

University of Richmond

## UR Scholarship Repository

---

Honors Theses

Student Research


---

4-29-2022

# Halogen Bonding: A Computational Chemistry Investigation of the Interaction between Thyroid Hormone and Deiodinase

William Rice  
*University of Richmond*

Follow this and additional works at: <https://scholarship.richmond.edu/honors-theses>

 Part of the [Biology Commons](#), and the [Molecular Biology Commons](#)

---

### Recommended Citation

Rice, William, "Halogen Bonding: A Computational Chemistry Investigation of the Interaction between Thyroid Hormone and Deiodinase" (2022). *Honors Theses*. 1598.  
<https://scholarship.richmond.edu/honors-theses/1598>

This Thesis is brought to you for free and open access by the Student Research at UR Scholarship Repository. It has been accepted for inclusion in Honors Theses by an authorized administrator of UR Scholarship Repository. For more information, please contact [scholarshiprepository@richmond.edu](mailto:scholarshiprepository@richmond.edu).

Halogen Bonding: A Computational Chemistry Investigation of  
the Interaction between Thyroid Hormone and Deiodinase

by

William Rice

Honors Thesis

Submitted to:

Biology Department  
University of Richmond  
Richmond, VA

April 29, 2022

Advisor: Dr. Kelling Donald

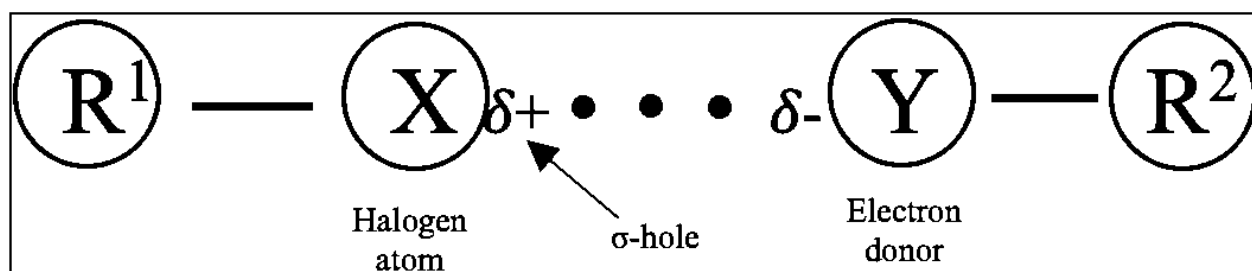
## Abstract

Halogen bonding is a noncovalent interaction that continues to garner interest among the scientific community. Investigation of halogen bonds in biological contexts typically revolves around rational drug design for developing therapeutics. However, halogen bonding may be occurring naturally in our body every day. Thyroid hormone and its regulating enzyme, iodothyronine deiodinase, show promising results for a halogen bonding interaction that happens during catalysis. Prior work has examined the interaction between the iodine of thyroid hormone and the selenium of iodothyronine deiodinase. However, this study is the first of its kind to use computational chemistry methods to analyze the halogen bond using the complete enzyme/ligand complex. While halogen bonding between thyroid hormone and deiodinase cannot be confirmed from this study, a more complex picture has been painted that alludes to the existence of such electrostatic interaction.

## Introduction

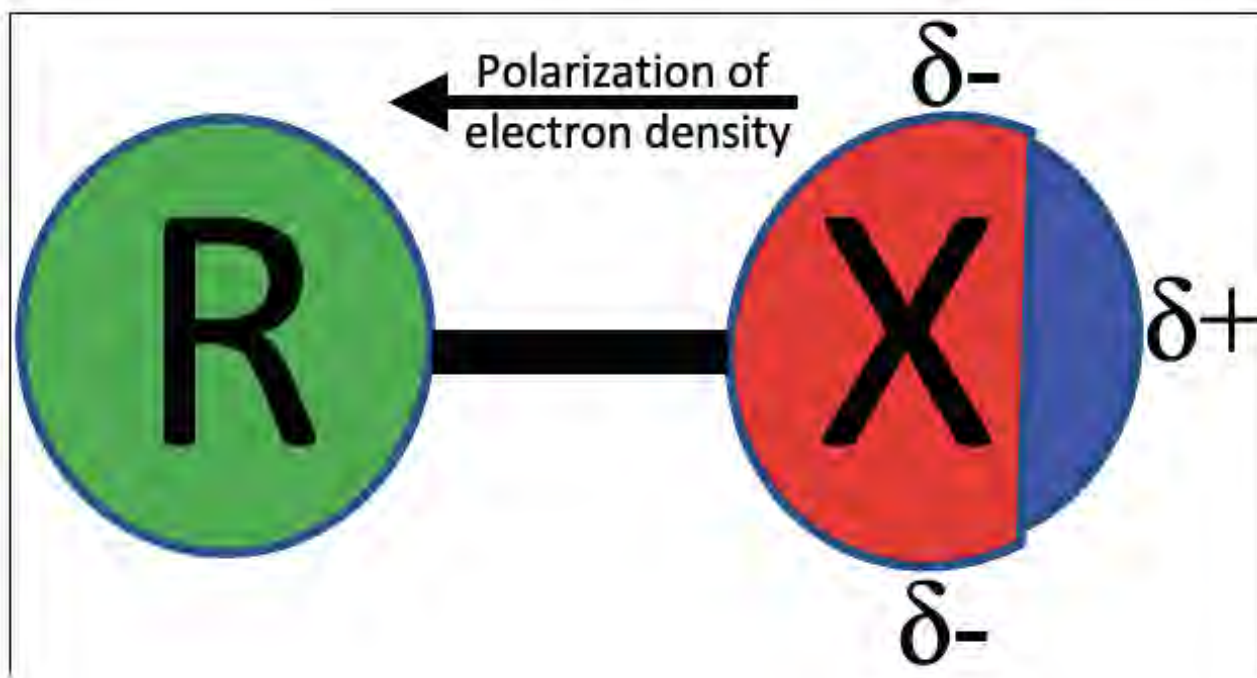
### Defining Halogen Bonding

Halogen bonding is an electrostatic chemical bonding phenomenon resulting from the attraction between a partially positive region on a halogen atom and a negatively charged electron-donor or base (Figure 1).<sup>1</sup>



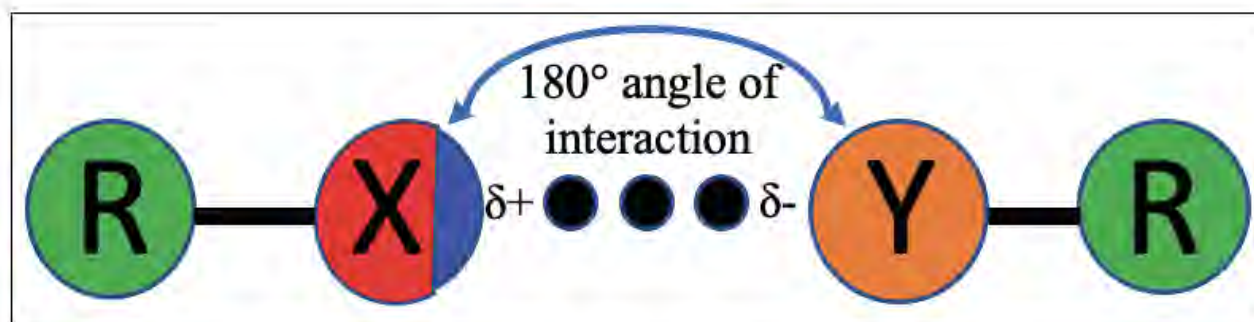
**Figure 1. Depiction of halogen bonding.** The halogen atom is shown as X and an electron donor species is shown as Y. Partial positive potential is shown as  $\delta^+$  and partial negative potential is shown as  $\delta^-$ . The halogen bonding interaction between X and Y is depicted as the three dots.

In accordance with their large electronegativity, halogen atoms are strong electron-withdrawing species, so the existence of a partially positive region on these atoms is counterintuitive. However, the partially positive region, referred to as the  $\sigma$ -hole, comes about due to an electron withdrawing group to which it is bonded pulling electron density from the halogen parallel to the R-X axis. Subsequently, the halogen atom has a nonuniform distribution of electrostatic potential, with negative electrostatic potential building up around the atom in a “belt”, and positive electrostatic potential at the distal tip, known as the  $\sigma$ -hole (Figure 2).



**Figure 2. Model of the nonuniform distribution of electron density on halogen atom.** The blue region of the halogen atom represents the partially positive potential of the  $\sigma$ -hole. Note that the  $\sigma$ -hole is opposite of the R-X axis. The filled lone pairs account for the negative electrostatic potential on the halogen perpendicular to the R-X axis.

As a result, the halogen atom can act as a Lewis acid that undergoes bonding interactions with a Lewis base. Analogous to the hydrogen bond, the halogen bond is a non-covalent interaction between an atom with a partial positive potential and an atom with a partial negative potential. In the case of the hydrogen bond, the interaction results from a partially positive hydrogen attracting a partially negative atom, such as an oxygen or a nitrogen. The size and strength of the  $\sigma$ -hole is directly related to the size of the halogen, with the very polarizable iodine exhibiting the most partially positive region, followed by bromine, chlorine, and lastly fluorine. Furthermore, the location of the  $\sigma$ -hole on the tip of the halogen, opposite of the R-X bond, requires a roughly  $180^\circ$  angle of interaction with the partially negative atom (Figure 3).



**Figure 3. 180° angle of interaction for halogen bonding.** The  $\sigma$ -hole only exists opposite of the R-X bond axis and therefore only attracts a partially negative atom at roughly 180°.

While there are important differences between halogen and hydrogen bonding, the electrostatic basis for both interactions draw stark comparisons. Since hydrogen bonding plays a crucial role in the conformation of important large molecules, such as the double helix in DNA or the secondary structure of proteins, investigating the impact of halogen bonding is undoubtedly warranted.

