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

Mst. Noorjahan Begum,1,2 Md Tarikul Islam,1 Shekh Rezwan Hossain,1,3 Golam Sarower Bhuyan,4 Mohammad A. Halim,5 Imrul Shahriar,5 Suprovath Kumar Sarker,1,2 Shahinur Haque,6 Tasnia Kawsar Konika,6 Md. Sazzadul Islam,1,7 Asifuzzaman Rahat,4 Syeda Kashfi Qadri,8 Rosy Sultana,1,9 Suraiya Begum,10 Sadia Sultana,6 Narayan Saha,11 Mizanul Hasan,6 M. A. Hasanat,12 Hurjahan Banu,12 Hossain Uddin Shekhar12,3 Emran Kabir Chowdhury,3 Abu A. Sajib,2 Abul B. M. M. K. Islam,2 Syed Saleheen Qadri,1,4 Firdausi Qadri,1,4,13 Sharif Akhteruzzaman,2 and Kaisissar Mannoor1,4

1 Laboratory of Genetics and Genomics, Institute for Developing Science and Health Initiatives (idesHi), Mohakhali, Dhaka 1212, Bangladesh
2 Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh
3 Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka 1000, Bangladesh
4 Infectious Diseases Laboratory, Institute for Developing Science and Health Initiatives (idesHi), Mohakhali, Dhaka 1212, Bangladesh
5 The Red-Green Research Centre (RGRC), 218 Elephant Road, Dhaka 1205, Bangladesh
6 Nuclear Medicine and Allied Sciences, Bangabandhu Sheikh Mujib Medical University (BSMMU), Shahbag, Dhaka 1000, Bangladesh
7 Department of Biochemistry and Molecular Biology, Jagannath University, Dhaka, Bangladesh
8 Department of Paediatric Medicine, KK Women’s and Children’s Hospital, 100 Bukit Timah Road, Singapore
9 Bangladesh University of Health Sciences, Bangladesh
10 Department of Paediatrics Endocrinology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Shahbag, Dhaka 1000, Bangladesh
11 Pediatric Neurology, National Institute of Neurosciences & Hospital, Dhaka 1207, Bangladesh
12 Department of Endocrinology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Shahbag, Dhaka 1000, Bangladesh
13 Department of Enteric and Respiratory Infectious Diseases, Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Mohakhali, Dhaka, Bangladesh

Correspondence should be addressed to Kaisissar Mannoor; kaisissar@ideshi.org

Received 7 July 2018; Revised 3 January 2019; Accepted 10 January 2019; Published 24 February 2019

Academic Editor: Gerald J. Wyckoff

Copyright © 2019 Mst. Noorjahan Begum et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Although thyroid dyshormonogenesis (TDH) accounts for 10-20% of congenital hypothyroidism (CH), the molecular etiology of TDH is unknown in Bangladesh. Thyroid peroxidase (TPO) is most frequently associated with TDH and the present study investigated the spectrum of TPO mutations in Bangladeshi patients and analyzed the effects of mutations on TPO protein structure through in silico approach. Sequencing-based analysis of TPO gene revealed four mutations in 36 diagnosed patients with TDH including three nonsynonymous mutations, namely, p.Ala373Ser, p.Ser398Thr, and p.Thr725Pro, and one synonymous mutation p.Pro715Pro. Homology modelling-based analysis of predicted structures of MPO-like domain (TPO142-738) and the full-length TPO
protein (TPO\textsubscript{3-93}) revealed differences between mutant and wild type structures. Molecular docking studies were performed between predicted structures and heme. TPO\textsubscript{3-93} predicted structure showed more reliable results in terms of interactions with the heme prosthetic group as the binding energies were -11.5 kcal/mol, -3.2 kcal/mol, -11.5 kcal/mol, and -7.9 kcal/mol for WT, p.Ala373Ser, p.Ser398Thr, and p.Thr725Pro, respectively, implying that p.Ala373Ser and p.Thr725Pro mutations were more damaging than p.Ser398Thr. However, for the TPO\textsubscript{42-738} predicted structures, the binding energies were -11.9 kcal/mol, -10.8 kcal/mol, -2.5 kcal/mol, and -5.3 kcal/mol for the wild type protein, mutant proteins with p.Ala373Ser, p.Ser398Thr, and p.Thr725Pro substitutions, respectively. However, when the interactions between the crucial residues including residues His239, Arg396, Gln399, and His494 of TPO protein and heme were taken into consideration using both TPO\textsubscript{3-93} and TPO\textsubscript{42-738} predicted structures, it appeared that p.Ala373Ser and p.Thr725Pro could affect the interactions more severely than the p.Ser398Thr mutation. Validation of the molecular docking results was performed by computer simulation in terms of quantum mechanics/molecular mechanics (QM/MM) and molecular dynamics (MD) simulation. In conclusion, the substitutions mutations, namely, p.Ala373Ser, p.Ser398Thr, and p.Thr725Pro, had been involved in Bangladeshi patients with TDH and molecular docking-based study revealed that these mutations had damaging effect on the TPO protein activity.

1. Introduction

Congenital hypothyroidism (CH) is defined as insufficient production of thyroid hormone at birth [1]. The global frequency of CH is 1 in 3000-4000 whereas a pilot study in Bangladesh reported an incidence rate of 1 in 1300 [2, 3]. Thyroid dyshormonogenesis (TDH) which results due to defect in the pathway of thyroid hormone biosynthesis accounts for 10 to 20% cases of CH [4]. Till now, mutations in seven genes, namely, NIS (sodium iodide symporter, SLC5A5), PDS (Pendrin or SLC26A4), TPO (thyroid peroxidase), TG (thryoglobulin), IYD (iodotyrosine deiodinase, DEHALI), DUOX2, and DUOX2 (dual oxidase), have been reported to be involved in pathogenesis of TDH [5]. However, mutations in TPO gene which may result in total iodide organification defect (TIOD) or partial iodide organification defect (PIOD) have been described as the most common forms of TDH, and to date more than 60 different mutations have been described [6, 7]. TPO is the major enzyme in thyroid hormone biosynthesis and it catalyzes both iodination and coupling of iodothyrosine residues in TG (thyroglobulin). It is a glycosylated hemeprotein of 110 kDa bound at the apical membrane of thyrocytes [8]. The single gene encoding TPO is located on chromosome 2p25 containing 17 exons and it spans at least 150 kb [9]. It is a homodimer protein (each monomer consists of 933 amino acid residues) and contains a peroxidase domain, three additional extracellular domains, a transmembrane helix, and a short C-terminal intracellular tail [10]. Although a low-resolution crystal structure of TPO has been reported, its high-resolution structure remains to be determined [11, 12]. The closest known homologues of TPO are lactoperoxidase (LPO), myeloperoxidase (MPO), and eosinophil peroxidase (EPO) with a sequence identity of 48%, 47%, and 47%, respectively. X-ray crystallographic structures have been determined for MPO and LPO [13, 14], providing a platform for investigating the structural basis of TPO. Nevertheless, Sarah et al. investigated plausible modes of TPO structure and dimer organization through in silico approach [15]. However, how certain mutations affect the TPO protein structure and its enzyme activity is still unknown.

Investigation of mutational spectrum is paramount for genetic disorder study as it gives insight into the disease pathogenicity as well as severity. Screening and identification of mutational spectrum in the TPO gene of patients with TIOD and PIOD have been reported in different countries of the world like Argentina, Netherlands, Japan, Portugal, and China [6, 16–19]. In addition, changes in the TPO enzyme were assayed in vitro to compare mutant and wild type activities by several study groups and demonstrated mild to severe TPO enzyme inactivity for some mutations [20, 21]. Though the prevalence of CH is more than twice the global incidence rate in Bangladesh, molecular basis of CH is still unknown in this country. Moreover, the effects of different mutations in the TPO gene on enzyme activity have not been investigated. In this study, we investigated the spectrum of mutations in TPO gene of patients with TDH and explored the possible effect of these mutations on the structure of TPO protein and TPO’s MPO-like domain through in silico approach such as homology modelling, molecular docking followed by quantum mechanics/molecular mechanics (QM/MM) and molecular dynamics (MD) simulation.

