Genetic Spectrum and Metabolic Profiling of Congenital Hypothyroid Children in Hospital Care Settings in Bangladesh

PhD Thesis

Mst. Noorjahan Begum

Registration Number: 38/2016-17

DEPARTMENT OF GENETIC ENGINEERING AND BIOTECHNOLOGY
FACULTY OF BIOLOGICAL SCIENCES
UNIVERSITY OF DHAKA
DHAKA, BANGLADESH

MARCH 2020

A DISSERTATION SUBMITTED TO THE UNIVERSITY OF DHAKA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN GENETIC ENGINEERING AND BIOTECHNOLOGY

CERTIFICATE

This is to certify that Mst. Noorjahan Begum (Registration Number: 38/2016-17) conducted her PhD thesis work entitled **"Genetic Spectrum and Metabolic Profiling of Congenital Hypothyroid Children in Hospital Care Settings in Bangladesh"** under my supervision for the fulfillment of the degree of 'Doctor of Philosophy in Genetic Engineering and Biotechnology' from University of Dhaka, Dhaka-1000, Bangladesh. Furthermore, her works are also co-supervised by Dr. Kaiissar Mannoor, Institute for Developing Science and Health Initiatives, Bangladesh. The thesis work or any part of this has not been submitted anywhere for any other degree.

S A Zaman 01/03/2020
.....

(Supervisor)

Sharif Akhteruzzaman, PhD

Professor, Department of Genetic Engineering and Biotechnology, University of Dhaka

Dhaka, Bangladesh

Mobile: +88-01714111267

Email: sazaman@du.ac.bd

01/03/2020
.....

(Co-supervisor)

Kaiissar Mannoor, PhD

Head & Senior Scientist, Institute for Developing Science & Health Initiatives, Dhaka, Bangladesh

Mobile: +88-01797440713

Email: kaiissar@ideshi.org

DECLARATION

I hereby declare that this thesis entitled **"Genetic Spectrum and Metabolic Profiling of Congenital Hypothyroid Children in Hospital Care Settings in Bangladesh"** contains no material which has been accepted for the award of any other degree or diploma in any university or equivalent institution. To the best of my knowledge, the thesis contains no material previously published or being considered for publication, except where due reference is made in the text of the thesis.

01.03.2020

Mst. Noorjahan Begum
Doctoral Fellow
Registration No. 38/ 2016-17
Department of Genetic Engineering and Biotechnology
University of Dhaka
Bangladesh

DEDICATION

Dedicated to

My Family; Parents, Siblings, Husband &

Daughters “Mersiha Rashid & Zuhayra Rashid”

My Teachers & My Country “Bangladesh”

ACKNOWLEDGEMENTS

First of all, I am thankful to Almightily Allah for His blessings to find out a new way to explore my research works.

I would like to express my deep sense of gratitude and profound respect to my academic supervisor Professor Dr. Sharif Akhteruzzaman, Department of Genetic Engineering and Biotechnology, University of Dhaka, Bangladesh, who has assisted me and supported me in all the stages of my PhD work with great patience and enormous care. His support was a great inspiration for me to balance every ups and downs in my PhD period. I am really proud to work under his supervision utilizing the world class facilities of Molecular Biotechnology Laboratory established at the Department of Genetic Engineering and Biotechnology funded by HEQEP (CP-4029) under University Grants Commission.

I am grateful to my co-supervisor Dr. Kaiissar Mannoor, Head & Senior Scientist, Institute for Developing Science and Health Initiatives (ideSHi) for his valuable and caring guidance from study design to conduct the experiments using the facilities of ideSHi laboratory.
I wish to express my deep gratitude to Dr. Firdausi Qadri, Chairperson, Institute for Developing Science and Health Initiatives (ideSHi) and Senior Scientist, Infectious Disease Division (IDD), International Center for Diarrheal Disease Research, Bangladesh (icddr,b) for her continuous monitoring and inspiration in every steps of my PhD work. I am pleased to express my honor to Professor Syed Saleheen Qadri, Scientific Coordinator of ideSHi for his continuous positive support and valuable suggestions to improve myself as a researcher.

My sincere thanks to Professor Dr. Mohammad Nazmul Ahsan, Chairperson, Department of Genetic Engineering and Biotechnology, University of Dhaka, Bangladesh for supporting me with all research facilities in the department.

I am also grateful to Prof. Dr. Md. Aftab Uddin, Prof. Dr. Md. Anwarul Azim Akhand, Prof. Dr. Jesmin, and Prof. Dr. Sabina Yeasmin, Department of Genetic Engineering and Biotechnology, University of Dhaka for their respected recommendations on my research. I am also thankful to Assoc. Prof. Dr. ABM Md. Khademul Islam and Assoc. Prof. Dr. Abu Ashfaqur Sajib, Department of Genetic Engineering and Biotechnology, University of Dhaka

for their suggestions during my PhD works. I am grateful to Dr. Mustak Ibn Ayub, Dr. S.M. Mahbubur Rashid, Tarana Sharmin, Abira Khan, Ahmed Faisal, Mahmudul Hasan Akash and Wahid Murad Jitu for their support. I am also thankful to Md. Mahamud Hasan for his help and also grateful to Dr. Rokeya Begum, Md. Shawkat Hossain, Mahbub, Mizan and Drishty.

My special thanks to Dr. Mohammad A. Halim, Assistant Professor, Arkansas University, USA, for his constant support to conduct all the computational analysis using the resources of Red-Green Research Center, Bangladesh. I am also grateful to Imrul Shahriar for his support.

A very special thanks to the whole team of ideSHi to support me in works. I express my thanks to Tarikul Islam, Suprovath Sarker and Golam Sarower Bhuyan for their support to conduct the study. I am thankful to Dr. Shahinur Haque, Dr. Tasnia Kawsar Konika, Dr. Nusrat Sultana and Dr. Rosy Sultana for their support in specimen collection. I am thankful to Dr. Md. Sazzadul Islam, Dr. Md. Alauddin and Dr. Polash Karmaker for their suggestions. I am grateful to Asifuzzaman Rahat, Shekh Rezwan, Antora Jahan, Rumana Trisha, Shohel Maruf, Mahmud, Sarjana, Tuba, Jumanah, Ema, Konica, Shagoofa, Shahriar and Aurna for their cooperation.

I would like to thank and acknowledge all the clinicians from Bangabandhu Sheikh Mujib Medical University (BSMMU), specially to Dr. Mizanul Hasan, Professor, National Institute of Nuclear Medicine and Allied Sciences (NINMAS) and Dr. M A Hasanat, Professor, Dept. of Endocrinology (BSMMU) for their support in providing clinical specimens. I am thankful to Dr. Hurjahan Banu, Dr. Suraiya Begum and Dr. Shyamoly for their cooperation in BSMMU.

I am thankful to Prof. Zeba Islam Seraj, Dept. of Biochemistry & Molecular Biology, University of Dhaka and Prof. Mohammad Nasir Uddin, Texas A&M University Health Science Center College of Medicine, USA, for their motivational communication to conduct my work. I also thank Dr. Sabrina Moriom Elias and Dr. Richard Malo for their support.

Finally, I would like to thank my family members (specially My Mother, late Mother-in-law, Siblings, Husband-Md. Mamunur Rashid and Daughters-Mersiha Rashid and Zuhayra Rashid) to support me all the times to accomplish my PhD work.

Sincerely,
Mst. Noorjahan Begum

ABSTRACT

Congenital Hypothyroidism (CH) is the deficiency of thyroid hormones at birth and a major cause of mental retardation if left untreated. The frequency of CH is much higher in Bangladesh than that of the global incidences and late-diagnosis is a common scenario. There is limited data on genetic etiology of Congenital Hypothyroidism in Bangladesh and patients diagnosed at late-stage suffer from clinical complications even under treatment with levothyroxine drug. The present study aimed to investigate the genetic spectrum and metabolic profiling of Congenital Hypothyroid children of Bangladesh in order to correlate the clinical complications of CH with the blood acylcarnitines and amino acids levels.

A total of 65 confirmed cases of Congenital Hypothyroid children under follow-up treatment were enrolled in this study. Among them 44 participants with thyroid Dyshormonogenesis were considered for TPO gene analysis; and 21 cases with thyroid Dysgenesis were studied for TSHR and PAX8 gene. Moreover, 107 healthy subjects were enrolled for metabolic profiling. Molecular analysis was performed using PCR followed by Sanger sequencing of TPO, TSHR and PAX8 genes. The effect of mutations on the 3D structure of the TPO monomer and dimer structures along with TSHR proteins were analyzed by Molecular Docking and Molecular Dynamics Simulations using Autodock vina/PyRx protocol and YASARA computational software. Finally, High Resolution Melt (HRM) curve analysis method was set up to detect the common mutations found in the TPO gene. For metabolic profiling, blood acylcarnitines and amino acids levels of patients were compared with healthy controls using LC-MS/MS technique.

Analysis of TPO gene identified 12 mutations. In almost all patients, 4 mutations belonging to Exon 8 and Exon 12 (hotspot region) were commonly detected. Analysis of the effect of the three non-synonymous mutations c.1117G>T (p.Ala373Ser); c.1193G>C (p.Ser398Thr); and c.2173A>C (p.Thr725Pro) on the 3D structure of TPO protein revealed alterations in the binding affinities of the ligands (Heme, H_2O_2 and I^-) for the wild type cases compared to mutant cases. A HRM (High Resolution Melt) curve analysis method was established to detect these 3 mutations in homozygous, heterozygous and wild type states. Analysis of TSHR gene on the other hand, found a mutation c.1523C>T (p.Ser508Leu) in one patient and 20 patients had

c.2181G>C (p.Asp727Glu) mutation in Exon 10 of the gene. The effect of these two mutations on 3D structure of small molecules drug (MS437 and MS438) binding site was investigated and we found that the binding affinities of the small molecules were decreased in case of mutant TSHR protein compared to wild type TSHR protein. No mutation was detected in PAX8 gene.

Profiling of acylcarnitines showed that the mean concentration of long chain acylcarnitines was significantly lower in the patients group than that of the Healthy Controls (HCs) group (with a $P=0.0014$). Also, the indicator ratio of β-oxidation rate, namely C0/(C16+C18) was significantly higher in the patient group than that of the HC group ($P<0.0001$), suggesting a low activity of CPT-I enzyme in the patients. Furthermore, the mean plasma TG level of the patients was significantly higher than that of the HC group ($P=0.02$) suggesting that the patients might have been suffering from fatty liver. Moreover, in terms of the analyzed amino acids profile, 9 out of 15 amino acids were significantly decreased in the patients compared to the healthy controls ($P<0.0008$). Serum albumin levels were also significantly ($P<0.0001$) lower in the patients group compared to the HCs, suggesting that there might be a decreased rate in overall protein metabolism.

This study is the first of its kind that demonstrated the genetic basis of Congenital Hypothyroidism in Bangladeshi patients. The correlation between mutation and functional activity of specific proteins were also investigated using several *in silico* tools. In conclusion, metabolic profiling combined with mutation analysis might help to better understand the disease severity and to modify the treatment strategy of congenital hypothyroid children in Bangladesh.

Table of Contents

Title **Page**

CERTIFICATE ..... **i**

DECLARATION ..... **ii**

DEDICATION ..... **iii**

ACKNOWLEDGEMENTS ..... **iv**

ABSTRACT ..... **vi**

Table of Contents ..... **viii**

List of Tables ..... **xii**

List of Figures ..... **xiv**

List of Abbreviation ..... **xvi**

CHAPTER 1: INTRODUCTION **1**

1.1. Background **1**

1.1.1. Hypothyroidism 1

1.1.2. Congenital Hypothyroidism (CH) and the Global Incidence 1

1.1.3. Clinical Features 1

1.1.4. Confirmatory Test of Congenital Hypothyroidism 1

1.1.5. Treatment and Prevention 2

1.2. Molecular Genetics of Congenital Hypothyroidism **2**

1.2.1. Major Features of TPO Gene and Mechanism of Action of TPO Protein 3

1.2.2. Effect of Mutation in TPO Gene and the 3D Structure of TPO Protein 4

1.3. Significance of *in silico* Approach to Study 3D Structures of Proteins **5**

1.3.1. Modelling of 3D Structures using Bioinformatics Tools 5

1.3.2. Basic Parameters to Calculate the Energy Level of 3D Structures of Protein-Ligand Complexes (Quantum Mechanical Calculations and Density Functional Theory) 6

1.3.3. Study of Protein Structures using Molecular Docking Approach 6

1.3.4. Molecular Dynamics Simulation 7

1.3.5. *In silico* Approach for the Study of Dimerization of TPO Protein Structures 7

1.3.6. The Sequential Docking Approach for TPO Dimer Protein 8

1.4. High Resolution Melt (HRM) Curve Analysis to Detect Mutations in TPO Gene **8**

1.5. Molecular Investigation of TSHR and PAX8 Gene 9
1.5.1. Structure and Function of TSHR Gene 9
1.5.2. PAX8 (Paired Box Eight) Gene 10
1.6. Metabolic Profiling of Congenital Hypothyroid Children 10
1.6.1. Profiling of Acylcarnitines 10
1.6.2. Profiling of Amino Acids 11
1.7. Rationale of the Research Work 13
1.8. Objectives of the Research Work 14
General objectives 14
Specific objectives 14
CHAPTER 2: MATERIALS AND METHODS 15
2.1. Study Design, Clinical Settings and Ethical Clearance 15
2.1.1. Study Participants (Visits at BSMMU) 16
2.1.2. Healthy Subjects 16
2.1.3. Collection and Processing of Blood Specimens 16
2.2. Molecular Analysis of TPO, TSHR and PAX8 Gene 17
2.2.1. Genomic DNA Isolation to Perform PCR 18
2.2.2. Polymerase Chain Reaction (PCR) Amplification of TPO Gene 18
2.2.3. PCR Amplification of TSHR Gene 20
2.2.4. PCR Amplification of PAX8 Gene 21
2.2.5. Agarose Gel Electrophoresis and Gel Documentation 22
2.2.6. Purification of PCR Products 22
2.2.7. Cycle Sequencing Reaction Followed by Sanger Sequencing 23
2.2.8. Purification of Cycle Sequencing Product 23
2.2.9. Analysis of Sanger Sequencing Data 23
2.3. Study of the Effect of Non-Synonymous Mutations on 3D Structure of TPO Protein 24
2.3.1. Modelling of 3D Structure of TPO protein through *in silico* Approach 24
2.3.2. Validation of the 3D Structures 25
2.3.3. Optimization of Heme using Quantum Mechanical (QM) Calculations 25
2.3.4. Molecular Docking Analysis of Heme with 3D Structures of TPO protein 25
2.3.5. Visualization and Analysis of Molecular Docking Results 26
2.3.6. Validation of Molecular Docking Results using QM/MM and Molecular Dynamics (MD) Simulation 27
2.4. Modelling and MD Simulation of TPO Dimer Structures 28

2.4.1. Analysis of Binding Energies After Sequential Docking Approach 28
2.4.2. Molecular Docking of Hydrogen Peroxide (H_2O_2) and Iodide (I^-) with TPO WT-HD, TPO MT1-HD and TPO MT3-HD Homodimer Structures by PyRx 28
2.4.3. Molecular Dynamics (MD) Simulation After Sequential Docking 28
2.5. Setting Up and Validation of High Resolution Melt (HRM) Curve Analysis 29
2.6. Effect of Mutation on Predicted 3D Structure of Transmembrane Region of TSHR Protein 30
2.7. Profiling of Acylcarnitines and Amino Acids using LC-MS/MS 31
2.7.1. HPLC and MS/MS Settings 31
2.7.2. Validation of Method 32
2.7.3. Preparation of Specimens from Dried Blood Spots 32
2.7.4. Data Collection and Data Processing 32
2.7.5. Measurement of Plasma Triglycerides (TG) 34
2.7.6. Analysis of Gene Expression Level of CPT1 (Carnitine Palmitoyltransferase1) 34
2.7.7. Quantification of Plasma Albumin 34
2.7.8. Statistical Analysis 35
CHAPTER 3: RESULTS 36
3.1. Baseline Information of the Study Participants 36
3.2. Identification of Mutation in TPO Gene 38
3.2.1. Modelling of 3D Structure of Myeloperoxidase (MPO)-Like Domain $TPO_{142-738}$ and Full-Length TPO_{1-933} Protein 40
3.2.2. Validation of the 3D Structures of MPO Like Domain and Full-Length TPO protein 42
3.2.3. Optimization of Heme, Hydrogen Peroxide (H_2O_2) and Iodide (I^-) 43
3.2.4. Comparison of Molecular Docking Results Between Heme and 3D Structures of MPO like domain and full-length TPO protein (Wild Type and Mutant TPO Cases) 43
3.2.5. Validation of Molecular Docking Analysis through QM/MM and Molecular Dynamics (MD) Simulation 49
3.3. Modelling of TPO Dimer Structures (TPO WT-HD, TPO MT1-HD & TPO MT3-HD) and MD Simulation 54
3.3.1. Analysis of Molecular Docking Results (Sequential Approach) 55
3.3.2. Molecular Docking of Hydrogen Peroxide (H_2O_2) and Iodide (I^-) With Predicted 3D Structures of TPO Dimer Proteins (TPOWT-HD, TPOMT1-HD and TPOMT3-HD) 55
3.3.3. MD Simulation of TPO Dimer Structures after Sequential Docking 58
3.4. Setting Up of High Resolution Melt (HRM) Curve Analysis 58
3.4.1. Screening of C.1117G>T Variant 58
3.4.2. Screening of C.1193G>C Variant 60

3.4.3. Screening of C.2173A>C Variant 62
3.5. Investigation of Mutation in TSHR Gene 63
3.5.1. Modelling of the Transmembrane Region of TSHR Protein 64
3.5.2. Molecular Docking of MS437 and MS438 with TM Region of TSHR Proteins (Wild Type and Mutant) 65
3.5.3. Molecular Dynamics Simulation of TSHR Protein Structures 68
3.6. Genetic Study on PAX 8 Gene 69
3.7. Profiling of Acylcarnitines using LC-MS/MS 70
3.7.1. Free Carnitine, Total Carnitine and Total Acylcarnitine Profile 70
3.7.2. Long Chain (LC)-Acylcarnitine Profile 70
3.7.3. Indicator Ratios of Metabolic Process and Measurement of Triglyceride (TG) Level 72
3.7.4. Measurement of CPT1 Gene Expression Levels 74
3.8. Profiling of Amino Acids using LC-MS/MS technique 75
3.8.1. Levels of Blood Amino Acids 75
3.8.2. Metabolic and Nutritional State Indicator Amino Acid Ratios and Plasma Albumin Level 76
CHAPTER 4: DISCUSSION 78
4.1. Analysis of Genetic Spectrum of Congenital Hypothyroid Patients 78
4.2. Metabolic Profiling of Congenital Hypothyroid Children 87
CHAPTER 5: CONCLUSION 91
CHAPTER 6: REFERENCES 92
Supplementary Information 108
Ethical Clearance 124

List of Tables

Table No. and Title **Page**

Table 2. 1. List of primers for PCR amplification and Sanger sequencing of TPO gene 19

Table 2. 2. List of primers for PCR amplification and Sanger sequencing of TSHR gene ... 20

Table 2. 3. List of primers for PCR amplification and Sanger sequencing of PAX8 gene.... 21

Table 2. 4. Grid box center and grid box size for full length TPO1-933 26

Table 2. 5. List of primers used in HRM curve analysis..... 29

Table 3. 1. Baseline information of the hypothyroid patients and healthy participants..... 36

Table 3. 2. Distribution of hypothyroid patients among different types of hypothyroidism . 37

Table 3. 3. Clinical complications manifested by hypothyroid patients. 38

Table 3. 4. Mutation detection in the TPO gene of hypothyroid patients 40

Table 3. 5. Features of 3D structures predicted by I-TASSER server. 41

Table 3. 6. 3D Structure Validation Data Obtained from Verify3D and RAMPAGE 42

Table 3. 7. Comparison of Binding energies and non-bond interactions of Heme with MPO domain ($TPO_{142-738}$) and full length TPO protein (TPO_{1-933}) after flexible docking 46

Table 3. 8. SymmDock server predicted structural comparison for TPO WT HD, TPO MT1 HD, TPO MT3 HD 54

Table 3. 9. Binding energies of H2O2 and iodine (I-) with TPO WT HD, TPO MT1 HD, TPO MT3 HD protein after flexible sequential docking..... 56

Table 3. 10. Mutation detection in the TSHR gene of hypothyroid patients and analysis of the effect of mutation using different bioinformatics tools 64

Table 3. 11. Summary of the corresponding model numbers, C–score, TM–score and the RMSD–score of the predicted 3D structures of TSHR 368-764 WT, TSHR 368-764 MT1 and TSHR 368-764 MT2. 65

Table 3. 12. Binding energy and non-bond interactions of MS437 with TSHR368-764WT, TSHR368-764 MT1 and TSHR368-764 MT2 proteins after flexible docking 67

Table 3.13. Binding energy and non-bond interactions of MS438 with TSHR368-764WT, TSHR368-764 MT1 and TSHR368-764 MT2 proteins after flexible docking 67

Table 3. 14. Comparison of free carnitine, long chain acylcarnitines, total carnitines and total acylcarnitines between the patients and the healthy controls 71

Table 3. 15. Comparison of concentrations of amino acids between hypothyroid patients and healthy controls. 76

Supplementary Tables

Table No. and Title **Page**

Table S 1. Binding energies and non-bond interactions of Heme with TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, TPO142-738 MT3 after flexible docking. 111

Table S 2. Binding energies and non-bond interactions of Heme with TPO1-933WT, TPO1-933MT1, TPO1-933 MT2 and TPO1-933 MT3 after flexible docking. 115

Table S 3. Summarized data for the bond distances and bond angles of specific atoms of H2O2 after optimization. 118

Table S 4. Sequenced samples that were used for HRM method setup 119

Table S 5. Unknown samples used for HRM method validation 121

Table S 6. Comparison of 8 short chain and 7 medium chain acylcarnitines between patients and healthy controls 123

List of Figures

Figure No. and Title **Page**

Figure 1. 1. Chromosomal location of TPO gene 3

Figure 1. 2. Chromosomal location of TSHR gene 9

Figure 1. 3. Chromosomal location of PAX8 gene 10

Figure 2. 1. A Representative flowchart of the present study 15

Figure 2. 2. Overview of Molecular and Bioinformatics analysis 17

Figure 2. 3. Analysis of the effect of mutation on the 3D structure of TSHR protein 31

Figure 2. 4. Outline of metabolic profiling 33

Figure 3. 1. A representative chromatogram showing a heterozygous mutation in Exon 8 of TPO gene 39

Figure 3. 2. Alignment of BLAST result presenting two mutations in Exon8 of TPO gene . 39

Figure 3. 3. The representative predicted 3D structures of the TPO proteins. 41

Figure 3. 4. Non-bond interactions of Heme with 3D structures of full length TPO protein using BIOVIA Discovery Studio 2017 47

Figure 3. 5. Non-bond interactions of Heme with 3D structures of MPO like domain 48

Figure 3. 6. Non-bond interactions of Heme with corresponding predicted structures as obtained using a BIOVIA Discovery Studio 2017, after QM/MM calculations 50

Figure 3. 7. Superimposed structures of protein-ligand (Heme) complexes before and after 5000 ps MD simulation 51

Figure 3. 8. Non-bond interactions of Heme with corresponding predicted structures as obtained after docking with the structures retrieved from the 5000 ps simulation using a BIOVIA Discovery Studio 2017 53

Figure 3. 9. A representative dimer structure of TPO protein predicted by Symmdock server 54

Figure 3. 10. Graphical presentation of (a) van der Waals surface area, (b) Molecular surface area, (c) Solvent accessible surface area 55

Figure 3. 11. Non-bond interactions after sequential docking with H2O2 and iodine (I-) with both of left and Right monomer 57

Figure 3. 12. Normalized melt curves for the specimens targeting the c.1117G>T variant . 59

Figure 3. 13. Difference curves generated by specimens targeting the c.1117G>T variant. . 60
Figure 3. 14. Normalized melt curves generated by specimens targeting the c.1193G>C variant 61
Figure 3. 15. Difference curves for specimens targeting the c.1193G>C variant..... 61
Figure 3. 16. Normalized melt curves for specimens targeting the c.2173A>C variant 62
Figure 3. 17. Differential curves for specimens targeting the c.2173A>C variant 63
Figure 3. 18. A representative chromatogram showing a mutation in Exon10 of TSHR gene 64
Figure 3. 19. The predicted 3D structures of transmembrane region of TSHR proteins (368-764) 65
Figure 3. 20. Non-bond interactions of MS437 with corresponding predicted structures as obtained using a BIOVIA Discovery Studio 2017 68
Figure 3. 21. Non-bond interactions of MS438 with corresponding predicted structures as obtained using a BIOVIA Discovery Studio 2017. 68
Figure 3. 22. After MD simulation, RMSD calculation for both of the ligands (A)MS437 and (B)MS438 69
Figure 3. 23. Comparison of CPT-I enzyme deficiency indicator and plasma TG concentration between the patients and the healthy controls..... 73
Figure 3. 24. Comparison of CPT-I gene expression level with GAPDH housekeeping gene between the patients and the healthy controls..... 74
Figure 3. 25. Comparison of blood Gly/ΣBCAA indicator of protein intake ratios and plasma Albumin between the hypothyroid patient and the healthy controls. 77

Supplementary Figures

Figure No. and Title **Page**
Figure S 1. Structural differences of heme before and after optimization..... 108
Figure S 2. The predicted 3D structures of the proteins. (A) TPO1-933 WT, (B) TPO1-933 MT1, (C) TPO1-933 MT2, (D) TPO1-933 MT3. 109
Figure S 3. QM/MM ONIOM calcualtion set up for protein-ligand complexes, ... 110
Figure S 4. Structure of H2O2 and iodide (I-) after optimization..... 110
Figure S 5. Bond angles of specific atoms of MS437 and MS438 after optimization..... 110

List of Abbreviation

Abbreviation	
3D	3-Dimensional
BCAA	Branched Chain Amino Acids
BMI	Body Mass Index
BSMMU	Bangabandhu Shaikh Mujib Medical University
CCP	Complement control protein
CH	Congenital Hypothyroidism
CPT1	Carnitine Palmitoyltransferase1
CPT1A	Carnitine Palmitoyltransferase1-A
CPT1B	Carnitine Palmitoyltransferase1-B
C-score	Confidence Score
CV	Coefficient of Variance
DUOX2	Dual Oxidase
DFT	Density Functional Theory
EPO	Eosinophil Peroxidase
EGF	Epidermal Growth Factor
FT_4	Free-T_4
G6PD	Glucose 6-Phosphate Dehydrogenase
HRM	High Resolution Melt
HD	Homodimer
I-TASSER	Iterative Threading ASSEbmly Refinement
IQR	Interquartile Range
IYD	Iodotyrosine Deiodinase
LT_4	Levothyroxine
LC	Long Chain
LC-MS/MS	Liquid Chromatography-Mass Spectrometry/Mass Spectrometry
LPO	Lactoperoxidase
MD	Molecular Dynamics
MPO	Myeloperoxidase

Abbreviation

MT1	Mutant1
MT2	Mutant2
MT3	Mutant3
NAFLD	Nonalcoholic Fatty Liver Disease
NIS	Sodium Iodine Symporter
NINMAS	National Institute of Nuclear Medicine and Allied Sciences
OMIM	Online Mendelian Inheritence in Men
PAX8	Paired Box-8
PDS	Pendrin
PIOD	Partial Iodide Organification Defect
PCR	Polymerase Chain Reaction
QM/MM	Quantum Mechanics/Molecular Mechanics
RMSD	Root Mean Square Deviation
SASA	Solvent Accessible Surface Area
SD	Standard Deviation
SEM	Standard Error of Mean
TDH	Thyroid Dyshormonogenesis
TPO	Thyroid Peroxidase
Tg	Thyroglobulin
TG	Triglycerides
TSH	Thyroid Stimulating Hormone
TSHR	Thyroid Stimulating Hormone Receptor
TIOD	Total Iodide Organification Defect
TM	Templates Score
WT	Wild Type
XB	Halogen Bonding

CHAPTER 1: INTRODUCTION

1.1. Background

1.1.1. Hypothyroidism

Hypothyroidism is the most common endocrine disorder in which thyroid hormone deficiency occurs. Thyroid hormones (T_4 and T_3) play a significant role in regulating the metabolism, overall growth of the body and improvement of neurodevelopmental milestones. Hypothyroidism can be classified into different clusters based on the time of onset, malfunction of the thyroid gland and severity of the disorder. A number of crucial factors including autoimmune disease, treatment for hyperthyroidism, thyroid surgery and radiation therapy can play a significant role to develop the disorder. Sometimes, it may result from congenital disease, pituitary and hypothalamic disorder, pregnancy and Iodine deficiency.

1.1.2. Congenital Hypothyroidism (CH) and the Global Incidence

Congenital Hypothyroidism is the most common preventable endocrine disorder in which the thyroid hormone deficiency occurs at birth [1]. The global data showed that the frequency of CH is 1 in 3000-4000 children, whereas a pilot study reported an incidence rate of 1 in 1300 in Bangladesh [2,3]. The most common cause of CH is a shortage of iodine in the diet of the mother and the affected infants. CH is classified as permanent or transient in which permanent CH requires lifelong treatment and transient CH requires short-term treatment.

1.1.3. Clinical Features

Congenital hypothyroid children experience various clinical complications such as prolonged jaundice at birth, dry skin, constipation, delayed developmental milestone, lethargy, umbilical hernia, weight gain, open fontanelle, protruded tongue, hoarse voice and idiotic face, etc.

1.1.4. Confirmatory Test of Congenital Hypothyroidism

To diagnose CH at early age after birth, biochemical tests are performed. At first blood TSH (Thyroid Stimulating Hormone) level is measured using Heel prick and if blood TSH is ≥20

mU/L, neonate's peripheral blood is analyzed for FT_4 (Free T_4) to determine whether the baby is at risk of hypothyroidism. When Initial TSH > 30 mU/L and Initial T_4 < 10th percentile then it is suggestive of Congenital Hypothyroidism [1]. Moreover, Anti-TPO antibody and T_3 level are also measured to categorize the disorder. Ultra sonogram of thyroid gland is also required to classify patients either as thyroid Dyshormonogenesis or thyroid Dysgenesis.

1.1.5. Treatment and Prevention

Newborn screening is important to detect Congenital Hypothyroidism so that early treatment can be initiated to avoid mental retardation of the patients. Treatment should be started within the first 1-2 weeks of life with a daily dose of thyroxine hormone replacement therapy. The most commonly recommended dose range is 10-15 μg/kg/day [4,5].

1.2. Molecular Genetics of Congenital Hypothyroidism

Congenital Hypothyroidism can be caused by a variety of factors and only some of which are genetic in origin. Genetic causes account for about 15 to 20 percent of cases of Congenital Hypothyroidism. Although CH is a genetically heterogeneous disorder, the candidate genes divide the disorder into two main groups namely thyroid Dysgenesis and thyroid Dyshormonogenesis. Different studies including online databases such as Genetics Home Reference and Online Mendelian Inheritance in Men (OMIM) suggested that about 10-20 percent of total cases with CH were associated with thyroid Dyshormonogenesis that would result from mutations in one of several genes involved in the biosynthesis of thyroid hormones [6]. Till now, mutations in seven genes, namely NIS (sodium iodine symporter, SLC5A5), PDS (Pendrin or SLC26A4), TPO (thyroid peroxidase), Tg (thyroglobulin), IYD (iodotyrosine deiodinase, DEHAL1), DUOX2 and DUOXA2 (Dual oxidase) have been reported to be involved in pathogenesis of Thyroid Dyshormonogenesis (TDH) [7].

The above-mentioned databases also described that about 80-85% of CH cases are associated with disorders of thyroid gland development (Dysgenesis) which is characterized by agenesis (absent of thyroid gland, 40%), ectopic (located in a distant region, 40%) and the remaining cases are associated with hypoplasia (small size). Although the actual cause of thyroid dysgenesis is still under investigation, some studies have suggested that 4 major genes which play roles in the proper growth and development of the thyroid gland, such as TSHR (Thyroid

stimulating hormone receptor) and three transcription factors- TTF-1, TTF-2, and PAX8 (paired box-8, transcription factor) [8,9]. Mutations in these genes prevent or disrupt normal development of the gland.

1.2.1. Major Features of TPO Gene and Mechanism of Action of TPO Protein

TPO is the major enzyme in thyroid hormone biosynthesis pathway and it catalyzes both iodination and coupling of iodotyrosine residues in thyroglobulin and resulting iodotyrosines to form the thyroid hormones T_3 and T_4 [10]. TPO gene is located on chromosome 2p25 containing 17 Exons which code for 933 amino acids [11,12]. TPO is a glycosylated Heme binding extracellular protein of 110 kDa anchored to the apical membrane of thyroid follicular cells by a transmembrane region [13]. Mature TPO exists as homodimers *in vivo*, however, the localization of dimerization process in the cell is still unknown [14]. Figure 1.1 shows the chromosomal location of the TPO gene.

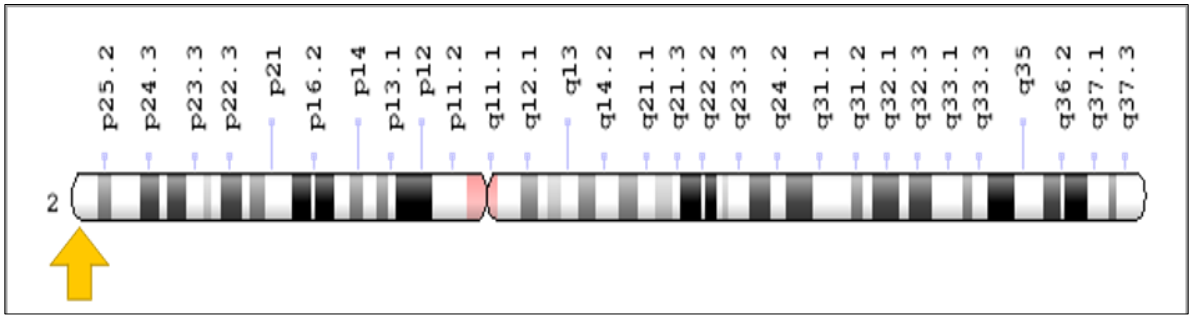
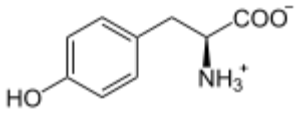
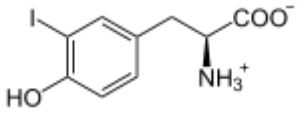


Figure 1. 1. Chromosomal location of TPO gene (source: Genetics Home References)

The chemical reactions catalyzed by thyroid peroxidase occur on the outer apical membrane surface and are mediated by hydrogen peroxide and iodine. The tyrosine residues of Thyroglobulin become iodinated by sequential reactions catalyzed by TPO as stated below.

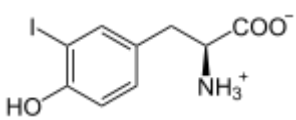


$$+ I^- + H^+ + H_2O_2 \rightarrow$$

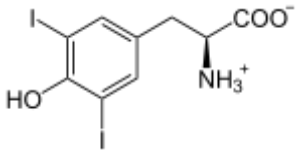


$$+ 2\ H_2O$$

Iodide is oxidized to iodine radical which immediately reacts with tyrosine.



$$+ I^- + H^+ + H_2O_2 \rightarrow$$



$$+ 2\ H_2O$$

Different studies showed that Asp238, His239, Arg396, Glu399 and His494 are the crucial amino acids found in the catalytic site of TPO and also involved in the formation of active complex which is known as "Compound 1". The proximal histidine His494 and distal histidine His239 interact with the heme molecule, whereas Arg396 participates in the cleavage of peroxide O–O bond and Asp238 and Glu399 interact covalently with the heme molecule. These amino acid residues are conserved in TPO and other mammalian peroxidases [15,16].

1.2.2. Effect of Mutation in TPO Gene and the 3D Structure of TPO Protein

Global data shows that there are more than 60 mutations found in TPO gene which are associated with total iodide organification defect (TIOD) or partial iodide organification defect (PIOD) and these phenomena have been described as the most common form of thyroid Dyshormonogenesis [17,18]. The high resolution crystallographic structure of TPO protein is under determined since only a low resolution structure is available [19-21]. The very close homologous proteins of TPO are mentioned as lactoperoxidase (LPO), myeloperoxidase (MPO), and eosinophil peroxidase (EPO) with a sequence identity of 48%, 47% and 47% respectively, and the crystallographic structure of MPO and LPO gave an depth insight of exploring the structural basis of TPO protein [20,22].

However, Sarah *et al*. determined the possible modes of TPO monomer and dimer organization using *in silico* approach [23]. Conversely, the effect of mutation on 3D structure of TPO protein and its enzyme activity are still under examination. A very recent study showed that mutations can affect at least less than or equal to 75 percent of activity of TPO dimer protein [24]. Analysis of mutation spectrum is essential to study the genetic disorders since it explores the actual reason behind disease pathogenicity as well as severity. Countries like Argentina, Netherlands, Japan, Portugal, and China have examined and reported different mutations in TPO gene through screening of patients with TIOD and PIOD [17,25-28].

Moreover, the correlation of mutation and the TPO enzyme activity were studied in vitro by different group of scientists and mild to severe TPO enzyme inactivity were found due to specific mutations compared to wild type cases [29,30]. Even though, the incidence rate of Congenital Hypothyroidism is much higher in Bangladesh than the global ratio, genetic spectrum of this disorder is still under investigation. Moreover, only a small data on the effect

of mutations on 3D structure of the major protein TPO was published very recently among the Bangladeshi patients [31].

In the present study, we investigated the mutational spectrum in TPO gene among the patients with thyroid Dyshormonogenesis to explore the possible effect of specific non-synonymous mutations on the 3D structure of TPO protein (both of monomer and dimer structures) and MPO-like domain of TPO through *in silico* approach, such as homology modelling, molecular docking followed by QM/MM (Quantum Mechanics/Molecular Mechanics) and Molecular Dynamics (MD) Simulation.

1.3. Significance of *in silico* Approach to Study 3D Structures of Proteins

Since a mutation can affect the biological activity of a protein, *in silico* approach has the potential of speeding up the process of classifying these effects into pathogenic or neutral through computer-based techniques. Currently, computational methods have become essential for studying different complex biological systems such as protein-ligand, protein-protein, drug-protein interactions that play important roles in signal transduction, cell regulation and macromolecular assemblies. Molecular docking simulations is one of the major computational methods for virtual screening of biological compounds in protein-ligand interactions for known structural and functional domains of protein molecules [32].

1.3.1. Modelling of 3D Structures using Bioinformatics Tools

3D structures of proteins are modelled from amino acid sequences using computer algorithms which can give different folding pattern and domain information of a specific protein [33]. In many cases, 3D structures provide valuable information for the large fraction of sequences whose structures are hard to determine by crystallography [34]. 3D structures of proteins are usually important in the field of drug screening process [35]. The I-TASSER server is an integrated platform for automated generation of protein structure and function prediction tool based on the sequence-to-structure-to-function paradigm with 3 main basics such as C-score (confidence score), TM (Templates score) score and RMSD (root mean square deviation) value. The cutoff value of C-score is -1.5 to select models of correct topology [36]. The 3D structures must be validated by different bioinformatics tools such as Varify3D and RAMPAGE to investigate the accuracy of basic structural features of protein.

1.3.2. Basic Parameters to Calculate the Energy Level of 3D Structures of Protein-Ligand Complexes (Quantum Mechanical Calculations and Density Functional Theory)

Quantum mechanics (QM) has been used in the study of biomolecular structures and functions which provides information about the energy levels of the proteins, the charges of their amino acid side chains, their electrostatic potentials and permanent dipole moments (μ) [37,38]. QM methods analyze various protein-ligand interactions such as hydrogen bonds, hydrophobic interactions and electrostatic interactions. The combined QM/MM (quantum mechanics/molecular mechanics) method is one of the important tools in computational biology especially in the study of enzyme chemistry. This method utilizes quantum mechanics and MM force field to calculate the electronic structure of the active site and modeling of the protein environment that deals with the long-range electrostatics and steric effects on the enzyme reactivity [39].

Density functional theory (DFT) is the most common quantum mechanical methodology used in the field of chemistry, physics and biological systems. A basis set is a set of functions in theoretical and computational chemistry that is used to represent the electronic wave function in the Hartree–Fock method or density-functional theory. B3LYP is the most popular DFT model used for optimization of ligands in Gaussian 9 software package [40]. The SDD, 6-31G and MIDIX basis sets are also used to optimize different ligands such as the Heme, H_2O_2 and I^- [41].

1.3.3. Study of Protein Structures using Molecular Docking Approach

Molecular docking approach is the most common computer simulation procedure to predict the predominant binding mode(s) of ligands with a known 3D structure of protein in a protein-ligand complex [42]. Since molecular docking serves as a virtual screening tool, it selects the most desired compound from a database or library of molecules. The process describes the interactions between protein and ligands through the lowest energy pathway [43]. Molecular docking consists of a number of mathematical methods used to predict the strength of the non-covalent interaction, such as the binding affinities based on force-field, empirical, knowledge-based and consensus scoring. The protein-ligand docking procedure can be performed as rigid body docking and flexible docking [44]. Several types of noncovalent bonds such as hydrogen bond, electrostatic bond, hydrophobic bond are critical in maintaining the three-dimensional

structures of macromolecules. All types of these weak interactions are effective only over a short range and require close contact between the reacting groups [38].