### History of Halogen Bonding in the Literature

Although hydrogen bonding has been of special interest to chemists and biologists for the last century, research on halogen bonding has been relatively limited. The phenomenon of halogen bonding has been observed since the 1800s as iodide was shown to react with ammonia to form iodammonium.<sup>2</sup> In the 1950s, chemists began to describe “charge-transfer” complexes in which activation barriers for chemical reactions were reduced in a quasi-dative bonding interaction.<sup>3</sup> Further examinations of charge-transfer bonding brought observations of polymerization of organic compounds with halogenated molecular bridges. A paper published in 1962 described repeating p-xylene/carbon-tetrabromide molecules, attributing the aggregation to charge-transfer bonding between the  $\pi$  electrons of the aromatic ring and the bromine atoms.<sup>4</sup> In 1970, a paper identified dioxane-bromine molecules arranged in a linear fashion with bond

lengths shorter than the van der Waals radii.<sup>5</sup> The strength of the charge-transfer bonds corresponded directly with the size of the halogen, as iodine displayed the strongest interaction while chlorine displayed the weakest.<sup>5</sup> Alternatively, the same dioxane-bromine interaction was described where the halogen molecule acted as an electron acceptor.<sup>6</sup> Shorter intermolecular distances between a halogen and a partially negative atom, along with elongated intramolecular bonds in the halogen-containing molecule, were documented in a larger range of compounds, including acetone, methanol, pyridine, and benzyl sulfide.<sup>6</sup>

In the early 1990s, the term “halogen bonding” began to circulate in the literature with reference to charge-transfer bonding. Specifically, in 1992, Bruggermann and Kochi referred to the activation of aromatic hydrocarbons by stannic chloride as “formal halogen bonding”.<sup>7</sup> In 1998, Alkorta et al. wrote about charge-transfer complexes; however, they labeled the complexes as halogen bonded complexes in an analogy to hydrogen bonded complexes.<sup>8</sup> The use of the term halogen bonding gradually became more popular with the expansion in knowledge of the halogen’s anisotropic electrostatic potential distribution and the “sigma hole”.<sup>9</sup> After the identification of the  $\sigma$ -hole, research on halogen bonding exploded. It showed promise as a strong non-covalent intermolecular force, analogous to and even out competing hydrogen bonding in some scenarios.<sup>10</sup> The halogen bonding phenomenon graduated from its earlier anonymous categorization under the general heading of charge-transfer complexes due to arguments for the contributions of electrostatics and polarity in regards to the strength of interactions.<sup>11</sup> An official IUPAC definition was finalized and accepted in 2013, stating that “a halogen bond occurs when there is evidence of net attraction between an electrophilic region associated with a halogen atom in a molecular entity and a nucleophilic region in another, or the

same, molecular entity.”<sup>12</sup> Applications of the interaction under the newly minted halogen bond label began to surface in supramolecular chemistry, crystal engineering, and biological settings.<sup>13</sup>

### **Halogen Bonding in a Biological Context**

When you think of halogens in biological systems, maybe the fluoride in toothpaste or the chloride in table salt comes to mind. Accordingly, when thinking of common elements within the body, halogens are never discussed as a fundamental part of terrestrial architecture. Unlike carbon, nitrogen, oxygen, phosphorus, and calcium, halogens are not used by nature to constitute biological organisms and are only present in the body in trace amounts. For this reason, biologists did not react to the alarm bells that went off in the 1990s in the same way that chemists did regarding the phenomenon of halogen bonding. However, roughly a decade later, the halogen bonding shock wave finally made its way to the biological realm. In 2004, a survey of the Protein Data Bank was published that searched for biological molecules containing at least one C-X bond, finding 226 single-crystal structures and 87 halogenated nucleic acids.<sup>14</sup> From here, the authors narrowed their search for interactions that fit the current halogen bond definition. The results found 66 proteins and 6 nucleic acid structures having X-O interactions that fit the criteria, totaling 113 unique halogen bonding interactions.<sup>14</sup> The list continued to grow, and biologists began to take special interest in exploring the potential benefits of halogen bonding in biological systems.

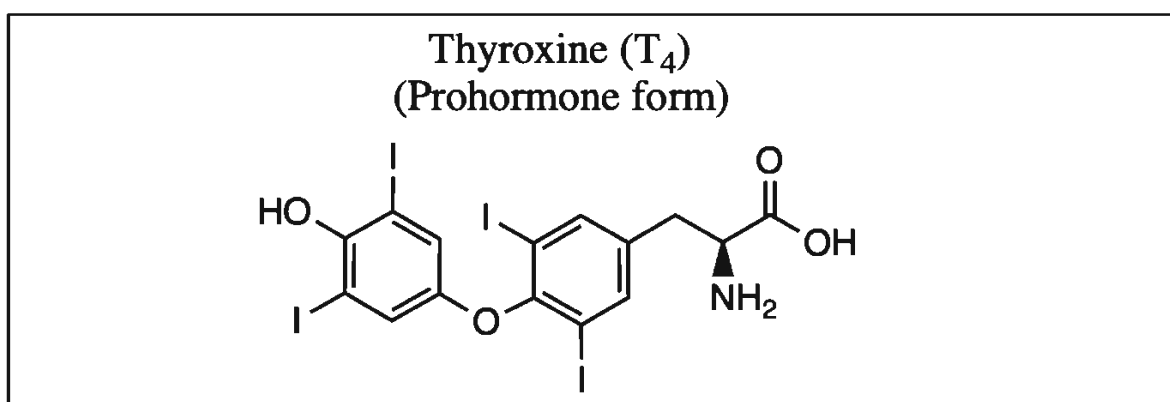
The halogen bond's strict conformation, strength of interaction, and tunability make it an attractive prospect for drug designers. By strategically incorporating a halogen atom into a molecular design, researchers can increase interaction strength between receptors and ligands.<sup>13</sup> The stirrings from the 2004 PDB Survey have produced a wide variety of publications within



medicinal chemistry, particularly oncology. An article from 2012 published promising results on reversing the inactivation of the p53 tumor suppressor through use of small molecules with incorporated halogen bonding sites.<sup>15</sup> Another article published in 2007 described the incorporation of halogen bonding into the molecular engineering of protein kinase inhibitors.<sup>16</sup> Halogen bonding was shown to cause a 20-fold increase in the binding affinity of the potential drug to the protein kinase ATP-binding domain.<sup>16</sup> While rational drug designers have utilized halogens as a way to modify receptor-ligand interactions, other researchers have identified innate halogen bonding interactions within the human body that may have large-scale implications on human health.<sup>13</sup>

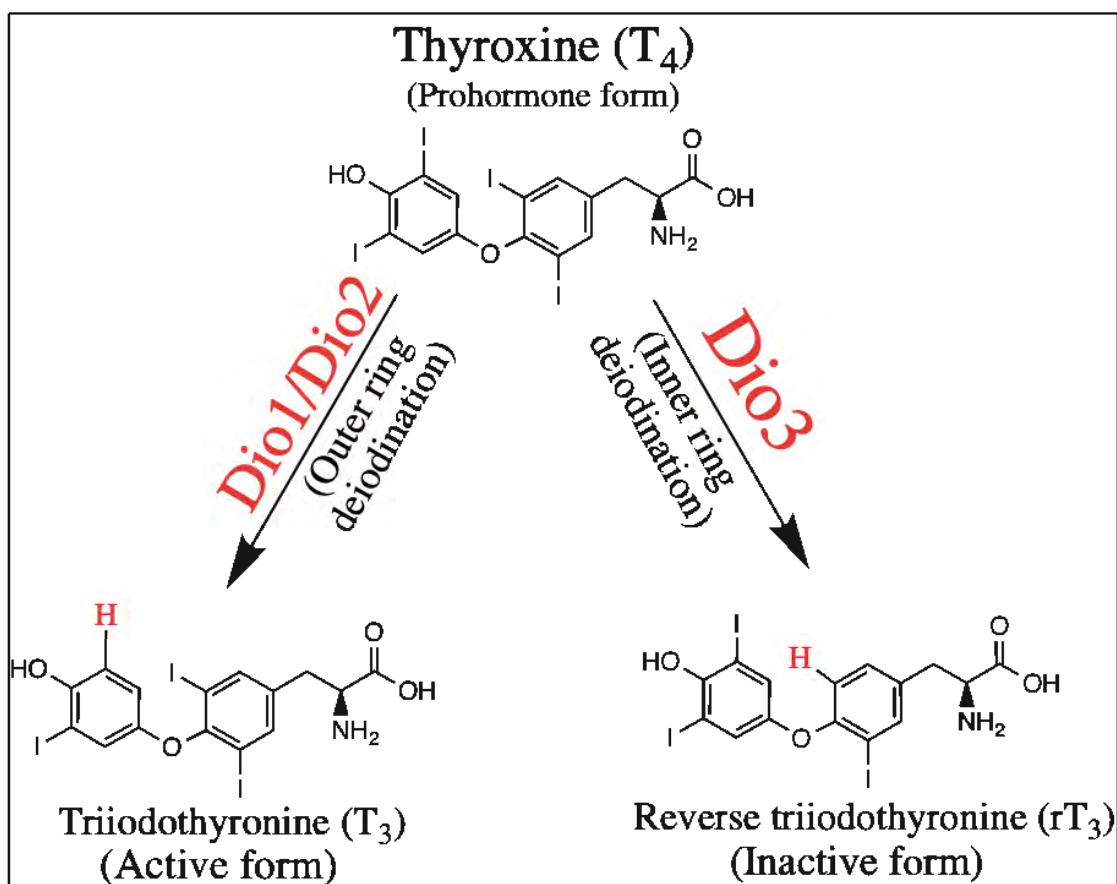
### Thyroid Hormones and Iodothyronine Deiodinase

An interesting place to search for halogen bonding in the human body is with thyroid hormones. Thyroid hormones are the only identified molecules in the body to incorporate iodine into their structures. Thyroxine (T<sub>4</sub>) contains four iodine atoms, hence the subscript in its abbreviation (Figure 4).



**Figure 4. Structure of thyroxine.** Thyroxine has an inner ring and an outer ring that each contain two iodine atoms. The inner ring iodine atoms are referred to the 3-iodine and 5-iodine, while the outer ring iodine atoms are the 3'-iodine and 5'-iodine. Thyroxine is the prohormone version of thyroid hormone because it is secreted by the thyroid glands and must be activated to produce its effects on tissue.

Activation of  $T_4$  occurs by removing the 5'-iodine from the outer ring. The resulting compound is called triiodothyronine, or  $T_3$ . However,  $T_4$  can also be permanently inactivated by removing the 5-iodine from the inner ring. The permanently inactive compound is called reverse triiodothyronine, or  $rT_3$ . The enzymes that remove iodine atoms from thyroid hormone are called iodothyronine deiodinases. Three types of iodothyronine deiodinases exist in humans: type 1 deiodinase, type 2 deiodinase, and type 3 deiodinase. Type 1 and 2 deiodinase perform outer ring deiodination, or activation of  $T_4$  to  $T_3$ . Type 3 deiodinase strictly performs inner ring deiodination, or inactivation of  $T_4$  to  $rT_3$  (Figure 5).