2. Methods and Materials

2.1. Study Participants. A total of 36 confirmed cases of congenital hypothyroid patients with dyshormonogenesis were enrolled at the clinical care settings of National Institute of Nuclear Medicine and Allied Sciences (NINMAS) and department of Endocrinology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Dhaka, for their follow-up examinations. Written informed consent was signed by the parents or guardians of the patients. 3 mL of blood was collected in an EDTA tube from each patient. Ethical clearance was obtained from the ethical committee of BSMMU and University of Dhaka.

2.2. DNA Isolation and Polymerase Chain Reaction (PCR) Amplification. Genomic DNA was isolated from the EDTA blood by using a Qiagen DNAeasy mini kit. The isolated DNA was then amplified by PCR using TPO gene specific primers that together covered from 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. The primer sequences are listed in Table 1.

| Primer Name | Forward (5'-3') | Reverse (5'-3') |
|-------------|----------------|----------------|
| TPO-F1 | CTGCAGTTGGCTTGTAGTAG | TGGTTTGGCCGTTTTTGGT |
| TPO-R1 | TGGTTTGGCCGTTTTTGGT | CTGCAGTTGGCTTGTAGTAG |
| TPO-F2 | TGGTTTGGCCGTTTTTGGT | CTGCAGTTGGCTTGTAGTAG |
| TPO-R2 | CTGCAGTTGGCTTGTAGTAG | TGGTTTGGCCGTTTTTGGT |

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 1μL 10X PCR buffer, 0.3 μL 25 mM MgCl₂, 2μL 5X Q-solution, 1.6 μL 25 mM dNTPs mixture, 0.2 μL 10 mM Forward and 0.2 μL Reverse primers, 0.05 μL Taq DNA Polymerase, and 50 ng of genomic DNA, and total reaction volume was made up to 10 μL by addition of nuclease free water. The thermal cycling condition included (a) initial denaturation at 95°C for 5 minutes, (b) 35 cycles of denaturation at 95°C for 40 seconds 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.

2.3. Sanger Sequencing of PCR Products. Prior to sequencing, PCR product was purified using a Qiagen PCR purification kit (Qiagen) according to manufacturer's instruction. The cycle sequencing PCR was then conducted using a BigDye Chain Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, USA) following manufacturer's instructions. The thermal cycling profile included (a) initial denaturation at 94°C for 1 minute, (b) 25 cycles of denaturation at 94°C for 10 seconds, 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. Following completion of cycle sequencing PCR, purification of the product was performed using a BigDye XTerminator® Purification Kit (Applied Biosystems). Finally, sequencing of the purified cycle sequencing product was performed on the ABI PRISM 310 automated sequencer (Applied Biosystems, USA).

2.4. Sequencing Data Analysis. Sequencing 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 (Accession number; NC_000002.12 retrieved from the NCBI database) using the basic local alignment search tool (BLAST). ExPASy translate tool was used to convert nucleotides sequence into corresponding amino acids.

2.5. Analysis of Effect of Three Nonsynonymous Mutations in the 3D Structure of TPO Protein

2.5.1. Prediction of 3D Structure through In Silico Approach. We found 4 mutations including three nonsynonymous mutations, namely, p.Ala373Ser; p.Ser398Thr and p.Thr725Pro and a synonymous mutation p.Pro725Pro in the TPO₁₋₉₃₃ gene. One of our primary goals was to investigate the effect of the nonsynonymous mutations on the 3D structure of MPO-like domain of the TPO protein. In addition, we aimed at seeing whether the mutations had any effects on the heme interactions with specific amino acid residues. The corresponding positions of the mutations, including p.Ala373Ser, p.Ser398Thr, and p.Thr725Pro in the MPO-like domain of TPO₁₋₇₃₈ were p.Ala232Ser, p.Ser257Thr, and p.Thr584Pro, respectively. Since the crystallographic structure of TPO protein was available at lower resolution, to understand the effect of mutations in the structure of MPO-like domain of TPO₁₋₇₃₈, amino acid sequences of wild type and mutant proteins containing different nonsynonymous mutations were submitted to I-TASSER server in order to obtain the 3D structures [22–24]. We obtained 1 model for each structure based on C-score, Template Modelling (TM) score, and Root Mean Square Deviation (RMSD) score. In addition, we investigated the effects of the mutations on full-length TPO protein structure and functions in order to compare with the results with MPO-like domain to see if any changes were found. For this purpose, we also predicted the structures of whole TPO protein (TPO₁₋₉₃₃) for the wild type (TPO WT) and the mutants (TPO MT1, TPO MT2, and TPO MT3) by using the I-TASSER server and obtained 5 models for each. 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 published article [15].

2.5.2. Validation of the 3D Structures. The predicted 3D structures of TPO₁₋₇₃₈ WT, TPO₁₋₇₃₈ MT1, TPO₁₋₇₃₈ MT2, TPO₁₋₇₃₈ MT3, TPO₁₋₉₃₃ WT, TPO₁₋₉₃₃ MT1, TPO₁₋₉₃₃ MT2, and TPO₁₋₉₃₃ MT3 were validated using 2 different web servers, namely, Verify3D server, and RAMPAGE server [25–27]. To validate the structure, the PDB format of the predicted

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

| Primer Name | Sequence (5’-3’) | Product size (base pair) |
|-------------|-----------------|-------------------------|
| TPO_Ex8F    | TCCAGAGCTCTACAACAGGTTGC | 679 |
| TPO_Ex8R    | GTACCTGAGAGAGGCCAC | 658 |
| TPO_Ex9F    | GAGGTCTCCATGCACTGTGACCT | 512 |
| TPO_Ex9R    | GGAAGGATTTGACGGGAGGCCAC | 529 |
| TPO_Ex10F   | TAGAACTGTGCCAAGGCTGTC | 363 |
| TPO_Ex10R   | CTAGCAGCCAGTTGTGACCTG | 292 |
| TPO_Ex11F   | GCCATGGGCATGAGGATGAGG | 270 |
| TPO_Ex11R   | CTGCCCTGCTGCAAGACGTG | 363 |
| TPO_Ex12F   | GGTTCATCCATGCAGTGACCT | 1044 |
| TPO_Ex13F   | GATTCACATGCAGTGACCT | 1044 |
| TPO_Ex14F   | CCTCATCCTCTTTTCGAGATGAC | 512 |
| TPO_Ex14R   | CAGACAGCCAGCAGCAGAAGTG | 512 |
Table 2: 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         |
| TPO1-933 WT    | 109.7751         | 128.9632        | 119.5649| 24.5891       | 20.8773       | 31.4612   |
| TPO1-933 MT1   | 105.160          | 102.301         | 98.7798 | 25.0          | 25.0          | 25.0      |
| TPO1-933 MT2   | 99.6151          | 102.226         | 100.5057| 25.0          | 25.0          | 25.0      |
| TPO1-933 MT3   | 102.196          | 124.345         | 113.7515| 25.0          | 25.0          | 25.0      |

3D structure was submitted to both the servers. The result included the percentage of the amino acids having the average 3D-1D score of ≥ 0.2 for Verify3D server. The RAMPAGE webserver performs Ramachandran Plot Analysis by providing results including 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.

2.5.3. Optimization of Heme Using Quantum Mechanical (QM) Calculations. Since TPO is a heme containing protein, 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 conducted using density functional theory (DFT) method employing Becke’s (B) exchange functional combining Lee, Yang, and Parr’s (LYP) correlation functional, widely known as B3LYP density functional in Gaussian 09 program package [28–30]. SDD basis set had been used for optimization of the heme [31, 32].

