Exploring the role of canonical and non-canonical linkers in Cadherin-23

Pritam Saha

MS15196

A dissertation submitted for the partial fulfilment of

BS-MS dual degree in Science

Department of Biological Sciences

Indian Institute of Science Education and Research Mohali

India

May 2020

CERTIFICATE OF EXAMINATION

This is to certify that the dissertation titled **"**Exploring the role of canonical and non-canonical linkers in Cadherin-23**"** submitted by Mr. Pritam Saha (Registration Number: MS15196) for the partial fulfilment of BS-MS dual degree programme of Indian Institute of Science Education and Research, Mohali has been examined by the thesis committee duly appointed by the institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Sabyasachi Rakshit (Supervisor) Dr. Sudip Mandal Dr. Lolitika Mandal

Date:

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr. Sabyasachi Rakshit at the Indian Institute of Science Education and Research, Mohali. This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Pritam Saha

April, 2020

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Sabyasachi Rakshit

(Supervisor)

ACKNOWLEDGEMENT

I heartily convey my regards to Dr. Sabyasachi Rakshit under whom I did my last summer and my MS thesis. He is co-operative in making things understandable to his students. He lets his students to come up with new ideas and encourage the students to pursue their interest. Above all, he gave me an excellent opportunity to work in this interesting field of biophysics, through which I have learnt a newer aspect of scientific mindset. I also thank my committee members, Dr. Sudip Mandal and Dr. Lolitika Moandal for their critical comments and suggestion on my project and my presentations.

My special thanks to Jagadish, Gaurav and Nisha, with whom I worked for my final year project. I really thank them for their tireless efforts and teaching me molecular biology and other biophysical techniques.

I am thankful to my lab members, Varinder, Gayathri, Naimat, Ankush, Surbhi, Sai, Gaurav, Tanuja, Vishavdeep, Dimple, Simarpreet, Veerpal, Shweta, Nilesh and Anjana with whom I have spent my last year here at IISER Mohali.

I am thankful to Dr. N.G.Prasad, Dr. R Vijaya Anand, Dr. Purnananda Guptasarma for supports during summer projects in their laboratory. I learned many interesting topics through these interdisciplinary projects.

Last, but not least, I would love to thank my batchmates and my seniors, Paresh, Ankur, Ayush, Trirupa, Ajit, Nilangshu, Debjit, Ishan, Archit, Sanjib, Prithwish, who were always there for me when I needed them the most.

There are many more, who are not in this list, but they are an integral part of this Journey and I am thankful to all of them.

CONTENT

List of figures

List of tables

Annotations

Canonical Linker – CL

Non-canonical linker – NCL

Cadherin-23 – CDH23

Protocadherin-15 – PCDH15

Circular Dichroism – CD

Maltose Binding Protein - MBP

Extracellular - EC

Preface

Most of the prokaryotic and eukaryotic proteins are constituted of multi-domains. Although domains are considered independent folded unit, the inter-domain linker (IDLs) plays an important role in maintaining structural stability and conformational flexibility of entire protein. One such multidomain protein is Cadherin-23, which constitutes the tip-link in inner ear and acts as gating spring in relaying sound induced mechanical signal to ion channel. The IDLs of Cadherin-23 binds to calcium ions, which provide rigidity and structural stability to the protein. Depending on the extent of Ca^{2+} binding, the linkers are divided into two types: Canonical linker (CL) and Non-canonical linker (NCL). In this study, we have designed a NCL and compared it to the CL. Ca^{2+} binding affinity; thermodynamic stability and force mediated unfolding pattern for both the proteins have been investigated. Differential effect of NCL and CL on the force propagation through tip-link is the focus of this study. Information from this investigation will influence the subject area of studying the mechanism of inner ear mechanotransduction along with the mystery behind hearing diseases and deafness.

CHAPTER 1

1.1. Introduction

Hearing is a mechano-electrical response. Sound wave enters and creates mechanical signal in the inner ear. This mechanical response deflects stereocillia of inner ear hair-cells and generates tension in tip-links. Tip links is formed by hetero-tetramer of two non-classical cadherin family proteins[9], Cadherin-23[1] (CDH23) and Protocadherin-15[2] (PCDH15). The tension in tip-links is conveyed to the mechano-sensitive ion channels present at the bottom of stereocilia and is transduced into electrical signal[3]. Overall, CDH23 and PCDH15 are, therefore, essential for maintaining normal hearing activity. Studies have shown that mutations in these protein coding genes can cause balance disorder and deafness[5, 6, 7, 8].

Cadherin proteins are calcium binding proteins which are mainly cell adhesion molecule. They are broadly divided into two types: classical and non-classical cadherin. Both CDH23 and PCDH15 are non-classical cadherin family proteins. CDH23 has 27 and PCDH15 has 11 extracellular (EC) domain repeats. EC1 and EC2 from these two proteins are involved to form the tip-link[4]. The tip-link is about 150-200 nm long in low calcium concentration (~50 μM) of endolymph[10].

The domains of all cadherin family proteins are connected with inter-domain linkers, which consist of 6-10 amino acids. These linkers bind to Ca^{2+} ions and maintain the structural rigidity. Based on the affinity to Ca^{2+} and evolutionary conservations, these linkers are divided into two types: CLs and NCLs. NCLs have less affinity towards calcium ion or it has lesser number of Ca^{2+} bound to it than the CLs. Being fully, or partially calcium free, a NCL is expected to be more flexible and induces more entropic conformations. Overall, NCLs do carry different properties than CLs and thus expected to fulfil very different purpose in the protein than the CLs.

1.2. Objective

1. Elucidating the conformational differences between CLs and NCLs and correlate the differential binding ability of calcium-ion (Ca^{2+}) with these linkers.

2. Observing the effect of linker in the thermodynamic stability of domains for both NCL and CL.

1.3. Evolution of Hypothesis

The length of the tip-links varies with the change in calcium concentration in endolymph[10]. The change in length is also manifested with the spring-like properties of tip-links, indicating that the Ca^{2+} ions, by controlling the conformations of the tip-link proteins, maintain the rigidity of tip-links. CDH23 is the longest protein in tip-links. It has 27 EC domains that mediate the tip-links. In between the EC-domains, there are 26 inter-domain linker (IDLs) regions for CDH23. These IDLs are the primary calcium binding motifs in CDH23 that are formed mostly by negatively charged and polar residues. The number of Ca^{2+} ion that binding to the CDH23 varies with the concentration of calcium present in the cellular or extracellular environment. Generally, when there are three Ca^{2+} bound to the residues of linker and ECs; with very high affinity, it is called CL. And any deviation in number of Ca^{2+} ions or a significant decrease in the Ca^{2+} binding affinity, due to changes only in these motif residues; is called as NCLs.

It is not fully understood how changing the residues in canonical calcium binding motif will affect the integrity of CDH23 protein or the tip-link. However, not all mutations, which are causing changes from canonical to non-canonical ones, overlap with the disease causing residues. In this study, we have inspected the role of NCL in structural change and tried to correlate this change with the differential binding ability as a functional component. However, before starting experimental design the reason behind investigating this unique problem lies in the structural information and the evolutionary conserved sequences in cadherin family proteins. Cadherin family proteins are conserved through evolutionary timescale in terms of both sequence and structure across all species.

Table 1: Sequence similarity matrix for Cadherin-23 gene amongst different species.

Sequence similarity matrix	***Mus musculus***	***Homo sapiens***	***Danio rerio***	***Rattus norvegicus***	***Pan troglodytes***	***Gallus gallus***
Mus musculus	100					
Homo sapiens	94	100				
Danio rerio	68	67	100			
Rattus norvegicus	98	94	68	100		
Pan troglodytes	94	99	67	94	100	
Gallus gallus	78	78	66	78	78	100

Table 1 represents similarity in Cadherin-23 amongst different species. The percentage value is of similar number of data points (one letter code of amino acid) in an array (Genbank sequence of CDH 23) among the categories (species). Using BioEdit Sequence Alignment Editor Software, multiple sequence alignment was performed[18]. CDH-23 amino acid sequences were collected from NCBI database.

It is not just amongst species but for different types of cadherin family proteins, the sequence similarity is significant. Due to the chain-bead-chain model like structure, it was possible to collect single domain wise structure and elucidating the sequence using PyMol.

A typical topology of seven-β strand structure in the form of Greek key motif is common for all cadherin domains. There are flexible linker regions, which are mostly α helical in nature. In case of CDH23, it is 3_{10} α helix. The representative structure of CDH23 is given below in Figure 1.

