Supporting Information (SI)

Structural studies of thyroid peroxidase show the monomer interacting with autoantibodies in thyroid autoimmune disease

Daniel E. Williams, Sarah N. Le, David E. Hoke, Peter G. Chandler, Monika Gora, Marlena Godlewska, J. Paul Banga and Ashley M. Buckle
**Figure S1 - Purification of the TPO construct ΔproTPOe-GCN4** (A) A chromatogram from a Superdex S200 16/60 column, showing ΔproTPOe-GCN4 eluting as a single major peak at 71.7mL, consistent with a 110 kDa protein. No other major large species appears to be present. (B) Reducing SDS-PAGE analysis of purified ΔproTPOe-GCN4 shows a major band at ~110 kDa.
Figure S2 – Mass spectrometry analysis of ΔproTPOe-GCN4. Sequence coverage was reported as 70% with a protein score of 19294, making ΔproTPOe-GCN4 the most abundant species in the sample. Full length TPOe-GCN4 is in black lettering, with detected peptides highlighted in red. Note that residues 1 through 108 comprise the signal peptide and propeptide that are not incorporated into full length ΔproTPOe-GCN4, though are included here to demonstrate their successful non-inclusion in our construct.
Figure S3 – Mass spectrometry analysis of suspected degraded ΔproTPOe-GCN4 fragment. Sequence coverage was reported as 63% with a protein score of 8710, making a degraded form of ΔproTPOe-GCN4 the most abundant species in the sample. Full length ΔproTPOe-GCN4 is in black lettering, with detected peptides highlighted in red. Note that residues 1 through 108 comprise the signal peptide and propeptide that are not incorporated into full length ΔproTPOe-GCN4, though are included here to demonstrate their successful non-inclusion in our construct.
Figure S4 – Bio-layer interferometry (BLI) sensorgram data of ΔproTPOe-8His binding to Fab.

Sensorgram curves according to a TR1.9 Fab concentration range of between 0 and 500 nM. ΔproTPOe-8His is immobilised on the biosensor surface. The data has been normalised against a blank run of buffer (1x PBS, pH 7.4). Vertical line at 450 s represents the end of the association phase. Dotted lines in black represent the fit calculated using a 1:1 binding model with global fitting within the BLItz Pro software. $R^2$ values for the calculated fit were reported as 0.97. $K_D$ was calculated as 20 nM.
Figure S5 – Snapshots from the trans ΔproTPOe MD trajectory show TPO changing conformation from extended to more compact structure. Snapshots from the trans ΔproTPOe MD trajectory as presented in Figure 6. (A) Representation of the starting model from Le and co-workers 1, as well as trans ΔproTPOe after 200, 300 and 400 ns of simulation. (B) Structural superpositions of the above snapshots with the starting trans ΔproTPOe model in red. Orange indicates trans ΔproTPOe after 100 ns of simulation, yellow after 200 ns, green after 300 ns and blue after 400 ns.
**Figure S6** — IDRs in context of the MD simulations with starting structures. (A) TPO models (simulation starting structures) with IDRs highlighted. (B) Representative structures taken from the MD simulations for each of the trans, cis and extended forms of the ΔproTPOe monomer, as in Figure 7. IDR-A residues are highlighted by red spheres, and IDR-B residues by blue spheres. The MPO-like
domain, CCP-like domain and EGF-like domain are coloured in forest green, light teal and marine blue respectively (as in Figure 1).

**Table S1** – Published residues involved in IDRs of TPO

<table>
<thead>
<tr>
<th>Antibody Involved</th>
<th>Number of Reported Epitopes</th>
<th>Epitopes</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IDR-A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T13</td>
<td>4</td>
<td>H353-Y363, P377-R386, K713-S720, Y766-Q775</td>
<td>2-6</td>
</tr>
<tr>
<td>ICA1</td>
<td>1</td>
<td>H353-Y363</td>
<td>2,3</td>
</tr>
<tr>
<td>TR1.9</td>
<td>2</td>
<td>K713, K713-S720</td>
<td>2,4,7</td>
</tr>
<tr>
<td>126TO10</td>
<td>3</td>
<td>R225, R646, D707</td>
<td>8,9</td>
</tr>
<tr>
<td>126TP1</td>
<td>3</td>
<td>R225, R646, D707</td>
<td>8,9</td>
</tr>
<tr>
<td>126TP7</td>
<td>1</td>
<td>R225</td>
<td>9</td>
</tr>
<tr>
<td><strong>IDR-B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>126TP5</td>
<td>5</td>
<td>D620, D624, K627, D630, F597-E604</td>
<td>8-10</td>
</tr>
<tr>
<td>126TP14</td>
<td>5</td>
<td>D620, D624, K627, D630, F597-E604</td>
<td>8-10</td>
</tr>
<tr>
<td>131TP7</td>
<td>1</td>
<td>K627</td>
<td>9</td>
</tr>
<tr>
<td>SP1.4</td>
<td>1</td>
<td>F597-E604</td>
<td>10</td>
</tr>
<tr>
<td>TR1.8</td>
<td>1</td>
<td>T611-V618</td>
<td>10</td>
</tr>
<tr>
<td>WR1.7</td>
<td>1</td>
<td>F597-E604</td>
<td>10</td>
</tr>
</tbody>
</table>