2.5.4. Molecular Docking of Heme with the Predicted 3D Structures of TPO. For molecular docking, Autodock Vina [33, 34] protocol was employed. The molecular docking approach involved the prediction of the interaction between a small molecule and a protein at the atomic level, which provided us the opportunity to investigate the behavior of small molecules in the binding site of target proteins and to elucidate the fundamental biochemical processes. Before docking processes, knowing the location of ligand binding sites in the target protein could significantly increase the docking efficiency [35]. 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 TPO1-933, i.e. MPO-like domain and these amino acid residues were involved with heme interactions and thus the catalytic activity of TPO [36, 37]. But in the case of MPO-like domain of TPO142-738, 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 TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, and TPO142-738 MT3. The center of the grid box for TPO142-738 WT was set to 71.337 Å, 73.748 Å, and 72.888 Å; whereas they were set to 71.281 Å, 73.813 Å, and 73.314 Å for TPO142-738 MT1; 73.115 Å, 74.888 Å, and 73.156 Å for TPO142-738 MT2; and 73.273 Å, 75.433 Å, and 74.247 Å for TPO142-738 MT3 in the x, y, and z directions, respectively. The grid box size was set to 25.0 Å, 25.0 Å, and 25.0 Å for TPO142-738 WT, MT1, MT2, and MT3 in the x, y, and z directions, respectively. Grid box value center and grid box size was also optimized for full-length TPO1-933 (Table 2).

2.5.5. Visualization and Analysis of Docking Results. The molecular docking results of heme with TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, and TPO142-738 MT3 and also for the whole protein TPO1-933 were visualized and analyzed using the PyMol (version 2.0) and BIOVIA Discovery Studio 2017 software [38, 39]. The binding affinities of heme with wild type TPO (TPO WT) compared to 3 mutant proteins (TPO MT1, TPO MT2, and TPO MT3) were observed. Moreover, the corresponding non-bond interactions of amino acids with heme were also studied for wild type and mutant 3D structures.

2.5.6. QM/MM and Molecular Dynamics (MD) Simulation. To validate the molecular docking results, further analysis by quantum chemical method QM/MM and molecular dynamics (MD) simulations were employed for full-length TPO proteins (wild type, MT1, and MT3) since published data showed MT1 and MT3 had severe effect on enzyme activity than MT2. The QM/MM calculations of the selected protein-ligand complexes were performed by a two-layer ONIOM method available in the Gaussian09 software package [40–46]. QM and MM regions have been shown in supplementary Figure 3. The heme molecule was included in the QM region and semiempirical PM6 level of theory [47] was considered due to large structure of the heme molecules. 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 is as follows: (see [48])

$$E_{ONIOM} = E_{real,MM} + E_{model,QM} - E_{model,MM}$$  \hspace{1cm} (1)

The real system consists of all the atoms and is calculated only at the MM level. The model system consists of the part
3.1. Demographic Information of the Study Participants.

A total of 36 confirmed cases of congenital hypothyroidism with dysmorphogenesis were enrolled in this study. Among 36 patients, 15 (41.67%) and 21 (58.33%) were female and male, respectively, with an average age of 7.58 ± 4.56 years.

3.2. Analysis of TPO Gene for Identification of Molecular Basis of Hypothyroidism.

As mutations in the TPO gene are commonly associated with thyroid dysmorphogenesis and related complications [18, 51], we opted to analyze the TPO gene to find whether there were mutations in this gene of the study participants. Upon Sanger sequencing of specimens from the patients with thyroid dysmorphogenesis targeting exon-8 to exon-14 which are commonly reported in TPO-associated thyroid dysmorphogenesis, mutations were detected in all 36 samples. A total of four mutations, namely, c.1117G>T (p.Ala373Ser), c.1193G>C (p.Ser398Thr), c.2145C>T (p.Pro715Pro), and c.2173A>C (p.Thr725Pro) were identified in the study participants (Table 3). The first two of these four mutations were detected in exon-8, whereas the remaining two mutations were detected in exon-12. Even though c.2145C>T (p.Pro715Pro) of the four mutations is a synonymous point mutation and is innocuous, the other three nonsynonymous mutations, namely, p.Ala373Ser, p.Ser398Thr, and p.Thr725Pro, had previously been reported in the patients with thyroid dysmorphogenesis and the reaction kinetics catalyzed by the mutant TPO enzyme proved to be similar with nonenzymatic reaction rates by several other studies [20, 52, 53].

3.3. Prediction of 3D Structures of Myeloperoxidase (MPO)-Like Domain (TPO142-738) and Full-Length TPO Protein (TPO1-933).

The crystallographic structure of TPO protein is available with low resolution and the catalytic domain of TPO142-738 is similar to human myeloperoxidase (MPO). Since the mutations that were identified in this present study were confined in the MPO-like domain of TPO protein, we investigated the effect of mutations on the 3D structure of MPO-like domain (TPO142-738) of TPO protein. Also, we wanted to see whether the mutations caused any changes in the interactions of various amino acid residues with heme prosthetic group. We submitted the amino acid TPO142-738 sequence for the wild type (WT), mutant 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 TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, and TPO142-738 MT3, respectively (Table 4 and Figure 1). To investigate the effects of the mutations on the full-length TPO protein structure and functions, we also predicted the 3D structures for

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

| Sl no. | Exon | Nucleotide position | Amino acid position | Functional change (polarity) | Reference |
|-------|------|---------------------|---------------------|-----------------------------|-----------|
| 1     | 8    | c.1117G>T           | p.Ala373Ser         | Similar to non-enzymatic reaction rate | [20, 52] |
| 2     | 8    | c.1193G>C           | p.Ser398Thr         | Low enzymatic reaction rate  | [20, 52] |
| 3     | 12   | c.2145C>T           | p.Pro715Pro         | Not applicable              | [20, 52, 53] |
| 4     | 12   | c.2173A>C           | p.Thr725Pro         | Similar to non-enzymatic reaction rate | [20, 52, 53] |

### Table 4: Summary of the corresponding model numbers, C-score, TM-score, and the RMSD-score of the predicted 3D structures of TPO142-738, TPO1-933, TPO142-738 MT1, TPO142-738 MT2, TPO142-738 MT3, TPO1-933 WT, TPO1-933 MT1, TPO1-933 MT2, and TPO1-933 MT3.

| Features | TPO142-738 | TPO1-933 | TPO142-738 MT1 | TPO142-738 MT2 | TPO142-738 MT3 | TPO1-933 WT | TPO1-933 MT1 | TPO1-933 MT2 | TPO1-933 MT3 |
|----------|------------|----------|---------------|---------------|---------------|------------|-------------|-------------|------------|
| Model no. | 01         | 01       | 01            | 01            | 01            | 02         | 01          | 05          | 01         |
| C-score  | 2          | 2        | 1.99          | 1.99          | –             | –          | 0.35        | 0.13        | 0.38        |
| 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, 5]; 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), and MT3 = mutant 3 (p.Thr725Pro).

For molecular dynamics simulation, YASARA dynamics program was employed and [48, 49]. AMBER14 force field was considered for all calculations. The size of the cubic simulation box was 167.17 Å × 167.17 Å × 167.17 Å and 151,343 water molecules were added to maintain a solvent density of 1.0 g/ml. 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 [50]. 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 [50]. 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 docked structures were analyzed using the same protocol as stated before.