1.3.4. Molecular Dynamics Simulation

Molecular Dynamics (MD) simulation is a powerful method to investigate the dynamic properties of proteins at atomic level for discovering the functional mechanisms in biological systems. MD simulation can be applied to generate a set of docking convenient structures and can additionally be used to estimate the stability of a ligand-receptor complex proposed by molecular docking. In the MD simulation process, RMSD value predicts the stability of ligand-receptor complex. When the RMSD value is more than a cut-off value, the protein-ligand complex is considered as unstable. Now a-days super computers permit MD simulations for several nanoseconds to determine the atomic fluctuations of protein-ligand complexes in biological system. In this process, the flexible regions of proteins are also determined for stability enhancement that are influenced by a number of interactions such as hydrophobic, electrostatic interactions and Hydrogen bonding. The increasing protein stability is one of the main target of protein engineering [45]. In case of absence or low crystallographic structures for a particular molecular target, MD simulations can produce alternative conformational states corresponding to these ligand-induced structures based on force field such as CHARMM, AMBER, and GROMACS [46]. MD simulations are important in the study of membrane proteins, for which experimental characterization of structural dynamics tends to be challenging [47]. In the present study, we used YASARA dynamics program applying AMBER14 force field to calculate and to visualize possible interactions of TPO wild type and mutant protein structures with selective ligands. The premade MD-scripts of YASARA simplified the complex analysis of membrane protein TPO by creating a membrane in the simulation process [48].

1.3.5. *In silico* Approach for the Study of Dimerization of TPO Protein Structures

Since TPO acts as a homodimer protein in biological systems, we tried to study the effect of mutation on TPO dimer structure and its biological function. Dimer structure can be formed by using bioinformatics tools such as PatchDock, FireDock and Symmdock web servers. Among them, Symmdock server provides symmetry-based result using homology modeled monomeric unit for the prediction of cyclically symmetric homomultimers describing the basic

parameters such as score, area and atomic contract energy values for each structure to select the desired structure of protein [49].

1.3.6. The Sequential Docking Approach for TPO Dimer Protein

In biological systems, most of the reactions occur in an organized networking or in a cascade pathway such as signal transduction, regulation of gene expression and sequential enzymatic reactions [50]. When a macromolecule consists of multiple binding sites, the binding affinity for a specific ligand could be altered by the occupancy of other binding sites by the same or different ligands [51]. In this circumstance, the reaction events occur sequentially through the one directional flow of signal or information [50]. Following this cascade pathway, the present study aimed to perform sequential docking approach for H_2O_2 and I^- with TPO dimer structures of wild type and mutant cases to investigate the effect of non-synonymous mutation on 3D structures and function of the protein.

1.4. High Resolution Melt (HRM) Curve Analysis to Detect Mutations in TPO Gene

A recent study from our lab showed that only genetic cause accounts for all of the Congenital Hypothyroid patients with Dyshormonogenesis in hospital settings in Bangladesh and didn't find any other etiology [31]. So we tried to set up an easy-to-use method to detect the common mutations in TPO gene. Since CH is easily treatable, it is important to investigate the etiology which would help to determine how long the patients need hormone replacement therapy and the CH patients with genetic etiology need lifelong hormone therapy [52]. Sometimes neonatal CH screening using biochemical tests becomes difficult due to the presence of maternal TSH in the specimens of neonates and this problem can be overcome by genetic screening at an early age [52].

High Resolution Melt (HRM) curve analysis is a real time PCR based molecular technique which is able to differentiate the genetic variations such as homozygous or heterozygous state for specific mutation compared to the wild type state in various genetic diseases including autosomal recessive, autosomal dominant and X-linked recessive disorders [53-56]. In this study, we aimed to detect the common mutations in the TPO gene associated with thyroid Dyshormonogenesis in Bangladeshi patients and targeted 3 non-synonymous mutations to

establish HRM curve analysis as a supplementary screening approach that could be faster, cheaper and reliable.

1.5. Molecular Investigation of TSHR and PAX8 Gene

Genetic Home Reference suggested that TSHR and PAX8 genes are the major genes responsible for thyroid Dysgenesis. In the present study we tried to explore the effect of mutations in these genes and investigated the small molecular drug binding affinity on proteins 3D structure.

1.5.1. Structure and Function of TSHR Gene

TSHR gene is primarily associated with thyroid Dysgenesis, since most of the mutations occurs in the gene in CH patients [8]. TSHR is a G protein coupled transmembrane receptor which is present on the surface of thyroid follicular cells. TSH, secreted by the anterior pituitary, mediates its effect through TSHR which is crucial for thyroid gland development and function. TSHR gene is located on chromosome 14q31 and contains 11 Exons code for a receptor protein of 764 amino acid residues [57,58]. Figure 1.2 shows the chromosomal location of TSHR gene.

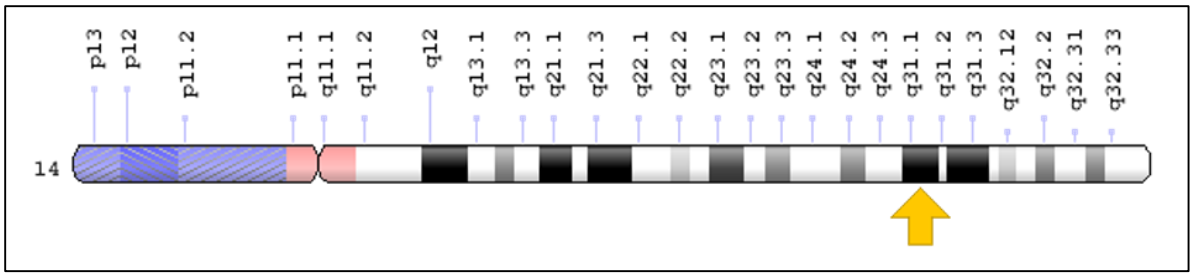


Figure 1. 2. Chromosomal location of TSHR gene (Genetics Home References)

TSHR has high affinity binding sites for TSH. Mutations in TSHR gene results in mutant TSHR protein which lacks its binding affinity to TSH or loses its ability to activate adenylate cyclase. And thus, mutant TSHR protein disrupts thyroid gland development and proper functioning. TSHR mutation may also be present with a normally placed thyroid gland. TSHR gene mutation is reported to be inherited as autosomal recessive manner and Exon 10 is known to carry majority of the mutations [59].

1.5.2. PAX8 (Paired Box Eight) Gene

The PAX8 gene is located on chromosome 2q12-q14 and it is consisted of 11 Exons. PAX-8 gene plays an important role in thyroid cell differentiation as well as in the maintenance of the differentiated state [9]. Heterozygous PAX8 mutations are found to be inherited as autosomal dominant manner in some of the patients. These patients either had hypoplastic or ectopic thyroid gland. Figure 1.3 presents the location of PAX8 gene in chromosome 2.

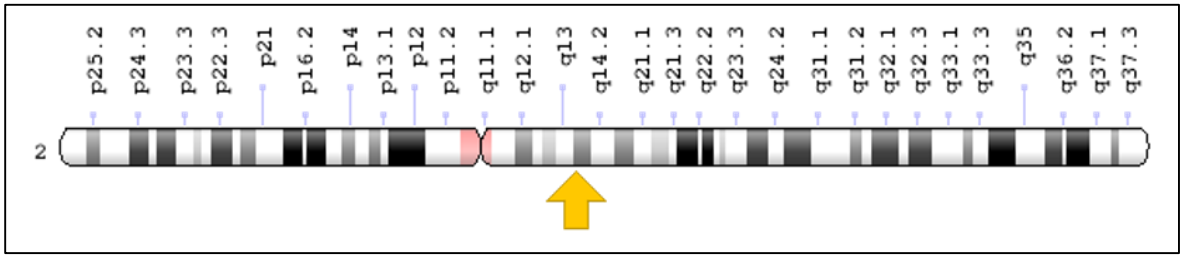


Figure 1. 3. Chromosomal location of PAX8 gene (Genetics Home References)

PAX8 mutation results in mutant PAX8 protein which has reduced or lost transcriptional activation function. As TPO is somehow dependent upon PAX8 for its efficient transcription, PAX8 mutation may cause partial iodine organification defect due to reduced TPO expression [60]. A study conducted by Macchia and team showed that common sites for PAX8 mutation are in Exon 2 and Exon 3 [9].

1.6. Metabolic Profiling of Congenital Hypothyroid Children

Hypothyroid patients suffer from many clinical complications which are associated with body's overall metabolism. Thus metabolic profiling for amino acids and acylcarnitines levels can also be described as a marker of fatigueness or energy crisis of the patients.

1.6.1. Profiling of Acylcarnitines

In the present study, all the patients were under treatment with levothyroxine drug. An elevated serum TSH and low T_4 or free T_4 (FT_4) level indicates a hypothyroid condition which is reversible with appropriate LT_4 treatment. Although LT_4-mediated management of hypothyroidism has been considered comparatively straightforward, data suggested that 5-10% of LT_4-treated patients suffer from persistent clinical complications of hypothyroidism despite maintenance of normal blood levels of TSH and FT_4 [61-63]. Among the physiological distresses in these patients, fatigue and fatigue-related symptoms are the most commonly

documented manifestations. However, the mechanism causing persistent fatigue in these patients remains unclear [61,62].

L-carnitine is an endogenous compound and its major biological role is in the transport of fatty acids across the inner mitochondrial membrane for fatty acid oxidation via the reversible binding of acyl groups from Coenzyme A (CoA) [64]. Carnitine has also a role in regulating the cellular to mitochondrial ratio of free CoA to acyl-CoA [65]. Alterations in carnitine homeostasis can have a harmful effect on human health and are associated with muscle weakness, progressive cardiomyopathy and encephalopathy [66-68]. Carnitine supplementation increases fatty acid oxidation and has been used in management of patients suffering from muscular hypoxia, peripheral vascular disease, angina pectoris, congestive heart failure, and hemodialysis [69-71].

In this study, whole blood carnitine-acylcarnitines profiling of 65 late-diagnosed hypothyroid patients was analyzed using liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS). Although these patients were kept on prescribed treatment regimen with LT_4 to maintain normal ranges of serum TSH and FT_4 levels, they still had been suffering from various fatigue-related complications. We hypothesized that the fatigue-related symptoms in these hypothyroid patients may be associated with altered carnitine-acylcarnitines homeostasis. Thus, this study aimed to profile carnitine-acylcarnitines levels of the hypothyroid children and to investigate whether any alterations of blood carnitine-acylcarnitines levels had correlation with their ongoing fatigue and fatigue-related complications.

1.6.2. Profiling of Amino Acids

Since thyroid hormones are essential for body's overall metabolism and energy production, we wanted to correlate the energy crisis of late-diagnosed congenital hypothyroid children with amino acids and acylcarnitines levels. Normally, thyroid hormones accelerated the biosynthesis of proteins by altering various amino acid pools [72]. In case of thyroid hormone deficiency, the amino acid concentration might be altered and manifests some clinical complications such as fatigue, even under treatment with LT_4. The functions of all types of amino acids are associated with different mechanism in human body such as in energy production, neurotransmission, immune function and regulations of growth hormones, etc.

[73]. In the present study, almost all the patients were late-diagnosed and they manifested some major complications like fatigue, delayed developmental milestone, dry skin, neonatal jaundice, etc. even though they were under treatment and their blood TSH and FT_4 levels were normal during the study period.

1.7. Rationale of the Research Work

Congenital Hypothyroidism (CH) is the most common preventable endocrine disorder that is associated with partial or complete loss of functions of the thyroid gland which affects infants from birth (Genetics Home References). The iodinated thyroid hormones are biosynthesized from thyroid gland and they play a significant role in the development of brain, regulation of overall growth and they control the metabolic reactions of the body. A previous study showed that the frequency of CH is 1 in 1300 children in Bangladesh, which is much higher than the global incidence [3]. There is also a lack of newborn screening program of CH in Bangladesh, so late diagnosis is a common phenomenon in Bangladesh. Early diagnosis and initiation of treatment is necessary to prevent the clinical complications experienced by these patients. If treatment is initiated within the first two weeks after birth, growth of infants develops normally. If the infants remain untreated for several months after birth, severe Congenital Hypothyroidism can lead to growth failure and permanent intellectual disability. Due to late diagnosis, patients experience many clinical complications even with levothyroxine treatment. There is limited data regarding Congenital Hypothyroidism in Bangladesh and the genetic etiology of CH is still under investigation. Moreover, the metabolic profiling is important to correlate the clinical complications of hypothyroid children to improve the quality life. In this study, we aimed to investigate the mutations in specific genes such as TPO, TSHR and PAX8; associated with Congenital Hypothyroidism in Bangladeshi patients as well as to correlate the clinical complications with metabolic profiling of these participants.

1.8. Objectives of the Research Work

General objectives

a. To investigate the genetic spectrum of congenital hypothyroid children
b. Metabolic profiling in terms of amino acids and carnitine/acylcarnitines of hypothyroid children

Specific objectives

a. Mutation detection of TPO gene responsible for thyroid dyshormonogenesis and investigation of the effect of mutations on 3D structures for both cases of monomer and dimer TPO proteins
b. Translations of the laboratory findings into an easy-to use approach (High Resolution Melt curve analysis) for screening of common mutations in TPO gene
c. Mutation detection of TSHR and PAX8 gene associated with thyroid dysgenesis
d. Profiling of Acylcarnitines and Amino acids in hypothyroid patients

CHAPTER 2: MATERIALS AND METHODS

2.1. Study Design, Clinical Settings and Ethical Clearance

The study was designed as a cross-sectional one with purposive sampling and carried out on 65 confirmed cases of Congenital Hypothyroid children in the Department of Endocrinology and National Institute of Nuclear Medicine and Allied Sciences (NINMAS) of Bangabandhu Shaikh Mujib Medical University (BSMMU). Ethical permission was obtained from the Ethical Review Committee of University of Dhaka and the study was collaborated with NINMAS and Dept. of Endocrinology, BSMMU for specimen collection. Prior to enrollment of study participants, a written informed consent along with the clinical information was collected from the parent(s) or legal guardian(s) of each patient. Figure 2.1 shows a representative workflow of the study.

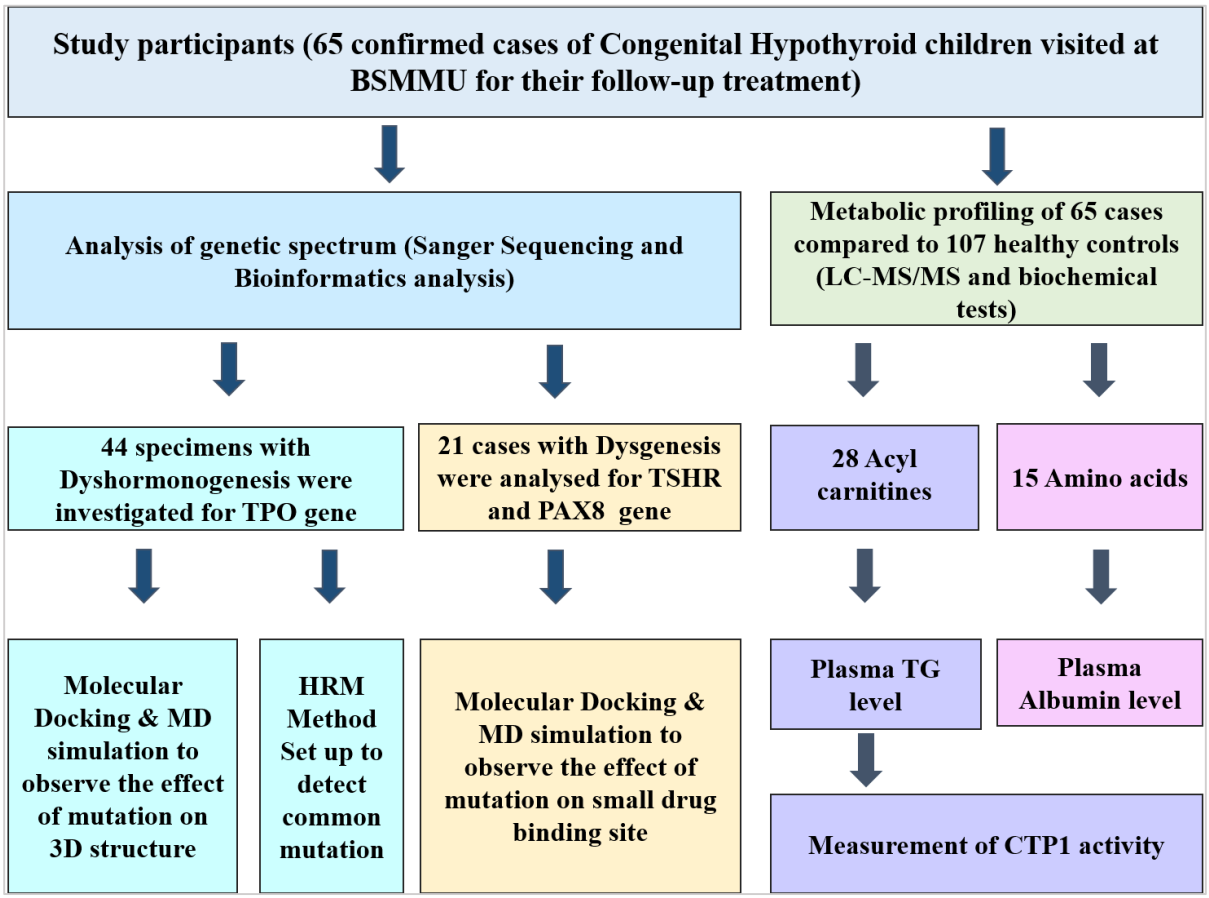


Figure 2. 1. A Representative flowchart of the present study

2.1.1. Study Participants (Visits at BSMMU)

Since the study was cross sectional, we recruited patients from BSMMU for over 3 years and collected specimens during 2016 to 2019. The frequency of Congenital Hypothyroidism (CH) is 1 in 3000-4000 children worldwide and we enrolled a total of 65 participants (aged less than 18 years) which represents a total of 227,500 children in Bangladesh. Every visits were performed twice a week. In the OPD (Outpatient Department) of Pediatrics Endocrinology and NINMAS, about 50-60 children visited per day with different endocrine disorders. Among them 1 or 2 children were confirmed cases for Congenital Hypothyroidism who visited BSMMU for their follow-up examinations. Almost all the patients were late-diagnosed and kept under treatment of Levothyroxine (LT_4) drug. If we wanted to collect more samples for CH, then we had to visit more times and enrolled more populations, that was actually not possible in this study period.

2.1.2. Healthy Subjects

A total of 107 participants (aged 2- 18 years) without any thyroid complications and fatigue were enrolled in this study as the healthy controls. All the healthy participants were the residents of Mirpur, Dhaka. A written informed consent was obtained from a parent or a legal guardian of each healthy participant.

2.1.3. Collection and Processing of Blood Specimens

Blood Specimens were collected from the participants to conduct the molecular, biochemical and metabolic profiling tests. A total of 5ml blood was collected from each participant. Then it was aliquoted in different tubes for different types of analysis, such as (1) 3 mL blood in EDTA tube for DNA isolation, (2) 75 µl blood to prepare Dried Blood Spots (DBS, Whatman® 903 generic multipart filter paper, GE Healthcare, Westborough, USA) for metabolic profiling, (3) 250 µl of blood in tube containing 750 µl Trizol® LS (Ambion, Life Technology, Cat. No. 10296-028) for conducting genetic study using RNA and (4) rest of the blood was used for separation of plasma-serum for biochemical test. All the samples were transported to the laboratory immediately. The DBS cards were dried for at least 4 hours at room temperature and once dried, they were stored at -70°C freezer with desiccant for protection from moisture. All the plasma, serum and Trizol containing specimens

were stored at -20°C freezer. After the genomic DNA isolation, EDTA containing blood was stored at -70°C freezer.

2.2. Molecular Analysis of TPO, TSHR and PAX8 Gene

Now-a-days, gene-based study plays the key role to explore the actual cause of a particular disease. The present study was designed to perform the molecular analysis in various steps. Figure 2.2 presents the overall procedures of molecular analysis.

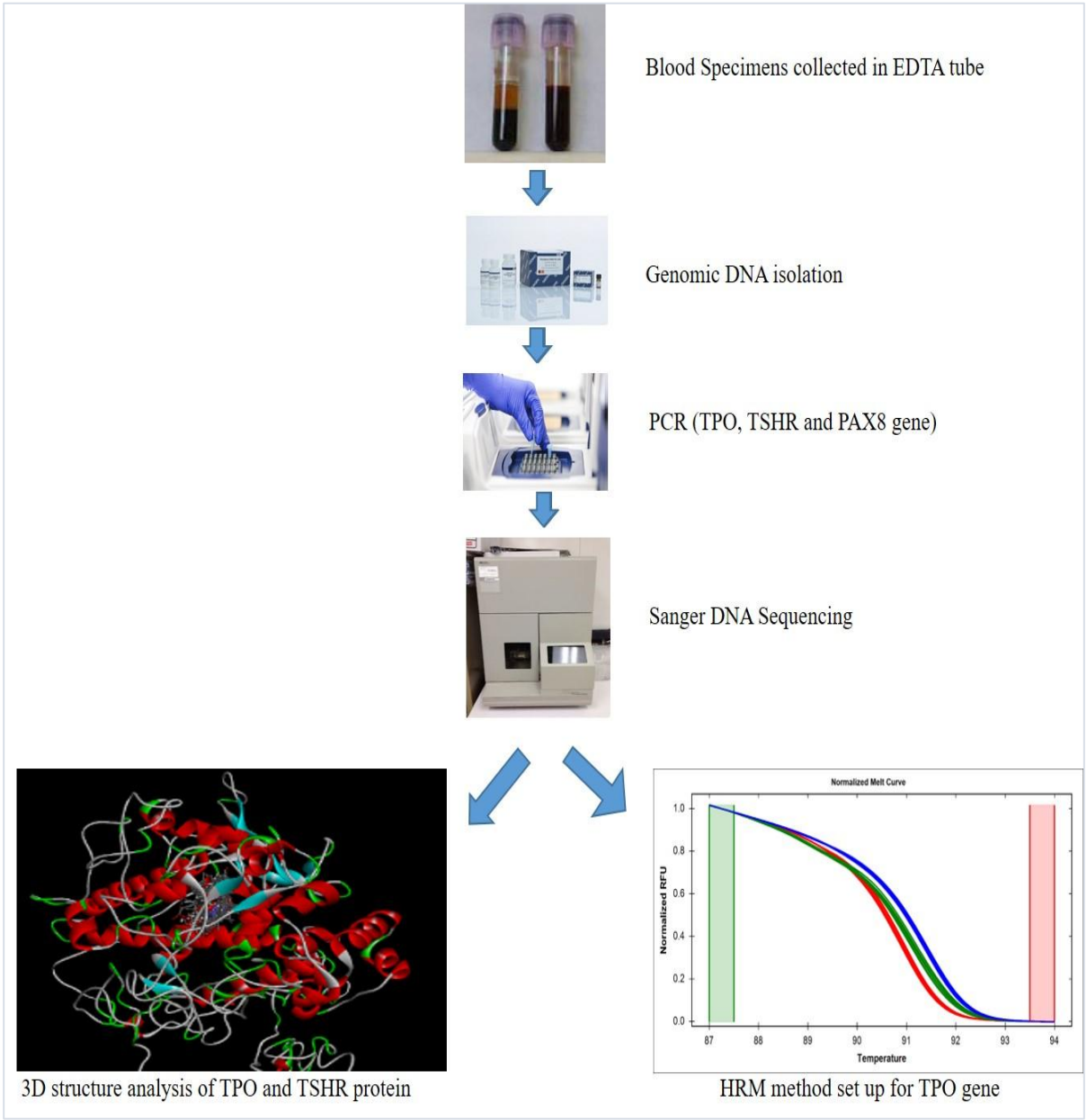


Figure 2. 2. Overview of Molecular and Bioinformatics analysis

2.2.1. Genomic DNA Isolation to Perform PCR

Genomic DNA was isolated from the EDTA blood by using Qiagen DNAeasy mini kit according to the manufacturer's instruction. 500 µl of FG1 buffer was taken in a 1.5 ml micro-centrifuge tube. 200 µl of whole blood was added to the FG1 buffer and mixed by inverting the tube 5-10 times. The mixture was then centrifuged at 10,000 x g for 5 minutes in fixed angle rotor. The supernatant was carefully removed so that the pellet remained in the tube. 1µl of QIAGEN protease was added to 100 µl of FG2 buffer and mixed by vortex in a fresh Eppendorf tube. Then 100 µL of FG2/QIAGEN protease was added to the pellet and vortexed immediately until the pellet was completely dissolved and the color was changed into olive green so that all the protein components were degraded. The mixture was then incubated in a water bath or heat block at 65^0C for 5 minutes. After incubation, 100 µl of isopropanol (100%) was added and mixed by inversion until DNA was precipitated as visible threads. The tube was centrifuged for 5 minutes at 10000×g. The supernatant was discarded and the pellet was dried by keeping the tube inverted state on a clean tissue paper for one minute. 100 µl of 70% ethanol was added and vortexed for 5 seconds. The tube was centrifuged for 5 minutes at 10000×g. The supernatant was carefully aspirated using a micropipette and keeping the micro-centrifuge tube in the inverted state on the tissue paper to allow the pellet to air dry for at least 5 minutes. Over-drying was avoided as the process can make it difficult to dissolve the DNA. Depending on the pellet size, 25-50 µl of nuclease free water was added and the tube was vortexed for 5 seconds and the mixture was incubated at 65^o C for one hour in water bath for dissolving DNA or left overnight at room temperature. Nuclease free water was used instead of FG3 buffer (Elution buffer). Finally, the concentration and the purity of the DNA was measured using a Nano drop machine and adjusted the concentration for PCR.

2.2.2. Polymerase Chain Reaction (PCR) Amplification of TPO Gene

The isolated DNA was then amplified by PCR using TPO gene-specific primers. At first, we performed PCR using primers set that could flank the sequence between Exon 8 to Exon 14, since global data showed that most of the common mutations in the TPO gene of the patients with Congenital Hypothyroidism were confined in this region. Next, we conducted PCR for other regions of the TPO gene. The primer sequences are listed in the Table 2.1 as follows.

Table 2. 1. List of primers for PCR amplification and Sanger sequencing of TPO gene

Primer Name	Sequence (5'-3')	Product size(base pair)
TPO_Ex1F	CTCGACTTCCTAGCATCTTGAC	306
TPO_Ex1R	CACTCATCAGTCAGTCAGTCAG	
TPO_Ex2F	TCTGTAACCTCCTGACATGGACG	384
TPO_Ex2R	GCCTAAAGTGACTCTCAGGAGC	
TPO_Ex3F	CTGTCAGTGAATCGCTGAACTG	306
TPO_Ex3R	CAGTTCTGTCTGCTGCTTGCAC	
TPO_Ex4F	CAGTGCCTGTCACATTGTCTGG	365
TPO_Ex4R	CCTGCACAAAGTCAAGGTGTCC	
TPO_Ex5F	GCTGGAGTCACTTTTGAGACTGC	289
TPO_Ex5R	TGATGGCATCTAGTCATGTTCAG	
TPO_Ex6F	CTTGAAGCCAGTCATGACCTGG	433
TPO_Ex6R	GAGGATCTGAAGGGATTGCACG	
TPO_Ex7F	CACAGGGTCCTCCTATGTCCTG	392
TPO_Ex7R	TGCACCAGATATTCACAACTGC	
TPO_Ex8F	TCCAGAGTCTTACAAAGGGTGC	679
TPO_Ex8R	GTACCTGGGAGAGAGAAGCCAC	
TPO_Ex9F	GAGGTGCTGTCTCTTGCCACTG	568
TPO_Ex9R	GGAAGAGTTCATGGGGACCAGC	
TPO_Ex10F	TAGAACTGAGCCAAGAGCTGTC	292
TPO_Ex10R	CTAGCAGCAGGTTGCTAGCTCG	
TPO_Ex11F	GACCATGGCATGAGTGAGATGG	363
TPO_Ex11R	CTGCCTCTGTGCAGAACGTG	
TPO_Ex12F	GGTTCTCCATGCACTGTGACCT	1044
TPO_Ex13R	GATTCCACGTGCCTGTCCTGAG	
TPO_Ex14F	CCTCATCACCTTTTCGGATGTGC	512
TPO_Ex14R	CAGACAGCAGGCACACGAAGTG	
TPO_Ex15F	CACGTCACATTCGGATCTGGAC	404
TPO_Ex15R	CCAGTAACCTGTGTTAGCTTCG	
TPO_Ex16F	GAGACAGGGTCCTCTTGGCTGC	557
TPO_Ex16R	AGGCTCACATTCTGGCTCACTG	
TPO_Ex17F	GCAGATCCCACTCTGTGAGTAG	636
TPO_Ex17R	GAGCACGGCATCAGAGATGC	

To amplify the desired target sequence of TPO gene, PCR amplification was conducted on a thermal cycler (Bio-Rad, USA). The final reaction volume was 10 µl for each of the reactions

which contained 1 µL 10X PCR buffer, 0.3 µL 25mM $MgCl_2$, 2 µL 5X Q-solution, 1.6 µL 2.5 mM dNTPs mixture, 0.2 µL 10mM Forward primer and 0.2 µL Reverse primer, 0.05 µL Taq DNA Polymerase, 50 ng of genomic DNA and total reaction volume was made up to 10µL by addition of nuclease free water. The thermal cycling condition included (a) initial denaturation at 95°C for 5 minutes, (b) cyclic denaturation at 95°C for 40 seconds and annealing at 58°C for 35 seconds and extension at 72°C for 40 seconds; and (c) final extension at 72°C for 5 minutes for 35 cycles.

2.2.3. PCR Amplification of TSHR Gene

The isolated DNA from Congenital Hypothyroid patients with Dysgenesis was then amplified by PCR using TSHR gene-specific primers listed below in Table 2.2.

Table 2. 2. List of primers for PCR amplification and Sanger sequencing of TSHR gene

Primer name	Primer sequence	Tm	Product size
TSHR_Ex1F	GGCATCTAAACTAGGCTTTGGAG	62.6 ºC	646 bp
TSHR_Ex1R	CTTCGGGCTGTTATTGAGCTGC	65.2 ºC	
TSHR_Ex2F	AGTGTGATGCGAGGCAAGAC	64.3 ºC	645 bp
TSHR_Ex2R	CAGCTAAGGTTTTGCCATATCCC	63.1 ºC	
TSHR_Ex3F	GTGGAACATTCCACAGGGTGAC	64.7 ºC	622 bp
TSHR_Ex3R	CTTCCAACCATGGAATTGAGGTG	63.3 ºC	
TSHR_Ex4F	AAAGTGGACAGAAACCAAGCC	62.9 ºC	470 bp
TSHR_Ex4R	GATCATTTCACCCGATACCTTGC	63 ºC	
TSHR_Ex5F	CCGAGCAGATGTATTGACACCAG	64.3 ºC	577 bp
TSHR_Ex5R	TACCCAAGTCTCTCTTGAGCC	62.7 ºC	
TSHR_Ex6F	TCCAGGTGCATGTCATCTAGG	63.1 ºC	556 bp
TSHR_Ex6R	GGTTGCATGGTCTGTAATGCC	63.4 ºC	
TSHR_Ex7F	AGAGACTGCAGCTGCTCCTCC	67.2 ºC	646 bp
TSHR_Ex7R	AGCTTTGGAACTTACCATTGGAG	62.7 ºC	

Primer name	Primer sequence	Tm	Product size
TSHR_Ex8F	GAATGTTTTAAGTGCTCAAGCCAG	62.4 ºC	834 bp
TSHR_Ex8R	GGCAATGATACAGAGGCTTCAGG	65.2 ºC	
TSHR_Ex9F	AGCATTTGTACTACTGGATACTGG	61.8 ºC	533 bp
TSHR_Ex9R	CTTCCAATTTCCTCTCCACCTG	62.3 ºC	
TSHR_Ex10F1	AGGAATGATGTCACAGAAACAGGC	64.7 ºC	1151 bp
TSHR_Ex10R1	GTGATGGCATACCAGCGCTCCAG	68.4 ºC	
TSHR_Ex10F2	CGCTTTCTCATGTGCAACCTGGC	67.7 ºC	1030 bp
TSHR_Ex10R2	GGGTGTCATGGGATTGGAATGC	65.2 ºC	

To amplify the desired target sequence of TSHR gene, the master mix preparation was same like the master mix for TPO gene amplification except the annealing temperature of the primer was 60 to 62 ºC. All the downstream procedures were same which were used for TPO gene PCR.

2.2.4. PCR Amplification of PAX8 Gene

The PAX8 gene was amplified to investigate whether there was any mutation found in CH patients with problem in thyroid gland formation (Dysgenesis). A set of primers that could flank the hotspot region of the PAX8 gene was used to amplify the target region. PCR master mix was prepared as mentioned earlier for TPO gene. Table 2.3 shows the primes used in case of PAX8 gene.

Table 2. 3. List of primers for PCR amplification and Sanger sequencing of PAX8 gene

Primer Name	Sequence (5'-3')	Product size(base pair)
PAX8_Ex3_F	AGCATCTTGGGAGTGAGAACTGG	404 bp
PAX8_Ex3_R	GATCAGCTGGAGAAGTCAAGC	

The optimized conditions used for PCR were pre-denaturation at 95°C for 15 minutes; 35 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 45 seconds and extension at 72°C for 1 minute; and a final extension at 72°C for 10 minutes. All the downstream

processes for Sanger sequencing and data analysis of PAX8 gene were similar to TPO gene sequencing.

2.2.5. Agarose Gel Electrophoresis and Gel Documentation

All the PCR products were analyzed in 1% (w/v) agarose gel to visualize the signals of PCR band of specific product size. Gel documentation was performed by Fusion Plus and Biorad documentation systems. 1kb plus ladder was used in the gel electrophoresis process.

2.2.6. Purification of PCR Products

Prior to sequencing, All the PCR products (TPO, TSHR and PAX8) were purified using Qiagen PCR purification kit according to the manufacturer's instruction. The PCR products were spun and then transferred to an Eppendorf tube. 1:250 volume pH Indicator-I was added to the Buffer PB. pH indicator I was added to the entire buffer contents. Ethanol (96-100%) was added to the Buffer PE. Five volumes of Buffer PB were added to 1 volume of the PCR products and mixed properly. It was ensured so that the color of the mixture was yellow (similar to Buffer PB). If the color of the mixture was orange/violet, 10 µL 3 M sodium acetate (pH 5.0) was added and mixed well by pipetting up and then inverting. The mixture was transferred to a MinElute Column that was placed in a 2 mL collection tube. The preparation was allowed to stand at room temperature for 5 mins and then it was centrifuged at 17,900 × g for 1 min. The flow-through was discarded, and the MinElute Column was placed back into the same collection tube. 750 µL Buffer PE was used to wash the column. The tube was allowed to stand at room temperature for 5 min and then centrifuged for 1 min. The flow-through was discarded and the column was placed back into the same collection tube followed by repetition of the centrifugation. Each column was then placed in a new 1.5 mL micro-centrifuge tube. Depending on the amount of PCR products, nuclease-free water was applied to the center of the column membrane for complete elution of the bound DNA and was allowed to stand at room temperature for 5 mins and then centrifuged at 17,900 x g for 1 min. The purity of the purified PCR product was then checked and the concentration was adjusted for cycle sequencing reaction.

2.2.7. Cycle Sequencing Reaction Followed by Sanger Sequencing

Cycle sequencing was performed for all the purified PCR products using BigDye Chain Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, USA) according to manufacturers' instruction. For each reaction, a master mix was prepared by adding 2.0 µL of 5X sequencing buffer, 0.5 µL of BigDye® chain terminator, 0.3 µL of individual primer (either forward or reverse) and nuclease free water. The template was spun and concentration was checked in a Nanodrop machine and the concentration was set depending on the PCR product size. In this reaction, 9 µL master mix and 1 µL purified PCR product was added to make the total volume 10 µL. The PCR tubes were spun at 4000 rpm for 3 minutes and then placed in the PCR machine. The thermal cycling profile included (a) initial denaturation at 94ºC for 1 minute, (b) cyclic denaturation at 94°C for 10 seconds and annealing at 58°C for 5 seconds and extension at 60°C for 4 minutes; and (c) final extension at 60°C for 10 minutes for 25 cycles.

2.2.8. Purification of Cycle Sequencing Product

The PCR product obtained from cycle sequencing PCR was then purified using a BigDye XTerminator® Purification Kit (Applied Biosystems). The products were centrifuged for 2 mins at 4100g. Purification of cycle sequencing products were performed by mixing with 45 µL of SAM solution and 10 µL of X-terminator solution in an automated shaker for 30 minutes. After centrifugation at 8000 X g for 10 minutes, 10 µL supernatants from each tube were transferred to a new 0.2 mL PCR tube strip. Finally, sequencing of the purified cycle sequencing product was performed on the ABI PRISM 310 automated sequencer (Applied Biosystems, USA).

2.2.9. Analysis of Sanger Sequencing Data

After completion of sequencing run, all the data were collected using ABI PRISM 310 data collection software version 3.1.0 (Applied Biosystems). Collected FASTA format of sequencing data was used to identify substitution or deletion mutations in the TPO gene by alignment with the reference sequence (for TPO gene, Accession number; NC_000002.12 retrieved from the NCBI database) using the basic local alignment search tool (BLAST). The accession numbers for TSHR and PAX8 genes were NG_009206.1 and NG_012384.1,

respectively. Finally, ExPASy translate tool was used to convert nucleotides sequence into corresponding amino acids and analyzed whether there was any substitution of amino acids.

2.3. Study of the Effect of Non-Synonymous Mutations on 3D Structure of TPO Protein

2.3.1. Modelling of 3D Structure of TPO protein through *in silico* Approach

The present study found three common non-synonymous mutations such as p.Ala373Ser, p.Ser398Thr and p.Thr725Pro in the hotspot region of the full length TPO_{1-933} protein. So that we set our first goal to consider the effect of these mutations on the 3D structure of MPO-like domain of the TPO protein. Additionally, we expected to understand whether the mutations had any effects on the Heme interactions with specific amino acid residues in the catalytic site of TPO protein. The corresponding positions of the mutations including p.Ala373Ser, p.Ser398Thr and p.Thr725Pro in the MPO-like domain of $TPO_{142-738}$ were p.Ala232Ser, p.Ser257Thr and p.Thr584Pro and designated as $TPO_{142-738}$MT1, $TPO_{142-738}$MT2 and $TPO_{142-738}$MT3, respectively as mentioned previously [31]. Since the high resolution crystallographic structure of TPO protein was not available and to understand the effect of mutations in the structure of MPO-like domain of $TPO_{142-738}$; amino acid sequence of wild type and sequences containing different non-synonymous mutations were submitted to I-TESSER server in order to obtain the 3D structures [33,36,74]. Based on C-score, Template Modelling (TM) score and Root Mean Square Deviation (RMSD) score, we obtained 1 model for each structure. Moreover, we studied the effects of the mutations on full length TPO protein structure and functions with the structural features of MPO like domain in order to compare the basic variation. For this reason, we also predicted the structures of full length TPO protein (TPO_{1-933}) of the wild type (TPO WT) and mutant cases (TPO MT1, TPO MT2 and TPO MT3) through the I-TASSER server and obtained five models for each described in our study [31]. From the models, we selected the suitable ones based on the organization of its various domains such as myeloperoxidase (MPO)-like domain (residues 142–738), Complement control protein (CCP)-like domain (residues 740–795), and Epidermal growth factor (EGF)-like domain (residues 796–846) of the whole protein according to the published article [23].

2.3.2. Validation of the 3D Structures

The predicted 3D structures of full length TPO protein and MPO like TPO protein such as TPO_{1-933}WT, TPO_{1-933}MT1, TPO_{1-933}MT2 and TPO_{1-933}MT3; and $TPO_{142-738}$ WT, $TPO_{142-738}$MT1, $TPO_{142-738}$ MT2 , $TPO_{142-738}$ MT3, were validated using different web-servers, namely Verify3D server and RAMPAGE server [75-77]. The PDB format of the predicted 3D structures were submitted to both the servers to validate the structures. From Verify3D server, we obtained the results including the percentage of the amino acids having the average 3D-1D score of ≥ 0.2. The RAMPAGE webserver implements Ramachandran Plot Analysis by providing results comprising the percentages of the amino acid residues within the favored, allowed and outlier regions. The results of the Verify3D and the RAMPAGE webservers were summarized in the results section.

2.3.3. Optimization of Heme using Quantum Mechanical (QM) Calculations

Since TPO is dependent on Heme for its activity, we investigated the effect of mutation on binding affinity of Heme with TPO and its interactions with specific amino acid residues. The initial structure of Heme was obtained from the Protein Data Bank (PDB) database, (PDB ID: HEM). Quantum mechanics (QM) calculation had been used to optimize the Heme. The QM calculation was performed in Gaussian 09 program based on density functional theory (DFT) employing Becke's (B) exchange functional combining Lee, Yang and Parr's (LYP) correlation functional, known as B3LYP density functional [78-80]. SDD basis set had been used for optimization of the Heme [81,82].

