.20**	**0.44**	0.012

2.4 Discussions

Wheat gliadins are large families of proteins with similar amino-acid sequences. They are classified into three main groups (α-, γ- and ω-gliadins) based on their electrophoretic mobility (Wrigley & Shepherd, 1973). Wheat α-gliadins contain the major immunodominant epitopes glia-α1a (PFPQPQLPY), glia-α1b (PYPQPQLPY) and glia-α2 (PQPQLPYPQ) present in the 33-residue peptide and an additional immunogenic DQ2.5-glia-α3 (FRPQQPYPQ) epitope (Ozuna et al., 2015).

These epitopic regions are the minimal amino acid sequence, which are highly reactive to human $CD4^+$ T-cells inducing inflammatory reactions in the small intestine causing celiac disease in predisposed individuals (Mitea *et al*., 2008). These proline and glutamine rich repetitive epitopes and peptides comprising the gliadin protein are rich in hydrophobic amino acids (phenylalanine, leucine and tyrosine) and are highly resistant to digestive enzymes presenting themselves as the true precipitating factors of celiac disease. Therefore, targeting the peptides containing these epitopes will allow for more accurate assessment of gluten toxicity. Moreover these epitopes are shown to have homologous sequences in hordein, secalin and avenin proteins which will further help in the assessment of more and more potentially harmful gluten and related epitopes in different members of the cereal family (Mitea *et al*., 2008). So in the present study we have selected these native non-deamidated peptide sequences (P16 and P17), keeping in mind the fact that they are highly desirable targets of gluten detection.

The main objective of this chapter was to synthesize and purify these selected P16 and P17 α-gliadin peptides as target for aptamer selection. The peptides were synthesized in-house manually by Fmoc-chemistry and purified by semi-preparative reversed-phase HPLC. The purity of the peptides were >90%, which is highly desired for generation of specific aptamers. The purified peptides were then characterized by mass spectrometry analysis for the presence of correct mass species. The mass analysis revealed the presence of species with mass corresponding to the cationic adducts of sodium and hydrogen, hence giving higher values of calculated monoisotopic masses for P16 and P17. These higher m/z corresponding to their

respective adducts does not signify impurity in the peptide, with the peptides having their correct molecular mass intact.

The second objective of this chapter is to study the secondary structural characteristics of the target peptides by circular dichroism. Based on secondary structural studies of α-gliadin proteins, it is predicted that the N-terminal repetitive domians comprises of a mixture of polyproline-II structure and β-reverse turns, while the non-repetitive C-terminal domains are rich in α-helix (Hsia & Anderson, 2001; Matsuo, Kohno & Morita, 2005; Matsushima, Creutz & Kretsinger, 1990; Tatham & Shewry, 1995). A recent study by Urade *et al.* (2018) have reported the presence of more random structure and small amount of α-helices in their study of recombinat α-gliadin protein (unpublished data). In the present study, the CD spectra analysis of the peptides revealed mostly random structure with the presence of turns and far lesser amount of beta-structure with insignificant or no helix formation (Table 2.2). This finding seems to be in aggrement with the earlier predicted data from other published works. So it is assumed that the random structure present in our peptides can be attributed to the presence of polyproline-II (PPII) conformation whose spectral properties are not yet being well differentiated from that of random structures. Moreover, the sequences of the peptides P16 and P17 are also rich in repetitive prolines, which further establish the assumption for the presence of PPII structure. Hence, the experimental results of the CD spectra for the peptide sequences P16 and P17 (which are a part of the amino-terminal repetitive region of the α-gliadins) are in conformity with the software calculated result that reveals mostly the presence

of random conformation with the presence of turns and a lesser amount of β-structure as shown in Table 2.2.

2.5 Conclusion

As an initial step towards the aptamer selection, two α-gliadin peptide sequences P16 (PFPQPQLPYPQPQLPY) and P17 (PQPFRPQQPYPQSQPQY) were synthesized and purified by reverse-phase HPLC with >90% purity as desired for both the peptides. The mass characterization of the peptides was subsequently done in a MALDI TOF/TOF mass analyzer to reveal the presence of correct molecular weight species. Finally, the structural integrity and conformation of the purified peptides were confirmed by circular dichroism spectra revealing mostly random structure attributed to the presence of polyproline II conformation, turns and a lesser amount of beta-structure.

3

Selection and characterization of specific DNA aptamers against the α-gliadin peptide PFPQPQLPYPQPQLPY

3.1 Introduction and overview

Detection of gluten is mainly done by antibody based immunoassays and kits which are considered as the gold standard in terms of current conventions given by Codex Alimentarius (Commission for International Food Standards). However, the instability of antibodies at high temperature, batch to batch variations, denaturation in the presence of reducing agents used in gluten extraction and obvious high cost of production make the whole process of gluten assessment very expensive. In order to circumvent the inherent disadvantages of the antibodies, alternative biocapture agents or biorecognition probes called aptamers are developed. These aptamer molecules can be either nucleic acids (DNA or RNA) or peptides. Nucleic acid aptamers are a group of highly specific and sensitive single stranded oligonucleotide ligands that bind to their targets with high binding affinity. Ever since its inception in 1990, the field of aptamers has undergone immense developments and improvements. The advantages and numerous applications of aptamers as an emerging biorecognition probe in the field of food safety have caught the attention of researchers with the potential to replace antibody based detection of gluten. Aptamers are selected *in vitro* by the process of SELEX. Aptamers offer numerous inherent advantages over antibodies like stability at high temperature and in a range of ionic conditions, regeneration capability, comparable binding affinity like antibodies with target specific binding and discriminating subtle structural

differences, non-immunogenic, ease of chemical modification, rapid and high scale-up capacity with low cost of production compared to antibodies.

In this chapter we describe the *in vitro* selection of specific DNA Aptamers against the synthetic α-gliadin peptide PFPQPQLPYPQPQLPY (P16) (61-76 amino acids, molecular weight 1910.17 Daltons) containing three overlapping DQ2.5 restricted epitopes glia-α1a (DQ2-α-I, α9), glia-α1b (DQ2-α-III) and glia-α2 (DQ2- α-II, α2). These epitopes are considered as the most immunodominant and toxic factors causing celiac disease. The peptide P16 is a part of the 33-mer immunodominant peptide which is strongly implicated in celiac disease. *In vitro* aptamer selection was carried out involving magnetic bead based SELEX using the peptide P16 which was synthesized and purified in the laboratory. The selection process identified two aptamer candidates that were highly enriched with the target peptide. The secondary structure characterization studies revealed different structural composition of the aptamers. Moreover, the binding studies revealed the strong binding affinity of the aptamer AG1 towards its target.

3.2 Material and Methods

3.2.1 Chemicals, reagents and kits

Primers and single-stranded DNA library (aptamer library) were procured from Integrated DNA Technologies (USA). Dynabeads M270 carboxylic acid, SYBR Gold nucleic acid stain, Electrophoretic Mobility Shift Assay (EMSA) kit and InsTA clone

PCR cloning kit were purchasedx from Invitrogen Thermo Scientific (USA). Streptavidin magnetic beads and Low molecular weight (LMW) DNA ladder were procured from New England Biolabs (USA). IPTG (Isopropyl *β*-D-thiogalactopyranoside) and X-Gal (5-bromo-4-chloro-3-indolyl-*β*-d-galactopyranopside) were purchased from Himedia (India). Miniprep plasmid isolation kit and, Sigma Readymix and agarose were procured from Sigma Aldrich (India). All other chemicals were analytical grade and cited wherever necessary.

3.2.2 Methods

3.2.2.1 Aptamer library design and PCR optimization

A single stranded, degenerate, oligodeoxynucleotide library containing a 40-mer nucleotide random region flanked by two primer annealing sites on both 5' and 3' ends was taken as the starting DNA pool for *in vitro* selection of DNA aptamers. The design of the aptamer library is **5'-ATACCAGTCTATTCAATT-N40-AGATAGTATGTGCAATCA-3'** (76-mer). This combinatorial library was synthesized in a 1 µmole scale and reconstituted in Tris-EDTA buffer, pH 7.5.

The amplification of the obtained aptamer library was optimised by the PCR cycle as mentioned in Table 3.1 using unmodified forward primer, F1: 5'-ATACCAGTCTATTCAATT-3', and unmodified reverse primer R1: 5'-TGATTGCACATACTATCT-3' for 14 cycles of amplification.

Table 3.1 Thermal profile of the optimised PCR cycles used in the amplification of aptamer library

PCR cycles		**Temperature**	**Time**
1. Initial Denaturation		94° C	5 minutes
2. 14 cycles of	Denaturation	94˚C	30 seconds
	Annealing	45˚C	30 seconds
	Extension	72˚C	45 seconds
3. Final extension		72˚C	5 minutes

3.2.2.2 Immobilization of the peptide on Dynabeads

Activated Dynabeads 100µl and an excess amount of the peptide as calculated from the bead binding capacity (30 mg/ml) were immobilized in the magnetic beads (according to the manufacturer) and incubated for 2 hours at room temperature. The peptide immobilization is based on carbodimide chemistry (Fig 3.1). After immobilization the unbound peptides were removed by washing the Dynabeads according to the manufacturer's guidelines. The coated Dynabeads were checked by mass spectrometry analysis on a MALDI-TOF/TOF mass analyzer for confirming the immobilization of the peptide on the beads. Then the immobilized beads were subsequently used for the SELEX rounds.

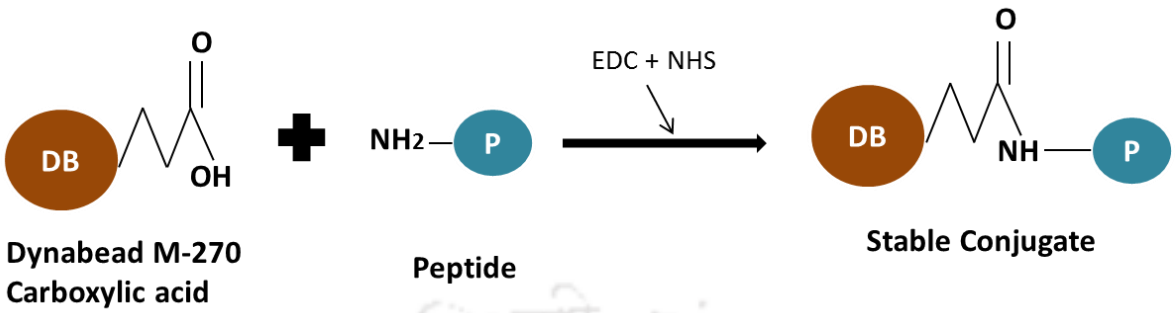


Fig. 3.1 Diagrammatic illustration of the reactions involved in peptide immobilization on Dynabeads M270 Carboxylic acid via carbodimide chemistry

3.2.2.3 *In vitro* selection of DNA aptamers by SELEX

500 picomoles of the starting DNA library 5'-ATACCAGTCTATTCAATT-N40-AGATAGTATGTGCAATCA-3' was dissolved in 500 µl of SELEX binding buffer (20 mM Tris HCl, pH 7.5, 250 mM NaCl, 5 mM MgCl2 and 5 mM KCl) and activated by heating at 95°C for 8 min, immediately cooled in ice for 8-10 min and then brought to room temperature for 10 min. Every time at the start of a new SELEX round, the aptamer library was heat activated and cooled to room temperature prior to the use. Before starting the first positive SELEX round, a negative SELEX cycle was carried out with 50µl of activated Dynabeads without immobilized peptide for 1 hour. The unbound DNA pool from the negative SELEX round was then used for the following rounds of aptamer selection. The unbound DNA pool from the negative SELEX cycle was retrieved and incubated with the peptide immobilized Dynabeads for 60 min in the first positive round. After completion of the first binding round, the unbound DNA pool was removed by washing with SELEX washing buffer (20mM

Tris HCl, pH 7.5, 250mM NaCl, 5mM MgCl2 and 5mM KCl, 0.01% Tween 20) and the bound sequences were heat eluted from the peptide immobilized beads and amplified by the optimized PCR protocol (Table 3.1) using unmodified forward primer F1 and 5'-biotin conjugated reverse primer R2: 5'-biotin-TGATTGCACATACTATCT-3'.

In order to separate the sense strand from the double-stranded PCR product, the PCR amplified aptamer pool was incubated with 20 µl of streptavidin magnetic beads for 1 hour at room temperature. The beads were then washed with streptavidin magnetic bead washing buffer (20mM Tris-HCl pH 7.5, 0.5M NaCl, 1mM EDTA) and then the sense strands were separated from the biotinylated strands by alkaline denaturation using 100µl of 100mM NaOH. The pH of the solution was subsequently adjusted to 7.4 using HCl. The separated single-stranded DNA pool was used for the selection rounds. A total of 16 iterative SELEX cycles were carried out to obtain the final enriched pool of aptamers. A total of 4 negative SELEX rounds were carried out to eliminate non-specific DNA sequences. The stringency of the SELEX rounds was increased by decreasing the binding time of the DNA pool with the immobilized Dynabeads, increasing the washing cycles and decreasing the bead volume used for binding with the DNA pool with progressive SELEX rounds. Table 3.2 describes in detail the SELEX conditions used for the DNA aptamer selection against peptide P16.

Table 3.2 SELEX conditions used in DNA aptamer selection against peptide P16

Rounds	Beads (in µl) + Target (pmoles)	Binding Time (in min)	Washing Steps	Negative SELEX
1st	100 µl, 2100 pmoles	60	1 x 200 ul	Yes
2nd	100 µl, 2100 pmoles	60	1 x 200 ul	---
3rd	100 µl, 2100 pmoles	60	2x 200 ul	---
4th	100 µl, 2100 pmoles	60	2x 200 ul	---
5th	100 µl, 2100 pmoles	60	2x 200 ul	Yes
6th	100 µl, 2100 pmoles	60	2x 200 ul	---
7th	100 µl, 2100 pmoles	60	2x 200 ul	---
8th	100 µl, 2100 pmoles	60	2x 200 ul	---
9th	100 µl, 2100 pmoles	60	2x 200 ul	Yes
10th	100 µl, 2100 pmoles	60	2x 200 ul	---
11th	100 µl, 2100 pmoles	60	3x 200 ul	---
12th	100 µl, 2100 pmoles	30	3x 200 ul	---
13th	100 µl, 2100 pmoles	30	3x 200 ul	Yes
14th	50 µl, 1050 pmoles	30	3x 200 ul	---
15th	50 µl, 1050 pmoles	15	3x 200 ul	---
16th	50 µl, 1050 pmoles	15	3x 200 ul	---

3.2.2.4 Cloning and sequencing of aptamers

The screened aptamer pool after the last round (16th round) of SELEX was PCR amplified with unmodified forward and reverse primers (F1 and R1 primers). These PCR products were purified by electro-elution. The PCR products were separated in 2% agarose gel electrophoresis using 0.1X TAE (Tris-acetate EDTA)

buffer. The correct size (76bp) product from the gel was excised and placed in a dialysis tube filled with 500µl of 0.1X TAE. The bag was then placed in the electrophoresis tank for 30-45 min applying a power of 80 Volts. The purified aptamer pool was then ligated using the reaction setup as mentioned in Table 3.3. The ligation reaction was incubated for 16 hours at 4°C. The ligated construct was then cloned into specialized TA cloning vector pTZ57R/T by InsTA Clone PCR Cloning kit and transformed into competent DH5α cells. Transformed clones were selected on the basis of blue/white screening. 80 white colonies were picked, grown overnight in liquid LB (Luria Bertani) media and plasmid was isolated by miniprep plasmid isolation kit. The isolated plasmid DNA was verified for the presence of aptamers with PCR amplification using primers F1 and R1. Subsequently, the positive plasmids were sent for Sanger's sequencing using M13 forward and reverse primers to AgriGenome Labs Pvt. Ltd., India.

The amount of insert (in ng) to be used in ligation reaction is calculated by the standard formula:

Amount of Insert (ng) = {Amount of vector (ng) x size of insert (kb)/Size of vector (kb)} x molar ratio of Insert:vector. The details of the reaction set up for the ligation is described in Table 3.3.

