

Molecular Evolution and Structural Insights of the Podocyte Slit-diaphragm Proteins: Emphasis on Podocin

A thesis submitted during 2020 to the University of Hyderabad
In partial fulfillment of the award of a Ph.D. degree in the Department of Biochemistry,
School of Life Science

by

SANDEEP KUMAR MULUKALA NARASIMHA



Department of Biochemistry
School of Life Sciences

University of Hyderabad
(P.O.) Central University, Gachibowli, Hyderabad – 500 046
Telangana



CERTIFICATE

This is to certify that the thesis entitled "**Molecular Evolution and Structural Insights of the Podocyte Slit-diaphragm Proteins: Emphasis on Podocin**", submitted by **Sandeep Kumar Mulukala Narasimha**, bearing Reg. No **17LBPH13** in partial fulfillment of the requirements for the award of Doctor of Philosophy in Biochemistry is a bonafide work carried out by him under my supervision and guidance.

The thesis has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

P. Anil Kumar

Signature of the Supervisor

Dr. P. ANIL KUMAR

Assistant Professor

Department of Biochemistry

University of Hyderabad

Hyderabad-500 046. India.

K. S. T. L. N.

Head of the Department

HEAD

Dept. of Biochemistry
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046.

P. Anil Kumar

Dean of the School

DEAN

School of Life Sciences
University of Hyderabad
Hyderabad - 500 046.

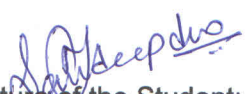


DECLARATION

I, **Sandeep Kumar Mulukala Narasimha**, hereby declare that this thesis entitled **"Molecular Evolution and Structural Insights of the Podocyte Slit-diaphragm Proteins: Emphasis on Podocin"**, submitted by me under the supervision of **Dr. Anil Kumar Pasupulati**, is a bonafide research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma

Date: 29th OCTOBER 2020

Name: **Sandeep Kumar Mulukala Narasimha**


Signature of the Student:

Regd. No. 17LBPH13



CERTIFICATE
(for Ph.D. Dissertation)

This is to certify that the thesis entitled “**Molecular Evolution and Structural Insights of the Podocyte Slit-diaphragm Proteins: Emphasis on Podocin**”, submitted by **Sandeep Kumar Mulukala Narasimha**, bearing Reg. No **17LBPH13** in partial fulfillment of the requirements for the award of Doctor of Philosophy in Biochemistry is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other university or institution for award of any degree or diploma.

Further, the student has the following publications before the submission of thesis/monograph for adjudication and has produced evidence for the same in the form of reprint in the relevant area of his research.

1. **N.S.K Mulukala**, K. Krishan, S.S. Irukuvajjula, G. Kanchan; P. Venkatesu, V. Ramakrishna, A.K. Pasupulati: Podocin forms Higher-Order Oligomers via Dynamic Subunit Exchange - Biochem Biophys Rep, 2020. 23: p. 100774.
2. **N.S.K. Mulukala***, P.P. Kar*, V. Ramakrishna, A.K. Pasupulati: Intrinsically disordered regions mediate macromolecular assembly of the Slit diaphragm proteins associated with Nephrotic syndrome - Molecular Simulation., 2019; 45:8: 603-613
3. **N.S.K. Mulukala**, R. Nishad, L.P. Kolligundla, M.A. Saleem, N.P. Prabhu, A.K. Pasupulati: *In Silico* Structural Characterization of Podocin and Assessment of Nephrotic Syndrome-Associated Podocin Mutants – IUBMB Life., 2016; 68 (7): 578-88.

The student has attended the following conferences during the course of his Ph.D.:

1. “XI International conference on the biology of yeasts and filamentous fungi” (2019): School of Life Sciences, University of Hyderabad, Hyderabad. (27th November - 29th November).
2. National Seminar on “Biomolecular interactions in development and disease” (2019): Department of Biochemistry, University of Hyderabad, Hyderabad. (26th September - 28th September).

3. "Workshop on animal models in cancer research" (2017): School of Medical Sciences, University of Hyderabad, Hyderabad. (24th October - 25th October).
4. "Workshop on protein structure and drug discovery" (2017): School of Life Sciences, University of Hyderabad, Hyderabad. (27th August - 5th September).
5. GIAN workshop on "Latest methods in structural biology" (2016): Jawaharlal Nehru University, Delhi. (15th November – 25th November).
6. IKMC "Accelerating innovation" (2016): IKP Knowledge Park, Hyderabad. (24th and 25th October).
7. 44th National seminar on crystallography (2016): IISER Pune, Maharashtra. (10th July – 13th July).
8. Use of radiation technology in agriculture, health and research (2016): UoH-BARC, School of Life Sciences, University of Hyderabad, Hyderabad. (13th November).
9. BioQuest (2015): School of Life Sciences, University of Hyderabad, Hyderabad (23rd and 24th September).

Further, the student has passed the following courses towards fulfillment of coursework requirement for Ph.D.

Course code	Name	Credits	Pass/Fail
BC801	Research Methodology/ Analytical Techniques	4	Pass
BC802	Research Ethics, Biosafety, Data Analysis, and Biostatistics	4	Pass
BC803	Scientific Writing and Research Proposal	4	Pass

P. Anil Kumar

Supervisor

Dr. P. ANIL KUMAR

Assistant Professor
Department of Biochemistry
University of Hyderabad
Hyderabad-500 046, India.

K. S. Chandra

Head of the Department

HEAD
Dept. of Biochemistry
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046.

S. N. Srinivas

Dean of the School

DEAN
School of Life Sciences
University of Hyderabad
Hyderabad - 500 046.

Dedicated to

Keshava Phary. &

Radhika Dasoji,

Triveni Kanchana,

and

Ramak Mulukala

Acknowledgments

Finally, the time has come to thank all the people who have been very instrumental in helping me reach this stage and help shape my career.

First and foremost, I would like to express my deep gratitude to my supervisor Dr. P. Anil Kumar who is the main reason that I have got an opportunity to pursue my dream of doing a Ph.D. He was a constant support and a source of motivation both in my professional and personal life. He has been very instrumental in my growth as a researcher. I don't think words would make enough justice to say how grateful and in debt, I am to him. He is like a father figure to me and I will forever be indebted to him.

I offer my sincerest thanks to my doctoral committee members Prof. S. Rajagopal and Dr. Mohd. Akif for their valuable suggestions and also in some cases permitting to use their instruments to gain valuable data.

I thank the current and former heads of the Biochemistry department, Prof. Krishnaveni Mishra and Prof. Mrinal Kanti Bhattacharyya for their immense help before and during my Ph.D. Further, I also would like to thank Prof. Sharmistha Banerjee and Prof. Naresh Babu Sepuri along with the HoDs for their support and help in enrolling for the Ph.D. I will forever remember their assistance.

I am particularly grateful for the assistance given by Prof. Ramakrishna Vadreu (BITS-Hyderabad), Mr. Shivkumar Sharma (BITS-Hyderabad), Prof. Venkatesu Pannuru (DU), Mr. Krishan Kumar (DU), Dr. Kanchan Garai (TIFR-Hyderabad), Dr. Rajashekar Kadmuri, and Mr. Naresh (BITS-Hyderabad). Their support, guidance, and readiness to help me in collecting data, analyzing data, and guidance on how to improve my research is exemplary. Their willingness to give their time so generously is very much appreciated.

I would like to offer my special thanks to Dr. Eswar Reddem, Dr. Adinarayana Madasu, Dr. Veera Babu, and Dr. Sunil Parthasarathy for taking the time to teach, help, and answer all of my questions with great patience. Their support in the initial stages of my research career will never be forgotten.

I would also like to thank the non-teaching staff of the Department of Biochemistry Mr. Chary and Mr. Prabhu for providing administrative support. Further, I would like to offer my thanks to current and former labmates Prasanna, Dhanunjay, Krishna, Asish, and Vaishnavi, and Aishwarya for their support and help throughout my Ph.D. I would like to

extend my special thanks to Mr. Rajkishor Nishad, Ms. Prajakta, and Mr. Abrar. I have had some amazing moments with them. Without their presence, I think my Ph.D. journey would have been incomplete. I will miss the days we spent on white rocks.

Although thanking family members comes in the last, I know you all have been very instrumental in my journey. First and foremost I would like to extend my heartfelt thank to my spouse (Rajashree). Without her help, support, sacrifice, and motivation I don't think I would be able to reach this stage. She has made many sacrifices so that I can pursue research without any hinderance. I guess the term better half was not coined just like that if it weren't for people like her. I also, extend my deepest thanks to my parents for their undying support and encouragement to pursue my goal of becoming a scientist. If it weren't for them I don't think I would been able to pursue Ph.D. Without their financial assistance and constant support pursuing Ph.D would been forever a dream.

Last but not least I wish to thank Ms. Radhika Dasoju and Ms. Triveni Kancha. This thesis is dedicated to them. They are the people who have taught me everything that I know. If it weren't for their teaching and encouragment in everything I did. I would not have had the dream to do a Ph.D. They are the master sculptors of my career.

I wish to thank the almighty God for without your help without which I would have achieved nothing.

Sandeep Mulukala N.K.

Table of Contents

List of Abbreviations	i
Table of Figures	ii
List of Tables.....	v
Abstract.....	1
Chapter 1: Introduction.....	2
1.1. Kidneys: Structure and Function	3
1.2. Nephron: The structural and functional unit of the kidney.....	4
1.3. Podocytes.....	6
1.4. Nephrotic syndrome.....	7
1.5. The slit-diaphragm: Development and composition.....	8
1.5.1. Nephrin	11
1.5.2. CD2-associated proteins (CD2AP).....	12
1.5.3. TRPC6.....	13
1.5.4. Podocin.....	14
Chapter 2: Molecular evolution of slit-diaphragm proteins.....	18
2.1. Prelude	19
2.2. Materials and Methods.....	20
2.2.1. Identifying the orthologues	20
2.2.2. Protein alignment and phylogenetic analysis.....	20
2.2.3. Domain Analysis	20
2.3. Results.....	22
2.3.1. SD proteins are confined to vertebrates with fewer exceptions	22
2.3.2. SD proteins and its orthologs share conserved domains.....	25
2.3.3. IURs are conserved motifs in SD proteins and its orthologs.....	30
2.4. Discussion	32
2.5. Ancillary data	36
Chapter 3: IURs mediate the assembly of SD proteins	42
3.1. Prelude	43
3.2. Methodology	44

3.2.1. Identifying homologs and sequence analysis of the SD proteins	44
3.2.2. Model building, refinement, and stereochemical quality analysis.....	44
3.2.3. Building the macromolecular complexes	45
3.2.4. Molecular dynamics simulations and analysis	46
3.3. Results.....	47
3.3.1. SD proteins consist of conserved structured and unstructured motifs	47
3.3.2. Prediction, validation, and insights into the built models.....	49
3.3.3. IURs may mediate complex formation.....	51
3.3.4. Mutations induced subtle changes in the mutant models	55
3.3.5. Mutations in the SD proteins disrupted the complex formation	57
3.4. Discussion	61
3.5. Ancillary data	65
Chapter 4: Structural features and oligomeric nature of human podocin domain	71
4.1. Prelude	72
4.2. Material and Methods	73
4.2.1. Protein cloning, expression, and purification	73
4.2.2. Size exclusion chromatography and multi-angle light scattering analysis (SEC-MALS).....	74
4.2.3. Fluorescence spectroscopy.....	74
4.2.4. Circular dichroism (CD) spectroscopy	75
4.2.5. Calorimetric analysis	75
4.3. Results.....	76
4.3.1. Purification of Podocin domain.....	76
4.3.2. Oligomeric nature of the podocin domain	78
4.3.3. Structural features of the homo-oligomers of podocin domain.....	80
4.3.4. Structural stability of the podocin domain	81
4.3.5. Calorimetric analysis of podocin domain	82
4.4. Discussion	84
4.5. Ancillary data	88
Chapter 5: Summary	91
References....	94

List of Abbreviations

GFB	– Glomerular filtration barrier
GBM	– Glomerular basement membrane
SD	– Slit-diaphragm
NS	– Nephrotic syndrome
SRNS	– Steroid resistant nephrotic syndrome
ESRD	– End-stage renal disease
FSGS	– Focal segmental glomerulosclerosis
<i>NPHS1</i>	– Nephrosis 1
<i>NPHS2</i>	– Nephrosis 2
CD2AP	– Cluster differentiation 2 associated protein
TRP	– Transient receptor potential
ZO-1	– Zonula occludens 1
KIRREL1	– Kin of IRRE like 1
CINDR	– CIN85 and CD2AP related
TRPC	– Transient receptor potential cation channel
SPFH	– Stomatin, Prohibitin, Flotillin, and HflC/K
PHB	– Prohibitin
Pfam	– protein family
BLAST	– Basic local alignment search tool
IURs	– Intrinsically unstructured regions
IUR-BD	– Intrinsically unstructured region – binding domain\
Ank	– Ankyrin repeats
Ig	– Immunoglobulin
FN3	– Fibronectin 3
DSC	– Differential scanning calorimetry
DLS	– Dynamic light scattering
CD	– Circular dichroism
FL	– Fluorescence
SEC	– Size exclusion chromatography
SEC-MALS	– size exclusion chromatography- Multi-angle light scattering
MALDI-TOF	– Matrix-assisted laser desorption/ionization-time of flight
WT	– Wild-type
RMSD	– root mean square deviation

Table of Figures

Figure 1: Schematic representation of the blood flow in and out of a nephron and the process involved in the formation of urine.....	4
Figure 2: A diagrammatic representation of the glomerular filtration barrier showing the capillary endothelium with fenestrations, the glomerular basement membrane, and the podocytes.	5
Figure 3: Diagrammatic representation of the glomerular filtration barrier showing three layers of filtration namely; glomerular epithelial cells, glomerular basement membrane, and podocytes.....	6
Figure 4: Transmission electron micrograph representing the secondary foot-processes, GBM, and capillary lumen. The thin line connecting the secondary foot-process is the slit-diaphragm.....	8
Figure 5: Schematic representation showing the associations among the key SD proteins leading to the formation of macromolecular complexes.....	11
Figure 6: Distribution of key slit-diaphragm proteins and its orthologs across the vertebrates and invertebrates: A) nephrin, B) CD2AP, C) podocin, and D) TRPC6. Organisms with complete genome sequences were considered for the study.	23
Figure 7: Domain organization in nephrin and its orthologs. Different colors are used to represent the subsets of Ig domains namely: Ig3 as green, CD80-like C2-set as red, and IgV as blue.....	26
Figure 8: Domain organization in CD2AP and CINDR sequences.	27
Figure 9: Domain organization of podocin sequences of the chordate phylum	28
Figure 10: Domain organization in TRPC6, TRPC3, and TRPC7 sequences. The ankyrin repeats in the sequences are represented as pink rectangle, similarly yellow rectangle represents TRP domains, and teal represents ion transport domain.....	29
Figure 11: Intrinsically disordered regions (IURs) predicted by PSI-PRED for the slit-diaphragm proteins and its orthologs in various phyla: A) IURs in nephrin and orthologs, B) IURs in CD2AP and orthologs, C) IURs in podocin, D) IURs in TRPC6 and orthologs. The Y- axis been adjusted for enhanced clarity of the residues that predicted as IURs. A prediction value of ≥ 0.5 for a residue indicates a residue to be intrinsically unstructured.	31
Figure 12: Secondary structure content representation of the key SD proteins: A) nephrin sequence comprising of Immunoglobulin like, Fibronectin type III (FN3), and transmembrane domains, B) CD2AP sequence exhibiting SH3 domains and a transmembrane pore lining region, C) Position of the SPFH domain and a transmembrane region in the podocin	

sequence, and D) TRPC6 sequence exhibiting ANK repeats, TRP domain, transmembrane, and pore lining helices.....	47
Figure 13: Predicted models of key SD proteins: A) Nephrin threading model displaying immunoglobulin-like domains and a fibronectin domain, B) The CD2AP model displaying SH3 domains was predicted using hybrid approach encompassing homology and ab-initio predictions, C) Predicted podocin model with the central SPFH domain and the characteristic long helical tail, and D) TRPC6 <i>ab-initio</i> predicted model. The IURs in the predicted models are represented in pink color, whereas the ordered segments are represented as blue (nephrin), cyan (CD2AP), red (podocin), and orange (TRPC6). The pore lining regions and the transmembrane segments in the models are represented as gray in color.	50
Figure 14: Plots representing the stability of the predicted WT complex after Molecular dynamic simulations: A) C-alpha backbone RMSD of the complex, and B) radius of gyration plot indicating the compactness of the complex.....	52
Figure 15: Predicted macromolecular complex of major SD proteins: A) Ribbon model and B) space fill and ribbon models showing protein-protein docking of nephrin (blue), CD2AP (cyan), podocin (red) and TRPC6 (orange) models. The interfacial contacts between interacting proteins are represented by green color. The insets in the figures are close-up views of each interacting partners and their respective interfacial residues. C) Nephrin-CD2AP, D) CD2AP-Podocin, E) Nephrin-TRPC6, and F) CD2AP-TRPC6.	53
Figure 16: Alignments of the WT and their counterpart mutant models: A) nephrin (blue) and V822M nephrin mutant, B-E) WT podocin (red) with R138Q podocin mutant (pink), R229Q podocin mutant (green), and R229Q-A284V podocin double mutant (cyan), F) WT CD2AP (cyan) with K301M CD2AP mutant, and G) WT TRPC6 (orange) with R895L TRPC6 mutant (pink).	56
Figure 17: Surface models representing each mutant complex: A) V822M nephrin mutant complex, B) K301M CD2AP mutant complex, C) R138Q podocin mutant complex and D) R822L TRPC6 mutant complex.....	58
Figure 18: Cloning of podocin domain: A) The NPHS2 gene that encodes podocin is located on chromosome 1 at the locus q25.2. B) The region 376bp – 1050bp of the gene was amplified and cloned into pET22b bacterial expression vector.	76
Figure 19: Expression, and purification of podocin domain: A & B) Coomassie blue staining and immunoblotting His-Tag HRP-conjugated antibody were done to identify the expression and solubility of the podocin domain. C) The SDS-PAGE analysis of samples after affinity chromatography	

purification of the urea solubilized cell lysate. D) Trypsinization and MALDI-TOF/TOF analysis of the 27KDa band excised from the earlier SDS-PAGE gel confirmed the presence of the podocin domain. UI-uninduced culture, I-Isopropyl β, D - thiogalactopyranoside Induced culture, P-Pellet fraction, S-Soluble fraction, F/T- flow through, E1-E6: elution fractions with 0.4M imidazole, 1-4: elution fractions with 1M imidazole.....	77
Figure 20: Podocin domain forms higher-order oligomers: SEC-MALS analysis of the podocin domain for molecular mass determination. The molecular weight on Y-axis and the refractive index on the secondary axis and were plotted against elution time.	78
Figure 21: Polydispersity and hydrodynamic radius of the podocin domain homo-oligomers: The DLS was in tandem with the SEC-MALS. DLS analysis of the samples corresponding to peak observed in SEC-MALS (1-5) is represented as a number percentage vs size in nanometers curve plots marked in different colors (A-E). The corresponding correlograms of the samples (1-5) represented as correlation coefficient vs time in microseconds.	79
Figure 22: Podocin domain exhibits secondary structural elements and tertiary packing: A. Far-UV CD spectrum of Podocin domain at 25oC, pH 8 in 10mM potassium phosphate buffer supplemented with 150mM NaCl and 2mM β-mercaptoethanol. B. Near-UV CD spectrum representing the tertiary packing of the podocin domain. C. Fluorescence emission of the podocin domain recorded by exciting the protein at 287 nm.	80
Figure 23: Effect of temperature on podocin domain: A. Intrinsic tryptophan fluorescence for podocin domain was measured as a function of temperature (20oC to 95oC) by exciting the protein at 287 nm. B. Peak emission of the podocin domain at 335 nm is plotted over a temperature range of 20oC to 95oC. C. FarUV CD spectra (250 – 200 nm) of the podocin domain was acquired as a function of temperature (20oC to 95oC). D. The MRE values of the podocin domain at 222nm is plotted as a function of temperature (range: 20oC to 95oC) to monitor the changes in the protein structure.	82
Figure 24: DSC analysis of podocin domain: The plots represent the endothermic and exothermic transitions of the podocin domain over a temperature range from 296K/23oC to 368K/95oC. The initial DSC artifact near 293K/20oC in the DSC data is not represented in the plot. The analysis data represented as a dotted line in the plot after baseline subtraction. Each transition peak (1-5) in the plot is represented by a color line which is a result of peak deconvolution function.	83
Figure 25: Podocin domain majorly adopts 16-mer homo-oligomer: SEC of the purified podocin domain at 5oC showing different podocin domain oligomeric conformations. A pre-packed HiLoad 16/600 Superdex	

200pg (GE Healthcare) column was equilibrated with 10 mM potassium phosphate, 150 mM NaCl, and 2 mM β -mercaptoethanol (pH 8.0). 1ml of protein at 9 μ M concentration was applied to the column at a constant flow rate of 0.5ml/min. The elution volume of the protein was detected by monitoring the UV absorbance at 280 nm.....	88
Figure 26: Podocin domain exhibits secondary structural elements: Far-UV CD spectrum of Podocin domain at 5oC, in 10mM potassium phosphate buffer supplemented with 150mM NaCl and 2mM β -mercaptoethanol (pH 8.0).....	88
Figure 27: Effect of temperature on podocin domain: A) Intrinsic tryptophan fluorescence for podocin domain was measured as a function of temperature (5oC to 95oC) by exciting the protein at 287 nm. B) Peak emission of the podocin domain at 335 nm is plotted over a temperature range of 5oC to 95oC. C) FarUV CD spectra (250 – 200 nm) of the podocin domain was acquired as a function of temperature (5oC to 95oC). D) The MRE values of the podocin domain at 222nm is plotted as a function of temperature (range: 5oC to 95o) to monitor the changes in the protein structure.....	89
Figure 28: DSC analysis of podocin domain: The plots represent the endothermic and exothermic transitions of the podocin domain over a temperature range from 5oC to 95oC. The initial DSC artifact near 273K in the DSC data is not represented in the plot. The analysis data after baseline subtraction is represented as a dotted line in the plots, whereas the transition peaks in the data are represented as solid color lines.	90

List of Tables

Table 1: List of some of the key genes that encode the SD proteins, their function, and associated disorders caused due to mutations or knock(out/down) of the genes.	10
Table 2: Human slit-diaphragm proteins and their orthologs proteins NCBI accession IDs identified by reciprocal best hit method. Note: '\$' indicates the ortholog of the respective SD protein identified from the organism mentioned in the braces.	24
Table 3: IUR predictions for nephrin, CD2AP, TRPC6, and podocin by Genesilico-meta disorder and PSI-PRED servers.	48
Table 4: Ramachandran outliers for Nephrin, CD2AP, Podocin, and TRPC6 models.	51
Table 5: Predicted salt bridge and hydrogen bonds at the protein-protein interfaces of nephrin, CD2AP, podocin, and TRPC6 complex. The annotation in parenthesis indicates, C- Charged, P- Polar, and AP- Apolar, and the yellow highlighted residues are predicted to be part of intrinsically	

disordered regions, whereas the residues in bold and italics are predicted IDR-binding domains.	54
Table 6: RMSD deviations showed by the mutant models when compared to the WT nephrin, CD2AP, podocin, and TRPC6 models respectively.	57
Table 7: Binding energies, number of salt bridges, hydrogen bonds, and non-bonded contacts predicted at the protein-protein interfaces of wild type and mutant complexes.	59
Table 8: MALDI TOF/TOF analysis of the purified protein: Trypsinization of the purified band at 27 kDa and subsequence analysis by MALDI-TOF/TOF showed 5 peptide sequences. BLAST analysis of these sequences against the non-redundant proteins database of NCBI showed 100% similarity with the human podocin sequence. Note: '▲' indicates the site of digestion by trypsin at arginine and lysine residues in the sequence.	78
Table 9: Enthalpy values and the transition temperatures as noticed in the dynamic scanning calorimetry.	83

Abstract

Podocytes are crucial cells of the glomerular filtration unit and playing a vital role at the interface of the blood-urine barrier. Podocyte slit-diaphragm is a modified tight junction that facilitates size and charge-dependent permselectivity. Several proteins including podocin, nephrin, CD2AP, and TRPC6 form a macromolecular assembly and constitute the slit-diaphragm (SD). The integrity of the SD depends on the interactions among the proteins involved in the macromolecular complexes with the key players being nephrin, CD2AP, podocin, and TRPC6. Interestingly, the orthologs of the key SD proteins are also involved in the similar complex formation and related functions suggesting an evolutionary relationship between the invertebrate orthologs and vertebrate SD proteins. We, therefore, studied the evolution of nephrin, CD2AP, podocin, and TRPC6 proteins. Several studies indicate that mutations cause severe damage to the SD structure which leads to severe proteinuria. But the structural insights of how these proteins interact to form complex assembly and how mutations alter these interactions remain poorly understood. It has long been an interest of our laboratory to decipher the structural features and thermodynamic properties of the key SD proteins as the structural details of these proteins remain elusive. Greater understanding of SD architecture is limited by lack of structural details of the proteins that constitute SD. Trying to specifically address these intriguing questions, the following observations have been made as a part of my doctoral work. An intense investigation has been carried out to study the structure of major SD proteins, but we were successful in detailing podocin only.

Chapter 1: Introduction

1.1. Kidneys: Structure and Function

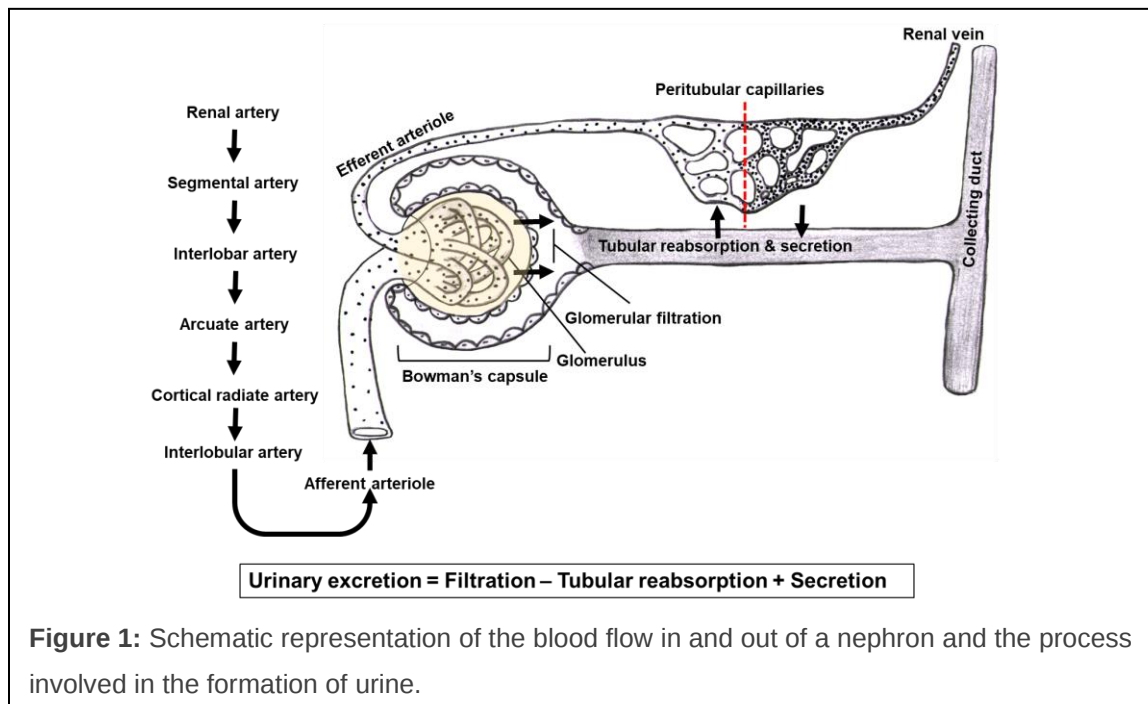
Vertebrate kidneys are two bean-shaped organs located on either side of the spinal cord in the retroperitoneal space. Kidneys contribute to the body homeostasis by regulating the water, acid-base, and electrolyte balance in addition to getting rid of toxic metabolic by-products [1]. The asymmetric placement of the kidneys can be attributed to the displacement caused by the liver over the right kidney due causing it to be placed lower in comparison to the left kidney present around T12 to L3 vertebrae. The dorsal side of the kidneys is partially protected by the 11th and the 12th ribs. Each kidney is about 110-140 mm long, 60mm wide, 40 cm thick, and weighs around 115-175 g depending on the gender [2]. Each kidney is covered by a dense, irregular connective tissue called a fibrous capsule. The kidneys are further covered in the renal fat pad (adipose tissue) which acts as a shock-absorbing layer and this is further encompassed by renal fascia [2]. Renal fascia helps the kidneys to firmly attach to the posterior abdominal wall.

Kidneys have convex (lateral) and concave (medial) borders. The renal hilum consisting of blood vessels, nerves, lymphatics, and ureters attaches to the concave border [2]. The transverse view of the kidney shows two zones, the outer cortex, and the inner medulla. These two regions are arranged into 5-8 pyramidal-shaped renal lobes, with each lobe separated by a renal column made up of connective tissue [2]. A part of the renal medulla and renal cortex make up a renal lobe that collects urine and terminates into renal papilla. The renal papilla from each lobe drains into a minor calyx and several minor calyces connect to form a major calyx. The major calyces join to form a single renal pelvis which later connects to the ureter. The kidneys are highly vascularized organs. The blood enters the kidney via the renal arteries branching from the descending aorta. The arteries divide and subdivide as represented in Fig.1 and enter into the functional unit of

the kidney called the nephron as afferent arteriole. The process of urine formation in the kidneys involves filtration, reabsorption, and secretion (Fig.1).

1.2. Nephron: The structural and functional unit of the kidney

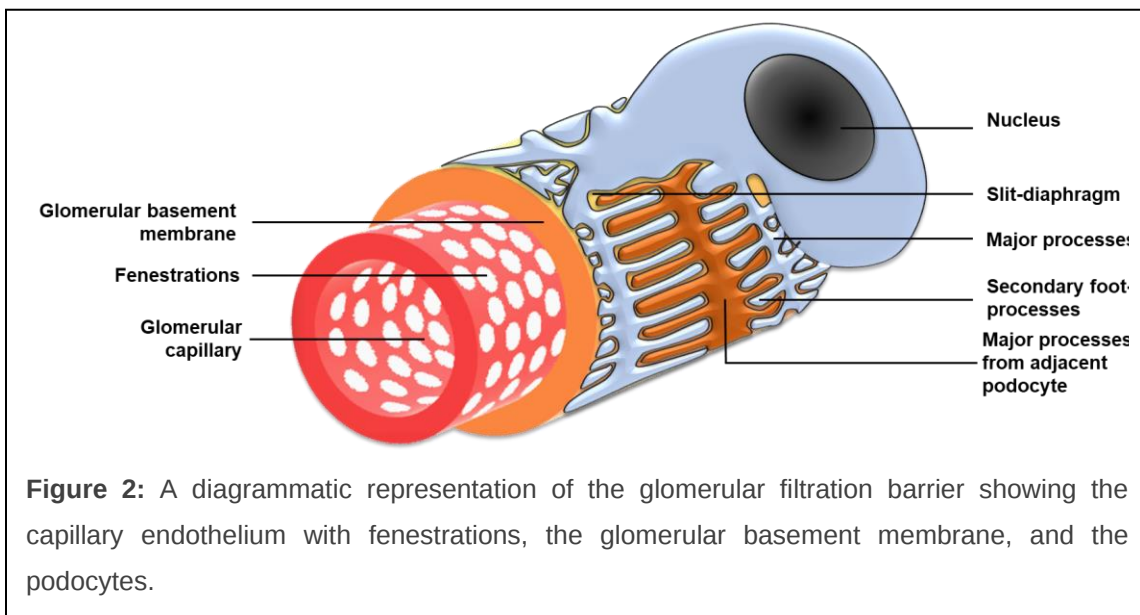
Each kidney comprises about 1-1.3 million nephrons spanning the cortex and the medulla regions. The renal corpuscle, the initial filtering portion of the nephron lies in the cortex while the renal tubule in of the nephron, which helps in the selective absorption of the filtrate starts from the cortex and passes deep into the medullary region (Fig.1). The renal corpuscle consists of the glomerulus and the bowman's capsule (Fig.1). The glomerulus and the renal tubule work in unison and regulate the final composition of the



urine [1]. The glomerulus consists of various cell types namely, the capillary endothelial cells, parietal epithelial cells, mesangial cells, and podocytes. The mesangial cells provide structural integrity to the glomerular capillary tuft, whereas the parietal epithelial cells are known to line the inner surface of the Bowman's capsule. Recent findings suggest that parietal epithelial cells play an important role in replacing the lost podocytes due to noxious

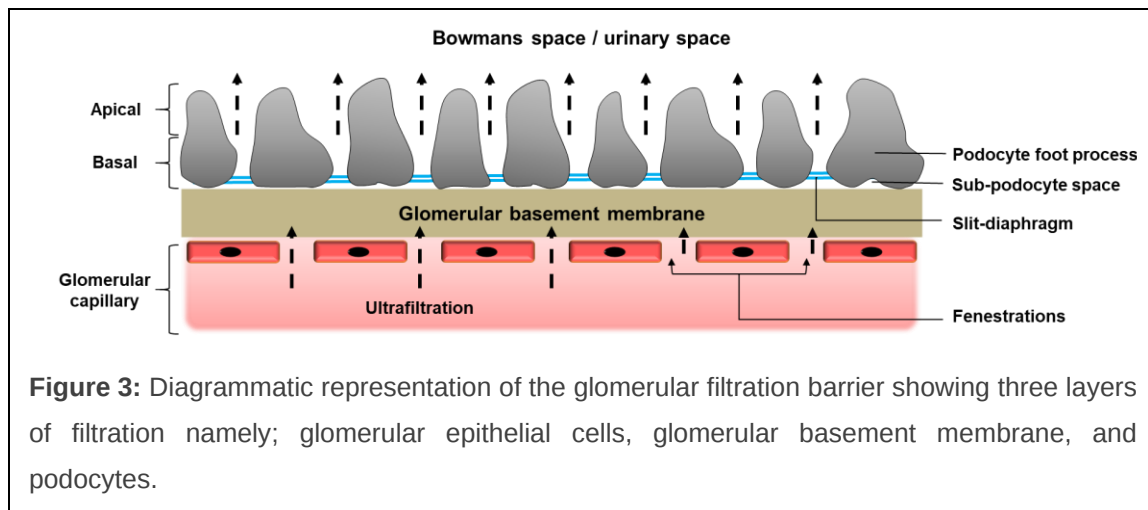
stimuli and also forms the basis of extra-capillary proliferative lesions and sclerosis in glomerular disease [3].

The glomerular filtration barrier (GFB) that acts as size, shape, and charge selective filtration barrier is anatomically comprised of three layers namely, capillary endothelium, glomerular basement membrane (GBM), and podocytes [4]. The glomerular endothelium or capillary endothelium consists of small pores called fenestrae of approximately 60-80 nm which facilitates the passage of fluid, blood plasma solutes, and protein while retaining red blood cells, white blood cells, and platelets. The GBM is made up of laminin-521 ($\alpha 5\beta 2\gamma 1$), collagen $\alpha 3\alpha 4\alpha 5(IV)$, nidogens-1 and -2, and agrin which are contributed by podocytes and capillary endothelium [5]. The typical GBM is about 250-400 nm thick and plays an important role in preventing the passage of proteins such as albumin and immunoglobulins [5]. The GBM is sandwiched between the capillary endothelium and the podocyte cells (Fig.2).



1.3. Podocytes

Podocytes are visceral epithelial cells of the glomerulus and they seek greater attention owing to their unique localization and their crucial role in glomerular biology. Podocyte maturation initially starts as cuboidal epithelial cells that undergo differentiation and give rise to several extensions thereby attaining nearly an octopus-like structure [6]. Each podocyte consists of a large cell body consisting of nucleus and organelles, finger-like projections called major process to arise from the cell body. The major process is majorly comprised of microtubules and intermediary filaments [7]. The major process further sub-divide into the secondary process which is made up of actin filaments [8]. The secondary process of adjacent podocytes interweaves to form a ladder-like structure called the slit-diaphragm (SD) [9-11]. Depressions of around 20-30 nm of unknown nature are also observed on the podocyte plasma membrane [12].



Similar to epithelial cells, the podocytes are highly polar and consist of apical and basal surfaces (Fig.3) [13]. The basal surface of the foot process is oriented towards the GBM while the apical side facing the urinary space is negatively charged due to the Glycocalyx coat enriched with sialoglycoproteins, with podocalyxin being the most abundant [14]. The negatively charged podocalyxin molecules create an electrostatic repulsion causing the neighboring foot-process to repel each other and keep the SD open

[14]. Thus, the GFB acts as a three-layer molecular sieve which is a size, shape, and charge selective structure that allows only positive ions, and minute-mid-sized solutes while retaining the anionic and macromolecules [15, 16].

The branched phenotype of the podocyte is largely held in place owing to the structural stability provided by the actin cytoskeleton. It is through the podocyte actin cytoskeleton; the podocytes provide support to the glomerular capillaries and ensure the pressure is required for filtration. The podocyte actin cytoskeleton consists of α -actinin, actin, myosin, talin, vinculin, and paxillin that act as a podocyte backbone [17]. The actin cytoskeleton of the podocyte foot processes is dynamic since they rapidly reorganize to meet any changes in filtration requirements [17]. It was earlier believed this process to be passive, however, increasing evidence suggested that actin cytoskeleton remodeling or breakdown is extensively dependent on signaling events [12, 17]. The actin cytoskeleton remodeling is subtle and limited however if podocytes undergo stress or injuries above a certain threshold, such as in cases of chronic hypertension or minimal change disease the podocyte foot process retracts (foot process effacement) leading to proteinuria [18, 19]. This pathological condition was later observed in hereditary podocytopathies [20].

1.4. Nephrotic syndrome

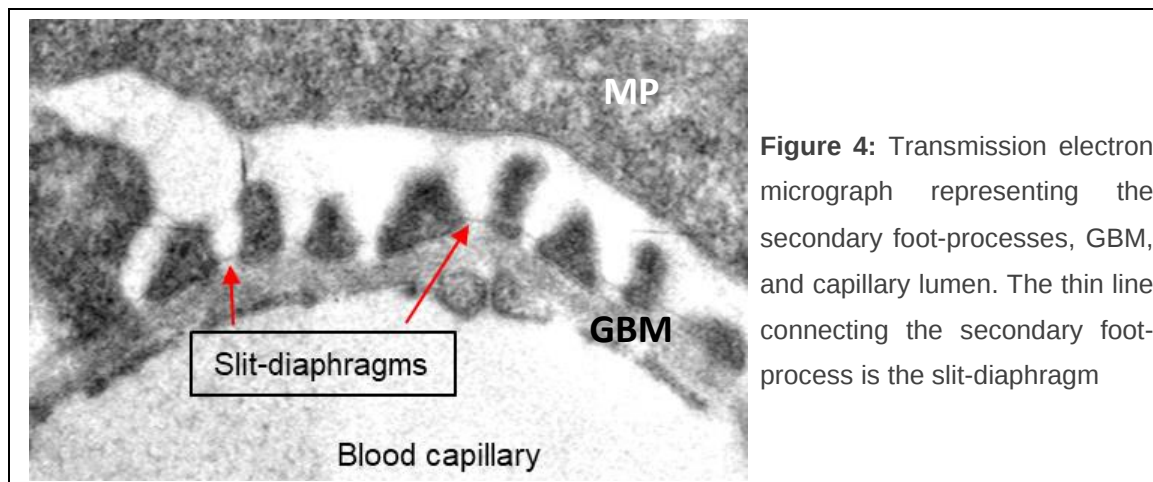
Altered podocyte morphology (effacement) leads to excess proteinuria, hypoalbuminemia, and edema which are hallmark symptoms of nephrotic syndrome (NS) [21]. The incidence of NS is about 20-40 per million worldwide but in India, it is around 90-100 per million [22]. Due to sustained loss of protein from the body ($>1\text{g/m}^2/24\text{hr}$), the NS patients have lower oncotic pressure resulting in edema initially in the lower body such as legs followed by anasarca in advanced stages. Other abnormalities in NS patients include increased levels of aldosterone and antidiuretic hormone [22]. Nephrotic syndrome can

occur either due to glomerular diseases (primary NS) or due to an existing condition such as diabetes that affects the kidneys along with other body organs (secondary NS).

The majority of the NS cases in children are primary with no known causative factors. These patients show minimal abnormalities in the glomerulus under a light microscope (minimal change NS). However, patients with congenital nephropathies present with lesions in the glomerulus which is a hallmark phenotype. In the former case, Corticoid therapy is the usual recourse to abate NS. However, patients with congenital nephropathy usually do not respond to the corticoid therapy leading to a condition called the steroid-resistant form of NS (SRNS). Proteinuria in SRNS patients typically onsets in infants during 0-3 months of age which eventually progresses to irreversible kidney failure within a decade. Congenital nephropathy patients display mutations in the proteins that constitute the SD.

1.5. The slit-diaphragm: Development and composition

The SD serves as size, charge, and shape-selective barrier preventing the plasma proteins to filter into urine averting proteinuria (Fig.4). The SD formation occurs in several stages wherein, it develops initially as a tight junction during the formation of comma and S-shape stages of the glomerular development [23]. Eventually, as the glomerular development progresses, the primary SD structure evolves into a modified tight-adherens



junction (Fig.4) [24]. It is observed that proteins from both tight and adherens junction co-localize at the SD alongside neuronal junction proteins such as nephrin, Kin of IRRE like-1 (KIRREL1), etc. [24]. The SD width ranges from 20-50 nm which is just sufficient to curb the passage of proteins from blood into the urinary space [25]. The SD cytoarchitecture is maintained by a complex interplay of protein-protein associations and unique signaling pathways [26]. Although the functional relevance of all the proteins at the SD remains elusive, the unique ensemble of different junction proteins enables them to perform the following tasks: 1) act as macromolecular filters, 2) anchor the SD to the GBM, 3) maintain the SD architecture via adaptor proteins that are connected to the foot-process actin cytoskeleton, and 4) participate in signaling events that regulate the plasticity of foot-process [24]. It has been shown that the integrity of the SD cannot be attributed solely to the structural proteins and unique signaling events but the activity of the associated ion channels also dictates SD integrity [27].

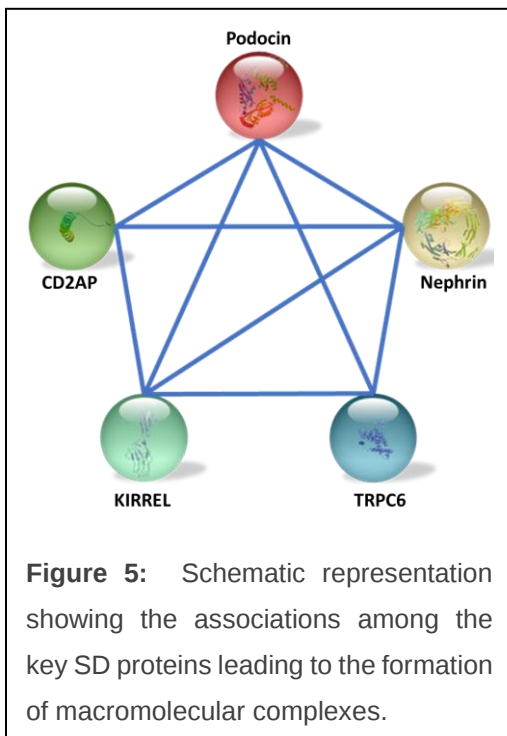
Remarkably, the SD structure is found to be well conserved in Chordates ranging from pronephros of *Danio rerio* to mammalian kidneys. Further, in some invertebrate organisms' like the nephrocytes of *Drosophila melanogaster* structure similar to the SD was observed [24]. Nephrocytes consist of homologs of the SD proteins which associate into intracellular diaphragms that filter large hemolymph constituents from the labyrinthine channels [28, 29]. Unlike the podocytes, the filtrate in the nephrocytes is endocytosed and stored within the nephrocytes [24]. Despite thorough research efforts, the protein composition at the SD remains poorly understood nevertheless table.1 list some of the key SD proteins and their role. Studies have shown that several proteins belonging to tight and adherent junctions have been identified at the SD [30-36].

Table 1: List of some of the key genes that encode the SD proteins, their function, and associated disorders caused due to mutations or knock(out/down) of the genes.

Gene name (chromosome location)	Protein encoded	Protein size (kDa)	Protein type and function	Disorders
<i>NPHS1</i> (Nephrosis 1) (19q 13.12)	Nephrin	185 to 200	Transmembrane protein and The backbone of the Slit diaphragm	Congenital nephrotic syndrome of the Finnish type and rapid progression to end-stage renal disease
<i>NPHS2</i> (1q 25.2)	Podocin	42	Transmembrane protein Helps in maintaining the Slit diaphragm integrity. Protein-protein interactions	Steroid-resistant nephrotic syndrome and rapid progression to end-stage renal disease.
<i>CD2AP</i> (Cluster differentiation associated protein 2) (6P 12.3)	CD2AP	74	Adapter protein crucial for the SD architecture development Attachment of nephrin podocin to the actin filaments of the podocyte cytoskeleton	Sporadic congenital nephrotic syndrome and focal segmental glomerulosclerosis (FSGS).
<i>TRPC6</i> (Transient receptor potential cation channel-6) (11q 22.1)	TRPC6	106	Transmembrane protein; calcium flux in the podocytes.	Increased calcium influx leads to the podocyte depletion due to the gain of function by mutations in TRPC6. Hyperactivation of the TRPC6 has deleterious effects on the podocytes not only in the FSGS but also in diabetic nephropathy.
<i>FAT1</i> (4q 35.2)	FAT 1	506	An adherent junction protein and an essential component of the SD required for maintaining the SD architecture.	Knockout mice showed severe proteinuria.
<i>KIRREL1</i> (Kin of IRRE like-1) (1q 23.1)	KIRREL 1	67	Transmembrane protein belonging to the Ig superfamily similar to nephrin. Protein-protein interactions and maintenance of the podocyte cytoskeleton.	<i>KIRREL1</i> gene knockdown resulted in the foot process effacement leading to proteinuria

<i>TJP1</i> (15q 13.1)	Zonula occludens-1	220	Cytosolic protein expressed on foot processes. Connects SD with the cytoskeleton. ZO-1 belongs to the family of the membrane-associated guanylate kinase (MAGUK) molecule. molecular scaffolds.	Injury by ischemia led to the loss of interactions between ZO-1 and KIRREL 1, resulting in the foot process effacement and proteinuria.
<i>Jam4</i> (21q 22.2)	Junctional adhesion molecule 4 (JAM4),	93	Transmembrane protein helps in cell adhesion. JAM4 is expressed in podocytes and detected at the slit diaphragm. MAG1 interactions with nephrin and JAM4 proteins of the SD lead to the formation of the tripartite structure at the SD and thus contributes effectively towards the construction of the SD.	Not known

Although several proteins constitute the SD nephrin, CD2AP, TRPC6, and podocin form the crux of SD architecture by forming a large macromolecular complex that provides



structural stability to the SD and also participates in signaling events that dictate the podocyte morphology (Fig.5).

1.5.1. Nephrin

In humans, nephrin is 1241 amino acids transmembrane protein encoded by *NPHS1*. Nephrin was the first SD protein to be identified as a product of the mutated gene in patients with Finnish-type NS [25, 37]. Nephrin is a large molecule consisting of a long extracellular domain (~35 nm) containing eight

immunoglobulins (Ig)-like modules, a transmembrane segment, followed by a cytoplasmic fibronectin type III (FN3) domain [25, 38]. Initially, nephrin was believed to be a podocyte-specific protein, however recent studies showed that nephrin helps in the development of coronary arteries in human and mice embryos [37, 39]. Furthermore, nephrin expression was also observed in the pancreatic islet cells, beta cells, and lymphoid tissues [40-42].

The majority of Ig domains in nephrin are C2 types which are characteristic for proteins mediating cell-cell or cell-matrix interactions [14]. The KIRREL family (formerly NEPH/nephrin like) are structurally similar to nephrin, composed of five Ig domains [43]. KIRREL 1 and 2 are localized to SD exclusively since, injection of anti-KIRREL antibodies in mice led to proteinuria and foot process effacement [14, 44]. Nephrin along with KIRREL1 are critical structural components of the SD that bridge the gap between the adjacent foot process [14, 44]. It was shown that nephrin forms dimers with nephrin/KIRREL1 from adjacent podocyte foot process [44-46]. Apart from providing integrity to the SD, nephrin along with KIRREL1 also participates in signaling transduction [47]. The cytoplasmic tail of nephrin contains well conserved tyrosine residues that undergo phosphorylation by Fyn (an Src family non-receptor protein tyrosine kinase) [48, 49]. This step is essential for the proper functioning of the SD, since deletion Fyn encoding gene resulted in foot process effacement, and proteinuria. Apart from binding with KIRREL1, nephrin also interacts with CD2AP, podocin and TRPC6 via its cytoplasmic C-terminus [50-53]. Furthermore, it was reported that podocin interaction with cytoplasmic tail of nephrin augments nephrin signaling that dictates the SD integrity [49, 54].

1.5.2. CD2-associated proteins (CD2AP)

Human CD2AP is a 639 residue protein identified primarily as an actin-binding cytoplasmic ligand for CD2 in T-cells and natural killer cells [55-57]. Nevertheless, localization studies showed CD2AP also expresses at the SD along with nephrin and

podocin towards the cytoplasmic face of the SD [52, 58]. CD2AP is known to play a key role in SD integrity since CD2AP knockout mice developed nephrotic syndrome with heavy proteinuria [59]. CD2AP is a multidomain adapter protein that mediates the SD interactions with the actin cytoskeleton of the podocyte foot-process [57, 60]. The N-terminus of the CD2AP consists of three SH3 domains followed by a coiled-coil leucine zipper [61]. Wherein, the SH3 domains in CD2AP show a high affinity towards proline-rich (PXPXPR) regions in its binding partners and the coiled-coil regions facilitate its homo-oligomerization [38, 62]. Co-immunoprecipitation studies revealed that the C-terminal segment of the CD2AP has an affinity towards nephrin and podocin [52, 58]. However, the stoichiometry of the CD2AP-nephrin-podocin interactions is still not clear. Further CD2AP was also shown to assist in podocyte signaling since it mediates the interaction between nephrin and podocyte actin cytoskeleton [51].

1.5.3. TRPC6

TRPC6 is another significant molecule that localizes to the SD. TRPC6 belongs to the larger family of transient receptor potential (TRP) proteins [63]. The TRPC consists of 7 related proteins namely, TRPC1 to TRPC7 out of which except TRPC2 the rest of the channels expresses in podocytes [64, 65]. TRPC6 in humans is a 931 residue transmembrane protein that majorly helps in calcium influx into the podocytes when compared to the other channels [66].

TRPC6 is a tetrameric structure with cytosolic N- and C- terminus. Each monomeric unit of TRPC6 is made up of helices S1 to S6 of which helices S1-S4 bundle together and act as voltage sensor whereas, the S5 and the S6 loops are connected by the P-loop [67]. The N-terminus of the TRPC6 consists of highly conserved ankyrin (Ank) repeats followed by a coiled-coil domain [67]. The carboxyl terminus consists of a plethora of conserved domains namely, TRP domain, proline-rich motif, calmodulin domain, IP3

domain, and a coiled-coil domain [67]. The presence of Ank repeats the coiled-coil domain, and proline-rich regions help TRPC6 to interact with neighboring SD proteins. Coimmunoprecipitation has shown that TRPC6 interacts with the dynamin, a protein that helps in trafficking, cytoskeleton dynamics, and regulation [68]. Furthermore, TRPC6 also interacts with nephrin, podocin, and α -actinin-4. It is reported that TRPC6 interaction with the above partner proteins is crucial for the normal functioning of podocytes [53, 69-71].

Unlike the other SD proteins mutations in TRPC6 are seldom noted, nevertheless, the mutations in TRPC6 caused FSGS. Interestingly, mutations in other genes that encode SD proteins reportedly hyper activates the TRPC6 channel leading to deleterious effects on the podocytes [72]. Altered expression of TRPC6 was also noted in some cases such as membranous glomerulonephritis and in the minimal change disease which led to increased calcium influx into the podocytes [73]. The increased calcium influx leads to the devastating effects on the podocytes that are lost either by apoptosis or by detachment [74].

1.5.4. Podocin

Podocin is 383 amino acids protein belonging to the stomatin family of proteins. Positional cloning and expression analysis of the Podocin suggests that it is encoded by *NPHS2*, and is exclusively present at the podocyte SD [75]. Podocin was observed to be a frequently mutated gene accounting for ~18% of SRNS cases [75]. Podocin attaches to the inner surface of the podocyte plasma membrane via a short transmembrane segment (101-125 residues) thus adapting a hair-pin/hook-like structure with cytosolic N- and C-terminuses [75]. Several reports show that podocin consists of a well-conserved domain from 164-325 residues called the Stomatin, Prohibitin, Follitin, and HflC (SPFH) aka Prohibitin (PHB) domain [75]. The Stomatin family consists of evolutionarily conserved SPFH domain found from bacteria to mammals [76]. The homology exhibited by the

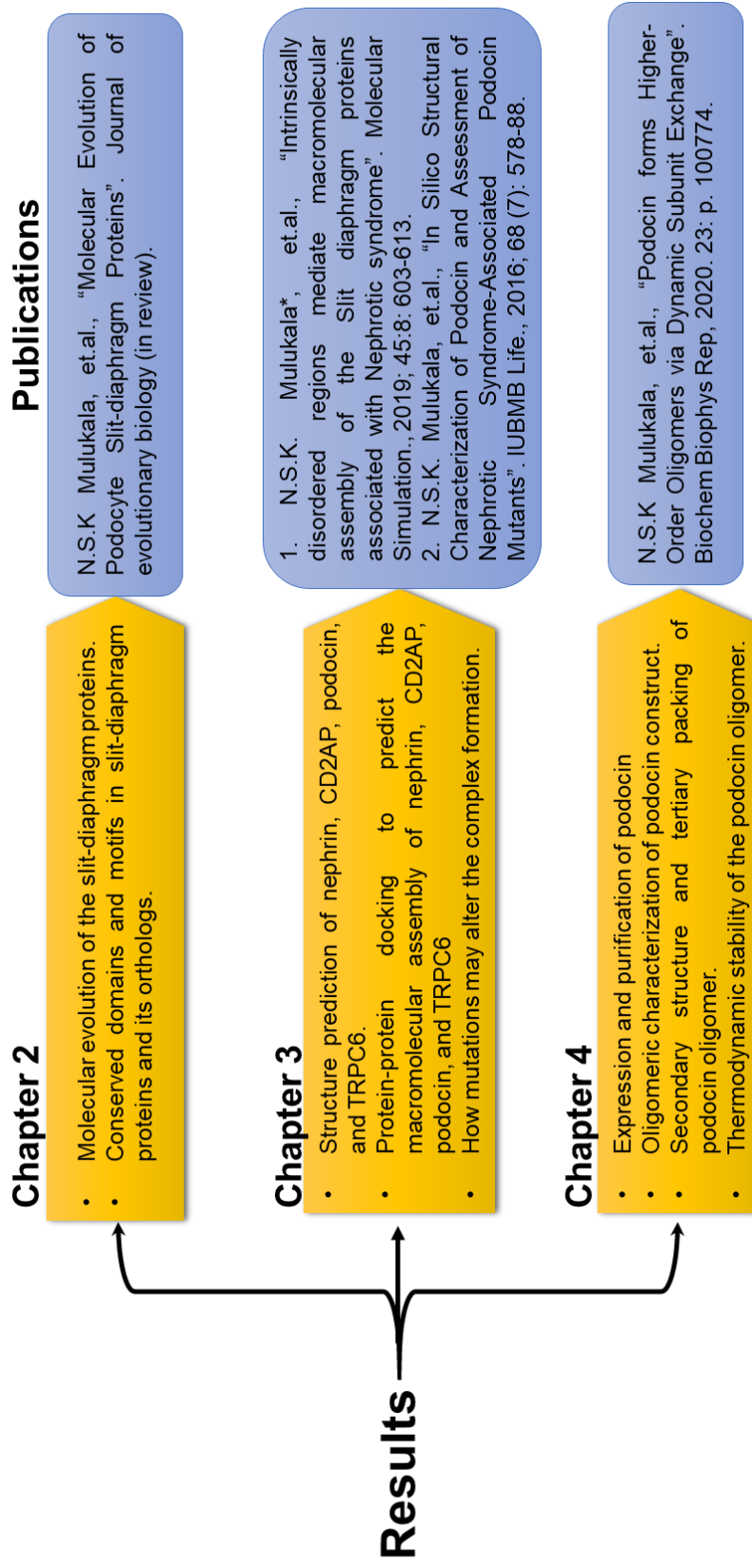
podocin is about 44% homology with the other members of the stomatin family via the SPFH domain and the rest of the sequence did not show homology with any known protein sequences [75]. The stomatin family members consist of highly conserved cysteine residues upstream of the transmembrane segment that undergoes palmitoylation [77]. It is suggested that the palmitoyl moiety facilitates the membrane insertion of the stomatin family proteins [78].

