4.1 Introduction

BRCA1 interacts with numerous molecules and consequently plays an essential role in multiple pathways such as the DNA damage repair, cell-cycle regulation and transcription regulation [13, 19, 20][2, 8, 89, 162, 163]. It acts as a tumor suppressor and has been found to be associated with hereditary breast and ovarian cancers [3, 164]. Furthermore, BRCA1 BRCT deletion mutant showed perturbed behavior towards sub-nuclear co-localization with H2AX, thus highlighting its importance in DNA damage repair [85, 116, 147, 165].

Figure 4.1: Schematic representation for structure of MERIT40. Willebrand Factor A (vWFA) motifs are named for the prototype member of the family, von Willebrand factor A, a platelet glycoprotein with critical roles in blood clotting.

MERIT40 (MEdiator of RAP80 Interaction and Targeting 40) is an essential molecule of the BRCA1 complex and helps to maintain its protein complex integrity (**Figure 4.1**). Knockdown of MERIT40 significantly reduces the RAP80 and ABRAXAS protein levels, consequently affecting the integrity of BRCA1 complex [13, 19]. MERIT40 downregulated cells shown lower the stability of BRCA1 complex compared to any other member of this complex [13]. To elucidate its role in BRCA1 complex, structure of MERIT40 and its correlation with functional properties were analyzed.

The aim here is to use the biophysical, biochemical, *in-silico* and structural biology tools to dissect the structure aspect of MERIT40. We have performed cloning, expression and purification of MERIT40 and bio-physicochemical characterization using multi-model approach.

4.2 Material and Methods

In the present study, a structural stability, protein folding and thermodynamic analysis of MERIT40 was carried out. MERIT40 was sub-cloned into bacterial expression vector pGEX-kT using one step cloning method, as described in chapter 3 (**Figure 3.3**). The nucleotide sequence of primers used in PCR reaction for cloning are shown in **Figure 4.2**. The gene of interest was amplified using following PCR condition: 95 °C denaturation (5 minutes), 95 °C denaturation (45 seconds), annealing 65°C for 35 seconds, Extension 72°C at 0.5kb/min, final extension 72°C for 10 minutes, 25 cycles. The PCR amplified product was digested with EcoR1 and BamH1 and ligated in the pGEX-kT vector. The ligation mixture was transformed into *E.coli DH5α* cells. Colony screening was done by insert release from the ligated vectors using EcoR1 and BamH1 restriction enzymes, and final confirmation was done by DNA sequencing. Sequence verified clones were then proceeded for protein preparation.

GTC	BamH1	TEV	ATG ATGGAAGTGGCAGAGCCCAGCAG
GTC	EcoR1	STOP	CCATTGAGGTTGAGGCCACTGTCTGA

Figure 4.2: Primer used during PCR amplification of MERIT40.

4.2.1 Protein expression: MERIT40 was expressed in bacterial system using *E.coli* Rosetta (2DE3) strain. For protein expression, 100 µg/µl plasmid construct was transformed into *E.coli* Rosetta (2DE3) cells and grown on LB agar plate containing antibiotic ampicillin (100 µg/ml). Further protocol is described below.

Protocol for protein expression:

1. **Inoculation:** Pick-up a single transformed colony from antibiotic resistant LB agar plate and inoculate it in 100 ml LB broth containing 100 µg/ml of ampicillin. Incubate on shaker incubator at 37^{o}C over night.
2. **Dilution:** Inoculate 10 ml of starting culture to 1000 ml (1: 100 ratios) of autoclaved LB broth containing the 100 µg/ml of ampicillin. Incubate the flasks on shaker incubator at 37 ^{o}C until it has reached mid-log phase i.e. A_{600} between 0.6-0.8.
3. **Induction:** Cool down the flasks and add 400 µl IPTG (stock 1M) and incubate on shaker incubator at 24^{o}C for 16 hours.
4. **Harvesting:** Transfer the culture to centrifuge bottles and centrifuge for 10 minutes at 6000 rpm at 4^{o}C. Resuspend the pellet in a small amount of supernatant and centrifuge for 15 minutes at 5000 rpm, 4 ^{o}C.
5. **Storage:** The pellet obtained can be stored at -80 ^{o}C until further use.

Protein purification was carried out using affinity chromatography followed by size exclusion chromatography by AKTA-FPLC to get pure and homogenous protein.

4.2.2 Purification of MERIT40:

Purification buffer composition: 300 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, 10 mM β-ME, 5% glycerol, pH 7.5

FPLC buffer composition: 300 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, 10 mM BME, pH 7.5

Protocol for purification of MERIT40:

1. **Re-suspension**: Re-suspend the pellet of MERIT40 in 40 ml of purification buffer supplemented with 200 µl of 200 mM PMSF and 20 µl of protease inhibitor.

2. **Ultra sonication**: Transfer the suspension into centrifuge tube and sonicate at 50 pulse rate and 50 power with 1.45 minutes of duty cycle.
3. **Centrifugation:** After sonication the suspension is subjected to centrifugation at 18000 rpm for 45 minutes at 4^{o}C.
4. **Equilibration of GST beads**: Give two column volume washes with distilled water to remove the traces of ethanol as the beads are stored in 20% ethanol and then 3 column washes with purification buffer.
5. **Binding**: The soluble fraction obtained is brought to room temperature and then mixed with affinity resin properly and incubated at room temperature for 1 hour.
6. **Washing**: After binding, the column is given 4 column washes with wash buffer so as to remove non specific protein. Take 40µl of beads to load on gel to observe the bound protein.
7. **Cleavage**: Add 400 µ of TEV protease enzyme, 40 µl of protease inhibitor cocktail and 100 µl of PMSF in 20 ml of purification buffer and perform the cleavage step for 3 hours by passing the solution through column at interval of 1 hour. This step is performed to cleave the GST tag from fusion protein to get the purified native protein. Take out 40ul of beads to observe the cleavage of protein.
8. **Elution**: After TEV cleavage the protein is eluted in 30 ml of purification buffer.
9. **Calibration of Ni-NTA resin**: Give 2 column washes with double distilled water and then 5 to 6 column wash with washing Buffer.
10. **Metal Ion Chelate Affinity Chromatography**: After calibration of Ni-NTA resin, pass the eluted fractions through them. This step was done to remove his-tagged TEV enzyme from protein of interest.

