ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

FUNCTIONAL ANALYSIS OF VUS (VARIANT OF UNCERTAIN SIGNIFICANCE) OF HUMAN MUTS HOMOLOG 2/6 (hMSH2/6) PROTEINS

M.Sc. THESIS

Celil Mert GÜL

Department of Molecular Biology-Genetics & Biotechnology

Molecular Biology-Genetics & Biotechnology Programme

AUGUST 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

FUNCTIONAL ANALYSIS OF VUS (VARIANT OF UNCERTAIN SIGNIFICANCE) OF HUMAN MUTS HOMOLOG 2/6 (hMSH2/6) PROTEINS

M.Sc. THESIS

Celil Mert GÜL
(521211103)

Department of Molecular Biology-Genetics & Biotechnology

Molecular Biology-Genetics & Biotechnology Programme

Thesis Advisor: Prof. Dr. Gizem DİNLER DOĞANAY

AUGUST 2023

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

İNSAN MUTS HOMOLOG 2 VE 6 (HMSH2/6) PROTEİNLERİNDE BULUNAN KLİNİK ÖNEMİ BELİRSİZ VARYANTLARIN FONKSİYONEL ANALİZİ

YÜKSEK LİSANS TEZİ

Celil Mert GÜL
(521211103)

Moleküler Biyoloji-Genetik & Biyoteknoloji Anabilim Dalı

Moleküler Biyoloji-Genetik & Biyoteknoloji Programı

Tez Danışmanı: Prof. Dr. Gizem DİNLER DOĞANAY

AĞUSTOS 2023

Celil Mert GÜL, a M.Sc. student of İTU Graduate School student ID 521211103, successfully defended the thesis/dissertation entitled "FUNCTIONAL ANALYSIS OF VUS (VARIANT OF UNCERTAIN SIGNIFICANCE) OF HUMAN MUTS HOMOLOG 2/6 (hMSH2/6) PROTEINS", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor:	**Prof. Dr. Gizem DİNLER DOĞANAY** İstanbul Technical University	.....
Jury Members:	**Assoc. Prof. Ceren ÇIRACI MUĞAN** Istanbul Technical University	.....
	Assoc. Prof. Timuçin AVŞAR Bahcesehir University	.....

Date of Submission : 25 July 2023
Date of Defense : 18 August 2023

To all believers,

FOREWORD

First of all, I would like to sincerely thank to Gizem DİNLER DOĞANAY, who has been always supportive and sensible to me and gave me a chance to study under her supervision by trusting me.

I appreciate all GLAB members specially Nisan Denizce CAN, Tuğba KIZILBOĞA, Özge ANAÇ, Ceren ÇİFTÇİ and Ecenur ÇEBİ for their collaboration and support during my studies.

I also want to thank my loving family members Akif Eren GÜL, Aynur KOZAN, Günnur GÜL, and Yaşar GÜL who have been always supportive and understanding, and with me throughout this study and to the beautiful people and my friends for their compassion, and confidence throughout this adventure.

This project is funded by TUBITAK project numbers 318S127 and 318S129.

August 2023 Celil Mert GÜL

TABLE OF CONTENTS

ABBREVIATIONS

ACMG : The American College of Medical Genetics and Genomics
CRC : Colorectal Cancer
MMR : Mismatch Repair
MSI : Microsatellite Instability
NGS : Next Generation Sequencing
VUS : Variant of Uncertain Significance
IB : Immunoblotting
WB : Western Blotting
SDS-PAGE : Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis
LS : Lynch Syndrome
HNPCC : Hereditary Non-Polyposis Colorectal Cancer
IPTG : Isopropyl-β-D-Thiogalactopyranoside
DNA : Deoxyribonucleic Acid
kDa : Kilodalton
PCR : Polymerase Chain Reaction
qPCR : Quantitative Polymerase Chain Reaction
WT : Wild Type
RNA : Ribonucleic Acid
PMSF : Phenylmethanesulfonyl fluoride
IMAC : Immobilized Metal Affinity Chromatography
AEX : Anion Exchange Chromatography

SYMBOLS

°C	**:** Celcius Degree
g	**:** Gram
M	**:** Molar
µM	**:** Micromolar
mM	**:** Minimolar
t	**:** Time
min	**:** Minute
OD	**:** Optical Density
Sec	**:** Seconds
h	: hours
RT	**:** Room Temperature
rpm	**:** Round per minute
v/v	**:** Volume to volume
w/v	**:** Weight to volume
kDa	**:** Kilodalton
mL	**:** Milliliter
V	**:** Voltage
µL	**:** Microliter
bp	**:** Basepair
kb	**:** Kilobase

LIST OF TABLES

Page

LIST OF FIGURES

Page

FUNCTIONAL ANALYSIS OF VUS (VARIANT OF UNCERTAIN SIGNIFICANCE) OF HUMAN MUTS HOMOLOG 2/6 (hMSH2/6) PROTEINS

SUMMARY

Lynch syndrome (LS) or hereditary non-polyposis colorectal cancer (HNPCC) is a genetic condition that raises the risk of colorectal cancer and related cancers. Germline genetic variants of the DNA mismatch repair genes *MSH2*, *MSH6*, *MLH1*, and *PMS2* are primarily responsible for its occurrence. In eukaryotes, MSH2 and MSH6 combine to form the MutSα complex, which is responsible for recognizing mismatches and assembling the necessary proteins for mismatch repair. Defining the functional consequences of variants is crucial in enrolling LS patients in appropriate surveillance programs to reduce morbidity and mortality.

Herein, the mutation profile of hereditary colorectal cancer in the Turkish population was determined by analyzing the variation spectrum of 26 cancer susceptibility genes in 371 patients with colorectal cancer using next-generation sequencing technology. The detected variants were interpreted based on The American College of Medical Genetics and Genomics (ACMG) recommendations. The MSH2 and MSH6 loci were screened for variants of unknown significance (VUS) to determine the effect of nucleotide substitution on protein function. Mismatch repair deficient cell line (LoVo) was transiently transfected with hMSH2 wild-type (MSH2-WT) and hMSH6 wild-type (MSH6-WT) genes and selected mutated subclones (MSH6-R577C, MSH6-S1279N, and MSH2-A733T). Regulation of mRNA expression and protein expression of interacting proteins were investigated, which showed hMSH6 gene expression was decreased when LoVo cells were transfected with MSH2-A733T compared to MSH2-WT. Expression levels of the downstream targets and interaction partner proteins were not affected significantly due to mutated forms of proteins. *In-vitro* binding assays showed that mutations did not affect the interaction between MSH2 and MSH6. Purified MSH2-WT and MSH2-A733T proteins that were produced from HEK293T cells were not obtained in stable forms. Bacterial co-expression of genes exhibited soluble protein production, and purification method for wild type proteins was promoted. However, due to the low stability of high molecular weighted proteins, it was decided that this method could be used in domain-specific studies with improvements. In order to reclassify clinical importance of the selected VUS, functional studies are needed to be improved. Reclassification of VUS in *MSH2/MSH6* has the potential to improve variant classification accuracy, increase risk assessment, facilitate recommended clinical decision making, and provide more accurate genetic counseling for affected individuals and their families. Consequently, implementing personalized management strategies can effectively reduce cancer mortality and morbidity in Lynch syndrome populations.

İNSAN MUTS HOMOLOG 2 VE 6 (HMSH2/6) PROTEİNLERİNDE BULUNAN KLİNİK ÖNEMİ BELİRSİZ VARYANTLARIN FONKSİYONEL ANALİZİ

ÖZET

Lynch sendromu (LS) veya kalıtsal polipsiz kolorektal kanser (HNPCC), kolorektal kanser ve ilgili kanserler riskini artıran genetik bir semptomdur. Lynch sendromunun ortaya çıkmasında DNA uyuşmazlığı onarım genlerinde MSH2, MSH6, MLH1 ve PMS2'de gözlemlenen germ hattı genetik varyantları birincil dereceden rol oynamaktadır. Ökaryotlarda MSH2 ve MSH6 proteinleri, DNA'daki uyumsuzlukları tanımaktan ve uyumsuzluk onarımı için gerekli proteinleri bir araya getirmekten sorumlu olan MutSα kompleksini oluşturmaktadır. Belirsiz öneme sahip varyantların (VUS) fonksiyonel sonuçlarının tanımlanması, morbidite ve mortaliteyi azaltmak için LS hastalarını uygun sürveyans programlarına kaydetmede çok önemlidir. Bu çalışmada kalıtsal kolorektal kanserin Türk popülasyonundaki mutasyon profili, kolorektal kanserli 371 hastada 26 kansere yatkınlık geninin varyasyon spektrumunun, yeni nesil dizileme teknolojisi kullanılarak analiz edilmesiyle belirlenmiştir. Tespit edilen varyantlar, Amerikan Tıbbi Genetik ve Genomik Koleji (ACMG) standartlarına göre yorumlanmıştır. MSH2 ve MSH6 lokusları, nükleotit değişiminin protein fonksiyonu üzerindeki etkisini belirlemek için önemi bilinmeyen varyantlar (VUS) açısından taranmıştır. Uyumsuzluk onarımı bakımından eksik hücre hattı (LoVo), hMSH2 vahşi tip (MSH2-WT) ve hMSH6 vahşi tip (MSH6-WT) genleri ve seçilmiş mutasyona uğramış alt klonlar (MSH6-R577C, MSH6-S1279N ve MSH2-A733T) ile geçici olarak transfekte edilmiştir. Etkileşen proteinlerin mRNA ifadesindeki değişiklikler ve protein ifadesi araştırılmıştır. Sonuç olarak, MSH2-WT'ye kıyasla LoVo hücreleri MSH2-A733T ile transfekte edildiğinde *hMSH6* gen ekspresyonunun azaldığı gözlemlenmiştir. Aşağı akış ve etkileşim ortağı proteinlerinin ekspresyon seviyelerinin, mutasyona uğramış protein formları varlığında önemli ölçüde etkilenmediği görülmüştür. *İn-vitro* bağlanma deneyleri, mutasyonların MSH2 ve MSH6 arasındaki etkileşimin bozulmadığını göstermiştir. HEK293T hücrelerinden üretilen saflaştırılmış MSH2-WT ve MSH2-A733T proteinleri stabil formlarda elde edilememiştir. Bakteri konağının yabanıl tip gen dizilerini taşıyan iki vektörle eş zamanlı transforme edilmesi sonucu proteinlerin çözünür formda üretilmesi sağlanmış ve saflaştırma protokolü oluşturulmuştur. Ancak proteinlerin boyutlarının büyüklüğünden kaynaklanan düşük verimden dolayı bu metodun domain spesifik çalışmada kullanılabileceği ve optimize edilmesi gerektiğine karar verilmiştir. Seçilen klinik önemi bilinmeyen varyantların yeniden sınıflandırılabilmesi için fonksiyonel çalışmaların geliştirilmesine ihtiyaç duyulmaktadır. *MSH2/MSH6*'da tespit edilen VUS'ların yeniden sınıflandırılması, varyant sınıflandırma doğruluğunu artırma, risk değerlendirmesini artırma, tavsiye edilen klinik karar vermeyi kolaylaştırma ve etkilenen bireyler ve aileleri için daha doğru genetik danışmanlık sağlama potansiyeline sahiptir. Kişiselleştirilmiş yönetim stratejilerinin uygulanması, Lynch sendromu popülasyonlarında kanser mortalitesini ve morbiditesini etkili bir şekilde azaltabilir.

1. INTRODUCTION

1.1 Hereditary Cancer Syndromes

Most hereditary cancer syndromes—which make up 5–10% of all malignancies—are inherited in an autosomal dominant fashion, meaning that only one copy of a faulty gene is necessary for the disease to manifest. A higher risk of early-onset tumor development exists in those with genetic cancer syndromes (Rahner and Steinke, 2008).

The prevalence of colorectal cancers is the consequence of sporadic genetic events, yet 2-5% of colorectal cancers are due to germline pathogenic variants (Jasperson et al., 2010).

1.1.1 Colorectal cancer statistics

According to the World Health Organization (WHO), with 1.9 million new cases and 935,000 fatalities in 2020, colorectal cancer becomes the third most often diagnosed cancer and the second major cause of cancer-related deaths worldwide (Sung et al., 2021). With a total of 21,191 new cases recorded in 2020, colorectal cancer was the third most often diagnosed cancer in both men and women in Turkey (Yanik & Bakar-Ates, 2023). In Turkey, the incidence of colorectal cancer was 17.5 per 100,000 people, while the fatality rate was 7.5 per 100,000 people. Furthermore, over the past ten years, it has been noted that both the incidence and mortality rates of colorectal cancer have been rising in Turkey (Yanik & Bakar-Ates, 2023; Sung et al., 2021).

1.1.2 Lynch syndrome

Lynch syndrome, also known as Hereditary Non-Polyposis Colorectal Cancer (HNPCC), is a genetic condition associated with an increased likelihood of developing several cancers, most notably colorectal and endometrial cancer as well as stomach, small intestine, kidney, ureter, and ovarium cancers (Jass, 2006). It is a rare condition, inherited in autosomal dominant pattern and caused by the germline mutations that found in DNA Mismatch Repair (MMR) genes (*MLH1, MSH2, MSH6,* and *PMS2*). As

a result of these alterations, the MMR genes' ability to protect against DNA damage and mutations may be compromised. Mutations found in *MLH1* and *MSH2* genes are responsible for 70-80% of all instances of Lynch syndrome, whereas *MSH6* and *PMS2* mutations only account for 10-20% of such cases (Baglietto et al., 2010; Jass, 2006).

Genetic testing is frequently used to diagnose Lynch syndrome and find the alterations in the MMR genes. Lynch syndrome patients can benefit from screening and surveillance procedures like routine colonoscopies and other imaging tests to either discover cancer early or inhibit it from progressing (Baglietto et al., 2010).

1.2 MMR Genes

MMR genes are a group of genes that belong to genetic susceptibility genes isolated from hereditary non-polyposis colorectal cancer (HNPCC), and *hMSH2* gene is the first cloned MMR gene by Fishel et al. (1993). In total, 9 MMR genes have been identified; MutS homologs (*MSH2*, *MSH3*, *MSH4*, *MSH5*, and *MSH6*), MutL homologs (*MLH1*, *MLH3*) and post-meiotic segregation increased (*PMS1*, *PMS2*). The primary objective of these genes is to sustain the robustness and stability of the genetic material by rectifying mismatched bases. Effectively functioning MMR genes direct impaired DNA to apoptosis and avert the generation of malignant cells with mutations. Due to the presence of defective mutations in MMR genes, genetic stability may be impaired, and the cell progresses towards tumorigenesis formation (He et al., 2022).

Malfunction of the mismatch repair (MMR) genes might cause either directly or indirectly the inability to repair altered coding areas, which can activate proto-oncogenes and deactivate tumor suppressor genes. Subsequently, this can lead to the unrestrained proliferation and differentiation of cells, resulting in tumorigenesis. Thus, MMR genes are accepted as the third class of genes that are strongly linked to carcinogenesis in addition to oncogenes and tumor suppressor genes (He et al., 2022; Da Silva Calixto et al., 2018).

1.2.1 Microsatellite instability (MSI)

Approximately 3% of the human genome is composed of microsatellites, which are repeating units of 1 to 6 base pairs (Kelkar et al., 2010) Tumor microsatellite instability (MSI) is caused by pathogenic variants in the MMR genes, which lead to repair errors in repetitive sequences. MSI occurs in numerous tumor tissues, but it is characteristic

of colorectal cancer. Approximately 15% of all colorectal malignancies and 90% of colorectal cancers in patients with Lynch syndrome are associated with MSI (Kurzawski et al., 2004; Gian, 2018).

High levels of MSI (MSI-H) are typically associated with mutations within the *hMLH1* and *hMSH2* genes (Martin-Lopez & Fishel, 2013). A low level of MSI (MSI-L) appears to be primarily caused by mutations in the *hMSH6* (10%) and *hPMS2* (5%) genes (Gian, 2018). Cause of approximately 5% of MSI malignancies is unknown. MSI is present in greater than 95% of LS/HNPCC tumors, but only 10–15% of sporadic colorectal malignancies (Gian, 2018). Importantly, once reliable markers were established, MSI has become a reliable indicator of MMR defects in human malignancies. In addition, 15% of sporadic cases of colorectal cancer have been identified in association with epigenetic inactivation of MLH1 (Martin-Lopez & Fishel, 2013).

1.2.2 DNA damage

DNA is an organic molecule that carries genetic information from generation to generation. DNA damage can be defined as any change that alters the coding properties, functions or structure of DNA. DNA, which can be damaged by various environmental factors such as radiation, chemicals, or ultraviolet light, can also be damaged by endogenous factors such as reactive oxygen species and reactive nitrogen species produced by cellular mechanisms. Moreover, DNA can be affected during natural processes such as DNA replication, homologous recombination, or DNA repair.

DNA lesions can take many diverse forms: single-strand breaks (SSBs), double-strand breaks (DSBs), bulky adducts, UV-induced photoproducts, DNA-protein cross-links, and insertion/deletion mismatches (Martin, 2008).

1.2.3 MMR proteins in DNA damage response

Halting the cell cycle is a crucial defense mechanism against DNA damage-induced genomic instability. Proteins necessary for cell cycle arrest, such as ATM, ATR, p53, p73, Chk1, and Chk2, have been found and described by a considerable amount of research (Li, 2008). Cell cycle arrest in response to many types of DNA damaging agents is faulty in hMutS- and hMutL-deficient cells. It has been observed that cells

lacking MMR are unable to phosphorylate p53 and p73 in response to DNA damage (Li, 2008). During the response to DNA damage, the kinases ATM, ATR, and/or c-Abl phosphorylate p53 and p73, and physically in the interaction with hMutS and hMutL. These results suggest that hMutS and hMutL play a role in the signaling pathway that initiates cell cycle arrest and/or apoptosis in response to DNA damage (Duckett et al., 1996).

1.3 DNA Mismatch Repair Mechanism

MMR is the highly conserved cellular post-replication process that maintains DNA homeostasis and thus plays a major role in ensuring genomic stability (Pećina-Šlaus et al., 2020) (Figure 1.1). The main purpose of the mismatch repair mechanism is to correct mispaired nucleotides or small insertion-deletion loops. Correction of the errors depends on the mismatch repair mechanism proteins; if they cannot properly function, accumulation of the errors increases the mutational rate of the cells. The performance of the DNA polymerase is not entirely flawless, DNA polymerase might result in 1 error in 10000 nucleotides during duplication of double strands even though DNA polymerases have a proofreading activity (Pećina-Šlaus et al., 2020). The second and most efficient defense mechanism to correct a misincorporation in the strand is the mismatch repair mechanism.

The MMR system is an assurance that takes charge of not transmitting any anomalies in DNA to future generations. Therefore, spontaneous mutation rate increases by virtue of any deficiency in the system. MMR deficiency in humans is related to hereditary and sporadic human cancers since the system is required for cell cycle arrest or programmed cell death in the response of certain types of DNA damage (Li, 2008).

MMR deficiency or dysfunction in humans is linked to a variety of hereditary and sporadic cancers, particularly colorectal, gastrointestinal, and endometrial cancers since the system is necessary for cell cycle arrest or programmed cell death upon exposure to various types of DNA damage (Li, 2008).

1.3.1 In E. coli

Escherichia coli was the first organism whose mismatch repair machinery was identified (Lahue et al., 1989). It has been discovered that the mechanism is composed of 4 main steps: recognition, excision, resynthesis, and ligation. Recognition is

performed by the MutS complex, which binds to mismatched base and recruits MutL into MutS-mismatched base complex by undergoing conformational changes. Main role of the MutL is to provide the link between MutS and downstream proteins. Next, MutH having endonuclease activity binds to hemimethylated GATC sequence and creates a nick in the newly synthesized unmethylated strand. The excision starting point that determined by MutH continues over UvrD which is a helicase that unwinds the duplex DNA enhancing nick towards the mismatch. It generates single-stranded DNA that is expeditiously protected by single-stranded DNA binding proteins (SSB), making it a target for DNA polymerase III to resynthesize. It is followed by DNA ligase to seal the gaps between nucleotides (Lahue et al., 1989; Modrich, 1996).

1.3.2 In humans

Evolutionarily conserved MMR mechanism in humans is similar to the indicated *E. coli* MMR in some aspects, including nick-directed strand specificity and bidirectional exonuclease activity, as well as substrate specificity. Mammalian MMR system proteins have been specified regarding their homology to *E. coli* MMR (Lahue et al., 1989).

MSH proteins or MutS homologues and MLH proteins or MutL homologues act as heterodimers, contrary to homodimer formation in *E. coli*, providing mismatch recognition and initiation of repair. hMSH2 heterodimers with hMSH6 to form hMutSα complex that is used for the recognition of base-base mismatches or small insertion-deletion loops in 1-2 nucleotide lengths, while hMutSβ is composed of hMSH2 and hMSH3 that preferentially recognize unpaired insertion-deletion loops larger than 2 or more nucleotides (Li, 2008). Human MutL homologs hMLH1 heterodimerize with hMLH3, hPMS1 and hPMS2 to form MutLα, MutLβ, and hMutLγ complexes, respectively (Pećina-Šlaus et al., 2020). However, only MutLα complex plays a role in human MMR as a molecular intermediary and conducts mismatch induced excision. Another player is Proliferating Cell Nuclear Antigen (PCNA) which is a homotrimeric complex that conveys the processivity of DNA polymerase by encircling it as a sliding clamp (Devakumar et al., 2019).

PCNA interacts with MSH3 and MSH6 through their PIP (PCNA-Interacting Peptide) motifs, which tether MutSα to PCNA on replication forks. Activation of MutLα complex is induced by the orientation of PCNA (Devakumar et al., 2019). It establishes

strand specificity and controls the excision of mismatched bases from newly synthesized DNA. PCNA must be present for 3’ nick-directed MMR, however, it is not fundamental for 5’ nick-directed MMR.

On the other hand, RFC or replication factor C is involved in the loading of PCNA into the proximity of mismatched DNA during repair (Devakumar et al., 2019).

A 5’-3’ exonuclease, Exonuclease 1 (EXO1), is one of the officials of MMR for both 3’ and 5’ nick-directed MMR. The nick, which is created by the MLH1:PMS2 complex (MutLα) which is activated by PCNA and RFC, is enlarged to a broadened gap by EXO1. EXO1 interacts with MutS complexes and RPA; it is not essential for it to form a link with MutL during 5' directed excision, yet it does so throughout 3' directed excision, where MutL's endonuclease activity functions as a cryptic agent (Li, 2008; Martin-Lopez & Fishel, 2013).

RPA is a protein that binds to single-stranded DNA and participates in every stage of MMR. Prior to the formation of the MutS and MutL complexes, it attaches to nicked heteroduplex DNA and boosts mismatch-induced excision, shields ssDNA in the nuclease-open area, and promotes DNA synthesis (Pećina-Šlaus et al., 2020).

High mobility group box 1 (HMGB1) protein is a mismatch binding protein that possesses DNA-unwinding activity and collaborates with the MutSα complex (Martin-Lopez & Fishel, 2013).

Then, DNA polymerase δ is recruited to the gapped region, where it begins to resynthesize DNA by taking template strand as a guide. The MMR process in humans is ended by DNA Ligase, which seals the gaps between added newly synthesized nucleotides (Martin-Lopez & Fishel, 2013).

1.4 MSH Proteins

hMSH2 protein, the primary regulator of the MMR system in humans, is encoded by the *hMSH2* gene, which is located at chromosome 2p21 (Pećina-Šlaus et al., 2020). It contains 934 amino acids, which weigh around 105 kDa. It heterodimerizes with hMSH3 or hMSH6 to form complexes hMutSα or hMutSβ, respectively. Since there is 10 times higher expression rate of *hMSH6* compared to *hMSH3*, hMutSα is more abundant in human cells.

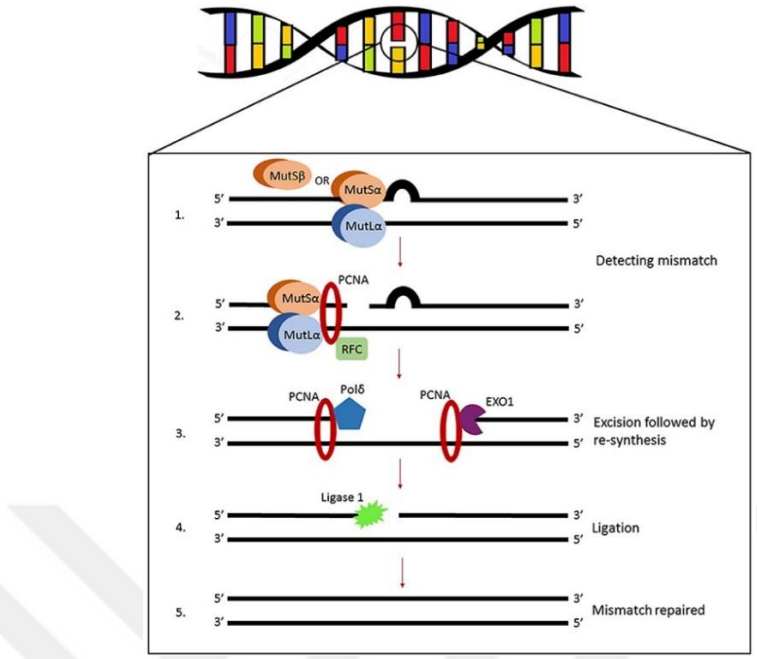

Figure 1.1: Illustration of mismatch repair mechanism (Pećina-Šlaus et al., 2020).

hMSH6 protein, which is encoded by hMSH6 gene, contains 10 exons that cover 24 kb of genomic sequence and is located on the 2p16.3 chromosome (Pećina-Šlaus et al., 2020). The molecular weight of the hMSH6 protein is approximately 160 kDa, and it includes 1360 amino acids in total. It is unstable in the absence of its heterodimer partner hMSH2, therefore occupies 80-90% of the free hMSH2 in the cell.

hMSH2 and hMSH6 share 5 conserved domains: mismatch binding, connector, lever, clamp region, and ATPase. In addition to the common domains, hMSH6 has an additional N-terminal disordered domain. The MutS dimer is formed via the juxtaposition of two asymmetrical mirror images, each containing a mismatch binding and ATPase domain (Edelbrock et al., 2013).

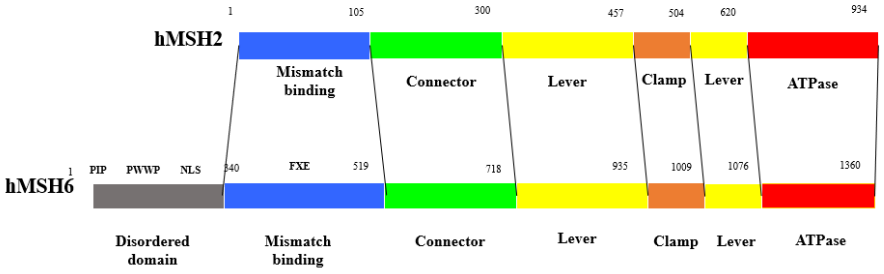

Figure 1.2: hMSH2/hMSH6 structural domain homology (Created with BioRender.com).

MutS has been observed to bind asymmetrically to bases with mismatched nucleotides. Consistent residues in the mismatch binding domain interact with the DNA duplex. Compared to the analogous region of hMSH2 (amino acids 1-124), the region of hMSH6 containing amino acids 362–518 allows for a stronger association with the DNA. Furthermore, the mismatched nucleotide forms a robust DNA bond with hMSH6 in this region. Unlike hMSH2 or hMSH3, eukaryotic hMSH6 retains a specific Phe-X Glu motif that confers mismatch binding affinity in bacterial MutS (Edelbrock et al., 2013; Warren et al., 2007).

hMutSα has been found to be able to hydrolyze ATP in addition to its DNA recognition abilities. There are two non-equivalent essential ATP hydrolytic sites that are distinct from one another as a member of ATP binding cassette (ABC)-transporter superfamily at the C termini of hMSH2 and hMSH6. Both sites have residues that are interacting with one another to produce ATPase sites. The ATPase site of hMSH2 is located between the 620th and the 934th amino acids, whereas the ATPase site of hMSH6 is located between the 1076th and the 1360th amino acids. There are nine distinct potential combinations of ADP, ATP, or vacancies for each protein; as a consequence, MutS may take on a variety of structural conformations and perform a wide range of functions. On the other hand, hMSH2 has a higher affinity for ADP, whereas hMSH6 has a higher affinity for ATP when it is in an unbound condition to DNA. Although MutS possesses an intrinsic ATPase activity, *in-vitro* MMR cannot proceed without the hMutSalpha-driven hydrolysis of ATP (Warren et al., 2007; Gorman et al., 2007).