2.3.4. Molecular Docking Analysis of Heme with 3D Structures of TPO protein

Autodock vina protocol was employed to perform molecular docking approach [83,84]. This method is associated with the prediction of the interactions between a small molecule and a protein in atomic levels delivering us the opportunity to inspect the behavior of small molecules in the binding site of target proteins and to explain the fundamental biochemical processes. The information about the position of ligand binding sites in the target protein could significantly increase the docking efficiency during the molecular docking analysis [85]. The binding sites of Heme with TPO had been identified. It was found that Asp238, His239, Arg396, Glu399 and His494 were present in the active site of TPO_{1-933} i.e., MPO-like domain

and these amino acid residues were involved with Heme interactions and thus the catalytic activity of TPO [15,16]. But in case of MPO-like domain of $TPO_{142-738,}$ the corresponding position of Asp238, His239, Arg396, Glu399 and His494 would be Asp97, His98, Arg255, Glu258 and His353, respectively. Optimal confined search space had been selected for successful flexible molecular docking of Heme with $TPO_{142-738}$ WT, $TPO_{142-738}$ MT1, $TPO_{142-738}$ MT2 and $TPO_{142-738}$ MT3. The center of the grid box for $TPO_{142-738}$ WT was set to71.337Å, 73.748Å and 72.888Å, whereas they were set to 71.281Å, 73.813Å and 73.314Å for $TPO_{142-738}$ MT1, 73.115Å, 74.888Å and 73.156Å for $TPO_{142-738}$ MT2 and 73.273Å, 75.433Å and 74.247Å for $TPO_{142-738}$ MT3 in the x, y and z directions, respectively as mentioned in our previous study [31]. The grid box size was set to 25.0Å, 25.0Å and 25.0Å for $TPO_{142-738}$ WT, MT1, MT2 and MT3 in the x, y and z directions, respectively. Grid box value center and grid box size were also optimized for full length TPO_{1-933} (Table 2.4).

Table 2. 4. Grid box center and grid box size for full length TPO1-933 WT, MT1, MT2 and MT3

Proteins	Grid box center (Å)			Grid box size (Å)		
	X	Y	Z	X	Y	Z
TPO_{1-933} WT	109.7751	128.9632	119.5649	24.5891	20.8773	31.4612
TPO_{1-933} MT1	105.160	102.301	98.7798	25.0	25.0	25.0
TPO_{1-933} MT2	99.6151	102.226	100.5057	25.0	25.0	25.0
TPO_{1-933} MT3	102.196	124.345	113.7515	25.0	25.0	25.0

2.3.5. Visualization and Analysis of Molecular Docking Results

The molecular docking results of Heme with the MPO like domain of TPO such as $TPO_{142-738}$ WT, $TPO_{142-738}$ MT1, $TPO_{142-738}$ MT2 and $TPO_{142-738}$ MT3 and also for the full length TPO protein TPO_{1-933} were visualized and analyzed using the PyMol (version 2.0) and BIOVIA Discovery Studio 2017 software [86,87]. The binding affinity of Heme with wild type TPO (TPO WT) were compared to 3 mutants (TPO MT1, TPO MT2 and TPO MT3) proteins. Furthermore, the corresponding non-bond interactions of amino acids in the catalytic pocket of TPO protein and Heme were also considered for wild type and mutant 3D structures as described in the published article [31].

2.3.6. Validation of Molecular Docking Results using QM/MM and Molecular Dynamics (MD) Simulation

Since published data showed that MT1 and MT3 had severe effect on enzyme activity than MT2 then we further validated the molecular docking results by quantum chemical method QM/MM and Molecular Dynamics (MD) simulations for full length TPO proteins (Wild type, MT1, MT3). The QM/MM calculations of the selected protein-ligand complexes were performed by a two-layer ONIOM method available in the Gaussian09 software package [88-94]. The QM and MM regions have been shown in supplementary figure 03. The Heme molecule was included in the QM region and semi-empirical PM6 level of theory was considered due to large structure of the Heme molecule that was depicted in the previously published data [31,95]. The regions of full-length protein were computed in the MM region and the Universal Force Field (UFF) was used for the energy minimization. The total ONIOM energy of the entire system was described as, $E^{ONIOM} = E^{real,MM} + E^{model,QM} - E^{model,MM}$ [96]. The real system consists of all the atoms and was calculated only at the MM level depicted in Supplementary Figure 3. The model system consists of the part of the system (such as Heme) that was treated at the QM level [88].

To perform Molecular Dynamics simulation, YASARA dynamics program was employed and AMBER14 force field was considered for all calculations [96,97]. The size of the cubic simulation box was 167.17Å* , 167.17 Å* and 167.17 Å; and 151,343 water molecules were added to maintain a solvent density of 1.0 g/ml as mentioned in our article [31]. The total number of atoms in the system was 469,240. Due to the large size of the protein, we have performed 5000 ps MD simulation [98]. For short-range van der Waals and Coulomb interactions, a cut-off radius of 8.0 Å was considered. Long-range electrostatic interactions were taken into consideration using the Particle Mesh Ewald algorithm [98]. MD simulation was performed for 5000 ps at 310K having a time step of 1.25 fs and the simulation snapshots were saved at every 100 ps. To check whether the ligand stayed in the same binding pocket after 5000 ps of simulation, the Heme ligands were retrieved from the simulated ligand-protein complexes and molecular docking was performed again using Autodock Vina protocol as mentioned earlier. The Molecular docked structures were analyzed using the same protocol as stated previously.

2.4. Modelling and MD Simulation of TPO Dimer Structures

To investigate the effect of mutation on TPO dimer structures, we selected the previously predicted monomer structures of TPO proteins for wild type, MT1 and MT3 to model dimer structures (TPO WT-HD, TPO MT1-HD, and TPO MT3-HD) using SymmDock web server and BIOVIA Discovery Studio version 4.5. Then the molecular dynamics simulation was performed using YASARA software to stabilize those dimer structures. As TPO is a membrane bound protein, membrane was attached during the simulation process. Then the dynamics result was analyzed by comparing van der Waals surface area, solvent accessible surface area (SASA) and molecular surface area for TPO wild type and mutant dimer proteins.

2.4.1. Analysis of Binding Energies After Sequential Docking Approach

2.4.2. Molecular Docking of Hydrogen Peroxide (H_2O_2) and Iodide (I^-) with TPO WT-HD, TPO MT1-HD and TPO MT3-HD Homodimer Structures by PyRx

As TPO dimer protein exerts its effect through a sequential reaction with H_2O_2 and I^- to synthesize the thyroid hormones, we optimized these ligands structures and depicted in Supplementary Figure 4 and Supplementary Table 3. In the sequential docking process, H_2O_2 and I^- were treated as ligands for Heme-bound dimer structures of TPO protein, such as TPO WT-HD, TPO MT1-HD and TPO MT3-HD. The docking procedures were performed step by step for both of the "L (left)-monomer" and "R (right)-monomer") of dimer structures. At first H_2O_2 was docked with the heme-bound TPO dimer structure and finally I^- was docked with the Heme and H_2O_2 bound TPO dimer structures. In the biological systems, the binding of H_2O_2 with the TPO dimer initiates the catalytic reaction by forming the "Compound I" in which iodide binds and oxidizes to form "Compound II" and finally reacts with tyrosine residue of Tg (thyroglobulin) to synthesize thyroid hormones. After flexible docking, we analyzed the binding energies of H_2O_2 and I^- with the predicted dimer structures and also visualized the interactions with specific amino acids by using the PyMol (version 2.3.2) software packages and BIOVIA Discovery version 4.5 respectively.

2.4.3. Molecular Dynamics (MD) Simulation After Sequential Docking

YASARA dynamics program was used to conduct molecular dynamics simulation where AMBER14 force field was considered for all calculations. As TPO is a membrane bound

protein, a membrane was attached with the dimer structures during simulation process. MD simulation was performed for 5000 ps at 310K having a time step of 1.25 fs and the simulation snapshots were saved at every 100 ps. To observe whether the ligands H_2O_2 and I^- stayed in the same binding pocket after 5000 ps of simulation, the H_2O_2 and I^- ligands were retrieved from the simulated ligand-protein complexes and molecular docking was performed again using Autodock Vina protocol by PyRx version 0.8 software. The binding affinities and corresponding position of ligands were analyzed from the re-docked structures.

2.5. Setting Up and Validation of High Resolution Melt (HRM) Curve Analysis

For analysis of TPO gene mutations by HRM method, we targeted 3 common non-synonymous mutations in the TPO gene identified by Sanger sequencing. For this purpose, we designed 3 sets of primers for detecting c.1117G>T and c.1193G>C mutations in Exon 8 and c.2173A>C mutation in Exon 12. Firstly, 20 pre-sequenced specimens with known mutations were used to set up the HRM method. Homozygous, heterozygous and wild type specimens for the specific mutations were subjected to HRM curve analysis. Finally, 16 unknown samples were run to validate the method (Supplementary Table 4 and 5). These 16 samples were further tested by Sanger sequencing to confirm the mutation and validate the HRM approach. Table 2.5 shows the primers which were designed for HRM study.

Table 2. 5. List of primers used in HRM curve analysis

Primer name	**Mutation (nucleotide position)**	**Primers sequences** (5'-3')	**Product size (base pair)**
TPO_G1117T_Ex8	c.1117G>T	CGCCTACCTGCCCTTCGTGC CGTCTCCGGCCAGGAAGCAG	101
TPO_G1193C_Ex8	c.1193G>C	CTGCTTCCTGGCCGGAGACG ACAGCGTGTGCAGTGCCGTCAG	65
TPO_A2173C_Ex12	c.2173A>C	AGACTTTGAGTCTTGTGACAGC GTGAGAGGAGACCGAACTTCACC	90

To amplify the target sequence, master mix was prepared following the protocol provided with Precision Melt Supermix kit (Bio-Rad). To prepare a reaction mixture of 5 µL of 2X precision melt super mix, 0.2 µL of each forward and reverse primer and 1 µL of DNA (50ng); 3.6 µL

nuclease free water were added to make the reaction volume up to 10µL. Moreover, for detection of c.1193G>C mutation by HRM, 8mM MgCl2 was added to the reaction mixture and the reaction volume was adjusted accordingly. The cyclic condition was divided into 2 steps, namely real time PCR amplification followed by melt curve analysis using a single program. The real time PCR-based HRM was performed on a CFX96 Touch™ Real-Time PCR machine (Bio-Rad). The real time PCR cyclic condition was as follows: initial denaturation at 95°C for 3 min; 40 cycles of denaturation at 95°C for 10 s, annealing at 60 °C for 15 s and extension at 72°C for 15 s. After completion of the real time PCR, the subsequent melt curve program was initiated through cycles of denaturation at 95°C for 30 s, renaturation at 60°C for 1 min, and then melting at 65°C to 95°C with an increment of 0.1°c per 5 s for c.1117G>T, c.2173A>C mutation. And notably, an increment of 0.2°c per 5 s was used for c.1193G>C mutation. After completion of the real time PCR-HRM, the data were analyzed using a Precision Melt Analysis™ Software (BioRad).

The melt curve shape sensitivity for cluster detection was set to 100%. The difference in Tm threshold for the cluster detection was set to 0.1 to 0.2 and the normalized and temperature-shifted views were used for analysis. The results of pre-sequenced samples were compared with HRM analysis and additional 16 unknown samples were subjected to validate the HRM method using the same procedure that was followed for the pre-sequenced known samples. Finally, the normalized melt curve and the difference curves for both wild type and mutant specimens (homozygous and heterozygous) were calculated and analyzed to detect the mutations in the TPO gene.

2.6. Effect of Mutation on Predicted 3D Structure of Transmembrane Region of TSHR Protein

After performing the Sanger Sequencing technique, we detected 2 mutations in Transmembrane (TM)-region of TSHR protein. Since the small molecules drug binds to the TM region of TSHR protein, we targeted this region to see the effect of mutations in that particular site. TSHR protein is composed of a total of 764 amino acids and the TM region belongs to 368 to 764 amino acids of the full length TSHR protein. We used i-TASSER tool to predict the 3D structures of TM-region namely $TSHR_{368-764}$ by following the procedures as mentioned earlier for TPO protein. We selected two promising small molecules ligands

(MS437 and MS 438) which act as agonist to TSHR protein. The optimized structure of MS437 and MS438 were depicted in Supplementary Figure 5. Finally, the molecular docking study was performed using the same tools mentioned earlier for TPO study. Non bond interactions were also observed both for of MS437 and MS438 molecules. Figure 2.3 describes the protocol followed for TSHR study.

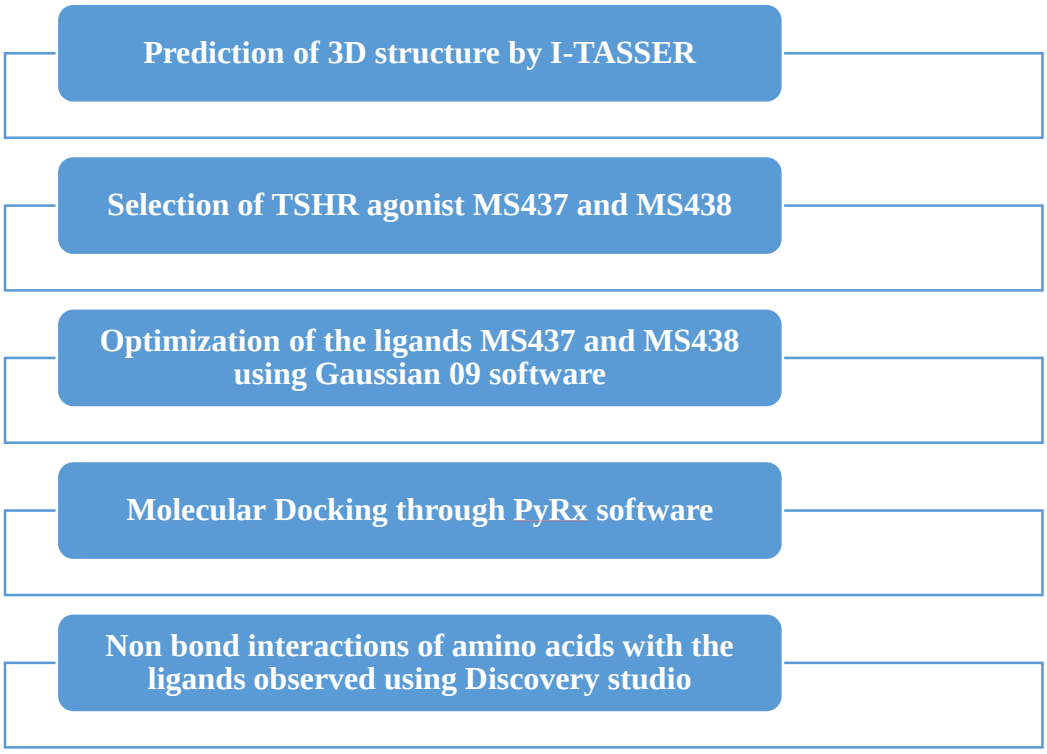


Figure 2. 3. Analysis of the effect of mutation on the 3D structure of TSHR protein

2.7. Profiling of Acylcarnitines and Amino Acids using LC-MS/MS

2.7.1. HPLC and MS/MS Settings

A total of 28 acylcarnitines and 15 Amino acids were measured in this study. Quantification of the levels were done by using ABSCIEX-3200MD-QTRAP and Shimadzu LCMS-8050 mass spectrometer (Shimadzu Corporation, Kyoto, Japan) equipped with electrospray ionization interface in the positive ion mode and multiple reaction monitoring (MRM) system. The overall settings for data acquisition were as follows: The interface voltage at 4.5 kV, interface temperature at 250°C, block temperature at 250°C, heat block temperature at 400°C,

nebulizing gas flow at 3.0 L/min, drying gas flow at 15.0 L/min, collision gas (argon) pressure at 230 kPa and a well time of 5 msec.

2.7.2. Validation of Method

Before testing the samples of patients, the LC-MS/MS method validation was performed for both of ABSCIEX-3200MD-QTRAP and Shimadzu-8050 LC-MS/MS machines. The LC-MS/MS method validation was performed using three different levels of quality control DBS spots, namely low-, medium- and high- level controls that were provided with the NeoMass AAAC kit (Lab systems Diagnostics Oy, Vantaa, Finland). Method validation included inter-assay and intra-assay accuracy, precision, linearity and recovery.

2.7.3. Preparation of Specimens from Dried Blood Spots

Extraction and quantification of acylcarnitines from the DBS cards were done using a NeoMass AAAC kit (Labsystems Diagnostics Oy, Vantaa, Finland) according to the manufacturer's instructions. Briefly, the lyophilized internal standards were reconstituted according to the manufacturer's specifications and the daily working solution was prepared by diluting the reconstituted internal standards 1:100 (v/v) using the extraction solution. A blood spot with a diameter of 3.2 mm was punched into a U-bottomed microplate well and 100 µL extraction solution with internal standard was added. The 96-well microtiter plate was then covered with a piece of adhesive film and incubated for 20 min at room temperature with shaking speed of 650 rpm. After incubation time was over, 70 µL of the content from each well was transferred to a V-bottomed analysis plate. The plate was covered with aluminum foil to minimize evaporation of the solution. Finally, the analysis plate was placed to the auto-sampler of the instrument and 1 µL sample was injected to perform the analysis.

2.7.4. Data Collection and Data Processing

The settings for data acquisition and data processing has been performed using analysis software MD analyst and ChemoView. The concentration of each analyte was measured using the formula: Concentration of an analyte (nmol/mL) = the area of the analyte × the concentration of the internal standard/ the area of the internal standard. Figure 2.4 describes the overall process of LC-MS/MS technique.

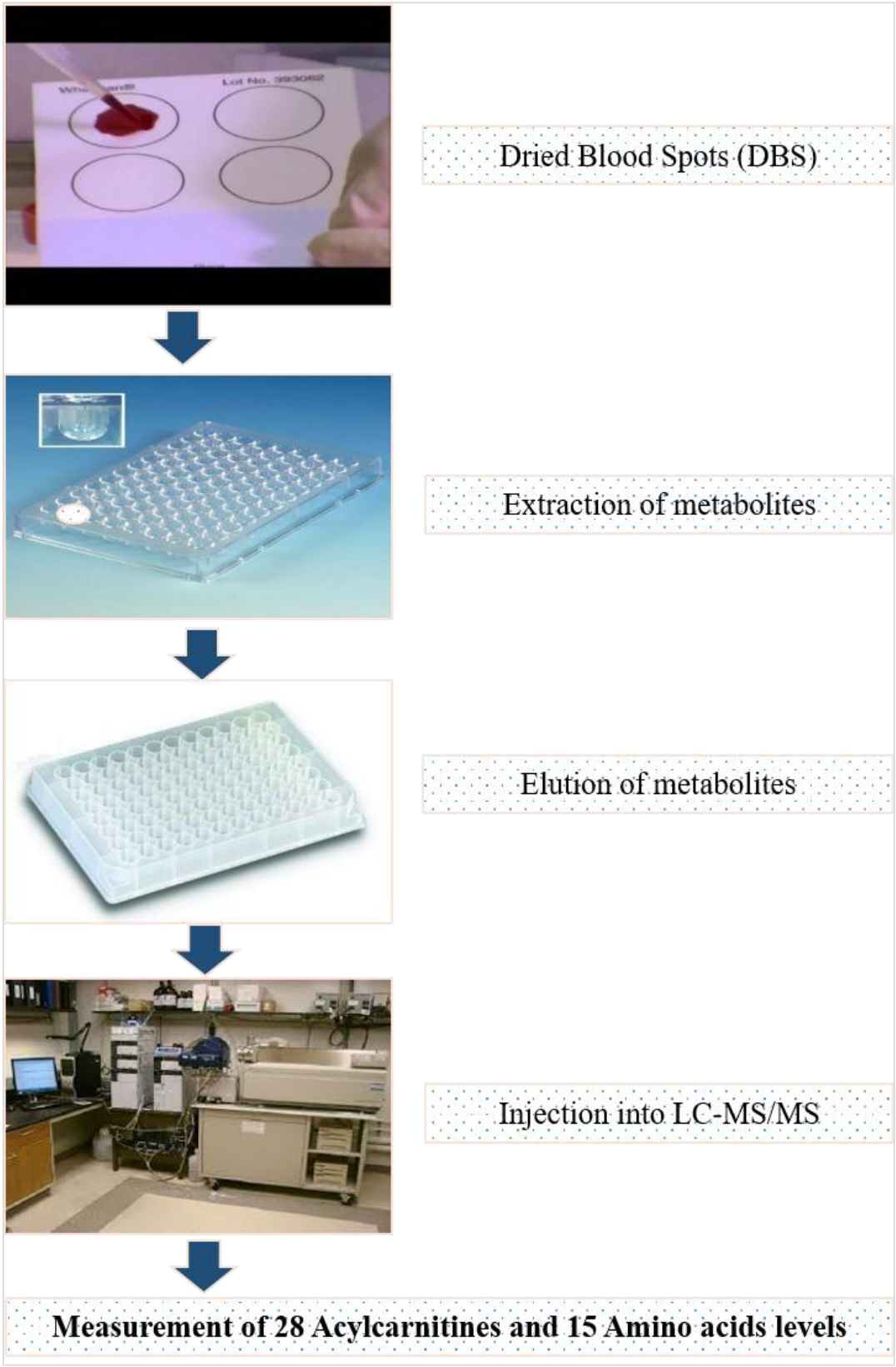


Figure 2. 4. Outline of metabolic profiling

2.7.5. Measurement of Plasma Triglycerides (TG)

The plasma triglycerides (TG) levels of both patients and healthy controls were quantified by enzymatic colorimetric assay using Randox Triglycerides assay kit (Randox Laboratories Ltd., Crumlin, UK). The test was performed following the instructions of the manual of the kit. The absorbance of the sample was read on a DR-7000D semi-automatic analyzer (Dirui, Jilin, China).

2.7.6. Analysis of Gene Expression Level of CPT1 (Carnitine Palmitoyltransferase1)

From Trizol containing blood samples, total RNA was isolated and the RNA from each eluate was used in separate reverse transcription reactions using a real time PCR machine (BIORAD, USA). RNA concentrations were quantified and triplicate of each sample was compared with a housekeeping reference gene GAPDH. CPT1A (Carnitine Palmitoyltransferase1-A) and CPT1B (Carnitine Palmitoyltransferase1-B) gene expression profiles were compared with the control gene. The average fold change of gene expression was calculated using the formula, ΔCt = Ct[Target]-Ct[Housekeeping] and $\Delta\Delta Ct$ = (ΔExp.)-(ΔControl) and got the -$\Delta\Delta Ct$ log-fold-change.

2.7.7. Quantification of Plasma Albumin

The plasma Albumin levels of both patients and healthy controls were measured by enzymatic colorimetric assay using Randox Albumin assay kit (Randox Laboratories Ltd., Crumlin, UK). The test was performed following the procedure provided in the manual of the kit. 3 µL of plasma samples was added to 1000 µL of diluted BCG (bromocresol green, BCG) solution and incubated for 20 min at room temperature (between 20 - 25°C). Plasma albumin bound to BCG and the absorbance of the Albumin-BCG complex was measured at 630 nm in a DR-7000D semi-automatic analyzer (Dirui, Jilin, China). The concentration of Albumin in plasma was measured by this equation: Conc. of Albumin in plasma sample = (Abs. of plasma sample/Abs. of Standard)*Conc. of standard and the measured unit of albumin was g/dL.

2.7.8. Statistical Analysis

Mean, standard deviation (SD), coefficient of variance (CV), interquartile range (IQR) and SEM (Standard Error of Mean) were calculated using Microsoft Excel (2016). Unpaired t-tests with Welch's correction for patients and healthy controls were performed using GraphPad Prism 7.0 software. A P-value <0.05 was considered statistically significant.

CHAPTER 3: RESULTS

3.1. Baseline Information of the Study Participants

Among 65 patients, 44 participants were diagnosed with thyroid Dyshormonogenesis and 21 was diagnosed as thyroid Dysgenesis. Table 3.1 summarizes the baseline information of all study participants including patients and healthy controls. The enrolled hypothyroid patients (N=65, 30 males; 35 females) had an average age of 7.97 ± 4.29 years and a body mass index (BMI) of 17.0±4.4 kg/m^2. The healthy subjects (N=107, 54 males & 53 females) had the average age of 7.24 ± 4.27 years and a BMI of 15.78±2.62 kg/m^2. As shown in Table 3.1, the enrollment was age- and sex-matched because there were no significant differences in age and gender distribution between the patients and the healthy control groups. The mean BMI, too, did not differ between the two groups.

Table 3. 1. Baseline information of the hypothyroid patients and healthy participants

Parameters	Patients	Healthy Controls
Male/ Female	30/35	54/53
Age (years) mean±SD	7.97±4.29	7.24±4.27
BMI (kg/m^2) mean±SD	17.0±4.4	15.78±2.62
Serum TSH (mIU/L) before treatment initiation, mean±SD	66.07±57.38	NA
Serum TSH (mIU/L) during treatment, mean±SD	3.61±4.41	NA
*Serum FT_4 (pmol/L) during treatment, mean±SD	17.68±5.09	NA
Average LT_4 treatment initiation time in months, mean±SD	19.7 ± 4.72	NA

*** FT4 level before treatment initiation was not recorded for majority of the patients. NA= Not Applicable.**

All the hypothyroid patients received thyroid hormone replacement therapy on daily basis and LT_4 dosages were adjusted based on age, sex and BMI of the patients. Dose-adjustment was

also needed for the hypothyroid patients who had clinical manifestations atypical of hypothyroidism. The mean serum TSH level of the patients before initiation of hormone replacement therapy was recorded 66.07±57.38 mIU/L, whereas the baseline TSH level following LT_4 therapy was 3.61±4.41, indicating that there was as much as 18.3-fold higher level of serum TSH before treatment initiation. All the patients enrolled in the study could maintain a normal serum free T_4 (FT_4) level (17.68±5.09 pmol/L) following LT_4 therapy. However, the study could not compare the FT_4 levels of pre-treatment period with that of post-treatment period because the former was not recorded for majority of the patients. All the patients received only LT_4 treatment and none of them had a history of therapy with triiodothyronine (T_3) or combination therapy with LT_4 and T_3.

Table 3. 2. Distribution of hypothyroid patients among different types of hypothyroidism

Type of thyroid disorder (N=65)	No. of Patients
Dyshormonogenesis:	44
Dysgenesis: Agenesis	15
Ectopic	1
Hypoplasia	5

Among 65 enrolled hypothyroid patients, the majority were suffering from Dyshormonogenesis (67.69%, n=44), whereas about 32.30% (n=21) had thyroid dysgenesis with variable degree of thyroid anatomy defects including thyroid gland agenesis (23.07%), ectopic (1.5%) and hypoplasia of thyroid gland (7.6%), respectively (Table 3.2). However, all the patients could persistently maintain the normal TSH and FT_4 levels upon hormone replacement therapy, although they did have various complications due to late diagnosis and delayed treatment initiation (19.7 ± 4.72 months) (Table 3.1).

Among the hypothyroid patients, 39 (60%) had delayed developmental milestone which could be due to delayed diagnosis and treatment initiation. In addition, 22 (33.84%), 20 (30.76%), 7 (10.76%), 6 (9.23%) and 9 (13.84%) patients had dry or rough skin, ill-looking or lethargic, thyromegaly or enlarged thyroid gland, constipation and other complications like decreased activity, increased sleepiness or poor school performance, respectively (Table 3.3).

Table 3. 3. Clinical complications manifested by hypothyroid patients.

Clinical complications	No. (No. %)
Delayed developmental milestone	39 (60%)
Dry/Rough skin	22 (33.84%)
Ill looking/Lethargic	20 (30.76%)
Thyromegaly/Enlarged thyroid gland	7 (10.76%)
Constipation	6 (9.23%)
Other complications[a]	9 (13.84%)

[a] Decreased activity, increased sleepiness, Poor school performance)

3.2. Identification of Mutation in TPO Gene

Since mutations in the TPO gene are commonly associated with thyroid Dyshormonogenesis and related complications, we analyzed the TPO gene to find whether there were mutations in this gene of the study participants [27,99]. First, we performed Sanger sequencing of specimens from the patients targeting the hotspot region from Exon-8 to Exon-14 which are commonly reported in TPO-associated thyroid Dyshormonogenesis and mutations were detected in all 44 samples. After investigating the hotspot region, we also performed the detection of mutations in other regions of the TPO gene. Figure 3.1 and Figure 3.2 show a representative chromatogram and BLAST result of specific mutations.

A total of 12 mutations detected in TPO gene are summarized in Table 3.4. Among these mutations, four were commonly identified in the hotspot region of TPO gene of the study participants. A total of 18 patients had mutations in Exon 8, seven patients had mutations in Exon 12 and twenty-two patients had mutations in both Exon 8 and Exon 12. Since almost all of the patients had mutations in the hotspot region that included Exon 8 and Exon 12, we further tried to investigate the effect of selective 3 non-synonymous mutations namely c.1117G>T (p.Ala373Ser), c.1193G>C (p.Ser398Thr) and c.2173A>C (p.Thr725Pro) on 3D structures of the TPO protein which were previously reported in the patients with thyroid Dyshormonogenesis and the reaction kinetics catalyzed by the mutant TPO enzyme proved to be similar with non-enzymatic reaction rates by several other studies [29,31,100,101].

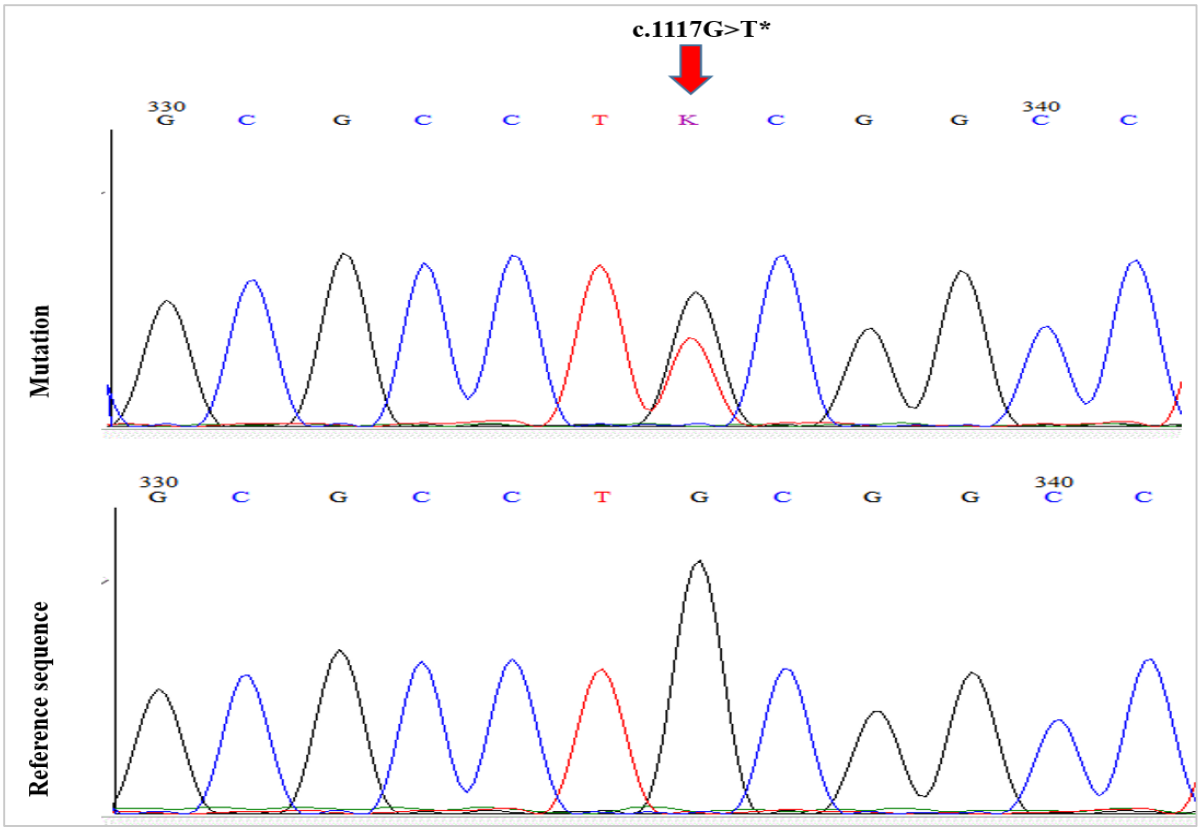


Figure 3. 1. A representative chromatogram showing a heterozygous mutation in Exon 8 of TPO gene

Homo sapiens thyroid peroxidase (TPO), RefSeqGene on chromosome 2

Sequence ID: NG_011581.1 Length: 136265 Number of Matches: 1

Range 1: 68656 to 69098 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
808 bits(437)	0.0	441/443(99%)	0/443(0%)	Plus/Plus

```
Query  1      CGCCTGTCTGCCCTTCTACCGCTCTTCGGCCGCCTGCGGCACCGGGGACCAAGGCGCGCT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  68656  CGCCTGTCTGCCCTTCTACCGCTCTTCGGCCGCCTGCGGCACCGGGGACCAAGGCGCGCT  68715

Query  61     CTTTGGGAACCTGTCCACGGCCAACCCGCGGCAGCAGATGAACGGGTTGACCTCGTTCCT  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  68716  CTTTGGGAACCTGTCCACGGCCAACCCGCGGCAGCAGATGAACGGGTTGACCTCGTTCCT  68775

Query  121    GGACGCGTCCACCGTGTATGGCAGCTCCCCGGCCCTAGAGAGGCAGCTGCGGAACTGGAC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  68776  GGACGCGTCCACCGTGTATGGCAGCTCCCCGGCCCTAGAGAGGCAGCTGCGGAACTGGAC  68835

Query  181    CAGTGCCGAAGGGCTGCTCCGCGTCCACGCGCGCCTCCGGGACTCCGGCCGCGCCTACCT  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  68836  CAGTGCCGAAGGGCTGCTCCGCGTCCACGCGCGCCTCCGGGACTCCGGCCGCGCCTACCT  68895

Query  241    GCCCTTCGTGCCGCCACGCGCGCCTTCGGCCTGTGCGCCCGAGCCCGGCATCCCCGGAGA  300
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  68896  GCCCTTCGTGCCGCCACGCGCGCCTGCGGCCTGTGCGCCCGAGCCCGGCATCCCCGGAGA  68955

Query  301    GACCCGCGGGCCCTGCTTCCTGGCCGGAGACGGCCGCGCCACCCAGGTCCCCTCCCTGAC  360
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Sbjct  68956  GACCCGCGGGCCCTGCTTCCTGGCCGGAGACGGCCGCGCCAGCCAGGTCCCCTCCCTGAC  69015

Query  361    GGCACTGCACACGCTGTGGCTGCGCGAGCACAACCGCCTGGCCGCGGCGCTCAAGGCCCT  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  69016  GGCACTGCACACGCTGTGGCTGCGCGAGCACAACCGCCTGGCCGCGGCGCTCAAGGCCCT  69075

Query  421    CAATGCGCACTGGAGCGCGGACG  443
              |||||||||||||||||||||||
Sbjct  69076  CAATGCGCACTGGAGCGCGGACG  69098
```

Figure 3. 2. Alignment of BLAST result presenting two mutations in Exon 8 of TPO gene

Table 3. 4. Mutation detection in the TPO gene of hypothyroid patients

Exon No.	Nucleotide position	Amino acid position	Exon No.	Nucleotide position	Amino acid position
2	c.103C>G	p.35Glu>Glu	10	c.1728G>A	p.Ala576Ala
5	c.404C>A	p.Pro135His	11	c.1998C>T	p.Asp666Asp
7	c.860G>T	p.Leu287Leu	11	c.2120G>A	p.Asp707Asp
7	c.769G>T	p.Ala257Ser	12	c.2145C>T	p.Pro715Pro
8	c.1117G>T	p.Ala373Ser	12	c.2173A>C	p.Thr725Pro
8	c.1193G>C	p.Ser398Thr	15	c.2540T>C	p.Val847Ala

3.2.1. Modelling of 3D Structure of Myeloperoxidase (MPO)-Like Domain $TPO_{142-738}$ and Full-Length TPO_{1-933} Protein

The low resolution crystallographic structure of TPO protein is available and the catalytic domain of $TPO_{142-738}$ is similar to human Myeloperoxidase (MPO). Since the commonly identified three non-synonymous mutations in the present study were confined in the MPO-like domain of TPO protein, we explored the effect of mutations on the 3D structure of MPO-like domain ($TPO_{142-738}$) of TPO protein. Additionally, we desired to investigate whether the mutations caused any alteration in the interactions of various amino acid residues in the catalytic pocket of TPO protein with the ligand (Heme). We submitted the amino acid $TPO_{142-738}$ sequence for the wild type (WT), mutants- MT1, MT2 and MT3 to the I-TASSER server and obtained 3D structures with C-scores of 2, 2, 1.99, and 1.99 for $TPO_{142-738}$ WT, $TPO_{142-738}$ MT1, $TPO_{142-738}$ MT2 and $TPO_{142-738}$ MT3, respectively (Table 3.5) (Figure 3.3 and Supplementary Figure 2). To explore the effects of the mutations on the full-length TPO protein structure and functions, we also predicted the 3D structures for TPO_{1-933} WT, TPO_{1-933} MT1, TPO_{1-933} MT2 and TPO_{1-933} MT3 using the I-TASSER server and compared the results with the MPO-like domain. Finally, we obtained 5 models from the server for each of TPO_{1-933} WT, TPO_{1-933} MT1, TPO_{1-933} MT2 and TPO_{1-933} MT3 and we chose the best model by analyzing the organization of the MPO-like domain (residues 142–738); complement control protein (CCP)-like domain (residues 740–795) and Epidermal growth factor (EGF)-like

domain (residues 796–846) if they were in correct arrangement [102-106] (Figure 3.3). The corresponding C-score, TM-score and RMSD score of TPO_{1-933} WT, TPO_{1-933} MT1, TPO_{1-933} MT2 and TPO_{1-933} MT3 were summarized in Table 3.5 as mentioned in published article [31].

Table 3. 5. Features of 3D structures predicted by I-TASSER server.

Features	$TPO_{142-738}$ WT	$TPO_{142-738}$ MT1	$TPO_{142-738}$ MT2	$TPO_{142-738}$ MT3	TPO_{1-933} WT	TPO_{1-933} MT1	TPO_{1-933} MT2	TPO_{1-933} MT3
Model no.	01	01	01	01	02	01	05	01
C–score	2	2	1.99	1.99	–3.35	–3.23	–3.27	–2.9
TM–score	0.99±0.04	0.99±0.04	0.99±0.04	0.99±0.04	–	0.35±0.12	–	0.38±0.13
RMSD (Å)	3.5±2.4	3.4±2.4	3.6±2.5	3.6±2.5	–	17.2±2.7	–	16.2±3.1

C–score=Confidence score range: [-5, 2]; TM–score=Template Modelling score, TM–score < 0.17 indicates random similarity and TM–score > 0.5 indicates correct similarity; RMSD=Root Mean Square Deviation, WT = Wild type, MT1 = Mutant 1 (p.Ala373Ser), MT2 = Mutant 2 (p.Ser398Thr), MT3 = Mutant 3 (p.Thr725Pro).

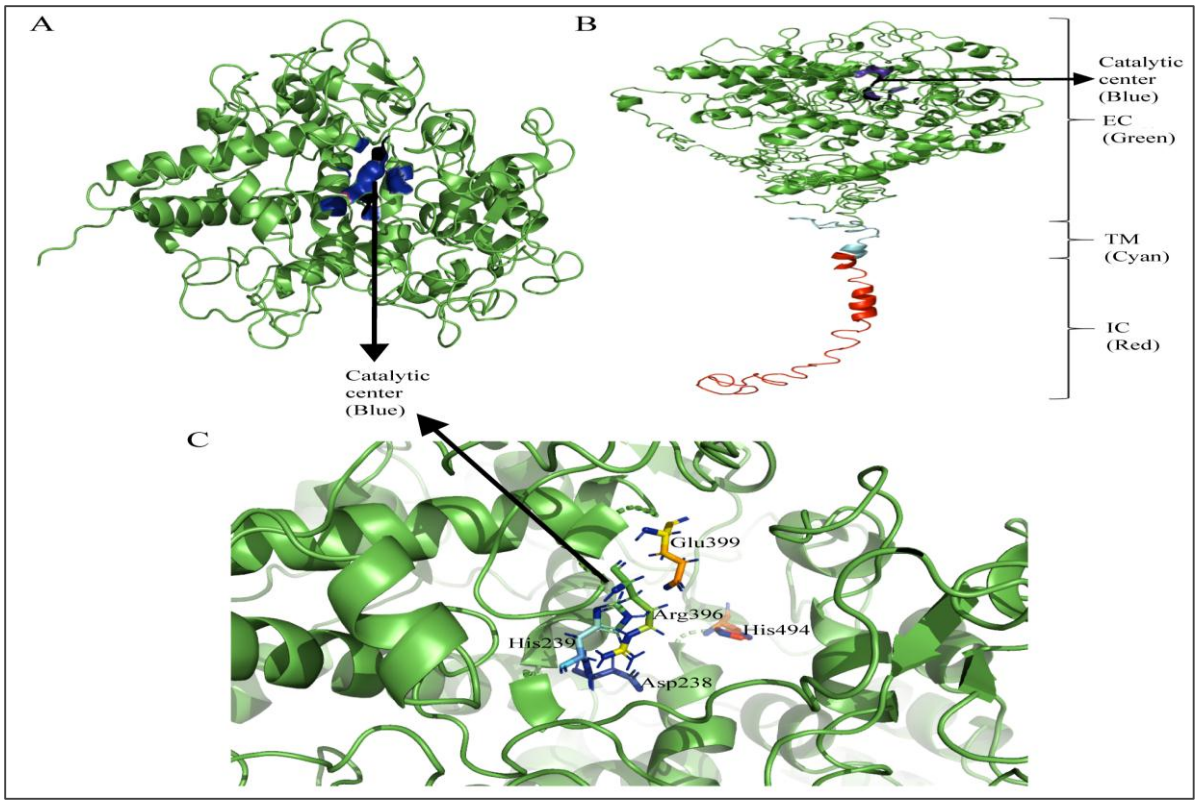


Figure 3. 3. The representative predicted 3D structures of the TPO proteins. (A) TPO142-738 WT, (B) TPO1-933 WT, (C) Catalytic site of TPO with crucial amino acids Asp238, His239, Arg396, Glu399 and His494. The specific regions of the TPO1-933 WT structure are indicated with the corresponding color written in the brackets. EC=Extracellular region shown in green color, TM=Transmembrane region shown in cyan color, IC=Intracellular region shown in red color.