11. **Concentrating the protein:** Transfer the eluted protein in pre-equilibrated 10 kDa centricon and concentrate the protein upto 2 ml by centrifuging it at 4500 rpm for 10 minutes at 4^{o}C. Check the concentration on Nanodrop spectrophotometer (280 nm). Centrifuge for 10 minutes at 5000 rpm at 4 ^{o}C to remove any soluble aggregates or precipitates.
12. **Gel filtration:** Inject 2 ml of concentrated protein on AKTA- FPLC against FPLC buffer.
13. **Fraction collection**: Collect the purified protein obtained through FPLC in 1.7 ml eppendorf at its elution volume according to gel filtration spectra profile of sample.
14. **Loading on SDS-PAGE 12% gel**: Load 20 µl of FPLC fractions on SDS-PAGE, stain with commassie dye, and then destain to visualize the protein of interest.
15. **Concentrate the protein**: The fractions which showed purified protein bands of interest were further concentrated as per requirement.

The purified proteins were used in various bio-physicochemical experiments. The complete details of protocol have been discussed earlier in chapter 3 (material and methods).

4.3 Results and discussion

4.3.1 Cloning, expression and purification of MERIT40: Selected potential clones when digested with the EcoR1 and BamH1 restriction enzymes showed the insert release of appropriate size (**Figure 4.3 B**). Sequencing result confirmed the presence of ligated gene of interest in the vector with desired frame of codon sequence.

MERIT40 eluted from superdex 200 column into the monomeric and dimeric form with some proportion of aggregated protein (**Figure 4.4**). The elution profile of MERIT40

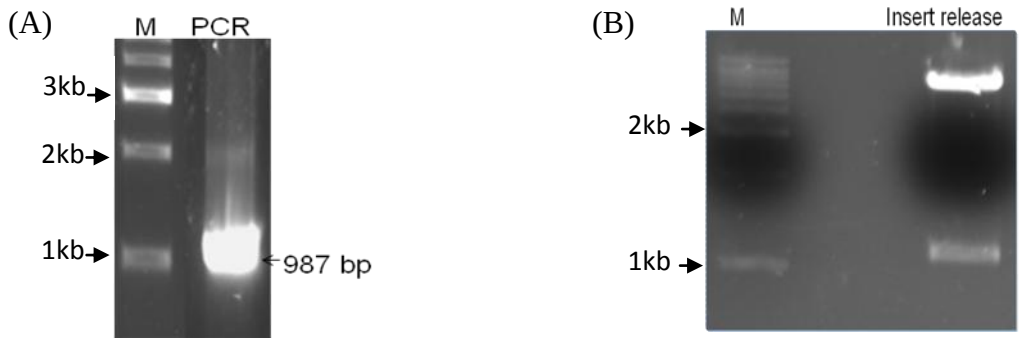

Figure 4.3: Agarose gel showing the PCR amplified product (A) and insert release from potential clones' plasmids (B) of pGEX-kT-MERIT40.

suggests that most of the protein is monomeric in nature, while dimeric proportion was limited to ~40%. The monomeric fractions were proceeded for biophysical characterization and crystallization.

4.3.2 Molecular modeling: Predication of protein function often requires knowledge of structure and physiochemical properties [166, 167]. Till date, MERIT40 is the least characterized member of BRCA1- complex, and literature search could not find any structure that exists in the protein data bank. It showed least homology with existing proteins in database and hence could not be modeled using homology modeling.

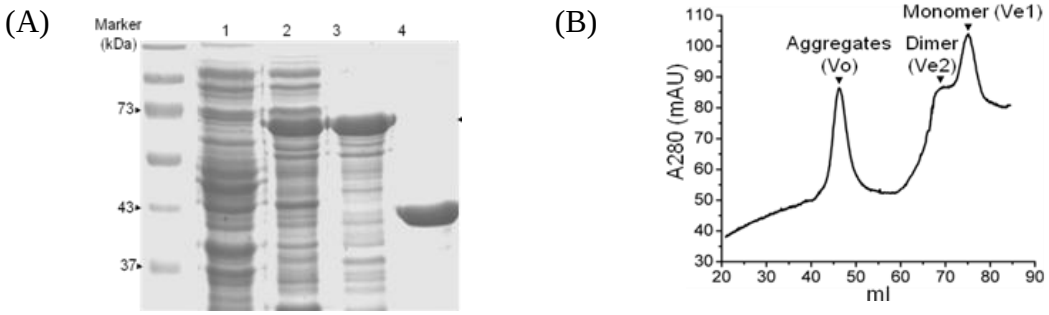

Figure 4.4: Expression and purification of MERIT40. (A) SDS-PAGE showing expression and purification profile of MERIT40. Lane 1-uninduced, 2-soluble protein, 3-protein bound on affinity resin, 4- purified protein. Single arrow- fusion protein. (B) Gel filtration chromatography elution profile of MERIT40 (AKTA-FPLC Superdex 200). V_e represents the elution volume and V_o is for void volume [168].

Considering the fact of non-availability of suitable template for homology modeling, we decided to perform ab-inito modeling using Robetta server which employs combinatorial approaches for model generation. [136] The model built for the full-length MERIT40 sequence is shown in **Figure 4.5**. The overall quality and reliability of the model was assessed based on the energy and stereo chemical geometry of the structure (PROVE_Plot Z-score RMS: 30.936, overall quality factor (ERRAT): 84.936, Verify_3D (of 87.88% residues) score>0.2).