The connector domains of hMutSalpha are necessary for MutL to interact with hMutSalpha. Although the structurally conserved connector domain of hMSH2 (between the 125th and 300th amino acid) is necessary for the formation of the MutS-MutL ternary complex between MSH2-MSH6-PMS1-MLH1, the connector domain of hMSH6 (between the 519th and 718th amino acid) interacts with the HMGB1 protein. This interaction takes place not only through hMSH6, but also within other members of the MutS family, such as MSH3 and MSH5 (Warren et al., 2007).

Lever domains (between 300^{th}-457^{th} and 554^{th}-620^{th} amino acid for hMSH2, and between 718^{th}-935^{th} and 1009^{th}-1076^{th} amino acid for hMSH6) and the clamp domains (between 457^{th} and 554^{th} amino acid for hMSH2, 935^{th} and 1009^{th} for hMSH6) of MutSalpha complex is responsible to contact bent DNA backbone. It plays an

important role for exhibiting conformational changes during DNA contact (Warren et al., 2007; Gorman et al., 2007).

In the disordered domain of hMSH6 (between 1-389 amino acids), PWWP sequence is present, which serves non-specific DNA–protein interactions and interaction with chromatin associated proteins. Also, conserved PCNA interaction protein (PIP) motif (QXX(L/I) XXFF is found in the N-terminal end, which is necessary for coordination of MMR at the replication fork and co-localization of PCNA with hMSH6. Moreover, N-terminal region of hMSH6 includes 3 nuclear localization signals (NLSs), where the nuclear import is triggered owing to phosphorylation events. In addition, there are 22 distinct phosphorylation sites on hMSH6 even though the post-translational regulation of MutSalpha is still not clear (Kaliyaperumal et al., 2011). On these sites, Cyclin Dependent Kinase (CDK) and Ataxia Telangiectasia Mutated (ATM)/ ATM- and Rad3-Related (ATR) recognition motifs have been found, which may be related with cell cycle and DNA damage responses. It has been detected that the after phosphorylation by Protein Kinase C (PKC) and Casein Kinase 2 (CK2) *in-vivo*, it resulted with increased MMR, nuclear translocation, and chromosomal binding (Warren et al., 2007).

hMSH6 and hMSH2 as well as other MMR proteins such as MLH1, MSH3, and PMS2 are constitutively expressed in G1 phase, with the increasing expression rate during S and G2 phases. MSH2 and MSH6 transcriptional regulation studies showed that constitutive mRNA expression is regulated by SP1 transcription factor from GC-rich binding sites of promoter region. During and after entry into S phase, gene expression of *MSH2* and *MSH6* is upregulated by E2F1 and E2F3 transcription factors, while E2F7 promoter element represses transcription of *MSH2* and *MSH6* at the end of the S phase. MutSalpha is tough to be degraded by ubiquitin-proteasome system after G2 phase (Gorman et al., 2007; Mendillo et al., 2009).

MutSalpha not only recognizes base-base mismatches but also acknowledges specific lesions including O6meG, complex pyrimidine dimers, halogenated pyrimidines, bulky adducts such as benzo[c]phenanthrene dihydrodiol epoxide, and cisplatin adducts, 8-oxoGuanine.

Crystal structures of hMutSalpha complex were obtained bounded with G: T, single base T insert or O6-meG-T, or G: U mismatch forms of oligomers. The structure of

full length MSH2 and 5 domains of the MSH6 without its disordered domain are available. According to the structure, it is determined that the MutSα and the DNA binding is not altered depending on the structure of DNA, meaning that it doesn't signal different downstream pathways (Warren et al., 2007; Gorman et al., 2007).

1.5 Multigene Panel Testing and Variant Interpretation

1.5.1 Multigene panel testing by next generation sequencing

Next-generation sequencing, frequently referred to as NGS, is a high-throughput sequencing technology that enables feasible determination of larger segments of an organism's genome from multiple specimens at the same time. Despite producing amplicons, it is generating millions of DNA sequences in parallel. After immobilizing fragments on a solid surface and initiating the sequencing reaction, next-generation sequencers can read sequences with a length ranging from 35 to 500 base pairs. As a part of the advancements in molecular biology and other relevant genomic research, sequencing technologies have been utilized in clinical studies to examine whole-genome sequencing, whole-exome sequencing, prenatal testing, and multi-gene analysis (Seifi et al. 2017).

To accurately forecast the risk of a person developing cancer, genetic testing is extremely crucial. According to Plon et al. (2008) and Genetic Alliance (2009), doing an analysis of the genomic DNA of individuals using sequence-based approaches can uncover mutations that have the potential to interfere with the normal gene functioning.

The genetic screening technique known as "multigene panel testing" is capable of simultaneously analyzing numerous genes for the purpose of locating probable genetic alterations that may be contributing factors in the development of a certain illness or condition.

The sequencing of five to eighty or more genes is the goal of multi-gene panel testing (Kurian & Ford, 2015). Decisions about which genes should be included in the test are made by the physician based on the patient's medical history and the history of the patient's family.

The multigene panels are intended to conduct tests for genes that are known to be linked to a specific ailment or a set of disorders that are closely connected to one

another. Clinicians decide germline testing based on the kind of disease being treated and the gene panels that are available. These panels have the potential to include genes connected to hereditary cancer syndromes, cardiovascular diseases, neurological problems, and a wide variety of other ailments. Common multigene panel tests can be listed as Color Hereditary Cancer Test (DynaCare): 30 genes associated with hereditary cancer syndromes, CancerNext-Expanded (Ambry Genetics): 80 cancer predisposition genes, Genomic Unity Comprehensive Cancer Panel (Fabric Genomics): Over 300 cancer-related genes.

Multigene panel testing by using NGS strategy has various benefits over conventional testing of a single gene, including an enhanced diagnostic yield, cost-effectiveness, and the capability to find mutations in genes that may not have been suspected based on the clinical presentation or family history of the patient.

1.5.2 Variant interpretation

The interpretation of genetic variations is extremely important since it is utilized to establish the clinical significance of variants throughout the diagnostic or preventative processes. Standard guidelines are utilized during the process of variant classification, such as those established by the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP). The ACMG categorizes variations in a progressive manner using a five-stage system: pathogenic, likely pathogenic, unknown significance, likely benign, and benign. Pathogenic variations are directly responsible for the development of illness, although the disease does not necessarily have to be completely penetrant. In conditions that are recessive and X-linked, a single pathogenic variant may not cause disease; however, this does not alter the classification of the condition. Variants that have disease-causing effects that are roughly 90–100% of the time or higher are considered to be likely pathogenic. In situations like this, further scraps of evidence are required to establish that the variation in question has a pathogenic effect. According to Genetic Alliance (2009), a variation of uncertain significance indicates that there is no information that can be supported by pieces of evidence about the issue of whether the variant is pathogenic or not. In addition, according to the research that has been conducted up to this point, it has been determined that presumably benign variations do not have a remarkable influence on disease; nevertheless, further data is required to corroborate this assertion.

On the other hand, benign variations are tolerated since they do not result in any adverse effects or contribute in any way to the development of illness. The ACMG criteria are dependent on demographic data, computational and predictive data, functional data, segregation data, de novo data, and allelic data. The presence of very strong, strong, moderate, or supportive evidence is taken into consideration when determining the stage of clinical significance of a variation. According to Richards et al. (2015) and Plon et al. (2008), the information that may be gathered from genetic testing can have potentially favorable consequences for the diagnosis of a problem, the beginning of a new therapy, or the use of preventative techniques.

In addition, the interpretation of test findings might be difficult, and it is necessary to have competence in genetics and genetic counseling in order to guarantee that decisions on management and treatment will be suitable. Therefore, individuals who are considering multigene panel testing are strongly encouraged to seek genetic counseling first.

1.6 Purpose of the Study

Interpretation of variants is of great importance for both the LS patient and their families. In the study, cancer susceptible genes of 371 colorectal cancer patients were investigated by multi-gene panel testing. After variant interpretation was performed to classify pathogenicity level of variants, variants of uncertain significance (VUS) of MSH2 and MSH6 genes were selected for functional studies. The purpose of this study is to investigate the effects of MSH2 and MSH6 VUSs on the formation and stability of the MutSalpha complex, a key component of the MMR mechanism whose effect on the progression of colorectal cancer is known.

2. MATERIALS & METHODS

1.6 Materials

2.1.1 Equipment

Equipment that was used in the study was listed in Table A.1 in Appendix A.

2.1.2 Chemicals

Chemicals utilized in the study were illustrated in Table A.2 in Appendix A.

2.1.3 Molecular kits and commercial enzymes

Molecular kits, and enzymes or enzyme mixes were listed in Table A.3 and Table A.4 in Appendix B.

2.1.4 Antibodies

Antibodies that were used in the study were illustrated in Table A.5 in Appendix B.

2.1.5 Buffers and solutions

Buffers and their content within solutions that were used in the study were shown in Table A.6 and Table A.7 in Appendix C.

2.1.6 Gene and amino acid sequences

Target gene and amino acid sequences were illustrated in Appendix D.

1.7 Methods

2.2.1 Variant interpretation

The genomic DNA of 371 individuals diagnosed with colorectal cancer was extracted from their blood samples using the ZeeSan DNA isolation kit. Then, genetic testing was performed using the Sophia Hereditary Cancer Solution kit (Sophia Genetics) targeting the splicing junctions and coding regions of 26 genes (*ATM, APC, BARD1, BRCA1, BRCA2, BRIP1, CDH1, CHEK2, EPCAM, MLH1, MRE11, MSH2, MSH6, MUTYH, NBN, PALB2, PIK3CA, PMS2, PMS2CL(1), PTEN, RAD50, RAD51C,*

RAD51D, STK11, TP53, XRCC2) related to Lynch and intestinal polyposis syndromes and breast and ovarian cancers through sequencing on a Illumina's NextSeq550 instrument.

Following the sequencing, the variant calling was carried out with the assistance of SOPHIA DDM, a bioinformatics analysis program. The variations were analyzed by inserting dbSNP codes into variant databases including ClinVar, InterVar, Franklin, and Varsome.

Variants were classified using a five-tier segmentation protocol presented by American College of Medical Genetics and Genomics (ACMG) guideline regarding evidence of pathogenicity, as shown in Table 2.1 according to variant interpretation data and patient background information.

2.2.2 Cloning

2.2.2.1 Cloning into bacterial expression vector by in-fusion cloning

h*MSH2* and h*MSH6* genes were cloned into pET30a (+) and pET15b (+) bacterial expression vectors, respectively. First, full-length h*MSH2* (pFB1_hMSH2) and h*MSH6* (pFB1_hMSH6) were purchased individually from AddGene. Bacteria containing target plasmids were taken from vials using pipette tips. Stripping technique was applied in order to get one colony on LB agar plates containing 100 µg/mL ampicillin and incubated overnight at 37 °C. Single colonies were picked from plates, and put in 50 mL falcon tubes containing 10 mL LB + 10 µL ampicillin, incubated at 37 °C on a shaker for overnight (16 hours). Cell culture was centrifuged at 4000 rpm, at 37°C for 20 minutes, cell pellet was taken without LB and isolated by using Nucleospin Plasmid/Plasmid (NoLid protocol) by following the isolation of high-copy plasmid DNA from *E. coli* protocol according to manufacturer's instructions.

Table 2.1: ACMG Classification criteria for pathogenic variants (Richards et al. 2015).

Data Type	Degree of evidence of pathogenicity			
	Supporting	Moderate	Strong	Very Strong
Population		Absent in population databases PM2	Prevalence in affected statistically increased over controls PS4	
Computational and predictive	Multiple lines of computational evidence support a deleterious effect on the gene /gene product PP3	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before PM5 Protein length changing variant PM4	Same amino acid change as an established pathogenic variant PS1	Predicted null variant in a gene where LOF is a known mechanism of disease PVS1
Functional	Missense in gene with low rate of benign missense variants and path. missenses common PP2	Mutational hot spot or well-studied functional domain without benign variation PM1	Well-established functional studies show a deleterious effect PS3	
Segregation	Cosegregation with disease in multiple affected family members PP1			
De novo		De novo (without paternity & maternity confirmed) PM6	De novo (paternity and maternity confirmed) PS2	
Allelic		For recessive disorders, detected in trans with a pathogenic variant PM3		
Another database	Reputable source = pathogenic PP5			
Other	Patient's phenotype or FH highly specific for gene PP4			

2.2.2.1.1 Gene amplification

h*MSH2* gene was cloned into the pET30a expression vector, C-terminal end 6X histidine-tag (HHHHHH) of pET30a was used. Forward (5'-GGAGATATACATATGGCGGTGCAG CCGAAGGA-3') and reverse (5'-GAAATCATTTCACGAATAAAAGTTACTACGAGTCTCGAGCACCACCACCA C-3') primers were designed, including vector homolog, restriction enzyme recognition site, and gene specific sites.

h*MSH6* gene was cloned into the pET15b expression vector, by the addition of a FLAG-tag (DYKDDDDK) at C-terminal end of target gene. Forward (5'-TAACAATTCCCC TCTAGAATGTCTCGTCAGAGCACCCTGTAC-3') and reverse (5'-CCATAAATTGCTGACTTTGATTAAGGAATTAGATTACAAGGATGACGATG ACAAGTGAGGATCCGGCTGCTAA-3') primers were designed, including vector homolog, restriction enzyme recognition site and gene specific sites.

In order to amplify gene targets, a reaction was set up with 2x CloneAmp HiFi PCR Premix, 0.25 µM forward and 0.25 µM reverse primers, 100 ng DNA template, and up to 50 µL dH_2O. Cycling conditions of the hMSH2 amplification were 98 °C-15 sec., [98 °C-10 sec., 64 °C -10 sec., 72 °C-150 sec. (45 sec/kb)] *35 cycles, 12 °C- hold, while cycling conditions of the h*MSH6* amplification were 98 °C-15 sec., [98 °C-10 sec., 64.5 °C -10 sec., 72 °C-180 sec. (45 sec/kb)] *35 cycles, 12 °C- hold.

2.2.2.1.2 Restriction digestion and ligation

The restriction reaction was settled before incubating at 37 °C for 3 hours under the following conditions.

For pET30a, 1,5 µg vector, 10 units of NdeI and 10 units of Xho1 restriction enzymes, and 10X CutSmart buffer, up to 50 µL dH_2O.For pET15b, 1,5 µg vector, 10 units of XbaI and 10 units of BamH1 restriction enzymes, and 10X CutSmart buffer, up to 50 µL dH_2O. Following both reactions, products were mixed with 6X DNA loading dye and loaded on 0.8% agarose gel at 100V for 40 minutes. Bands were visualized under UV light, and cut out from agarose gel, and NucleoSpin PCR clean-up, gel extraction kit was applied according to manufacturer's instructions.

In-fusion reaction was set up with 100 ng insert, 50 ng vector, 5X In-fusion master mix, and up to 10 µL dH_2O, incubated at 50 °C for 15 minutes.

2.2.2.1.3 Transformation

2 µL of in-fusion product was mixed with 50 µL of DH5α chemically competent *E. coli* cells, and incubation on ice for 30 minutes was provided. Heat shock was applied for 45 seconds at 42 °C, following immediate transfer to ice for 5 minutes. 500 µL SOC medium/1 mL LB was added onto tubes and incubated on a shaker at 37 °C for an hour. Cells were centrifuged for 5 minutes under 4000 rpm and pellet was resuspended with 100 µL non-discarded medium, then spread on agar plates either supplemented with kanamycin (30 µg/mL) or ampicillin (100 µg/mL) antibiotic. Plates were incubated overnight at 37 °C.

2.2.2.2 Cloning into mammalian expression vector

2.2.2.2.1 Gene amplification

pcDNA3.1 (+) was used for cloning of h*MSH2* and h*MSH6* genes into a mammalian expression vector.

Forward primer (5'- CTAGGATCCATGGCAGTGCAGCCGAAGGA-3') and reverse primer (5'-ACCACCACCACCACCACTGATCTAGACA- 3') for h*MSH2* cloning into pcDNA3.1 and forward primer (5'-ATCGGATCCATGTCTCGTCAGAGCACCCT-3') and reverse primer (5'-CAAGGATGACGATGACAAGTGACTCGAGGTGT-3') for h*MSH6* cloning into pcDNA3.1 included restriction enzyme specific recognition site and gene specific sites.

Gene targets were amplified using PCR, including Q5 High-Fidelity 2X Mastermix, 0.25 µM forward and 0.25 µM reverse primers, 100 ng DNA template, and up to 25 µL dH_2O.

Cycling conditions of the h*MSH2* amplification were 98 °C-30 sec., [98 °C-10 sec., 72 °C -20 sec., 72 °C-140 sec. (45 sec/kb)] *35 cycles, 72 °C-2 min., 4 °C- hold, while cycling conditions of the h*MSH6* amplification were 98 °C-30 sec., [98 °C-10 sec., 65 °C -20 sec., 72 °C-180 seconds (45 sec/kb)] *35 cycles, 72 °C-2 min, 4 °C- hold.

After amplification of targeted DNA fragments, products that were mixed with 6X DNA loading dye were loaded on a 0.8% agarose gel and run under 100V for 40 minutes. Bands were observed under UV light, and cut out from agarose gel, and NucleoSpin PCR clean-up, gel extraction was applied according to the manufacturer's instructions.

2.2.2.2.2 Restriction and ligation

In order to restrict amplified gene products and vector backbones, restriction reaction was set in following conditions;

For h*MSH2* and pcDNA3.1, 1 µg vector/insert, 10 units of BamHI and 10 units of XbaI restriction enzymes, and 10X CutSmart buffer, up to 50 µL dH_2O.

For h*MSH6* and pcDNA3.1, 1 µg vector/insert, 10 units of BamHI and 10 units of XhoI restriction enzymes, and 10X CutSmart buffer, up to 50 µL dH_2O.

After that, restricted gene targets were cleaned up using PCR clean-up protocol, while restricted vector backbones were loaded onto agarose gel as in previous experiments and in gel digestion was performed.

Ligation conditions for *MSH2*-pcDNA3.1 and *MSH6*-pcDNA3.1 were calculated by using NEBioCalculator platform, and reaction was incubated at 16 °C, overnight for efficient ligation.

Ligation conditions for h*MSH2*-pcDNA3.1 was 10X T4 DNA Ligase Buffer, 1 µL T4 DNA Ligase, 80 ng vector DNA, 115,5 ng insert DNA (1:3 vector/insert DNA ratio), up to 20 µL dH_2O, whereas for h*MSH6*-pcDNA3.1 was 10X T4 DNA Ligase Buffer, 1 µL T4 DNA Ligase, 80 ng vector DNA, 192 ng insert DNA (1:3 vector/insert DNA ratio), up to 20 µL dH_2O.

Following ligation, transformation was applied as previously described.

2.2.3 Colony PCR

Single colonies were selected from plates containing desired clones. Colony PCR reaction was composed of OneTaq 2X Master Mix, 0.2 µM forward primer (T7 promoter for both; 5'-TAATACGACTCACTATAGGG-3'), 0.2 µM reverse primer (T7 terminator for bacterial vectors; 5'-GCTAGTTATTGCTCAGCGG-3'; bovine growth hormone terminator (BGH) for mammalian vectors; 5'-

TAGAAGGCACAGTCGAGG-3'), 2 µL template DNA (colony dissolved in µL dH_2O), up to 25 µL dH_2O.

Cycling conditions of the h*MSH2* amplification were 94 °C-1 min., [94 °C-20 sec., 55 °C -30 sec., 68 °C-140 seconds (45 sec/kb)] *30 cycles, 68 °C-5 min., 4 °C- hold, while cycling conditions of the h*MSH6* amplification were 94°C-1 min., [94 °C-20 sec., 55 °C -30 sec., 68 °C-180 sec. (45 sec/kb)] *30 cycles, 68 °C-5 min., 4 °C- hold, 4 °C- hold.

2.2.4 Site directed mutagenesis (SDM)

Primers for SDM implementation were designed for selected mutations; MSH2-A733T, MSH6-R577C, and MSH6-S1279N. Site-directed mutagenesis was applied on MSH2-pcDNA3.1 and MSH6-pcDNA3.1 plasmid by using two-sided mutation containing primers (Table 2.2). Reaction was set up in total of 20 µL final volume with 2x CloneAmp HiFi PCR Premix, 0.25 µM forward and 0.25 µM reverse primers, 25 ng DNA template, and up to 20 µL dH_2O.

Table 2.2: Site directed mutagenesis primers used in the study.

Primer Name	Primer Sequence (5'-3')
MSH2-A733T-Forward	GTTGGAAACTACTTCTATCCTC
MSH2-A733T-Reverse2	CCTGAGGATAGAAGTAGTTTCCA
MSH6-R577C-Forward	TTCAGATGATTGCCATTGTTC
MSH6-R577C-Reverse2	CAATGGCAATCATCTGAAAACTG
MSH6-S1279N-Forward	GAAGACCCCAACCAGGAGACT
MSH6-S1279N-Reverse2	TCCTGGTTGGGGTCTTCACATTCA

PCR conditions for MSH6-R577C and MSH6-S1279N mutation applications were 98 °C-30 sec., [98 °C-10 sec., 61 °C (for R577C)/ 69 °C (for S1279N) -25 sec., 72 °C-240 sec. (25 sec/kb)] *25 cycles, 72 °C-2 min., 4 °C- hold.

PCR conditions for MSH2-A733T mutation application were 98 °C-30 sec., [98 °C-10 sec., 60 °C -25 sec., 72 °C-210 sec. (25 sec/kb)] *25 cycles, 72 °C-2 min., 4°C-hold.

Following PCR, KLD reaction was performed containing 0.5 µL PCR product, 2X KLD Reaction Buffer, 10X KLD Enzyme Mix, and up to 5 µL dH_2O, and incubated at room temperature for 30 minutes. Then, transformation of products into competent *E. coli* Dh5α cells was provided as previously described.

2.2.5 Sequencing

Selected colonies were grown in 10 mL LB and 10 µL proper antibiotic (ampicillin/kanamycin), by incubating overnight at 37 °C. Plasmids were isolated using MN's NucleoSpin Plasmid, Mini kit for plasmid DNA. The resulting plasmids were sent to Sanger sequencing with universal T7-promoter, and T7-terminator/BGH-terminator primers, as well as primers designed for efficient sequencing of h*MSH2* and h*MSH6* (Table 2.3)

Table 2.3: Sequencing primers used in the study.

Primer Name	Primer Sequence (5'-3')
MSH2-S1-Forward	GCAAAAAGGGAGAGCAGATG
MSH2-S2-Reverse	CCTTTCTTCACCTGATAAAGCAT
MSH6-S1-Forward	AAAGAGAAGAGATGAGCACAGG
MSH6-S2-Forward	ATATTCCCTTGGATTCTGACACAG

2.2.6 Mammalian cell culture

Human colorectal carcinoma cells, HCT116 (ATCC® CCL-247™), and human embryonic kidney cells, HEK293T (ATCC® CRL-3216™), were both cultured in DMEM full (DMEM (Dulbecco's modified Eagle's medium) with 10% fetal bovine serum (FBS), 10 units/ml penicillin-G and 10 mg/ml streptomycin). LoVo colorectal carcinoma cells, as a kind gift from Gozuacik Lab (KUTTAM Research Center for Translational Medicine, Koc University Hospital) were cultured in Roswell Park Memorial Institute (RPMI) 1640 full medium (RPMI 1640 with 20% FBS, 10 units/ml penicillin-G and 10 mg/ml streptomycin) at 37 °C, in a 5% CO_2 incubator.

2.2.6.1 Transfection

HEK293T, HCT116 and LoVo cells were seeded into 6-well plates containing 0.5 x 10^6 cells per well, or 3 x 10^6 alive cells were seeded into 100-mm cell culture dishes. After 48 hours of incubation, cells were transiently transfected with 0.5-2.5 µg plasmids (MSH2-WT-pcDNA3.1, MSH2-A733T-pcDNA3.1, MSH6-WT-pcDNA3.1, MSH6-R577C-pcDNA3.1, MSH6-S1279N-pcDNA3.1, mock (pcDNA3.1) per well of 6-well plates or 9-12 µg plasmids per one 100-mm cell culture dish and PEI transfection reagent (1:3 DNA/PEI), with only DMEM or RPMI 1640 media. The transfected cells were incubated at 37 °C, in a 5% CO_2 incubator for 24-72 h before total RNA or protein isolation.

2.2.6.2 Total RNA isolation and reverse transcriptase PCR (RT-PCR)

48 hours after transfection, medium of the cell culture dishes was discarded. Cells were scraped from the surface of the plates with the addition of ice cold 1x TBS. Collected cells were centrifuged at 500 xg, for 10 minutes at 4°C. Total RNA Purification Mini Spin Kit was applied according to the manufacturer's instructions. Isolated RNA concentrations were measured using NanoDrop™ 2000 Spectrophotometer. Isolated total RNA was converted to total cDNA by using reaction components; oligo dT, 6N random hexamer, 5X reaction buffer, 0.2 mM dNTP, 0,33 ng M-MLV Reverse Transcriptase, 25 ng RNase Inhibitor, dH_2O, total RNA (500 ng). Reaction was performed at 37 °C for 30 minutes, followed by incubation at 70 °C for 5 minutes. Concentration of converted ssDNA's was measured using NanoDrop™ 2000 Spectrophotometer.

2.2.6.2 Quantitative real time PCR analysis (qRT-PCR)

Quantitative real-time reaction was operated using 2X PowerTrack SYBR Green Master Mix, 500 ng total cDNA, 0.5 µM forward and 0.5 µM reverse primers (Table 2.4), and up to 20 µL dH_2O. Primers targeting Beta-actin was used for internal control mRNA expression.

RT-PCR cycling conditions were 95 °C-10 sec., [98 °C-10 sec., 54 °C -10 sec., 72 °C-10 sec.] *45 cycles, [95 °C-5 sec., 65 °C -1 min., 97 °C-continuous]

Table 2.4: qRT-PCR primers used for MSH6 or Beta-actin expression in the study.

Primer Name	Primer Sequence (5'-3')
MSH6-S1-Forward	AAAGAGAAGAGATGAGCACAGG
MSH6-qPCR-Reverse	GTAGAACCATTTGTTCCATTCAG
Beta-Actin-Forward	CACCATTGGCAATGAGCGGTTC
Beta-Actin-Reverse	AGGTCTTTGCGGATGTCCACGT

2.2.6.3 Cell viability assay (MTT)

HCT116 and LoVo cells were seeded into 96-well plates containing 1.0 x 10^4 cells per well, in 100 µL DMEM full or RPMI 1640 full medium. Cells were transfected with 0.15 µq of plasmid (1:3 DNA/PEI) per well. 24, 48, and 72 hours after transfection, 10 µL of 3-4,5-Dimethyl-2-thiaxolyl-2,5-diphenyl-2H-tetrazolium bromide (MTT) reagent (5 mg/mL) was added to medium and incubated at 37 °C for 4 h. Later, media containing MTT was discarded, and 100 µL DMSO was added onto each well,

incubated in dark at room temperature for 5 minutes. Absorbance values were measured at 570 and 655 nm wavelengths using a microplate reader.

2.2.7 Total protein isolation

To isolate total protein, cells were scraped from cell culture plates in the presence of ice-cold 1x TBS after 48 hours of transient transfection. Collected cells were centrifuged at 500 xg, 4 °C for 10 minutes. Each pellet that was collected from one well of 6-well plates was resuspended with 80 µL RIPA or NP-40 buffer (supplemented with fresh 3 mM PMSF, and 1X PI cocktail (cOmplete™, EDTA-free Protease Inhibitor Cocktail), while cell pellets collected from one cell culture dish were resuspended with 500 µL RIPA or NP-40 buffer. After vortexing and incubation on ice for 30 minutes, cells were centrifuged at 25000 xg, at 4 °C for 30 minutes and supernatants were collected as total protein samples. Isolated total protein concentration was measured using Bradford Assay and bovine serum albumin (BSA) standard at different concentrations (0.25 mg/mL-2 mg/mL).