3. Result

3.1. Demographic Information of the Study Participants.

The crystallographic structure of TPO protein is available with low resolution and the catalytic domain of TPO142-738 is similar to human myeloperoxidase (MPO). Since the mutations that were identified in this present study were confined in the MPO-like domain of TPO protein, we investigated the effect of mutations on the 3D structure of MPO-like domain (TPO142-738) of TPO protein. Also, we wanted to see whether the mutations caused any changes in the interactions of various amino acid residues with heme prosthetic group. We submitted the amino acid TPO142-738 sequence for the wild type (WT), mutant 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 TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, and TPO142-738 MT3, respectively (Table 4 and Figure 1). To investigate 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. From the server, we obtained 5 models 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 [54–58] (Figure 1 and Supplementary Figure 2). 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 4.
Table 5: Summary of the Verify3D and RAMPAGE webservice results for the predicted 3D structures of TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, TPO142-738 MT3, TPO-933 WT, TPO-933 MT1, TPO-933 MT2, and TPO-933 MT3 proteins.

| Protein     | Verify3D | Favored region | Allowed region | Outlier region |
|-------------|----------|----------------|----------------|---------------|
| TPO142-738 WT | 96.15%   | 89.78%         | 90.64%         | 94.14%        |
| TPO142-738 MT1 | 94.14%   | 85.20%         | 10.40%         | 4.40%         |
| TPO142-738 MT2 | 94.64%   | 81.20%         | 14.30%         | 4.50%         |
| TPO142-738 MT3 | 73.10%   | 68.40%         | 17.20%         | 5.50%         |
| TPO-933 WT   | 70.74%   | 69.60%         | 16.20%         | 14.20%        |
| TPO-933 MT1  | 76.10%   | 66.90%         | 18.40%         | 12.70%        |
| TPO-933 MT2  | 76.63%   | 69.80%         | 18.80%         | 11.40%        |
| TPO-933 MT3  | 76.73%   | 76.80%         | 18.90%         | 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), and MT3 = mutant 3 (p.Thr725Pro).

3.4. Validation of the 3D Structures of TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, TPO142-738 MT3, TPO-933 WT, TPO-933 MT1, TPO-933 MT2, and TPO-933 MT3 Proteins. We validated the 3D structures of TPO WT, MT1, MT2, and MT3 proteins by Verify3D server which measures the accuracy of the predicted 3D structure model with its respective residues (1D) by assigning a structural class based on its location and environment. 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 TPO142-738 WT (96.15%), TPO142-738 MT1 (89.78%), TPO142-738 MT2 (94.14%), and TPO142-738 MT3 (94.64%) (Table 5). Moreover, validation of the structures by the RAMPAGE webserver also provided the percentages of the amino acid residues within favored, allowed, and outlier regions. In case of TPO142-738 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 TPO142-738 MT1, 84.4% residues were in the favored regions, 11.1% in the allowed region, and 4.4% in the outlier regions. On the one hand, for TPO142-738 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 TPO142-738 MT3, 81.2% residues were confined in the favored region, 14.3% in the allowed region and 4.5% in the outlier regions (Table 5). The Verify3D results showed that the percentages of amino acid residues having average 3D-1D score of ≥ 0.2 for TPO-933 WT, TPO-933 MT1, TPO-933 MT2, and TPO-933 MT3 were 73.10%, 70.74%, 76.10%, and 76.63%, respectively. The results provided by RAMPAGE server for TPO-933 WT, TPO-933 MT1, TPO-933 MT2, and TPO-933 MT3 were summarized in Table 5.

3.5. Optimization of Heme. Density functional theory using B3LYP/SDD was employed for the optimization of the heme prosthetic group. Slight changes in bond distances and bond angles were observed between crystal and DFT structures which are presented in Table 6 and Supplementary Figure 1.

3.6. Molecular Docking, Visualization and Analysis of the Docking of Heme with the Predicted Structures of TPO142-738 WT, TPO142-738 MT1, TPO142-738 MT2, TPO142-738 MT3, TPO-933 WT, TPO-933 MT1, TPO-933 MT2, and TPO-933 MT3 Proteins. From the analysis, it was found that heme interacted with TPO-933 WT through a total of 21 non-bond interactions including interactions with Arg491, and Arg582 through hydrogen bonds, interactions with His239, Val400, Phe490, Arg491, His494, Ile497, Phe523, Leu560, Leu564, and Leu575 through hydrophobic interaction, and interactions with Arg396, Glu399, and Arg491 through electrostatic interactions (Table 7 and Figure 2). Moreover, similar to TPO-933 WT, when MPO-like domain (TPO142-738) WT predicted structure was used for docking, 21 non-bond interactions of amino acids with heme were observed as well (Figure 3). However, not all of these 21 interactions for TPO142-738 and TPO-933 were common. The TPO142-738 WT interactions with heme included Arg582 and Arg586 through hydrogen bonds, His239, Val400, His494, Ile497, Phe523, Leu560, Leu564, Leu566, Leu567, and Leu575 through hydrophobic interactions and Arg396, and Glu399 through electrostatic interactions. The common 11 interactions for TPO-933 WT and TPO142-738 WT included Arg396, Glu399, and Arg491 through electrostatic interactions (Table 7 and Figure 2). Moreover, similar to TPO142-738 WT, when TPO142-738 MT1 predicted structure was used for molecular docking, a total of 19 residues were found to interact with heme. Among those residues, 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 interactions (Table 7, Figure 2 and Supplementary Table 1). Total number of interactions decreased significantly for TPO142-738 MT1 compared to the TPO-933 WT and the Glu399 residue which is crucial for interaction was also absent in the TPO142-738 MT1. On the other hand, when TPO142-738 MT1 predicted structure was used for molecular docking, a total of 19 residues were found to interact with heme (Table 7, Figure 3 and Supplementary Table 1). Among those 19 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 interactions, and the residue Glu399 interacted through electrostatic interaction with the heme (Table 7). However, the crucial interactions of heme with His239, Arg396, and His494 were absent in the TPO142-738 MT1. The structure-based docking of both TPO-933 MT1 and TPO142-738 MT1 suggested that MT1 mutation was damaging for the TPO enzyme activity.

For TPO-933 MT2, a total of 20 amino acid residues were found to interact with the heme. Among those residues, 3
Table 6: Summarized data for the bond distances and bond angles of specific atoms of heme before and after optimization.

| Atoms     | X-ray Structure Bond distance (nm) | DFT structure Atoms Bond distance (nm) | X-ray Structure Bond angle (degrees) | DFT Structure Atoms Bond angle (degrees) |
|-----------|-----------------------------------|----------------------------------------|-------------------------------------|----------------------------------------|
| 39N..43Fe | 1.977                             | 13N..43Fe                              | 2.013                               | 82.23°                                 |
| 40N..43Fe | 2.393                             | 14N..43Fe                              | 2.008                               | 80.66°                                 |
| 41N..43Fe | 2.088                             | 15N..43Fe                              | 2.008                               | 94.79°                                 |
| 42N..43Fe | 1.845                             | 42N..43Fe                              | 2.011                               | 98.39°                                 |
Table 7: Binding energy and non-bond interactions of heme with TPO\textsubscript{142-738} WT, TPO\textsubscript{142-738} MT1, TPO\textsubscript{142-738} MT2, TPO\textsubscript{142-738} MT3, TPO\textsubscript{1-933} WT, TPO\textsubscript{1-933} MT1, TPO\textsubscript{1-933} MT2, and TPO\textsubscript{1-933} MT3 proteins after flexible docking.

