

## RESEARCH ARTICLE

# Informatics investigations into anti-thyroid drug induced agranulocytosis associated with multiple HLA-B alleles

Kerry A. Ramsbottom<sup>1</sup>, Daniel F. Carr<sup>2</sup>, Daniel J. Rigden<sup>1</sup>, Andrew R. Jones<sup>1\*</sup><sup>1</sup> Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom, <sup>2</sup> Institute of Translational Medicine, University of Liverpool, Liverpool, United Kingdom\* [Andrew.Jones@liverpool.ac.uk](mailto:Andrew.Jones@liverpool.ac.uk)

## Abstract

### Introduction

Adverse drug reactions have been linked with HLA alleles in different studies. These HLA proteins play an essential role in the adaptive immune response for the presentation of self and non-self peptides. Anti-thyroid drugs methimazole and propylthiouracil have been associated with drug induced agranulocytosis (severe lower white blood cell count) in patients with B\*27:05, B\*38:02 and DRB1\*08:03 alleles in different populations: Taiwanese, Vietnamese, Han Chinese and Caucasian.

### Methods

In this study, informatics methods were used to investigate if any sequence or structural similarities exist between the two associated HLA-B alleles, compared with a set of “control” alleles assumed not be associated, which could help explain the molecular basis of the adverse drug reaction. We demonstrated using MHC Motif Viewer and MHCcluster that the two alleles do not have a propensity to bind similar peptides, and thus at a gross level the structure of the antigen presentation region of the two alleles are not similar. We also performed multiple sequence alignment to identify polymorphisms shared by the risk but not by the control alleles and molecular docking to compare the predicted binding poses of the drug-allele combinations.

### Results

Two residues, Cys67 and Thr80, were identified from the multiple sequence alignments to be unique to these risk alleles alone. The molecular docking showed the poses of the risk alleles to favour the F-pocket of the peptide binding groove, close to the Thr80 residue, with the control alleles generally favouring a different pocket. The data are thus suggestive that Thr80 may be a critical residue in HLA-mediated anti-thyroid drug induced agranulocytosis, and thus can guide future research and risk assessment.

## OPEN ACCESS

**Citation:** Ramsbottom KA, Carr DF, Rigden DJ, Jones AR (2020) Informatics investigations into anti-thyroid drug induced agranulocytosis associated with multiple HLA-B alleles. PLoS ONE 15(2): e0220754. <https://doi.org/10.1371/journal.pone.0220754>

**Editor:** Natalia Sizochenko, Dartmouth College, UNITED STATES

**Received:** July 12, 2019

**Accepted:** January 22, 2020

**Published:** February 6, 2020

**Copyright:** © 2020 Ramsbottom et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** The research was part-funded by the Medical Research Council grant for the Centre for Drug Safety Science, University of Liverpool (Grant Number: MR/L006758/1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. There was no additional external funding received for this study.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Adverse drug reactions have been linked to Human Leukocyte Antigens (HLA) in multiple different studies, where an individual carrying a specific risk allele has a higher risk of developing a reaction to that drug, including skin conditions like Stevens-Johnson syndrome and toxic epidermal necrolysis (SJS/TEN) and drug induced liver injury [1–3]. HLA proteins play a role in the adaptive immune response, presenting peptides to the T-cell receptors. Non-self peptides are then recognised and elicit an immune response where appropriate [4]. Occasionally, unnatural interaction of drugs during this process results in an adverse drug reaction. The role of HLA in these adverse drug reactions has been hypothesised in three main ways: the Hapten model, the Pharmacological Interaction model and the Altered Peptide Repertoire model. The *Hapten* model predicts the drug binds covalently to a self-protein and is processed via HLA molecules; this drug-protein combination is presented and recognised as being non-self, initiating an immune response [5]. The *Pharmacological Interaction (PI)* model predicts the drugs bind directly to the TCR or via the formation of HLA-drug complexes which activate T cells and thus initiate an immune response without the need for a specific peptide ligand [6]. Except where noted below, the results presented here work under the assumption that the drugs we are investigating here will follow the *Altered Peptide Repertoire* model, with the drug interacting non-covalently with the HLA molecule directly within the antigen presentation site. This leads to a difference in the self-peptide set that is presented to the T-cell receptors and thus, initiating an immune response [5]. For more details on the different mechanisms, and schematic diagrams, see Yun *et al.* (2016). Currently, the most widely investigated HLA-ADR association is that of abacavir with B\*57:01. For this association, the crystal structure of the drug bound in complex with the risk allele is available [7, 8]. Illing *et al.* [7] and Ostrov *et al.* [8] have demonstrated the Altered Peptide Repertoire model with high confidence for abacavir, including the crystal structure for abacavir bound in the peptide binding groove of the associated risk allele, B\*57:01, along with proteomics evidence for different peptides being presented in the bound and unbound cases [7].

The peptide binding groove of the HLA is a long hydrophobic cleft formed between the  $\alpha$ -helices and  $\beta$ -sheet platform. This cleft is much larger than the naturally involved binding sites that proteins have for small organic molecules. The peptide binding groove contains six subsites (Figure D in S1 File). The size and stereochemistry of the subsites are determined by the polymorphic residues along the cleft [9]. The specificity of peptide binding is determined, in part, by the interactions between anchor residues on the peptide side chains at two or more of these subsites [10].

Anti-thyroid drugs are used to treat hyperthyroidism as they normalise thyroid function through binding to the thyroid peroxidase enzyme [11]. These drugs are thioamides containing a thiocarbonyl group and a thiourea moiety within a heterocyclic structure. The common agents used are methimazole, carbimazole and propylthiouracil [12]. Agranulocytosis has been defined as absolute neutrophil count below 500/ $\mu$ l of blood and includes not only neutrophil count but also the absolute number of eosinophils, basophils and mast cells [13]. Patients with severe neutropenia are likely to experience infections which may be life-threatening or even fatal. The mechanism of anti-thyroid induced agranulocytosis is through either direct toxicity or immune-mediated toxicity [13]. The incidence of agranulocytosis in England and Wales has been estimated at around 7 cases per million people per year, with an adjusted odds ratio for neutropenia of 34.7 for users of thyroid inhibitors (i.e. anti-thyroid drugs) [14]. Most cases of agranulocytosis are idiosyncratic reactions to drugs or their metabolites, including anti-thyroid drugs. Other causes include splenic sequestration, nutritional deficiencies, infections, immune neutropenia, haematological disease and primary congenital or chronic neutropenia

Table 1. HLA associations seen carbimazole, methimazole and propylthiouracil across populations.

Study	Allele(s) Associated	Drug(s) Associated	Cohort Ethnicity	Cohort size		Odds Ratio	P-value
				Risk	Control		
He <i>et al.</i> [16]	B*27:05 B*38:02 DRB1*08:03	methimazole propylthiouracil	Han Chinese	27	135	66.24 7.525 4.316	9.24x10 <sup>-5</sup> 8.68x10 <sup>-4</sup> 2.8x10 <sup>-3</sup>
Chen <i>et al.</i> [17]	B*38:02 DRB1*08:03	methimazole carbimazole propylthiouracil	Taiwanese	42	1208	21.48 6.13	6.28x10 <sup>-18</sup> 1.35x10 <sup>-8</sup>
Thao <i>et al.</i> [18]	B*38:02	methimazole propylthiouracil	Vietnamese	21	81	28.6	5.2x10 <sup>-7</sup>
Hallberg <i>et al.</i> [19]	B*27:05	methimazole carbimazole propylthiouracil	Caucasian	234	5170	3.24	1.20x10 <sup>-11</sup>

<https://doi.org/10.1371/journal.pone.0220754.t001>

[15]. An increased risk of agranulocytosis has been associated with anti-thyroid drugs carbimazole, methimazole and propylthiouracil for three different HLA alleles in Asian and Caucasian populations, these associations are shown in Table 1.

The structures of methimazole and propylthiouracil are shown in Figure F in S1 File. It can be seen that these two associated drugs share a common thiocarbonyl group, which is also seen in the experimental anti-thyroid drugs. The associated anti-thyroid drugs share a common target, thyroid peroxidase, in their normal mechanism of action [20]. A study by Pradhan *et al.* conducted molecular docking of methimazole and propylthiouracil with thyroid peroxidase. The results of this study predicted both drugs to bind in the same position, with the sulphur group forming a hydrogen bond with Arg491 of the thyroid peroxidase, resulting in inhibition of thyroid hormone production [11]. As these drugs show similar binding to the same target, it is reasonable to hypothesise that both drugs might also have a shared mechanism for the adverse drug reaction, going some way to explaining why multiple drugs have been associated with the same alleles. It can be noted that methimazole is more commonly associated than propylthiouracil, although neither drug has been studied in isolation [16, 17, 19]. In the He *et al.* study, there were 26 methimazole agranulocytosis cases compared to 3 for propylthiouracil. For the Chen *et al.* study, of the 42 thioamide-induced agranulocytosis cases, it was stated that 9 of these were patients taking carbimazole, 9 propylthiouracil and 23 methimazole. For the Hallberg *et al.* study, 39 cases were induced by anti-thyroid agents, of these 29 were methimazole, 5 carbimazole and 5 propylthiouracil. Although all three anti-thyroid drugs have been incorporated into the association studies, our study focuses on the associations seen with methimazole and propylthiouracil. Carbimazole is the pro-drug of methimazole, responsible for the antithyroid activity, and has a short half-life of 5.3–5.4 hours with peak plasma concentrations of methimazole being present after 1 or 2 hours [21, 22]. We are working under the assumption that the mechanisms involved in the adverse drug reaction follow the altered peptide repertoire model, the active drug methimazole would be most likely to interact with the HLA during this process and so carbimazole is excluded from this analysis.