2.2.8 SDS-PAGE and immunoblotting

Samples containing 40 µg of total protein or same volume of samples were mixed with 4x Laemmli sample buffer and incubated at 95°C for 10 minutes. Then samples were loaded on 8% denaturing polyacrylamide gels and run under constant 95-125 V for 1-2 hours. After SDS-PAGE was completed, gels were either dyed with Coomasie blue dye for direct imaging or used to transfer the proteins onto a nitrocellulose membrane for immunoblotting. Gels were transferred to nitrocellulose membranes in a wet-transfer system at constant 100 V for 90 minutes or at constant 16 V for 16 hours. Following the transfer, membranes were incubated with a 5% blocking solution, and then blotted with targeted primary antibodies overnight on a rotator at 4°C. After that, membranes were washed with 1x TBS-T, and incubated with corresponding anti-mouse or anti-rabbit IgG-HRP conjugated secondary antibodies at room temperature on a shaker for 2 hours. After the membranes were washed with 1x TBS-T again, Clarity ECL Western Blotting Substrate was used to visualize proteins on the ChemiDoc imaging system.

2.2.9 Co-immunoprecipitation

Protein G coupled dynabeads (50 µL per one condition) were washed with 200 µL TBS-T, two times. Supernatants were discarded using magnetic stand, and the beads

were incubated with 2 µL primary antibody in 200 µL TBS-T on shaker for 30 minutes at room temperature to allow binding of antibody and beads. Following the discernment of the supernatant, 250 µg total cell lysate was added onto beads for each condition and incubated at 4 °C overnight on a rotator. Then, using magnetic stand, unbounded proteins were collected in supernatant form. Beads were washed with 1x TBS (200 µL*2 times). To elute proteins, the beads were incubated with 20 µL elution buffer (50 mM Glycine, pH: 2.8) 2 minutes, and then 10 µL 4x laemmli was added before boiling proteins at 70 °C for 10 minutes. Unbound and load samples were prepared as containing 15 µg total protein per lane according to values that were measured earlier, with 4x laemmli and incubated at 95°C for 10 minutes.

2.2.10 Pull-down assay

Total protein content of the cells transfected with MSH2-WT or MSH2-A733T was measured by Bradford assay, and total 1 mg protein was incubated with 150 µL Ni-NTA resin on a rotator at 4 °C for 3 hours (0.5 mL total volume). Then, resin-protein mix was loaded on a column, and flow-through section was collected. Ni-NTA resin-protein complex lying on the column was washed with wash buffer 3 times (1 mL). Same amount of cell lysate of cells transfected with MSH6-WT, MSH6-R577C, or MSH6-S1279N was added onto Ni-NTA-protein mix for 3 hours at 4 °C. Then, flow-through II was collected through loading mixture on a column. Column was washed with wash buffer 3 times (1 mL). After the addition of 100 µL elution buffer, tip of the column was covered and incubated for 2 minutes, and elution fractions were collected.

2.2.11 Purification

To purify MSH2-WT and MSH2-A733T, tandem purification process was applied using the ÄKTA Avant 25 FPLC system (GE Healthcare). In immobilized metal affinity chromatography, a 1 mL HisTrap FF column was used to purify target proteins. The column was equilibrated with the binding buffer for 5 Colume Volume (CV). After the column was properly equilibrated, the sample was inoculated. The column was then washed for 5 CV with a wash buffer containing 20 mM imidazole. To separate the target protein from the column, the elution buffer including 250 mM imidazole was passed through the column until it reached 100 percent using a linear increase and 15 CV step gradient. To determine the purity of fractions, SDS-PAGE and immunoblotting analyses were performed. The collected proteins were then diluted with Anion exchange chromatography (AEX) binding buffer at a ratio of 1:3.

It was then centrifuged at 4000 rpm at 4 °C until the filter level dropped to a sufficient level. The buffer exchange was carried on until it had been diluted roughly 1000-fold.

Same system was used for the elimination of degradation products by using Anion exchange chromatography, 1 mL HiTrap Q XL column. Column was equilibrated with a binding buffer for 5 CV. After that, column was washed with binding buffer, and proteins were eluted with elution buffer containing 1 M NaCl, with 10 CV linear gradient and 15 CV step gradient. To determine the purity of fractions, SDS-PAGE and immunoblotting analyses were performed. Pure protein-containing fractions were combined and diluted with 2x storage buffer (1:3). Centrifugation at 4000 rpm at 4 °C was applied until the filter level dropped sufficiently. Buffer exchange continued until it was diluted approximately 1000-fold.

2.2.12 Bacterial expression

2.2.12.1 Expression optimization

Competent *E. coli* BL21 and *E. coli* GroEL/GroES-Rosetta cells were cotransformed with both MSH2-pET30a and MSH6-pET15b plasmids as previously described. Single colonies formed by co-transformed MSH2-pET30a and MSH6-pET15b plasmids into both BL21 and GroEL/GroES-Rosetta competent cells were transferred into 50 mL starter cultures (50 mL LB+ 50 µL Amp or 50 µL Kan).

Scale up from preculture to the large culture was achieved by inoculating 1 mL of precultures into the 50 mL LB + 50 µl Amp/Kan containing Erlenmeyer flasks. To optimize the growth conditions that ended up with soluble target proteins with the highest yield, three different temperatures (17 °C, 30 °C, 37 °C) were tried after induction by addition of 0.5 mM IPTG when OD_{600} reached 0.6. For the cultures at 30 °C and 37 °C, induction was provided for 4 and 3 hours, respectively. Then, samples were centrifuged at 4000 rpm, at 4 °C for 20 minutes, and pellets were stored at -20 °C. Culture at 17 °C incubated overnight (16 h) an after that pellet was obtained following same steps.

2.2.12.2 Cell lysis

In expression optimization, pellets were lysed with 10 mL of bacterial lysis buffer per 1 gram of cell dry weight. Sonication was applied with MS72 probe at 40% amplitude, for 15 seconds on ice, and centrifuged at 24000 xg, at 4 °C for 30 minutes. Supernatants were collected and pellets were dissolved in the same amount of pellet

dissolvation buffer and centrifuged in the same condition. SDS/WB samples were prepared, containing 30 µL sample and 10 µL 4x laemmli dye, and loaded on 8% polyacrylamide gels, and overnight transfer was applied.

2.2.12.3 Purification

The ÄKTA Avant 25 FPLC system (GE Healthcare) was utilized for the purification. Anion exchange chromatography (AEX) was used to purify target proteins by using a HiTrap Crude 5 mL column. Equilibration was applied with binding buffer for 5 CV. After the sample was loaded onto a column, to eliminate nonspecific bindings, the column was washed with binding/wash buffer containing 150 mM KCl. An elution buffer was applied for 10 CV linear and 15 CV step gradient. Elution fractions were collected and concentrated according to the chromatographic results. To determine the purity of fractions, SDS-PAGE and immunoblotting analyses were performed. Pure protein-containing fractions were combined for further purification. Collected fractions were concentrated, and buffer content was changed depending on the further purification method.

To increase the purity of protein, heparin affinity chromatography was implemented by using HiTrap Heparin HP column in ÄKTA Avant system. Equilibration, washing, and elution steps were applied in the amounts previously described. The same amount of fractions was separated and analyzed through SDS-PAGE and immunoblotting.

3. RESULTS

3.1 Colorectal Cancer Statistics

Alterations in the cancer predisposition genes of 371 colorectal cancer patients were investigated by using next-generation sequencing technology. The study group was selected according to cancer patients that were consulted at Umraniye Teaching and Research Hospital, in Istanbul/Turkey. Distribution of pathogenic, VUS, or benign variant carrying patients were classified according to ACMG guideline, demonstrated in Figure 3.1. 19% (n= 72) of the colorectal cancer patients carried pathogenic variants, while 30% (n= 111) of patients had VUS among their investigated cancer susceptibility genes. 51% (n= 188) of the patients resulted with only benign or likely benign variants. As a note, six of the 371 colorectal cancer patients were also diagnosed with breast cancer.

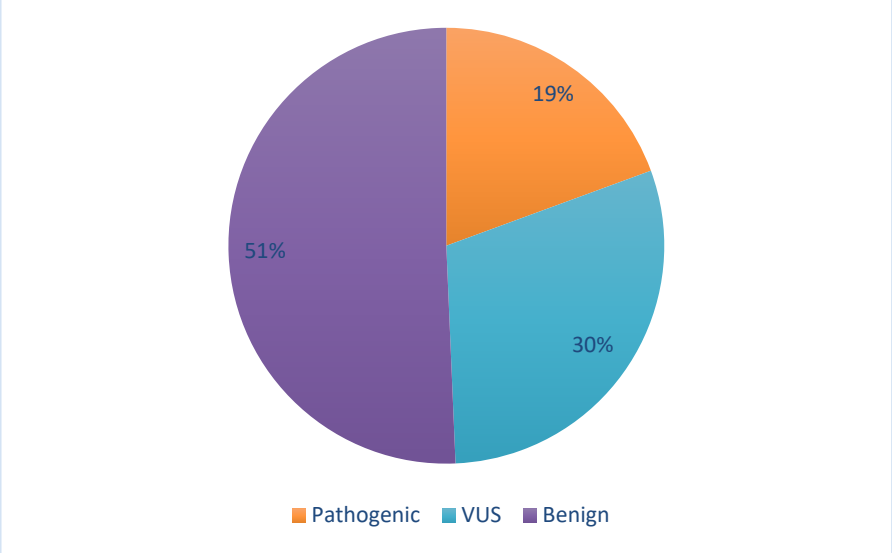

Figure 3.1: Variant distribution among colorectal cancer patients (n=371).

66 different rare pathogenic variants were identified in 72 colorectal cancer patients. *MUTYH* (22,72 %) gene was the most pathogenic variant including gene, followed by *MLH1* (21,21%), *MSH2* (13,63%), and *CHEK2* (6,06%) genes. Moreover, 4,54 % of the pathogenic variants were found both in the important Lynch syndrome related genes *MSH6* and *PMS2*. Pathogenic variants were distributed according to their

mutation types in Figure 3.2. Most of the variants' molecular effects were identified as missense or in frame insertion/deletions.

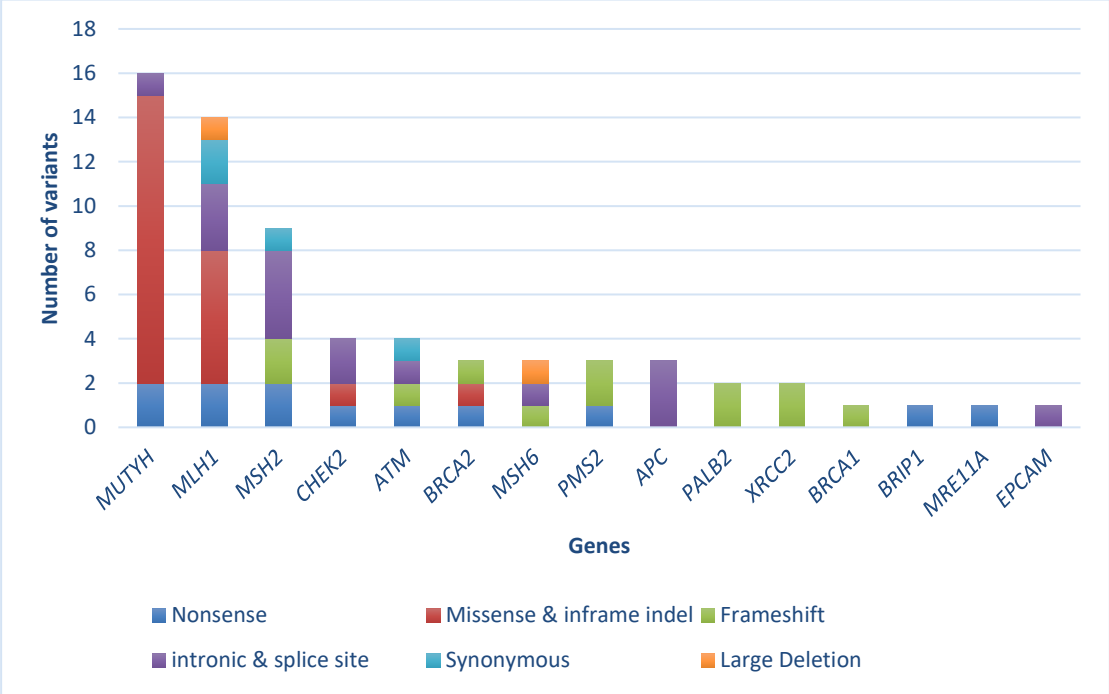

Figure 3.2: Pathogenic variant distribution among genes based on their molecular effects.

In 371 colorectal cancer patients, 158 variants with uncertain significance (VUS) were detected. Majority of missense variants were identified with 94,93% among VUS. *ATM* was the most VUS carrying gene with 12,02%, followed by *MSH2*, *CHEK2* and *APC* genes with 8,86%. To facilitate the determination of the actual effects of VUS, the numbers of variations were normalized, accomplished by dividing each number of variants by the total length of the coding sequence in the gene and then multiplying the result by 1000. Figure 3.3. presents a comparison of the values across the different genes. *CHEK2* exhibited the highest VUS with 8,57 across the board after normalization was applied, and followed by *RAD51D (6,07)*, *MSH2 (4,99)* and *MLH1 (4,84)* genes. When other Lynch syndrome related genes were examined, *MSH6* and *PMS2* resulted in 2,69 and 0,38 calibrated VUS numbers, respectively.

In total, 10 different *MSH2* and 8 different *MSH6* variants were identified in colorectal patients, which were then classified depending on the localization of the mutation on the protein. The domain specific distribution of identified VUS and the list of the variants were illustrated in Figure 3.4 and Table 3.1.

Among these *MSH2* and *MSH6* inconclusive variants, *MSH2* c.2197G>A (p. Ala733Thr) variant, which was localized at the ATPase domain of MSH2, *MSH6* c.3836G>A (p. Ser1279Asn), and *MSH6* c.1729C>T (p. Arg577Cys) variants that were found in the ATPase and the connector domain of MSH6, respectively, were selected for further investigation.

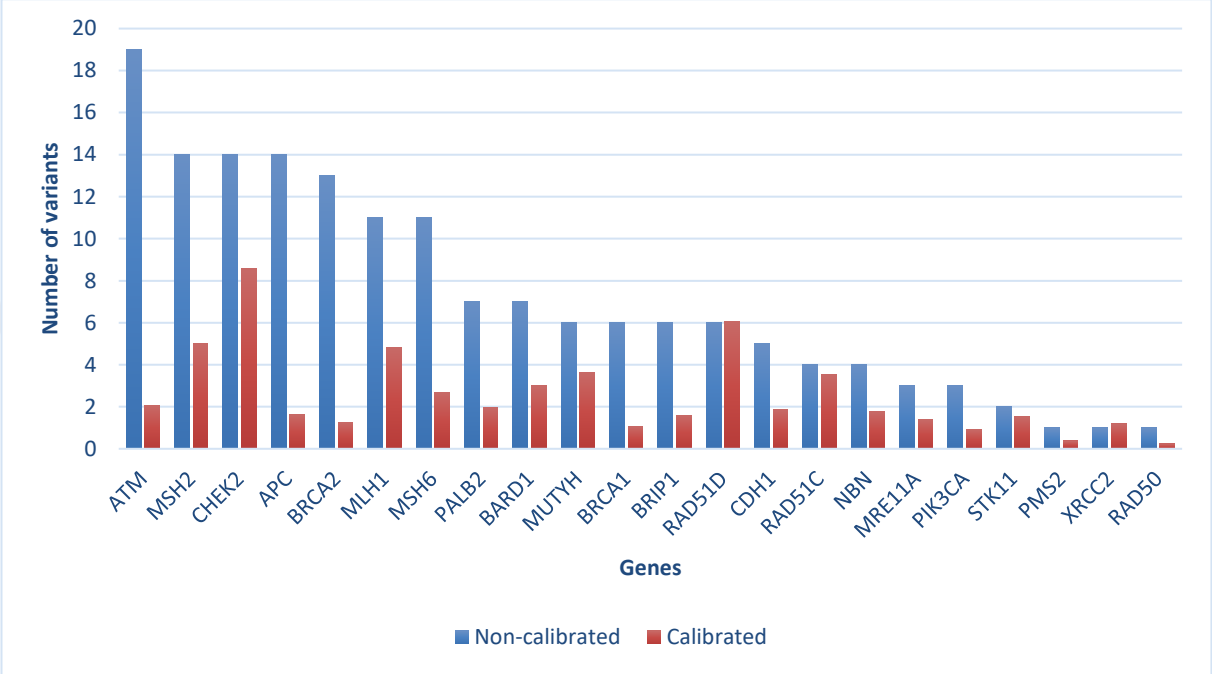

Figure 3.3: Comparison of calibrated and non-calibrated numbers of all uncertain significant variants in study group.

In total, 10 different *MSH2* and 8 different *MSH6* variants were identified in colorectal patients, which were then classified depending on the localization of the mutation on the protein. The domain specific distribution of identified VUS and the list of the variants were illustrated in Figure 3.4 and Table 3.1. Among these *MSH2* and *MSH6* inconclusive variants, *MSH2* c.2197G>A (p. Ala733Thr) variant, which was localized at the ATPase domain of MSH2, *MSH6* c.3836G>A (p. Ser1279Asn), and *MSH6* c.1729C>T (p. Arg577Cys) variants that were found in the ATPase and the connector domain of MSH6, respectively, were selected for further investigation.

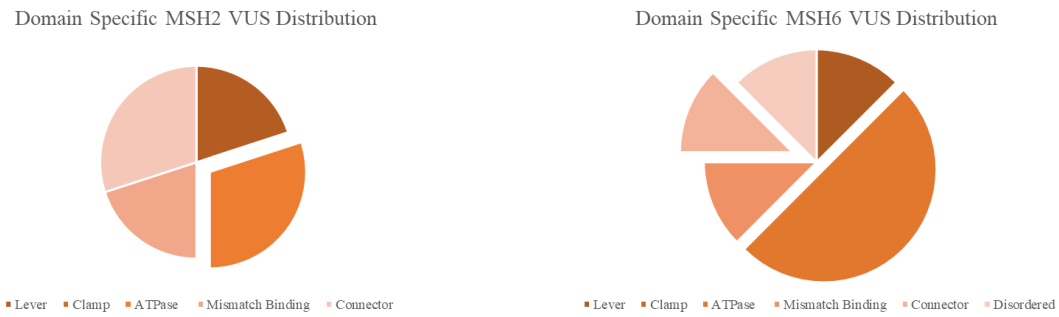

Figure 3.4: Domain specific distribution of *MSH2* and *MSH6* VUS (left to right).

Table 3.1: List of the identified *MSH2* and *MSH6* VUS, including location on gene, nucleotide, and amino acid changes (Selected variants were depicted as *).

Variants	Domains
MSH2 c.2197G>A (p. Ala733Thr) *	ATPase
MSH2 c.2606C>A (p. Ala869Glu)	ATPase
MSH2 c.2333G>T (p. Cys778Phe)	ATPase
MSH2 c.580A>G (p. Ile194Val)	Connector
MSH2 c.727C>T (p. Arg243Trp)	Connector
MSH2 c.775C>T (p. Pro259Ser)	Connector
MSH2 c.440T>G (p. Val147Gly)	Mismatch Binding
MSH2 c.382C>G (p. Leu128Val)	Mismatch Binding
MSH2 c.1774A>G (p. Met592Val)	Lever
MSH2 c.1198A>G (p. Asn400Asp)	Lever
MSH6 c.3299C>T (p. Thr1100Met)	ATPase
MSH6 c.4057T>G (p. Leu1353Val)	ATPase
MSH6 c.3287A>G (p. His1096Arg)	ATPase
MSH6 c.3836G>A (p. Ser1279Asn) *	ATPase
MSH6 c.1149G>T (p. Arg383Ser)	Mismatch Binding
MSH6 c.2215A>G (p. Thr739Ala)	Disordered
MSH6 c.998C>T (p. Thr333Ile)	Lever
MSH6 c.1729C>T (p. Arg577Cys) *	Connector

3.2 Cloning

3.2.1 Cloning of hMSH2/hMSH6 into pET30a/ pET15b

In order to study selected variants on the protein level, first full length of hMSH2 and hMSH6 genes acquired by AddGene were cloned into pET30a, and pET15b expression vectors for bacterial expression of targeted proteins, respectively.

Figure 3.5 shows the representative map of the plasmids; cloning of the genes was confirmed by colony PCR and afterwards by sequencing. It was affirmed that the sequencing result was matched with target gene sequences with 99% similarity with each primer read according to the BLAST multiple sequence alignment tool.

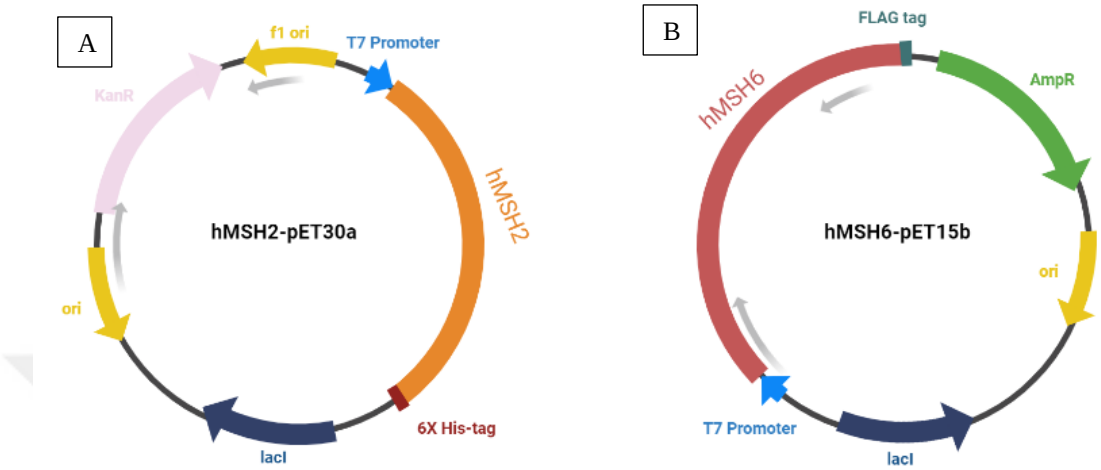

Figure 3.5: Representation of bacterial expression vectors; hMSH2-pET30a and hMSH6-pET15b (Created with BioRender.com).

3.2.2 Cloning of hMSH2/hMSH6 into pcDNA3.1

Gene targets of hMSH2 and hMSH6 which were amplified through previously obtained bacterial expression plasmids, were cloned into mammalian expression vector pcDNA3.1, for eukaryotic protein production, illustrated in Figure 3.6. Sequencing results were confirmed by using both the BLAST multiple sequence alignment tool and the EMBOSS-Needle sequence alignment tool, demonstrating that the cloning of the target genes into desired expression vector was successfully achieved. Figure 3.7 shows one part of the chromatogram of the sequenced gene targets by either T7 or BGH primers, for both hMHS2-pcDNA3.1 and hMSH6-pcDNA3.1.

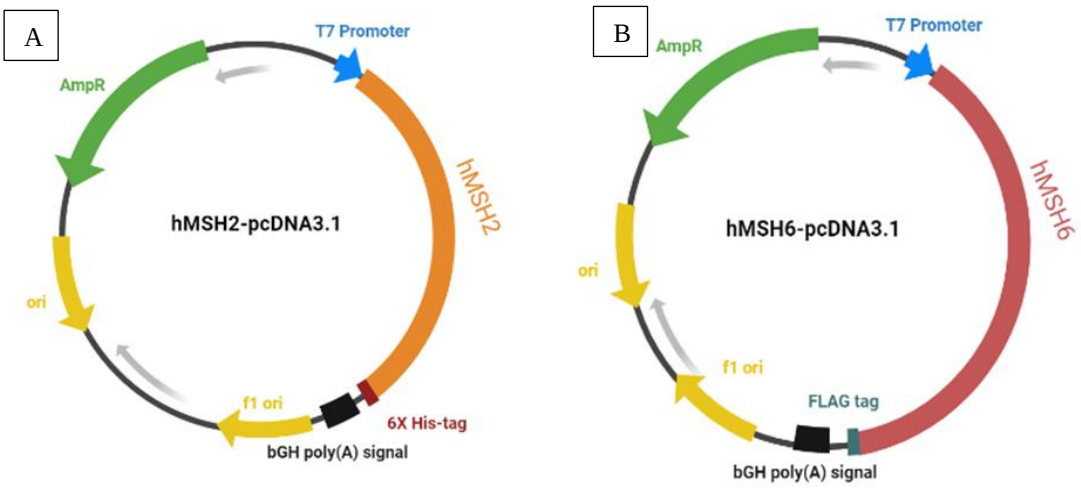

Figure 3.6: Representative map of mammalian expression vectors; hMSH2-pcDNA3.1 and hMSH6-pcDNA3.1 (Created with BioRender.com).

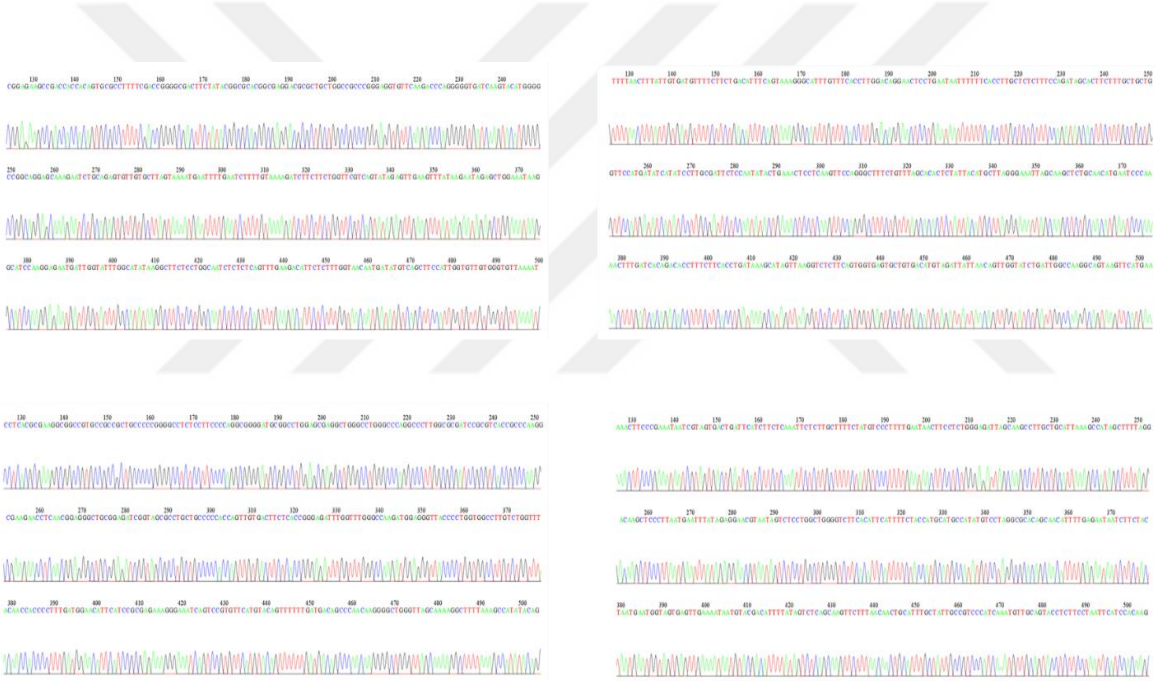

Figure 3.7: Chromatograph of the sequencing data of hMSH2-pcDNA3.1 (upper left and right) and hMSH6-pcDNA3.1 (lower left and right) read by T7-promoter and BGH-reverse primers.

