Human Thyroid Peroxidase:

A structural and kinetic overview of wild type and mutant variants

Abstract: Human thyroid peroxidase is associated with severe thyroid malfunction where it's mainly involved as an organ specific autoantigen. Nonetheless, the proteins' three-dimensional structure remains largely unstudied. Computational analysis of its mutant variants and implications on enzyme kinetics can prove a useful tool to understand underlying pathways on thyroid associated conditions.

Author: Guillermo Whitney

Affiliations: No declared affiliations.

Background

Human thyroid peroxidase (hTPO) plays a critical role in the normal functioning of the thyroid gland. It is a membrane bound glycoprotein found in the apical membrane of thyroid follicular cells and is responsible for the catalysis of L-thyroxine and triiodo-L-thyronine (Belforte et al., 2015). The reaction starts with the iodination of tyrosine residues in thyroglobulin to form mono and di-iodinated forms which are coupled to form either T3 or T4. The enzymatic process requires the presence of iodine, hydrogen peroxide, and Tg (Ruf & Carayon, 2006).

The enzyme is encoded by the hTPO gene which spans over 150kb and is found in chromosome 2p25. The gene encodes 933 amino acids that make up a large extracellular domain, a transmembrane domain, and a short intracellular tail. The extracellular domain is responsible for catalytic activity. The biosynthesis of this enzyme is complex because of it possessing two active sites and its requirement of a heme group for functionality. Despite its importance as a hormone regulatory enzyme, hTPO's three-dimensional structure is yet to be determined experimentally.

hTPO is one of the main thyroid autoantigens. Its malfunction is associated with Grave's disease and Hashimoto's thyroiditis which respectively cause hyperthyroidism and hypothyroidism. These two conditions affect about 5% of the worldwide population making them the most common organ specific autoimmune disease in humans. Both of these conditions result from autoimmune surveillance of hypersecreting mutants whereby autoreactive T cells attack mutants in the thyroid with the cost of triggering a humoral response (Milo, Korem Kohanim, Toledano, & Alon, 2023). Similarly, in cases of congenital hyperthyroidism, mutations in the hTPO gene are the main cause of hormone dysregulation and iodine organification defects that lead to mental retardation in about 1:3000 newborns. Such is the case of the previously characterized missense mutants p.C808R, p.G387R and p.P499L that showed decreased enzyme activity, lower efficiencies, and higher Km's (Belforte et al., 2015).

When looking to understand an enzyme's functionality, Michaelis-Menten kinetics remains as one of the simplest and best-known models to describe kinetic parameters. The values of the equation relate reaction velocity to substrate concentration in a system where substrate (S) and enzyme (E) bind reversibly to form an enzyme substrate complex (ES) which then react irreversibly to form product (P) and release the enzyme (E):

$$E + S \leftrightarrow ES \rightarrow E + P$$

The kinetic values of this generalized equation are expressed in a Michaelis Menten equation through:

$$v = \frac{Vmax[S]}{Km[S]}$$

Where Vmax represents the maximum velocity of the system, and Km is the substrate concentration at which the reaction is 50% of the Vmax.

Not surprisingly, single point missense mutations are shown to affect enzymatic activity by reshaping not only active sites of enzymes but also their ability to biochemically interact with substrates. Hydrophilic/hydrophobic balance and charge distribution are shown to affect physiochemical and catalytic properties of proteins. Surface charge also appears to play and important role in interactions between charged groups that arise during the binding of substrates to enzymes and throughout thermal or pH deactivation of proteins (Longo & Combes, 1997). Therefore, hydrophobicity of a protein ultimately plays an essential role in determining a protein three-dimensional structure and therefore its kinetic activity.

All of this considered, computationally analyzing the effects of known missense mutations on the structure of thyroid peroxidase can yield important information regarding molecular weight, amino acid composition, and hydrophobicity. Associating these changes to the resulting Michaelis Menten kinetics for each mutant relative to wild type hTPO can provide valuable insight into the biochemistry underlying thyroid hormone dysregulation.

Computational methods:

In their paper "Kinetic characterization of human thyroperoxidase. Normal and pathological enzyme expression in Baculovirus System: A molecular model of functional expression."(Belforte et al., 2015), authors discuss the impact of three missense mutations on the kinetic parameters of hTPO and its possible implications on congenital hypothyroidism. The first step on the computational analysis performed therefore aims to compare the WT hTPO to the three reported mutants.

In this first coding chunk, alpha fold was used to visualize the WT of hTPO consisting of 933 amino acids. This is an AI system developed by DeepMind to predict a protein three-dimensional structure from its amino acid sequence. The database contains about 200 million entries to cover UniProt and SwissProt files. The WT hTPO enzyme reported in this paper was downloaded as a .pdb file designated in UniProt as P07202.