B\*38:02 has a very similar sequence to B\*38:01 with only one residue difference (B\*38:02 Thr80Ile B\*38:01). It therefore is highly possible that B\*38:01 could also be a risk allele in this context. B\*38:02 (Figure G in S1 File) is most commonly found in Asian populations (0.69%–6.6%, frequencies taken from Allele Frequency Net Database (AFND) [23]), with Caucasian (0.003–0.2%) and African populations (0.007–0.06%) having much lower frequencies. B\*38:01 is more commonly found in Caucasian populations (0.6–6.7%), compared to Asian populations (0.15–1.0%) and therefore allele frequencies for B\*38:01 in the Asian populations studies are likely not detected (Figure H in S1 File). The Hallberg *et al.* study (which found B\*27:05 to

be significantly associated) was based on Caucasian populations with a high proportion of Swedish patients. Comparing the allele frequencies from AFND [23] for B\*27:05 (Figure I in S1 File) and B\*38:01 in Swedish populations, B\*27:05 shows frequencies between 10.5 and 20% with B\*38:01 showing no entries—usually indicating that the allele was not detected in these populations. It could therefore be plausible that the B\*38:01 allele could be an ADR-associated allele but this is not seen in the association study due to B\*38:01 being a low frequency allele in Swedish populations. However, without case-tolerant association studies showing patients with B\*38:01, this association cannot be confirmed or rejected. For the purpose of this study, the B\*38:01 is considered as a possible risk allele.

A previous study conducted by Chen *et al.* completed molecular docking with methimazole and propylthiouracil docked with B\*38:02, B\*38:01 and DRB1\*08:03. This study showed poses favouring both the B- and F-pockets along the peptide binding groove. Four suspected key residues were identified: Cys67, Asn77, Thr80 and Thr123. Their study focussed on the associations seen in Taiwan individuals and therefore does not include investigations into the B\*27:05 alleles seen in Caucasian populations. Due to the differences in general structure between Class I and Class II HLA, it is difficult to compare docking predictions between the two allele classes. The B\*38:02 and B\*38:01 alleles were modelled using the same template structures, with the DRB1\*08:03 structure modelled using only DRB1\*01:01 as a template. The quality of the models used for molecular docking can impact the docking results and therefore the selection of templates is very important.

The purpose of this study is to investigate the associated alleles, in particular looking for commonalities between the two associated HLA-B alleles in order to look for similarities in these alleles that might shed light on the underlying mechanisms of the adverse drug reactions seen.

## Methods

### Sequence and structural analysis

In order to conduct a reliable analysis between cases and controls, the control alleles must first be carefully selected to ensure they can be reliably assumed to be non-associated. The case and control frequencies from He *et al.* [16], Chen *et al.* [17] and Hallberg *et al.* [19] were compared alongside healthy individual frequency data obtained from AFND [23]. Alleles where the study control allele frequency or healthy individual frequencies sourced from AFND were similar to or greater than the case allele frequencies were selected as controls (S1 File, i. Control allele selection). These alleles could safely be assumed to be non-associated as they do not show enrichment in the case groups.

Structural differences were assessed between the risk and selected control alleles. Firstly, the peptide binding regions were compared, this is likely where the drug binding would occur and so differences here would be important for the mechanism of interaction involved in the adverse drug reaction. The peptide binding regions were compared using MHC motif viewer [24] and MHCcluster [25]. MHC motif viewer was used to compare the predicted binding motifs for the B\*27:05 and B\*38:02 risk alleles as well as the B\*38:01 possible risk and the selected control alleles. MHCcluster was similarly used to compare the global similarities of peptide binding predictions for the risk, possible risk and selected control alleles. Both the MHC motif viewer and MHCcluster post-process NetMHCpan scores to give predictions of motifs or similarities between peptides. NetMHCpan uses artificial neural networks to predict the peptide binding of HLA molecules based on IEDB experimental data of peptides known to be presented by given alleles, including data for the binding of peptides to B\*27:05 and B\*38:01 [26, 27]. Therefore, the predictions generated will be based largely on experimental data.

The protein molecules were further compared using multiple sequence alignment to view differences across the whole protein and also individual residue changes within sub-pockets along the peptide binding groove between the risk and control alleles. Firstly, the sequences for the risk, possible risk and control alleles were aligned. The alignments were then extended to look at common alleles selected from AFND and NMDP (National Marrow Donor Program) [19, 23] Asian and Caucasian populations and extended further again to consider common alleles found in all populations. When considering the common alleles, it is important to note that the definitive risk profile (association status) of these alleles is unknown. Although, due to the high frequency and the fact that they have not been seen to be associated with agranulocytosis, it can be assumed that these alleles are likely not associated. These multiple sequence alignments allow us to look for residue similarities unique to the risk alleles and therefore identify residues which may be involved in the mechanism of binding for the adverse drug reactions.

### Molecular docking

Molecular docking was used to compare the predicted binding sites between risk and control alleles. Crystal structures were obtained from the PDB database where available. For those alleles where the crystal structures are not available, the structures were predicted using Modeller [28, 29]. Crystal structures, given the suffix ‘\_S’, were available for B\*27:05 (1OGT [30]), B\*15:01 (1XR9 [31]) and B\*51:01 (1E27 [32]). Homology models, given the suffix ‘\_M’, were created for B\*38:02, B\*38:01, B\*40:06, B\*46:01 and B\*54:01 (Table B in S1 File, S1 Models). Methimazole, the active form of carbimazole, and propylthiouracil were used to dock with the B\*27:05\_S and B\*38:02\_M risk alleles, B\*38:01\_M possible risk allele and selected control alleles; B\*15:01\_S, B\*40:06\_M, B\*46:01\_M, B\*51:01\_S and B\*54:01\_M.

Structures and sequences of three similar alleles were obtained searching the PDB database using BLAST-P [33]. The structures for these similar alleles, with high sequence identity (e.g. 95–98% identity for the B\*38:02 templates), were then used as templates for Modeller. Target and template sequences were aligned with ClustalX [34] and ten models were made for each structure, using Modeller 9.9 automodel class [35]. The model with the lowest objective function was chosen for the docking. Table B in S1 File summarises the structures and models obtained. Drug structures for methimazole and propylthiouracil were both obtained from the PDB database (5FF1 [36] and 5HPW [37] respectively).

AutoDockFR [38] was used, with default parameters, to dock both methimazole and propylthiouracil with the B\*27:05 and B\*38:02 risk alleles, B\*38:01 possible risk allele and the selected control alleles. Structural PDBQT files were prepared using AutoDock Tools [39] for both the alleles and drug structures. AutoGrid [40] was used to map the target allele structures and select grid points in order to search both the peptide binding groove (PBG) and the top three largest pockets (Top3). AutoGrid is a graphical user interface for specifying the docking box. This was used to identify pockets within the target structures, the pockets of interest were then selected and the search box calculated to encompass these pockets. The box sizes and positions were therefore automatically generated, centred around the areas of interest. The top three largest pockets were selected by pocket volume. Searching the largest pockets on the protein increases the search space to cover more of the protein and allows an alternate binding pose away from the peptide binding region. This helps identify if the peptide binding groove is in fact the most favourable binding region, ensuring the most favourable poses are obtained. A total of 10 poses were obtained for each allele for each drug and each search space, resulting in 20 poses per allele for each drug. CSM-lig [41] was used to compare the predicted binding affinities of each pose for each drug-allele combination. These predictions were used to

compare the binding affinities between poses seen in the B-pocket and the F-pocket in risk and control alleles. These predicted results were used alongside the docking scores from AutoDockFR to help identify the most favourable pocket for each of the drug-allele combinations. Further analysis was completed conducting 100 runs for each drug-allele combination (S1 File, ii. Docking analysis), this allowed further analysis of the favouring of binding pockets. These poses were then automatically assigned positions (i.e. binding in B or F pocket) using k-mean analysis and used to visualise the favouring of binding poses for each drug-allele combination. Visual inspection of the docking poses showed two distinct groups within the B and F pockets (Figure B in S1 File), therefore k was defined as 2. The co-ordinates of the C1 atom of each pose were used for the clustering, allowing clustering of dissimilar ligands with differing number of atoms. LigPlot [42] was used to visualise the interactions for each of the lowest scoring poses for each drug-allele combination, in order to compare the residues involved with binding.