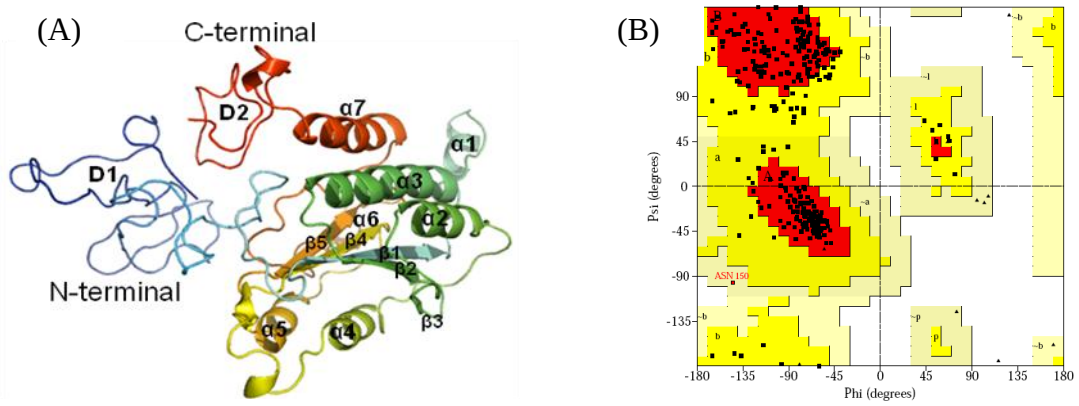

Figure 4.5: Modeled structure of MERIT40 and its validation. Full length protein structure was modeled from Robetta server (A); Modeled structure having 5-β sheets, 7-α helices and disordered region at N-terminal and C-terminal (α1-7: alpha helices, β1-5 : beta sheets, D1-Diordered 1, D2-Diordered 2). (B) Ramachandran plot of modeled structure [168].

The Ramachandran plot has shown 99.7% of the residues to be present in the allowed region [169] and none were in disallowed region **(Figure 4.5)** [170, 171].The observed coordinates in Procheck analysis were found to be satisfactory (>80%), Verify 3D profile revealed that the average score remained above 0.2 and never dropped to below zero, which is significant [172] [173]. Furthermore, structural geometry was evaluated using Mol Probity web server (http://molprobity.biochem.duke.edu/) which showed all the parameters in required range **(Table 4.1).** The model quality was found reliable, and can be used for further studies such as docking and domain organization.

Table 4.1: Protein geometry evaluation using Mol Probity web server [168]

Structural parameter	Value (%)
Poor rotamers	00.00
Ramachandran outliers	00.00
Ramachandran favored	92.66
Cβ deviations >0.25Å	00.00
Bad backbone bonds	00.00
Bad backbone angles	00.00
Asn's, Gln's, and His's orientation (correctly)	100.0

4.3.3 Domain Organization of MERIT40: Intrinsic dis-orderedness is a fairly common phenomenon in many proteins, especially those related to cancer and testis [174]. It has been reported that intrinsically disordered proteins lack the rigid three dimensional structures either for entire length or in localized regions [175]. Due to the difference in conformations and interactions, they play an important biological role in major cellular processes such as transcription, cell cycle regulation, signal transduction and DNA repair [176]. Since MERIT40 is a member of BRCA1-complex, and RAP80 has reported association with retinoid-related, testis associated receptor (RTR), it would be more likely that MERIT40 has disordered regions [108]. To assess this possibility and map the disordered region, the structure of MERIT40 was analyzed using *in-silico* and *in-vitro* approaches. From the model of MERIT40, it has been observed that N- and C-terminal regions are dis-ordered while its middle domain exhibits a well folded *β/α* conformation. Disordered regions of the N-terminal (~ 70 amino acids) and C-terminal (~ 30 amino acids) displayed a random structure pattern (**Figure 4.5**). The ordered region of the MERIT40 (~71–299 residues) encompasses four parallel, one anti-parallel β-strand and seven α-helices. The modeled structure provides the crucial information for disorderness, hence it would be interesting to see whether the same scenario exists in an *in-vitro* condition as well. Limited proteolysis is one of the widely used approach for

identification and mapping of disordered regions of proteins in solution [177]. A compact globular domain of protein should significantly resist the protease digestion, whereas disordered region undergoes rapid digestion due to more accessibility of protease cleavage sites [177].

To confirm the dis-orderness of the N- and C-terminal regions and identify the stable domain, MERIT40 was purified using affinity chromatography (**Figure 4.4**) and subjected to limited proteolysis assay. Trypsin and chymotrypsin proteases are commonly used to locate the compact globular domain(s) in solution [178]. It was observed that both proteases rapidly digested MERIT40 (amino acids 1–329) to a proteolytically stable fragment. MALDI-TOF/TOF analysis indicated that chymotrypsin digested fragment lacked the first 70 amino acids from the N-terminal region and a short stretch of 31 amino acids from the C-terminal (**Figure 4.6**). Furthermore, trypsin digestion generated similar pattern at the N-terminal leaving the C-terminal intact (**Figure 4.6**). Mass spectrometry identification of the stable fragment obtained after limited proteolysis suggested that it belongs to a region from ~ 71–298 residues (**Figure 4.6, Figure 4.7**). This stable fragment is likely defined in the range of MERIT40 vWFA-like domain (95–298), which has sequence homology with the von Willebrand factor (vWA) domain of the Rpn10 proteasome subunit [20, 179]. Based on the structural similarity, it seems that the stable fragment of MERIT40 (71–298) may have proteasome-like activity, thus raising the possibility of its involvement in the ubiquitination and deubiquitination processes along with RAP80 UIM.