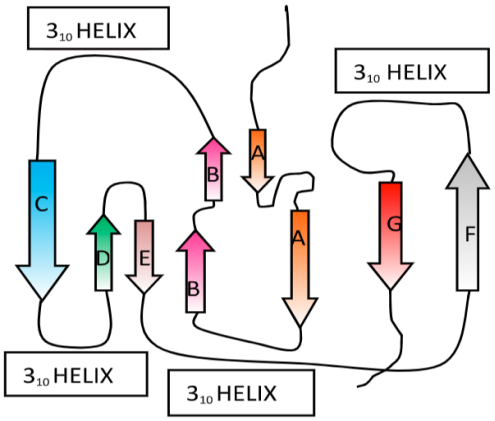

Figure 1: The model representation of a single EC domain unit of cadherin family proteins. A, B, C, D, E, F and G are the 7 different β-strands connected with α and 3_{10} helix linker.

This common Greek key motif like structure is common in both classical and non-classical cadherin and it is conserved in cadherin family proteins across all species. Although the structure is conserved, the similarity amongst the domain is not very significant. Observation was made for inter domain sequence similarity for CDH-23 and multiple sequence alignment was done for highly conserved sequence in the inter-domain linker region along with flexible calcium binding region from domains in PCDH-15.

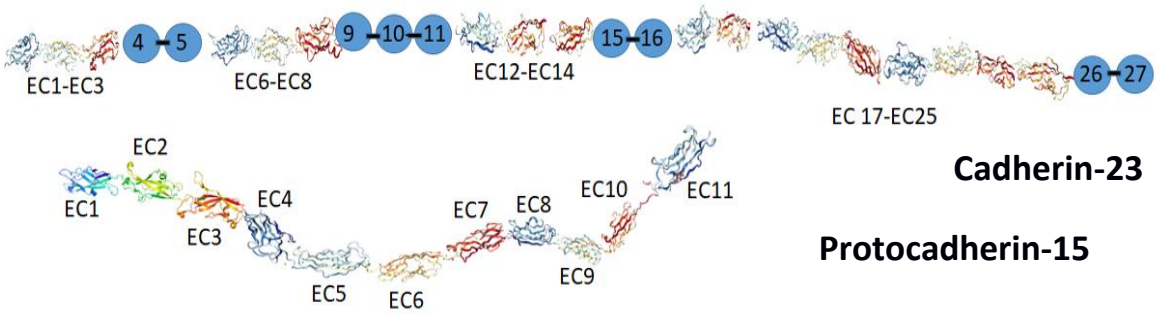

Figure 2: Cartoon representation of structure of Cadherin-23 and Protocadherin-15. Out of 27 domains in Cadherin-23, crystal structure is solved for 18 domains (dotted in blue sphere).

Amino acid sequence similarity was calculated using the SIAS tool (http://imed.med.ucm.es/Tools/sias.html).

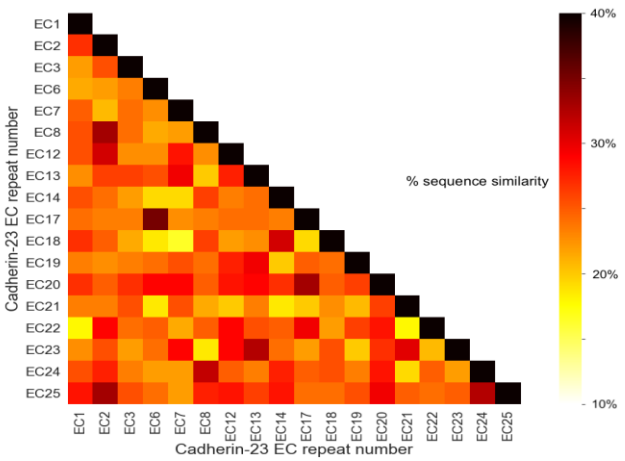

Figure 3: Sequence similarity matrix showing 15-35% sequence similarity derived from multiple sequence alignment using BioEdit and SIAS server.

Although the domain sequence similarity is less, but sequence and nature of the linker region are much conserved across all cadherin family proteins in all species. Those are mostly negatively charged and polar amino acid and known as calcium-ion binding motif.

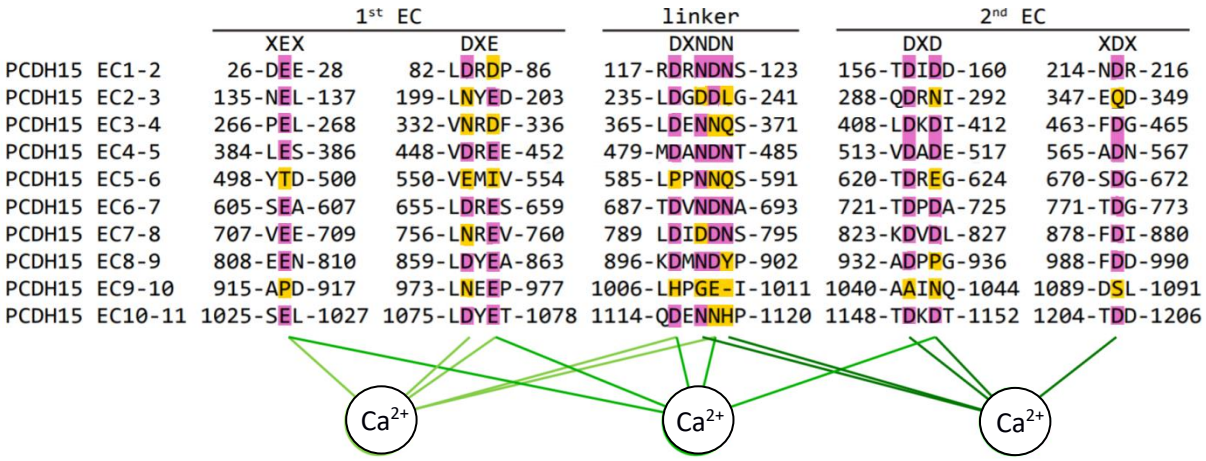

Figure 4: A list of conserved amino acid in the inter-domain linker region and flexible loops from neighbouring domains. These regions act as a calcium binding motif in Protocadherin-15. This sequence alignment is done from human Protocadherin15 (NCBI: NP_001136243.1)[15]. Yellow and magenta is the NCL and CL, respectively.

This high sequence similarity of linkers, conserved Ca^{2+} binding motif position and structural similarity among domains imposes a conformational dependence of whole protein on linkers.

CHAPTER 2

Experimental design, Cloning, Expression and Purification

2.1. Experimental design

The common calcium binding residue model in wild-type cadherin is given below in Table 2. This is also known as canonical model.

Table 2: The canonical model of Calcium ion binding motif. (X is any amino acid)

Motif Position	**EC_X**	**EC_X**	**LINKER**	**EC_{X+1}**	**EC_{X+1}**
Amino acid	X-E-X	D-(R/Y)-E	D-X-N-D-N	D-X-D	X-D-X

Any different amino acid residue in this model is called as NCL model. Mutation in non-CL is also responsible for causing hearing diseases. The focus of my study is in CDH23 EC1-2 due to its direct involvement in tip link formation. These are some disease mutants that are found in the calcium binding motif of EC1 and EC2 of CDH23. The D101G mutation has shown to have a 6 fold decrease[17] in Ca^{2+} binding affinity and cooperativity.

Table 3: A list of disease mutations in Ca^{2+} binding motif of Cadherin-23 EC1-2.

EC number	**Mutation**	**Ca^{2+} binding Motif**	**Diagnosis**
EC1[11]	D71G	**D**-R-E	Non-syndromic
EC1[12]	D101G	**D**-X-N-D-N	Non-syndromic
EC2[13]	E120Q	X-**E**-X	Unknown
EC2[14]	D137N	D-X-**D**	Non-syndromic

2.2. Cloning

The target construct was EC1-2 of Cadherin23. The mutated amino acids are the residues that were in the CDH23 EC1-2 construct, which help binding the three Calcium ions. Taking into account the composition of the natural NCL linker in PCDH-15 EC3-4 and amino acids that are involved in calcium binding in EC1-2, construct was designed for CDH23 EC1-2. **D**R**E** and DXN**DN** calcium binding motif in CL were mutated to **N**R**D** and DXN**NQ** making the linker non-canonical.

Table 4: Ca^{2+} binding residues of CL and NCL model.