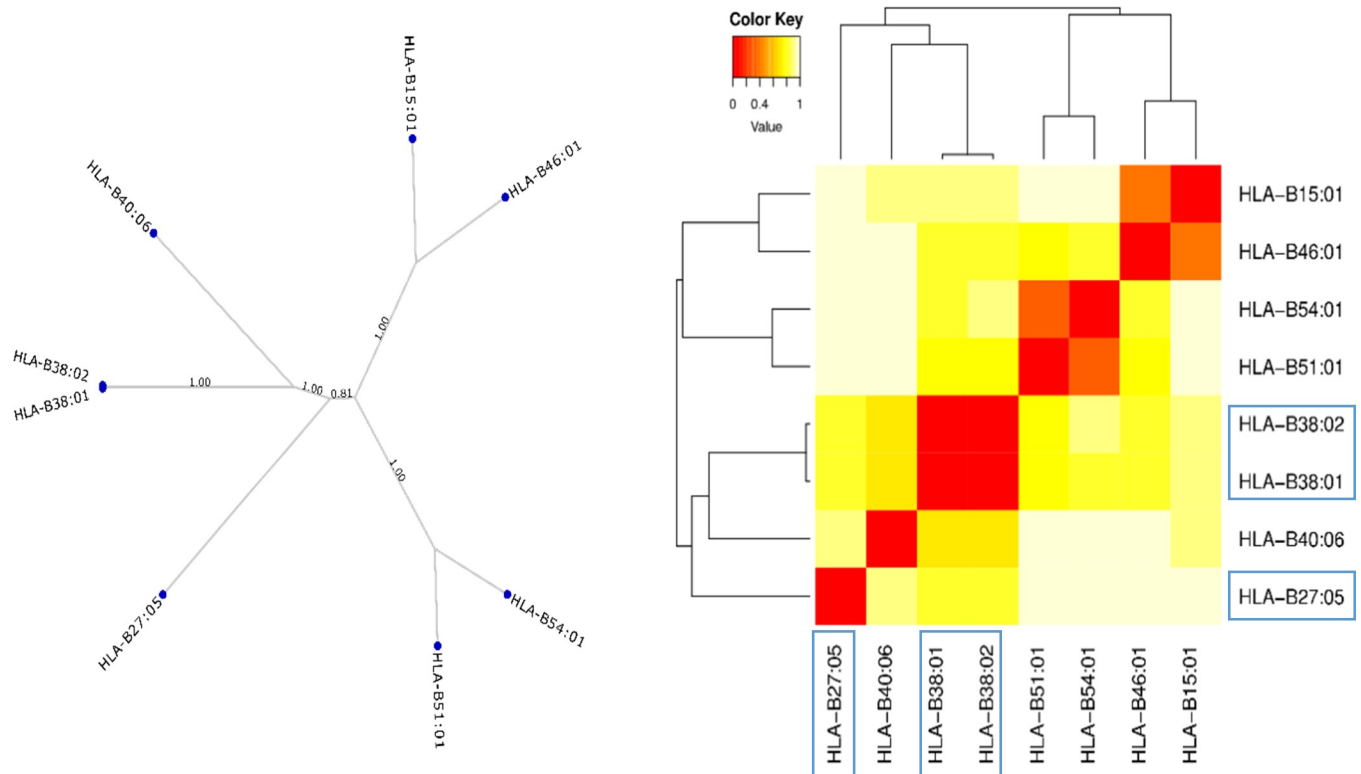
In order to investigate how the size and structure of the ligands and pockets could be having an impact on the molecular docking results, structurally similar molecules were docked to each of the risk and control alleles. The PDB database was searched for ligands with  $\geq 50\%$  similarity to methimazole and propylthiouracil, using the chemical components search to compare chemical descriptors of ligands within the chemical component dictionary [43]. Four ligands were selected: MZY (1,3-dihydroimidazole-2-thione) and TUL (2-thioxo-2,3-dihydropyrimidin-4(1H)-one) have been used as experimental antithyroid drugs and include a thiocarbonyl group, DMI (2,3-Dimethylimidazolium Ion) and EV0 (2-amino-6-propylpyrimidin-4(3H)-one) do not include the thiocarbonyl group, (Figure F in S1 File). These structures were prepared in the same way as the associated anti-thyroid drugs, using AutoDock Tools [39]. AutoDockFR [38] was used to dock the non-associated compounds with B\*27:05 and B\*38:02 risk alleles, B\*38:01 possible risk allele and selected control alleles. The binding poses of these were then compared to those of the associated drugs, in order to deduce if the size and structures of the ligands and pockets could be having an impact on the molecular docking results.

## Results

### Sequence and structural analysis

We firstly investigated whether allele frequencies for controls were representative of larger populations available in similar regions. This allows us to test for potential biased sampling in source studies, especially as controls have been combined from different countries, as well as to determine our own control (non-associated) alleles for further comparison. Case and control frequencies were calculated from the data provided for the He *et al.* and Hallberg *et al.* studies [16, 19]. Alleles with study control frequency over 3% were investigated. Five alleles were selected to be used as controls: B\*15:01, B\*40:06, B\*46:01, B\*51:01 and B\*54:01. It can be reasonably assumed that these alleles are non-associated alleles based on the frequency data of both the Han Northern China and European populations (Figure A in S1 File). The control allele selection is covered in more detail within S1 File (i. Control allele selection).

The MHC Motif viewer [24] displays the preference for given HLA alleles to bind peptides with amino acids at given positions within an n-mer peptide e.g. 9mer for HLA class I. The motifs have been generated via the NetMHCpan [26] prediction method being run over a large selection of natural peptides, which has been trained originally with experimental data (peptides presented by given HLA alleles) from the IEDB database [27]. While the predicted motifs cannot give us a direct measure of the likelihood of a drug to bind in the cleft of a given HLA allele, they can be indicative of whether different alleles share similar peptide binding regions. If the hapten hypothesis of HLA-mediated ADRs was true for anti-thyroid drugs, then

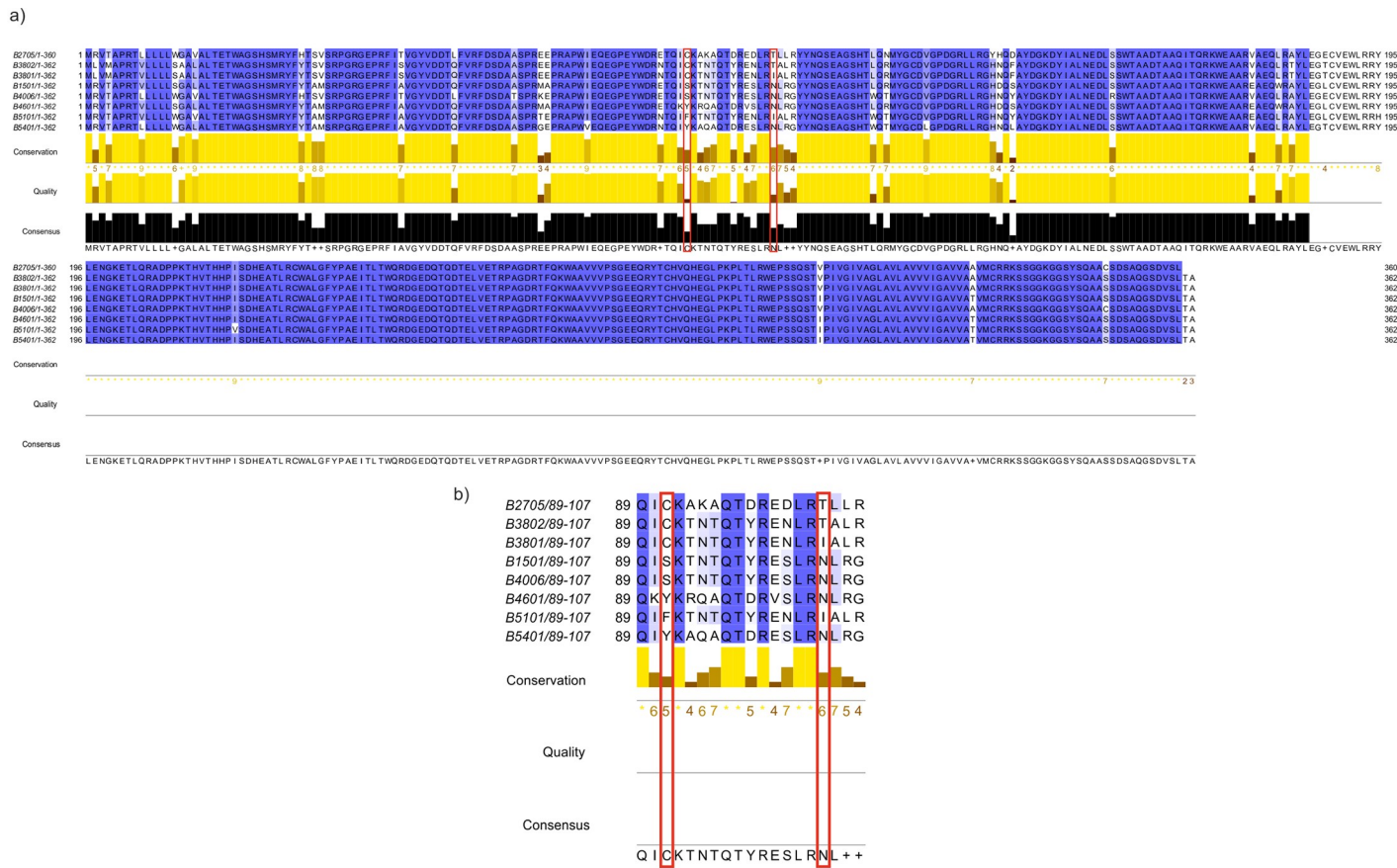


**Fig 1. MHCcluster output for risk and control alleles.** Specificity tree shows clustering of alleles and heat map shows the similarity between the binding motifs of each B allele, comparing known risk and possible risk (B\*38:01) alleles (highlighted in blue) with selected controls. The scale shows the distance between the alleles, with red (0) showing very similar binding motifs and white (1) showing dissimilar binding motifs. Trees to the left and above the matrix show the hierarchical clustering of the different B alleles.