Table 3.3 Ligation reaction set-up for TA cloning of enriched aptamer pool

Reaction components	**Volume**
10X Ligation buffer	1.0 µl
Vector pTZ57R/T (50ng)	1.0 µl (55 ng)
Purified Aptamer pool	Added in ratio 3:1 (insert/vector ratio)
T4 DNA ligase (3 units/µl)	1.0 µl
Nuclease free water	Volume made to 10 µl
Total volume	10 µl

3.2.2.5 *In silico* structure prediction of aptamers

In silico secondary structure prediction and thermodynamic properties of the enriched sequences were analyzed by a free energy minimization algorithm using Mfold web-server (Zuker, 2003). The ionic conditions for Mfold study were kept same as the SELEX binding buffer with temperature fixed at 25°C. The QGRS (Quadruplex forming G-rich sequences) mapper software was used to analyze the presence of G-quadruplexes forming regions in the candidate sequences (Kikin *et al.* 2006). The parameters were set at maximum QGRS length of 40 with G-group number and loop size kept as default (i.e. minimum G-group of 2 and loop size between 0 and 36 nucleotides). The authors have also devised a scoring system to evaluate a QGRS in its likelihood to form a stable G-quadruplex structure. This scoring method scores

sequences based on its capability to form better G-quadruplex forming candidates given by their G-scores. Higher G-scores are provided to a sequence that will make better G-quadruplex forming candidates.

3.2.2.6 Structural analyses of aptamers by circular dichroism

The structural integrity of the aptamer candidates was analyzed using circular dichroism (CD) spectroscopy. For structural characterization a total of 4µM of the candidate aptamers both in binding buffer and deionized water was activated by heating at 95°C for 8 min, immediately cooled in ice for 8-10 min and then brought to room temperature for 10 min. Spectral readings were recorded in a Jasco J-1500 spectropolarimeter at 25°C. Far-UV scanning was obtained in the range 200-340 nm, with an average of 4 scans at a rate of 100 nm/min, bandwidth 1 nm, wavelength step size of 0.5 nm and response of 2 seconds. Baseline was corrected and solvent contributions were subtracted from the spectra and smoothed by Savitsky-Golay algorithm (Savitzky & Golay, 1964).

3.2.2.7 Binding affinity studies of aptamers

3.2.2.7.1 Circular Dichroism study

The effect of binding of the peptide with the aptamer candidates was studied using circular dichroism (CD). For binding analyses of the aptamer-peptide complex, a total of 4µM of aptamer was heat activated and cooled to room temperature

followed by addition of 20 µM of the target peptide and incubating for 1 hour prior to recording of the CD spectra. The CD spectra were recorded with the same parameters as mentioned in section 3.2.6. Baseline was corrected and solvent contributions were subtracted from the spectra and smoothed by Savitsky-Golay algorithm for all the experiments.

3.2.2.7.2 Isothermal Titration Calorimetry (ITC) study

Thermodynamic parameters and binding affinity of the aptamer-target interaction was studied by isothermal titration calorimetry (ITC). Calorimetry experiments were performed at 25°C using a MicroCal iTC200 (GE Healthcare, UK). A total of 30 µM aptamer in SELEX binding buffer was activated by heating at 95°C for 8 min, cooled in ice for 8-10min and then to room temperature prior to the ITC experiment. The aptamer (12514 Daltons) was loaded into the calorimetric cell (350 µl) and titrated by sequential addition of 1.5 µL-aliquots of the peptide P16 (1910.17 Daltons) from the stirring syringe (180 rpm) at a concentration of 900 µM at a time interval of 150 seconds and for 26 injections. The volume of the first injection was 0.4µl and it was not considered in the fitting. The control experiment without the aptamer in the cell was performed to correct the heat of dilution. ITC study with the related peptide P17 was also performed to check for specificity of the aptamer towards its target peptide using the same experimental conditions.

3.2.2.7.3 Electrophoretic mobility shift assay (EMSA) study

To further ascertain the formation of the aptamer-peptide complex, EMSA was conducted. A fixed amount of 10 nM of the aptamer was heated at 95°C for 8 min, immediately cooled in ice for 8-10 min and equilibrated to room temperature followed by the addition of increasing concentration of the target peptide (P16: 0, 25 µM, 50 µM, 100 µM and 150 µM) in a total of 25 µl reaction volume. The mixture was incubated at room temperature for one hour and then separated using 15% native (29:1 acrylamide/bisacrylamide) TBE (Tris-borate EDTA) gel for about 1.5 hour at 180 V and 4°C. The gels were subsequently stained with SYBR Gold dye and visualized under UV. A control lane with only aptamer was loaded for reference.

3.3 Results

3.3.1 Aptamer library design and PCR optimization

The design of the aptamer library plays a critical role in determining the success of a SELEX process. The most important feature of the aptamer library is its stability and capacity to denature easily at 95°C during PCR amplification. The N40 aptamer library used in the present study has a mean melting temperature of 69.1°C ensuring that it is easily denatured during PCR. Moreover, the random region of the aptamer library was chosen in such a way that it provides optimum diversity which is required for the successful evolution of aptamers. The structure of the aptamer

library was predicted by Mfold which revealed that the random regions were mostly involved in the three-dimensional diversity of the library (Fig 3.2).

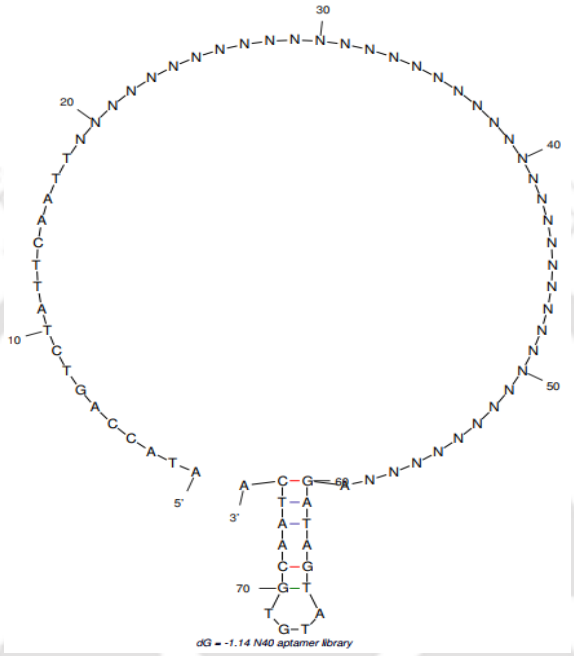


Fig. 3.2 Structure prediction of the N40 aptamer library by Mfold

The amplification of the N40 aptamer library with the optimised PCR cycle for 14 cycles of amplification yielded the correct size product that shows bands of 76bp confirming the successful optimisation of the oligonucleotide pool (Fig 3.3).

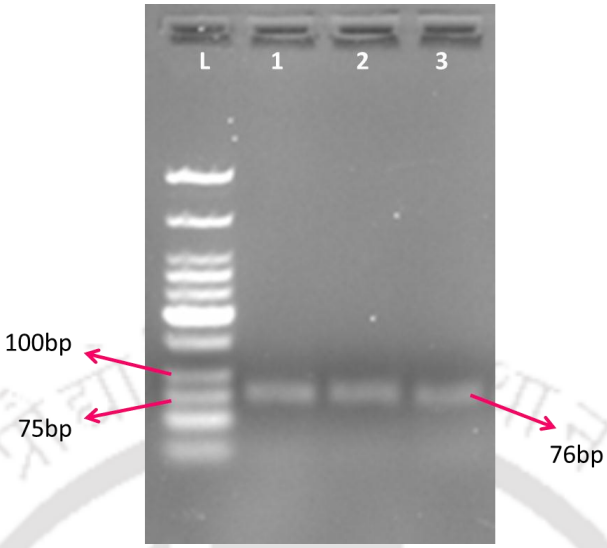


Fig. 3.3 PCR optimization of N40 aptamer library for 14 cycles of amplification (L denotes low molecular weight DNA ladder and the lanes 1, 2, 3 shows the amplified aptamer library)

3.3.2 Immobilization of the peptide on Dynabeads

The immobilization of the peptides on Dynabeads were successfully optimized and confirmed by matrix-assisted laser desorption/ionization mass spectrometry analysis (MALDI-TOF) for the presence of the correct mass of the peptide P16 before using it in SELEX process.

3.3.3 *In vitro* selection of DNA aptamers by SELEX

A total of 16 rounds of SELEX with 4 negative SELEX cycles were carried out. The negative rounds were carried out to eliminate non-specific candidate sequences

while the positive cycles amplified the best binders in each round. With each of the progressive positive cycles of SELEX a gradual enrichment of the candidate aptamers was occurring. The enrichment of the aptamer pool in SELEX rounds were monitored by 2% agarose gel electrophoresis (Fig 3.4) showing successful generation of enriched pool of aptamers at the end of the 16th cycle. Then these candidates were cloned and sequenced.

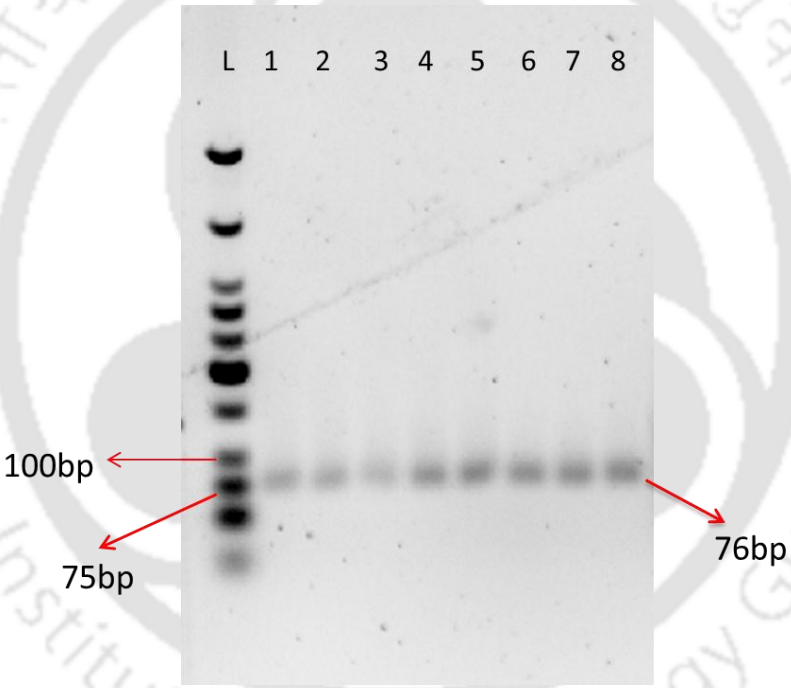


Fig. 3.4 Enrichment of the aptamer pool in SELEX cycles against peptide P16; 2% agarose gel electrophoresis image of PCR amplification in various SELEX rounds showing increase in intensity of the bands with progress in SELEX cycle (Lane L: ladder, Lane 1-8 shows the 2nd, 4th, 6th, 8th, 10th , 12th, 14th and the 16th rounds of SELEX cycle)

3.3.4 Cloning and sequencing of enriched aptamers

A total of 80 clones were screened, plasmid DNA isolated and amplified by specific primers which revealed 19 clones carrying the aptamer inserts. The clones were subjected to sequencing which revealed only 13 sequences with the proper aptamer inserts. Among them two aptamer sequences AG1 and AG2 were highly enriched (Table 3.4). Among both the sequences AG1 had the highest copy number.

Table 3.4 Enriched aptamer sequences against target P16 after 16 rounds of SELEX

Aptamer name	Sequence
AG1	CAGGTGTGTCAGAGCGGGTGTCCGAGGGAGATAGGGTCCT
AG2	AACGCGCCACCCCCTCTAGAACACTAACGAGTATCTTCCA

3.3.5 *In silico* structure prediction of aptamers

Secondary structure of the aptamers AG1 and AG2 were predicted using the Mfold web server (Zuker, 2003). The δG values for AG1 and AG2 were deduced as -11.87 and -7.43 kcal mol^{-1}, respectively (Table 3.5). The structure of AG1 contains two conjoined stem loop structures and aptamer AG2 forms and extended arm with two small loops (Fig 3.5). Further analysis of the sequences by QGRS Mapper

software (Kikin *et al.*, 2006) revealed the formation of one G-quadruplex structure in AG1 (Table 3.6).

Table 3.5 Thermodynamic profile of enriched sequences as deduced by Mfold

Aptamer name	Sequence of the enriched Aptamer	Free Gibbs energy value (ΔG) (kcal/mol, 25˚C)	Enthalpy change (ΔH) (kcal/mol)	Entropy change (ΔS) (cal/k.mol)
AG1	CAGGTGTGTCAGAG CGGGTGTCCGAGGG AGATAGGGTCCT	-11.87	-59.90	-161
AG2	AACGCGCCACCCCC TCTAGAACACTAACG AGTATCTTCCA	-7.43	-43.30	-120.3

Table 3.6 The QGRS sequence prediction for aptamer AG1 by QGRS Mapper software

Aptamer name	Position	Length	QGRS	G-score
AG1	3	34	GGTGTGTCAGAGCGGGTGT CCGAGGGAGATAGGG	27

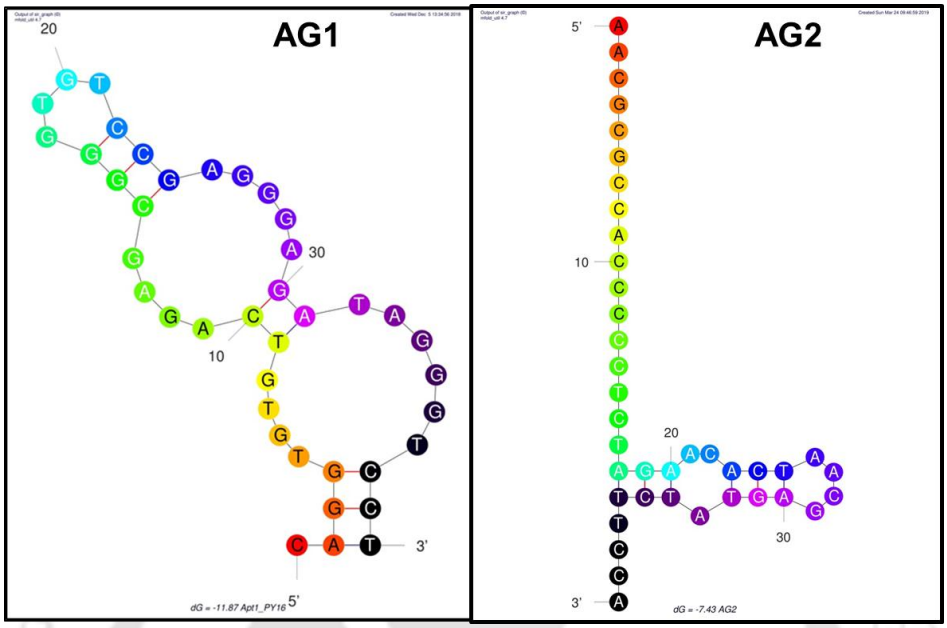
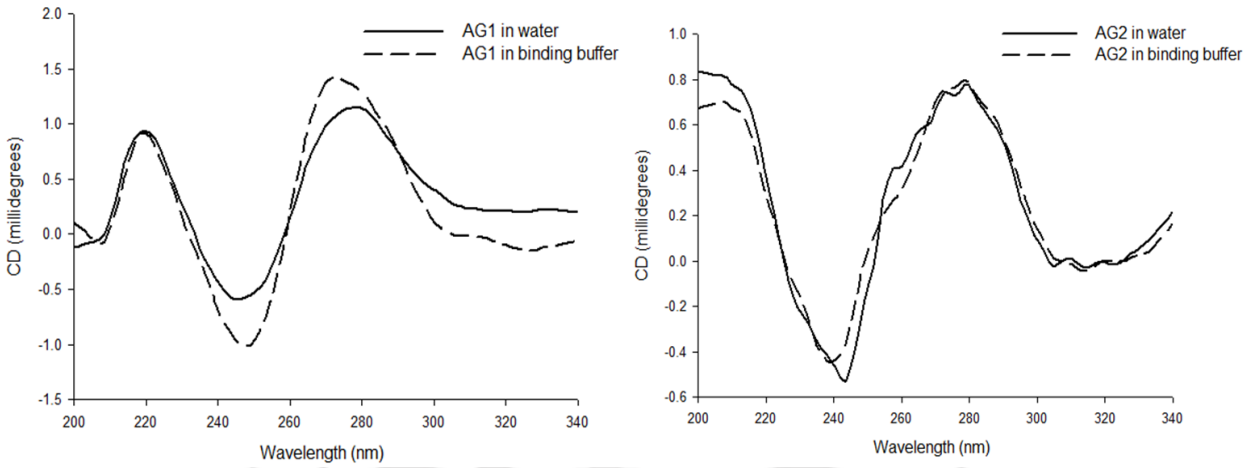


Fig. 3.5 Secondary structure prediction of the enriched aptamers by Mfold; (A) aptamer AG1 forms two conjoined stem loop structures and (B) aptamer AG2 forms an extended arm with two small loops

3.3.6 Structural analyses of aptamers by circular dichroism

To confirm the integrity and structure of the aptamers as predicted by Mfold and QGRS Mapper, circular dichroism spectra of both the aptamers AG1 and AG2 were studied in SELEX binding buffer and deionized water. The signature peaks corresponding to a G-quadruplex structure were not seen; however peaks corresponding to duplex B-form DNA with CD maxima near 280nm and a negative peak near 245nm (Fig 3.6) was found for both the aptamers suggesting that they correctly hybridized to form double helical regions as deduced by Mfold predictions.