EC number	EC1	EC1	LINKER	EC2	EC2
CL Model	SED	DRE	DVNDN	DPD	TDQ
NCL Model	SED	NRD	DVNNQ	DPD	TDQ

To design the construct, four primers with inserted mutation and two universal forward and reverse primers with the respective tag, were designed. The list of the primers is given below in Table 5.

Table 5: The primer sequences for cloning.

Primer name	Primer sequence
CDH-linker1-FP	CAGCCGCTGAATCGCGATACC
CDH-linker1-RP	TATAGGTATCGCGATTCAGCGGC
CDH-linker2-FP	GGTGACGTTAACAATCAGGCG
CDH-linker2-RP	TATACGTCGGCGCCTGATTGTTAACG
CDH-EC1-FP	CTCGAGGGATCCCGCATTATCGTTAACGTCACC
CDH-EC2-RP	CAATAATCTCGAGTCATTAGCTGCTACCCGTTTCCGGCAG

The mutant construct was cloned using principle of site-directed mutagenesis polymerase chain reaction. To make a tetra-mutant construct, a double mutant construct was made and from that after three PCR reactions later, final product was constructed. This was site directed mutagenesis followed by overlap extension PCR. The schematics of this PCR for cloning is given below in Figure 5.

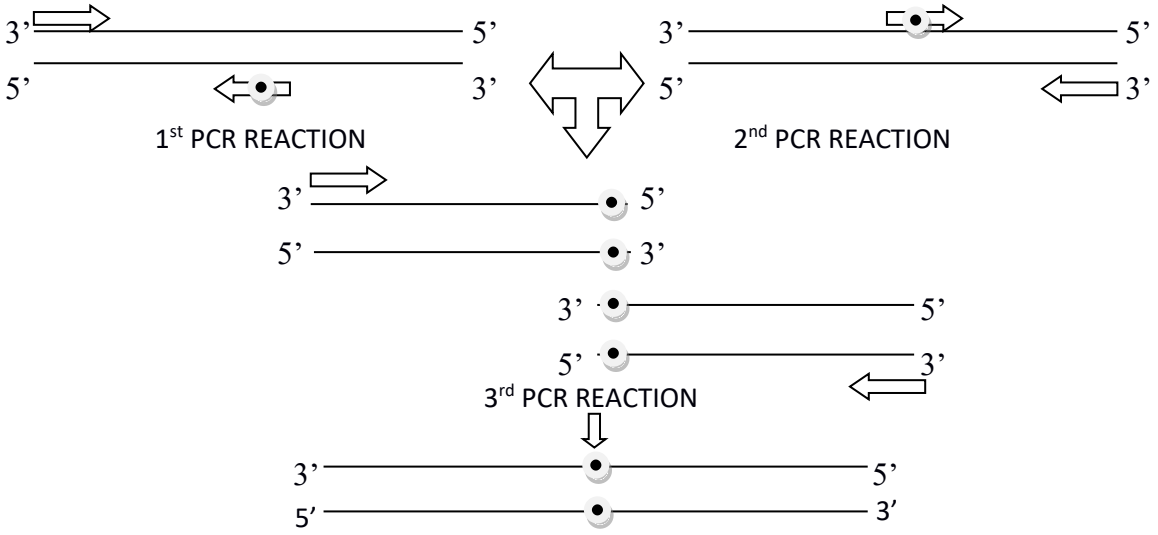

Figure 5: Schematics of overlap extension PCR for the cloning.

For cloning the PCR master-mix composition and reaction stepwise protocol is given below in Table 6 and Table 7, respectively.

Table 6: The PCR Reaction Mixture composition for cloning.

Components	Volume (in µl)
10µM Forward Primer	0.5
10µM Reverse Primer	0.5
dNTPs mix	0.5
Plasmid (25-30 ng)	Volume according to concentration
Standard Taq buffer (10x)	2.5
Taq DNA Polymerase	0.1
Autoclaved Water	Make up total volume to 25 µl

Table 7: The PCR reaction protocol for cloning.

Step	Temperature	Time	
Initial denaturation	95°C	5 minutes	
Denaturation	95°C	30 secs	30 cycles
Annealing	61°C - 65°C	30 secs	
Extension	72°C	1 kb/min	
Final extension	72°C	5 Minutes	
Hold	4°C	∞	

For constructing NCL CDH23 EC1-2 construct, three sets of PCR were done for the first two mutations and then another three reactions for the remaining two of four mutations. Taq polymerase was used for site-specific mutation causing PCR and the overlap extension PCR.

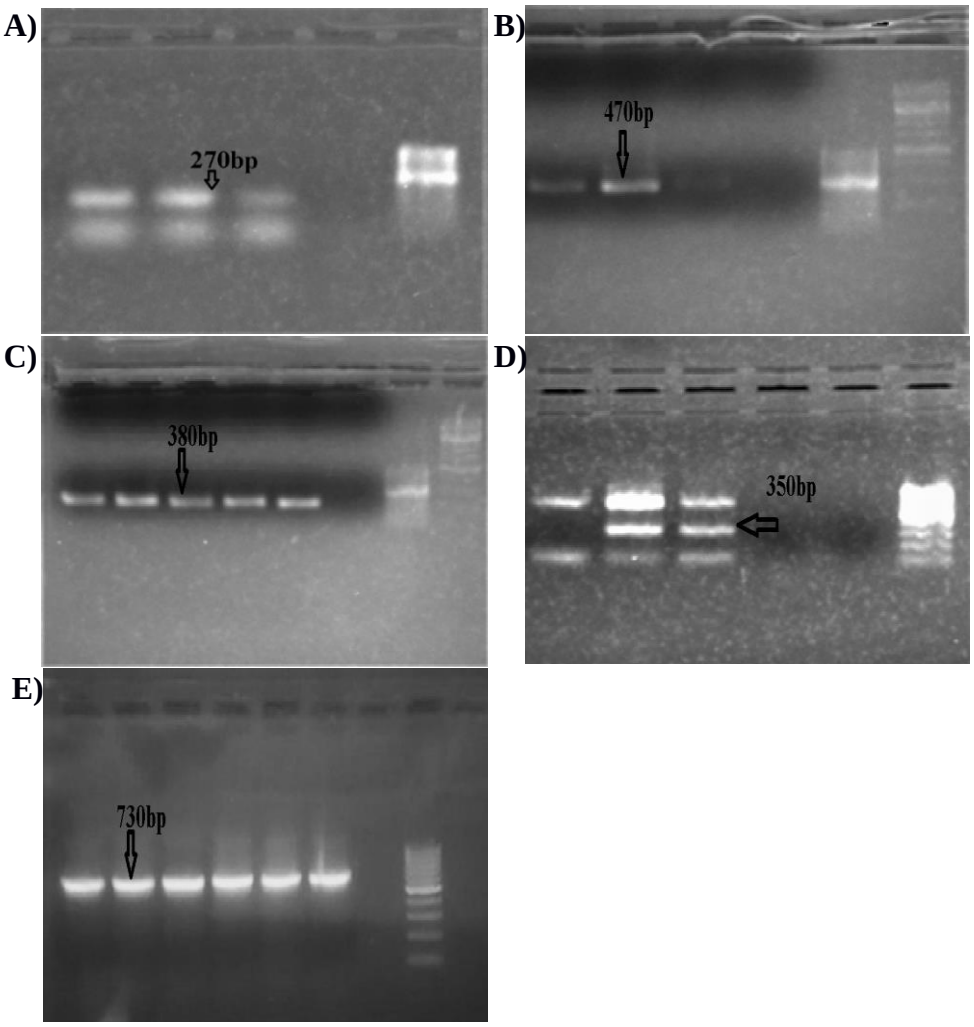

Figure 6: PCR amplification for NCL construct.

All images shown are of agarose gel, post electrophoresis. Before electrophoresis, EtBr was added to the gel. DNA fragments were separated at 90 V for 40 minutes. The agarose gel was exposed at UV light and the gel image was captured using Gel documentation system[19].

(A). PCR reaction with CDH EC1 FP and CDH linker1 RP resulting 270 bp fragments.
(B). PCR reaction with CDH EC2 RP and CDH linker1 FP resulting 460 bp fragments.
(C). PCR reaction with CDH EC1 FP and CDH linker2 RP resulting 380 bp fragments.
(D). PCR reaction with CDH EC2 RP and CDH linker2 FP resulting 350 bp fragments.
(E). Final product of Cadherin-23 EC1-2 tetra-mutant of 730 bp.