3.3 Site Directed Mutagenesis

Amino acid alterations belonging to the selected VUSes, (*MSH2* c.2197G>A (p. Ala733Thr), *MSH6* c.3836G>A (p. Ser1279Asn), and *MSH6* c.1729C>T (p. Arg577Cys)) of MSH2 and MSH6 were applied into either hMSH2-pcDNA3.1 or hMSH6-pcDNA3.1 vector by site directed mutagenesis. Applied substitutions were first confirmed by colony PCR, and the reaction was set by using mutation containing primers and sequencing primers to amplify a region containing mutation. PCR

products run on 1% agarose gel, showed that amplification of the desired regions was obtained from 5 of the 6, as shown in Figure 3.8. MSH2-WT and MSH6-WT plasmids were used as a control to demonstrate the effectiveness of the reaction.

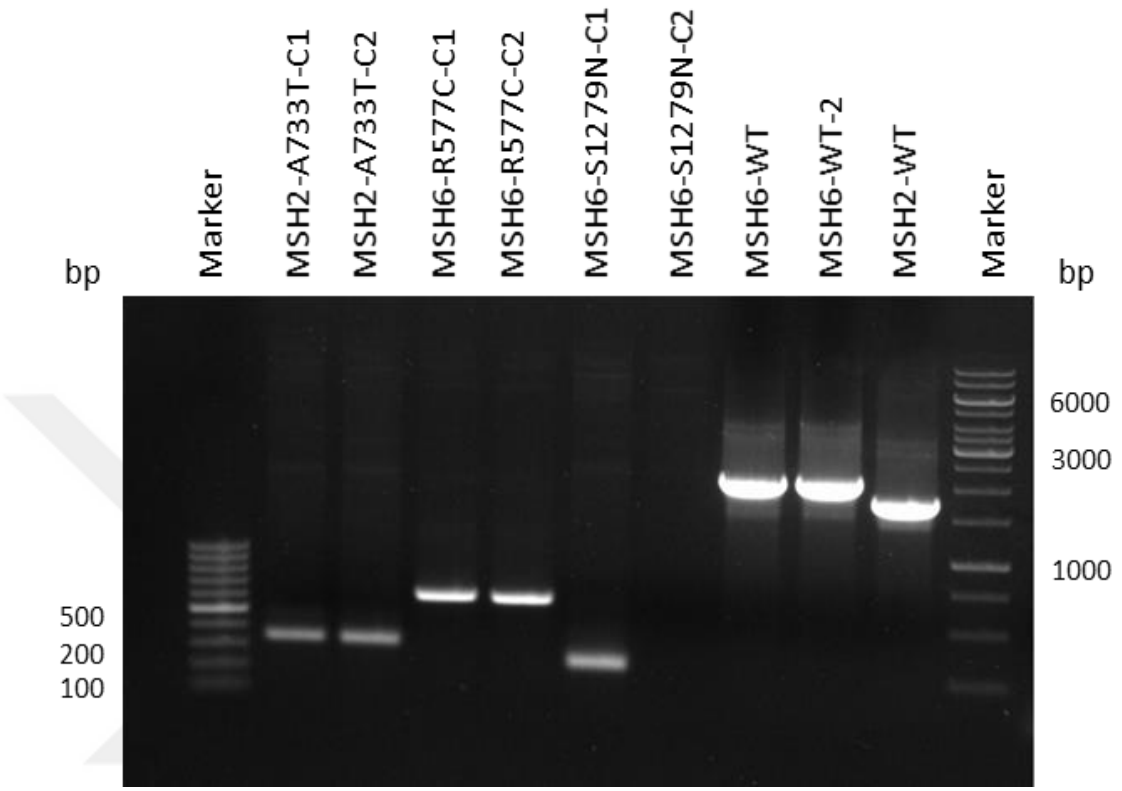

Figure 3.8: Amplified gene regions run on 1% agarose gel.

After that, the mutation application was confirmed by Sanger sequencing. The obtained reads were aligned with WT forms of the genes, which revealed that the plasmids included a single nucleotide substitution in the desired position. Figure 3.9 demonstrates the multiple alignment results of selected VUS acquired by the EMBOSS-Needle multiple sequence alignment tool.

3.4 Expression Optimization of pcDNA3.1 in HEK293T Cells

Overexpression of the wild-type MSH2 and MSH6 proteins was optimized in transiently transfected HEK293T cells. After transfection of HEK293T cells, different incubation times were conducted, and it was found that protein production after 48 hours of transfection was the most efficient condition for both individual expression and co-expression of target proteins. Band intensities were observed according to anti-His, anti-FLAG, and anti-GAPDH primary antibodies concerning FLAG-tagged MSH6, His-tagged MSH2, and internal control GAPDH levels (Figure 3.10).

Figure 3.9: Multiple alignment result of the target gene (MSH2-WT for A, MSH6-WT for B and C) and sequencing data as a result of site directed mutagenesis implementation, nucleotide substitutions were marked in red spheres (A) MSH2 c.2197G>A (p. Ala733Thr), B) MSH6 c.1729C>T (p. Arg577Cys), c) MSH6 c.3836G>A (p. Ser1279Asn)).

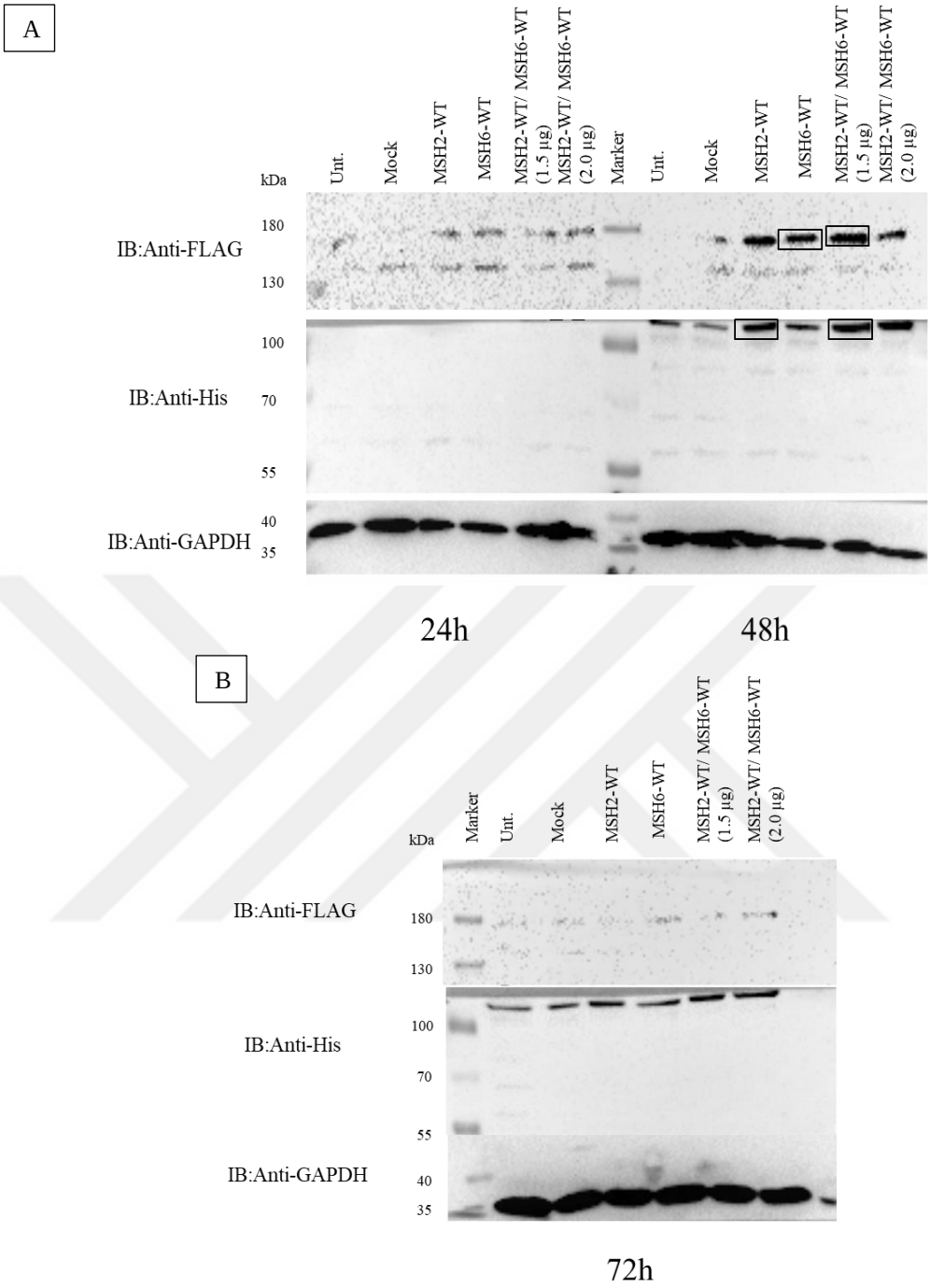

Figure 3.10: Immunoblotting result of the whole protein content of the HEK293T cells transfected/co-transfected A) after 24 and 48 hours of transient transfection, B) after 72 hours of transient transfection with MSH2/MSH6 plasmids in the presence of FLAG primary antibody (1:1000), His-tagged primary antibody (1:1000), and GAPDH primary antibody (1:1000), optimum band intensities marked in rectangles.

After incubation time optimization, different plasmid DNA amounts for WT-containing plasmids were optimized. A range of different plasmid amounts were investigated, starting from 0.5 µg plasmid (1:3 DNA/PEI) up to 2.5 µg plasmid (1:3 DNA/PEI), for individual expression of plasmids (Figure 3.11). Analyzed co-expression amounts for plasmids were; MSH2-WT 1.0 µg/ MSH6-WT 1.0 µg, MSH2-

WT 1.0 µg/ MSH6-WT 1.5 µg, MSH2-WT 1.5 µg/ MSH6-WT 1.5 µg, MSH2-WT 1.5 µg/ MSH6-WT 2.0 µg. Results showed that even though there are non-specific bands shown in untransfected and mock transfected wells, most stable protein production was observed as a result of 1.5 µg MSH2-WT transfection. Moreover, the clearest band was detected in wells containing HEK293T cell lysates transfected with 1.5 µg MSH6-WT plasmids. In addition, co-transfection trials depicted that the same amount of MSH2-WT and MSH6-WT plasmids in 1.5 µg resulted in higher protein expression.

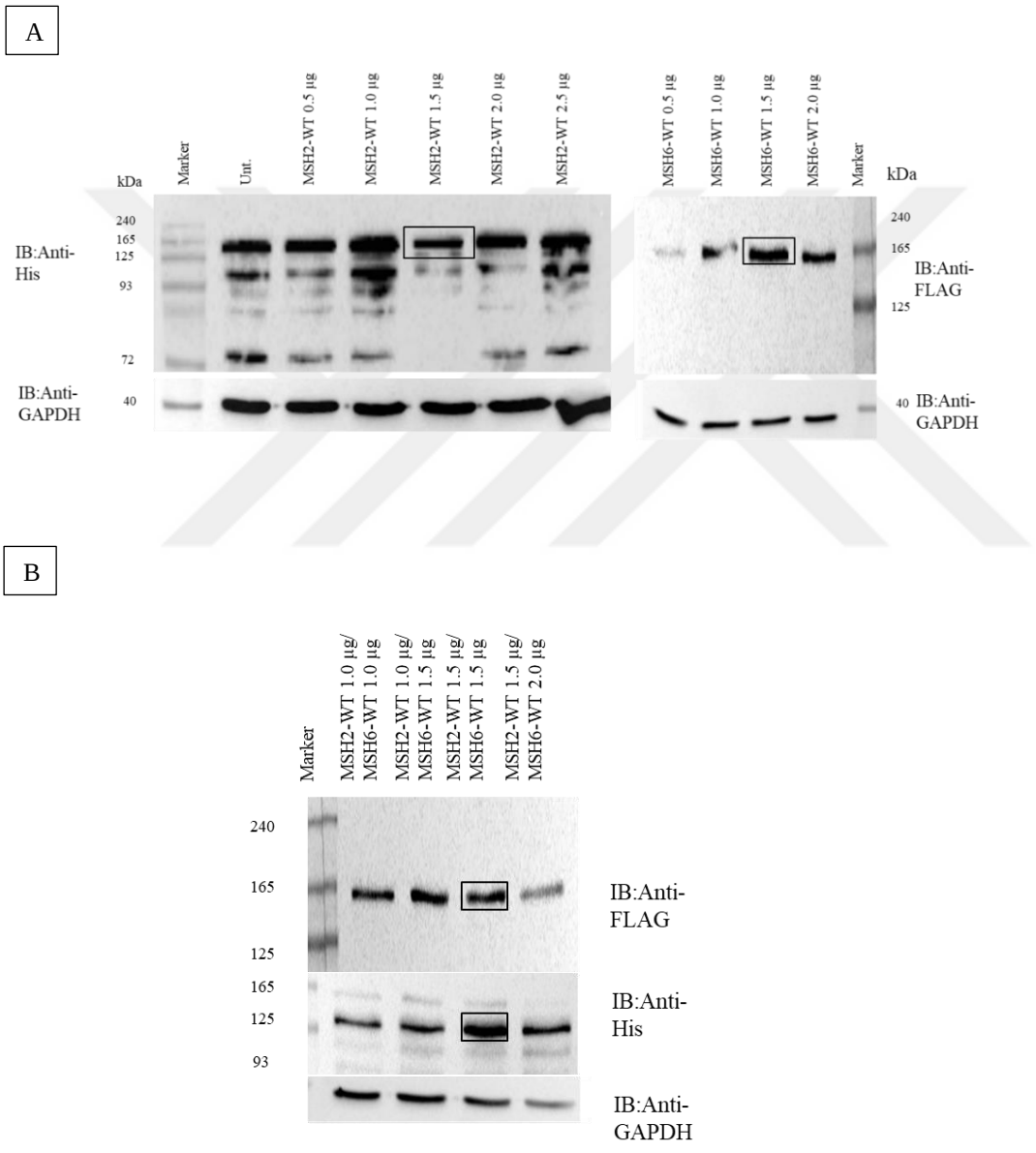

Figure 3.11: Immunoblotting result of the whole protein content of the HEK293T cells A) transfected or B) co-transfected (DNA amount optimization) with MSH2/MSH6 plasmids in the presence of FLAG primary antibody (1:1000), His-tagged primary antibody (1:1000), and GAPDH primary antibody (1:1000), optimum band intensities marked in rectangles.

3.5 Cytotoxicity Assay (MTT)

Cytotoxic effects of plasmids in MSH2-deficient LoVo or MLH1-deficient HCT116 colorectal cancer cells were measured using the MTT assay. Absorbance measurements were performed for three independent biological and technical replicates at 24, 48 and 72 hours after transfection. Results were normalized to untransfected conditions at 24 hours and compared with wild-type and mutated form of plasmids. It was determined that there are no significant changes when cells transfected with plasmids in both LoVo or HCT116 cells (Figure 3.12).

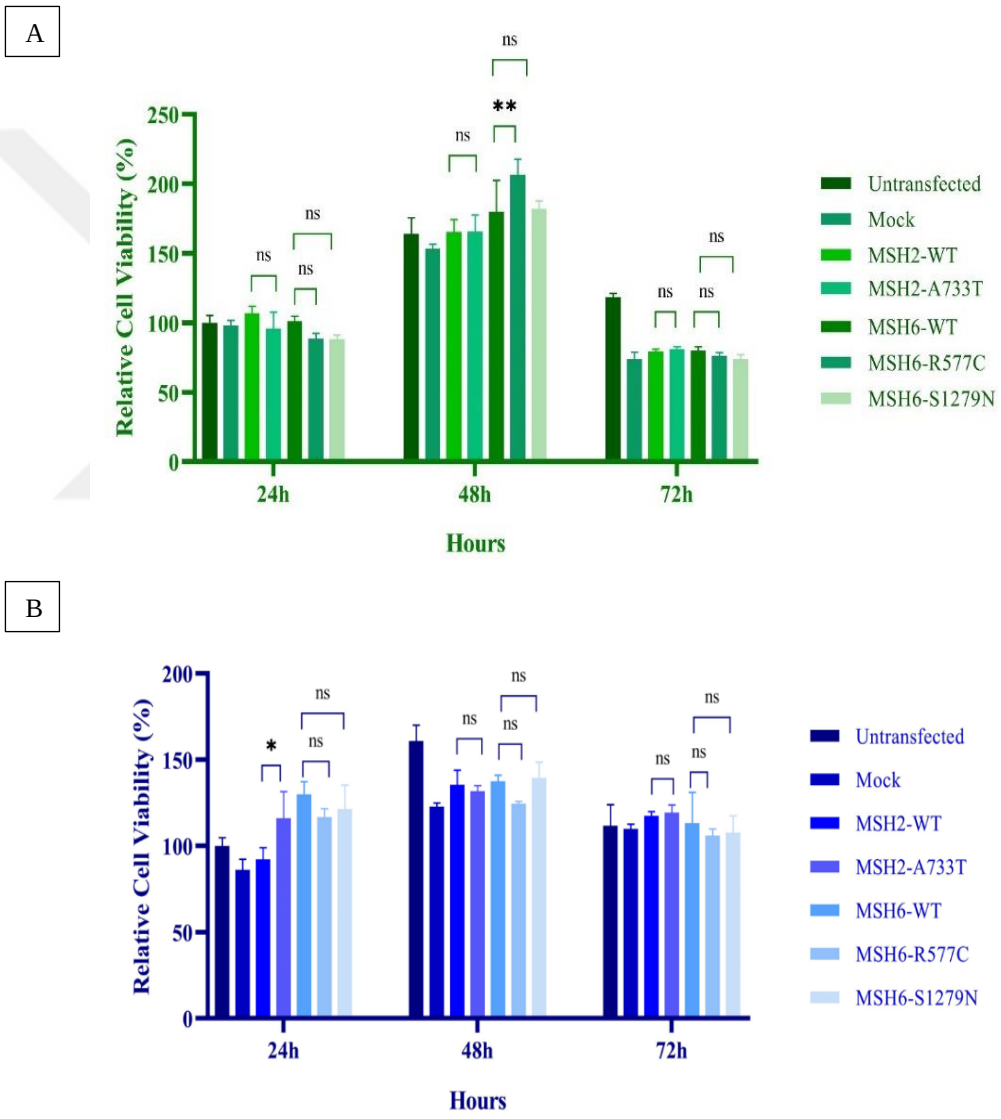

Figure 3.12: Cell viability assay (MTT) results of A) LoVo colorectal cancer cells and B) HCT-116 colorectal cancer cells in the time intervals of 24, 48, and 72 hours after transfection (normalized to untransfected condition at 24 h). Analyses were performed three times for verification of obtained results. ($p<0.05$, $**p<0.01$, $***p<0.001$, $****p<0.0001$, Two-way ANOVA, t-test, 95% Confidence Interval)

3.6 Quantitative Real Time PCR Analysis (qRT-PCR)

To investigate the effect of MSH2 overexpression on the regulation of MSH6 expression, mRNA levels of MSH6 were quantified in transiently transfected LoVo cells with either mock, MSH2-WT or MSH2-A733T. Quantification of hMSH6 mRNA showed a significant decrease in hMSH6 levels when cells were transfected with MSH2-A733T compared to MSH2-WT. Figure 3.13 demonstrates ΔΔCT-factor of hMSH6 depending on the transfection conditions.

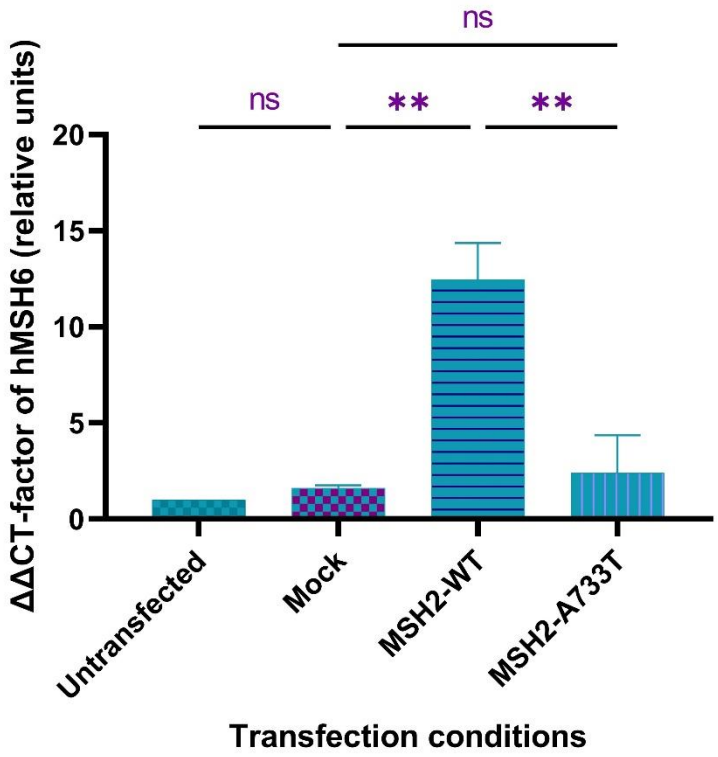

Figure 3.13: ΔΔCT- factor of hMSH6 gene expression pattern as a result of transfection of LoVo cells normalized to Beta-Actin gene expression levels ($p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$, $^{****}p<0.0001$, One-way ANOVA, t-test, 95% Confidence Interval).

3.7 Changes in Protein Expression Levels

Expression levels of direct or indirect interaction partners of hMSH2/hMSH6 were investigated in LoVo cells transfected with wild-type or mutant form containing plasmids (Figure 3.14). MMR proteins, such as MLH1, PMS2 and PCNA were investigated as well as ATM and APC as being interacting partners of MutSalpha complex in other cellular pathways such as homologous recombination and cell cycle checkpoint. As a result, overexpression of target genes with or without mutation

presenting form did not cause any significant change in expression levels of interacting proteins.

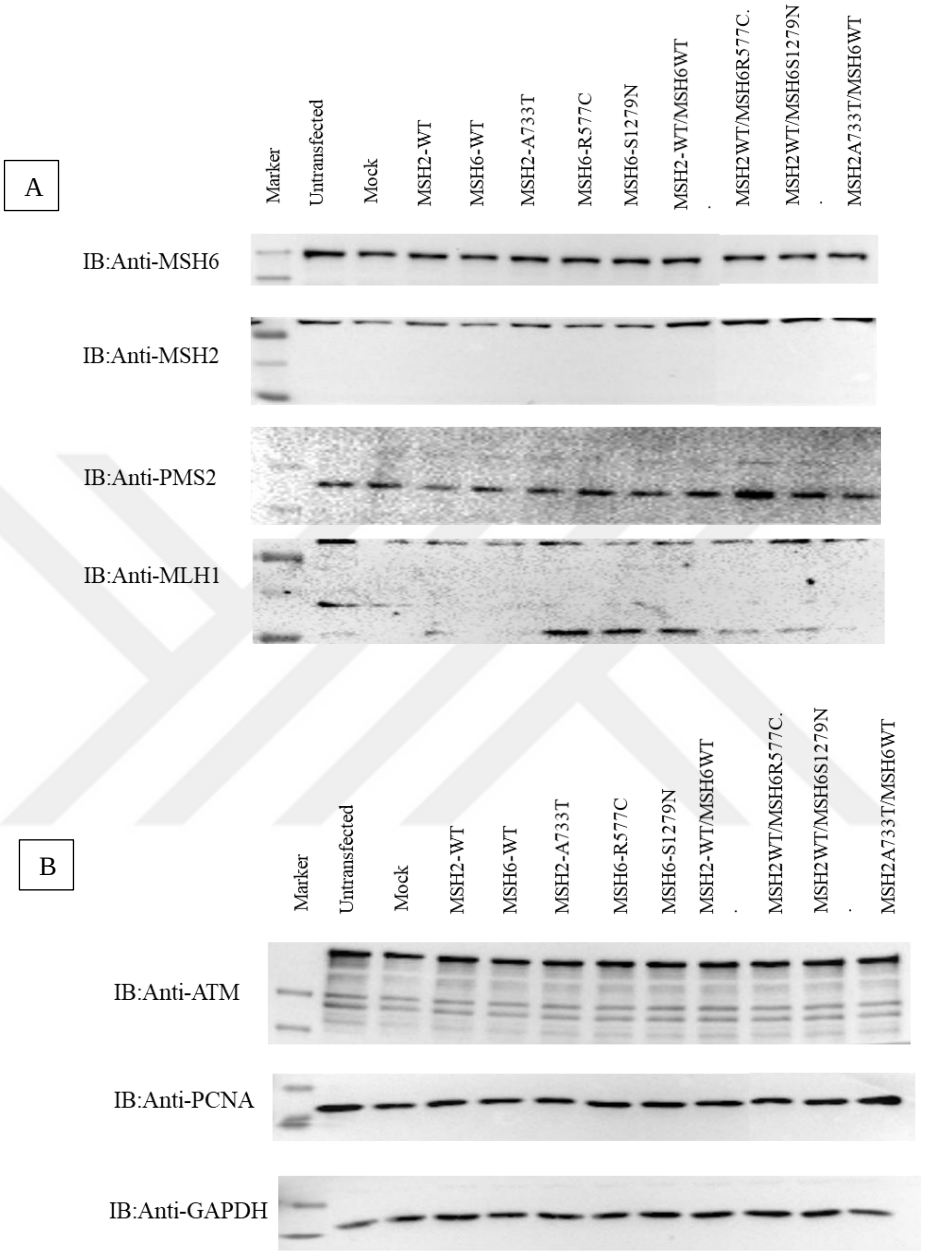

Figure 3.14: Immunoblotting result of the whole protein content of the LoVo cells transfected/co-transfected with MSH2/MSH6 plasmids in the presence of primary complex protein antibodies; A) MSH2 pAb (1:1000), MSH6 pAb (1:1000), and PMS2 pAb (1:100), MLH1 pAb (1:100), B) interaction partners and internal control protein antibodies; ATM pAb (1:1000), PCNA pAb (1:1000), GAPDH pAb (1:1000).

3.8 Co-Immunoprecipitation

In order to examine whether selected VUSes disrupt the MutSalpha complex formation by blocking the interaction of MSH2 and MSH6, MSH2 was immunoprecipitated from LoVo cells transfected with either mock, MSH2-WT, or MSH2-A733T by using a monoclonal anti-MSH2 primary antibody. Immunoblotting showed that the immunoprecipitation of MSH2 was successful, with slightly increased band intensities in MSH2-WT and MSH2-A733T transfected LoVo cell lysates (Figure 3.15). In both conditions, including untransfected and mock transfected LoVo cells, MSH6 was co-immunoprecipitated with both MSH2-WT and MSH2-A733T. This result indicated that the overexpressed MSH2 proteins interact with hMSH6 *in-vitro* to form MutSalpha complex, and mutant hMSH2 did not show significant effect on the binding of the two proteins.

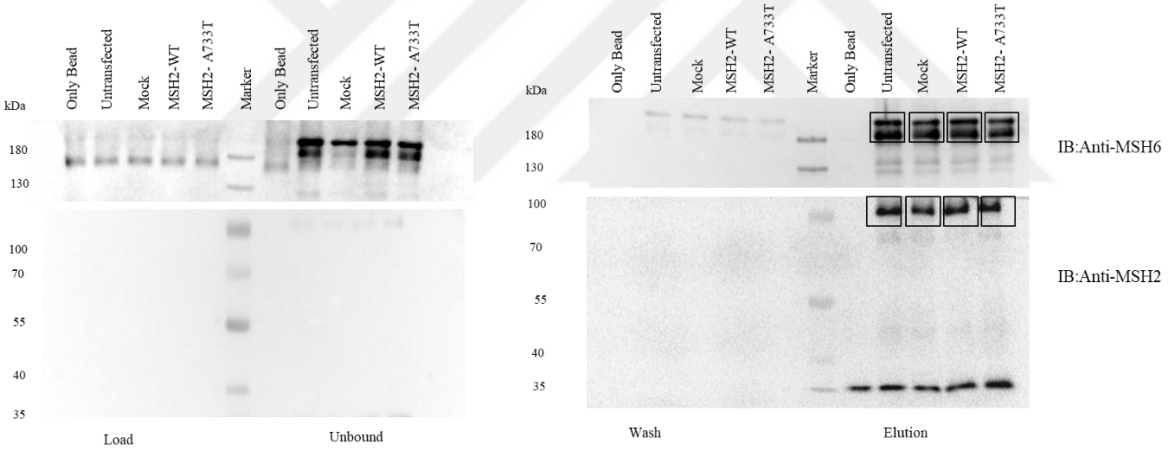

Figure 3.15: Immunoblotting result of the co-immunoprecipitation of MSH2 proteins from transfected LoVo cells, in the presence of MSH2 pAb (1:1000) and MSH6 pAb (1:1000).