| Protein type | Binding Affinity (kcal/mol) | Hydrogen bond | Hydrophobic bond | Electrostatic bond | Total interactions |
|--------------|-------------------------------|----------------|------------------|--------------------|--------------------|
| TPO\textsubscript{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\textsubscript{142-738} MT1 | -10.8 | Met90(Met231), Gly93(Gly234), Ser261(Ser402), Gly352(Gly493) | Gly93(Gly234), Gly94(Gly235), Val259(Val400), Phe349(Phe490), Arg350(Arg491), Ile356(Ile497), Phe382(Phe523), Leu419(Leu560), Leu423(Leu564), Leu426(Leu567), Leu434(Leu575) | Glu258(Glu399) | 19 |
| TPO\textsubscript{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\textsubscript{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) | Glu258(Glu399) | 16 |
| TPO\textsubscript{1-933} WT | -11.5 | Arg491, Arg582 | His239, Val400, Phe490, Arg491, His494, Ile497, Phe523, Leu560, Leu564, Leu575 | Arg396, Glu399, Arg491 | 21 |
| TPO\textsubscript{1-933} MT1 | -3.2 | His239, His494 | His239, Phe243, Arg396, Phe524, Leu567 | Not found | 12 |
| TPO\textsubscript{1-933} MT2 | -11.5 | His239, Arg491, Arg582 | His239, Val400, Arg491, His494, His497, Phe523, Leu564, Val566, Leu575 | Arg396, Glu399 | 20 |
| TPO\textsubscript{1-933} MT3 | -7.9 | Gln246, Arg491, Ser568 | Phe243, Val400, Arg491, His494, His497, 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\textsubscript{142-738} the amino acid outside the first bracket indicates the position in predicted structure for TPO\textsubscript{142-738} and the amino acid residues in first bracket indicates the real position in TPO\textsubscript{1-933} protein; WT = wild type; MT1 = mutant 1 (p.Ala373Ser); MT2 = mutant 2 (p.Ser398Thr); MT3 = mutant 3 (p.Thr725Pro).
residues 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 interactions, and 2 residues including Arg396 and Glu399 interacted through electrostatic interactions (Table 7, Figure 2 and Supplementary Table 2). All four crucial interacting residues, namely, His239, Arg396, Glu399, and His494, which are important for MPO-like domain activity in TPO enzyme, were found to interact with heme in TPO1-933 MT2. On the other hand, when TPO142-738 MT2 predicted structure was used for docking, a total of 20 residues were found to interact with heme. Among those 20 residues, 3 residues including Arg491, Asn579, and Arg582 interacted with heme through hydrogen bonds, the residues Phe243, His494, Phe523, Phe524, Leu564, Phe565, Val566, Leu567, and Leu575 interacted through hydrophobic interactions and the residue Glu399 interacted through electrostatic interaction (Table 7, Figure 3 and Supplementary Table 1). However, the crucial interactions of heme with residues His239 and Arg396 were absent in the TPO142-738 MT2. The TPO1-933 structure-based docking showed that the mutations did not result in any major changes in interactions and all major interacting residues were present, whereas the TPO142-738 structure-based docking suggested result which was similar to TPO1-933 except that the major interacting residues His239 and Arg396 were absent. Together, the docking results are indicative of the fact that this mutation might have an association with PIOD.

For TPO1-933 MT3, a total of 21 residues were found to interact with the heme. Among those residues, 3 residues including Gln246, Arg491, and Ser568 interacted with heme through hydrogen bonds, 11 residues including Phe243, Arg491, His494, Ile497, Phe523, Phe524, Leu564, Val566, Leu567, and Leu575 interacted through hydrophobic interactions and the residue Glu399 interacted through electrostatic interaction (Table 7, Figure 2 and Supplementary Table 1). However, the crucial interactions of heme with residues His239 and Arg396 which are important for MPO-like domain activity...
in the TPO enzyme, were absent in TPO_{1-933} MT3. On the other hand, when TPO_{142-738} MT3 predicted structure was used for docking, a total of 16 residues were found to interact with heme. Among those 16 residues, 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 interactions and the residue Glu399 interacted through electrostatic interaction (Table 7, Figure 3 and Supplementary Table 1). However, the crucial interactions of heme with His494 were absent in the TPO_{142-738} MT3. Both TPO_{1-933} and TPO_{142-738} structure-based docking with heme suggested that MT3 mutation might have damaging effect to the TPO protein activity.

3.7. QM/MM and Molecular Dynamics (MD) Simulation. According to Guria et al. TPO_{1-933} MT1 (p.Ala373Ser) and TPO_{1-933} MT3 (p.Thr725Pro) showed more damaging effect on the catalytic activity of TPO protein [20] and our molecular docking-based study showed that full-length TPO_{1-933} MT2 (p.Ser398Thr) structure interacted to all the crucial amino acids in the catalytic site of TPO; thus we selected these TPO_{1-933} MT1 (p.Ala373Ser) and TPO_{1-933} MT3 (p.Thr725Pro) mutant cases for further analysis to validate the molecular docking results and to compare the amino acid interactions with the wild type structure. 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 groups were further studied by QM/MM and the mode of interactions is depicted in Figure 4. It was observed that the interacting residues were the same as obtained from the initial docking results.

To investigate the structural changes during molecular dynamics simulation, the protein-ligand complex after 5000 ps of simulation was superimposed on the initial docked protein-ligand complex. The superimposed structures of

Figure 3: Non-bond interactions of heme with corresponding predicted structures as obtained using a BIOVIA Discovery Studio 2017. (a) TPO_{142-738} WT, (b) TPO_{142-738} MT1, (c) TPO_{142-738} MT2, and (d) TPO_{142-738} MT3. The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding numbers.
Figure 4: 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) TPO\textsubscript{1-933} WT, (b) TPO\textsubscript{1-933} MT1, and (c) TPO\textsubscript{1-933} MT3. The amino acid residues and their positions are designated as the three letter abbreviations and the corresponding numbers.

4. Discussion

Though congenital hypothyroidism (CH) is the most common preventable disorder, newborn screening is not a regular practice in Bangladesh. Due to late initiation of treatment, many late-diagnosed hypothyroid patients experience various typical signs and symptoms of hypothyroidism even though they receive regular levothyroxine treatment. Although several genes have been reported to be involved in thyroid dysshormonogenesis (TDH), mutation in the TPO gene is frequently described with mild to severe repercussions resulting in partial iodine organification defect (PIOD) to total iodine organification defect (TIOD) [21]. In this present study, we investigated (a) the mutational spectrum in the TPO gene of TDH patients and (b) the influence of specific mutation on TPO protein structure by means of in silico approach. To the best of our knowledge, this is the first molecular investigation on genetic etiology of TDH in Bangladesh.

In this study, we analyzed exons 8-14 of TPO gene of the TDH patients as the previous studies had reported that this region was crucial for the enzymatic activity and mutations in this region may result in absence or reduction in TPO activity [20, 59]. Analysis of 36 specimens revealed three non-synonymous mutations including p.Ala373Ser (TPO MT1), p.Ser398Thr (TPO MT2), and p.Thr725Pro (TPO MT3) and one synonymous mutation p.Pro715Pro. The identified non-synonymous mutations had previously been reported to be pathogenic or disease-causing mutations [51, 53]. Moreover, a cloning-based study involving aforementioned mutations
by Guria et al. reported that these mutations could result in low expression of TPO mRNAs as well as a reduction in TPO enzyme activity [20]. They also demonstrated that mutation p.Ala373Ser (TPO MT1) was more damaging than mutation p.Ser398Thr (TPO MT2). This phenomenon could be explained by a change in aliphatic amino acid to hydroxyl-containing amino acid at 373th position of TPO protein, whereas the other mutation (p.Ser398Thr) did not result in such a shift from one group of amino acids to another. However, the mutation in exon-12, namely, p.Thr725Pro (TPO MT3), could result in a failure of TPO protein to shift to its active state as it is well reported that threonine is the phosphorylation site for protein activation [60, 61]. Moreover, p.Thr725Pro had been reported to be associated with autoimmune hypothyroidism [62]. As CH is genetically and phenotypically diverse, molecular studies may provide additional information for diagnosis, classification, and prognosis of the disease. Particularly, in patients with normal thyroid gland morphology, it could be difficult to distinguish between thyrotropin resistance and dyshormonogenesis; and molecular genetic studies may reveal true etiology of the disease in these cases. In this study, at least one of this three damaging mutations was found in a homozygous state or two of them were in a heterozygous state in each of the study participants. We previously reported a mutational hot-spot in the HBB gene and established a cost-effective molecular method (high resolution melting curve analysis) for screening of HBB gene mutations in Bangladesh [63]. Such a cost-effective approach can be adopted for TDH patients and carrier screening targeting the TPO genes in Bangladesh.