2.3. Expression and purification

For future experiµent, both the CL WT and tetra-mutant NCL gene of mouse CDH23 EC1–2 (Q24 to D228 in NP_075859.2) were recombinantly modified with 6×His-tag at N-terminal for affinity-purification, V2C mutation and sort-tag (LPETG) at C-terminus for poly-protein construct formation and surface attachment respectively[20].

This was cloned into Nde1 and Xho1 sites of PET21a vector (Novagen, Merck) and expressed in BL21 (DE3) (Stratagene) as soluble proteins with 1 mM of L-Rhamnose (MP Biomedicals) induction at an optical density of 0.4 and with 1000 mM of IPTG at an optical density of 0.6 in the culture media[20]. Then the media was kept at 37°C for 4 hours. Buffer composition was 25 mM HEPES, 50 mM KCl, 100 mM NaCl, and 2 mM $CaCl_2$. This composition was maintained throughout our experiments except the Ca^{2+} content in the buffer and CD buffer. All the proteins were obtained in cell inclusion bodies and subsequently purified using Ni-NTA (Qiagen) affinity-columns and superdex 200 increase 10/300 GL (GE Lifesciences) size-exclusion columns. The refolding was done by gradually lowering Urea concentration from 8M to 0M over 72 hours at 4°C[21]. 1mM of Dithiothreitol (DTT) was used to restrict di-sulfide bond formation and oligomer formation based precipitation[21]. The purity of the proteins was monitored from SDS–PAGE gel electrophoresis[20].

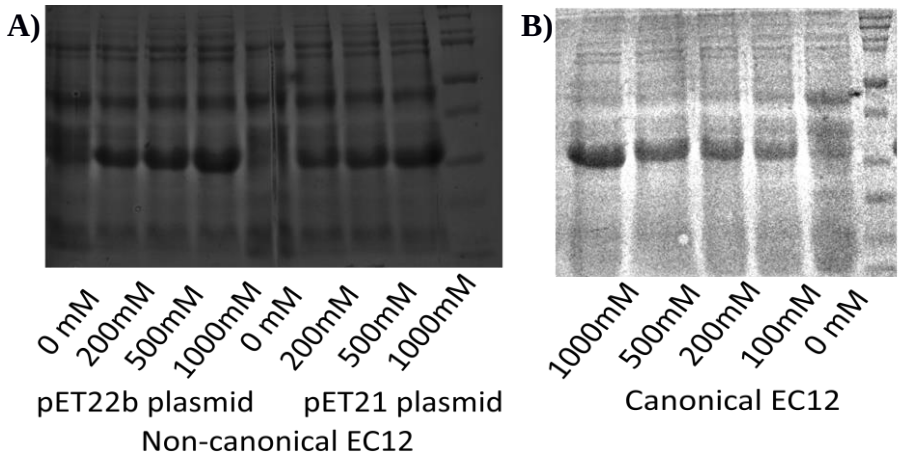

Figure 7: SDS-PAGE analysis showing expression of both CL and NCL CDH23 EC1-2 in *Escherichia Coli* with different IPTG induction concentration. (A). NCL protein IPTG standardization was done using both pET21a and pET22b plasmid, whereas (B). CL protein was expressed using pET21a plasmid.

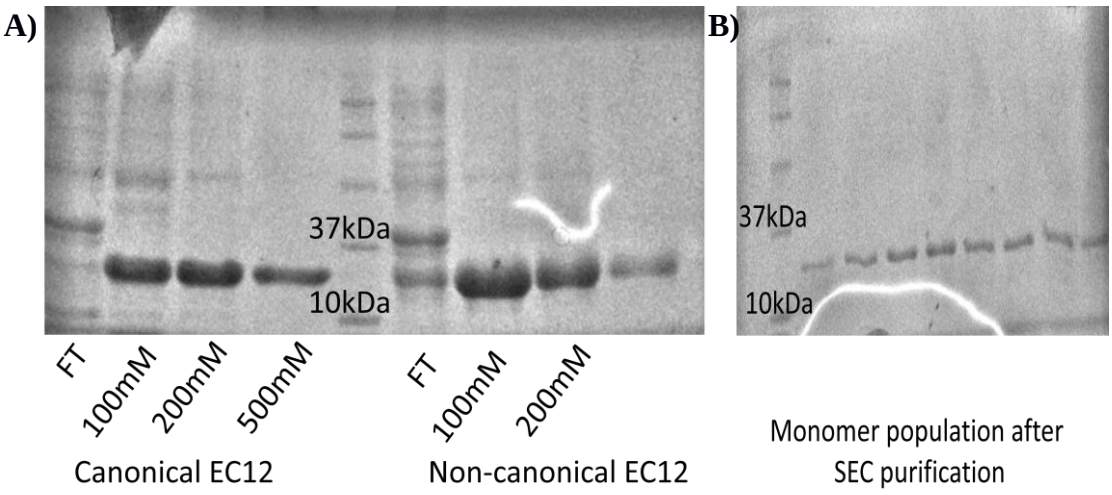

Figure 8: SDS-PAGE analysis showing purification of both CL and NCL CDH23 EC1-2. (A). Both NCL and CL were purified using Ni-NTA affinity columns and (B). Purification of CL CDH23 EC1-2 using Superdex 200 increase 10/300 GL (GE Lifesciences) size-exclusion columns respectively.

CHAPTER 3

Spectroscopic comparison of calcium binding affinity between CL and NCL linker model CDH23 EC1-2 protein

3.1. Aim

Compare the calcium binding affinity of CL and NCL between CDH23 EC1-2.

3.2. Introduction

Cadherin family proteins are calcium adhesion protein, which means calcium has a huge role in maintaining the structure and function of these proteins. One of such protein is Cadherin-23, which is a non-classical cadherin and present in inner ear. CDH23 is directly related to diseases of hearing loss and many disease mutations has been found in calcium binding residues present in the CDH23. Although, the endolymph calcium concentration is near 50 µM but there is continuous influx of Ca^{2+}, K^{+} ions near the ion channels. It is obvious that the structure and function of CDH23 are dependent on the calcium concentration of inner ear. To understand the role of calcium in inner ear mechanotransduction, it is highly important to quantify the calcium binding affinity of this protein.

There are various optical spectroscopic techniques, which can be used to measure Ca^{2+} binding constants. This is mostly used for the proteins which display a difference in the Ca^{2+} bound and free forms, for example, CD or steady-state fluorescence spectrum.

Calcium binding affinity is calculated by various methods such as differential scanning calorimetry (DSC), isothermal calorimetric titration (ITC) and using MD simulations. If the protein shows change upon induction of quantifiable Ca^{2+}, it is possible to start from a Ca^{2+} free form and then stepwise titrate the Ca^{2+}. In each step a measurement is taken of spectra or intensity and from that binding constant can be measured. But mostly the protein is not able to show those changes with different calcium concentration in the environment or the instrument is not sensitive enough to record the changes.

One of the popular indirect approaches is to take a chelator which shows quantifiable spectroscopic change with binding to calcium ions. To elaborate this, if a chelator is present in a buffer solution, free of Ca^{2+} and by stepwise addition of Ca^{2+}, it shows changes in the intensity or spectra. The same chelator if mixed with protein in same concentration, which provides same probability to bind to Ca^{2+}, while Ca^{2+} stepwise addition, will show lesser rate of changes due to the fact that the protein is also chelating the Ca^{2+} ions. Because the protein sites have threshold-binding affinity, at some point or several points of titration the changes in measured spectroscopic unit for the chelator will decrease or stop.

3.3. Experimental Procedure

3.3.1 Chemicals and materials needed

(1) UV-Visible spectrophotometer.
(2) Quartz cuvette.
(3) Chromophoric calcium chelator: Quin-2
(4) Ca^{2+} free buffer (25 mM HEPES, 50 mM KCl, 100 mM NaCl).
(5) 1mM $CaCl_2$
(6) 5mM EDTA
(7) CL and NCL CDH23 EC1-2 protein.

3.3.2 Procedure

(1) A Ca^{2+} free solution of 20 µM quin-2 is prepared in calcium free buffer (25 mM HEPES, 50 mM KCl, 100 mM NaCl).
(2) The cuvette is filled with 5mM EDTA and let sit for 1 hour. After this step cuvette is rinsed with ddH_2O several times.
(3) The protein solution was made in Ca^{2+} free buffer and kept in rotor with chelex for 2 hours.
(4) 99 µl of 20 µM protein was mixed with 1 µl of 2mM quin2, so that both have equal concentration.
(5) After the mixing with pipette is done, the absorbance at 263nm was measured.