3.9 Pull-down Assay

In order to investigate the *in-vitro* bindings of hMSH2 and hMSH6, a pull-down assay was performed. His-tagged MSH2-WT and MSH2-A733T proteins that were overexpressed in HEK293T cells, were targeted with Ni-NTA resins to be purified from cell lysates. Bounded proteins were incubated with Flag-tagged MSH6-WT, MSH6-R577C, or MSH6-S1279N transfected HEK293T cell lysates. Results showed that the mutant forms of the MSH6 interact with the overexpressed MSH2-WT

proteins, indicating that mutations have no significant effect on binding of hMSH2 and hMSH6 *in-vitro* (Figure 3.16).

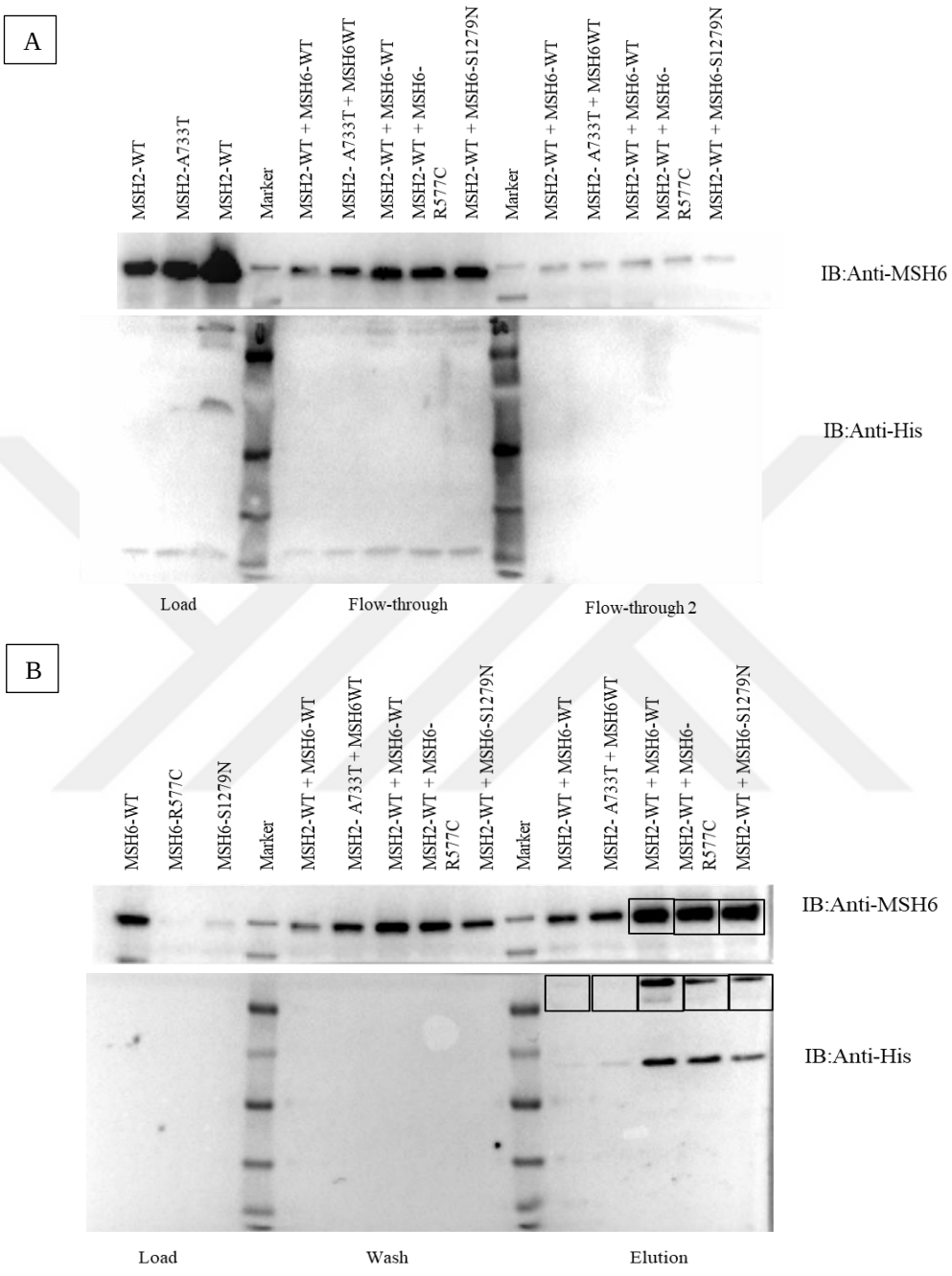

Figure 3.16: Immunoblotting result of the pull-down assay of the His-tagged MSH2 proteins and Flag-tagged MSH6 proteins produced in HEK293T or LoVo cells by using Ni-NTA resin, in the presence of A) and B) MSH6 pAb, His pAb, and C) MSH2 pAb, FLAG pAb.

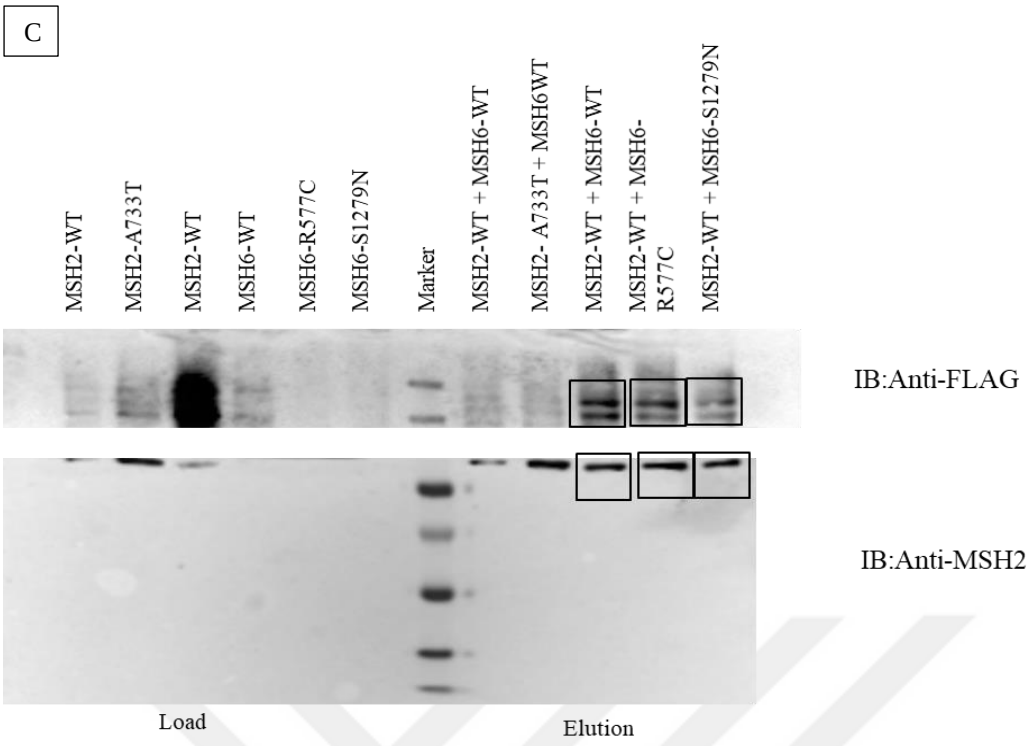

Figure 3.16 (continued): Immunoblotting result of the pull-down assay of the His-tagged MSH2 proteins and Flag-tagged MSH6 proteins produced in HEK293T or LoVo cells by using Ni-NTA resin, in the presence of A) and B) MSH6 pAb, His pAb, and C) MSH2 pAb, FLAG pAb.

3.10 Proteasome Inhibition on mRNA and Protein Expression

hMSH6 is known to be unstable in the absence of hMSH2, and the expression of hMSH6 is tightly regulated by hMSH2 expression. To see the effect of hMSSH2 on MSH6 regulation when the ubiquitin dependent proteasome pathway is inhibited, MG132 was implemented in LoVo cells transfected with either mock, MSH2-WT, or MSH2-A733T. mRNA expression level of hMSH6 was quantified using qRT-PCR. Moreover, protein expression levels of hMSH2 and hMSH6 were assessed by immunoblotting. Transcriptional expression levels of hMSH6 were decreased when LoVo cells were transfected with MSH2-A733T even ubiquitin-proteasome system is inhibited by MG132 compared to LoVo cells transfected with MSH2-WT (Figure 3.17). When protein expression profiles of hMSH2 and hMSH6 examined, it was determined that MSH6 levels were decreased when cells were treated with MG132, while MSH2 levels did not significantly changed (Figure 3.18). The membranes were blotted for ubiquitin to confirm that the MG132 treatment was successful.

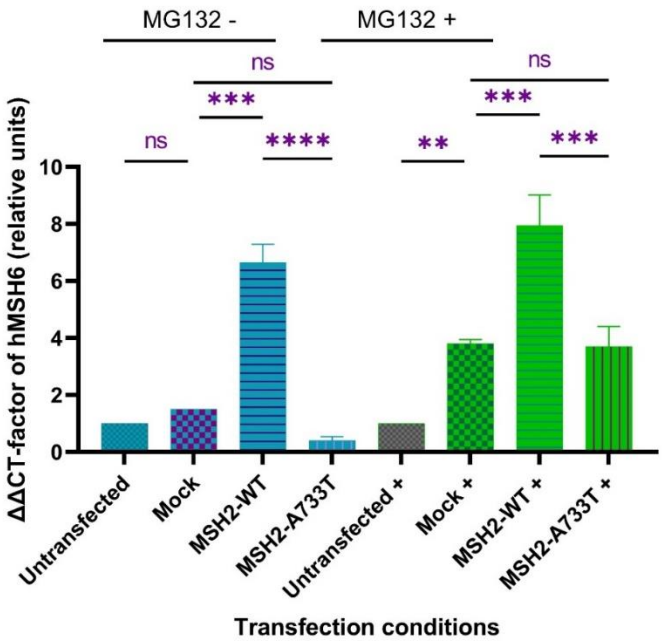

Figure 3.17: ΔΔCT- factor of hMSH6 gene expression pattern as a result of transfection of LoVo cells normalized to Beta-Actin gene expression levels no treated or treated with MG132 ($p<0.05$, $**p<0.01$, $***p<0.001$, $****p<0.0001$, One-way ANOVA, t-test, 95% Confidence Interval).

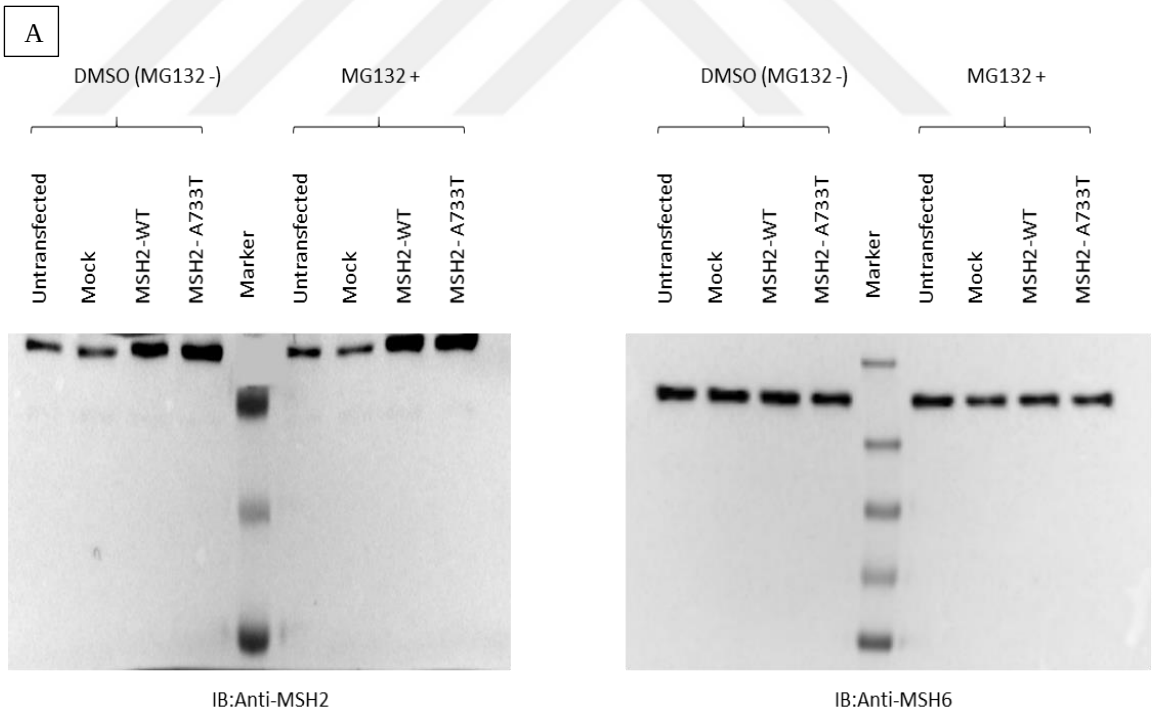

Figure 3.18: Immunoblotting result of the whole protein content of the LoVo cells transfected with MSH2-WT or MSH2-A733T, treated with DMSO (negative control) or MG132 (proteasome inhibitor), in the presence of A) MSH2 pAb (1:1000), MSH6 pAb (1:1000), B) Ubiquitin pAb (1:1000), and GAPDH pAb (1:1000).

B

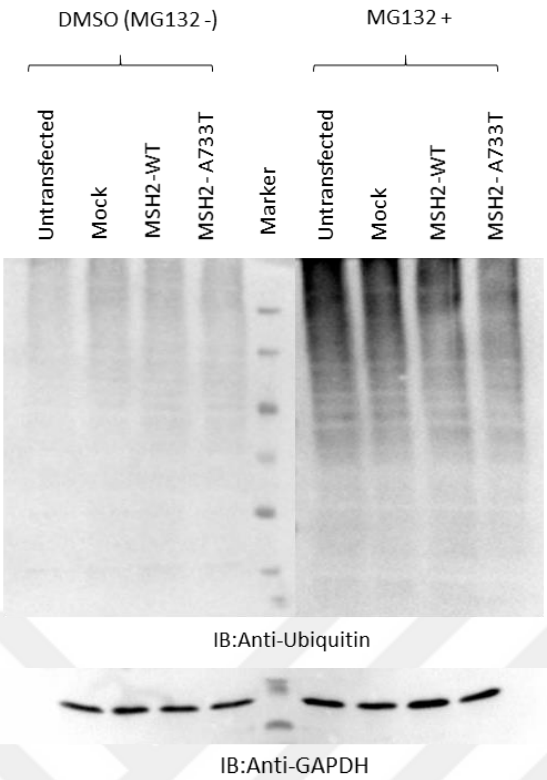

Figure 3.18 (continued): Immunoblotting result of the whole protein content of the LoVo cells transfected with MSH2-WT or MSH2-A733T, treated with DMSO (negative control) or MG132 (proteasome inhibitor), in the presence of A) MSH2 pAb (1:1000), MSH6 pAb (1:1000), B) Ubiquitin pAb (1:1000), and GAPDH pAb (1:1000).

3.11 Purification of hMSH2-WT and hMSH2-A733T

Immobilized metal ion affinity chromatography (IMAC) was preferred to purify the His-tagged hMSH2 proteins. hMSH2-WT and hMSH2-A733T proteins that were overexpressed in HEK293T cells were purified from cell lysates by using HisTrap FF crude 1 mL column. The chromatogram of the run was visualized with the assistance of Unicorn 7.0 software (Figure 3.19, 3.20 and 3.23). Peaks detected in elution fractions were collected and pooled for Anion exchange chromatography (AEX) as 2nd step purification. After each run, purification fractions were collected and run on gels as well. Although, hMSH2-WT could not be visualized on both SDS gel and membrane, hMSH2-A733T fractions were shown in SDS gel and immunoblotting after IMAC Figure 3.21. Even if the SDS gel and blotting bands were not observed for elution fractions of hMSH2-WT and hMSH2-A733T, when the fractions were pooled and concentrated, they were detected in SDS gel (Figure 3.23). The elution fractions that corresponded to the peak on chromatograph (Figure 3.22) were concentrated and run on 8% polyacrylamide gels with BSA standards (Figure 3.23). Yet, the bands were

detected at lower molecular weight than expected, indicating that the purified proteins were unstable and degraded.

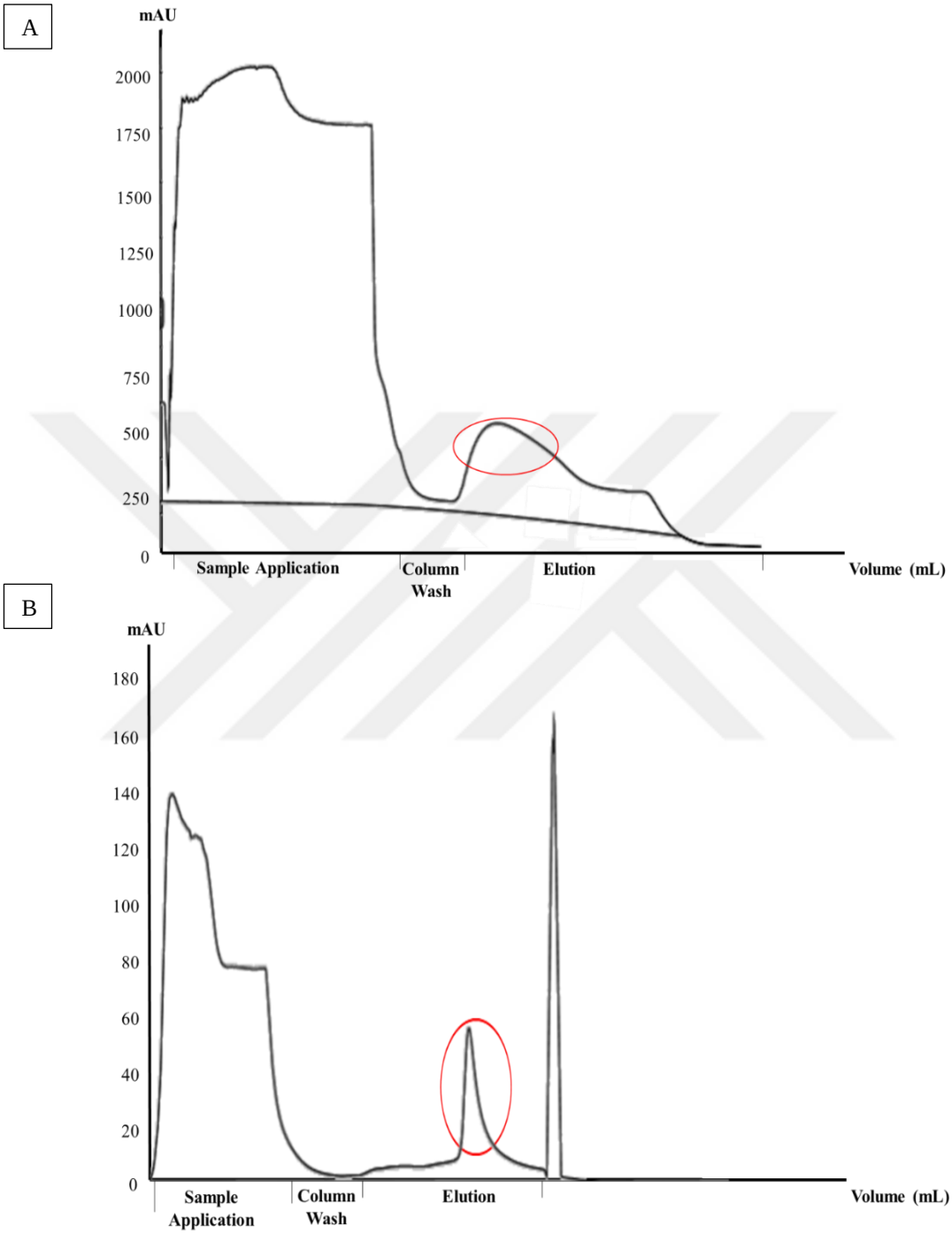

Figure 3.19: Chromatograph of the purification of hMSH2-WT produced in HEK293T cells, after A) IMAC and B) AEX.

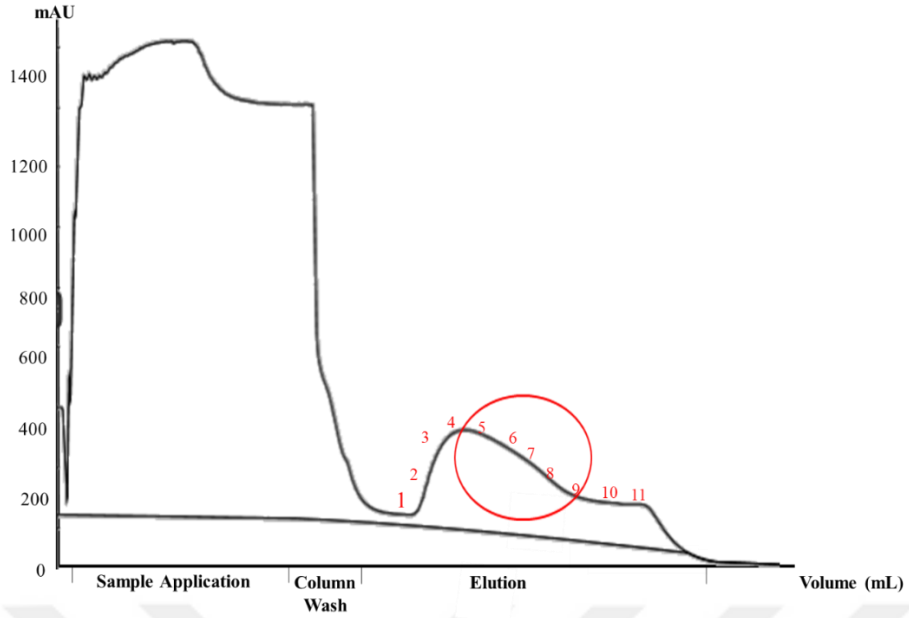

Figure 3.20: Chromatograph of the IMAC purification of hMSH2-A733T produced in HEK293T cells.

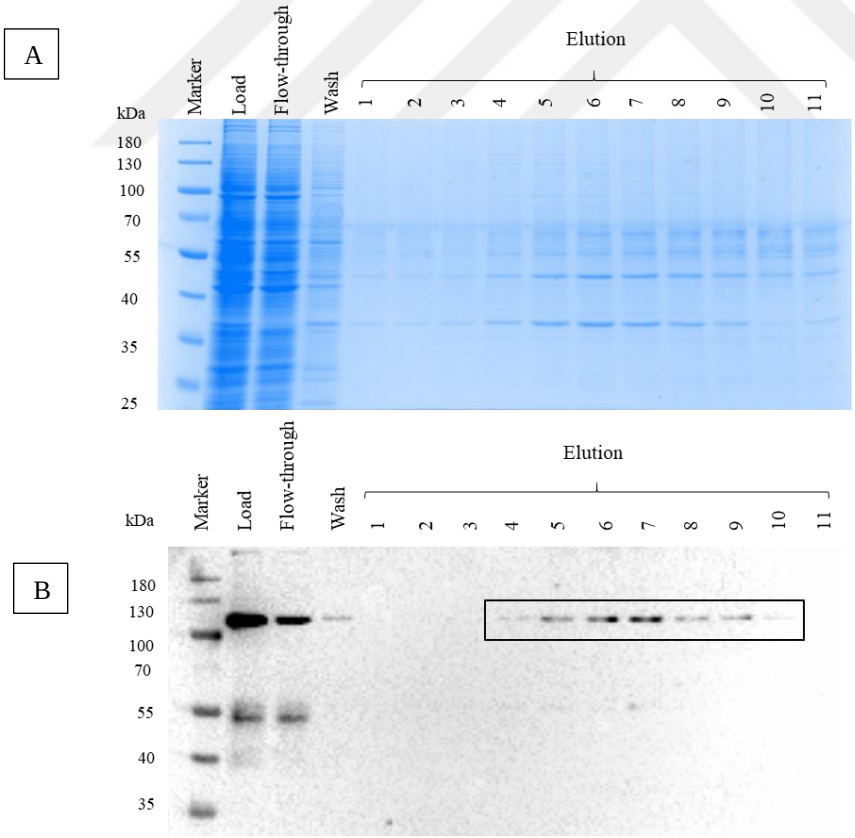

Figure 3.21: A) SDS-PAGE and B) immunoblotting result of IMAC purification of hMSH2-A733T produced in HEK293T cells, in the presence of His pAb (1:1000).

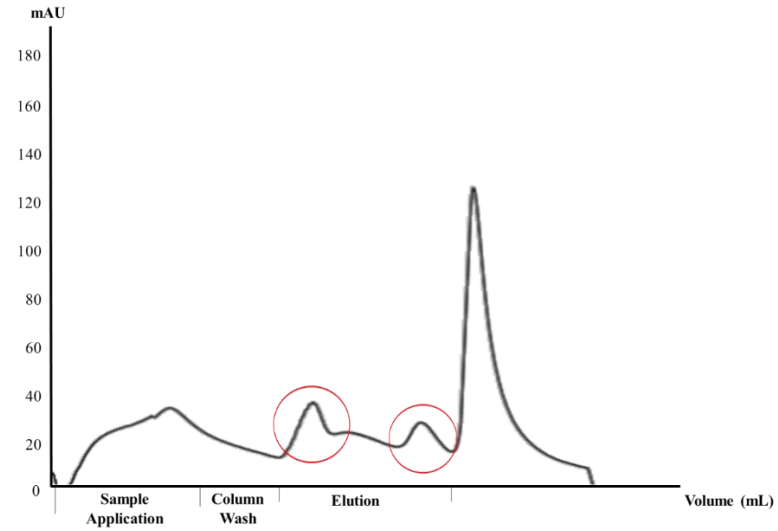

Figure 3.22: Chromatograph of the AEX purification of hMSH2-A733T produced in HEK293T cells, following IMAC purification.

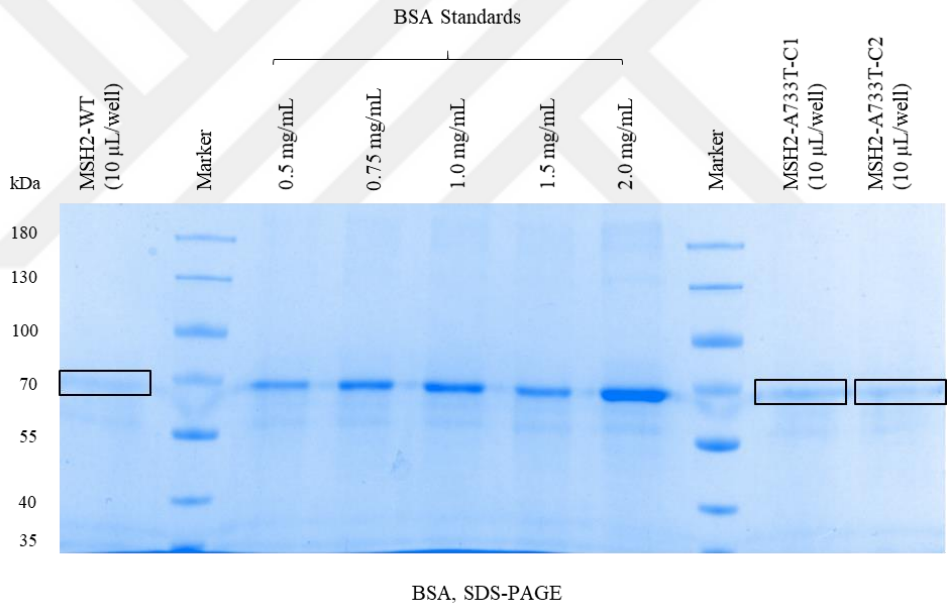

Figure 3.23: SDS-PAGE image of the concentrated purified MSH2-WT and MSH2-A733T that run with BSA standards.