Substitution of an amino acid affects the shape, function, or binding properties of a given protein. With growing importance of genetics and genomics in health sector, considerable efforts have been devoted to linking human phenotypes to genotypic variations at the nucleotide level and associated changes in 3D protein structure [64, 65]. Since the identified mutations were pathogenic, we wanted to investigate how these mutations were related to dyshormonogenesis by affecting the structural integrity and function of TPO protein.
To date, there is no high resolution X-ray crystallographic structure for any of the TPO proteins (wild type or mutated) in the public databases. However, there were structures for closely related proteins, namely, myeloperoxidases (MPO) and lactoperoxidases (LPO) [11, 12]. Thus we predicted and validated the three dimensional (3D) structures of TPO (WT and MT) protein using various bioinformatics tools. We studied the effects of mutations on the structural integrity, arrangement of various domains, and folding patterns of the TPO protein. To further study mutational effects, we performed molecular docking of heme in the catalytic site of TPO to investigate how these mutations could affect the activity of TPO enzyme. We obtained the binding energies of heme for 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 analyzed the detailed results of molecular docking by observing the non-bond interactions of heme with specific amino acid residues to understand the effects of mutations on the functions of TPO protein.

In our study, we applied our bioinformatics approaches on the MPO-like domain of TPO (TPO142-738) and full-length TPO protein (TPO1-933) targeting the non-bond interactions between the heme group and amino acid residues as the mutations identified in this study were found in this region and this MPO-like domain was the catalytic site of TPO enzyme [54]. We obtained the predicted structures for the wild type and the mutant proteins from the I-TASSER server for TPO142-738 and all the structures had significant confidence score (C-score) suggesting that the structures were valid for further studies [22]. However, I-TASSER predicted structures for TPO1-933 had lower C-score value and the reasons may be due to prediction for a very large protein, because I-TASSER predicts structures which are based on iterative threading and also homology modelling and as TPO has no crystallographic structure, this still remains a challenge [22, 24]. We also verified the structures using the Verify3D and the RAMPAGE web servers and both of the servers gave satisfactory results [27, 66]. As heme is crucial for the catalytic activity of TPO, 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. We observed a decrease in heme binding energies in the cases of mutant TPO proteins suggesting that the catalytic activity of mutant TPO might have been hampered. Further investigation gave us information about the heme interactions with specific amino acid residues of TPO protein. The heme prosthetic group of wild type structure interacted with all the important amino acid residues including His239, Arg396, Glu399 and His494 through non-bond interactions.
But in the cases of mutant structures, some of the important amino acid interactions were absent, suggesting that the mutations might have damaging effects on heme interactions and thereby affecting the catalytic activity of TPO enzyme. Guria et al. demonstrated that MT1 (p.Ala373Ser), MT2 (p.Ser398Thr), and MT3 (p.Thr725Pro) had damaging effect on TPO mRNA expression and enzyme activity [20]. However, MT1 (p.Ala373Ser) and MT3 (p.Thr725Pro) showed iodination reactions similar to the nonenzymatic 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 [20]. This finding is consistent with our molecular docking-based study targeting both the MPO-like domain of TPO (TPO<sub>142-738</sub>) and the full-length TPO protein (TPO<sub>1-933</sub>) predicted structures. MT1 and MT3 had more enhanced influence on the interactions between several crucial residues and the heme group, whereas MT2 had less influence on the interaction between the TPO protein and the heme prosthetic group. Different studies showed that quantum chemical methods such as DFT or QM/MM can also be employed for investigating the intermolecular interactions [67–69]. Thus, to validate the results of molecular docking, we performed QM/MM calculations on the full-length wild type TPO protein and two mutant proteins (MT1 and MT3) that had severe damaging effect on the catalytic activity of TPO.

The heme group was treated using a semiempirical approach of calculations and the protein was treated using an approach of molecular mechanical calculations. The structures obtained after performing QM/MM calculations were taken into consideration for further analysis of the presence of non-bond interactions of the active site residues with the heme group. The interactions observed were exactly similar to those obtained from molecular docking, suggesting that the docking results obtained were valid.

In cellular environment, protein structures are always in a dynamic condition. To mimic the cellular environment, we performed molecular dynamics simulation in a cubic simulation box containing water molecules and NaCl and a pH of 7.4 was maintained. After performing 5000 ps simulation under the influence of AMBER14 force field, the final protein structures were superimposed on the corresponding docked protein structure. The results infer that the protein-ligand complexes studied were considerably stable over time as there was very small change in the coordinates of the heme group after the simulation. To observe the effect of protein dynamics, the protein structures were retrieved from the 5000 ps simulation snapshot and again docked with the heme ligand. The finding was rather interesting as the wild type protein was still found to be interacting with the crucial amino acid residues. However, these important interactions were found to be absent for the mutant proteins that might be a potential cause of a decrease in the enzymatic activity of TPO protein.

To understand the in-depth characteristics of TPO protein of both wild type and mutant forms, we performed in silico approach to mimic cellular environment. However, various computational chemistry-based approaches like IR and Raman spectroscopy can be used to identify the changes in molecular level with higher specificity [70, 71]. Such compositional analysis in various cellular conditions has become very popular for characterizing the biochemical changes in various disease conditions and also for the study on characterizing the spectrum of various hormones such as corticosteroids [70–72]. The study by Claudio et al. analyzed the molecular vibrational spectrum of thyroid tissues from normal and disease conditions which could ultimately represent the characteristics of secondary structures of proteins [73]. Although use of IR and Raman spectroscopy could offer more insights into the physiological conditions of TDH patients carrying mutations in the TPO gene, the present study was not subjected to such approaches because we did not have such facilities.

In this present study, we had used computational approaches such as molecular docking, QM/MM and molecular dynamics simulation to investigate the effect of mutations on TPO protein interactions. The present study could be useful for future studies including study of the effect of mutations on TPO dimer organization and how mutations in the TPO gene could lead to a change in the TPO protein conformation.

5. Conclusion

Our study investigated the genetic etiology of Bangladeshi patients with TDH, which may further help us to screen and categorize the disease. Three pathogenic mutations were observed in the patients with TDH including p.Ala373Ser, p.Ser398Thr, and p.Thr725Pro. In silico based study revealed that p.Ala373Ser and p.Thr725Pro mutations resulted in a significant change in interactions between the amino acid residues of TPO protein and the heme group, whereas the p.Ser398Thr mutation could affect the TPO protein to a lesser extent. The results of this study may help better understand the correlation between specific mutation in the TPO gene and altered biological activities of the TPO protein as well as disease severity among the TDH patients. Furthermore, future study on dimer formation of TPO protein and functional activity study of TPO enzyme in physiological condition is expected to shed light on how mutation in the TPO gene can affect thyroid hormone biosynthesis pathway and find a mutation-based better treatment strategy for individual patients.

Data Availability

The data used to support the findings of this study are included within the article.

Ethical Approval

This study was approved by the Ethical Review Board for Human Studies of BSMMU and University of Dhaka.

Conflicts of Interest

The authors declare that they have no conflicts of interest.
Authors’ Contributions

Mst. Noorjahan Begum, Md Tarikul Islam, Golam Sarower Bhuyan, Suprovath Kumar Sarker, Shekh Rezwan Hossain, and Imrul Shahriar conducted the laboratory work, data analysis, and manuscript draft writing. Shahinur Haque, Tasnia Kawsar Konika, Asifuzzaman Rahat, Rosy Sultana, Hurjan Banu, Narayan Saha, M. A. Hasanat, Mizanul Hasan, Abu A. Sajib, Emran Kabir Chowdhury, Hossain Uddin Shekhhar, Abul B. M. K. Islam, Suraiya Begum, Md. Sazzadul Islam, and Sadia Sultana assisted in specimen and clinical information collection. Kaissar Mannoor, Firdausi Qadri, Syed Saleheen Qadri, Mohammad A. Halim, Syeda Kashfi Qadri, and Sharif Akhteruzzaman designed the study plan and supervised the overall project. Md Tarikul Islam, Shekh Rezwan Hossain, and Golam Sarower Bhuyan contributed equally to this work. All authors have read and approved the paper for publication.

Acknowledgments

The authors are thankful to UGC for its generous support and also grateful to ideSHi, RGRC, and BSMMU. This study was supported by grant (CP-4029) from the Higher Education Quality Enhancement Project of the University Grant Commission (UGC) of Bangladesh.

Supplementary Materials

Supplementary Figure 1: Structural differences of heme before and after optimization from three different points of view.