4.3.4 Role of MERIT40 in DNA damage repair: Unraveling the function of a novel protein using multidisciplinary approach is a challenging task. Structure based function

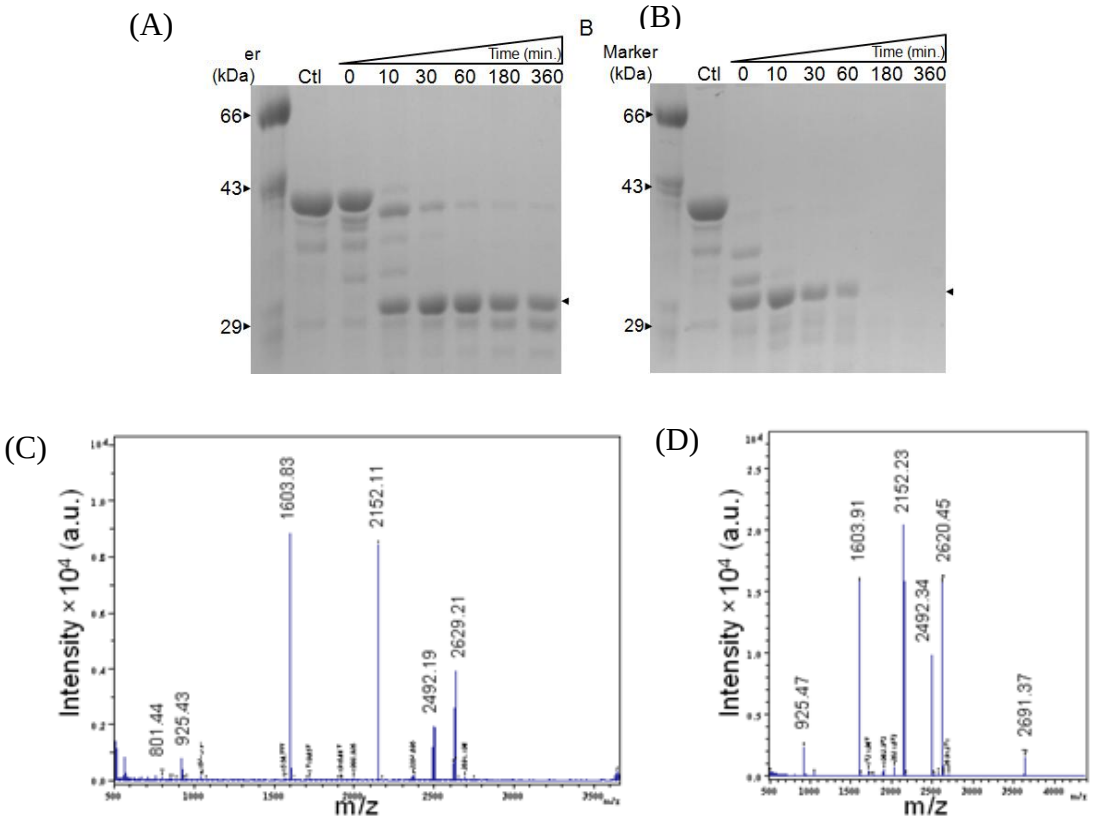

Figure 4.6: Limited proteolysis followed by mass spectrometry of MERIT40. Time dependent digestion of purified MERIT40 with (A) chymotrypsin and (B) trypsin (single arrow-digestion product). Peptides generated in mass spectrometry after digestion with (C) chymotrypsin and (D) trypsin digestion [168].

Figure 4.7: Matched peptides of proteolytically stable region in case of (E) trypsin and (F) chymotrypsin digestion.

prediction is one of the approaches for identifying the function of proteins in biological processes [135]. DALI server predicts the probable structural homologs of the protein (Holm & Rosenstrom 2010).

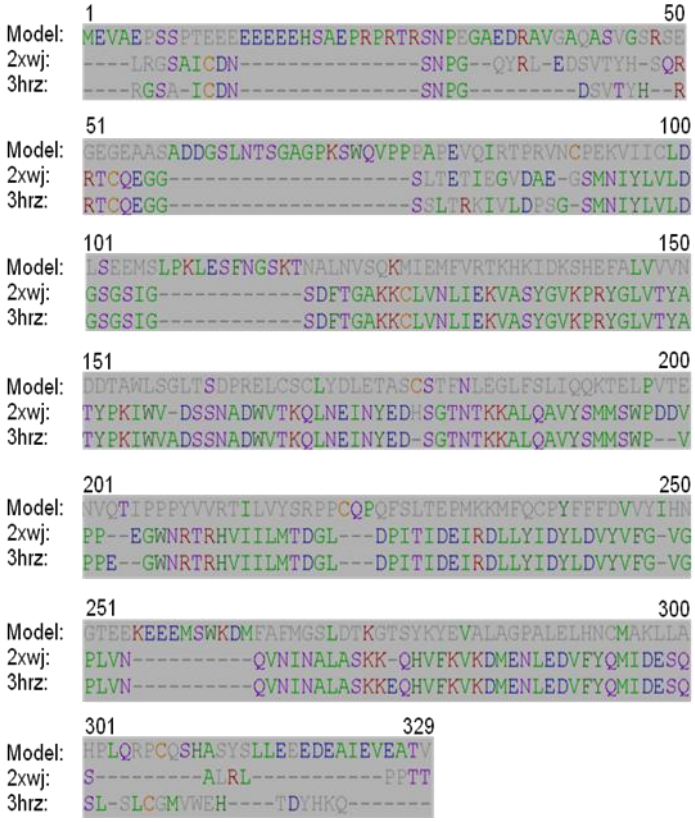

Figure 4.8: Multiple structural alignment of MERIT40 model with DALI searched templates [168].