(6) Again absorbance at 263 nm is measured and 0.5 µl of 1mM $CaCl_2$ stock aliquot solution also made in same buffer is added and mixed along with the solution step by step still absorbance shows the saturation.

3.3.3. Precaution

(1) The concentration of chelator and the protein should be exactly same so that the probability of binding to calcium ions is equal for both.

(2) To get the exact absorbance of the chelator, absorbance should be measured at a wavelength where protein shows no absorption or very poor absorbance with change in Ca^{2+} ion concentration.

(3) The calcium ion concentration in the buffer should be below 1µM otherwise the buffer can be passed through chelex column to get it free from calcium ions.

(4) The addition of Ca^{2+} ion solution should be accurate and exactly same for each step to get a precise value of calcium binding constant.

(5) After each step of $CaCl_2$ solution addition, the mixture needs to have a incubation period to reach equilibrium. It can be measured by taking spectra after a period of 30 second starting from 0^{th} min to 10^{th} min after $CaCl_2$ addition.

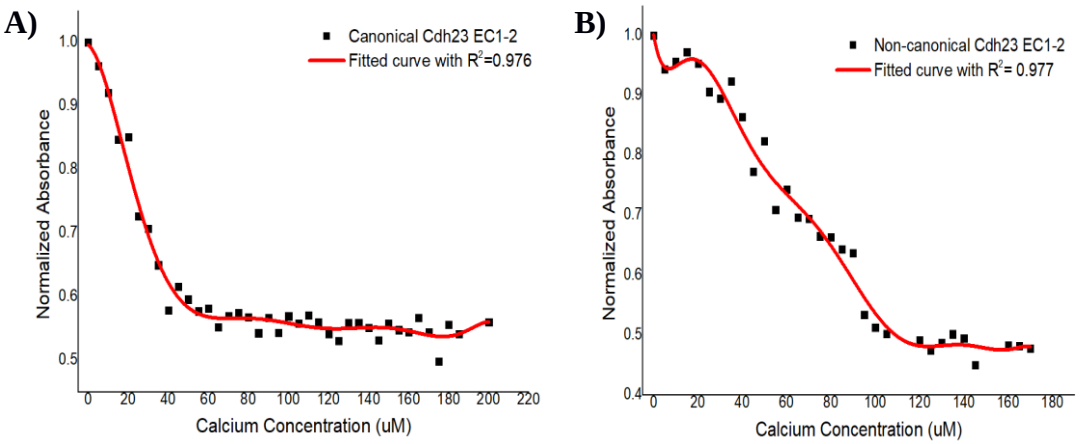

Figure 9: Representation of quin2-protein mixture solution absorbance at 262nm wavelength with varying calcium concentration. (A). CL CDH23 EC1-2 fitted with 4 Ca^{2+} model with Calcium Binding affinity of 79.432 μM^{-1}, 1.2516 nM^{-1}, 45.008 μM^{-1}, 40.218 μM^{-1} (B). NCL CDH23 EC1-2 fitted with 3 Ca^{2+} model with Calcium Binding affinity of 6 μM^{-1}, 46 μM^{-1}, 116 μM^{-1}.

CHAPTER 4

Quantitative estimation of thermodynamic stability between CL CDH23 EC1-2 and NCL CDH23 EC1-2

4.1. Aim

To study conformation changes with temperature melting and chemical denaturation using CD and Tryptophan fluorescence change with temperature melting using Steady-state Fluorescence.

4.2. Introduction

Various proteins including cadherin family have multiple domains in their strcture joined via inter-domain linker. Unlike globular protein due to the linker, domain boundary is very precise in this sort of structured protein. Whether these linkers provide conformational stability or increase the flexibility depends on the composition and nature of the amino acid in the linkers[24]. Also in the absence of linker, it has been reported to alter thermodynamic stability in case of a cadherin family protein [25].

For NCL cadherin-23 protein model the linker is more flexible than WT as shown in the molecular dynamics simulation. In addition, calcium chelator experiment has indicted that there is one less calcium in the inter-domain region. The four mutations and the removed calcium might change the thermodynamic stability of the protein containing two EC domain segments.

A)

B)

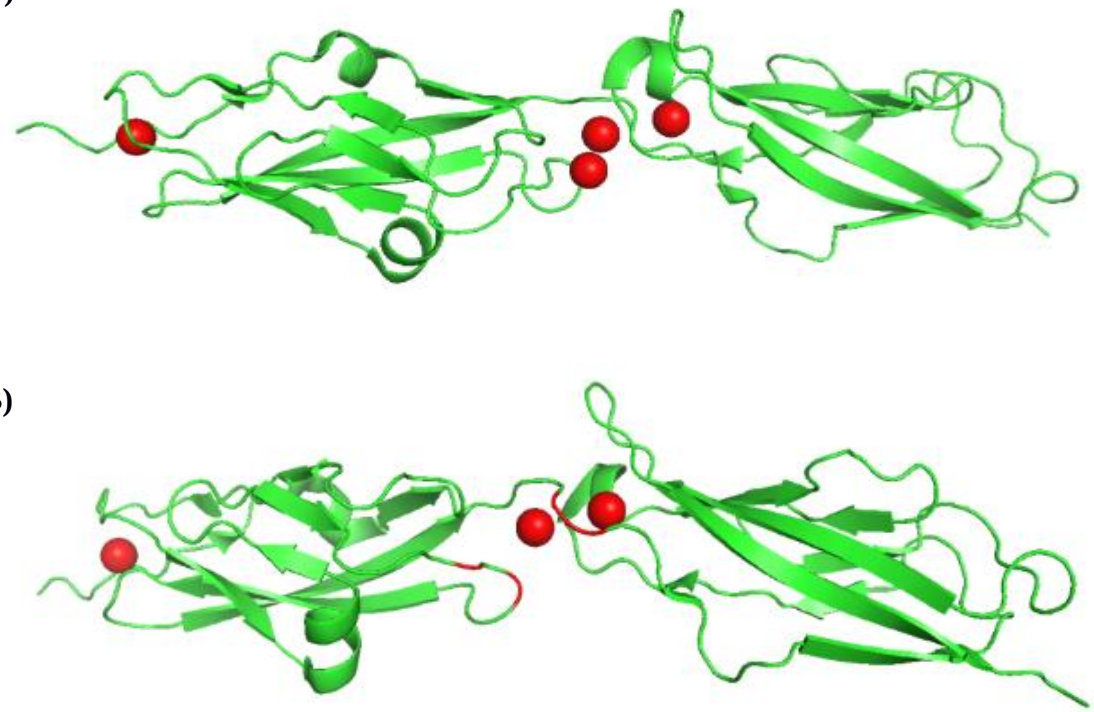

Figure 10: (A) CL CDH23 EC1-2 representative structure with 4 Ca^{2+} (B). NCL CDH23 EC1-2 representative structure with 3 Ca^{2+}. The 4 mutations in the calcium binding region for NCL are highlighted with red.

4.3. Lower thermodynamic stability of NCL protein studied via Circular Dichroism

The far-UV CD spectra were recorded on a J-810 Jasco CD spectrometer[26]. Changes were recorded in the range of 200 to 280 nm and final spectrum was averaged by setting up three scans for each sample. For CD temperature melting study at higher temperature and CD chemical denaturation study at higher denaturant concentration, the starting wavelength of recording was changed continuously to avoid damaging the instrument. For CD denaturant and CD temperature melting study the concentration was kept 10 µM so that the HT voltage won't cross 800. Beyond 700 or 750 the data becomes very disproportionate to the signal and noise becomes very high. CD experiments were conducted at a room temperature. The path-length of the quartz cuvette was 1 mm. Buffer baseline subtraction was done for each set of spectra[28]. CD temperature study was done

from 20°C to 83°C with 3°C interval and 2 minutes incubation heating period after it reaches the target temperature.