**Figure 5. Different types of deiodinases and their functions.** Dio1, Dio2, and Dio3 refer to type 1, type 2, and type 3 deiodinase, respectively. Dio1 and Dio2 perform outer ring deiodination of the 5'-iodine, effectively activating thyroxine. Dio3 only performs inner ring deiodination of the 5-iodine, permanently inactivating thyroxine.

Activation and inactivation of thyroid hormone functions to maintain thyroid homeostasis. Thyroid hormone is heavily involved in controlling metabolic rates and development in almost every organ system in the body.<sup>17</sup> Pathologically, overactive thyroid hormone can result in hyperthyroidism and underactive thyroid hormone can result in hypothyroidism. Therefore, the deiodinase pathway has major implications for human health and may be used as a tool to treat various thyroid disorders.

Bayse and Rafferty suggested that halogen bonding could be the basis for iodothyronine deiodinase activity.<sup>18</sup> All three types of deiodinases contain a very rare selenocysteine residue in their active site that performs the deiodination. The selenium atom of the selenocysteine, proposed by Bayse and Rafferty, may undergo a halogen bonding interaction with the iodine on thyroxine. They also suggested that the rendezvous of these two very rare trace elements in such a consequential biochemical pathway could be the result of selection due to halogen bonding. The strong nucleophilicity of selenium, in combination with the robust  $\sigma$ -hole of iodine, is the ideal scenario for halogen bonding to arise. To investigate the potential for halogen bonding, Bayse and Rafferty used computational chemistry methods to model the interaction between Dio1 and thyroxine. However, their model was very basic, using methyl selenol and an aryl iodide to represent the selenocysteine residue and thyroxine, respectively. Their results did indicate that halogen bonding does occur between iodine and selenium. Nevertheless, their oversimplified approach serves as a barrier to the validity of any conclusions on whether a halogen bonding occurs in the interaction between deiodinase and thyroxine.

## Taking a Different Approach

Expanding on Bayse and Rafferty's work, I also used computational chemistry methods to evaluate the potential halogen bonding interaction between iodine and selenium in the active site of iodothyronine deiodinase. However, my analysis utilized the full deiodinase enzyme and thyroxine ligand instead of truncated representations of each. Additionally, my investigation strictly focused on the inactivating deiodinase, Dio3. I hypothesized that a halogen bonding interaction was occurring in the active site of Dio3 between the inner ring 5-iodine and the selenium. To test this hypothesis, I first needed to ensure that a significant  $\sigma$ -hole existed on the 5-iodine. Then, I needed to put thyroxine in complex with Dio3 to evaluate the proximity and geometry of the 5-iodine and the selenium atom.

## Computational Methods

### Thyroid Hormone Optimization

The geometries of the different forms of thyroid hormone were optimized using the Gaussian 09 suite at the B3LYP level of theory.<sup>19</sup> Harmonic vibrational frequency analyses were completed in conjunction with the optimization at the B3LYP level of theory to ensure that the optimized geometries corresponded to the lowest potential surface energy of the molecule. Correlation-consistent triple- $\zeta$  (cc-pVTZ) basis sets were utilized for all atoms in thyroid hormone except for iodine. For iodine, an energy-consistent pseudopotential from the Stuttgart/Cologne was utilized to model the electrons of the atom.<sup>20</sup>

The quantitative analysis of the surface maxima of thyroid hormone was accomplished using the wavefunction analyzer Multiwfn.<sup>21</sup> A ".wfn" output file was coded for in the route section of the Gaussian input file. Surface maxima energy values corresponding to the different

$\sigma$ -holes were organized into an excel spreadsheet. Associated molecular pictures produced by Multiwfn were included as reference to the quantitative data.

Electrostatic potential surfaces and HOMO-LUMOs of the different thyroid hormones were visualized using GaussView 5.0.9.<sup>22</sup> All potentials were computed on the 0.001 au isodensity surface and visualized on a scale from  $-2.0e^{-2}$  to  $2.0e^{-2}$  atomic units.

### **Protein Optimization**

The PDB file 4TR3 was used as the starting point for Deiodinase type 3.<sup>23</sup> However, the crystal structure had to replace the typical selenocysteine with a cysteine during expression of the protein. To get a true model of the protein, Cys170 was swapped for selenocysteine, and the protein was computationally optimized using the ONIOM functionality from the Gaussian 09 suite. ONIOM stands for Our own N-layered Integrated molecular Orbital and molecular Mechanics. This approach allows for the optimization of large systems, such as proteins, by using a less computationally demanding method for the whole protein, referred to as the Low Layer, and a more computationally demanding method for the area of interest, such as the active site, referred to as the High Layer. The Low Layer in my calculation used the Amber molecular mechanics method while the High Layer used the Hartree Fock method with a 6-31g(d) basis set. The optimized protein was displayed using GaussView 5.0.9 with the High Layer in ball-and-stick format and the Low Layer in wire format.

### **Docking Study**

An open-source program called AutoDock Vina was used for the molecular docking study.<sup>24</sup> The Gaussian output files for both the protein and thyroxine had to be converted to a

.pdb file and then to a .pdbqt file in order to be compatible with AutoDock Vina. A grid box was centered around the active site with coordinates -9, -3, and 20 for the x, y, and z directions. The size of the box was 22, 14, and 26 Å in the x, y, and z dimensions. AutoDock computed the lowest energy conformation and produced a .pdb file that was analyzed using Pymol software.<sup>25</sup>

### Sequence Alignment

NCBI Homologene was used to locate the Dio3 sequence in humans.<sup>26</sup> NCBI BLAST was then used to compare the Dio3 gene in humans to the Dio3 gene in mice.<sup>27</sup> An alignment of the two sequences was then performed using Clustal Omega.

## Results and Significance

### Dio3 Gene Sequence Alignment between *Homo sapiens* and *Mus musculus*

To investigate the similarities in the Dio3 proteins between humans and mice, I searched for the Dio3 gene in NCBI's Homologene system to locate the sequence in *Homo sapiens*. NCBI BLAST provided that the Dio3 gene in mice is 95.05% identical to the gene in humans. Additionally, I aligned the two sequences using Clustal Omega (Figure 6). Most importantly from this analysis, the active site amino acids (S167, U170, E200, H202, H219, F258, and R275) were conserved between the two species. The selenocysteine residues (U170) are highlighted in yellow. The high degree of similarity between the two genes allows the analysis done on the mouse protein to be extrapolated to the human protein.

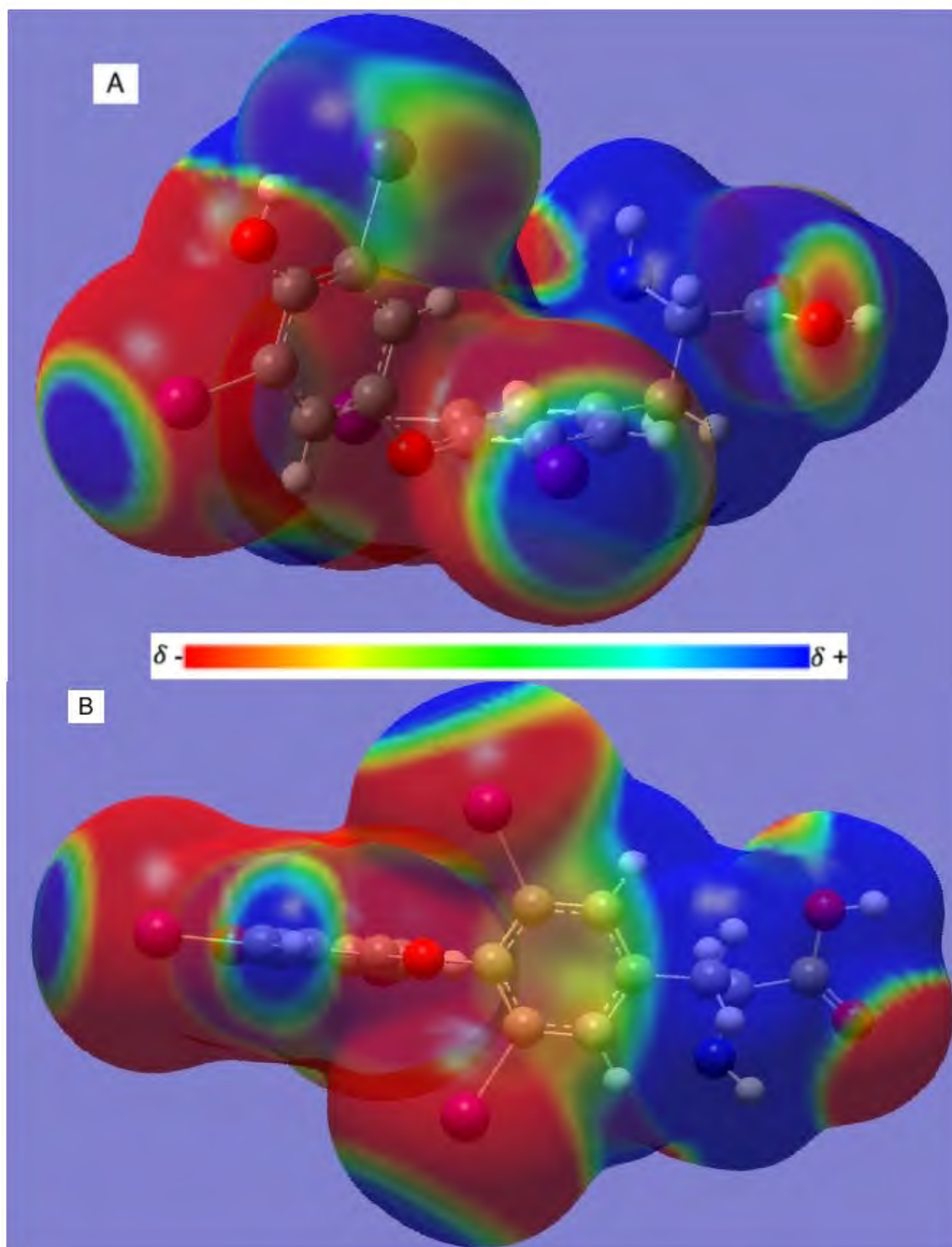
1	mprqaasrlvvgegegppgasgpaatmlrslllhsirlcagtasclvifp	50
	.     ..	
1	mprqatsrlvvgegegsggasgpaatmlrslllhsirlcagtasclvifp	50
51	rflgtafmlwlldfcirkhflrrrhpdhpepevelnsegeempddppi	100
	.. ..... :	
51	rflgtafmlwlldfcirkhflgrrrrggpepevelnsegeevppddppi	100
101	cvstdnrcltiasikavwhgqkldffkqaheggpapnsevvrpdgfgsqr	150
	.. ..... :	
101	cvstdnrcltiasikavwhgqkldffkqaheggpapnsevvlpdggfgsqh	150
151	ildyaggtrplvlnfgsctuppfmarmsafqrlvskyqrdrvdfliiyiee	200
	.     ..... :	
151	ildyaggtrplvlnfgsctuppfmarmsafqrlvskyqrdrvdfliiyiee	200
201	ahpsdgwvtttdspyvipqhrlsledrvsaarvlqggapgcalvldtmanss	250
	:     ..... :	
201	ahpsdgwvtttdspyvipqhrlsledrvsaarvlqggapgcalvldtmanss	250
251	ssaygayferlyvigsqgtimyyqggpdyqvselrtwlyrdeqlhgr	300
	..... :	
251	ssaygayferlyvigsqgtimyyqggpdyqvselrtwlyrdeqlhgar	300
301	phrf	304
	. . .	
301	prrv	304