It is reported that podocin plays a crucial role as a scaffolding molecule at the SD due to its ability to interact with different podocyte proteins and mutations in the *NPHS2* gene led to the early onset of SRNS and complete loss of kidney function within a decade [75]. Co-immunoprecipitation studies have shown that the C-terminus of podocin binds to CD2AP and the cytoplasmic tail of nephrin [52]. Further, co-localization experiments have suggested that TRPC6 and podocin localized to the SD suggesting podocin-TRPC6 interactions [53, 70, 78]. Apart from the hetero-oligomers formed by podocin with its neighboring proteins, podocin like its family of proteins was shown to associate into an amalgam of homo-oligomers [79]. It is shown that different constructs of podocin associate with different oligomeric states [79]. Understanding the precise oligomeric conformation attained by the full-length podocin may help us understand its role in the formation of the macromolecular complex at the SD. Mutation in the *NPHS2* gene often leads to the improper targeting of the mutant podocin to the SD but rather was found to retain either in the endoplasmic reticulum, intracellular vesicles, or leads to the formation of the inclusion bodies [80]. It is yet to be elucidated how mutation affects the morphology of the podocin and thus is responsible for proteinuria.

Despite considerable research efforts, only the details regarding the localization and possible interactions between SD proteins have been elucidated. It is yet to be understood how these proteins though contain diverse domains work in unison to regulate the structure of the SD. Therefore, understanding the evolution of these proteins may

perhaps explain this question. Also, the structural information of these proteins, the information on how these proteins associate to form the macromolecular complex, and how mutations interfere with the complex formation are burning questions that need to be answered as providing vital clues to design therapeutic targets for preventing proteinuria. Our lab has since longed an interest to understand these details. We have therefore formulated the following objectives to specifically address these burning questions.

Overview of the thesis



Chapter 2: Molecular evolution of slit-diaphragm proteins

2.1. Prelude

Invertebrate organisms do not possess typical nephrons but possess nephron-like components. However, nephron-like components share several similarities to that of vertebrate nephrons. These observations suggest that the complexity of the vertebrate excretory system was inherited from their invertebrate systems. For example, the insect nephrocytes and the nephrons in the human kidney share several similarities [81, 82]. Furthermore, the orthologues of the major constituents of the SD were expressed in the nephrocytes of *Drosophila melanogaster*, and form a complex that closely mirrors the vertebrate SD complex [82]. The similarities between invertebrate nephrocytes and vertebrate podocytes suggest that these cells and structures are evolutionarily related that tingles the interest to investigate the evolution of the SD proteins and find the relevant orthologs in different metazoans. Our study was aimed to identify the orthologs of the nephrin, CD2AP, podocin, and TRPC6 across metazoans. We analyzed the domain composition and intrinsically unstructured regions (IURs) of the identified proteins to assess the evolutionary relationship of these orthologous proteins with the human SD proteins.

2.2. Materials and Methods

2.2.1. Identifying the orthologues

Four human SD proteins namely nephrin (NCBI: NP_004637.1), CD2AP (NCBI: NP_036252.1), podocin (NCBI: NP_055440.1), and TRPC6 (NCBI: NP_004612.2) were selected and investigated for their orthologs in metazoan organisms for which complete genome sequence is available. The reciprocal best hit approach with default settings in the pBLAST tool of the NCBI database was used to identify the orthologs. In cases where an ortholog of a protein could not be identified, the human ortholog identified from the next closest organism or phylum was used as a template to identify potential orthologs. The analysis was performed with the annotated protein sequences from 27 completely sequenced genomes.

2.2.2. Protein alignment and phylogenetic analysis

Multiple sequence alignment (MSA) of the retrieved orthologous proteins FASTA sequences from the NCBI database was performed using Multiple Sequence Comparison by Log-Expectation (MUSCLE) tool of the MEGA-X software. Default parameters such as gap penalty (2.90), hydrophobicity multiplier (1.20), and UPGMA cluster method were used to perform MSA. Phylogenetic trees were also derived for nephrin, CD2AP, podocin, and TRPC6 using the results from the respective MSAs. The maximum-likelihood method was used for constructing the phylogenetic trees using the default parameters in the MEGA-X software.

2.2.3. Domain Analysis

Various domains in the identified orthologs were predicted using the protein families (Pfam) database with an E-value cut-off of 1.0. Visual representation of the domain in the sequences was done using the illustrator of biological sequences (IBS) software ver. 1.0. Further, we analyzed the domain sequences using the PSI-PRED server

which uses a two-stage neural network to predict the secondary structure using data generated by BLAST. Furthermore, the server also extrapolated intrinsically unstructured regions (IUR's). According to the PSI-PRED server, prediction confidence of ≥ 0.5 indicates the residues as intrinsically unstructured. The IURs in the organisms were plotted as scattered plots using the OriginPro 2020 software.

2.3. Results

2.3.1. SD proteins are confined to vertebrates with fewer exceptions

We used the reciprocal best hit method to find the orthologs for human SD proteins in both the invertebrate and vertebrate phyla. The list of human SD proteins and their orthologs in various invertebrate and vertebrate phyla is tabulated along with the respective sequence accession numbers (Table.2). Our analysis revealed that nephrin is present in several organisms ranging from Tardigrada to higher vertebrates suggesting a diverse distribution (Fig.6A). In the case of nematodes and arthropods, we observed synaptogenesis protein-2 (*Caenorhabditis*) and sticks and stones protein (*Drosophila*) as the orthologs for nephrin. However, we could not identify the nephrin orthologs in the phyla Onychophora, Nemertea, Phoronida, Cyclostomata, and Chondrichthyes. It is interesting to note that Aves despite possessing the SD structure do not have nephrin or any orthologs of nephrin. We observed nephrin orthologs in the phyla Platyhelminths, Annelida, and Cephalochordata. Nevertheless, due to the limited genome sequence data availability, we could not identify the nephrin ortholog names in the above-mentioned phyla. Unlike nephrin; CD2AP, podocin, and TRPC6 were majorly restricted to the vertebrate phylum with few exceptions. We noticed that only CINDR (CIN85 and CD2AP related) protein in arthropods showed homology with CD2AP (Fig.6B). Similarly, the TRPC3 protein in Mollusca and Echinodermata and TRPC7-like protein in Hemichordates were found to be closely related to TRPC6 (Fig.6C). Although TRPC6 was predominantly restricted to vertebrates, it is also identified in Rotifers. Our analysis revealed that podocin was present only in the vertebrate organisms and we did not find podocin or its related proteins outside the vertebrate phylum (Fig.6D). These results suggest that although the distribution of nephrin is diverse, CD2AP, podocin, and TRPC6 are majorly restricted to vertebrates.

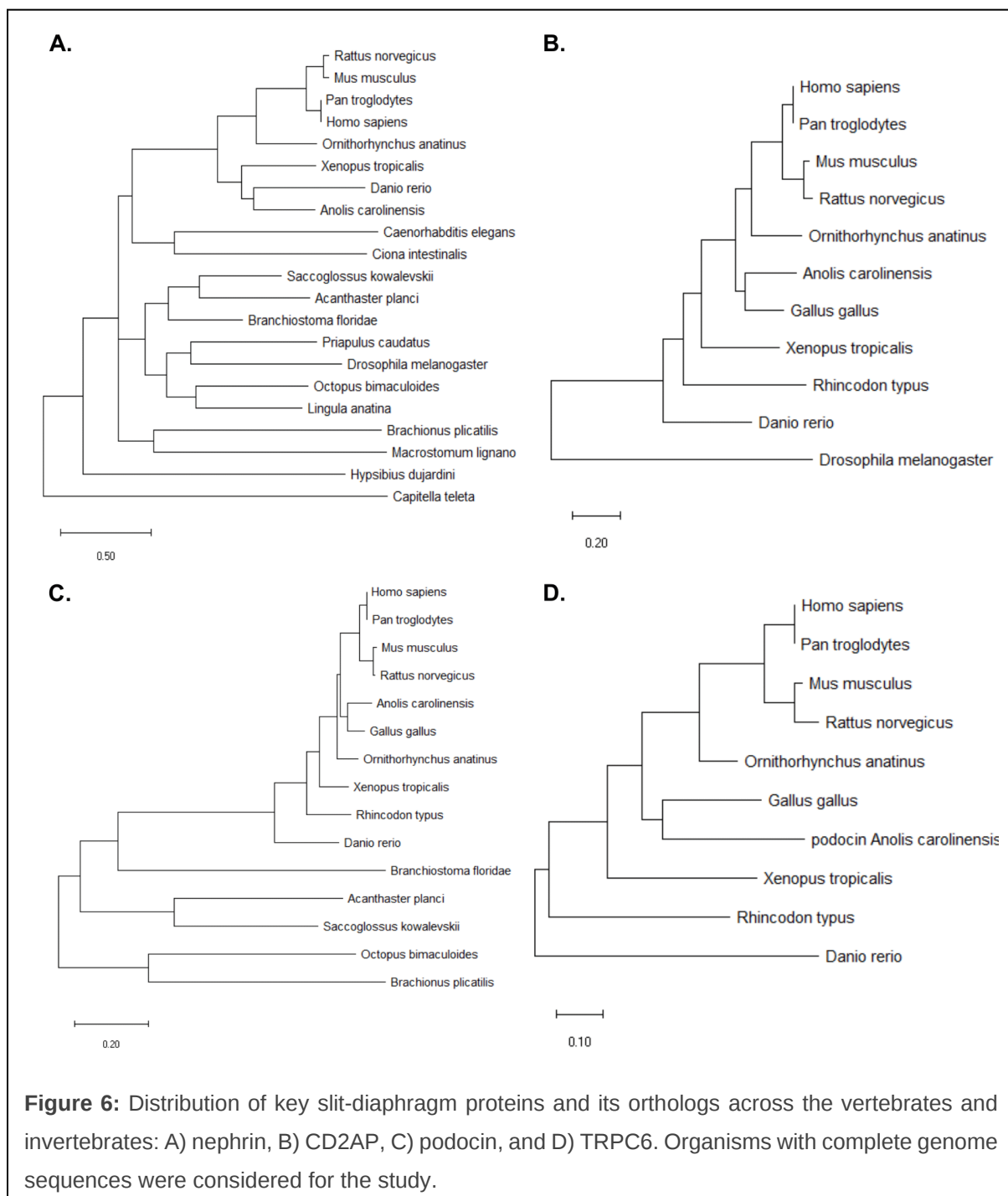


Table 2: Human slit-diaphragm proteins and their orthologs proteins NCBI accession IDs identified by reciprocal best hit method. Note: '\$' indicates the ortholog of the respective SD protein identified from the organism mentioned in the braces.

Phylum	Organism	Tax ID	Nephrin	CD2AP	Podocin	TRPC6
Priapulida	<i>Priapulid caudatus</i>	37621	XP_014667088.1	No relevance	No relevance	No relevance
Nematode	<i>Caenorhabditis elegans</i>	6239	NP_001309674.1	No relevance	No relevance	No relevance
Tardigrada	<i>Hypsibius dujardini</i>	232323	OQV25698.1	No relevance	No relevance	No relevance
Onychophora	<i>Euperipatoides rowelli</i>	49087	No relevance	No relevance	No relevance	No relevance
Arthropoda	<i>Drosophila melanogaster</i>	7227	NP_788286.1	NP_001263129.1	No relevance	No relevance
Rotifera	<i>Brachionus plicatilis</i>	10195	RNA22476.1	No relevance	No relevance	RNA42410.1 ^{\$} (Xenopus tropicalis)
Platyhelminthes	<i>Macrostomum lignano</i>	282301	PAA62947.1	No relevance	No relevance	No relevance
Mollusca	<i>Octopus bimaculoides</i>	37653	XP_014776688.1	No relevance	No relevance	XP_014783218.1 ^{\$} (Xenopus tropicalis)
Annelida	<i>Capitella teleta</i>	283909	ELU18150.1	No relevance	No relevance	No relevance
Nemertea	<i>Notospermus geniculatus</i>	416868	No relevance	No relevance	No relevance	No relevance
Brachiopoda	<i>Lingula anatina</i>	7574	XP_013380213.1	No relevance	No relevance	No relevance
Phoronida	<i>Phoronis australis</i>	115415	No relevance	No relevance	No relevance	No relevance
Hemichordate	<i>Saccoglossus kowalevskii</i>	10224	XP_006819647.1	No relevance	No relevance	XP_002730569.1 ^{\$} (Danio rerio)
Echinodermata	<i>Acanthaster planci</i>	133434	XP_022110423.1	No relevance	No relevance	XP_022103606.1
Cephalochordata	<i>Branchiostoma floridae</i>	7739	XP_002590121.1	No relevance	No relevance	XP_002607434.1 ^{\$} (Xenopus tropicalis)
Urochordata	<i>Ciona intestinalis</i>	7719	XP_002122747.1	No relevance	No relevance	No relevance
Cyclostomata	<i>Petromyzon marinus</i>	7757	No relevance	No relevance	No relevance	No relevance
Chondrichthyes	<i>Rhincodon typus</i>	259920	No relevance	XP_020370685.1	XP_020382509.1	XP_020383938.1

Osteichthyes	<i>Danio rerio</i>	7955	XP_01720 6503.1	NP_0010 08583.2	NP_0010 18155.2	XP_00516 1247.1
Amphibia	<i>Xenopus tropicalis</i>	8364	XP_03176 1552.1	NP_0011 21435.1	XP_01794 9096.1	XP_00293 5616.2
Reptilia	<i>Anolis carolinensis</i>	28377	XP_01685 1514.1	XP_00811 4796.1	XP_00322 5506.1	XP_00810 6273.1
Aves	<i>Gallus gallus</i>	9031	Absent	NP_0013 05332.1	XP_42226 5.3	XP_41718 4.4
Mammalia	<i>Ornithorhynchus anatinus</i>	9258	XP_02892 1034.1	XP_02892 8027.1	XP_00151 5734.2	XP_02890 3565.1
	<i>Rattus norvegicus</i>	10116	NP_0721 50.1	AAM4702 9.1	NP_5708 41.2	NP_4460 11.1
	<i>Mus musculus</i>	10090	NP_0623 32.2	AAI38375. 1	NP_5697 23.1	XP_00650 9912.1
	<i>Pan troglodytes</i>	9598	XP_01679 1216.2	XP_00944 9690.1	XP_01678 8875.1	XP_01677 7341.2
	<i>Homo sapiens</i>	9606	NP_0046 37.1 (Gene ID: 4868)	NP_0362 52.1 (Gene ID: 23607)	NP_0554 40.1 (Gene ID: 7827)	NP_0046 12.2 (Gene ID: 23607)

2.3.2. SD proteins and its orthologs share conserved domains

As we observed the presence of the SD proteins and their orthologs in various phyla predominantly from vertebrates, we next assessed the evolutionary accumulation and conservation of unique domains in these proteins. Human nephrin consists of eight immunoglobulins (Ig) domains and a fibronectin-3 (FN-3) domain [37]. Based on the sequence and number of strands in the β -sandwich of the Greek key motif (of Ig domain) we observed that human nephrin consists of one Ig5, five-CD80-like C2-set Ig2, two-Ig3 domains (Fig.7). Our analysis revealed that in the invertebrates both nephrin and its orthologs consist of more Ig5 and Ig3 domains as compared to the vertebrates wherein, CD80-like C2-set Ig2 domains are significantly more. For example, the sticks and stones protein in *Drosophila* is composed of two-Ig5, four-Ig3, and only three-CD80-like C2-set Ig2 domains (Fig.7). Interestingly, we noticed that in both *Octopus* (Mollusca) and *Branchiostoma* (Cephalochordate) nephrin sequence is devoid of the FN3 domain, though it has Ig domains (Fig.7).

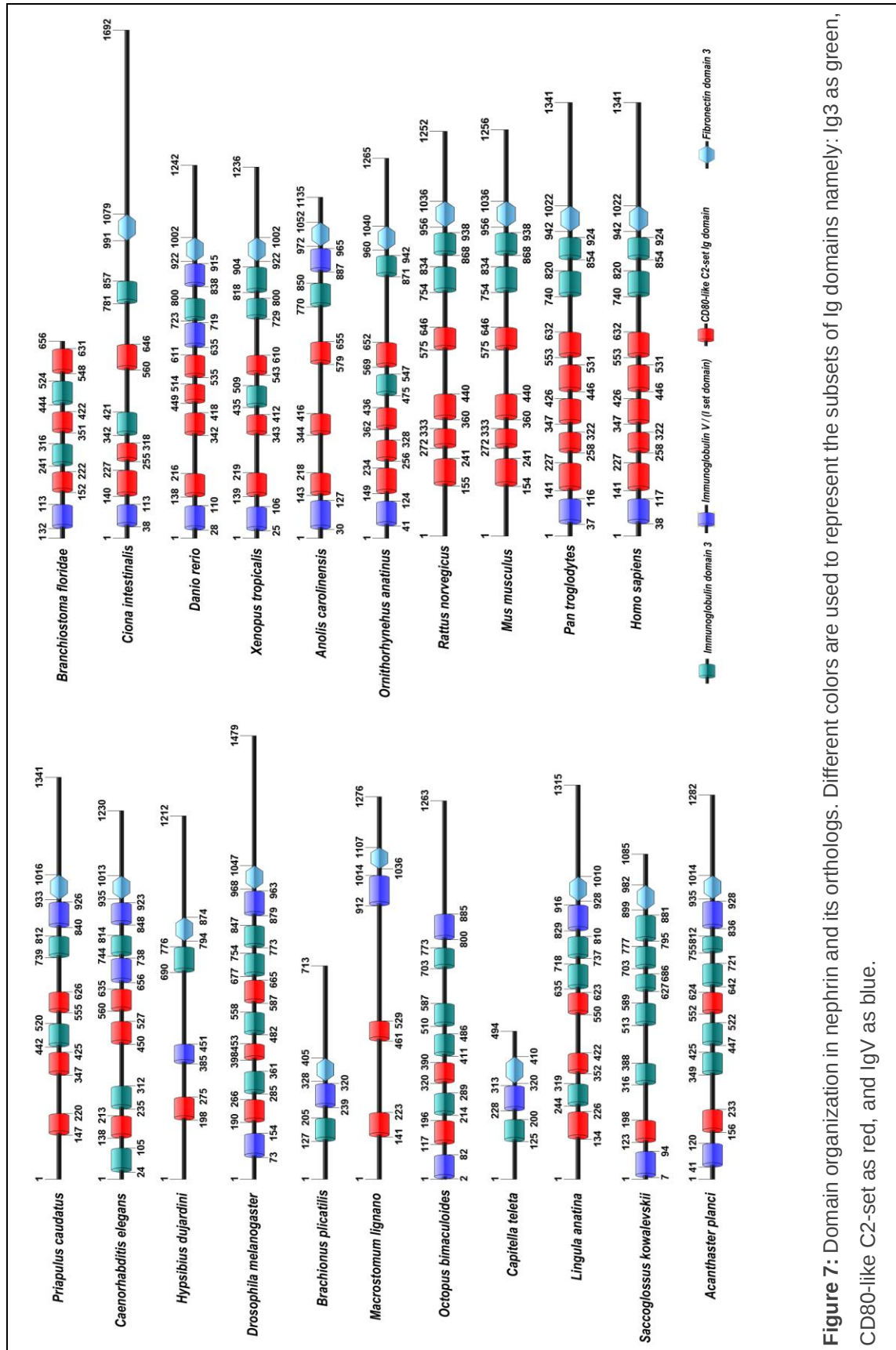
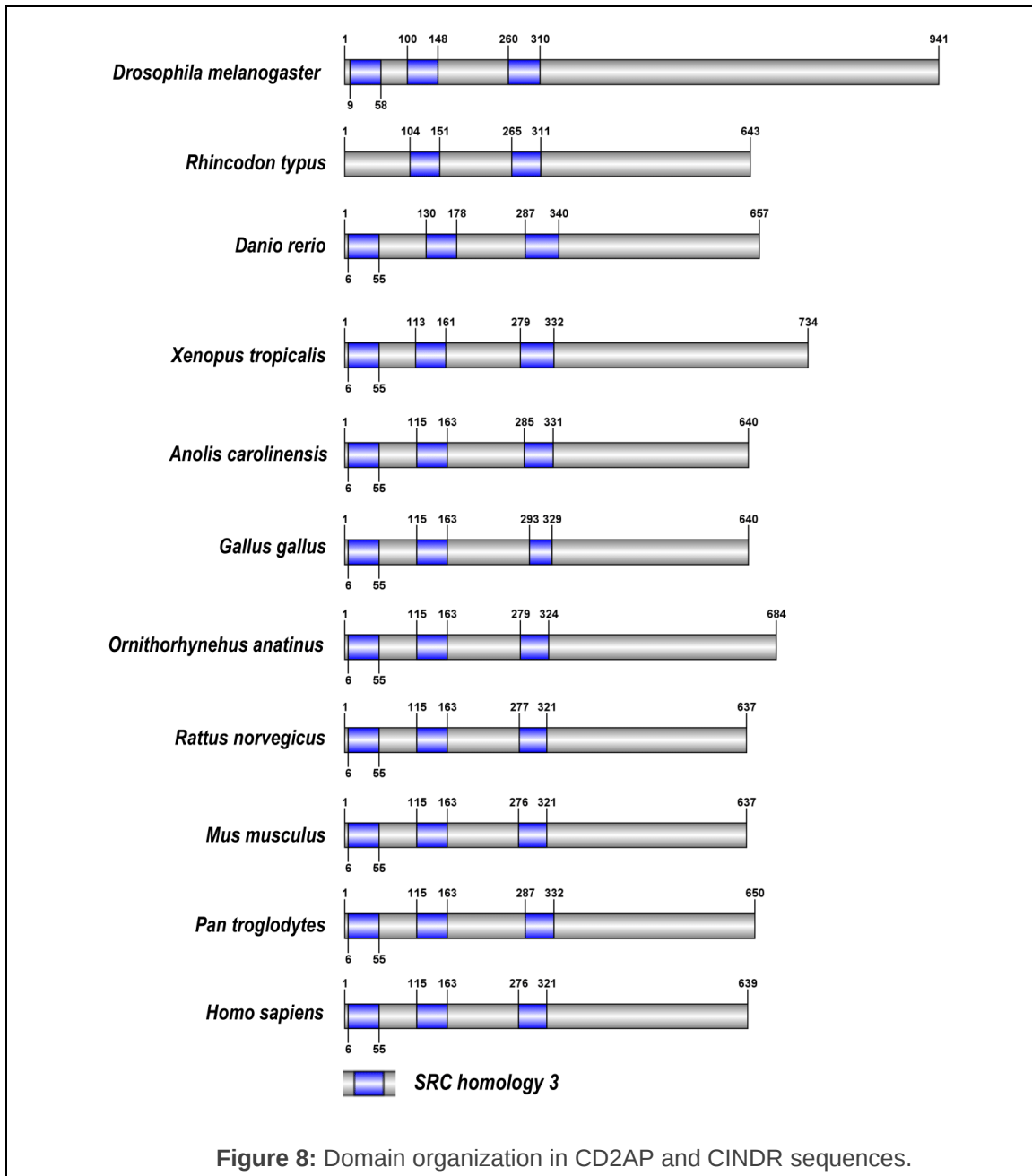


Figure 7: Domain organization in nephrin and its orthologs. Different colors are used to represent the subsets of Ig domains namely: Ig3 as green, CD80-like C2-set as red, and IgV as blue.



The human CD2AP protein consists of three Src Homology-3 (SH3) domains [83]. Pfam analysis of the CD2AP and the CINDR sequences showed that the three SH3 domains are highly conserved (Fig.8). Our analysis could not predict any other domains in either CD2AP or the orthologs of CD2AP (Fig.8). Podocin in humans consists of only an SPFH domain while the rest of the sequence does not contain any known domains (Fig.9) [75]. Similar to the human podocin sequence, the podocin sequences from the

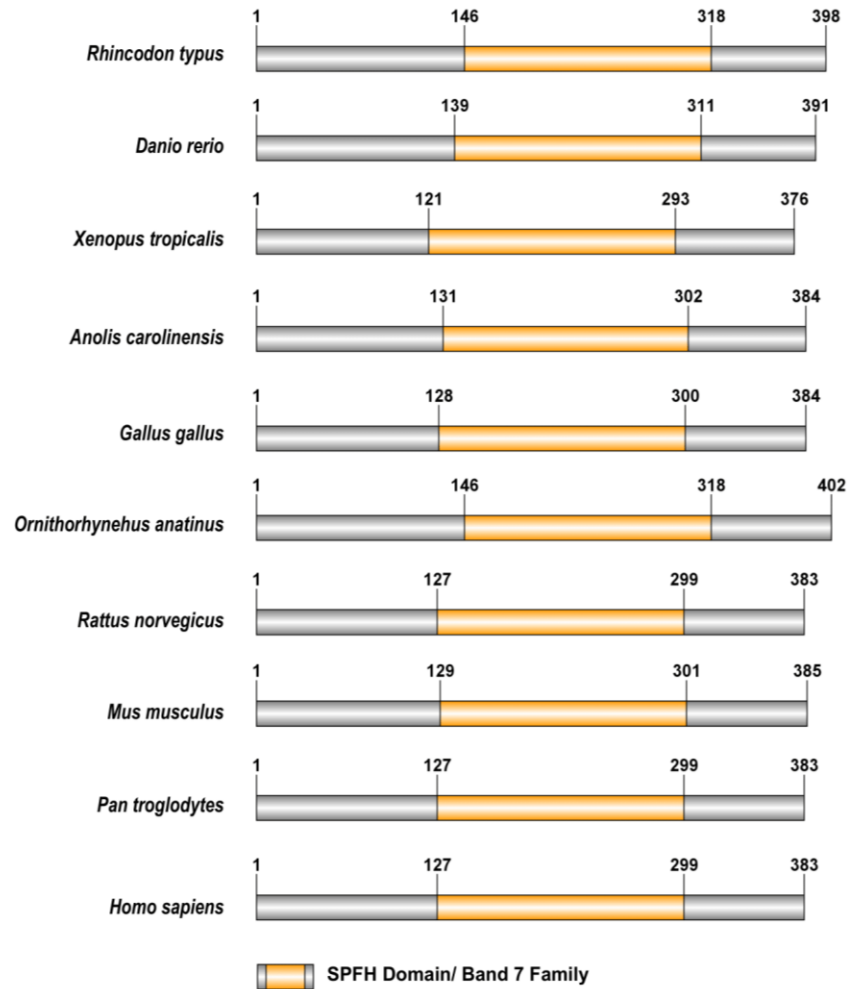


Figure 9: Domain organization of podocin sequences of the chordate phylum

other vertebrate organisms also had the SPFH domain suggesting that this domain is highly conserved among vertebrates (Fig.9). The human TRPC6 sequence consists of three conserved ankyrin repeats, a TRP domain, and an ion transport domain [38]. Our analysis of the TRPC6, TRPC3, and TRPC7 sequences showed that Ankyrin repeats, TRP, and ion transport domains were well conserved with a few exceptions (Fig.10). In the TRPC6 sequence of *Rhincodon typus*, we observed only the ion transport domain but not the ankyrin repeats and TRP domain, whereas, in *Acanthaster planci* the TRPC3

protein lacked the ankyrin repeats but possessed the ion transport and the TRP domains (Fig.10).

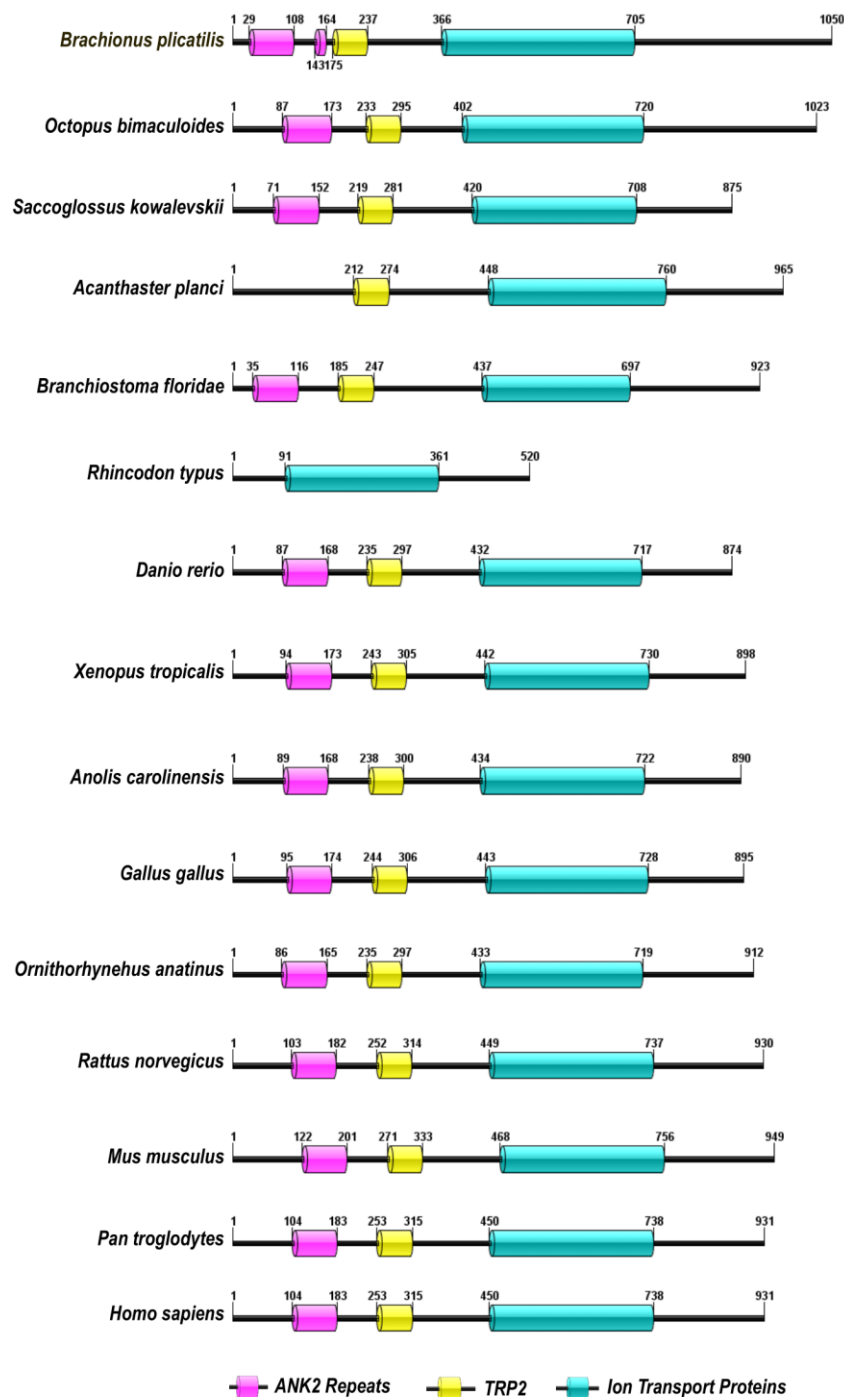


Figure 10: Domain organization in TRPC6, TRPC3, and TRPC7 sequences. The ankyrin repeats in the sequences are represented as pink rectangle, similarly yellow rectangle represents TRP domains, and teal represents ion transport domain.

2.3.3. IURs are conserved motifs in SD proteins and its orthologs

SD proteins are known to form large complexes via homo and heterophilic interactions which are believed to be mediated by IURs [21, 26, 38]. Since the SD proteins and its orthologs share similar domains and appreciable sequence similarity, we investigated if IUR motifs are conserved in the SD proteins and its orthologs. Our results showed that the majority of the nephrin and its orthologous sequences have IURs at both the N- and C-terminus as well as in few Ig domains (Fig.11A). However, the nephrin orthologs in Rotifera, Mollusca, and Annelida had IURs only at the C-terminus and in the Ig domains but not at the N-terminus (Fig.11A). Also, our analysis revealed that the nephrin sequence from amphibians had IURs even in the FN3 domain. In the invertebrates, Ig3 and Ig5 domains consist of most of the IURs while CD80-like C2-set Ig2 domains consist of few IURs. In the case of vertebrates, the vice versa was observed. In the case of CD2AP and CINDR sequences, IURs form the majority of the sequence (>50%) except for the regions that are a part of the first and the third SH3 domains (Fig.11B). But in the CD2AP of *Rhincodon typus*, all the three SH3 domains were devoid of IURs but the rest of the sequence had IURs. Analysis of the podocin sequences from the Chondrichthyes to Mammalia showed IURs at both the N- and the C-terminuses (Fig.11C). The distribution pattern of IURs in TRPC6 and its orthologs were similar to nephrin and its orthologs. Our results indicate that TRPC6 and TRPC6 orthologs possess IURs at both the N- and C-terminus with some intermittent regions in the sequences (Fig.11D). Nevertheless, the TRPC6 sequences in vertebrates (Amphibia-Mammalia) had IURs only at the N- and the C-terminus. These results suggest that IURs are conserved motifs among the SD proteins and its orthologs. Table.A1 in the ancillary section lists the detailed list of the residues predicted as IURs/IUR-BD in the nephrin, CD2AP, podocin, and TRPC6 sequences.

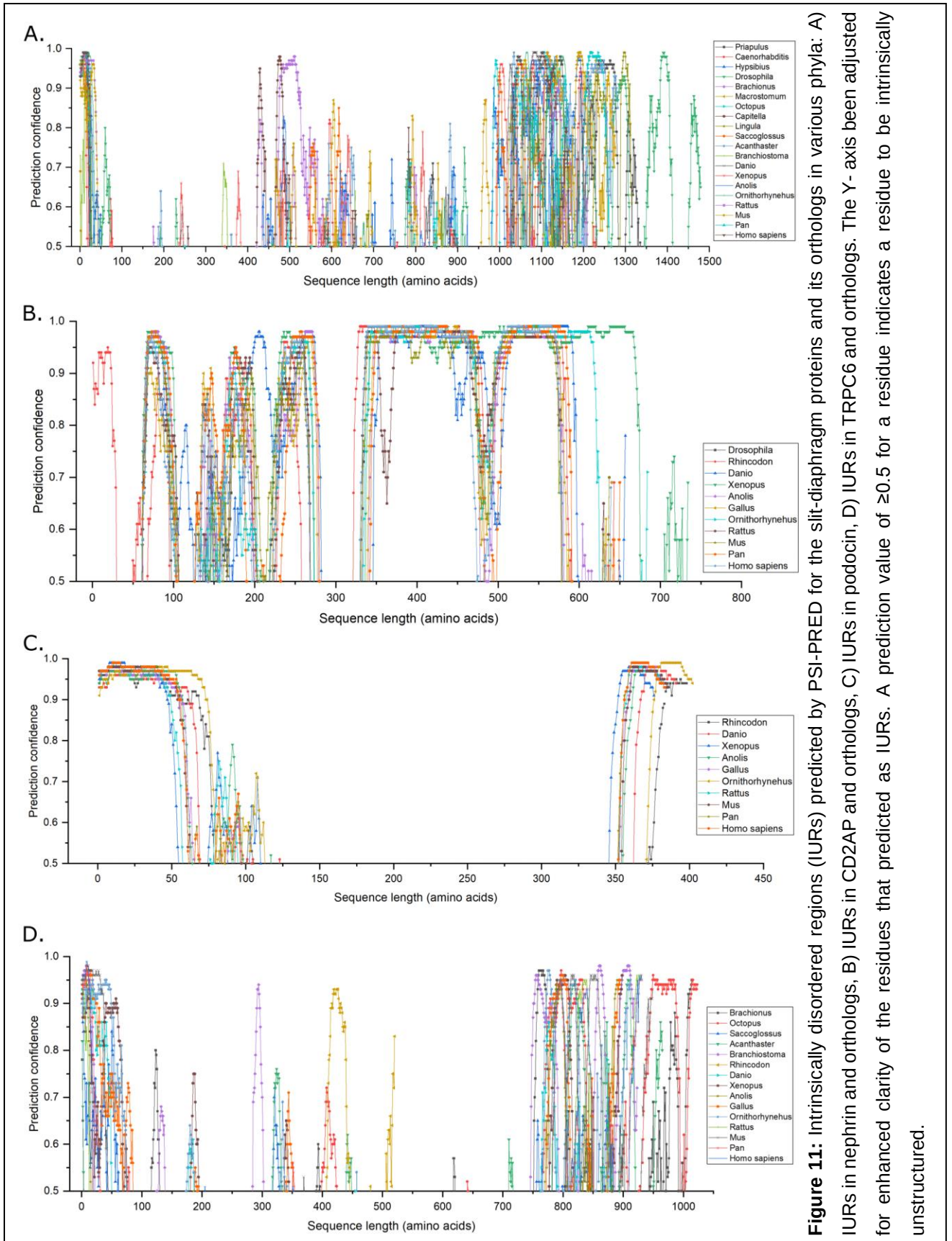


Figure 11: Intrinsically disordered regions (IURs) predicted by PSI-PRED for the slit-diaphragm proteins and its orthologs in various phyla: A) IURs in nephrin and orthologs, B) IURs in CD2AP and orthologs, C) IURs in podocin, D) IURs in TRPC6 and orthologs. The Y- axis been adjusted for enhanced clarity of the residues that predicted as IURs. A prediction value of ≥ 0.5 for a residue indicates a residue to be intrinsically unstructured.

2.4. Discussion

Primitive nephron-like structures were identified in many invertebrate phyla suggesting an evolutionary link between the vertebrate and invertebrate excretory units which further suggests that orthologous proteins similar to the vertebrate SD proteins may be present in other metazoans. Therefore, we investigated the molecular evolution of the major human podocyte SD proteins namely nephrin, CD2AP, podocin, and TRPC6 proteins that play a crucial role in aiding SD integrity. Our study showed that though nephrin is distributed across several phyla in both vertebrates and invertebrates, while the remaining three proteins we investigated are largely confined to vertebrates. SD proteins and its orthologs share several conserved domains and IURs indicating that the invertebrate SD proteins and its invertebrate orthologs are evolutionarily related.

In humans, nephrin is 1241 amino acids transmembrane protein made up of eight Ig domains and an FN3 domain [38]. Nephrin along with KIRREL1 forms the characteristic zipper-like structure bridging the gap between the adjacent foot process [44]. Mutations or knockdown of the gene encoding the nephrin caused Finnish-type congenital NS and improper development of coronary arteries in human and mice embryos [37, 39]. Nephrin expression was also observed in the pancreatic islet cells, beta cells, and lymphoid tissues [40-42]. Therefore, the divergent presence of nephrin in metazoans is observed in our analysis. Nevertheless, while analyzing the domains of the orthologs of nephrin, we noticed that in invertebrates the CD80-like C2-set Ig domains were less as compared to the nephrin sequences in vertebrates. It is known that nephrin forms dimers with nephrin/KIRREL1 from adjacent podocyte foot processes [44-46]. We speculate that the increased number of CD80-like C2-set Ig domains in vertebrates may facilitate stronger homophilic interactions between neighboring nephrin/KIRREL1 molecules when compared to Ig3 or Ig5 domains. Apart from providing integrity to the SD, nephrin along

with KIRREL1 also participates in signaling transduction. The cytoplasmic tails of both nephrin and KIRREL1 consist of several conserved tyrosine residues that undergo phosphorylation by Fyn (an Src family nonreceptor protein tyrosine kinase) [49]. This phosphorylation step is essential for the proper function of the SD, since deletion of the gene encoding Fyn resulted in abnormal filtration, podocyte foot process effacement, and proteinuria. Furthermore, the binding of podocin with nephrin augments nephrin signaling [49, 54].

Human CD2AP is a 639 residue protein and is primarily identified as an actin-binding cytoplasmic ligand for CD2 in T-cells and natural killer cells [55-57]. Localization studies showed CD2AP expression at the SD along with nephrin and podocin [52, 58]. CD2AP acts as an acting binding adaptor protein and helps in nephrin/KIRREL1 signaling in podocytes [51]. CD2AP knockout mice developed NS with heavy proteinuria suggesting the importance of CD2AP in SD integrity and podocyte permselectivity [59]. Although CD2AP shares ~50% similarity with CIN85, which belong to SH3 domain-containing kinase-binding protein 1 (SH3KBP1), our analysis identified only one ortholog for CD2AP in *Drosophila*, and it is known as CINDR. The adaptor molecules characteristically consist of three SH3 domains, proline-rich motif, and a coiled-coil region [83]. Similarly, we observed the signature SH3 domains in CD2AP. Interestingly, CINDR also consists of SH3 domains. Similar to CD2AP, CINDR is involved in cell adhesion, cytoskeleton dynamics, and synaptic vesicle trafficking in *Drosophila* [84, 85].

Positional cloning identified *NPHS2* encodes a podocyte exclusive 383 residues integral membrane protein called podocin [75]. Podocin acts as a scaffolding molecule and provides structural integrity to the SD by forming a macromolecular complex with nephrin, CD2AP, TRPC6, and KIRREL1 [52, 78]. It was reported that besides forming a heteromeric complex, podocin like its family members associates into higher-order oligomers [21, 79]. It is suggested that the numerous interactions formed by podocin with

its client proteins may be attributed to the podocin's oligomeric nature since each molecule in the podocin homo-oligomer can interact with one of its client proteins [38]. Podocin despite showing close homology (~40%) with the stomatin family proteins, our study could not identify the orthologs of podocin outside the vertebrate phylum. Further, the homologous region is restricted only to the SPFH domain, while the rest of the sequence did not correlate with any known domains.

The human TRPC6 is a 931 residues calcium ion transport channel protein associated with smooth muscle contraction, pulmonary endothelial permeability, neuronal protection against ischemia, and also the structure and function of podocytes [86]. TRPC6 and its related TRPC channels are a part of a larger family of TRP proteins involved majorly in chemo- and mechanosensation [87]. It was shown that the mutations in TRPC6 caused the late onset of the focal segmental glomerulosclerosis. In some cases, the mutations increased calcium ion influx into the cell leading to altered podocyte morphology, however, the mechanistic insights of how this occurs are still poorly understood [53]. Based on the sequence similarity, the TRPC6, TRPC3, and TRPC7 proteins share appreciable homology. Furthermore, the TRPC proteins share several conserved regions namely; a) Ankyrin (ANK) repeats, b) coiled-coil domain, c) a 25 residues TRP domain, d) proline-rich sequence, followed by e) a calmodulin and IP3 receptor-binding region (CIRB region), and f) C-terminal coiled-coiled domain [53]. Therefore, it is predictable that the reciprocal best-hit method retrieved TRPC3 and TRPC7 proteins as the orthologs of TRPC6 and also that these sequences share considerably conserved domains. TRPC6 besides being a component of the SD is a critical controller of the actin cytoskeleton. Activation of TRCP6 leads to the surge in cytoplasmic Ca^{2+} levels therefore TRPC6 overexpression leads to loss of actin stress fibers in podocytes, disruption of focal adhesions, and proteinuria in mice [88, 89]. In the case of nephrin, podocin, CD2AP mutations are associated with diminished function and

proteinuria. Decreased expression of nephrin and podocin manifest in pathology phenotype. Alternatively, elevated TRPC6 expression is associated with glomerulopathy [89].

In the vertebrates, the SD proteins are known to co-localize at the SD while interact and form a large macromolecular complex [70]. Further, it was suggested that SD proteins consist of IURs through which these proteins may be forming complexes. IURs are areas in the protein sequences that do not adopt any secondary structure conformation in isolation, however, in the presence of an interacting partner or a suitable ligand, IURs adopt ordered structures [90]. Furthermore, they are known to play a crucial role in mediating protein-protein interactions and signaling events [91]. Since SD proteins are predicted to have IURs, we were interested to know whether IURs are also conserved in the SD orthologs. Our results have shown that IURs are also conserved motifs across the SD orthologs which further ascertains the evolutionary relationship between SD proteins and its orthologs. Further, based on the location of IURs in nephrin, podocin, and TRPC6 IURs may promote both homomeric and heteromeric interactions.

2.5. Ancillary data

Table A1: Detailed list of amino acid residues predicted as IURs and IUR-binding domain (BD) by the PSI-PRED server in the SD proteins and its orthologs in various metazoans.

Intrinsically unstructured regions (IURs) predicted in nephrin and its orthologs			
Names of the organisms	NCBI Accession ID	IURs	IUR- (Binding domains)
<i>Priapulius caudatus</i>	XP_014667088.1	833;835-839;882-883;889;1030-1036;1078-1157;1197-1283;1300-1314;1316;1318-1322;1325-1326;1330-1331	01-24;834;840-844;1029;1037-1044;1315;1317;1323-1324;1327-1328;1335
<i>Caenorhabditis elegans</i>	NP_001309674.1	74-75;345;482-492;544-551;589-593;595-600;754-756;843;886-887;891-892;1026-1028;1082-1126;1224	01-17;71-73;76-78;594;601-604;883-885;889-890;1025;1029-1040;1073-1081;1184-1201;1125;1228-1230
<i>Hypsibius dujardini</i>	OQV25698.1	06-30;32-33;36;42-46;459;461;623-629;696-699;702-704;739;741-746;880-882;889-891;894;981-988;990-999;1027-1200	01-05;31;34-35;37-41;48-50;451-456;630-641;883-888;892-893;895;989;1000-1012
<i>Euperipatoides rowelli</i>	No relevance		
<i>Drosophila melanogaster</i>	NP_788286.1	775-781;912-922;1053-1074;1111-1112;1130;1132-1133;1138-1139;1141-1167;1178-1180;1187;1189-1191;1208-1213;1226;1266-1310;1345;1349-1412;1452-1456;1461-1472	01-25;47-51;53-73;228-232;782-787;923;1075;1110;1113-1129;1137;1140;1173-1176;1188;1214-1225;1227-1232;1450-1451;1457-1460;1473-1479
<i>Brachionus plicatilis</i>	RNA22476.1	175-176;421-423;470-573	424-434;576-580;618-624;630-633
<i>Macrostomum lignano</i>	PAA62947.1	24;477-481;554;557;559;596-612;614-616;618;628-630;687-694;697;785-797;846-847;849-857;862-863;956-975;1118-1119;1121;1145-1150;1178-1276	01-22;463-474;555-556;558;560-563;695-696;798-800;858-861;864;1123-1144;1151
<i>Octopus binaculoides</i>	XP_014776688.1	551-795-797;986-987;1034-1106;1111-1112;1116;1126-1127;1129;1195-1260	983-985;988-1000;1132-1141;1261
<i>Capitella teleta</i>	ELU18150.1	426;449	421-425;427-442;466-494
<i>Notospermus geniculatus</i>	No relevance		
<i>Lingula anatina</i>	XP_013380213.1	580;582;585;587-588;681-688;877;1020-1024;1073-1104;1124-1126;1151-1153;1164;1222;1229;1232-1236;1239-1257;1287-1288;1298;1301-1309	01-24;581;583-584;586;589-590;1025-1033;1066-1072;1128-1133;1142-1149;1163;1165-1166;1223-1228;1230-1231;1237-1238;1289-1297;1299-1300;1310-1315

<i>Phoronis australis</i>	No relevance		
<i>Saccoglossus kowalevskii</i>	XP_006819647.1	01-02;548-561;611-616;843;992-1000;1008	562;617-626;1001-1007;1009-1020;1039-1062;1083-1085
<i>Acanthaster planci</i>	XP_022110423.1	186-188;192-197;360-361;586-594;645-653;657;659-690;695;749;751;832-839;876-878;1029-1030;1078-1183;1197;1204-1281	01-31;190-191;595-596;654-655;840-841;879-889;890;1026-1028;1031-1042;1186-1188;1198-1201;1282
<i>Branchiostoma floridae</i>	XP_002590121.1	04-29;350-351;653-656	01;339-349;386
<i>Ciona intestinalis</i>	XP_002122747.1	Sequence too long	
<i>Petromyzon marinus</i>	No relevance		
<i>Rhincodon typus</i>	No relevance		
<i>Danio rerio</i>	XP_017206503.1	240-244;581-584;823-828;870-871;874;1017-1018;1076-1129;1185-1230	01-22;245-247;829-832;872-873;875-880;1019-1035;1070-1075;1132-1135;1140
<i>Xenopus tropicalis</i>	XP_017944912.1	73;75-76;236;239-242;373-375;382;475-480;483-485;575-576;580-581;633-637;811-818;823;1000-1001;1003-1005;1018-1019;1021-1023;1029;1075;1079-1080;1082-1083;1085	01-20;243-246;376-381;383;638-645;819-821;994-999;1002;1006-1017;1020
<i>Anolis carolinensis</i>	XP_016851514.1	478-482;488;491-493;621-624;627;811-817;820;922-923;1059-1060;1064-1100;1102-1107	01-23;483-487;489-490;617-620;1061;1101;1108-1113;1115-1135
<i>Gallus gallus</i>	Sequence not available		
<i>Ornithorhynchus anatinus</i>	XP_028921034.1	06-08;191-193;196;457-459;614-615;617-618;674-677;679;861-863;867-870;909-912;916-917;1108-1109;1111-1165	01-05;09-35;1055-1075;1110
<i>Rattus norvegicus</i>	NP_072150.1	06-09;39;41-42;502-514;607;795;800-804;1050-1052;1054;1075-1076;1115-1117;1120-1150;1197	01-05;10-38;40;796-799;1053;1055-1071;1103-1114;1118-1119;1152-1156;1159-1173;1206-1209
<i>Mus musculus</i>	NP_062332.2	06-09;13-15;17;501-515;604;606-608;610-611;669-670;795-796;799-804;1052;1070-1071;1112-1157	01-05;10-12;16;18-41;797-798;1050-1051;1053-1069;1104-1111;1159-1178
<i>Pan troglodytes</i>	XP_016791216.2	250;495-500;590-593;650;786-791;843-847;891-894;898-900;1029-1030;1034-1039;1042-1044;1052-1059;1099-1136	01-25;781-785;848-851;1040-1041;1045-1051;1091-1098;1139-1143;1146-1164
<i>Homo sapiens</i>	NP_004637.1	257-258;493-502;589-593;596-597;649-655;787-790;891-893;898-899;1028-1031;1033-1044;1055-1057;1101-1128;1131-1137	01-27;656-657;780-786;1032;1045-1054;1090-1100;1129-1130;1141-1143;1146-1165

Intrinsically unstructured regions (IURs) predicted in CD2AP and its orthologs			
<i>Priapulus caudatus</i>	No relevance		
<i>Caenorhabditis elegans</i>	No relevance		
<i>Hypsibius dujardini</i>	No relevance		
<i>Euperipatoides rowelli</i>	No relevance		
<i>Drosophila melanogaster</i>	XP_016866130.1	62-68;85-97;127-132;160-198;219-267;333-476	69-84;133-154
<i>Brachionus plicatilis</i>	No relevance		
<i>Macrostomum lignano</i>	No relevance		
<i>Octopus bimaculoides</i>	No relevance		
<i>Capitella teleta</i>	No relevance		
<i>Notospermus geniculatus</i>	No relevance		
<i>Lingula anatina</i>	No relevance		
<i>Phoronis australis</i>	No relevance		
<i>Saccoglossus kowalevskii</i>	No relevance		
<i>Acanthaster planci</i>	No relevance		
<i>Branchiostoma floridae</i>	No relevance		
<i>Ciona intestinalis</i>	No relevance		
<i>Petromyzon marinus</i>	No relevance		
<i>Rhincodon typus</i>	XP_020370685.1	13-14;51-52;58;157-205;227-234;239-257;322-587;589-590;636	01-12;15-29;50;54-57;59-62;68-95;235-238;642-643
<i>Danio rerio</i>	NP_001008583.2	62-124;139-142;159-169;172-282;346-599;649-650	143-158;655-657
<i>Xenopus tropicalis</i>	NP_001121435.1	61-105;155-204;206-208;211-212;218-273;337-675;705;712;726-727	141-153;708-711;713-724;733-734
<i>Anolis carolinensis</i>	XP_008114796.1	62-105;145-199;224-278;341-477;483-484;486;488-581;604-607;611;615;632-633	133-143;639-640
<i>Gallus gallus</i>	NP_001305332.1	62-106;128-203;222-277;339-584;586-587;632-633	639-640
<i>Ornithorhynchus anatinus</i>	XP_028928027.1	62-103;133-135;137-139;146-147;151-203;224-272;334-624;677	140-142;148;683-684
<i>Rattus norvegicus</i>	AAM47029.1	62-106;137-138;166-201;203;224-268;331-578;629;630	140-149;636-637
<i>Mus musculus</i>	AAI38375.1	62-104;144;161-210;213;215-267;331-578;629-630	146-151;636-637

<i>Pan troglodytes</i>	XP_009449690.1	62-104;126-190;210-211;231-280;342-483;493-494;499-591;642-643	649-650
<i>Homo sapiens</i>	NP_036252.1	62-102;147;149;170-198;223-268;331-478;481-580;631-632	127-142;148;638-640
Intrinsically unstructured regions (IURs) predicted in podocin and its orthologs			
<i>Priapulus caudatus</i>		No relevance	
<i>Caenorhabditis elegans</i>		No relevance	
<i>Hypsibius dujardini</i>		No relevance	
<i>Euperipatoides rowelli</i>		No relevance	
<i>Drosophila melanogaster</i>		No relevance	
<i>Brachionus plicatilis</i>		No relevance	
<i>Macrostomum lignano</i>		No relevance	
<i>Octopus bimaculoides</i>		No relevance	
<i>Capitella teleta</i>		No relevance	
<i>Notospermus geniculatus</i>		No relevance	
<i>Lingula anatina</i>		No relevance	
<i>Phoronis australis</i>		No relevance	
<i>Saccoglossus kowalevskii</i>		No relevance	
<i>Acanthaster planci</i>		No relevance	
<i>Branchiostoma floridae</i>		No relevance	
<i>Ciona intestinalis</i>		No relevance	
<i>Petromyzon marinus</i>		No relevance	
<i>Rhincodon typus</i>	XP_020382509.1	02;05-78;105	01;03-04;100-104;372-398
<i>Danio rerio</i>	NP_001018155.2	3-18;21;24-69;101;104-105;123	1-2;19-20;22-23;93-98;363-391
<i>Xenopus tropicalis</i>	XP_017949096.1	09-17;31-34;36-54;348	01-08;18-30;35;75-87;103-110;346-347;349-376
<i>Anolis carolinensis</i>	XP_003225506.1	06-11;13;17-26;30-39;46-64;117	01-05;12;14-16;27-29;40-45;83-98;355-384
<i>Gallus gallus</i>	XP_422265.3	08-11;13;18-19;21-25;34-37;39-65;83;95;353-358	01-07;12;14-17;20;26-33;38;88-94;96;359-384
<i>Ornithorhynchus anatinus</i>	XP_001515734.2	01-79;81-82;84;383-384	85-86;98-112;371-382;385-402
<i>Rattus norvegicus</i>	NP_570841.2	18;29-32;34-57;92-94	01-17;19-28;33;76-90;353-

			383
<i>Mus musculus</i>	NP_569723.1	07-08;16-23;28-62;97	01-06;09-15;24-27;354-385
<i>Pan troglodytes</i>	XP_016788875.1	06-13;68;80-83;18-24;26-61;94	01-05;14-17;66-67;86-93;95-96;352-383;25
<i>Homo sapiens</i>	NP_055440.1	06-13;16;18-24;26-61;67-68;94;352;364-366	01-05;14-15;17;25;66;80-83;86-93;95-96;353-363;367-383
Intrinsically unstructured regions (IURs) predicted in TRPC6 and its orthologs			
<i>Priapulus caudatus</i>	No relevance		
<i>Caenorhabditis elegans</i>	No relevance		
<i>Hypsibius dujardini</i>	No relevance		
<i>Euperipatoides rowelli</i>	No relevance		
<i>Drosophila melanogaster</i>	No relevance		
<i>Brachionus plicatilis</i>	RNA42410.1	01-02,117,122-129,390-396,618-620,754-777,803,805-807,810,943,963,997-998	116,118-121,752-753,808-809,945-950,956-961,967-992,1003-1015
<i>Macrostomum lignano</i>	No relevance		
<i>Octopus bimaculoides</i>	XP_014783218.1	07-08,12,15-19,23-26,28-29,31-76,78-79,399-424,641-642,766-840,866,870,926,928,932-994,1000-1001	01-06,09-11,13-14,20-22,27,30,897-907,1004-1023
<i>Capitella teleta</i>	No relevance		
<i>Notospermus geniculatus</i>	No relevance		
<i>Lingula anatina</i>	No relevance		
<i>Phoronis australis</i>	No relevance		
<i>Saccoglossus kowalevskii</i>	XP_002730569.1	06-07,18-41,319-328,764-765,799,846	01-05,08-17,43-61,756-759,830-843,869-875
<i>Acanthaster planci</i>	XP_022103606.1	02-03,317-333,440-448,810-841,869,871-873,879-880,906	01,710-716,807-809,863-868,870,874-877,893-905,907-920,950-965
<i>Branchiostoma floridae</i>	XP_002607434.1	125,133,135-138,284-302,746-793,820-821,824-826,828,887-888,890,898-906,909,914-916	01-21,128-132,134,827,851-874,889,891-897,907-908,910-913,917-923
<i>Ciona intestinalis</i>	No relevance		
<i>Petromyzon marinus</i>	No relevance		
<i>Rhincodon typus</i>	XP_020383938.1	410-441,443-445,477	409,442,509-520
<i>Danio rerio</i>	XP_005161247.1	06-49,178-186,455,778-788,790-791,820-821,823-825,867	01-05,763-777,826-828,872-874

<i>Xenopus tropicalis</i>	XP_002935616.2	01-68,73-75,181-194,337-344,777-807,811-816,843-844,846-848,851	71-72,808-810,849-850,881-898
<i>Anolis carolinensis</i>	XP_008106273.1	07-08,186,336,773-793,832,834-840,845-846,848	01-06,09-15,770-772,794-806,841-844,872-890
<i>Gallus gallus</i>	XP_417184.4	01-67,188-194,339,341-351,780-813,841-842,844-847,849	70-85,773-779,848,874-895
<i>Ornithorhynchus anatinus</i>	XP_028903565.1	02-74,174-187,330,333-337,341,457,772-773,777-788,827,829,854,859,867	01,764-771,774-776,801-826,828,856-858,860-865,893-912
<i>Rattus norvegicus</i>	NP_446011.1	350,789-792,795-800,802-807,812,847,874-879,881-883	01-12,782-788,793-794,801,813-816,818-846,910-930
<i>Mus musculus</i>	XP_006509912.1	06-13,15-26,28-29,31-43,369,808-810,827,830-836,838-839,848-866,905	01-05,14,27,30,804-807,811-826,837,840-847,893-902,930-949
<i>Pan troglodytes</i>	XP_016777341.2	351-353,804-807,814-816,818,827-837,845-846,848	01-17,783-803,819-826,838-844,847,875-884,910-931
<i>Homo sapiens</i>	NP_004612.2	17,205,351-352,804-808,812-818,820-848,876-880,882,884	01-16,18-19,783-803,819,875,883,910-931

Chapter 3: IURs mediate the assembly of SD proteins

3.1. Prelude

Nephrin, CD2AP, podocin, and TRPC6 are suggested as key proteins in aiding structural stability of the SD since, mutations in any of these proteins led to severe proteinuria [50, 52, 53, 58, 92]. Immunoprecipitation studies suggest that the C-terminus of podocin interacts with CD2AP and nephrin, whereas, CD2AP interacts with the cytoplasmic domain of nephrin [52, 58]. It is suggested that interactions between podocin and nephrin are crucial for nephrin signaling required for the proper maintenance of the foot-process actin cytoskeleton [50]. Similarly, co-immunoprecipitation experiments revealed that TRPC6 interacts with nephrin and podocin but not with CD2AP [53]. Although it was suggested that these four proteins colocalize as a complex, contributions of these proteins to the integrity and stability of SD are largely unknown. It is yet to be elucidated how mutations in these proteins elicit structural alterations to the SD of podocytes, which eventually manifests in moderate to severe proteinuria.