To perform this analysis, the Bio.PDB package was installed. This module focuses on working with crystal structures of biological macromolecules and creates structures from pdb files to analyze models, chains, residues, and atoms as needed. With this in mind, the second step of the analysis focused on analyzing some lower-level structures. The first step was therefore obtaining the residues (amino acids) in the WT protein from which we can determine what the molecular weight is. This was performed with the biopython polypeptide builder as follows:

```
ppb=PPBuilder()
for pp in ppb.build_peptides(structure):
        seq_info['Sequence'] = seq # store BioPython Seq() object
        seq_info['Molecular Weight'] = analyzed_seq.molecular_weight()
```

Like molecular weight, we can obtain the GRAVY (Grand Average of Hydropathy) value. This value is based on a computer program which evaluates the hydropathic character of a protein by taking into consideration hydrophobicity/hydrophilicity of each of the 20 amino acid side chains (Kyte & Doolittle, 1982). We can also use bio python to define amino acid compositions/percetages and then take all these values to determine secondary structures amino acid fractions (Carpenter, 2021).

The second part of this initial structure analysis would be to consider the single point mutations: p.C808R, p.G387R and p.P499L. These mutants were generated by the authors and expressed in a baculovirus system; therefore, they aren't available as .pdb files to analyze on alpha fold. However, the "MODELL" package can be used to perform specific residue mutations. This is a comparative protein structure modeling program that predicts the three-dimensional structure of a given template based on its alignment to proteins of a known structure (Webb & Sali, 2016). In this case the prediction would be on the mutants and the template would be the WT hTPO .pdb file.

After generating the mutant structures, the lower-level analysis to determine molecular weight, amino acid composition, GRAVY value, and secondary structure amino acid fractions was repeated to perform a comparative analysis to the WT hTPO.

In the second part of the computational analysis, Michalis-Menten kinetics were evaluated for the WT, and each of the three mutants. To determine these values, guaiacol was used as a chromogenic substrate in the oxidative reaction where after addition of hydrogen peroxide the substrate is broken down into colorless tetra-guaiacol by TPO. First, Michaelis-Menten dynamics were simulated with the "*scipy solve_ivp library*" where each of the Kcat, Km, and E_0 values were inputted with an initial guaiacol concentration of 50mM. Then the MM velocity curve was plotted as a function of the guaiacol concentration to observe ES activity differences. This curve was defined as: $v = Kcat\ x\ E0\ x\ \frac{[A]}{Km+[A]}$. Finally, the Lineweaver Burk plot was generated to visualize the reciprocals of Vmax, Km.

Results

When visualizing the WT hTPO we see that the protein is 933 amino acids in length with some distinguishable characteristics, including an extracellular domain, transmembrane domain, and short intracellular tail. It's of interest to note that the intracellular tail and transmembrane domain are formed by helixes and significantly extend from the protein core (Fig1).

The amino acid composition, hydrophobicity, and secondary structures of the WT were then considered. The enzyme is composed of one sequence, with a molecular weight of 102961.53. The GRAVY value for the enzyme was determined at -0.33912 which indicates it's a hydrophilic enzyme. Secondary structure amino acid fractions were determined at 0.2647 for helix, 0.2529 for turns, and 0.2743 for beta sheets.

Restructuring of pdb files to generate mutants was confirmed by looking at the amino acid compositions changes. For C808R, the molecular weight of the protein increases to 103060.6, and its GRAVY values decreases to -0.343529, indicating its more hydrophilic. The secondary structure proportions are helix:0.2647, turn:0.2518, and sheet:0.274. For G387R, a molecular weight of 103060.7, a GRAVY value of -0.343515, and secondary structure proportions of helix: 0.2647, turn: 0.2518, and sheet: 0.274 were observed. Finally, for P499L, a molecular weight of 103060.6, a GRAVY value of -0.3435152, and secondary structure proportions of helix: 0.2647, turn: 0.2518, sheet: 0.2743 were observed.

When visualizing kinetic parameters according to MM parameters we observe significant changes in mutants relative to the WT which presents a steep slope and takes about 3s to break down the 50mM of guaiacol. The C808R shows the highest decrease in activity, taking up to four times longer to break down 50mM of guaiacol. The slope of this reaction is significantly less steep than the other ones and appear to even out gradually. The G387R and P499L mutants are more closely related, taking about double the time than the WT to break down 50mM of guaiacol (Fig. 2).