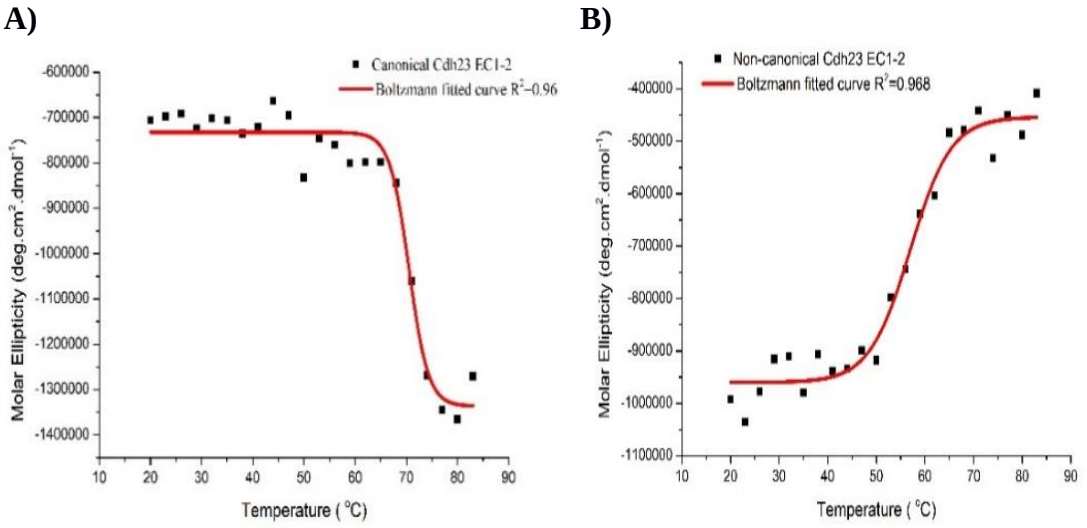

Figure 11: Representation of secondary structural changes in protein structure monitored by far-UV CD spectra and change in molar ellipticity as a function of temperature. (A). CL CDH23 EC1-2 temperature melting study at 215 nm. (B). NCL CDH23 EC1-2 temperature melting study at 208 nm.

Table 8: Melting temperature of CL and NCL from thermal denaturation CD study.

Protein	Melting Temperature (°C)
CL CDH23 EC1-2	70.52 ± 0.53 °C
NCL CDH23 EC1-2	56.97 ± 0.87 °C

For chemical denaturation study, changes were recorded in the range of 212 to 280 nm and final spectrum was averaged by setting up three scans for each sample. Due to the nature of the experiment the HT voltage was crossing the safety threshold value of 800 beyond 210 nm at higher denaturant concentration. A stock solution of Urea (10M) was made in the same buffer just before the experiment. The concentration of NCL protein and CL protein were 60 µM and 30 µM. The samples were made according to the list in the figure 12.

CL PROTEIN (30uM)	Urea (10M)	CONCENTRATION (10M)	BUFFER (ul)	Total Volume	NCL PROTEIN (60uM)	Urea(10M)	CONCENTRATION (10M)	BUFFER (ul)	Total Volume
53.33	0	0	106.67	160	26.66	0	0	133.34	160
53.33	4	0.25	102.67	160	26.66	4	0.25	129.34	160
53.33	8	0.5	98.67	160	26.66	8	0.5	125.34	160
53.33	12	0.75	94.67	160	26.66	12	0.75	121.34	160
53.33	16	1	90.67	160	26.66	16	1	117.34	160
53.33	20	1.25	86.67	160	26.66	20	1.25	113.34	160
53.33	24	1.5	82.67	160	26.66	24	1.5	109.34	160
53.33	28	1.75	78.67	160	26.66	28	1.75	105.34	160
53.33	32	2	74.67	160	26.66	32	2	101.34	160
53.33	36	2.25	70.67	160	26.66	36	2.25	97.34	160
53.33	40	2.5	66.67	160	26.66	40	2.5	93.34	160
53.33	44	2.75	62.67	160	26.66	44	2.75	89.34	160
53.33	48	3	58.67	160	26.66	48	3	85.34	160
53.33	52	3.25	54.67	160	26.66	52	3.25	81.34	160
53.33	56	3.5	50.67	160	26.66	56	3.5	77.34	160
53.33	60	3.75	46.67	160	26.66	60	3.75	73.34	160
53.33	64	4	42.67	160	26.66	64	4	69.34	160
53.33	68	4.25	38.67	160	26.66	68	4.25	65.34	160
53.33	72	4.5	34.67	160	26.66	72	4.5	61.34	160
53.33	76	4.75	30.67	160	26.66	76	4.75	57.34	160
53.33	80	5	26.67	160	26.66	80	5	53.34	160
53.33	84	5.25	22.67	160	26.66	84	5.25	49.34	160
53.33	88	5.5	18.67	160	26.66	88	5.5	45.34	160
53.33	92	5.75	14.67	160	26.66	92	5.75	41.34	160
53.33	96	6	10.67	160	26.66	96	6	37.34	160
53.33	100	6.25	6.67	160	26.66	100	6.25	33.34	160
53.33	104	6.5	2.67	160	26.66	104	6.5	29.34	160

Figure 12: For CD chemical denaturation study samples were made separately according to the given chart. Denaturation study was performed from 0 M to 6.5 M Urea denaturation with 3 min incubation time for chemical denaturation.

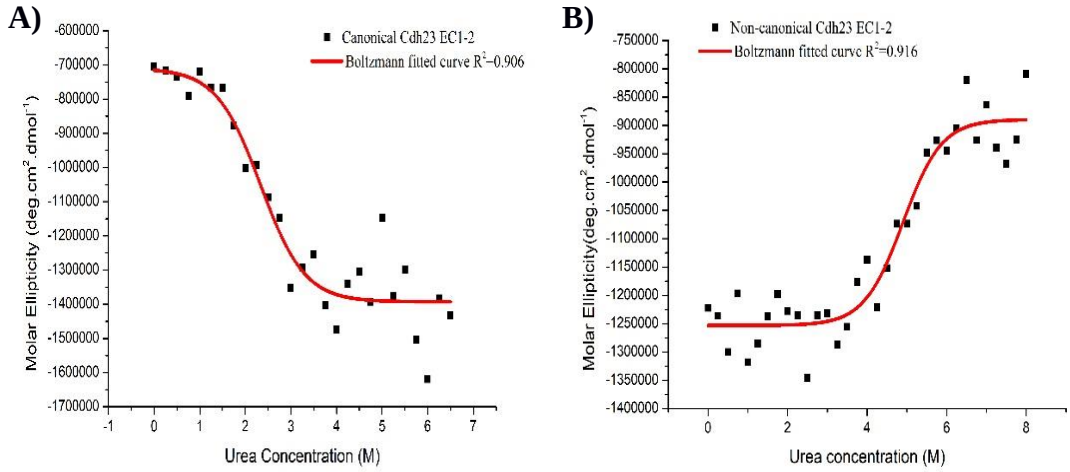

Figure 13: Representation of secondary structural changes in protein structure monitored by far-UV CD spectra and change in molar ellipticity as a function of Urea concentration. (A). CL CDH23 EC1-2 Urea denaturation study at 212 nm. (B). NCL CDH23 EC1-2 Urea denaturation study at 211 nm.

Table 9: Critical denaturant concentration of CL and NCL from chemical denaturation CD study.

PROTEIN	Critical DenaturantConcentration
CL CDH23 EC1-2	2.33±0.15 M
NCL CDH23 EC1-2	4.89±0.14 M

The CD buffer (1 Mm HEPES, 10 Mm NaCl) has very low salt concentration. NCL protein was very unstable and forming aggregates at high temperature. CD temperature melting plot was shifting towards 220 nm at high temperature. Also it is a intensity plot at single wavelenth and it is possible that due to aggregation the NCL aggregates have masked some signal for which the signal intensity is lowered. And, thus forming a different pattern than CL model in temperature melting plot.

Also, the higher critical urea concentration needed to denature the structure in case of NCL protein does not necessariliy mean it has higher higher thermodynamic stability. Because it is observed that at higher stress thermal or chemical, NCL model protein is forming aggregates. It is possible that the accessibility of the polar and charged amino-acid are being lowered due to formation of structured aggregates and thus more concentration of Urea is needed to denature the structure. Also, the similar pattern of both experiment for NCL protein confirms the masking of signal which in term causes lower intensity at higherer chemical denaturant stress at higher concentration of urea.

4.4. Lower thermodynamic stability of NCL protein studied via Steady-state Fluorescence

The thermal stability at higher salt concentration where NCL protein does not aggregate, tryptophan fluorescence with temperature melting was measured using steady state fluorescence spectrometer. Due to the presence of only one tryptophan in CDH23 EC1, it was used as intrinsic probe to measure protein conformational change with temperature.