3.2.2. Validation of the 3D Structures of MPO Like Domain and Full-Length TPO protein

The 3D structures of TPO WT, MT1, MT2 and MT3 proteins were validated by Verify3D server which measures the compatibility of the predicted 3D structures. In verify3D, more than 80% amino acid residues had the average 3D-1D score of ≥ 0.2 which confers the validity of the 3D structures of $TPO_{142-738}$ WT (96.15%), $TPO_{142-738}$ MT1 (89.78%), $TPO_{142-738}$ MT2 (94.14%) and $TPO_{142-738}$ MT3 (94.64%) (Table 3.6). In addition, validation of the structures by the RAMPAGE web server also provided the percentages of the amino acid residues within favored, allowed and outlier regions. In case of $TPO_{142-738}$ WT, 84.5% residues were within the favored region, 11.6% were within the allowed region and 3.9% were within the outlier region. For $TPO_{142-738}$ MT1, 84.4% residues were in the favored regions, 11.1% in the allowed region and 4.4% in the outlier regions.

Differently, for $TPO_{142-738}$ MT2, 85.2% residues were confined in the favored regions, 10.4% in the allowed region and 4.4% in the outlier regions and on the other hand, for $TPO_{142-738}$ MT3, 81.2% residues were confined in the favored region, 14.3% in the allowed region and 4.5% in the outlier regions (Table 3.6). The Verify3D results revealed that the percentages of amino acid residues having average 3D-1D score of ≥ 0.2 for TPO_{1-933} WT, TPO_{1-933} MT1, TPO_{1-933} MT2 and TPO_{1-933} MT3 were 73.10%, 70.74%, 76.10% and 76.63%, respectively. The results provided by RAMPAGE server for TPO_{1-933} WT, TPO_{1-933} MT1, TPO_{1-933} MT2 and TPO_{1-933} MT3 were summarized in Table 3.6.

Table 3. 6. 3D Structure Validation Data Obtained from Verify3D and RAMPAGE Webserver

	Tools	$TPO_{142-738}$ WT	$TPO_{142-738}$ MT1	$TPO_{142-738}$ MT2	$TPO_{142-738}$ MT3	TPO_{1-933} WT	TPO_{1-933} MT1	TPO_{1-933} MT2	TPO_{1-933} MT3
	Verify3D	96.15%	89.78%	94.14%	94.64%	73.10%	70.74%	76.10%	76.63%
RAMPAGE	Favored region	84.5%	84.4%	85.2%	81.2%	68.40%	69.60%	66.90%	69.80%
RAMPAGE	Allowed region	11.6%	11.1%	10.4%	14.3%	17.20%	16.20%	18.40%	18.80%
RAMPAGE	Outlier region	3.9%	4.4%	4.4%	4.5%	14.40%	14.20%	14.70%	11.40%

Verify3D: Percentages of amino acids having the average 3D-1D score ≥ 0.2; RAMPAGE: Percentages of the amino acid residues within the favored, allowed and outlier regions. WT = Wild type, MT1 = Mutant 1 (p.Ala373Ser), MT2 = Mutant 2 (p.Ser398Thr), MT3 = Mutant 3 (p.Thr725Pro).

3.2.3. Optimization of Heme, Hydrogen Peroxide (H_2O_2) and Iodide (I^-)

Density functional theory using B3LYP/SDD was employed for the optimization of the Heme prosthetic group, H_2O_2 and I^- using Gaussian 9 software. Slight changes in bond distances and bond angles were observed between crystal and DFT structures mentioned in Supplementary Figure 1 and 4.

3.2.4. Comparison of Molecular Docking Results Between Heme and 3D Structures of MPO like domain and full-length TPO protein (Wild Type and Mutant TPO Cases)

Molecular docking analysis showed that Heme interacted with TPO_{1-933} WT through a total of 21 non-bond interactions including Arg491 and Arg582 through hydrogen bonds, interactions with His239, Val400, Phe490, Arg491, His494, Ile497, Phe523, Leu560, Leu564, and Leu575 through hydrophobic bond and interactions with Arg396, Glu399, and Arg491 through electrostatic bond (Table 3.7 and Figure 3.4). On the other hand, 21 non-bond interactions of amino acids with Heme were also observed for MPO-like domain ($TPO_{142-738}$ WT) of wild type TPO protein (Figure 3.5). However, not all of these 21 interactions for $TPO_{142-738}$ and TPO_{1-933} were common. The $TPO_{142-738}$ WT interactions with Heme included Arg582, and Arg586 through hydrogen bonds, His239, Val400, His494, Ile497, Phe523, Leu560, Leu564, Val566, Leu567, and Leu575 through hydrophobic bonds and Arg396, and Glu399 through electrostatic bonds. The common 11 interactions for TPO_{1-933} WT and $TPO_{142-738}$ WT included His239, Arg396, Glu399, Val400, His494, Ile497, Phe523, Leu560, Leu564, Leu575, and Arg582. Furthermore, of the aforementioned 11 residues, 4 residues including His239, Arg396, Glu399 and His494 had been reported to be crucial for enzyme activity presented within the MPO like domain of full length TPO protein [15,16].

A total of 12 residues were found to interact with Heme for TPO_{1-933} MT1 case. Among them, 2 residues namely, His239 and His494 interacted with Heme through hydrogen bonds and 5 residues including His239, Phe243, Arg396, Phe524, and Leu567 interacted through hydrophobic bonds (Table 3.7, Figure 3.4 and Supplementary Table 1 and 2). In case of TPO_{1-933} MT1, total number of interactions decreased significantly compared to the TPO_{1-933} WT and the crucial amino acid residue Glu399 was absent in the TPO_{1-933} MT1 interactions.
Differently, molecular docking results showed that when $TPO_{142-738}$ MT1 predicted structure was docked, a total of 19 residues were found to interact with Heme (Table 3.7 and Figure 3.5)

(Supplementary Table 1 and 2). Among the 19 amino acid residues, 4 residues including Met231, Gly234, Ser402, and Gly493 interacted with Heme through hydrogen bonds, the residues Gly234, Gln235, Val400, Phe490, Arg491, Ile497, Phe523, Leu560, Leu564 and Leu575 interacted through hydrophobic bonds and the residue Glu399 interacted through electrostatic bond with the Heme ligand (Table 3.7). Nevertheless, the crucial interactions of Heme with His239, Arg396 and His494 were absent in the mutant case $TPO_{142-738}$ MT1. The 3D structure-based molecular docking data for both of TPO_{1-933} and $TPO_{142-738}$ suggested that MT1 mutation was damaging for the TPO enzyme activity due to the absence of the crucial amino acids in the catalytic pocket of the TPO enzyme.

In case of TPO_{1-933} MT2, Heme ligand was interacted with a total of 20 amino acid residues. Among them, 3 amino acids including His239, Arg491, and Arg582 interacted with Heme through hydrogen bonds, 9 residues including His239, Val400, Arg491, His494, Ile497, Phe523, Leu564, Val566 and Leu575 interacted through hydrophobic bonds and 2 residues including Arg396 and Glu399 interacted through electrostatic bonds (Table 3.7 and Figure 3.4) (Supplementary Table 1 and 2). All the four crucial amino acid residues, namely His239, Arg396, Glu399 and His494 which are important for MPO- like domain activity in TPO enzyme were found to be interacted with Heme in mutant protein TPO_{1-933} MT2. On the other hand, MPO like TPO protein $TPO_{142-738}$ MT2 interacted with 20 residues with Heme through molecular docking study. Among the 20 interacted amino acid residues, 3 residues including Arg491, Asn579 and Arg582 interacted with Heme by hydrogen bonds, the residues Phe243, His494, Phe523, Phe524, Leu564, Phe565, Val566, Leu567 and Leu575 interacted through hydrophobic bonds and the residue Glu399 interacted through electrostatic bond (Table 3.7, Figure 3.5 and Supplementary Table 1 and 2). However, the crucial interactions of Heme with the important residues His239 and Arg396 were absent in the mutant case $TPO_{142-738}$ MT2. The full length TPO_{1-933} structure-based molecular docking analysis revealed that this mutation did not result in any major changes in interactions and all crucial interacting residues were present, while the $TPO_{142-738}$ structure-based docking result showed that almost all the interactions were similar to TPO_{1-933} except that the crucial amino acid residues His239 and Arg396 were absent. In the molecular docking approach, MT2 for both of full length and MPO like domain presented that the results are suggestive of the fact that this mutation might have an association with PIOD.

On the other hand, Heme was interacted with 21 amino acid residues in case of TPO_{1-933} MT3. Among them, 3 residues including Gln246, Arg491 and Ser568 interacted with Heme through hydrogen bonds, 11 residues including Phe243, Val400, Arg491, His494, Ile497, Phe523, Phe524, Leu560, Leu564, Val566 and Leu575 interacted through hydrophobic interactions and the residue Glu399 interacted through electrostatic bond (Table 3.7 and Figure 3.4) (Supplementary Table 1 and 2). Two crucial amino acid residues, namely His239 and Arg396 which are important for TPO enzyme activity were absent in the mutant case TPO_{1-933} MT3. Alternatively, when docking was performed for $TPO_{142-738}$ MT3 predicted structure, a total of 16 residues were found to be interacted with Heme. Among them, 4 residues including His239, Arg396, Arg491 and Arg582 interacted with Heme through hydrogen bonds, the residues Phe243, Phe523, Phe524, Leu564, Phe565, Leu567 and Leu575 interacted through hydrophobic bonds and the residue Glu399 interacted through electrostatic bond (Table 3.7 and Figure 3.5) (Supplementary Table 1 and 2). However, the crucial interactions of Heme with His494 was absent in the $TPO_{142-738}$ MT3. The docking result of Heme with mutant 3 cases for both of TPO_{1-933} and $TPO_{142-738}$ domain suggested that this MT3 might have damaging effect to the TPO enzyme activity.

Table 3. 7. Comparison of Binding energies and non-bond interactions of Heme with MPO domain ($TPO_{142-738}$) of TPO protein and full length TPO protein (TPO_{1-933}) after flexible docking

Protein type	Binding Affinity (kcal/mol)	Hydrogen bond	Hydrophobic bond	Electrostatic bond	Total interactions
$TPO_{142-738}$ WT	–11.9	Arg441 (Arg582), Arg445 (Arg586)	His98(His239), Val259(Val400), His353(His494), Ile356(Ile497), Phe382(Phe523), Leu419(Leu560), Leu423(Leu564), Val425(Val566), Leu426(Leu567), Leu434(Leu575)	Arg255(Arg396), Glu258(Glu399)	21
$TPO_{142-738}$ MT1	–10.8	Met90(Met231), Gly93(Gly234), Ser261(Ser402), Gly352(Gly493)	Gly93(Gly234), Gln94(Gln235), Val259(Val400), Phe349(Phe490), Arg350(Arg491), Ile356(Ile497), Phe382(Phe523), Leu419(Leu560), Leu423(Leu564), Leu434(Leu575)	Glu258(Glu399)	19
$TPO_{142-738}$ MT2	–2.5	Arg350(Arg491), Asn438(Asn579), Arg441(Arg582)	Phe102(Phe243),His353(His494), Phe382(Phe523), Phe383(Phe524), Leu423(Leu564), Phe424(Phe565), Val425(Val566), Leu426(Leu567), Leu434(Leu575)	Glu258(Glu399)	20
$TPO_{142-738}$ MT3	–5.3	His98(His239), Arg255(Arg396), Arg350(Arg491), Arg441(Arg582)	Phe102(Phe243), Phe382(Phe523), Phe383(Phe524), Leu423(Leu564), Phe424(Phe565), Leu426(Leu567), Leu434(Leu575)	Glu258(Glu399)	16
TPO_{1-933} WT	–11.5	Arg491, Arg582	His239, Val400, Phe490, Arg491, His494, Ile497, Phe523, Leu560, Leu564, Leu575	Arg396, Glu399, Arg491	21
TPO_{1-933} MT1	–3.2	His239, His494	His239, Phe243, Arg396, Phe524, Leu567	Not found	12
TPO_{1-933} MT2	–11.5	His239, Arg491, Arg582	His239, Val400, Arg491, His494, Ile497, Phe523, Leu564, Val566, Leu575	Arg396, Glu399	20
TPO_{1-933} MT3	–7.9	Gln246,Arg491, Ser568	Phe243, Val400, Arg491, His494, Ile497, Phe523, Phe524, Leu560, Leu564, Val566, Leu575	Glu399	21

The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding number; in case of $TPO_{142-738}$ the amino acid outside the first bracket indicates the position in predicted structure for $TPO_{142-738}$ and the amino acid residues in first bracket indicates the real position in TPO_{1-933} protein; WT = Wild type; MT1 = Mutant 1 (p.Ala373Ser); MT2 = Mutant 2 (p.Ser398Thr); MT3 = Mutant 3 (p.Thr725Pro).

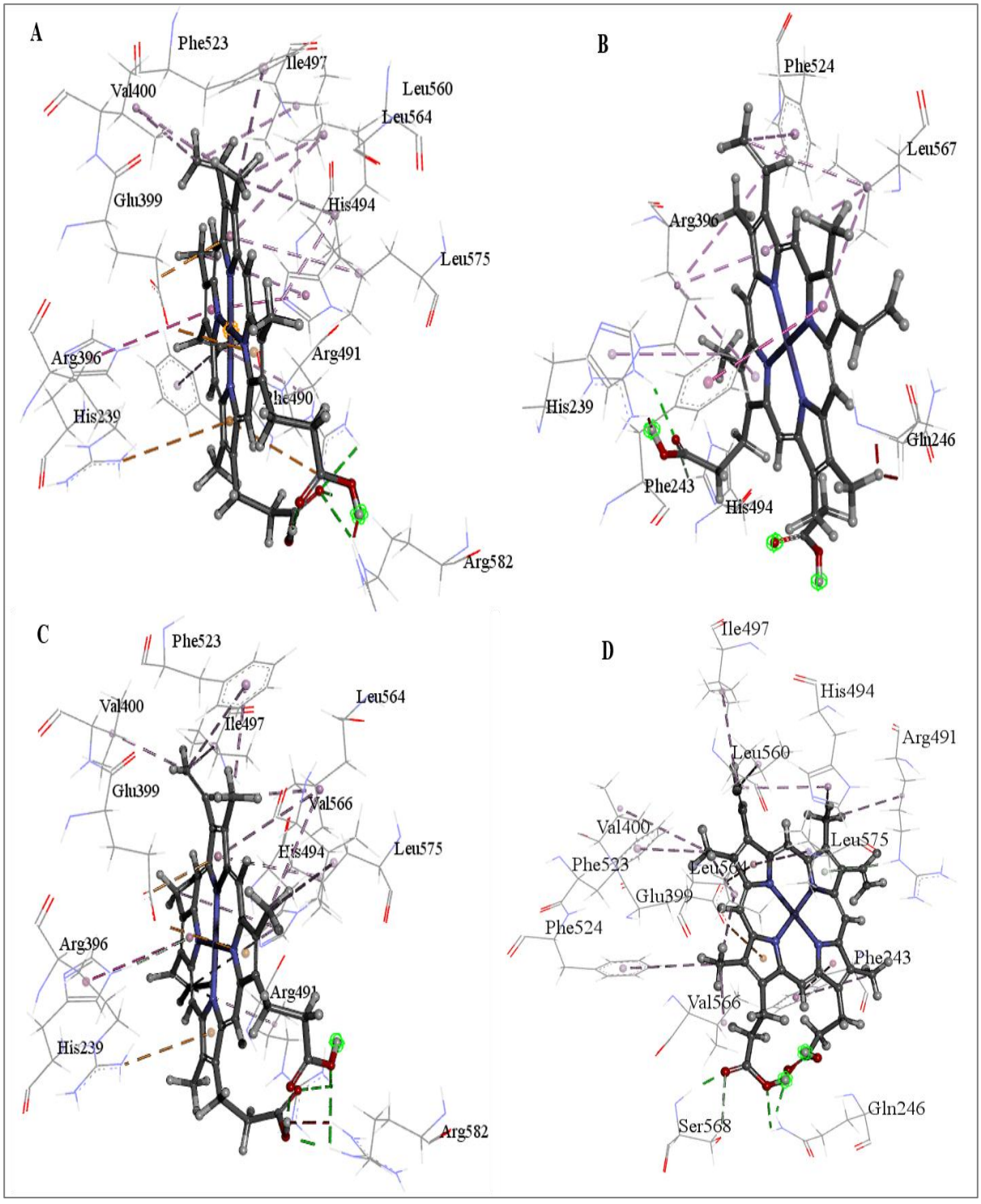


Figure 3. 4. Non-bond interactions of Heme with 3D structures of full length TPO protein using BIOVIA Discovery Studio 2017. (A) TPO1-933 WT, (B) TPO1-933 MT1, (C) TPO1-933 MT2, and (D) TPO1-933 MT3. The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding numbers

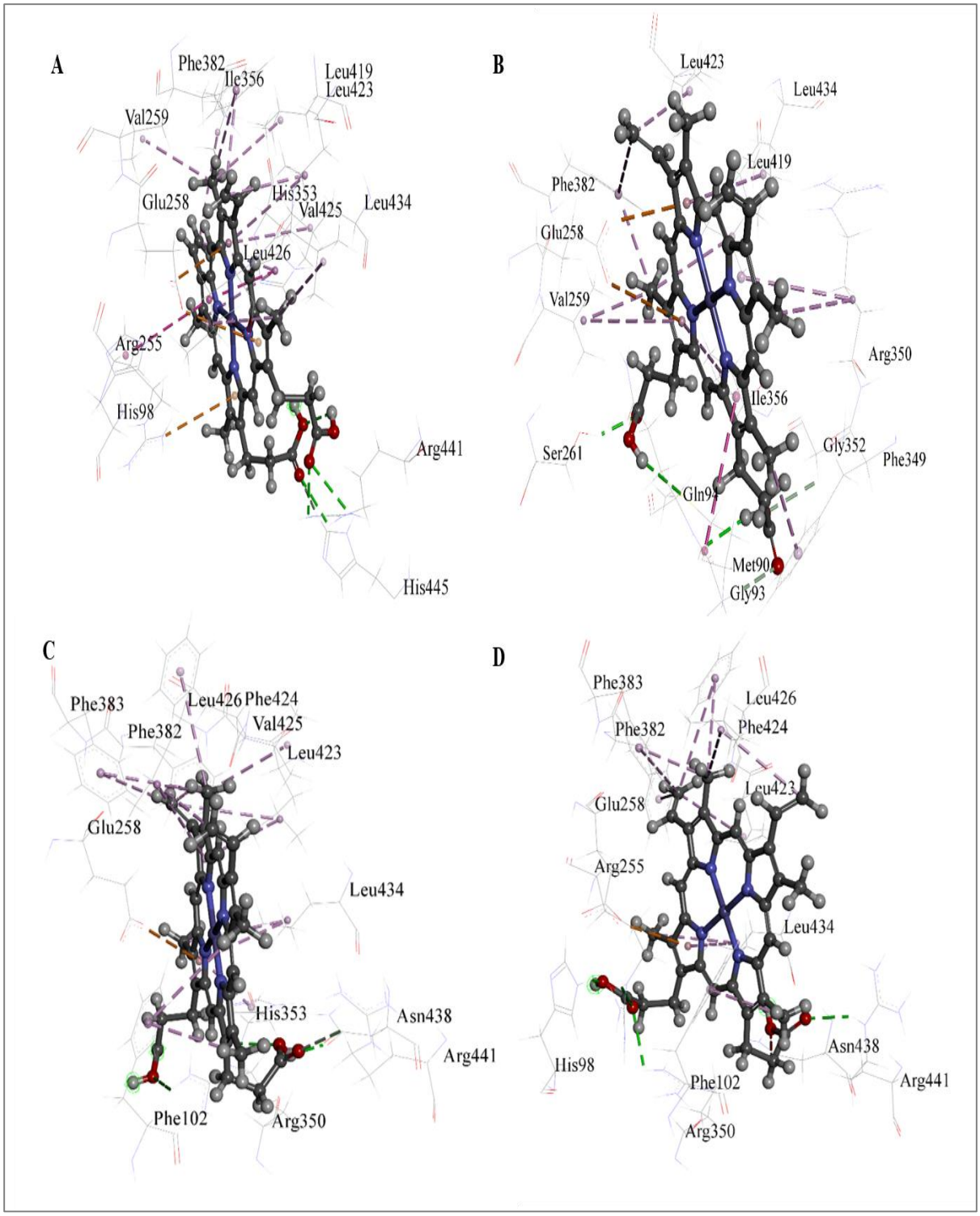


Figure 3. 5. Non-bond interactions of Heme with 3D structures of MPO like domain (A) TPO142-738 WT, (B) TPO142-738 MT1, (C) TPO142-738 MT2, (D) TPO142-738 MT3. The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding numbers.

3.2.5. Validation of Molecular Docking Analysis through QM/MM and Molecular Dynamics (MD) Simulation

To validate the molecular docking results and to compare the amino acid interactions of Heme with the wild type structure and the mutant cases, we performed QM/MM and Molecular Dynamics Simulation. A study by Guria *et. al.* explored that TPO_{1-933} MT1 (p.Ala373Ser) and TPO_{1-933}MT3 (p.Thr725Pro) showed more damaging effect on the catalytic activity of TPO protein [29]. Moreover, the present molecular docking-based study specified that the full length TPO_{1-933} MT2 (p.Ser398Thr) structure interacted to all the crucial amino acids in the catalytic site of TPO, suggesting that this mutation might have less damaging effect on TPO enzyme activity [31]. Hence for further analysis, we selected the docked structures of wild type TPO protein (TPO_{1-933} WT) and mutant cases 1 and 3 namely, TPO_{1-933} MT1 (p.Ala373Ser) and TPO_{1-933}MT3 (p.Thr725Pro) respectively. The interactions between the amino acid residues of TPO_{1-933} WT, TPO_{1-933} MT1 and TPO_{1-933} MT3 proteins and the Heme group was studied by QM/MM and the type of interactions are illustrated in Figure 3.6. It is identified that the interacting residues were same as acquired from the initial docking results.

During molecular dynamics simulation, the protein-ligand complex after 5000 ps of simulation was superimposed on the initial docked protein-ligand complex to examine the structural deviations. The superimposed structures of TPO_{1-933} WT, TPO_{1-933}MT1 (p.Ala373Ser) and TPO_{1-933}MT3 (p.Thr725Pro) proteins are represented in Figure 3.7. From the analysis, it was observed that the Heme ligand was found within the catalytic sites of TPO protein.

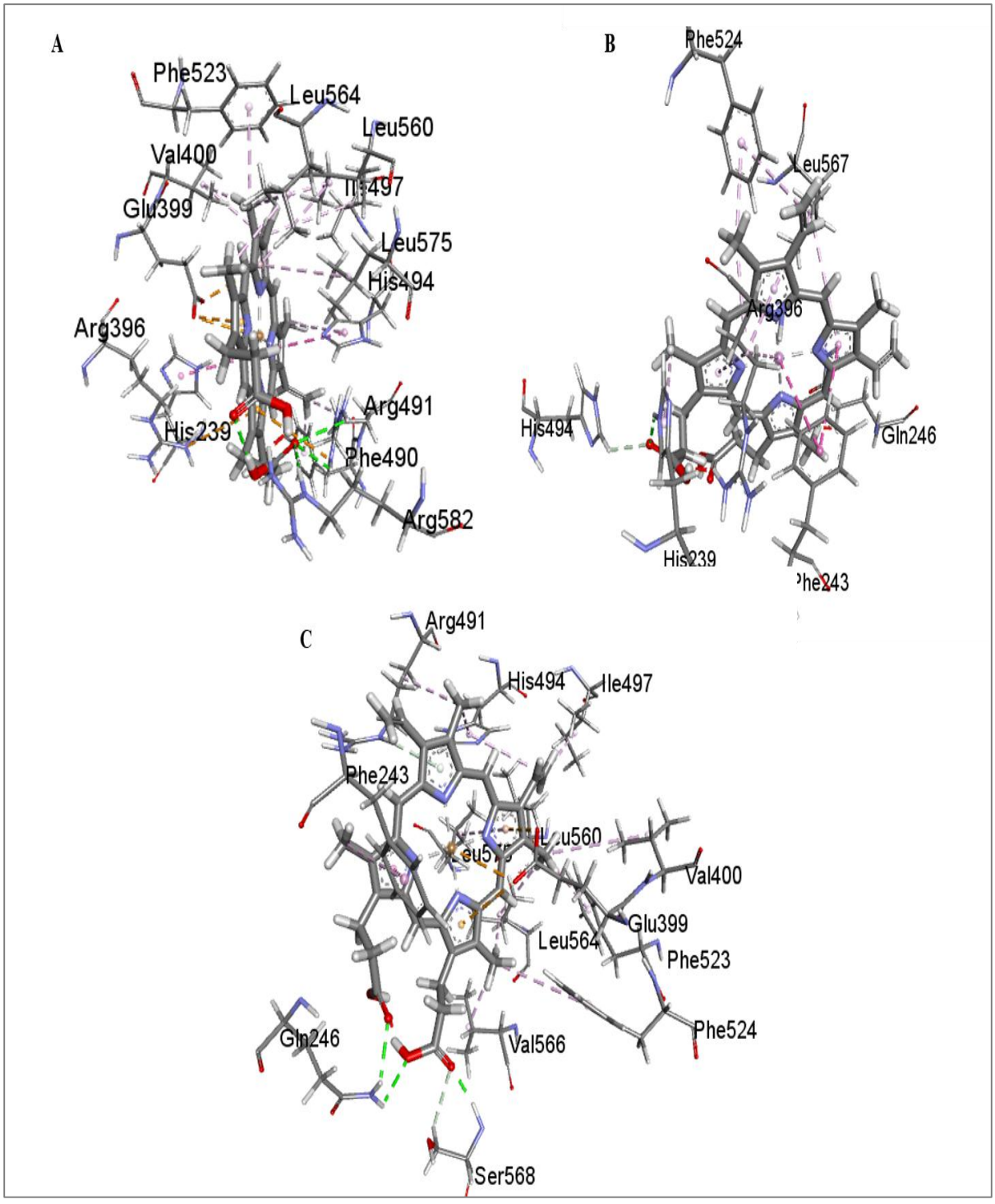


Figure 3. 6. Non-bond interactions of Heme with corresponding predicted structures as obtained using a BIOVIA Discovery Studio 2017, after QM/MM calculations of the structures with Heme molecule in the active site performed at the ONIOM- (PM6: UFF) level of theory. (A) TPO1-933 WT, (B) TPO1-933 MT1, (C) TPO1-933 MT3. The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding numbers.

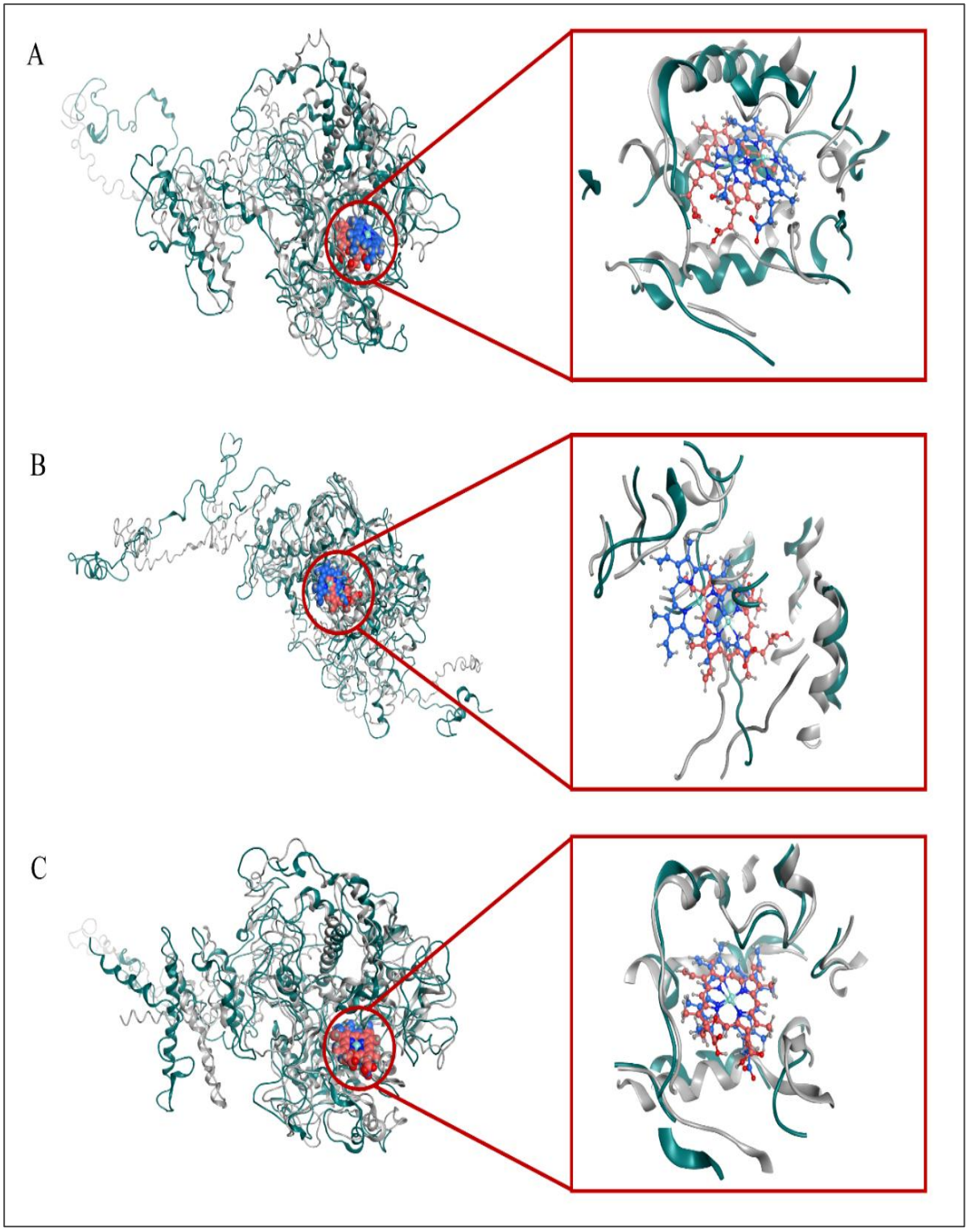


Figure 3. 7. Superimposed structures of protein-ligand (Heme) complexes before and after 5000 ps MD simulation. (A) TPO1-933 WT, (B) TPO1-933 MT1 and(C) TPO1-933 MT3. Grey & Blue colors indicate results= before MD simulation whereas Green & Pink indicate results=after MD simulation.

The retrieved protein structure after 5000 ps simulation was again docked with the Heme ligand since protein flexibility can give rise to difference in binding interaction of a ligand with its target protein. The re-docked results after simulation was depicted in Figure 3.8. The molecular docking analysis after implementing molecular dynamics simulation of the protein structures revealed that Heme interacted with $TPO_{1\text{-}933}$WT through all the crucial amino acids including Asp238, His239, Arg396, Glu399 and His494, whereas it interacted through Glu399 and His494 residues in case of $TPO_{1\text{-}933}$MT1 and $TPO_{1\text{-}933}$MT3 cases. Therefore, there were no interactions with the other 3 crucial amino acid residues, namely Asp238, His239 and Arg396 in the mutant cases 1 and 3. Since these 5 amino acid residues are essential for the catalytic activity of the TPO protein, the lack of interactions with one of these crucial amino acids could affect the structural as well as functional activity of the protein.

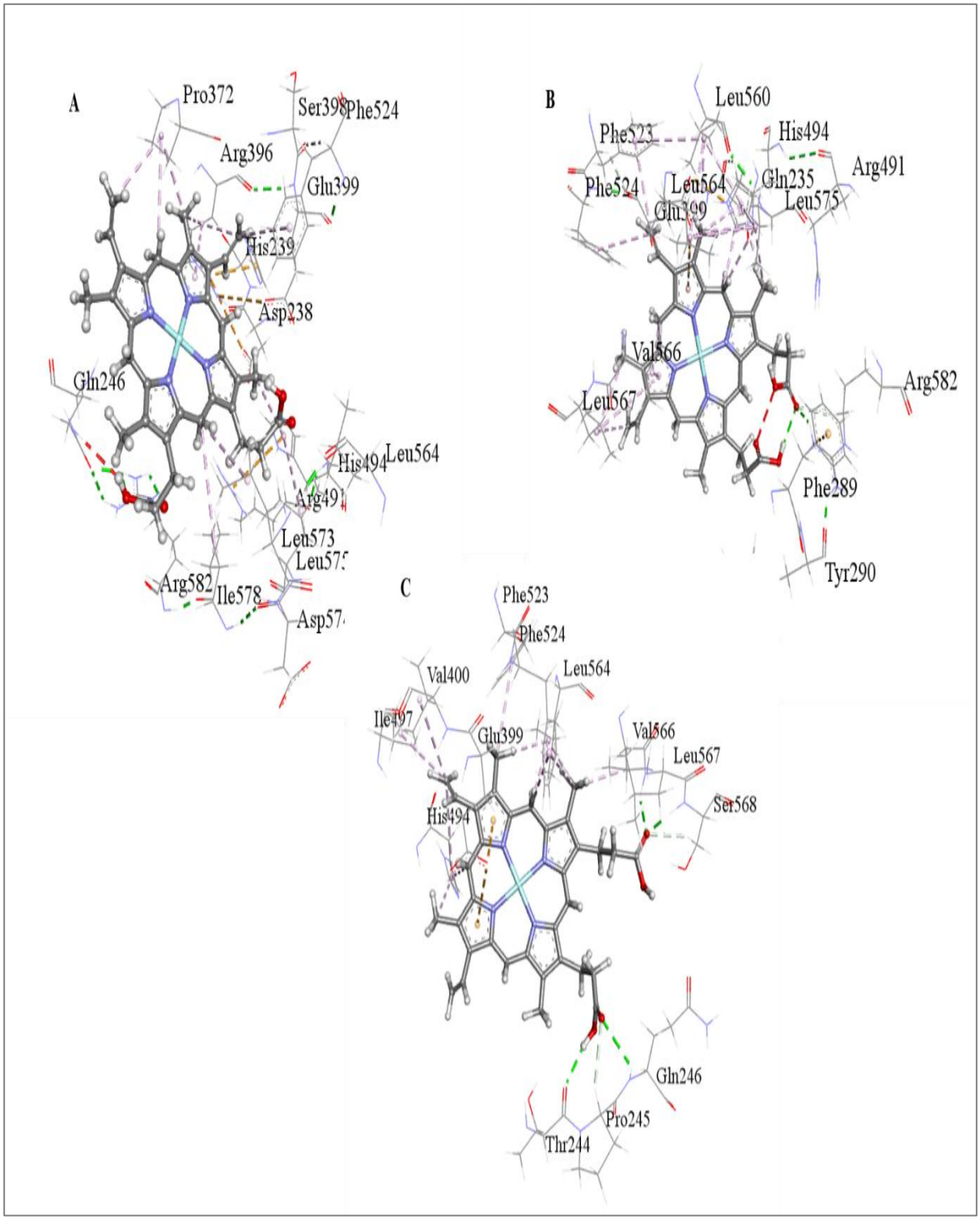


Figure 3. 8. Non-bond interactions of Heme with corresponding predicted structures as obtained after docking with the structures retrieved from the 5000 ps simulation using a BIOVIA Discovery Studio 2017. (A) TPO1-933 WT, (B) TPO1-933 MT1 and (C) TPO1-933 MT3. The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding numbers.

3.3. Modelling of TPO Dimer Structures (TPO WT-HD, TPO MT1-HD & TPO MT3-HD) and MD Simulation

Since TPO exists as a homodimer in biological systems, we wanted to see the effect of mutations (MT1 and MT3) in TPO dimer structures and thus the structures were modelled by SymmDock server based on Score, Area and ACE values (Table 3.8). Then MD Simulation was performed to analyze the van der Waals surface area, solvent accessible surface area (SASA) and molecular surface area for all the dimer structures and found that the wild type dimer (TPO WT-HD) structure was more stable than mutant structures (TPO MT1-HD and TPO MT3-HD) (Figure 3.9).

Table 3. 8. SymmDock server predicted structural comparison for TPO WT HD, TPO MT1 HD, TPO MT3 HD

Proteins	Score	Area	ACE
TPO WT HD	22554	3691.70	-361.26
TPO MT1 HD	21964	2683.40	271.65
TPO MT3 HD	25706	3988.80	-249.80

ACE=Atomic contact energies to global energies

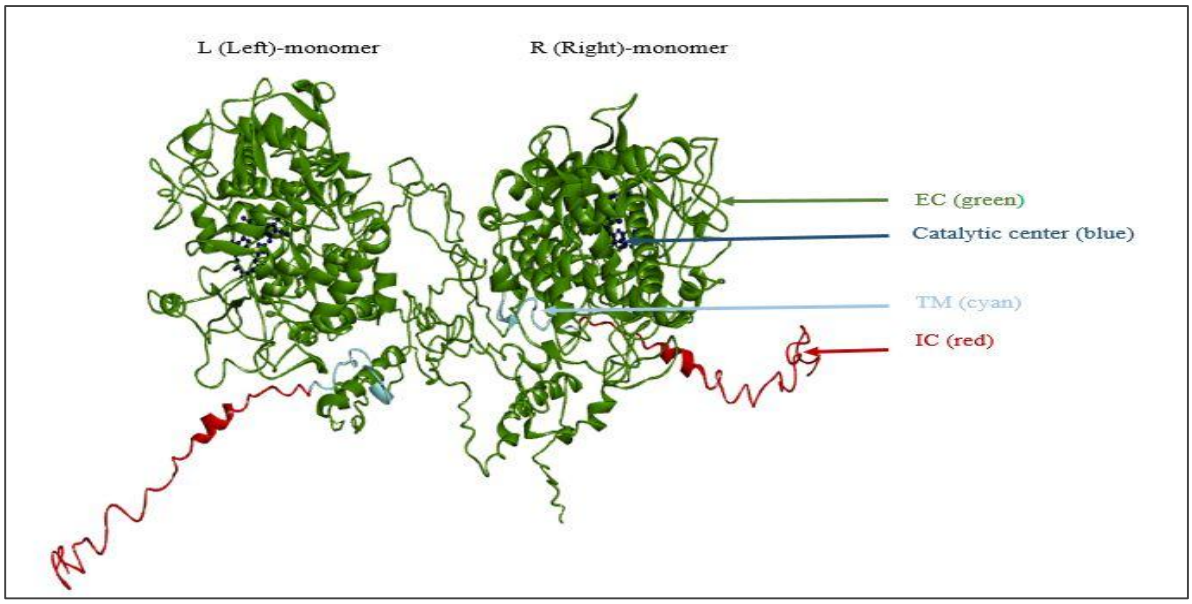


Figure 3. 9. A representative dimer structure of TPO protein predicted by Symmdock server. The specific regions of the structures are indicated with the corresponding color written in the brackets. EC=Extracellular region shown in green color, TM=Transmembrane region shown in cyan color, IC=Intracellular region shown in red color, Catalytic center shown in blue color

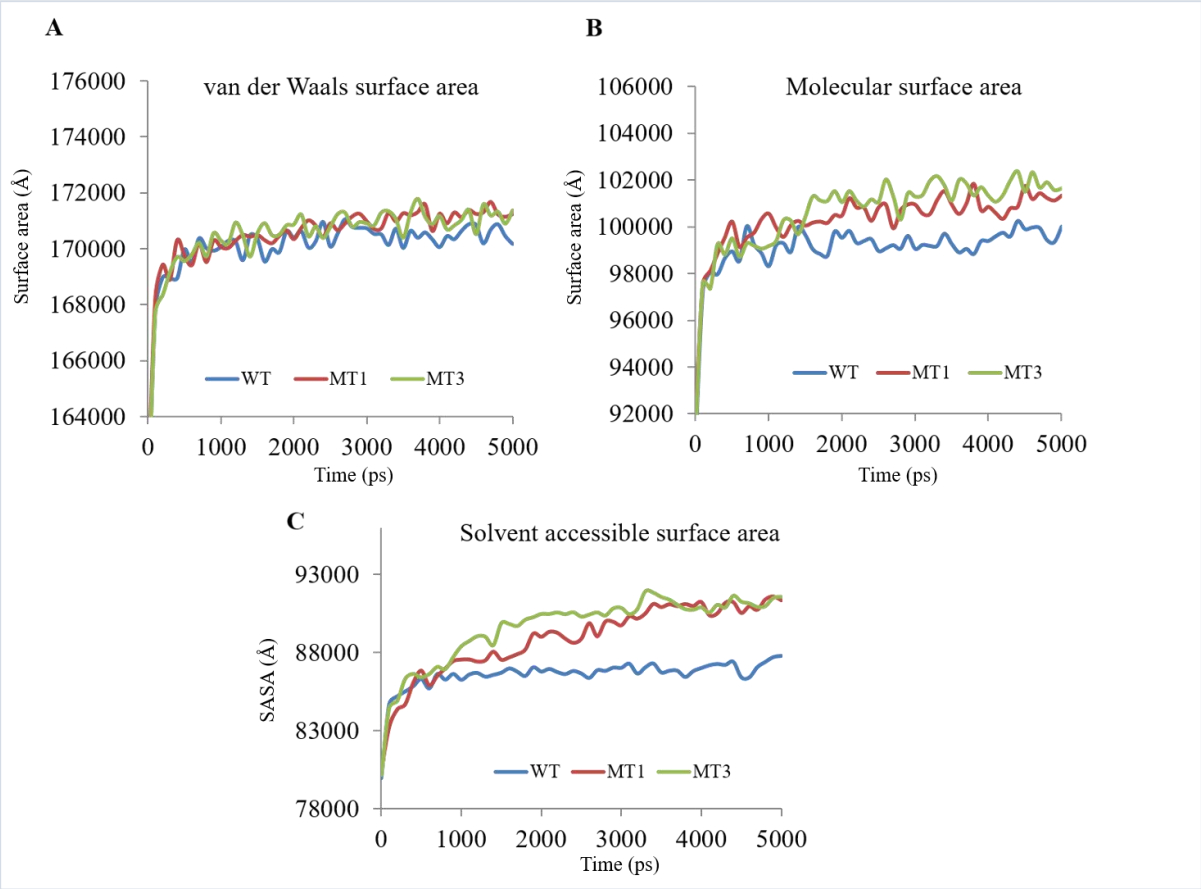


Figure 3. 10. Graphical presentation of (a) van der Waals surface area, (b) Molecular surface area, (c) Solvent accessible surface area of TPO WT HD, TPO MT1 HD, TPO MT3 HD after MD simulation.