A **B**

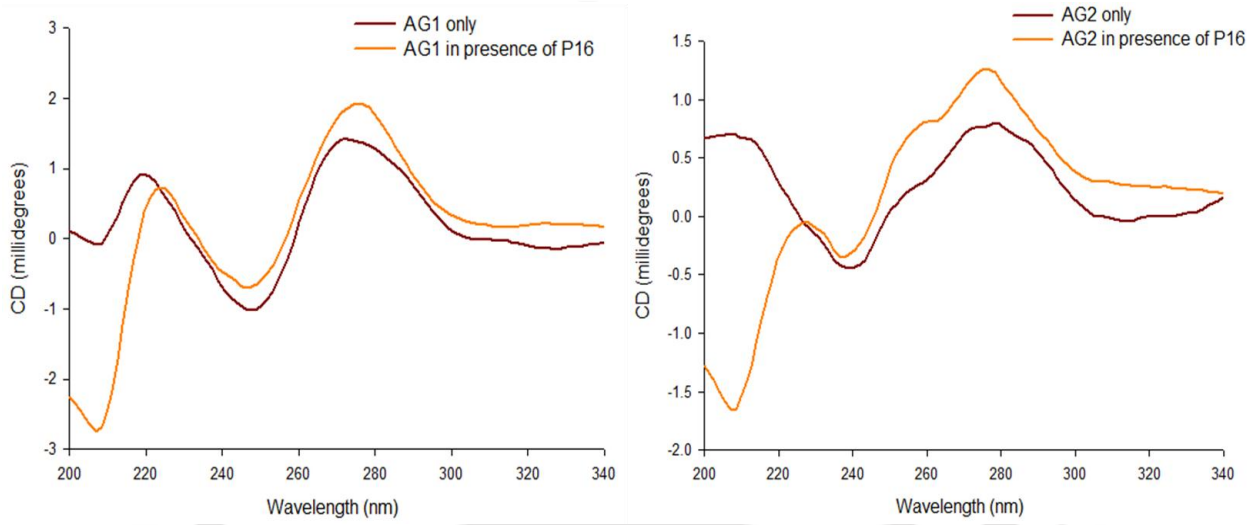
Fig. 3.6 Circular dichroism spectra of aptamers AG1 and AG2 in deionized water and SELEX binding buffer; the CD spectra of both (A) AG1 and (B) AG2 show maxima near 280nm and minima near 245nm indicating the presence of B-form DNA with the formation of double-stranded regions as predicted by Mfold

3.3.7 Binding affinity studies of aptamers

3.3.7.1 Circular dichroism studies

As a preliminary binding analysis of the aptamer-target complex, circular dichroism spectrum of the two aptamers in the presence of the target peptide was recorded. When the aptamers were allowed to bind with the peptide P16, the ellipticity of both the aptamers AG1 and AG2 were markedly increased as evident from the increase in peak intensity, confirming the conclusive binding functions of the aptamers with their target. The aptamer upon binding to the target acquires an induced CD (ICD) spectrum through electron rearrangements and these ICD spectra

can be used to confirm the aptamer-target binding interaction. Thus both the aptamers AG1 and AG2 had a binding interaction with the target peptide as shown by their respective CD spectra for both the aptamers (Fig 3.7).



A **B**

Fig. 3.7 Binding study of the aptamer-peptide complex by circular dichroism, (A) the AG1-peptide complex and (B) the AG2-peptide complex shows marked increase in ellipticity of the CD spectra on binding to its target peptide

3.3.7.2 Isothermal Titration Calorimetry (ITC) studies

After confirming the binding event between the aptamer and its target peptide by circular dichroism, we further analyzed the thermodynamic parameters and binding affinity by homogenous aptamer-target binding interaction using isothermal titration calorimetry. The dissociation constant of the binding event was also determined from the isotherm. ITC was carried out in the same conditions as

SELEX. ITC study elaborates the binding isotherm of the aptamer-peptide complex. The molar ratio versus heat exchange during the specific binding event of the aptamer AG1 with P16 is presented in a saturation curve (Fig.3.8). The amount of power required to maintain the reaction cell at constant temperature after each injection was monitored as a function of time. Hence, each injection generated a heat-burst curve (ucal/sec) versus time (min). The isotherm for AG1-P16 complex was fitted to one-site binding model and further by sigmoidal fitting in Origin 5.0 software (Microcal, Inc). ITC studies between AG1 and P16 revealed a dissociation constant (Kd) value of ~1.54μM. The ITC results for the aptamer AG2 did not give a considerable dissociation constant and was not considered for further analysis (Fig 3.9, A). ITC was also used to study the specificity of the AG1 aptamer, by allowing it to bind with the related peptide P17, which gave no significant heat change suggesting that the aptamer AG1 is specific towards its target (Fig 3.9, B). The binding mechanism between AG1 and its target P16 reveals a favorable enthalpy of reaction (ΔH) and unfavorable entropy (ΔS) at 25°C as shown in Fig 3.8. These thermodynamic data reveal a local conformational change caused by hydrophobic effect which probably drives the binding process. The peptide P16 has a hydrophobic nature, and when a hydrophobic molecule is involved in an interaction, hydration becomes entropically unfavorable, leading to order in the system with the price of an unfavorable entropy which drives the reaction (Amaya-González *et al.*, 2015).

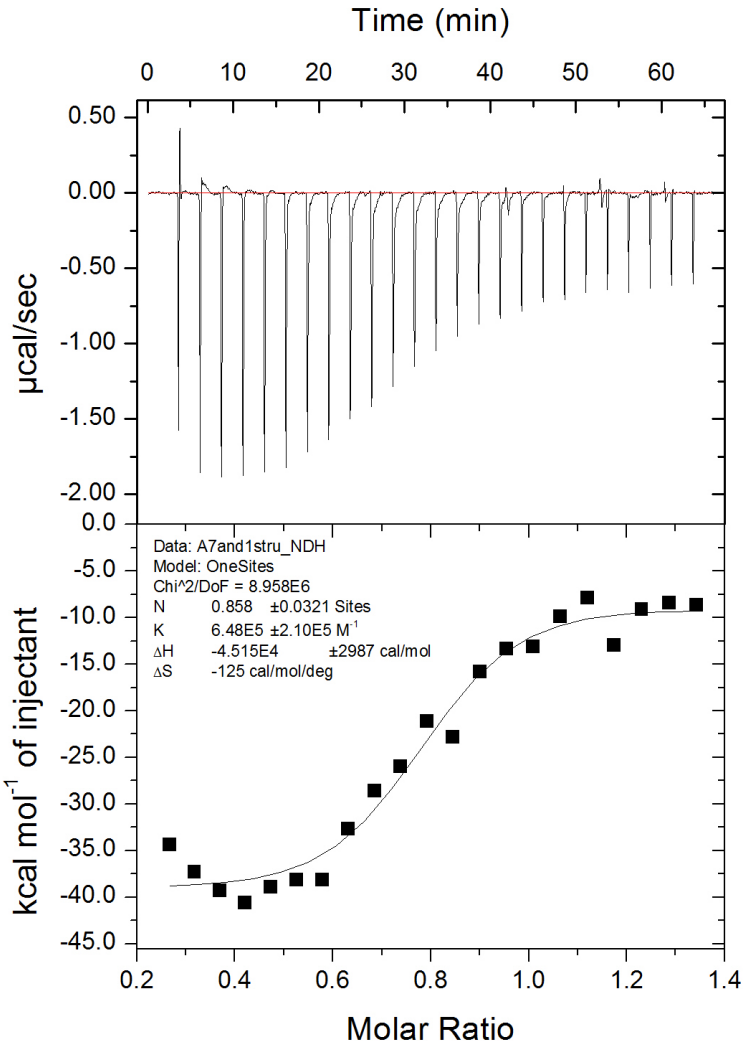


Fig. 3.8 ITC isotherms for AG1 aptamer-peptide complex; binding reaction was carried out with 30 µM aptamer and 900 µM target peptide at 25°C in aptamer binding buffer. Top image shows power as a function of time and bottom image depicts the heat change per injection fitted to a single-site binding model for both the aptamers. Kd value of 1.54 µM was calculated for the aptamer AG1

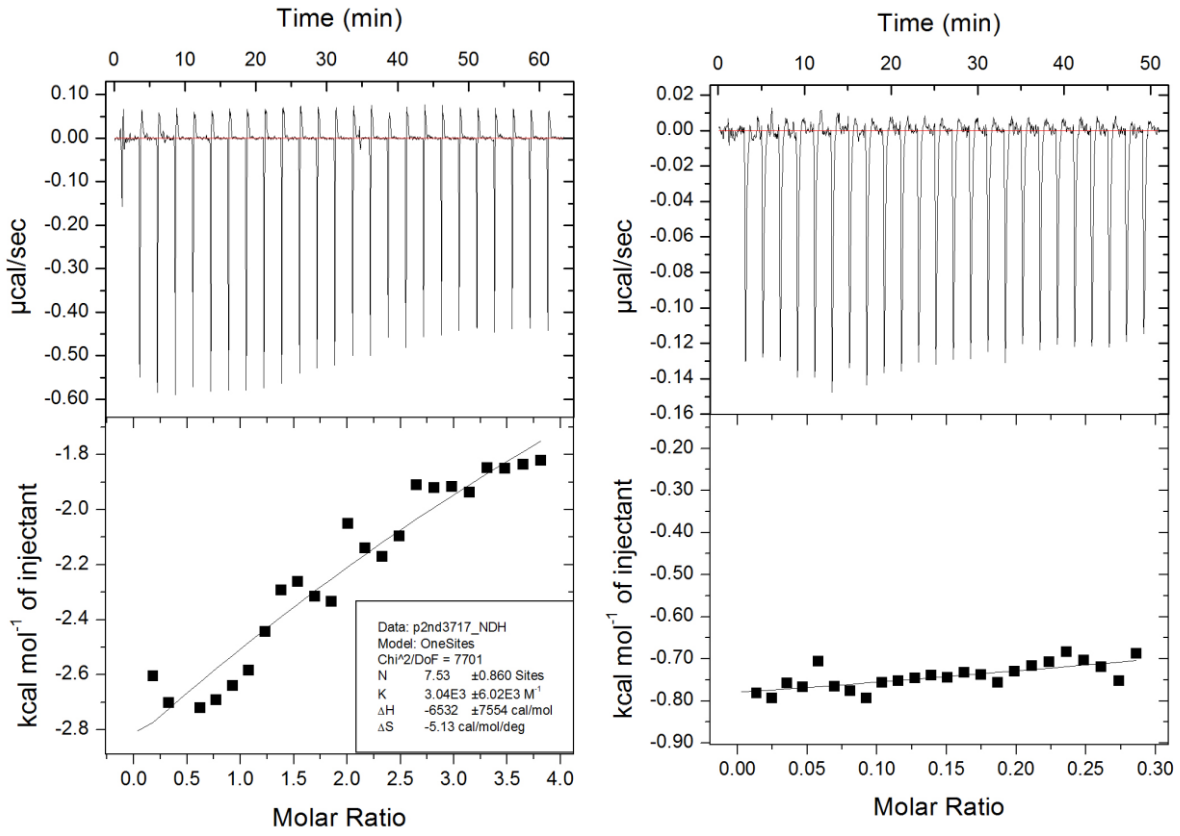


Fig. 3.9 ITC isotherms for (A) apatmer AG2 with target peptide P16, and (B) aptamer AG1 with the related peptide P17

3.3.7.3 Electrophoretic mobility shift assay (EMSA) study

Electrophoretic mobility shift assay (EMSA) was also done to further confirm the formation of aptamer AG1-peptide complex using 15% Native PAGE (29:1) in TBE (Tris-borate EDTA) gel for 1.5-2 hour at 180 V and 4°C (Fig 3.8). The aptamer-peptide complex was studied with an increasing concentration of peptide P16 (0, 25µM, 50µM, 100µM and 150µM). The EMSA results displayed a clear band-shift of the bound complex with the increasing concentration of the peptide compared with

the band position of only aptamer AG1 lane. Thus the EMSA result further validated the binding event between the aptamer AG1 and peptide P16.

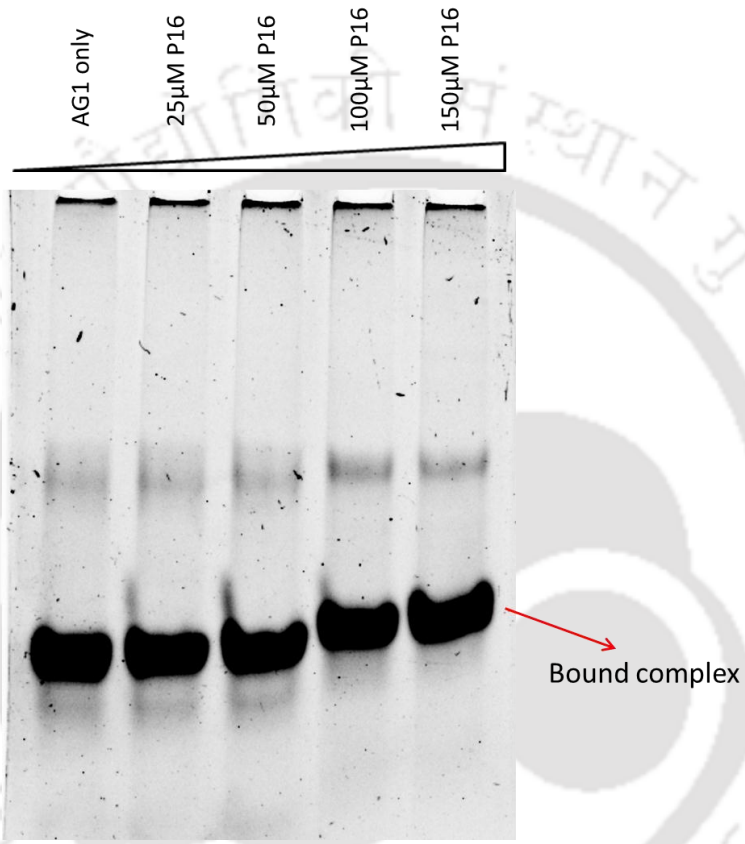


Fig. 3.10 Electrophoretic mobility shift assay (EMSA) for binding interaction study of the apatmer AG1-peptide complex with increasing concentration of the peptide

3.4 Discussions

Nucleic acid aptamers are highly specific ligands binding to their targets with high binding affinity. They are obtained from a random sequence pool of RNA or DNA against a target by SELEX. Successful selection of aptamer against a target is

dependent on many factors like the design of the library, target properties and the selection conditions. Before starting any SELEX experiment the aptamer library has to be judiciously selected so that it provides maximum diversity given by the length of the random region. The central random region length detremines the abundance and likely complexity of the selected structures (Knight, Birmingham & Yarus, 2004). The complex secondary structures like hairpins, loops, pseudoknots, G-quadruplex etc are associated with the binding affinity of the aptamers towards the target. Hence, longer random region imparts greater diversity of sequences resulting in complexity of the secondary structures providing better opportunities for the identification of high affinity aptamers (Marshall & Ellington, 2000). In the study we have selected a DNA aptamer library containing 40 mer long central random region was selected with theoretical diversity of 10^{14}-10^{15} sequences. The DNA library was choosen over RNA because it is more stable and cheaper.

The second important point is the target properties which govern the success of a selection process. It is well understood from the literature studies on aptamers that, not all targets yied succesful aptamers. In the present study, the target peptides of α-gliadin proteins are found to be hydrophobic in nature and these are less likely for a successful SELEX. Herein a charge-balanced peptide was selected for aptamer selection. A successful generation of specific DNA aptamer against the wheat α-gliadin peptide PFPQPQLPYPQPQLYP after 16 rounds of SELEX cycle is presented. Two highly enriched sequences AG1 and AG2 were obtained in the aptamer pool. Secondary structure prediction of the selected aptamer candidates by Mfold showed compex structure and was confirmed by CD studies. The QGRS

mapper predicts the capability of the aptamer AG1 to form G-quadruplex stucture. G-quadruplex forming ability indicates that the aptamer AG1 has higher thermal stability.