3.12 Bacterial Expression

3.12.1 Co-expression optimization

As one of the target proteins is unstable in the absence of the other, co-production of both proteins in bacterial cells was tried by simultaneously transforming BL21 (DE3) and Rosetta (GroEL/ES containing) cells with hMSH2-pET30a and hMSH6-pET15b. Co-expression of proteins were tried in 3 different temperatures (17 °C, 30 °C, and 37 °C) after they were induced by 0.5 mM IPTG. Cells were lysed with lysis buffer (25

mM Tris-HCl, 300 mM NaCl 1.0 mM EDTA, 1.0 mM PMSF, 5% glycerol, pH: 8.0.) and sonicated using the KE76 probe for a minute at intervals of 5 seconds on and 10 seconds off pulses at 40% amplitude. Fractions were run on 8% polyacrylamide gels. Figure 3.24 illustrates that hMSH2/hMSH6 proteins were soluble in BL21 cells that were grown only at 30 °C in the presence of 0.5 mM IPTG.

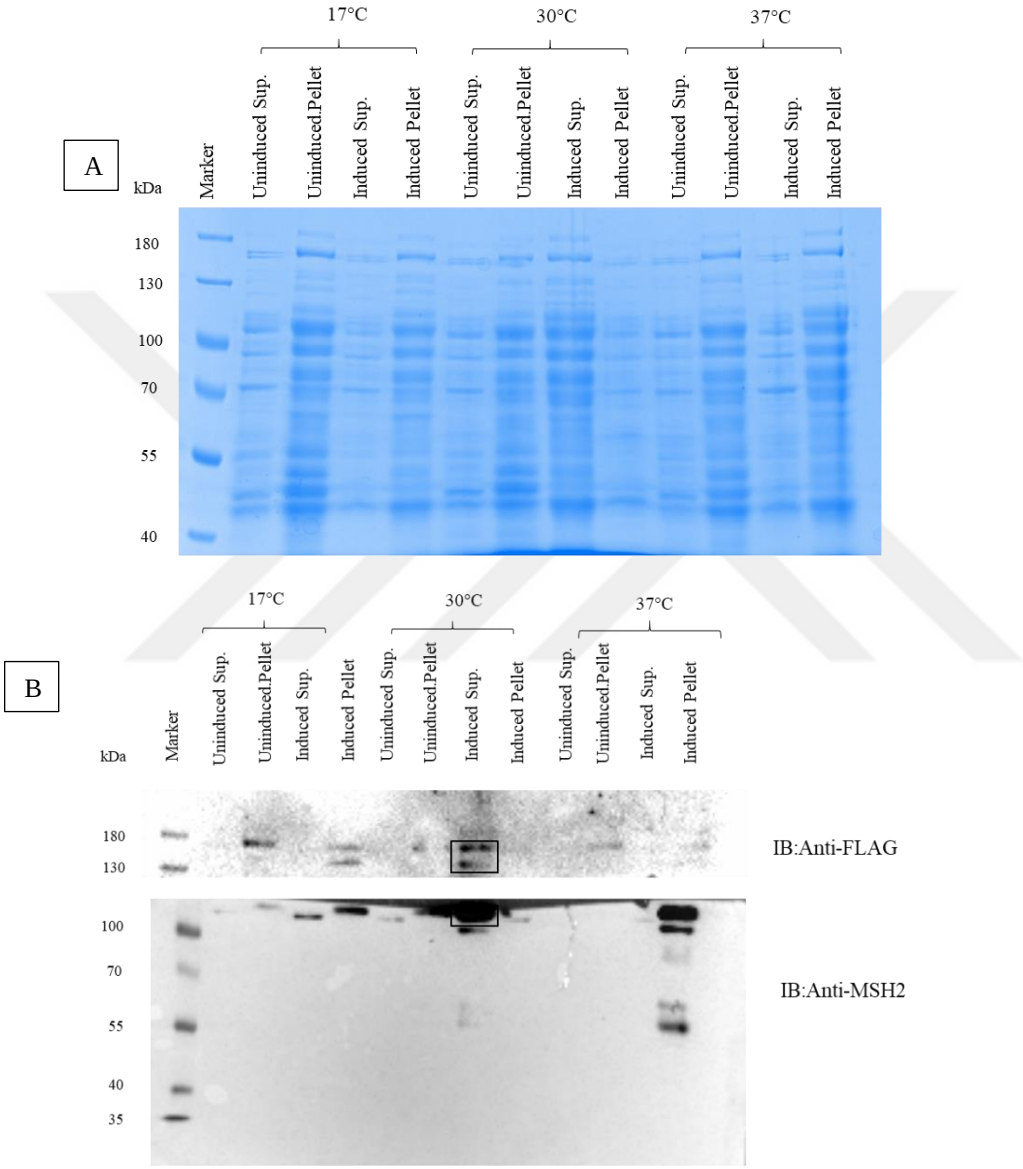

Figure 3.24: A) SDS-PAGE and B) immunoblotting results of whole cell lysate of the co-expression of hMSH2-pET30a/hMSH6-pET15b in *E. coli* BL21 cells at 17 °C, 30 °C and 37 °C in the presence of 0,5 mM IPTG, blotted with FLAG pAb (1:1000) or MSH2 pAb (1:1000).

Moreover, soluble hMSH2/hMSH6 proteins were expressed only at 30 °C in the presence of 0.5 mM IPTG in Rosetta cells *(E. coli*). When the performance of both *E.*

coli cells was compared, BL21 was best for hMSH2 protein and Rosetta was better for soluble hMSH6 protein in the presence of their co-expressed partners (Figure 3.25).

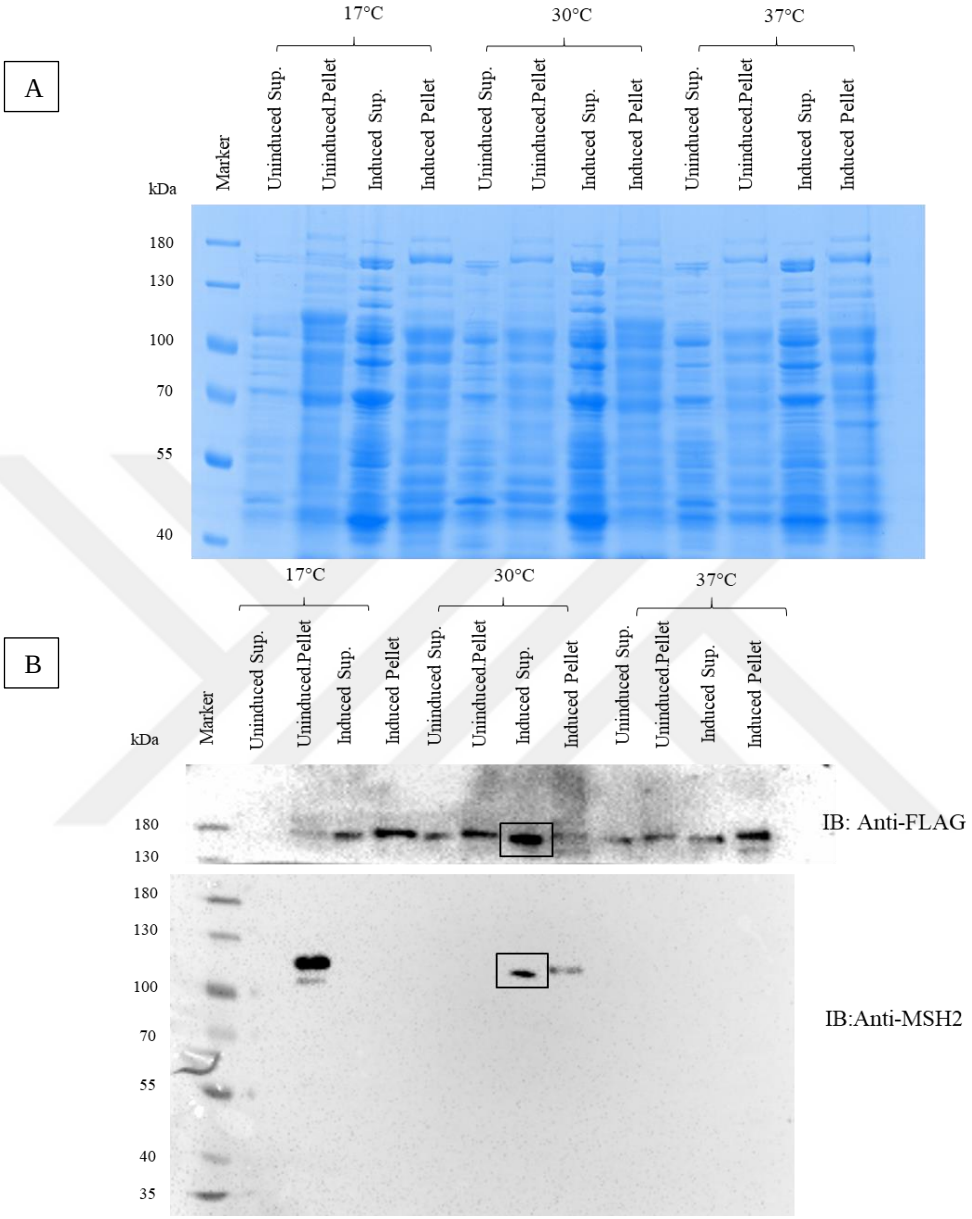

Figure 3.25: SDS-PAGE and immunoblotting results of whole cell lysate of the co-expression of hMSH2-pET30a/hMSH6-pET15b in *E. coli* Rosetta (GroEL/ES) cells at 17 °C, 30 °C and 37 °C in the presence of 0,5 mM IPTG, blotted with FLAG pAb (1:1000) or MSH2 pAb (1:1000).

3.12.2 Purification optimization

In order to purify hMSH2 and hMSH6 proteins as complex from Rosetta/GroEL/ES cells that were co-transfected with hMSH2-pET30a, and hMSH6-pET15b, anion exchange chromatography (AEX) was performed using a HiTrap Crude, 5 mL column. Figure 3.26 demonstrates the chromatograph result of the first step purification, and

purification fractions were run on 8% polyacrylamide gels, both SDS-PAGE and immunoblotting results were obtained (Figure 3.27).

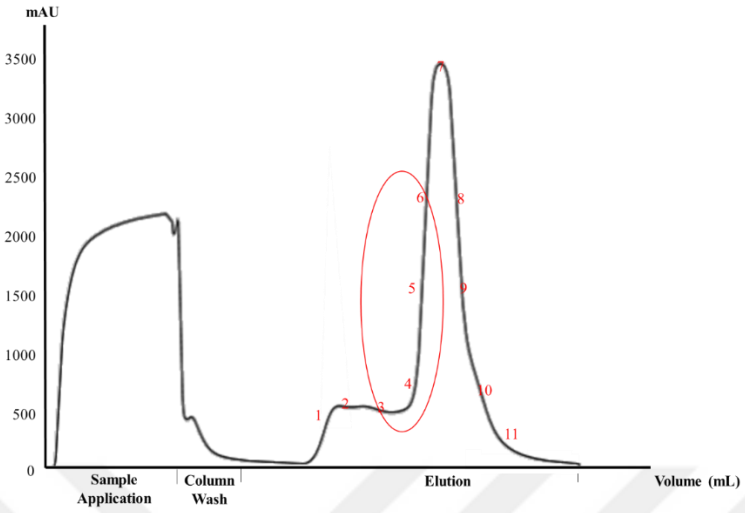

Figure 3.26: Chromatograph of the AEX purification of hMSH2-WT and hMSH6-WT expressed in *E. coli* Rosetta (GroEL/ES) cells.

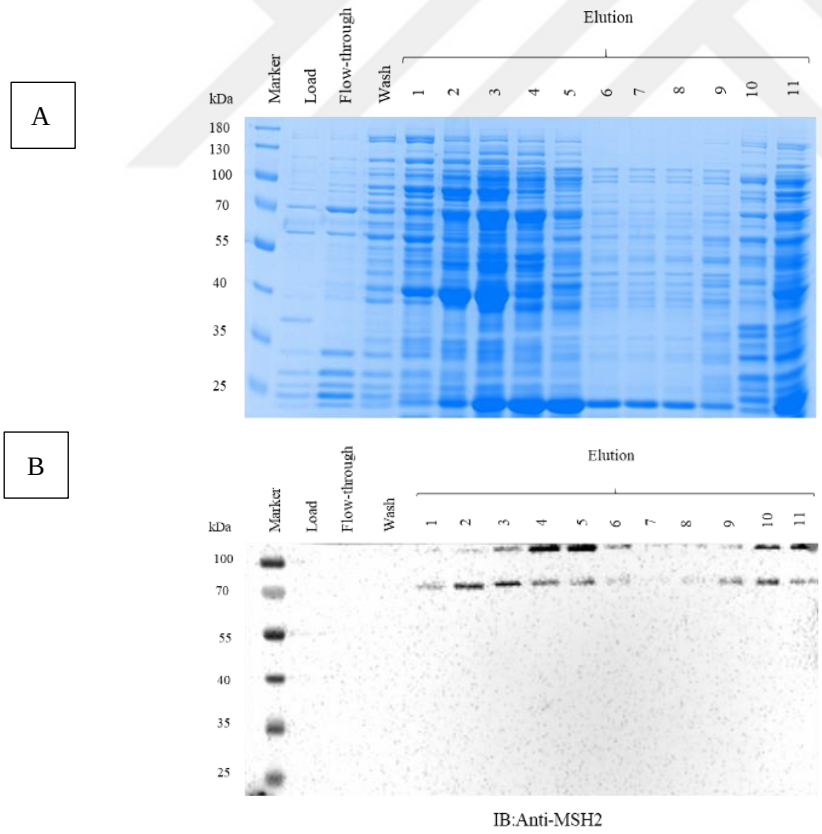

Figure 3.27: A) SDS-PAGE and B) Immunoblotting results of anion exchange (AEX) purification of hMSH2-WT and hMSH6-WT expressed in *E. coli* Rosetta (GroEL/ES) cells, in the presence of MSH2 pAb (1:1000).

Peaks detected in elution fractions were collected and pooled for 2nd step Heparin affinity chromatography (Figure 3.26). Chromatograph result of the 2nd step purification was illustrated on Figure 3.28. SDS-PAGE and immunoblotting results showed that both hMSH2 and hMSH6 furtherly purified by heparin affinity chromatography (Figure 3.29).

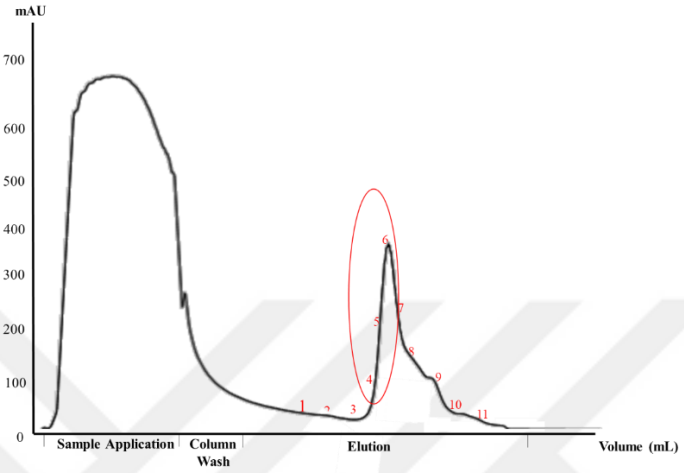

Figure 3.28: Chromatograph of the Heparin affinity purification of hMSH2-WT and hMSH6-WT expressed in *E. coli* Rosetta (GroEL/ES) cells, after AEX chromatography.

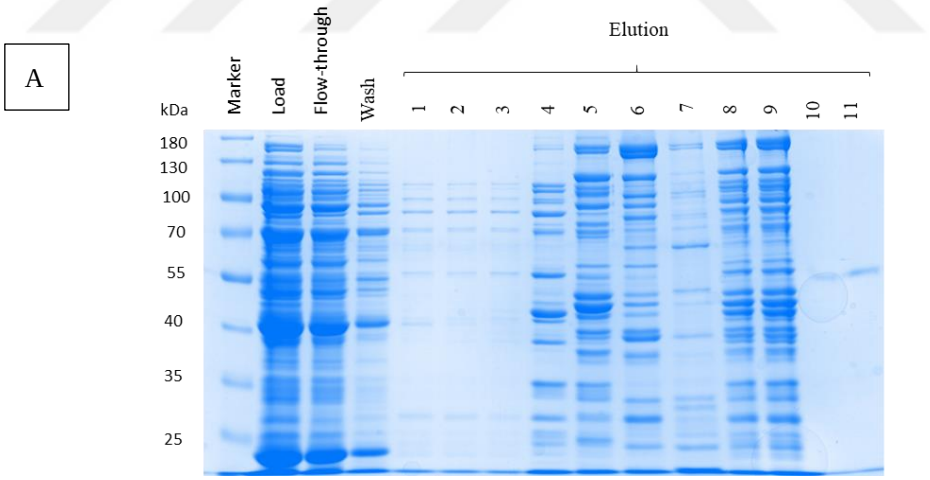

Figure 3.29: SDS-PAGE and immunoblotting results of Heparin affinity purification of hMSH2-WT and hMSH6-WT expressed in *E. coli* Rosetta (GroEL/ES) cells after AEX purification, in the presence of FLAG pAb (1:1000) and MSH2 pAb (1:1000).

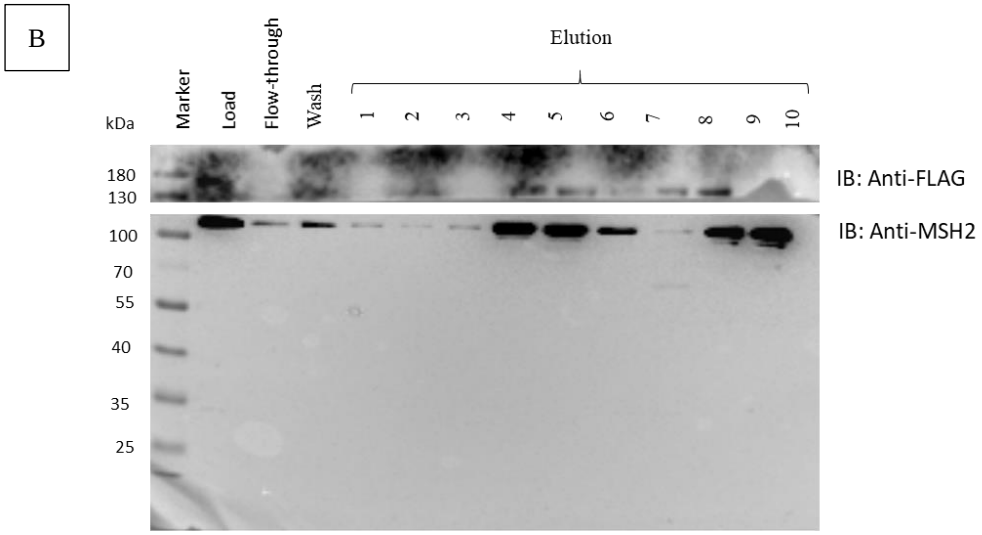

Figure 3.29 (continued): SDS-PAGE and immunoblotting results of Heparin affinity purification of hMSH2-WT and hMSH6-WT expressed in *E. coli* Rosetta (GroEL/ES) cells after AEX purification, in the presence of FLAG pAb (1:1000) and MSH2 pAb (1:1000).

4. DISCUSSION

Colorectal cancer is the third most commonly diagnosed cancer and the second leading cause of cancer-related mortality worldwide (GLOBOCAN, 2020). Moreover, 2- 5% of all incidences are hereditary cancer syndromes caused by inherited germline pathogenic variants (Jasperson et al., 2010). Lynch syndrome or hereditary non-polyposis colorectal cancer is predominantly caused by alterations in the MMR genes *(MLH1, MSH2, MSH6,* and *PMS2)* and is characterized by an elevated risk of notably developing colorectal and endometrial cancers (Jass, 2006). Having a clear genetic diagnosis is crucial for the clinical management of LS patients, as it enables the implementation of screening, preventive, and personalized therapeutic measures that significantly benefit the patients. Therefore, identifying pathogenic variants in families can offer testing to family members, sparing non-carriers from frequent surveillance and including carriers in cancer surveillance programs. The utilization of next-generation sequencing technology has enabled the identification of alterations in the MMR genes and the diagnosis of Lynch syndrome through the use of multigene panel testing. This study enabled the screening of the variant spectrum of Turkish colorectal cancer patients in such a large population for the first time (Akcay et al., 2010; Erdem and Bahsi, 2020).

Investigation of variant analyses revealed that 19% (n = 72) of the patients were pathogenic variant carriers, while 30% (n = 111) of them were carrying variants of uncertain significance (VUS). Most pathogenic variant carrying gene was *MUTYH,* yet since only biallelic pathogenic variants are considered positive for having MUTYH-associated polyposis (MAP), it was excluded (Akcay et al., 2010). Among identified pathogenic MMR gene variants, *MLH1* and *MSH2* were the most pathogenic variant containing genes with 21,21% and 13,63%, respectively, while 4,54% of all pathogenic variants was accounted by both *MSH6* and *PMS2*. When only MMR related genes were considered, 43,92% of all pathogenic variants were detected, indicating a wide range of prevalence. The VUS number distribution among genes showed that the normalized variant numbers were higher in *MSH2* (4,99) and *MLH1* (4,84) compared to *MSH6* (2,69) and *PMS2* (0,38). However, since they both function as heterodimers,

it is important to select heterodimer partners for further studies. Normalized MutSalpha complex VUS numbers were 7,68, while MutL complex contained 5,22. Furthermore, a pre-study conducted by Akcay et al. (2010) indicated that *MSH2* and *MSH6* were the most VUS containing mismatch repair mechanism genes found in a cohort including colorectal cancer patients, breast cancer patients and healthy individuals in Turkish population. Therefore, VUS found in the MutSalpha complex was selected for functional studies at transcriptional and translational levels.

3 variants were selected concerning non-conflicting evidence of pathogenicity, meaning that there was no conflicting interpretation on databases among the 10 different *MSH2*, and 8 different *MSH6* variants detected. Moreover, functional studies have not been reported for these selected variants. *MSH2* c.2197G>A (p. Ala733Thr) and *MSH6* c.3836G>A (p. Ser1279Asn) were selected for their location on the ATPase domain and may have an important effect on functionality of complex. MSH2 and MSH6 ATPase domains are required for DNA lesion recognition, MMR process activation, interaction with other MMR proteins, and repair machinery coordination (Gorman et al., 2007). ATP binding and hydrolysis activities trigger conformational changes, which permit the rapid and accurate repair of DNA mismatches and IDLs. *MSH6* c.1729C>T (p. Arg577Cys) is located on connector domain of MSH6. A conserved hydrophobic core in the connector domain establishes an interaction with the homologous domain in MSH2, allowing for stable dimerization (Mendillo et al., 2009). Furthermore, the connector domain comprises a flexible loop area that is critical for recognizing certain DNA structures like loops and bulges, which are frequently associated with DNA mismatches. In addition, changing the positively charged amino acid group into a neutral amino acid may have affected the binding interactions.

To study structure and function of the selected VUS, first, the full length of coding sequences of wild-type (WT) hMSH2 and hMSH6 were taken from NCBI (CCDS1834.1 and CCDS1836.1) and custom-cloned into cloning vectors by AddGene to minimize genomic instability and maximize accuracy (Smith, 2007). hMSH2-WT and hMSH6-WT sequences were amplified from plasmids containing gene targets and cloned with C terminal end modifications in both bacterial and mammalian expression vectors to facilitate purification.

HEK293T cells were used to overexpress target proteins since it is used for eukaryotic protein production due to their high transfection efficiency, human origin, scalability,

suspension culture capacity, and compatibility with transient transfection (Iuchi et al., 2020). Overexpression optimization results showed that the 48h after transfection was the best condition for protein production. There was no study on overexpression of hMSH2 and hMSH6 in HEK293T cells. MSH2 and MSH6 are found in the highest amounts in the cell when the cells are in G1 phase, and expression of the MMR proteins increases when cells are exposed to DNA damage (Gorman et al., 2007). So, overexpression time was expected when the doubling time of HEK293T cells, which is about 12-20 hours, was considered (Lomba et al., 2021). Detection of proteins through anti-FLAG and anti-His primary antibodies confirmed that the proteins were overexpressed. Proteins were observed at their expected molecular weights as MHS6 in the 160 kDa range, and MSH2 in the 105 kDa range (Pećina-Šlaus et al., 2020).

Even if there were non-specific bonds on the membrane by antibodies, it was shown that the overexpression of target proteins increased the band intensities. Optimized amount of plasmid, selected in terms of the most steady and abundant protein production, was 1.5 μg for both proteins.

The LoVo cells were selected since they lack expression of one or more MMR proteins, primarily MSH2 (Brieger et al., 2002), and are useful to study MMR deficiency and its implication on DNA repair and cancer development (Kantelinen et al., 2012). Overexpression of hMSH2 in LoVo cells presented an increase in protein expression level concordant with previous studies in the presence of both forms of protein, WT or A733T mutated version (Figure 3.15 and Figure 3.18) (Kantelinen et al., 2010).

MTT assay which was used to measure cytotoxic effects of plasmids containing WT or mutated forms of genes in LoVo or HCT116 cells, revealed that there was no significant change between plasmids normalized to untransfected cells except for 48 h in LoVo cells. Proliferative activity of the transiently transfected cells and controls decreased after 48 hours for LoVo cells, which did not indicate that overexpression of WT or mutated forms of genes whether affected or not proliferation. The inconclusive result was caused by the instability of LoVo cells, which was not consistent with the studies conducted by Brieger et al. (2002).

hMSH6 (human MutS homolog 6) expression is tightly regulated at the transcriptional level, as illustrated in Figure 4.1. Numerous cis-acting sites in the hMSH6 promoter

region bind to multiple transcription factors and influence gene expression. Sp1 (specificity protein 1) binds to the GC-rich areas of the hMSH6 promoter, and is required for basal transcriptional activity, modulating hMSH6 expression positively (Gorman et al., 2007; Mendillo et al., 2009). Other transcription factors that bind to the hMSH6 promoter include AP-1, AP-2, E2F, and NF-kB (Mendillo et al., 2009). hMSH6 expression is also influenced by epigenetic modifications such as DNA methylation and histone acetylation, in addition to transcription factors (Mendillo et al., 2009). Moreover, it has been shown that MSH2 interacts with transcription factors, co-activators, and co-repressors (such as CTCF and EZH2) to change the activity of the MSH6 gene promoter (Gorman et al., 2007; Mendillo et al., 2009). Depending on the context and cellular conditions, these interactions can either increase or inhibit MSH6 transcription.

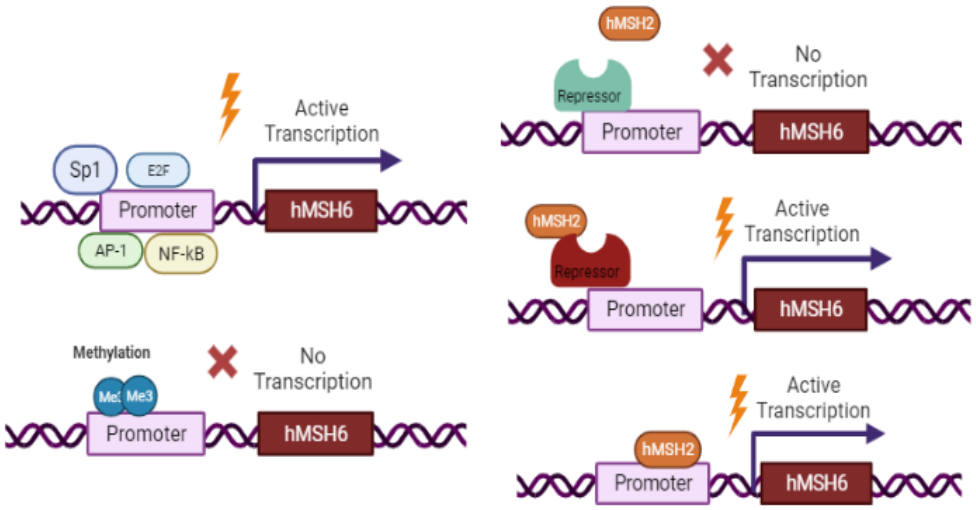

Figure 4.1: Representation of hMSH6 transcriptional regulation (Created with BioRender.com).