DALI search was carried out to find the structural homologs of MERIT40 to predict its possible function [14]. Our search could find only three structural homologs, (i) Cobra Venom Factor (PDB ID: 3hrz), (ii) Complement Factor B (PDB ID: 2ok5) and (iii) Complement C3 Beta Chain (PDB ID: 2xwj) (**Figure 4.8**). Cobra Venom Factor (CVF) is a complement-activating protein in Cobra venom which plays a crucial role in host defense [180]. It resembles structurally and functionally with Complement C3 that plays an important role in complement activation and immunity (**Table 4.2**) [181]. The β-chain of Complement C3 is homologous to the α-chain of CVF, and the C-terminal and N-

terminal portions of Complement C3 α-chain are homologous to the α- and β- chains of CVF, respectively [182, 183]. Complement fragment (C3b)/CVF interacts with Factor B (FB), a single-domain serine protease which circulates in plasma, and forms the pro-convertase complex during complement activation. FB consists of a central vWA domain and a C-terminal trypsin like serine protease (SP) domain that form the protease segment. Structural similarity of MERIT40 vWFA like domain with a protease complex highlights its possible role in proteosome-like activity and involvement in ubiquitination processes.

Table 4.2: Structural comparison of MERIT40 with structurally similar protein using DALI [168]

Protein	PDB ID	Z-score	Sequence identity (%)	RMSD
Complement C3 Beta Chain	2XWJ	15.2	11	3.7
Cobra Venom Factor	3HRZ	14.9	10	4.1
Complement Factor B	2OK5	13.7	10	4.2
Complement C2a Fragment	2I6S	13.2	10	3.8
Complement C2	2ODP	13.0	10	3.9

4.3.5 Oligomeric properties of MERIT40: Oligomerization is a common property found among disordered and hydrophobic patch containing proteins [138]. These proteins exist in several oligomeric form which may or may not have biological significance. To determine the magnitude of oligomerization and types of different oligomeric species, MERIT40 has been characterized using size exclusion chromatography, native gel electrophoresis, glutaraldehyde cross-linking, mass spectrometry, Dynamic Light Scattering (DLS) and different *in-silico* tools. Gel filtration profile showed that the protein predominantly exists as a monomer, however, a second peak corresponding to dimer was also observed (**Figure 4.4**). Furthermore, to confirm the presence of dimeric population and to delineate the multimeric behavior of MERIT40, native gel electrophoresis was performed. The presence of two distinct populations on native gel

corresponding to monomer and dimer provided evidence that MERIT40 also exists in a dimeric form. In order to evaluate the proportion of dimeric population and to confirm whether it occurs in a concentration-dependent manner, different concentrations of MERIT40 were analyzed on native gel. We observed that MERIT40 exhibit a concentration-independent formation of dimers and the dimeric proportion remained same irrespective of increasing concentration from 0.1mg/ml to 10mg/ml (**Figure 4.9**). To support these finding and re-evaluate the concentration independent dimer formation, different concentrations of protein (0.1-10 mg/ml) were treated with glutaraldehyde and run on SDS-PAGE. Dimeric fractions remained same, irrespective of the changes in concentration, thus validating MERIT40 dimerization is concentration independent (**Figure 4.9**).

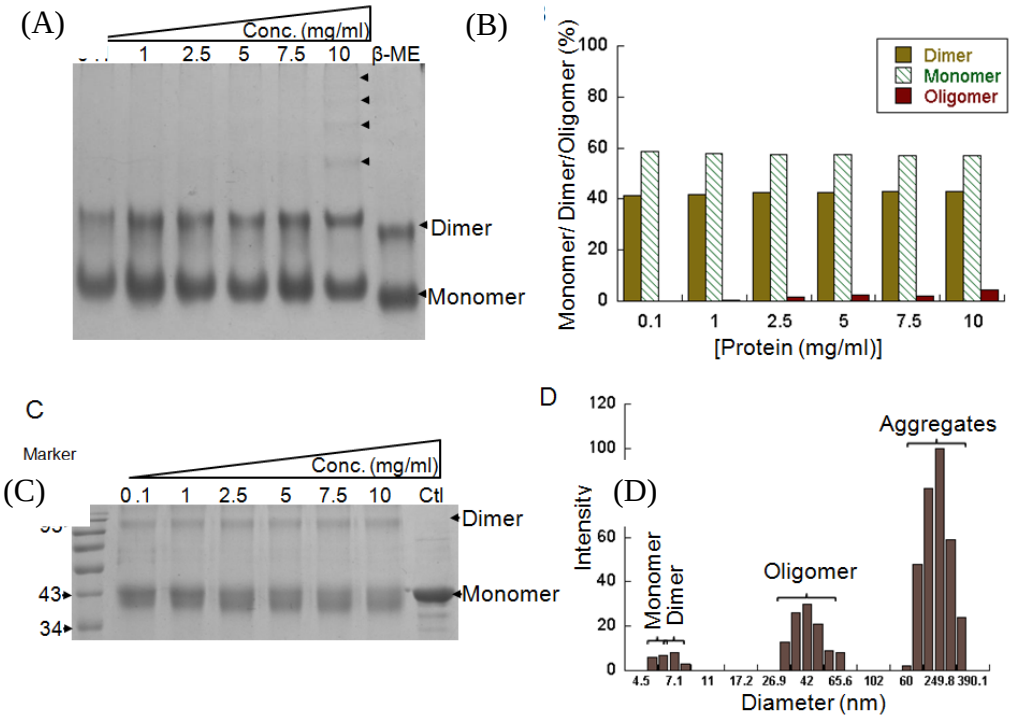

Figure 4.9: Oligomeric characterization of MERIT40. (A) Native gel (non denaturing), (B) densitometry profile native gel, (C) Glutaraldehyde Cross-linking reaction product on SDS-PAGE and (D) Dynamic light scattering. Last lane in native gel: β-ME treated sample Ctl: control as untreated with glutaraldehyde [168].

To further comprehend these findings, mass spectrometry MALDI-TOF was performed which revealed the presence of monomeric and dimeric population (**Figure 4.10**). It is often observed that molecules undergo dimerization due to inter-molecular disulfide linkage [184]. To investigate further whether dimer formation is occurring due to disulfide linkage, the purified MERIT40 protein was treated with β-ME and loaded onto native page. There was no effect found on the dimeric population even after β-ME treatment, indicating dimer formation probably does not occur due to inter-molecular disulfide linkage (**Figure 4.9**). Densitometric analysis suggested that the monomeric and dimeric protein fractions were found to be approximately 60% and 40% of the total protein, respectively, in a range of 0.1–10 mg/ml concentrations (**Figure 4.9**). Similar proportion was also observed on gel filtration chromatography (**Figure 4.4**). Hence MERIT40 exists as a dimer but no relative concentration dependency exists for dimer formation.