The CD studies for both the aptamers AG1 and AG2 with their target peptide have shown an increase in the ellipticity of the CD spectra following the binding event with the target confirming the conclusive binding functions in both the aptamers. The ITC results for AG1-target complex deduced a Kd value of 1.54µM which shows the high binding affinity of the aptamer AG1 with its target. The homogenous aptamer-target interaction from ITC helped us to understand the dynamics of the interaction involving a hydrophobic target. The highly unfavourable entropy of reaction drives the binding process of aptamer-target complex by bringing order in the system with the penalty of an unfavourable entropy. The ITC binding isotherms for aptamer AG2 did not give any conclusive results and hence was not considered further in this study. An ITC experiment was also conducted to determine the specificity of the aptamer AG1 with the related peptide P17 (PQPFRPQQPYPQSQPQY) which showed no significant enthalpy change suggesting that the aptamer AG1 is highly specific to its target. Finally, the EMSA results further ascertained the formation of the binding-complex. There is a clear shift in the band of the bound complex with the increasing concentration of the peptide. With increase in peptide concentration more and more amount of aptamer was bound with the peptide offering resistance in electrophoretic movement and hence shift in band. The EMSA experiment was repeated several times and everytime the same pattern was obtained in the gel. Hence, it is concluded that the shift in the

band observed with the increasing concentration of the peptide is beacause of the increasing population of the bound complex offering resistance in the electrophoretic movement and hence showing a band shift.

3.5 Conclusion

This chapter reports successful development of aptamers against the α-gliadin peptide P16. This study describes the generation of two parallelly evolved novel aptamers AG1 and AG2 with highly variable sequence and structure from the same DNA pool. Though both the aptamers bind to their target peptide, AG1 offers high binding affinity towards its target as revealed from the ITC study. The AG1 aptamer offered a dissociation constant (Kd value) of 1.54μM for its target. Hence, it can be concluded that this novel aptamer AG1, is a prospective candidate for the development of aptamer based gluten detection system.

4

Selection and characterization of DNA aptamers against the α-gliadin peptide PQPFRPQQPYPQSQPQY

4.1 Introduction and overview

As discussed earlier, gluten detection has become important in the present scenario because of the increasing cases of celiac disease, not only in Europe and the USA but to a great extent in the North-Indian population (Ramakrishna *et al.*, 2016). For a developing nation like India, screening and detection of gluten with antibody based platforms becomes an expensive affair. Also antibodies are having various inherent disadvantages, which pave the ways for exploring new technologies like aptamer based detection systems. Aptamers with many advantages over antibodies are a potential biorecognition platform. Studies on aptamer based gluten detection are being carried out by a research group in Spain, and apart from them, no such study is being carried out elsewhere till now. So, with India and other parts of Asia with increasing prevalence of celiac disease cases, needs a sensitive and specific monitoring of gluten levels food products and drugs. Here, aptamer based detection offers a sensitive way of detecting gluten comparable to antibodies, with the added advantage of low cost of production and with the ease of fabrication into a biosensing device.

In this chapter we describe the *in vitro* selection of DNA Aptamers against another synthetic wheat α-gliadin peptide PQPFRPQQPYPQSQPQY (P17) (86-102 amino acids and molecular weight 2086.26 Daltons) containing one immunogenic DQ2.5 restricted epitope glia-α3 (glia-α20) has been reported for the first time. Specificity studies of antibodies have shown that the antibody developed against the

epitope also reacts with gluten fractions from other cereal families (Mitea *et al.,* 2008). Also the sequence motif PQQPYP present in the glia-α3 epitope has homologous sequences in other fractions of gliadin protein. This shows that the gluten proteins share a high degree of homology. Hence, this peptide sequence was selected as a potential target for the development of aptamers which in turn help in the identification of as many gluten epitopes present in the related proteins (hordeins, secalins and avenins). With this goal in mind, i*n vitro* SELEX was carried out involving magnetic bead based SELEX using the peptide P17 synthesized and purified in lab as described in chapter 2. The selection process identified four aptamer candidates that were highly enriched with the target peptide. The secondary structure characterization studies revealed different structural composition of the aptamers. Moreover, the binding studies revealed the strong binding affinity of the aptamers AG75 and AG122 towards its target.

4.2 Materials and Methods

4.2.1 Chemicals, reagents and kits

Primers and single-stranded DNA library (aptamer library) were procured from Integrated DNA Technologies (USA). Dynabeads M270 carboxylic acid, SYBR Gold nucleic acid stain, Electrophoretic Mobility Shift Assay (EMSA) kit and InsTA clone PCR cloning kit were purchased from Invitrogen Thermo Scientific (USA). Streptavidin magnetic beads and Low molecular weight (LMW) DNA ladder were purchased from New England Biolabs (USA). IPTG (Isopropyl *β*-D-

thiogalactopyranoside) and X-Gal (5-bromo-4-chloro-3-indolyl-*β*-d-galactopyranopside) were purchased from Himedia (India). Miniprep plasmid isolation kit and Agarose was purchased from Sigma Aldrich (India). All other chemicals were analytical grade and cited wherever necessary.

4.2.2 Methods

4.2.2.1 *In vitro* selection of DNA aptamers against P17 by SELEX

Immobilisation of the peptide P17 on Dynabeads was done as discussed in chapter 3. To start with the SELEX cycle, the combinatorial N40 library (**5'-**ATACCAGTCTATTCAATT-N40-AGATAGTATGTGCAATCA-3') with a theoretical diversity of 10^{14}-10^{15} different sequences was used as the starting DNA pool. A total of 2 nanomoles of the starting DNA pool was taken from the library stock and was dissolved in 500 µl of SELEX binding buffer. The pool was subsequently activated by heating at 95°C for 8 min, immediately cooled in ice for 8-10 min and then brought to room temperature for 10 min. Before starting the positive SELEX rounds, firstly a negative SELEX cycle was carried out as discussed in chapter 3. The unbound DNA pool from the negative SELEX cycle was incubated with P17 peptide immobilized Dynabeads. After the completion of the binding round, the unbound DNA pool was removed by washing with SELEX washing buffer and the bound sequences were heat eluted and amplified using unmodified forward primer, F1 and 5'-biotin conjugated reverse primer, R2: 5'-biotin-TGATTGCACATACTATCT-3'. Single strand separation of the PCR product was done using streptavidin magnetic beads as

discussed earlier in chapter 3. The single-stranded DNA pool was then used for the following rounds of selection. The cycle of binding, separation and amplification is carried out until an enriched pool of DNA is obtained as monitored in a 2% agarose gel (Fig 4.1). A total of 14 iterative SELEX cycles were carried out to obtain the final enriched pool of aptamers. A total of 4 negative SELEX rounds with only Dynabeads without peptide immobilization were carried out to eliminate non-specific DNA sequences. The stringency of the SELEX rounds was increased by decreasing the binding time, increasing the washing cycles and decreasing the bead volume. Table 4.1 describes in detail the SELEX conditions used for the DNA aptamer selection against peptide P17.

Table 4.1 SELEX conditions used in selection of aptamer against peptide P17

Rounds	Beads (in µl) + Target (pmoles)	Binding Time (in min)	Washing Steps	Negative SELEX
1st	100 µl, 2100 pmoles	60	1 x 200 ul	Yes
2nd	100 µl, 2100 pmoles	60	1 x 200 ul	---
3rd	100 µl, 2100 pmoles	60	2x 200 ul	---
4th	100 µl, 2100 pmoles	60	2x 200 ul	---
5th	100 µl, 2100 pmoles	60	2x 200 ul	Yes
6th	100 µl, 2100 pmoles	60	2x 200 ul	---
7th	100 µl, 2100 pmoles	60	2x 200 ul	---
8th	100 µl, 2100 pmoles	60	2x 200 ul	Yes
9th	100 µl, 2100 pmoles	60	2x 200 ul	---
10th	100 µl, 2100 pmoles	60	3x 200 ul	---
11th	100 µl, 1050 pmoles	30	3x 200 ul	Yes
12th	100 µl, 1050 pmoles	30	3x 200 ul	---
13th	50 µl, 1050 pmoles	30	3x 200 ul	---

14th	50 µl, 1050 pmoles	15	3x 200 ul	---

4.2.2.2 Cloning and sequencing of enriched aptamers

The screened aptamer pool after the 14th round of SELEX was PCR amplified with unmodified forward and reverse primers (F1 and R1 primers) and purified by electro-elution as discussed earlier in chapter 3. The purified DNA was cloned into specialized vector pTZ57R/T by InsTA Clone PCR Cloning kit and transformed into competent DH5α cells. Transformed clones were selected on the basis of blue/white screening. 130 white colonies were randomly picked, grown in LB media and plasmid isolated by miniprep plasmid isolation kit. The isolated plasmid DNA was checked for the presence of aptamer candidates by PCR amplification using specific primers. Subsequently, the positive plasmid DNA having the correct size aptamer insert was sequenced by Sanger's sequencing using M13 forward and reverse primers. The obtained aptamer sequences were then aligned using Clustal Omega.

4.2.2.3 *In silico* structure analysis of the enriched candidates

In silico secondary structure predictions and thermodynamic properties of the enriched sequences were analyzed using Mfold web-server. The ionic conditions for Mfold study were kept same as the SELEX binding buffer with temperature fixed at 25°C. G-quadruplex mapping of the enriched sequences were done by the online QGRS (Quadruplex forming G-rich sequences) mapper software. The parameters

were set at maximum QGRS length of 40 with G-group number and loop size kept as default (i.e. minimum G-group of 2 and loop size between 0 and 36 nucleotides).

4.2.2.4 Structural analyses of aptamers by circular dichroism

Structural integrity of the aptamers was studied using circular dichroism (CD). For structural characterization a total of 4 µM of the candidate aptamers were heat activated and cooled to room temperature. Readings were then recorded in a Jasco J-1500 spectropolarimeter at 25°C. Far UV scanning were obtained in the range 200-340nm, with an average of 4 scans at a rate of 100nm/min, bandwidth 1nm, wavelength step size of 0.5nm and response of 2 seconds. The baseline was corrected and solvent contributions were subtracted from the spectra and smoothed by Savitsky-Golay algorithm for all the experiments.

4.2.2.5 Binding affinity Studies of aptamers

4.2.2.5.1 Circular Dichroism Study

CD spectra of the aptamer-peptide complex were recorded to determine the binding interaction between them. A total of 4µM of aptamer was heat activated and cooled to room temperature followed by addition of 20µM of the target peptide and incubating for 1 hour prior to recording of the CD spectra. The spectra were recorded with the same parameters as mentioned in section 4.2.2.4. The baseline was

corrected and solvent contributions were subtracted and smoothed by Savitsky-Golay algorithm.

4.2.2.5.2 Isothermal Titration Calorimetry (ITC) study

Isothermal titration calorimetry (ITC) studies were carried out to study the thermodynamic parameters and binding affinity of the aptamers. Calorimetry experiments were performed at 25°C using a MicroCal iTC200 (GE Healthcare, UK). ITC studies for the aptamers were carried out with a total of 30µM aptamer in SELEX binding buffer was heat activated and cooled to room temperature prior to the ITC experiment. 350 µl of the aptamer (~12000 Daltons) was loaded into the calorimetric cell of the instrument at a concentration of 30 µM and titrated by adding 1 µL-aliquots from the syringe stirring at 180 rpm, where the peptide P17 (2086.26 Daltons) was loaded at a concentration of 900 µM. The time interval between each injection was 120 seconds and a total of 30 injections were set to attain saturation. The volume of the first injection was 0.4 µl and was not considered in the fitting. Control experiment without the aptamer in the cell was performed to correct the heat of dilution for all the experiments.

4.2.2.5.3 Electrophoretic mobility shift assay (EMSA)

To further ascertain the binding of the aptamer-peptide complex EMSA was conducted in the presence of the target peptide. A fixed amount of 1 nM of the aptamer was heated at 95°C for 8 min, cooled in ice for 8 min and equilibrated to

room temperature followed by the addition of increasing concentration of the target peptide (P17: 0, 10 µM, 20 µM, 30 µM, 40 µM, 60 µM, 80 µM) in a total of 25 µl reaction volume. The mixture was incubated at room temperature for one hour and then separated using 15% native (29:1) TBE (Tris-borate EDTA) gel for about 1.5 to 2 hour at 180 V and 4°C. The gels were subsequently stained with SYBR Gold dye and visualized under UV. A control lane with only aptamer was also loaded for reference.

4.3 Results

4.3.1 *In vitro* selection of DNA aptamers against P17 by SELEX

Total 14 SELEX cycles containing 4 rounds of negative selection was carried out to screen potential aptamer candidates. The SELEX conditions were chosen such that it results in enrichment of the best binders. The conditions were relaxed at the start of the process with increase in stringency towards the end. A gradual enrichment of the aptamer pool was seen with the progressive cycles of SELEX as visualized in 2% agarose gel electrophoresis (Fig 4.1). The 14^{th} round of enriched candidates were PCR amplified, cloned and sequenced.

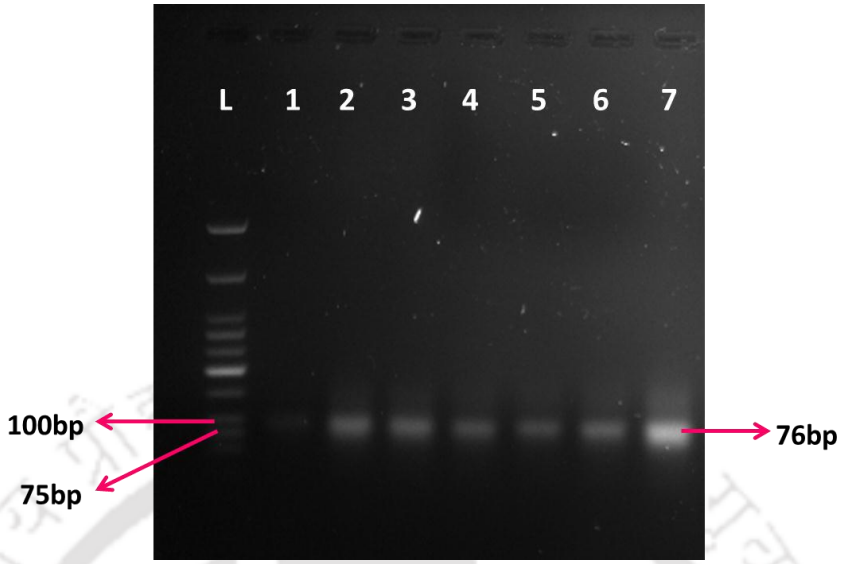


Fig. 4.1 A 2% agarose gel electrophoresis picture of the PCR products showing the enrichment of the aptamer pool in various SELEX rounds against peptide P17 (Lane L: ladder, Lane 1-7 shows the 2nd, 4th, 6th, 8th, 10th , 12th and 14th round aptamer pool)

4.3.2 Cloning and sequencing of enriched aptamers

Total 130 white colonies were screened, plasmid DNA isolated and amplified by F1 and R1 unmodified primers. Amplification revealed 23 positive clones carrying the aptamer inserts. The clones were further sequenced to reveal 18 sequences with the desired aptamer sequence out of which four candidates AG5, AG75, AG110 and AG122 were highly enriched (Table 4.2). The aptamer AG122 had the highest copy number followed by AG75. The alignment studies of the enriched aptamers did not reveal any strong conserved regions. The sequence similarity among the enriched candidates was shown as a neighbor joining tree (Fig 4.2).