So, the first question was how MSH6 expression is affected by either overexpression of MSH2-WT or MSH2-A733T. LoVo cells were overexpressed with target plasmids, and transcriptional level of MSH6 was investigated by targeting the Exon 4 region of MSH6 by qRT-PCR. Results demonstrated that overexpression of MSH2-WT increased the MSH6 expression level, while MSH2-A733T overexpression caused a slight increase compared to mock transfection and a sharp decrease compared to MSH2-WT overexpression (Figure 3.13). Considering the effect of only MSH2 expression in the regulation of MSH6 expression, some scenarios emerge that affect

MSH6 expression level. Lower expression levels of MSH2, decreased levels of binding affinity of MSH2 to MSH6, an increase in the degradation target of expressed MSH6, or an unstable form of MSH2-A733T may have caused a decrease in MSH6 expression levels.

Investigation of protein expression levels, depending on overexpression of MSH2, MSH6 or MSH2/MSH6 for WT and mutated forms of each protein showed that there was no significant difference between untransfected and transfected versions of cells, only MSH2 transfected LoVo cells showed a slight increase in MSH2 protein production. It was expected that the downstream interaction partner complex (MutL) proteins MLH1 and PMS2 levels were not altered as well as the studies conducted by Kantelinen et al. (2010; 2012).

The ubiquitin-proteasome system, often known as the UPS, is a process for the breakdown of proteins in cells that is highly controlled mechanism play role in the regulation of protein turnover, the management of the quantity of proteins, and the preservation of protein homeostasis (Gong et al., 2016). It is well established that hMSH6 cannot function properly in the absence of hMSH2, and the expression of hMSH6 is stringently controlled by the expression of hMSH2. It was purposed to see the amount of mRNA and protein expression levels of hMSH6 when UPS is inhibited by proteasome inhibitor MG132 (Tarjanyi et al., 2022). Transcriptional expression of hMSH6 was increased when MG132 was applied, yet the expression profile was not altered compared to the negative control. Moreover, MSH2 proteins were accumulated as expected while MSH6 levels were decreased when MG132 was applied, which may be caused by MSH6 proteins being degraded by other degradation processes such as autophagy-associated and xenophagy-associated proteasome systems (Kocaturk & Gozuacik, 2018).

The next step was to investigate the binding of produced MSH2 and MSH6 proteins *in-vitro* by co-immunoprecipitation. LoVo cells were transfected with MSH2-WT or MSH2-A733T only to see the effect of binding with endogenous MSH6 proteins. It was examined that mutation did not influence binding of MSH2 and MSH6 proteins *in-vitro*.

Interaction of MSH2 and MSH6 proteins were also investigated *in-vitro* through pull-down assay. Histidine-tagged MSH2-WT and MSH2-A733T proteins were purified

by Ni-NTA resins, then MSH6-WT, MSH6-R577C, and MSH6-S1279N including cell lysates were incubated with Ni-NTA bounded His-tagged proteins. Results showed that the mutations did not affect the binding of proteins in vitro. As a result of the pull-down assay experiment, it was found that the mutation in the connector area of MSH6 (R577C), which was presumed to have an anticipated effect on binding and stability due to its location, did not result in a change in binding in the predicted region. Therefore, investigation of the binding ability of MSH2 with MSH6 revealed that there was no effect on binding depending on the presence of mutation.

The effect of mutations on protein stability was investigated by producing MSH2-WT and MSH2-A733T proteins in HEK293T cells and purifying by using their affinity tags. Degradation products were eliminated by applying anion exchange chromatography after immobilized metal ion affinity chromatography (IMAC). When concentrated final products were run on an 8% gel with BSA standards, it was seen that the proteins were degraded and observed to be around 70 kDa which was expected to be around 105 kDa. Stability of the recombinant protein may be affected by the intrinsic protein properties, experimental conditions, and storage methods. Moreover, histidine tag addition may have disrupted the folding characteristics of natural proteins, destabilizing the final product.

According to a study conducted by Antony and Hingorani (2003), co-expression of MSH2 and MSH6 plasmids in a bacterial expression system was successful, and both proteins were purified by using some purification steps. So, bacterial expression vectors containing hMSH2 and hMSH6 were separately co-transformed into BL21 and Rosetta GroEL/ES containing *E. coli* strains. BL21 strains are engineered to lack the lon and ompT proteases, which decrease recombinant protein breakdown, but they also lack post-translational modifications and the eukaryotic cellular environment required for optimum folding and function (Jeong et al., 2015).Transfecting Rosetta cells with GroEL/ES carrying plasmids modified the Rosetta strain, which adds a feature of the system that is a molecular chaperone found in bacteria that helps to fold freshly generated proteins and refold misfolded or aggregated proteins (Gupta et al., 2014).

Expression of both hMSH2 and hMSH6 was provided in both strains, in different incubation temperatures. 30 °C was the optimum temperature for protein production for both strains. Since Rosetta cells expressed proteins in higher amounts, they were selected. Purification was applied to cell lysate of Rosetta GroEL/ES cells expressing

hMSH2 and hMSH6. Firstly, AEX was applied since theoretical isoelectric points (pI) of hMSH2 (5.79) and hMSH6 (6.31) were lower than the selected buffer pH. After collecting a wide range of proteins, the second step of purification was selected as heparin affinity chromatography since DNA-binding proteins have strong affinity against highly sulfated glycosaminoglycan (GAG) structure (Xiong et al., 2008). Purification of both proteins were successful, however; there were nonspecific proteins purified under the same conditions. Hence, purification methods need to be optimized for both stable and correctly folded proteins in purified forms.

Full length MutSalpha complex proteins are used to study individual MSH2/MSH6 characterization via complex. MutSalpha is a heterodimer form of MSH2 and MSH6, accounting for approximately 260 kDa protein. Studies suggest that the expression of MutSalpha complex is most suitable in SF9 cells which are a cell line derived from the ovaries of the fall armyworm (*Spodoptera frugiperda*) (Kantelinen et al., 2010). They supply the essential cellular machinery for post-translational modifications, enable the creation of complex proteins, allow for native-like folding, give high protein yields, and keep MutS functional, which makes them suitable host for hMSH2/hMSH6 expression (Schneider & Seifert, 2010).

To conclude, a total of 3 VUS of MSH2 and MSH6 were selected to compare protein interaction, stability, and function with the wild-type proteins. Results showed that MSH2-A733T overexpression decreases transcriptional expression of hMSH6 compared to MSH2-WT. Selected mutations did not affect the interaction of MSH2/MSH6. However, since the proteins could not be obtained in full-length, folded, and stable form, stability and functional studies were not completed. Bacterial co-expression of hMSH2/hMSH6 demonstrated that soluble forms of proteins were successful. The purification method should be optimized, yet it may not fulfill the requirements of the complex concerning post-translational modifications. Domain-specific studies will be useful to study the functionality of the wild-type and mutated forms of proteins.

In order to be able to reclassify VUS of MSH2/MSH6, PS3 or BS3 evidence that "well-established functional studies show a (no) deleterious effect" needs to be confirmed. Thus, additional functional studies should be applied to make a commitment about the effect of mutation. *In-vitro* mismatch repair assay, which was discovered by Modrich (1987) and developed by Holmes et al. (1990), includes a mismatch-containing

construct, purified MutSalpha complex, MutSalpha-deficient nuclear extract, and other elements. It is one of the fundamental functional studies used for MMR. Moreover, MMR activity can be studied by using rapid cell-free assays discovered by Drost et al. (2010–2012), nuclear localization assays (Belvederesi et al., 2012), RNA splicing analysis (Tournier et al., 2008), in vivo stability assays (Lutzen et al., 2008), and 6-thioguanine selection assays in murine embryonic stem cells (mESCs) (Wielders et al., 2011). In conclusion, functional assays can provide valuable insight into the impact of a variant on the protein or mRNA level, making it crucial to reclassify VUS.

REFERENCES

Akcay, I. M., Celik, E., Agaoglu, N. B., Alkurt, G., Kizilboga Akgun, T., Yildiz, J., Enc, F., Kir, G., Canbek, S., Kilic, A., Zemheri, E., Ezberci, F., Ozcelik, M., Dinler Doganay, G., & Doganay, L. (2020). Germline pathogenic variant spectrum in 25 cancer susceptibility genes in Turkish breast and colorectal cancer patients and elderly controls. International Journal of Cancer, 148(2), 285–295. https://doi.org/10.1002/ijc.33199

Antony, E., & Hingorani, M. M. (2003). Mismatch Recognition-Coupled Stabilization of Msh2-Msh6 in an ATP-Bound State at the Initiation of DNA Repair. *Biochemistry*, *42*(25), 7682–7693. https://doi.org/10.1021/bi034602h

Baglietto, L., Lindor, N. M., Dowty, J. G., White, D. J., Wagner, A., Garcia, E. B. G., Vriends, A. H. J. T., Cartwright, N., Barnetson, R. A., Farrington, S. M., Tenesa, A., Hampel, H., Buchanan, D. D., Arnold, S., Young, J. P., Walsh, M., Jass, J. R., Macrae, F. A., Antill, Y., . . . Jenkins, M. A. (2010). Risks of Lynch Syndrome Cancers for MSH6 Mutation Carriers. *Journal of the National Cancer Institute*, *102*(3), 193–201. https://doi.org/10.1093/jnci/djp473

Belvederesi, L., Bianchi, F., Loretelli, C., Bracci, R., Cascinu, S., & Cellerino, R. (2012). Sub-cellular localization analysis of MSH6 missense mutations does not reveal an overt MSH6 nuclear transport impairment. *Familial Cancer*, *11*(4), 675–680. https://doi.org/10.1007/s10689-012-9558-y

Brieger, A., Trojan, J., Raedle, J., Plotz, G., & Zeuzem, S. (2002). Transient mismatch repair gene transfection for functional analysis of genetic hMLH1 and hMSH2 variants. *Gut*, *51*(5), 677–684. https://doi.org/10.1136/gut.51.5.677

Da Silva Calixto, P., Lopes, O. S., Maia, M. D. S., Herrero, S. S. T., Longui, C. A., De Melo, C. G. F., De Carvalho Filho, I. R., Soares, L. R., De Medeiros, A. C., Delatorre, P., Khayat, A. S., Burbano, R. R., & Lima, E. M. (2018). Single-Nucleotide Polymorphisms of the MSH2 and MLH1 Genes, Potential Molecular Markers for Susceptibility to the Development of Basal Cell Carcinoma in the Brazilian Population. *Pathology & Oncology Research*, *24*(3), 489–496. https://doi.org/10.1007/s12253-017-0265-8

Devakumar, L. J. P. S., Gaubitz, C., Lundblad, V., Kelch, B. A., & Kubota, T. (2019). Effective mismatch repair depends on timely control of PCNA retention on DNA by the Elg1 complex. *Nucleic Acids Research*, *47*(13), 6826–6841. https://doi.org/10.1093/nar/gkz441

Drost, M., Zonneveld, J. B. M., Van Hees, S., Rasmussen, L. J., Hofstra, R. M., & De Wind, N. (2012). A rapid and cell-free assay to test the activity of

lynch syndrome-associated MSH2 and MSH6 missense variants. *Human Mutation*, *33*(3), 488–494. https://doi.org/10.1002/humu.22000

Drost, M., Zonneveld, J. É. B., Van Dijk, L., Morreau, H., Tops, C. M. J., Vasen, H. F. A., Wijnen, J. T., & De Wind, N. (2010). A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. *Human Mutation*, *31*(3), 247–253. https://doi.org/10.1002/humu.21180

Duckett, D. R., Drummond, J. R., Murchie, A. I., Reardon, J. T., Sancar, A., Lilley, D. M., & Modrich, P. (1996). Human MutSalpha recognizes damaged DNA base pairs containing O6-methylguanine, O4-methylthymine, or the cisplatin-d(GpG) adduct. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(13), 6443–6447. https://doi.org/10.1073/pnas.93.13.6443

Edelbrock, M. A., Kaliyaperumal, S., & Williams, K. J. (2013). Structural, molecular and cellular functions of MSH2 and MSH6 during DNA mismatch repair, damage signaling and other noncanonical activities. *Mutation Research: Fundamental and Molecular Mechanisms of Mutagenesis*, *743–744*, 53–66. https://doi.org/10.1016/j.mrfmmm.2012.12.008

Genetic Alliance; The New York-Mid-Atlantic Consortium for Genetic and Newborn Screening Services. (2009) Understanding Genetics: A New York, Mid-Atlantic Guide for Patients and Health Professionals. Washington (DC): Genetic Alliance; 8. APPENDIX G, GENETIC TESTING. Available from: https://www.ncbi.nlm.nih.gov/books/NBK115571/

Gian, L. D. (2018). Microsatellite instability in colorectal cancer. *PubMed Central (PMC)*. https://doi.org/10.23750/abm.v89i9-S.7960

Gong, B., Radulovic, M., Figueiredo-Pereira, M. E., & Cardozo, C. (2016). The Ubiquitin-Proteasome System: Potential Therapeutic Targets for Alzheimer's Disease and Spinal Cord Injury. *Frontiers in Molecular Neuroscience*, *9*. https://doi.org/10.3389/fnmol.2016.00004

Gorman, J., Chowdhury, A. G., Surtees, J. A., Shimada, J., Reichman, D. R., Alani, E., & Greene, E. C. (2007). Dynamic Basis for One-Dimensional DNA Scanning by the Mismatch Repair Complex Msh2-Msh6. *Molecular Cell*, *28*(3), 359–370. https://doi.org/10.1016/j.molcel.2007.09.008

Gupta, A., Haldar, S., Miličić, G., Hartl, F. U., & Hayer-Hartl, M. (2014). Active Cage Mechanism of Chaperonin-Assisted Protein Folding Demonstrated at Single-Molecule Level. *Journal of Molecular Biology*, *426*(15), 2739–2754. https://doi.org/10.1016/j.jmb.2014.04.018

He, Y., Zhang, L., Zhou, R., Wang, Y., & Chen, H. (2022). The role of DNA mismatch repair in immunotherapy of human cancer. *International Journal of Biological Sciences*, *18*(7), 2821–2832. https://doi.org/10.7150/ijbs.71714

Holmes, J., Clark, S., & Modrich, P. (1990). Strand-specific mismatch correction in nuclear extracts of human and Drosophila melanogaster cell lines.

Proceedings of the National Academy of Sciences, *87*(15), 5837–5841. https://doi.org/10.1073/pnas.87.15.5837

Iuchi, K., Oya, K., Hosoya, K., Sasaki, K., Sakurada, Y., Nakano, T., & Hisatomi, H. (2020). Different morphologies of human embryonic kidney 293T cells in various types of culture dishes. *Cytotechnology*, *72*(1), 131–140. https://doi.org/10.1007/s10616-019-00363-w

Jasperson, K. W., Tuohy, T. M., Neklason, D. W., & Burt, R. W. (2010). Hereditary and familial colon cancer. Gastroenterology, 138(6), 2044-2058.

Jass, J. R. (2006). Hereditary non-polyposis colorectal cancer: The rise and fall of a confusing term. *World Journal of Gastroenterology*, *12*(31), 4943. https://doi.org/10.3748/wjg.v12.i31.4943

Jeong, H., Kim, H. J., & Lee, S. Y. (2015). Complete Genome Sequence of Escherichia coli Strain BL21. *Genome Announcements*, *3*(2). https://doi.org/10.1128/genomea.00134-15

Kaliyaperumal, S., Patrick, S. M., & Williams, K. J. (2011). Phosphorylated hMSH6: DNA mismatch versus DNA damage recognition. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 706(1–2), 36–45. https://doi.org/10.1016/j.mrfmmm.2010.10.008

Kantelinen, J., Kansikas, M., Candelin, S., Hampel, H., Smith, B., Holm, L., Kariola, R., & Nyström, M. (2012). Mismatch repair analysis of inherited MSH2 and/or MSH6 variation pairs found in cancer patients. *Human Mutation*, *33*(8), 1294–1301. https://doi.org/10.1002/humu.22119

Kantelinen, J., Kansikas, M., Korhonen, M. K., Ollila, S., Heinimann, K., Kariola, R., & Nyström, M. (2010). MutSβ exceeds MutSα in dinucleotide loop repair. *British Journal of Cancer*, *102*(6), 1068–1073. https://doi.org/10.1038/sj.bjc.6605531

Kelkar, Y. D., Strubczewski, N., Hile, S. E., Chiaromonte, F., Eckert, K. A., & Makova, K. D. (2010). What is a microsatellite: a computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC repeats. Genome biology and evolution, 2, 620-635.

Kocaturk, N. M., & Gozuacik, D. (2018). Crosstalk Between Mammalian Autophagy and the Ubiquitin-Proteasome System. *Frontiers in Cell and Developmental Biology*, *6*. https://doi.org/10.3389/fcell.2018.00128

Kurzawski, G., Suchy, J., Debniak, T., Kladny, J., & Lubinski, J. (2004). Importance of microsatellite instability (MSI) in colorectal cancer: MSI as a diagnostic tool. Ann Oncol, 15(Suppl IV), 283-284.

Lahue, R. S., Au, K. G., & Modrich, P. (1989). DNA Mismatch Correction in a Defined System. *Science*, *245*(4914), 160–164. https://doi.org/10.1126/science.2665076

Li, G. (2008). Mechanisms and functions of DNA mismatch repair. *Cell Research*, *18*(1), 85–98. https://doi.org/10.1038/cr.2007.115

Lomba, A. L. O., Tirapelle, M. C., Biaggio, R. T., Abreu-Neto, M. S., Covas, D. T., Picanço-Castro, V., Swiech, K., & Mizukami, A. (2021). Serum-Free Suspension Adaptation of HEK-293T Cells: Basis for Large-Scale

Biopharmaceutical Production. *Brazilian Archives of Biology and Technology*, *64*. https://doi.org/10.1590/1678-4324-2021200817

Lützen, A., De Wind, N., Georgijevic, D., Nielsen, F. C., & Rasmussen, L. J. (2008). Functional analysis of HNPCC-related missense mutations in MSH2. *Mutation Research: Fundamental and Molecular Mechanisms of Mutagenesis*, *645*(1–2), 44–55. https://doi.org/10.1016/j.mrfmmm.2008.08.015

Martin, L. J. (2008). DNA Damage and Repair. *Journal of Neuropathology and Experimental Neurology*, *67*(5), 377–387. https://doi.org/10.1097/nen.0b013e31816ff780

Martin-Lopez, J., & Fishel, R. (2013). The mechanism of mismatch repair and the functional analysis of mismatch repair defects in Lynch syndrome. *Familial Cancer*, *12*(2), 159–168. https://doi.org/10.1007/s10689-013-9635-x

Mendillo, M. L., Hargreaves, V., Jamison, J. W., Mo, A. O., Li, S., Putnam, C. D., Woods, V. L., & Kolodner, R. D. (2009). A conserved MutS homolog connector domain interface interacts with MutL homologs. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22223–22228. https://doi.org/10.1073/pnas.0912250106

Modrich, P. (1996). Mismatch Repair in Replication Fidelity, Genetic Recombination, and Cancer Biology. *Annual Review of Biochemistry*, *65*(1), 101–133. https://doi.org/10.1146/annurev.biochem.65.1.101

Pećina-Šlaus, N., Kafka, A., Salamon, I., & Bukovac, A. (2020). Mismatch Repair Pathway, Genome Stability and Cancer. *Frontiers in Molecular Biosciences*, *7*. https://doi.org/10.3389/fmolb.2020.00122

Plon, S. E., Eccles, D. M., Easton, D., Foulkes, W. D., Genuardi, M., Greenblatt, M. S., Hogervorst, F. B. L., Hoogerbrugge, N., Spurdle, A. B., & Tavtigian, S. V. (2008). Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Human Mutation, 29(11), 1282–1291. https://doi.org/10.1002/humu.20880

Rahner, N., & Steinke, V. (2008). Hereditary cancer syndromes. Deutsches Ärzteblatt International, 105(41), 706

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine, 17(5), 405–423. https://doi.org/10.1038/gim.2015.30

Schneider, E. A., & Seifert, R. (2010). Sf9 cells: A versatile model system to investigate the pharmacological properties of G protein-coupled receptors. *Pharmacology & Therapeutics*, *128*(3), 387–418. https://doi.org/10.1016/j.pharmthera.2010.07.005

Seifi, M., Ghasemi, A., Raeisi, S., & Heidarzadeh, S. (2017). Application of Next generation Sequencing in Clinical Molecular Diagnostics. Brazilian Archives of Biology and Technology, 60.

Shimodaira, H., Yoshioka-Yamashita, A., Kolodner, R. D., & Wang, J. Y. J. (2003). Interaction of mismatch repair protein PMS2 and the p53-related transcription factor p73 in apoptosis response to cisplatin. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(5), 2420–2425. https://doi.org/10.1073/pnas.0438031100

Smith, C. A. (2007). Cloning and mutagenesis: tinkering with the order of things. *Nature Methods*, *4*(5), 455–461. https://doi.org/10.1038/nmeth0507-455

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209–249. https://doi.org/10.3322/caac.21660

Tarjanyi, O., Haerer, J. +., Vecsernyés, M., Berta, G., Stayer-Harci, A., Balogh, B., Farkas, N., Boldizsár, F., Szeberényi, J., & Sétáló, G. (2022). Prolonged treatment with the proteasome inhibitor MG-132 induces apoptosis in PC12 rat pheochromocytoma cells. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-09763-z

Tournier, I., Vezain, M., Martins, A., Charbonnier, F., Baert-Desurmont, S., Olschwang, S., Wang, Q., Buisine, M., Soret, J., Tazi, J., Frebourg, T., & Tosi, M. (2008). A large fraction of unclassified variants of the mismatch repair genes*MLH1*and*MSH2*is associated with splicing defects. *Human Mutation*, *29*(12), 1412–1424. https://doi.org/10.1002/humu.20796

Warren, J. L., Pohlhaus, T. J., Changela, A., Iyer, R., Modrich, P., & Beese, L. S. (2007). Structure of the Human MutSα DNA Lesion Recognition Complex. *Molecular Cell*, *26*(4), 579–592. https://doi.org/10.1016/j.molcel.2007.04.018

Wielders, E. a. L., Dekker, R., Holt, I., Morris, G. E., & Riele, H. T. (2011). Characterization of *MSH2* variants by endogenous gene modification in mouse embryonic stem cells. *Human Mutation*, *32*(4), 389–396. https://doi.org/10.1002/humu.21448

Xiong, S., Zhang, L., & He, Q. (2008). Fractionation of Proteins by Heparin Chromatography. In *Methods in molecular biology* (pp. 213–221). Springer Science+Business Media. https://doi.org/10.1007/978-1-60327-064-9_18

Yanik, B., & Bakar-Ates, F. (2023). The Effects of Orcinol on Proliferation and Apoptosis of SW480 Human Colorectal Cancer Cells. *Ankara Üniversitesi Eczacılık Fakültesi Dergisi*. https://doi.org/10.33483/jfpau.1228071

APPENDICES

APPENDIX A

Equipment and chemicals used in the study were shown in Table A.1 and Table A.2

Table A.1: List of equipment that used in the study.

Equipment	Brand
NanoDrop™ 2000 Spectrophotometer	Thermo Fisher
Autoclave	Zealway
Balance	Precisa
Balance	Precisa
Centrifuge	VWR
Centrifuge	Thermo Fisher
Deepfreeze (-20ºC)	Bosch
Electronic Pipet Controller	Santa Cruz
Electrophoresis Power Supply	GE Healthcare
Electrophoresis Power Supply	Bio-Rad
ÄKTA avant system	GE Healthcare
Freezer (-80°C)	Aucma
HisTrap FF Crude	Cytiva
HiTrap Q XL column	Cytiva
Ice Flaker	Scotsman
Imaging System	Bio-Rad
Laminar Flow	Faster
Magnetic Stirrer	DragonLab
Microcentrifuge	ScanSpeed
Micropipette (0.2µl-2µl)	Thermo Fisher
Micropipette (2µl-20µl)	Thermo Fisher
Micropipette (20µl-200µl)	Thermo Fisher
Micropipette(100µl-1000µl)	Thermo Fisher
Mini Centrifuge	Inovia
pH Meter	Mettler Toledo
Platform Shaker	Heidolph
Refrigerator (+4ºC)	Bosch
Rotator Mixer	FinePCR
Semi-Dry Blotting System	Bio-Rad
Shaker	Sartorius
Spectrophotometer	Schimadzu
Thermal Cycler	Bio-Rad
Ultrapure Water System	Merck
Vacuum Concentrator	Eppendorf
Vortex	DragonLab

Table A.2: List of chemicals used in the study.

Chemical	Brand
Dynabeads™ Protein G for Immunoprecipitation	Invitrogen
Acetic Acid	Sigma-Aldrich
Acrylamide/Bis-Acrylamide (40%)	Bio-Rad
Ammonium Persulfate (APS)	Sigma-Aldrich
Bacteriological Peptone	WISENT
Blotto, Non-Fat Dry Milk	Santa Cruz
Bovine Serum Albumin (BSA)	Sigma-Aldrich
Bradford Reagent	Bio-Rad
Bromophenol Blue	Sigma-Aldrich
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Roche
Coomassie Brilliant Blue R-250	Fluka
Dithiothreitol (DTT)	Fermentas
DMEM (1X)	Thermo Scientific
DMSO	Biomatik
dNTP	Genaxxon
Gel Loading Dye, Purple (6X)	NEB
Glycerol	MultiCell
Glycine	Chem Cruz
HisPur™ Ni-NTA Resin	Thermo Scientific
Hydrochloric Acid (HCl)	Sigma-Aldrich
Isopropanol	Merck
Methanol	Merck
Nitrocellulose Membrane (0.45µm)	Bio-Rad
Oligo dT	Genaxxon
PBS	Thermo Scientific
Phosphate-Buffered Saline (PBS)	Thermo Scientific
PMSF	Genaxxon
Prestained Protein Ladder (PageRuler)	Thermo Scientific
Random Hexamer	Genaxxon
RPMI 1640 Medium	Gibco
Sodium Chloride (NaCl)	Chem Cruz
Sodium Dodecyl Sulfate (SDS)	Sigma-Aldrich
Tetramethylethylenediamine (TEMED)	Sigma-Aldrich
Tris Base	MultiCell
Trpsin-EDTA	Thermo Scientific
Tween-20	BioShop
Urea	Sigma-Aldrich
Yeast Extract	Pronadisa
methoxy PEG propionaldehyde	

APPENDIX B

Molecular kits, enzyme or enzyme mixes and antibodies were illustrated in the following tables; Table A.3, Table A.4, Table A.5.

Table A.3: Molecular kits that were used in the study.

Equipment	Brand
Clarity Western ECL Substrate	Bio-Rad
In-Fusion Cloning Kit	Takara Bio
NucleoSpin PCR clean-up, gel extraction kit	Mackerey-Nagel
NucleoSpin Plasmid, Midi kit for plasmid DNA	Mackerey-Nagel
NucleoSpin Plasmid, Mini kit for plasmid DNA	Mackerey-Nagel
Total RNA Purification Mini Spin Kit	Genaxxon

Table A.4: List of enzyme/ enzyme mixes used in the study.

Enzyme	Brand
OneTaq® 2X Master Mix with Standard Buffer	NEB
10X CutSmart Buffer	NEB
10X T4 DNA Ligase Buffer	NEB
BamHI-HF® Restriction Enzyme	NEB
CloneAmp™ HiFi PCR Premix	TakaraBio
NdeI® Restriction Enzyme	NEB
Q5® High-Fidelity 2X Master Mix	NEB
T4 DNA Ligase	NEB
Vita 5X Reaction Buffer	GDD Biolabs
Vita MML-V Reverse Transcriptase	GDD Biolabs
Vita RNase Inhibitor	GDD Biolabs
Vita Taq DNA Polymerase	GDD Biolabs
XbaI® Restriction Enzyme	NEB
XhoI® Restriction Enzyme	NEB

Table A.5: List of antibodies and their dilution rates that were used in the study.