**Figure 6.** Dio3 protein sequence alignment between *Homo sapiens* and *Mus musculus*. Sequences are 95.05% identical. Selenocysteine residue highlighted at the 170 positions in both sequences.

### $\sigma$ -Hole Analysis on Thyroid Hormone

An analysis of the electrostatic surface potentials on thyroid hormone was performed to investigate the size and strength of the  $\sigma$ -holes. The analysis consisted of generating qualitative electrostatic surface potential maps (Figure 7) that reflect the quantitative data extracted from the wavefunction file (Table 1).

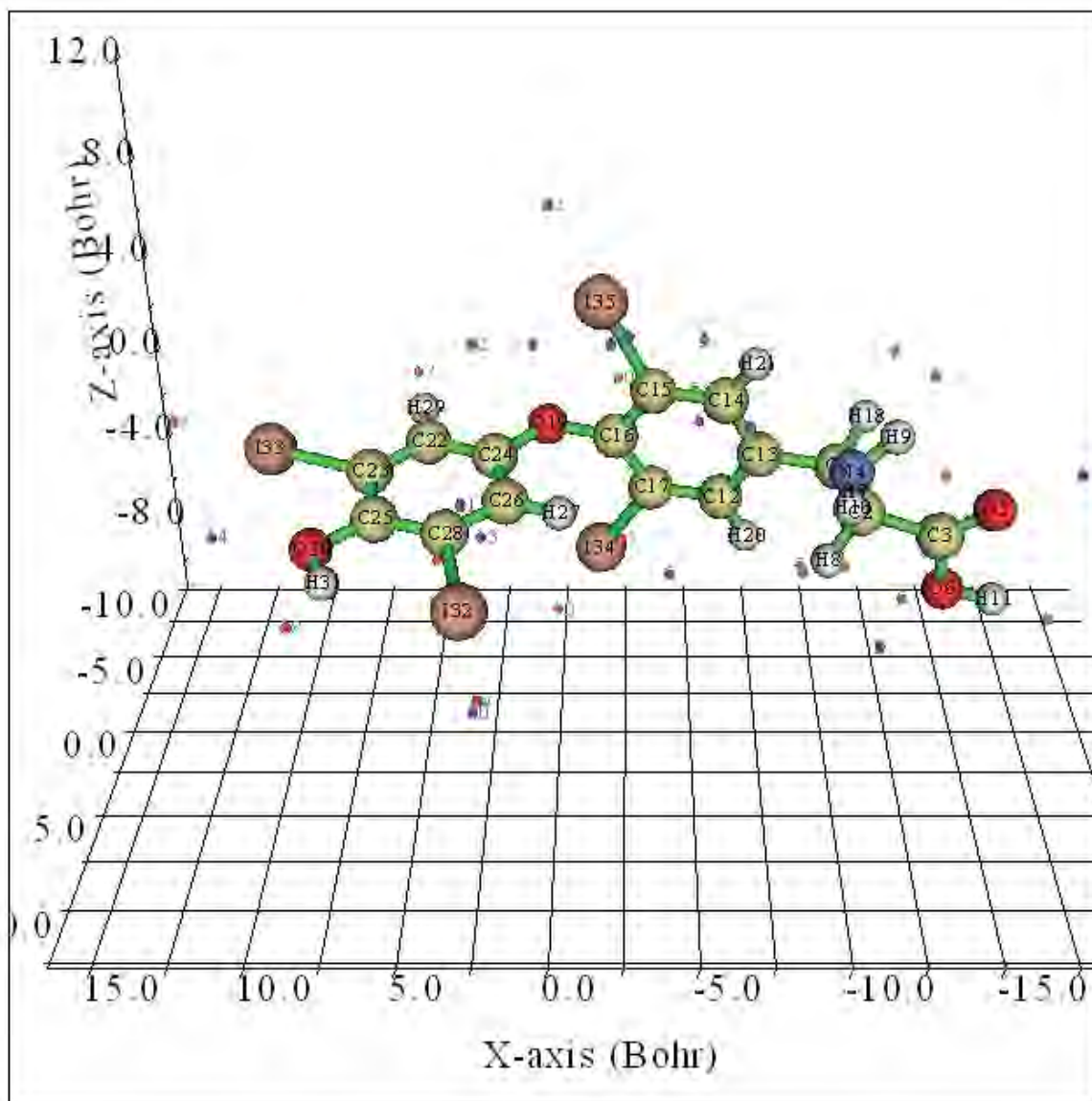
The electrostatic surface potential map shows the existence of four different  $\sigma$ -holes on thyroxine, each corresponding to one of the iodine atoms (Figure 7). Importantly, the inner ring 5-iodine atom, which is the target iodine for Dio3 deiodination, displays a significant positive electrostatic potential 180° from the C-I bond, consistent with the identification of a  $\sigma$ -hole.



**Figure 7A and B. Electrostatic surface potential maps of thyroxine.** Red indicates the areas of most negative electrostatic potential and blue indicates areas of most positive electrostatic potential. Each of the four iodine atoms have their own  $\sigma$ -hole. The target inner ring 5-iodine is shown in A as the atom coming straight out of the page and in B on the bottom of the molecule.



The quantitative results of the  $\sigma$ -hole analysis are provided in the form of a readout of the wavefunction output file using the Multiwfn software. The software provides a graph of the molecule with different points representing the local electrostatic potential minima and maxima of the molecule (Figure 8). The numbers on the purple and red dots correspond to the data points in the table produced by Multiwfn (Figure 9).



**Figure 8. Multiwfn output graph example.** Purple dots indicate local electrostatic potential minima and red dots indicate local electrostatic potential maxima. The target 5-iodine atom is labeled as I34, and its corresponding electrostatic potential maximum is 13.

Number of surface minima:		14		
#	a.u.	eV	kcal/mol	X/Y/Z coordinate(Angstrom)
1	-0.03878192	-1.055310	-24.336043	-9.446580 1.908287 1.177581
2	-0.01216442	-0.331011	-7.633297	-6.269628 0.816192 -3.253679
3	-0.03252822	-0.885138	-20.411781	-3.557580 1.706987 2.036503
4	-0.00327926	-0.089233	-2.057770	-2.852813 -3.162268 -0.424118
5	-0.00407752	-0.110955	-2.558687	-2.902820 -2.797732 1.573458
6	0.00494809	0.134644	3.104978	-2.182560 0.557083 -1.743172
7	-0.01273338	-0.346493	-7.990324	-1.407753 1.460146 3.679945
8	-0.01274957	-0.346933	-8.000484	-1.099151 1.522988 3.566520
9	-0.02692397	-0.732638	-16.895061	0.503328 -3.126468 1.310559
10	-0.00680674	-0.185221	-4.271298	1.432707 3.251870 -3.011709
11	-0.00746497	-0.203132	-4.684345	1.386790 4.016477 1.812280
12	-0.01604394	-0.436578	-10.067730	1.482413 -0.487463 2.597712
13	-0.01511596	-0.411326	-9.485414	1.554252 -1.388523 -2.118538
14	-0.03975328	-1.081742	-24.945582	6.182894 1.652945 -0.268199
Number of surface maxima:		19		
#	a.u.	eV	kcal/mol	X/Y/Z coordinate(Angstrom)
1	0.08777910	2.388591	55.082264	-9.378814 1.045231 -2.461788
2	0.03413299	0.928806	21.418792	-7.516526 -0.664428 -0.308077
3	0.03909422	1.063808	24.532011	-6.774675 1.386886 2.889624
4	0.02984614	0.812155	18.728754	-6.401923 -1.032003 2.185504
5	0.04388212	1.194093	27.536471	-6.268801 2.834959 -0.807373
6	0.05290529	1.439626	33.198598	-5.051871 3.889828 0.553699
7	0.03460231	0.941577	21.713296	-4.867602 -1.295710 -2.713171
8	0.03066225	0.834362	19.240866	-4.769848 0.223592 -1.938720
9	0.00015215	0.004140	0.095473	-2.755249 -2.782455 0.484129
10	0.00047790	0.013004	0.299888	-1.199092 -3.197544 0.553117
11	0.03077721	0.837490	19.313004	-1.277233 1.640824 -0.294966
12	0.03266525	0.888867	20.497768	0.052288 -0.753939 5.224569
13	0.03421968	0.931165	21.473191	0.092809 -2.576549 -4.625340
14	0.04043750	1.100360	25.374937	1.109742 5.687492 -0.954007
15	-0.00634595	-0.172682	-3.982145	2.301954 -0.004835 -1.792028
16	-0.00666845	-0.181458	-4.184517	2.318689 0.578481 1.688955
17	0.02061773	0.561037	12.937829	2.868957 -3.505631 0.523782
18	0.05548898	1.509932	34.819889	4.515332 3.958949 -0.675546
19	0.02786513	0.758249	17.485650	7.524725 -1.798828 0.267821

**Figure 9. Multiwfn output table example.** Each of the points from the output graph has energy and position data provided by the output table such as the one here. For the sake of my research, I only used the kcal/mol energy values to be consistent with previous literature on  $\sigma$ -holes.