The measurement started from 25 °C to 95 °C with 5 °C temperature difference and 3 minutes incubation heating period after reaching the target temperature. Counts per

second (CPS) was monitored closely so that it does not cross 10^6 cps, otherwise the photodiode would be damaged.

$$f_D = (Y - Y_N) / (Y_D - Y_N) \qquad \text{(equation 1)}$$

Where f_D is the unfolded fraction, Y_N and Y_D are the value for native and denatured state.

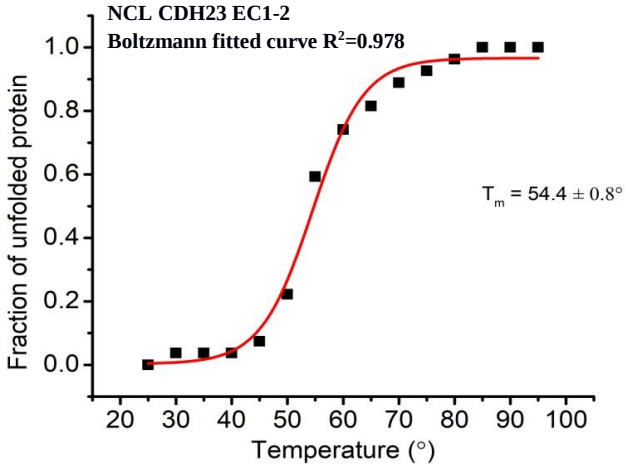

Figure 14: Representative study of thermal unfolding curve of NCL CDH23 EC1-2 obtained from the equation 1, mention above.

To elucidate a conclusion from the results of all these experiments would be- NCL CDH23 EC1-2 has lower thermodynamic stability and is more aggregation prone than CL CDH23 EC1-2.

CHAPTER 5

Characterization of mechano-responsive behaviour of both CL and NCL protein through single-molecule force spectroscopy using AFM

5.1. Aim

Deciphering the unfolding behaviour of CL and NCL CDH23 EC1-2 using AFM based single-molecule force spectroscopy.

5.2. Experimental setup

For the single-molecule force-spectroscopy experiment Atomic Force Microscope (AFM) (Nano wizard 3, JPK Instruments, Germany) was used[21]. Sortagging protocol was used to immobilize the C-termini of the protein molecules on the freshly cleaned (with surface preparation protocol) glass-coverslips and Si_3N_4 cantilevers (Olympus, OMCL-TR400PSA-1)[21]. 10% bifunctional PEG (NHS-PEG-maleimide) in mono-functional PEG (NHS-PEG) was used so that a higher density of molecules on the surface can be achieved. This mixture will allow to create a distance of $\sim$150 nm between two neighbouring molecule[22].

A chimeric construct was made using Maltose Binding Protein (MBP) and CDH23 EC1-2. MBP was used due to its known signature-unfolding pattern. MBP was modified with GG and CC in different ends. Cantilever was coated with streptavidin to interact with the cysteine amino acid in V2C mutation, which was exposed with TEV protease enzyme cut in the TEV site just before V2C mutation site. CDH23 EC1-2 was modified with sort tag (LPETG) which would react with the GGG modification in signature MBP. Surface was coated with NHS-PEG-Maleimide, which would interact with CC modification in MBP.

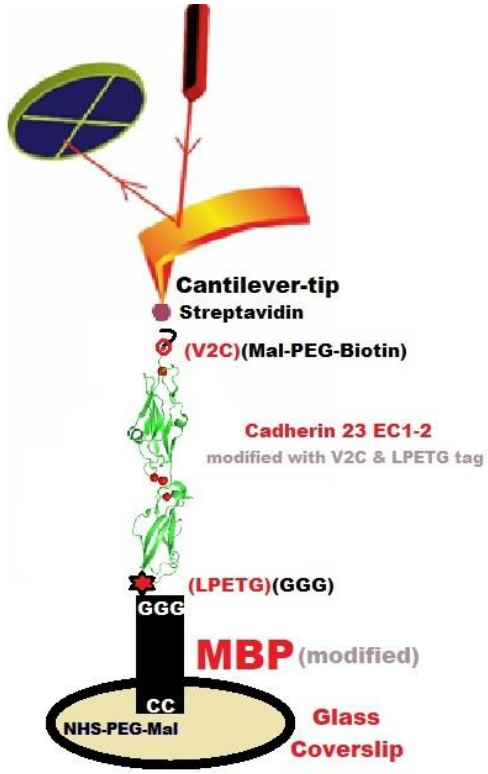

Figure 15: Schematics of the Chimeric construct of force-ramp experiment.

Both CL and NCL model protein was modified with V2C mutation and sort tag (LPETG) in N and C terminus respectively. MBP protein is modified with GGG and CC at different end. Using principle of Maleimide and cysteine reaction, LPETG and GGG reaction and streptavidin biotin interaction; this experiment was designed.

5.3. Procedure

1. This dynamic force ramp experiment is done to observe the unfolding pattern of a construct. There is only two quantifiable variable in this experiment Force and height which corresponds to Unfolding force (pN) and contour length (nm) respectively.
2. The cantilever was brought down to the surface at 2000 nm s^{-1}, a incubation time interval of 0.5 s was given for every time the experiment was repeated for the proteins to interact.
3. The cantilever was retracted at velocities varying from 1000, 2000, 5000, and 10000 nm s^{-1}. At each pulling velocity, 4590 force-curves were measured.

4. After each measurement at a single velocity, the thermal noise was calculated so that the spring constant of the modified cantilever can be measured. The spring constant of the modified cantilevers falls in the range of 50-54 pN nm^{-1}.

5.4. Results and Discussion

The saw-tooth pattern was observed due to domain unfolding in force curves. These force curves were fitted using worm like chain (WLC) model of polymer elasticity (equation 2).

Equation for the WLC model is as follows:

$$F = \frac{k\beta T}{Lp}\left[4\left(1-\frac{x}{Lc}\right)^{-2} - \frac{1}{4} + \frac{x}{Lc}\right] \qquad \text{(Equation 2)}$$

Here, unfolding force is F, contour length is Lc, and persistence length is Lp, Bolzmann constant is kβ, temperature is T and the extension is represented by x[21]. Unfolding parameters like unfolding force (F) and contour length (Lc) were calulated.

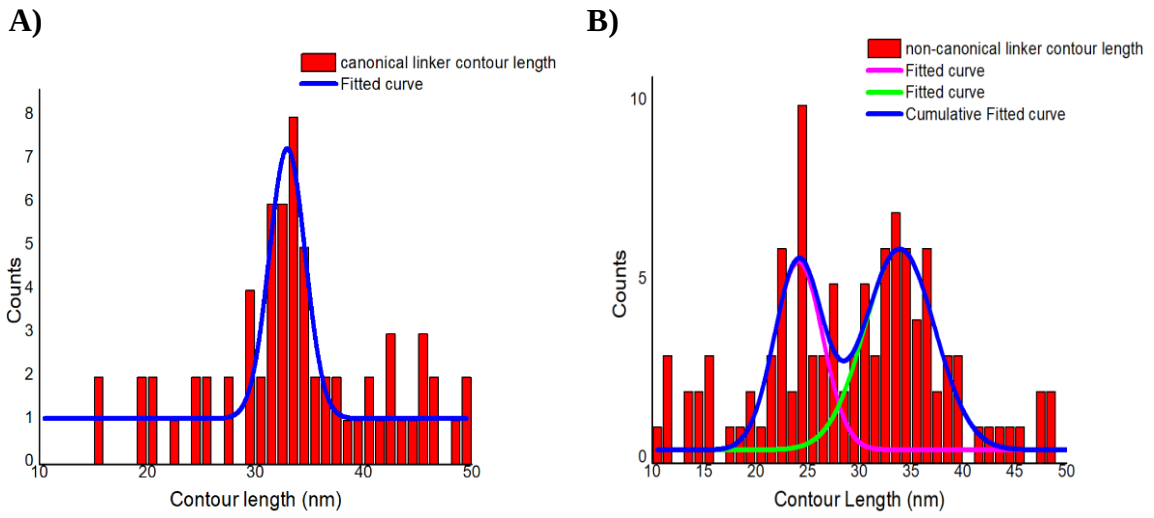

Figure 16: (A) Contour length histogram for CL depicting a 33 nm peak. (B) Contour length histogram of NCL depicting 24 and 34 nm peaks.

NCL exhibits smaller force induced extension along with entire domain extension. It is also prone to smaller extension in the range of 10-16 nm than CLcontaining CDH23 EC1-2.