<https://doi.org/10.1371/journal.pone.0220754.g001>

the peptide binding ability of associated alleles might be related, with the caveat that drugs binding to peptides would likely change their affinity for particular alleles. Comparing the peptide binding motifs from MHC Motif Viewer (Figure J in [S1 File](#)) it can be seen that although, as expected due to the high sequence similarity, the B\*38:02 and B\*38:01 alleles show very similar binding motifs, the B\*27:05 allele shows a very different motif. This shows that there are differences between the favoured peptides and thus the peptide binding grooves of the B\*27:05 and B\*38:02 risk alleles. A similar observation can be made from the output for MHCcluster (Fig 1), with differences seen between B\*27:05 and B\*38:02. MHCcluster is also based on predictions from random human peptides processed by NetMHCpan and generates a distance measure between the peptide binding specificity scores generated by the software to create a tree representation. Similar to MHC Motif Viewer, the results can tell us about overall relatedness of predicted peptide binding by different alleles, but not directly about the likelihood of drug binding. From the tree-based output, it can be seen that the B\*38:02 and B\*38:01 alleles show very similar clustering, whereas the B\*27:05 allele shows clustering differing from this and is not more closely related to the other risk allele than any of the selected control alleles. We can conclude therefore that B\*38:02 and B\*38:01 alleles bind highly similar pools of peptides, whereas B\*27:05 binds a largely distinct set of peptides. More generally, the differences seen between the favoured peptides of the risk alleles allows us to conclude that these alleles are not structurally similar.

Multiple sequence alignments were used to investigate individual residue differences between alleles. Comparing the B\*27:05 and B\*38:02 risk alleles with the B\*38:01 possible risk



**Fig 2. Multiple sequence alignment for risk and control alleles.** (a) Multiple sequence alignment for risk alleles B\*27:05 and B\*38:02, possible risk allele B\*38:01 and selected controls B\*15:01, B\*40:06 B\*46:01, B\*51:01 and B\*54:01. (b) Focusing on positions 65–83. Positions 67 and 80 highlighted in red.

<https://doi.org/10.1371/journal.pone.0220754.g002>

and selected controls, it can be seen that there are two residues which are seen to be unique to the risk alleles: Cys67 and Thr80 (Fig 2). These are both residues that were identified as potentially important in the Chen *et al.* study [17]. When the comparisons were extended to look at common alleles, obtained from searching AFND for Caucasian and Asian populations and also common alleles found in the NMDP database [19] (Figure K in S1 File), a similar pattern can be seen. These two residues are rarely found in these common alleles, although it must be noted that the association status of these common alleles is unknown.

### Molecular docking

**Methimazole.** Methimazole was docked with each of the risk, possible risk and selected control alleles, using AutoDockFR [39], in order to compare the favourable binding poses in each case. Table 2 summarises the poses seen for methimazole, showing the pocket of the lowest scoring pose, the number of poses in each pocket and the median scores for the poses in those pockets, for each search space. From this and Fig 3, showing the predicted binding poses for the top scoring pose for each allele, it can be seen that the risk alleles favour the F-pocket for drug binding, close to the position 80 identified. Both scores as well as the number of poses in each pocket, while searching the peptide binding region, favour this pocket. Extending the search space to cover the top three largest pockets, B\*27:05\_S favours a pocket outside of the



Table 2. Summary of molecular docking positions for methimazole.

Status	Allele	Mutation	Peptide Binding Groove					Extended Search Space						
			Lowest	B	Median	F	Median	Lowest	B	Median	F	Median	O	Median
Risk	B*27:05_S	N80T	F	5	-3.20	5	-3.26	O	2	-3.15	0	N/A	8	-3.31
Risk	B*38:02_M	N80T	F	1	-3.10	9	-4.39	F	1	-3.22	7	-4.39	2	-3.88
Possible risk	B*38:01_M	N80I	F	0	N/A	10	-3.66	O	6	-3.62	0	N/A	4	-3.89
Control	B*15:01_S	N80	B	8	-3.60	2	-3.58	O	2	-3.55	0	N/A	8	-4.14
Control	B*40:06_M	N80	F	0	N/A	10	-3.65	F	0	N/A	10	-3.74	0	N/A
Control	B*46:01_M	N80	B	10	-3.55	0	N/A	B	7	-3.43	3	-3.32	0	N/A
Control	B*51:01_S	N80I	B	5	-3.58	5	-3.54	F	2	-3.54	8	-3.55	0	N/A
Control	B*54:01_M	N80	F	1	-3.29	9	-3.47	F	1	-3.48	9	-3.45	0	N/A

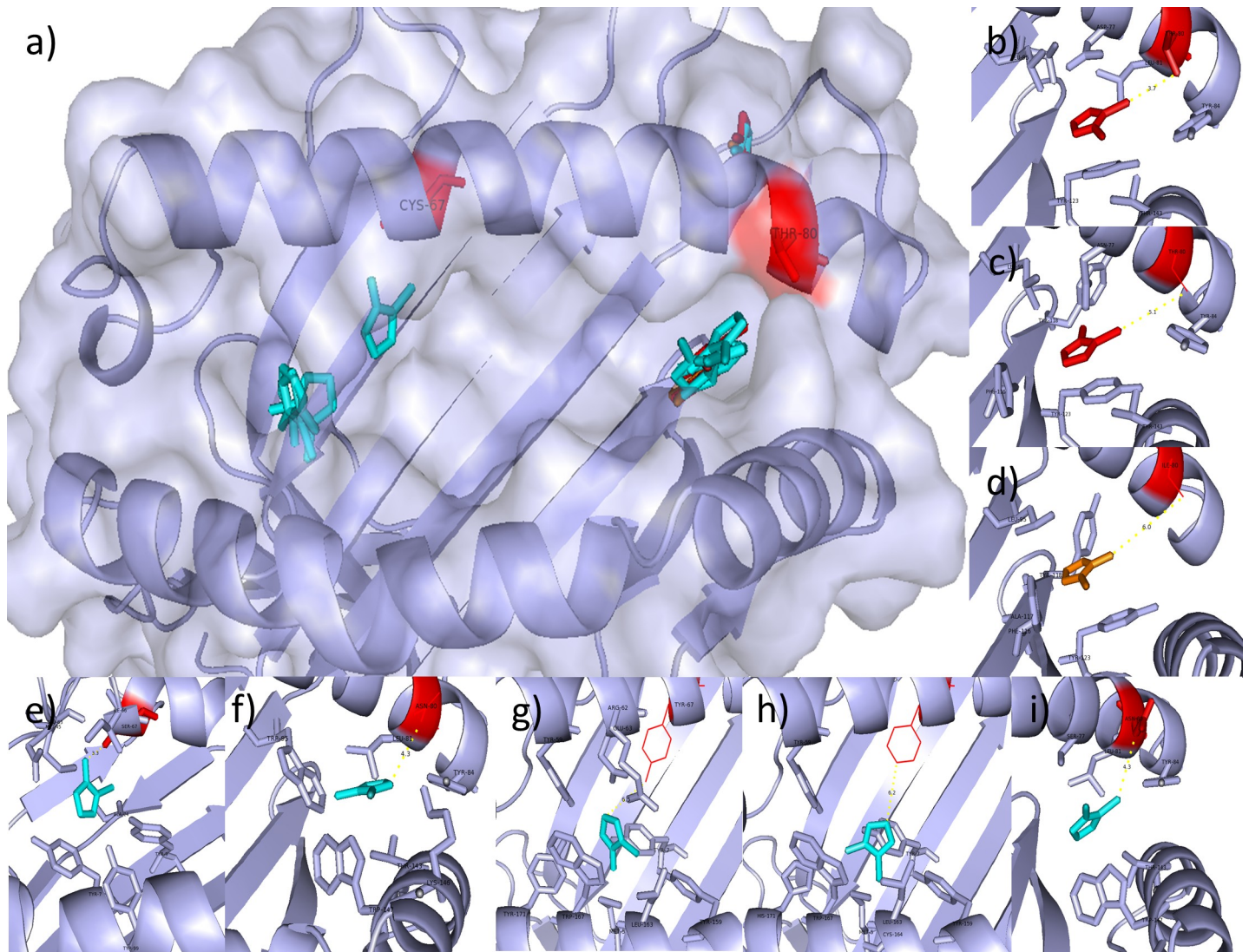
Position of the lowest scoring methimazole pose along with the number of poses in each pocket and the median of the pose scores in each pocket for each of the alleles using the search space covering the peptide binding groove and the top 3 pockets identified on the protein (Extended search space). Scores given as kcal/mol. 'O' refers to pockets other than the B and F, i.e. outside of the binding groove.