Plotting these values to show Michaelis Menten velocity vs substrate concentration shows that while all variants increase in activity at similar rates, the mutants have significantly reduced Vmax values. C808R reaches a maximum velocity at about 9 mM/minute, while the second and third mutants show values of about 14 and 16 mM/minute (Fig. 3).

Finally, a Lineweaver-Burke plot was used to plot the velocity and substrate reciprocals. The graph shows us that all the mutants have reduced Km and Vmax values. When looking at the y-intercepts that describe 1/V, one can observe that M1 has the lowest enzyme velocity, followed by M2, M3, and WT. The x-intercept shows -1/Km which shows similar dynamics as 1/V, however the close association between M2 & M3 is more noticeable (Fig. 4).

Discussion:

Overall, computational analysis shows the mutants having lower kinetic activity which is visualized by seeing they take significantly longer to catalyze the conversion of guaiacol to tetra guaiacol. Noticeably, M2 and M3 are more closely related in kinetic values than M1 which is significantly affected by the C808R mutation. While three-dimensional large-scale changes where not visualized in the single missense mutations, charge and hydropathy parameters were altered which could be correlated to some of the kinetic activity variations.

Normal kinetic values can be associated with the WT hTPO expression which is shown as a 933 amino acid enzyme with an extracellular domain, transmembrane domain, and short intracellular tail which are observed through alpha fold predictions. Parallelly, the protein was computationally determined to be hydrophobic with an even distribution of amino acids found in helix, coil, and beta sheet structures. Through the guaiacol reaction, hTPO was shown to break down guaiacol into tetra-guaiacol at a relatively fast rate with a Vmax value of 26.040. Meanwhile, the three enzyme mutants showed decreased enzyme kinetics with Vmax values being lowest for the M1, followed by M2, and M3. These mutations are also associated with changes in molecular weight, secondary structure fractions, and increased hydrophilicity. While these point mutations might have had not affected the two active sites dramatically, they did affect the overall charge and structure of the protein. Because surface charge plays a significant role between charged groups in the formation of the ES group and its later irreversible transition into E+P, it's not surprising that overall Michalis-Menten dynamics were significantly altered. These results evidence the potential of using comparative protein analysis to understand the implications of variant hTPO mutants on health disorders which is even more significant with an enzyme such as this one that has an undetermined three-dimensional structure.

However, while alpha fold can predict the location of the different enzyme domains, it's still important to note that the structure has only been predicted through its homology to other proteins and not yet experimentally. Therefore, our understanding of the hTPO protein and the significance of its domains might change as more is deduced about this complex enzyme. Furthermore, while useful for a quick analysis, comparative modeling with programs such as MODELLER already imply a limitation in that the template is not the same as the protein being analyzed. This leads to errors in side chain packing, errors in regions without a template, errors due to misalignments, or even errors due to incorrect template utilization (Eswar et al., 2006).

To better understand the implication of mutations on hTPO kinetic activity one could look at bigger mutants in which whole exons are affected. This would be the case observed in Hashimoto's and Grave's disease and could help understand the severity of these single organ autoimmune disorders. At the same time, more accurate protein folding predictions could be done on the single point mutations studied in this analysis if alpha fold was used instead of MODELLER. The database used by alpha fold is much wider than the one template chosen in the later program and thus would enable one to better understand why such significant changes in Michalis-Menten dynamics are seen for the different variants. Nevertheless, with complex enzymes such as hTPO using comparative modelling remains a good technique to get an initial understanding of the biochemical nature of different variants.

Figures:

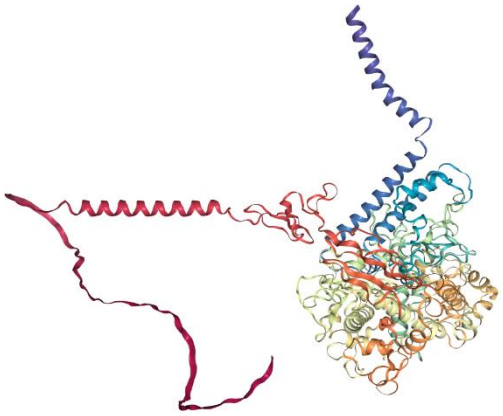

Figure 1: Alpha fold prediction of human thyroid peroxidase hormone visualized in bio python.