In this part of our study, we built models for nephrin, CD2AP, podocin, and TRPC6 and refined them using energy minimizations. Further, we predicted the macromolecular assembly of these four proteins by protein-protein docking. We also performed molecular dynamics simulations of the complex. Our studies revealed that these four proteins possess IURs and IUR-binding domains (IDR-BD), which may actively involve in both homodimerization (nephrin-nephrin) and macromolecular complex. Further, our study provides insights on how mutations in SD proteins associate with morphological changes that can lead to proteinuria

3.2. Methodology

3.2.1. Identifying homologs and sequence analysis of the SD proteins

The *Homo sapiens* nephrin (accession ID: O60500), CD2AP (accession ID: Q9Y5K6), TRPC6 (accession ID: Q9Y210), and podocin (accession ID: Q9NP85) sequences were obtained from the UniProt database (www.uniprot.org). NCBI Protein BLAST search was done for the above protein sequences against the PDB database to identify homologs or motifs to guide model building. PSI-PRED server was used to predict the secondary structure makeup of the proteins [93-95]. In addition to the PSI-PRED IUR predictions in nephrin, CD2AP, podocin, and TRPC6 sequences, Critical Assessment of Protein Structure Prediction 9 (CASP9) validated server namely the Genesilico-meta disorder server (Meta Disorder MD2) was also used. This server predicts IURs based on the optimization of the predictions made by 15 different primary IUR prediction servers.

3.2.2. Model building, refinement, and stereochemical quality analysis

There were no structural templates for nephrin to perform homology modeling. We, therefore, used RaptorX (<http://raptorx.uchicago.edu/StructurePrediction/predict>), a threading-based modeling server to predict a model for nephrin. However, this server was unable to predict a model for the full-length sequence, owing to which we have considered part sequences with overlapping residues i.e., 1-1241, and 1140-1241 residues to predict the model. The templates from RaptorX were stitched using modeller 9.17 to predict the full-length model. *Ab-initio* predictions were done for the proteins that could not be built by either homology modeling or by threading modeling. Although, the crystal structures of the SH3 domains of the CD2AP are solved [61, 96, 97], the rest of the sequence showed less homology to build a complete model. Therefore, a hybrid method encompassing homology modeling and *ab-initio* predictions was used to predict the full-length CD2AP model. I-

TASSER, an *ab initio* server was used for predicting structure for the non-homologous regions of CD2AP. The models obtained from the I-TASSER server and the x-ray crystallography data of the SH3 domains were used as templates in modeller 9.17 software to build a model for CD2AP. NCBI pBLAST search revealed that the crystal structure of the core domain of stomatin from *Pyrococcus horikoshii* (Chain A, PDB ID-3BK6) showed an identity score of 39% with podocin. However, the sequence coverage of the template with podocin was only 44% (residues 165-331). Therefore, I-TASSER predictions were done for the residues 1-164 and 332-383 of podocin [90]. The models from the I-TASSER and structure of the 3BK6 were used as templates in modeler 9.17 to predict a full-length model for podocin. In the case of TRPC6, there were no suitable homologs and the threading-based modeling was unable to predict a structure. Therefore, *ab-initio* predictions were done to obtain the full-length TRPC6 model. Once suitable models were built for the SD proteins, energy minimizations were carried out using ModRefiner software.

Furthermore, the models were thoroughly subjected to stereochemistry analysis using PROCHECK software. V822M nephrin, K301M CD2AP, R138Q podocin, R229Q podocin, R229Q-A284V double mutant podocin, and R895L TRPC6 mutant models were generated by homology modeling using the wild-type models (WT) as templates and the models thus obtained were minimized using ModRefiner. The WT and mutant models were then analyzed using (Protein Interactions Calculator) PIC server. This server details intra-protein hydrophobic, main chain-main chain, main chain-side chain, side chain-side chain, aromatic-aromatic, ionic, aromatic-sulfur, and cation-pi interactions [98].

3.2.3. Building the macromolecular complexes

Protein-protein docking was carried out using PatchDock, a validated Critical Assessment of Prediction of Interactions (CAPRI) server [99, 100]. Initially, rigid docking

was done for the predicted nephrin and CD2AP models. To this complex, podocin followed by TRPC6 models were blind-docked since the interacting regions of TRPC6 with nephrin, podocin, and CD2AP are not known. In the case of building mutant complexes, all the proteins were blind docked sequentially since, the effect of mutations on the structure of each protein is not known. The results obtained from PatchDock were refined through FireDock and the best-docked complex was then selected for molecular dynamics simulations.

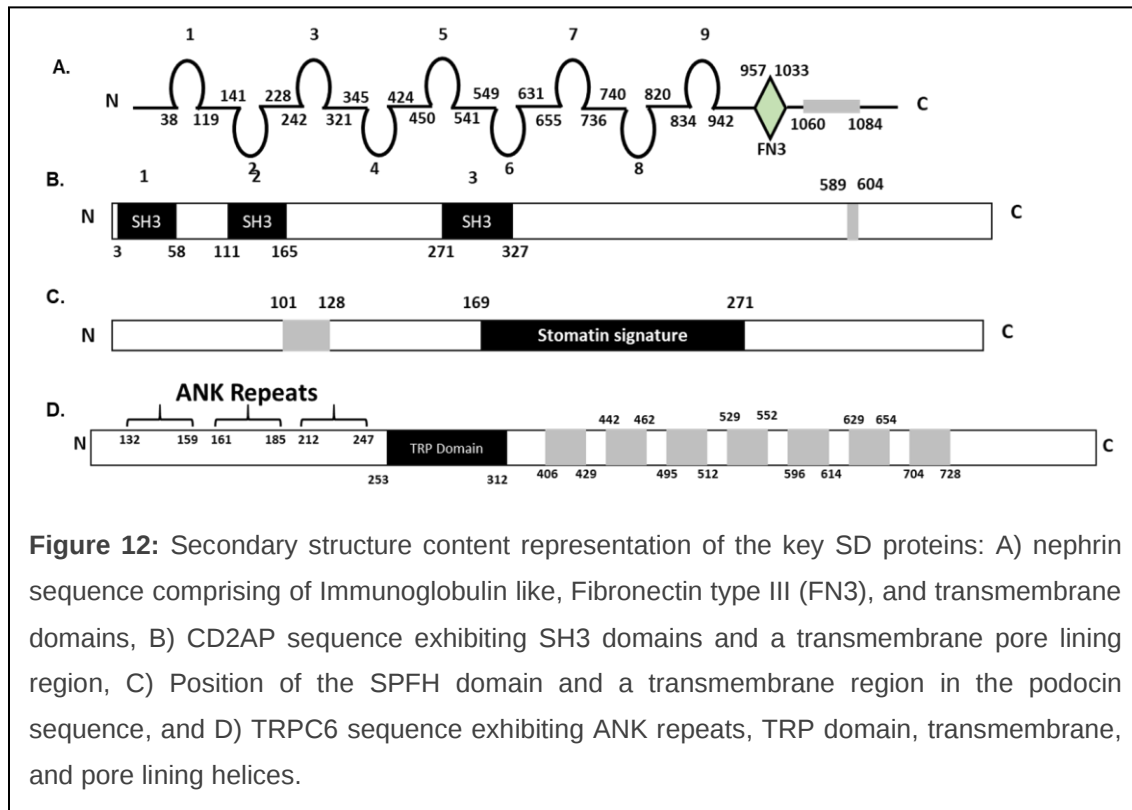
3.2.4. Molecular dynamics simulations and analysis

To ascertain the stability of the macromolecular assembly, we performed molecular dynamics (MD) simulations using a modified approach from the studies by Kirubakaran et al. [101, 102]. Molecular dynamics simulations were performed using a Gromos96-43a force field. The complex was centered in a 35.0 x 30.0 x 64.0Å box, solvated with an SPC216 water system. 41 Na⁺ ions were added to neutralize the charge of the system. Initial energy minimization and equilibration steps were carried out by steepest descent minimization. The system was equilibrated at a constant temperature and pressure of 300K and 1atm. The long-range electrostatics were treated with the particle-mesh Ewald method. PRODIGY webserver was used to extrapolate the intermolecular contacts and types of interactions such as charged-charged, charged-polar, charged-apolar, polar-polar, polar-apolar, apolar-apolar, percentage of a non-interacting surface (NIS) charged residue, and percentage of NIS apolar residues in the complex. PDBsum server was also used to reaffirm the intermolecular contacts besides analyzing the number of salt bridges, hydrogen bonds, and non-bonded contacts present at the protein-protein interface.

3.3. Results

3.3.1. SD proteins consist of conserved structured and unstructured motifs

The native sequences taken from the UniProt database are nephrin comprising of 1241 residues, CD2AP with 639 residues, podocin made of 383 residues, and TRPC6 consist of 931 residues. A pBLAST search of the above sequences was done against the non-redundant database revealed the presence of putative conserved domains in all the sequences. In nephrin, the residues 38-119 belonged to the immunoglobulin (Ig) V-set domain, residues 141-228, 345-424, 450-541, 549-631, 655-736, and 834-942 were identified as Ig-like-C2 domains and residues 242-321 and 740-820 were identified as the third Ig domain. Residues 957-1033 resembles a fibronectin type III domain in the nephrin sequence (Fig.12A). In the case of the CD2AP sequence, three SH3 domains were identified at residues at 3-58, 111-165, and 271-327 (Fig.12B). Similarly, analysis of the



podocin sequence revealed the presence of SPFH (stomatin) family signature motif at

residues 169-271 (Fig.12C). pBLAST of TRPC6 search revealed 3 “Ankyrin” (ANK) repeats at 132-159, 161-185, and 212-247 followed by a TRP domain at 253-312 residues (Fig.12D).

Table 3: IUR predictions for nephrin, CD2AP, TRPC6, and podocin by Genesilico-meta disorder and PSI-PRED servers.

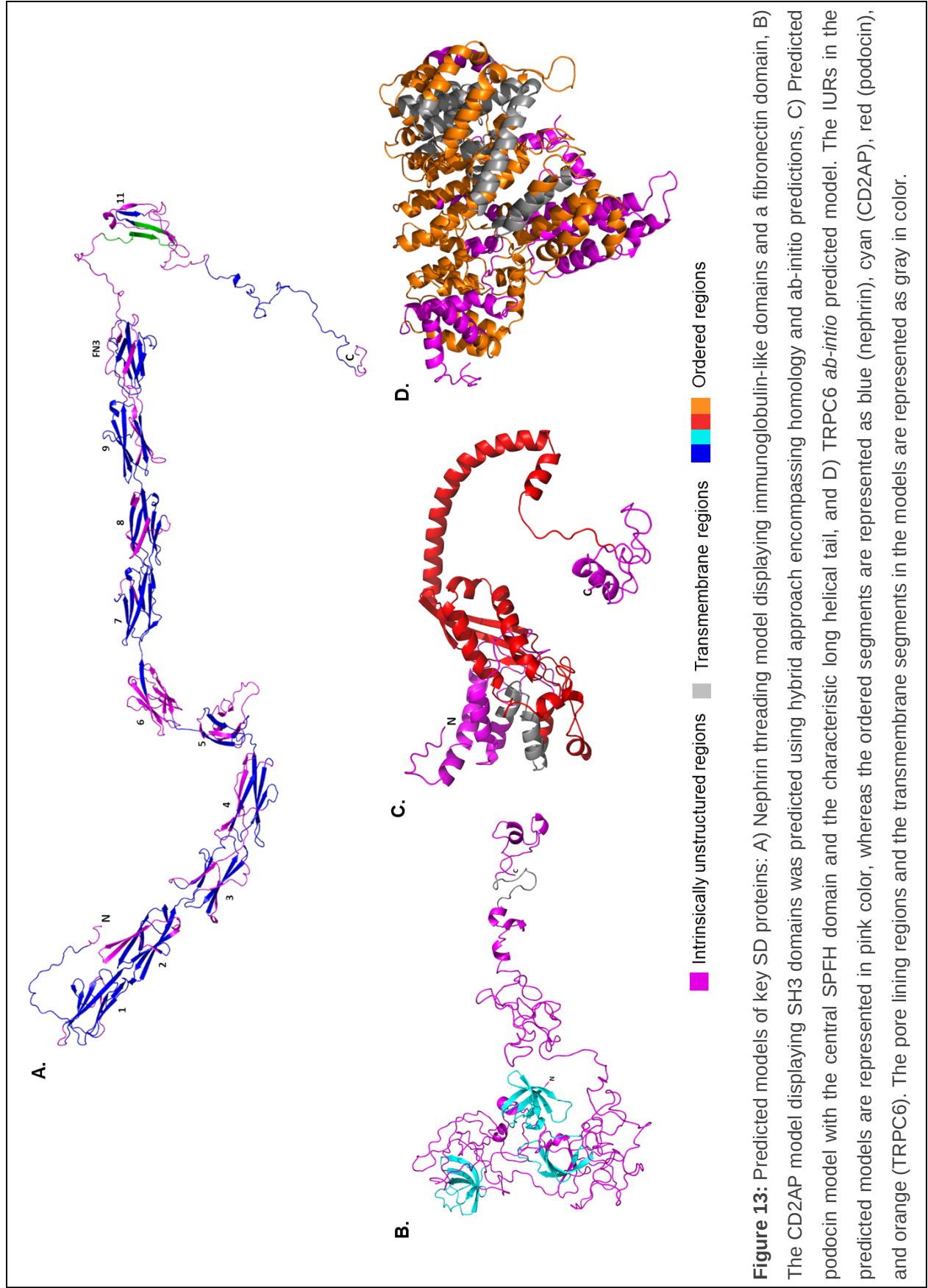
Server	Genesilico- Meta disorder server (IURs)	PSI-PRED	
		IURs	Transmembrane region predictions
Nephrin	1-8, 75-89, 117-121, 179-213, 243-255, 285-295, 323-365, 383-396, 407-409, 445-457, 474-529, 547-636, 689-696, 742-755, 779-798, 840-857, 886-900, 938-952, 992-1006, 1031-1059, 1095-1164, 1228-1241	1, 2, 349, 485, 487 (BD), 488, 489- 491 (BD), 494, 783- 786 (BD), 788, 790 (BD), 791, 1037, 1038-1055 (BD), 1092-1109 (BD), 1111 (BD), 1116-1166, 1178-1179 (BD), 1181-1228	1060-1084
TRPC6	1-40, 62-94, 189-204, 258-276, 347-363, 468-482, 790-931	1-84, 196-205 (BD), 790-848, 876-889 (BD), 894 (BD), 900 (BD), 908-930 (BD), 931	406-429, 442-462, 495-512 (PL), 529-552, 596-614 (PL), 629-654 (PL), 704-728 (PL)
CD2AP	1, 58-110, 165-274, 283-290, 320-590, 594-639	62 (BD), 63-79, 80 (BD), 81-101, 169-201, 224-263, 264-266 (BD), 267-270, 292, 296-299 (BD), 305-581, 603, 606-607 (BD), 610, 612, 614 (BD), 615, 617-622 (BD), 624, 632(BD), 638-639 (BD)	589-604 (PL)
Podocin	1-100, 336-383	1-2, 3-6 (BD), 7-94, 353-356 (BD), 357-361, 362-376 (BD), 377-381, 382-383 (BD)	101-128
Remarks	Meta – Disorder MD2	BD – Binding Domain	Pore lining (PL) regions

Upon initial screening with the PSI-PRED server, IURs were identified in nephrin, CD2AP, podocin, and TRPC6 sequences. Further screening with the Genesilico-meta

disorder server (Metadisorder MD2) confirmed the presence of IURs in all four proteins (Table.3). Apart from predicting IURs, PSI-PRED predicted certain regions in the IURs containing binding hotspots called IUR-binding domains. In nephrin the IUR-BD were predicted at residues 487, 489- 491, 783- 786, 790, 1038-1055, 1092-1109, 1111 and 1178-1179. Similarly, in CD2AP we found IUR-BD at residues 62, 80, 264-266, 296-299, 606-607, 614, 617-622, 632 and 638-639 and in podocin at residues 3-6, 353-356, 362-376 and 382-383. In TRPC6, IUR-BD were identified at residues 196-205, 876-889, 894, 900, and 908-930. The PSI-PRED server also predicted transmembrane regions in all the sequences (Table.3). Our analysis indicated that transmembrane segments were present both in nephrin at residues 1060-1084 and in podocin at residues 101-128. In the case of CD2AP, a transmembrane pore-lining (PL) region was predicted at residues 589-604. Analysis of the TRPC6 sequence revealed three transmembrane helices at residues 406-429, 442-462, 529-552, and four trans-membrane pore-lining regions at residues 495-512, 596-614, 629-654, and 704-728 (Table.3).

3.3.2. Prediction, validation, and insights into the built models

Nephrin and TRPC6 had no acceptable homologs to build models. Therefore, threading and *ab-initio* modeling were used to build models for these proteins. Whereas, *ab-initio* in conjunction with homology modeling were used to build models for CD2AP and podocin. The models were refined to obtain stereo-chemically acceptable structures. The Ramachandran plots for predicted structures revealed that outlier values are in the acceptable range (Table.4). Upon analyzing the predicted models, we observed that the nephrin model is composed of Greek key motifs at residues 41-132, 138-237, 244-333, 343-438, 445-541, 548-640, 647-737, 741-835, and 839-937 (Fig. 13A). We identified a new domain with three antiparallel sheets at residues 1068-1136 (Fig.13A). An earlier study did not report about this unique domain in nephrin [37].



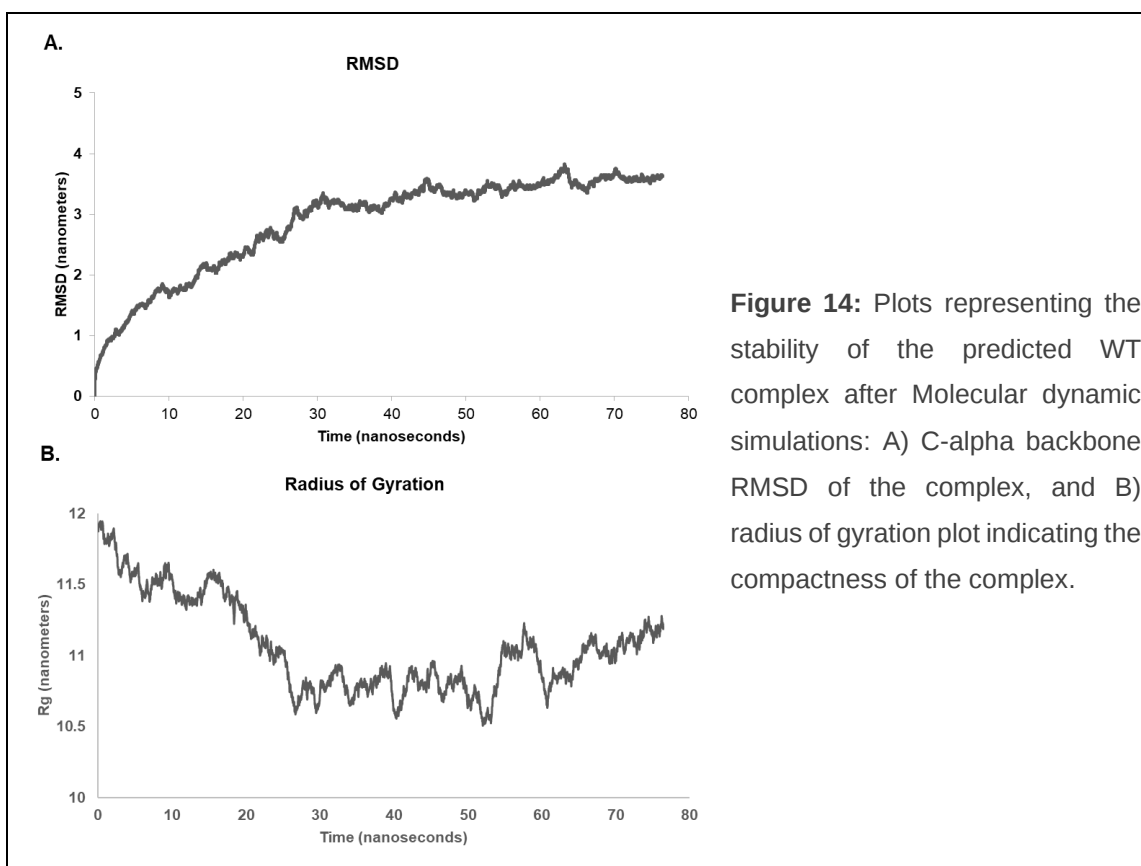
The residues 956-1033 were identified as the fibronectin type III domain in the nephrin model. In the case of the CD2AP model, three SH3 domains are localized to the residues 2-57, 113-158, and 272-328 (Fig.13B). Two noticeable features were observed in the CD2AP model; (a) the presence of long random coils in-between the SH3 domains and (b) several IURs across the protein sequence. The podocin model displayed SPFH family motif at residues 169-271 and long random coils at the N- and C- terminals, which are also identified to be IURs (Fig.13C). Similarly, upon analysis of the TRPC6 model, we observed signature ANK family repeats consisting of antiparallel helices joined by β -turns (Fig.13D). PDBsum analysis of the TRPC6 model revealed the presence of a pore with dimensions 1.49Åx17.2Å (radius x length).

Table 4: Ramachandran outliers for Nephrin, CD2AP, Podocin, and TRPC6 models.

Protein	Core	Allowed	Generously allowed	Dis-allowed
Nephrin	88.1%	9.9%	1.4%	0.6%
CD2AP	72.1%	25.1%	1.3%	1.5%
Podocin	81.8%	13.6%	2.1%	2.4%
TRPC6	81.7%	13.7%	3.2%	1.4%

3.3.3. IURs may mediate complex formation

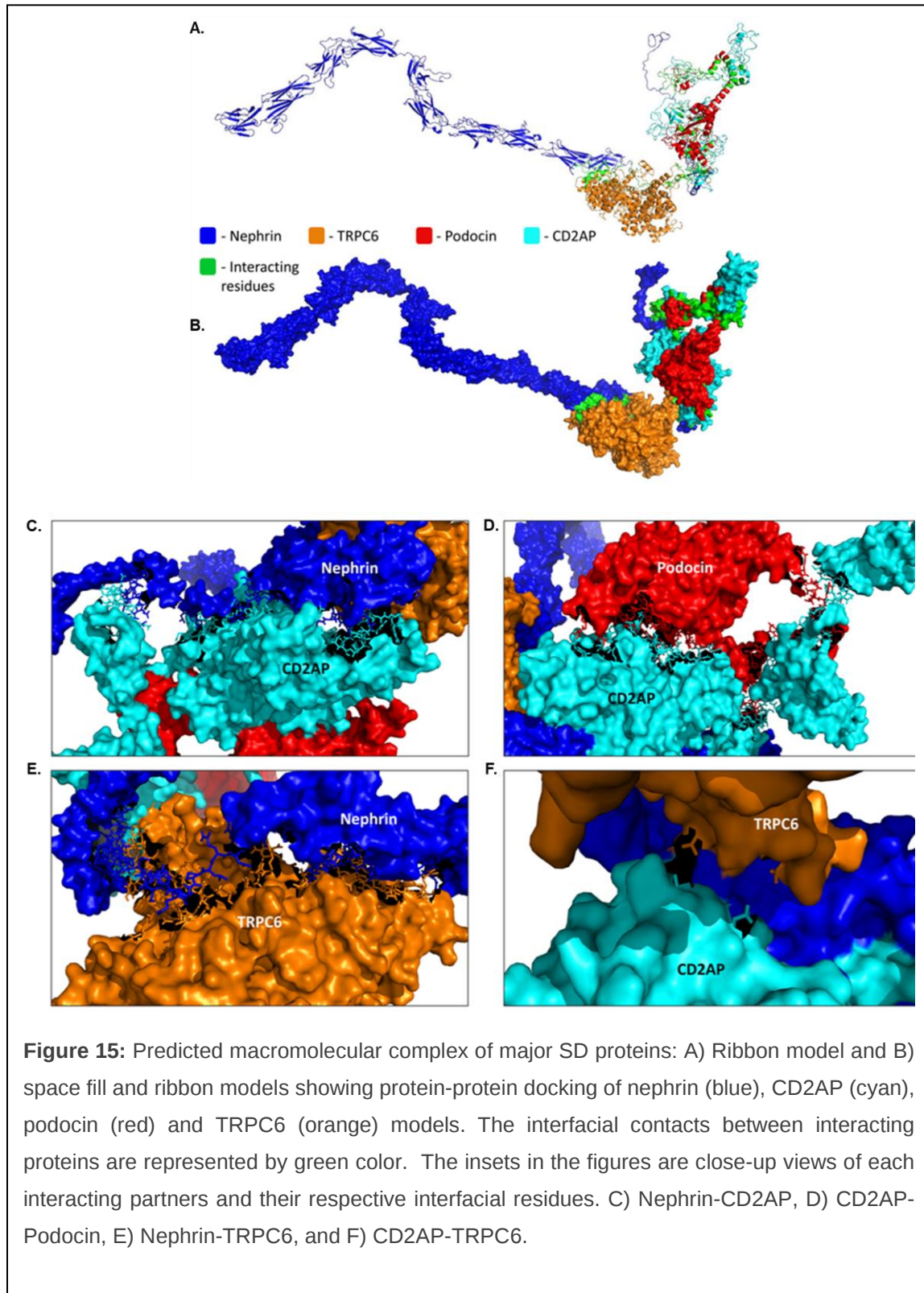
Experimental evidence did not provide adequate information about the interacting regions or the residues between these proteins when they constitute a multimeric complex. Based on the available information on the role of IURs, we speculated that IURs/IUR-BD of these four proteins may play a key role in the formation of a multimeric protein complex [103]. Therefore, to gain insights on the interfacial residues participating in the complex formation, we docked nephrin and CD2AP based on regions specified in



the literature [50, 52, 53, 92] to obtain nephrin-CD2AP complex. To this complex, we docked podocin followed by TRPC6 sequentially. The macromolecular complex thus obtained was subjected to MD simulations to understand the stability of the macromolecular assembly. The root mean square deviation (RMSD) of the complex stabilized at ~30ns and did not change much after that indicating that the complex was stable (Fig.14A). Further stability of the complex was inspected by the radius of gyration. This plot indicates the compactness of the protein and for a stably folded protein, it will not fluctuate much and vice versa for the unstable protein. The radius of gyration plot for our protein complex showed an initial fluctuation until 30ns post which that fluctuation was considerably less (Fig. 14B).

We predicted that the C-terminal residues 1072-1201 and residues 851-1140 of nephrin interact with 116-447 residues of CD2AP (Fig.15C & Table.5) and with 154-903 residues of TRPC6 (Fig.15E & Table.5) with binding energies (ΔG) of -18 kcal/mol and ΔG

of -20.3kcal/mol. Residues 1-592 of CD2AP interacted with both the N- and the C-terminus of podocin (10-381) with a ΔG of -21.8kcal/mol (Fig.15D & Table.5). Interestingly, we found



that residues 114 and 131 of CD2AP formed non-bonded contacts with residues 887 and 883 of TRPC6 with ΔG of -4.2kcal/mol (Fig.15F & Table.5). It is noteworthy that our docking studies could not predict interactions between nephrin-podocin and podocin-TRPC6. Our analysis suggests that most of the interfacial contacts between nephrin, CD2AP, podocin, and TRPC6 are contributed by IURs/IUR-BD (Table.5, highlighted in yellow and bold-italics), which reiterates the importance of IURs in the macromolecular assembly. A complete list of protein-protein interactions including the non-bonded contacts responsible for the complex formation is provided in the ancillary table. A2. In addition to the binding energy calculations, the number of contacts responsible for the complex formation such as salt bridges, hydrogen bonds, and non-bonded contacts has been listed (Table.5).

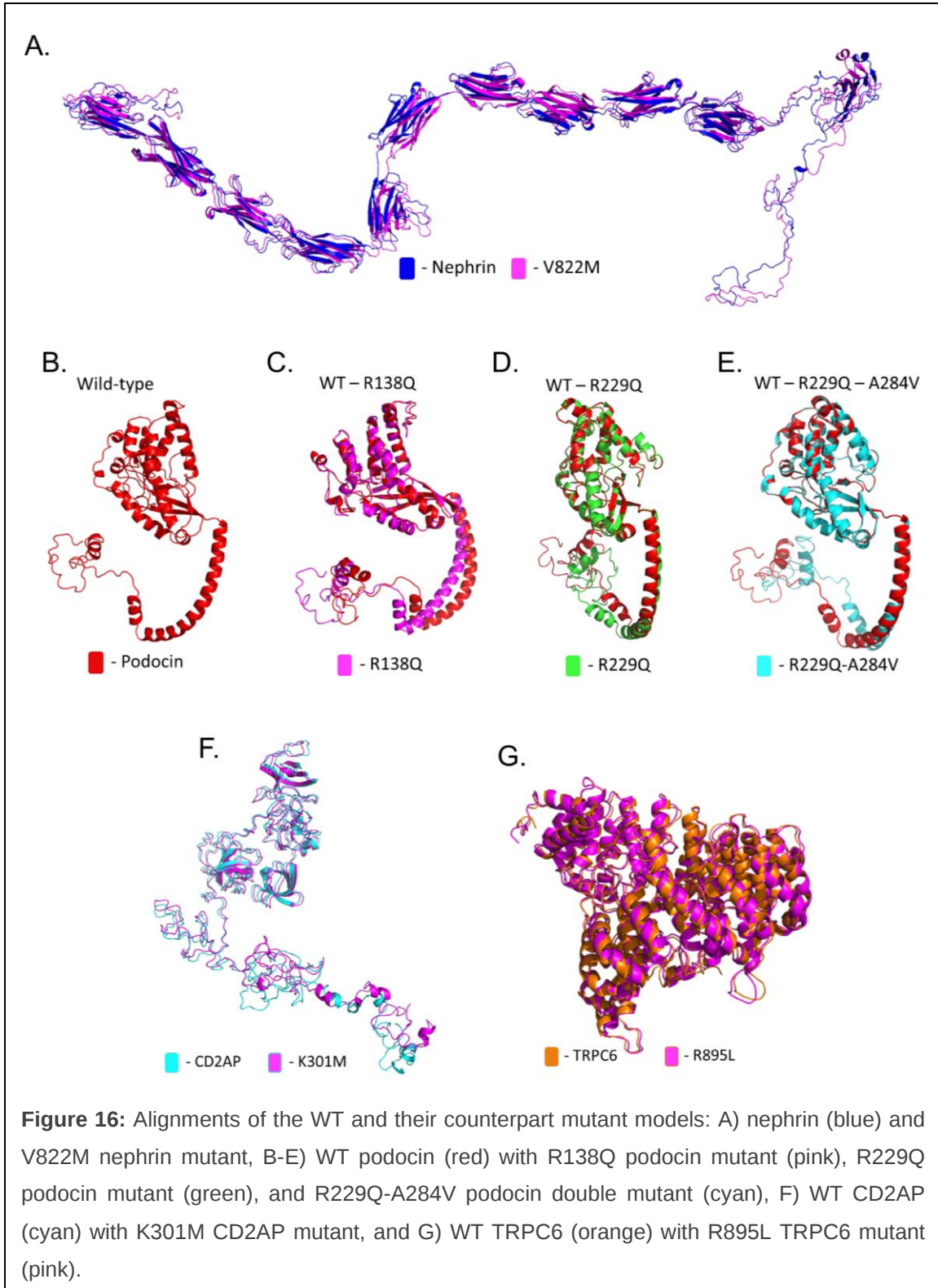
Table 5: Predicted salt bridge and hydrogen bonds at the protein-protein interfaces of nephrin, CD2AP, podocin, and TRPC6 complex. The annotation in parenthesis indicates, C- Charged, P- Polar, and AP- Apolar, and the yellow highlighted residues are predicted to be part of intrinsically disordered regions, whereas the residues in bold and italics are predicted IDR-binding domains.

Type of interaction	(C)	(P)	(AP)
Nephrin-CD2AP interacting residues ΔG: -18 kcal mol-1			
Nephrin	CD2AP		
Salt Bridges:			
HIS 1174 (P)	ASP 286	-	-
ASP 1194 (C)	LYS 436	-	-
Hydrogen bonds:			
ARG 1143 (C)	-	ASN 160	VAL 162
SER 1146 (P)	-	SER 189	GLY 186
GLN 1148 (P)	GLU 258	-	VAL 257, ILE 259
SER 1157 (P)	LYS 261	-	-
TYR 1158 (P)	LYS 261	-	-
SER 1159 (P)	LYS 261	-	-
ARG 1160 (C)	ASP 263	THR 260, THR 262	-
PHE 1162 (AP)	-	THR 262	-
THR 1163 (P)	ASP 263	-	-
HIS 1174 (P)	GLU 285, ASP 286, GLU 317	-	-
ALA 1188 (AP)	LYS 436	-	-
GLY 1190 (AP)	ARG 412	SER 434	-
TYR 1193 (P)	GLU 414	-	-
ASP 1194 (C)	LYS 444	-	-
VAL 1196 (AP)	-	SER 443	VAL 442
GLN 1197 (P)	-	SER 443	-

GLY 1199 (AP)	GLU 440		
Nephrin- TRPC6 interacting residues ΔG : -20.3 kcal mol ⁻¹			
Nephrin	TRPC6		
Salt Bridges:			
ASP 1047 (C)	-	HIS 264	-
ARG 1091 (C)	GLU 886	-	-
Hydrogen bonds			
GLY 853 (P)	ASP 478		-
SER 910 (P)	ASP 478	-	-
ALA 911 (AP)	ARG 437, ASP 478	-	-
ALA 912 (AP)	ARG 437, ASP 478	-	-
TYR 915 (P)	ARG 437	ASN 479	-
ASP 942 (C)	-	SER 430	-
ALA 1025 (AP)	-	-	ALA 421
GLY 1028 (AP)	-	-	PHE 419
SER 1042 (P)	ASP 357, ARG 360	-	-
SER 1055 (P)	ASP 250	-	-
GLY 1056 (AP)	ASP 250	-	-
LEU 1062 (AP)	-	-	LEU 899
TYR 1139 (P)	GLU 883	-	-
CD2AP - Podocin Interacting residues ΔG : -21.8 kcal mol ⁻¹			
CD2AP	Podocin		
Salt Bridges:			
GLU 34 (C)	LYS 377	-	-
ASP 51 (C)	LYS 148	-	-
LYS 579 (C)	GLU 303	-	-
GLU 592 (C)	ARG 306	-	-
Hydrogen bonds:			
MET 1 (AP)	ASP 250	-	-
GLN 33 (P)	ASP 379	-	-
ARG 59 (C)	-	SER 205	-
GLU 60 (C)	-	SER 205	-
LYS 64 (C)	-	SER 86, SER 120	-
SER 399 (P)	-	-	LEU 373
THR 407 (P)	-	-	PRO 369
THR 468 (P)	LYS 368	-	-
LEU 487 (AP)	-	-	LEU 324
THR 490 (P)	-	HIS 325	-
ALA 562 (AP)	ARG 322	-	-
PHE 563 (AP)	ARG 322	-	-
THR 575 (P)	LYS 299	-	-
VAL 578 (AP)	ARG 306	-	-
LYS 579 (C)	GLU 303,	-	-
CD2AP - TRPC6 Interacting residues G : -4.2 kcal mol ⁻¹			

3.3.4. Mutations induced subtle changes in the mutant models

To garner clues on how a mutation in each protein influenced the assembly of the complex, we selected one mutant for each of these four proteins (Nephrin V822M, CD2AP K301M, Podocin R138Q, and TRPC6 R895L). An additional recessive mutant (R229Q)



which did not cause NS independently but caused severe SNRS in conjunction with A284V were also considered in the case of podocin. We considered these mutations

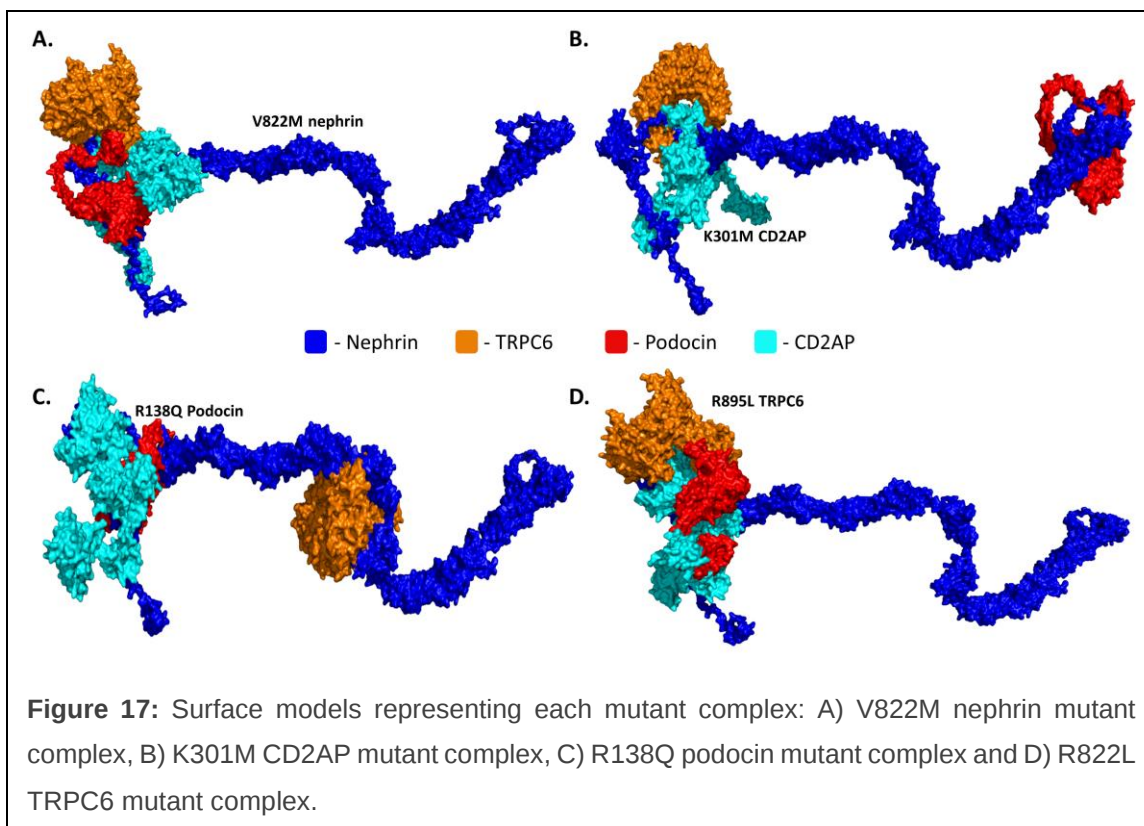
based on their association with the severity of NS [71, 104-106]. Although these mutations did not fall into the interfacial contacts predicted in the WT complex, we would like to elucidate how a mutation influenced the tertiary structure of a protein and compromise protein-protein interactions. We compared the secondary and tertiary structure predictions of each of the mutants with their WT predictions using PSI-PRED and PIC server. Our analysis revealed nephrin, CD2AP, and TRPC6 mutants did not have any changes in their secondary and tertiary structure predictions in the mutated region. However, the secondary structure predictions for R138Q podocin showed the mutation region as a part of the beta-sheet, whereas in the 3D model this region was observed as a loop. This discrepancy could be associated with the podocin's ability to form homo-oligomers like other members of the protein's family. Overall tertiary structure comparison of mutant and WT models revealed subtle alterations to the protein stereochemistry (Fig. 16A-G) (Table.6).

Table 6: RMSD deviations showed by the mutant models when compared to the WT nephrin, CD2AP, podocin, and TRPC6 models respectively.

Mutants	RMSD (Å)
V822M nephrin	3.780
K301M CD2AP	1.797
R138Q Podocin	0.581
R229Q podocin (recessive mutant)	0.231
R229Q-A284V podocin double mutant	0.339
R895L TRPC6	1.030

3.3.5. Mutations in the SD proteins disrupted the complex formation

Next, we docked these mutant proteins with their WT neighbors to check aberrations in the macromolecular assembly. In the case of nephrin mutant (V822M), the interactions between nephrin-CD2AP, nephrin-TRPC6, and CD2AP-podocin reduced



significantly along with decreased binding propensity, while the CD2AP-TRPC6 interactions increased considerably when compared to the WT complex. Interestingly, in the nephrin mutant, we found new interactions between nephrin-podocin and podocin-TRPC6 (Fig.17A & Table.7). In the case of CD2AP (K301M) mutant, the interactions between CD2AP-podocin are compromised, since the podocin is bound to the N-terminus of nephrin in striking contrast to the WT complex. Also, we observed that the interaction propensities between nephrin-CD2AP and between CD2AP-TRPC6 increased considerably. Furthermore, due to the K301M CD2AP mutation, only the N-terminus of TRPC6 docked downstream to the FN3 domain of nephrin. This result is in contrast to the WT complex wherein, TRPC6 docked to the FN3 domain (Fig.17B & Table.7).

In the podocin mutant (R138Q), the number of interactions and the binding propensities of nephrin-CD2AP and CD2AP-podocin decreased along with complete loss of CD2AP-TRPC6 interactions. Further, in this complex, we noticed TRPC6 docked to the

5th Ig domain of nephrin (Fig.17C & Table.7). The R895L TRPC6 mutation severely distorted the TRPC6 structure, due to which its interactions with nephrin are severely compromised and also caused TRPC6 to interact strongly with CD2AP and vice versa with podocin. Also, the binding propensity of CD2AP-podocin in this complex although remained similar, the interfacial contacts between them changed significantly. We also noticed that the binding energy between nephrin-CD2AP decreased considerably in the R895L TRPC6 mutant (Fig.17D & Table.7). Our analysis suggests that mutations in each protein significantly distorted the complex assembly.

Table 7: Binding energies, number of salt bridges, hydrogen bonds, and non-bonded contacts predicted at the protein-protein interfaces of wild type and mutant complexes.

Interactions	Nephrin -CD2AP	Nephrin - podocin	Nephrin - TRPC6	CD2AP - podocin	CD2AP - TRPC6	Podocin - TRPC6
Wild-type complex						
Binding energy ΔG (kcal/mol)	-18.0	-	-20.3	-21.8	-4.2	-
No. of Interface residues	39-48	-	51-41	60-58	2-2	-
Salt Bridges	2	-	2	4	-	-
Hydrogen bonds	30	-	23	18	-	-
Non-bonded contacts	364	-	343	313	2	-
V822M Nephrin complex						
Binding energy ΔG (kcal/mol)	-17.6	-15.7	-11.6	-10.1	-11.5	-5.3
Interface residues	37-38	32-36	15-17	20-21	20-20	8-3
Salt Bridges	2	2	-	2	-	-
Hydrogen bonds	9	7	4	1	6	1
Non-bonded contacts	595	515	210	237	245	69
K301M CD2AP complex						
Binding energy ΔG (kcal/mol)	-26.0	-21.0	-11.9	-	-15.4	-
Interface residues	54-60	41-40	22-26	-	49-41	-
Salt Bridges	2	-	1	-	3	-
Hydrogen bonds	10	4	7	-	8	-
Non-bonded contacts	1219	570	348	-	657	-
R138Q Podocin complex						
Binding energy ΔG (kcal/mol)	-15.4	-20.1	-20.1	-10.4	-	-

Interface residues	34-32	33-43	44-52	17-15	-	-
Salt Bridges	4	2	7	3	-	-
Hydrogen bonds	4	5	14	5	-	-
Non-bonded contacts	413	496	676	227	-	-
	R895L TRPC6 complex					
Binding energy ΔG (kcal/mol)	-15.4	-	-9.6	-21.1	-20.5	-4.7
Interface residues	34-32	-	15-15	52-62	42-48	6-5
Salt Bridges	4	-	-	2	2	-
Hydrogen bonds	4	-	2	8	10	1
Non-bonded contacts	413	-	293	921	699	31

3.4. Discussion

The details of the macromolecular assembly of SD proteins and how mutations in each protein distort interactions and compromise the permselectivity of the SD is of interest to the nephrologists. The present study is aimed to elucidate the coordinated interactions of the complex through computationally derived models for nephrin, CD2AP, podocin, and TRPC6. Using sequence analysis, docking, and molecular dynamics simulations we shed light on the interfacial residues responsible for the complex formation and discuss how the dynamic states of IURs and IUR-BD in each protein play a crucial role in protein-protein interactions. We also discuss, how mutations in each protein can significantly affect each protein in the complex thereby causing alterations in SD architecture leading to proteinuria.

It is known that IURs typically involve in the formation of multiprotein complexes and signaling pathways [103]. The flexibility offered by the IURs allows the sequence to adopt diverse conformations to bind to its receptor or ligand [107]. Both the IUR prediction servers identified IURs in all these proteins. Our results suggest that IURs/IUR-BD predicted in the Ig domains of adjacent nephrin molecules may aid in interactions facilitating the formation of dimers thus, imparting the characteristic zipper-like structure to the SD. Besides being involved in the formation of homodimers, our results show that the IURs and IUR-BD of nephrin also involve in the formation of the heteromeric complex (nephrin-CD2AP-podocin-TRPC6).

To delineate the role of IURs in the formation of heteromeric complex, we performed docking and simulations studies with nephrin, CD2AP, podocin, and TRPC6 models. From our predictions, it is evident that most of the interacting residues between nephrin, CD2AP, podocin, and TRPC6 are either IURs/or IUR-BD. An earlier study showed that the 3rd SH3 domain of CD2AP interacts with nephrin and it was speculated

that the C-terminal Tyr¹¹⁷⁶, Tyr¹¹⁹³, and Tyr¹²¹⁰ of nephrin might act as possible docking sites for the SH3 domains [50, 92]. Our results suggest that non-IUR Tyr¹¹⁷⁶ and Tyr¹¹⁹³ residues of nephrin interacted with non-IUR residues 313-316 of the 3rd SH3 domain and IUR residues 412,414,430,437,441-447 of CD2AP. Therefore, our observations are consistent with the experimental observations. Further, our studies also show that the 2nd SH3 domain of CD2AP is involved in nephrin-CD2AP interactions. It was reported earlier that CD2AP and the C-terminal 81 residues of nephrin bind to C-terminus of podocin [50]. Although we were able to predict IURs/IUR-BD in nephrin and podocin, our docking method could not predict nephrin-podocin interactions in the complex. We speculate that nephrin may independently bind to another podocin and not necessarily interact with the podocin bound to CD2AP; this condition may be possible as podocin is known to form homo-oligomers at the SD [105]. On the other hand, it is known that SH3 domains possess an affinity for proline-rich regions (PXXP) [108]. Sequence analysis revealed that residues 362-376 of podocin are proline-rich and our results indicate that IUR residues 398-408, 468, and 482 of CD2AP interact with 362, 364, 369, and 372 IUR-BD prolines of podocin. These results confirm the affinity of SH3 towards proline residues and also suggest the important role of IURs in protein-protein interactions. Further, our results also show that all the three SH3 domains of CD2AP interacted with podocin.

The earliest evidence of a macromolecular assembly came from co-immunoprecipitation experiments which showed that TRPC6 interacts with nephrin and podocin [53]. Our results suggest that TRPC6 interacted with the C-terminal residues of nephrin. Interestingly, the docking studies also predicted CD2AP-TRPC6 interactions in the complex with significantly less binding energy when compared to its interacting neighbors. We speculate that these weak TRPC6-CD2AP interactions may elicit functions that are yet to be comprehended. Nevertheless, our docking method was unable to predict interactions between TRPC6 and podocin, a condition that may be similar to the missing

nephrin-podocin interactions as stated above. All these results suggest the involvement of IURs/IUR-BD in facilitating interactions between nephrin, CD2AP, podocin, and TRPC6.