$\sigma$ -holes are areas of positive electrostatic potential, so they are local maxima. I organized the  $\sigma$ -hole energy data for thyroxine ( $T_4$ ), triiodothyronine ( $T_3$ ), and reverse triiodothyronine

( $rT_3$ ) (Table 1).

**Table 1. Energy data for the  $\sigma$ -holes on T<sub>4</sub>, T<sub>3</sub>, and rT<sub>3</sub>.**

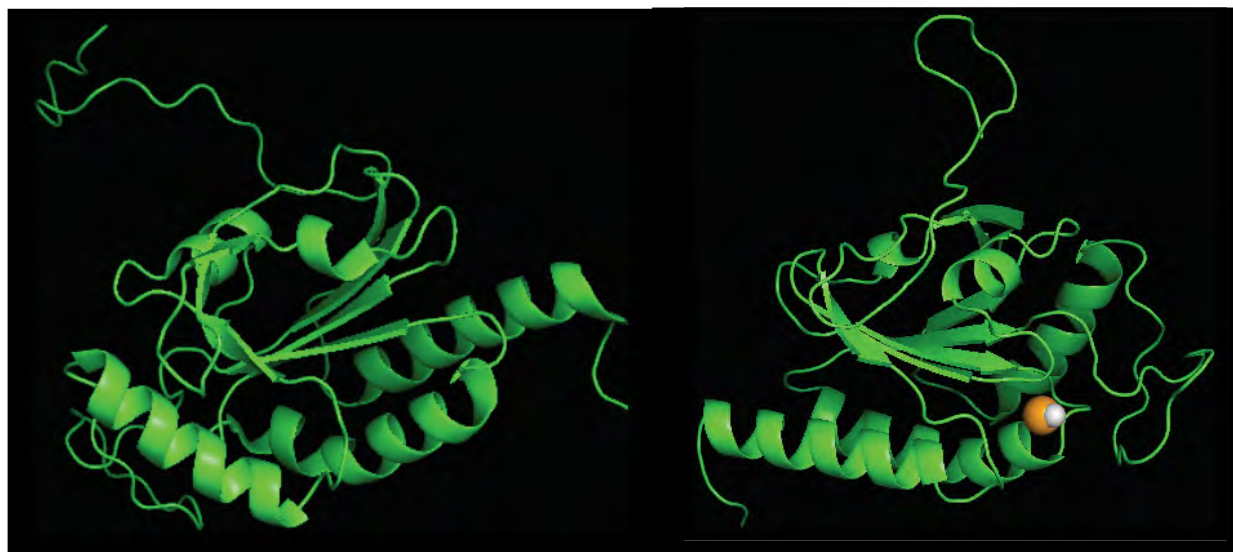
<b>Thyroid Hormone WFN Analysis</b>			
<b>Atom</b>	<b>Energy (kcal/mol)</b>		
	<b>(T<sub>4</sub>)</b>	<b>(T<sub>3</sub>)</b>	<b>(rT<sub>3</sub>)</b>
5-Iodine	21.47	19.20	
3-Iodine	20.50	20.22	15.42
5'-Iodine	25.37		20.34
3'-Iodine	17.49	14.85	27.52
Carboxyl Hydrogen	55.08	54.85	52.26
4'-Hydroxyl Hydrogen	34.82	53.50	37.92

Interestingly, the strongest  $\sigma$ -hole on T<sub>4</sub> is actually the outer ring 5'-iodine with an energy value of 25.37 kcal/mol. However, the inner ring 5-iodine is the second strongest, with an energy value of 21.47 kcal/mol. For comparison, I included the energy values for two different hydrogens that would exhibit typical hydrogen bonding activity. The energy values for the carboxyl hydrogen and the 4'-hydroxyl hydrogen are 55.08 kcal/mol and 34.82 kcal/mol, respectively. Therefore, the relative interaction strength potential of the halogen bonds in this molecule are roughly half that of typical hydrogen bonds. In summary, four different iodine-centered  $\sigma$ -holes exist on thyroxine, and the  $\sigma$ -hole at the 5-iodine position has significant halogen bonding potential.

### **Optimization of Dio3 with Selenocysteine in Active Site**

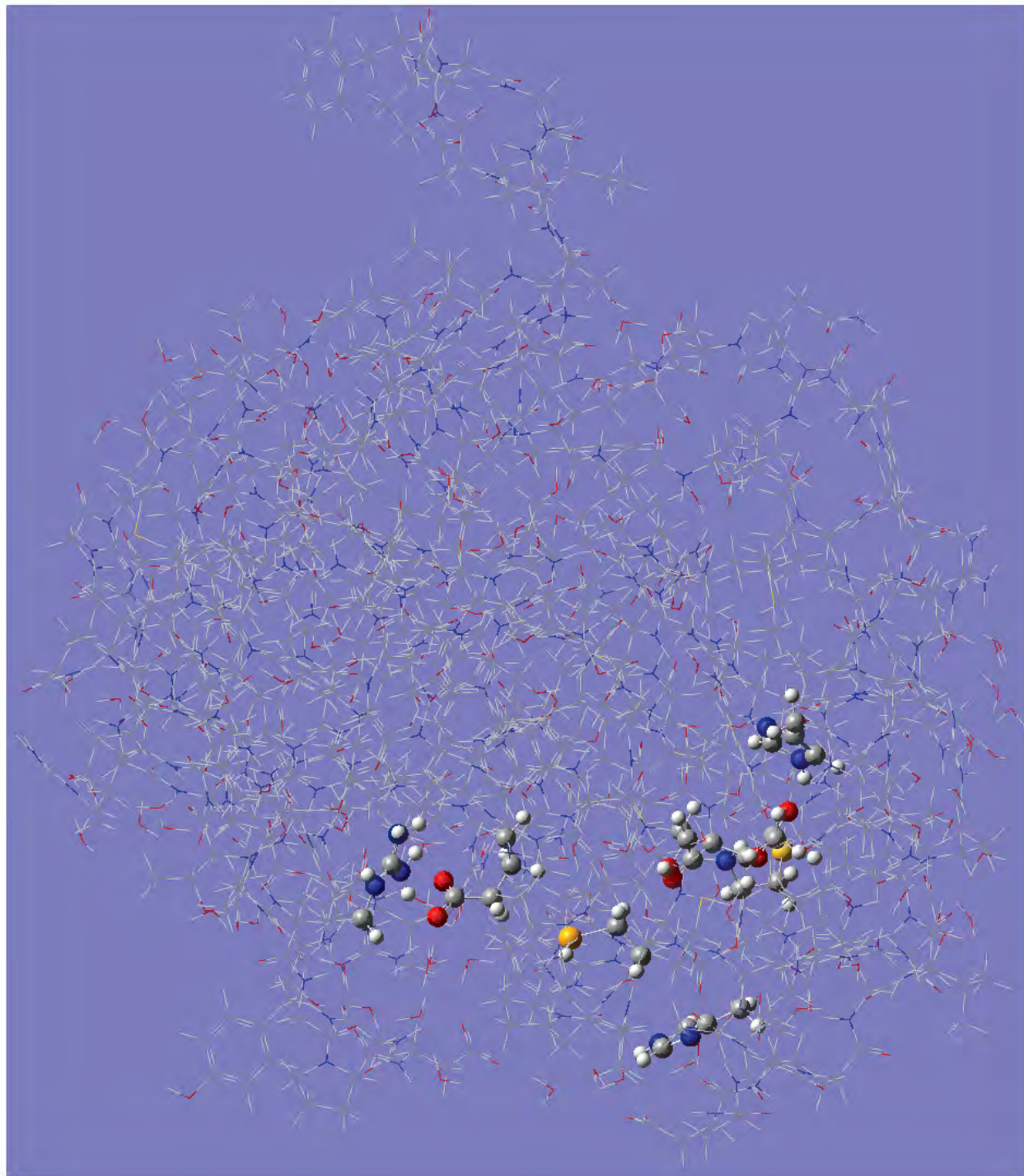
The PDB file 4TR3 provided the crystal structure of Dio3. However, the critical selenocysteine residue in the active site was replaced with a typical cysteine residue by

Schweizer et al. for expression and purification purposes. To obtain an accurate representation of the true Dio3 enzyme, I replaced Cys170 with selenocysteine. Since I modified the chemical identity of the PDB file, I needed to optimize the structure of the whole protein to get a reliable representation of the protein with the selenocysteine residue in place. The optimization was done using the ONIOM functionality from Gaussian 09 (see methods for more details on ONIOM). After roughly two months of trial and error, the ONIOM job finished with a normal termination, meaning that the program converged on a minimum energy structure for the protein containing the selenocysteine residue. Optimized Dio3 has very similar secondary and tertiary structure compared to the crystal structure provided by PDB file 4TR3. Both structures contain the core  $\alpha$ -helices and  $\beta$ -sheets in roughly the same position (Figure 10).



**Figure 10. Secondary and tertiary structure of 4TR3 and optimized Dio3.** 4TR3 is shown on the left and optimized Dio3 is shown on the right. The selenium and its hydrogen from the selenocysteine residue are shown in the optimized Dio3 picture as the orange and grey spheres. Visualization of the two proteins using Pymol's cartoon format shows very similar secondary and tertiary structure. In both structures, the five  $\beta$ -sheets in the middle of the protein are sandwiched by three longer  $\alpha$ -helices on the bottom of the protein and two shorter  $\alpha$ -helices towards the top.

