
Structural Bioinformatics

pyHVis3D: Visualising Molecular Simulation deduced H-bond networks in 3D: Application to T-cell receptor interactions

Bernhard Knapp^{1,2*}, Marta Alcala², Hao Zhang³, Clare West¹, P. Anton van der Merwe³, Charlotte M. Deane¹

¹Department of Statistics, University of Oxford, UK, ²Department of Basic Sciences, International University of Catalonia, Spain, ³Sir William Dunn School of Pathology, University of Oxford, UK.

*To whom correspondence should be addressed.

Associate Editor: XXXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: Hydrogen bonds (H-bonds) play an essential role for many molecular interactions but are also often transient, making visualizing them in a flexible system challenging.

Results: We provide pyHVis3D which allows for an easy to interpret 3D visualisation of H-bonds resulting from molecular simulations. We demonstrate the power of pyHVis3D by using it to explain the changes in experimentally measured binding affinities for three T-cell receptor / peptide / MHC complexes and mutants of each of these complexes.

Availability: pyHVis3D can be downloaded for free from <http://opig.stats.ox.ac.uk/resources>

Contact: science.bernhard.knapp@gmail.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Hydrogen bonds (H-bonds) are non-covalent interactions (5 to 30 kJ/mol) that play an essential role in the stabilisation of protein, DNA, and RNA structures as well as interfaces between them. An H-bond is an electrostatic attraction occurring between a hydrogen atom (H) bound to an electronegative atom (donor) and another electronegative atom (acceptor). Distance and angle constraints are commonly used to determine the presence of an H-bond (e.g. DSSP (Kabsch and Sander, 1983) or GROMACS (Pronk et al., 2013)). For example, a frequently used distance cut-off between acceptor and donor is $< 3.5 \text{ \AA}$ and an angle between hydrogen – donor – acceptor $< 30^\circ$ (Pronk et al., 2013).

As proteins are dynamic structures the existence of an H-bond can be transient and is not a binary decision. Computational techniques such as Molecular Dynamics (MD) simulations can now provide us with dynamic views of biological systems but visualising H-bond information from these techniques is currently not straightforward.

Here we present pyHVis3D, a convenient way to illustrate H-bonds deduced from molecular simulations. Our approach allows the visualisation of individual trajectories as well as the comparison of mutants. We illustrate the usability of our approach by explaining experimentally

measured binding affinities in the T-cell receptor (TCR), peptide, and Major Histocompatibility Complexes (MHC) interface.

2 Methods

Any simulation package can be used to create all atom molecular simulations of the system of interest. The output trajectory format needs to be *.xtc, or *.pdb.

We have implemented a Python 3 based software package to calculate pair-wise H-bonds between all atoms of all frames of the simulation trajectory. A grid-based algorithm calculates an $n \times n$ matrix where n is the number of donor/acceptor atoms of the simulation and each matrix element contains the average presence of an H-bond between two atoms over time. The matrix can be compressed on a per residue basis to save memory and inflated for all atoms for representation purposes. Alternatively only user specified chains and/or residues can be analysed.

If the aim is to compare two simulations then the above procedure is carried out for both trajectories and a difference matrix can be calculated. For each matrix element above a user specified threshold a 3-dimensional cylinder is drawn by Python's matplotlib. The diameter of the cylinder represents the presence of an H-bond over all frames i.e. a

value of 0 Å means that an H-bond is never present and is not visualised and 1 Å means that an H-bond is present in 100% of all frames. The value can exceed 1 Å if visualisation is done on a per residue instead of per atom basis and two amino acids have on average more than one H-bond between them.

In the case of a difference matrix, a blue cylinder indicates more H-bonds for simulation one (e.g. wild-type) while a red cylinder indicates more H-bonds for the other simulation (e.g. mutant). The relative difference between wild-type and mutant is illustrated i.e. the cylinder can only be red or blue but not both. The radius of the cylinder is proportional to the amount of difference. The radius is 0 Å if an H-bond is equally present in the wild-type and mutation simulation and the radius is 1 Å if one H-bond is present 100% of the time in one of the sets but 0% in the other set.