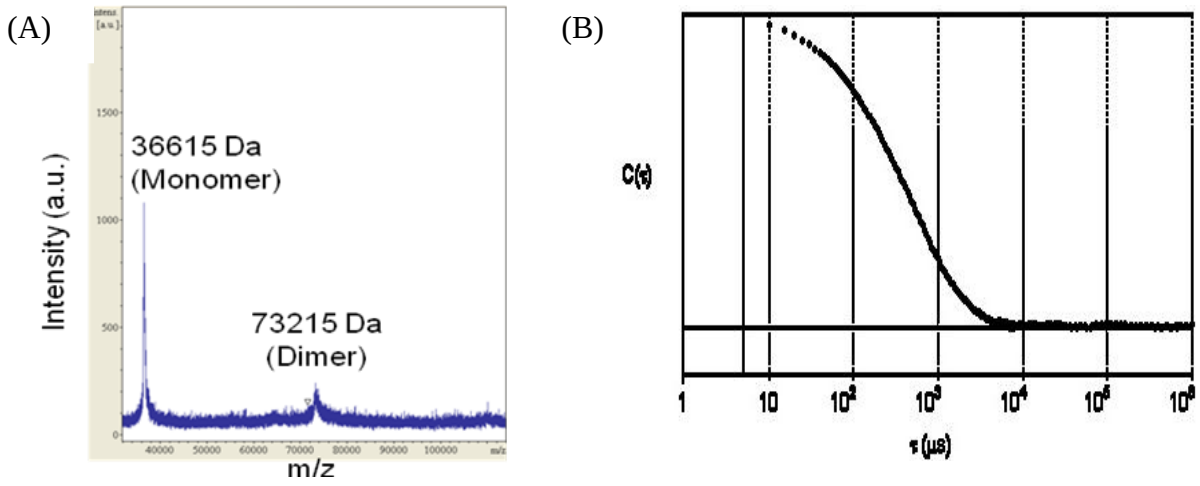

Figure 4.10: (A). MALDI-TOF mass spectrometry profile of full length MERIT40 showing presence of monomer and (B) dimer Correlation coefficient obtained in Dynamic Light Scattering analysis of MERIT40 for molecular size measurement [168].

To determine the heterogenity present in purified MERIT40 protein, Dynamic Light Scattering (DLS) experiment was performed. It was observed that majority of samples are composed of three different clusters. The first cluster corresponds to the monomeric and dimeric populations, while the second and third clusters correspond to higher-order oligomers or aggregates (**Figure 4.9, Figure 4.10**). The diameter of the monomer determined from modeled structure using in-silico approach was average 5–6 nm and the effective hydrodynamic diameter determined by the DLS was found in 5–6 nm range, which showed good concurrence with each other (**Figure 4.5, Figure 4.9**).

4.3.6 Dimer Interface of MERIT40: In order to determine the *in-silico* symmetrical dimer formation, blind docking of MERIT40 monomeric molecules was performed after molecular dynamic simulation (MDS) using Patchdock and Symmdock servers [137]. Only dimeric (diameter: 6–8 nm) form was observed in the docked structure (**Figure 4.11**). The dimeric structure was stabilized by involvement of Gly115, Gln125, Lys126 and Arg133 residues at the interface. To validate these findings, *in-silico* substitution of these residues with bulky and charge repulsive (Gln115, Ile125, Glu126 and Glu133 respectively), as well as Ala residues were made, followed by docking of monomers. Substitution of either of the residue in the monomers, disfavored the *in-silico* dimerization of MERIT40 (docking scores [185] of various substitution: wild-type 14640, Gly115Gln 13430, Gly115Ala 13408, Gln125Ile 14620, Gln125Ala 14614, Lys126Glu 14596, Lys126Ala 14308, Arg133Glu 14110, Arg133Ala 13392). The energy (kJ/mol) of dimeric complex due to Gly115Ala (-19.56), Gln125Ala (-6.78), Lys126Ala (0.84) and Arg133Ala (-19.43) was found higher as compared to wild- type (-47.92),

indicating relatively decreased stability of dimers due to substitution with the alanine residue.

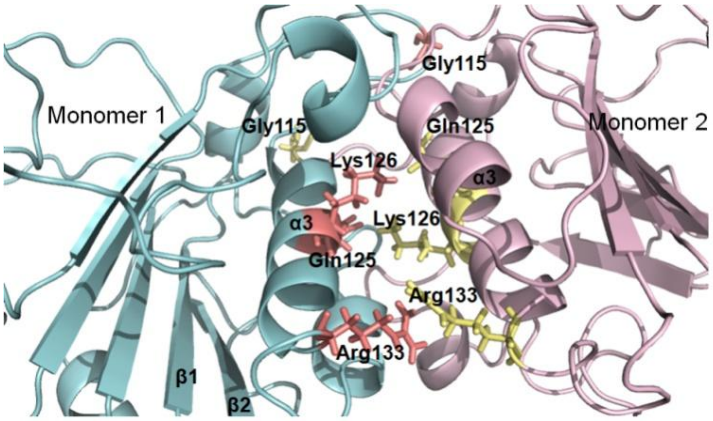

Figure 4.11: Dimer interface of MERIT40. Monomer1 interface residues are with light red color and monomer 2 with light yellow [168].

4.3.7 Structural analysis of MERIT40: To characterize the secondary structural constituents of MERIT40, FPLC purified MERIT40 protein was used to record far-UV Circular- Dichroism (CD) spectrum (**Figure 4.12**). The CD spectra were analyzed using DichroWeb server. It was observed that the α-helices and β-sheets are approximately 20 and 16.6%, respectively [186]. Furthermore, secondary structural components of modeled structure was determined using the PROSS server (http://roselab.jhu.edu/utils/runpross.html), which showed the percentage of α-helices and β-sheets to be ~ 23 and 15%, respectively. These results from the two independent approaches were in good agreement.