A mutation in each protein was selected from the human gene mutation database [109]. V822M nephrin mutant formed small distinct puncta along the plasma membrane as opposed to the large distinct puncta formed by the WT protein [106]. The V822M mutation in nephrin does not fall into the predicted interfacial contacts but falls into the 8th Ig-like-C2 domain, which is the extracellular segment of the protein. According to our results, the V822M mutation-induced morphological changes in nephrin and affecting its ability to appropriately interact with CD2AP, podocin, and TRPC6. From our results, we suggest that this mutation may also disrupt the nephrin-nephrin homodimer formation. These alterations may perhaps explain the small distinct nephrin puncta due to V822M mutation.

The K301M mutation in CD2AP was shown to disrupt interactions between CD2AP and its neighbors leading to the early onset of SRNS [104]. K301M is a non-conservative mutation that falls into the non-interacting residues of the third SH3 domain. We show that mutation of lysine to methionine in CD2AP drastically alters the CD2AP structure and its ability to form a complex with the tyrosine residues of nephrin and also with proline-rich regions of podocin. All these changes may ultimately lead to compromised complex formation thus disturbing the SD architecture.

The R138Q podocin mutant localized at the endoplasmic reticulum and interferes with the recruitment of nephrin to the lipid rafts [105]. In our complex model, the R138 residue of podocin was observed to be buried inside the protein. Although, both the arginine and glutamine have similar charge groups, their side chains differ considerably. From our analysis, it is evident that the R138Q mutation altered the essential interactions required for maintaining the tertiary structure of podocin, which may affect its scaffolding

ability and its potential to recruit nephrin and CD2AP to the SD leading to alterations in SD architecture.

The R895 residue in TRPC6 is evolutionarily conserved and the mutation R895L caused NS at an early age of 1 year [71]. Further, this mutation caused enhanced calcium influx into the podocyte disrupting the podocyte cytoskeleton [71]. In our WT complex, the positively charged R895 residue in TRPC6 interacted with the leu¹⁰⁶², pro¹⁰⁶⁴, and leu¹⁰⁶⁶ of the transmembrane region of nephrin forming charged–charged, charged-polar and charged-apolar contacts. Mutation of R895 to non-polar leucine in TRPC6, severely changed its structure thus losing the above interactions and also causing significant loss of interfacial contacts between nephrin and TRPC6. This loss of contacts may induce conformational changes in cytoskeleton-bound nephrin, which consequently affects the SD architecture leading to proteinuria. Also, from our results we speculate the alteration of the TRPC6 structure may lead to protein aggregation since, TRPC6 showed interactions with podocin and CD2AP in the TRPC6 mutant complex.

3.5. Ancillary data

Table A2: Complete list of predicted interactions at the protein-protein interfaces of nephrin, CD2AP, podocin, and TRPC6 complex. The annotation in parenthesis indicates, C- Charged, P- Polar, and AP- Apolar, and the yellow highlighted residues are predicted to be part of intrinsically disordered regions, whereas the residues in bold and italics are predicted IDR-binding domains.

Nephrin-CD2AP interacting residues ΔG : -18 kcal mol ⁻¹			
Type of interaction	(C)	(P)	(AP)
Nephrin	CD2AP		
Salt Bridges:			
HIS 1174 (P)	ASP 286	-	-
ASP 1194 (C)	LYS 436	-	-
Hydrogen bonds:			
ARG 1143 (C)	-	ASN 160	VAL 162
SER 1146 (P)	-	SER 189	GLY 186
GLN 1148 (P)	GLU 258	-	VAL 257, ILE 259
SER 1157 (P)	LYS 261	-	-
TYR 1158 (P)	LYS 261	-	-
SER 1159 (P)	LYS 261	-	-
ARG 1160 (C)	ASP 263	THR 260, THR 262	-
PHE 1162 (AP)	-	THR 262	-
THR 1163 (P)	ASP 263	-	-
HIS 1174 (P)	GLU 285, ASP 286, GLU 317	-	-
ALA 1188 (AP)	LYS 436	-	-
GLY 1190 (AP)	ARG 412	SER 434	-
TYR 1193 (P)	GLU 414	-	-
ASP 1194 (C)	LYS 444	-	-
VAL 1196 (AP)	-	SER 443	VAL 442
GLN 1197 (P)	-	SER 443	-
GLY 1199 (AP)	GLU 440	-	-
Non-Bonded Contacts:			
LEU 1072 (AP)	-	-	LEU 116, PHE 117
ILE 1097 (AP)	-	-	PHE 117
GLU 1099 (C)	-	ASN 160	PHE 161
GLU 1102 (C)	LYS 221	-	-
PRO 1137 (AP)	GLU 118	-	-
TYR 1139 (P)	GLU 118	TYR 119	LEU 116, PHE 117, VAL 131
ARG1140 (C)	-	-	LEU 116, PHE 117
SER 1141 (P)	-	-	PHE 117, PHE 161
LEU 1142 (AP)	-	-	LEU 116, PHE 117
			LEU 116, PHE 117, PHE 161, VAL 162, LEU 184, ALA 185, GLY 186, PRO 187
ARG 1143 (C)	LYS 163, GLU 164	SER 159, ASN 160	ALA 185, GLY 186, PRO 187
ASP 1144 (C)	-	-	ALA 185, GLY 186, PRO 187
PHE 1145 (AP)	LYS 221	ASN 160, SER 189, SER 193,	ALA 185, GLY 186, PRO 187, PRO 192
SER 1146 (P)	LYS 221	THR 188, SER 189, SER 193,	GLY 186, PRO 187, PRO 192
PRO 1147 (AP)	GLU 258	THR 188, SER 189, SER 193	GLY 186, PRO 187, PRO 190, PRO 192, LEU 194, VAL 257, ILE 259

GLN 1148 (P)	GLU 258	THR 188, SER 189, SER 256, THR 260	VAL 257, ILE 259
LEU 1149 (AP)	-	SER 193	LEU 194, VAL 257, ILE 259
PRO 1150 (AP)	-	THR 260	ILE 259,
SER 1157 (P)	LYS 261	THR 264	GLY 266, ILE 268
TYR 1158 (P)	LYS 261, ASP 263	THR 264	ILE 268
SER 1159 (P)	LYS 261, ASP 263	THR 264	
ARG 1160 (C)	LYS 261, ASP 263	THR 260, THR 262	ILE 259
GLY 1161 (AP)	THR 262, ASP 263	-	
PHE 1162 (AP)	ASP 263	THR 200, THR 262	ALA 201
THR 1163 (P)	ASP 263, LYS 269	-	-
PHE 1171 (AP)	GLU 317	-	-
GLY 1173 (AP)	LYS 316	ASN 314,	GLY 315,
	GLU 285, ASP 286, GLU 287, ARG 310, LYS 316, GLU 317	-	GLY 315, GLY 318
HIS 1174 (P)			
LEU 1175 (AP)	GLU 285, LYS 316	-	GLY 315,
TYR 1176 (P)	LYS 316	ASN 314	LEU 313, GLY 315
PRO 1184 (AP)	LYS 436	-	-
PRO 1185 (AP)	LYS 436, GLU 438, GLU 440,	THR 439	PHE 437, VAL 442
SER 1186 (P)	LYS 436, GLU 438	-	PHE 437
GLY 1187 (AP)	LYS 436	-	PHE 437
ALA 1188 (AP)	LYS 436	SER 434, SER 435	-
TRP 1189 (P)	ARG 412, LYS 436	SER 434, SER 435	-
GLY 1190 (AP)	ARG 412, LYS 436	SER 434, SER 435	-
PRO 1191 (AP)	ARG 412, GLU 414,	SER 434	PRO 413
LEU 1192 (AP)	ARG 412, GLU 414, LYS 436, LYS 444	-	-
			VAL 430, PHE 437, PRO 441, VAL 442, LEU 445, LEU 447
TYR 1193 (P)	ARG 412, GLU 414, LYS 444, LYS 446	SER 443	VAL 442, LEU 445
ASP 1194 (C)	LYS 436, LYS 444	SER 443	VAL 417, VAL 442
GLU 1195 (C)	LYS 444	SER 443	PRO 441, VAL 442
VAL 1196 (AP)	GLU 440, LYS 444	SER 443	ILE 426, GLY 428, PRO 441, VAL 442
GLN 1197 (P)	GLU 440	ASN 427, SER 443	VAL 442
MET 1198 (AP)	GLU 440, PRO 441	ASN 427	
GLY 1199 (AP),			
PRO 1200 (AP),	GLU 440	-	-
TRP 1201(P)			
Nephrin- TRPC6 interacting residues ΔG: -20.3 kcal mol⁻¹			
Nephrin	TRPC6	Salt Bridges:	
ASP 1047 (C)	-	HIS 264	-
ARG 1091 (C)	GLU 886	-	-
Hydrogen bonds			
GLY 853 (P)	ASP 478		-
SER 910 (P)	ASP 478	-	-
ALA 911 (AP)	ARG 437, ASP 478	-	-
ALA 912 (AP)	ARG 437, ASP 478	-	-
TYR 915 (P)	ARG 437	ASN 479	-
ASP 942 (C)	-	SER 430	-
ALA 1025 (AP)	-	-	ALA 421
GLY 1028 (AP)	-	-	PHE 419
SER 1042 (P)	ASP 357, ARG 360	-	-
SER 1055 (P)	ASP 250	-	-
GLY 1056 (AP)	ASP 250	-	-

LEU 1062 (AP)	-	-	LEU 899
TYR 1139 (P)	GLU 883	-	-
Non-Bonded Contacts			
ALA 851 (P)	ASP 478	-	-
ALA 852 (P)	GLU 474, ASP 478	THR 475, ASN 479	-
GLY 853 (P)	GLU 474, ASP 478	THR 475	-
ASP 854 (C)	GLU 474	THR 475	LEU 471,
THR 856 (P)	ASP 478	-	-
VAL 909 (AP)	ASP 478	-	-
SER 910 (P)	ARG 437, ASP 478, LYS 481	-	-
ALA 911 (AP)	ARG 437, ASP 478	THR 475, ASN 479	ALA 480
ALA 912 (AP)	ARG 437, ASP 478, LYS 481,	ASN 479, GLN 482	ALA 480, LEU 483
GLN 913 (P)	ARG 437, ASP 478,	ASN 479, GLN 482,	MET 486
ASP 914 (C)	ARG 437	-	-
TYR 915 (P)	ARG 437	ASN 479	-
ALA 916 (AP)	ARG 437	-	-
SER 937 (P)	-	ASN 479	-
ARG 940 (C)	-	TRP 425, THR 475	LEU 422, PRO 472,
PRO 941 (AP)	-	TRP 425	LEU 422
ASP 942 (C)	-	TYR 424, TRP 425, SER 430	LEU 422, ILE 423, PHE 426, LYS 434
PRO 943 (AP)	-	TYR 424, TRP 425	LEU 422, ILE 423, PHE 426
PRO 944 (AP)	LYS 371	TYR 424	ILE 423, PHE 426
SER 945 (P)	-	-	ILE 423
GLY 946 (AP)	-	-	LEU 366, ILE 423
LYS 948 (C)	GLU 311	-	-
VAL 949 (AP)	GLU 311, LYS 367	SER 364,	LEU 307, LEU 363, LEU 366
VAL 950 (AP)	GLU 311, LYS 367	-	-
ARG 976 (C)	GLU 351	GLN 354	-
TYR 1009 (P),	GLU 351	-	-
ARG 1010 (C)	-	-	-
ALA 1015 (AP)	-	-	LEU 422
LEU 1019 (AP),	-	THR 475	-
GLY 1020 (AP)	-	-	LEU 422
SER 1022 (P)	-	-	ALA 421, LEU 422
GLY 1023 (AP)	-	-	ALA 421, LEU 422, ILE 423
LEU 1024 (AP)	-	-	LEU 366, PHE 419, LEU 420, ALA 421, LEU 422, ILE 423
ALA 1025 (AP)	-	-	PHE 419, ALA 421
ASP 1026 (C)	LYS 469	-	PRO 418, PHE 419, LEU 420, ALA 421
LYS 1027 (C)	GLU 466	ASN 460	PHE 419, LEU 420, ALA 421
GLY 1028 (AP)	-	-	LEU 363, LEU 366, PHE 419, LEU 420
THR 1029 (P)	ASP 357	-	LEU 363, LEU 366
GLN 1030 (P)	-	-	LEU 363
PRO 1032 (AP)	-	-	-
THR 1034 (P),	ARG 360	-	-
GLN 1040 (P)	-	-	GLY 356
PRO 1041 (AP)	-	-	GLY 356, GLY 359,
SER 1042 (P)	ASP 357, ARG 360	HIS 358	GLY 356
GLY 1043 (AP)	ARG 360	-	-
GLU 1044 (C)	ARG 360	-	-
ASP 1047 (C)	-	HIS 264	-
GLN 1048 (P)	-	TYR 251	PHE 252
LEU 1049 (AP)	GLU 200, LYS 261	HIS 264	PHE 252

PRO1050 (AP)	GLU 200	TYR 251	PHE 252
THR 1051 (P)	GLU 200	-	PRO 196, LEU 201
GLU 1052 (C)	GLU 200	-	PRO 196, LEU 201
PRO 1053 (AP)	GLU 200 , ASP 250	SER 195 , TYR 251	PRO 196 , PHE 252
PRO 1054 (AP)	ASP 250	TYR 251	PRO 196
SER 1055 (P)	LYS 154, ASP 250	HIS 249	PRO 196
GLY 1056 (AP),			
PRO 1057 (AP),	ASP 250	-	-
SER 1058 (P)			
GLY 1059 (AP)	GLU 900	SER 903	-
LEU 1060 (AP)	GLU 900	-	LEU 899
PRO 1061 (AP)	-	-	LEU 899
PRO 1061 (AP)	GLU 900	-	-
LEU 1062 (AP)	ARG 895, GLU 900	-	LEU 898, LEU 899
LEU 1063 (AP)	-	TYR 896	ILE 891, LEU 899
PRO 1064 (AP)	ARG 895	TYR 896	ILE 891, LEU 899
VAL 1065 (AP)	-	-	ILE 891
LEU 1066 (AP)	LYS 888 , ASP 890, ARG 895	GLN 889	ILE 887 , ILE 891
PHE 1067 (AP)	-	GLN 889	-
PHE 1067 (AP)	ILE 891		-
ALA 1068 (AP)	LYS 888 ,	GLN 889	-
LEU 1069 (AP)	GLN 889	-	-
GLY 1070 (AP)	GLU 881, GLU 883	GLN 889	GLY 882
GLY 1071 (AP)	GLU 881, GLU 883	GLN 889	GLY 882
LEU 1072 (AP)	GLU 881, GLU 883	-	GLY 882
LEU 1073 (AP)	VAL 879, GLU 881	-	
VAL 1081 (AP)	GLU 881	-	
VAL 1084 (AP)	-	GLN 889	-
TRP 1086 (P)	GLU 886 , ASP 890	GLN 889	GLY 882, ILE 887, ILE 891
ARG 1123 (C)	LYS 888	GLN 889	-
TYR 1139 (P)	GLU 883	-	-
ARG 1140 (C)	GLU 881, GLU 883, GLU 886	-	GLY 882
CD2AP - Podocin Interacting residues ΔG: -21.8 kcal mol⁻¹			
CD2AP	Podocin		
		Salt Bridges:	
GLU 34 (C)	LYS 377	-	-
ASP 51 (C)	LYS 148	-	-
LYS 579 (C)	GLU 303	-	-
GLU 592 (C)	ARG 306	-	-
		Hydrogen bonds:	
MET 1 (AP)	ASP 250	-	-
GLN 33 (P)	ASP 379	-	-
ARG 59 (C)	-	SER 205	-
GLU 60 (C)	-	SER 205	-
LYS 64 (C)	-	SER 86, SER 120	-
SER 399 (P)	-	-	LEU 373
THR 407 (P)	-	-	PRO 369
THR 468 (P)	LYS 368	-	-
LEU 487 (AP)	-	-	LEU 324
THR 490 (P)	-	HIS 325	-
ALA 562 (AP)	ARG 322	-	-
PHE 563 (AP)	ARG 322	-	-
THR 575 (P)	LYS 299	-	-
VAL 578 (AP)	ARG 306	-	-
LYS 579 (C)	GLU 303,	-	-
		Non-bonded contacts:	
MET 1 (AP)	ASP 250	CYS 254, TRP 256	ILE 255
VAL 2 (AP)	-	CYS 254, TRP 256,	ILE 255, GLY 257

ASP 3 (C)	-	TRP 256	ILE 255
TYR 4 (P)	-	-	ILE 255
GLN 33 (P)	ASP 379	SER 380	PRO 381
GLU 34 (C)	LYS 377, ASP 379		
ASP 51 (C)	LYS 148		
ILE 57 (AP)	-	TRP 256	LEU 207, ILE 255
LYS 58 (C)	-	SER 205, SER 206, TRP 256	ALA 147, LEU 207, GLY 257
ARG 59 (C)	-	SER 205, SER 206	ILE 113, LEU 204, LEU 207
GLU 60 (C)	-	SER 201, SER 205, SER 206	PRO 150, PHE 154, LEU 204,
THR 61 (P)	-	THR 116, SER 205	ILE 113, PHE 155, LEU 204, LEU 207
GLU 62 (C)	-	THR 116,	PHE 119, PHE 154, PHE 155, LEU 204
PHE 63 (AP)	-	THR 116	PHE 119, PHE 123
LYS 64 (C)	-	SER 86, THR 116, SER 120	GLU 87, PRO 89, PHE 117, PHE 119, PHE 123
ASP 66 (C)	GLU 87	-	PHE 123
ILE 70 (AP)	-	SER 86	ALA 82, LEU 83
LYS 71 (C)	GLU 79	-	-
HIS 75 (P)	-	-	LEU 207
SER 86 (P)	ARG 13	-	-
LEU 90 (AP)	-	-	-
PRO 91 (AP)	-	TRP 122	PHE 119, PHE 123, PHE 155, PRO 157
ALA 92 (AP)	-	-	PHE 119
PRO 99 (AP)	-	-	GLY 149
GLN 100 (P)	-	-	GLY 149, PRO 150, LEU 152, PHE 154
THR 101 (P)	-	-	LEU 152
LYS 102 (C)	LYS 148,	-	GLY 149, LEU 152
ASN 138 (P)	ARG 13	-	-
GLU 139 (C)	ARG 10	-	-
GLU 175 (C), GLN 177 (P)	ARG 13	-	-
ASN 323 (P), ALA 349 (AP), PRO 350 (AP), LYS 351 (C)	LYS 148	-	-
LYS 351 (C)	-	-	GLY 149, LEU 152
PRO 352 (AP)	LYS 148	-	LEU 152
GLU 353 (C)	-	-	LEU 152
THR 396 (P)	LYS 377, ASP 379		
LYS 397 (C)	-	ASN 374	-
ALA 398 (AP)	-	ASN 374	PRO 372, LEU 373
SER 399 (P)	-	ASN 374	PRO 372, LEU 373
ASN 400 (P)	GLU 371,	-	PRO 372, LEU 373
LEU 401 (AP)	GLU 371	-	VAL 338, PRO 362, VAL 370, PRO 372, LEU 373
ARG 403 (C)	-	-	VAL 370
SER 404 (P)	GLU 371	-	PRO 369, VAL 370
SER 405 (P)	-	-	PRO 369
GLY 406 (AP)	LYS 368	-	PRO 369
THR 407 (P)	LYS 368	-	PRO 369, VAL 370
VAL 408 (AP)	-	-	PRO 369
PHE 461 (AP)	LYS 368	-	-
VAL 466 (AP)	LYS 368	-	-
ARG 467 (C)	LYS 368, GLU 371	SER 367	-
THR 468 (P)	LYS 368	-	PRO 369
SER 469 (P)	LYS 368	SER 367	PRO 366, PRO 369
LYS 470 (C)	LYS 368	SER 367	PRO 366

GLU 471 (C)	-	SER 365, SER 367	PRO 366
THR 472 (P)	-	SER 365, SER 367	LEU 340, PHE 363, PRO 366
ALA 481 (AP)	-	SER 365	-
SER 482 (P)	-	SER 365	LEU 342, PRO 343, PHE 363, PRO 364,
SER 483 (P)		THR 332	LEU 330, PRO 343, PHE 344
GLU 484 (C)	-	-	LEU 324, LEU 327, LEU 330, PHE 344
ASN 485 (P)	-	-	LEU 324, LEU 327, LEU 330, PRO 343, PHE 344
LEU 486 (AP)	-	HIS 325, THR 326, GLN 328	LEU 324, LEU 327 PHE 344
LEU 487 (AP)	-	HIS 325	LEU 321, LEU 324, LEU 327
THR 490 (P)	ARG 322	HIS 325	LEU 321
ALA 491 (AP)	-	HIS 325	-
HIS 513 (P), SER 514 (P), GLU 516 (C), THR 561 (P)	-	-	LEU 321
ALA 562 (AP)	ARG 322	-	ALA 308, ILE 311, LEU 312, ALA 318, VAL 319, LEU 321, SER 304, LEU 305, MET 307, ALA 308, ILE 311, LEU 312
PHE 563 (AP)	GLU 303, ARG 322	-	-
LEU 564 (AP)	ARG 322	HIS 325	-
THR 565 (P)	-	SER 304	-
LEU 567 (AP)	-	SER 304	ALA 300, ALA 301,
GLU 568 (C)	GLU 296	-	-
LYS 570 (C)	-	-	ALA 300
ALA 571 (AP)	GLU 296	-	ALA 297, ALA 300
LYS 572 (C)	GLU 296	-	-
THR 575 (P)	GLU 296, LYS 299	-	ALA 295, ALA 297, ALA 300
ASP 576 (C), ASP 577 (C)	LYS 299	-	-
VAL 578 (AP)	GLU 303, ARG 306	-	-
LYS 579 (C)	GLU 298, LYS 299, GLU 303, ARG 306	SER 302	ALA 300
LYS 580 (C)	GLU 298, LYS 299, ARG 306	SER 302	ALA 294
ASN 581 (P)	LYS299	-	-
SER 582 (P)	GLU 298, LYS 299	-	ALA 294, ALA 295
LEU 583 (AP)	ARG 291		VAL 290, MET 292, ALA 294, ALA 295
GLU 592 (C)	ARG 306	-	-
CD2AP - TRPC6 Interacting residues ΔG: -4.2 kcal mol⁻¹			
CD2AP	TRPC6	Non bonded contacts:	
LYS 114 (C)	ASP 877	-	-
VAL 131 (AP)	GLU 883	-	-

Chapter 4: Structural features and oligomeric nature of human podocin domain

4.1. Prelude

Podocin is a 383 amino acid protein localizing to the lipid rafts along with other SD proteins [50, 52, 105]. Podocin shares 44% homology and several structural similarities with stomatin family proteins due to the presence of a highly conserved Prohibitin (PHB) domain [38, 90]. Podocin adopts a hook-like structure with its cytoplasmic N- and C-terminals and attaches to the inner side of the plasma membrane via a transmembrane domain at 100-125 residues [75, 90, 110]. Structural characterization of stomatin revealed that different truncations of the protein associated with different oligomeric states and the C-terminus of the protein are crucial for homo-oligomerization [111-113]. Studies with truncated C-terminal human podocin revealed that it forms a dimer [79]. Though it was proposed that longer constructs of podocin were capable of associating into higher-order oligomers, it was never demonstrated [79]. Additionally, co-immunoprecipitation studies with nephrin, CD2AP, TRPC6, and KIRREL1 indicated that these proteins interact majorly with the C-terminus of podocin [50, 52, 53, 58, 92, 105, 114]. We reported earlier (in chapter 3) that these interactions are mediated by the intrinsically unstructured regions (IURs) present in these proteins [38].

Although a large body of evidence suggests a central role for the podocin in the SD protein complex, the precise mechanism by which podocin oligomerizes and acts as a scaffolding molecule remains to be elucidated. Importantly, it is not known whether all the other SD proteins interact with a single podocin or with homo-oligomers? Therefore, in this study, we attempted to understand the oligomeric nature of protein using a truncated construct (residues: 126-350), which encompasses the PHB domain, C-terminal oligomerization site, and 4 out of the 6 cysteines present in the native podocin sequence.

4.2. Material and Methods

4.2.1. Protein cloning, expression, and purification

The codon-optimized podocin gene (1152bp) was purchased from Gene Art (Life Technologies, USA). The regions encoding the amino acids 126-350 (podocin domain) was amplified with the primers 5' CCC GAA TTC G AAA GTG GTG CAA GAA 3' (forward) and 5' GAA CTC GAG CAG ACA ATT CAG CAG ATC 3' (reverse) and cloned into pET22b at EcoRI/XhoI sites. The recombinant construct was transformed into Arctic express (DE3) competent cells (Agilent Technologies, USA). The transformed cells were grown at 37°C in LB media supplemented with 100 µg/ml ampicillin and 10 µg/ml gentamycin. Protein expression was induced with 0.2 mM IPTG and cultured further for 16hrs at 14°C. The cells were harvested by centrifugation (13300 x g, 20 mins, and 4°C) and sonicated in the opening buffer (50 mM potassium phosphate (pH 8.0), 0.3 M NaCl, 5 mM β-mercaptoethanol, and 0.1% Triton X-100) followed by clarification by centrifugation at 18000 x g for 45 mins at 4°C. The inclusion bodies were solubilized in 50 mM potassium phosphate (pH 8.0), 0.3 M NaCl, 5 mM β-mercaptoethanol and 8 M Urea followed by clarification at 18000 x g for 1hr at room temperature. The solubilized protein was then purified using Ni-NTA agarose (Qiagen). The purity was confirmed (>98%) on a 12% SDS-PAGE. From the SDS-PAGE gel, the band corresponding to the podocin domain (27 kDa) was excised and subjected to tryptic digestion and analyzed using MALDI-TOF/TOF (Bruker Autoflex III smart beam, Bruker Daltonics, Bremen, Germany) to confirm the protein sequence.

The purified podocin was renatured by rapid dilution at 1:10 into 10mM potassium phosphate buffer with 150 mM NaCl and 2 mM β-mercaptoethanol (pH 8.0) followed by dialysis in the same buffer to remove traces of urea and imidazole. The same pH and buffer composition are uniformly used in all the subsequent experiments. An

extinction coefficient of the $12950 \text{ M}^{-1}\text{cm}^{-1}$ was used for determining the protein concentration on Jasco V-630 UV-Vis spectrophotometer

4.2.2. Size exclusion chromatography and multi-angle light scattering analysis (SEC-MALS)

We performed SEC-MALS to estimate the oligomeric nature of the podocin domain. SEC-MALS was performed at room temperature by passing 500 μl protein ($12\mu\text{M}$) at 0.3 ml/min flow rate through a Superdex S200 SEC column (GE Healthcare) pre-equilibrated with 10mM potassium phosphate buffer supplemented with 150mM NaCl and 2mM β -mercaptoethanol (pH 8.0). This column was attached to the MALS system (AF2000- Postnova) for analyzing the molar mass of the protein. The protein sample from SEC-MALS was next passed through the flow cell equipped Zetasizer Nano ZS90 dynamic light scattering (DLS) device (Malvern Instruments Ltd, UK) equipped with a 4 mW He-Ne laser. The backscattering was measured at 173 nm for analyzing the polydispersity index (PDI) and the hydrodynamic radius of the protein.

4.2.3. Fluorescence spectroscopy

Intrinsic tryptophan fluorescence was measured using Jasco FP-6300 (Japan) equipped with an intense xenon flash lamp as the light source. 287 nm was used as the excitation wavelength and the 300-450 nm spectral range was used for obtaining the emission spectrum of the sample ($12 \mu\text{M}$). For accessing the stability of the podocin domain thermal-induced unfolding was performed. Fluorescence emission at 335 nm as a function of increasing temperature was recorded in triplicates at a bandwidth of 2.5 nm and a scan speed of 200 nm/min for every degree rise in temperature. All the spectra were buffer corrected. The effect of temperature on the protein stability was analyzed by plotting fluorescence intensity at 335 nm using Origin (pro)-version2020b (Origin Lab Corporation, Northampton, MA).

4.2.4. Circular dichroism (CD) spectroscopy

The CD spectroscopy measurements were recorded on Jasco J-1500 spectropolarimeter (Japan) equipped with a thermoelectric cell holder. The Far-UV (260-195 nm) CD measurements of podocin (12 μ M) were recorded using a 0.2 cm path length cell at 2.5 nm bandwidth and a scan speed of 50 nm/min. The Near-UV CD measurements were also recorded for the protein sample at 25 μ M concentration using a 0.5cm pathlength cell at the bandwidth of 2.5nm and a scan speed of 100 nm/min. Both for far-UV and near-UV the data was recorded in triplicates. To assess the effect of temperature and thus the stability of the domain, the sample was subjected to a steady increase in temperature, and spectra were recorded at an interval of 5°C over a spectral range of 200 nm-250 nm. The data were plotted using origin lab software after buffer correction.

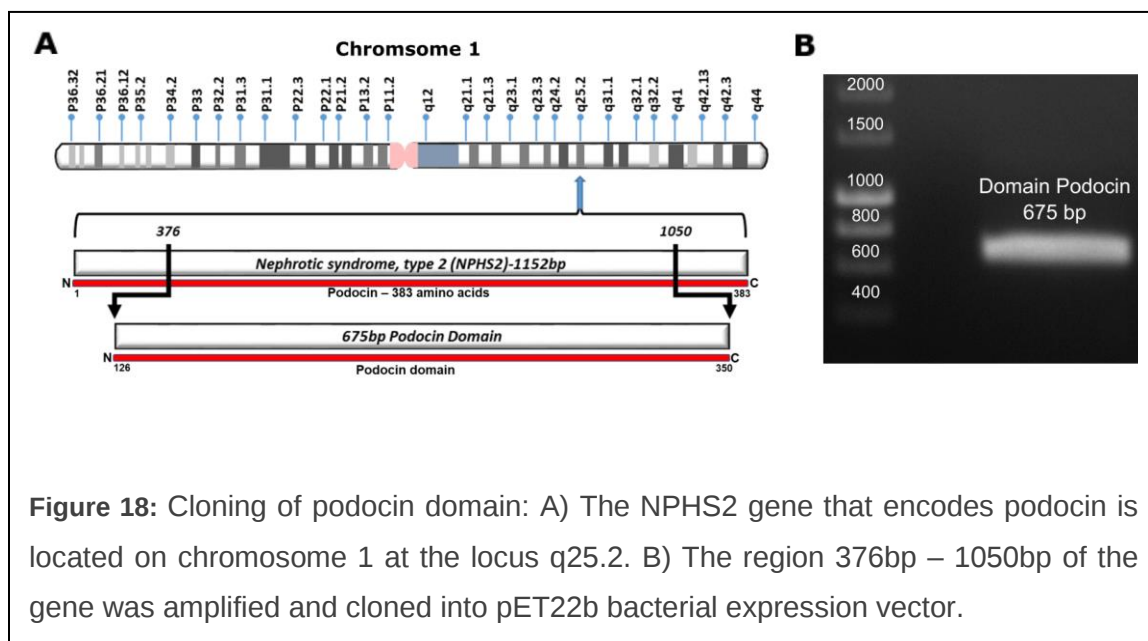
4.2.5. Calorimetric analysis

The transition temperature (T_g) of the podocin domain were obtained from measurements using NANO DSC (TA Instruments, USA). Sample containing protein concentration of 12 μ M and a volume of 0.650 μ L was loaded into the sample capillary and change in heat flow was recorded against reference buffer at a constant pressure of 3 atm, over a temperature range of 293K to 368K, with a scan rate of 1 K/min and a 300-sec cell equilibration time. Buffer scans were first performed before loading protein for baseline reproducibility. The obtained data was buffer corrected and the analysis data was plotted for peak integration with the peak analyzer option in Origin pro 2020b software. From the peaks, the T_g and the enthalpy of transition (ΔH_{cal}) were calculated.

4.3. Results

4.3.1. Purification of Podocin domain

The region from 376bp to 1050bp of human *NPHS2* gene encoding 126-350 amino acid residues of podocin was PCR amplified and cloned into the pET22b expression vector (Fig. 18A&B). IPTG-induced expression of the construct resulted in the protein to form inclusion bodies which were confirmed by Coomassie blue staining and immunoblotting



with anti-His antibody (Fig. 19A&B). The protein was purified from inclusion bodies by Ni-NTA affinity chromatography after solubilization in 8M urea (Fig. 19C). Tryptic digestion of the band corresponding to the podocin domain from the SDS-PAGE gel (Fig. 19C-lane E2) and subsequent analysis of the digested products by MALDI-TOF/TOF revealed five peptide fragments (Fig. 19D and Table. 8). NCBI BLAST search of these peptide sequences against the non-redundant database confirmed the purified protein as human podocin covering the region 126-350 amino acids.

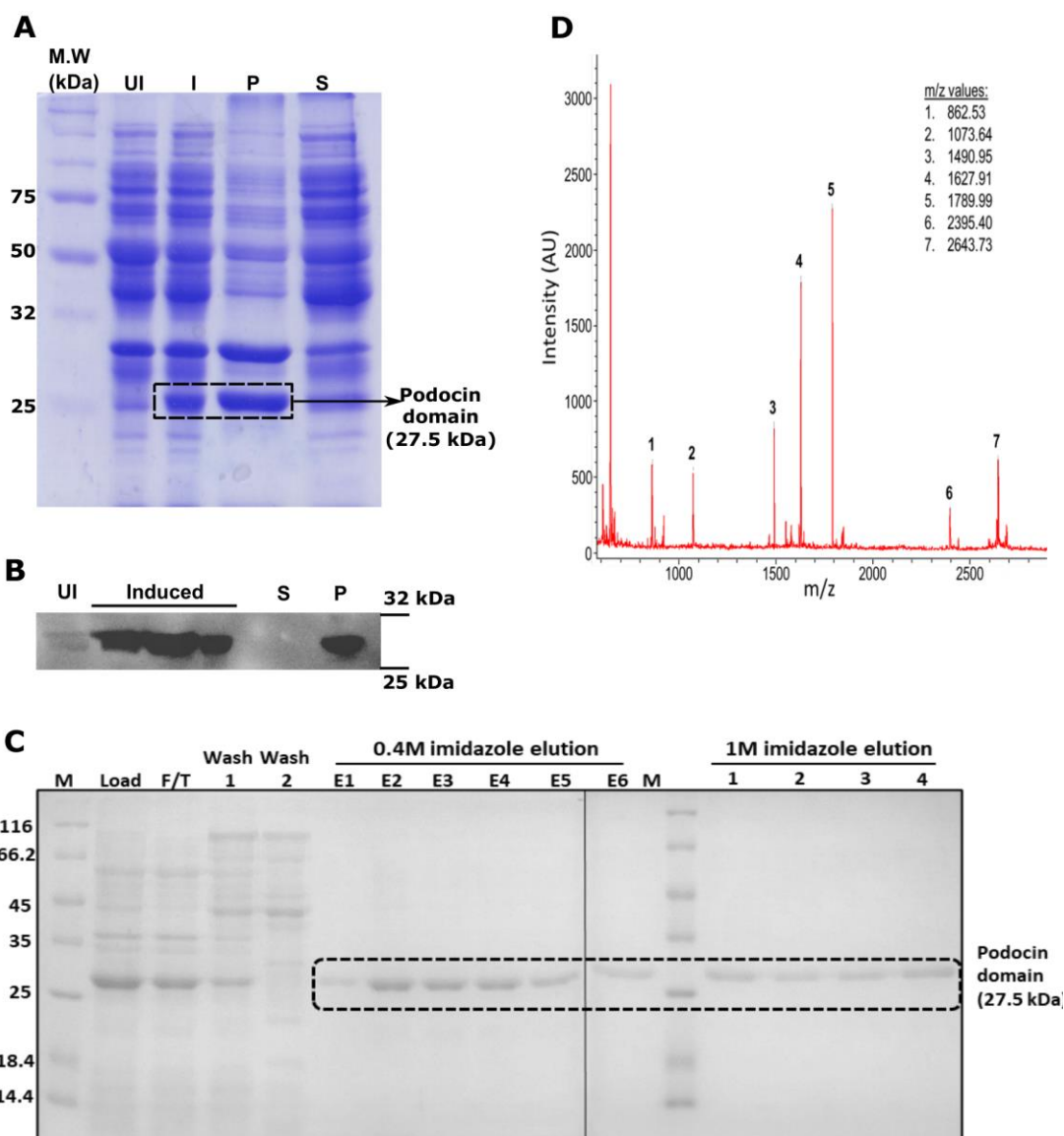


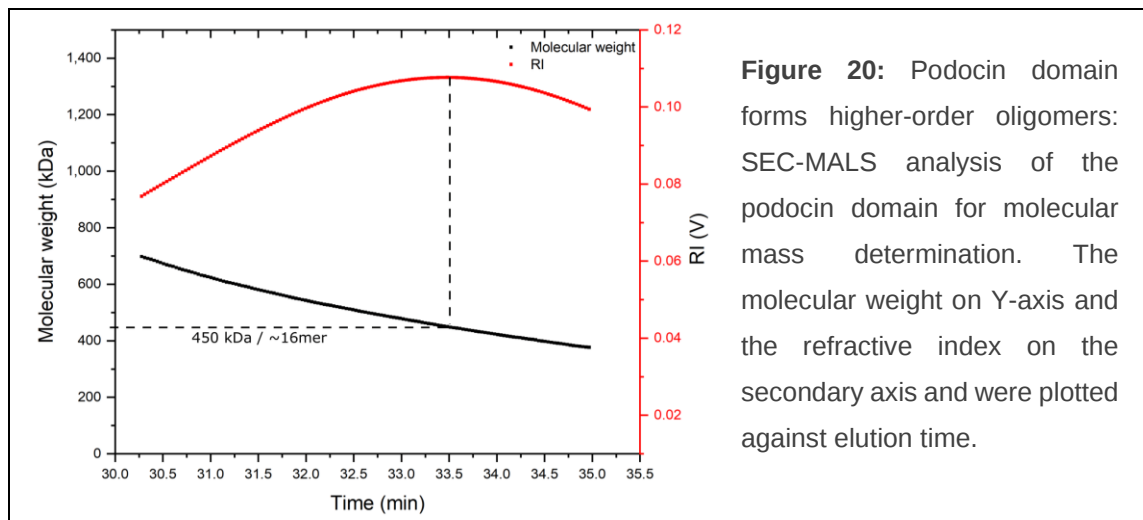
Figure 19: Expression, and purification of podocin domain: A & B) Coomassie blue staining and immunoblotting His-Tag HRP-conjugated antibody were done to identify the expression and solubility of the podocin domain. C) The SDS-PAGE analysis of samples after affinity chromatography purification of the urea solubilized cell lysate. D) Trypsinization and MALDI-TOF/TOF analysis of the 27KDa band excised from the earlier SDS-PAGE gel confirmed the presence of the podocin domain. UI-uninduced culture, I-Isopropyl β , D - thiogalactopyranoside Induced culture, P-Pellet fraction, S-Soluble fraction, F/T- flow through, E1-E6: elution fractions with 0.4M imidazole, 1-4: elution fractions with 1M imidazole.

Table 8: MALDI TOF/TOF analysis of the purified protein: Trypsinization of the purified band at 27 kDa and subsequence analysis by MALDI-TOF/TOF showed 5 peptide sequences. BLAST analysis of these sequences against the non-redundant proteins database of NCBI showed 100% similarity with the human podocin sequence. Note: '▲' indicates the site of digestion by trypsin at arginine and lysine residues in the sequence.

Sequence		Observed	Mr (expt.)	Mr (calc.)	Expect	Peptide
Start	End					
134	146	1490.95	1489.95	1489.93	2.3e-03	R▲VIIFRLGHLLPGR▼A
149	168	2395.40	2394.40	2394.22	9.1e-03	K▲GPGFLFFFLPCLDTYHKVDLR▼L
263	286	2643.73	2642.72	2642.45	6.1e-04	R▲IEIKDVRLPAGLQHS�AVEAEAQR▼Q
270	286	1789.99	1788.99	1788.95	2.4e-06	R▲LPAGLQHS�AVEAEAQR▼Q
307	322	1627.91	1626.90	1626.88	3.5e-06	R▲MAAEILSGTPAAVQLR▼Y

4.3.2. Oligomeric nature of the podocin domain

It was reported that stomatin family members exist as homo-oligomers [111-113]. Since podocin shares significant homology with stomatin, we analyzed the oligomeric nature of the podocin which includes the PHB domain. SEC-MALS data is represented as a combinatorial plot of refractive index, and molecular weight versus elution time (Fig. 20). A maximum refractive index value of 0.11 was observed which corresponds to a molecular weight of 450kDa, suggesting that the podocin domain is a 16-mer oligomer (monomer =



27.5kDa; therefore; 450 kDa/ 27.5 kDa = ~16-mer). In addition to the predominant 16-mer species, other oligomeric conformations of the podocin domain ranging from 25-mer

(refractive index: 0.08, molecular weight: 697kDa) to 13-mer (refractive index: 0.10, molecular weight: 376kDa) were also observed, but to a lesser extent. The DLS, which was in tandem with the SEC-MALS analyzed the eluted samples for hydrodynamic radius and polydispersity.

A hydrodynamic radius range of 13.39 - 9.37 nm corresponding to elution fractions from 25-mer to 13-mer was observed respectively (Fig. 21A-E). The average PDI of the

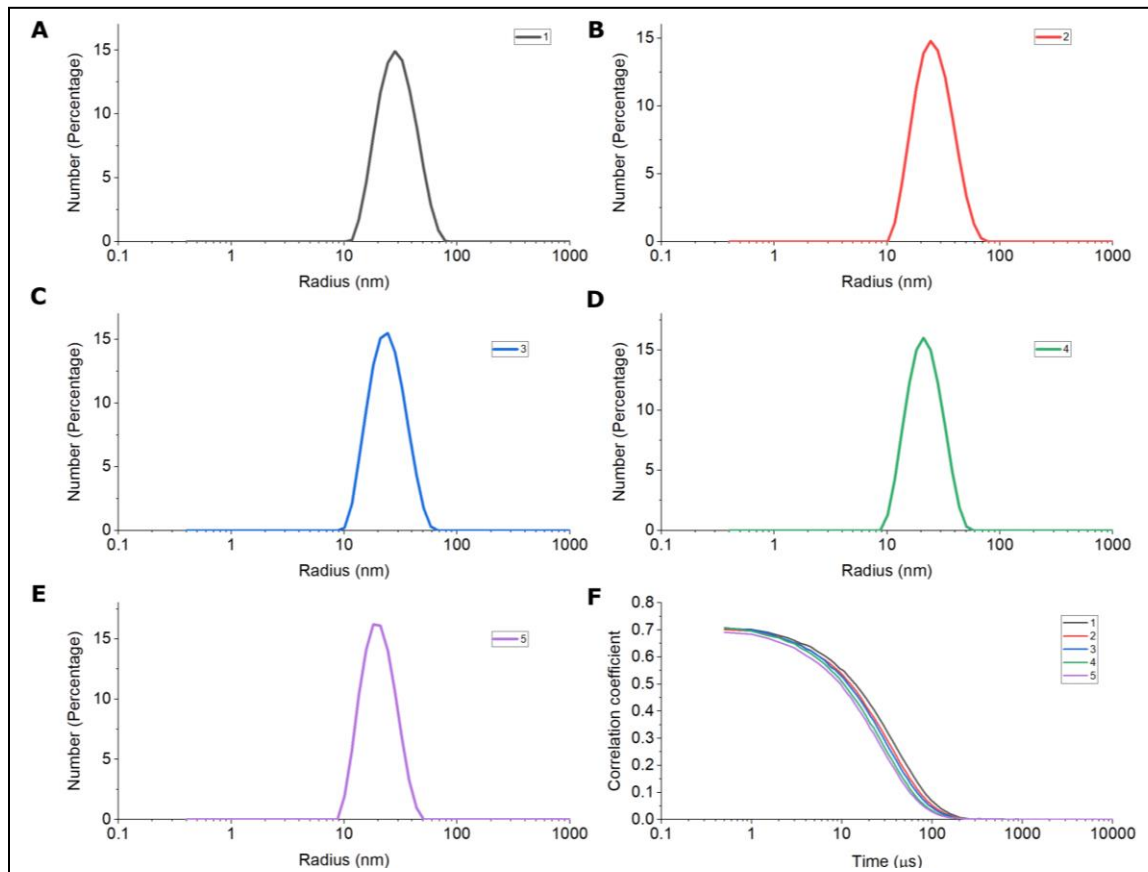


Figure 21: Polydispersity and hydrodynamic radius of the podocin domain homo-oligomers: The DLS was in tandem with the SEC-MALS. DLS analysis of the samples corresponding to peak observed in SEC-MALS (1-5) is represented as a number percentage vs size in nanometers curve plots marked in different colors (A-E). The corresponding correlograms of the samples (1-5) represented as correlation coefficient vs time in microseconds.

sample was found to be 0.14, which suggests that the sample consists of one major population by volume, however, a broad size range within the population implies the presence of different size species. From these results, it is evident that the major species

of the podocin domains is a 16-mer oligomer while a minor population of other higher-order oligomers also appears to exist in solution.

4.3.3. Structural features of the homo-oligomers of podocin domain

The CD spectra in the far UV region (250-195 nm) provides information on the secondary structure while the near UV region (310-250 nm) on the tertiary packing. The far UV CD signal for the truncated podocin domain indicates that the protein adopts a secondary structure (Fig. 22A). The shape of the spectrum suggests the presence of both α -helices and β -sheet structures. The near UV spectrum of the sample shows a broad

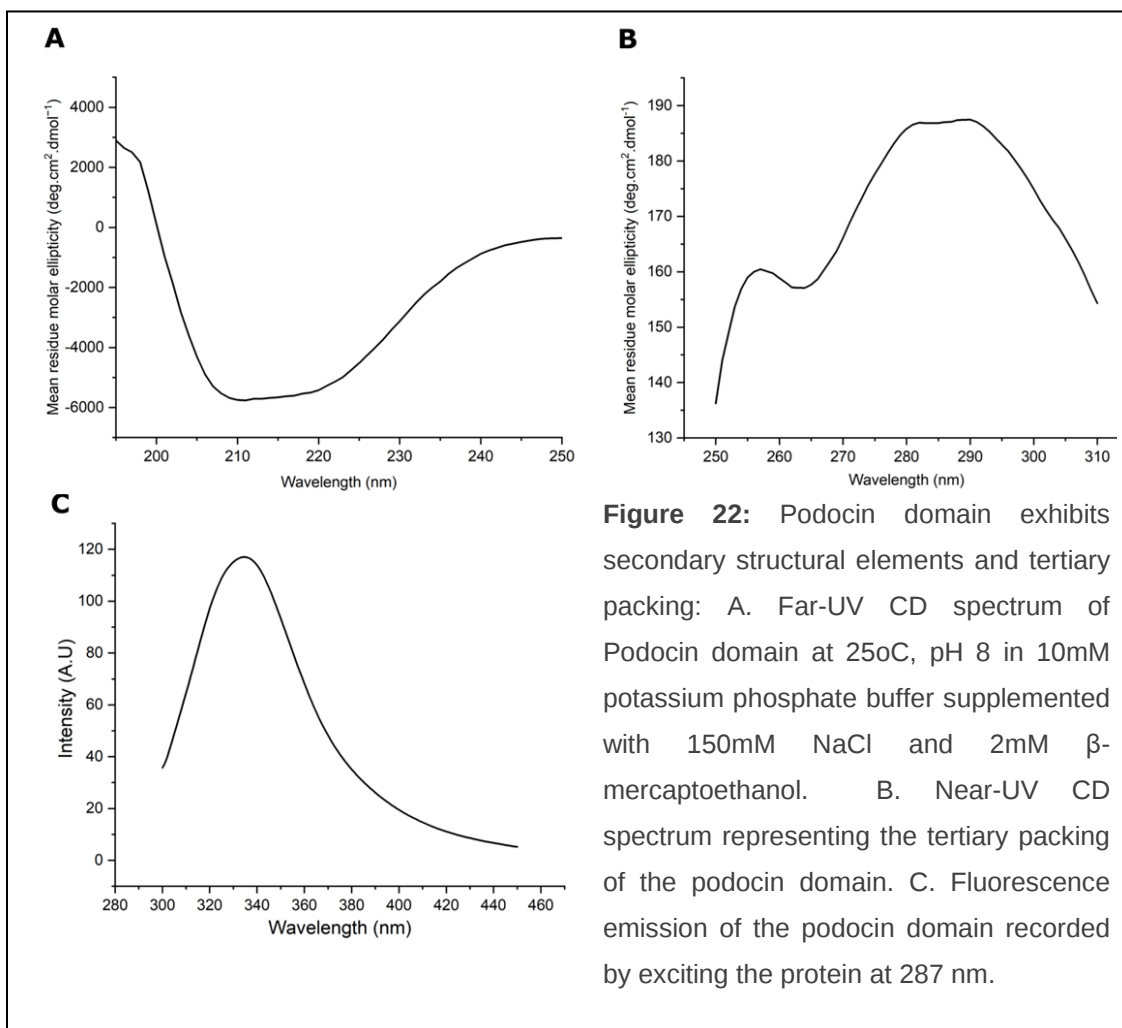


Figure 22: Podocin domain exhibits secondary structural elements and tertiary packing: A. Far-UV CD spectrum of Podocin domain at 25oC, pH 8 in 10mM potassium phosphate buffer supplemented with 150mM NaCl and 2mM β -mercaptoethanol. B. Near-UV CD spectrum representing the tertiary packing of the podocin domain. C. Fluorescence emission of the podocin domain recorded by exciting the protein at 287 nm.

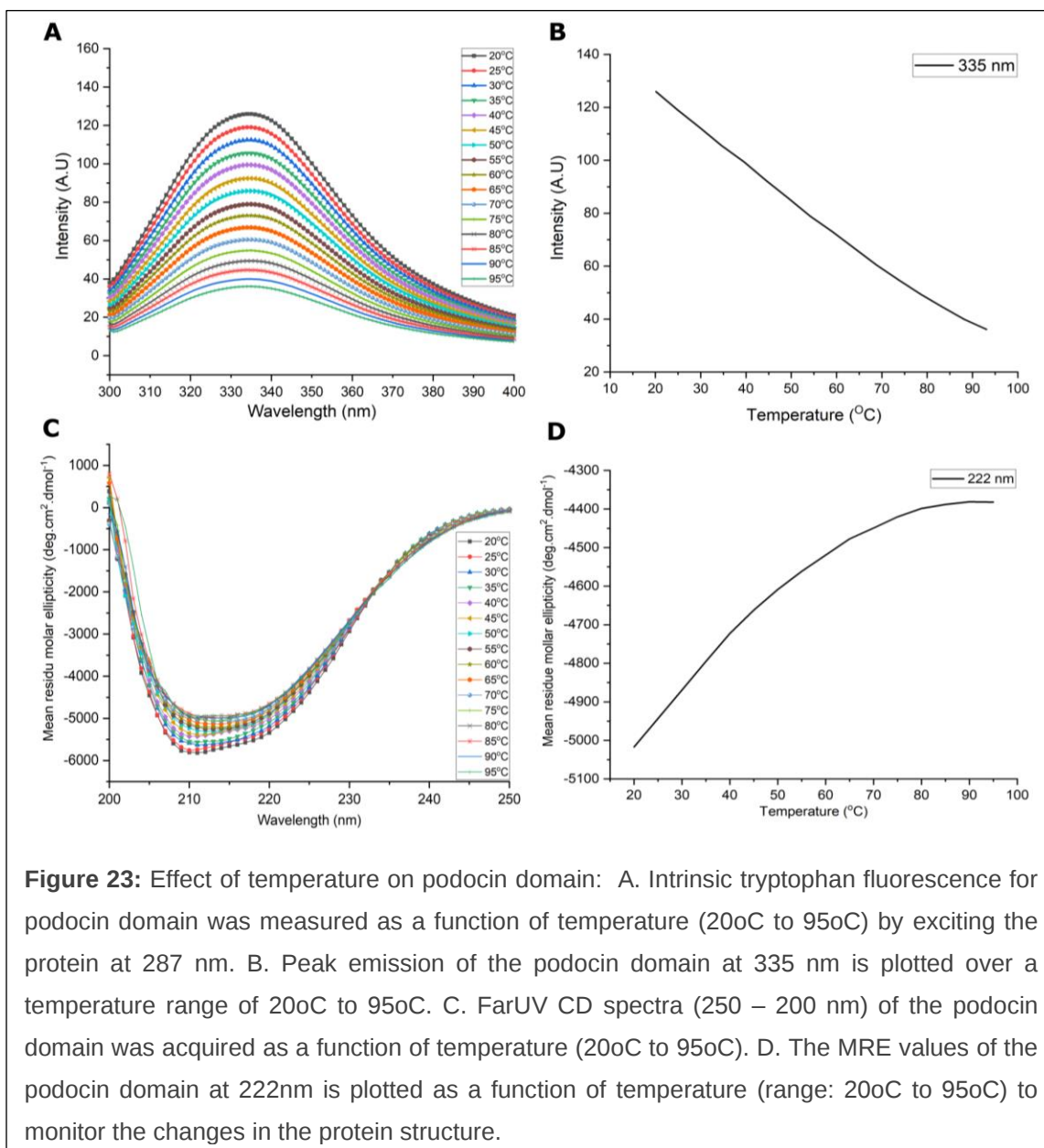
peak at 280nm typically contributed from tryptophan residue (W256), and a minor peak from ~250 – 260 nm could be contributed by five-tyrosine and eight-phenylalanine

residues (Fig. 22B). The CD data in the near UV regions suggests the arrangement of aromatic amino acids in a restricted environment and thus implying a folded structure adopted by the polypeptide chain. The intrinsic tryptophan emission spectrum of the podocin domain shows λ_{max} value at 335 nm (Fig. 22C). The λ_{max} of native proteins is related to the polarity of the environment of the tryptophan residue and typically can range from 308-350 nm. Unfolded forms with residues in the apolar microenvironment show a blue shift of the λ_{max} [115].

4.3.4. Structural stability of the podocin domain

Although, the intrinsic fluorescence may not unequivocally provide structural information, the changes in the folded state can be followed by monitoring the emission intensity which conveys the changes in the native state tryptophan environment. The changes in the CD signal and the perturbations of the intensity of the intrinsic fluorescence emission were estimated to assess the stability of the folded podocin domain. Temperature-induced unfolding was monitored using the fluorescence emission spectra with increasing temperature in the range of 20°C-95°C. The spectra showed a uniform decrease in λ_{max} intensity without either a bathochromic or hypsochromic shift (Fig. 23A). The change in fluorescence intensity as a function of temperature (Fig. 23B) shows a linear transition and does not show a sigmoidal shape typically observed for proteins with a tightly packed tertiary core. The absence of a folded baseline suggests a not so tightly packed tertiary core. Far UV CD spectral changes were also monitored as a function of temperature. With the gradual increase in temperature a gradual loss of signal intensity (Fig. 23C), loss of secondary structure, and the absence of a native baseline was observed

(Fig. 23D) implying that the podocin domain possesses secondary and tertiary interactions but lack a tightly packed core.