Table 4.2 Enriched aptamer sequences for target peptide P17 after 14 rounds of SELEX

Aptamer name	Sequence
AG5	ATCAACAAAAAAAGGTTGTATAGTCCTCGGTTAATAGGCA
AG75	ACGGTGCACTAGACTGAGGCGATGCCGGGTGGGCTCCACT
AG110	AGTGGTGATACAGTACGTAGACAGCCTATTGATCCGTAGT
AG122	TCGAGGTTACGAGTGGGGGCTGCAGTTTCTCCCCGCGGAG

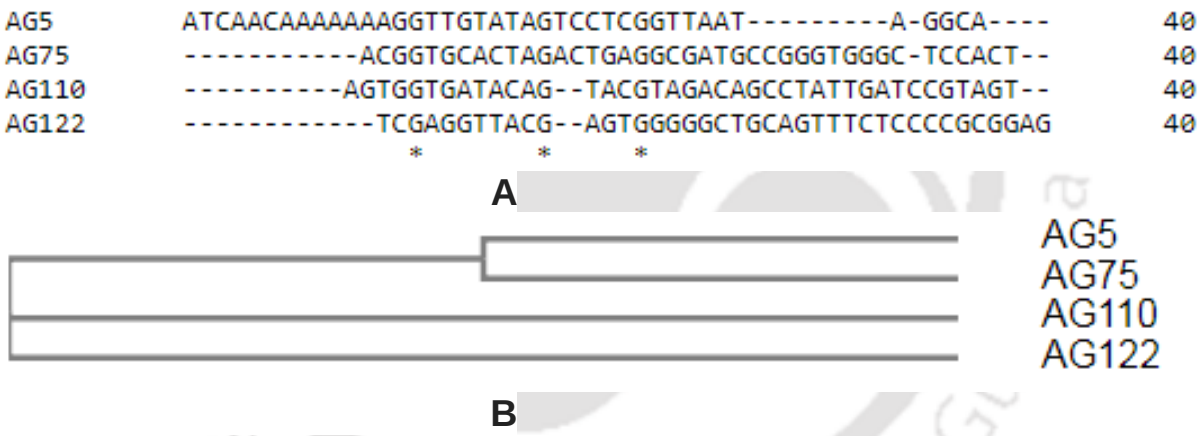
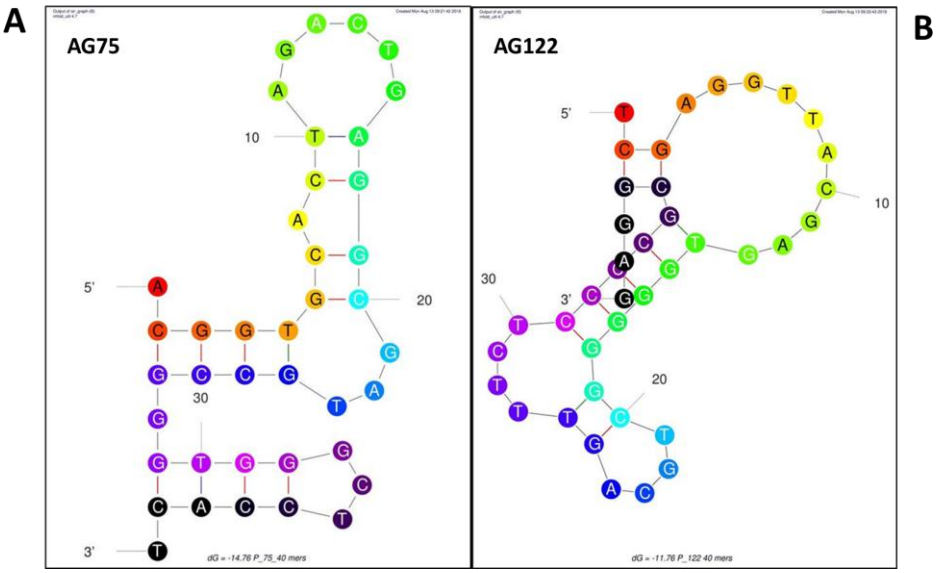


Fig. 4.2 Alignment studies, (A) sequence alignment of the random region of AG5, AG75, AG110 and AG122 by Clustal Omega and (B) neighbor joining tree for the aligned sequences

4.3.3 *In silico* structure analysis of the enriched candidates

Secondary structures of the aptamers were predicted using the Mfold. The δG values for AG75, AG122, AF110 and AG5 were deduced as -14.76, -11.76, -10.57

and -9.03 Kcal mol^{-1}, respectively. Other thermodynamic parameters are depicted in Table 4.3. The structure of the aptamer AG75 forms three hairpin like structure, AG122 forms one bigger stem-loop structure with an extended arm joining two small loops, AG110 forms a similar structure like AG75 with three stem-loop structures and AG5 forms two stem-loop structure joined by a long arm (Fig 4.3). Visual analyses of the aptamers reveal almost similar secondary structures for all the candidates. Further analysis of the sequences by QGRS Mapper software revealed the formation of G-quadruplex structure in AG75 and AG122 (Table 4.4). The lower δG values and the G-quadruplex forming capability of the aptamers AG75 and AG122 reveal that they acquire higher thermal stability compared to AG5 and AG110. Moreover, G-quadruplex DNAs are prospective candidates for a good binder with the added advantage of protection from nucleases. Hence, the aptamers AG75 and AG122 were undertaken for further binding studies.



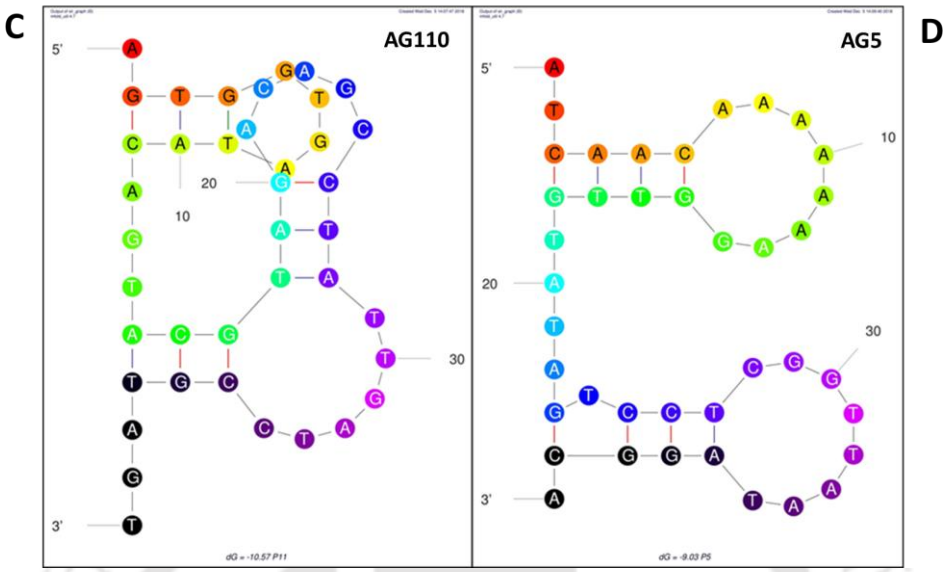


Fig. 4.3 Secondary structure prediction of the enriched aptamers by Mfold, (A) aptamer AG75 form three hairpin like structure; (B) AG122 forms one bigger stem-loop structure with an extended arm joining two small loops; (C) AG110 forms a similar structure like AG75 with three stem-loop structures; and (D) AG5 forms two stem-loop structure joined by a long arm

Table 4.3 Thermodynamic profile of enriched sequences as deduced by Mfold

Aptamer name	Free Gibbs energy value (ΔG) (kcal/mol, 25˚C)	Enthalpy change (ΔH) (kcal/mol)	Entropy change (ΔS) (cal/K.mol)
AG75	-14.76	-89.10	-249.3
AG122	-11.76	-52.60	-136.9
AG110	-10.57	-60.40	-167.1
AG5	-9.03	-74.10	-218.2

Table 4.4 QGRS prediction of candidates AG75 and AG122 against peptide P17

Aptamer name	Position	Length	QGRS	G-score
AG75	3	31	GGTGCACTAGACTGAGG CGATGCCGGGTGGG	21
AG122	5	34	GGTTACGAGTGGGGGCT GCAGTTTCTCCCCGCGG	15

4.3.4 Structural analyses of aptamers by circular dichroism

To confirm the structure of the aptamers as predicted by Mfold and QGRS Mapper softwares the circular dichroism spectra of both the aptamers were studied in deionized water and SELEX binding buffer. Peaks corresponding to duplex B-form DNA with CD maxima near 280 nm and a negative peak near 245 nm was seen for both AG75 and AG122 (Fig. 4.4, A and B) suggesting the correct hybridization of the aptamer to form double helical regions as deduced by Mfold predictions. Moreover, the signature peaks corresponding to a G-quadruplex structure was not seen for both the peptide sequences.

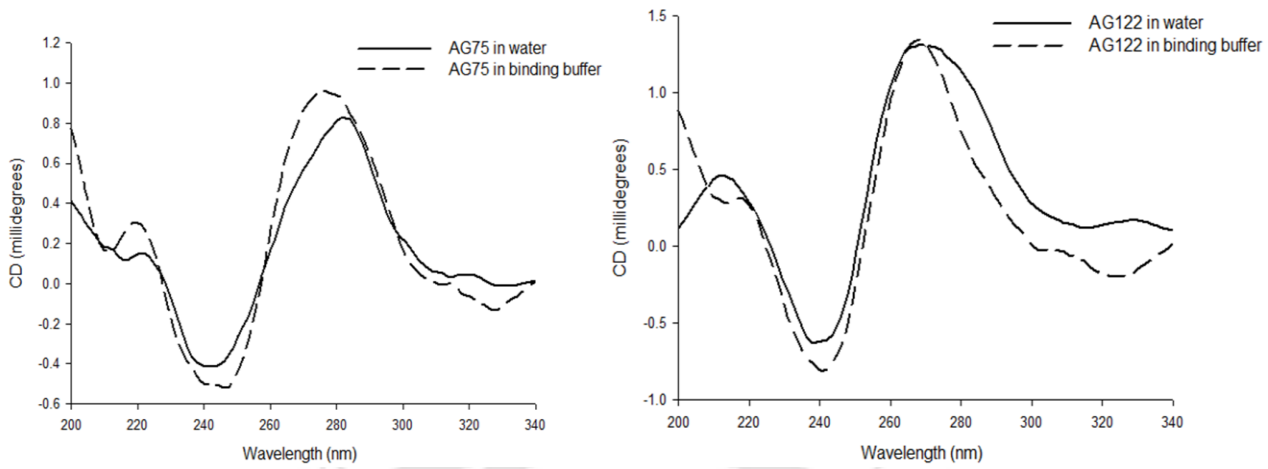


Fig. 4.4 Circular dichroism spectra of aptamers AG75 and AG122 in water and SELEX binding buffer; CD spectra of both (A) AG75 and (B) AG122 shows positive peak near 280nm and a negative peak near 245 nm

4.3.5 Binding affinity studies of aptamers

4.3.5.1 Circular dichroism study

As a preliminary binding analysis and to study the conformations of the aptamers-target complex, circular dichroism spectrum of the two aptamers after incubation with the target peptide was recorded. The spectra for AG75 aptamer-target complex show marked increase in ellipticity of the CD spectra compared to spectra of AG75 aptamer alone, as evident from the increase in peak intensity (Fig. 4.5, A). This is due to induced CD spectrum occurring as a result of the aptamer binding to its target, confirming the conclusive binding function of the aptamer with

the target peptide. The CD spectra for the aptamer AG122-peptide complex did not show the increase in ellipticity of the CD spectra on binding to its target (Fig 4.5, B).

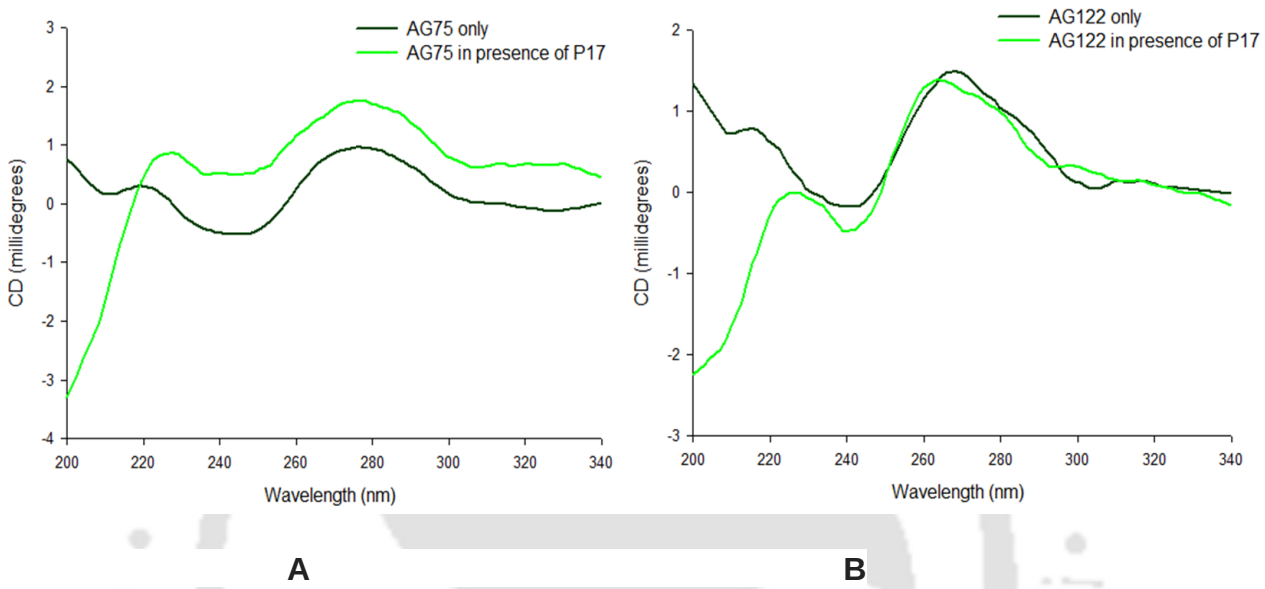
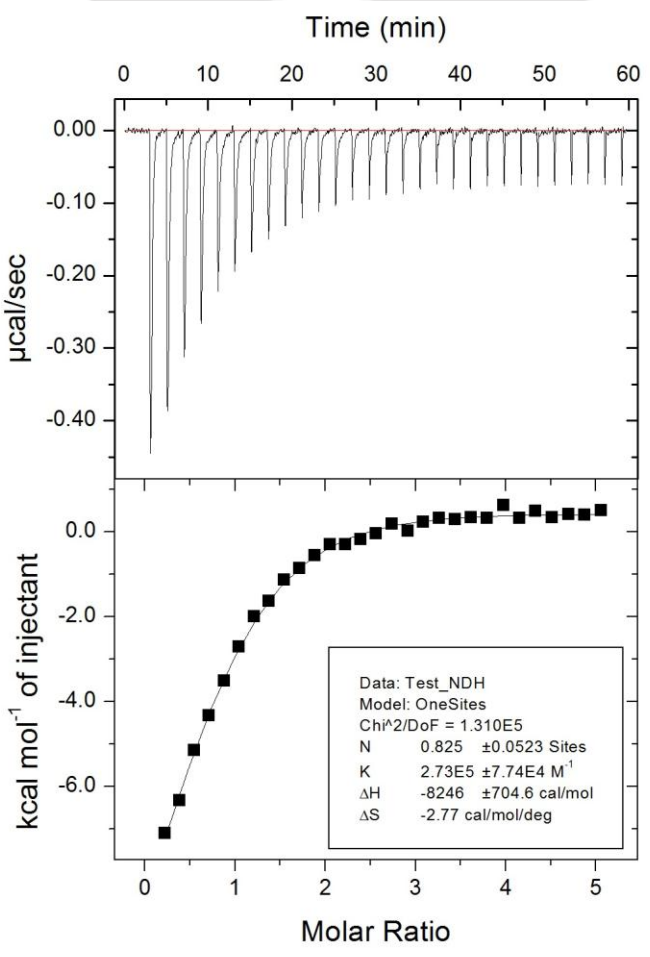


Fig. 4.5 Binding study of the aptamer-peptide complex by circular dichroism; (A) the AG75-peptide complex shows marked increase in ellipticity of the CD spectra on binding to its target peptide, while (B) AG122-peptide complex did not

4.3.5.2 Isothermal Titration Calorimetry (ITC) study

Isothermal titration calorimetry studies were performed to confirm the the binding of the aptamer-target complex and deduce the binding constant. The buffer conditions used in the ITC experiment was maintained same as SELEX buffer conditions. The molar ratio versus heat exchanged during the specific binding event of the aptamers with their target was presented in a saturation curve (Fig.4.6). Power required in maintaining the reaction cell at a fixed temperature after every injection is

monitored as a function of time. Hence, each injection generated a heat-burst curve (ucal/sec) versus time (min). Isotherms were fitted to one-site binding model. ITC studies of aptamers AG75 and AG122 with target peptide P17 revealed a dissociation constant (Kd value) of 3.7 µM and 10.9 µM respectively. The binding mechanism reveals a favorable enthalpy of reaction (ΔH) and unfavorable entropy (ΔS) for AG75. Thus, the binding of AG75-P17 complex is entropy driven reaction as discussed in chapter 3. However, the isotherm for AG122 reveals a highly favorable enthalpy and entropy of reaction. This is mostly because of the local conformational change in the aptamers due to target binding (Pagano, Mattia & Giancola, 2009).