Tryptophan and tyrosine residues are the most frequently used intrinsic fluorophores to study the conformational changes induced in the protein due to external agents like temperature and chaotrophs [187]. These aromatic residues are generally buried inside the protein and are extremely sensitive to modifications around their micro-environment.

To study the three–dimensional structure of MERIT40, Trp and Tyr microenvironment were monitored using fluorescence spectroscopy.

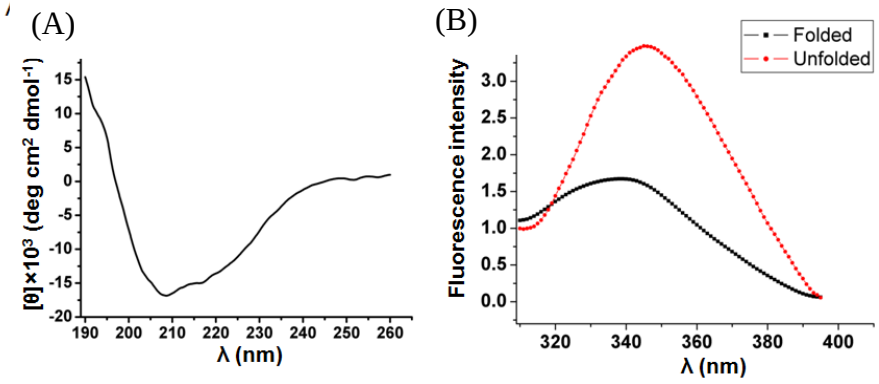

Figure 4.12: Far-UV CD spectra (A), and urea unfolding (B) using fluorescence, of MERIT40 [168].

An emission maximum at λ 338 nm was obtained for the native protein, while a red shift at λ 346 nm was observed for the unfolded protein (**Figure 4.12**). This indicates that most of the tryptophan and tyrosine residues of the native protein are buried inside the hydrophobic core.

4.3.8 Assessment of stability and unfolding pathway: To obtain the stability parameters and folding pattern of MERIT40, thermal denaturation was studied using CD and fluorescence spectroscopic probes. In Far-UV CD data, maximum change in ellipticity at different temperature was observed at λ208 nm, which was therefore used to monitor the changes in protein's secondary structure at varied temperatures (15-80°C) (**Figure 4.13**). Upon data fitting the T_m (melting temperature) was calculated to be 48±4.1°C, and thermodynamic parameters were $\Delta G°_{H2O}$ 4.8±1.0 kcal/mol, ΔH 144 ±2.5 kcal/mol. To further validate these findings, protein was unfolded in a temperature gradient of 15-80°C and unfolding was monitored by recording fluorescence emission for Trp, Tyr upon excitation at λ280 nm . The change in emission maxima of protein at various temperatures was used to determine its folding pathway and thermodynamics

parameters. Fluorescence spectroscopy studies showed a T_m value of 52±3.2°C and energy parameter ($\Delta G°_{H2O}$ 4.6±1.5 kcal/mol, ΔH 142±2.3 kcal/mol (**Figure 4.13)**. These parameters indicate that MERIT40 possessed moderate stability, and its enthalpy and free energy values are similar to globular protein [188].

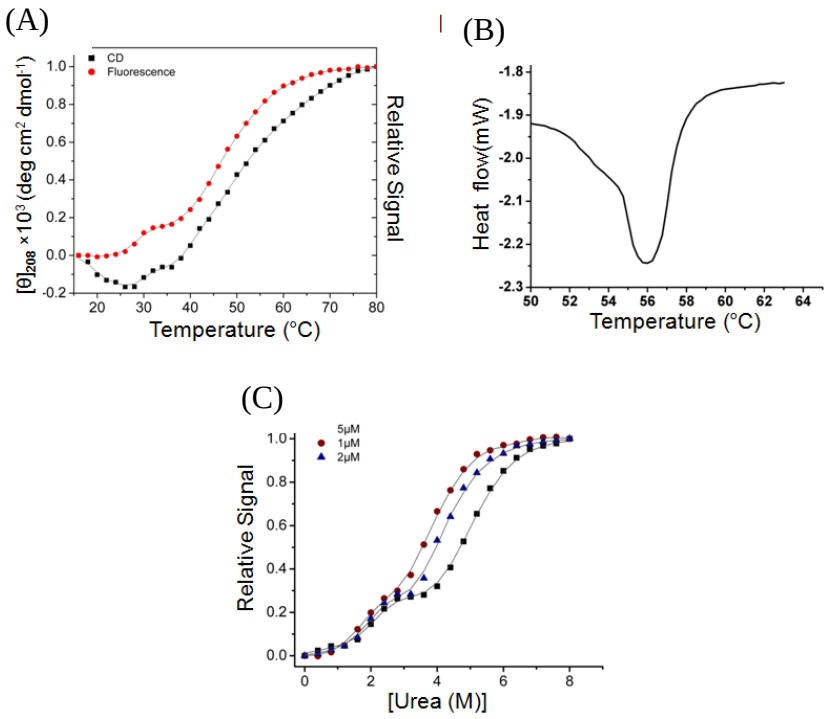

Figure 4.13: Denaturation profile of MERIT40 using CD, DSC and fluorescence spectroscopy. (A) Overlay of protein fractions unfolded at different temperature in CD and fluorescence spectroscopy, (B) DSC thermogram and (C) Chemical unfolding using fluorescence at different urea concentration. Relative signal represent fraction unfolded [168].