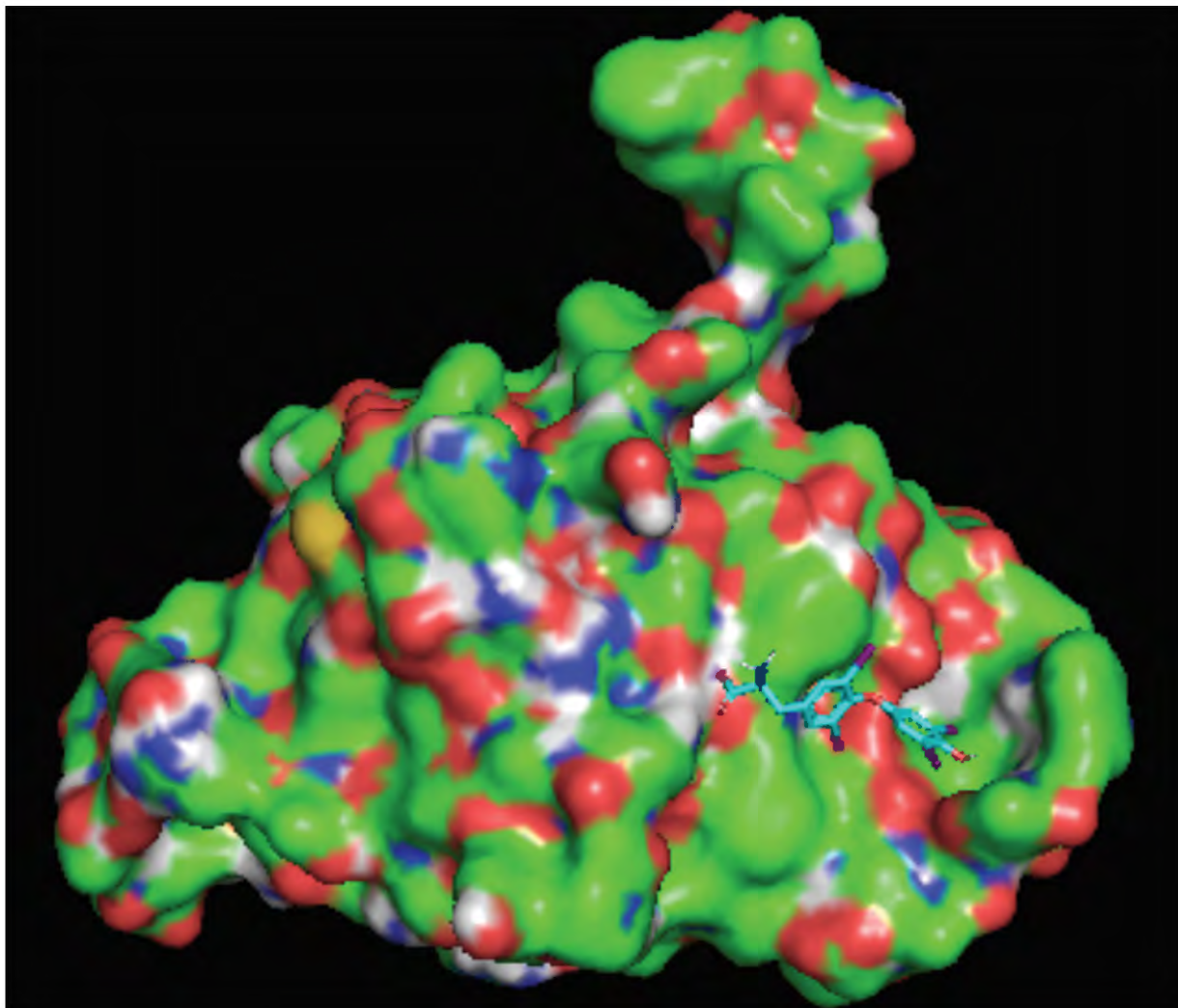




**Figure 11. Optimized structure of Dio3 with selenocysteine.** This visualization of the optimized protein, provided by GaussView, shows the active site amino acids residues in Ball and Stick format and the rest of the protein in Wire format. The selenium atom of the selenocysteine residue is shown in orange.

### Docking Study with Thyroxine in Complex with Updated Dio3

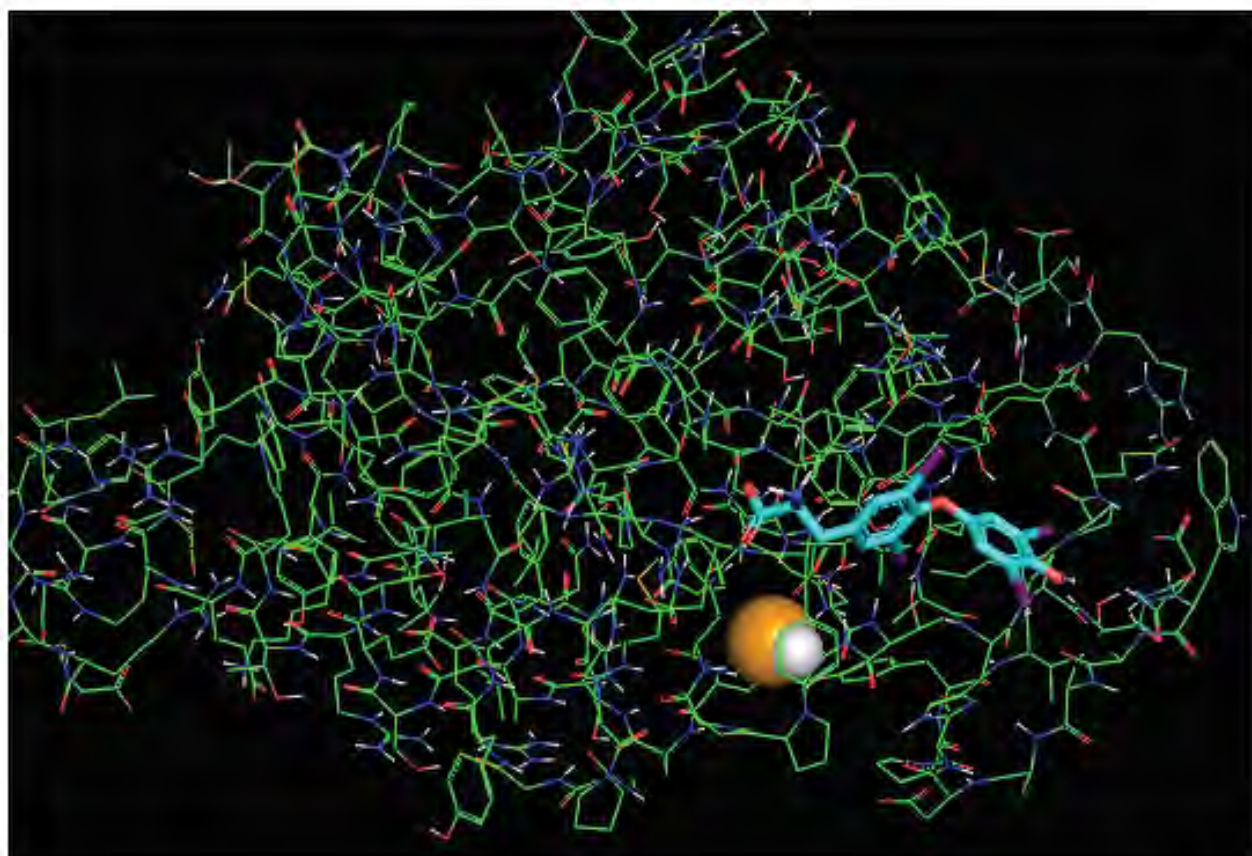
The completion of the Dio3 optimization job allowed for the analysis of the position of thyroxine in complex with Dio3. Using a program called Autodock Vina, I was able to find the lowest energy binding conformation of thyroxine on Dio3. Autodock Vina placed thyroxine in the active site binding cleft (Figure 12).



**Figure 12. Thyroxine binding in active site pocket of Dio3.** Optimized Dio3 is shown in the Molecular Surface format and thyroxine is shown in Stick format. Thyroxine is shown binding in the active site cleft of optimized Dio3 on the right side of the protein.

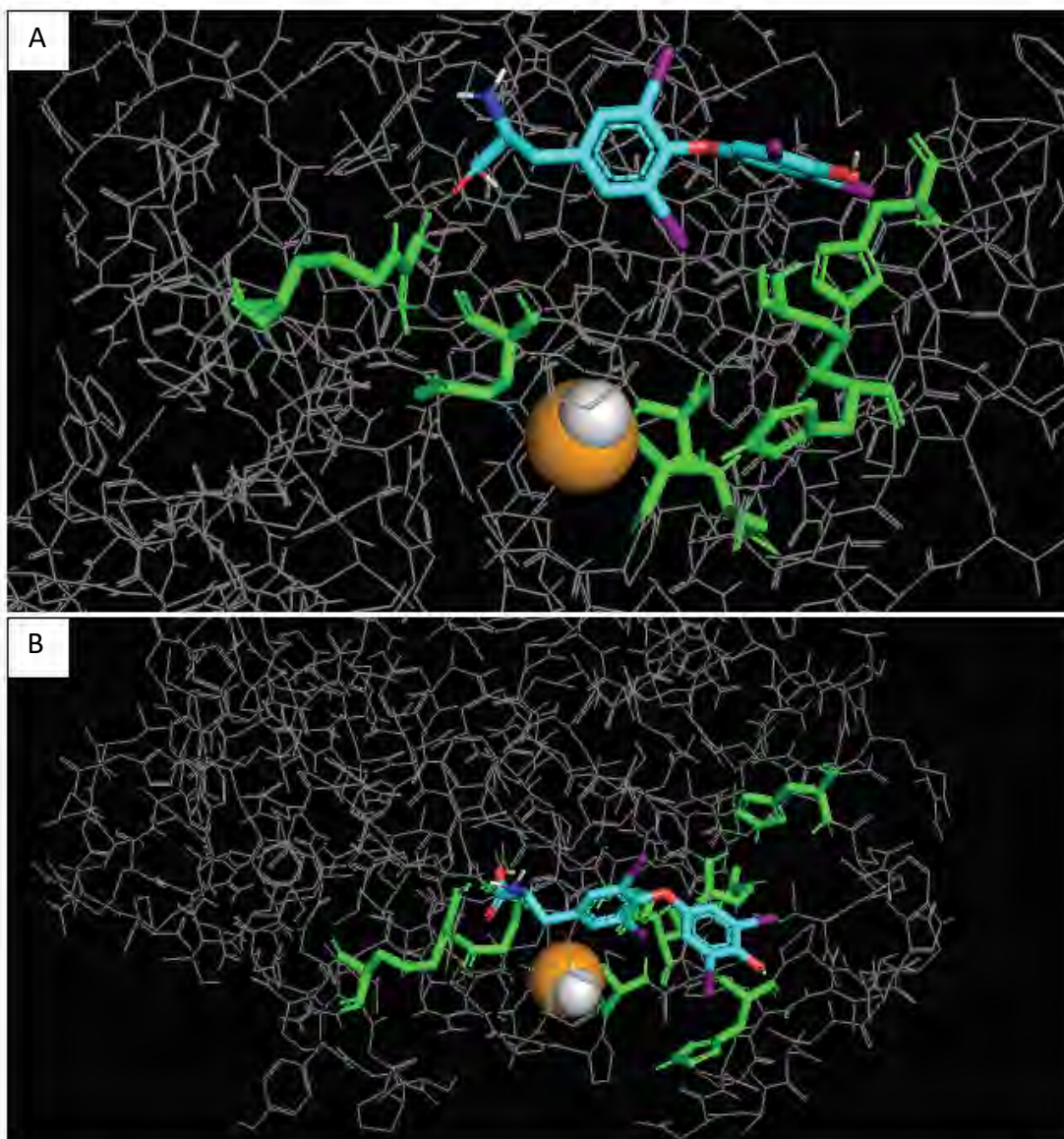


AutoDock Vina decided that pointing the target 5-iodine towards the interior of the active site was the lowest energy binding conformation (Figure 13). This is of great significance because it confirmed that the ONIOM optimization of the protein was accurate, since Dio3 strictly performs inner ring deiodination at the 5-iodine position. If the location and conformation of thyroxine was somewhere other than the active site, or not pointing the 5-iodine in the correct position to undergo deiodination, then the reimplementation of the selenocysteine and subsequent ONIOM optimization would be invalid.