Supplementary Figure 2: The predicted 3D structures of the proteins. (A) TPO1-933 WT, (B) TPO1-933 MT1, (C) TPO1-933 MT2, and (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, and IC = intracellular region shown in red color, catalytic center shown in blue color.

Supplementary Figure 3: QM/MM ONIOM calculation set up for protein-ligand complexes, high level (ball and stick), and low level (line). Supplementary Table 1: Binding energies and non-bond interactions of heme with TPO1-933 WT, TPO1-933 MT1, TPO1-933 MT2, and TPO1-933 MT3 after flexible docking. Supplementary Table 2: Binding energies and non-bond interactions of heme with TPO1-933 WT, TPO1-933 MT1, TPO1-933 MT2, and TPO1-933 MT3 after flexible docking. (Supplementary Materials)

References

[1] M. V. Rastogi and S. H. LaFranchi, “Congenital hypothyroidism,” Orphanet Journal of Rare Diseases, vol. 5, no. 1, article 17, 2010.
[2] D. A. Fisher, “Second international conference on neonatalthyroid screening: progress report,” Journal of Pediatrics, vol. 102, no. 5, pp. 653–654, 1983.
[3] M. Hasan, N. Nahar, A. Ahmed, and F. Moslem, “Screening for congenital hypothyroidism-a new era in Bangladesh,” The Southeast Asian Journal of Tropical Medicine and Public Health, vol. 34, pp. 162–164, 2003.
[4] C. C. Lee, F. Harum, M. Y. Jalaludin, C. Y. Lim, K. L. Ng, and S. Mat Junit, “Functional analyses of c.2268dup in thyroid peroxidase gene associated with goitrous congenital hypothyroidism,” BioMed Research International, vol. 2014, Article ID 370538, 11 pages, 2014.
[5] H. Cangul, Z. Ayacan, A. Olivera-Nappa et al., “Thyroid dyshor- monogenesis is mainly caused by TPO mutations in consan-guineous community,” Clinical Endocrinology, vol. 79, no. 2, pp. 275–281, 2013.
[6] B. Bakker, H. Bikker, T. Vulsma, J. S. E. De Randamie, B. M. Wiedijk, and J. J. M. De Vijlder, “Two decades of screening for congenital hypothyroidism in the Netherlands: TPO gene mutations in total iodide organisation defects (an update),” The Journal of Clinical Endocrinology & Metabolism, vol. 85, no. 10, pp. 3708–3712, 2000.
[7] F. S. Belforte, M. B. Miras, M. C. Olcese et al., “Congenital goitrous hypothyroidism: Mutation analysis in the thyroid peroxidase gene,” Clinical Endocrinology, vol. 76, no. 4, pp. 568–576, 2012.
[8] A. Taurog, M. Dorris, and D. R. Doerge, “Evidence for a radical mechanism in peroxidase-catalyzed coupling. I. Steady-state experiments with various peroxidases,” Archives of Biochemistry and Biophysics, vol. 315, no. 1, pp. 82–90, 1994.
[9] Y. Endo, S. Onogi, K. Umeki et al., “Regional localization of the gene for thyroid peroxidase to human chromosome 2p25 and mouse chromosome 12C,” Genomics, vol. 25, no. 3, pp. 760–761, 1995.
[10] J. P. Banga, D. Mahadevan, G. J. Barton et al., “Prediction of domain organisation and secondary structure of thyroid peroxidase, a human autoantigen involved in destructive thyroiditis,” FEBS Letters, vol. 266, no. 1-2, pp. 133–141, 1990.
[11] A. Gardas, M. K. Sohi, B. J. Sutton, A. M. McGregor, and J. P. Banga, “Purification and crystallisation of the autoantigen thyroid peroxidase from human Graves’ thyroid tissue,” Biochemical and Biophysical Research Communications, vol. 234, no. 2, pp. 366–370, 1997.
[12] E. Hendry, G. Taylor, K. Ziemnicka, F. Gunnern Jones, J. Furmaniak, and B. Rees Smith, “Recombiant human thyroid peroxidase expressed in insect cells is soluble at high concentration and fits the autoantigen,” Journal of Endocrinology, vol. 160, no. 3, pp. R13–R15, 1999.
[13] T. J. Fiedler, C. A. Davey, and R. E. Fenna, “X-ray crystal structure and characterization of halide-binding site of human myeloperoxidase at 1.8 Å resolution,” The Journal of Biological Chemistry, vol. 275, no. 16, pp. 1964–1971, 2000.
[14] A. K. Singh, N. Singh, S. Sharma et al., “Crystal Structure of Lactoperoxidase at 2.4 Å Resolution,” Journal of Molecular Biology, vol. 376, no. 4, pp. 1060–1075, 2008.
[15] S. N. Le, B. T. Poerbski, J. McCoey et al., “Modelling of thyroid peroxidase reveals insights into its enzyme function and autoantigenicity,” PLoS ONE, vol. 10, no. 12, Article ID e0142615, 2015.
[16] C. M. Rivolta, S. A. Esperante, L. Gruñeiro-Papendieck et al., “Partial iodide organification defect caused by a novel mutation of the thyroid peroxidase gene in three siblings,” Clinical Endocrinology, vol. 59, no. 2, pp. 198–206, 2003.
[18] C. Rodrigues, P. Jorge, J. Pires Soares et al., “Mutation screening of the thyroid peroxidase gene in a cohort of 55 Portuguese patients with congenital hypothyroidism,” *European Journal of Endocrinology*, vol. 152, no. 2, pp. 193–198, 2005.

[19] J.-Y. Wu, S.-G. Shu, C.-F. Yang, C.-C. Lee, and F.-J. Tsai, “Mutation analysis of thyroid peroxidase gene in Chinese patients with total iodide organization defect: identification of five novel mutations,” *Journal of Endocrinology*, vol. 172, no. 3, pp. 627–635, 2002.

[20] S. Guria, B. Bankura, N. Balmiki et al., “Functional analysis of thyroid peroxidase gene mutations detected in patients with thyroid dyshormonogenesis,” *International Journal of Endocrinology*, vol. 2014, Article ID 390121, 8 pages, 2014.

[21] H. Bäker, F. Baas, and J. J. M. De Vijlder, “Molecular analysis of mutated thyroid peroxidase detected in patients with total iodide organization defects,” *The Journal of Clinical Endocrinology & Metabolism*, vol. 82, no. 2, pp. 649–653, 1997.

[22] Y. Zhang, “I-TASSER server for protein 3D structure prediction,” *BMC Bioinformatics*, vol. 9, no. 1, article 40, 2008.

[23] A. Roy, A. Kucukural, and Y. Zhang, “I-TASSER: a unified platform for automated protein structure and function prediction,” *Nature Protocols*, vol. 5, no. 4, pp. 725–738, 2010.

[24] J. Yang, Y. Yan, A. Roy, D. Xu, J. Pison, and Y. Zhang, “The I-TASSER Suite: protein structure and function prediction,” *Nature Methods*, vol. 12, no. 1, pp. 7–8, 2015.

[25] J. U. Bowie, R. Luthy, and D. Eisenberg, “A method to identify protein sequences that fold into a known three-dimensional structure,” *Science*, vol. 253, no. 5016, pp. 164–170, 1991.

[26] R. Luthy, J. U. Bowie, and D. Eisenberg, “Assessment of protein models with three-dimensional profiles,” *Nature*, vol. 356, no. 6364, pp. 83–85, 1992.

[27] S. C. Lovell, I. W. Davis, W. B. Arendall et al., “Structure validation by Cα geometry: φ, ψ and Cβ deviation,” *Proteins: Structure, Function, and Bioinformatics*, vol. 50, no. 3, pp. 437–450, 2003.

[28] A. D. Becke, “Density-functional exchange-energy approximation with correct asymptotic behavior,” *Physical Review A: Atomic, Molecular and Optical Physics*, vol. 38, no. 6, pp. 3098–3100, 1988.