Finally, MD simulation analysis showed that the ligand (Heme) binging pocket of TPO dimer structures was consistent with the TPO monomer structures.

3.3.1. Analysis of Molecular Docking Results (Sequential Approach)

3.3.2. Molecular Docking of Hydrogen Peroxide (H_2O_2) and Iodide (I^-) With Predicted 3D Structures of TPO Dimer Proteins (TPOWT-HD, TPOMT1-HD and TPOMT3-HD)

Sequential docking approach suggested that at first H_2O_2 was docked with the Heme bound TPO dimer structures (MD simulated) and then iodide(I^-) was docked on that structures bound with Heme and H_2O_2. The docking procedures were performed individually on each of the monomers ("L- monomer" and "R-monomer") of TPO dimer structures of wild type and mutant cases. Table 3.9 described that the average binding energies of H_2O_2 with TPO WT-HD, TPO MT1-HD and TPO MT3-HD were -2.85, -2.65 and -2.75 (kcal/mole), respectively,

which suggested that there was a slight alteration of binding energies between the wild type and the mutant structures.

In case of sequential docking of iodide (I^-) with TPO WT-HD, TPO MT1-HD & TPO MT3-HD, we found that the average binding energies for TPO WT-HD was -1.2 kcal/mole (-1.3 kcal/mole for L- monomer and -1.1 kcal/mole for R-monomer). Analysis of TPO MT1-HD protein showed the binding energies for the left and the right monomers were -1.1 kcal/mole and -1.3 kcal/mole, respectively and the average binding energy was same as wild type dimer protein. TPO MT3-HD showed the binding energies for iodide (I^-) was -1.1 kcal/mole both for of the left and right monomers. This result demonstrated that there was no significant change in the binding energies for iodide (I^-) with both of wild type and mutant TPO dimer structures. Finally, it was analyzed that in most cases, L-monomers of wild type and mutant dimer structures showed attachment of iodide (I^-) with carbon atom. R- monomer of TPO WT-HD interacted with iodide (I^-) through His239, Arg396, Ala397 and Thr404; whereas, R-monomers of mutant cases (TPO MT1-HD and TPO MT3-HD) did not interacted properly with those amino acids (Figure3.11).

Table 3. 9. Binding energies of H2O2 and iodine (I-) with TPO WT HD, TPO MT1 HD, TPO MT3 HD protein after flexible sequential docking

TPO dimer	Protein Monomer	Binding energies for H_2O_2 (kcal / mole)	Binding energies for iodide (I^-) (kcal / mole)	Average Binding energies for H_2O_2 (kcal / mole)	Average Binding energies for iodide (I^-) (kcal / mole)
TPO WT-HD	L	-2.8	-1.3	-2.85	-1.2
	R	-2.9	-1.1		
TPO MT1-HD	L	-2.7	-1.1	-2.65	-1.2
	R	-2.6	-1.3		
TPO MT3-HD	L	-2.6	-1.1	-2.75	-1.1
	R	-2.9	-1.1		

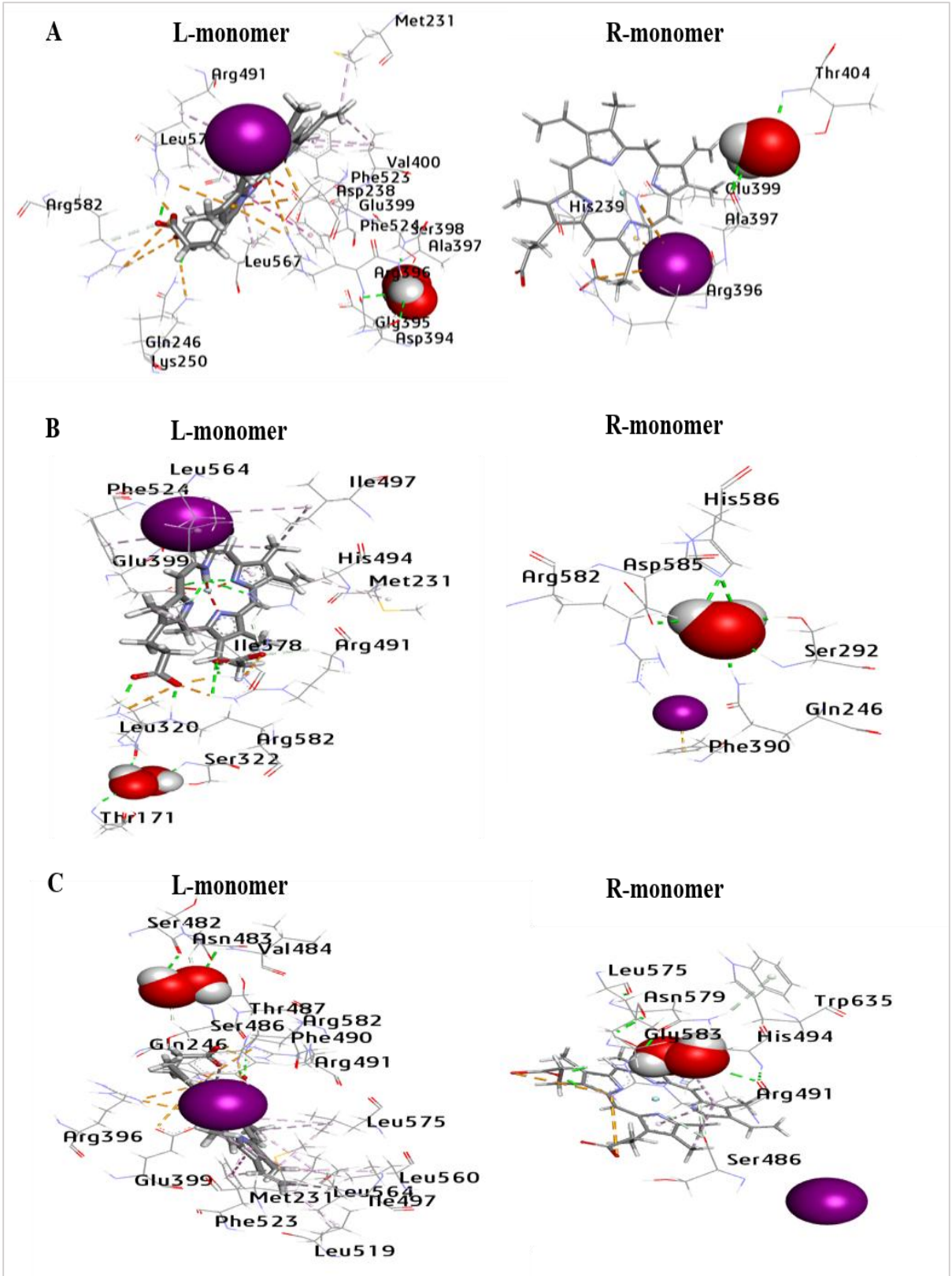


Figure 3. 11. Non-bond interactions after sequential docking with H_2O2 and iodide (I-) with both of left and Right monomer of (a) TPO WT-HD, (b) TPO MT1-HD and (c) TPO MT3-HD.

3.3.3. MD Simulation of TPO Dimer Structures after Sequential Docking

Molecular Dynamics simulation was performed for 5000 ps using YASARA software to observe whether the ligands H_2O_2 and iodide (I^-) stayed in the same binding pocket after simulation. To do so, the ligands were retrieved from the simulated protein-ligand complexes and molecular docking was performed again and found that the binding energies for both of H_2O_2 and iodide (I^-) were same as stated in Table 3.9. Thus, the dynamics results showed that the binding pocket of TPO protein for essential ligands was consistent with the monomer study by observing the presence of crucial amino acids of the catalytic site.

3.4. Setting Up of High Resolution Melt (HRM) Curve Analysis

In the first phase of the experiments, 36 participants were recruited for HRM study and among them 20 specimens were sequenced to find out the mutational hotspot region from Exon 8 to Exon 14 of TPO gene to get the pre-genotyped samples with known mutations. The ultimate goal of this Sanger sequencing experiment was to define the mutational hotspot region of TPO gene in the hypothyroid patients that could be targeted for HRM primer design for generation of known HRM curve patterns. Furthermore, additional 16 samples were recruited to validate and check the specificity and sensitivity of the HRM method. Upon sequencing the specimens of initially recruited 20 participants, four different substitution mutations were identified such as c.1117G>T (p.Ala373Ser), c.1193G>C (p.Ser398Thr), c.2145C>T (p.Pro715Pro) and c.2173A>C (p.Thr725Pro) from which we targeted the three non-synonymous mutations to develop HRM method.

3.4.1. Screening of C.1117G>T Variant

To establish a rapid HRM-based screening approach targeting the c.1117G>T variant, a pair of primers (TPO_G1117T_Ex8) was designed that flanked the c.1117G>T variant. Then the pre-genotyped samples were subjected to real-time PCR followed by HRM analysis. The HRM curve analysis generated three distinctive clusters which corresponded to wild type, homozygous and heterozygous c.1117G>T variants (Figure 3.12 and Figure 3.13). Figure 3.12 represents the Normalized melt curves targeting the c.1117G>T variant. As we could see here, the c.1117G>T variant resulted in a decrease in melting temperature in the homozygous and heterozygous specimens compared to the wild type allele. Even the heterozygous samples were

distinguishable from the homozygous samples, as the fluorescence started to drop quickly at the initial phase of melting for heterozygous samples. However, in the later phase of melting, the homozygous samples started to lose fluorescence intensity quicker than the heterozygous samples and as a consequence they crossed each other at a certain point of melting making them distinguishable from each other. Similar to the normalized melting curve, the temperature shifted difference curve could generate distinctive melting patterns for the wild type, homozygous and heterozygous specimens (Figure 3.13). The reliability of the method was further validated by analyzing 16 unknown samples with Dyshormonogenesis. At first, these samples were tested by HRM and then Sanger sequencing was done to check the sensitivity and specificity of the method. A total of 5 specimens had c.1117G>T variant in heterozygous states, 6 specimens were in a homozygous state and the remaining 5 had the wild type allele. The HRM result was consistent with the sequencing data, which implies that sensitivity and specificity for the detection of c.1117G>T variant was 100% for both homozygous and heterozygous alleles.

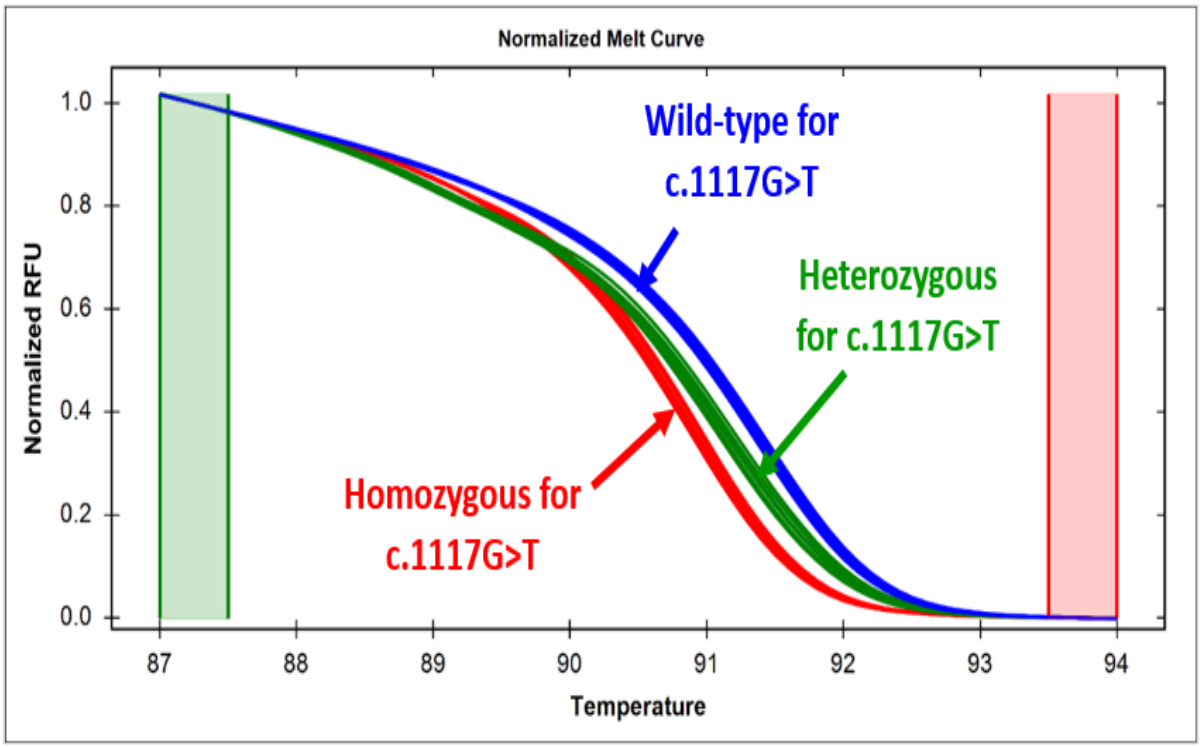


Figure 3. 12. Normalized melt curves for the specimens targeting the c.1117G>T variant (Exon-8). Normalized melt curves showing that the specimens with homozygous and heterozygous states are clearly distinguishable from the wild type specimens, as manifested by the difference in relative fluorescence unit.

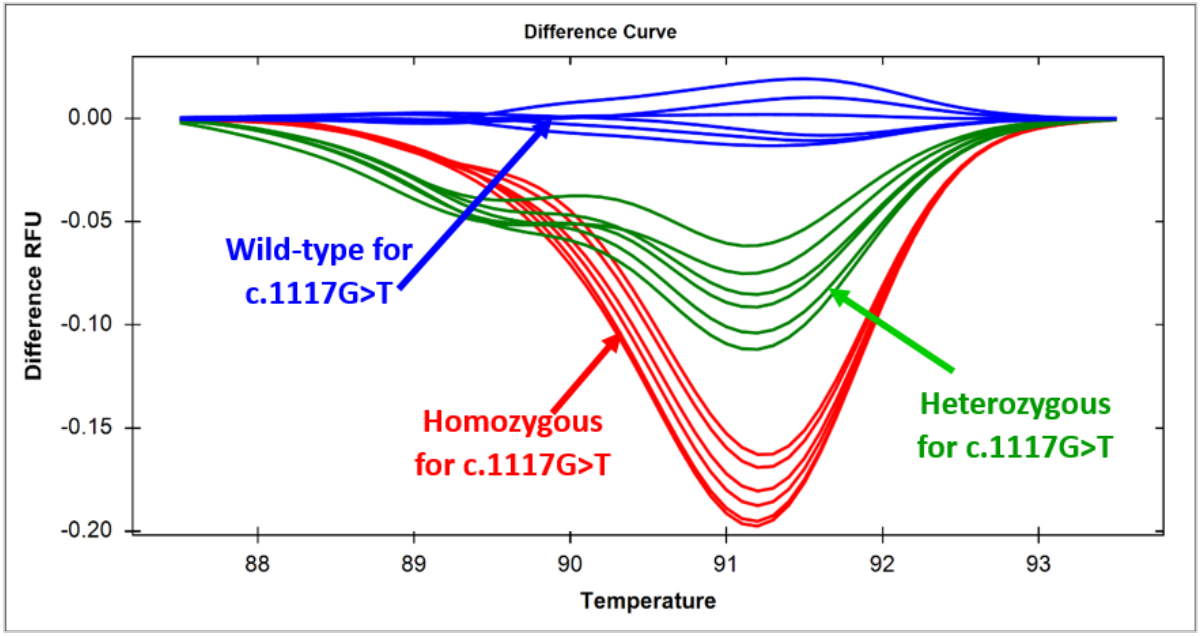


Figure 3. 13. Difference curves generated by specimens targeting the c.1117G>T variant (Exon-8). Discernable changes in three difference curves showing that the specimens with homozygous and heterozygous states are clearly distinguishable from the wild type allele, as manifested by the difference in relative fluorescence unit.

3.4.2. Screening of C.1193G>C Variant

The second set of primers, namely TPO_G1193C_Ex8 was used for analysis of c.1193G>C variant by the HRM approach. When the pre-genotyped samples were subjected to HRM analysis, three different clusters were observed in the melt curve analysis. One of the clusters corresponded to the heterozygous samples, other two were for the homozygous samples and for the wild type samples (Figure 3.14). However, in the difference curve analysis, 3 different clusters were clearly observed for homozygous, heterozygous and wild type variant of c.1193G>C (Figure 3.15). Similar observation was observed with 16 unknown samples that were subjected to HRM analysis. Four out of 16 samples came out as heterozygous by HRM analysis and this result was consistent with the sequencing data. However, 8 homozygous and 4 wild type samples formed different clusters in this case. This observation implies that both heterozygous and homozygous states for c.1193G>C variant could be detected with 100% sensitivity and specificity.

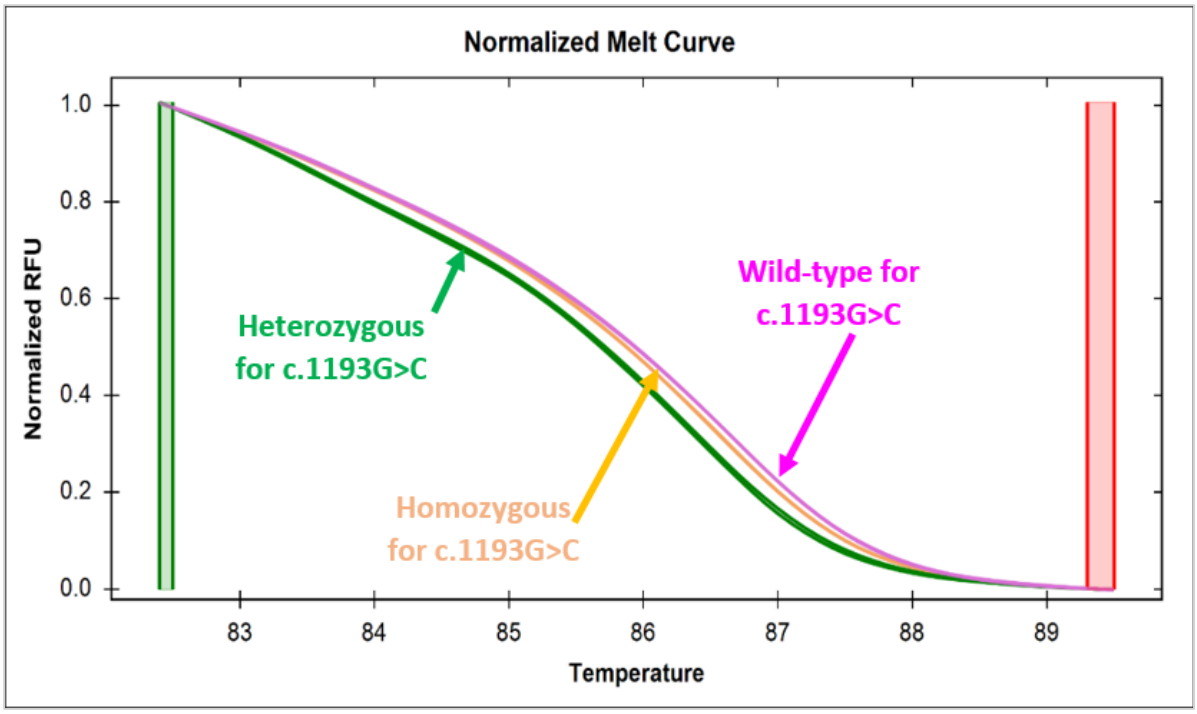


Figure 3. 14. Normalized melt curves generated by specimens targeting the c.1193G>C variant (Exon-8). Discernable changes in normalized melt curves showing that the specimens with homozygous (orange color) and heterozygous (green color) states are clearly distinguishable from the wild type (purple color) alleles, as manifested by the difference in relative fluorescence unit.

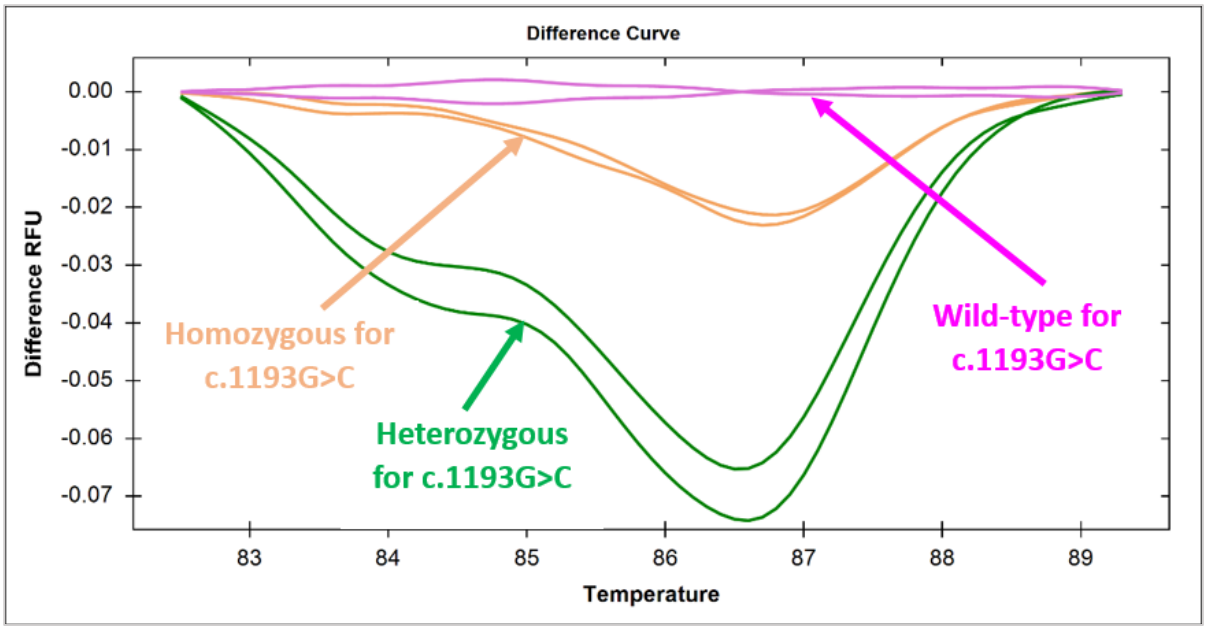


Figure 3. 15. Difference curves for specimens targeting the c.1193G>C variant (Exon-8). Discernable changes in difference curves showing that the specimens with homozygous and heterozygous states are clearly distinguishable from the wild type states, as manifested by the difference in relative fluorescence unit.

3.4.3. Screening of C.2173A>C Variant

The third set of primers, namely TPO_A2173C_Ex12 was used to analyze another TPO gene mutation designated as c.2173A>C. The pre-genotyped wild type, homozygous c. 2173A>C and heterozygous c. 2173A>C specimens were subjected to HRM analysis. The wild type, homozygous and heterozygous variants formed distinct clusters (Figure 3.16 and Figure 3.17). The homozygous c. 2173A>C substitution resulted in an increase in melting temperature and thus the specimens with homozygous c. 2173A>C variant had higher fluorescence intensity than that with the wild type allele during melting, as manifested by the relative fluorescence unit (Figure 3.16). On the other hand, although the specimens with the heterozygous c. 2173A>C variant followed a pattern of melting that had initially lower fluorescence intensity compared to the wild type, the melting curve patterns almost overlapped with each other in later stage (upper panel of Figure 3.16). Thus the homozygous c. 2173A>C, heterozygous c. 2173A>C and the wild type alleles were discernable form each other. Similar to the normalized melting curve, the difference curve analysis could also distinguish different states involving c. 2173A>C variant (Figure 3.17). HRM analysis of 16 unknown samples targeting the c. 2173A>C variant was also 100% sensitive and specific. The HRM approach showed that 5 out of 16 unknown samples were wild type, 6 were homozygous and the rest 5 were heterozygous. Sequencing of those 16 unknown samples revealed that real time PCR-HRM-based result was consistent with the sequencing result.

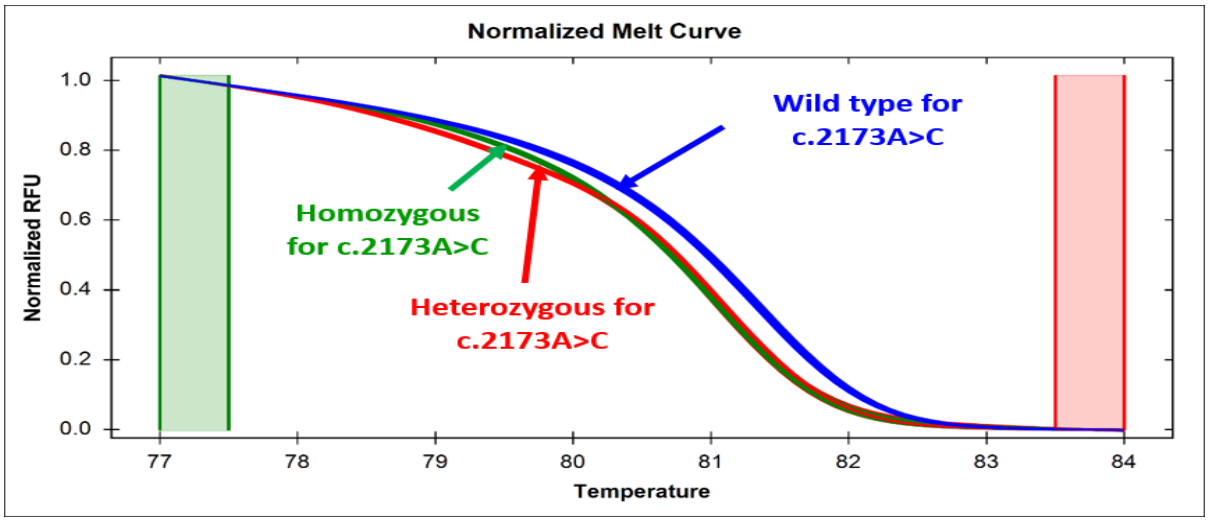


Figure 3. 16. Normalized melt curves for specimens targeting the c.2173A>C variant (Exon-12). Discernable changes in normalized melt curves showing that the specimens with homozygous and heterozygous states are clearly distinguishable from the wild type alleles, as manifested by the difference in relative fluorescence unit.

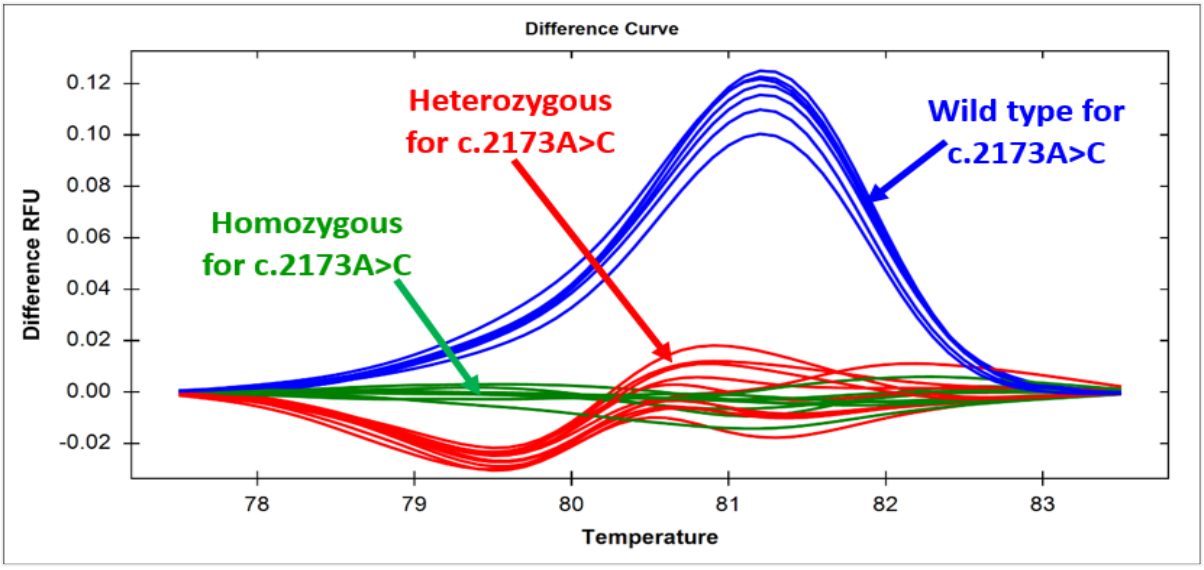


Figure 3. 17. Differential curves for specimens targeting the c.2173A>C variant (Exon-12). Discernable changes in difference curves showing specimens with homozygous and heterozygous states are clearly distinguishable from the wild type, as manifested by the difference in relative fluorescence unit.

3.5. Investigation of Mutation in TSHR Gene

All the 21 patients with Dysgenesis had mutation in Exon 10 among a total of 11 Exons in TSHR gene. The mutations we found namely, c.1523C>T (p.Ser508Leu) and c.2181G>C (p.Asp727Glu). Among the 21 patients, only one patient had mutation c.1523C>T (p.Ser508Leu) and 20 patients had the other variant c.2181G>C (p.Asp727Glu). The variants were then analyzed by bioinformatics tools to explore the pathogenic effect. Firstly, the mutations were tested by Polyphen 2, Mutation taster and PROVEAN bioinformatics tools to see whether they possessed damaging effect or not. We found that the mutation c.1523C>T probably had damaging effect and c.2181G>C variant showed benign effect. Figure 3.18 represents a chromatogram showing the mutation (c.1523C>T) for the specific participant and Table 3.10 shows the mutation found in TSHR gene.

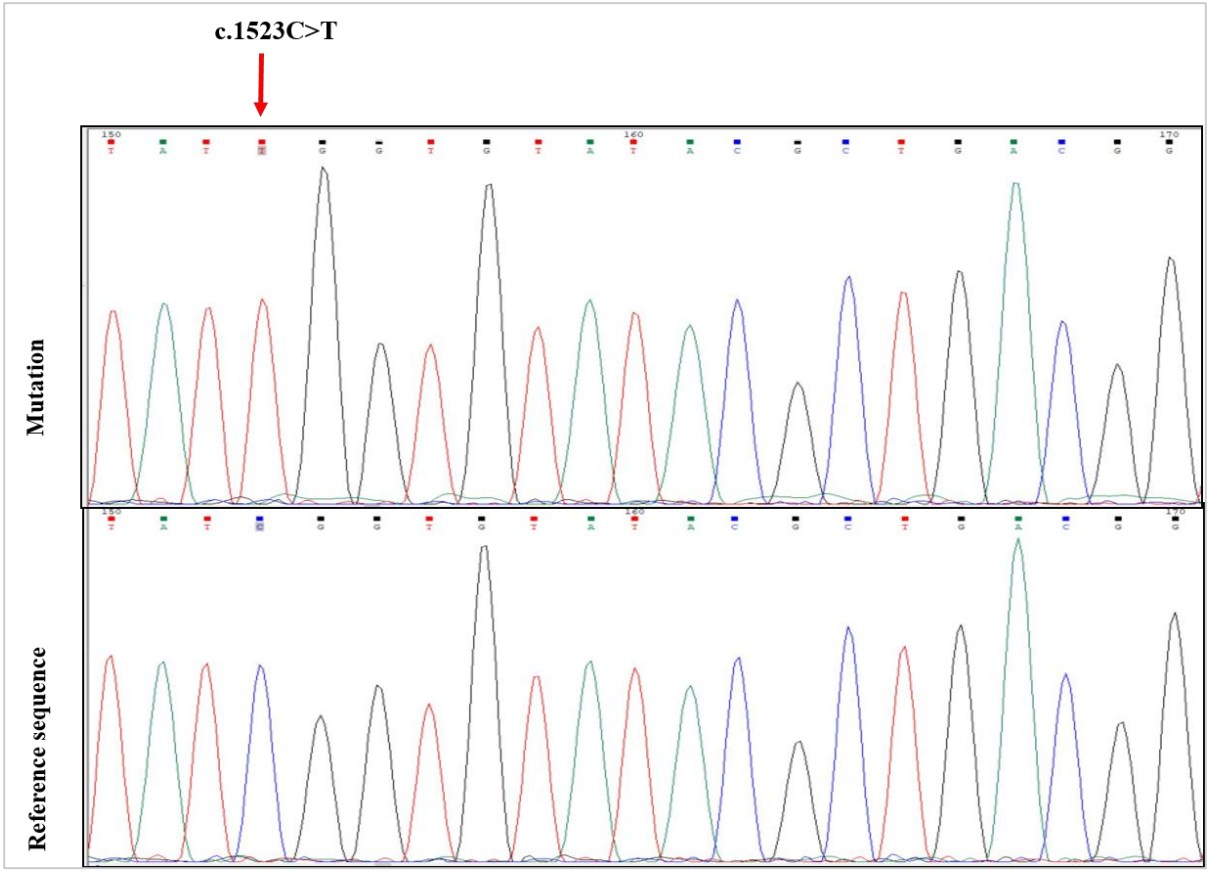


Figure 3. 18. A representative chromatogram showing a mutation in Exon 10 of TSHR gene

Table 3. 10. Mutation detection in the TSHR gene of hypothyroid patients and analysis of the effect of mutation using different bioinformatics tools

Nucleotide position	Amino acid position	Polyphen 2	Mutation Taster	PROVEAN
c.1523C>T	p.Ser508Leu	Probably Damaging	Disease causing	Deleterious
c.2181G>C	p.Asp727Glu	Benign	Benign	Neutral

3.5.1. Modelling of the Transmembrane Region of TSHR Protein

Since the small molecules drug bind to the transmembrane region of the TSHR protein, we predicted the structure by using I-TASSER tool according to previously described TPO protein. The predicted structures were designated as wild type TSHR $^{368\text{-}764}$ WT, p.Ser508Leu variant as TSHR $^{368\text{-}764}$ MT1 and p.Asp727Glu variant as TSHR $^{368\text{-}764}$ MT2. Table 3.11 shows the C-score for wild type, MT1 and MT2 was -0.71, -0.61 and -0.63, respectively. The TM

score and RMSD were 0.62±0.14, 8.4±4.5Å for TSHR $_{368-764}$ WT; 0.64±0.13 and 8.2±4.5Å TSHR $_{368-764}$ MT1; 0.63±0.13 and 8.2±4.5Å for TSHR $_{368-764}$ MT2. Figure 3.19 depicts the 3D structures of the TM-region of TSHR protein and the small molecules MS437 and MS438.

Table 3. 11. Summary of the corresponding model numbers, C–score, TM–score and the RMSD–score of the predicted 3D structures of TSHR 368-764 WT, TSHR 368-764 MT1 and TSHR 368-764 MT2.

Features	TSHR $_{368-764}$ WT	TSHR $_{368-764}$ MT1	TSHR $_{368-764}$ MT2
Model no.	01	01	01
C–score	-0.71	-0.61	-0.63
TM–score	0.62±0.14	0.64±0.13	0.63±0.13
RMSD (Å)	8.4±4.5Å	8.2±4.5Å	8.2±4.5Å

C–score=Confidence score range: [-5, 2]; TM–score=Template Modelling score, TM–score < 0.17 indicates random similarity and TM–score > 0.5 indicates correct similarity; RMSD=Root Mean Square Deviation. WT = Wild type, MT1 = Mutant 1 (p.Ser508Leu) and MT2 = Mutant 2 (p.Asp727Glu).

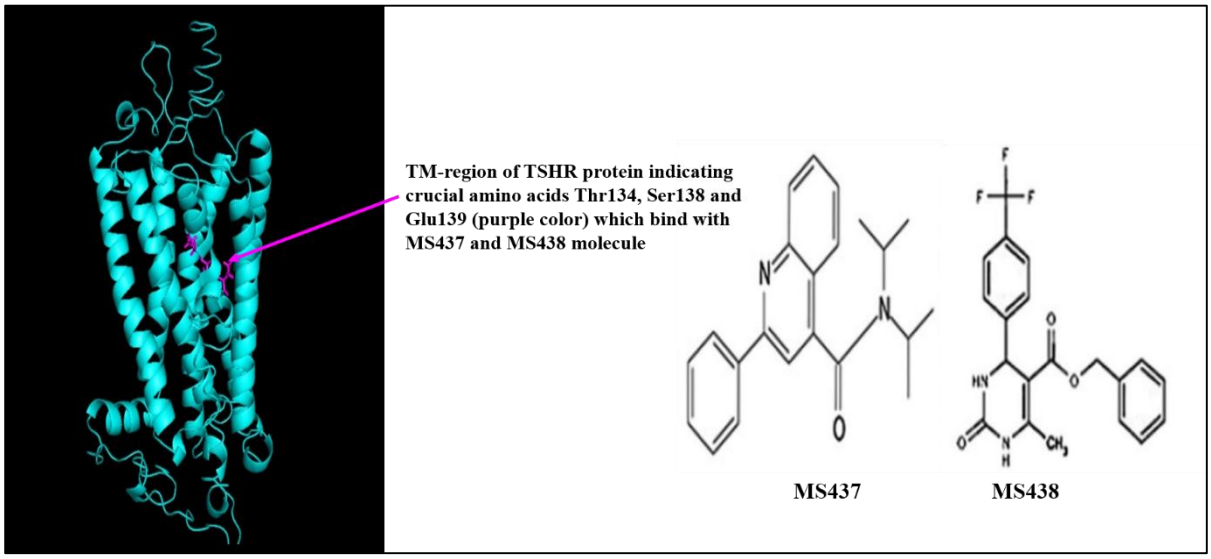


Figure 3. 19. The predicted 3D structures of transmembrane region of TSHR proteins (368-764) showing crucial amino acids responsible for binding with MS437 and MS438 molecules.

3.5.2. Molecular Docking of MS437 and MS438 with TM Region of TSHR Proteins (Wild Type and Mutant)

The structures of the small molecules MS437 and MS438 were optimized and molecular docking was performed using PyRx software. The Table 3.12 showed that the binding affinity of the wild type TSHR protein (-6, kcal/mol, $TSHR_{368-764}$WT) was higher compared to the mutant cases (-4.8 kcal/mol for $TSHR_{368-764}$ MT1 and -5.7 kcal/mol for $TSHR_{368-764}$ MT2).

Total non-bond interactions were 11, 19 and 12 for wild type, MT1 and MT2, respectively. MS437 binds to 501 Threonine and MS438 binds to Serine 505 and Glutamic acid 506 of transmembrane helix3 (TMH3) in full length TSHR protein with corresponding amino acid position Thr134, Ser138 and Glu139, respectively in TM-region [107]. We tried to investigate whether these crucial amino acids could interact with small molecules thyrogenic drug. We found that in case of MS437 none of these three amino acids could interact with either wild type and mutant cases. On the other hand, MS438 interacted with all the crucial amino acids including Thr134, Ser138 and Glu139 for wild type case and for the mutant cases ($TSHR_{368-764}$MT1 and $TSHR_{368-764}$MT2), it could interact only with Ser138. The binding affinities were -7.1 kcal/mol, -5.4 kcal/mol and -2.6 kcal/mol for $TSHR_{368-764}$WT, $TSHR_{368-764}$ MT1 and $TSHR_{368-764}$MT2, respectively. The binding affinity of wild type protein was much higher for both of the ligands (MS437 and MS438). All the non-bond interactions were depicted in Figure 3.20 and Figure 3.21 for MS437 and MS438 respectively.