In addition a heatmap of the H-bond matrix is shown and a text file containing numerical values (including a significance test; see SI) about each H-bond is given. For high quality 3D representations a VMD (Humphrey et al., 1996) readable file containing 3D plotting commands is created.

PyHVis3D can be run via an easy-to-use graphical user interface (GUI) or directly via python commands to allow batch processing.

3 H-bonds in the TCR/peptide/MHC interface

The interaction between T-cells and antigen presenting cells is a crucial process in the human immune system. Antigen presenting cells use MHCs to present fragments of potentially harmful proteins on their cell surface. These peptide/MHC (pMHC) complexes are scanned and bound by TCRs. Depending on the combination of TCR, peptide, and MHC (TCRpMHC) an immune response against a peptide can take place. The binding between pMHC and TCR is of low affinity (K_D 1-100 μ M) and somewhat degenerate in that one TCR can recognise multiple pMHCs and one pMHC can be recognised by multiple TCRs.

Here we used pyHVis3D to understand the effects of single point mutations in the MHC class I HLA-A2 on its binding affinity to three different TCRs. We extracted A6/LLFGYPVYV/HLA-A*02:01 (1A07; K_D =0.88 μ M), JM22/GILGFVFTL/HLA-A*02:01 (1OGA; K_D =5.29 μ M), and 1G4/SLLMWITQC/HLA-A*02:01 (2BNR; K_D =3.19 μ M) from the Protein Data Bank (PDB). We computationally introduced the mutation R65_{MHC} into the MHC α -chain of all three wild-type complexes using DeepView (Guex and Peitsch, 1997).

The R65_{MHC} mutation leads to a decreased binding affinity for the 1G4 TCR (+1.81 $\Delta\Delta G$ kcal/mol) and A6 TCR (>+3.16 $\Delta\Delta G$ kcal/mol) but to an increased binding affinity for the JM22 TCR (-0.36 $\Delta\Delta G$ kcal/mol) (Zhang et al., 2016). These different effects on different TCRs make 65_{MHC} an interesting case for investigation.

We ran MD simulations for all six complexes including their constant TCR domains (3 TCRpMHC wild-type complexes and the R65_{MHC} mutant of each) for 100 ns each. Ten replicas (identical parameters but different seeds) of each complex were performed. This led to 60 simulations with a total runtime of 6 000 ns. The detailed simulation protocol is described in the supplementary material.

In JM22 the R65_{MHC} mutation increases the binding affinity ($\Delta\Delta G$ -0.36 kcal/mol) (Zhang et al., 2016). It is not clear why changing a positively charged side-chain that is not in contact with the peptide and not closely packed (Fig. 1A) to an alanine increases the binding affinity between pMHC and TCR.

Using pyHVis3D to analyse the H-bond network of the wild-type and the mutant allows us to shed light on this question. R65_{MHC} binds unstably to several different residues of CDR2 β (N55-K59; blue hedgehog-like structure in the inset of Fig. 1B) and thereby destabilises the CDR2 β

loop increasing the entropy and lowering the binding affinity between TCR and MHC. In the R65A mutant this destabilisation is not present allowing Q52_{CDR2 β} to act as stabilisation centre (Fig. 1C).

The disruptive interaction between R65_{MHC} and CDR2 β is not present for the 1G4 TCR (Fig. S1). Here a stable H-bond is formed between R65_{MHC} and D55_{CDR2 β} . This H-bond is not possible in the A65_{MHC} mutant explaining the lower binding affinity of the R65_{MHC} mutant ($\Delta\Delta G$ +1.81 kcal/mol).