<https://doi.org/10.1371/journal.pone.0220754.t002>

peptide binding groove, although still close to the position 80 identified. B\*38:02\_M still favours the F-pocket and B\*38:01\_M favours the B-pocket, with some poses seen outside of the peptide binding groove. Control alleles B\*15:01\_S, B\*46:01\_M and B\*51:01\_S all favour the B-pocket, both with scores and number of poses, while B\*40:06\_M and B\*54:01\_M favour the F-pocket, searching the peptide binding groove. Extending the search space to cover the top three largest poses for each of these alleles, B\*15:01\_S now favours a pocket outside of the binding groove, B\*46:01\_M still shows favouring of the B-pocket and B\*40:06\_M, B\*51:01\_S and B\*54:01\_M favour the F-pocket. Comparing the interactions seen from the LigPlot diagrams (Figure L in S1 File), it can be seen that B\*27:05\_S and B\*54:01\_M both are predicted to have potential interactions with methimazole at residue 80. The residue at position 80 of B\*27:05\_S forms hydrophobic interactions with the thiocarbonyl group of the methimazole, where the position 80 residue for B\*52:01\_M interacts with one of the nitrogen atoms.

It has previously been shown that the process and parameterisation of homology modelling may have an impact on the molecular docking results, compared to docking within a crystal structure [44]. It can be seen here that the control alleles favouring the F-pocket are generally the modelled structures, with the B\*15:01 and B\*51:01 known crystal structures showing favouring of the B-pocket. Similarly, B\*38:02 and B\*38:01 are both modelled structures. Due to the very high sequence similarity between these alleles, it may be that these alleles show similar binding due to the similarity of the modelled structures. Two alleles that differ little will often show similar models, although larger differences would likely have an effect. These models were shown to be structurally similar with an RMSD of 0.460Å (221 to 221 atoms).

**Propylthiouracil.** Both methimazole and propylthiouracil have been associated with drug induced agranulocytosis with B\*27:05 and B\*38:02. Propylthiouracil was therefore also docked with the risk, possible risk and selected control alleles using AutoDockFR [39]. The docking results were then compared between alleles and with the methimazole results to identify difference in favourable binding poses between the drug-allele combinations. Table 3 summarises the poses seen for propylthiouracil, showing the pocket of the lowest scoring pose, the number of poses in each pocket and the median scores for the poses in those pockets for each search space. From this and Fig 4, showing the predicted binding poses for the top scoring pose for each allele, similar patterns to those seen for methimazole can be seen. The B\*27:05\_S and B\*38:02\_M risk alleles favour the F-pocket searching the peptide binding groove and favour other pockets outside of the peptide binding groove when extending the search, with poses



**Fig 3. Molecular docking poses for methimazole.** a) Top scoring docking poses of methimazole for B\*38:02\_M and B\*27:05\_S risk alleles (red), B\*38:01\_M non-associated allele (orange) and the control alleles (blue) B\*15:01\_S, B\*40:06\_M, B\*46:01\_M, B\*51:01\_S and B\*54:01\_M, using peptide binding groove search space and top3 pockets search space for AutoDockFR. b-i) Docking poses showing interacting residues as sticks for B\*27:05 (b) and B\*38:02 (c) risk alleles, B\*38:01 (d) suspected risk allele and control alleles B\*15:01 (e), B\*40:06 (f), B\*46:01 (g), B\*51:01 (h) and B\*54:01 (i). Positions 67 and 80 are shown in red, with measurements between ligands and side chain shown in yellow (Å).

<https://doi.org/10.1371/journal.pone.0220754.g003>

shown to lie close to Thr80 and B\*38:02\_M showing the lowest scoring pose within the F-pocket. The B\*38:01\_M possible risk shows favouring of the F-pocket searching both the peptide binding groove and extending to cover the top three largest pockets. Looking at the control alleles, B\*46:01\_M favours the B-pocket with both search spaces, B\*15:01\_S favours the B-pocket with scores but not with number of poses when searching the peptide binding groove and favours pockets outside of the groove when extending the search space, including a pocket close to the Thr80 position. B\*40:06\_M, B\*51:01\_S and B\*54:01\_M all favour the F-pocket with both search spaces. Comparing the LigPlot poses for the best scoring propylthiouracil poses (Figure M in S1 File), searching the peptide binding groove, it can be seen that B\*27:05\_S, B\*38:02\_M, B\*38:01\_M, B\*40:06\_M, B\*51:01\_S and B\*54:01\_M all show hydrophobic interactions with propylthiouracil at the position 80 residue. B\*27:05 and B\*38:01 show

Table 3. Summary of molecular docking positions for propylthiouracil.

Status	Allele	Mutation	Peptide Binding Groove					Extended Search Space						
			Lowest	B	Median	F	Median	Lowest	B	Median	F	Median	O	Median
Risk	B*27:05_S	N80T	F	3	-5.84	7	-6.23	O	1	-5.48	0	N/A	9	-5.69
Risk	B*38:02_M	N80T	F	4	-5.66	6	-6.60	F	0	N/A	3	-6.58	7	-5.62
Possible risk	B*38:01_M	N80I	F	0	N/A	10	-6.48	F	0	N/A	10	-6.45	0	N/A
Control	B*15:01_S	N80	B	2	-6.34	8	-6.30	O	0	N/A	0	N/A	10	-6.65
Control	B*40:06_M	N80	F	0	N/A	10	-5.74	F	0	N/A	10	-5.80	0	N/A
Control	B*46:01_M	N80	B	10	-6.43	0	N/A	B	10	-6.41	0	N/A	0	N/A
Control	B*51:01_S	N80I	F	0	N/A	10	-6.45	F	3	-6.00	7	-6.44	0	N/A
Control	B*54:01_M	N80	F	0	N/A	10	-6.42	F	0	N/A	10	-6.41	0	N/A

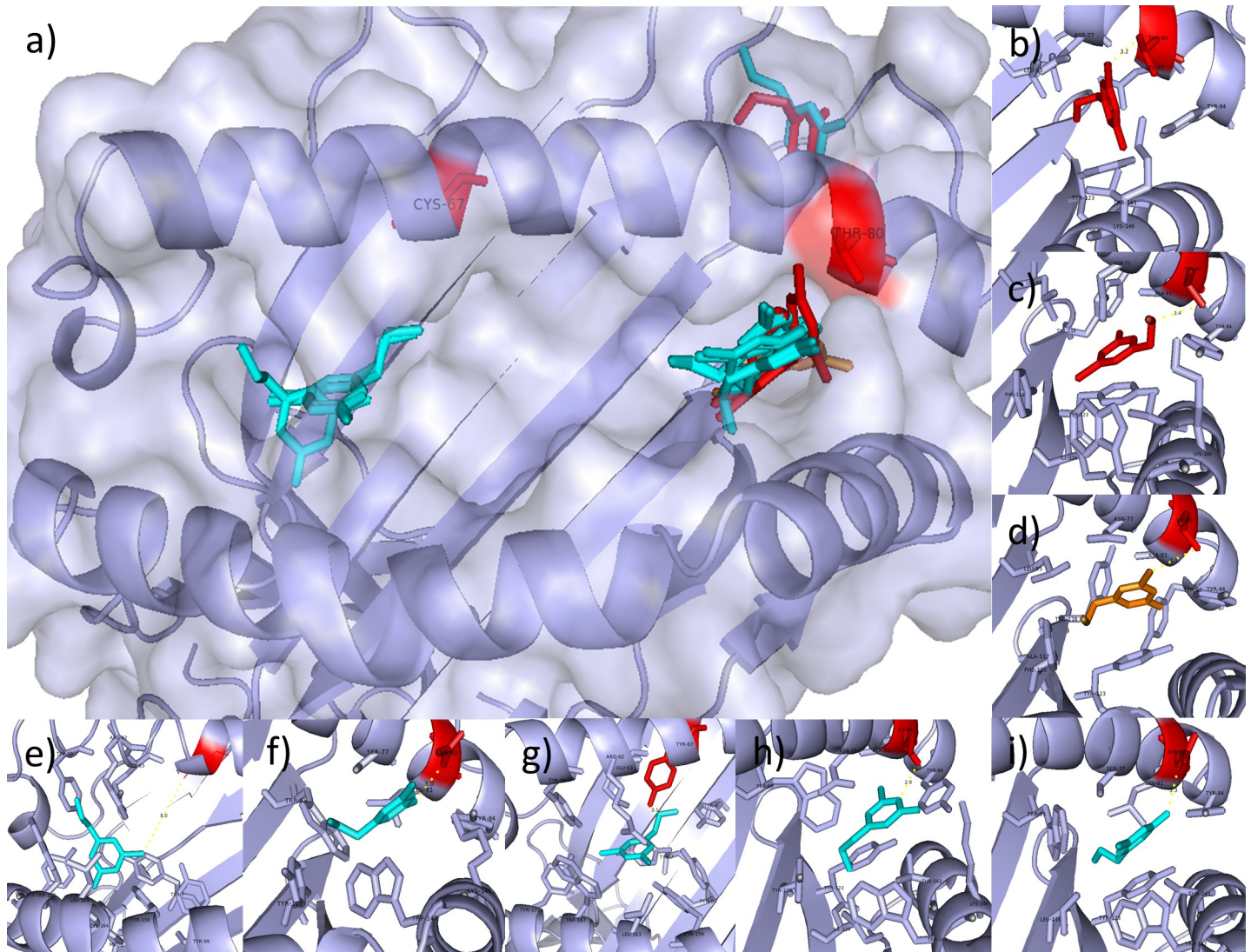
Position of the lowest scoring propylthiouracil pose along with the number of poses in each pocket and the median of the pose scores in each pocket for each of the alleles using the search space covering the peptide binding groove and the top 3 pockets identified on the protein (Extended search space). Scores given as kcal/mol. ‘O’ refers to pockets other than the B and F, i.e. outside of the binding groove.