4.3.5. Calorimetric analysis of podocin domain

Differential scanning calorimetry (DSC) gives the overall enthalpy value (ΔH_{cal}) for each structural transition. Therefore, we performed DSC to calculate the thermodynamic parameters such as T_g and the ΔH_{cal} associated with structural changes of the homooligomer. Peak deconvolution of the acquired data revealed 5 transition states, out of

which four are endothermic transitions (316 K, 325 K, 330 K, and 353 K) and one is an exothermic transition (345 K) (Fig. 24). The respective values for ΔH_{cal} are mentioned in Table 9. The DSC profile suggests that the oligomers in the mixture undergo dissociation via three transition temperatures namely 316 K, 325 K, and 330 K and the presence of exothermic transition at 345 K suggests possible hydrophobic interactions among the constituent oligomers before complete dissociation at 353 K.

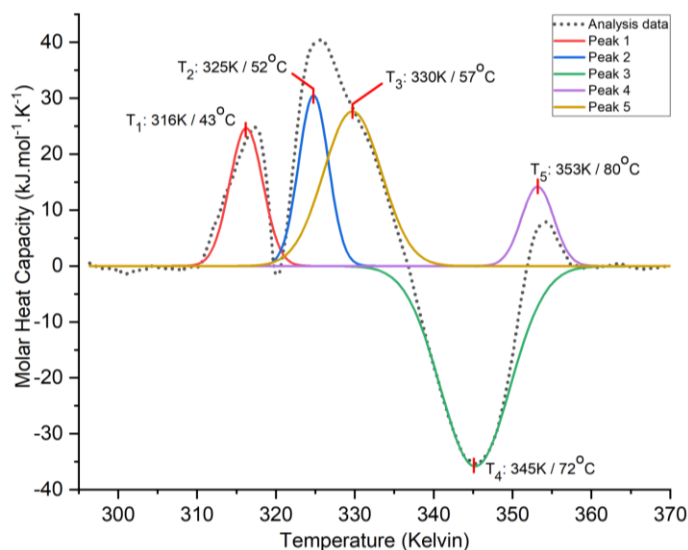


Figure 24: DSC analysis of podocin domain: The plots represent the endothermic and exothermic transitions of the podocin domain over a temperature range from 296K/23oC to 368K/95oC. The initial DSC artifact near 293K/20oC in the DSC data is not represented in the plot. The analysis data represented as a dotted line in the plot after baseline subtraction. Each transition peak (1-5) in the plot is represented by a color line which is a result of peak deconvolution function.

Table 9: Enthalpy values and the transition temperatures as noticed in the dynamic scanning calorimetry.

Temperature range	Transition temperatures (T _g)		Enthalpy change (ΔH_{cal}) (kJ/mol)
293K to 368K / (20°C – 95°C)	Endothermic	316 K / 43°C	130
		325 K / 52°C	145
		330 K / 57°C	73
		353 K / 80°C	254
	Exothermic	345 K / 72 °C	-397

4.4. Discussion

Podocin selectively expresses in the glomerular podocytes and it is instrumental for preserving the structural integrity of the SD. Though several mutations in the protein are associated with proteinuria in humans, the structural details of this protein are unclear. Here we report for the first time, the stoichiometry of oligomerization of the truncated human podocin construct. Our investigation indicates that at ambient temperature, and in a reduced environment the podocin domain predominantly adopts a 16-mer oligomeric state. However, other oligomeric conformations ranging from 25-mer to 13-mer were also observed nevertheless, the population of these states was comparatively less. The polydispersity index we report adds evidence to the presence of multiple oligomeric species. Additionally, CD and fluorescence spectra revealed podocin homo-oligomers have considerable secondary structure and tertiary packing.

Since stomatin proteins and podocin share significant homology, it is expected they may share several structural similarities. Crystallization studies of stomatin protein (*Pyrococcus horikoshii*; residues 56–224) revealed that it exists as a trimer and NMR studies of the same protein with a different truncation size (residues 66–174) associated with an amalgam of oligomers [111]. Though earlier reports suggested that podocin forms homo-oligomers the specificity of oligomerization was not known [50, 52, 78, 105]. Even though our study revealed truncated podocin forms predominantly a 16-mer, we yet to demonstrate whether full-length podocin also assembles as a 16-mer which is a major limitation of our study. We were unsuccessful in purifying the full-length podocin that could be due to several reasons including podocin may not be stable outside its native environment, due to the presence of IURs, or interference of the transmembrane segment with protein expression [90]. We, therefore, performed our studies with truncated podocin (126-350 residues) that encompasses the PHB domain and the oligomerization sites. The

regions '283-313' and '332-348' of podocin facilitate intermolecular interactions [79]. Huber et. al, reported that podocin (R138X) was unable to form oligomers suggesting the importance of this region in podocin-podocin interactions [105]. We included the 126-164 residues upstream of the PHB domain as it was reported that the high-affinity binding among homo-oligomers requires larger parts of the C-terminal region besides the PHB domain [105]. Accumulating evidence suggests that the 126-350 region of podocin is crucial for homo-oligomerization, it can be expected that mutations in these regions likely distort the intermolecular interactions. Mutations that cause NS could distort the innate ability of podocin to form oligomers and compromise its ability to act as a scaffolding molecule. Together these molecular deformations at the level of podocin oligomeric assembly may lead to altered permselectivity of SD and manifest in significant proteinuria. As it was shown that podocin interacts with several SD proteins, it is less likely a single podocin interacts with all proteins whereas different monomer of podocin 16-mer may interact with the neighboring proteins in the SD complex. It should be noted that the podocin domain predominantly forms a 16-mer, and to a lesser extent, 25-mer to 13-mer are also observed. These results are further justified by the hydrodynamic radius and the polydispersity index (PDI) measured by the DLS. The reported PDI for the podocin domain falls in the IUR range value (0.08 – 0.5) suggesting that although it has one major oligomer species by volume, the presence of multiple species could be possible [116]. We were unable to obtain crystals of the podocin domain for performing X-ray diffraction possibly due to its dynamic oligomeric nature.

Far UV-CD spectra revealed that the podocin domain consists of a considerable content of α -helices and β -sheets. We could not record far UV-CD spectra beyond 195nm possibly because of the presence of buffer components like β -mercaptoethanol that manifested in increased HT voltage. However, β -mercaptoethanol in the buffer helped in solubilizing the podocin domain and contributed towards the protein stability. It is

noteworthy that crystal structures of stomatin protein are devoid of disulfide linkages [113, 117]. In a study, Huber et.al reported that Cys126 and Cys160 residues of mouse podocin undergo palmitoylation and participate in membrane insertion [78]. Similarly, the Cys158 in the human podocin domain is also expected to be palmitoylated instead of forming disulfide linkages. This suggests that the disulfide bridges between cysteines either do not exist or are not contributing to protein stability.

While assessing the effect of temperature on the protein by probing its intrinsic tryptophan fluorescence and secondary structure content via measuring optical rotation by using the CD, the baseline was observed only in the case of the CD spectrum beyond 80°C. A significant linear decrease in the fluorescence intensity at 335 nm with increasing temperature may be due to the partial exposure of the lone Trp256 and other aromatic residues as observed in the predicted model for podocin and the crystal structure of stomatin protein (3BK6) (Fig. S1) [38, 113]. Also, there was no bathochromic shift observed, which ascertains the fact that the oligomer might not have dissociated completely. Similarly, far UV-CD spectra at 222nm were used to monitor secondary structural changes. A small yet noticeable change in the MRE value and the shape of the spectra implicates that most of the secondary structure was retained, however, a slight rearrangement of structure could be possible, owing to the increased hydrophobic effect. With the increasing temperature, the prominent double minima smoothen out indicating that there might be a slight increase in the β -sheet content. The stable MRE value observed beyond 80°C might suggest some residual local ordering that still prevails at high temperatures.

Calorimetric analysis of the oligomeric species revealed multiple transition points for the podocin domain. At 20°C the protein was shown to attain an amalgam of oligomeric states and a steady increase in temperature may promote dissociation of these different oligomers via three different transition temperatures since each oligomeric state may not

have the same transition point. Nevertheless, when the temperature was further increased significant exothermic transition state was observed. We assume the exothermic transition state may be due to the association of hydrophobic cores of the constituent oligomers. It is well known that lower temperatures do not favor hydrophobic interactions whereas, an increase in temperature to a certain degree promotes hydrophobic interactions among the protein molecules [118]. Subsequently, when the temperature is further increased, we speculate that the oligomer may invariably disassociate into lower oligomeric species. It is noteworthy that the podocin domain did not re-trace the path to its initial state when the protein was cooled. This peculiar behavior of the podocin domain starkly correlates with the oligomeric propensities as observed in the stomatin proteins [111, 112].

We performed the above-discussed experiments at 5°C and found that 16-mer seems to be the major conformer at both lower and ambient temperature (Ancillary data Fig.25-28). Interestingly in the calorimetric analysis, multiple transition points were not observed which suggest the presence of 16-mer species as the predominant conformer at lower temperatures and that was further confirmed by SEC analysis (Fig.27-28).

4.5. Ancillary data

Figure 25: Podocin domain majorly adopts 16-mer homo-oligomer: SEC of the purified podocin domain at 50C showing different podocin domain oligomeric conformations. A pre-packed HiLoad 16/600 Superdex 200pg (GE Healthcare) column was equilibrated with 10 mM potassium phosphate, 150 mM NaCl, and 2 mM β -mercaptoethanol (pH 8.0). 1ml of protein at 9 μ M concentration was applied to the column at a constant flow rate of 0.5ml/min. The elution volume of the protein was detected by monitoring the UV absorbance at 280 nm.

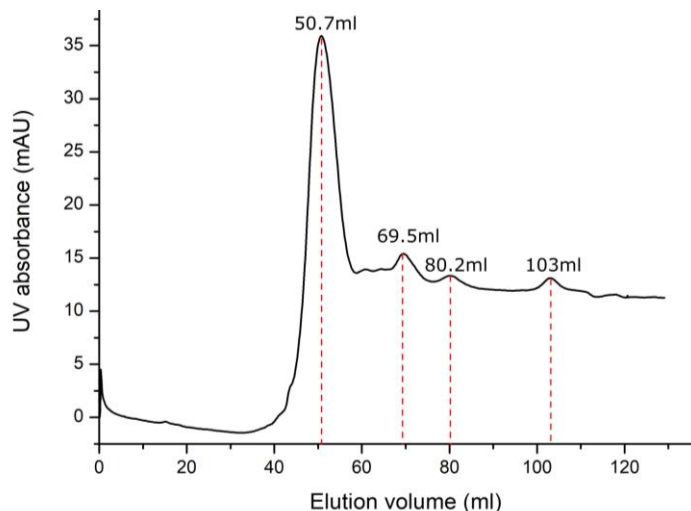


Figure 26: Podocin domain exhibits secondary structural elements: Far-UV CD spectrum of Podocin domain at 50C, in 10mM potassium phosphate buffer supplemented with 150mM NaCl and 2mM β -mercaptoethanol (pH 8.0).

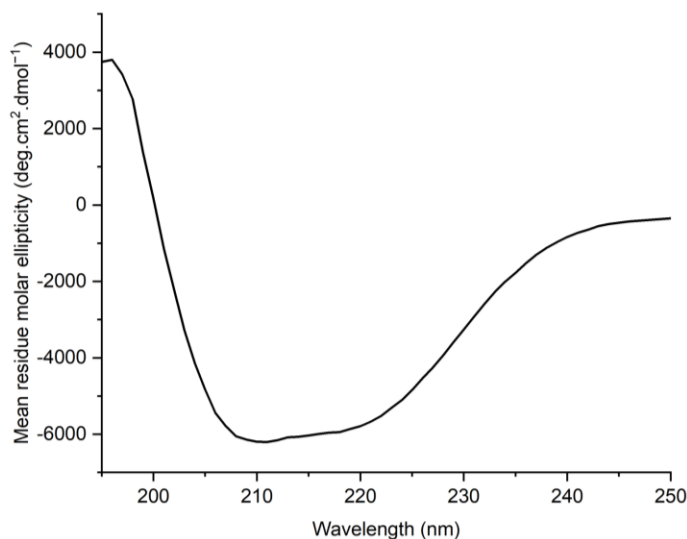


Figure 27: Effect of temperature on podocin domain: A) Intrinsic tryptophan fluorescence for podocin domain was measured as a function of temperature (5oC to 95oC) by exciting the protein at 287 nm. B) Peak emission of the podocin domain at 335 nm is plotted over a temperature range of 5oC to 95oC. C) FarUV CD spectra (250 – 200 nm) of the podocin domain was acquired as a function of temperature (5oC to 95oC). D) The MRE values of the podocin domain at 222nm is plotted as a function of temperature (range: 5oC to 95o) to monitor the changes in the protein structure.

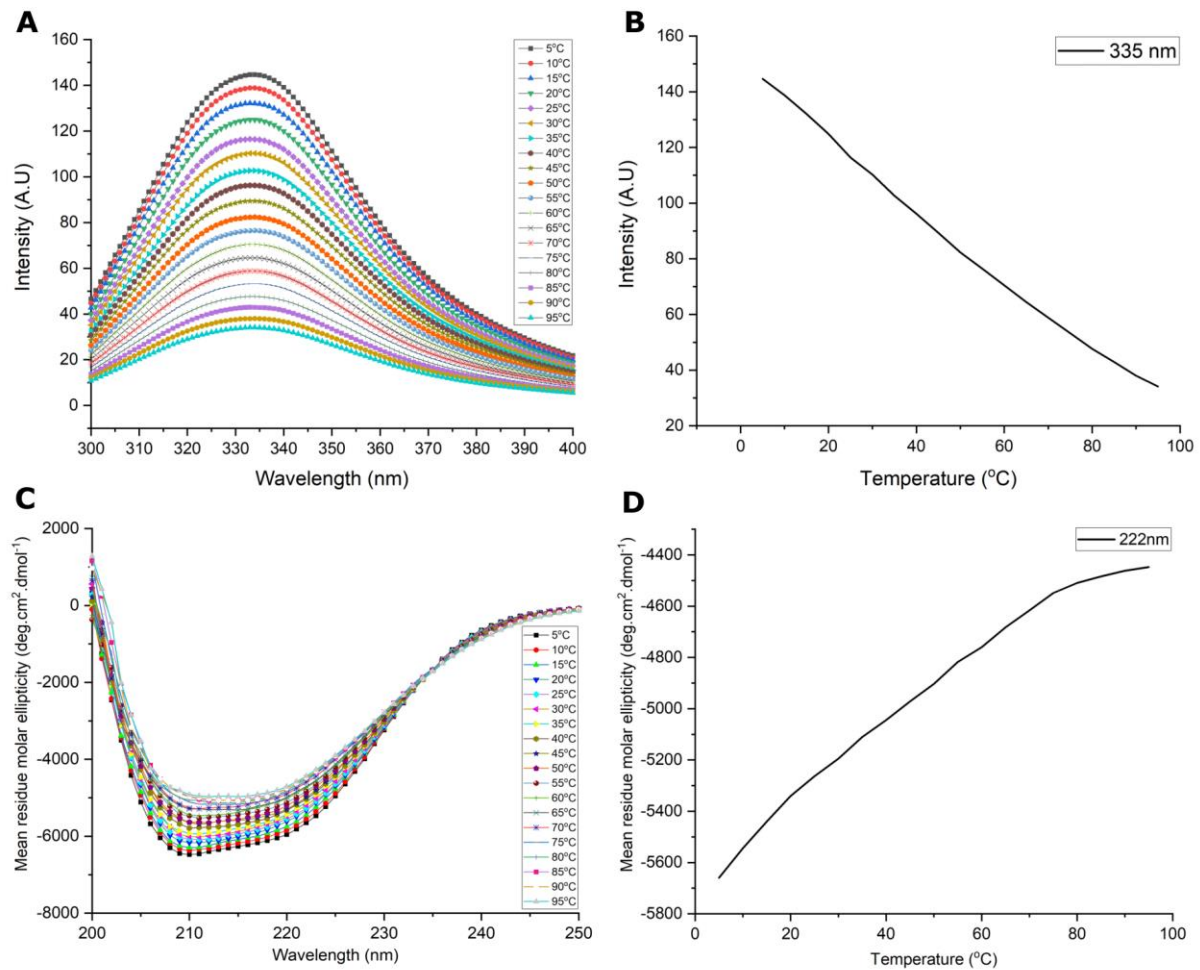
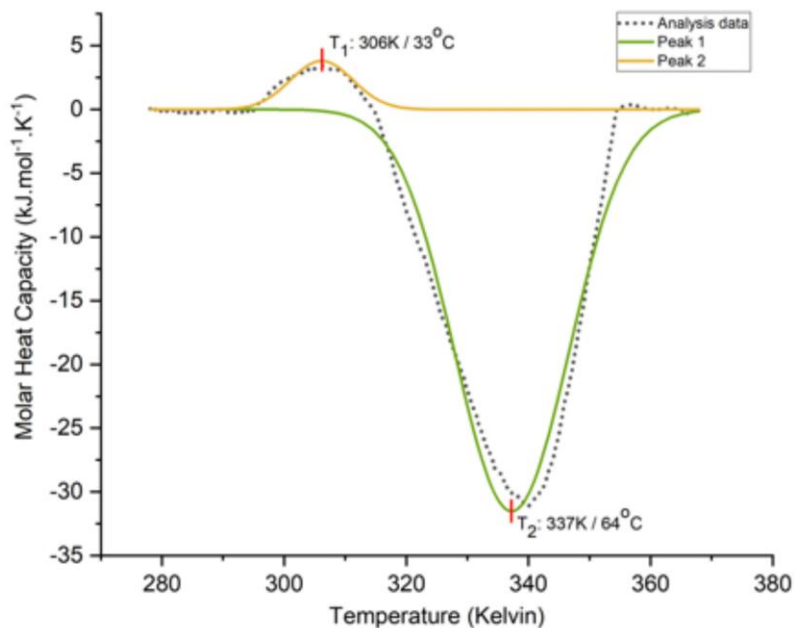


Figure 28: DSC analysis of podocin domain: The plots represent the endothermic and exothermic transitions of the podocin domain over a temperature range from 50C to 95oC. The initial DSC artifact near 273K in the DSC data is not represented in the plot. The analysis data after baseline subtraction is represented as a dotted line in the plots, whereas the transition peaks in the data are represented as solid color lines.



Chapter 5: Summary

The molecular evolution and the structural details of the SD proteins are still unclear and how they help in maintaining the structure of the SD is a question that many researchers are currently pursuing. During my doctoral work, we have attempted to specifically address these questions.

1. Molecular evolution of Slit-diaphragm proteins.

The evolutionary studies provide novel insights into the evolutionary relationships between the vertebrate SD proteins and the invertebrate orthologs. We propose that the SD proteins may have evolved from the orthologs sequence identified in the invertebrate phyla. We also show that the unique domains present in the SD proteins are highly conserved. Further, our study shows that IURs are highly conserved motifs among the vertebrate and the invertebrate sequence which further adds evidence for the role of IURs in the SD complex formation. Though, studies reveal these four proteins interact and maintain the SD architecture, the precise stoichiometry of these proteins yet to be unraveled. Also, how the mutation alters the structure and functional relationship needs to be explained.

2. Intrinsically unstructured regions mediate the assembly of Slit-diaphragm proteins.

Characterizing the structural details of the SD proteins is still an ongoing and challenging project for many researchers due to difficulties in obtaining stable constructs. Therefore, to gain preliminary structural insights of the key SD proteins and how they can organize into a complex, we have built models for nephrin, CD2AP, podocin, and TRPC6 using a combination of *in-silico* methods. Our study suggests that IURs and IUR-BD play a major role in mediating the macromolecular assembly of nephrin, CD2AP, podocin, and TRPC6. Further, we suggest that IURs and IUR-BD may also help in the homodimer interactions between adjacent nephrin molecules imparting a zipper-like structure to the

SD. We speculate that mutations in any of these proteins although subtly alters the protein structure, drastically influence the mutant protein's ability to form a proper complex with its WT neighboring proteins.

3. Structural features and oligomeric nature of human podocin domain

Next, we chose to express and purify podocin in an attempt to understand its stability and thermodynamic properties. Our study, to the best of our knowledge, is the first to report the cloning, expression, and purification of the human podocin domain (126-350). We showed that the podocin domain (126-350) is majorly a 16-mer homo-oligomer. However, to a lesser extent, the protein is capable of associating with other oligomeric states. CD and FL analysis of the protein indicated that the podocin domain in isolation attains considerable secondary structure and tertiary packing. However, the significance of the podocin oligomerization in the formation of the large macromolecular complex and how a mutation in a podocin monomer will affect its oligomerization status need to be addressed. Further structural characterization of podocin and other slit diaphragm proteins is greatly warranted to understand the mechanism of pathogenesis of the NS

References

References:

1. Anil Kumar, P., et al., *Molecular and cellular events mediating glomerular podocyte dysfunction and depletion in diabetes mellitus*. Front Endocrinol (Lausanne), 2014. **5**: p. 151.
2. Lindsay M. Biga, S.D., Amy Harwell, Robin Hopkins, Joel Kaufmann, Mike LeMaster, Philip Matern, Katie Morrison-Graham, Devon Quick, and Jon Runyeon, *Anatomy & Physiology*. first ed. 2020: OpenStax/Oregon State University.
3. Miesen, L., E. Steenbergen, and B. Smeets, *Parietal cells-new perspectives in glomerular disease*. Cell Tissue Res, 2017. **369**(1): p. 237-244.
4. Mukhi, D., et al., *Novel Actions of Growth Hormone in Podocytes: Implications for Diabetic Nephropathy*. Front Med (Lausanne), 2017. **4**: p. 102.
5. Miner, J.H., *Glomerular basement membrane composition and the filtration barrier*. Pediatr Nephrol, 2011. **26**(9): p. 1413-7.
6. Quaggin, S.E. and J.A. Kreidberg, *Development of the renal glomerulus: good neighbors and good fences*. Development, 2008. **135**(4): p. 609-20.
7. Drenckhahn, D. and R.P. Franke, *Ultrastructural organization of contractile and cytoskeletal proteins in glomerular podocytes of chicken, rat, and man*. Lab Invest, 1988. **59**(5): p. 673-82.
8. Andrews, P.M. and S.B. Bates, *Filamentous actin bundles in the kidney*. Anat Rec, 1984. **210**(1): p. 1-9.
9. Rodewald, R. and M.J. Karnovsky, *Porous substructure of the glomerular slit diaphragm in the rat and mouse*. J Cell Biol, 1974. **60**(2): p. 423-33.
10. Karnovsky, M.J. and G.B. Ryan, *Substructure of the glomerular slit diaphragm in freeze-fractured normal rat kidney*. J Cell Biol, 1975. **65**(1): p. 233-6.
11. Yamada, E., *The fine structure of the renal glomerulus of the mouse*. J Biophys Biochem Cytol, 1955. **1**(6): p. 551-66.
12. Garg, P., *A Review of Podocyte Biology*. Am J Nephrol, 2018. **47 Suppl 1**: p. 3-13.
13. Kumar, P.A., et al., *Molecular and cellular events mediating glomerular podocyte dysfunction and depletion in diabetes mellitus*. Frontiers in Endocrinology, 2014. **5**.
14. Patrakka, J. and K. Tryggvason, *Nephrin--a unique structural and signaling protein of the kidney filter*. Trends Mol Med, 2007. **13**(9): p. 396-403.
15. Tryggvason, K. and J. Wartiovaara, *How does the kidney filter plasma?* Physiology (Bethesda), 2005. **20**: p. 96-101.
16. Menon, M.C., P.Y. Chuang, and C.J. He, *The glomerular filtration barrier: components and crosstalk*. Int J Nephrol, 2012. **2012**: p. 749010.
17. Mundel, P., et al., *Rearrangements of the cytoskeleton and cell contacts induce process formation during differentiation of conditionally immortalized mouse podocyte cell lines*. Exp Cell Res, 1997. **236**(1): p. 248-58.
18. Abu Hamad, R., et al., *Response of Renal Podocytes to Excessive Hydrostatic Pressure: a Pathophysiologic Cascade in a Malignant Hypertension Model*. Kidney Blood Press Res, 2017. **42**(6): p. 1104-1118.
19. Robson, A.M., et al., *Normal glomerular permeability and its modification by minimal change nephrotic syndrome*. J Clin Invest, 1974. **54**(5): p. 1190-9.
20. Wiggins, R.C., *The spectrum of podocytopathies: a unifying view of glomerular diseases*. Kidney Int, 2007. **71**(12): p. 1205-14.
21. Mulukala, S.K.N., et al., *Structural features and oligomeric nature of human podocin domain*. Biochem Biophys Rep, 2020. **23**: p. 100774.

22. Srivastava, R. and A. Bagga, *Pediatric Nephrology*. Sixth edition ed. 2016: Jaypee Brothers, Medical Publishers Pvt. Limited. 618.
23. Reeves, W., J.P. Caulfield, and M.G. Farquhar, *Differentiation of epithelial foot processes and filtration slits: sequential appearance of occluding junctions, epithelial polyanion, and slit membranes in developing glomeruli*. *Lab Invest*, 1978. **39**(2): p. 90-100.
24. Grahammer, F., C. Schell, and T.B. Huber, *The podocyte slit diaphragm--from a thin grey line to a complex signalling hub*. *Nat Rev Nephrol*, 2013. **9**(10): p. 587-98.
25. Ruotsalainen, V., et al., *Nephrin is specifically located at the slit diaphragm of glomerular podocytes*. *Proc Natl Acad Sci U S A*, 1999. **96**(14): p. 7962-7.
26. Benzing, T., *Signaling at the slit diaphragm*. *J Am Soc Nephrol*, 2004. **15**(6): p. 1382-91.
27. Pierchala, B.A., M.R. Munoz, and C.C. Tsui, *Proteomic analysis of the slit diaphragm complex: CLIC5 is a protein critical for podocyte morphology and function*. *Kidney Int*, 2010. **78**(9): p. 868-82.
28. Zhuang, S., et al., *Sns and Kirre, the Drosophila orthologs of Nephrin and Neph1, direct adhesion, fusion and formation of a slit diaphragm-like structure in insect nephrocytes*. *Development*, 2009. **136**(14): p. 2335-44.
29. Weavers, H., et al., *The insect nephrocyte is a podocyte-like cell with a filtration slit diaphragm*. *Nature*, 2009. **457**(7227): p. 322-6.
30. Reiser, J., et al., *The glomerular slit diaphragm is a modified adherens junction*. *J Am Soc Nephrol*, 2000. **11**(1): p. 1-8.
31. Schnabel, E., J.M. Anderson, and M.G. Farquhar, *The tight junction protein ZO-1 is concentrated along slit diaphragms of the glomerular epithelium*. *J Cell Biol*, 1990. **111**(3): p. 1255-63.
32. Shono, A., et al., *Podocin participates in the assembly of tight junctions between foot processes in nephrotic podocytes*. *J Am Soc Nephrol*, 2007. **18**(9): p. 2525-33.
33. Kurihara, H., et al., *The altered glomerular filtration slits seen in puromycin aminonucleoside nephrosis and protamine sulfate-treated rats contain the tight junction protein ZO-1*. *Am J Pathol*, 1992. **141**(4): p. 805-16.
34. Hirabayashi, S., et al., *MAGI-1 is a component of the glomerular slit diaphragm that is tightly associated with nephrin*. *Lab Invest*, 2005. **85**(12): p. 1528-43.
35. Lehtonen, S., et al., *Cell junction-associated proteins IQGAP1, MAGI-2, CASK, spectrins, and alpha-actinin are components of the nephrin multiprotein complex*. *Proc Natl Acad Sci U S A*, 2005. **102**(28): p. 9814-9.
36. Nagai, M., et al., *Coxsackievirus and adenovirus receptor, a tight junction membrane protein, is expressed in glomerular podocytes in the kidney*. *Lab Invest*, 2003. **83**(6): p. 901-11.
37. Kestila, M., et al., *Positionally cloned gene for a novel glomerular protein--nephrin--is mutated in congenital nephrotic syndrome*. *Mol Cell*, 1998. **1**(4): p. 575-82.
38. Mulukala Narasimha, S.K., et al., *Intrinsically disordered regions mediate macromolecular assembly of the Slit diaphragm proteins associated with Nephrotic syndrome*. *Molecular Simulation*, 2019. **45**(8): p. 603-613.
39. Wagner, N., et al., *The podocyte protein nephrin is required for cardiac vessel formation*. *Hum Mol Genet*, 2011. **20**(11): p. 2182-94.
40. Astrom, E., et al., *Nephrin in human lymphoid tissues*. *Cell Mol Life Sci*, 2006. **63**(4): p. 498-504.
41. Fornoni, A., et al., *Nephrin is expressed on the surface of insulin vesicles and facilitates glucose-stimulated insulin release*. *Diabetes*, 2010. **59**(1): p. 190-9.

42. Zanone, M.M., et al., *Expression of nephrin by human pancreatic islet endothelial cells*. Diabetologia, 2005. **48**(9): p. 1789-97.
43. Sellin, L., et al., *NEPH1 defines a novel family of podocin interacting proteins*. FASEB J, 2003. **17**(1): p. 115-7.
44. Liu, G., et al., *Neph1 and nephrin interaction in the slit diaphragm is an important determinant of glomerular permeability*. The Journal of clinical investigation, 2003. **112**(2): p. 209-221.
45. Khoshnoodi, J., et al., *Nephrin promotes cell-cell adhesion through homophilic interactions*. Am J Pathol, 2003. **163**(6): p. 2337-46.
46. Wartiovaara, J., et al., *Nephrin strands contribute to a porous slit diaphragm scaffold as revealed by electron tomography*. J Clin Invest, 2004. **114**(10): p. 1475-83.
47. Hartleben, B., et al., *Neph-Nephrin proteins bind the Par3-Par6-atypical protein kinase C (aPKC) complex to regulate podocyte cell polarity*. J Biol Chem, 2008. **283**(34): p. 23033-8.
48. Putaala, H., et al., *Primary structure of mouse and rat nephrin cDNA and structure and expression of the mouse gene*. J Am Soc Nephrol, 2000. **11**(6): p. 991-1001.
49. Verma, R., et al., *Fyn binds to and phosphorylates the kidney slit diaphragm component Nephrin*. J Biol Chem, 2003. **278**(23): p. 20716-23.
50. Huber, T.B., et al., *Interaction with podocin facilitates nephrin signaling*. J Biol Chem, 2001. **276**(45): p. 41543-6.
51. Huber, T.B., et al., *Nephrin and CD2AP associate with phosphoinositide 3-OH kinase and stimulate AKT-dependent signaling*. Mol Cell Biol, 2003. **23**(14): p. 4917-28.
52. Schwarz, K., et al., *Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin*. J Clin Invest, 2001. **108**(11): p. 1621-9.
53. Dryer, S.E. and J. Reiser, *TRPC6 channels and their binding partners in podocytes: role in glomerular filtration and pathophysiology*. Am J Physiol Renal Physiol, 2010. **299**(4): p. F689-701.
54. Yu, C.C., et al., *Lupus-like kidney disease in mice deficient in the Src family tyrosine kinases Lyn and Fyn*. Curr Biol, 2001. **11**(1): p. 34-8.
55. Dustin, M.L., et al., *A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts*. Cell, 1998. **94**(5): p. 667-77.
56. Ma, Y.T., et al., *CD2AP is indispensable to multistep cytotoxic process by NK cells*. Molecular Immunology, 2010. **47**(5): p. 1074-1082.
57. Lehtonen, S., F. Zhao, and E. Lehtonen, *CD2-associated protein directly interacts with the actin cytoskeleton*. Am J Physiol Renal Physiol, 2002. **283**(4): p. F734-43.
58. Shih, N.Y., et al., *CD2AP localizes to the slit diaphragm and binds to nephrin via a novel C-terminal domain*. Am J Pathol, 2001. **159**(6): p. 2303-8.
59. Shih, N.Y., et al., *Congenital nephrotic syndrome in mice lacking CD2-associated protein*. Science, 1999. **286**(5438): p. 312-5.
60. Li, C., et al., *CD2AP is expressed with nephrin in developing podocytes and is found widely in mature kidney and elsewhere*. Am J Physiol Renal Physiol, 2000. **279**(4): p. F785-92.
61. Ortega Roldan, J.L., et al., *The high resolution NMR structure of the third SH3 domain of CD2AP*. J Biomol NMR, 2007. **39**(4): p. 331-6.
62. Adair, B.D., et al., *Structure of the kidney slit diaphragm adapter protein CD2-associated protein as determined with electron microscopy*. J Am Soc Nephrol, 2014. **25**(7): p. 1465-73.

63. Abramowitz, J., E. Yildirim, and L. Birnbaumer, *The TRPC Family of Ion Channels: Relation to the TRP Superfamily and Role in Receptor- and Store-Operated Calcium Entry*, in *TRP Ion Channel Function in Sensory Transduction and Cellular Signaling Cascades*, W.B. Liedtke and S. Heller, Editors. 2007: Boca Raton (FL).
64. Wilson, C. and S.E. Dryer, *A mutation in TRPC6 channels abolishes their activation by hypoosmotic stretch but does not affect activation by diacylglycerol or G protein signaling cascades*. *Am J Physiol Renal Physiol*, 2014. **306**(9): p. F1018-25.
65. Goel, M., et al., *Identification and localization of TRPC channels in the rat kidney*. *Am J Physiol Renal Physiol*, 2006. **290**(5): p. F1241-52.
66. Ilatovskaya, D.V., et al., *Angiotensin II has acute effects on TRPC6 channels in podocytes of freshly isolated glomeruli*. *Kidney Int*, 2014. **86**(3): p. 506-14.
67. Azumaya, C.M., et al., *Cryo-EM structure of the cytoplasmic domain of murine transient receptor potential cation channel subfamily C member 6 (TRPC6)*. *J Biol Chem*, 2018. **293**(26): p. 10381-10391.
68. Zhu, M.X., *Multiple roles of calmodulin and other Ca(2+)-binding proteins in the functional regulation of TRP channels*. *Pflugers Arch*, 2005. **451**(1): p. 105-15.
69. Ilatovskaya, D.V. and A. Staruschenko, *TRPC6 channel as an emerging determinant of the podocyte injury susceptibility in kidney diseases*. *Am J Physiol Renal Physiol*, 2015. **309**(5): p. F393-7.
70. Reiser, J., et al., *TRPC6 is a glomerular slit diaphragm-associated channel required for normal renal function*. *Nat Genet*, 2005. **37**(7): p. 739-44.
71. Gigante, M., et al., *TRPC6 mutations in children with steroid-resistant nephrotic syndrome and atypical phenotype*. *Clin J Am Soc Nephrol*, 2011. **6**(7): p. 1626-34.
72. Winn, M.P., et al., *A mutation in the TRPC6 cation channel causes familial focal segmental glomerulosclerosis*. *Science*, 2005. **308**(5729): p. 1801-4.
73. Moller, C.C., et al., *Induction of TRPC6 channel in acquired forms of proteinuric kidney disease*. *J Am Soc Nephrol*, 2007. **18**(1): p. 29-36.
74. Winn, M.P., et al., *Unexpected role of TRPC6 channel in familial nephrotic syndrome: does it have clinical implications?* *J Am Soc Nephrol*, 2006. **17**(2): p. 378-87.
75. Boute, N., et al., *NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome*. *Nat Genet*, 2000. **24**(4): p. 349-54.
76. Green, J.B. and J.P. Young, *Slipins: ancient origin, duplication and diversification of the stomatin protein family*. *BMC Evol Biol*, 2008. **8**: p. 44.
77. Snyers, L., E. Umlauf, and R. Prohaska, *Cysteine 29 is the major palmitoylation site on stomatin*. *FEBS Lett*, 1999. **449**(2-3): p. 101-4.
78. Huber, T.B., et al., *Podocin and MEC-2 bind cholesterol to regulate the activity of associated ion channels*. *Proc Natl Acad Sci U S A*, 2006. **103**(46): p. 17079-86.
79. Straner, P., et al., *C-terminal oligomerization of podocin mediates interallelic interactions*. *Biochim Biophys Acta Mol Basis Dis*, 2018. **1864**(7): p. 2448-2457.
80. Tory, K., et al., *Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome*. *Nat Genet*, 2014. **46**(3): p. 299-304.
81. Ruppert, E.E., *Evolutionary Origin of the Vertebrate Nephron*. *American Zoologist*, 1994. **34**(4): p. 542-553.
82. Weavers, H., et al., *The insect nephrocyte is a podocyte-like cell with a filtration slit diaphragm*. *Nature*, 2009. **457**(7227): p. 322-U5.
83. Dikic, I., *CIN85/CMS family of adaptor molecules*. *FEBS Lett*, 2002. **529**(1): p. 110-5.

84. Johnson, R.I., M.J. Seppa, and R.L. Cagan, *The Drosophila CD2AP/CIN85 orthologue Cindr regulates junctions and cytoskeleton dynamics during tissue patterning*. J Cell Biol, 2008. **180**(6): p. 1191-204.
85. Ojelade, S.A., et al., *cindr, the Drosophila Homolog of the CD2AP Alzheimer's Disease Risk Gene, Is Required for Synaptic Transmission and Proteostasis*. Cell Rep, 2019. **28**(7): p. 1799-1813 e5.
86. Dietrich, A. and T. Gudermann, *TRPC6: physiological function and pathophysiological relevance*. Handb Exp Pharmacol, 2014. **222**: p. 157-88.
87. Eid, S.R. and D.N. Cortright, *Transient receptor potential channels on sensory nerves*. Handb Exp Pharmacol, 2009(194): p. 261-81.
88. Chen, S., et al., *Calcium entry via TRPC6 mediates albumin overload-induced endoplasmic reticulum stress and apoptosis in podocytes*. Cell Calcium, 2011. **50**(6): p. 523-9.
89. Nakuluri, K., et al., *Cerebral ischemia induces TRPC6 via HIF1alpha/ZEB2 axis in the glomerular podocytes and contributes to proteinuria*. Sci Rep, 2019. **9**(1): p. 17897.
90. Mulukala, S.K., et al., *In silico Structural characterization of podocin and assessment of nephrotic syndrome-associated podocin mutants*. IUBMB Life, 2016. **68**(7): p. 578-88.
91. Linding, R., et al., *Protein disorder prediction: implications for structural proteomics*. Structure, 2003. **11**(11): p. 1453-9.
92. Palmen, T., et al., *Interaction of endogenous nephrin and CD2-associated protein in mouse epithelial M-1 cell line*. J Am Soc Nephrol, 2002. **13**(7): p. 1766-72.
93. Nugent, T. and D.T. Jones, *Transmembrane protein topology prediction using support vector machines*. BMC Bioinformatics, 2009. **10**(1): p. 159.
94. Jones, D.T., W.R. Taylor, and J.M. Thornton, *A model recognition approach to the prediction of all-helical membrane protein structure and topology*. Biochemistry, 1994. **33**(10): p. 3038-49.
95. Jones, D.T. and D. Cozzetto, *DISOPRED3: precise disordered region predictions with annotated protein-binding activity*. Bioinformatics, 2015. **31**(6): p. 857-63.
96. Yao, B., et al., *Solution structure of the second SH3 domain of human CMS and a newly identified binding site at the C-terminus of c-Cbl*. Biochim Biophys Acta, 2007. **1774**(1): p. 35-43.
97. Moncalian, G., et al., *Atypical polyproline recognition by the CMS N-terminal Src homology 3 domain*. J Biol Chem, 2006. **281**(50): p. 38845-53.
98. Tina, K.G., R. Bhadra, and N. Srinivasan, *PIC: Protein Interactions Calculator*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W473-6.
99. Duhovny, D., R. Nussinov, and H.J. Wolfson, *Efficient Unbound Docking of Rigid Molecules*, in *Algorithms in Bioinformatics: Second International Workshop, WABI 2002 Rome, Italy, September 17–21, 2002 Proceedings*, R. Guigó and D. Gusfield, Editors. 2002, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 185-200.
100. Schneidman-Duhovny, D., et al., *PatchDock and SymmDock: servers for rigid and symmetric docking*. Nucleic Acids Res, 2005. **33**(Web Server issue): p. W363-7.
101. Kirubakaran, P., et al., *In silico structural and functional analysis of the human TOPK protein by structure modeling and molecular dynamics studies*. J Mol Model, 2013. **19**(1): p. 407-19.
102. Kirubakaran, P., et al., *Homology modeling, molecular dynamics, and molecular docking studies of Trichomonas vaginalis carbamate kinase*. Medicinal Chemistry Research, 2012. **21**(8): p. 2105-2116.
103. Wright, P.E. and H.J. Dyson, *Intrinsically disordered proteins in cellular signalling and regulation*. Nat Rev Mol Cell Biol, 2015. **16**(1): p. 18-29.

104. Gigante, M., et al., *CD2AP mutations are associated with sporadic nephrotic syndrome and focal segmental glomerulosclerosis (FSGS)*. *Nephrol Dial Transplant*, 2009. **24**(6): p. 1858-64.
105. Huber, T.B., et al., *Molecular basis of the functional podocin-nephrin complex: mutations in the NPHS2 gene disrupt nephrin targeting to lipid raft microdomains*. *Hum Mol Genet*, 2003. **12**(24): p. 3397-405.
106. Shono, A., et al., *Predisposition to relapsing nephrotic syndrome by a nephrin mutation that interferes with assembly of functioning microdomains*. *Hum Mol Genet*, 2009. **18**(16): p. 2943-56.
107. Collins, M.O., et al., *Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder*. *Mol Cell Proteomics*, 2008. **7**(7): p. 1331-48.
108. Mayer, B.J., *SH3 domains: complexity in moderation*. *J Cell Sci*, 2001. **114**(Pt 7): p. 1253-63.
109. Stenson, P.D., et al., *Human Gene Mutation Database (HGMD): 2003 update*. *Hum Mutat*, 2003. **21**(6): p. 577-81.
110. Roselli, S., et al., *Podocin localizes in the kidney to the slit diaphragm area*. *Am J Pathol*, 2002. **160**(1): p. 131-9.
111. Kuwahara, Y., et al., *Unusual thermal disassembly of the SPFH domain oligomer from *Pyrococcus horikoshii**. *Biophys J*, 2009. **97**(7): p. 2034-43.
112. Snyers, L., E. Umlauf, and R. Prohaska, *Oligomeric nature of the integral membrane protein stomatin*. *J Biol Chem*, 1998. **273**(27): p. 17221-6.
113. Yokoyama, H., S. Fujii, and I. Matsui, *Crystal structure of a core domain of stomatin from *Pyrococcus horikoshii* illustrates a novel trimeric and coiled-coil fold*. *J Mol Biol*, 2008. **376**(3): p. 868-78.
114. Mulukala, N.S.K., et al., *Intrinsically disordered regions mediate macromolecular assembly of the Slit diaphragm proteins associated with Nephrotic syndrome*. *Molecular Simulation*, 2019. **45**(8): p. 603-613.
115. Eftink, M.R., *The use of fluorescence methods to monitor unfolding transitions in proteins*. *Biophys J*, 1994. **66**(2 Pt 1): p. 482-501.
116. Asmawati, et al., *Characteristics of cinnamaldehyde nanoemulsion prepared using APV-high pressure homogenizer and ultra turrax*. 2014. **1614**(1): p. 244-250.
117. Brand, J., et al., *A stomatin dimer modulates the activity of acid-sensing ion channels*. *EMBO J*, 2012. **31**(17): p. 3635-46.
118. van Dijk, E., A. Hoogeveen, and S. Abeln, *The hydrophobic temperature dependence of amino acids directly calculated from protein structures*. *PLoS Comput Biol*, 2015. **11**(5): p. e1004277.

Publications and Conferences

In Silico Structural Characterization of Podocin and Assessment of Nephrotic Syndrome-Associated Podocin Mutants

Sandeep Kumar
Narasimha Mulukala¹
Rajkishor Nishad¹
Lakshmi
Prasanna Kolligundla¹
Moin A. Saleem²
Nagu Prakash Prabhu³
Anil Kumar Pasupulati^{1*}

¹Department of Biochemistry, School of Life Sciences, University of Hyderabad, Hyderabad, Telangana, India

²Academic Renal Unit, School of Clinical Sciences, University of Bristol, Bristol, United Kingdom

³Department of Biotechnology and Bioinformatics, University of Hyderabad, Hyderabad, Telangana, India

Abstract

Nephrotic syndrome (NS) is manifested by hyperproteinuria, hypoalbuminemia, and edema. *NPHS2* that encodes podocin was found to have most mutations among the genes that are involved in the pathophysiology of NS. Podocin, an integral membrane protein belonging to stomatin family, is expressed exclusively in podocytes and is localized to slit-diaphragm (SD). Mutations in podocin are known to be associated with steroid-resistant NS and rapid progression to end-stage renal disease, thus signifying its role in maintaining SD integrity and podocyte function. The structural insights of podocin are not known, and the precise mechanism by which podocin con-

tributes to the architecture of SD is yet to be elucidated. In this study, we deduced a model for human podocin, discussed the details of transmembrane localization and intrinsically unstructured regions, and provide an understanding of how podocin interacts with other SD components. Intraprotein interactions were assessed in wild-type podocin and in some of its mutants that are associated with idiopathic NS. Mutations in podocin alter the innate intraprotein interactions affecting the native structure of podocin and its ability to form critical complex with subpodocyte proteins. © 2016 IUBMB Life, 68(7):578–588, 2016

Keywords: nephrotic syndrome; proteinuria; podocytes; podocin; slit diaphragm; molecular modeling

Introduction

The kidneys are vital organs that help to maintain body homeostasis by regulating blood pressure, acid–base, electrolyte, and water balance. Human kidney constitutes a million nephrons that collectively perform three key events including (a)

glomerular filtration of water and small molecules from renal plasma; (b) tubular reabsorption of glomerular filtrate; and (c) tubular secretion of metabolic waste products into the filtrate. Thus, glomerulus in concert with tubular region of the nephron tightly regulates the composition of glomerular filtrate and ensures almost protein-free ultrafiltrated urine. Glomerulus, where initiation of filtration occurs, contains a tuft of capillaries and several resident cell types that include mesangial cells, endothelial cells, and glomerular visceral epithelial cells, also known as podocytes.

Proteinuria is a hallmark of renal damage in several glomerular diseases due to the alterations in glomerular filtration barrier (GFB; refs. 1 and 2). The three components that constitute GFB include fenestrated glomerular endothelial cells, glomerular basement membrane (GBM), and glomerular visceral epithelial cells, known as podocytes. A wealth of literature highlighted that podocytes are critical for glomerular filtration (3,4). Podocytes are terminally differentiated epithelial cells

© 2016 International Union of Biochemistry and Molecular Biology
Volume 68, Number 7, July 2016, Pages 578–588

*Address correspondence to: Anil Kumar Pasupulati, Department of Biochemistry, School of Life Sciences, University of Hyderabad, Gachibowli, Hyderabad 500046, Andhra Pradesh, India.
Tel: +91-40-23134519. Fax: +91-40-23010120.
E-mail: pasupulati.anilkumar@gmail.com

Received 14 April 2016; Accepted 3 May 2016

DOI 10.1002/iub.1515

Published online 18 May 2016 in Wiley Online Library
(wileyonlinelibrary.com)

with large cell body consisting of organelles. These cells possess a main cell body, major foot process made of microtubules and intermediate filaments, and numerous secondary foot process that play a predominant role in attaching podocytes to the GBM. The secondary foot process of neighboring podocytes forms an adherent junction called slit-diaphragm (SD) that forms a sole contact between neighboring podocytes and ensures size, shape, and charge-selective permeability to the GFB (1). SD is made up of proteins like nephrin, podocin, CD2-associated protein (CD2AP), ZO-1, and P-cadherin. Mutations in the SD proteins alter podocyte architecture and permselectivity, which eventually abate the function of GFB culminating in proteinuria (5). Mutations in podocyte proteins could target the function of podocyte by affecting the structure of the SD, by perturbing the intricate podocyte cytoskeleton, by breaking cell-matrix interactions, and by blocking important signaling pathways. All these events manifest in the effacement of podocyte foot processes, disruption of GFB, and proteinuria (6).

Nephrotic syndrome (NS) is a nonspecific renal disorder characterized by heavy proteinuria, hypoalbuminemia, and edema. Idiopathic NS (INS) that occurs most frequently in children represents >90% of NS cases aged between 1 and 10 years and 50% of NS cases aged above 10 years (7). Electron microscopy of renal biopsies from INS subjects revealed diffuse foot process effacement, minimal change disease (MCD), and focal segmental glomerulosclerosis (FSGS). INS that occurs mainly in children is classified into steroid-sensitive NS (SSNS) and steroid-resistant NS (SRNS), depending on their response to corticosteroid therapy. SRNS has a poor prognosis and often leads to end-stage renal disease (ESRD). Mutations in more than 20 genes that predominantly encode podocyte proteins have been identified in monogenic forms of SRNS (8). Among those genes, *NPHS2*, encoding podocin, is the most frequently mutated gene responsible for up to 18% of total SRNS cases (9,10). Podocin consists of 383 amino acids and is expressed exclusively in podocytes, wherein it localizes to SD as an oligomer (10–12). Podocin belongs to stomatin family and is hypothesized to form hairpin-like intramembrane loop and intracellular C-terminus (12). It was reported that podocin interacts with other proteins of SD such as CD2AP and nephrin with its C-terminus (13). The interactions between podocin and other proteins are considered to be critical for the maintenance of intact SD architecture. Mutation in podocin led to the early onset of SRNS typically before 6 years and rapidly progresses to ESRD within a decade, indicating the instrumental role of podocin in podocyte biology and in maintaining the integrity of GFB.

Podocin mutants such as R138Q and V180M were observed to be restricted to endoplasmic reticulum and intracellular vesicles or form inclusion bodies thus preventing its localization to SD (14). However, it is yet to be known how mutations in podocin are associated with morphological changes ranging from MCD and FSGS that manifest in heavy proteinuria in SRNS. Further studies are warranted to under-

stand how mutations in podocin will distort the innate interactions with its sub SD proteins. In this study we attempted to generate 3D models for podocin and its mutants that are associated with SRNS to gain insights about how mutations affect the protein structure and contribute toward altered SD architecture and podocyte function.

Methodology

BLAST Analysis of Podocin Sequence and Sequence Alignment

Human podocin sequence (accession ID: Q9NP85) was obtained from ExPASy database. As crystal structure for podocin is not available, a reliable model was built based on sequence similarity search using protein BLAST (pBLAST) tool against predetermined protein structures deposited in Protein Data Bank (PDB). A comparative sequence alignment of the identified homologous sequences was done using the tools ClustalW and ClustalX.

Structure Prediction Using I-TASSER Server

Three-dimensional structure prediction for podocin sequence bits (1–164 and 362–383) that showed less (<25%) homology was performed using the I-TASSER web server (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>). I-TASSER, an *ab initio*-based protein structure prediction software works on sequence-to-structure-to-function pattern (15). The server generated five models, which are ranked based on the structure density of the SPICKER clustering. A confidence score (*C*-score) was calculated based on the statistical significance of the threading profile–profile alignment as well as structure convergence of the Monte Carlo simulations. This *C*-score denotes the quality of the models generated by the server.

Model Prediction and Stereochemistry Analysis of the Model

The predicted structure from I-TASSER server and the homolog sequences obtained from the BLAST search were taken as templates to generate a model for podocin. A multiple sequence-based modeling was done using Modeller 9.15 software that generates models by satisfying spatial restraints via automated comparative modeling (16). The multiple sequence alignment file in Modeller 9.15 was manually tweaked to match the query and template sequences. The generated model was then analyzed in detail for its protein structure stereochemistry using protein parameters analysis tool known as PROCHECK.

Predicting Intrinsically Unstructured, Transmembrane Regions and Assessing Intraprotein Contacts

Because of the anomalous N- and C-terminal random coils observed in the generated model, a secondary structure analysis was done using Position-Specific Iterated (PSI)-PRED server (<http://bioinf.cs.ucl.ac.uk/psipred/>). This server uses a two-stage neural network to predict the secondary structure using data generated by PSI-BLAST (17). Furthermore, the server also extrapolated intrinsically unstructured regions (IURs;

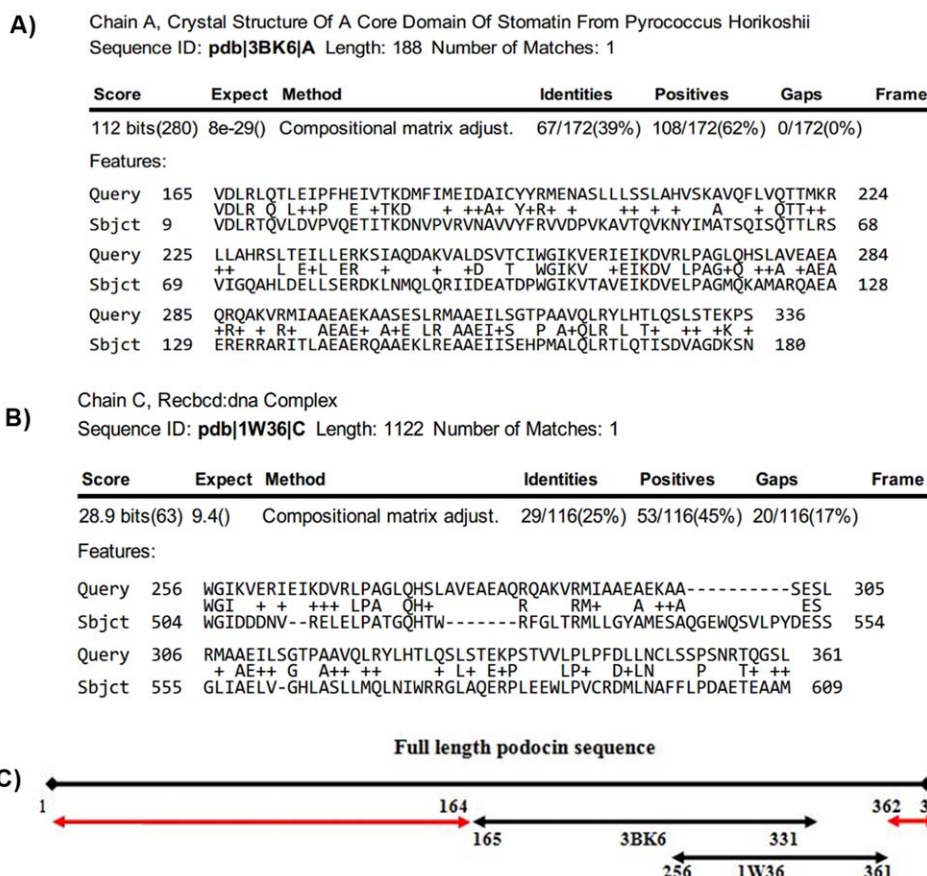


FIG 1

Templates used for generating full-length podocin model: (A) BLAST alignment of podocin with chain A, crystal structure of a core domain of stomatin from *Pyrococcus horikoshii* (PDB ID: 3BK6); and (B) BLAST alignment of podocin with chain C, RecBCD:DNA complex (*Escherichia coli*; PDB ID: 1W36). C: Templates used to generate full-length 3D model for podocin. The arrows in red indicate the parts of sequence for which 3D models were predicted using I-TASSER.