The epitopes that have been identified as making up the immunodominant regions (IDRs) of TPO, named IDR-A and IDR-B.
### Table S2 – Melting point data of ΔproTPOe-8His in different buffer conditions

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH</th>
<th>$T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 mM HEPES, 250 mM NaCl</td>
<td>8.0</td>
<td>52.3</td>
</tr>
<tr>
<td>50 mM HEPES, 250 mM NaCl</td>
<td>7.0</td>
<td>55.2</td>
</tr>
<tr>
<td>50 mM Sodium Phosphate, 250 mM NaCl</td>
<td>6.0</td>
<td>54.7</td>
</tr>
<tr>
<td>50 mM Sodium Acetate, 250 mM NaCl</td>
<td>5.5</td>
<td>53.9</td>
</tr>
<tr>
<td>50 mM Glycine, 250 mM NaCl</td>
<td>4.0</td>
<td>53.6</td>
</tr>
</tbody>
</table>
Table S3 – Theoretical and calculated Stokes Radii ($R_s$) of TPO

<table>
<thead>
<tr>
<th>Reference Dataset</th>
<th>Equation</th>
<th>Stokes Radius (Å)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Monomer</td>
<td>Dimer</td>
<td></td>
</tr>
<tr>
<td>Globular Folded Proteins</td>
<td>$R_s = (4.75)N^{0.29}$</td>
<td>32.52</td>
<td>39.75</td>
<td></td>
</tr>
<tr>
<td>Denatured Unfolded Proteins</td>
<td>$R_s = (2.21)N^{0.57}$</td>
<td>96.93</td>
<td>143.89</td>
<td></td>
</tr>
<tr>
<td>Analytical SEC of TPO with no TM domain</td>
<td></td>
<td>51.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUC of ΔproTPOe-GCN4</td>
<td></td>
<td>75.7</td>
<td>91.4</td>
<td></td>
</tr>
<tr>
<td>AUC of ΔproTPOe-8His</td>
<td></td>
<td>64.9</td>
<td>77.1</td>
<td></td>
</tr>
<tr>
<td>AUC of ΔproTPOe-GCN4 with TR1.9</td>
<td></td>
<td>52.7</td>
<td>66.4</td>
<td></td>
</tr>
<tr>
<td>AUC of ΔproTPOe-8His with TR1.9</td>
<td></td>
<td>57.4</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

AUC, analytical ultracentrifugation; SEC, size exclusion chromatography; TM, transmembrane domain.
## Table S4 – Model Fit Percentages

<table>
<thead>
<tr>
<th>Model</th>
<th>Percentage of Molecules within the EM Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trans Monomer</td>
<td>73</td>
</tr>
<tr>
<td>Cis Monomer</td>
<td>72</td>
</tr>
<tr>
<td>Trans Dimer</td>
<td>54</td>
</tr>
<tr>
<td>Cis Dimer</td>
<td>59</td>
</tr>
<tr>
<td>Trans Monomer with Fab</td>
<td>53</td>
</tr>
<tr>
<td>Cis Monomer with Fab</td>
<td>55</td>
</tr>
<tr>
<td>Curled Monomer with Fab</td>
<td>58</td>
</tr>
<tr>
<td>Curled Monomer with scFv format of TR1.9</td>
<td>73</td>
</tr>
<tr>
<td>Trans Monomer with Fab sequentially fit*</td>
<td>68</td>
</tr>
<tr>
<td>Cis Monomer with Fab sequentially fit*</td>
<td>70</td>
</tr>
<tr>
<td>Trans Dimer with Fab</td>
<td>33</td>
</tr>
<tr>
<td>Cis Dimer with Fab</td>
<td>37</td>
</tr>
</tbody>
</table>

Fit percentages of various TPO models within the electron microscopy (EM) map. Asterisks (*) indicates configurations where TR 1.9 Fab was fitted into the available space in the envelope without regard to its epitope’s location, rather than in a realistic orientation in which the complementarity determining regions (CDR) face the published epitope of K713-S720.
References