CHAPTER 6

6.1. Conclusion

In this dissertation, we investigated the difference between CL and NCL. In the first chapter we introduced the significant sequence similarity in the IDLs and classified these linkers into two types. In next chapter we designed a construct by mutating a CL to a NCL, based on the natural NCL composition and essential amino acids present in CL to bind to the Ca^{2+} and we expressed CDH23 EC1-2 protein with both CL and NCL. In the third chapter, we compared the differential binding of Ca^{2+} using calcium chelator titration method and found out that unlike CL which binds to 4 Ca^{2+}, the NCL protein binds to 3 Ca^{2+}. In the 4^{th} chapter tryptophan fluorescence study in Steady-state fluorescence; CD with temperature melting study and chemical denaturation study confirmed that the NCL model protein is thermodynamically less stable than CL model protein and more prone to aggregation in higher stress. In the subsequent chapter, procedure and results are shown from single molecule experiments, which were performed using AFM. NCL linker model showed both smaller and bigger unfolding contour length distribution unlike CL. This means upon force induction, NCL linker unfolds in smaller contour length and might play a crucial role in force propagation through tip-link.

6.2. Future Works

Unfolding and refolding rate using electro-magnetic tweezer at a single molecular level will give us insight about the kinetic stability of this linker. Simulations are being performed using NAMD, Gromacs to elucidate the conformational characteristics and the possible force propagation pathway through NCL and CL model. In addition, hetero-FRET experiments will be performed to observe the flexibility of the protein to see the effect of the bent conformation in the linker.

BIBLIOGRAPHY

1. Jan Siemens, Concepcion Lillo, Rachel A. Dumont, Anna Reynolds & Ulrich Müller. Cadherin 23 is a component of the tip link in hair-cell stereocilia. *Naturevolume 428*, pages950–955 (2004)
2. Zubair M. Ahmed, Richard Goodyear, Saima Riazuddin, Ayala Lagziel, and Thomas B. Friedman. The Tip-Link Antigen, a Protein Associated with the Transduction Complex of Sensory Hair Cells, Is Protocadherin-15. *Journal of Neuroscience* 28 June 2006, 26 (26) 7022-7034; DOI: https://doi.org/10.1523/JNEUROSCI.1163-06.2006
3. Yoshie Narui and Marcos Sotomayor* Tuning Inner-Ear Tip-Link Affinity Through Alternatively Spliced Variants of Protocadherin-15. *Biochemistry*. 2018 March 20; 57(11): 1702–171. doi:10.1021/acs.biochem.7b01075
4. Ruishuang Geng1, Marcos Sotomayor2, Kimberly Kinder1, Suhasini R. and Kumar N. Alagramam1,6,7,* Noddy, a mouse harboring a missense mutation in protocadherin-15, reveals the impact of disrupting a critical interaction site between tip-link cadherins in inner ear hair cells. *J Neurosci.* 2013 March 6; 33(10): 4395–4404. doi:10.1523/JNEUROSCI.4514-12.2013.
5. Alagramam KN[1], Yuan H, Kuehn MH, Murcia CL & Smith RJ. Mutations in the novel protocadherin PCDH15 cause Usher syndrome type 1F.PMID: 11487575 DOI: 10.1093/hmg/10.16.1709
6. Ahmed ZM[1], Riazuddin S, Bernstein SL, Ahmed Z, Wilcox ER. Mutations of the protocadherin gene PCDH15 cause Usher syndrome type 1F.PMID: 11398101 PMCID: PMC1226045 DOI: 10.1086/321277
7. Di Palma F[1], Holme RH, Bryda EC, Belyantseva IA, Pellegrino R, Kachar B, Steel KP, Noben-Trauth K. Mutations in CDH23, encoding a new type of cadherin, cause stereocilia disorganization in waltzer, the mouse model for Usher syndrome type 1D.PMID: 11138008 DOI: 10.1038/83660
8. Bork JM[1], Peters LM, Riazuddin S, Bernstein SL, Morell RJ. Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23.PMID: 11090341 PMCID: PMC1234923 DOI: 10.1086/316954

9. Kazmierczak P[1], Sakaguchi H, Tokita J, Wilson-Kubalek EM, Kachar B. Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. PMID: 17805295 DOI: 10.1038/nature06091

10. Furness DN[1], Katori Y, Nirmal Kumar B, Hackney CM. The dimensions and structural attachments of tip links in mammalian cochlear hair cells and the effects of exposure to different levels of extracellular calcium. *Neuroscience*. 2008 Jun 12;154(1):10-21. doi: 10.1016/j.neuroscience.2008.02.010. Epub 2008 Feb 19.

11. Sloan-Heggen CM[1,2], Bierer AO[1], Shearer AE[1], Kolbe DL[1], Smith RJH[3,4,5]. Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss. *Hum Genet.* 2016 Apr;135(4):441-450. doi: 10.1007/s00439-016-1648-8. Epub 2016 Mar 11.

12. Astuto LM[1], Bork JM, Weston MD, Askew JW, Kimberling WJ. CDH23 mutation and phenotype heterogeneity: a profile of 107 diverse families with Usher syndrome and nonsyndromic deafness. PMID:12075507 PMCID:PMC379159 DOI:10.1086/341558

13. Baux D[1], Faugère V, Larrieu L, Le Guédard-Méreuze S, Roux AF. UMD-USHbases: a comprehensive set of databases to record and analyse pathogenic mutations and unclassified variants in seven Usher syndrome causing genes. *Hum Mutat.* 2008 Aug;29(8):E76-87. doi: 10.1002/humu.20780.

14. Maiko Miyagawa, Shin-ya Nishio, Shin-ichi Usami. Prevalence and Clinical Features of Hearing Loss Patient with CDH23 Mutation: A Large Cohort Study. Published: August 10, 2012 https://doi.org/10.1371/journal.pone.0040366

15. Powers RE[1], Gaudet R[2], Sotomayor M[3]. A Partial Calcium-Free Linker Confers Flexibility to Inner-Ear Protocadherin-15. 7;25(3):482-495. doi: 10.1016/j.str.2017.01.014. Epub 2017 Feb 23.

16. Raul Araya-Secchi, Brandon L. Neel & Marcos Sotomayor. An elastic element in the protocadherin-15 tip link of the inner ear. *Nature Communications* volume7, Article number: 13458 (2016)

17. M. Sotomayor, W.A. Weihofen, R. Gaudet, D.P. Corey. Structural determinants of Cadherin-23 function in hearing and deafness. *Neuron*, 66 (2010).

18. Thompson JD[1], Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-

specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994 Nov 11;22(22):4673-80.

19. Lee PY[1], Costumbrado J, Hsu CY, Kim YH. Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp.* 2012 Apr 20;(62). pii: 3923. doi: 10.3791/3923.
20. Jagadish P. Hazra1 , Amin Sagar1 , Nisha Arora1 , Debadutta Deb1 and Sabyasachi Rakshit1,3. Broken force dispersal network in tip-links by the mutations at the Ca^{2+}-binding residues induces hearing-loss. *Biochemical Journal* (2019) 476 2411–2425 https://doi.org/10.1042/BCJ20190453
21. Anupam Singh, Vaibhav Upadhyay, [...], and Amulya Kumar Panda. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microb Cell Fact* 2015 Mar 25;14.41. doi: 10.1186/s12934-015-0222-8
22. ShwethaSrinivasan[a1], Jagadish.P.Hazra[a1], Gayathri S.Singaraju[a], DebaduttaDeb[a], SabyasachiRakshit[ab]. ESCORTing proteins directly from whole cell-lysate for single-molecule studies. *Anal Biochem* 2017 Oct 15;535:35-42. Doi: 10.1016/j.ab.2017.07.022
23. Valentina Cappello, Laura Marchetti, Paola Parlanti, Silvia Landi, Mauro Gemmi. Ultrastructural Characterization of the Lower Motor System in a Mouse Model of Krabbe Disease. *Scientific Reports* volume 6, Article number: 1 (2016)
24. Richard A. George, Jaap Heringa. An analysis of protein domain linkers: their classification and role in protein folding. Protein Engineering, Design and Selection https://doi.org/10.1093/protein/15.11.871
25. Alka Prasad , Huaying Zhao , John M. Rutherford , Nicole Housley, Susan Pedigo. Effect of linker segments on the stability of epithelial cadherin domain 2. https://doi.org/10.1002/prot.20657
26. Mily Bhattacharya, Neha Jain, and Samrat Mukhopadhyay* Insights into the Mechanism of Aggregation and Fibril Formation from Bovine Serum Albumin. https://doi.org/10.1021/jp111528c