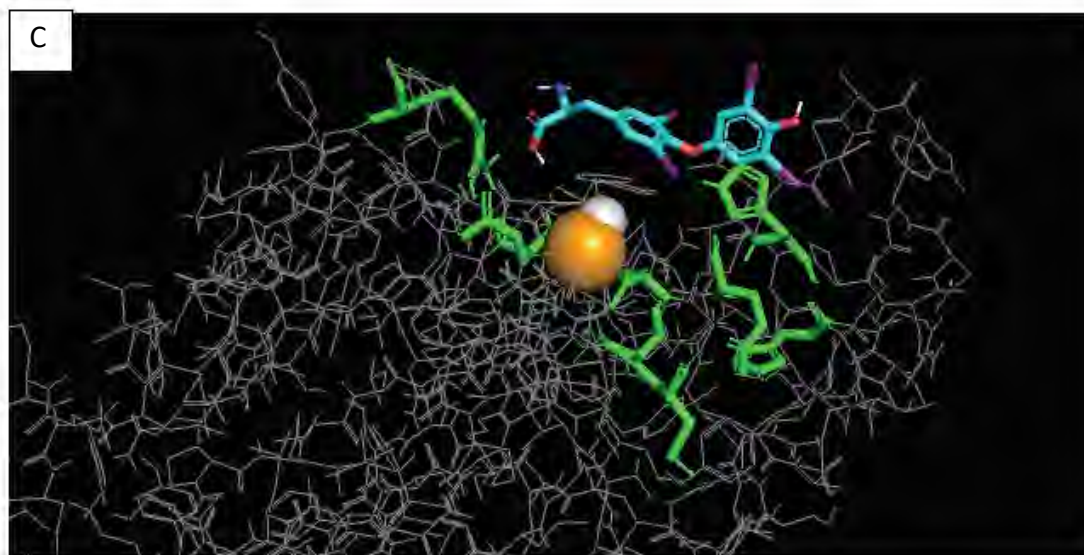


**Figure 13. Thyroxine conformation points 5-iodine to interior of enzyme.** Using Pymol to visualize, thyroxine, shown in Stick format, is oriented with the purple 5-iodine towards the interior of the enzyme, which is shown in Wire format. The selenium atom is displayed as the large orange sphere for reference.

Zooming in on the target 5-iodine within the active site, it is apparent that the iodine is in close relation to the selenium atom. However, for a halogen bonding interaction to exist, there needs to be a  $180^\circ$  angle between the 5-iodine and the selenium. The 5-iodine is not directed at a  $180^\circ$  angle to the selenium and, therefore, I cannot conclude that there is a halogen bonding interaction occurring in this scenario (Figure 14).



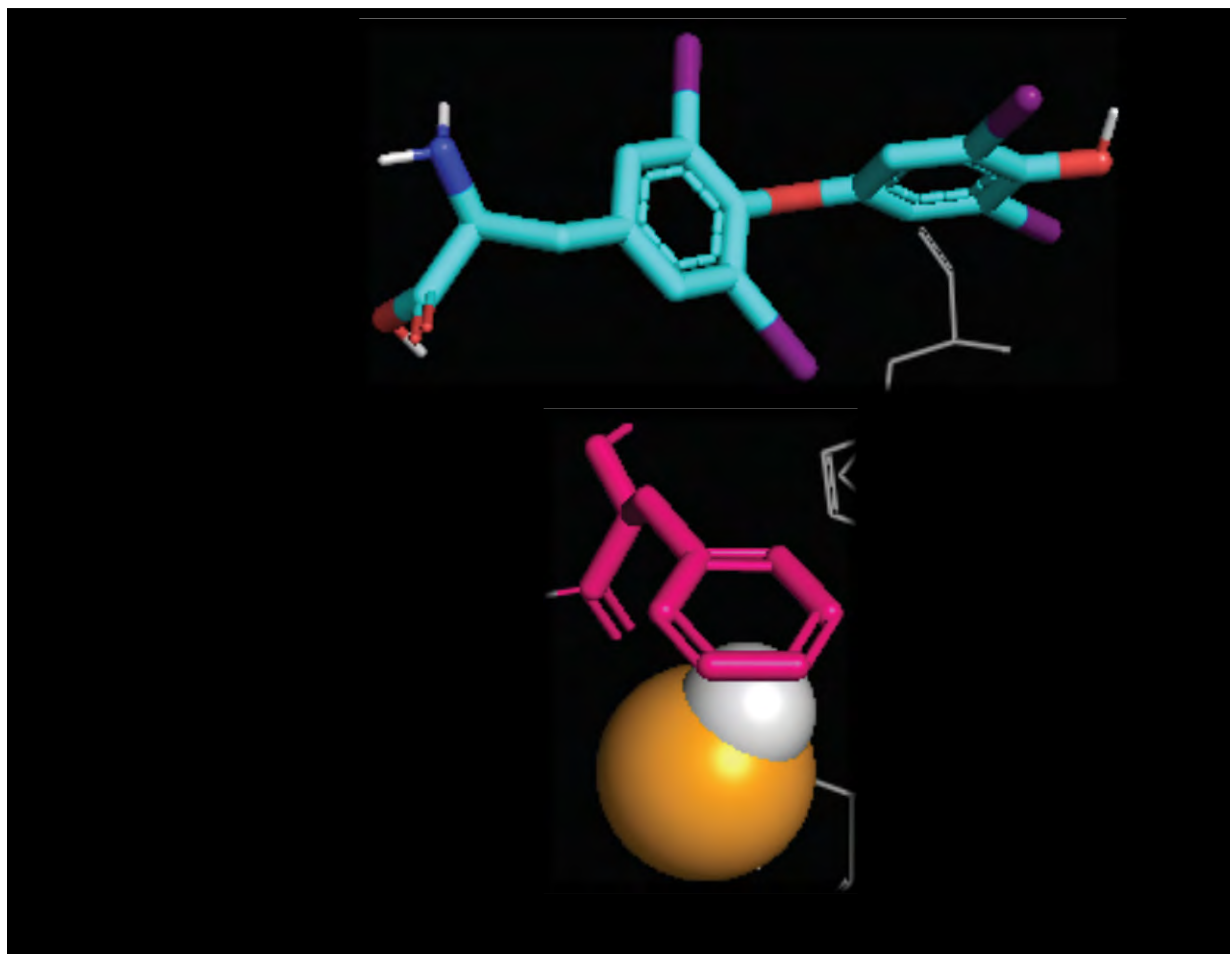




**Figure 14 A, B, and C. Positioning of 5-iodine within Dio3 active site.** Active site amino acids are highlighted in green. Thyroxine is colored accordingly: carbon = cyan; iodine = purple; oxygen = red; nitrogen = blue. The selenium atom is shown as the large orange sphere. (A) Face-on view of thyroxine complexed with Dio3. The 5-iodine is on the bottom of thyroxine and pointed towards the middle of the active site. (B) Bird's-eye view of thyroxine complexed with Dio3. The 5-iodine is shown pointing into active site pocket. (C) Worm's-eye view of thyroxine complexed with Dio3. This view serves as the best perspective for visualizing the active site cleft and position of 5-iodine in relation to the selenium. The 5-iodine is close in proximity to the selenium but not pointing directly at the atom.

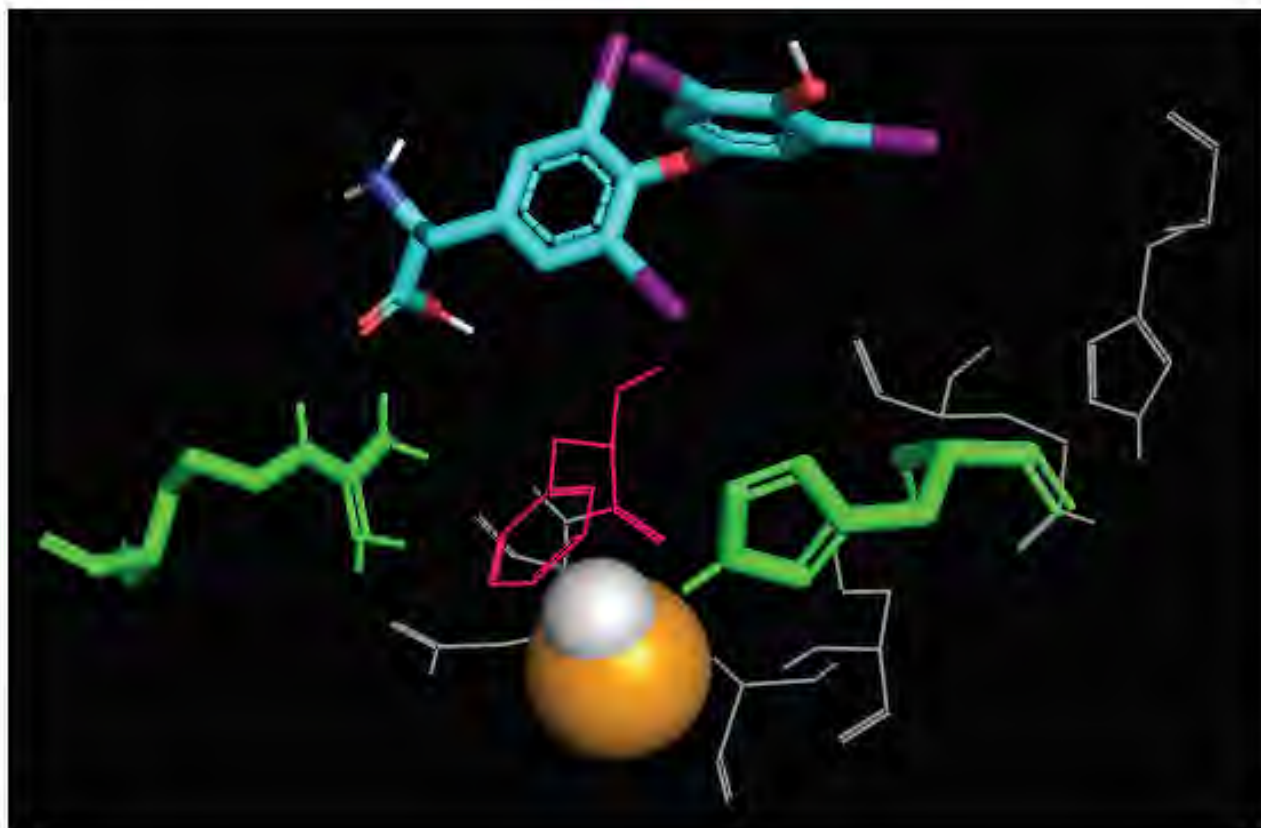
### Dimerization and Removal of Phe258 Shield