Both methods demonstrate that the protein most likely unfolds via the formation of an intermediate species as indicated by abrupt increase of signal in the unfolding curve. In order to validate these findings, the unfolding profile of native protein was assessed using differential scanning calorimetry (DSC). Purified MERIT40 protein in the concentration of 1 mg/ml was allowed to unfold in a temperature gradient of 15-80°C in calorimetry, and unfolding transition was obtained. A peak maximum during unfolding transition was considered as melting temperature (T_m). DSC thermogram suggested overlapping of two

transitions, which could be corresponding to intermediate species formed during the unfolding process **(Figure 4.13).** The melting temperature (55±0.9°C) and enthalpy parameters (ΔH 150±2.0 kcal/mol) were derived using Calisto software. The T_m value determined by DSC and fluorescence were similar and unfolding profile indicated that MERIT40 protein forms an intermediate species. Considering the dimeric behavior, it can be concluded that MERIT40 forms the dimeric intermediate during unfolding process. To test this hypothesis, a concentration dependent urea denaturation study was performed to find the type(s) of intermediate species formed during unfolding process. Protein was incubated in different urea concentration (0-8M) till equilibrium was achieved and emission maximum using fluorescence were recorded. Change in average emission wavelength at different urea concentrations was plotted to calculate unfolded fraction. A biphasic unfolding curve with two transitions was observed and protein concentration dependency in the second transition indicated that dissociation of MERIT40 subunits occurs due to a dimeric intermediate [189] (**Figure 4.13**). The chemical denaturation using urea as the unfolding agent substantiates the thermal denaturation findings.

4.3.9 Clinical implication of MERIT40 variants: Mutations or variants are known to exist almost in every member of BRCA1–complex. However, predicting the pathogenicity of mutations discovered in the cohort of patients is still challenging and is a subject of extensive studies [5, 190]. Mutations reported on BRCA1 has been reported to have role in cancer susceptibility [191]. Disease implication caused by these variants arise either due to alteration of whole domain structure or localized alteration in weak inter- or intra-molecular interactions [191]. To characterize the pathogenicity of sequence variants identified in the patient cohort and further evaluate the importance of weak intra-

molecular interactions in disease susceptibility, seven recently reported sequence variants of MERIT40 (c.342C > T, c.344 + 41A > T, c.393C > T, c.787 − 6C > T, c.821A > G, c.837G > A, c.87G > A) were analyzed.

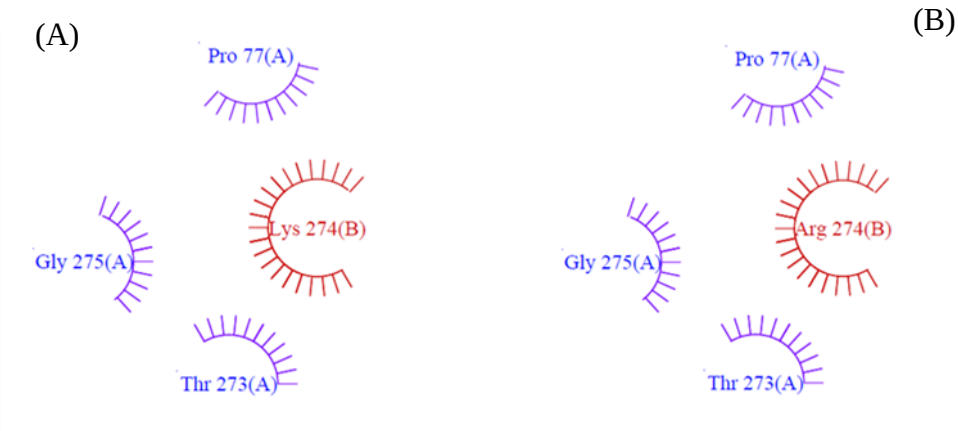

Figure 4.14: Mutational analysis of MERIT40. (A) ligplot of Lys 274. (B) ligplot of Arg 274 [168].

It is reported that none of the above variants have any role in predisposing the individual to disease [190]. All the variants are silent in nature except of c.821A > G in exon 9, which lead to the substitution of Arg to Lys at 274 position. However, change in Lys to Arg doesn't bring any pathological implication as observed in cohort of patients [190]. We decided to explore the cause of MERIT40 K274R variant neutrality. The molecular environment around the Lys274 residue and changes as a result of substitution with Arg was analyzed. Structural analysis of modeled structure unraveled that Lys274 is buried in the hydrophobic core of the protein which probably restricts its binding with other intra amino acids **(Figure 4.14)**. Lys274 forms several non-bonded interactions with Pro77, Gly75 and Thr273. There were no changes found in the structural conformation and hydrophobic environment around Lys274 residue of MERIT40 as a result of substitution with Arg and majority of interactions with adjacent residues remained unaltered **(Figure 4.14)**. Since intramolecular interactions around the mutant remained unaffected, it could

be assumed that MERIT40 K274R mutant is clinically insignificant. This assumption is in agreement with the findings of a previous report which states that mutations predisposing to breast cancer are either very rare or absent in the coding region of MERIT40 [190].

4.4 Crystallization of MERIT40

Initial crystallization trials for MERIT40 protein was set with 15 mg/ml protein concentration. The protein and mother liquor solution was mixed in 1:1 ratio (1 µl+1µl) and allowed to crystallize at 22°C with 500 µl reservoir solution in a closed system. A clear drop or light precipitation was observed in most of the drops.

4.5 Conclusion

MERIT40 can be classified as an intrinsically disordered protein due to the presence of N-and C-terminal disordered region. However, its middle region showed well defined compact structure which indicates a good possibility of crystallization of middle region for X-ray diffraction study. Structural homologs of MERIT40 suggests its plausible role in complement activation pathway, nevertheless, definite function determination needs further experimental evidences. Its central region shows remarkable stability towards protease digestion and has structural similarity with vWA-like region, a domain mainly present in complement activation factors. MERIT40 undergoes a three-state unfolding transition pathway with a dimeric intermediate. It exhibits concentration independent dimerization and unfolds through a dimeric intermediate. MERIT40 could be a multifunction molecule having role in DNA damage repair and complement activation.