A

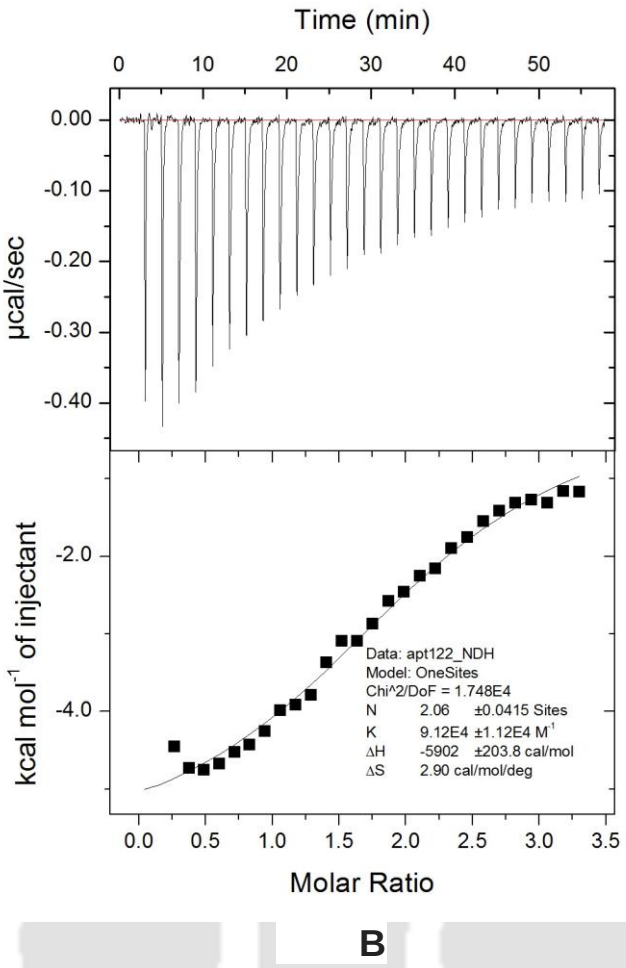


Fig. 4.6 ITC isotherms for (A) AG75 and (B) AG122 aptamer-peptide complex; binding reaction was carried out with 30µM aptamer and 900µM target peptide at 25°C in aptamer binding buffer (top image shows power as a function of time and lower image depicts the change in heat per injection, the Kd values for AG75 and AG122 are 3.7µM and 10.9µM,respectively)

4.3.5.3 Electrophoretic mobility shift assay (EMSA)

EMSA was also done to further confirm the formation of aptamer-peptide complex using 15% Native PAGE (29:1 acrylamide/bisacrylamide) in TBE (Tris-

borate EDTA) gel for 1.5-2 hours at 180 V and 4°C. The AG75-P17 complex was studied with an increasing concentration of the peptide (0, 10µM, 20µM, 30µM, 40µM, 50µM, 60µM and 80µM) with fixed concentration of the aptamer. EMSA results displayed a clear band-shift of the aptamer-peptide complex (compared to only aptamer), with the increasing concentration of the peptide (Fig 4.7). The EMSA results for AG122 did not give conclusive results and therefore not considered in the study. Thus, the EMSA result for the aptamer AG75 further validated the binding event between the aptamer and peptide P17.

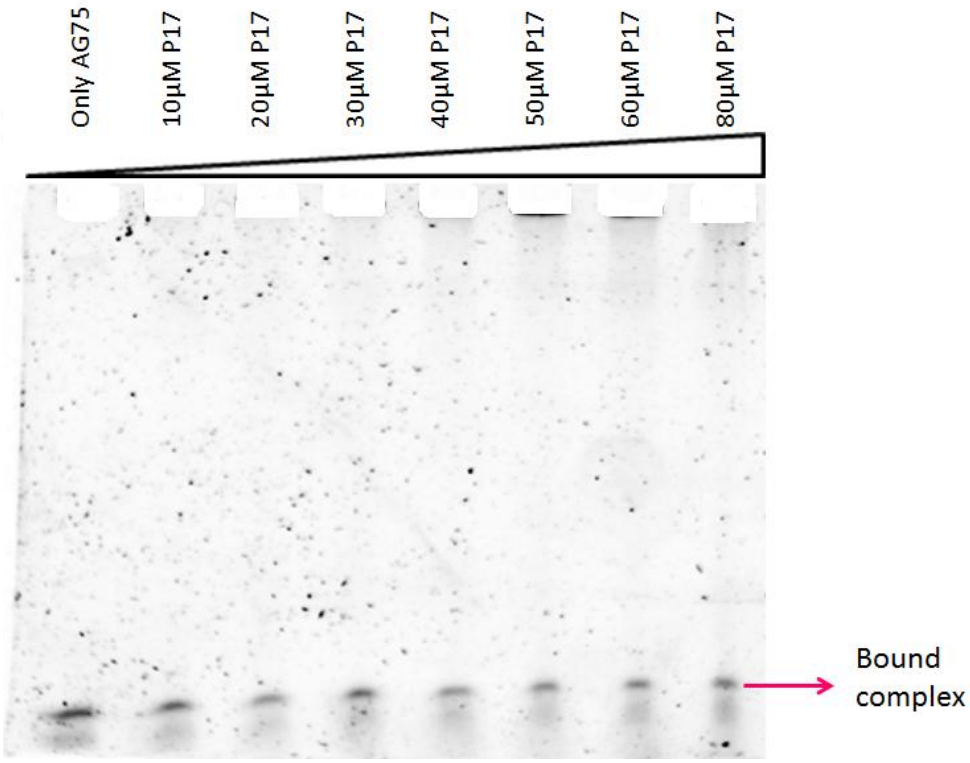


Fig. 4.7 Electrophoretic mobility shift assay (EMSA) for the binding interaction study of the apatmer AG75-peptide complex with increasing concentration of the peptide

4.4 Discussions

The goal of the present investigation was the development of high affinity aptamers against the peptide PQPFRPQQPYPQSQPQY (P17) containing an immunogenic epitope DQ2.5 glia-α3. The generated DNA aptamers were found to bind with their target peptide with micromolar (μM) sensitivity. Hence, we demonstrate a successful generation of highly sensitive aptamers against the wheat α-gliadin peptide.

The first objective of the chapter was to screen potential aptamers against the target P17 by SELEX. For this, an aptamer library with 40 mer random region offering a huge diversity of sequences as the initial pool, considered to be optimum for obtaining high affinity candidates was selected. The success of a SELEX process is highly dependent on the starting aptamer library, because a library with greater sequence diversity results in the selection of better candidate sequences with higher structural diversity and better binding affinity (Stoltenburg, Reinemann & Strehlitz, 2007). Secondly, a successful SELEX process is also dependent on the SELEX conditions which have to be optimised according to the target. Herein, we describe the conditions to be neither very stringent nor relaxed to allow for gradual enrichment of the best binders while removing the non-specific and loose binders. Finally, the target properties are also an important consideration for a successful SELEX. Hence, the peptide was selected in a way such that it is positively charged at the pH used for aptamer binding buffer, so that the DNA which is negatively charged offers

good binding affinity towards the target and help in the evolution of high affinity binders. The enrichment of the SELEX process was completed at the 14th cycle, which was monitored by 2% agarose gel electrophoresis. As the SELEX progress, the diversity of the aptamer pool keeps on decreasing and starts enriching the best candidates. The last round candidates are the strongest binders, which were then cloned and sequenced by Sanger's sequencing and aligned for sequence similarity among the obtained pool. Analysis of the sequencing data revealed that four sequences AG5, AG75, AG110 and AG122 were highly enriched with AG122 having the highest copy number. Alignment studies of the random region of the enriched aptamers did not reveal any strong conserved regions.

Further *in silico* analyses such as free energy comparison and structural stability was performed to determine the aptamers with best binding affinity towards the target. The candidates AG75 and AG122 with the most thermodynamically stable structure and the capability to form G-quadruplex structure was chosen and synthesized for further analysis. After obtaining the synthesized sequences the structural integrity and conformation of the aptamers were studied by circular dichroism which confirmed the findings of *in silico* analysis by Mfold. After confirming the structural conformations of AG75 and AG122, the preliminary binding analysis of these aptamers with P17 was also determined by CD. CD results demonstrated the formation of the aptamer-target binding complex in case of aptamer AG75, showing an increase in ellipticity of the CD spectra on binding to its target, while the aptamer AG122 did not form any bound complex with its target. These findings formed the basis for further analysis.

Next binding analysis was done by ITC experiments. ITC enumerates the binding stoichiometry with high precision, which comes from the fact that the analyses of the binding molecules are done in their native form in a homogenous system. Moreover the estimation of correct binding stoichiometry is highly significant in the characterization of the binding mechanism of the aptamer and its target. The thermodynamic parameters like ΔG, ΔH, ΔS is important in further understanding of the underlying molecular mechanisms of interaction. The ITC isotherm of the aptamer AG75 shows a favorable enthalpy of reaction and unfavorable entropy of reaction. So we can assume that the binding-interaction between AG75 and its target is an entropy driven reaction (similar to the aptamer AG1 against the peptide P16 discussed in chapter 3). The ITC isotherm for the other aptamer AG122 revealed a high enthalpy of reaction, which is driving the molecular interaction of the aptamer with the target. The thermodynamic data for AG122 indicates the local conformation change of the aptamer upon target binding, resulting in solvent rearrangement and high heat of enthalpy (Pagano *et al.*, 2009). The isotherms were fitted to a one-site binding model and the heat of dilution was subtracted for all the experiments. Thus, a high binding affinity of the aptamers AG75 and AG122 against P17 was confirmed.

A binding analyses based on EMSA was done for both the aptamers, but aptamer AG122 did not give much convincing results. The other aptamer AG75 target complex demonstrated better results. The EMSA technique helped in the better understanding of interaction of the binding complex depicting a shift in the molecular size bands of the bound complex with increasing concentration of the

ligand molecule. The peptide as a ligand is very small (~2kDa) and the aptamer is much bigger (~12kDa), so the shift in the band of the bound complex is very less, but can be clearly distinguished from the lane where only aptamer is present. Hence, the EMSA findings further validated the binding between the aptamer and the target peptide.

4.5 Conclusion

The chapter reports successful development of aptamers against the α-gliadin peptide P17. Four enriched DNA aptamers were generated against the peptide P17 following a systematic evolutionary approach and then finally validating the target binding capabilities of the enriched aptamers by binding affinity analyses. Two aptamers AG75 and AG122 were considered for further studies based on their stability as deduced by *in silico* and circular dichroism studies. The aptamer AG75 showed a higher binding affinity towards the target compared to AG122 as revealed from ITC study. The K_d values of 3.7 µM and 10.9 µM were calculated for AG75 and AG122, respectively. Hence, the developed aptamer AG75 can be considered as a potential candidate for developing a sensitive gluten detection platform offering a low cost biorecognition probe.

5

Summary and Future Prospects

5.1 Summary

The present study demonstrates the selection of aptamers specific to two wheat α-gliadin peptides PFPQPQLPYPQPQLPY (P16) and PQPFRPQQPYPQSQPQY (P17) containing toxic and immunogenic DQ2.5 restricted epitopes glia-α1a, glia-α1b, glia-α2 and glia-α3. Development of aptamers against these particular wheat α-gliadin peptides as a potential target for gluten detection was accomplished for the first time. Recently, a research group from Spain has demonstrated the development of two aptamers against the recombinant 33-mer α-gliadin peptide showing high binding affinity with Kd values in nanomolar range (González *et al*.,2014). Herein, we have described the development of few aptamers with binding affinity in the micromolar (μM) sensitivity range, which is acceptable for a smaller peptide molecule and could be developed into highly sensitive aptamer-based detection system for gluten assessment.

5.1.1 Selection and characterization of specific DNA aptamers against the α-gliadin peptide PFPQPQLPYPQPQLPY (P16)

A successful synthesis and purification of the target peptides by reverse-phase HPLC, followed by mass spectrometry and secondary structure analyses of the target peptide using circular dichroism spectroscopy was obtained. The secondary structure analysis of the peptide revealed mostly random structure in the presence of turns and lesser amount of beta-structure with insignificant or no helix

formation which is found to be similar to the earlier reported data. Next, the successful development of high affinity specific aptamer candidate against the purified target peptide P16 was achieved employing SELEX technology *in vitro*. *In vitro* selection of specific aptamers was followed by TA cloning, screening and sequencing of the enriched pool of candidates. After screening of 80 clones two highly enriched aptamer sequences in the DNA pool were obtained. The structure analyses of the candidate aptamers done by circular dichroism revealed the correct structural conformation of the aptamers. The interaction study done by CD also confirmed the binding between the aptamers and target peptides confirming the formation of the binding complex. The aptamers and the respective binding complexes as seen in the CD spectra exhibited a stable behavior. Also the CD spectra confirmed the aptamers fold into distinct stem loop structures as shown by Mfold predictions. CD studies were taken as a basis for further binding analysis by ITC and EMSA. The dissociation constant for the AG1 binding complex was deduced from the ITC studies, which validated the successful generation of highly efficient aptamer AG1 against the target peptide. The specificity study of the aptamer AG1 with the related peptide P17 by ITC did not show any significant heat change, confirming AG1 being highly specific towards its target peptide P16. The dissociation constant of the aptamer AG1 as deduced from ITC experiment was 1.54µM. The ITC studies were also further supported by EMSA which proved the formation of stable AG1-peptide complex in the gel.

Hence, the developed anti-gliadin DNA aptamer has the potential to be developed into a highly sensitive probe against gluten detection.

5.1.2 Selection and characterization of DNA aptamers against α-gliadin peptide PQPFRPQQPYPQSQPQY (P17)

In brief it is demonstrated that the aptamer candidates selected after 14 iterative rounds of SELEX against the second target peptide PQPFRPQQPYPQSQPQY, showed high binding affinity towards it. Two aptamers AG75 and AG122 were highly enriched in the pool as compared to the other candidates. They demonstrated higher thermodynamic and structural stability as predicted by Mfold and confirmed by CD studies. The binding studies of these aptamers by CD confirmed the formation of stable binding complex in case of AG75 and ITC studies also deduced a high binding affinity for the aptamer AG75 towards its target as given by the Kd value. Aptamers AG75 and AG122 showed dissociation constant of 3.7µM and 10.9µM, respectively. This binding data was further supported by EMSA study of the AG75 target complex. Hence, the successful development of high affinity target-binding aptamer AG75 against the peptide P17 was demonstrated.

The developed potential aptamer can also be exploited as a detection probe against gluten and its products. The aptamer AG75 is novel and the sequence of this aptamers don't show any similarity with the aptamers so far developed elsewhere by other groups working on aptamer development against any other targets of gluten. So, these novel aptamers are prospective candidates increasing the aptamer tool-box against gluten offering higher sensitivity of detection.

5.2 Future direction of work

The present study is believed to be the first step towards developing an aptamer-based detection system for gluten assessment. Hence, this study needs further validation and exploration on few important points to augment the work involving aptamer target interactions and their practical feasibility to fabricate them in a sensing device. Firstly, truncation studies can be performed to determine the optimum sequence length of these aptamers capable of binding to the target peptides. Secondly, validation studies based on enzyme-linked aptamer assays (ELAA) need to be performed using real samples (gluten containing cereals and foodstuffs) and compare it to the commercially available antibody-based immunoassays and kits. By doing this the sensitivities of both the methods could be compared. Thirdly, obtaining further information on the three-dimensional structure of the aptamers and aptamer-target complex may help in the understanding of its detailed interaction pattern with the hydrophobic target like gluten and the underlying mechanisms involved in binding. Also further research into the three dimensional structure of the alpha-gliadin protein has to be done to have a concrete experimental insight into the epitope structure of the alpha-gliadin protein. These insights will help in the modification of the SELEX technique so that aptamer candidates with higher binding affinities are obtained. Hence, these understandings will be vital for the effective detection and will be an important basis for next level of investigation.

In conclusion, it can be stated that the present study is a preliminary investigation towards the development of a highly sensitive platform for the detection of gluten, based on aptamers as detection probes. These aptamer-based probes can further be modified chemically to enhance their existing properties and can be easily fabricated into cheap and cost effective biosensing devices assuring health safety for celiac patients.

Bibliography

Abadie, V., Sollid, L. M., Barreiro, L. B., Jabri, B. (2011). Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annual Review of Immunology*, 29, 493–525.

Alves, T. O., D'Almeida, C. T., Ferreira, M. S. (2017). Determination of Gluten Peptides Associated with Celiac Disease by Mass Spectrometry. In *Celiac Disease and Non-Celiac Gluten Sensitivity*. IntechOpen.

Amaya-González, S., de-los-Santos-Álvarez, N., Miranda-Ordieres, A. J., Lobo-Castañón, M. J. (2014). Aptamer binding to celiac disease-triggering hydrophobic proteins: a sensitive gluten detection approach. *Analytical Chemistry,* 86(5), 2733–2739.

Amaya-González, S., de-los-Santos-Álvarez, N., Miranda-Ordieres, A. J., Lobo-Castañón, M. J. (2015). Sensitive gluten determination in gluten-free foods by an electrochemical aptamer-based assay. *Analytical and Bioanalytical Chemistry*, 407(20), 6021–6029.

Amaya-González, S., López-López, L., Miranda-Castro, R., de-los-Santos-Álvarez, N., Miranda-Ordieres, A. J., Lobo-Castañón, M. J. (2015). Affinity of aptamers binding 33-mer gliadin peptide and gluten proteins: Influence of immobilization and labeling tags. *Analytica Chimica Acta*, 873, 63–70.

Anderson, R. P., Degano, P., Godkin, A. J., Jewell, D. P., Hill, A. V. (2000). *In vivo* antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nature Medicine*, 6(3), 337.