While scanning the literature, I learned that Dio3 requires dimerization to become activated. The crystal structure from PDB file 4TR3 represents the inactive monomeric form. According to Schweizer et al., Phe258 acts as a shield that blocks any reactivity with the selenium (Figure 15). Upon dimerization, the  $\alpha 2/\beta 3$ -loop, which contains Phe258, relaxes, and removes the Phe258 shield from the active site.



**Figure 15. Phe258 shield in active site.** The Phe258 residue, shown in pink, blocks the selenium atom, shown as the orange sphere, from the 5-iodine atom, shown in purple and protruding into the active site. The other active site amino acids are shown in grey.

Removal of the Phe258 shield would open the possibility for a  $180^\circ$  angle between the selenium and the 5-iodine. Additionally, Schweizer et al. alluded to the probability of His202 binding with the 4'-hydroxyl of thyroxine. With Arg275 already shown to be bound to the carboxyl end of thyroxine, the removal of Phe258 from the active site by dimerization would allow the 4'-hydroxyl of thyroxine to swing down and bind with His202 (Figure 16). Visualizing in three dimensions, if the Phe258 shield was removed, and the carboxyl end and 4-hydroxyl of thyroxine were bound to Arg275 and His202, respectively, then the 5-iodine and the selenium would be interacting at a  $180^\circ$  angle.



**Figure 16.** The potential for His202 to clamp thyroxine in the active site would put the 5-iodine interacting with the selenium at a  $180^\circ$  angle. The carboxyl end of thyroxine was already observed to bind to Arg275 from the docking study. If Phe258 was removed from the active site, then the 4'-hydroxyl could bind to His202 and lock a roughly  $180^\circ$  angle between the 5-iodine and selenium.

## Conclusion and Future Work

My data indicate that a halogen bonding interaction is unlikely to occur between the 5-iodine and the selenium in the inactive monomeric state. Additionally, while I do suspect that a halogen bonding interaction can occur in the active dimeric state, more computational analysis of the dimeric state will be necessary to address this hypothesis. Future work on this project should examine the conformation of thyroxine in the active site of Dio3 in the active dimeric state. Approaching this question could be done in a couple of ways. The Phe258 residue could simply be replaced by a less bulky amino acid, such as alanine, mimicking the conformational change

that occurs with dimerization. This would remove the phenol group from the side chain, which acts as the shield blocking access to selenium. The new Dio3 protein containing the F258A mutation could be optimized using ONIOM, and a subsequent docking study could be run to analyze the position of the 5-iodine in relation to the unguarded selenium. Alternatively, a completely different approach could be utilized, such as Molecular Dynamics, to simulate the true conformational changes observed with dimerization of Dio3. This approach would allow for a more complete picture of the structure of Dio3 as a dimer rather than just altering one active site amino acid residue to mimic the dimerization process.

## Acknowledgements

I would like to thank Dr. Donald for his willingness to take a biology kid into his chemistry research group. He took a chance on me as I did not have any prior experience with computational chemistry nor research in general coming into my senior year. The personal growth that I have achieved during this year is all devoted to his commitment to learning. He is truly passionate for teaching, and I am extremely grateful for the opportunity he provided me. I would also like to acknowledge our postbac, Ulrick Gaillard, and the rest of the Donald group for giving me advice on how to best perform computational chemistry jobs. Additionally, I would like to thank Dr. Dattelbaum for serving as my second reader on this paper. Lastly, a special thanks to Dr. Fernando Clemente of Gaussian Help for assisting me with the ONIOM optimization of Dio3.

## References

---

- <sup>1</sup> Parker, A. J., Stewart, J., Donald, K. J., & Parish, C. A. (2012). Halogen bonding in DNA base pairs. *Journal of the American Chemical Society*, *134*(11), 5165-5172.
- <sup>2</sup> Guthrie, F. (1863). XXVIII.—On the iodide of iodammonium. *Journal of the Chemical Society*, *16*, 239-244.
- <sup>3</sup> Mulliken, R. S. (1952). Molecular compounds and their spectra. II. *Journal of the American Chemical Society*, *74*(3), 811-824.
- <sup>4</sup> Strieter, F. J., & Templeton, D. H. (1962). Crystal Structure of the Carbon Tetrabromide p-Xylene Complex. *The Journal of Chemical Physics*, *37*(1), 161-164.
- <sup>5</sup> Hassel, O. (1970). Structural aspects of interatomic charge-transfer bonding. *Science*, *170*(3957), 497-502.
- <sup>6</sup> Bent, H. A. (1968). Structural chemistry of donor-acceptor interactions. *Chemical Reviews*, *68*(5), 587-648.
- <sup>7</sup> Bruggermann, K., & Kochi, J. K. (1992). Charge-transfer activation of aromatic hydrocarbons with stannic chloride. *The Journal of Organic Chemistry*, *57*(10), 2956-2960
- <sup>8</sup> Alkorta, I., Rozas, I., & Elguero, J. (1998). Charge-transfer complexes between dihalogen compounds and electron donors. *The Journal of Physical Chemistry A*, *102*(46), 9278-9285.
- <sup>9</sup> Clark, T., Hennemann, M., Murray, J. S., & Politzer, P. (2007). Halogen bonding: the  $\sigma$ -hole. *Journal of molecular modeling*, *13*(2), 291-296.
- <sup>10</sup> Robertson, C. C., Wright, J. S., Carrington, E. J., Perutz, R. N., Hunter, C. A., & Brammer, L. (2017). Hydrogen bonding vs. halogen bonding: the solvent decides. *Chemical science*, *8*(8), 5392-5398.
- <sup>11</sup> Brinck, T., Borrfors, A.N. Electrostatics and polarization determine the strength of the halogen bond: a red card for charge transfer. *J Mol Model* *25*, 125 (2019).
- <sup>12</sup> Desiraju, G., Ho, P., Kloo, L., Legon, A., Marquardt, R., Metrangolo, P., Politzer, P., Resnati, G. & Rissanen, K. (2013). Definition of the halogen bond (IUPAC Recommendations 2013). *Pure and Applied Chemistry*, *85*(8), 1711-1713.
- <sup>13</sup> Gabriella Cavallo, Pierangelo Metrangolo, Roberto Milani, Tullio Pilati, Arri Priimagi, Giuseppe Resnati, and Giancarlo Terraneo. *Chemical Reviews* **2016** *116* (4), 2478-2601

- 
- <sup>14</sup> Auffinger, P., Hays, F. A., Westhof, E., & Ho, P. S. (2004). Halogen bonds in biological molecules. *Proceedings of the National Academy of Sciences*, *101*(48), 16789-16794.
- <sup>15</sup> Wilcken, R., Liu, X., Zimmermann, M. O., Rutherford, T. J., Fersht, A. R., Joerger, A. C., & Boeckler, F. M. (2012). Halogen-enriched fragment libraries as leads for drug rescue of mutant p53. *Journal of the American Chemical Society*, *134*(15), 6810-6818.
- <sup>16</sup> Liao, J. J. L. (2007). Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *Journal of medicinal chemistry*, *50*(3), 409-424.
- <sup>17</sup> Armstrong, M., Asuka, E., & Fingeret, A. (2020). Physiology, thyroid function. *StatPearls [Internet]*.
- <sup>18</sup> Bayse, C. A., & Rafferty, E. R. (2010). Is halogen bonding the basis for iodothyronine deiodinase activity?. *Inorganic chemistry*, *49*(12), 5365-5367.
- <sup>19</sup> Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.
- <sup>20</sup> K.A. Peterson, B.C. Shepler, D. Figgen, H. Stoll, *J. Phys. Chem. A* *110*, 13877 (2006).
- <sup>21</sup> Tian Lu, Feiwu Chen, *J. Comput. Chem.*, *33*, 580-592 (2012).
- <sup>22</sup> Dennington, R.D., Keith, T.A. and Millam, J.M. (2008) GaussView 5.0.9, Gaussian.
- <sup>23</sup> Schweizer, U., Schlicker, C., Braun, D., Köhrle, J., & Steegborn, C. (2014). Crystal structure of mammalian selenocysteine-dependent iodothyronine deiodinase suggests a peroxiredoxin-like catalytic mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(29), 10526–10531. <https://doi.org/10.1073/pnas.1323873111>
- <sup>24</sup> Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, *31*(2), 455-461.
- <sup>25</sup> The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

---

<sup>26</sup> Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, D. C., Funk, K., Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S., Lu, Z., Madden, T. L., O'Leary, N., Phan, L., Rangwala, S. H., Schneider, V. A., Skripchenko, Y., ... Sherry, S. T. (2021). Database resources of the National Center for Biotechnology Information. *Nucleic acids research*, *49*(D1), D10–D17. <https://doi.org/10.1093/nar/gkaa892>

<sup>27</sup> Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)