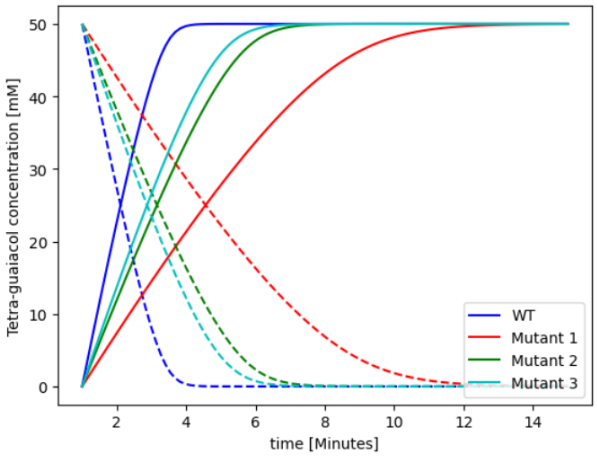

Figure 2: Michalis-Menten modelling of hTPO and three mutant types in guaiacol reaction.

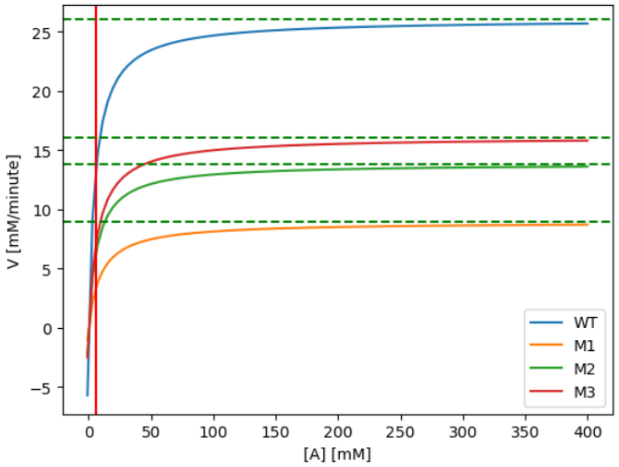

Figure 3: MM plot correlating enzyme variant velocity and guaiacol concentration.

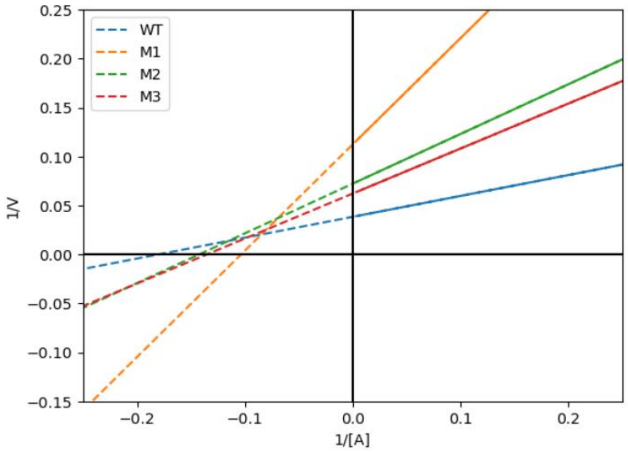

Figure 4: Lineweaver-Burke plot of hTPO and mutant variants showing 1/V, -1/Km, and Km/V.

References:

Belforte, F. S., Targovnik, A. M., González-Lebrero, R. M., Osorio Larroche, C., Citterio, C. E., González-Sarmiento, R., . . . Rivolta, C. M. (2015). Kinetic characterization of human thyroperoxidase. Normal and pathological enzyme expression in Baculovirus System: A molecular model of functional expression. *Molecular and Cellular Endocrinology, 404*, 9-15. doi:https://doi.org/10.1016/j.mce.2014.12.026

Carpenter, A. (2021). Visualizing and Analyzing Proteins in Python: Proteomics with Biopython. Retrieved from https://towardsdatascience.com/visualizing-and-analyzing-proteins-in-python-bd99521ccd

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., . . . Sali, A. (2006). Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics, Chapter 5*, Unit-5.6. doi:10.1002/0471250953.bi0506s15

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology, 157*(1), 105-132. doi:https://doi.org/10.1016/0022-2836(82)90515-0

Longo, M. A., & Combes, D. (1997). Influence of surface hydrophilic/hydrophobic balance on enzyme properties. *J Biotechnol, 58*(1), 21-32. doi:10.1016/s0168-1656(97)00120-x

Milo, T., Korem Kohanim, Y., Toledano, Y., & Alon, U. (2023). Autoimmune thyroid diseases as a cost of physiological autoimmune surveillance. *Trends in Immunology, 44*(5), 365-371. doi:10.1016/j.it.2023.03.007

Ruf, J., & Carayon, P. (2006). Structural and functional aspects of thyroid peroxidase. *Arch Biochem Biophys, 445*(2), 269-277. doi:10.1016/j.abb.2005.06.023

Webb, B., & Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics, 54*, 5.6.1-5.6.37. doi:10.1002/cpbi.3