<https://doi.org/10.1371/journal.pone.0220754.t003>

interactions of this position with the thiocarbonyl group, B\*38:02 with the carbon tail, B\*54:01 with the oxygen atom and both B\*40:06 and B\*51:01 with one of the nitrogen atoms, with the interaction seen in B\*40:06 being a hydrogen bond rather than the usual hydrophobic interaction seen.

CSM-lig [41] was used to predict the binding affinity for each pose of each drug-allele combination, where higher predictions indicate more favourable binding. Figure E in S1 File shows some differences between the favouring seen from the docking scores. Looking at the risk alleles, CSM-lig shows slight favouring of the B pocket for B\*27:05 and B\*38:02 with propylthiouracil. With methimazole, favouring of the B is seen with B\*27:05 and favouring of the F with B\*38:02. B\*38:01 shows constant favouring of the F pocket, with all poses being predicted to bind in the F pocket. The control alleles show a mixture of B and F favouring. CSM-lig has previously been demonstrated to differentiate between binding in risk and control alleles for the abacavir association seen with B\*57:01 (unpublished analysis). No clear patterns can be seen between the risk and control groups (Figure E in S1 File) comparing the predicted binding affinities, this may be due to a weaker association being seen than with the abacavir drug-allele associations, resulting in less favourable binding. We next expanded the docking analysis to 100 runs to explore whether further trends in poses found for B and F pockets could be observed (Fig 5). From this we can see consistent favouring of the F pocket for the risk alleles. Although the docking is not able to fully distinguish between the risk and controls, with the controls showing mixed favouring of the B and F pockets, it is evident that docking against the risk alleles strongly favours binding in the F pocket for each of the drugs.

Similar molecules to the investigated methimazole and propylthiouracil were identified: MZY, TUL, DMI and EV0. The similar molecules were docked to each of the risk and control alleles in order to deduce if the size and structures of the ligands and pockets could be having an impact on the molecular docking results (Figure F in S1 File). From these investigations, it was found that the similar molecules that have been used as experimental anti-thyroid drugs and included the thiocarbonyl group (MZY and TUL), showed similar binding patterns to the associated anti-thyroid drugs with the risk alleles favouring the F-pocket. The possible risk B\*38:01\_M allele can be seen to favour the B-pocket by scores but the F by number of poses for MZY and favour the F for TUL (Table C in S1 File). Control alleles B\*15:01\_S and B\*46:01\_M both favour the B via lower scores and more poses for both drugs. With control B\*51:01\_S also favouring the B for TUL considering scores but not poses. For the other

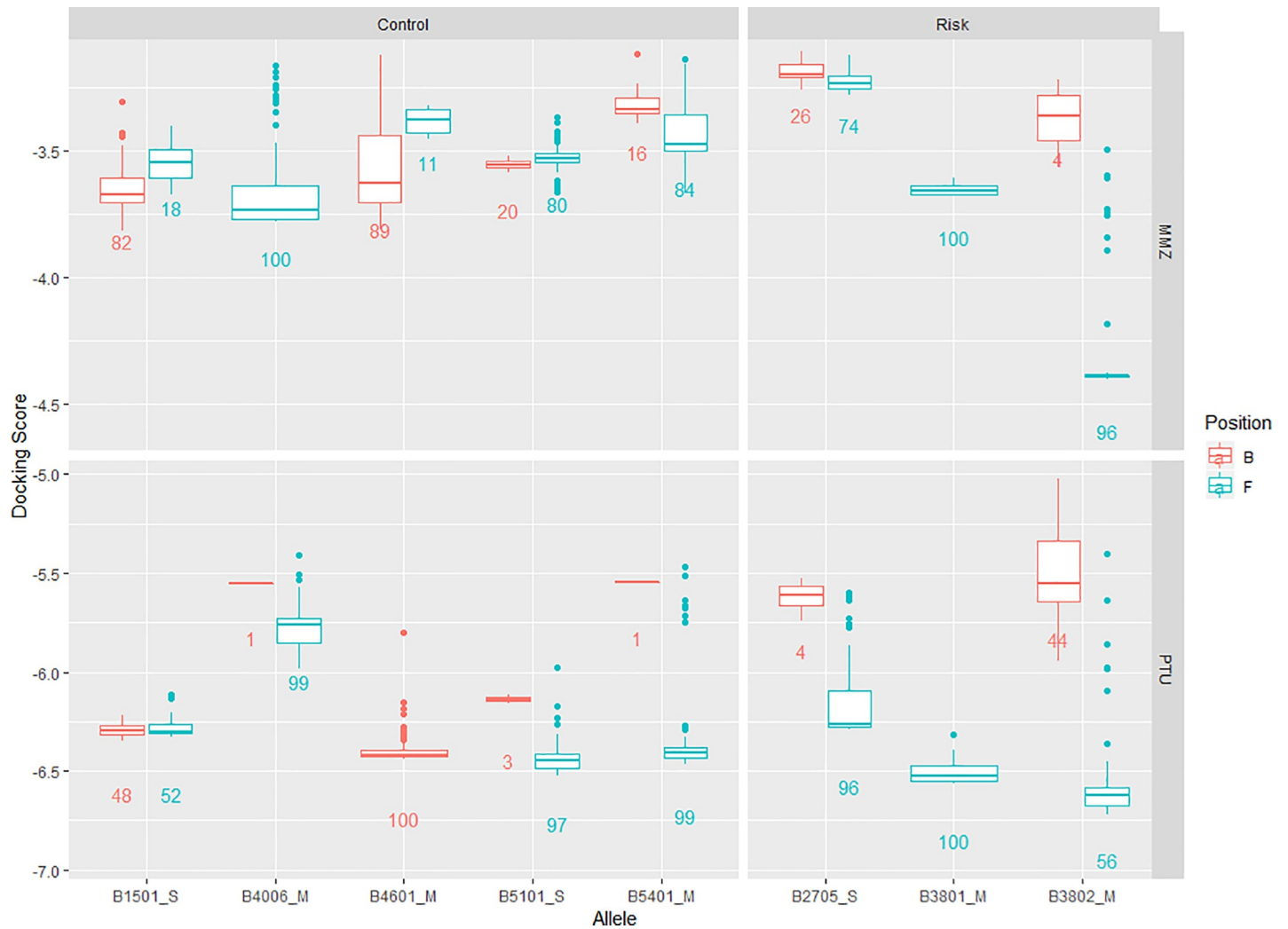


**Fig 4. Molecular docking poses for propylthiouracil.** a) Top scoring docking poses of propylthiouracil for B\*38:02\_M and B\*27:05\_S risk alleles (red), B\*38:01\_M non-associated allele (orange) and the control alleles (blue) B\*15:01\_S, B\*40:06\_M, B\*46:01\_M, B\*51:01\_S and B\*54:01\_M, using peptide binding groove search space and top3 pockets search space for AutoDockFR. b-i) Docking poses showing interacting residues as sticks for B\*27:05 (b) and B\*38:02 (c) risk alleles, B\*38:01 (d) suspected risk allele and control alleles B\*15:01 (e), B\*40:06 (f), B\*46:01 (g), B\*51:01 (h) and B\*54:01 (i). Positions 67 and 80 are shown in red, with measurements between ligands and side chain shown in yellow (Å).

<https://doi.org/10.1371/journal.pone.0220754.g004>

molecules without the sulfhydryl group (DMI and EV0), a difference in trends is seen with the B\*27:05\_S favouring the B-pocket for both drugs and B\*38:02\_M showing the lowest scoring pose for DMI favouring the B-pocket but still favouring the F based on poses and for EV0. B\*38:01\_M possible risk still shows favouring of the F-pocket for these drugs. The controls all show favouring of the B through scores and number of poses for both drugs, except B\*54:01\_M, which shows favouring of the F-pocket for EV0. Docking poses are shown in Figure N in [S1 File](#).

[Table 4](#) summarises the number of poses making hydrophobic interactions with position Thr80 of both risk alleles for each of the investigated drugs. It can be seen that for B\*27:05 the associated drugs often make Thr80 interactions with interactions with the thiocarbonyl group being most favourable. The experimental drugs show similar interactions with Thr80 as those



**Fig 5. Boxplots for methimazole and propylthiouracil.** Boxplots showing docking scores for 100 poses searching the peptide binding groove, using both methimazole (MMZ) and propylthiouracil (PTU) for each of the alleles.

<https://doi.org/10.1371/journal.pone.0220754.g005>

seen with the associated drugs (Figure O in [S1 File](#)). For B\*38:02, fewer interactions are made between the drugs and the position 80 Thr residue than seen for B\*27:05 (Figure P in [S1 File](#)). In both risk alleles, the most favourable poses commonly form interactions between the thio-carbonyl group of methimazole, propylthiouracil, MZY and TUL, and positions 77, 80, 81, 84 and 123. Interactions are also seen with positions 95, 116, 124, 143 and 147. All these residues seen making interactions surround the F-pocket. Positions 77, 80 and 123 were identified as of interest by the Chen *et al.* study [17]. This adds strength to the hypothesis that the Thr80 position could be involved in the mechanism here, especially for B\*27:05. The structure of the ligands, mainly the 'S' group found in the associated and experimental anti-thyroid drugs, could therefore be potentially important for the predicted binding poses seen here.