DISOPRED3), transmembrane (TM) topology prediction (MEMSAT3), and TM helix prediction (MEMSAT3VM; refs. 18–20). However, to gain further evidence on IURs in podocin, the sequence was trained exhaustively in various Critical Assessment of Protein Structure Prediction (CASP) validated servers such as Genesilico metadisorder service (Metadisorder, MetadisorderMD, MetadisorderMD2, Metadisorder3D; ref. 21), PONDR (22), DisProt (23), IUPred (24), Dis-EMBL (disorder by loops/coils definition, disorder by HOT-loops, Remark-465; ref. 25), SPINE-D (26), MFDp (DISOclust, DISOpred, IUPredL, and IUPredS; ref. 27), and PredictProtein (28). The PSI-PRED server and the PredictProtein server extrapolated TM helices in addition to predicting IURs (29). The intraprotein interactions such as main chain–main chain, main chain–side chain, and side chain–side chain, and hydrophobic interactions for a given set of 3D coordinates of the protein were assessed using protein interactions calculator (PIC) server (30).

Building Models for Podocin Mutants

We have selected six mutants of podocin (R3G, P89T, R322Q, R322P, H325Y, and V370G) that were shown to associate with SRNS from HGMD database (31,32). The criterion for selecting

these particular mutants was that they occur in predicted IURs as described above. Models for these mutants were built using Modeller 9.15 software considering wild-type podocin as a template. Furthermore, stereochemistry of mutant models was analyzed using PROCHECK software.

Results

Sequence Alignment and Homology Modeling of Podocin

Podocin shows highest homology with stomatin from *H. sapiens* (47% identity and 67% similarity) and with mechanosensory protein-2 from *Caenorhabditis elegans* (44% identity and 65% similarity; ref. 10). Considering this homology as a base, a pBLAST search was performed with human podocin sequence against PDB database. The search yielded two prominent templates: 1) chain A, crystal structure of a core domain of stomatin from *Pyrococcus horikoshii* (PDB ID: 3BK6; ref. 33), and 2) chain C, RecBCD:DNA complex from *Escherichia coli* (PDB ID: 1W36; ref. 34). These two templates covered the podocin sequence from 165 to 361 amino acids.

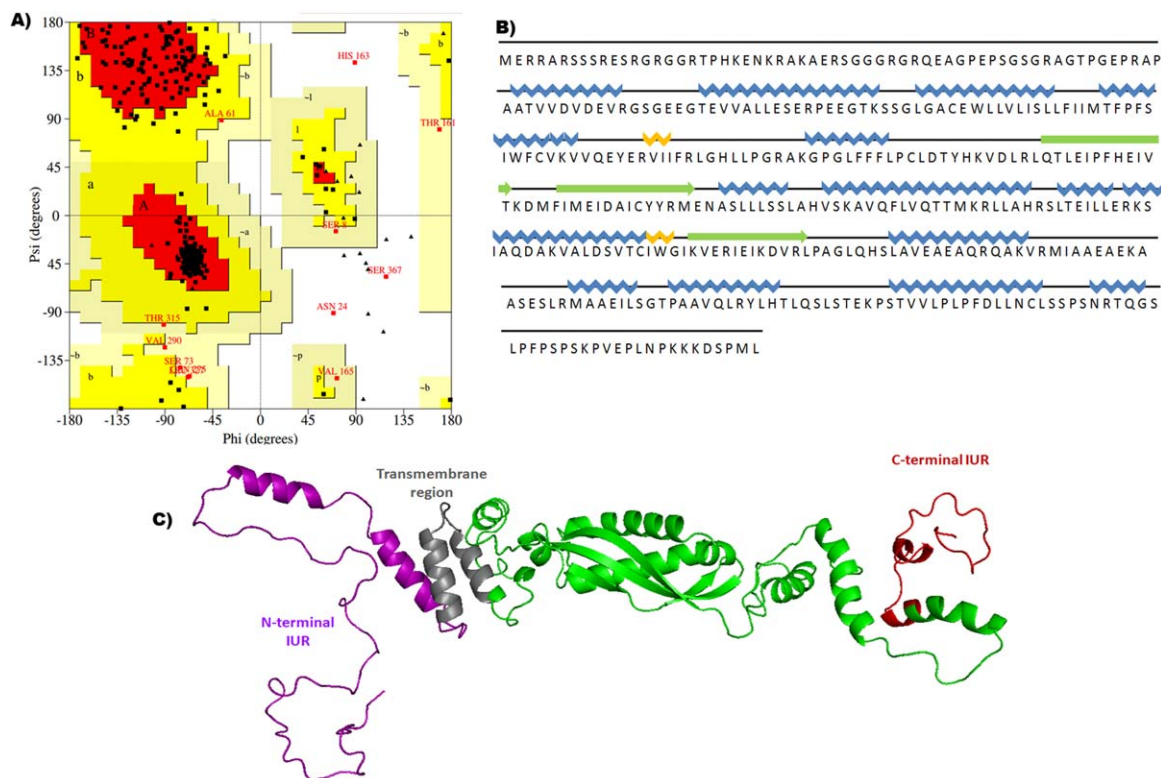


FIG 2

Wild-type podocin model: (A) Ramachandran plot for the wild-type podocin; (B) secondary structure of the generated podocin model, (C) modeled podocin structure based on a multiple sequence-based homology modeling using templates from pBLAST search and I-TASSER predictions.

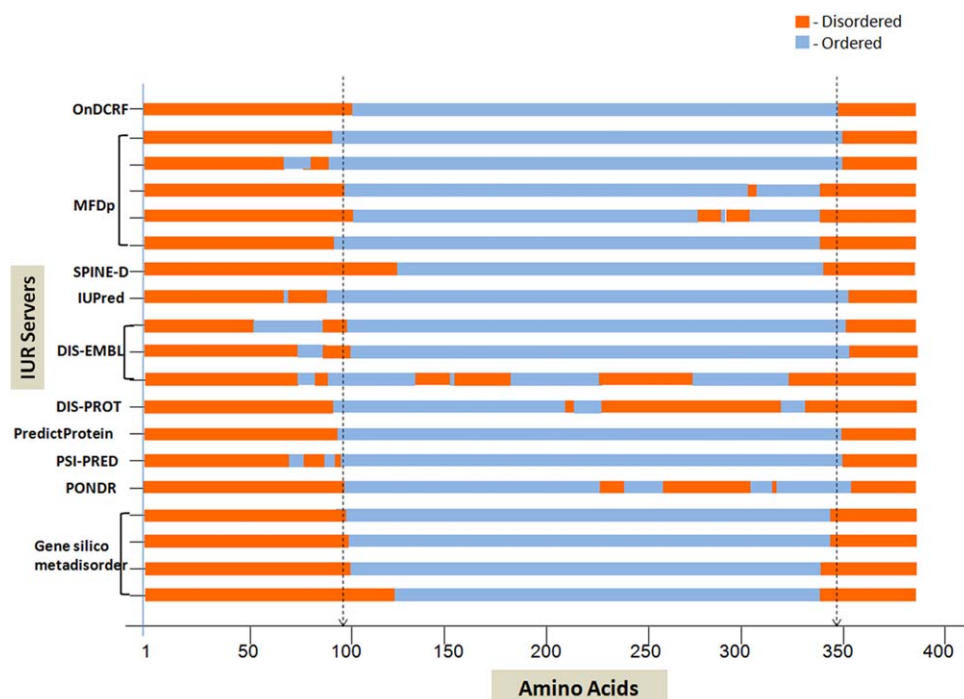


FIG 3

Localization of intrinsically unstructured regions (IURs) in podocin. A consensus data from different IUR prediction servers identified IURs in podocin spanning 1–98 and 347–383 residues at the N- and C-terminus, respectively.

TABLE 1

Prediction of Intrinsically Unstructured Regions (IURs) in Podocin by Critical Assessment of Protein Structure Prediction (CASP) Validated Servers

Server	Residues	Remarks
Genesilico metadisorder server	1–121, 337–383	Metadisorder 3D
	1–100, 338–383	Metadisorder MD2
	1–98, 347–383	Metadisorder MD
	1–97, 347–383	Metadisorder
PONDR	1–96, 230–242, 268–315, 322–325, 357–383	–
PSI-Pred	1–7 (BD), 8–61, 63–69 (BD), 77–89 (BD), 90–94, 353–356 (BD), 357–362, 363 (BD), 364–366, 367–377 (BD), 378–383	BD: Binding domains
PredictProtein	1–94, 351–383	–
DIS-PROT	1–94, 216–226, 277–323, 336–383	–
DIS-EMBL	1–78, 86–99, 141–163, 168–183, 225–276, 327–383	Disordered by loops/coils definition
	1–76, 84–100, 359–383	Disordered by HOT-loops
	1–59, 86–96, 355–383	Remark: 465
IUPred	1–70, 78–88, 355–383	–
SPINE-D	1–125, 344–383	–
MFDp	1–96, 340–383	MFDp
	1–101, 279–290, 296, 298–316, 342–383	DISOclust
	1–98, 308–310, 342–383	DISOPred
	1–70, 78–88, 350–383	IUPredL
	1–88, 352–383	IUPredS
OnDCRF	1–100, 347–383	Predicting ordered and disordered regions using conditional random fields

Although two homologous templates were identified from the pBLAST search, the N- and C-terminal regions (1–164 and 362–383 amino acids) of podocin showed <25% homology to build a reliable full-length model. Hence, an *ab initio* structure prediction was done using I-TASSER server for these regions of podocin. The templates thus obtained from the BLAST search and I-TASSER prediction were used to build a multiple sequence-based model (Fig. 1). The Ramachandran plot for podocin model revealed 85.5% of amino acids in most favorable regions; 10.9% amino acids in additionally allowed regions; 2.7% of amino acids in generously allowed regions; and 0.9% amino acids in disallowed regions (Fig. 2A). It is

noteworthy that the prediction revealed the presence of random coils at both N- and C-terminals (Figs. 2B and 2C). We speculated that these random coils could be IURs, as BLAST search with N-terminus podocin sequence showed homology with influenza A virus nucleoprotein that showed unstructured regions. Therefore, we further continued our quest of gaining insights on podocin structure by analyzing whether these coils are intrinsically unstructured. As IURs form anchors and signaling motifs, we speculate that IURs in podocin may involve in the formation of scaffolding complex with CD2AP and TRPC6 ion channel besides facilitating signaling via nephrin (13,35).

TABLE 2 Prediction of intrinsically unstructured regions (IURs) in stomatin family proteins

Proteins	Residues
Erythrocyte band 7 integral membrane protein	1–25, 203–241, 278–288
Hflc	1–3, 134–197, 228–273, 318–334
Hflk	1–79, 210–223, 249–287, 330–419
Protein UNC-1	1–25, 196–239, 279–285
Stomatin-1	1–36, 202–249, 286–330
Stomatin-2	1–117, 297–340, 353–374
Stomatin-3	1–7, 193–209, 259–267
Stomatin-like protein-1	1–49, 221–292, 395–398
Stomatin-like protein-3	1–22, 197–242, 259–263, 274–288

Predicting Intrinsically Unstructured and Transmembrane Regions

Podocin sequence was screened using seven standalone and three meta-IUR prediction servers, which predicted an N-terminal IUR (1–98 residues) and C-terminal IUR (347–383 residues; Table 1 and Fig. 3). Furthermore, we have used Genesilico metadisorder prediction server to assess the presence of IURs, if any in other stomatin family proteins. Interestingly, all stomatin family proteins showed IURs at various regions in their respective sequences (Table 2). These results strongly suggest the presence of IURs at both N- and C-terminus of podocin. The PSI-PRED server identified the presence of protein-binding regions between residues 1–7, 63–69, 77–89, 353–356, 363, and 367–377 in the identified IURs, which perhaps explains scaffolding complex formation of podocin with

CD2AP, nephrin (13), and extended complex with TRPC6 (35). The server also predicted a TM region in podocin spanning from ~101 to 125 amino acids.

It has been reported that stomatin family members exhibit cytoplasmic C-and N-terminus (36), and thus, podocin is also expected to show similar structural properties to that of other stomatin family members. Contrastingly, PSI-PRED analysis showed two possible orientations for N- and C-terminals of podocin. The prediction suggests that N-terminus is cytoplasmic and C-terminus is extracellular if the TM region is restricted between 98 and 125 residues (Fig. 4A). Alternatively, it was also predicted that N-terminus is extracellular and C-terminus is cytoplasmic if the TM region is localized between 103 and 127 residues (Fig. 4B).

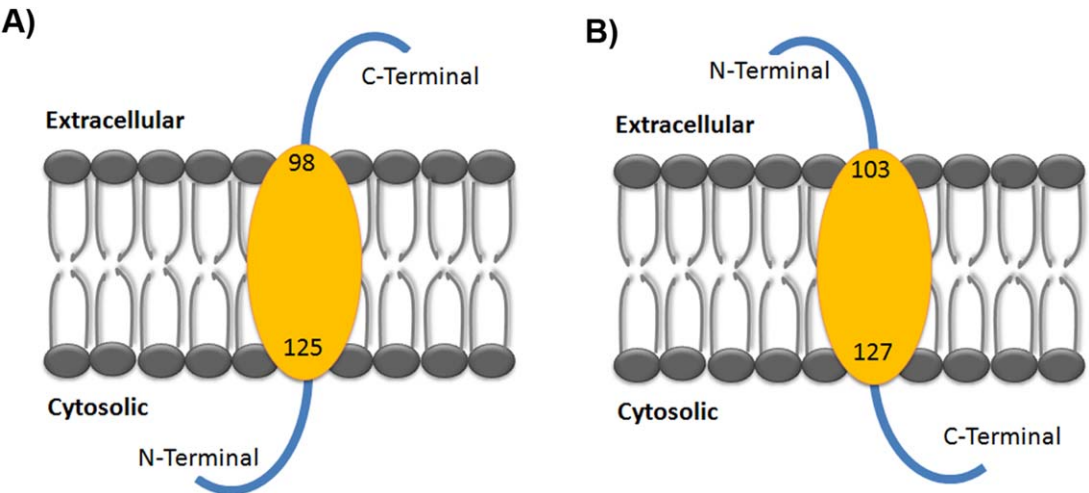


FIG 4 Transmembrane (TM) region predictions for podocin: PSI-PRED predicted two possible conformations for podocin transmembrane region: (A) N-terminus is cytoplasmic and C-terminus is extracellular if TM region is located between 98 and 125 residues, and (B) N-terminus is extracellular and C-terminus is cytoplasmic if the TM region spans between 103 and 127 residues.

TABLE 3
Distribution of backbone dihedral angles of identified mutants

S.No	Mutants	Ramachandran plot analysis			
		Most favored (%)	Additionally allowed (%)	Generously allowed (%)	Disallowed (%)
1	R3G	86.6	10.3	2.4	0.6
2	P89T	85.8	11.5	1.8	0.9
3	R322Q	87.0	10.6	1.8	0.6
4	R322P	85.4	11.9	2.1	0.6
5	H325Y	86.1	10.9	2.4	0.6
6	V370G	86.0	11.2	1.8	0.9

Mutations in Podocin Alter Intraprotein Interactions

To analyze how mutations in podocin, specifically in the predicted IURs, alter its structure and perturb its interactions with subpodocyte proteins, six podocin mutants were selected, and models for these mutants were generated using wild-type podocin model as a template. The Ramachandran plots for the mutants showed minimal variation with respect to that of wild-type podocin despite the fact that these podocin mutants are associated with NS (Table 3).

The mutants were then analyzed to identify the changes if any in the intraprotein interactions with the help of PIC server. This server calculates various interactions based on the given coordinates of 3D structure of a protein (30). Substantial differences in “main chain–main chain,” “main chain–side chain,” “side chain–side chain,” and “hydrophobic interactions” were observed when compared with wild-type podocin.

These altered intraprotein interactions in podocin mutants showed considerable distortions in their secondary structures. All podocin mutants we analyzed showed the formation of an extended α -helix between residues 307 and 325 except in mutant R3G, wherein T315 residue was found to be a part of the α -helix (residues 316–322) that was observed in wild-type protein (Fig. 5). Two β -sheets (170–181 and 185–197 residues) in the wild-type podocin were trimmed (170–180 and 186–197 residues) in the case of H325Y mutant. Furthermore, in mutants R3G, P89T, R322Q, and V370G, stable α -helices (residues 151–156) were replaced with 3_{10} helices (residues 153–155; Fig. 5).

Root-mean-square deviations of podocin mutants with that of podocin wild type showed that R3G mutant deviates by 1.23 Å when compared with wild-type podocin. Similarly, P89T, R322Q, R322P, H325Y, and V370G mutants showed 1.01, 1.04, 1.12, 0.68, and 0.55 Å deviations when compared with the wild-type podocin (Fig. 6).

Discussion

In this study, we have provided insights of podocin structure that was generated using a multiple template-based structure

prediction analysis. This study predicted the orientation of TM region and the presence of random coils, which we have identified as IURs at N- and C-terminal regions of podocin. We have also identified protein-binding regions in these IURs, which could play a critical role in the complex formation with subpodocyte SD proteins. Intraprotein interactions among various amino acid residues were assessed in wild type and in some of the podocin mutants that were associated with steroid-resistant form of INS. Mutations in podocin alter innate intraprotein interactions affecting the native structure of podocin and its ability to form critical complex with subpodocyte proteins.

The structure of human podocin was designed based on the multiple sequence homology modeling, and the model thus generated was evaluated for protein stereochemistry using PROCHECK software. Interestingly, it was observed that both N- and C-terminal regions of podocin showed random coils. pBLAST search with the N-terminal region of podocin (1–156) showed homology with the nucleoprotein of influenza A virus (PDB ID: 3ZDP; identity: 33%, query coverage: 39%), which displays a flexible disordered region (402–428 residues). This region was identified to connect the tail loop and the main body of the nucleoprotein that points toward the RNA-binding surface (37). Therefore, we have speculated the presence of intrinsically unstructured regions (IURs) in the podocin.

IURs do not form a fixed 3D structure under physiological conditions either in their entirety or they may contain intrinsically disordered regions. They resemble the denatured states of ordered proteins and are described as an ensemble of rapidly interconverting alternative structures (38). IURs take up different structures on binding to different targets and thereby exhibit functional flexibility through the formation of fuzzy complexes (39). Comprehensive analysis by various CASP validated servers revealed that podocin consists of both N-terminal (1–98 amino acids) and C-terminal (347–383 amino acids) IURs. Our predictions also identified protein-binding regions in these IURs, which makes us speculate that these IURs may help podocin to serve as an anchor, signaling motif

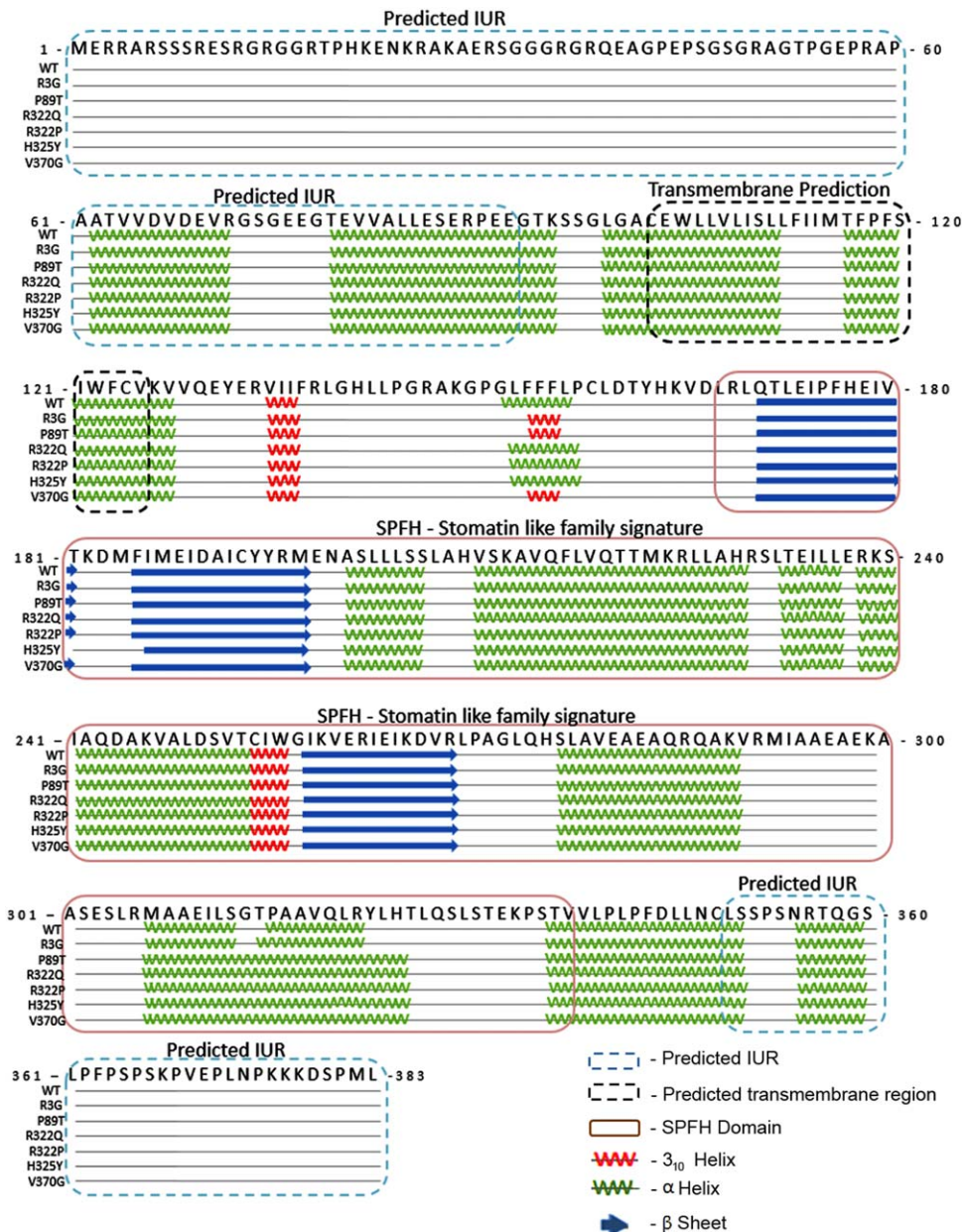


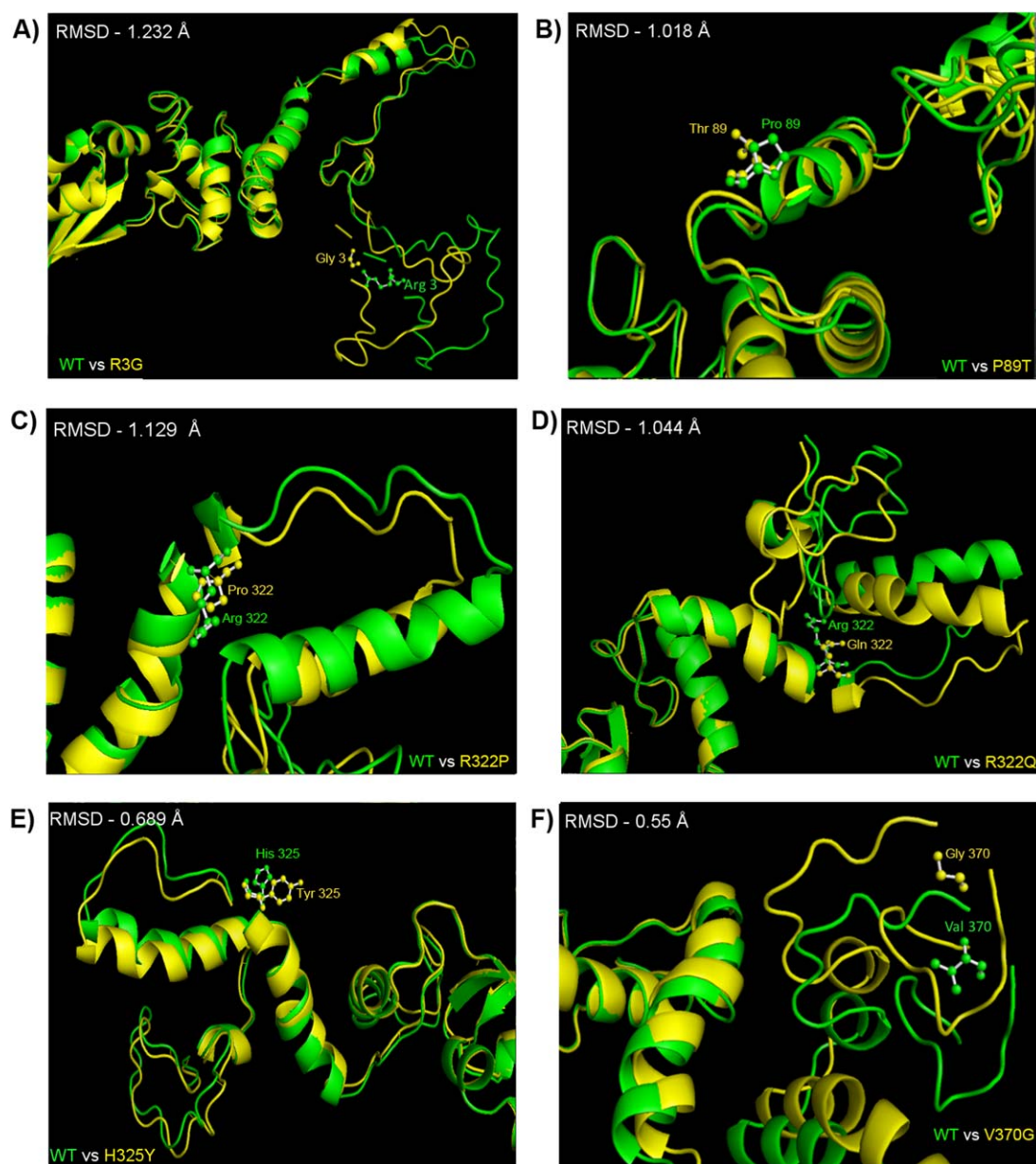
FIG 5

Mutations in podocin distorted their secondary structure. An extended α -helix between residues 307 and 325 was noticed in all the mutants apart from R3G mutant, wherein T315 residue was found to be a part of the α -helix (residues 316–322) when compared with WT. In mutants R3G, P89T, R322Q, and V370G, stable α -helices (residues 151–156) were replaced with 3_{10} helices (residues 153–155). However, in mutation H325Y, additional structural changes such as trimming of two β -sheets to residues 170–180 and 186–197 were also observed.

or help in interacting with neighboring proteins to form a complex, thus maintaining podocyte SD integrity (13,35,40). However, a detailed elucidation of how nephrin-CD2AP-podocin complex is formed is greatly warranted, which could be accomplished by solving the high-resolution crystal structure of these protein(s) complexes.

When we were analyzing IURs in podocin, the PSI-PRED server predicted the presence of a TM region (100–125 amino acids) in podocin, which correlated with the results from an

earlier study (10). Interestingly, TM region that was identified in podocin is homologous to that present in other stomatin family members (10). It was not known whether N-terminal region is cytoplasmic or extracellular; however, it was reported that C-terminus is cytoplasmic (13). Nevertheless, the possibility of podocin to form a hairpin loop to interact with the membrane from cytoplasmic side cannot be ruled out as other stomatin family proteins possess cytoplasmic C-terminus and forms hairpin loops (10). Furthermore, it was shown that


FIG 6

Structural alignment of podocin and its mutants: Each figure (A–F) represents structural alignment of wild-type podocin model with a mutant. The ball and stick representation in each figure (A: WT vs. R3G; B: WT vs. P89T; C: WT vs. R322P; D: WT vs. R322Q; E: WT vs. H325Y; and F: WT vs. V370G) shows the amino acid that was mutated. Root-mean-square deviation for each mutant was calculated using Pymol.

N- and C-terminal-specific podocin antibodies colocalized at the cytoplasmic face of the plasma membrane suggesting the presence of hairpin loop structure (41).

Mutations in *NPHS2* cause early onset of SRNS and ensue rapid progression to ESRD. We have selected six mutations from HDMG database and generated models for them. We attempted to understand how mutations specifically in the intrinsically unstructured regions affected the macromolecular assembly. It was found that mutations elicit significant changes in intraprotein interactions such as main chain–main chain, main chain–side chain, hydrophobic, and side chain–side chain

contacts. These changes showed considerable differences in protein secondary structure. It could be speculated that alterations in intraprotein interactions bring about changes in microenvironment and compromise stereochemistry and macromolecular assembly of podocin with its subpodocyte components. Furthermore, the structural changes observed at distant regions of mutations may be viewed along with the larger changes observed in the intermittent residues as well. For instance, R3G mutant shows change of secondary structure from α -helix to 3_{10} helix in the region of 151–156, thus bringing about distant structural changes via the changes in

intermittent residues, which could be observed in other mutant forms as well. This could possibly explain how mutations in podocin cause detrimental changes in a larger scale to the podocyte SD structure and the loss of podocyte function.

It is noteworthy that podocin mutants that were selected in this study such as R3G, R322Q, R322P, H325Y, and V370G were manifested in FSGS and that the mutation P89T resulted in MCD (31,32,42). It was shown that podocin mutant H325Y progresses to ESRD (43). From the observed data, we speculated that even the slightest variation in the conformation of the protein secondary structure could alter the macromolecular assembly of podocin with sub-SD proteins because of the alterations of podocin backbone, which in turn could eventually manifest in altered architecture of SD and proteinuria in NS.

Acknowledgements

The authors thank the Department of Science and Technology, India, for providing research grants IFA11-LSBM-02 and SB/FT/LS-307/2012 (to A. K. P.). The authors acknowledge the resources provided by Biotechnology Infrastructure Facility (BIF) of the University of Hyderabad.

Conflict of Interest

The authors do not have any conflict of interest to declare.

References

- [1] Anil Kumar, P., Welsh, G. I., Saleem, M. A., and Menon, R. K. (2014) Molecular and cellular events mediating glomerular podocyte dysfunction and depletion in diabetes mellitus. *Front. Endocrinol. (Lausanne)* 5, 151.
- [2] Bazzi, C., Bakoush, O., and Gesualdo, L. (2012) Proteinuria: from molecular to clinical applications in glomerulonephritis. *Int. J. Nephrol.* 2012, 424968.
- [3] Kumar, P. A., Frank, C. B., and Ram, K. M. (2011) The glomerular podocyte as a target of growth hormone action: implications for the pathogenesis of diabetic nephropathy. *Curr. Diabetes Rev.* 7, 50–55.
- [4] Greka, A., and Mundel, P. (2012) Cell biology and pathology of podocytes. *Annu. Rev. Physiol.* 74, 299–323.
- [5] Somlo, S., and Mundel, P. (2000) *Getting a foothold in nephrotic syndrome.* *Nat. Genet.* 24, 333–335.
- [6] Gigante, M., Piemontese, M., Gesualdo, L., Iolascon, A., and Aucella, F. (2011) Molecular and genetic basis of inherited nephrotic syndrome. *Int. J. Nephrol.* 2011, 792195.
- [7] International Study of Kidney Disease in Children. (1978) Nephrotic syndrome in children: prediction of histopathology from clinical and laboratory characteristics at time of diagnosis. *Kidney Int.* 13, 159–165.
- [8] Sadowski, C. E., Lovric, S., Ashraf, S., Pabst, W. L., Gee, H. Y., et al. (2015) A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. *J. Am. Soc. Nephrol.* 26, 1279–1289.
- [9] Benoit, G., Machuca, E., and Antignac, C. (2010) Hereditary nephrotic syndrome: a systematic approach for genetic testing and a review of associated podocyte gene mutations. *Pediatr. Nephrol.* 25, 1621–1632.
- [10] Boute, N., Gribouval, O., Roselli, S., Benessy, F., Lee, H., et al. (2000) NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome. *Nat. Genet.* 24, 349–354.
- [11] Miner, J. H. (2002) Focusing on the glomerular slit-diaphragm: podocin enters the picture. *Am. J. Pathol.* 160, 3–5.
- [12] Schurek, E. M., Völker, L. A., Tax, J., Lamkemeyer, T., Rinschen, M. M., et al. (2014) A disease-causing mutation illuminates the protein membrane topology of the kidney-expressed prohibitin homology (PHB) domain protein podocin. *J. Biol. Chem.* 289, 11262–11271.
- [13] Schwarz, K., Simons, M., Reiser, J., Saleem, M. A., Faul, C., et al. (2001) Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin. *J. Clin. Invest.* 108, 1621–1629.
- [14] Nishibori, Y., Liu, L., Hosoyamada, M., Endou, H., Kudo, A., et al. (2004) Disease-causing missense mutations in NPHS2 gene alter normal nephrin trafficking to the plasma membrane. *Kidney Int.* 66, 1755–1765.
- [15] Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40.
- [16] Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815.
- [17] Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195–202.
- [18] Jones, D. T., and Cozzetto, D. (2015) DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* 31, 857–863.
- [19] Jones, D. T., Taylor, W. R., and Thornton, J. M. (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* 33, 3038–3049.
- [20] Nugent, T., and Jones, D. T. (2009) Transmembrane protein topology prediction using support vector machines. *BMC Bioinformatics* 10, 159.
- [21] Kozłowski, L. P., and Bujnicki, J. M. (2012) MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. *BMC Bioinformatics* 13, 1–11.
- [22] Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., et al. (2003) Predicting intrinsic disorder from amino acid sequence. *Proteins* 53(Suppl.6), 566–572.
- [23] Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., Dunker, A. K., et al. (2005) Optimizing long intrinsic disorder predictors with protein evolutionary information. *J. Bioinform. Comput. Biol.* 3, 35–60.
- [24] Dosztanyi, Z., Csizmek, V., Tompa, P., and Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21, 3433–3434.
- [25] Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., et al. (2003) Protein disorder prediction: implications for structural proteomics. *Structure* 11, 1453–1459.
- [26] Zhang, T., Faraggi, E., Xue, B., Dunker, A. K., Uversky, V. N., et al. (2012) SPINE-D: accurate prediction of short and long disordered regions by a single neural-network based method. *J. Biomol. Struct. Dyn.* 29, 799–813.
- [27] Mizianty, M. J., Stach, W., Chen, K., Kedarisetti, K. D., Disfani, F. M., et al. (2010) Improved sequence-based prediction of disordered regions with multilayer fusion of multiple information sources. *Bioinformatics* 26, i489–i496.
- [28] Schlessinger, A., Punta, M., Yachdav, G., Kajan, L., and Rost, B. (2009) Improved disorder prediction by combination of orthogonal approaches. *PLoS One* 4, e4433.
- [29] Rost, B., Fariselli, P., and Casadio, R. (1996) Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci.* 5, 1704–1718.
- [30] Tina, K. G., Bhadra, R., and Srinivasan, N. (2007) PIC: protein interactions calculator. *Nucleic Acids Res.* 35, W473–W476 (Web Server Issue).
- [31] Berdeli, A., Mir, S., Yavascan, O., Serdaroglu, E., Bak, M., et al. (2007) NPHS2 (podocin) mutations in Turkish children with idiopathic nephrotic syndrome. *Pediatr. Nephrol.* 22, 2031–2040.
- [32] Caridi, G., Gigante, M., Ravani, P., Trivelli, A., Barbano, G., et al. (2009) Clinical features and long-term outcome of nephrotic syndrome associated with heterozygous NPHS1 and NPHS2 mutations. *Clin. J. Am. Soc. Nephrol.* 4, 1065–1072.
- [33] Yokoyama, H., Fujii, S., and Matsui, I. (2008) Crystal structure of a core domain of stomatin from *Pyrococcus horikoshii* illustrates a novel trimeric and coiled-coil fold. *J. Mol. Biol.* 376, 868–878.
- [34] Singleton, M. R., Dillingham, M. S., Gaudier, M., Kowalczykowski, S. C., and Wigley, D. B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* 432, 187–193.
- [35] Huber, T. B., Schermer, B., Muller, R. U., Hohne, M., Bartram, M., et al. (2006) Podocin and MEC-2 bind cholesterol to regulate the activity of associated ion channels. *Proc. Natl. Acad. Sci. USA* 103, 17079–17086.

- [36] Lapatsina, L., Brand, J., Poole, K., Daumke, O., and Lewin, G. R. (2012) Stomatin-domain proteins. *Eur. J. Cell Biol.* 91, 240–245.
- [37] Chenavas, S., Estrozi, L. F., Slama-Schwok, A., Delmas, B., Di Primo, C., et al. (2013) Monomeric nucleoprotein of influenza A virus. *PLoS Pathog.* 9, e1003275.
- [38] Sickmeier, M., Hamilton, J. A., LeGall, T., Vacic, V., Cortese, M. S., et al. (2007) DisProt: the database of disordered proteins. *Nucleic Acids Res.* 35, D786–D793 (Database Issue).
- [39] Tompa, P., and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.* 33, 2–8.
- [40] Huber, T. B., Kottgen, M., Schilling, B., Walz, G., and Benzing, T. (2001) Interaction with podocin facilitates nephrin signaling. *J. Biol. Chem.* 276, 41543–41546.
- [41] Roselli, S., Gribouval, O., Boute, N., Sich, M., Benessy, F., et al. (2002) Podocin localizes in the kidney to the slit diaphragm area. *Am. J. Pathol.* 160, 131–139.
- [42] Sharma, S., Kabra, M., Hari, P., Dinda, A. K., and Bagga, A. (2008) NPHS2 and WT1 mutations in Indian children with steroid-resistant nephrotic syndrome. *Genomic Med.* 2, 241–252.
- [43] Caridi, G., Bertelli, R., Scolari, F., Sanna-Cherchi, S., Di Duca, M., et al. (2003) Podocin mutations in sporadic focal-segmental glomerulosclerosis occurring in adulthood. *Kidney Int.* 64, 365–368.



Intrinsically disordered regions mediate macromolecular assembly of the Slit diaphragm proteins associated with Nephrotic syndrome

Sandeep K. Mulukala Narasimha^{a*}, Prajna Parimita Kar^b, Ramakrishna Vadrevu^c and Anil K. Pasupulati^a

^aDepartment of Biochemistry, University of Hyderabad, Hyderabad, India; ^bLaboratory of Molecular Interactions, National Institute of Animal Biotechnology, Hyderabad, India; ^cDepartment of Biological Sciences, BITS-Pilani, Hyderabad Campus, Hyderabad, India

ABSTRACT

The glomerular filtration barrier (GFB) of the kidney plays an instrumental role in preventing the excretion of large molecules and ensuring the formation of ultra-filtrated urine. Podocytes are essential components of GFB that provide epithelial coverage to the fine glomerular capillaries. Slit-diaphragm (SD) that forms the sole contact between adjacent foot-processes of the podocytes consists of multimeric protein assemblies. SD serves as a molecular sieve and confers size and charge-selective barrier. Nephrin, podocin, TRPC6, and CD2AP are some of the key proteins that constitute the SD. Mutations in these proteins are implicated in nephrotic syndrome and congenital nephropathies which are characterised by heavy proteinuria. The mechanism of how mutations in these proteins predispose to proteinuria is not known. Furthermore, the structural details of proteins that constitute SD are largely unknown. In this study, we built models for nephrin, CD2AP, podocin, and TRPC6 followed by docking and molecular dynamics simulations of the complex of these proteins. We speculate that the interfacial residues of SD proteins form a macromolecular complex through intrinsically disordered regions thereby conferring architectural stability to the SD, which is critical for glomerular permselectivity.

ARTICLE HISTORY

Received 31 August 2018
Accepted 10 January 2019

KEYWORDS

Podocyte; nephrotic syndrome; slit-diaphragm; protein-protein interactions; intrinsically disordered regions

1. Introduction

The kidneys play an instrumental role in maintaining the body homeostasis by regulating acid-base, electrolyte and water balance. Each kidney is composed of about a million nephrons and the two units of a nephron; glomerulus and renal tubule work in concert and regulate the final composition of the urine. The glomerular filtration barrier (GFB), which ensures glomerular permselectivity constitutes of capillary endothelium, glomerular basement membrane, and specialised epithelial cells, called podocytes [1]. Among the three constituents of GFB; podocytes play a critical role in regulating glomerular filtration. Podocytes are highly branched terminally differentiated cells consisting of a major cell body, primary processes, and secondary foot processes. The secondary foot processes of adjacent podocyte(s) interlace to form a zipper-like structure called the slit diaphragm (SD) [2]. SD is ~40 nm wide and is sufficient to curb the passage of albumin and allows only water and small molecules to pass through it [3]. The SD is a negatively charged modified tight junction composed of several proteins including nephrin, CD2-associated protein (CD2AP), podocin, transient receptor potential cation channel, subfamily c, member 6 (TRPC6), zona occludens 1 (ZO-1), and nephrin-like 1 (NEPH1) [4]. It is considered that SD offers size, shape and charge-selective barrier at the plasma-urine interface, thus dictating the glomerular permselectivity.

Nephrotic syndrome (NS) is a nonspecific renal disorder characterised by heavy proteinuria, hypoalbuminemia, and oedema. The idiopathic NS is classified into steroid-sensitive

(SSNS) and steroid-resistant (SRNS), depending on their response to corticosteroid therapy. SRNS has a poor prognosis and often leads to end-stage renal disease (ESRD). Mutations in about 20 genes that majorly encode podocyte proteins were identified to be associated with SRNS [5]. The most severe congenital NS of the Finnish type is caused due to mutations in Nephrotic syndrome, type 1 gene (*NPHS1*) that encodes nephrin. Mutations in *NPHS1* cause massive proteinuria in utero and nephrosis at birth [6]. Mutations in Nephrotic syndrome, type 2 (*NPHS2*) gene that encodes podocin occur frequently and are responsible for ~18% of reported SRNS cases [7,8]. Mutations in podocin typically lead to the early onset of SRNS, usually before 6 years of age, which later progresses to ESRD by 10 years of age [9]. Mutations in CD2AP, which is another SD protein, were also associated with congenital NS [10]. TRPC6 is also a constituent of SD and mutations in the conserved residues of TRPC6 increased calcium influx leading to podocyte dysfunction and proteinuria [11].

Increasing evidence suggests that nephrin, CD2AP, podocin and TRPC6 form a complex, which is believed to play a major role in maintaining the SD architecture [12,13,14,15,16]. Immunoprecipitation studies suggest that the C-terminus of podocin interacts with CD2AP and nephrin [15]. Furthermore, CD2AP was also found to interact with the cytoplasmic domain of nephrin [16]. It has been shown that the interaction between podocin and nephrin is crucial for nephrin signalling [13]. Similarly, co-immunoprecipitation experiments revealed that TRPC6 interacts with nephrin and podocin but not with

CD2AP [12]. Although it was suggested that these four proteins colocalize as a complex, contributions of these proteins to the integrity and stability of SD is not known. It is yet to be elucidated how mutations in these proteins contribute towards the structural alterations to the SD of podocytes, which eventually manifests in severe proteinuria.

In this study, we built models for nephrin, CD2AP, podocin, and TRPC6 and refined them using energy minimizations. Further, we predicted the macromolecular assembly of these four proteins by protein–protein docking. We also performed molecular dynamics simulations of the complex. Our studies revealed that these four proteins possess intrinsically disordered regions (IDRs) and IDR-binding domains (IDR-BD), which were actively involved in forming both homodimers (nephrin–nephrin) and macromolecular complex. Further, our study provides insights on how mutations in SD proteins associate with morphological changes that can lead to proteinuria

2. Methodology

2.1. Identifying homologs, motifs, IDRs and transmembrane regions in SD proteins

The *Homo sapiens* nephrin (accession ID: O60500), CD2AP (accession ID: Q9Y5K6), TRPC6 (accession ID: Q9Y210) and podocin (accession ID: Q9NP85) sequences were obtained from UniProt database (www.uniprot.org). Sequence analysis was performed to identify homologs or motifs to guide model building. An NCBI Protein BLAST search for all these sequences was done against the PDB database. We next wanted to understand the secondary structure details of these proteins, for which we employed PSI-PRED server. This server also predicted transmembrane regions and IDRs and predicting disordered protein binding residues [17,18,19]. To further confirm the presence of IDRs in nephrin, CD2AP, podocin and TRPC6 sequences, we used a Critical Assessment of Protein Structure Prediction 9 (CASP9) validated Genesilico-meta disorder server (Meta Disorder MD2). This server predicts IDRs in a sequence, based on the optimisation of the predictions made by 15 different primary prediction servers.

2.2. Model building, refinement and stereochemical quality analysis

There were no structural templates for nephrin to perform homology modelling. We, therefore, used RaptorX (<http://raptorx.uchicago.edu/StructurePrediction/predict>), a threading-based modelling server to predict a model for nephrin. However, this server was unable to predict a model for the full-length sequence, owing to which we have considered part sequences with overlapping residues i.e. 1–1241, and 1140–1241 residues to predict the model. Next, the templates from RaptorX were stitched using modeller 9.17 to predict the full-length model. We performed *ab-initio* predictions for proteins that could not be built by either homology modelling or by threading modelling. In the case of CD2AP, the crystal structures of the SH3 domains were already solved [20,21,22]. However, the rest of the CD2AP sequence showed less homology to

build a complete model. Therefore, we used a hybrid method encompassing homology modelling and *ab-initio* predictions. I-TASSER, an *ab initio* server was used for predicting structure for the non-homologous regions of CD2AP. The models obtained from the I-TASSER server and the solved structures from pBLAST search were used as templates in modeller 9.17 software to build a model for CD2AP. NCBI pBLAST search revealed that the crystal structure of the core domain of stomatin from *Pyrococcus horikoshii* (Chain A, PDB ID-3BK6) showed an identity score of 39% with podocin. However, the sequence coverage of the template with podocin was only 44% (residues 165–331). Therefore, I-TASSER predictions were done for the residues 1–164 and 332–383 of podocin [4]. The models from the I-TASSER and the template from the pBLAST search were used as templates in modeller 9.17 to predict a full-length model for podocin. In the case of TRPC6, there were no suitable homologs and the threading based modelling was unable to predict a structure. Therefore, *ab-initio* predictions were done to obtain the full-length TRPC6 model. Once suitable models were built for the SD proteins, energy minimizations were carried out using ModRefiner software.

Furthermore, the models were thoroughly subjected to stereochemistry analysis using PROCHECK software. The models for V822M nephrin, K301M CD2AP, R138Q podocin and R895L TRPC6 mutants were generated by homology modelling using the wildtype models as templates and the models thus obtained were minimised using ModRefiner. The wild-type and mutant models were then analysed using (Protein Interactions Calculator) PIC server. This server details intra-protein hydrophobic, main chain-main chain, main chain-side chain, side chain-side chain, aromatic-aromatic, ionic, aromatic-sulphur, and cation-pi interactions [23].

2.3. Building the macromolecular complexes

Protein–protein docking was carried out using PatchDock, a validated Critical Assessment of Prediction of Interactions (CAPRI) server [24,25]. Initially, rigid docking was done for the predicted nephrin and CD2AP models. To this complex, podocin followed by TRPC6 models were blind docked since the interacting regions of TRPC6 with nephrin, podocin and CD2AP are not known. In the case of building mutant complexes, all the proteins were blind docked sequentially since, the effect of mutations on the structure of each protein is not known. The results obtained from PatchDock were refined through FireDock and the best-docked complex was then selected for molecular dynamics simulations.

2.4. Molecular dynamics simulations and analysis

In order to ascertain the stability of the macromolecular assembly we performed molecular dynamics (MD) simulations using a modified approach from the studies by Kirubakaran et al. [26,27]. Molecular dynamics simulations were performed using a Gromos96–43a force field. The complex was centred in a 35.0 × 30.0 × 64.0 Å box, which was solvated with an SPC216 water system. The charge of the system was neutralised by 41 Na⁺ ions following which, initial energy minimisation

and equilibration steps were carried out by steepest descent minimisation. The system was further equilibrated at a constant temperature and pressure of 300 K and 1 atm. The long-range electrostatics were treated with the particle-mesh Ewald method. We next used PRODIGY web server to extrapolate the intermolecular contacts and types of interactions such as charged-charged, charged-polar, charged-apolar, polar-polar, polar-apolar, apolar-apolar, percentage of a non-interacting surface (NIS) charged residue and percentage of NIS apolar residues in the complex. We employed PDBsum server to reaffirm the intermolecular contacts besides analysing the number of salt bridges, hydrogen bonds, and non-bonded contacts present at the protein–protein interface.

3. Results

3.1. Sequence analysis of nephrin, CD2AP, podocin, and TRPC6

The native sequences taken from UniProt database are nephrin comprising of 1241 residues, CD2AP with 639 residues, podocin made of 383 residues and TRPC6 consist of 931 residues. A pBLAST search of the above sequences was done against the non-redundant database revealed the presence of putative conserved domains in all the sequences. In nephrin, the residues 38–119 belonged to immunoglobulin (Ig) V-set domain, residues 141–228, 345–424, 450–541, 549–631, 655–736, and 834–942 were identified as Ig-like-C2 domains and residues 242–321 and 740–820 were identified as the third Ig domain. Residues 957–1033 resembles a fibronectin type III domain in the nephrin sequence (Figure 1A). In the case of CD2AP

sequence, three SH3 domains were identified at residues at 3–58, 111–165, and 271–327 (Figure 2A). Similarly, analysis of the podocin sequence revealed the presence of SPFH (stomatins) family signature motif at residues 169–271 (Figure 3A). pBLAST of TRPC6 search revealed 3 ‘Ankyrin’ (ANK) repeats at 132–159, 161–185 and 212–247 followed by a TRP domain at 253–312 residues (Figure 4A).

3.2. SD proteins consist of intrinsically disordered regions (IDRs)

Upon initial screening with the PSI-PRED server, IDRs were identified in nephrin, CD2AP, podocin, and TRPC6 sequences. Further screening with Genesilico-meta disorder server (Meta-disorder MD2) confirmed the presence of IDRs in all four proteins (Table 1). Apart from predicting IDRs, PSI-PRED predicted certain regions in the IDRs containing binding hot-spots called as IDR-binding domains. In nephrin the IDR-BD were predicted at residues 487, 489–491, 783–786, 790, 1038–1055, 1092–1109, 1111 and 1178–1179. Similarly, in CD2AP we found IDR-BD at residues 62, 80, 264–266, 296–299, 606–607, 614, 617–622, 632 and 638–639 and in podocin at residues 3–6, 353–356, 362–376 and 382–383. In TRPC6, IDR-BD were identified at residues 196–205, 877, 879–889, 894, 900 and 908–930. The PSI-PRED server also predicted transmembrane regions in all the sequences (Table 1). Our analysis indicated that transmembrane segments were present both in nephrin at residues 1060–1084 and in podocin at residues 101–128. In the case of CD2AP, a transmembrane pore-lining (PL) region was predicted at residues 589–604. Analysis of the TRPC6 sequence revealed three transmembrane helices at residues

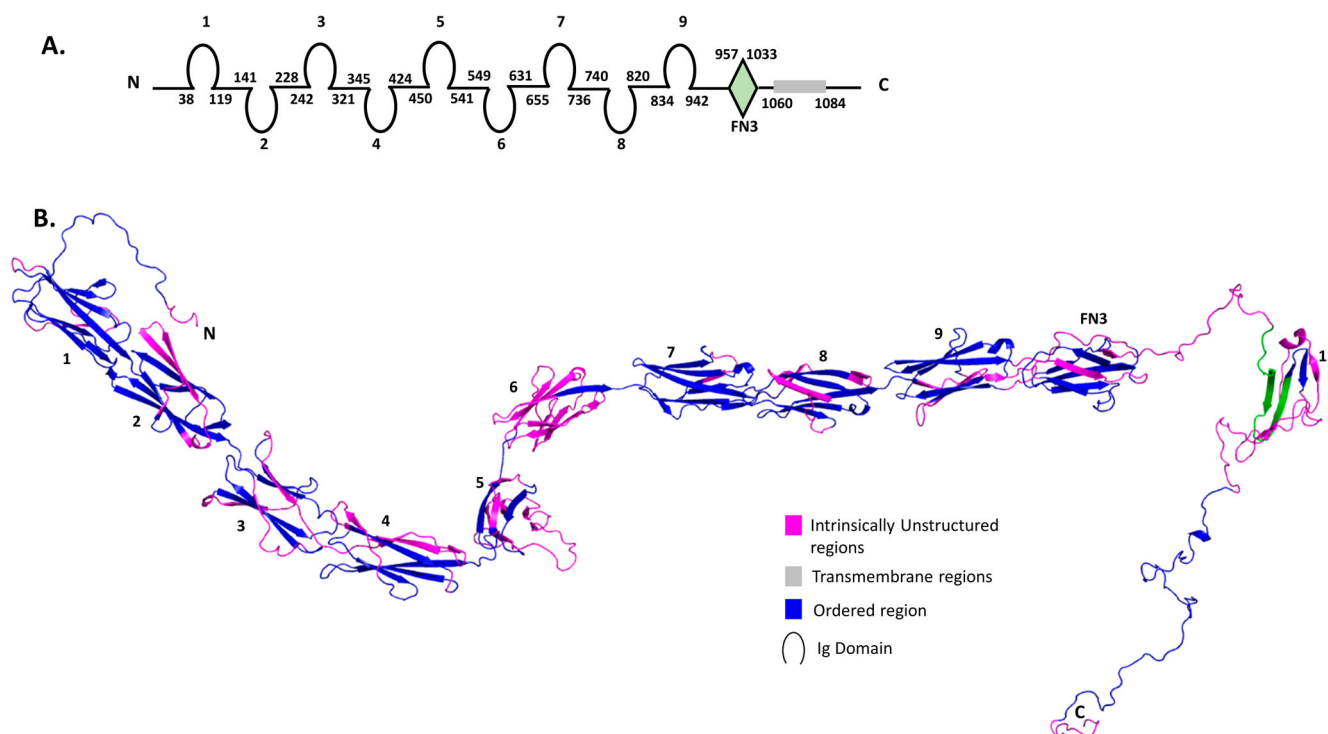


Figure 1. (Colour online) Predicted structure of nephrin: A) Schematic representation of nephrin sequence comprising of Immunoglobulin-like, Fibronectin type III (FN3), and transmembrane domains, B) Threading modelling predicted structure showing 9 Immunoglobulin-like domains and a fibronectin domain in nephrin. In the figure, a 11th domain covering the transmembrane segment is a new domain prediction that was not reported earlier. The IURs (pink) in the model may facilitate the formation of homodimers (nephrin-nephrin) or heteromers (nephrin-CD2AP-podocin-TRPC6).