Antibodies	Brand and Model	Dilution in 5% BSA
PMS2 Antibody (B-3)	SantaCruz	1:100
GAPDH (14C10)	CST	1:1000
Anti-His Antibody	Sigma Aldrich	1:1000-1:2000
Anti-Mouse IgG (7076S)	CST	1:3000
Anti-Rabbit IgG (7074S)	CST	1:3000
APC Antibody (ALi 12-28)	SantaCruz	1:1000
ATM Antibody (G-12)	SantaCruz	1:1000
Beta-Actin (4970S)	CST	1:1000
FLAG M2 Antibody	Sigma Aldrich	1:1000
GTBP Antibody (E-8)	SantaCruz	1:1000
MLH1 Antibody (B-12)	SantaCruz	1:100
MSH2 Antibody (D-6)	SantaCruz	1:1000
PCNA Antibody (F-2)	SantaCruz	1:1000

APPENDIX C

Common buffer/solutions were listed in Table A.6, while purification or lysis buffers were listed in Table A.7

Table A.6: List of common buffer/solutions and their content or preparations.

Common Buffer/Solutions	Content/Preparation
10% SDS	10 grams of SDS diluted into 100 mL dH_2O
Tris-HCl (1.5 M) (pH:6.8 or 8.8)	181,65 grams of Tris base dissolved in 1L dH_2O, arrage pH by adding 10 M HCl
10% w/v APS	1 gram ammonium persulphate dissolved in 10 mL dH_2O
10X Running Buffer (pH:8.3)	30 grams of Tris base, 144 grams of Glycine and 10 grams of SDS dissolved in 1L dH_2O
4X Laemmlie dye	2,4 mL 1 M Tris (pH:6.8), 4 mL glycerol, 0,8-gram SDS, 4 mg bromophenol blue, 0,5 mL Betamercaptoehtanol (BME), 3,1 mL dH_2O
Coomasie stain	0.1% Coomasie R250 (1 gr), 100 mL Acetic acid, 400 mL Methanol, 500 mL dH_2O
Destaining solution	100 mL Acetic acid, 400 mL Methanol, 500 mL dH_2O
10X TBS (pH:7.6)	87,66 grams of NaCl, 24,22 grams of Tris base dissolved in 1L dH_2O, arrage pH by adding 10M HCl
1X TBS-T	100 mL 10X TBS, 900 mL dH_2O, 1 mL Tween-20
10X Transfer Buffer	30 grams of Tris base, 144 grams of Glycine dissolved in 1L dH_2O
1X Transfer Buffer	100 mL 10X Transfer buffer, 200 mL methanol, 700 mL dH_2O
5% Blocking solution	1,5 grams of milk powder dissolved in 30 mL filtered 1X TBS-T
5% BSA	1,5 grams of BSA dissolved in 30 mL filtered 1X TBS-T
Primary antibody	3 mL 5% BSA, 6 µL primary antibody, 0.02% sodium azide (60 µL)
Secondary antibody	30 mL 5% blocking solution, 10 µL secondary antibody, 0.02% sodium azide (60 µL)
8% separating gel	5,35 mL dH_2O, 2 mL 40% bis-acrylamide, 2,5 mL Tris-HCl (pH:8.8), 100 µL 10% SDS, 100 µL 10% APS, 10 µL TEMED
10% separating gel	4,79 mL dH_2O, 2,5 mL 40% bis-acrylamide, 2,5 mL Tris-HCl (pH:8.8), 100 µL 10% SDS, 100 µL 10% APS, 10 µL TEMED
5% stacking gel	1,40 mL dH2O, 0,27 mL 40% bis-acrylamide, 0,25 mL Tris-HCl (pH:6.8), 20 µL 10% SDS, 20 µL 10% APS, 2 µL TEMED

Table A.7: List of Lysis/Purification buffers used in the study and their content.

Lysis/Purification Buffers	Content
Mammalian Expression	
RIPA	10mM Tris-HCl, pH 8.0, 1mM EDTA, 0.5mM EGTA, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS, 140mM NaCl
Pull-down assay wash	150 mM NaCl, 50 mM Tris-HCl, 20 mM Imidazole, pH:8.0
Pull-down assay elution	150 mM NaCl, 50 mM Tris-HCl, 250 mM Imidazole, pH:8.0
Immunoprecipitation wash	1X TBS
Immunoprecipitation elution	100 mM Glycine pH:2.8
IMAC binding	50 mM Tris-HCl, pH:8.0, 5% glycerol, 150 mM NaCl, 0.1% Triton-100X, 0.05% NP-40, 5 mM $MgCl_2$
IMAC wash	50 mM Tris-HCl, pH:8.0, 5% glycerol, 150 mM NaCl, 0.1% Triton-100X, 0.05% NP-40, 5 mM $MgCl_2$, 20 mM Imidazole
IMAC elution	50 mM Tris-HCl, pH:8.0, 5% glycerol, 150 mM NaCl, 0.1% Triton-100X, 0.05% NP-40, 5 mM $MgCl_2$, 250 mM Imidazole
AEX binding/wash	25 mM Tris-HCl, pH: 8.0, 5% glycerol, 5 mM $MgCl_2$
AEX elution	25 mM Tris-HCl, pH: 8.0, 5% glycerol, 5 mM $MgCl_2$, 1 M NaCl
Storage	50 mM Tris-HCl, 150 mM NaCl, pH:8.0, 10% glycerol, 0.05% NP-40, 1 mM DTT, 5 mM $MgCl_2$
Bacterial Expression	
Bacterial Lysis	25 mM Tris-HCl, pH:8.0, 1 mM EDTA, 5% glycerol, 300 mM NaCl, 1 mM PMSF
Pellet dissolvation	8 M Urea
Bacterial Lysis for purification	25 mM HEPES-KOH, pH:7.5, 10% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT (fresh), 1 mM PMSF (fresh), after lyse adjust to 150 mM KCl
AEX Binding/Wash	25 mM HEPES-KOH, pH:7.5, 10% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA, 150 mM KCl
AEX Elution	25 mM HEPES-KOH, pH:7.5, 10% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 M KCl
Buffer Exchange	25 mM HEPES-KOH, pH:7.5, 10% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA
Heparin Binding/Wash	25 mM HEPES-KOH, pH:7.5, 10% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA, 150 mM KCl
Heparin Elution	25 mM HEPES-KOH, pH:7.5, 10% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 M KCl
Storage	25 mM HEPES-KOH, 150 mM KCl, pH:7.5, 10% glycerol, 1 mM DTT, 5 mM $MgCl_2$

APPENDIX D

Target amino acid sequence of modified hMSH6 with additional C-terminal FLAG tag

MSRQSTLYSFFPKSPALSDANKASARASREGGRAAAAPGASPSPGGDAAWSEAGPGPRPLAR
SASPPKAKNLNGGLRRSVAPAAPTSCDFSPGDLVWAKMEGYPWWPCLVYNHPFDGTFIREK
GKSVRVHVQFFDDSPTRGWVSKRLLKPYTGSKSKEAQKGGHFYSAKPEILRAMQRADEALN
KDKIKRLELAVCDEPSEPEEEEEMEVGTTYVTDKSEEDNEIESEEEVQPKTQGSRRSSRQIKKR
RVISDSESDIGGSDVEFKPDTKEEGSSDEISSGVGDSESEGLNSPVKVARKRKRMVTGNGSLK
RKSSRKETPSATKQATSISSETKNTLRAFSAPQNSESQAHVSGGGDDSSRPTVWYHETLEWLK
EEKRRDEHRRRPDHPDFDASTLYVPEDFLNSCTPGMRKWWQIKSQNFDLVICYKVGKFYEL
YHMDALIGVSELGLVFMKGNWAHSGFPEIAFGRYSDSLVQKGYKVARVEQTETPEMMEAR
CRKMAHISKYDRVVRREICRIITKGTQTYSVLEGDPSENYSKYLLSLKEKEEDSSGHTRAYGV
CFVDTSLGKFFIGQFSDDRHCSRFRTLVAHYPPVQVLFEKGNLSKETKTILKSSLSCSLQEGLIP
GSQFWDASKTLRTLLEEEYFREKLSDGIGVMLPQVLKGMTSESDSIGLTPGEKSELALSALGG
CVFYLKKCLIDQELLSMANFEEYIPLDSDTVSTTRSGAIFTKAYQRMVLDAVTLNNLEIFLNG
TNGSTEGTLLERVDTCHTPFGKRLLKQWLCAPLCNHYAINDRLDAIEDLMVVPDKISEVVEL
LKKLPDLERLLSKIHNVGSPLKSQNHPDSRAIMYEETTYSKKKIIDFLSALEGFKVMCKIIGIME
EVADGFKSKILKQVISLQTKNPEGRFPDLTVELNRWDTAFDHEKARKTGLITPKAGFDSDYD
QALADIRENEQSLLEYLEKQRNRIGCRTIVYWGIGRNRYQLEIPENFTTRNLPEEYELKSTKKG
CKRYWTKTIEKKLANLINAEERRDVSLKDCMRRLFYNFDKNYKDWQSAVECIAVLDVLLCL
ANYSRGGDGPMCRPVILLPEDTPPFLELKGSRHPCITKTFFGDDFIPNDILIGCEEEEQENGKAY
CVLVTGPNMGGKSTLMRQAGLLAVMAQMGCYVPAEVCRLTPIDRVFTRLGASDRIMSGEST
FFVELSETASILMHATAHSLVLVDELGRGTATFDGTAIANAVVKELAETIKCRTLFSTHYHSL
VEDYSQNVAVRLGHMACMVENECEDPSQETITFLYKFIKGACPKSYGFNAARLANLPEEVIQ
KGHRKAREFEKMNQSLRLFREVCLASERSTVDAEAVHKLLTLIKELDYKDDDDK*

Target coding sequence of modified hMSH6 with additional C-terminal FLAG tag

ATGTCGCGACAGAGCACCCTGTACAGCTTCTTCCCCAAGTCTCCGGCGCTGAGTGATGCC
AACAAGGCCTCGGCCAGGGCCTCACGCGAAGGCGGCCGTGCCGCCGCTGCCCCCGGGGC
CTCTCCTTCCCCAGGCGGGGATGCGGCCTGGAGCGAGGCTGGGCCTGGGCCCAGGCCCTT
GGCGCGATCCGCGTCACCGCCCAAGGCGAAGAACCTCAACGGAGGGCTGCGGAGATCGG
TAGCGCCTGCTGCCCCCACCAGTTGTGACTTCTCACCGGGAGATTTGGTTTGGGCCAAGA
TGGAGGGTTACCCCTGGTGGCCTTGTCTGGTTTACAACCACCCCTTTGATGGAACATTCAT
CCGCGAGAAAGGGAAATCAGTCCGTGTTCATGTACAGTTTTTTGATGACAGCCCAACAAG
GGGCTGGGTTAGCAAAAGGCTTTTAAAGCCATATACAGGTTCAAAATCAAAGGAAGCCC
AGAAGGGAGGTCATTTTTACAGTGCAAAGCCTGAAATACTGAGAGCAATGCAACGTGCA
GACGAAGCCTTAAATAAAGACAAGATTAAGAGGCTTGAATTGGCAGTTTGTGATGAGCC
CTCAGAGCCAGAAGAGGAAGAAGAGATGGAGGTAGGCACAACTTACGTAACAGATAAG

AGTGAAGAAGATAATGAAATTGAGAGTGAAGAGGAAGTACAGCCTAAGACACAAGGAT
CTAGGCGAAGTAGCCGCCAAATAAAAAAACGAAGGGTCATATCAGATTCTGAGAGTGAC
ATTGGTGGCTCTGATGTGGAATTTAAGCCAGACACTAAGGAGGAAGGAAGCAGTGATGA
AATAAGCAGTGGAGTGGGGGATAGTGAGAGTGAAGGCCTGAACAGCCCTGTCAAAGTTG
CTCGAAAGCGGAAGAGAATGGTGACTGGAAATGGCTCTCTTAAAAGGAAAAGCTCTAGG
AAGGAAACGCCCTCAGCCACCAAACAAGCAACTAGCATTTCATCAGAAACCAAGAATAC
TTTGAGAGCTTTCTCTGCCCCTCAAAATTCTGAATCCCAAGCCCACGTTAGTGGAGGTGG
TGATGACAGTAGTCGCCCTACTGTTTGGTATCATGAAACTTTAGAATGGCTTAAGGAGGA
AAAGAGAAGAGATGAGCACAGGAGGAGGCCTGATCACCCCGATTTTGATGCATCTACAC
TCTATGTGCCTGAGGATTTCCTCAATTCTTGTACTCCTGGGATGAGGAAGTGGTGGCAGA
TTAAGTCTCAGAACTTTGATCTTGTCATCTGTTACAAGGTGGGGAAATTTTATGAGCTGTA
CCACATGGATGCTCTTATTGGAGTCAGTGAACTGGGGCTGGTATTCATGAAAGGCAACTG
GGCCCATTCTGGCTTTCCTGAAATTGCATTTGGCCGTTATTCAGATTCCCTGGTGCAGAAG
GGCTATAAAGTAGCACGAGTGGAACAGACTGAGACTCCAGAAATGATGGAGGCACGATG
TAGAAAGATGGCACATATATCCAAGTATGATAGAGTGGTGAGGAGGGAGATCTGTAGGA
TCATTACCAAGGGTACACAGACTTACAGTGTGCTGGAAGGTGATCCCTCTGAGAACTACA
GTAAGTATCTTCTTAGCCTCAAAGAAAAAGAGGAAGATTCTTCTGGCCATACTCGTGCAT
ATGGTGTGTGCTTTGTTGATACTTCACTGGGAAAGTTTTTCATAGGTCAGTTTTCAGATGA
TCGCCATTGTTCGAGATTTAGGACTCTAGTGGCACACTATCCCCCAGTACAAGTTTTATTT
GAAAAAGGAAATCTCTCAAAGGAAACTAAAACAATTCTAAAGAGTTCATTGTCCTGTTCT
CTTCAGGAAGGTCTGATACCCGGCTCCCAGTTTTGGGATGCATCCAAAACTTTGAGAACT
CTCCTTGAGGAAGAATATTTTAGGGAAAAGCTAAGTGATGGCATTGGGGTGATGTTACCC
CAGGTGCTTAAAGGTATGACTTCAGAGTCTGATTCCATTGGGTTGACACCAGGAGAGAA
AAGTGAATTGGCCCTCTCTGCTCTAGGTGGTTGTGTCTTCTACCTCAAAAAATGCCTTATT
GATCAGGAGCTTTTATCAATGGCTAATTTTGAAGAATATATTCCCTTGGATTCTGACACA
GTCAGCACTACAAGATCTGGTGCTATCTTCACCAAAGCCTATCAACGAATGGTGCTAGAT
GCAGTGACATTAAACAACTTGGAGATTTTTCTGAATGGAACAAATGGTTCTACTGAAGGA
ACCCTACTAGAGAGGGTTGATACTTGCCATACTCCTTTTGGTAAGCGGCTCCTAAAGCAA
TGGCTTTGTGCCCCACTCTGTAACCATTATGCTATTAATGATCGTCTAGATGCCATAGAAG
ACCTCATGGTTGTGCCTGACAAAATCTCCGAAGTTGTAGAGCTTCTAAAGAAGCTTCCAG
ATCTTGAGAGGCTACTCAGTAAAATTCATAATGTTGGGTCTCCCCTGAAGAGTCAGAACC
ACCCAGACAGCAGGGCTATAATGTATGAAGAAACTACATACAGCAAGAAGAAGATTATT
GATTTTCTTTCTGCTCTGGAAGGATTCAAAGTAATGTGTAAAATTATAGGGATCATGGAA
GAAGTTGCTGATGGTTTTAAGTCTAAAATCCTTAAGCAGGTCATCTCTCTGCAGACAAAA
AATCCTGAAGGTCGTTTTCCTGATTTGACTGTAGAATTGAACCGATGGGATACAGCCTTT
GACCATGAAAAGGCTCGAAAGACTGGACTTATTACTCCCAAAGCAGGCTTTGACTCTGAT
TATGACCAAGCTCTTGCTGACATAAGAGAAAATGAACAGAGCCTCCTGGAATACCTAGA
GAAACAGCGCAACAGAATTGGCTGTAGGACCATAGTCTATTGGGGGATTGGTAGGAACC
GTTACCAGCTGGAAATTCCTGAGAATTTCACCACTCGCAATTTGCCAGAAGAATACGAGT
TGAAATCTACCAAGAAGGGCTGTAAACGATACTGGACCAAAACTATTGAAAAGAAGTTG

GCTAATCTCATAAATGCTGAAGAACGGAGGGATGTATCATTGAAGGACTGCATGCGGCG
ACTGTTCTATAACTTTGATAAAAATTACAAGGACTGGCAGTCTGCTGTAGAGTGTATCGC
AGTGTTGGATGTTTTACTGTGCCTGGCTAACTATAGTCGAGGGGGTGATGGTCCTATGTG
TCGCCCAGTAATTCTGTTGCCGGAAGATACCCCCCCCTTCTTAGAGCTTAAAGGATCACG
CCATCCTTGCATTACGAAGACTTTTTTTGGAGATGATTTTATTCCTAATGACATTCTAATA
GGCTGTGAGGAAGAGGAGCAGGAAAATGGCAAAGCCTATTGTGTGCTTGTTACTGGACC
AAATATGGGGGGCAAGTCTACGCTTATGAGACAGGCTGGCTTATTAGCTGTAATGGCCCA
GATGGGTTGTTACGTCCCTGCTGAAGTGTGCAGGCTCACACCAATTGATAGAGTGTTTAC
TAGACTTGGTGCCTCAGACAGAATAATGTCAGGTGAAAGTACATTTTTTGTTGAATTAAG
TGAAACTGCCAGCATACTCATGCATGCAACAGCACATTCTCTGGTGCTTGTGGATGAATT
AGGAAGAGGTACTGCAACATTTGATGGGACGGCAATAGCAAATGCAGTTGTTAAAGAAC
TTGCTGAGACTATAAAATGTCGTACATTATTTTCAACTCACTACCATTCATTAGTAGAAG
ATTATTCTCAAAATGTTGCTGTGCGCCTAGGACATATGGCATGCATGGTAGAAAATGAAT
GTGAAGACCCCAGCCAGGAGACTATTACGTTCCTCTATAAATTCATTAAGGGAGCTTGTC
CTAAAAGCTATGGCTTTAATGCAGCAAGGCTTGCTAATCTCCCAGAGGAAGTTATTCAAA
AGGGACATAGAAAAGCAAGAGAATTTGAGAAGATGAATCAGTCACTACGATTATTTCGG
GAAGTTTGCCTGGCTAGTGAAAGGTCAACTGTAGATGCTGAAGCTGTCCATAAATTGCTG
ACTTTGATTAAGGAATTAGATTACAAGGATGACGATGACAAGTGA

Target amino acid sequence of modified hMSH2 with additional C-terminal 6-Histidine tag

MAVQPKETLQLESAAEVGFVRFFQGMPEKPTTTVRLFDRGDFYTAHGEDALLAAREVFKTQ
GVIKYMGPAGAKNLQSVVLSKMNFESFVKDLLLVRQYRVEVYKNRAGNKASKENDWYLY
KASPGNLSQFEDILFGNNDMSASIGVVGVKMSAVDGQRQVGVGYVDSIQRKLGLCEFPDND
QFSNLEALLIQIGPKECVLPGGETAGDMGKLRQIIQRGGILITERKKADFSTKDIYQDLNRLLK
GKKGEQMNSAVLPEMENQVAVSSLSAVIKFLELLSDDSNFGQFELTTFDFSQYMKLDIAAVR
ALNLFQGSVEDTTGSQSLAALLNKCKTPQGQRLVNQWIKQPLMDKNRIEERLNLVEAFVEAE
LRQTLQEDLLRRFPDLNRLAKKFQRQAANLQDCYRLYQGINQLPNVIQALEKHEGKHQKLLL
AVFVTPLTDLRSDFSKFQEMIETTLDMDQVENHEFLVKPSFDPNLSELREIMNDLEKKMQSTL
ISAARDLGLDPGKQIKLDSSAQFGYYFRVTCKEEKVLRNNKNFSTVDIQKNGVKFTNSKLTSL
NEEYTKNKTEYEEAQDAIVKEIVNISSGYVEPMQTLNDVLAQLDAVVSFAHVSNGAPVPYVR
PAILEKGQGRIILKASRHACVEVQDEIAFIPNDVYFEKDKQMFHIITGPNMGGKSTYIRQTGVI
VLMAQIGCFVPCESAEVSIVDCILARVGAGDSQLKGVSTFMAEMLETASILRSATKDSLIIIDE
LGRGTSTYDGFGLAWAISEYIATKIGAFCMFATHFHELTALANQIPTVNNLHVTALTTEETLT
MLYQVKKGVCDQSFGIHVAELANFPKHVIECAKQKALELEEFQYIGESQGYDIMEPAAKKCY
LEREQGEKIIQEFLSKVKQMPFTEMSEENITIKLKQLKAEVIAKNNSFVNEIISRIKVTTHHHHH
H*

Target coding sequence of modified hMSH2 with additional C-terminal 6-Histidine tag

ATGGCGGTGCAGCCGAAGGAGACGCTGCAGTTGGAGAGCGCGGCCGAGGTCGGCTTCGT
GCGCTTCTTTCAGGGCATGCCGGAGAAGCCGACCACCACAGTGCGCCTTTTCGACCGGGG
CGACTTCTATACGGCGCACGGCGAGGACGCGCTGCTGGCCGCCCGGGAGGTGTTCAAGA
CCCAGGGGGTGATCAAGTACATGGGGCCGGCAGGAGCAAAGAATCTGCAGAGTGTTGTG
CTTAGTAAAATGAATTTTGAATCTTTTGTAAAAGATCTTCTTCTGGTTCGTCAGTATAGAG
TTGAAGTTTATAAGAATAGAGCTGGAAATAAGGCATCCAAGGAGAATGATTGGTATTTG
GCATATAAGGCTTCTCCTGGCAATCTCTCTCAGTTTGAAGACATTCTCTTTGGTAACAATG
ATATGTCAGCTTCCATTGGTGTTGTGGGTGTTAAAATGTCCGCAGTTGATGGCCAGAGAC
AGGTTGGAGTTGGGTATGTGGATTCCATACAGAGGAAACTAGGACTGTGTGAATTCCCTG
ATAATGATCAGTTCTCCAATCTTGAGGCTCTCCTCATCCAGATTGGACCAAAGGAATGTG
TTTTACCCGGAGGAGAGACTGCTGGAGACATGGGGAAACTGAGACAGATAATTCAAAGA
GGAGGAATTCTGATCACAGAAAGAAAAAAAGCTGACTTTTCCACAAAAGACATTTATCA
GGACCTCAACCGGTTGTTGAAAGGCAAAAAGGGAGAGCAGATGAATAGTGCTGTATTGC
CAGAAATGGAGAATCAGGTTGCAGTTTCATCACTGTCTGCGGTAATCAAGTTTTTAGAAC
TCTTATCAGATGATTCCAACTTTGGACAGTTTGAACTGACTACTTTTGACTTCAGCCAGTA
TATGAAATTGGATATTGCAGCAGTCAGAGCCCTTAACCTTTTTCAGGGTTCTGTTGAAGA
TACCACTGGCTCTCAGTCTCTGGCTGCCTTGCTGAATAAGTGTAAAACCCCTCAAGGACA
AAGACTTGTTAACCAGTGGATTAAGCAGCCTCTCATGGATAAGAACAGAATAGAGGAGA
GATTGAATTTAGTGGAAGCTTTTGTAGAAGATGCAGAATTGAGGCAGACTTTACAAGAA
GATTTACTTCGTCGATTCCCAGATCTTAACCGACTTGCCAAGAAGTTTCAAAGACAAGCA
GCAAACTTACAAGATTGTTACCGACTCTATCAGGGTATAAATCAACTACCTAATGTTATA
CAGGCTCTGGAAAAACATGAAGGAAAACACCAGAAATTATTGTTGGCAGTTTTTGTGACT
CCTCTTACTGATCTTCGTTCTGACTTCTCCAAGTTTCAGGAAATGATAGAAACAACTTTAG
ATATGGATCAGGTGGAAAACCATGAATTCCTTGTAAAACCTTCATTTGATCCTAATCTCA
GTGAATTAAGAGAAATAATGAATGACTTGGAAAAGAAGATGCAGTCAACATTAATAAGT
GCAGCCAGAGATCTTGGCTTGGACCCTGGCAAACAGATTAAACTGGATTCCAGTGCACA
GTTTGGATATTACTTTCGTGTAACCTGTAAGGAAGAAAAAGTCCTTCGTAACAATAAAAA
CTTTAGTACTGTAGATATCCAGAAGAATGGTGTTAAATTTACCAACAGCAAATTGACTTC
TTTAAATGAAGAGTATACCAAAAATAAAACAGAATATGAAGAAGCCCAGGATGCCATTG
TTAAAGAAATTGTCAATATTTCTTCAGGCTATGTAGAACCAATGCAGACACTCAATGATG
TGTTAGCTCAGCTAGATGCTGTTGTCAGCTTTGCTCACGTGTCAAATGGAGCACCTGTTCC
ATATGTACGACCAGCCATTTTGGAGAAAGGACAAGGAAGAATTATATTAAAAGCATCCA
GGCATGCTTGTGTTGAAGTTCAAGATGAAATTGCATTTATTCCTAATGACGTATACTTTGA
AAAAGATAAACAGATGTTCCACATCATTACTGGCCCCAATATGGGAGGTAAATCAACAT
ATATTCGACAAACTGGGGTGATAGTACTCATGGCCCAAATTGGGTGTTTTGTGCCATGTG
AGTCAGCAGAAGTGTCCATTGTGGACTGCATCTTAGCCCGAGTAGGGGCTGGTGACAGTC
AATTGAAAGGAGTCTCCACGTTCATGGCTGAAATGTTGGAAACTGCTTCTATCCTCAGGT
CTGCAACCAAAGATTCATTAATAATCATAGATGAATTGGGAAGAGGAACTTCTACCTACG
ATGGATTTGGGTTAGCATGGGCTATATCAGAATACATTGCAACAAAGATTGGTGCTTTTT
GCATGTTTGCAACCCATTTTCATGAACTTACTGCCTTGGCCAATCAGATACCAACTGTTAA

TAATCTACATGTCACAGCACTCACCACTGAAGAGACCTTAACTATGCTTTATCAGGTGAA
GAAAGGTGTCTGTGATCAAAGTTTTGGGATTCATGTTGCAGAGCTTGCTAATTTCCCTAA
GCATGTAATAGAGTGTGCTAAACAGAAAGCCCTGGAACTTGAGGAGTTTCAGTATATTG
GAGAATCGCAAGGATATGATATCATGGAACCAGCAGCAAAGAAGTGCTATCTGGAAAGA
GAGCAAGGTGAAAAAATTATTCAGGAGTTCCTGTCCAAGGTGAAACAAATGCCCTTTACT
GAAATGTCAGAAGAAAACATCACAATAAAGTTAAAACAGCTAAAAGCTGAAGTAATAGC
AAAGAATAATAGCTTTGTAAATGAAATCATTTCACGAATAAAAGTTACTACGCACCACCA
CCACCACCACTG

CURRICULUM VITAE

Name Surname : Celil Mert Gül

EDUCATION :

- **B.Sc. :** 2021, Istanbul Technical University, Science and Letters, Molecular Biology and Genetics

PROFESSIONAL EXPERIENCE AND REWARDS:

- **2021-2023, TUBITAK, BIDEB 2210E**
- **2022-2023, TUBITAK 1001: 119Z261**

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Gül C.M., Can N.D., Çebi E., Kizilboga T., Doğanay H.L., Dinler Doğanay G. 2023:** Functional analysis of Variants of Unknown Significance (VUS) in MSH2 and MSH6 identified by Next-generation Sequencing of colorectal cancer patients in the Turkish population - International Graduate Research Symposium, 16-18 May, 2023 Istanbul, Turkey.