Table 3. 12. Binding energy and non-bond interactions of MS437 with TSHR368-764WT, TSHR368-764 MT1 and TSHR368-764 MT2 proteins after flexible docking

Protein type	Binding Affinity (kcal/mol)	Hydrogen bond	Hydrophobic bond	Total interactions
$TSHR_{368-764}$ WT	-6	SER274(641)	VAL219(586), ALA277(644), VAL135(502), MET270(637), ILE273(460), PRO204(571), VAL297(664)	11
$TSHR_{368-764}$ MT1	-4.8	-	VAL135(502), LEU222(589), VAL219(586), MET270(637), ILE273(640), VAL219(586), ALA277(644), LEU203(570), LYS293(660), VAL297(664), LYS293(660), VAL297(664), PHE218(585), TYR300(667)	19
$TSHR_{368-764}$ MT2	-5.7	-	VAL261(628), LEU310(677), ILE313(680), ALA306(673), LEU302(669), CYS305(672), ALA306(673), PHE309(676)	12

Table 3.13. Binding energy and non-bond interactions of MS438 with TSHR368-764WT, TSHR368-764 MT1 and TSHR368-764 MT2 proteins after flexible docking

Protein type	Binding Affinity (kcal/mol)	Hydrogen bond	Electrostatic bond	Hydrophobic bond	Total interactions
$TSHR_{368-764}$ WT	-7.1	ASN223(590), VAL135(502), SER138(505), GLU139(506), ASN223(590),	-	THR134(501), VAL142(509), MET270(637), VAL297(664)	13
$TSHR_{368-764}$ MT1	-5.4	SER138(505), ILE273(640), VAL135(502), SER274(641), CYS202(569),	LYS293 (660)	LEU100(467), LEU203(570), ALA277(640), LYS293(660), VAL297(664), ILE273(640)	14
$TSHR_{368-764}$ MT2	-2.6	SER138(505), LEU278(645), SER274(641), ALA277(644),	-	VAL219(586), ALA277(644), LEU278(645), LEU100(467), VAL297(664), PRO301(668), VAL219(586)	14

Table 3.12 and 3.13. the amino acid residues and their positions are designated as the three letter abbreviations and the corresponding number; in case of $TSHR_{368-764}$ the amino acid outside the first bracket indicates the position in predicted structure and the amino acid residues in first bracket indicates the real position in $TSHR_{1-764}$ protein; WT = Wild type; MT1 = Mutant 1 (p.Ser508Leu) and MT2 = Mutant 2 (p.Asp727Glu).

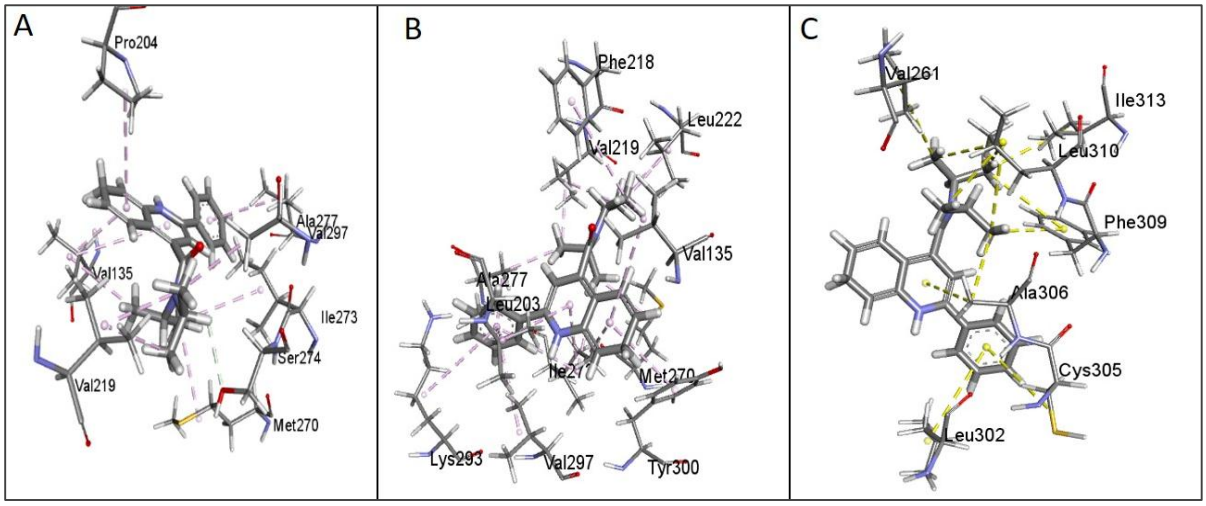


Figure 3. 20. Non-bond interactions of MS437 with corresponding predicted structures as obtained using a BIOVIA Discovery Studio 2017. (A) TSHR368-764WT and MS437, (B) TSHR368-764MT1 and MS437 and (C) TSHR368-764MT2 and MS437.

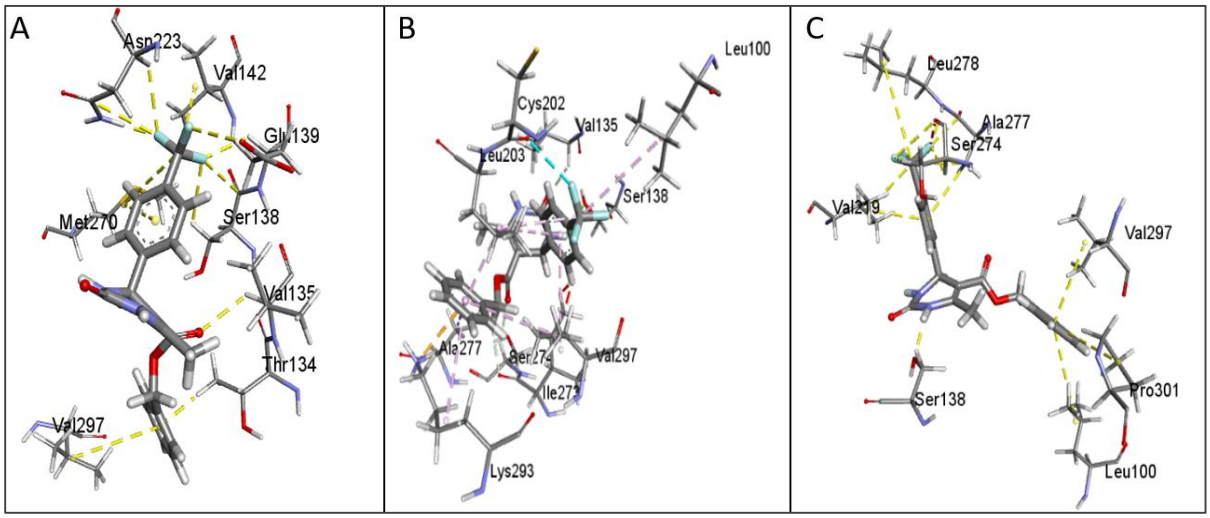


Figure 3. 21. Non-bond interactions of MS438 with corresponding predicted structures as obtained using a BIOVIA Discovery Studio 2017. A. TSHR368-764WT and MS438 B. TSHR368-764MT1 and MS438 and C. TSHR368-764MT2 and MS438

3.5.3. Molecular Dynamics Simulation of TSHR Protein Structures

To see the effect of mutation, MD simulation was performed for 5000ns and RMSD was calculated. Here the more RMSD value indicated the more possibility of instability of the 3D structures of protein. In case of MS437 molecule, the 3D structures of mutant cases were found more unstable compared to wild type TSHR protein. On the other hand, although MS438 showed that the mutant cases were unstable initially, the wild type case became unstable after completion of 5ns run during the MD simulation (Figure 3.22).

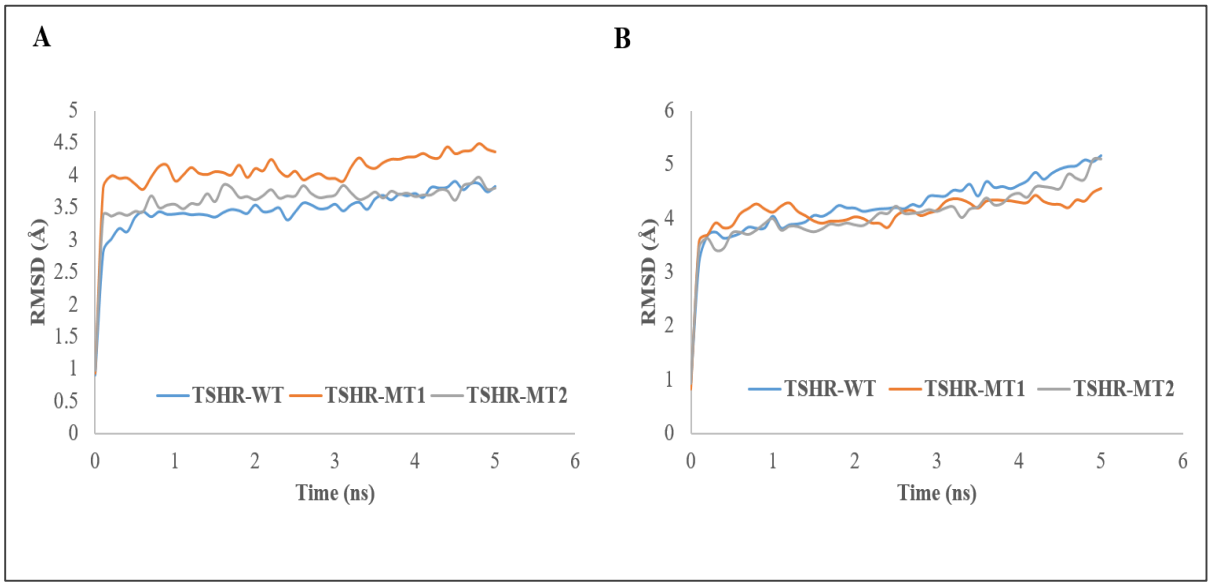


Figure 3. 22. After MD simulation, RMSD calculation for both of the ligands (A)MS437 and (B)MS438 among WT, MT1 and MT2 structures of TSHR protein.

3.6. Genetic Study on PAX 8 Gene

Since PAX8 gene is associated with thyroid Dysgenesis, we tried to investigate the association of mutation in this gene. However, we didn't find any mutation in the hotspot region (Exon 3) of PAX8 gene in the study participants. Study of the full gene along with other factors might be associated with our Bangladeshi patients.

3.7. Profiling of Acylcarnitines using LC-MS/MS

3.7.1. Free Carnitine, Total Carnitine and Total Acylcarnitine Profile

Despite daily thyroid hormone replacement therapy, the CH patients had been suffering from persistent fatigue. As carnitine plays a critical role in mitochondrial fatty acid oxidation, it has been suggested that fatigue-related complications may involve in altered carnitine-acylcarnitine homeostasis. Table 3.14 shows a comparison among L-carnitine (C0), total carnitines and 13 long chain acylcarnitines as well as total acylcarnitines between the hypothyroid patients with fatigue or fatigue-related symptoms and healthy controls. The mean (±SD**)** free carnitine concentration was 45.38±12.61 µmol/L in the patient group, whereas it was 41.54±9.85 µmol/L in the healthy control group and the difference between the 2 groups generated a P value of 0.049. On the other hand, the concentrations of total acylcarnitines and total carnitines of the patient group were 21.95±7.66 µmol/L and 67.33±18.27 µmol/L, respectively, which were higher than that of the control group (total acylcarnitine=20.96±5.61 µmol/L; total carnitine=62.51±14.13 µmol/L), although the difference did not generate a significant P value.

3.7.2. Long Chain (LC)-Acylcarnitine Profile

In this metabolomic profiling study, 13 long chain acylcarnitines were analyzed which included 6 saturated (C14, C16, C18, C14OH, C16OH and C18OH), 5 monounsaturated (C14:1, C16:1, C18:1, C16:1OH and C18:1OH) and 2 polyunsaturated (C14:2 and C18:2) acylcarnitines. The blood concentrations of acylcarnitines with saturated long chain fatty acids were 0.09±0.05 µmol/L, 1.02±0.32 µmol/L, 0.39±0.14 µmol/L, 0.01±.007 µmol/L, 0.001±0.002 µmol/L and 0.001±0.003 µmol/L for myristoylcarnitine (C14), palmitoylcarnitine (C16), stearoylcarnitine (C18), hydroxymyristoylcarnitine (C14OH), hydroxypalmitoylcarnitine (C16OH) and 3-hydroxystearoylcarnitine (C18-OH), respectively, in the patients group. The C16, C18 and C16OH acylcarnitine concentrations in the blood of hypothyroid patients were significantly lower than that of the healthy control participants with C16, C18 and C16OH acylcarnitine concentrations of 1.24±0.43, 0.47±0.17 and 0.01±0.003 µmol/L, respectively and the differences between the two groups produced significant P values of 0.0003, 0.002 and 0.001, respectively (Table 3.14).

Table 3. 14. Comparison of free carnitine, long chain acylcarnitines, total carnitines and total acylcarnitines between the patients and the healthy controls

Parameters	Patient µmol/L (mean±SD)	CV (%)	IQR	Healthy Control µmol/L (mean±SD)	CV (%)	IQR	*p* value
Free Carnitine (C0)	45.38±12.61	27.79	17.45	41.54±9.85	23.72	4.9	0.049*
Total acylcarnitines	21.95±7.66	36.1	11.54	20.96±5.61	25.48	6.02	0.4
Total Carnitines	67.33±18.27	29.43	24.92	62.51±14.13	29.14	15.95	0.08
Long chain acylcarnitines							
Myristoylcarnitine (C14)	0.09±0.05	53.0	0.04	0.09±0.05	59.65	0.03	0.95
Myristoleylcarnitine (C14:1)	0.07±0.03	54.7	0.05	0.07±0.03	46.29	0.04	0.76
Tetradecadienoylcarnitine (C14:2)	0.03±0.02	72.18	0.03	0.03±0.02	52.0	0.02	0.95
Palmitoylcarnitine (C16)	1.02±0.32	31.67	0.43	1.24±0.43	34.38	0.47	0.0003*
Palmitoyleylcarnitine (C16:1)	0.06±0.02	34.02	0.02	0.06±0.02	31.03	0.02	0.58
Stearoylcarnitine (C18)	0.39±0.14	37.95	0.19	0.47±0.17	35.5	0.25	0.002*
Oleylcarnitine (C18:1)	0.7±0.21	30.32	0.32	0.81±0.25	31.43	0.3	0.007*
Octadecadienyl carnitine (C18:2)	0.29±0.1	36.83	0.1	0.32±0.1	31.13	0.11	0.113
Hydroxymyristoylcarnitine (C14OH)	0.01±.007	66.37	0.008	0.01±.006	64.17	0.008	0.76
Hydroxypalmitoylcarnitine (C16OH)	0.001±0.002	71.58	0.004	0.01±0.003	42.13	0.005	0.001*
Hydroxypalmitoleylcarnitine (C16:1OH)	0.03±0.01	51.71	0.01	0.03±0.014	46.96	0.01	0.26
3-Hydroxystearoylcarnitine (C18-OH)	0.001±0.003	83.99	0.004	0.001±0.003	59.98	0.004	0.013*

Parameters	Patient µmol/L (mean±SD)	CV (%)	IQR	Healthy Control µmol/L (mean±SD)	CV (%)	IQR	*p* value
Hydroxyoleylcarnitine (C18:1OH)	0.01±0.006	54.72	0.008	0.01±0.004	38.2	0.006	0.91
Total long chain acylcarnitines	**2.67±0.87**	**31.02**	**1.13**	**3.15±0.93**	**29.36**	**1.09**	**0.0014***

CV= Coefficient of variance; IQR= Interquartile range; * P< 0.05 was considered significant, Total acylcarnitines= Total long chain acylcarnitines + Total medium chain acylcarnitines + Total short chain acylcarnitines; Total Carnitine= Free carnitine + Total acylcarnitines

However, C14, C14OH and C18OH acylcarntine levels remained almost unaltered in the patients and healthy control groups. On the other hand, the mean concentrations of all the monounsaturated long chain acylcarnitines were almost similar in the patients and the control participants except oleylcarnitine (C18:1). It is noticeable here that oleylcarnitine (C18:1) decreased significantly in the patients (0.7±0.21 µmol/L) compared to the healthy controls (0.81±0.25 µmol/L), generating a P-value of 0.007. Among the two polyunsaturated acylcarnitines, octadecadienyl carnitine (C18:2) level was found to be decreased in the patients compared to the controls without generating statistically significant difference (P= 0.113), whereas the concentration of tetradecadienoylcarnitine (C14:2) of the patients and the healthy controls did not differ. However, in contrast to the short and medium chain acylcarnitines (Supplementary Table 6), the mean concentration of these 13 long chain acylcarnitines was significantly lower in the patients (2.67±0.0.87 µmol/L) than that in the healthy controls (3.15±0.93 µmol/L) and the difference between the two groups generated a significant P-value of 0.0014.

3.7.3. Indicator Ratios of Metabolic Process and Measurement of Triglyceride (TG) Level

With a significant decrease in blood concentrations of long chain acylcarnitines (C16, C18, C18:1, C16OH and C18OH) along with a total decline in concentration of LC-acylcarnities and an increase in blood level of free carnitine (C0); C0/(C16+C18) ratio was found to be significantly higher in the patients (ratio= 34.55± 14.88, IQR=12.9, %CV=43.07) than that in

the healthy controls (ratio= 25.73± 6.87, IQR=8.48, %CV=26.71 (P< 0.0001) (Figure 23.A). All these results translate that the activity of CPT-I enzyme which plays a pivotal role in the β-oxidation of long chain fatty acids by converting them to acylcarnitines for translocation from the cytosol into the inner mitochondrial matrix is low in patients group [108,109].

The data demonstrating a disturbance in the metabolism of long chain acylcarnitines in the hypothyroid patients with fatigue due to an insufficient CPT-I activity prompted us to ask whether the long chain fatty acids had diverted use in the biosynthesis of TG. Accordingly, we quantified plasma TG levels. Figure 23.B demonstrates that the plasma TG level was significantly higher in the patients (88.92± 59.54 mg/dL) (IQR= 77.85, %CV= 66.96) than that in the healthy controls (58.33± 15.79 mg/dL) (IQR= 24.75, %CV= 27.06). The difference in TG concentration between the two groups generated a significant P-value of 0.02. Interestingly, although the TG level in the hypothyroid patient's blood was higher and none of the patients was overweight (average BMI was 17.0±4.4 kg/m^2), the findings indicate that the deposition of TG might not have occurred in a significant level in the adipose tissues.

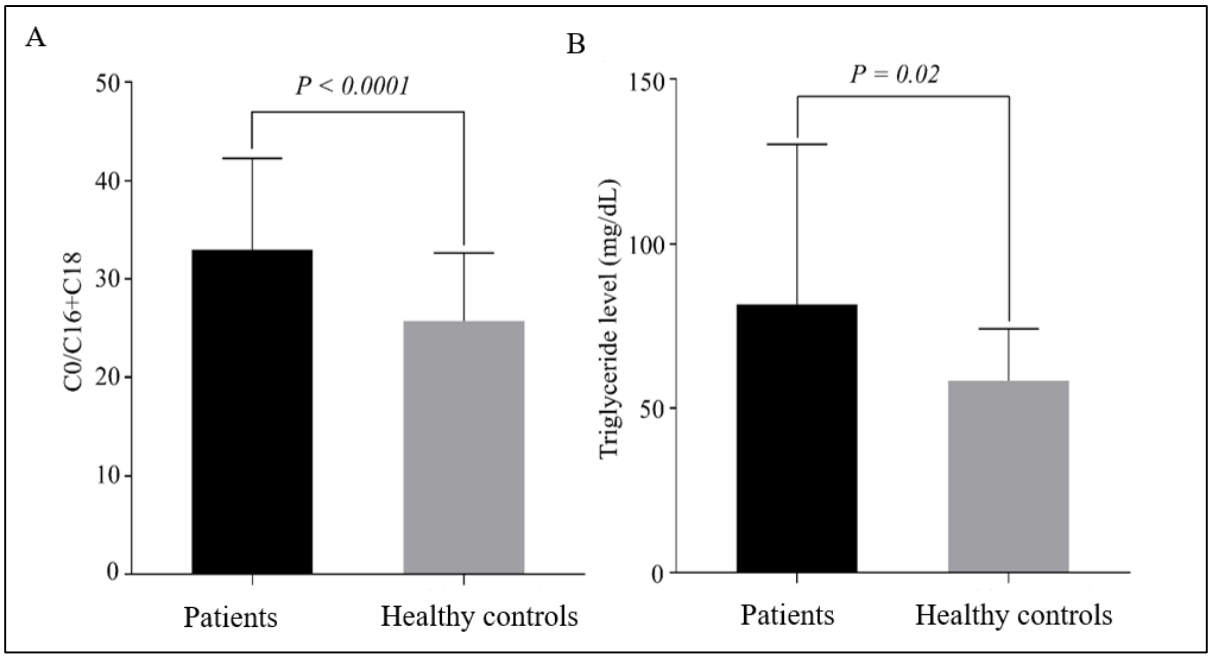


Figure 3. 23. Comparison of CPT-I enzyme deficiency indicator and plasma TG concentration between the patients and the healthy controls. (A) C0/(C16+C18) ratio, an important indicator of CPT-I enzyme activity, was significantly higher in the patients than in the control group (P<0.0001). (B) Plasma Triglyceride (TG) concentrations was found higher in hypothyroid patients compared to the healthy controls and the difference generated a significant P-value of 0.02. A P-value<0.05 was considered significant.

3.7.4. Measurement of CPT1 Gene Expression Levels

Since the indicator ratio C0/(C16+C18) showed that there might be defect in CPT1 activity, we quantified the mRNA level of CPT1A and CPT1B gene expression for both healthy controls and congenital hypothyroid children. GAPDH housekeeping gene expression was compared with CPT1A and CPT1B gene expression.

The average fold change of CPT1A and CPT1B was calculated and compared to GAPDH (Figure 3.24). In case of CPT1A activity, we found that the CPT 1A gene expression was higher in patient group (5.66 times of GAPDH) compared to the healthy control (2.66 times of GAPDH) group and the CPT1B was significantly lower in patient group (0.39 times of GAPDH) compared to healthy control group (3.67 times of GAPDH). Further, the ratio CPT1A/CPT1B was also calculated between CH patients with the non-hypothyroid participants. The primary result showed that the ratio was approximately 14 times higher in patients group (14.31) than that of the healthy controls (0.73) depicted in figure 3.24. (C).

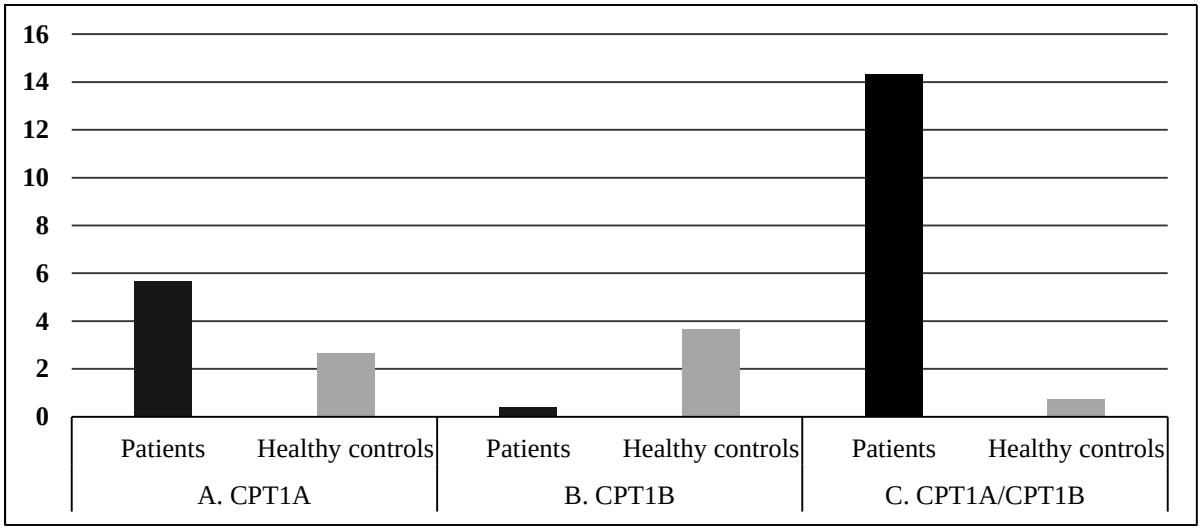


Figure 3. 24. Comparison of CPT-I gene expression level with GAPDH housekeeping gene between the patients and the healthy controls. (A) CPT-IA gene expression level. (B) CPT1B gene expression level from mRNA and (C) CPT1A/CPT1B ratio.

3.8. Profiling of Amino Acids using LC-MS/MS technique

Similar with the acylcarnitine levels, all the patients and healthy subjects without any signs symptoms of hypothyroidism were enrolled for further analysis of their amino acid levels using LC-MS/MS. All the patients were under levothyroxine (LT_4) treatment. None of the enrolled patients were treated using T3 or combination of LT_4 and T_3. During enrollment for this study, serum TSH and FT_4 level of the patients were within normal range. Although all the patients were under LT_4 treatment and their serum FT_4 levels were within normal range, majority of the patients were suffering from several clinical complications.

3.8.1. Levels of Blood Amino Acids

Several animal model study proves that alteration of blood concentration of amino acids occurs because thyroid hormone status influences the protein and amino acid metabolism [72,110]. As mentioned previously, all the enrolled hypothyroid patients were under LT_4 treatment but their treatment was delayed because of late diagnosis. In this study, we assessed whether the LT_4 treatment could have any effect on the concentration of blood amino acids of late-diagnosed hypothyroid patients compared to the healthy controls using dried blood spot specimens. As depicted in Table 3.15, among the essential amino acids; Phenylalanine, Methionine, Leucine and Valine were significantly decreased in the hypothyroid patients compared to the control group (Phe: 49.66 ± 1.626 µmol/L vs. 65.56 ± 1.483 µmol/L, P<0.0001*; Met: 20.36 ± 0.8299µmol/L vs. 24.38 ± 0.8278 µmol/L, P=0.0008*; Leu: 128 ± 4.539 µmol/L vs. 156.5 ± 3.399 µmol/L, P<0.0001*, Val: 125.3 ± 5.372 µmol/L vs. 147.9 ± 3.156 µmol/L, P=0.0005*) while Lysine did not reach statistical significance (Lys: 208.8 ± 5.539 µmol/L vs. 219.2 ± 3.706 µmol/L, P=0.1236). As compared to the control group, although non-essential amino acids Glycine, Tyrosine, Glutamate, Serine and Ornithine were significantly lower in the patient group (Gly: 321 ± 18.87 µmol/L vs. 465.4 ± 16.57 µmol/L, P<0.0001; Tyr: 38.75 ± 2.504µmol/L vs. 48.4 ± 2.16 µmol/L, P=0.0041*; Glu: 242.6 ± 11.09 µmol/L Vs. 334.1 ± 7.448 µmol/L, P<0.0001; Ser: 210.1 ± 8.911 µmol/L vs. 248.4 ± 5.744 µmol/L, P=0.0005* and Orn: 88.99 ± 3.153 µmol/L vs. 114.5 ± 2.441 µmol/L, P<0.0001), Alanine and Proline didn't generate a significant P value (Ala: 166.3 ± 7.607µmol/L vs. 183.8 ± 4.536 µmol/L, P=0.0508; Pro: 149 ± 7.613 vs. 163.2 ± 5.035, P=0.1226). Besides, blood levels of Arginine and Aspartate were significantly higher in the patient group than the control

group (Arg: 26.07 ± 1.985 µmol/L vs. 16.17 ± 1.211 µmol/L, P < 0.0001 and Asp: 194.5 ± 17.54 µmol/L vs. 148.9 ± 9.968 µmol/L, P = 0.0258*). In addition to, there was no significant difference in Citruline levels between the two groups.

Table 3. 15. Comparison of concentrations of amino acids between hypothyroid patients and healthy controls.

Amino acids	Patient (Mean±SEM) Conc.(µmol/L)	Healthy (Mean±SEM) Conc.(µmol/L)	*p* value
Phenyl alanine	49.66 ± 1.626	65.56 ± 1.483	<0.0001*
Methionine	20.36 ± 0.8299	24.38 ± 0.8278	0.0008*
Leucine	128 ± 4.539	156.5 ± 3.399	<0.0001*
Valine	125.3 ± 5.372	147.9 ± 3.156	0.0005*
Lysine	208.8 ± 5.539	219.2 ± 3.706	0.1236
Glycine	321 ± 18.87	465.4 ± 16.57	<0.0001*
Tyrosine	38.75 ± 2.504	48.4 ± 2.16	0.0041*
Alanine	166.3 ± 7.607	183.8 ± 4.536	0.0508
Glutamate	242.6 ± 11.09	334.1 ± 7.448	<0.0001*
Aspartate	194.5 ± 17.54	148.9 ± 9.968	0.0258*
Arginine	26.07 ± 1.985	16.17 ± 1.211	<0.0001*
Serine	210.1 ± 8.911	248.4 ± 5.744	0.0005*
Proline	149 ± 7.613	163.2 ± 5.035	0.1226
Ornithine	88.99 ± 3.153	114.5 ± 2.441	<0.0001*
Citruline	21.41 ± 0.8501	22.45 ± 0.6368	0.3295

*** P< 0.05 was considered significant, SEM=Standard Error of Mean**

3.8.2. Metabolic and Nutritional State Indicator Amino Acid Ratios and Plasma Albumin Level

Metabolic and nutritional state can be determined using ratios between certain amino acids or groups of amino acids. The phenylalanine to tyrosine (Phe/Tyr) and Glycine to sum of branched chain amino acids (Gly/ΣBCAA) ratios can be used as indicator of catabolic state and protein intake, respectively [111-113]. We compared these ratios between patient group and healthy control group to assess the metabolic and nutritional state of the patient group. Catabolic state indicator amino ratio (Phe/Try) did not differ significantly between the patient group and the healthy control group (1.747 ± 0.1764 vs. 1.677 ± 0.122, P=0.739). Figure 3.25 indicates that indicator ratio of protein intake (Gly/ΣBCAA) and plasma Albumin level were

significantly lower in the patient group compared to the healthy control group (Gly/ΣBCAA: 0.5301 ± 0.01718vs. 1.456 ± 0.05663, P <0.0001). Since the amino acids level decreased in patient group compared to the healthy group, we tried to investigate the actual cause of amino acid deficiency and Plasma albumin. The level of plasma albumin was significantly lower in the patient group (3.174 ± 0.09163) compared to the healthy controls (3.655 ± 0.1066) with a p-value of <0.0009.

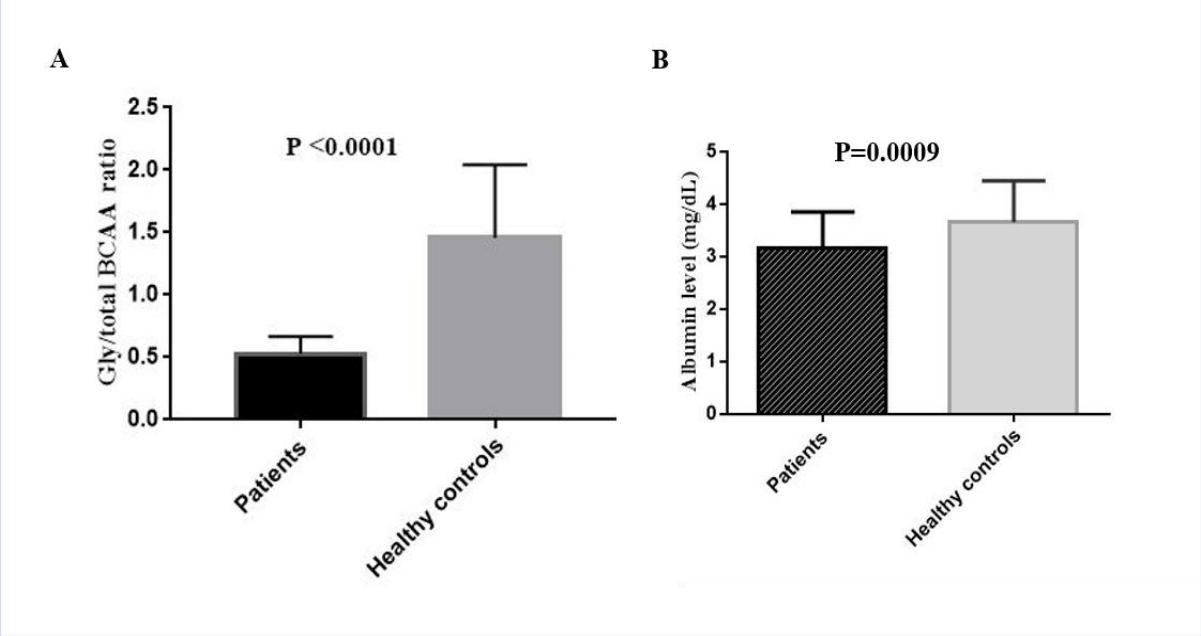


Figure 3. 25. Comparison of blood Gly/ΣBCAA indicator of protein intake ratios and plasma Albumin between the hypothyroid patient and the healthy controls. (A) Gly/ΣBCAA indicator ratio and (B) measurement of plasma Albumin level respectively

CHAPTER 4: DISCUSSION

4.1. Analysis of Genetic Spectrum of Congenital Hypothyroid Patients

Newborn screening for genetic and endocrine disorders is not commonly practiced in Bangladesh. Globally, Congenital Hypothyroidism (CH) is the most common preventable disorder but due to late diagnosis and initiation of treatment, hypothyroid patients experience various clinical complications even under levothyroxine drug treatment. Though there are at least seven genes reported to be involved in thyroid Dyshormonogenesis (TDH), TPO gene mutations are mostly associated with mild to severe consequences resulting in partial iodine organification defect (PIOD) to total iodine organification defect (TIOD) [30]. In this study, we explored (a) the mutational spectrum in the TPO gene of patients with Dyshormonogenesis and (b) the effect of specific mutations on 3D structures of monomer and dimer form of TPO protein using *in silico* approach. Further, we selected the patients with thyroid dysgenesis and analyzed TSHR and PAX8 gene. To the best of our knowledge, this is the first molecular analysis on genetic etiology of Congenital Hypothyroid children in Bangladesh.

In this study, we found a total of 12 mutations in TPO gene namely c.103C>G (p.35Glu>Glu) in Exon 2, c.404C>A (p.Pro135His) in Exon 5 c.860G>T (p.Leu287Leu) and c.769G>T (p.Ala257Ser) in Exon 7, c.1117G>T (p.Ala373Ser) and c.1193G>C (p.Ser398Thr) in Exon 8, c.1728G>A (p.Ala576Ala) in Exon 10, c.1998C>T (p.Asp666Asp) and c.2120G>A (p.Asp707Asp) mutations in Exon11, c.2145C>T (p.Pro715Pro) and c.2173A>C (p.Thr725Pro) in Exon12 and c.2540T>C (p.Val847Ala) in Exon 15. At first we analyzed Exons 8-14 of TPO gene of the TDH patients as the previous studies reported that this region was crucial for the enzymatic activity and mutations in this region may result in absence or reduction in TPO activity [29,114]. Next, we analyzed other regions of the TPO gene. Since 4 mutations in Exon 8 and Exon 12 were commonly presented in almost all the participants, we further proceeded for the computational analysis with three most common non-synonymous mutations including p.Ala373Ser (TPO MT1), p.Ser398Thr (TPO MT2) and p.Thr725Pro (TPO MT3) in the hotspot region of TPO Exons and we published the data very recently [31]. The non-synonymous mutations detected in this study had previously been reported to be

pathogenic or disease-causing [99,100]. Moreover, Guria *et. al.* [29] conducted a cloning-based study involving above-mentioned mutations and stated that these mutations could result in a low expression of TPO mRNAs as well as a reduction in TPO enzyme activity. It was also found that mutation p.Ala373Ser (TPO MT1) was more deleterious than mutation p.Ser398Thr (TPO MT2). This effect could be elucidated by a change in aliphatic amino acid to hydroxyl-group containing amino acid at 373th position of TPO protein, whereas the other mutation (p.Ser398Thr) was confined within the same group of amino acids substitution. Since threonine is the phosphorylation site for protein activation, the mutation p.Thr725Pro (TPO MT3) in Exon 12 could result in TPO protein inactivation [115,116]. Furthermore, p.Thr725Pro had been reported to be associated with autoimmune hypothyroidism [117]. Though the actual cause of CH is genetically and phenotypically heterogeneous, molecular studies may afford additional suggestions for diagnosis, classification and prognosis of the disease. Particularly, when the patients have normal thyroid gland, it's really hard to find out the reason of disorder whether it is due to thyrotropin resistance and Dyshormonogenesis. In this situation, genetic studies may expose the true etiology of the disease. In this study, at least one of these three damaging mutations was found in a homozygous state or two of them were in a heterozygous state in each of the study participants.

An amino acid substitution causing non-synonymous mutation can affect the shape, function or binding properties of a given protein. Now a days, much attention have been given in health sectors to study genetics and genomics and extensive efforts have been devoted to linking human phenotype to genotype variations at nucleotide level and associated changes on 3D structures of protein [118,119]**.** As the identified mutations were damaging, we desired to explore how these mutations were correlated with Dyshormonogenesis by affecting the structural integrity and function of TPO protein. Till now, there is no high resolution X-ray crystallographic structure for any of the wild type or mutant TPO proteins in the public databases. Conversely, there were structures for closely-related proteins, namely myeloperoxidases (MPO) and lactoperoxidases (LPO) [21,120]. Therefore, we modelled and validated the three dimensional (3D) structures of wild type and mutant TPO (WT and MT) proteins using different bioinformatics tools. The effects of mutations were considered on the structural features, specific domain arrangement and folding patterns of the TPO protein. Additionally, we performed molecular docking analysis to see the effect of mutations on Heme

binding catalytic site of TPO and to explore how these mutations could alter the binding affinity and activity of TPO protein. We analyzed the binding energies of Heme for both of the wild type and the mutant structures of TPO protein and found a reduction in binding affinities in the mutant structures compared to the wild type one. Further, we investigated the comprehensive non-bond interactions of Heme with specific amino acid residues in the catalytic site of TPO to understand the effects of mutations on the functions of TPO protein described in our earlier publication [31].

The present study aimed to apply the bioinformatics approaches on the full-length TPO protein (TPO_{1-933}) and the MPO-like domain of TPO ($TPO_{142-738}$) describing different non-bond interactions between the Heme group and amino acid residues since the identified mutations were found in the corresponding DNA sequence of this MPO-like catalytic domain of TPO protein [106]. When we obtained the predicted 3D structures of wild type and the mutant cases for $TPO_{142-738}$ from the I-TASSER server, all the structures showed significant confidence score (C-score), suggesting that the structures were valid for further studies [33]. However, the predicted 3D structures for full length TPO_{1-933} had lower C-score which might be due to large amino acid sequence of TPO protein because I-TASSER can predict structures based on iterative threading as well as homology modeling and TPO has no crystallographic structure which is still a challenge to be resolved [33,36]. All the 3D structures were validated using the Verify3D and the RAMPAGE web servers that provided acceptable outcomes [77,121].

As TPO depends on Heme exclusively for its catalytic activity, the molecular docking of Heme with the wild type and the mutant TPO structures could help us to understand the effects of mutations on the functions of TPO protein. The binding affinities of heme for the mutant cases were decreased compared to wild type TPO protein suggesting that the catalytic activity of mutant TPO might have been damaged. Further investigation provided the basic information about the Heme interactions with specific amino acid residues in the catalytic pocket of TPO protein. Molecular docking analysis showed that when the ligand (Heme) interacted with wild type structures, it worked together with all the crucial amino acid residues including His239, Arg396, Glu399 and His494 through non-bond interactions. But in the cases of mutant structures, some of the essential amino acid interactions were absent, suggesting that the mutations might have harmful effects on Heme interactions and thus distressing the catalytic

activity of TPO enzyme. Guria *et al*. [29] verified that MT1 (p.Ala373Ser), MT2 (p.Ser398Thr), and MT3 (p.Thr725Pro) had damaging effect on TPO mRNA expression and enzyme activity. However, MT1 (p.Ala373Ser) and MT3 (p.Thr725Pro) showed iodination reactions similar to the non-enzymatic reaction rate, suggesting that these two mutations were more damaging than the MT2 (p.Ser398Thr) which showed more efficient iodination reaction than MT1 and MT3 but less efficient than the wild type TPO protein [29]. This finding is consistent with our molecular docking-based study targeting the predicted structures of both MPO-like domain of TPO ($TPO_{142-738}$) and the full-length TPO protein (TPO_{1-933}) mentioned in previous data [31]. MT1 and MT3 had more influence on the interactions between several crucial residues and the Heme ligand, whereas MT2 had less influence on the interaction between the TPO protein and the Heme prosthetic group. Additionally, different studies showed that quantum chemical methods such as DFT or QM/MM could also be employed for investigating the intermolecular interactions of protein and ligands and other molecules [122-124]. So that, we performed QM/MM calculations to validate the molecular docking results of wild type TPO protein compared to mutant proteins (MT1 and MT3) that had severe damaging effect on the catalytic activity of TPO.

During the QM/MM calculations, the Heme ligand was treated using a semi-empirical approach of calculations and the protein was treated using molecular mechanical approach of calculations. After completing the QM/MM calculations, the structures were taken into consideration for further analysis of the presence of non-bond interactions of the active site residues with the Heme group. The identified interactions were exactly similar to those obtained from molecular docking analysis, proposing that the obtained docking results were valid.