[29] C. Lee, W. Yang, and R. G. Parr, “Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density,” *Physical Review B: Condensed Matter and Materials Physics*, vol. 37, no. 2, pp. 785–789, 1988.

[30] M. Frisch, G. Trucks, and H. Schlegel, *GAUSSIAN 09*, Gaussian 09, Revision D. 01, Gaussian Inc, Wallingford, Conn, USA, 2009.

[31] M. Dolg, U. Wedig, H. Stoll, and H. Preuss, "Energy-adjusted ab initio pseudopotentials for the first row transition elements," *The Journal of Chemical Physics*, vol. 86, no. 2, pp. 866–872, 1986.

[32] D. Andreu, U. Haussermann, M. Dolg, H. Stoll, and H. Preuss, "Energy-adjusted ab initio pseudopotentials for the second and third row transition elements," *Theoretical Chemistry Accounts*, vol. 77, no. 2, pp. 123–141, 1990.

[33] S. Dallakyan and A. J. Olson, "Small-molecule library screening by docking with PyRx," *Methods in Molecular Biology*, vol. 1263, pp. 243–250, 2015.

[34] O. Trotta and A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading," *Journal of Computational Chemistry*, vol. 31, no. 2, pp. 455–461, 2010.

[35] T. Lengauer and M. Rarey, "Computational methods for biomolecular docking," *Current Opinion in Structural Biology*, vol. 6, no. 3, pp. 402–406, 1996.

[36] P. G. Furtmüller, M. Zederbauer, W. Jantschko et al., “Active site structure and catalytic mechanisms of human peroxidases,” *Archives of Biochemistry and Biophysics*, vol. 445, no. 2, pp. 199–213, 2006.

[37] A. Taurog, “Molecular evolution of thyroid peroxidase,” *Biochimie*, vol. 81, no. 5, pp. 557–562, 1999.

[38] L. Schrödinger, "LLE2010The PyMOL Molecular Graphics System, Version 1.4. The PyMOL molecular graphics system," 2010.

[39] BIONIA DS, *Discovery Studio Modeling Environment*, Dassault Systèmes, San Diego, Calif, USA, 2017.

[40] T. Vreven, K. S. Byun, I. Komáromi et al., “Combining quantum mechanics methods with molecular mechanics methods in ONIOM,” *Journal of Chemical Theory and Computation*, vol. 2, no. 3, pp. 815–826, 2006.

[41] K. G. Vladimirovich, N. R. Nizamovich, T. N. Borisovna et al., “Improvement technology manufacturing poly radioprotective microbial poliantigen,” *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, vol. 7, no. 4, pp. 2193–2199, 2016.

[42] T. Vreven, M. J. Frisch, K. N. Kudin, H. B. Schlegel, and K. Morokuma, “Geometry optimization with QM/MM methods II: Explicit quadratic coupling,” *Molecular Physics*, vol. 104, no. 5-7, pp. 701–714, 2006.

[43] T. Vreven and K. Morokuma, “On the Application of the IMOMO (integrated molecular orbital + molecular orbital) Method,” *Journal of Computational Chemistry*, vol. 21, no. 16, pp. 1419–1432, 2000.

[44] S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma, and M. I. Frisch, "A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives," *Journal of Molecular Structure: THEOCHEM*, vol. 461-462, pp. 1–21, 1999.

[45] F. Maseras and K. Morokuma, "IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states," *Journal of Computational Chemistry*, vol. 16, no. 9, pp. 1170–1179, 1995.

[46] M. Frisch, G. Trucks, H. B. Schlegel et al., *Gaussian 09*, revision a. 02, Gaussian, Inc, Wallingford, Conn, USA, 2009.

[47] J. P. Stewart, “Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements,” *Journal of Molecular Modeling*, vol. 13, no. 12, pp. 1173–1213, 2007.

[48] E. Krieger, T. Darden, S. B. Nabuurs, A. Finkelstein, and G. Vriend, “Making optimal use of empirical energy functions: Force-field parameterization in crystal space,” *Proteins: Structure, Function, and Genetics*, vol. 57, no. 4, pp. 678–683, 2004.

[49] F. Krieger, B. Fierz, O. Bieri, M. Drewello, and T. Kiefhaber, “Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding,” *Journal of Molecular Biology*, vol. 332, no. 1, pp. 265–274, 2003.

[50] T. Darden, D. York, and L. Pedersen, “Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems,” *The Journal of Chemical Physics*, vol. 98, no. 12, pp. 10089–10092, 1993.

[51] M. Avbelj, H. Tahirowic, M. Debjak et al., “High prevalence if thyroid peroxidase gene mutations in patients with thyroid dyshormonogenesis,” *European Journal of Endocrinology*, vol. 156, no. 5, pp. 511–519, 2007.

[52] C. C. Lee, F. Harun, M. Y. Jalaludin, C. H. Heh, R. Othman, and S. Mat Junit, “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, vol. 2013, Article ID 987186, 7 pages, 2013.

[53] N. Balmiki, B. Bankura, S. Guria et al., “Genetic analysis of thyroid peroxidase (TPO) gene in patients whose hypothyroidism was found in adulthood in West Bengal, India,” Endocrine Journal, vol. 61, no. 3, pp. 289–296, 2014.

[54] F. Libert, J. Ruel, M. Ludgate et al., “Thyroperoxidase, an auto-antigen with a mosaic structure made of nuclear and mitochondrial gene modules,” EMBO Journal, vol. 6, no. 13, pp. 4193–4196, 1987.

[55] S. Kimura and M. Ikeda-Saito, “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 Genetics, vol. 3, no. 2, pp. 113–120, 1988.

[56] M. Huse and J. Kuriyan, “The conformational plasticity of the same gene family,” Proteins: Structure, Function, and Genetics, vol. 3, no. 2, pp. 113–120, 1988.

[57] H. B. Bikker, T. Vulsma, F. Baas, and J. J. M. de Vijlder, “Identification of five novel inactivating mutations in the human thyroid peroxidase gene by denaturing gradient gel electrophoresis,” Human Mutation, vol. 6, no. 1, pp. 9–16, 1995.

[58] V. Le Fourn, M. Ferrand, and J.-L. Franc, “Endoproteolytic cleavage of human thyroperoxidase: Role of the propeptide in the protein folding process,” The Journal of Biological Chemistry, vol. 280, no. 6, pp. 4568–4577, 2005.

[59] O. P. Cherkasova, M. M. Nazarov, D. A. Sapozhnikov et al., “Vibrational spectra of corticosteroid hormones in the terahertz range,” in Laser Applications in Life Sciences, vol. 7376, p. 73760P, International Society for Optics and Photonics, 2010.

[60] R. Prabhakar, P. E. M. Siegbahn, B. F. Minaev, and H. Ågren, “Activation of triplet dioxygen by glucose oxidase: Spin-orbit coupling in the superoxide ion,” The Journal of Physical Chemistry B, vol. 106, no. 14, pp. 3742–3750, 2002.

[61] B. F. Minaev and V. A. Minaeva, “Spin-dependent binding of dioxygen to heme and charge-transfer mechanism of spin-orbit coupling enhancement,” Ukrainska Bioorganica Acta, vol. 6, no. 2, pp. 56–64, 2008.

[62] V. A. Minaeva, B. F. Minaev, G. V. Baryshnikov et al., “Temperature effects in low-frequency Raman spectra of corticosteroid hormones,” Optics and Spectroscopy, vol. 118, no. 2, pp. 214–223, 2015.

[63] K. Eberhardt, C. Steibing, C. Matthäus, M. Schmitt, and J. Popp, “Advantages and limitations of Raman spectroscopy for molecular diagnostics: an update,” Expert Review of Molecular Diagnostics, vol. 15, no. 6, pp. 773–787, 2015.

[64] C. A. Soto, L. P. Medeiros-Neto, L. dos Santos et al., “Infrared and confocal Raman spectroscopy to differentiate changes in the protein secondary structure in normal and abnormal thyroid tissues,” Journal of Raman Spectroscopy, vol. 49, no. 7, pp. 1165–1173, 2018.