Arentz-Hansen, H., Körner, R., Molberg, Ø., Quarsten, H., Vader, W., Kooy, Y.M., Lundin, K.E., Koning, F., Roepstorff, P., Sollid, L.M., McAdam, S.N. (2000). The

intestinal T cell response to α-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *Journal of Experimental Medicine,* 191(4), 603–612.

Arentz–Hansen, H., Mcadam, S. N., Molberg, Ø., Fleckenstein, B., Lundin, K. E., Jørgensen, T. J., Jung, G., Roepstorff, P., Sollid, L.M. (2002). Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterology*, 123(3), 803–809.

Asai, R., Nishimura, S. I., Aita, T., Takahashi, K. (2004). *In vitro* selection of DNA aptamers on chips using a method for generating point mutations. *Analytical Letters*, 37(4), 645–656.

Balakireva, A., Zamyatnin, A. (2016). Properties of gluten intolerance: gluten structure, evolution, pathogenicity and detoxification capabilities. *Nutrients,* 8(10), 644.

Barak, S., Mudgil, D., Khatkar, B. S. (2015). Biochemical and functional properties of wheat gliadins: a review. *Critical Reviews in Food Science and Nutrition*, 55(3), 357–368.

Bergseng, E., Sidney, J., Sette, A., Sollid, L. M. (2008). Analysis of the binding of gluten T-cell epitopes to various human leukocyte antigen class II molecules. *Human Immunology*, 69(2), 94–100.

Bhattacharya, M., Dubey, A. P., Mathur, N. B. (2009). Prevalence of celiac disease in North Indian children. *Indian Pediatrics*, 46(5), 415.

Burke, D. H., Willis, J. H. (1998). Recombination, RNA evolution, and bifunctional RNA molecules isolated through chimeric SELEX. *Rna*, 4(9), 1165–1175.

Camafeita, E., Alfonso, P., Mothes, T., Méndez, E. (1997). Matrix-assisted laser

desorption/ionization time-of-flight mass spectrometric micro-analysis: the first non-immunological alternative attempt to quantify gluten gliadins in food samples. *Journal of Mass Spectrometry*, 32(9), 940–947.

Cataldo, F., Montalto, G. (2007). Celiac disease in the developing countries: a new and challenging public health problem. *World Journal of Gastroenterology: WJG*, 13(15), 2153.

Chirdo, F. G., Añón, M. C., Fossati, C. A. (1995). Optimization of a competitive ELISA with polyclonal antibodies for quantification of prolamins in foods. *Food and Agricultural Immunology*, 7(4), 333–343.

Cho, C. Y., Nowatzke, W., Oliver, K., Garber, E. A. (2015). Multiplex detection of food allergens and gluten. *Analytical and Bioanalytical Chemistry*, 407(14), 4195–4206.

Cho, E. J., Lee, J.W., Ellington, A. D. (2009). Applications of aptamers as sensors. *Annual Review of Analytical Chemistry*, 2, 241–264.

Ciccocioppo, R., Di Sabatino, A., Corazza, G. R. (2005). The immune recognition of gluten in coeliac disease. *Clinical & Experimental Immunology*, 140(3), 408–416.

Coulter, L. R., Landree, M. A., Cooper, T. A. (1997). Identification of a new class of exonic splicing enhancers by in vivo selection. *Molecular and Cellular Biology,* 17(4), 2143–2150.

Czaja-Bulsa, G. (2015). Non coeliac gluten sensitivity–A new disease with gluten intolerance. *Clinical Nutrition*, 34(2), 189–194.

Darmostuk, M., Rimpelova, S., Gbelcova, H., Ruml, T. (2015). Current approaches in SELEX: An update to aptamer selection technology. *Biotechnology Advances*,

33(6), 1141–1161.

Don, C., Halbmayr-Jech, E., Rogers, A., Koehler, P. (2014). AACCI Approved Methods Technical Committee report: collaborative study on the immunochemical quantitation of intact gluten in rice flour and rice-based products using G12 sandwich ELISA. *Cereal Foods World*, 59(4), 187–193.

Doudna, J. A., Cech, T. R. (2002). The chemical repertoire of natural ribozymes. *Nature*, 418(6894), 222.

Dunn, M. R., Jimenez, R. M., Chaput, J. C. (2017). Analysis of aptamer discovery and technology. *Nature Reviews Chemistry*, 1(10), 76.

Elli, L., Branchi, F., Tomba, C., Villalta, D., Norsa, L., Ferretti, F., Roncoroni, L., Bardella, M.T. (2015). Diagnosis of gluten related disorders: Celiac disease, wheat allergy and non-celiac gluten sensitivity. *World Journal of Gastroenterology: WJG*, 21(23), 7110.

Ellington, A. D., Szostak, J. W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, 346(6287), 818.

Ellis, H. J., Rosen-Bronson, S., O'Reilly, N., Ciclitira, P. J. (1998). Measurement of gluten using a monoclonal antibody to a coeliac toxic peptide of A gliadin. *Gut*, 43(2), 190–195.

Fasano, A., Catassi, C. (2001). Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum. *Gastroenterology*, 120(3), 636–651.

Fraser, J. S., Engel, W., Ellis, H. J., Moodie, S. J., Pollock, E. L., Wieser, H., Ciclitira, P. J. (2003). Coeliac disease: *in vivo* toxicity of the putative immunodominant epitope. *Gut*, 52(12), 1698–1702.

Freedman, A. R., Galfre, G., Gal, E., Ellis, H. J., Ciclitira, P. J. (1987). Monoclonal

antibody ELISA to quantitate wheat gliadin contamination of gluten-free foods. *Journal of Immunological Methods*, 98(1), 123–127.

Freeman, H. J., Chopra, A., Clandinin, M. T., Thomson, A. B. (2011). Recent advances in celiac disease. *World Journal of Gastroenterology: WJG*, 17(18), 2259.

Friis, S. U. (1988). Enzyme-linked immunosorbent assay for quantitation of cereal proteins toxic in coeliac disease. *Clinica Chimica Acta*, 178(3), 261–270.

G Chirdo, F., Arranz, E. (2015). Cereal proteins: immunostimulatory and toxic peptides. *OmniaScience Monographs*.

García, E., Llorente, M., Hernando, A., Kieffer, R., Wieser, H., Mendez, E. (2005). Development of a general procedure for complete extraction of gliadins for heat processed and unheated foods. *European Journal of Gastroenterology & Hepatology*, *17*(5), 529–539.

Gold, L., Brown, D., He, Y. Y., Shtatland, T., Singer, B. S., Wu, Y. (1997). From oligonucleotide shapes to genomic SELEX: novel biological regulatory loops. *Proceedings of the National Academy of Sciences*, 94(1), 59–64.

Golden, M. C., Collins, B. D., Willis, M. C., Koch, T. H. (2000). Diagnostic potential of PhotoSELEX-evolved ssDNA aptamers. *Journal of Biotechnology*, 81(2–3), 167–178.

Gopinath, S. C. (2006). Aptamers. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*, 1–33.

Green, P. H., Jabri, B. (2003). Coeliac disease. *The Lancet*, 362(9381), 383–391.

Gupta, R., Reddy, D. N., Makharia, G. K., Sood, A., Ramakrishna, B. S., Yachha, S. K., Thapa, B.R., Banerjee, R., Anuradha, S., Dutta, U., Puri, A.S. (2009). Indian

task force for celiac disease: current status. *World Journal of Gastroenterology: WJG*, 15(48), 6028.

Haraszi, R., Chassaigne, H., Maquet, A., Ulberth, F. (2011). Analytical methods for detection of gluten in food—method developments in support of food labeling legislation. *Journal of AOAC International*, 94(4), 1006–1025.

Hernando, A., Valdes, I., Méndez, E. (2003). New strategy for the determination of gliadins in maize-or rice-based foods matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: fractionation of gliadins from maize or rice prolamins by acidic treatment. *Journal of Mass Spectrometry*, 38(8), 862–871.

Herrera, M. G., Benedini, L. A., Lonez, C., Schilardi, P. L., Hellweg, T., Ruysschaert, J.M., Dodero, V. I. (2015). Self-assembly of 33-mer gliadin peptide oligomers. *Soft Matter*, 11(44), 8648–8660.

Hicke, B. J., Marion, C., Chang, Y.F., Gould, T., Lynott, C. K., Parma, D., Schmidt, P.G., Warren, S. (2001). Tenascin-C aptamers are generated using tumor cells and purified protein. *Journal of Biological Chemistry*, 276(52), 48644–48654.

Horvath, K., Mehta, D. I. (2000). Celiac disease—A worldwide problem. *The Indian Journal of Pediatrics,* 67(10), 757–763.

Hsia, C. C., Anderson, O. D. (2001). Isolation and characterization of wheat ω-gliadin genes. *Theoretical and Applied Genetics*, 103(1), 37–44.

Ilgu, M., Nilsen-Hamilton, M. (2016). Aptamers in analytics. *Analyst*, 141(5), 1551–1568.

Irvine, D., Tuerk, C., Gold, L. (1991). Selexion: Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. *Journal of Molecular Biology*, 222(3), 739–761.

Jayasena, S. D. (1999). Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clinical Chemistry*, 45(9), 1628–1650.

Jensen, K. B., Atkinson, B. L., Willis, M. C., Koch, T. H., Gold, L. (1995). Using in vitro selection to direct the covalent attachment of human immunodeficiency virus type 1 Rev protein to high-affinity RNA ligands. *Proceedings of the National Academy of Sciences*, 92(26), 12220–12224.

Kagnoff, M. F. (2007). Celiac disease: pathogenesis of a model immunogenetic disease. *The Journal of Clinical Investigation*, 117(1), 41–49.

Kang, J. Y., Kang, A. H., Green, A., Gwee, K. A., Ho, K. Y. (2013). Systematic review: worldwide variation in the frequency of coeliac disease and changes over time. *Alimentary Pharmacology & Therapeutics*, 38(3), 226–245.

Kang, K.N., Lee, Y.S. (2012). RNA aptamers: a review of recent trends and applications. *In Future Trends in Biotechnology*, 153–169. Springer, Berlin, Heidelberg.

Kawakami, J., Imanaka, H., Yokota, Y., Sugimoto, N. (2000). In vitro selection of aptamers that act with Zn2+. *Journal of Inorganic Biochemistry*, 82(1–4), 197–206.

Keefe, A. D., Pai, S., Ellington, A. (2010). Aptamers as therapeutics. *Nature Reviews Drug Discovery*, 9(7), 537.

Kikin, O., D'Antonio, L., Bagga, P. S. (2006). QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Research*, 34(suppl_2), W676–W682.

Klug, S. J., Famulok, M. (1994). All you wanted to know about SELEX. *Molecular Biology Reports*, 20(2), 97–107.

Klußmann, S., Nolte, A., Bald, R., Erdmann, V. A., Fürste, J. P. (1996). Mirror-image RNA that binds D-adenosine. *Nature Biotechnology*, 14(9), 1112.

Knight, R., Birmingham, A., Yarus, M. (2004). BayesFold: Rational 2 folds that combine thermodynamic, covariation, and chemical data for aligned RNA sequences. *Rna*, 10(9), 1323–1336.

Koehler, P., Schwalb, T., Immer, U., Lacorn, M., Wehling, P., Don, C. (2013). AACCI Approved Methods Technical Committee report: collaborative study on the immunochemical determination of intact gluten using an R5 sandwich ELISA. *Cereal Foods World*, 58(1), 36–40.

Koning, F., Gilissen, L., Wijmenga, C. (2005). Gluten: a two-edged sword. Immunopathogenesis of celiac disease. In *Springer Seminars in Immunopathology*, 27, 217–232. Springer-Verlag.

Lebenthal, E., Shteyer, E., Branski, D. (2008). The changing clinical presentation of celiac disease. In *Frontiers in celiac disease*, 12, 18–22. Karger Publishers.

Lebwohl, B., Sanders, D. S., Green, P. H. (2018). Coeliac disease. *The Lancet*, 391(10115), 70–81.

Li, S., Xu, H., Ding, H., Huang, Y., Cao, X., Yang, G., Li, J., Xie, Z., Meng, Y., Li, X., Zhao, Q. (2009). Identification of an aptamer targeting hnRNP A1 by tissue slide-based SELEX. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, 218(3), 327–336.

Lionetti, E., Castellaneta, S., Francavilla, R., Pulvirenti, A., Tonutti, E., Amarri, S., Barbato, M., Barbera, C., Barera, G., Bellantoni, A., Castellano, E. (2014). Introduction of gluten, HLA status, and the risk of celiac disease in children. *New England Journal of Medicine*, 371(14), 1295–1303.

Lionetti, E., Gatti, S., Pulvirenti, A., Catassi, C. (2015). Celiac disease from a global perspective. *Best Practice & Research Clinical Gastroenterology*, 29(3), 365–379.

Losowsky, M. S. (2008). A history of coeliac disease. *Digestive Diseases*, *26*(2), 112–120.

Maberley, D. (2005). Pegaptanib for neovascular age-related macular degeneration. *Issues in Emerging Health Technologies*, (76), 1–4.

Marshall, K. A., Ellington, A. D. (2000). [14] In vitro selection of RNA aptamers.

Martell, R. E., Nevins, J. R., Sullenger, B. A. (2002). Optimizing aptamer activity for gene therapy applications using expression cassette SELEX. *Molecular Therapy*, 6(1), 30–34.

Matsuo, H., Kohno, K., Morita, E. (2005). Molecular cloning, recombinant expression and IgE-binding epitope of ω-5 gliadin, a major allergen in wheat-dependent exercise-induced anaphylaxis. *The FEBS Journal*, 272(17), 4431–4438.

Matsushima, N., Creutz, C. E., Kretsinger, R. H. (1990). Polyproline, β-turn helices. Novel secondary structures proposed for the tandem repeats within rhodopsin, synaptophysin, synexin, gliadin, RNA polymerase II, hordein, and gluten. *Proteins: Structure, Function, and Bioinformatics*, 7(2), 125–155.

McAllister, B. P., Williams, E., Clarke, K. (2018). A comprehensive review of celiac disease/gluten-sensitive enteropathies. *Clinical Reviews in Allergy & Immunology*, 1–18.

McKillop, D. F., Gosling, J. P., Stevens, F. M., Fottrell, P. F. (1985). Enzyme immunoassay of gliadin in food. Portland Press Limited.

Mena, M., Sousa, C. (2015). Analytical tools for gluten detection: Policies and

regulation. *OmniaScience Monographs*.

Méndez, E., Vela, C., Immer, U., Janssen, F. W. (2005). Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *European Journal of Gastroenterology & Hepatology*, 17(10), 1053–1063.

Mendonsa, S. D., Bowser, M. T. (2004). *In vitro* evolution of functional DNA using capillary electrophoresis. *Journal of the American Chemical Society*, 126(1), 20–21.

Meresse, B., Malamut, G., Cerf-Bensussan, N. (2012). Celiac disease: an immunological jigsaw. *Immunity*, 36(6), 907–919.

Miranda-Castro, R., de-los-Santos-Álvarez, N., Miranda-Ordieres, A., Lobo-Castañón, M. (2016). Harnessing aptamers to overcome challenges in gluten detection. *Biosensors*, 6(2), 16.

Mitea, C., Kooy-Winkelaar, Y., van Veelen, P., de Ru, A., Drijfhout, J. W., Koning, F., Dekking, L. (2008). Fine specificity of monoclonal antibodies against celiac disease–inducing peptides in the gluteome. *The American Journal of Clinical Nutrition*, 88(4), 1057–1066.

Molberg, Ø., Uhlen, A. K., Jensen, T., Flæte, N. S., Fleckenstein, B., Arentz–Hansen, H., Raki, M., Lundin, K.E., Sollid, L.M. (2005). Mapping of gluten T-cell epitopes in the bread wheat ancestors: implications for celiac disease. *Gastroenterology*, 128(2), 393–401.

Morón, B., Cebolla, A., Manyani, H., Alvarez-Maqueda, M., Megías, M., Thomas, M.D., López, M.C., Sousa, C. (2008). Sensitive detection of cereal fractions that are toxic to celiac disease patients by using monoclonal antibodies to a main

immunogenic wheat peptide. *The American Journal of Clinical Nutrition*, 87(2), 405–414.

Morris, K. N., Jensen, K. B., Julin, C. M., Weil, M., Gold, L. (1998). High affinity ligands from in vitro selection: complex targets. *Proceedings of the National Academy of Sciences*, 95(6), 2902–2907.

Mujico, J. R., Lombardía, M., Mena, M. C., Méndez, E., Albar, J. P. (2011). A highly sensitive real-time PCR system for quantification of wheat contamination in gluten-free food for celiac patients. *Food Chemistry*, 128(3), 795–801.