In vivo, the protein structures always exist in a dynamic state. To create and mimic the cellular atmosphere, we performed molecular dynamics simulation with a cubic simulation box containing water molecules and NaCl having a pH of 7.4. After performing MD simulation for 5000 ps under the influence of AMBER14 force field, the ultimate protein structures were superimposed on the corresponding molecular docked protein structure. The results inferred that the studied protein-ligand complexes were considerably stable over time though there was very small change in the coordinates of the Heme ligand after the simulation.

The 3D structures of protein were retrieved from the 5ns simulation snapshot and re-docked with the Heme ligand in order to inspect the protein dynamics simulation. The analyzed results indicated that the wild type protein was found to be interacted with the crucial amino acids presented in catalytic domain of TPO protein whereas, mutants proteins had lack of these crucial interactions that might cause a reduction of TPO enzyme activity.

In our study, we performed *in silico* approach to mimic cellular environment for better understanding the features of wild type and mutant TPO protein structures. Nevertheless, different computational chemistry-based methods such as IR and Raman spectroscopy could detect the alterations of protein molecules in molecular level with higher specificity [125,126]. To observe the biochemical changes associated with different disease conditions and characterizing spectrums of several hormones like corticosteroid, these spectroscopic analyses become an essential part within cellular environment [125-127]. A study conducted by Claudio *et al.* [128] investigated the molecular vibrational spectrum of thyroid tissues from healthy subjects and disease conditions which could eventually signify the features of secondary structures of proteins. Even though the use of IR and Raman spectroscopy could offer more perceptions into the physiological conditions of patients having mutations in the TPO gene, the present study was not subjected to such approaches due to unavailability of these facilities. With our best capacities, we performed computational approaches such as molecular docking, QM/MM and molecular dynamics simulation to explore the effect of specific mutations on TPO protein interactions in monomer form.

Further, we tried to explore the effect of mutations on dimer structures of TPO wild type, MT1 and MT3 proteins. To do so, we used Symmdock server for dimer modeling as it provides cyclically symmetric result for homomultimers. We performed Molecular dynamics (MD) simulation for 5000 ps to understand the actual phenomena of structural and functional properties of TPO dimer protein and membrane was attached during simulation process since TPO is a membrane bound protein. The simulation was performed according to the procedure mentioned earlier for TPO monomer study. MD simulation showed that WT dimer was more stable than mutant dimers by analyzing the data from van der Walls, solvent accessible surface area and molecular surface area. The enzymatic processes for the biosynthesis of T_3 and T_4 hormone require the presence of H_2O_2, iodide and Thyroglobulin (Tg). To further investigate

the effect of TPO mutation on the enzymatic process in the thyroid hormone synthesis pathway, the study was designed to explore the interaction of H_2O_2 and Iodide (I^-) simultaneously with the wild type and mutant dimer structures. To understand the interaction between H_2O_2 and dimer structures, we performed sequential docking for each monomer (left- and right-side monomer of the dimer form). The Molecular docking analysis showed that though binding affinities for WT and MT structures were close but TPO WT homodimer (HD) showed higher binding affinity for H_2O_2 than mutant dimers and TPO MT1-HD showed lowest binding affinity for H_2O_2. We also observed that the binding affinities of Iodide (I^-) for wild type and mutant homodimers were close to each other and the TPO WT-HD and TPO MT1-HD showed the same binding affinity, whereas TPO MT3-HD showed relatively less binding affinity than those structures. Although the docking approaches are more efficient to understand how the halogenated ligand can bind to some specific targets, but almost all the docking scoring functions are not capable to model the XB (halogen bonding) in a correct way, leading to some errors that we can interpret as false positives or vice versa [129].

From the present study, we have found that the binding pocket of TPO protein for essential ligands was consistent with the monomer study and effect of mutations on TPO protein can severely affect the monomer structure compared to the dimer structure. The effect of mutation and ligand (H_2O_2, I^-) interactions with TPO dimer protein might be regulated by other cellular processes such as post translational processes, intracellular trafficking and targeting.

A literature search on the TPO gene in hypothyroid patients demonstrated that most of the mutations were confined between Exon 8 and Exon 14 and very few mutations had been identified outside this region [114]. Then we tried to set a method to detect the common mutations detected in TPO gene in Bangladeshi patients. Although the mutations had been detected as homozygous or heterozygous states, our study had confirmed that both alleles of the TPO gene of all 44 hypothyroid patients had been affected by mutations, further confirming the recessive pattern of this disease. This is why we confined our further study on TPO gene mutation analysis using the target region between Exon 8 to Exon 14. The identified non-synonymous mutations had previously been reported to be pathogenic or disease-causing [99,100]. The present study had identified a total of four mutations in the hotspot region from Exon 8 to Exon 12 of TPO gene by Sanger sequencing. Since we targeted this region to find

out mutations in Bangladeshi patients, we aimed to establish a supplementary method to screen this region. In Bangladesh, there is very little information about newborn screening and genetic etiology of CH is also unknown here. High resolution melt (HRM) curve analysis method represents a significant advancement in mutation detection over the years. Recently, HRM method has been established for detection of variants of beta-globin gene in thalassemia patients in Bangladesh [54]. Moreover, HRM approach was sensitive as well as easy-to-use molecular method for detection of another genetic disorder named G6PD (Glucose 6-phosphate dehydrogenase) deficiency in Bangladesh [130]. The present study aimed to establish a faster, reliable and cheaper molecular-based approach to detect the common mutations in the TPO gene responsible for thyroid Dyshormonogenesis in Bangladeshi CH patients and to see whether the approach could be used as a regular screening method to identify the genetic etiology of Dyshormonogenesis. Since TPO gene mutation is inherited in an autosomal recessive manner to cause Dyshormonogenesis, its carrier state can also be investigated by this HRM method.

The genetic study is important to investigate the cause of CH. There are some screening methods for diagnosis of CH, such as measurement of serum/blood TSH, T_3 and T_4. However, these approaches can only confirm the CH cases but not the actual etiology. That is, the conventional screening method of CH cannot say whether it is acquired or genetic. If the actual etiology is known, then duration of treatment can be defined based on the causes. If it is due to a genetic cause, the patients should be followed up because these patients are kept under levothyroxine treatment for their whole life. On the other hand, treatment should be continued for the first 3 years of life for acquired cause [52]. So, the treatment strategy will be different for CH cases with genetic etiology and other reasons of CH with Dyshormonogenesis. If genetic basis of CH is defined in the country, carrier screening is possible targeting the underlying genetic cause. If the parents are found to be carriers of CH involving the TPO gene, their children or newborns could be screened and appropriate measures can be taken, such as early initiation of treatment which would help preventing mental retardation.

In this study, we identified 3 non-synonymous common mutations in the hotspot region of TPO gene detected by Sanger sequencing in Bangladeshi CH patients with Dyshormonogenesis. The non-synonymous mutations were c.1117G>T, c.1193G>C and c.2173A>C. To establish

the HRM method, we targeted these three mutations and we designed primers targeting the hotspot region of TPO gene. Three sets of primers were designed to keep the product size between 65-101 base pair which were within the required product size for HRM strategy [131]. The first two mutations were within Exon 8 and the 3rd one was present in the Exon 12 of TPO gene. A study was conducted to see the functional activity of TPO enzyme due to mutations which described that c.1117G>T and c.2173A>C showed non-enzymatic reaction rate and c.1193G>C showed slightly reduced enzymatic reaction rate compared to the wild type TPO protein [29]. To conduct the HRM curve analysis, the samples with heterozygous, homozygous and wild type alleles were subjected to experimental analysis. For the first set of primers which targeted the mutation c.1117G>T; both homozygous and heterozygous states were clearly distinguishable from wild type allele (Figure 3.12 and 3.13).

However, c.1193G>C mutation was much more difficult to differentiate due to formation of similar number of hydrogen bonds for G>C substitution and thus similar level of bond energy was involved for both the wild type and mutant variants. To overcome the difficulties, the optimum concentration of $MgCl_2$ was determined to be 8mM for detection of a single G>C point mutation by HRM because 8mM $MgCl_2$ concentration could clearly distinguish among homozygous, heterozygous and wild type alleles. This showed that $MgCl_2$ could have an effect on HRM study to differentiate different states of mutation of G/C alleles. For c.1193G>C variant, the melt curve showed almost similar patterns among the samples with the wild type allele and also samples with homozygous and heterozygous alleles. However, temperature shifted curve could clearly differentiate all the states. Different studies demonstrated that purine to purine or pyrimidine to pyrimidine nucleotide substitution, such as A to T or G to C substitution could not be distinguished by HRM method due to the same melting temperature [132-134].

For the third mutation c.2173A>C, Adenine nucleotide was substituted by Cytosine nucleotide and due to the difference in bond energy between Purine and Pyrimidine group, melting temperature was shifted for both heterozygous and homozygous states compared to the wild type state. The temperature shifted pattern was differentiated in such a manner so that the wild type had lower Tm pattern compared to the heterozygous and homozygous states.

Although, Sanger sequencing is the gold standard for mutation detection, HRM can be a fast and less expensive supplemental approach with 100% sensitivity and specificity for screening and detection of mutations in the TPO gene in Bangladeshi patients where the frequency of genetic etiology of CH patients with Dyshormonogenesis seems to be almost 100 percent.

On the other hand, we also focused on the etiology of Dysgenesis types of CH patients having small glands or ectopic gland or agenesis (absent of thyroid gland). Different studies suggested that TSHR was the major gene responsible for growth and development of thyroid gland [8,135]. TSH binds to the receptor and creates the signaling pathway through G-protein coupled-receptor and Cyclic AMP-mediated adenylate cyclase. The full-length protein structure of TSHR is still under investigation through crystallography. Only a low small extracellular part of TSHR is described by crystallography. Mutations in the TSHR gene results from loss or gain of function of the protein that causes different phenotypic variations and lead to hyperthyrotopinemia to severe Congenital Hypothyroidism [136-138].

Analysis of TSHR gene showed that two mutations were found, namely c.1523C>T and c.2181G>C in the patients and we analyzed the effect of mutations by using different bioinformatics tools. Almost all the tools were very much popular to analyze the mutational effect such as Polyphen 2, Mutation Taster and PROVEAN. The mutation c.1523C>T was found to be damaging, disease causing or deleterious and c.2181G>C was found to be benign or neutral. Since MS437 and MS438 are thyrogenic potent molecules, we selected these two molecules as ligands for molecular docking with the wild type and mutant structures of TSHR protein [107]. The molecular docking analysis showed that the binding affinity for both of the ligands with mutant cases were decreased compared to the wild type TSHR protein. There is no published data based on c.1523C>T mutation but c.2181G>C was found as a normal variant. A study showed that although Asp727Glu was found as a normal variant it was also identified in follicular thyroid carcinomas [139]. The RMSD value from MD simulation indicated that when mutations in the transmembrane region of TSHR gene occurs, MS437-bound structures were more unstable compared to the MS438 molecule. Moreover, irrespective of wild type or mutant cases, MS437 didn't interact with the crucial amino acids. In mutant cases, MS438 interacted with all the crucial amino acids including Thr134, Ser138 and Glu139 which suggest

that mutation can affect more severely for MS437-bound than MS438-bound 3D structures of TSHR protein.

Since PAX8 is also associated with thyroid Dysgenesis, we tried to investigate the mutations in this gene but we didn't get any mutation in the PAX8 gene [9]. Due to limitation of resource, we only sequenced Exon 3 of PAX8 gene. Future studies on other region of the PAX 8 gene and the remaining genes responsible for CH might explore the full spectrum of the genetic etiology of Congenital Hypothyroidism in Bangladeshi patients.

4.2. Metabolic Profiling of Congenital Hypothyroid Children

This is the first LC-MS/MS-based metabolic profiling report from Bangladesh where carnitine-acylcarnitines and amino acids metabolites in the blood of hypothyroid patients under thyroid hormone replacement therapy. The study aimed to inquire whether any alterations of blood carnitine-acylcarnitine and amino acid levels in the hypothyroid patients under treatment had correlation with their fatigue-related complications.

Although the recommended time for diagnosis and therapeutic intervention should be initiated between the first two and three weeks of life, the average treatment initiation time for the patients of this study was 19.7 ± 4.72 months [140,141]. Clinical complications other than neurologic deficits, if any, in the hypothyroid patients are considered reversible upon hormone replacement therapy [1,142-144]. Unfortunately, although all the CH patients in the study had a normal blood TSH and FT_4 levels upon daily LT_4 treatment, they all reported at least one of the hypothyroidism-associated symptoms, with fatigue and fatigue-related symptoms as the most common complications.

Wong *et al.* [145] reported that there were no apparent differences in serum acylcarnitines profiles among hypo-, hyper- and euthyroid states and inferred that the acylcarnitines profiles were relatively unremarkable in thyroid disease. Interestingly, in our study, the levels of free carnitine, total carnitines and acylcarnitines were higher in the patient group than that of the healthy controls, although the difference was significant for free carnitine only. Particularly, the total concentration of long chain acylcarnitines along with the concentrations of six individual long chain acylcarnitines were significantly lower in the patients compared to the controls. The present study differs from the study by Wong *et al.* [145] that we measured

carnitine parameters in the whole blood using DBS specimens and the comparisons of the parameters were made between the late-diagnosed hypothyroid patients under treatment and the healthy controls. In addition, small sample size and incomplete follow-up were stated as important limitations of their study. However, Galland *et al*. [146] demonstrated that hepatic carnitine level in rat was modified due to changes in thyroid hormone status, while the carnitine concentrations of kidney and skeletal muscle cells remained unchanged. According to Pande *et al.* [147] dietary thyroxine could cause an increase in carnitine level in the liver and serum, while carnitine levels in the heart, skeletal muscle and the urinary excretion of carnitine remained unaffected in experimental rats. Similarly, thyroxine treatment had also been reported to cause an increase in carnitine in hearts of guinea pigs [148]. However, a direct comparison of these results with our findings of increased free carnitine in whole blood is difficult due to differences among species as well as variation in duration and dose of the thyroid hormone replacement therapy. Maebashi *et al.*[149] found a positive correlation between urinary carnitine excretion and serum thyroxine concentration, with hyperthyroid subjects having significantly elevated levels of carnitine excretion and hypothyroid subjects having markedly reduced carnitine excretion. Urinary carnitine excretion became normal in both groups with correction of thyroid status. Nevertheless, renal clearance and absorption of carnitine were not investigated in our study. The ratio of urinary carnitine/acylcarnitines excretion of the patients and healthy controls could have helped us to understand better the altered carnitine-acylcarnitines homeostasis in these late-diagnosed hypothyroid patients under treatment.

The present study showed that the mean concentration of total LC-acylcarnitines was significantly lower in the patients compared to the healthy group (P=0.0014). A study of chronic fatigue syndrome patients showed a significantly lower level of LC-acylcarnitines compared to the healthy subjects, although there was no significant differences in free carnitine, total carnitine or total acylcarnitines levels between the two groups [64]. Another study also reported a small differences in free carnitine, total carnitine and total acylcarnitines concentrations between the chronic fatigue syndrome patients and the healthy controls [150]. But the findings of An *et al.* [151] made it more paradoxical as they reported that carnitine supplementation could significantly lessen both fatigue severity scale and mental fatigue score of hypothyroid patients younger than 50 years of age and thus they concluded that carnitine

supplementation may be useful in alleviating fatigue symptoms in this group. Since fatty acid oxidation depends on the presence of carnitine, an increase in carnitine level may cause an increase in rate of fatty acid oxidation and thus contribute to energy supply. Thyroid hormone controls the hepatic conversion of γ-butyrobetaine to carnitine in rat and also increase carnitine bioavailability [146]. Elevated carnitine levels had been shown to stimulate fatty acid oxidation through mechanisms involving increased CPT-I activity. Interestingly, in this study, although the carnitine level was high in the blood of patients, a decrease in CPT-I activity was found in these hypothyroid patients with energy crisis having normal TSH and FT_4 levels after hormone replacement therapy. However, several studies reported carnitine as a peripheral antagonist of thyroid hormone action and may act as a peripheral inhibitor of thyroid hormone uptake [152-154]. This inhibitory effect may be part of a retro-control loop between thyroid hormone and carnitine and thus the reduced thyroid hormone at cellular level may hinder CPT-I activity [155]. Study of cellular hypothyroidism in different tissues is in requisite in order to minimize the discrepancy between findings of different study groups and also to devise a proper management strategy for hypothyroid patients with fatigue and related symptoms.

In this study, the levels of six LC-acylcarnitines as well as the mean concentration of total LC-acylcarnitines were significantly lower in the patients than in the controls, although the levels of short and medium chain acylcarnitines remained unaffected. The β-oxidation was found in a reduced rate, which was probably due to a decrease in CPT-I activity. These results motivated us to measure the plasma TG level of the patients. The hypothyroid patients had significantly higher levels of plasma TG compared to the healthy controls and the finding was supported by other studies [156-158]. The excess TG usually deposits inside the cells of adipose tissues and liver as lipid droplets [159]. However, none of the patients enrolled in this study was overweight (BMI= 17.0±4.4 kg/m^2). Also, the lean body mass did not differ between the hypothyroid patients and the healthy controls (data not shown), suggesting that the hepatocytes might be the primary target of TG deposition in these patients. A wide range of studies reported both overt and subclinical hypothyroidism as a risk factor of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH), the chronic liver diseases which might progress to cirrhosis and hepatocellular carcinoma [160-162], although some newly published studies failed to find such an association [162,163]. It had been demonstrated that acylcarnitines levels were significantly higher in the whole blood than in the plasma because

higher concentrations of LC-acylcarnitines in the erythrocyte membranes could contribute significantly to the whole blood acylcarnitines [164,165]. However, plasma acylcarnitines concentrations might have a pattern of variation similar to the one observed in the whole blood specimens. Finally, there are some limitations in the study, notably urinary excretion levels of acylcarnitines of the patients were not measured. Targeting more acylcarnitines metabolites might have helped to delve further into the patients' ongoing fatigue-related complications.

In the study of CPT1 gene expression, it was shown that CPT1B level was decreased in the patient group compared to the healthy control group, whereas CPT1A was higher in the patient group compared to the healthy controls group. The liver isoform, CPT1A expressed in almost all cells of the body whereas CPT1B is expressed in heart and skeletal muscles [166-168]. CPT1 activity is inhibited by malonyl-CoA and this is good target for treatment of metabolic disorder [169]. The higher level of CPT1A gene expression and the ratio CPT1A/CPT1B in patients might be due to other factors including old samples collected more than 2 years ago in a trizol containing tube and freshly collected samples can give the actual data of CTP1 activity. Future study on investigating mutations through sanger sequencing of CPT1 gene can reveal the defective situation appeared in our study participants.

Additionally, thyroid hormones control the amino acid metabolism in different ways [72,110]. In the present study, we found a decrease in levels of 9 amino acids in the CH patients compared to the healthy subjects, although their TSH and T_4 levels were normal under LT_4 treatment. Plasma albumin level was also significantly decreased in the patient group compared to the healthy group which suggested a decrease level of protein synthesis in the liver due to thyroid disorder [170]. This might be due to any disturbance related to exposure of the pathway of protein metabolism in hypothyroid condition. Catabolic state indicator amino acids ratio (Phe/Try) didn't differ between the two groups, whereas the Glycine to sum of branched chain amino acids (Gly/ΣBCAA) ratios indicated a decreasing trend in protein intake [113,171,172]. These alteration of amino acid level might be also affected by nutritional status of the study participants.

Finally, it can be discussed that hypothyroid condition affects the body's overall metabolism and explore a new era to investigate more factors associated with the clinical complications of the patients even under treatment.

CHAPTER 5: CONCLUSION

The present study investigated the genetic etiology of Congenital Hypothyroid children focusing on 3 major genes TPO, TSHR and PAX8 in conditions of both Dyshormonogenesis and Dysgenesis. In Bangladesh, due to late diagnosis, the patients experience clinical complications even under treatment. The metabolic profiling of the patients can reveal the correlation between complications and the level of amino acids and acylcarnitines. These analyses can help to better understand the clinical features as well as any alteration of treatment strategy for individual patients under the guidance of the physicians and ultimately it will focus on the importance of Newborn screening for Congenital Hypothyroidism.

CHAPTER 6: REFERENCES

1. Rastogi MV, LaFranchi SH (2010) Congenital hypothyroidism. Orphanet journal of rare diseases 5: 17.
2. Fisher DA (1983) Second International Conference on Neonatal Screening: progress Report. J Pediatr 102: 653-654.
3. Hasan M, Nahar N, Ahmed A, Moslem F (2004) Screening for congenital hypothyroidism-a new era in Bangladesh. SOUTHEAST ASIAN JOURNAL OF TROPICAL MEDICINE AND PUBLIC HEALTH 34: 162-164.
4. Selva KA, Harper A, Downs A, Blasco P, Lafranchi S (2005) Neurodevelopmental outcomes in congenital hypothyroidism: comparison of initial T4 dose and time to reach target T4 and TSH. The Journal of pediatrics 147: 775-780.
5. Selva KA, Mandel SH, Rien L, Sesser D, Miyahira R, et al. (2002) Initial treatment dose of L-thyroxine in congenital hypothyroidism. The Journal of pediatrics 141: 786-792.
6. Lee CC, Harun F, Jalaludin MY, Lim CY, Ng KL, et al. (2014) Functional analyses of c. 2268dup in thyroid peroxidase gene associated with goitrous congenital hypothyroidism. BioMed research international 2014.
7. Cangul H, Aycan Z, Olivera- Nappa A, Saglam H, Schoenmakers NA, et al. (2013) Thyroid dyshormonogenesis is mainly caused by TPO mutations in consanguineous community. Clinical endocrinology 79: 275-281.
8. Cangul H, Aycan Z, Saglam H, Forman JR, Cetinkaya S, et al. (2012) TSHR is the main causative locus in autosomal recessively inherited thyroid dysgenesis. Journal of Pediatric Endocrinology and Metabolism 25: 419-426.
9. Macchia PE, Lapi P, Krude H, Pirro MT, Missero C, et al. (1998) PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis. Nature genetics 19: 83.
10. Taurog A, Dorris ML, Doerge DR (1996) Mechanism of simultaneous iodination and coupling catalyzed by thyroid peroxidase. Archives of biochemistry and biophysics 330: 24-32.

11. Endo Y, Onogi S, Fujita T (1995) Regional localization of the gene for thyroid peroxidase to human chromosome 2p25 and mouse chromosome 12C. Genomics 25.
12. Banga JP, Mahadevan D, Barton GJ, Sutton BJ, Saldanha JW, et al. (1990) Prediction of domain organisation and secondary structure of thyroid peroxidase, a human autoantigen involved in destructive thyroiditis. FEBS letters 266: 133-141.
13. Taurog A, Dorris M, Doerge DR (1994) Evidence for a radical mechanism in peroxidase-catalyzed coupling. I. Steady-state experiments with various peroxidases. Archives of biochemistry and biophysics 315: 82-89.
14. Godlewska M, Gardas A, Góra M Biosynthesis, post-translational modifications, maturation, and physiological function of human thyroid peroxidase** Biosynteza, modyfikacje postranslacyjne, dojrzewanie i fizjologiczna funkcja ludzkiej peroksydazy tarczycowej.
15. Furtmüller PG, Zederbauer M, Jantschko W, Helm J, Bogner M, et al. (2006) Active site structure and catalytic mechanisms of human peroxidases. Archives of biochemistry and biophysics 445: 199-213.
16. Taurog A (1999) Molecular evolution of thyroid peroxidase. Biochimie 81: 557-562.
17. Bakker B, Bikker H, Vulsma T, de Randamie JS, Wiedijk BM, et al. (2000) Two decades of screening for congenital hypothyroidism in The Netherlands: TPO gene mutations in total iodide organification defects (an update). The Journal of Clinical Endocrinology & Metabolism 85: 3708-3712.
18. Belforte FS, Miras MB, Olcese MC, Sobrero G, Testa G, et al. (2012) Congenital goitrous hypothyroidism: mutation analysis in the thyroid peroxidase gene. Clinical endocrinology 76: 568-576.
19. Hendry E, Taylor G, Ziemnicka K, Grennan Jones F, Furmaniak J, et al. (1999) Recombinant human thyroid peroxidase expressed in insect cells is soluble at high concentrations and forms diffracting crystals. Journal of endocrinology 160: R13.
20. Fiedler TJ, Davey CA, Fenna RE (2000) X-ray crystal structure and characterization of halide-binding sites of human myeloperoxidase at 1.8 Å resolution. Journal of Biological Chemistry 275: 11964-11971.

21. Gardas A, Sohi M, Sutton B, McGregor A, Banga J (1997) Purification and crystallisation of the autoantigen thyroid peroxidase from human Graves' thyroid tissue. Biochemical and biophysical research communications 234: 366-370.
22. Singh AK, Singh N, Sharma S, Singh SB, Kaur P, et al. (2008) Crystal structure of lactoperoxidase at 2.4 Å resolution. Journal of molecular biology 376: 1060-1075.
23. Le SN, Porebski BT, McCoey J, Fodor J, Riley B, et al. (2015) Modelling of thyroid peroxidase reveals insights into its enzyme function and autoantigenicity. PloS one 10: e0142615.
24. Bergendahl LT, Gerasimavicius L, Miles J, Macdonald L, Wells JN, et al. (2019) The role of protein complexes in human genetic disease. Protein Science.
25. Kotani T, Umeki K, Kawano Ji, Suganuma T, Hishinuma A, et al. (2003) Partial iodide organification defect caused by a novel mutation of the thyroid peroxidase gene in three siblings. Clinical endocrinology 59: 198-206.
26. Rivolta CM, Esperante SA, Gruñeiro- Papendieck L, Chiesa A, Moya CM, et al. (2003) Five novel inactivating mutations in the thyroid peroxidase gene responsible for congenital goiter and iodide organification defect. Human mutation 22: 259-259.
27. Rodrigues C, Jorge P, Soares JP, Santos I, Salomao R, et al. (2005) Mutation screening of the thyroid peroxidase gene in a cohort of 55 Portuguese patients with congenital hypothyroidism. European Journal of Endocrinology 152: 193-198.
28. Wu J, Shu S, Yang C, Lee C, Tsai F-J (2002) Mutation analysis of thyroid peroxidase gene in Chinese patients with total iodide organification defect: identification of five novel mutations. Journal of Endocrinology 172: 627-635.
29. Guria S, Bankura B, Balmiki N, Pattanayak AK, Das TK, et al. (2014) Functional analysis of thyroid peroxidase gene mutations detected in patients with thyroid dyshormonogenesis. International journal of endocrinology 2014.
30. Bikker H, Baas F, de Vijlder JJ (1997) Molecular analysis of mutated thyroid peroxidase detected in patients with total iodide organification defects. The Journal of Clinical Endocrinology & Metabolism 82: 649-653.
31. Begum M, Islam MT, Hossain SR, Bhuyan GS, Halim MA, et al. (2019) Mutation Spectrum in TPO Gene of Bangladeshi Patients with Thyroid Dyshormonogenesis and

Analysis of the Effects of Different Mutations on the Structural Features and Functions of TPO Protein through In Silico Approach. BioMed research international 2019.

32. Mobaraki N, Hemmateenejad B, Weikl TR, Sakhteman A (2019) On the relationship between docking scores and protein conformational changes in HIV-1 protease. Journal of Molecular Graphics and Modelling 91: 186-193.
33. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC bioinformatics 9: 40.
34. Baker D, Sali A (2001) Protein structure prediction and structural genomics. Science 294: 93-96.
35. Zhang Y (2009) Protein structure prediction: when is it useful? Current opinion in structural biology 19: 145-155.
36. Yang J, Yan R, Roy A, Xu D, Poisson J, et al. (2015) The I-TASSER Suite: protein structure and function prediction. Nature methods 12: 7.
37. Karafyllidis IG (2008) Quantum mechanical model for information transfer from DNA to protein. Biosystems 93: 191-198.
38. Wang X, Li Y, Gao Y, Yang Z, Lu C, et al. (2018) A quantum mechanical computational method for modeling electrostatic and solvation effects of protein. Scientific reports 8: 5475.
39. Ainsley J, Lodola A, Mulholland AJ, Christov CZ, Karabencheva-Christova TG (2018) Combined Quantum Mechanics and Molecular Mechanics Studies of Enzymatic Reaction Mechanisms. Advances in protein chemistry and structural biology 113: 1-32.
40. Sousa SF, Carvalho ES, Ferreira DM, Tavares IS, Fernandes PA, et al. (2009) Comparative analysis of the performance of commonly available density functionals in the determination of geometrical parameters for zinc complexes. Journal of computational chemistry 30: 2752-2763.
41. Easton RE, Giesen DJ, Welch A, Cramer CJ, Truhlar DG (1996) The MIDI! basis set for quantum mechanical calculations of molecular geometries and partial charges. Theoretica chimica acta 93: 281-301.
42. Morris GM, Lim-Wilby M (2008) Molecular docking. Molecular modeling of proteins: Springer. pp. 365-382.

43. Sousa SF, Fernandes PA, Ramos MJ (2006) Protein–ligand docking: current status and future challenges. Proteins: Structure, Function, and Bioinformatics 65: 15-26.

44. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) A geometric approach to macromolecule-ligand interactions. Journal of molecular biology 161: 269-288.

45. Pikkemaat MG, Linssen AB, Berendsen HJ, Janssen DB (2002) Molecular dynamics simulations as a tool for improving protein stability. Protein engineering 15: 185-192.

46. Salsbury Jr FR (2010) Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. Current opinion in pharmacology 10: 738-744.

47. Dror RO, Jensen MØ, Borhani DW, Shaw DE (2010) Perspectives on: molecular dynamics and computational methods: exploring atomic resolution physiology on a femtosecond to millisecond timescale using molecular dynamics simulations. The Journal of General Physiology 135: 555.

48. Land H, Humble MS (2018) YASARA: a tool to obtain structural guidance in biocatalytic investigations. Protein Engineering: Springer. pp. 43-67.

49. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic acids research 33: W363-W367.

50. Eun C, Kekenes-Huskey PM, Metzger VT, McCammon JA (2014) A model study of sequential enzyme reactions and electrostatic channeling. The Journal of chemical physics 140: 03B607_601.

51. Cattoni DI, Chara O, Kaufman SB, Flecha FLG (2015) Cooperativity in binding processes: new insights from phenomenological modeling. PloS one 10: e0146043.

52. Büyükgebiz A (2006) Newborn screening for congenital hypothyroidism. Journal of Pediatric Endocrinology and Metabolism 19: 1291-1298.

53. Yan J-b, Xu H-p, Xiong C, Ren Z-r, Tian G-l, et al. (2010) Rapid and reliable detection of glucose-6-phosphate dehydrogenase (G6PD) gene mutations in Han Chinese using high-resolution melting analysis. The Journal of Molecular Diagnostics 12: 305-311.

54. Islam MT, Sarkar SK, Sultana N, Begum MN, Bhuyan GS, et al. (2018) High resolution melting curve analysis targeting the HBB gene mutational hot-spot offers a reliable screening approach for all common as well as most of the rare beta-globin gene mutations in Bangladesh. BMC genetics 19: 1.

55. Bono C, Nuzzo D, Albeggiani G, Zizzo C, Francofonte D, et al. (2011) Genetic screening of Fabry patients with EcoTILLING and HRM technology. BMC research notes 4: 323.

56. Bataille S, Berland Y, Fontes M, Burtey S (2011) High Resolution Melt analysis for mutation screening in PKD1 and PKD2. BMC nephrology 12: 57.

57. Nagayama Y, Kaufman KD, Seto P, Rapoport B (1989) Molecular cloning, sequence and functional expression of the cDNA for the human thyrotropin receptor. Biochemical and biophysical research communications 165: 1184-1190.

58. KAKINUMA A, NAGAYAMA Y (2002) Multiple messenger ribonucleic acid transcripts and revised gene organization of the human TSH receptor. Endocrine journal 49: 175-180.

59. De Roux N, Misrahi M, Brauner R, Houang M, Carel J, et al. (1996) Four families with loss of function mutations of the thyrotropin receptor. The Journal of Clinical Endocrinology & Metabolism 81: 4229-4235.

60. Park S, Chatterjee V (2005) Genetics of congenital hypothyroidism. Journal of medical genetics 42: 379-389.

61. Saravanan P, Chau WF, Roberts N, Vedhara K, Greenwood R, et al. (2002) Psychological well- being in patients on 'adequate'doses of l- thyroxine: results of a large, controlled community- based questionnaire study. Clinical endocrinology 57: 577-585.

62. Wekking EM, Appelhof BC, Fliers E, Schene AH, Huyser J, et al. (2005) Cognitive functioning and well-being in euthyroid patients on thyroxine replacement therapy for primary hypothyroidism. European journal of endocrinology 153: 747-753.

63. Wiersinga WM, Duntas L, Fadeyev V, Nygaard B, Vanderpump MP (2012) 2012 ETA guidelines: the use of L-T4+ L-T3 in the treatment of hypothyroidism. European thyroid journal 1: 55-71.

64. Reuter SE, Evans AM (2011) Long- chain acylcarnitine deficiency in patients with chronic fatigue syndrome. Potential involvement of altered carnitine palmitoyltransferase- I activity. Journal of internal medicine 270: 76-84.

65. Bieber L (1988) Carnitine. Annual review of biochemistry 57: 261-283.

66. Scholte H, Rodrigues RP, Luyt-Houwen I, Hedwig M, Verduin M, et al. (1990) Primary carnitine deficiency. Journal of clinical chemistry and clinical biochemistry Zeitschrift fur klinische Chemie und klinische Biochemie 28: 351-357.

67. Famularo G, Matricardi F, Nucera E, Santini G, De Simone C (1997) Carnitine deficiency: primary and secondary syndromes. Carnitine Today: Springer. pp. 119-161.

68. Pons R, De Vivo DC (1995) Primary and secondary carnitine deficiency syndromes. Journal of Child Neurology 10: 2S8-2S24.

69. Duranay M, Akay H, Yilmaz FM, Şeneş M, Tekeli N, et al. (2006) Effects of L-carnitine infusions on inflammatory and nutritional markers in haemodialysis patients. Nephrology Dialysis Transplantation 21: 3211-3214.

70. Rizos I (2000) Three-year survival of patients with heart failure caused by dilated cardiomyopathy and L-carnitine administration. American heart journal 139: s120-s123.

71. Brevetti G, Chiariello M, Ferulano G, Policicchio A, Nevola E, et al. (1988) Increases in walking distance in patients with peripheral vascular disease treated with L-carnitine: a double-blind, cross-over study. Circulation 77: 767-773.

72. NESS GC, TAKAHASHI T, LEE Y-P (1969) Thyroid hormones on amino acid and protein metabolism. I. Concentration and composition of free amino acids in blood plasma of the rat. Endocrinology 85: 1166-1171.

73. Wu G (2009) Amino acids: metabolism, functions, and nutrition. Amino acids 37: 1-17.

74. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols 5: 725.

75. Bowie JU, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science 253: 164-170.

76. Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356: 83.

77. Lovell SC, Davis IW, Arendall III WB, De Bakker PI, Word JM, et al. (2003) Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins: Structure, Function, and Bioinformatics 50: 437-450.

78. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. Physical review A 38: 3098.

79. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical review B 37: 785.

80. Frisch M, Trucks G, Schlegel H (2009) GAUSSIAN 09, Gaussian 09, Revision D. 01. Gaussian Inc, Wallingford, CT.

81. Dolg M, Wedig U, Stoll H, Preuss H (1987) Energy- adjusted abinitio pseudopotentials for the first row transition elements. The Journal of chemical physics 86: 866-872.

82. Andrae D, Haeussermann U, Dolg M, Stoll H, Preuss H (1990) Energy-adjustedab initio pseudopotentials for the second and third row transition elements. Theoretica chimica acta 77: 123-141.

83. Dallakyan S, Olson AJ (2015) Small-molecule library screening by docking with PyRx. Chemical Biology: Springer. pp. 243-250.

84. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry 31: 455-461.

85. Lengauer T, Rarey M (1996) Computational methods for biomolecular docking. Current opinion in structural biology 6: 402-406.

86. Schrödinger L (2010) Schrödinger, LLC2010The PyMOL Molecular Graphics System, Version 1.4. The PyMOL molecular graphics system, version 1.

87. BIOVIA DS (2017) Discovery Studio Modeling Environment, Release 2017, San Diego: Dassault Systèmes.

88. Vreven T, Byun KS, Komáromi I, Dapprich S, Montgomery Jr JA, et al. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM. Journal of Chemical Theory and Computation 2: 815-826.

89. Vladimirovich KG, Nizamovich NR, Borisovna TN, Mikhailovich VN, Raufovich GR, et al. (2016) Improvement technology manufacturing poly radioprotective microbial poliantigen. Research Journal Of Pharmaceutical Biological And Chemical Sciences 7: 2193-2199.

90. Vreven T, Frisch M, Kudin K, Schlegel H, Morokuma K (2006) Geometry optimization with QM/MM methods II: Explicit quadratic coupling. Molecular Physics 104: 701-714.

91. Vreven T, Morokuma K (2000) On the application of the IMOMO (integrated molecular orbital+ molecular orbital) method. Journal of Computational Chemistry 21: 1419-1432.

92. Dapprich S, Komáromi I, Byun KS, Morokuma K, Frisch MJ (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives1. Journal of Molecular Structure: THEOCHEM 461: 1-21.

93. Maseras F, Morokuma K (1995) IMOMM: A new integrated ab initio+ molecular mechanics geometry optimization scheme of equilibrium structures and transition states. Journal of Computational Chemistry 16: 1170-1179.

94. Frisch M, Trucks G, Schlegel HB, Scuseria GE, Robb MA, et al. (2009) Gaussian 09, revision a. 02, gaussian. Inc, Wallingford, CT 200.

95. Stewart JJ (2007) Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. Journal of Molecular modeling 13: 1173-1213.

96. Krieger E, Darden T, Nabuurs SB, Finkelstein A, Vriend G (2004) Making optimal use of empirical energy functions: force- field parameterization in crystal space. Proteins: Structure, Function, and Bioinformatics 57: 678-683.

97. Krieger F, Fierz B, Bieri O, Drewello M, Kiefhaber T (2003) Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. Journal of molecular biology 332: 265-274.

98. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. The Journal of chemical physics 98: 10089-10092.

99. Avbelj M, Tahirovic H, Debeljak M, Kusekova M, Toromanovic A, et al. (2007) High prevalence of thyroid peroxidase gene mutations in patients with thyroid dyshormonogenesis. European Journal of Endocrinology 156: 511-519.

100. Balmiki N, Bankura B, Guria S, Das TK, Pattanayak AK, et al. (2014) Genetic analysis of thyroid peroxidase (TPO) gene in patients whose hypothyroidism was found in adulthood in West Bengal, India. Endocrine journal 61: 289-296.

101. Lee CC, Harun F, Jalaludin MY, Heh CH, Othman R, et al. (2013) A novel, homozygous c. 1502T> G (p. Val501Gly) mutation in the thyroid peroxidase gene in Malaysian sisters with congenital hypothyroidism and multinodular goiter. International journal of endocrinology 2013.

102. Kimura S, Ikeda- Saito M (1988) Human myeloperoxidase and thyroid peroxidase, two enzymes with separate and distinct physiological functions, are evolutionarily related members of the same gene family. Proteins: Structure, Function, and Bioinformatics 3: 113-120.

103. Yokoyama N, Taurog A (1988) Porcine thyroid peroxidase: relationship between the native enzyme and an active, highly purified tryptic fragment. Molecular endocrinology 2: 838-844.

104. Godlewska M, Góra M, Buckle AM, Porebski BT, Kemp EH, et al. (2014) A redundant role of human thyroid peroxidase propeptide for cellular, enzymatic, and immunological activity. Thyroid 24: 371-382.

105. Le Fourn V, Ferrand M, Franc J-L (2005) Endoproteolytic Cleavage of Human Thyroperoxidase Role Of The Propeptide In The Protein Folding Process. Journal of Biological Chemistry 280: 4568-4577.

106. Libert F, Ruel J, Ludgate M, Swillens S, Alexander N, et al. (1987) Thyroperoxidase, an auto- antigen with a mosaic structure made of nuclear and mitochondrial gene modules. The EMBO journal 6: 4193-4196.

107. Latif R, Ali MR, Ma R, David M, Morshed SA, et al. (2015) New small molecule agonists to the thyrotropin receptor. Thyroid 25: 51-62.

108. Yamazaki N, Shinohara Y, Shima A, Yamanaka Y, Terada H (1996) Isolation and characterization of cDNA and genomic clones encoding human muscle type carnitine palmitoyltransferase I. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression 1307: 157-161.

109. Britton CH, Schultz RA, Zhang B, Esser V, Foster DW, et al. (1995) Human liver mitochondrial carnitine palmitoyltransferase I: characterization of its cDNA and chromosomal localization and partial analysis of the gene. Proceedings of the National Academy of Sciences 92: 1984-1988.

110. Müller M, Seitz H (1984) Thyroid hormone action on intermediary metabolism. Part III. Protein metabolism in hyper-and hypothyroidism. Klinische Wochenschrift 62: 97-102.