## Discussion

The purpose of this study was to investigate the associations seen between HLA and anti-thyroid alleles, focusing on the commonalities seen between the HLA-B associated alleles, to

Table 4. Summary of Thr80 interactions for investigated ligands.

Drug	B* 27:05				B* 38:02			
	No Thr80	Thr80-S	Other Thr80	Most favourable pose	No Thr80	Thr80-S	Other Thr80	Most favourable pose
MMZ	5	4	1	Thr80-S	10	0	0	No Thr80
PTU	3	5	2	Thr80-S	5	2	3	Other Thr80
MZY	1	2	7	Other Thr80	9	0	1	No Thr80
TUL	5	2	3	Thr80-S	10	0	0	No Thr80
DMI	9	0	1	No Thr80	10	0	0	No Thr80
EVO	10	0	0	No Thr80	5	0	5	Other Thr80

Summary counts showing the number of poses seen making interactions with the Thr80 residue of each of the associated risk alleles for the associated drugs (MMZ and PTU), the experimental anti-thyroid drugs containing the thiocarbonyl group (MZY and TUL) and the other investigated ligands. Where 'No Thr 80' relates to poses showing no interactions between the Thr80 residue and the drug, 'Thr80-S' shows interactions between the thiocarbonyl group and Thr80 and 'Other Thr80' indicates interactions made between Thr80 and the drug but not with the thiocarbonyl group. The 'Most favourable pose' column shows the interaction seen for the pose which showed the lowest docking score for each of the drug-allele combinations.

<https://doi.org/10.1371/journal.pone.0220754.t004>

identify a potential shared mechanism. This was done through comparing the peptide binding regions, including the whole peptide binding groove and specific residue changes alongside the predicted binding poses of the drugs with each of the risk and control alleles. It was found that the risk alleles favour different peptides and so we can conclude that the gross structures of their peptide binding grooves are rather dissimilar. However, when the multiple sequence alignments were used to focus on specific residue changes, it could be seen that two residues were found to be unique to the risk alleles. These two residues, Cys67 and Thr80, were therefore considered to be potentially important for the mechanism of action for the adverse drug reactions seen, confirming the results of a previous study by Chen *et al.* where the Cys67 and Thr80 were identified, amongst others, as potentially important for the binding of the associated drugs with the risk alleles [17]. From the results of the molecular docking, the risk alleles were shown to favour the F-pocket for both drugs. This pocket lies alongside the Thr80 residue which was identified as potentially important for the mechanism of action. It was seen that the Thr80 interacts hydrophobically with the thiocarbonyl group of both associated drugs in B\*27:05, with similar interactions also being seen, but to a lesser extent, with B\*38:02. The residue at position 80 of the control alleles was seen to only make interactions with the methimazole for B\*54:01, although these were different interactions from those seen for the risk alleles. For propylthiouracil, this position made interactions with the molecule for three of the five control alleles, although these were again seen as different interactions to the risk, with the B\*38:01 possible risk showing similar interactions to the B\*27:05 risk allele. It is therefore reasonable to conclude that Thr80 could be involved in the mechanism of interaction, if only indirectly by influencing the conformations of the surrounding residues located around the F-pocket (Figure Q in S1 File).

Although the docking results for the risk alleles showed consistent results, with the risk alleles always favouring the F-pocket, the predicted poses for the control alleles showed some variation with some drug-allele combinations favouring the B-pocket and some favouring the F-pocket. Docking has previously been shown to be imperfect when considering these complex HLA cases [44], it is therefore understandable that the docking was unable to distinguish completely between risk and control for this case. Docking results are commonly reported as the most favourable pose, here we show a comparison of multiple runs as well as comparisons with selected control alleles. The risk alleles are shown to continuously favour the F-pocket over many runs, providing evidence that this is more likely position of binding. The variation

seen between controls could be due to a number of factors including the homology modelling of the protein structures. As discussed in a previous study [44], homology modelling can impact the docking performance and inaccurate models could result in inaccurate docking results. Here, the similar alleles B\*38:02 (risk) and B\*38:01 (possible risk), with one mutated residue at position 80, were docked with both of the associated drugs. These alleles showed similar binding patterns with both alleles generally favouring the F-pocket. Since we have concluded here that the Thr80 could potentially be involved in the mechanism of action, it would be expected that the B\*38:01 would show different docking results to the risk alleles as this allele does not possess a Threonine residue at position 80. However, both the B\*38:02 and B\*38:01 structures were obtained through homology modelling. This could impact the results of the molecular docking as although the mutation has been modelled, it may not have been accurately represented here and the similar structures could produce similar models.

The results seen here confirm and build on those seen in the Chen *et al.* study [17], with Thr80 being identified as important for the mechanism of the adverse drug reaction. The Chen *et al.* study conducted docking of B\*38:02, B\*38:01 and DRB1\*08:03 with methimazole and propylthiouracil. Models were created for B\*38:02 and B\*38:01 using the same five templates (A\*24:02, C\*08:01 and A\*02:01 along with mouse MHC and human HLA-E) and the model for DRB1\*08:03 created using one template (DRB1\*01:01). The templates used for the B alleles show similarities of 83–86% with B\*38:02 and the DRB1\*01:01 template shows 92% identity with DRB1\*08:03. In this study, the B\*38:02 model was created using templates with 95–98% identity with differing templates with similar identity being used for the other modelled structures created. Model selection is a very important aspect of molecular docking and can greatly impact the docking results seen [44]. This study was able to recreate similar docking results seen previously for B\*38:02 and B\*38:01, using our own modelled structures, whilst also incorporating docking results for the B\*27:05 risk allele and comparisons with selected control alleles. Here, we also went further to investigate the structural differences between the associated risk alleles and selected controls, through sequence alignments and comparison of binding motifs.

In order to further our understandings of the mechanisms involved here and to test the hypothesis that the Thr80 is important for binding, it would be interesting to investigate B\*38:01 further as this is a good potential biological control due to its similarity with the associated B\*38:02. For this, further association studies to confirm the association status of B\*38:01 would be needed. Structural analysis through crystal structures would be needed to confirm the binding predictions of the drug allele combinations investigated and therefore the involvement of Thr80 in the predisposition to this adverse drug reaction.

## Supporting information

**S1 File. Supporting information.** Supporting figures and tables.  
(DOCX)

**S1 Models. Homology models.** Homology models used for alleles without crystal structures available.

(ZIP)