Murch, S. (2016). Recent advances in celiac disease. *The Indian Journal of Pediatrics*, 83(12–13), 1428–1435.

Nimjee, S. M., Rusconi, C. P., Sullenger, B. A. (2005). Aptamers: an emerging class of therapeutics. *Annu. Rev. Med.*, 56, 555–583.

Nimjee, S. M., White, R. R., Becker, R. C., Sullenger, B. A. (2017). Aptamers as therapeutics. *Annual Review of Pharmacology and Toxicology*, 57, 61–79.

Ohuchi, S. P., Ohtsu, T., Nakamura, Y. (2006). Selection of RNA aptamers against recombinant transforming growth factor-β type III receptor displayed on cell surface. *Biochimie*, 88(7), 897–904.

Osborne, T. B. (1907). The proteins of the wheat kernel. (No. 84). Carnegie institution of Washington.

Ozuna, C. V., Iehisa, J. C., Giménez, M. J., Alvarez, J. B., Sousa, C., Barro, F. (2015). Diversification of the celiac disease α-gliadin complex in wheat: A 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant Journal*, 82(5), 794–805.

Pagano, B., Mattia, C. A., Giancola, C. (2009). Applications of isothermal titration

calorimetry in biophysical studies of G-quadruplexes. *International Journal of Molecular Sciences*, 10(7), 2935–2957.

Pagratis, N., Gold, L., Shtatland, T., Javornik, B. Gilead Sciences Inc, 2005. Truncation selex method. *U.S. Patent Application* 11/058,726.

Park, S., Ahn, J.Y., Jo, M., Lee, D., Lis, J. T., Craighead, H. G., Kim, S. (2009). Selection and elution of aptamers using nanoporous sol-gel arrays with integrated microheaters. *Lab on a Chip*, 9(9), 1206–1212.

Parzanese, I., Qehajaj, D., Patrinicola, F., Aralica, M., Chiriva-Internati, M., Stifter, S., Elli, L., Grizzi, F. (2017). Celiac disease: From pathophysiology to treatment. *World Journal of Gastrointestinal Pathophysiology*, 8(2), 27.

Patris, S., Vandeput, M., Kauffmann, J.M. (2016). Antibodies as target for affinity biosensors. *TrAC Trends in Analytical Chemistry*, *79*, 239–246.

Penner, G. (2012). Commercialization of an aptamer-based diagnostic test. *IVD Technol*, 18(4), 31–37.

Pinto, A., Polo, P. N., Henry, O., Redondo, M. C., Svobodova, M., O'Sullivan, C. K. (2014). Label-free detection of gliadin food allergen mediated by real-time apta-PCR. *Analytical and Bioanalytical Chemistry*, 406(2), 515–524.

Poddar, U., Thapa, B. R., Nain, C. K., Prasad, A., Singh, K. (2002). Celiac disease in India: are they true cases of celiac disease? *Journal of Pediatric Gastroenterology and Nutrition*, 35(4), 508–512.

Qi, P. F., Wei, Y. M., Yue, Y. W., Yan, Z. H., Zheng, Y. L. (2006). Biochemical and molecular characterization of gliadins. *Molecular Biology*, 40(5), 713–723.

Qiao, S.W., Bergseng, E., Molberg, Ø., Jung, G., Fleckenstein, B., Sollid, L. M. (2005). Refining the rules of gliadin T cell epitope binding to the disease-

associated DQ2 molecule in celiac disease: importance of proline spacing and glutamine deamidation. *The Journal of Immunology*, 175(1), 254–261.

Ramakrishna, B. S., Makharia, G. K., Chetri, K., Dutta, S., Mathur, P., Ahuja, V., Amarchand, R., Balamurugan, R., Chowdhury, S.D., Daniel, D., Das, A. (2016). Prevalence of adult celiac disease in India: regional variations and associations. *The American Journal of Gastroenterology*, 111(1), 115.

Redondo, M. C. B., Griffin, P. B., Ransanz, M. G., Ellis, H. J., Ciclitira, P. J., O'sullivan, C. K. (2005). Monoclonal antibody-based competitive assay for the sensitive detection of coeliac disease toxic prolamins. *Analytica Chimica Acta*, 551(1–2), 105–114.

Rewers, M. (2005). Epidemiology of celiac disease: what are the prevalence, incidence, and progression of celiac disease? *Gastroenterology*, *128*(4), S47–S51.

Robertson, D. L., Joyce, G. F. (1990). Selection *in vitro* of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature*, 344(6265), 467.

Rubio–Tapia, A., Kyle, R. A., Kaplan, E. L., Johnson, D. R., Page, W., Erdtmann, F., Brantner, T.L., Kim, W.R., Phelps, T.K., Lahr, B.D., Zinsmeister, A.R. (2009). Increased prevalence and mortality in undiagnosed celiac disease. *Gastroenterology*, 137(1), 88–93.

Rubio, E. B., Homs, M. C., O'Sullivan, C. K. (2004). Aptamers: Powerful Molecular Tools for Therapeutics and Diagnostics. *Molecular Analysis and Genome Discovery*, 191.

Ruckman, J., Green, L. S., Beeson, J., Waugh, S., Gillette, W. L., Henninger, D. D., Claesson-Welsh, L., Janjic, N. (1998). 2′-Fluoropyrimidine RNA-based aptamers

to the 165-amino acid form of vascular endothelial growth factor (VEGF165) inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. *Journal of Biological Chemistry*, 273(32), 20556–20567.

Salentijn, E. M., Esselink, D. G., Goryunova, S. V, van der Meer, I. M., Gilissen, L. J. W., Smulders, M. J. (2013). Quantitative and qualitative differences in celiac disease epitopes among durum wheat varieties identified through deep RNA-amplicon sequencing. *BMC Genomics*, 14(1), 905.

Sandberg, M., Lundberg, L., Ferm, M., Yman, I. M. (2003). Real time PCR for the detection and discrimination of cereal contamination in gluten free foods. *European Food Research and Technology*, 217(4), 344–349.

Sapone, A., Lammers, K. M., Casolaro, V., Cammarota, M., Giuliano, M. T., De Rosa, M., Stefanile, R., Mazzarella, G., Tolone, C., Russo, M.I., Esposito, P. (2011). Divergence of gut permeability and mucosal immune gene expression in two gluten-associated conditions: celiac disease and gluten sensitivity. *BMC Medicine*, 9(1), 23.

Savitzky, A., Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627–1639.

Scanlon, S. A., Murray, J. A. (2011). Update on celiac disease–etiology, differential diagnosis, drug targets, and management advances. *Clinical and Experimental Gastroenterology*, 4, 297.

Scherf, K. A., Poms, R. E. (2016). Recent developments in analytical methods for tracing gluten. *Journal of Cereal Science*, 67, 112–122.

Shamir, R., Heyman, M. B., Koning, F., Wijmenga, C., Gutierrez-Achury, J., Catassi,

C., Gatti, S., Fasano, A., Discepolo, V., Korponay-Szabó, I.R., Zevit, N. (2014). Celiac disease: past, present, and future challenges: dedicated to the memory of our friend and colleague, Prof David Branski (1944-2013). S1-20.

Shan, L., Qiao, S.-W., Arentz-Hansen, H., Molberg, Ø., Gray, G. M., Sollid, L. M., Khosla, C. (2005). Identification and analysis of multivalent proteolytically resistant peptides from gluten: implications for celiac sprue. *Journal of Proteome Research*, 4(5), 1732–1741.

Shewry, P. R., Tatham, A. S., Forde, J., Kreis, M., Miflin, B. J. (1986). The classification and nomenclature of wheat gluten proteins: a reassessment. *Journal of Cereal Science*, 4(2), 97–106.

Shimada, T., Fujita, N., Maeda, M., Ishihama, A. (2005). Systematic search for the Cra-binding promoters using genomic SELEX system. *Genes to Cells*, 10(9), 907–918.

Simón, E., Larretxi, I., Churruca, I., Lasa, A., Bustamante, M. Á., Navarro, V., del Pilar Fernández-Gil, M., Miranda, J. (2017). *Nutritional and Analytical Approaches of Gluten-Free Diet in Celiac Disease*. Springer.

Singer, B. S., Shtatland, T., Brown, D., Gold, L. (1997). Libraries for genomic SELEX. *Nucleic Acids Research*, 25(4), 781–786.

Sjöström, H., Lundin, K. E. A., Molberg, O., Körner, R., McAdam, S. N., Anthonsen, D., Quarsten, H., Norén, O., Roepstorff, P., Thorsby, E., Sollid, L.M. (1998). Identification of a gliadin T-cell epitope in coeliac disease: general importance of gliadin deamidation for intestinal T-cell recognition. *Scandinavian Journal of Immunology*, 48, 111–115.

Skerritt, J. H., Hill, A. S. (1990). Monoclonal antibody sandwich enzyme

immunoassays for determination of gluten in foods. *Journal of Agricultural and Food Chemistry*, 38(8), 1771–1778.

Slot, I. D. B., Van Der Fels-Klerx, H. J., Bremer, M. G. E. G., Hamer, R. J. (2016). Immunochemical detection methods for gluten in food products: Where do we go from here? *Critical Reviews in Food Science and Nutrition*, 56(15), 2455–2466.

Smith, D., Kirschenheuter, G. P., Charlton, J., Guidot, D. M., Repine, J. E. (1995). *In vitro* selection of RNA-based irreversible inhibitors of human neutrophil elastase. *Chemistry & Biology*, 2(11), 741–750.

Sollid, L. M., Jabri, B. (2005). Is celiac disease an autoimmune disorder? *Current Opinion in Immunology*, 17(6), 595–600.

Sollid, L. M., Qiao, S.W., Anderson, R. P., Gianfrani, C., Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics,* 64(6), 455–460.

Sorell, L., López, J. A., Valdés, I., Alfonso, P., Camafeita, E., Acevedo, B., Chirdo, F., Gavilondo, J., Méndez, E. (1998). An innovative sandwich ELISA system based on an antibody cocktail for gluten analysis. *FEBS Letters*, 439(1–2), 46–50.

Sreerama, N., Woody, R. W. (2000). Analysis of protein CD spectra: comparison of CONTIN, SELCON3, and CDSSTR methods in CDPro software. In *Biophysical Journal*, 78, 334A–334A. BIOPHYSICAL SOCIETY 9650 ROCKVILLE PIKE, BETHESDA, MD 20814-3998 USA.

Stoltenburg, R., Nikolaus, N., Strehlitz, B. (2012). Capture-SELEX: selection of DNA aptamers for aminoglycoside antibiotics. *Journal of Analytical Methods in*

Chemistry, 2012.

Stoltenburg, R., Reinemann, C., Strehlitz, B. (2005). FluMag-SELEX as an advantageous method for DNA aptamer selection. *Analytical and Bioanalytical Chemistry*, 383(1), 83–91.

Stoltenburg, R., Reinemann, C., Strehlitz, B. (2007). SELEX—a (r) evolutionary method to generate high-affinity nucleic acid ligands. *Biomolecular Engineering*, 24(4), 381–403.

Sun, H., Zu, Y. (2015). A Highlight of recent advances in aptamer technology and its application. *Molecules*, *20*(7), 11959–11980.

Svigelj, R., Dossi, N., Toniolo, R., Miranda-Castro, R., de-los-Santos-Álvarez, N., Lobo-Castañón, M. J. (2018). Selection of anti-gluten DNA aptamers in a deep eutectic solvent. *Angewandte Chemie*, 130(39), 13032–13036.

Tack, G. J., Verbeek, W. H., Schreurs, M. W., Mulder, C. J. (2010). The spectrum of celiac disease: epidemiology, clinical aspects and treatment. *Nature Reviews Gastroenterology & Hepatology*, 7(4), 204.

Tatham, A. S., Shewry, P. R. (1995). The S-poor prolamins of wheat, barley and rye. *Journal of Cereal Science*, 22(1), 1–16.

Tok, J. B.-H., Fischer, N. O. (2008). Single microbead SELEX for efficient ssDNA aptamer generation against botulinum neurotoxin. *Chemical Communications*, (16), 1883–1885.

Troncone, R., Vitale, M., Donatiello, A., Farris, E., Rossi, G., Auricchio, S. (1986). A sandwich enzyme immunoassay for wheat gliadin. *Journal of Immunological Methods*, 92(1), 21–23.

Tuerk, C., Gold, L. (1990). Systematic evolution of ligands by exponential

enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249(4968), 505–510.

Tye-Din, J. A., Galipeau, H., Agardh, D. (2018). Celiac disease: a review of current concepts in pathogenesis, prevention and novel therapies. *Frontiers in Pediatrics*, 6, 350.

Tye-Din, J. A., Stewart, J. A., Dromey, J. A., Beissbarth, T., van Heel, D. A., Tatham, A., Henderson, K., Mannering, S.I., Gianfrani, C., Jewell, D.P., Hill, A.V. (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Science Translational Medicine*, 2(41), 41ra51-41ra51.

Urade, R., Sato, N., Sugiyama, M. (2018). Gliadins from wheat grain: An overview, from primary structure to nanostructures of aggregates. *Biophysical Reviews*, 10(2), 435–443.

Vader, W., Kooy, Y., van Veelen, P., de Ru, A., Harris, D., Benckhuijsen, W., Peña, S., Mearin, L., Drijfhout, J.W., Koning, F. (2002). The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. *Gastroenterology*, 122(7), 1729–1737.

Vader, W., Stepniak, D., Kooy, Y., Mearin, L., Thompson, A., van Rood, J. J., Spaenij, L., Koning, F. (2003). The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. *Proceedings of the National Academy of Sciences*, 100(21), 12390–12395.

Valdés, I., García, E., Llorente, M., Méndez, E. (2003). Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *European Journal of Gastroenterology &*

Hepatology, 15(5), 465–747.

Van Eckert, R., Berghofer, E., Ciclitira, P. J., Chirdo, F., Denery-Papini, S., Ellis, H. J., Ferranti, P., Goodwin, P., Immer, U., Mamone, G., Mendez, E. (2006). Towards a new gliadin reference material–isolation and characterisation. *Journal of Cereal Science*, 43(3), 331–341.

Vasilescu, A., Marty, J.L. (2016). Electrochemical aptasensors for the assessment of food quality and safety. *TrAC Trends in Analytical Chemistry*, 79, 60–70.

Weigand, J. E., Suess, B. (2009). Aptamers and riboswitches: perspectives in biotechnology. *Applied Microbiology and Biotechnology*, 85(2), 229.

Whitmore, L., Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 32(suppl_2), W668–W673.

Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiology*, 24(2), 115–119.

Wieser, H. (2008). Detection of gluten. In *Gluten-Free Cereal Products and Beverages*, 47–80. Elsevier.

WooáKim, D., BockáGu, M. (2012). Immobilization-free screening of aptamers assisted by graphene oxide. *Chemical Communications*, 48(15), 2071–2073.

Wrigley, C. W., Shepherd, K. W. (1973). Electrofocusing of grain proteins from wheat genotypes. *Annals of the New York Academy of Sciences*, 209(1), 154–162.

Wu, L., Curran, J. F. (1999). An allosteric synthetic DNA. *Nucleic Acids Research*, 27(6), 1512–1516.

Wu, X., Chen, J., Wu, M., Zhao, J. X. (2015). Aptamers: Active targeting ligands for cancer diagnosis and therapy. *Theranostics*, 5(4), 322–344.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization

prediction. *Nucleic Acids Research*, 31(13), 3406–3415.

Publications:

Kalita, J.J., **Buragohain, P.,** Mosahari, P.V., Bora, U. (2019). Food Allergens: Detection and Management. Food Bioactives: Functionality and Applications in Human Health.

Conference Proceedings:

1. **Buragohain, P.** Autophagy: A potent modulator and novel therapeutic strategy in Cancer. International Conference on Disease Biology and Therapeutics (ICDBT-2014), at Institute of Advanced Study in Science and Technology (IASST), Guwahati, Assam, India, 3rd to 5th December, 2014.
2. **Buragohain, P.,** Bora, U. Aptamers against celiac disease. Research Conclave 2015, Indian Institute of Technology Guwahati, Assam, March 23-25, 2015,
3. **Buragohain, P.,** Kabiraj, D., Bharali, B., Bora, U. Aptamer-nanoparticle complexes as powerful addition to biodiagnostics and novel nanobiomedical applications. International Conference on Advances in Nanotechnology (ICAN-2017), at Assam Don Bosco University (ADBU), Azara, Guwahati, Assam, India, 9th to 13th January, 2017.
4. Kalita, J., **Buragohain, P.,** Kabiraj, D., Mosahari, P.V., Bora, U. Future prospects of edible insects farming in North-East India. International Symposium on Biodiversity and Biobanking, BIODIVERSE 2018, at IIT Guwahati, North Guwahati, Assam, India, 27th to 29th January 2018.