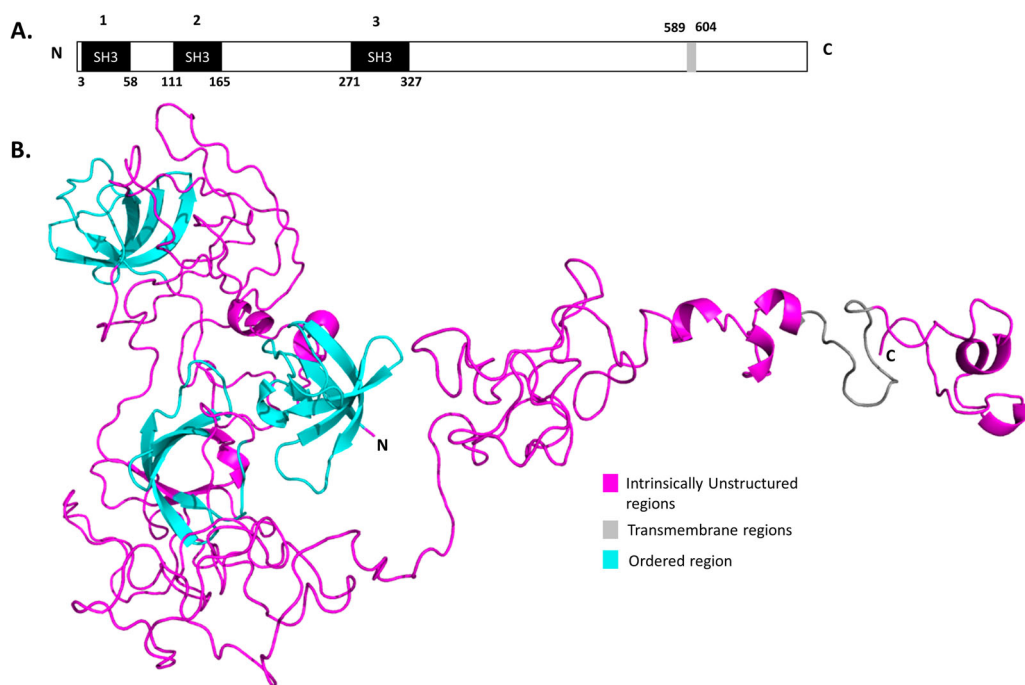


Figure 2. (Colour online) Predicted structure of CD2AP: A) Schematic representation of CD2AP sequence showing SH3 domains and a transmembrane pore lining region, B) The CD2AP model was predicted based on a combination of homology and *ab-initio* predictions. The model displays three SH3 domains. The IURs (Pink) shown in the model could act as possible sites for the formation of nephrin-CD2AP and CD2AP-podocin complexes.

406–429, 442–462, 529–552 and four trans-membrane pore-lining regions at residues 495–512, 596–614, 629–654 and 704–728 (Table 1).

3.3. Structure prediction, stereochemical analysis and insights into the predicted models

For nephrin and TRPC6 there are no acceptable homologs to build models. Therefore, threading modelling and *ab-initio*

were used to build models for these proteins. Whereas, *ab-initio* in conjunction with homology modelling were used to build models for CD2AP and podocin. These models were refined to obtain stereo-chemically acceptable structures. The Ramachandran plot analysis for predicted structures revealed that outlier values are in the acceptable range (Table 2). Upon analysing the predicted models, we observed that the nephrin model is composed of Greek key motifs at residues 41–132, 138–237, 244–333, 343–438, 445–541, 548–640,

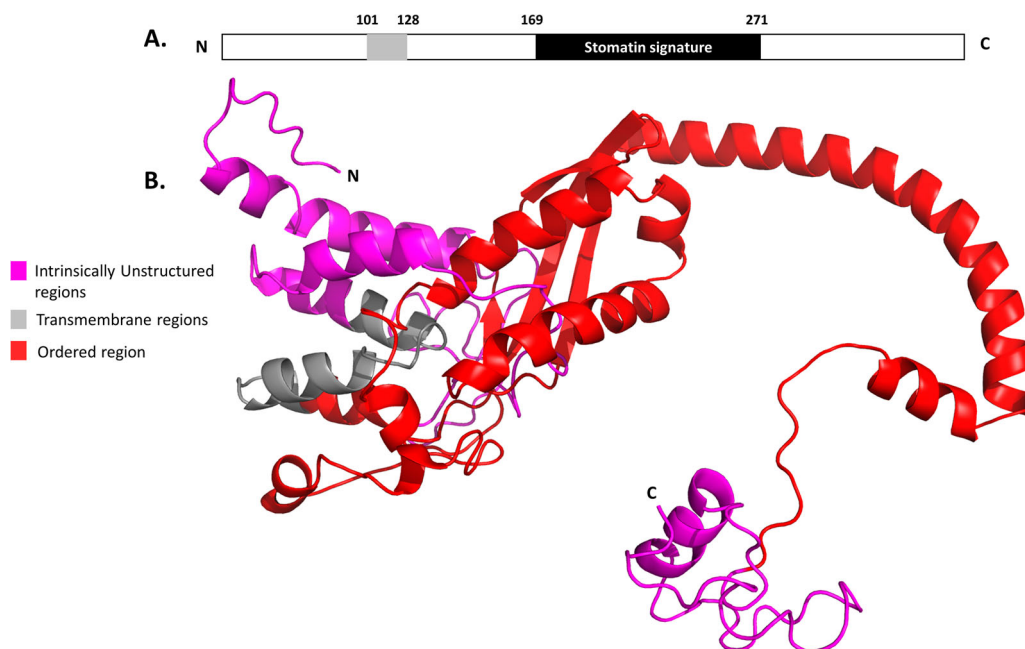


Figure 3. (Colour online) Predicted structure of podocin: A) Schematic representation of the podocin sequence showing SPFH domain and a transmembrane region, B) Podocin was built using a combination of homology and *ab-initio* predictions. The IURs (pink) shown in the model may act as a possible site for the formation of nephrin-podocin, podocin-TRPC6, and CD2AP-podocin complexes.

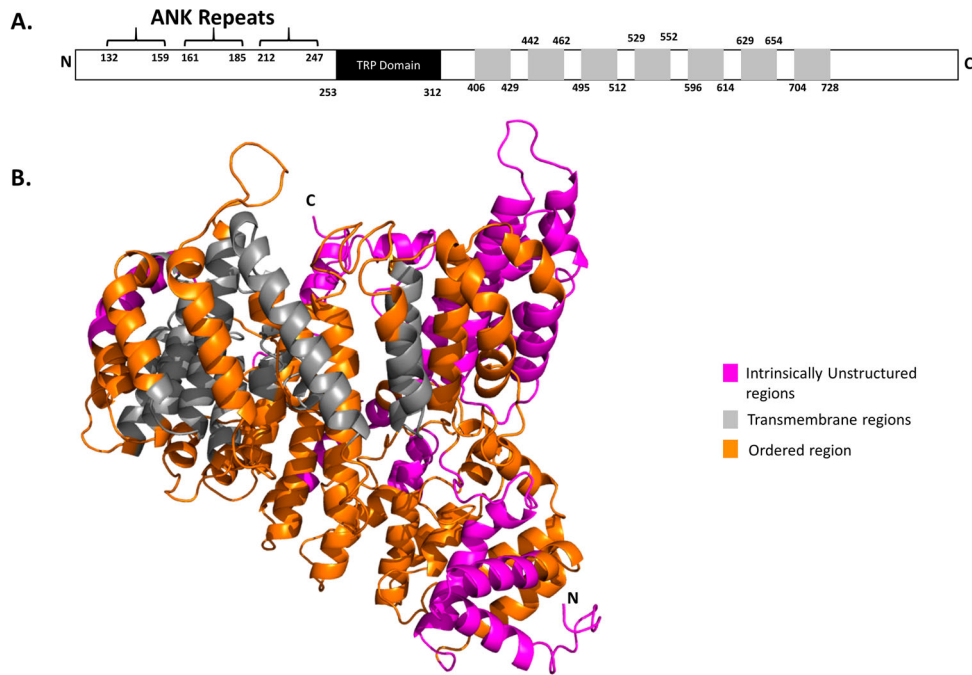


Figure 4. (Colour online) Predicted structure of TRPC6: A) Schematic representation of the TRPC6 sequence showing ANK repeats, TRP domain, transmembrane, and pore lining helices, B) The TRPC6 model was built using ab-initio predictions. The helices lining the pore are highlighted in grey and the IURs (pink) represent regions in the model that may facilitate the formation of the nephrin-TRPC6 and podocin-TRPC6 complexes.

Table 1. IUR predictions for nephrin, CD2AP, TRPC6 and podocin by Genesilico-meta disorder and PSI-PRED servers.

Server	Genesilico- Meta disorder server (IURs)	PSI-PRED	
		IURs	Transmembrane region predictions
Nephrin	1-8, 75-89, 117-121, 179-213, 243-255, 285-295, 323-365, 383-396, 407-409, 445-457, 474-529, 547-636, 689-696, 742-755, 779-798, 840-857, 886-900, 938-952, 992-1006, 1031-1059, 1095-1164, 1228-1241	1, 2, 349, 485, 487 (BD), 488, 489-491 (BD), 494, 783-786 (BD), 788, 790 (BD), 791, 1037, 1038-1055 (BD), 1092-1109 (BD), 1111 (BD), 1116-1166, 1178-1179 (BD), 1181-1228	1060-1084
TRPC6	1-40, 62-94, 189-204, 258-276, 347-363, 468-482, 790-931	1-84, 196-205 (BD), 790-848, 876, 877 (BD), 879-889 (BD), 894 (BD), 900 (BD), 908-930 (BD), 931	406-429, 442-462, 495-512 (PL), 529-552, 596-614 (PL), 629-654 (PL), 704-728 (PL)
CD2AP	1, 58-110, 165-274, 283-290, 320-590, 594-639	62 (BD), 63-79, 80 (BD), 81-101, 169-201, 224-263, 264-266 (BD), 267-270, 292, 296-299 (BD), 305-581, 603, 606-607 (BD), 610, 612, 614 (BD), 615, 617-622 (BD), 624, 632 (BD), 638-639 (BD)	589-604 (PL)
Podocin	1-100, 336-383	1-2, 3-6 (BD), 7-94, 353-356 (BD), 357-361, 362-376 (BD), 377-381, 382-383 (BD)	101-128
Remarks	Meta – Disorder MD2	BD – Binding Domain	Pore lining (PL) regions

647-737, 741-835, and 839-937 (Figure 1B). We identified a new domain with three antiparallel sheets at residues 1068-1136 (Figure 1B). An earlier study did not report about this unique domain in nephrin [6]. The residues 956-1033 was identified as fibronectin type III domain in the nephrin model. In the case of CD2AP model, three SH3 domains are localised to the residues 2-57, 113-158 and 272-328 (Figure 2B). Two noticeable features were observed in the CD2AP model; (a) presence of long random coils in-between the SH3 domains and (b) several IDRs across the protein sequence. The podocin model displayed SPFH family motif at residues 169-271 and long random coils at the N- and C-terminals, which are also identified to be IDRs (Figure 3B). Similarly, upon analysis the TRPC6 model, we observed signature ANK family repeats consisting of antiparallel helices joined by β -turns (Figure 4B). PDBsum analysis of the TRPC6 model revealed the presence of a pore with dimensions 1.49Åx17.2 Å (radius x length).

3.4. Protein-protein docking and molecular dynamic simulation analysis of the predicted complex

Experimental evidence did not provide adequate information about the interacting regions or the residues between these proteins when they constitute a multimeric complex. Based

Table 2. Ramachandran outlier's for Nephrin, CD2AP, Podocin and TRPC6 models.

Protein	Core	Allowed	Generously allowed	Dis-allowed
Nephrin	88.1%	9.9%	1.4%	0.6%
CD2AP	72.1%	25.1%	1.3%	1.5%
Podocin	81.8%	13.6%	2.1%	2.4%
TRPC6	81.7%	13.7%	3.2%	1.4%

on the available information on the role of IDRs, we speculated that IDRs/IDR-BD of these four proteins may play a key role in the formation of a multimeric protein complex [28]. Therefore, to gain the insights on the interfacial residues participating in the complex formation, we docked nephrin and CD2AP based on regions specified in the literature [12,13,14,15] to obtain nephrin-CD2AP complex. To this complex, we docked podocin followed by TRPC6 sequentially. The macromolecular complex thus obtained was subjected to MD simulations to understand the stability of the macromolecular assembly. We predicted that the C-terminal residues 1072–1201 and residues 851–1140 of nephrin interact with 116–447 residues of CD2AP (Figure 5A & Table 3) and with 154–903 residues of TRPC6 (Figure 5B & Table 3) with binding energies (ΔG) of -18 kcal/mol and ΔG of -20.3 kcal/mol. Residues 1–592 of CD2AP interacted with both the N- and the C-terminus of podocin (10–381) with a

ΔG of -21.8 kcal/mol (Figure 5C & Table 3). Interestingly, we found that residues 114 and 131 of CD2AP formed non-bonded contacts with residues 887 and 883 of TRPC6 with ΔG of -4.2 kcal/mol (Figure 5D & Table 3). It is noteworthy that our docking studies could not predict interactions between nephrin-podocin and podocin-TRPC6. Our analysis suggests that most of the interfacial contacts between nephrin, CD2AP, podocin and TRPC6 are contributed by IDRs/IDR-BD (Table 3, highlighted in yellow and bold-italics), which reiterates the importance of IDRs in the macromolecular assembly. A complete list of protein–protein interactions including the non-bonded contacts responsible for the complex formation is provided in the supplementary Table 1. In addition to the binding energy calculations, the number of contacts responsible for the complex formation such as salt bridges, hydrogen bonds and non-bonded contacts has been listed (Table 4).

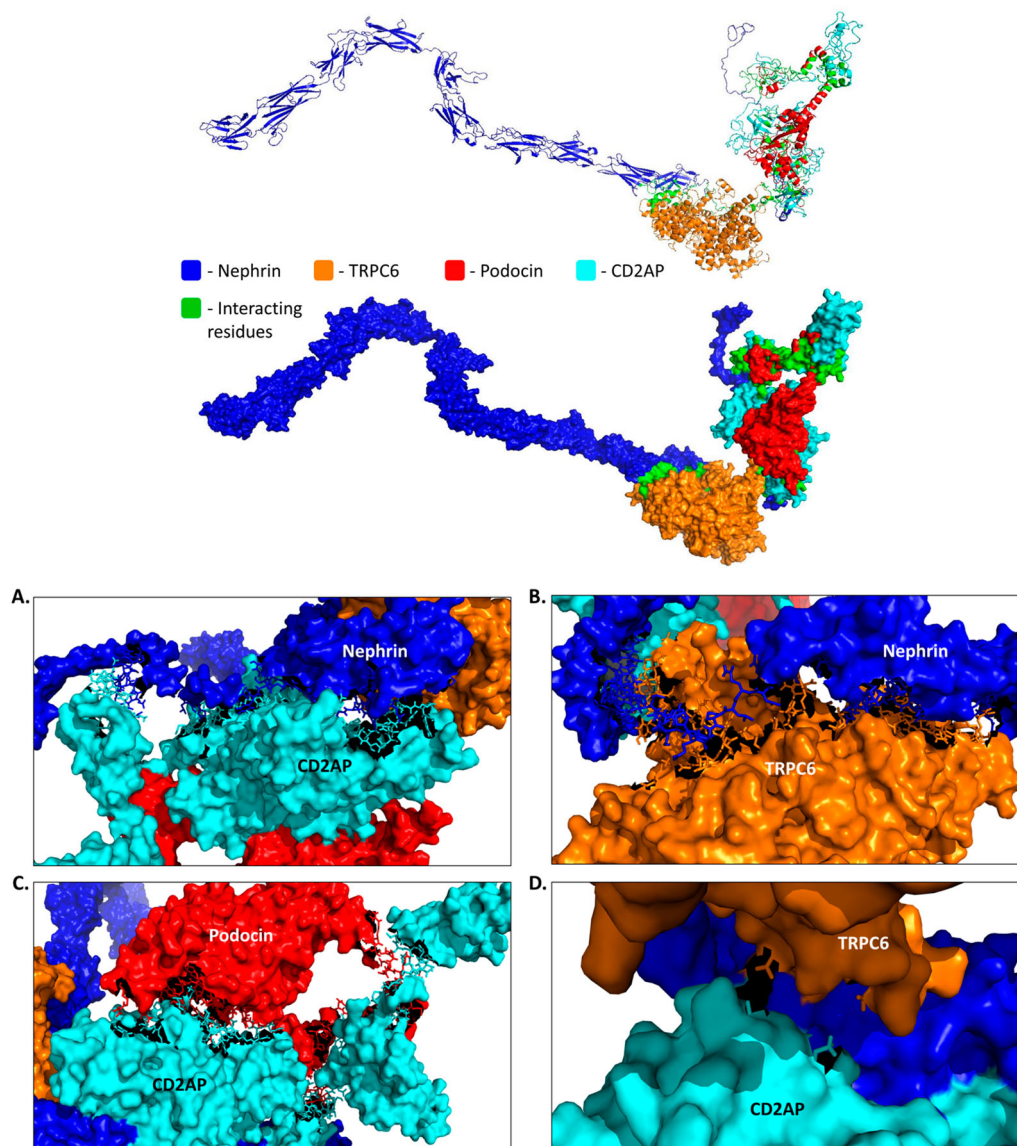


Figure 5. (Colour online) Space fill and ribbon models showing protein-protein docking of nephrin, CD2AP, podocin and TRPC6 models: In the figure, TRPC6 (Orange) and CD2AP (Light blue) are observed to bind at the C-terminus of nephrin (Blue). Further, the N-terminus of CD2AP can also be seen binding to both C-terminal residues of podocin (Red) and TRPC6. The interfacial contacts between interacting proteins are represented by green colour. The insets in the figures are close-up views of each interacting partners and their respective interfacial residues. A) Nephrin-CD2AP, B) Nephrin-TRPC6, C) CD2AP-Podocin, and D) CD2AP-TRPC6.

Table 3. Predicted salt bridge and hydrogen bonds at the protein-protein interfaces of nephrin, CD2AP, podocin and TRPC6 complex. The annotation in parenthesis indicates, C- Charged, P- Polar, and AP- Apolar, and the Gray highlighted residues are predicted to be part of intrinsically disordered regions, whereas the residues in bold and italics are predicted IDR-binding domains.

Type of interaction	(C)	(P)	(AP)
Nephrin-CD2AP interacting residues ΔG: -18 kcal mol⁻¹			
Nephrin	CD2AP		
Salt Bridges:			
HIS 1174 (P)	ASP 286	–	–
ASP 1194 (C)	LYS 436	–	–
Hydrogen bonds:			
ARG 1143 (C)	–	ASN 160	VAL 162
SER 1146 (P)	–	SER 189	GLY 186
GLN 1148 (P)	GLU 258	–	VAL 257, ILE 259
SER 1157 (P)	LYS 261	–	–
TYR 1158 (P)	LYS 261	–	–
SER 1159 (P)	LYS 261	–	–
ARG 1160 (C)	ASP 263	THR 260, THR 262	–
PHE 1162 (AP)	–	THR 262	–
THR 1163 (P)	ASP 263	–	–
HIS 1174 (P)	GLU 285, ASP 286, GLU 317	–	–
ALA 1188 (AP)	LYS 436	–	–
GLY 1190 (AP)	ARG 412	SER 434	–
TYR 1193 (P)	GLU 414	–	–
ASP 1194 (C)	LYS 444	–	–
VAL 1196 (AP)	–	SER 443	VAL 442
GLN 1197 (P)	–	SER 443	–
GLY 1199 (AP)	GLU 440	–	–
Nephrin-TRPC6 interacting residues ΔG: -20.3 kcal mol⁻¹			
Nephrin	TRPC6		
Salt Bridges:			
ASP 1047 (C)	–	HIS 264	–
ARG 1091 (C)	GLU 886	–	–
Hydrogen bonds:			
GLY 853 (P)	ASP 478	–	–
SER 910 (P)	ASP 478	–	–
ALA 911 (AP)	ARG 437, ASP 478	–	–
ALA 912 (AP)	ARG 437, ASP 478	–	–
TYR 915 (P)	ARG 437	ASN 479	–
ASP 942 (C)	–	SER 430	–
ALA 1025 (AP)	–	–	ALA 421
GLY 1028 (AP)	–	–	PHE 419
SER 1042 (P)	ASP 357, ARG 360	–	–
SER 1055 (P)	ASP 250	–	–
GLY 1056 (AP)	ASP 250	–	–
LEU 1062 (AP)	–	–	LEU 899
TYR 1139 (P)	GLU 883	–	–
CD2AP - Podocin interacting residues ΔG: -21.8 kcal mol⁻¹			
CD2AP	Podocin		
Salt Bridges:			
GLU 34 (C)	LYS 377	–	–
ASP 51 (C)	LYS 148	–	–
LYS 579 (C)	GLU 303	–	–
GLU 592 (C)	ARG 306	–	–
Hydrogen bonds:			
MET 1 (AP)	ASP 250	–	–
GLN 33 (P)	ASP 379	–	–
ARG 59 (C)	–	SER 205	–
GLU 60 (C)	–	SER 205	–
LYS 64 (C)	–	SER 86, SER 120	–
SER 399 (P)	–	–	LEU 373
THR 407 (P)	–	–	PRO 369
THR 468 (P)	LYS 368	–	–
LEU 487 (AP)	–	–	LEU 324
THR 490 (P)	–	HIS 325	–
ALA 562 (AP)	ARG 322	–	–
PHE 563 (AP)	ARG 322	–	–
THR 575 (P)	LYS 299	–	–
VAL 578 (AP)	ARG 306	–	–
LYS 579 (C)	GLU 303,	–	–
CD2AP - TRPC6 Interacting residues ΔG : -4.2 kcal mol⁻¹			

3.5. Mutations altered the complex formation

To garner clues on how a mutation in each protein influenced the assembly of the complex, we selected one mutant for each of these four proteins (Nephrin V822M, CD2AP K301M, Podocin R138Q and TRPC6 R895L). We considered these mutations based on their association with the severity of NS [10,11,29,30]. Although these mutations did not fall into the interfacial contacts predicted in the wild-type complex, we would like to elucidate how a mutation influenced the tertiary structure of a protein and compromise protein-protein interactions. We compared the secondary and tertiary structure predictions of each of the mutants with their wild-type predictions using PSI-PRED and PIC server. Our analysis revealed nephrin, CD2AP and TRPC6 mutants did not have any changes in their secondary and tertiary structure predictions at the mutation region. However, the secondary structure predictions for R138Q podocin showed the mutation region as a part of the beta sheet, whereas in the 3D model this region was observed as a loop. This discrepancy could be associated with the podocin's ability to form homo-oligomers like other members of its proteins family. Further, tertiary structure analysis of both mutant and wildtype nephrin, CD2AP, podocin and TRPC6 models revealed that each mutation caused significant alterations to the protein's stereochemistry (supplementary Table 2).

We next docked these mutant proteins with their wild-type neighbours to check aberrations in the macromolecular assembly. In the case of nephrin mutant (V822M), the interactions between nephrin-CD2AP, nephrin-TRPC6 and CD2AP-podocin reduced significantly along with decreased binding propensity, while the CD2AP-TRPC6 interactions increased considerably when compared to the wildtype complex. Interestingly, in the nephrin mutant, we found new interactions between nephrin-podocin and podocin-TRPC6 (Figure 6A & Table 4). In the case of CD2AP (K301M) mutant, the interactions between CD2AP-podocin are compromised, since the podocin bound to the N-terminus of nephrin in a striking contrast to the wild-type complex. Also, we observed that the interaction propensities between nephrin-CD2AP and between CD2AP-TRPC6 increased considerably. Furthermore, due to the K301M CD2AP mutation, only the N-terminus of TRPC6 docked downstream to the FN3 domain of nephrin. This result is in contrast to the wild-type complex wherein, TRPC6 docked to the FN3 domain (Figure 6B & Table 4).

In the podocin mutant (R138Q), the number of interactions and the binding propensities of nephrin-CD2AP and CD2AP-podocin decreased along with complete loss of CD2AP-TRPC6 interactions. Further, in this complex, we noticed TRPC6 docked to the 5th Ig domain of nephrin (Figure 6C & Table 4). The R895L TRPC6 mutation severely distorted the TRPC6 structure, due to which its interactions with nephrin are severely compromised and also caused TRPC6 to interact strongly with CD2AP and vice versa with podocin. Also, the binding propensity of CD2AP-podocin in this complex although remained similar, the interfacial contacts between them changed significantly. We also noticed that the binding

Table 4. Binding energies, number of salt bridges, hydrogen bonds and non-bonded contacts predicted at the protein-protein interfaces of wild type and mutant complexes.

Interactions	Nephrin -CD2AP	Nephrin -podocin	Nephrin -TRPC6	CD2AP -podocin	CD2AP -TRPC6	Podocin - TRPC6
Wild-type complex						
Binding energy ΔG (kcal/mol)	-18.0	-	-20.3	-21.8	-4.2	-
No. of Interface residues	39-48	-	51-41	60-58	2-2	-
Salt Bridges	2	-	2	4	-	-
Hydrogen bonds	30	-	23	18	-	-
Non-bonded contacts	364	-	343	313	2	-
V822M Nephrin complex						
Binding energy ΔG (kcal/mol)	-17.6	-15.7	-11.6	-10.1	-11.5	-5.3
Interface residues	37-38	32-36	15-17	20-21	20-20	8-3
Salt Bridges	2	2	-	2	-	-
Hydrogen bonds	9	7	4	1	6	1
Non-bonded contacts	595	515	210	237	245	69
K301M CD2AP complex						
Binding energy ΔG (kcal/mol)	-26.0	-21.0	-11.9	-	-15.4	-
Interface residues	54-60	41-40	22-26	-	49-41	-
Salt Bridges	2	-	1	-	3	-
Hydrogen bonds	10	4	7	-	8	-
Non-bonded contacts	1219	570	348	-	657	-
R138Q Podocin complex						
Binding energy ΔG (kcal/mol)	-15.4	-20.1	-20.1	-10.4	-	-
Interface residues	34-32	33-43	44-52	17-15	-	-
Salt Bridges	4	2	7	3	-	-
Hydrogen bonds	4	5	14	5	-	-
Non-bonded contacts	413	496	676	227	-	-
R895L TRPC6 complex						
Binding energy ΔG (kcal/mol)	-15.4	-	-9.6	-21.1	-20.5	-4.7
Interface residues	34-32	-	15-15	52-62	42-48	6-5
Salt Bridges	4	-	-	2	2	-
Hydrogen bonds	4	-	2	8	10	1
Non-bonded contacts	413	-	293	921	699	31

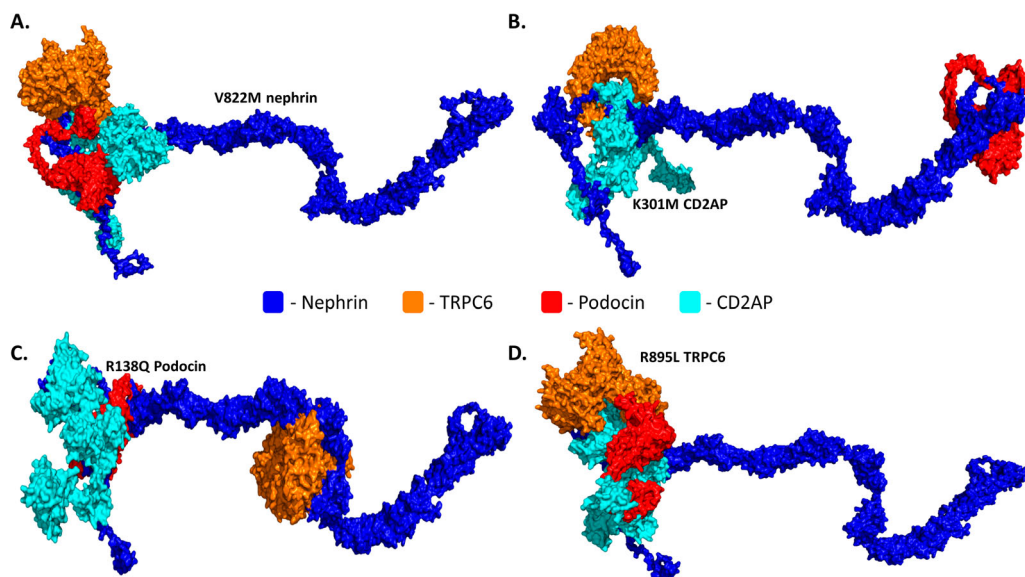
energy between nephrin-CD2AP decreased considerably in the R895L TRPC6 mutant (Figure 6D & Table 4). Our analysis suggests that mutations in each protein significantly distorted the complex assembly.

4. Discussion

The details of the macromolecular assembly of SD proteins and how mutations in each protein abolish interactions and compromise the permselectivity of the SD is of interest to the nephrologists. The present study is aimed to elucidate the coordinated

interactions of the complex through computationally derived models for nephrin, CD2AP, podocin, and TRPC6. Using sequence analysis, docking and molecular dynamic simulations we shed light on the interfacial residues responsible for the complex formation and discuss how the dynamic states of IDRs and IDR-BD in each protein play a crucial role in protein-protein interactions. We also discuss, how mutations in each protein can significantly affect each protein in the complex thereby causing alterations in SD architecture leading to proteinuria.

It is known that IDRs typically involve in the formation of multiprotein complexes and signalling pathways [28]. The

**Figure 6.** (Colour online) Surface models representing each mutant complex: A) V822M nephrin mutant complex, B) K301M CD2AP mutant complex, C) R138Q podocin mutant complex and D) R895L TRPC6 mutant complex.

flexibility offered by the IDRs allows the sequence to adopt diverse conformations to bind to its receptor or ligand [31]. Both the IDR prediction servers identified IDRs in all these proteins. Our results suggest that IDRs/IDR-BD predicted in the Ig domains of adjacent nephrin molecules may aid in interactions facilitating the formation of dimers thus, imparting the characteristic zipper-like structure to the SD. Besides being involved in the formation of homodimers, our results show that the IDRs and IDR-BD of nephrin also involve in the formation of the heteromeric complex (nephrin-CD2AP-podocin-TRPC6).

To delineate the role of IDRs in the formation of heteromeric complex, we performed docking and simulations studies with nephrin, CD2AP, podocin and TRPC6 models. From our predictions, it is evident that most of the interacting residues between nephrin, CD2AP, podocin, and TRPC6 are either IDRs or IDR-BD. An earlier study showed that the 3rd SH3 domain of CD2AP interacts with nephrin and it was speculated that the C-terminal Tyr¹¹⁷⁶, Tyr¹¹⁹³ and Tyr¹²¹⁰ of nephrin might act as possible docking sites for the SH3 domains [13,14]. Our results suggest that non-IDR Tyr¹¹⁷⁶ and Tyr¹¹⁹³ residues of nephrin interacted with non-IDR residues 313–316 of the 3rd SH3 domain and IDR residues 412,414,430,437,441–447 of CD2AP. Therefore, our observations are consistent with the experimental observations. Further, our studies also show that the 2nd SH3 domain of CD2AP is involved in nephrin-CD2AP interactions. It was reported earlier that CD2AP and the C-terminal 81 residues of nephrin bind to C-terminus of podocin [13]. Although we were able to predict IDRs/IDR-BD in nephrin and podocin, our docking method could not predict nephrin-podocin interactions in the complex. We speculate that nephrin may independently bind to another podocin and not necessarily interact with the podocin bound to CD2AP; this condition may be possible as podocin is known to form homooligomers at the SD [29]. On the other hand, it is known that SH3 domains possess an affinity for proline-rich regions (PXXP) [32]. Sequence analysis revealed that residues 362–376 of podocin are proline-rich and our results indicate that IDR residues 398–408, 468 and 482 of CD2AP interact with 362, 364, 369 and 372 IDR-BD prolines of podocin. These results confirm the affinity of SH3 towards proline residues and also suggest the important role of IDRs in protein–protein interactions. Further, our results also show that all the three SH3 domains of CD2AP interacted with podocin.

The earliest evidence of a macromolecular assembly came from co-immunoprecipitation experiments which showed that TRPC6 interacts with nephrin and podocin [12]. Our results suggest that TRPC6 interacted with the C-terminal residues of nephrin. Interestingly, the docking studies also predicted CD2AP-TRPC6 interactions in the complex with significantly less binding energy when compared to its interacting neighbours. We speculate that these weak TRPC6-CD2AP interactions may elicit functions that are yet to be comprehended. Nevertheless, our docking method was unable to predict interactions between TRPC6 and podocin, a condition that may be similar to the missing nephrin-podocin interactions as stated above. All these results clearly suggest the involvement of

IDRs/IDR-BD in facilitating interactions between nephrin, CD2AP, podocin, and TRPC6.

A mutation in each protein was selected from the human gene mutation database [33]. V822M nephrin mutant formed small distinct puncta along the plasma membrane as opposed to the large distinct puncta formed by the wild-type protein [30]. The V822M mutation in nephrin does not fall into the predicted interfacial contacts but falls into the 8th Ig-like-C2 domain, which is the extracellular segment of the protein. According to our results, the V822M mutation induced morphological changes in nephrin and affecting its ability to appropriately interact with CD2AP, podocin and TRPC6. From our results, we suggest that this mutation may also disrupt the nephrin-nephrin homodimer formation. These alterations may perhaps explain the small distinct nephrin puncta due to V822M mutation. The K301M mutation in CD2AP was shown to disrupt interactions between CD2AP and its neighbours leading to the early onset of SRNS [10]. K301M is a non-conservative mutation that falls into the non-interacting residues of the third SH3 domain. We show that mutation of lysine to methionine in CD2AP drastically alters the CD2AP structure and its ability to form a complex with the tyrosine residues of nephrin and also with proline-rich regions of podocin. All these changes may ultimately lead to compromised complex formation thus disturbing the SD architecture. The R138Q podocin mutant localised at the endoplasmic reticulum and interferes with the recruitment of nephrin to the lipid rafts [29]. In our complex model, the R138 residue of podocin was observed to be buried inside the protein. Although, both the arginine and glutamine have similar charge groups, their side chains differ considerably. From our analysis, it is evident that the R138Q mutation altered the essential interactions required for maintaining the tertiary structure of podocin, which may affect its scaffolding ability and its potential to recruit nephrin and CD2AP to the SD leading to alterations in SD architecture. The R895 residue in TRPC6 is an evolutionarily conserved and the mutation R895L caused NS at an early age of 1 year [11]. Further, this mutation caused enhanced calcium influx into the podocyte disrupting the podocyte cytoskeleton [11]. In our wild-type complex, the positively charged R895 residue in TRPC6 interacted with the leu¹⁰⁶², pro¹⁰⁶⁴, and leu¹⁰⁶⁶ of the transmembrane region of nephrin forming charged–charged, charged–polar and charged–apolar contacts. Mutation of R895 to non-polar leucine in TRPC6, severely changed its structure thus losing the above interactions and also causing significant loss of interfacial contacts between nephrin and TRPC6. This loss of contacts may induce conformational changes in cytoskeleton bound nephrin, which consequently affects the SD architecture leading to proteinuria. Also, from our results we speculate the alteration of the TRPC6 structure may lead to protein aggregation since, TRPC6 showed interactions with podocin and CD2AP in the TRPC6 mutant complex.

In conclusion, we report that IDRs and IDR-BD play a major role in mediating the macromolecular assembly of nephrin, CD2AP, podocin, and TRPC6. Further, we suggest that IDRs and IDR-BD may also help in the homodimer interactions between adjacent nephrin molecules imparting a zipper-like structure to the SD. We speculate that mutations in

any of these proteins can drastically alter the proteins structures or lead to the loss of intricate contacts with its client proteins causing distortions to the complex.

Acknowledgements

We thank the SERB, India for providing research grants (EMR/2015/002076 to AKP and also to Indian council of medical research (3/1/2(4)/Nephro/2015/NCD-II) for providing fellowship to Mulukala NSK. We acknowledge the resources provided by Biotechnology Infrastructure Facility of the University of Hyderabad and also the structural biology laboratory at BITS-Pilani Hyderabad.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Indian Council of Medical Research: [Grant Number 3/1/2(4)/Nephro/2015/NCD-II]; Science and Engineering Research Board: [Grant Number EMR/2015/002076].

Conflict of Interest

The authors declare no conflicts of interest.

References

- [1] Anil Kumar P, Welsh GI, Saleem MA, et al. Molecular and cellular events mediating glomerular podocyte dysfunction and depletion in diabetes mellitus. *Front Endocrinol (Lausanne)*. 2014;5(151). doi:10.3389/fendo.2014.00151. PubMed PMID: 25309512; PubMed Central PMCID: PMC4174857.
- [2] Patari-Sampo A, Ihalmio P, Holthofer H. Molecular basis of the glomerular filtration: nephrin and the emerging protein complex at the podocyte slit diaphragm. *Ann Med*. 2006;38(7):483–492. doi:10.1080/07853890600978149. PubMed PMID: 17101539.
- [3] Mukhi D, Nishad R, Menon RK, et al. Novel Actions of Growth Hormone in podocytes: Implications for Diabetic Nephropathy [Review]. *Front Med (Lausanne)*. 2017-July-12;4(102) English. doi:10.3389/fmed.2017.00102.
- [4] Mulukala SK, Nishad R, Kolligundla LP, et al. In silico structural characterization of podocin and assessment of nephrotic syndrome-associated podocin mutants. *IUBMB Life*. 2016 Jul;68(7):578–588. doi:10.1002/iub.1515. PubMed PMID: 27193387.
- [5] Sadowski CE, Lovric S, Ashraf S, et al. A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. *J Am Soc Nephrol: JASN*. 2015 Jun;26(6):1279–1289. doi:10.1681/ASN.2014050489. PubMed PMID: 25349199; PubMed Central PMCID: PMC4446877.
- [6] Kestila M, Lenkkeri U, Mannikko M, et al. Positionally cloned gene for a novel glomerular protein–nephrin–is mutated in congenital nephrotic syndrome. *Mol Cell*. 1998 Mar;1(4):575–582. PubMed PMID: 9660941.
- [7] Benoit G, Machuca E, Antignac C. Hereditary nephrotic syndrome: a systematic approach for genetic testing and a review of associated podocyte gene mutations. *Pediatr Nephrol*. 2010 Sep;25(9):1621–1632. doi:10.1007/s00467-010-1495-0. PubMed PMID: 20333530; PubMed Central PMCID: PMC2908444.
- [8] Boute N, Gribouval O, Roselli S, et al. NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome. *Nat Genet*. 2000 Apr;24(4):349–354. doi:10.1038/74166. PubMed PMID: 10742096.
- [9] Hinkes B, Vlangos C, Heeringa S, et al. Specific podocin mutations correlate with age of onset in steroid-resistant nephrotic syndrome. *J Am Soc Nephrol*. 2008 Feb;19(2):365–371. doi:10.1681/ASN.2007040452. PubMed PMID: 18216321; PubMed Central PMCID: PMC2396749.
- [10] Gigante M, Pontrelli P, Montemurno E, et al. CD2AP mutations are associated with sporadic nephrotic syndrome and focal segmental glomerulosclerosis (FSGS). *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*. 2009 Jun;24(6):1858–1864. doi:10.1093/ndt/gfn712. PubMed PMID: 19131354.
- [11] Gigante M, Caridi G, Montemurno E, et al. TRPC6 mutations in children with steroid-resistant nephrotic syndrome and atypical phenotype. *J Am Soc Nephrol*. 2011 Jul;6(7):1626–1634. doi:10.2215/CJN.07830910. PubMed PMID: 21734084.
- [12] Dryer SE, Reiser J. TRPC6 channels and their binding partners in podocytes: role in glomerular filtration and pathophysiology. *Am J Phys Renal Physiol*. 2010 Oct;299(4):F689–F701. doi:10.1152/ajprenal.00298.2010. PubMed PMID: 20685822; PubMed Central PMCID: PMC2957253.
- [13] Huber TB, Kottgen M, Schilling B, et al. Interaction with podocin facilitates nephrin signaling. *J Biol Chem*. 2001 Nov 09;276(45):41543–6. doi:10.1074/jbc.C100452200. PubMed PMID: 11562357.
- [14] Palmen T, Lehtonen S, Ora A, et al. Interaction of endogenous nephrin and CD2-associated protein in mouse epithelial M-1 cell line. *J AM SOC NEPHROL: JASN*. 2002 Jul;13(7):1766–1772. PubMed PMID: 12089372.
- [15] Schwarz K, Simons M, Reiser J, et al. Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin. *J Clin Invest*. 2001 Dec;108(11):1621–1629. doi:10.1172/JCI12849. PubMed PMID: 11733557; PubMed Central PMCID: PMC200981.
- [16] Shih NY, Li J, Cotran R, et al. CD2AP localizes to the slit diaphragm and binds to nephrin via a novel C-terminal domain. *Am. J. Pathol*. 2001 Dec;159(6):2303–2308. doi:10.1016/S0002-9440(10)63080-5. PubMed PMID: 11733379; PubMed Central PMCID: PMC1850607.
- [17] Nugent T, Jones DT. Transmembrane protein topology prediction using support vector machines [journal article]. *BMC Bioinformatics*. 2009 May 26;10(1):159. doi:10.1186/1471-2105-10-159.
- [18] Jones DT, Taylor WR, Thornton JM. A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry*. 1994 Mar 15;33(10):3038–3049. PubMed PMID: 8130217.
- [19] Jones DT, Cozzetto D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*. 2015 Mar 15;31(6):857–863. doi:10.1093/bioinformatics/btu744. PubMed PMID: 25391399; PubMed Central PMCID: PMC4380029.
- [20] Yao B, Zhang J, Dai H, et al. Solution structure of the second SH3 domain of human CMS and a newly identified binding site at the C-terminus of c-Cbl. *Biochim Biophys Acta*. 2007 Jan;1774(1):35–43. doi:10.1016/j.bbapap.2006.09.018. PubMed PMID: 17188587.
- [21] Moncalian G, Cardenas N, Deribe YL, et al. Atypical polyproline recognition by the CMS N-terminal Src homology 3 domain. *J Biol Chem*. 2006 Dec 15;281(50):38845–53. doi:10.1074/jbc.M606411200. PubMed PMID: 17020880.
- [22] Ortega Roldan JL, Romero Romero ML, Ora A, et al. The high resolution NMR structure of the third SH3 domain of CD2AP. *J Biomol NMR*. 2007 Dec;39(4):331–336. doi:10.1007/s10858-007-9201-7. PubMed PMID: 17922258.
- [23] Tina KG, Bhadra R, Srinivasan N. PIC: protein interactions Calculator. *Nucleic Acids Res*. 2007 Jul;35(Web Server issue):W473–W476. doi:10.1093/nar/gkm423. PubMed PMID: 17584791; PubMed Central PMCID: PMC1933215.
- [24] Duhovny D, Nussinov R, Wolfson HJ. Efficient Unbound docking of rigid molecules. In: Guigó R, Gusfield D, editor. *Algorithms in Bioinformatics: Second International Workshop, WABI 2002 Rome, Italy, September 17–21, 2002 Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2002. p. 185–200.
- [25] Schneidman-Duhovny D, Inbar Y, Nussinov R, et al. Patchdock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*. 2005 Jul 01;33(Web Server issue):W363–W367. doi:10.1093/nar/gki481. PubMed PMID: 15980490; PubMed Central PMCID: PMC1160241.

- [26] Kirubakaran P, Karthikeyan M, Singh Kh D, et al. In silico structural and functional analysis of the human TOPK protein by structure modeling and molecular dynamics studies. *J Mol Model*. 2013 Jan;19(1):407–419. doi:10.1007/s00894-012-1566-1. PubMed PMID: 22940854.
- [27] Kirubakaran P, Muthusamy K, Dhanachandra Singh K, et al. Homology modeling, molecular dynamics, and molecular docking studies of *Trichomonas vaginalis* carbamate kinase [journal article]. *Medicinal Chemistry Research*. 2012 August 01;21(8):2105–2116. doi:10.1007/s00044-011-9719-9.
- [28] Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Bio*. 2015 Jan;16(1):18–29. doi:10.1038/nrm3920. PubMed PMID: 25531225; PubMed Central PMCID: PMC4405151.
- [29] Huber TB, Simons M, Hartleben B, et al. Molecular basis of the functional podocin-nephrin complex: mutations in the NPHS2 gene disrupt nephrin targeting to lipid raft microdomains. *Hum Mol Genet*. 2003 Dec 15;12(24):3397–3405. doi:10.1093/hmg/ddg360. PubMed PMID: 14570703.
- [30] Shono A, Tsukaguchi H, Kitamura A, et al. Predisposition to relapsing nephrotic syndrome by a nephrin mutation that interferes with assembly of functioning microdomains. *Hum Mol Genet*. 2009 Aug 15;18(16):2943–2956. doi:10.1093/hmg/ddp232. PubMed PMID: 19443487.
- [31] Collins MO, Yu L, Campuzano I, et al. Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Molecular & Cellular Proteomics: MCP*. 2008 Jul;7(7):1331–1348. doi:10.1074/mcp.M700564-MCP200. PubMed PMID: 18388127.
- [32] Mayer BJ. SH3 domains: complexity in moderation. *J Cell Sci*. 2001 Apr;114(7):1253–1263. PubMed PMID: 11256992.
- [33] Stenson PD, Ball EV, Mort M, et al. Human gene mutation database (HGMD): 2003 update. *Hum Mutat*. 2003 Jun;21(6):577–581. doi:10.1002/humu.10212. PubMed PMID: 12754702.



Structural features and oligomeric nature of human podocin domain

Sandeep K.N. Mulukala^a, Shivkumar S. Irukuvajjala^b, Krishan Kumar^c, Kanchan Garai^d,
Pannuru Venkatesu^c, Ramakrishna Vadrevu^b, Anil K. Pasupulati^{a,*}

^a Department of Biochemistry, School of Life Sciences, University of Hyderabad, Hyderabad, 500046, India

^b Department of Biological Sciences, Birla Institute of Technology and Sciences, Pilani Hyderabad Campus, Hyderabad, 500078, India

^c Department of Chemistry, University of Delhi, New Delhi, 110 007, India

^d Tata Institute of Fundamental Research, Hyderabad, 500019, India

ARTICLE INFO

Keywords:

Podocin

Podocyte

Nephrotic syndrome

Slit-diaphragm

Proteinuria

ABSTRACT

Podocytes are crucial cells of the glomerular filtration unit and plays a vital role at the interface of the blood-urine barrier. Podocyte slit-diaphragm is a modified tight junction that facilitates size and charge-dependent permselectivity. Several proteins including podocin, nephrin, CD2AP, and TRPC6 form a macromolecular assembly and constitute the slit-diaphragm. Podocin is an integral membrane protein attached to the inner membrane of the podocyte via a short transmembrane region (101–125). The cytosolic N- and C-terminus help podocin to attain a hook-like structure. Podocin shares 44% homology with stomatin family proteins and similar to the stomatin proteins, podocin was shown to associate into higher-order oligomers at the site of slit-diaphragm. However, the stoichiometry of the homo-oligomers and how it partakes in the macromolecular assemblies with other slit-diaphragm proteins remains elusive. Here we investigated the oligomeric propensity of a truncated podocin construct (residues:126–350). We show that the podocin domain majorly homo-oligomerizes into a 16-mer. Circular dichroism and fluorescence spectroscopy suggest that the 16-mer oligomer has considerable secondary structure and moderate tertiary packing.

1. Introduction

Vertebrate kidneys regulate electrolyte and water balance to maintain body homeostasis. Each kidney is composed of about one million nephrons. The glomerulus and the renal tubule are the two major parts of a nephron that work in unison to ensure ultra-filtrated urine. The glomerular filtration barrier (GFB) offers permselectivity for the filtration of plasma components into the urine. The GFB consists of fenestrated capillary endothelium, glomerular basement membrane, and podocytes [1]. Podocytes are highly differentiated visceral epithelial cells that encase the glomerular capillaries. A typical podocyte cell consists of a protuberant cell body with primary processes made of actin and microtubules. The primary process further branches into secondary foot processes which interlaces with the neighboring foot processes forming a modified tight junction called the slit-diaphragm (SD) [2]. The SD is a negatively charged zipper-like structure bridging the 30–40nm gap between the adjacent foot processes. This structure curbs the

passage of albumin and other large molecules from the blood into primary urine thereby tightly regulating the composition of the glomerular filtrate [3,4]. The intricate structure of the SD is maintained by an array of protein assemblies. Proteins such as podocin, nephrin, CD-2 associated protein (CD2AP), transient receptor potential cation channel subfamily C member 6 (TRPC6), zonula occludens-1 (ZO-1), and Nephrin-like proteins 1, 2, and 3 (NEPH) interact to form macromolecular complexes and constitute the structure of SD [5–8].

Mutations in the SD proteins are associated with nephrotic syndrome (NS), which is presented with massive proteinuria, hypoalbuminemia, and edema [9]. Corticoid therapy is the usual recourse to abate NS and based on the patient's response to corticoid therapy NS is divided into steroid-sensitive NS (SSNS) and steroid-resistant NS (SRNS). Patients with congenital nephropathies usually belong to the SRNS group since they do not respond to corticoid therapy. Congenital nephropathy typically onsets in infants during 0–3 months of age which eventually progresses to irreversible kidney failure within a decade. Mutations in

Abbreviations: SD, slit-diaphragm; GFB, Glomerular filtration barrier; CD2AP, CD-2 associated protein; TRPC6, Transient receptor potential cation channel subfamily C member 6; ZO-1, Zonula occludens-1; NEPH, Nephrin-like protein; NS, Nephrotic syndrome; SRNS, steroid-resistant NS; *NPHS1* & 2, Nephrotic syndrome-type I and type II; IDRs, Intrinsically disordered regions; SEC, Size-exclusion chromatography; MALS, multi-angle light scattering; CD, Circular dichroism.

* Corresponding author. F73B, School of Life Sciences, University of Hyderabad, Gachibowli, Hyderabad, 500046, India.

E-mail address: pasupulati.anilkumar@gmail.com (A.K. Pasupulati).

<https://doi.org/10.1016/j.bbrep.2020.100774>

Received 15 April 2020; Received in revised form 31 May 2020; Accepted 2 June 2020

Available online 25 June 2020

2405-5808/© 2020 The Authors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

NPHS1 and *NPHS2* that encode for nephrin and podocin respectively result in the majority of congenital SRNS cases [10]. About 18% of the reported SRNS cases are due to mutations in podocin [11,12]. Mutations in other SD proteins such as TRPC6 and CD2AP have also been observed to cause NS but at a less frequent rate than nephrin and podocin [13–15].

Podocin is a 383 amino acid protein localizing to the lipid rafts along with other SD proteins [7,16,17]. Podocin shares 44% homology and several structural similarities with stomatin family proteins due to the presence of a highly conserved Prohibitin (PHB) domain [5,18]. Podocin adapts a hook-like structure with its cytoplasmic N- and C- terminus as it attaches to the inner side of the plasma membrane via a transmembrane domain at 100–125 residues [5,12,19]. Structural characterization of stomatin revealed that different truncations of the protein associated with different oligomeric states and the C-terminus of the protein is crucial for homo-oligomerization [20–22]. Studies with truncated C-terminal human podocin revealed that it forms a dimer [23]. Though it was proposed that longer constructs of podocin were capable of associating into higher-order oligomers, it was never demonstrated [23]. Additionally, co-immunoprecipitation studies with nephrin, CD2AP, TRPC6, and NEPH1 indicated that these proteins interact majorly with the C-terminus of podocin [6–9,16,17,24]. We reported earlier that these interactions are mediated by the intrinsically disordered regions (IDRs) present in these proteins [18].

Although a large body of evidence suggests a central role for the podocin in the SD protein complex, the precise mechanism by which podocin oligomerizes and acts as a scaffolding molecule remains to be elucidated. Importantly, it is not known whether all the other SD proteins interact with a single podocin or with homo-oligomers? Therefore, in this study, we attempted to understand the oligomeric nature of protein using a truncated construct (residues: 126–350), which encompasses the PHB domain, C-terminal oligomerization site, and 4 out of the 6 cysteines present in the native podocin sequence.

2. Material and methods

2.1. Protein cloning, expression, and purification

The codon-optimized podocin gene (1152bp) was purchased from Gene Art (Life Technologies, USA). The regions encoding the amino acids 126–350 (podocin domain) was amplified with the primers 5' CCC GAA TTC G AAA GTG GTG CAA GAA 3' (forward) and 5' GAA CTC GAG CAG ACA ATT CAG CAG ATC 3' (reverse) and cloned into pET22b at *EcoRI/XhoI* sites. The recombinant construct was transformed into Arctic express (DE3) competent cells (Agilent Technologies, USA). The transformed cells were grown at 37 °C in LB media supplemented with 100 µg/ml ampicillin and 10 µg/ml gentamycin. Protein expression was induced with 0.2 mM IPTG and cultured further for 16hrs at 14 °C. The cells were harvested by centrifugation (13300×g, 20 min, and 4 °C) and sonicated in the opening buffer (50 mM potassium phosphate (pH 8.0), 0.3 M NaCl, 5 mM β-mercaptoethanol and 0.1% Triton X-100) followed by clarification by centrifugation at 18000×g for 45 min at 4 °C. The inclusion bodies were solubilized in 50 mM potassium phosphate (pH 8.0), 0.3 M NaCl, 5 mM β-mercaptoethanol and 8 M Urea followed by clarification at 18000×g for 1h at room temperature. The solubilized protein was then purified using Ni-NTA agarose (Qiagen). The purity was confirmed (>98%) on a 12% SDS-PAGE. From the SDS-PAGE gel, the band corresponding to the podocin domain (27 kDa) was excised and subjected to tryptic digestion and analyzed using MALDI-TOF/TOF (Bruker Autoflex III smart beam, Bruker Daltonics, Bremen, Germany) to confirm the protein sequence.

The purified protein was renatured by rapid dilution at 1:10 into 10mM potassium phosphate buffer with 150 mM NaCl and 2 mM β-mercaptoethanol (pH 8.0) followed by dialysis in the same buffer to remove traces of urea and imidazole. The same pH and buffer composition are uniformly used in all the subsequent experiments. An

extinction coefficient of the 12950 M⁻¹ cm⁻¹ was used for determining the protein concentration on Jasco V-630 UV–Vis spectrophotometer.

2.2. Size exclusion chromatography and multi-angle light scattering analysis (SEC-MALS)

We performed SEC-MALS to estimate the oligomeric nature of the podocin domain. SEC-MALS was performed at room temperature by passing 500 µl protein (12µM) at 0.3 ml/min flow rate through a Superdex S200 SEC column (GE Healthcare) pre-equilibrated with 10mM potassium phosphate buffer supplemented with 150mM NaCl and 2mM β-mercaptoethanol (pH 8.0). This column was attached to the MALS system (AF2000- Postnova) for analyzing the molar mass of the protein. The protein sample from SEC-MALS was next passed through the flow cell equipped Zetasizer Nano ZS90 dynamic light scattering (DLS) device (Malvern Instruments Ltd, UK) equipped with a 4 mW He–Ne laser. The backscattering was measured at 173 nm for analyzing the polydispersity index (PDI) and the hydrodynamic radius of the protein.

2.3. Fluorescence spectroscopy

Intrinsic tryptophan fluorescence was measured using Jasco FP-6300 (Japan) equipped with an intense xenon flash lamp as the light source. 287 nm was used as the excitation wavelength and the 300–450 nm spectral range was used for obtaining the emission spectrum of the sample (12 µM). For accessing the stability of the podocin domain thermal-induced unfolding was performed. Fluorescence emission at 335 nm as a function of increasing temperature was recorded in triplicates at a bandwidth of 2.5 nm and a scan speed of 200 nm/min for every degree rise in temperature. All the spectra were buffer corrected. The effect of temperature on the protein stability was analyzed by plotting fluorescence intensity at 335 nm using Origin (pro)-version2020b (Origin Lab Corporation, Northampton, MA).