In the case of the A6 TCR the wild-type R65_{MHC} also forms a stable H-bond. In contrast to the 1G4 TCR, however, the H-bond is mainly formed between R65_{MHC} and D93_{CDR3 α} (Fig. S2). In the R65_{MHC} mutant this H-bond to D93 is not possible explaining the lower binding affinity of the R65A mutant ($\Delta\Delta G$ +3.16 kcal/mol).

Our computationally deduced H-bond networks visualised by pyHVis3D allows interesting insights as to why the mutation of MHC residue 65 from Arginine to Alanine increases the binding affinity to the JM22 TCR but decreases the binding affinity to 1G4 and A6. A full research article investigating the wild-type TCRs experimentally and computationally can be found in (Zhang et al., 2016).

Alternative allosteric network interaction analysers are discussed in (Stolzenberg et al., 2016). For example (Stolzenberg et al., 2015) was applied for MHC class II states (Wieczorek et al., 2016).

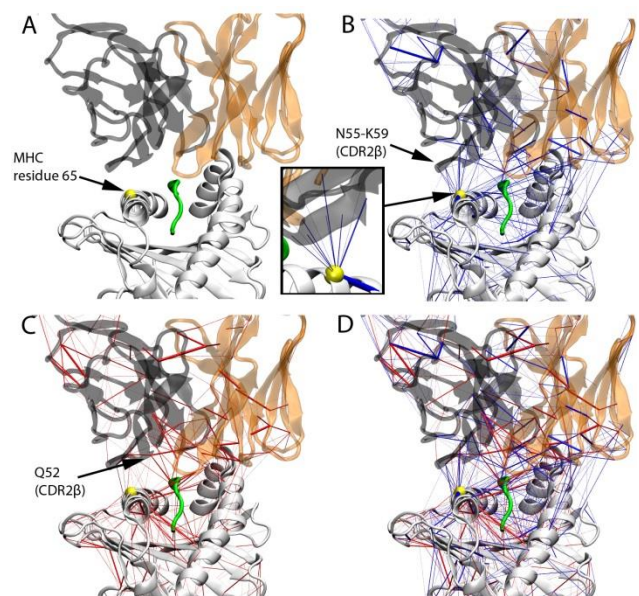


Fig 1: Visualisation of the JM22 H-bond network in reaction to the R65_{MHC} mutation. (A) Spatial arrangement between MHC (white), peptide (green), TCR α -chain (transparent orange), TCR β -chain (transparent black), and MHC residue 65 (yellow sphere). (B) H-bonds that are more frequently present in the wild-type simulation. The inset shows a magnified and rotated view of the hedgehog-like H-bond pattern between R65_{MHC} and CDR2 β (C) H-bonds that are more frequently present in the R65A mutant. (D) Overlay of (B) and (C).

Funding

EPSRC (EP/I017909/1), EPSRC&MRC CDT (EP/L016044/1), MRC Program Grant to AM (G9722488). The authors acknowledge the use of the Oxford Advanced Research Computing (ARC) facility. *Conflict of Interest:* none declared.

References

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 18, 2714-2723.

- Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22, 2577-2637.
- Pronk,S. et al. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *BIOINFORMATICS.*, 29, 845-854.
- Stolzenberg,S. et al. (2016) Computational approaches to detect allosteric pathways in transmembrane molecular machines. *Biochim. Biophys. Acta*, 1858, 1652-1662.
- Stolzenberg,S. et al. (2015) Mechanism of the Association between Na⁺ Binding and Conformations at the Intracellular Gate in Neurotransmitter:Sodium Symporters. *J. Biol. Chem.*, 290, 13992-14003.
- Wieczorek,M. et al. (2016) MHC class II complexes sample intermediate states along the peptide exchange pathway. *Nat. Commun.*, 7, 13224.
- Zhang,H. et al. (2016) The contribution of major histocompatibility complex contacts to the affinity and kinetics of T cell receptor binding. *Sci. Rep.*, 6, 35326.