## Author Contributions

**Investigation:** Kerry A. Ramsbottom.

**Supervision:** Daniel F. Carr, Daniel J. Rigden, Andrew R. Jones.

**Writing – original draft:** Kerry A. Ramsbottom.

**Writing – review & editing:** Kerry A. Ramsbottom.

## References

1. Negrini S, Becquemont L. HLA-associated drug hypersensitivity and the prediction of adverse drug reactions. *Pharmacogenomics*. 2017; 18(15):1441–57. <https://doi.org/10.2217/pgs-2017-0090> PMID: 29017379
2. Illing PT, Purcell AW, McCluskey J. The role of HLA genes in pharmacogenomics: unravelling HLA associated adverse drug reactions. *Immunogenetics*. 2017; 69(8–9):617–30. <https://doi.org/10.1007/s00251-017-1007-5> PMID: 28695285
3. Chen CB, Abe R, Pan RY, Wang CW, Hung SI, Tsai YG, et al. An Updated Review of the Molecular Mechanisms in Drug Hypersensitivity. *J Immunol Res*. 2018; 2018:6431694. <https://doi.org/10.1155/2018/6431694> PMID: 29651444
4. Coico R, Sunshine G. *Immunology. [electronic book] : a short course*: West Sussex, England : Wiley Blackwell, 2015. Seventh edition.
5. Yun J, Cai F, Lee FJ, Pichler WJ. T-cell-mediated drug hypersensitivity: immune mechanisms and their clinical relevance. *Asia Pac Allergy*. 2016; 6(2):77–89. <https://doi.org/10.5415/apallergy.2016.6.2.77> PMID: 27141480
6. Pavlos R, Mallal S, Phillips E. HLA and pharmacogenetics of drug hypersensitivity. *Pharmacogenomics*. 2012; 13(11):1285–306. <https://doi.org/10.2217/pgs.12.108> PMID: 22920398
7. Illing PT, Vivian JP, Dudek NL, Kostenko L, Chen Z, Bharadwaj M, et al. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature*. 2012; 486(7404):554–8. <https://doi.org/10.1038/nature11147> PMID: 22722860
8. Ostrov DA, Grant BJ, Pompeu YA, Sidney J, Harndahl M, Southwood S, et al. Drug hypersensitivity caused by alteration of the MHC-presented self-peptide repertoire. *Proc Natl Acad Sci U S A*. 2012; 109(25):9959–64. <https://doi.org/10.1073/pnas.1207934109> PMID: 22645359
9. Zeng Z, Castano AR, Segelke BW, Stura EA, Peterson PA, Wilson IA. Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. *Science*. 1997; 277(5324):339–45. <https://doi.org/10.1126/science.277.5324.339> PMID: 9219685
10. Johansen TE, McCullough K, Catipovic B, Su XM, Amzel M, Schneck JP. Peptide binding to MHC class I is determined by individual pockets in the binding groove. *Scand J Immunol*. 1997; 46(2):137–46. <https://doi.org/10.1046/j.1365-3083.1997.d01-102.x> PMID: 9583994
11. Pradhan S, Sarma H, Bharadwaz B, Mattaparthi VSK. Comparative Study on the Binding Affinity of Methimazole and Propylthiouracil to Thyroid Peroxidase as an Anti-Thyroid Drug: An Insilico Approach. *Journal of Molecular Imaging & Dynamics*. 2017; 07.
12. John M, Sundrarajan R, Gomadam SS. Anti-thyroid drugs in pediatric Graves' disease. *Indian J Endocrinol Metab*. 2015; 19(3):340–6. <https://doi.org/10.4103/2230-8210.152766> PMID: 25932387
13. Vicente N, Cardoso L, Barros L, Carrilho F. Antithyroid Drug-Induced Agranulocytosis: State of the Art on Diagnosis and Management. *Drugs R D*. 2017; 17(1):91–6. <https://doi.org/10.1007/s40268-017-0172-1> PMID: 28105610
14. van Staa TP, Boulton F, Cooper C, Hagenbeek A, Inskip H, Leufkens HG. Neutropenia and agranulocytosis in England and Wales: incidence and risk factors. *Am J Hematol*. 2003; 72(4):248–54. <https://doi.org/10.1002/ajh.10295> PMID: 12666135
15. Andres E, Zimmer J, Affenberger S, Federici L, Alt M, Maloisel F. Idiosyncratic drug-induced agranulocytosis: Update of an old disorder. *Eur J Intern Med*. 2006; 17(8):529–35. <https://doi.org/10.1016/j.ejim.2006.07.012> PMID: 17142169
16. He Y, Zheng J, Zhang Q, Hou P, Zhu F, Yang J, et al. Association of HLA-B and HLA-DRB1 polymorphisms with antithyroid drug-induced agranulocytosis in a Han population from northern China. *Sci Rep*. 2017; 7(1):11950. <https://doi.org/10.1038/s41598-017-12350-2> PMID: 28931918
17. Chen PL, Shih SR, Wang PW, Lin YC, Chu CC, Lin JH, et al. Genetic determinants of antithyroid drug-induced agranulocytosis by human leukocyte antigen genotyping and genome-wide association study. *Nat Commun*. 2015; 6:7633. <https://doi.org/10.1038/ncomms8633> PMID: 26151496
18. Thao MP, Tuan PVA, Linh LGH, Van Hoang L, Hen PH, Hoa LT, et al. Association of HLA-B (\*)38:02 with Antithyroid Drug-Induced Agranulocytosis in Kinh Vietnamese Patients. *Int J Endocrinol*. 2018; 2018:7965346. <https://doi.org/10.1155/2018/7965346> PMID: 30123266
19. Hallberg P, Eriksson N, Ibanez L, Bondon-Guitton E, Kreutz R, Carvajal A, et al. Genetic variants associated with antithyroid drug-induced agranulocytosis: a genome-wide association study in a European



- population. *Lancet Diabetes Endocrinol.* 2016; 4(6):507–16. [https://doi.org/10.1016/S2213-8587\(16\)00113-3](https://doi.org/10.1016/S2213-8587(16)00113-3) PMID: 27157822
20. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018; 46(D1):D1074–D82. <https://doi.org/10.1093/nar/gkx1037> PMID: 29126136
  21. Datapharm. The electronic Medicines Compendium (eMC) [Available from: <https://www.medicines.org.uk/emc/>].
  22. Kampmann JP, Hansen JM. Clinical pharmacokinetics of antithyroid drugs. *Clin Pharmacokinet.* 1981; 6(6):401–28. <https://doi.org/10.2165/00003088-198106060-00001> PMID: 6172233
  23. Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, da Silva AL, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res.* 2015; 43(Database issue):D784–8. <https://doi.org/10.1093/nar/gku1166> PMID: 25414323
  24. Rapin N, Hoof I, Lund O, Nielsen M. MHC motif viewer. *Immunogenetics.* 2008; 60(12):759–65. <https://doi.org/10.1007/s00251-008-0330-2> PMID: 18766337
  25. Thomsen M, Lundegaard C, Buus S, Lund O, Nielsen M. MHCcluster, a method for functional clustering of MHC molecules. *Immunogenetics.* 2013; 65(9):655–65. <https://doi.org/10.1007/s00251-013-0714-9> PMID: 23775223
  26. Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics.* 2009; 61(1):1–13. <https://doi.org/10.1007/s00251-008-0341-z> PMID: 19002680
  27. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* 2019; 47(D1):D339–D43. <https://doi.org/10.1093/nar/gky1006> PMID: 30357391
  28. Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics.* 2014; 47:5 6 1–32.
  29. Webb B, Madhusudhan. MS, Shen. M-Y, Dong. G, Marc A. Marti-Renom NE, Frank Alber, Maya Topf., Baldomero Oliva AF, Roberto Sánchez, Bozidar Yerkovich., et al. Method for comparative protein structure modeling by MODELLER 2015 [Release 9.15, r10497:[Available from: <https://salilab.org/modeller/9.15/manual/node11.html>].
  30. Hulsmeyer M, Fiorillo MT, Bettosini F, Sorrentino R, Saenger W, Ziegler A, et al. Dual, HLA-B27 subtype-dependent conformation of a self-peptide. *J Exp Med.* 2004; 199(2):271–81. <https://doi.org/10.1084/jem.20031690> PMID: 14734527
  31. Roder G, Blicher T, Justesen S, Johannesen B, Kristensen O, Kastrup J, et al. Crystal structures of two peptide-HLA-B\*1501 complexes; structural characterization of the HLA-B62 supertype. *Acta Crystallogr D Biol Crystallogr.* 2006; 62(Pt 11):1300–10. <https://doi.org/10.1107/S0907444906027636> PMID: 17057332
  32. Maenaka K, Maenaka T, Tomiyama H, Takiguchi M, Stuart DI, Jones EY. Nonstandard peptide binding revealed by crystal structures of HLA-B\*5101 complexed with HIV immunodominant epitopes. *J Immunol.* 2000; 165(6):3260–7. <https://doi.org/10.4049/jimmunol.165.6.3260> PMID: 10975842
  33. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25(17):3389–402. <https://doi.org/10.1093/nar/25.17.3389> PMID: 9254694
  34. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 1997; 25(24):4876–82. <https://doi.org/10.1093/nar/25.24.4876> PMID: 9396791
  35. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993; 234(3):779–815. <https://doi.org/10.1006/jmbi.1993.1626> PMID: 8254673
  36. Singh RP, Singh A, Sirohi HV, Singh AK, Kaur P, Sharma S, et al. Dual binding mode of antithyroid drug methimazole to mammalian heme peroxidases—structural determination of the lactoperoxidase-methimazole complex at 1.97 Å resolution. *FEBS Open Bio.* 2016; 6(7):640–50. <https://doi.org/10.1002/2211-5463.12051> PMID: 27398304
  37. Singh RP, Singh A, Kushwaha GS, Singh AK, Kaur P, Sharma S, et al. Mode of binding of the antithyroid drug propylthiouracil to mammalian haem peroxidases. *Acta Crystallogr F Struct Biol Commun.* 2015; 71(Pt 3):304–10. <https://doi.org/10.1107/S2053230X15001806> PMID: 25760705
  38. Ravindranath PA, Forli S, Goodsell DS, Olson AJ, Sanner MF. AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. *PLoS Comput Biol.* 2015; 11(12): e1004586. <https://doi.org/10.1371/journal.pcbi.1004586> PMID: 26629955

39. Sanner MF. Python: a programming language for software integration and development. *J Mol Graph Model*. 1999; 17(1):57–61. PMID: [10660911](#)
40. Sanner MF, Olson AJ, Spehner JC. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*. 1996; 38(3):305–20. [https://doi.org/10.1002/\(SICI\)1097-0282\(199603\)38:3%3C305::AID-BIP4%3E3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-0282(199603)38:3%3C305::AID-BIP4%3E3.0.CO;2-Y) PMID: [8906967](#)
41. Pires DE, Ascher DB. CSM-lig: a web server for assessing and comparing protein-small molecule affinities. *Nucleic Acids Res*. 2016; 44(W1):W557–61. <https://doi.org/10.1093/nar/gkw390> PMID: [27151202](#)
42. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng*. 1995; 8(2):127–34. <https://doi.org/10.1093/protein/8.2.127> PMID: [7630882](#)
43. Dimitropoulos D, Ionides J, Henrick K. Using MSDchem to search the PDB ligand dictionary. *Curr Protoc Bioinformatics*. 2006; Chapter 14:Unit14 3. <https://doi.org/10.1002/0471250953.bi1403s15> PMID: [18428761](#)
44. Ramsbottom KA, Carr DF, Jones AR, Rigden DJ. Critical assessment of approaches for molecular docking to elucidate associations of HLA alleles with adverse drug reactions. *Mol Immunol*. 2018; 101:488–99. <https://doi.org/10.1016/j.molimm.2018.08.003> PMID: [30125869](#)