2.4. Circular dichroism (CD) spectroscopy

The CD spectroscopy measurements were recorded on Jasco J-1500 spectropolarimeter (Japan) equipped with a thermoelectric cell holder. The Far-UV (260–195 nm) CD measurements of podocin domain (12 µM) were recorded using a 0.2 cm path length cell at 2.5 nm bandwidth and a scan speed of 50 nm/min. The Near-UV CD measurements were also recorded for the protein sample at 25 µM concentration using a 0.5cm pathlength cell at the bandwidth of 2.5nm and a scan speed of 100 nm/min. Both for far-UV and near-UV the data was recorded in triplicates. To assess the effect of temperature and thus the stability of the domain, the sample was subjected to a steady increase in temperature, and spectra were recorded at an interval of 5 °C over a spectral range of 200 nm–250 nm. The data were plotted using origin lab software after buffer correction.

2.5. Calorimetric analysis

The transition temperature (T_g) of the podocin domain was obtained from measurements using NANO DSC (TA Instruments, USA). Sample containing protein concentration of 12 µM and a volume of 0.650 µL was loaded into the sample capillary and change in heat flow was recorded against reference buffer at a constant pressure of 3 atm, over a temperature range of 293 K–368 K, with a scan rate of 1 K/min and a 300-sec cell equilibration time. Buffer scans were first performed before loading protein for baseline reproducibility. The obtained data was buffer corrected and the analysis data was plotted for peak integration with the peak analyzer option in Origin pro 2020b software. From the peaks, the T_g and the enthalpy of transition (ΔH_{cal}) were calculated.

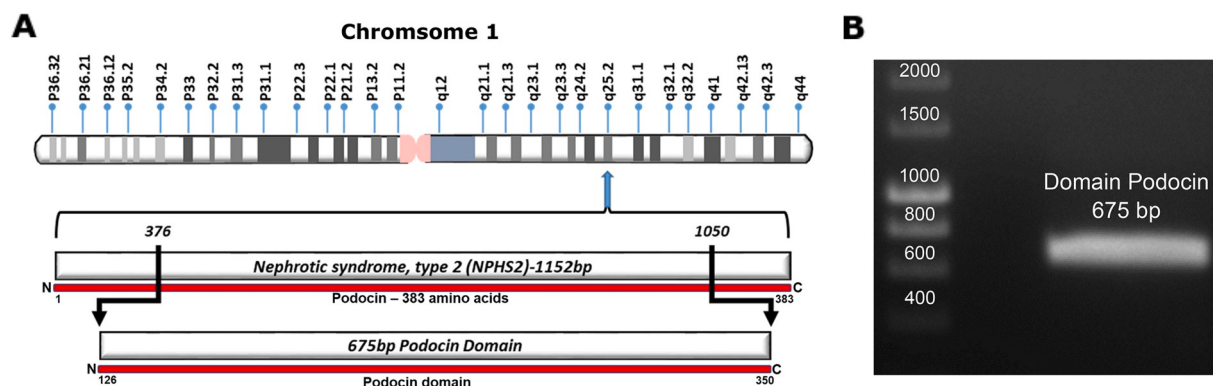


Fig. 1. Cloning of podocin domain: A. The NPHS2 gene that encodes podocin is located on chromosome 1 at the locus q25.2. B. The region 376bp – 1050bp of the gene was amplified and cloned into pET22b bacterial expression vector.

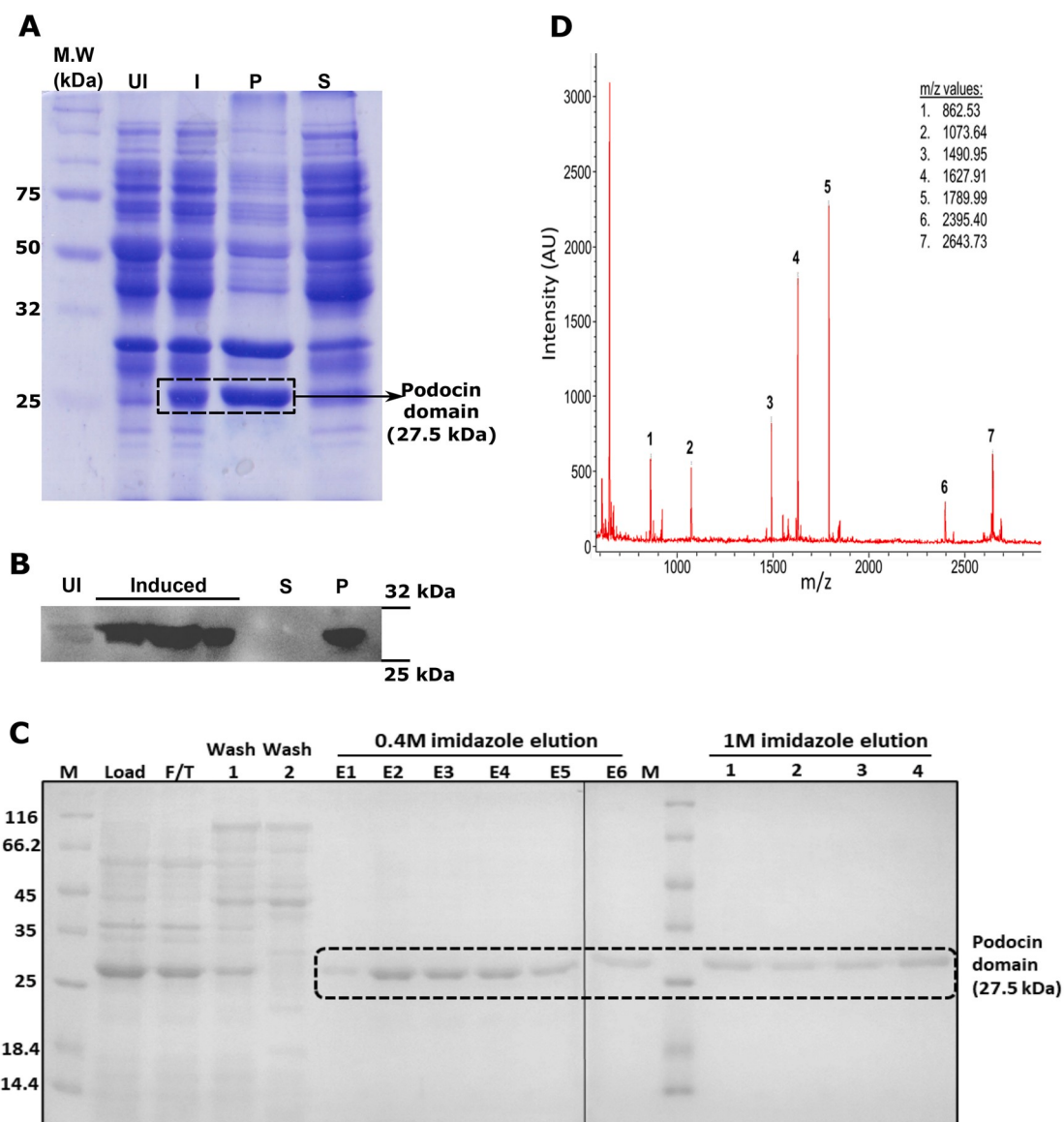


Fig. 2. Expression, and purification of podocin domain: A & B. Coomassie blue staining and immunoblotting with His-Tag HRP-conjugated antibody were done to identify the expression and solubility of the podocin domain. C. The SDS-PAGE analysis of samples after affinity chromatography purification of the urea solubilized cell lysate. D. Trypsinization and MALDI-TOF/TOF analysis of the 27KDa band excised from the earlier SDS-PAGE gel confirmed the presence of the podocin domain. UI-uninduced culture, I-Isopropyl β , D - thiogalactopyranoside Induced culture, P-Pellet fraction, S-Soluble fraction, F/T-flow through, E1-E6: elution fractions with 0.4 M imidazole, 1–4: elution fractions with 1 M imidazole. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1

MALDI TOF/TOF analysis of the purified protein: Trypsinization of the purified band at 27 kDa and subsequence analysis by MALDI-TOF/TOF showed 5 peptide sequences. BLAST analysis of these sequences against the non-redundant proteins database of NCBI showed 100% similarity with the human podocin sequence. Note: The black triangle indicates the site of digestion by trypsin at arginine and lysine residues in the sequence.

Sequence		Observed	Mr (expt.)	Mr (calc.)	Expect	Peptide
Start	End					
134	146	1490.95	1489.95	1489.93	2.3e-03	R $\blacktriangle$ VIIIFRLGHLPLGR $\blacktriangledown$ A
149	168	2395.40	2394.40	2394.22	9.1e-03	K $\blacktriangle$ GPGLFFFLPCLDTHKVDLR $\blacktriangledown$ L
263	286	2643.73	2642.72	2642.45	6.1e-04	R $\blacktriangle$ IEIKDVRLPAGLQHSLAVEAEAR $\blacktriangledown$ Q
270	286	1789.99	1788.99	1788.95	2.4e-06	R $\blacktriangle$ LPAGLQHSLAVEAEAR $\blacktriangledown$ Q
307	322	1627.91	1626.90	1626.88	3.5e-06	R $\blacktriangle$ MAEILSGTPAAVQLR $\blacktriangledown$ Y

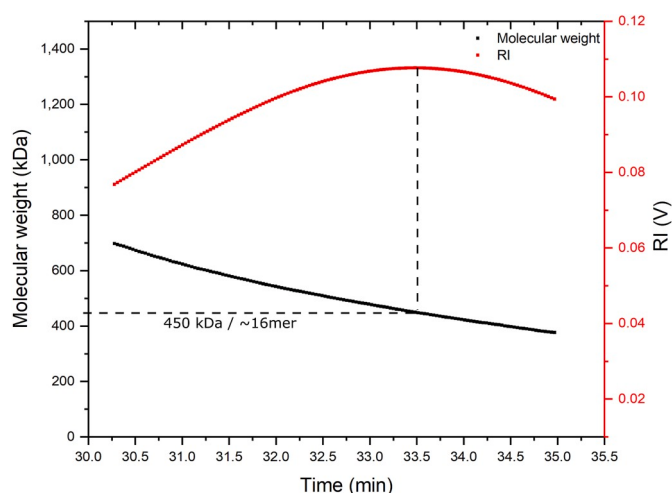


Fig. 3. Podocin domain forms higher-order oligomers: SEC-MALS analysis of the podocin domain for molecular mass determination. The molecular weight on Y-axis and the refractive index on the secondary axis and were plotted against elution time.

3. Results

3.1. Purification of podocin domain

The region from 376bp to 1050bp of human *NPHS2* gene encoding 126–350 amino acid residues of podocin was PCR amplified and cloned into the pET22b expression vector (Fig. 1A&B). IPTG induced expression of the construct resulted in the protein to form inclusion bodies which were confirmed by Coomassie blue staining and immunoblotting with anti-His antibody (Fig. 2A&B). The protein was purified from inclusion bodies by Ni-NTA affinity chromatography after solubilization in 8 M urea (Fig. 2C). Tryptic digestion of the band corresponding to the podocin domain from the SDS-PAGE gel (Fig. 2C-lane E2) and subsequent analysis of the digested products by MALDI-TOF/TOF revealed five peptide fragments (Fig. 2D and Table 1). NCBI BLAST search of these peptide sequences against the non-redundant database confirmed the purified protein as human podocin covering the region 126–350 amino acids.

3.2. Oligomeric nature of the podocin domain

It was reported that stomatin family members exist as homo-

oligomers [20–22]. Since podocin shares significant homology with stomatin, we analyzed the oligomeric nature of the podocin which includes the PHB domain. SEC-MALS data is represented as a combinatorial plot of refractive index, and molecular weight versus elution time (Fig. 3). A maximum refractive index value of 0.11 was observed which corresponds to a molecular weight of 450kDa, suggesting that the podocin domain is a 16-mer oligomer (monomer = 27.5kDa; therefore; 450 kDa/27.5 kDa = ~16-mer). In addition to the predominant 16-mer species, other oligomeric conformations of the podocin domain ranging from 25-mer (refractive index: 0.08, molecular weight: 697kDa) to 13-mer (refractive index: 0.10, molecular weight: 376kDa) were also observed, but to a lesser extent. The DLS, which was in tandem with the SEC-MALS analyzed the eluted samples for hydrodynamic radius and polydispersity. A hydrodynamic radius range of 13.39 - 9.37 nm corresponding to elutions from 25-mer to 13-mer was observed respectively (Fig. 4A–E). The average PDI of the sample was found to be 0.14, which suggests that the sample consists of one major population by volume, however, a broad size range within the population implies the presence of different size species. From these results, it is evident that the major species the podocin domain is a 16-mer while a minor population of other higher-order oligomers also appears to exist in solution.

3.3. Structural features of the homo-oligomers of podocin domain

The CD spectra in the far UV region (250–195 nm) provides information on the secondary structure while the near UV region (310–250 nm) on the tertiary packing. The far UV CD signal for the truncated podocin domain indicates that the protein adopts secondary structure (Fig. 5A). The shape of the spectrum suggests the presence of both α -helices and β -sheet structures. The near UV spectrum of the sample shows a broad peak at 280nm typically contributed by the tryptophan residue (W256), and a minor peak from ~250 – 260 nm could be contributed by five-tyrosine and eight-phenylalanine residues (Fig. 5B). The CD data in the near UV regions suggests the arrangement of aromatic amino acids in a restricted environment and thus implying a folded structure adopted by the polypeptide chain. The intrinsic tryptophan emission spectrum of the podocin domain shows $\lambda_{\max}$ value at 335 nm (Fig. 5C). The $\lambda_{\max}$ of native proteins is related to the polarity of the environment of the tryptophan residue and typically can range from 308–350 nm. Unfolded forms with residues in the apolar micro-environment show a blue shift of the $\lambda_{\max}$ [25].

3.4. Structural stability of the podocin domain

Although, the intrinsic fluorescence may not unequivocally provide

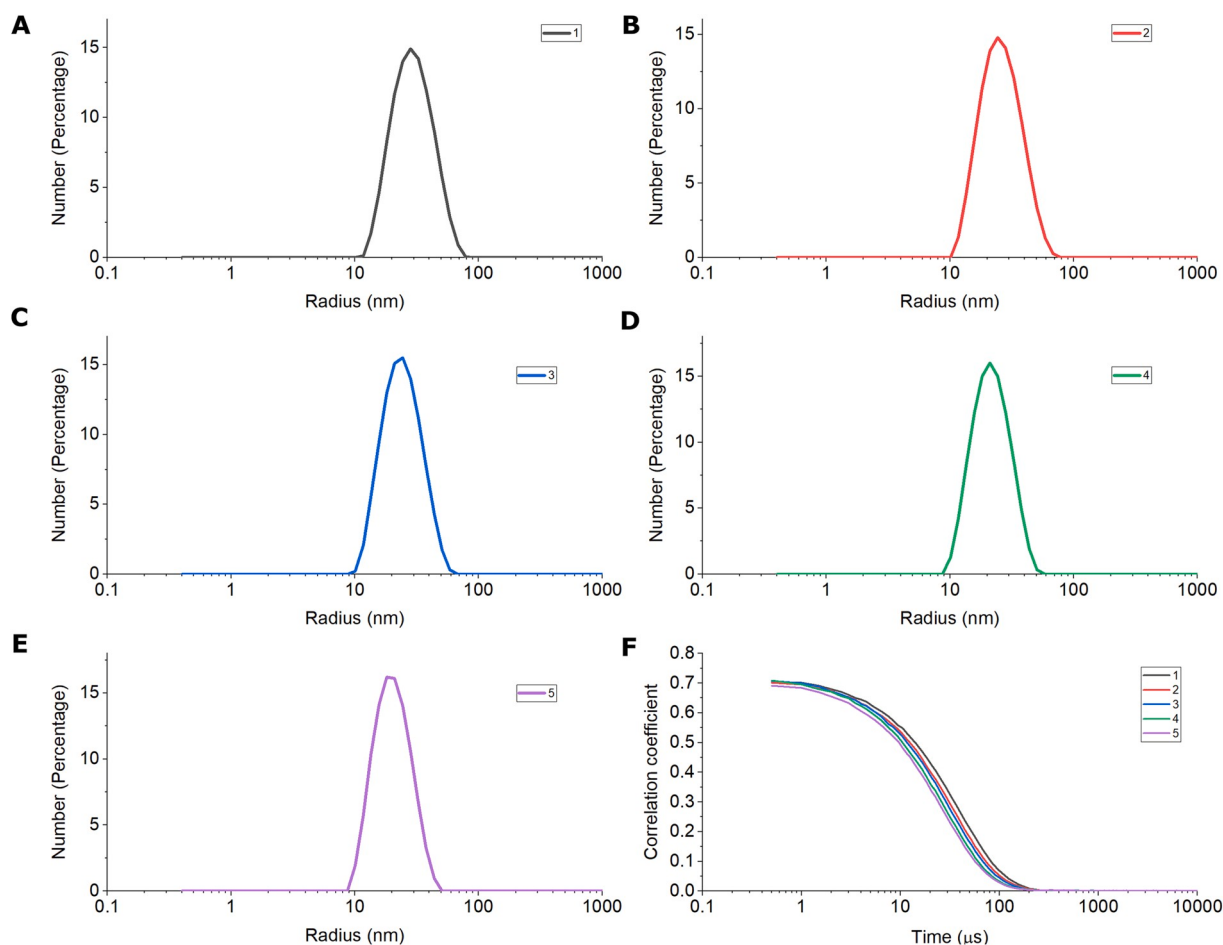


Fig. 4. Polydispersity and hydrodynamic radius of the podocin domain homo-oligomers: The DLS was in tandem with the SEC-MALS. DLS analysis of the samples corresponding to peak observed in SEC-MALS (1–5) is represented as a number percentage vs size in nanometers curve plots marked in different colors (A–E). The corresponding correlograms of the samples (1–5) represented as correlation coefficient vs time in microseconds. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

structural information, the changes in the folded state can be followed by monitoring the emission intensity which conveys the changes in the native state tryptophan environment. The changes in the CD signal and the perturbations of the intensity of the intrinsic fluorescence emission were estimated to assess the stability of the folded podocin domain. Temperature-induced unfolding was monitored using the fluorescence emission spectra with increasing temperature in the range of 20 °C–95 °C. The spectra showed a uniform decrease in λ_{max} intensity without either a bathochromic or hypsochromic shift (Fig. 6A). The change in fluorescence intensity as a function of temperature (Fig. 6B) shows a linear transition and does not show a sigmoidal shape typically observed for proteins with a tightly packed tertiary core. The absence of a folded baseline suggests a not so tightly packed tertiary core. Far UV CD spectral changes were also monitored as a function of temperature. With the gradual increase in temperature a gradual loss of signal intensity (Fig. 6C), loss of secondary structure, and also the absence of a native baseline was observed (Fig. 6D) implying that the podocin domain possesses secondary and tertiary interactions but lack a tightly packed core.

3.5. Calorimetric analysis of podocin domain

Differential scanning calorimetry (DSC) gives the overall enthalpy value (ΔH_{cal}) for each structural transition. Therefore, we performed DSC to calculate the thermodynamic parameters such as T_g and the ΔH_{cal} associated with structural changes of the homo-oligomer. Peak

deconvolution of the acquired data revealed 5 transition states, out of which four are endothermic transitions (316 K, 325 K, 330 K, and 353 K) and one is an exothermic transition (345 K) (Fig. 7). The respective values for ΔH_{cal} are mentioned in Table 2. The DSC profile suggests that the oligomers in the mixture undergo dissociation via three transition temperatures namely 316 K, 325 K, and 330 K and the presence of exothermic transition at 345 K suggests possible hydrophobic interactions among the constituent oligomers before complete dissociation at 353 K.

4. Discussion

Podocin selectively expresses in the glomerular podocytes and it is instrumental for preserving the structural integrity of the SD. Though several mutations in the protein are associated with proteinuria in humans, the structural details of this protein are unclear. Here we report for the first time, the stoichiometry of oligomerization of the truncated human podocin construct. Our investigation indicates that at ambient temperature, and in a reduced environment the podocin domain predominantly adopts a 16-mer oligomeric state. However, other oligomeric conformations ranging from 25-mer to 13-mer were also observed nevertheless, the population of these states was comparatively less. The polydispersity index we report adds evidence to the presence of multiple oligomeric species. Additionally, CD and fluorescence spectra revealed podocin homo-oligomers have considerable secondary structure and tertiary packing.

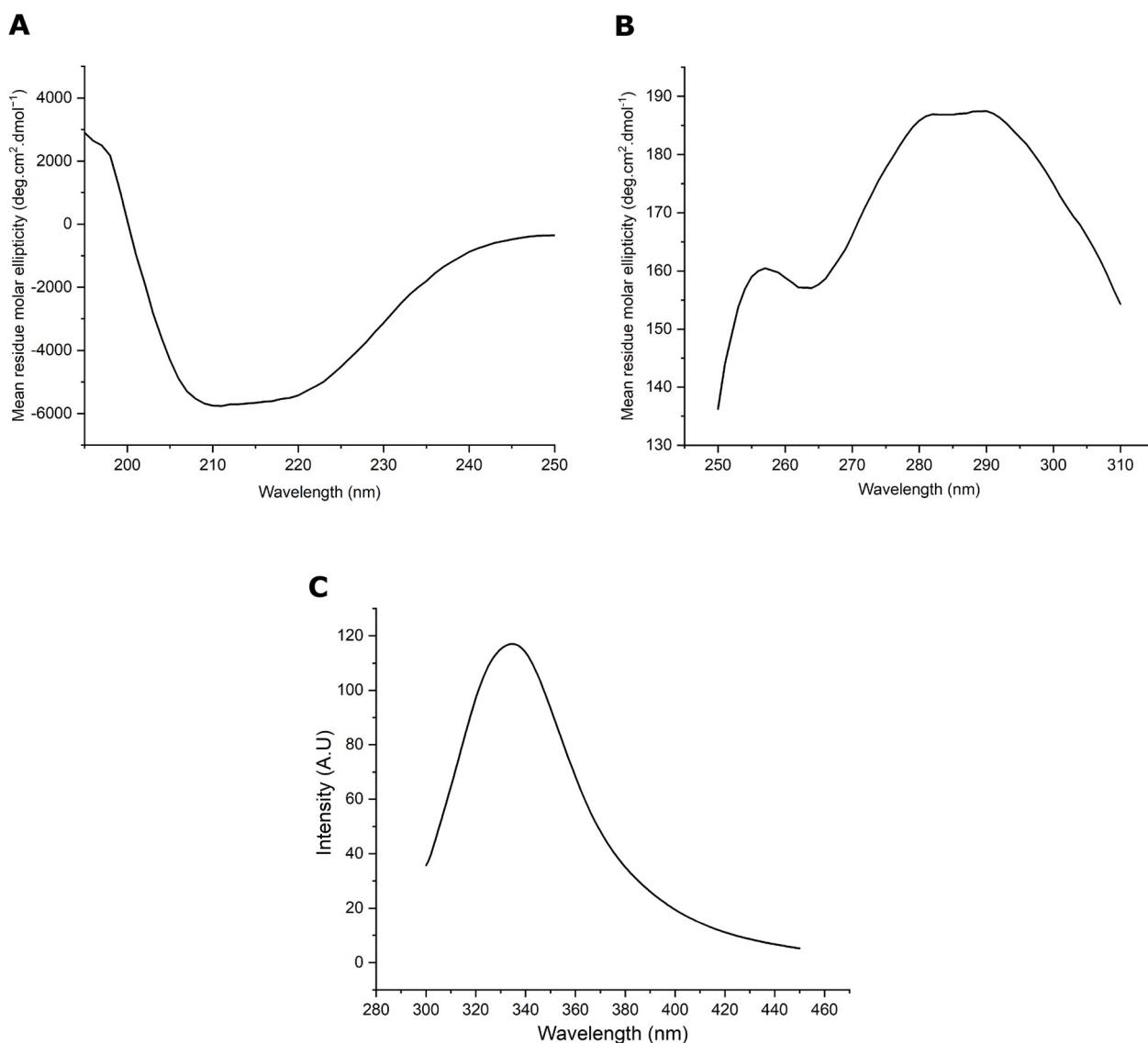


Fig. 5. Podocin domain exhibits secondary structural elements and tertiary packing: A. Far-UV CD spectrum of Podocin domain at 25 °C, pH 8 in 10mM potassium phosphate buffer supplemented with 150mM NaCl and 2mM β -mercaptoethanol. B. Near-UV CD spectrum representing the tertiary packing of the podocin domain. C. Fluorescence emission of the podocin domain recorded by exciting the protein at 287 nm.

Since stomatin proteins and podocin share significant homology, it expected they may share several structural similarities. Crystallization studies of stomatin protein (*Pyrococcus horikoshii*; residues 56–224) revealed that it exists as a trimer and NMR studies of the same protein with a different truncation size (residues 66–174) associated into an amalgam of oligomers [20]. Although earlier reports suggested that podocin forms homo-oligomers the specificity of oligomerization was not known [7,16,17,26]. Even though our study revealed truncated podocin forms predominantly a 16-mer, we are yet to demonstrate whether full-length podocin also assembles as a 16-mer which is a major limitation of our study. We were unsuccessful in purifying the full-length podocin that could be due to several reasons including podocin may not be stable outside its native environment, due to the presence of IDRs, or interference of the transmembrane segment with protein expression [5]. We, therefore, performed our studies with truncated podocin (126–350 residues) that encompasses the PHB domain and the oligomerization sites. The regions ‘283–313’ and ‘332–348’ of podocin facilitate intermolecular interactions [23]. Huber et. al, reported that podocin (R138X) was unable to form oligomers suggesting the importance of this region in podocin-podocin interactions [17]. We included the 126–164 residues

upstream of the PHB domain as it was reported that the high-affinity binding among homo-oligomers requires larger parts of the C-terminal region besides the PHB domain [17]. Accumulating evidence suggests that 126–350 region of podocin is crucial for homo-oligomerization, it can be expected that mutations in these regions likely distort the intermolecular interactions. Mutations that cause NS could distort the innate ability of podocin to form oligomers and compromise its ability to act as a scaffolding molecule. Together these molecular deformations at the level of podocin oligomeric assembly may lead to altered permselectivity of SD and manifest in significant proteinuria. As it was shown that podocin interacts with several SD proteins, it is less likely a single podocin interacts with all proteins whereas a different monomer of the podocin 16-mer may interact with the neighboring proteins in the SD complex. It should be noted that the podocin domain predominantly forms a 16-mer, and to a lesser extent 25-mer to 13-mer are also observed. These results are further justified by the hydrodynamic radius and the polydispersity index (PDI) measured by the DLS. The reported PDI for podocin domain falls in the midrange value (0.08–0.5) suggesting that although it has one major oligomer species by volume, the presence of multiple species could be possible [27]. We were unable to

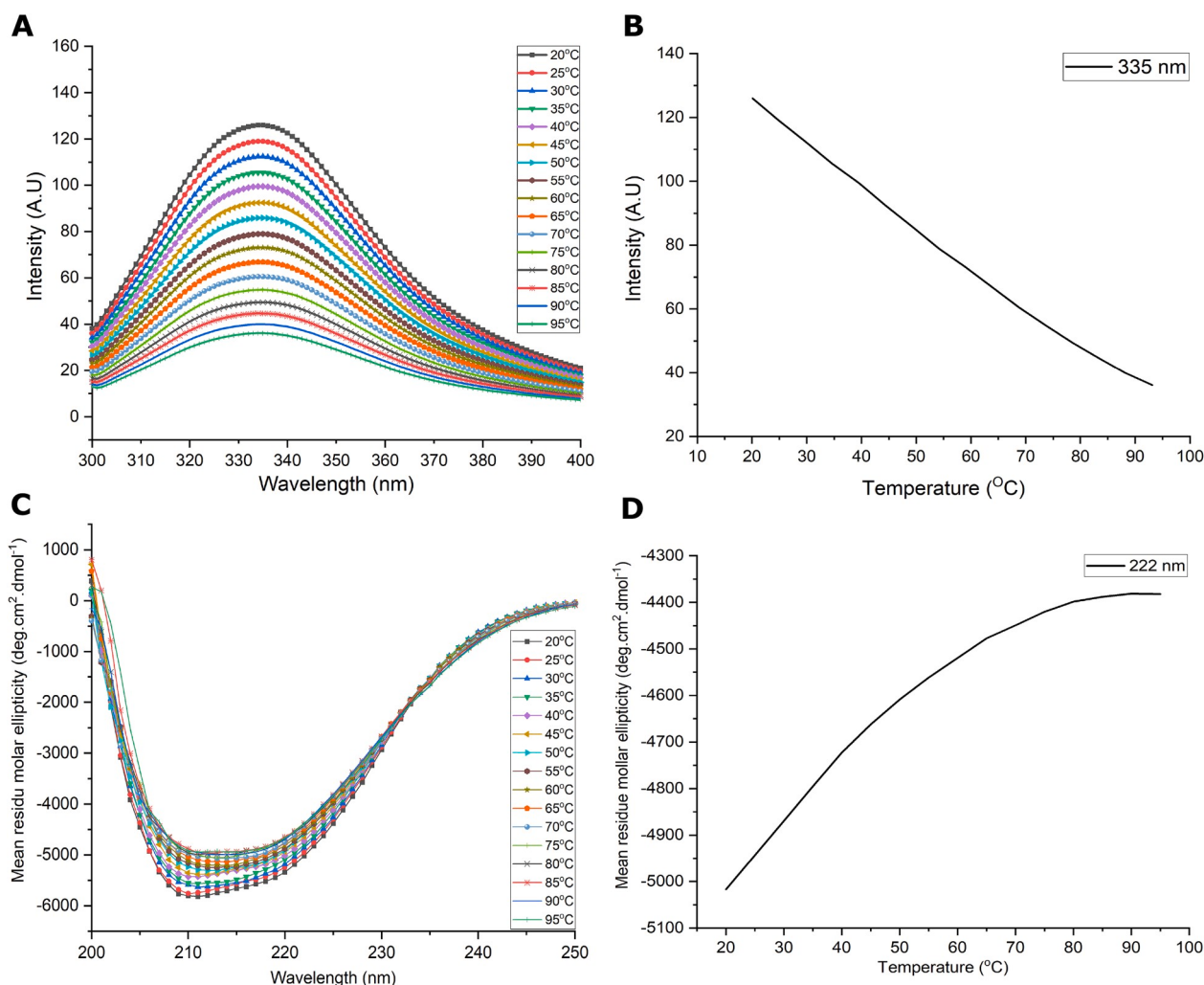


Fig. 6. Effect of temperature on podocin domain: A. Intrinsic tryptophan fluorescence for podocin domain was measured as a function of temperature (20 °C–95 °C) by exciting the protein at 287 nm. B. Peak emission of the podocin domain at 335 nm is plotted over a temperature range of 20 °C–95 °C. C. FarUV CD spectra (250–200 nm) of the podocin domain was acquired as a function of temperature (20 °C–95 °C). D. The MRE values of the podocin domain at 222nm is plotted as a function of temperature (range: 20 °C–95 °C) to monitor the changes in the protein structure.

obtain crystals of the podocin domain for performing X-ray diffraction possibly due to its dynamic oligomeric nature.

Far UV-CD spectra revealed that the podocin domain consists of a considerable amount of α -helices and β -sheets. We could not record far UV-CD spectra beyond 195nm possibly because of the presence of buffer components like β -mercaptoethanol that manifested in increased HT voltage. However, β -mercaptoethanol in the buffer helped in solubilizing podocin domain and contributed towards the protein stability. It is noteworthy that crystal structures of stomatin protein are devoid of disulfide linkages [22,28]. In a study, Huber et.al reported that Cys126 and Cys160 residues of mouse podocin undergo palmitoylation and participate in membrane insertion [26]. Similarly, the Cys158 in the human podocin domain is also expected to be palmitoylated instead of forming disulfide linkages. This suggests that the disulfide bridges between cysteines either do not exist or are not contributing to protein stability.

While assessing the effect of temperature on the protein by probing its intrinsic tryptophan fluorescence and secondary structure content via measuring optical rotation by using the CD, the baseline was observed only in the case of the CD spectrum beyond 80 °C. A significant linear decrease in the fluorescence intensity at 335 nm with increasing temperature may be due to the partial exposure of the lone Trp256 and other aromatic residues as observed in the predicted model for podocin and

the crystal structure of stomatin protein (3BK6) (Fig. S1) [18,22]. Also, there was no bathochromic shift observed, which ascertains the fact that the oligomer might not have dissociated completely. Similarly, far UV-CD spectra at 222nm were used to monitor secondary structural changes. A small yet noticeable change in the MRE value and the shape of the spectra implicates that most of the secondary structure was retained, however, a slight rearrangement of structure could be possible, owing to the increased hydrophobic effect. With the increasing temperature, the prominent double minima smoothen out indicating that there might be a slight increase in the β -sheet content. The stable MRE value observed beyond 80 °C might suggest some residual local ordering that still prevails at high temperatures.

Calorimetric analysis of the oligomeric species revealed multiple transition points for the podocin domain. At 20 °C the protein was shown to attain an amalgam of oligomeric states and a steady increase in temperature may promote dissociation of these different oligomers via three different transition temperatures since each oligomeric state may not have same transition point. Nevertheless, when the temperature was further increased significant exothermic transition state was observed. We assume the exothermic transition state may be due to the association of hydrophobic cores of the constituent oligomers. It is well known that lower temperatures do not favor hydrophobic interactions whereas, an increase in temperature to a certain degree promotes hydrophobic

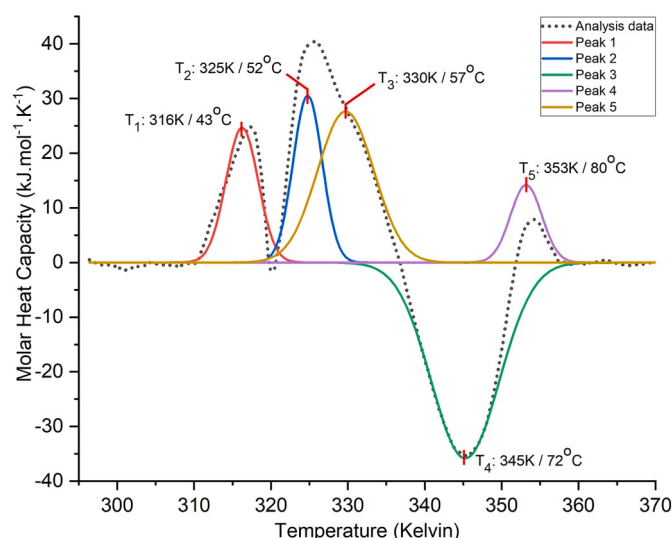


Fig. 7. DSC analysis of podocin domain: The plots represent the endothermic and exothermic transitions of the podocin domain over a temperature range from 296 K/23 °C to 368 K/95 °C. The initial DSC artifact near 293 K/20 °C in the DSC data is not represented in the plot. The analysis data represented as a dotted line in the plot after baseline subtraction. Each transition peak (1–5) in the plot is represented by a color line which is a result of peak deconvolution function. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2

Enthalpy values and the transition temperatures as noticed in the dynamic scanning calorimetry.

Temperature range	Transition temperatures (T _g)	Enthalpy change (ΔH _{cal}) (kJ/mol)
293 K–368 K/(20 °C–95 °C)	Endothermic	
	316 K/43 °C	130
	325 K/52 °C	145
	330 K/57 °C	73
	353 K/80 °C	254
	Exothermic	
	345 K/72 °C	-397

interactions among the protein molecules [29]. Subsequently, when the temperature is further increased, we speculate that the oligomer may invariably disassociate into lower oligomeric species. It is noteworthy that the podocin domain did not re-trace the path to its initial state when the protein was cooled. This peculiar behavior of the podocin domain starkly correlates with the oligomeric propensities as observed in the stomatin proteins [20,21]. Further, we performed above-discussed experiments at 5 °C and found that 16-mer seems to be the major conformer at both lower and ambient temperature (Figs. S2–S5). Interestingly in the calorimetric analysis, multiple transition points were not observed which suggest the presence of 16-mer species as the predominant conformer at lower temperatures and that was further confirmed by SEC analysis (Figs. S4–S5).

In conclusion, our study, to the best of our knowledge is the first to report the cloning, expression, and purification of the human podocin domain (126–350). We showed that the podocin domain is majorly a 16-mer homo-oligomer. However, to a lesser extent, the protein is capable of associating into other oligomeric states. CD and FL analysis of the protein indicated that the podocin domain in isolation attains considerable secondary structure and tertiary packing. However, the significance of the podocin oligomerization in the formation of the large

macromolecular complex and how a mutation in a podocin monomer will affect its oligomerization status need to be addressed. Further structural characterization of podocin and other slit diaphragm proteins is greatly warranted to understand the mechanism of pathogenesis of the nephrotic syndrome.

Author statement

Sandeep KMN: Conceptualization, Methodology, Visualization, Writing - Reviewing and Editing.

Shivkumar S Irukuvajjala: Methodology & Validation.

Krishan Kumar: Methodology.

Kanchan Garai: Methodology & Validation.

Pannuru Venkatesu: Methodology, Writing - original Draft.

Ramakrishna Vadrevu: Resources, Conceptualization, Formal Analysis, Writing - Reviewing and Editing.

Anil K Pasupulati: Conceptualization, Formal Analysis, Visualization, Supervision, Funding acquisition, Writing - Reviewing and Editing.

Funding

We acknowledge financial assistance from the Indian Council of Medical Research (Grant no: ICMR 2019/905).

Acknowledgments

We thank Dr. Akif, University of Hyderabad, for providing access to the AKTA-Purifier Gel-filtration system. We also acknowledge the Central Analytical laboratory of BITS - Pilani, Hyderabad campus for allowing access to Circular Dichroism Spectrophotometer. MNSK acknowledges ICMR for providing senior research fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbrep.2020.100774>.

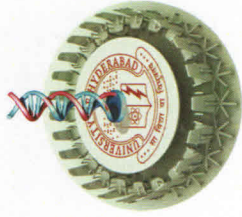
References

- [1] P. Anil Kumar, G.I. Welsh, M.A. Saleem, R.K. Menon, Molecular and cellular events mediating glomerular podocyte dysfunction and depletion in diabetes mellitus, *Front. Endocrinol.* 5 (2014) 151.
- [2] A. Patari-Sampo, P. Ihalmio, H. Holthofer, Molecular basis of the glomerular filtration: nephrin and the emerging protein complex at the podocyte slit diaphragm, *Ann. Med.* 38 (2006) 483–492.
- [3] D. Mukhi, R. Nishad, R.K. Menon, A.K. Pasupulati, Novel actions of growth hormone in podocytes: implications for diabetic nephropathy, *Front. Med.* 4 (2017).
- [4] R. Rodewald, M.J. Karnovsky, Porous substructure of the glomerular slit diaphragm in the rat and mouse, *J. Cell Biol.* 60 (1974) 423–433.
- [5] S.K. Mulukala, R. Nishad, L.P. Kolligundla, M.A. Saleem, N.P. Prabhu, A. K. Pasupulati, In silico Structural characterization of podocin and assessment of nephrotic syndrome-associated podocin mutants, *IUBMB Life* 68 (2016) 578–588.
- [6] S.E. Dryer, J. Reiser, TRPC6 channels and their binding partners in podocytes: role in glomerular filtration and pathophysiology, *Am. J. Physiol. Ren. Physiol.* 299 (2010) F689–F701.
- [7] K. Schwarz, M. Simons, J. Reiser, M.A. Saleem, C. Faul, W. Kriz, A.S. Shaw, L. B. Holzman, P. Mundel, Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin, *J. Clin. Invest.* 108 (2001) 1621–1629.
- [8] N.Y. Shih, J. Li, R. Cotran, P. Mundel, J.H. Miner, A.S. Shaw, CD2AP localizes to the slit diaphragm and binds to nephrin via a novel C-terminal domain, *Am. J. Pathol.* 159 (2001) 2303–2308.
- [9] N.S.K. Mulukala, P.P. Kar, R. Vadrevu, A.K. Pasupulati, Intrinsically disordered regions mediate macromolecular assembly of the Slit diaphragm proteins associated with Nephrotic syndrome, *Mol. Simulat.* 45 (2019) 603–613.
- [10] M. Kestila, U. Lenkkeri, M. Mannikko, J. Lamerdin, P. McCready, H. Putaala, V. Ruotsalainen, T. Morita, M. Nissinen, R. Herva, C.E. Kashtan, L. Peltonen, C. Holmberg, A. Olsen, K. Tryggvason, Positionally cloned gene for a novel glomerular protein–nephrin–is mutated in congenital nephrotic syndrome, *Mol. Cell* 1 (1998) 575–582.
- [11] G. Benoit, E. Machuca, C. Antignac, Hereditary nephrotic syndrome: a systematic approach for genetic testing and a review of associated podocyte gene mutations, *Pediatr. Nephrol.* 25 (2010) 1621–1632.

- [12] N. Boute, O. Gribouval, S. Roselli, F. Benessy, H. Lee, A. Fuchshuber, K. Dahan, M. C. Gubler, P. Niaudet, C. Antignac, NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome, *Nat. Genet.* 24 (2000) 349–354.
- [13] C.E. Sadowski, S. Lovric, S. Ashraf, W.L. Pabst, H.Y. Gee, S. Kohl, S. Engelmann, V. Vega-Warner, H. Fang, J. Halbritter, M.J. Somers, W. Tan, S. Shril, I. Fessi, R. P. Lifton, D. Bockenhauer, S. El-Desoky, J.A. Kari, M. Zenker, M.J. Kemper, D. Mueller, H.M. Fathy, N.A. Soliman, S.S. Group, F. Hildebrandt, A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome, *J. Am. Soc. Nephrol. : JASN (J. Am. Soc. Nephrol.)* 26 (2015) 1279–1289.
- [14] M. Gigante, G. Caridi, E. Montemurno, M. Soccio, M. d'Apolito, G. Cerullo, F. Aucella, A. Schirinz, F. Emma, L. Massella, G. Messina, T. De Palo, E. Ranieri, G. M. Ghiggeri, L. Gesualdo, TRPC6 mutations in children with steroid-resistant nephrotic syndrome and atypical phenotype, *Clin. J. Am. Soc. Nephrol. : CJASN* 6 (2011) 1626–1634.
- [15] M. Gigante, P. Pontrelli, E. Montemurno, L. Roca, F. Aucella, R. Penza, G. Caridi, E. Ranieri, G.M. Ghiggeri, L. Gesualdo, CD2AP Mutations Are Associated with Sporadic Nephrotic Syndrome and Focal Segmental Glomerulosclerosis (FSGS), *Nephrology, Dialysis, Transplantation*, vol 24, official publication of the European Dialysis and Transplant Association - European Renal Association, 2009, pp. 1858–1864.
- [16] T.B. Huber, M. Kottgen, B. Schilling, G. Walz, T. Benzing, Interaction with podocin facilitates nephrin signaling, *J. Biol. Chem.* 276 (2001) 41543–41546.
- [17] T.B. Huber, M. Simons, B. Hartleben, L. Sernetz, M. Schmidts, E. Gundlach, M. A. Saleem, G. Walz, T. Benzing, Molecular basis of the functional podocin-nephrin complex: mutations in the NPHS2 gene disrupt nephrin targeting to lipid raft microdomains, *Hum. Mol. Genet.* 12 (2003) 3397–3405.
- [18] S.K. Mulukala Narasimha, P.P. Kar, R. Vadrevu, A.K. Pasupulati, Intrinsically disordered regions mediate macromolecular assembly of the Slit diaphragm proteins associated with Nephrotic syndrome, *Mol. Simulat.* 45 (2019) 603–613.
- [19] S. Roselli, O. Gribouval, N. Boute, M. Sich, F. Benessy, T. Attie, M.C. Gubler, C. Antignac, Podocin localizes in the kidney to the slit diaphragm area, *Am. J. Pathol.* 160 (2002) 131–139.
- [20] Y. Kuwahara, S. Unzai, T. Nagata, Y. Hiroaki, H. Yokoyama, I. Matsui, T. Ikegami, Y. Fujiyoshi, H. Hiroaki, Unusual thermal disassembly of the SPFH domain oligomer from *Pyrococcus horikoshii*, *Biophys. J.* 97 (2009) 2034–2043.
- [21] L. Snyers, E. Umlauf, R. Prohaska, Oligomeric nature of the integral membrane protein stomatin, *J. Biol. Chem.* 273 (1998) 17221–17226.
- [22] H. Yokoyama, S. Fujii, I. Matsui, Crystal structure of a core domain of stomatin from *Pyrococcus horikoshii* illustrates a novel trimeric and coiled-coil fold, *J. Mol. Biol.* 376 (2008) 868–878.
- [23] P. Straner, E. Balogh, G. Schay, C. Arrondel, A. Miko, G. L'Aune, A. Benmerah, A. Perczel, K.M. D. C. Antignac, G. Mollet, K. Tory, C-terminal oligomerization of podocin mediates interallelic interactions, *Biochimica et biophysica acta, Mol. Basis Dis.* 1864 (2018) 2448–2457.
- [24] T. Palmen, S. Lehtonen, A. Ora, D. Kerjaschki, C. Antignac, E. Lehtonen, H. Holthofer, Interaction of endogenous nephrin and CD2-associated protein in mouse epithelial M-1 cell line, *J. Am. Soc. Nephrol. : JASN (J. Am. Soc. Nephrol.)* 13 (2002) 1766–1772.
- [25] M.R. Eftink, The use of fluorescence methods to monitor unfolding transitions in proteins, *Biophys. J.* 66 (1994) 482–501.
- [26] T.B. Huber, B. Schermer, R.U. Muller, M. Hohne, M. Bartram, A. Calixto, H. Hagmann, C. Reinhardt, F. Koos, K. Kunzelmann, E. Shirokova, D. Krautwurst, C. Harteneck, M. Simons, H. Pavenstadt, D. Kerjaschki, C. Thiele, G. Walz, M. Chalfie, T. Benzing, Podocin and MEC-2 bind cholesterol to regulate the activity of associated ion channels, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 17079–17086.
- [27] Asmawati, W.A.W. Mustapha, S.M. Yusop, M.Y. Maskat, A.F. Shamsuddin, Characteristics of Cinnamaldehyde Nanoemulsion Prepared Using APV-High Pressure Homogenizer and Ultra turrax 1614, 2014, pp. 244–250.
- [28] J. Brand, E.S. Smith, D. Schwefel, L. Lapatsina, K. Poole, D. Omerbasic, A. Kozlenkov, J. Behlke, G.R. Lewin, O. Daumke, A stomatin dimer modulates the activity of acid-sensing ion channels, *EMBO J.* 31 (2012) 3635–3646.
- [29] E. van Dijk, A. Hoogveen, S. Abeln, The hydrophobic temperature dependence of amino acids directly calculated from protein structures, *PLoS Comput. Biol.* 11 (2015), e1004277.



BioQuest 2015 23rd - 24th September



School of Life Sciences
University of Hyderabad
Hyderabad - 500 046, INDIA

Certificate of Participation

This is to certify that Prof./Dr./Mr./Ms..... ✓ **SANDEEP KUMAR MUKKALA NARASIMHA**.....

from Department of..... **Biochemistry**.....

has participated in the "**BioQuest 2015**" held on September 23rd and 24th, 2015.

PAZ
Convener

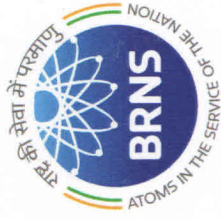
Deepp
Dean
School of Life Sciences



USE OF RADIATION TECHNOLOGY IN
AGRICULTURE, HEALTH AND RESEARCH
UNIVERSITY OF HYDERABAD - BHABHA ATOMIC RESEARCH CENTER

SCHOOL OF LIFE SCIENCES

UNIVERSITY OF HYDERABAD - 500046



Certificate of Participation

This is to certify that Dr./Mr./Ms

Vandee Kumar Mulukala Narasimha

from
Department of Biochemistry

has participated in a one day national seminar

on "THE USE OF RADIATION TECHNOLOGY" conducted by The University of Hyderabad

and Bhabha Atomic Research Center on 13th November, 2015.

P. Anil Kumar

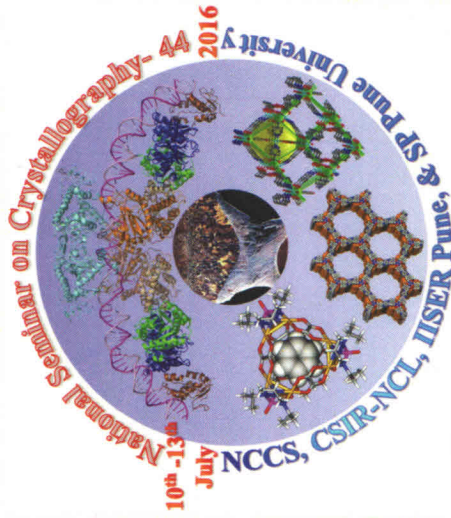
Organising Secretary

Department of Biochemistry

School of Life Sciences

44th National Seminar on Crystallography

NCCS



Certificate Of Participation

This is to certify that Dr./Mr./Ms./Mrs. Sandeep Kumar Mulukala Narasimha
Affiliated to School of Life Sciences, University of Hyderabad
participated / presented a paper (invited / oral / poster session) in the 44th National Seminar
on Crystallography held during 10th -13th July 2016, at Pune.

S. Mande

Dr. Shekhar C. Mande

Convener

IKMC 2016
ACCELERATING
INNOVATION

October 24 & 25, 2016, HICC, Hyderabad, India



Certificate

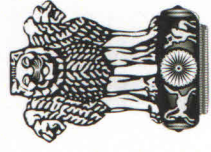
Awarded to Mr. M N Sandeep Kumar of Hyderabad University
in recognition of his Participation in the Technology Showcase organised on
the occasion of "IKMC2016: Accelerating Innovation"
on October 24, 2016 at HICC, Hyderabad, India.



Deepanwita

Deepanwita Chattopadhyay

Chairman & CEO, IKP Knowledge Park



सत्यमेव जयते
Department of Science and Technology
Government of India

MHRD-Global Initiative on Academic Network (GIAN)

Workshop on

Protein Structure And Drug Discovery

Certificate of Participation

This is to certify that Prof. / Dr. / Mr. / Ms. Sandeep Kumar Mbulukala & Narasimha
from University of Hyderabad participated in the course titled "Protein Structure And Drug
Discovery" during the period from **August 27- September 05, 2017**, School of Life Sciences,
University of Hyderabad, Hyderabad, INDIA.


Prof. Pallu Reddanna
Dean, School of Life Sciences


Dr. M. Rami Reddy
Coordinator



UKIERI
UK-India Education
and Research Initiative



UGC-UKIERI
Thematic Partnership

Workshop on Animal Models in Cancer Research

Organized by

School of Medical Sciences, University of Hyderabad, India

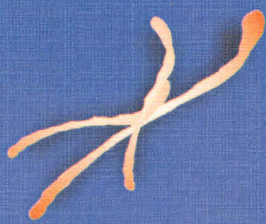
24th - 25th October 2017

CERTIFICATE

This is to certify that Dr./Mr./Ms./Shri. *Sandeep Kumar M.N.* has participated
in the workshop on **Animal Models in Cancer Research** conducted by School of Medical Sciences, University
of Hyderabad.

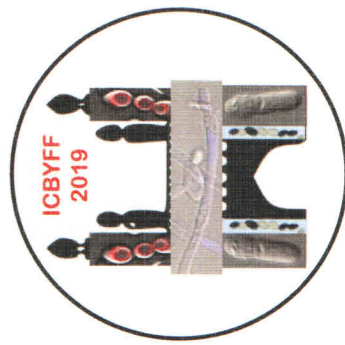
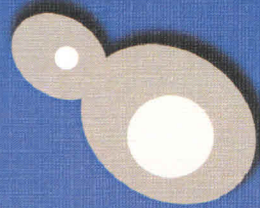
Prof. Geeta Vemuganti,
Dean, School of Medical Science,
University of Hyderabad

Prof. Sarah E Coupland
Department of Molecular & Clinical Cancer Medicine
University of Liverpool



ICBYFF

2019



Certificate of Participation

Sandeep Mulukala NK

This is to certify that

participated in the XI International Conference on Biology of Yeasts
and Filamentous Fungi organized at the School of Life Sciences,
University of Hyderabad, Hyderabad, India, from 27 Nov 2019 to
29 Nov 2019.

Krishna-Nukla

Prof. Krishnaveni Mishra

ORGANIZER



Dr. Rupinder Kaur

Dr. Rupinder Kaur

ORGANIZER

Plagiarism report

Molecular Evolution and structural insights of the podocyte slit-diaphragm proteins: Emphasis on Podocin

by Sandeep Kumar Mulukala Narasimha

Submission date: 29-Oct-2020 03:22PM (UTC+0530)

Submission ID: 1430018854

File name: For_plagiarism.pdf (5.86M)

Word count: 20836

Character count: 110529

Molecular Evolution and structural insights of the podocyte slit-diaphragm proteins: Emphasis on Podocin

ORIGINALITY REPORT

44%

SIMILARITY INDEX

23%

INTERNET SOURCES

43%

PUBLICATIONS

1%

STUDENT PAPERS

The similarity index from Primary sources 1, 2 and 3 are Sandeep's own publications and may be excluded from plagiarism. Leaving only ~2% of similarity with available literature.

1

www.tandfonline.com

Internet Source

P. H. Kumar

Dr. P. ANIL KUMAR

Assistant Professor

Department of Biochemistry

University of Hyderabad

Hyderabad-500 046. India.

22%

2

Sandeep K.N. Mulukala, Shivkumar S.

Irukuvajjula, Krishan Kumar, Kanchan Garai et

al. "Structural features and oligomeric nature of

human podocin domain". Biochemistry and

Biophysics Reports, 2020

Publication

P. H. Kumar

Dr. P. ANIL KUMAR

Assistant Professor

Department of Biochemistry

University of Hyderabad

Hyderabad-500 046. India.

16%

3

Sandeep KN Mulukala, Shivkumar S

Irukuvajjula, Krishan Kumar, Kanchan Garai et

al. "Structural Features and Oligomeric Nature

of Human Podocin Domain", Cold Spring Harbor

Laboratory, 2020

Publication

P. H. Kumar

Dr. P. ANIL KUMAR

Assistant Professor

Department of Biochemistry

University of Hyderabad

Hyderabad-500 046. India.

4%

4

Sandeep K. Mulukala Narasimha, Prajna

Parimita Kar, Ramakrishna Vadrevu, Anil K.

Pasupulati. "Intrinsically disordered regions

mediate macromolecular assembly of the Slit

diaphragm proteins associated with Nephrotic

syndrome", Molecular Simulation, 2019

<1%