111. Wannemacher R, Klainer A, Dinterman R, Beisel W (1976) The significance and mechanism of an increased serum phenylalanine-tyrosine ratio during infection. The American journal of clinical nutrition 29: 997-1006.
112. ARROYAVE G (1970) Comparative sensitivity of specific amino acid ratios versus" essential to nonessential" amino acid ratio. The American journal of clinical nutrition 23: 703-706.
113. Oberholzer V, Briddon A (1990) A novel use of amino acid ratios as an indicator of nutritional status. Amino Acids: Springer. pp. 1079-1083.
114. Bikker H, Vulsma T, Baas F, de Vijlder JJ (1995) Identification of five novel inactivating mutations in the human thyroid peroxidase gene by denaturing gradient gel electrophoresis. Human Mutation 6: 9-16.
115. Huse M, Kuriyan J (2002) The conformational plasticity of protein kinases. Cell 109: 275-282.
116. Krupa A, Preethi G, Srinivasan N (2004) Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases. Journal of molecular biology 339: 1025-1039.
117. Gomaa AEAM, Mohamed NM, Gaballah BA, El Shabrawy SM (2018) Assessment Of The Role Of Thyroid Peroxidase Gene Polymorphism As A Cause Of Hypothyroidism In Zagazig University Hospital Patients. Zagazig University Medical Journal 23.
118. Altshuler D, Daly MJ, Lander ES (2008) Genetic mapping in human disease. science 322: 881-888.
119. Bostein D, Risch N (2003) Discovering genotypes underlying human phenotypes: past success for Mendelian disease, future approaches for complex disease. Nat Genet 33: 228-237.
120. Hendry E, Taylor G, Ziemnicka K, Jones FG, Furmaniak J, et al. (1999) Recombinant human thyroid peroxidase expressed in insect cells is soluble at high concentrations and forms diffracting crystals. Journal of endocrinology 160: R13-R15.
121. Eisenberg D, Lüthy R, Bowie JU (1997) [20] VERIFY3D: assessment of protein models with three-dimensional profiles. Methods in enzymology: Elsevier. pp. 396-404.

122. Prabhakar R, Siegbahn PE, Minaev BF, Ågren H (2004) Spin transition during H2O2 formation in the oxidative half-reaction of copper amine oxidases. The Journal of Physical Chemistry B 108: 13882-13892.

123. Prabhakar R, Siegbahn PE, Minaev BF, Ågren H (2002) Activation of triplet dioxygen by glucose oxidase: spin− orbit coupling in the superoxide Ion. The Journal of Physical Chemistry B 106: 3742-3750.

124. Minaev B, Minaeva V (2008) Spin-dependent binding of dioxygen to heme and charge-transfer mechanism of spin-orbit coupling enhancement.

125. Cherkasova O, Nazarov M, Sapozhnikov D, Man'kova A, Fedulova E, et al. Vibrational spectra of corticosteroid hormones in the terahertz range; 2010. International Society for Optics and Photonics. pp. 73760P.

126. Minaeva VA, Minaev B, Baryshnikov G, Surovtsev N, Cherkasova O, et al. (2015) Temperature effects in low-frequency Raman spectra of corticosteroid hormones. Optics and Spectroscopy 118: 214-223.

127. Eberhardt K, Stiebing C, Matthaeus C, Schmitt M, Popp J (2015) Advantages and limitations of Raman spectroscopy for molecular diagnostics: an update. Expert review of molecular diagnostics 15: 773-787.

128. Téllez Soto CA, Medeiros- Neto LP, dos Santos L, Santos AB, Ferreira I, et al. (2018) Infrared and confocal Raman spectroscopy to differentiate changes in the protein secondary structure in normal and abnormal thyroid tissues. Journal of Raman Spectroscopy 49: 1165-1173.

129. Koebel MR, Schmadeke G, Posner RG, Sirimulla S (2016) AutoDock VinaXB: implementation of XBSF, new empirical halogen bond scoring function, into AutoDock Vina. Journal of cheminformatics 8: 27.

130. Islam MT, Sarker SK, Talukder S, Bhuyan GS, Rahat A, et al. (2018) High resolution melting curve analysis enables rapid and reliable detection of G6PD variants in heterozygous females. BMC genetics 19: 58.

131. Słomka M, Sobalska-Kwapis M, Wachulec M, Bartosz G, Strapagiel D (2017) High Resolution Melting (HRM) for High-Throughput Genotyping—Limitations and Caveats in Practical Case Studies. International journal of molecular sciences 18: 2316.

132. Vossen RH, Aten E, Roos A, den Dunnen JT (2009) High- Resolution Melting Analysis (HRMA)—More than just sequence variant screening. Human mutation 30: 860-866.

133. Hidalgo-Grass C, Strahilevitz J (2010) High-resolution melt curve analysis for identification of single nucleotide mutations in the quinolone resistance gene aac (6′)-Ib-cr. Antimicrobial agents and chemotherapy 54: 3509-3511.

134. Liew M, Pryor R, Palais R, Meadows C, Erali M, et al. (2004) Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clinical chemistry 50: 1156-1164.

135. Cangul H, Schoenmakers NA, Saglam H, Doganlar D, Saglam Y, et al. (2014) A deletion including exon 2 of the TSHR gene is associated with thyroid dysgenesis and severe congenital hypothyroidism. Journal of Pediatric Endocrinology and Metabolism 27: 731-735.

136. Refetoff S (2003) Resistance to thyrotropin. Journal of endocrinological investigation 26: 770-779.

137. Sunthornthepvarakul T, Gottschalk ME, Hayashi Y, Refetoff S (1995) Resistance to thyrotropin caused by mutations in the thyrotropin-receptor gene. New England Journal of Medicine 332: 155-160.

138. Beck-Peccoz P, Persani L, Calebiro D, Bonomi M, Mannavola D, et al. (2006) Syndromes of hormone resistance in the hypothalamic–pituitary–thyroid axis. Best Practice & Research Clinical Endocrinology & Metabolism 20: 529-546.

139. Cetani F, Tonacchera M, Pinchera A, Barsacchi R, Basolo F, et al. (1999) Genetic analysis of the TSH receptor gene in differentiated human thyroid carcinomas. Journal of endocrinological investigation 22: 273-278.

140. Aronson R, Ehrlich RM, Bailey JD, Rovef JF (1990) Growth in children with congenital hypothyroidism detected by neonatal screening. The Journal of pediatrics 116: 33-37.

141. Rovet J, Ehrlich R, Sorbara D (1987) Intellectual outcome in children with fetal hypothyroidism. The Journal of pediatrics 110: 700-704.

142. LaFranchi SH, Murphey WH, Foley TP, Larsen PR, Buist NR (1979) Neonatal hypothyroidism detected by the northwest regional screening program. Pediatrics 63: 180-191.

143. Grant D, Smith I, Fuggle P, Tokar S, Chapple J (1992) Congenital hypothyroidism detected by neonatal screening: relationship between biochemical severity and early clinical features. Archives of disease in childhood 67: 87-90.

144. Alm J, Hagenfeldt L, Larsson A, Lundberg K (1984) Incidence of congenital hypothyroidism: retrospective study of neonatal laboratory screening versus clinical symptoms as indicators leading to diagnosis. Br Med J (Clin Res Ed) 289: 1171-1175.

145. Wong S, Hannah-Shmouni F, Sinclair G, Sirrs S, Dahl M, et al. (2013) Acylcarnitine profile in thyroid disease. Clinical biochemistry 46: 180-183.

146. Galland S, Georges B, Le Borgne F, Conductier G, Dias JV, et al. (2002) Thyroid hormone controls carnitine status through modifications of γ-butyrobetaine hydroxylase activity and gene expression. Cellular and Molecular Life Sciences CMLS 59: 540-545.

147. Pande SV, Parvin R (1980) Clofibrate enhancement of mitochondrial carnitine transport system of rat liver and augmentation of liver carnitine and γ-butyrobetaine hydroxylase activity by thyroxine. Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism 617: 363-370.

148. Bressler R, Wittels B (1966) The effect of thyroxine on lipid and carbohdrate metabolism in the heart. The Journal of clinical investigation 45: 1326-1333.

149. Maebashi M, Kawamura N, Sato M, Imamura A, Yoshinaga K, et al. (1977) Urinary excretion of carnitine in patients with hyperthyroidism and hypothyroidism: augmentation by thyroid hormone. Metabolism 26: 351-356.

150. Jones MG, Goodwin CS, Amjad S, Chalmers RA (2005) Plasma and urinary carnitine and acylcarnitines in chronic fatigue syndrome. Clinica chimica acta 360: 173-177.

151. An JH, Kim YJ, Kim KJ, Kim SH, Kim NH, et al. (2016) L-carnitine supplementation for the management of fatigue in patients with hypothyroidism on levothyroxine treatment: a randomized, double-blind, placebo-controlled trial. Endocrine journal: EJ16-0109.

152. GILGORE SG, DeFELICE SL (1966) Evaluation of carnitine—an antagonist of thyroid hormone. Clinical pharmacology report. The Journal of Clinical Pharmacology 6: 349-350.

153. DeFELICE SL, GILGORE SG (1966) The antagonistic effect of carnitine in hyperthyroidism. Preliminary report. The Journal of Clinical Pharmacology 6: 351-353.

154. Benvenga S, Lakshmanan M, Trimarchi F (2000) Carnitine is a naturally occurring inhibitor of thyroid hormone nuclear uptake. Thyroid 10: 1043-1050.

155. Mynatt RL, Park EA, Thorngate FE, Das HK, Cook GA (1994) Changes in carnitine palmitoyltransferase-I mRNA abundance produced by hyperthyroidism and hypothyroidism parallel changes in activity. Biochemical and biophysical research communications 201: 932-937.

156. O'BRIEN T, DINNEEN SF, O'BRIEN PC, PALUMBO PJ. Hyperlipidemia in patients with primary and secondary hypothyroidism; 1993. Elsevier. pp. 860-866.

157. de Jesus Garduno-Garcia J, Alvirde-Garcia U, Lopez-Carrasco G, Mendoza MEP, Mehta R, et al. (2010) TSH and free thyroxine concentrations are associated with differing metabolic markers in euthyroid subjects. European Journal of Endocrinology 163: 273-278.

158. Åsvold BO, Vatten LJ, Nilsen TI, Bjøro T (2007) The association between TSH within the reference range and serum lipid concentrations in a population-based study. The HUNT Study. European journal of endocrinology 156: 181-186.

159. Pearce EN (2012) Update in lipid alterations in subclinical hypothyroidism. The Journal of Clinical Endocrinology & Metabolism 97: 326-333.

160. Chung GE, Kim D, Kim W, Yim JY, Park MJ, et al. (2012) Non-alcoholic fatty liver disease across the spectrum of hypothyroidism. Journal of hepatology 57: 150-156.

161. Ding WJ, Wang MM, Wang GS, Shen F, Qin JJ, et al. (2015) Thyroid function is associated with non- alcoholic fatty liver disease in chronic hepatitis B- infected subjects. Journal of gastroenterology and hepatology 30: 1753-1758.

162. Eshraghian A, Jahromi AH (2014) Non-alcoholic fatty liver disease and thyroid dysfunction: a systematic review. World journal of gastroenterology: WJG 20: 8102.

163. Jaruvongvanich V, Sanguankeo A, Upala S (2017) Nonalcoholic Fatty Liver Disease Is Not Associated with Thyroid Hormone Levels and Hypothyroidism: A Systematic Review and Meta-Analysis. European thyroid journal 6: 208-215.

164. Chace DH, Pons R, Chiriboga CA, McMahon DJ, Tein I, et al. (2003) Neonatal blood carnitine concentrations: normative data by electrospray tandem mass spectometry. Pediatric research 53: 823.

165. Shenai JP, Borum PR, Mohan P, Donlevy SC (1983) Carnitine status at birth of newborn infants of varying gestation. Pediatric research 17: 579.

166. Yamazaki N, Yamanaka Y, Hashimoto Y, Shinohara Y, Shima A, et al. (1997) Structural features of the gene encoding human muscle type carnitine palmitoyltransferase I 1. FEBS letters 409: 401-406.

167. BROWN NF, ESSER V, KIRKLAND JL, CORKEY BE, FOSTER DW, et al. (1997) Mouse white adipocytes and 3T3-L1 cells display an anomalous pattern of carnitine palmitoyltransferase (CPT) I isoform expression during differentiation: inter-tissue and inter-species expression of CPT I and CPT II enzymes. Biochemical Journal 327: 225-231.

168. Lee J, Ellis JM, Wolfgang MJ (2015) Adipose fatty acid oxidation is required for thermogenesis and potentiates oxidative stress-induced inflammation. Cell reports 10: 266-279.

169. Shi J, Zhu H, Arvidson DN, Woldegiorgis G (2000) The first 28 N-terminal amino acid residues of human heart muscle carnitine palmitoyltransferase I are essential for malonyl CoA sensitivity and high-affinity binding. Biochemistry 39: 712-717.

170. Griffin EE, Miller LL (1973) Effects of Hypothyroidism, Hyperthyroidism, and Thyroxine on Net Synthesis of Plasma Proteins by the Isolated Perfused Rat Liver Modulation Of The Response To Insulin Plus Cortisol In The Net Synthesis Of Albumin, Fibrinogen, A1-Acid Glycoprotein, A2-(Acute Phase) Globulin, And Haptoglobin. Journal of Biological Chemistry 248: 4716-4723.

171. Wannemacher Jr R, Klainer A, Dinterman R, Beisel W (1976) The significance and mechanism of an increased serum phenylalanine-tyrosine ratio during infection. The American journal of clinical nutrition 29: 997-1006.

172. ARROYAVE G (1970) Comparative sensitivity of specific amino acid ratios versus "essential to nonessential" amino acid ratio. Oxford University Press.

Supplementary Information

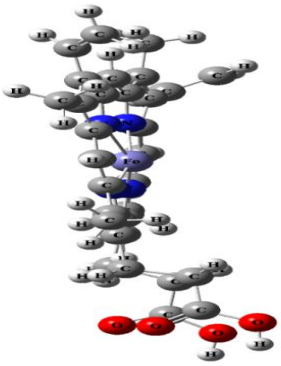
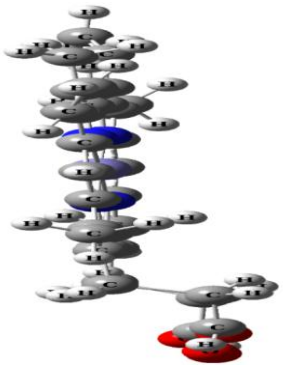
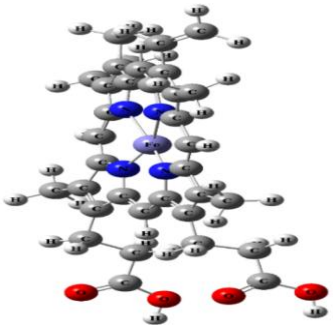
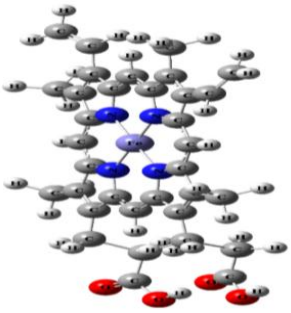
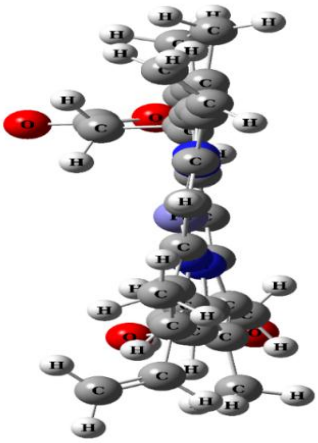
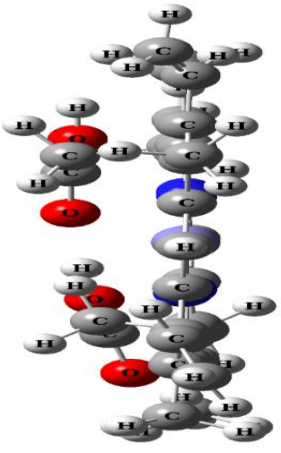
A) Before optimization	B) After optimization
 View 1	 View 1
 View 2	 View 2
 View 3	 View 3

Figure S 1. Structural differences of Heme before and after optimization from three different points of view.

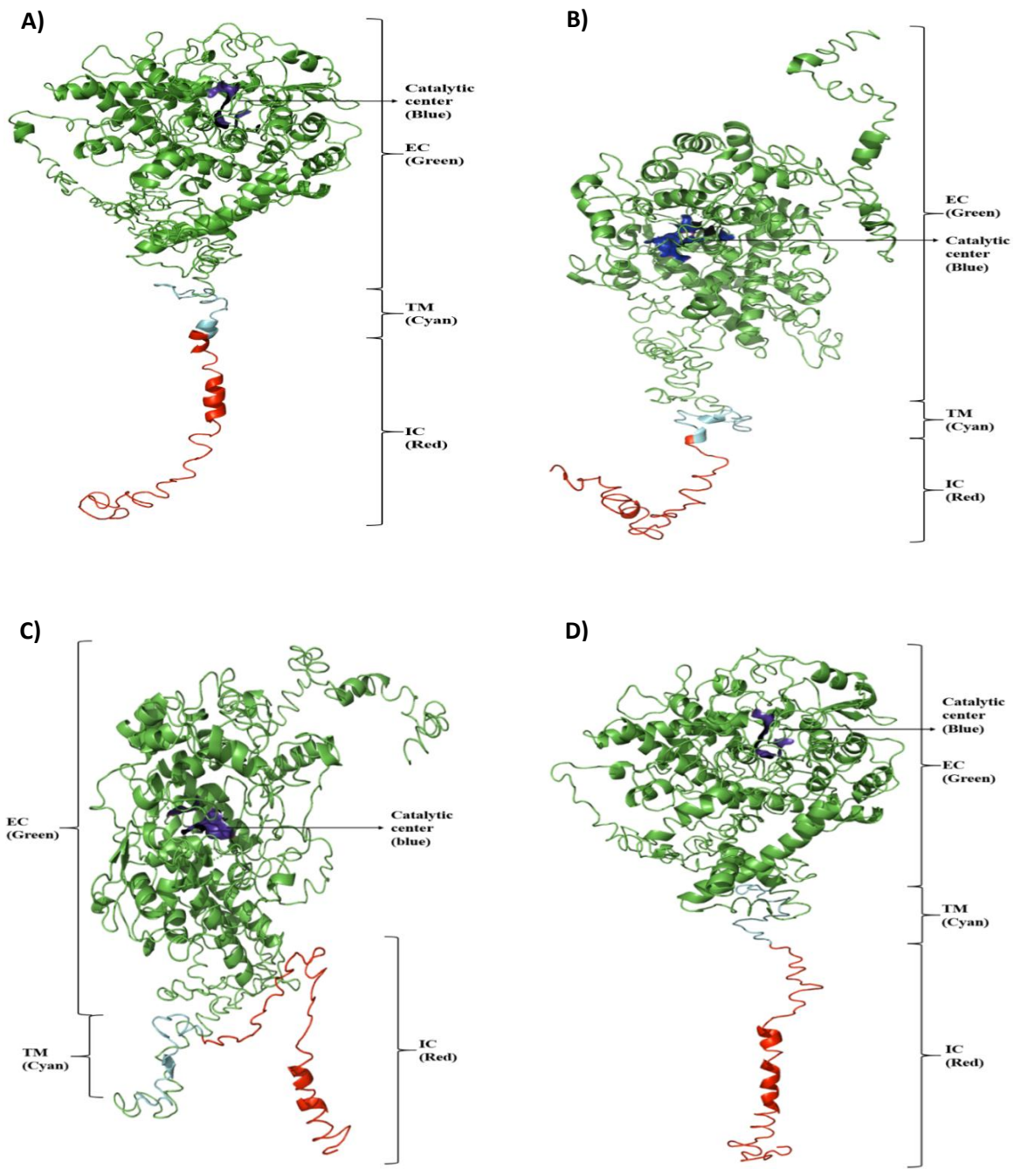


Figure S 2. The predicted 3D structures of the proteins. (A) TPO1-933 WT, (B) TPO1-933 MT1, (C) TPO1-933 MT2, (D) TPO1-933 MT3. The specific regions of the structures are indicated with the corresponding color written in the brackets. EC=Extracellular region shown in green color, TM=Transmembrane region shown in cyan color, IC=Intracellular region shown in red color, Catalytic center shown in blue color.

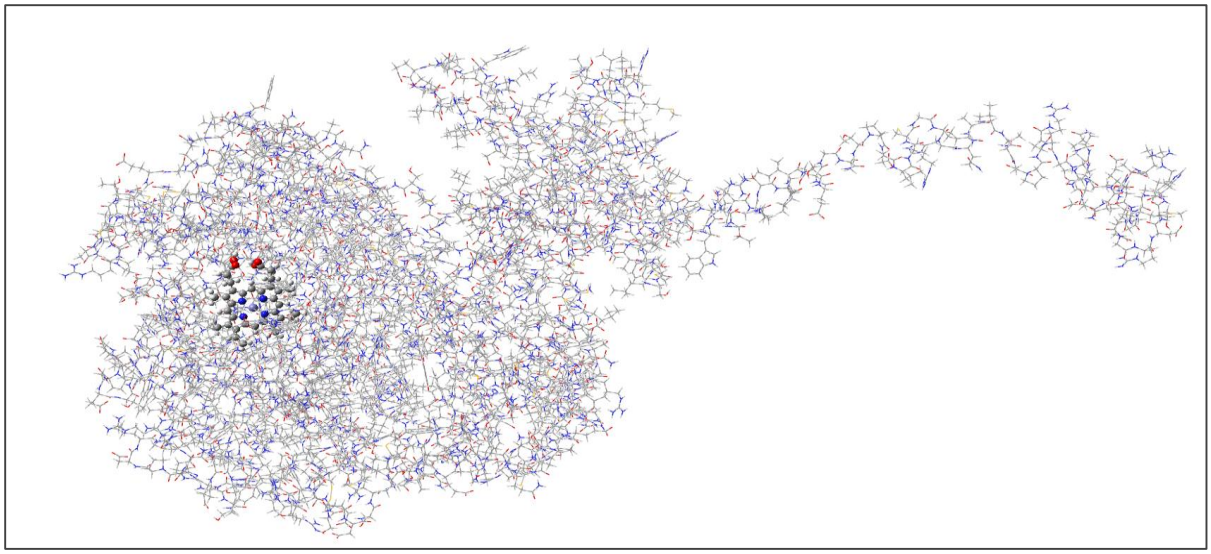


Figure S 3. QM/MM ONIOM calcualtion set up for protein-ligand complexes, high level (ball and stick) and low level (line).

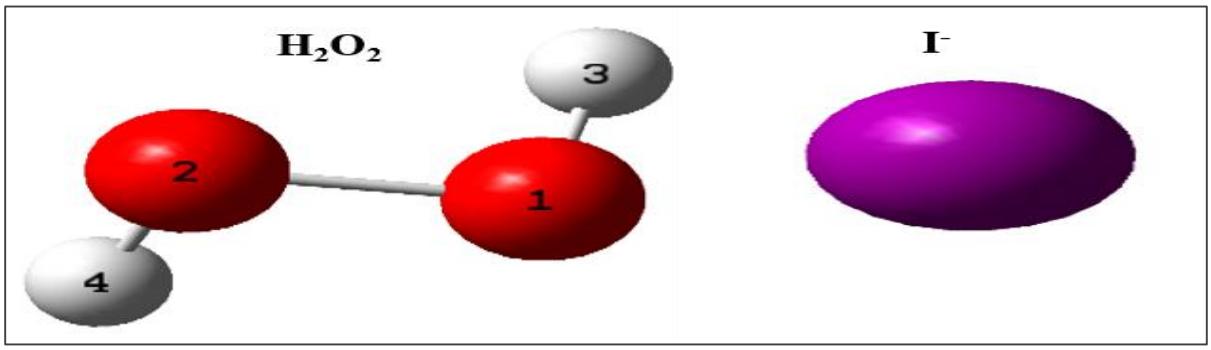


Figure S 4. Structure of H2O2 and iodide (I-) after optimization

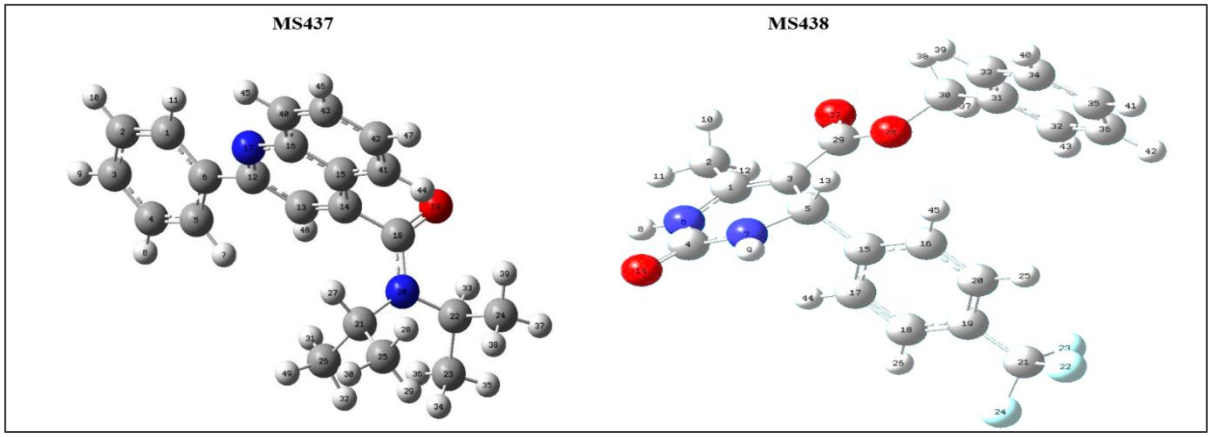


Figure S 5. Bond angles of specific atoms of MS437 and MS438 after optimization.

Table S 1. Binding energies and non-bond interactions of Heme with TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, TPO142-738 MT3 after flexible docking.

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
$TPO_{142-738}$ WT	–11.9	Hydrogen bond	Arg441(Arg582)	Conventional Hydrogen bond	2.668	21
					2.897	
					2.489	
			Arg445(Arg586)	Carbon Hydrogen bond	2.983	
		Hydrophobic interactions	His98(His239)	Pi-Pi T-Shaped	5.555	
			Val259(Val400)	Alkyl	4.305	
			His353(His 494)	Pi-Pi T-Shaped	4.356	
				Pi-Alkyl	5.235	
					4.876	
			Ile356(Ile497)	Alkyl	4.854	
			Phe382(Phe523)	Pi-Alkyl	5.023	
					4.113	
			Leu419(Leu560)	Alkyl	5.171	
			Leu423(Leu564)	Alkyl	4.517	
				Pi-Alkyl	5.288	
			Val425(Val566)	Alkyl	4.845	
			Leu426(Leu567)	Alkyl	5.458	
			Leu434(Leu575)	Pi-Alkyl	5.089	
		Electrostatic interactions	Arg255(Arg396)	Pi-Cation	4.461	
			Glu258(Glu399)	Pi-Anion	4.296	

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
					3.602	
$TPO_{142-738}$ MT1	–10.8	Hydrogen bond	Met90(Met231)	Conventional Hydrogen bond	2.327	19
					2.938	
			Gly93(Gly234)	Carbon Hydrogen bond	2.338	
			Ser261(Ser402)	Conventional Hydrogen bond	2.109	
			Gly352(Gly493)	Carbon Hydrogen bond	2.792	
		Hydrophobic interactions	Gly93(Gly234)	Amide Pi Stacked	5.0	
			Gln94(Gln235)			
			Val259(Val400)	Alkyl	3.566	
				Pi-Alkyl	5.471	
			Phe349(Phe490)	Pi-Alkyl	4.054	
			Arg350(Arg491)	Alkyl	4.666	
				Pi-Alkyl	5.031	
			Ile356(Ile497)	Pi-Alkyl	5.193	
			Phe382(Phe523)	Pi-Alkyl	3.95	
					4.787	
			Leu419(Leu560)	Alkyl	5.252	
			Leu423(Leu564)	Alkyl	4.464	
			Leu434(Leu575)	Pi-Alkyl	4.935	
			Glu258(Glu399)	Pi-Anion	3.914	

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
		Electrostatic interactions			4.651	
$TPO_{142-738}$ MT2	–2.5	Hydrogen bond	Arg350(Arg491)	Conventional Hydrogen bond	2.598	20
					2.96	
			Asn438(Asn579)		2.038	
			Arg441(Arg582)	Carbon Hydrogen bond	2.937	
		Hydrophobic interactions	Phe102(Phe243)	Pi-Alkyl	5.333	
					4.615	
			His353(His494)	Pi-Alkyl	4.462	
			Phe382(Phe523)	Pi-Alkyl	3.678	
			Phe383(Phe524)	Pi-Alkyl	4.897	
					4.558	
			Leu423(Leu564)	Alkyl	5.269	
				Pi-Alkyl	5.159	
			Phe424(Phe565)	Pi-Alkyl	5.105	
			Val425(Val566)	Alkyl	4.147	
			Leu426(Leu567)	Alkyl	3.903	
					3.824	
				Pi-Alkyl	5.226	
			Leu434(Leu575)	Alkyl	5.139	
					5.238	

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
		Electrostatic interactions	Glu258(Glu399)	Pi-Anion	3.338	
$TPO_{142-738}$ MT3	–5.3	Hydrogen bond	His98(His239)	Conventional Hydrogen bond	1.967	16
			Arg255(Arg396)		2.793	
			Arg350(Arg491)		2.686	
			Arg441(Arg582)		2.36	
		Hydrophobic interactions	Phe102(Phe243)	Pi-Alkyl	4.516	
			Phe382(Phe523)	Pi-Alkyl	3.675	
			Phe383(Phe524)	Pi-Alkyl	3.622	
					4.795	
			Leu423(Leu564)	Alkyl	4.628	
			Phe424(Phe565)	Pi-Alkyl	4.705	
					5.304	
			Leu426(Leu567)	Alkyl	4.406	
					4.555	
			Leu434(Leu575)	Alkyl	4.991	
				Pi-Alkyl	5.148	
		Electrostatic interactions	Glu258(Glu399)	Pi-Anion	3.633	

The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding number; in case of $TPO_{142-738}$ the amino acid outside the first bracket indicates the position in predicted structure for $TPO_{142-738}$ and the amino acid residues in first bracket indicates the real position in TPO_{1-933} protein; WT = Wild type; MT1 = Mutant 1 (p.Ala373Ser); MT2 = Mutant 2 (p.Ser398Thr); MT3 = Mutant 3 (p.Thr725Pro).

Table S 2. Binding energies and non-bond interactions of Heme with TPO1-933WT, TPO1-933MT1, TPO1-933 MT2 and TPO1-933 MT3 after flexible docking.

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
TPO_{1-933} WT	–11.5	Hydrogen bond	Arg491	Conventional Hydrogen bond	2.579	21
					2.765	
			Arg582		2.445	
		Hydrophobic interactions	His239	Pi-Pi T-Shaped	5.589	
			Val400	Alkyl	4.169	
					4.851	
			Phe490	Pi-Alkyl	4.875	
			Arg491	Alkyl	4.094	
			His494	Pi-Pi T-Shaped	4.139	
				Pi-Alkyl	5.076	
			Ile497	Alkyl	5.259	
			Phe523	Pi-Alkyl	4.176	
			Leu560	Pi-Alkyl	5.115	
				Alkyl	4.254	
			Leu564	Alkyl	4.066	
					4.406	
			Leu575	Pi-Alkyl	5.391	
		Electrostatic interactions	Arg396	Pi-Cation	4.894	
			Glu399	Pi-Anion	3.536	
					3.585	
			Arg491	Pi-Cation	4.526	

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
TPO_{1-933} MT1	–3.2	Hydrogen Bond	His239	Conventional Hydrogen bond	2.209	12
			His494	Carbon Hydrogen bond	2.809	
		Hydrophobic interactions	His239	Pi-Alkyl	4.919	
			Phe243	Pi-Pi T-Shaped	4.875	
			Arg396	Alkyl	3.979	
				Pi-Alkyl	3.680	
					4.706	
			Phe524	Pi-Alkyl	3.864	
					4.435	
			Leu567	Alkyl	5.123	
				Pi-Alkyl	4.565	
					5.397	
TPO_{1-933} MT2	–11.5	Hydrogen Bond	His239	Conventional Hydrogen bond	3.155	20
			Arg491		2.743	
			Arg582		2.999	
					1.899	
		Hydrophobic interactions	His239	Pi-Pi T-Shaped	5.261	
			Val400	Alkyl	3.558	
			Arg491	Alkyl	4.362	
			His494	Pi-Alkyl	4.944	
					5.043	

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
TPO_{1-933} MT3	–7.9		Ile497	Alkyl	5.116	21
			Phe523	Pi-Alkyl	4.003	
					5.001	
			Leu564	Alkyl	3.729	
					5.453	
				Pi-Alkyl	5.391	
			Val566	Alkyl	4.717	
			Leu575	Pi-Alkyl	4.676	
		Electrostatic interactions	Arg396	Pi-Cation	4.176	
			Glu399	Pi-Anion	3.596	
					3.979	
		Hydrogen Bond	Gln246	Conventional hydrogen bond	2.295	
					2.351	
			Arg491		3.085	
			Ser568	Carbon hydrogen bond	1.969	
					3.043	
		Hydrophobic interactions	Phe243	Pi-Pi T-Shaped	5.806	
				Pi-Alkyl	5.488	
			Val400	Alkyl	5.463	
			Arg491	Alkyl	4.445	
			His494	Pi-Alkyl	4.209	
					4.567	

Protein	Binding energy (kcal / mol)	Category	Interacting amino acids	Interaction types	Distance (Å)	Total no. of interactions
			Ile497	Alkyl	4.875	
			Phe523	Pi-Alkyl	4.544	
			Phe524	Pi-Alkyl	5.488	
			Leu560	Alkyl	4.702	
			Leu564	Alkyl	4.943	
					4.317	
			Val566	Alkyl	3.745	
			Leu575	Pi-Alkyl	4.767	
		Electrostatic interactions	Glu399	Pi-Anion	4.081	
					3.653	

The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding number.WT = Wild type; MT1 = Mutant 1 (p.Ala373Ser); MT2 = Mutant 2 (p.Ser398Thr); MT3 = Mutant 3 (p.Thr725Pro).

Table S 3. Summarized data for the bond distances and bond angles of specific atoms of H2O2 after optimization.

Atoms	Bond distance (nm)	Atoms	Bond angle (degrees)
O1…H3	0.97	O2…O1…H3	100.33°
O1…O2	1.48	O2…O1…H4	100.33°
O2…H4	0.97		

Table S 4. Sequenced samples that were used for HRM method setup

Sl no	Specimen	Mutation in Exon 08	Mutation in Exon 12
1	Specimen-01	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	No Mutation
2	Specimen-02	1193 G→C (398 Ser→Thr) *	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
3	Specimen-03	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr)	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
4	Specimen-04	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr)*	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
5	Specimen-05	1193 G→C (398 Ser→Thr) #	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
6	Specimen-06	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr)*	No Mutation
7	Specimen-07	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr) #	No Mutation
8	Specimen-08	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr) #	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
9	Specimen-09	No mutation	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
10	Specimen-10	1193 G→C (398 Ser→Thr) #	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
11	Specimen-11	No mutation	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
12	Specimen-12	1117 G→T (373 Ala→Ser)*	2145 C→T (715 Pro→Pro)*

Sl no	Specimen	Mutation in Exon 08	Mutation in Exon 12
		1193 G→C (398 Ser→Thr)*	2173 A→C (725 Thr→Pro)*
13	Specimen-13	1193 G→C (398 Ser→Thr) *	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
14	Specimen-14	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr)*	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
15	Specimen-15	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr) #	No mutation
16	Specimen-16	1193 G→C (398 Ser→Thr) *	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
17	Specimen-17	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	No mutation
18	Specimen-18	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr) #	No mutation
19	Specimen-19	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr)	No mutation
20	Specimen-20	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	No mutation

Note: * =heterozygous state, #=homozygous state

Table S 5. Unknown samples used for HRM method validation

Sl no	Specimen	Mutation in Exon 08	Mutation in Exon 12
1	Specimen-01	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr)*	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
2	Specimen-02	No mutation	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
3	Specimen-03	1117 G→T (373 Ala→Ser)* 1193G→C (398 Ser→Thr)*	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
4	Specimen-04	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	No Mutation
5	Specimen-05	No mutation	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
6	Specimen-06	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	No mutation
7	Specimen-07	1193 G→C (398 Ser→Thr)*	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
8	Specimen-08	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr) #	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
9	Specimen-09	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	2145 C→T (715 Pro→Pro) 2173 A→C (725 Thr→Pro)
10	Specimen-10	1193 G→C (398 Ser→Thr) #	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
11	Specimen-11	No mutation	2145 C→T (715 Pro→Pro) # 2173 A→C (725 Thr→Pro) #
12	Specimen-12	No mutation	2145 C→T (715 Pro→Pro) #

Sl no	Specimen	Mutation in Exon 08	Mutation in Exon 12
			2173 A→C (725 Thr→Pro) #
13	Specimen-13	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	No mutation
14	Specimen-14	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr)*	2145 C→T (715 Pro→Pro)* 2173 A→C (725 Thr→Pro)*
15	Specimen-15	1117 G→T (373 Ala→Ser)* 1193 G→C (398 Ser→Thr) #	No mutation
16	Specimen-16	1117 G→T (373 Ala→Ser) # 1193 G→C (398 Ser→Thr) #	No mutation

Note: * =heterozygous state, #=homozygous state

Table S 6. Comparison of 8 short chain and 7 medium chain acylcarnitines between patients and healthy controls

Category	Patient µmol/L (mean±SD)	CV (%)	IQR	Healthy Control µmol/L (mean±SD)	CV (%)	IQR	*P* Value
Short chain acylcarnitines							
Acetylcarnitine (C2)	16.56±6.5	39.26	8.72	15.30±4.52	29.54	4.9	0.19
Propionylcarnitine (C3)	1.52±0.59	38.91	0.98	1.38±0.43	30.91	0.51	0.11
Butyrylcarnitine (C4)	0.14±0.06	44.08	0.06	0.13±0.05	42.74	0.05	0.25
3-Hydroxybutyrylcarnitine (C4OH)	0.09±0.06	68.0	0.07	0.09±0.05	60.82	0.05	0.53
Isovalerylcarnitine (C5)	0.13±0.07	56.26	0.06	0.14±0.05	36.72	0.67	0.67
Tiglylcarnitine (C5:1)	0.01±0.004	68.24	0.007	0.01±0.005	78.21	0.006	0.61
3-OH isovalerylcarnitine (C5OH)	0.25±0.08	34.06	0.1	0.3±0.08	28.6	0.13	0.0012*
Glutarylcarnitine (C5DC)	0.08±0.04	48.73	0.05	0.07±0.02	39.97	0.04	0.22
Total short chain acylcarnitines	**18.77±6.95**	**37.07**	**10.19**	**17.40±4.84**	**27.85**	**5.13**	**0.195**
Medium chain acylcarnitines							
Hexanoylcarnitine (C6)	0.03±0.01	69.14	0.02	0.02±0.01	97.17	0.03	0.054
Octanoylcarnitine (C8)	0.02±0.01	132.3	0.04	0.02±0.02	194.5	0.02	0.22
Octenoylcarnitine (C8:1)	0.19±0.13	68.49	0.18	0.16±0.09	54.33	0.1	0.14
Decanoylcarnitine (C10)	0.08±0.05	67.85	0.06	0.07±0.04	59.97	0.04	0.24
Decenoylcarnitine (C10:1)	0.09±0.05	70.15	0.06	0.07±0.03	52.5	0.04	0.094
Decadienoylcarnitine (C10:2)	0.02±0.01	64.78	0.01	0.02±0.08	51.6	0.01	0.66
Lauroylcarnitine (C12)	0.05±0.02	50.05	0.03	0.05±0.02	41.74	0.02	0.42
Total medium chain acylcarnitines	**0.47±0.26**	**55.44**	**0.31**	**0.40±0.16**	**39.38**	**0.18**	**0.092**

Ethical Clearance

ডিন অফিস
জীববিজ্ঞান অনুষদ
ঢাকা বিশ্ববিদ্যালয়, ঢাকা-১০০০, বাংলাদেশ

Tel : 8613243
PABX : 9661900-59/4355, 7489
Fax : 880-2-865583
E-mail : deanbio@univdhaka.edu
mimdadul07@yahoo.com

নং - ---29---/ জীবঃ অনুঃ/২০১৬-২০১৭ তারিখ: ১৬/০৫/২০১৭

Professor Dr. Sharif Akhteruzzaman
Department of Genetic Engineering and Biotechnology
University of Dhaka
Dhaka-1000.

Sub: Ethical Clearance of Research Proposal entitled "Innovative approaches for development of diagnostic methods for detecting inborn errors of metabolism and genetic disorders using high-throughput metabolomic profiling as well as monoclonal antibody-based concept (CP-4029)".

Dear Dr. Akhteruzzaman,

I am happy to inform you that your proposal entitled **"Innovative approaches for development of diagnostic methods for detecting inborn errors of metabolism and genetic disorders using high-throughput metabolomic profiling as well as monoclonal antibody-based concept (CP-4029)"** was placed in the Ethical Clearance Certificate for Human Participants Committee meeting held on 16.05.2017 and has been approved for conducting your research project.

I wish for the success of your research project.

Professor Dr. M. Imdadul Hoque
Dean, Faculty of Biological Sciences
University of Dhaka
Dhaka-1000.