A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/89114

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk
Towards understanding chemical perception and selectivity for TIR1 and AFB5 auxin F-Box receptor family members in Arabidopsis

by

Mussa Quareshy

Thesis

Submitted to the University of Warwick
For partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Midlands Integrative Bioscience Training Partnership

September 2016
Table of Contents

Table of Contents ..............................................................................................................i
List of Figures..................................................................................................................vi
List of Tables....................................................................................................................xi
List of Schemes ................................................................................................................xi
Acknowledgements ......................................................................................................... xiii
Dedications ....................................................................................................................... xv
Declarations ...................................................................................................................... xvii
Abstract .......................................................................................................................... xix
Abbreviations ................................................................................................................... xxi

Chapter 1: Introduction ..................................................................................................... 1

1.1 Plant hormones .......................................................................................................... 3
1.2 Auxin/Indole-3-acetic acid (IAA) - the master regulator ............................................. 5
1.3 Other auxins .............................................................................................................. 6
1.4 Auxin homeostasis – regulating the master regulator ............................................... 7
1.5 Auxin transport – How does it get around? ............................................................... 8
1.6 Auxin receptors – where does IAA bind? ................................................................. 10
  1.6.1 F-box proteins - TIR1 ......................................................................................... 10
  1.6.2 ABP1 receptor .................................................................................................... 17
  1.6.3 SPK2a as a receptor ........................................................................................... 17
1.7 Kaethner’s auxin-binding site model ......................................................................... 19
1.8 IAA molecular interactions with the TIR1 binding pocket ...................................... 21
1.9 Synthetic auxins as herbicides .................................................................................. 23
1.10 Structure activity relationship studies of auxin-like compounds ............................ 26
1.10.1 Prior structure activity relationship studies ................................................. 26
1.11 Project approach .................................................................................................. 30
  1.11.1 Detailed SAR analysis and assay selection .................................................. 30
  1.11.2 \textit{In silico} docking approach .................................................................. 31
  1.11.3 Pharmacophore generation ...................................................................... 32
1.12 Project motivation ............................................................................................... 33
1.13 Project aims ......................................................................................................... 34

\textbf{Chapter 2: Methods and Materials} .................................................................. 35

2.1 Insect cell culture and protein expression ............................................................ 37
  2.1.1 Adherent and suspension insect cell cultures ............................................. 37
  2.1.2 Generation of recombinant baculoviruses ................................................. 37
  2.1.3 SDS-PAGE protocol ............................................................................... 43
  2.1.4 Western blot development ...................................................................... 44
  2.1.5 MOI titration estimation adherent cultures – with known virus titre ........ 46
  2.1.6 MOI titration estimation shaking cultures with known virus titre .......... 46
  2.1.7 Optimal protein expression time using virus to cell volume for baculovirus with unknown titre ................................................................. 46
  2.1.8 Large scale insect cell baculovirus induced expression ............................. 46
2.2 TEV protease expression and purification ............................................................ 47
  2.2.1 Expression ............................................................................................... 47
  2.2.2 Purification ............................................................................................... 47
2.3 F-box protein purification .................................................................................... 48
  2.3.1 Cell lysis and protein extraction ................................................................ 48
  2.3.2 Tandem Affinity Protein Purification ....................................................... 48
2.4 Small molecule library set up .............................................................................. 49
2.5 General root growth assay methodology ............................................................ 49
  2.5.1 Seed bulking ............................................................................................ 49
  2.5.2 Seedling germination ............................................................................... 50
  2.5.3 Seedling treatment with compounds ....................................................... 52
  2.5.4 Primary root measurements .................................................................... 54
  2.5.5 Root growth assay IC$_{50}$ determination .................................................. 56
  2.5.6 Lateral root density measurements ......................................................... 56
2.6 GFP confocal microscopy .................................................................................... 57
2.7 Pull down assays........................................................................................................57
2.8 Pharmacophore generation .......................................................................................57
2.9 Synthesis of iMQ (3-[(1H-1,2,3,4-tetrazol-5-yl)methyl]-1H-indole) ....................58
2.10 SPR methodology ....................................................................................................59
  2.10.1 Introduction ........................................................................................................59
  2.10.2 Binding interactions observed by SPR ...............................................................61
  2.10.3 Ligand immobilization .......................................................................................63
2.11 Auxin activity screening ..........................................................................................68
  2.11.1 Assay run order and sampling ..........................................................................68
  2.11.2 Sample preparation ...........................................................................................68
  2.11.3 Basis of the activity run ......................................................................................69
  2.11.4 SPR auxin activity assay reference subtraction ................................................69
2.12 Anti-auxin activity ..................................................................................................75

Chapter 3: IAA binding to TIR1 ..................................................................................77
  3.1 Introduction ...............................................................................................................79
   3.1.1 Techniques for studying small molecule binding ...............................................80
  3.2 Recombinant TIR1-ASK1 expression and purification ...........................................81
   3.2.1 Cell culture ........................................................................................................81
   3.2.2 Virus to cell ratio determination .......................................................................82
   3.2.3 Tandem Affinity Protein Purification ...............................................................83
  3.3 Observing direct binding of IAA to TIR1 by SPR ................................................88
   3.3.1 Biacore and introduction to SPR .......................................................................88
   3.3.2 TIR1-ASK1 immobilisation via amine coupling ...............................................88
   3.3.3 Attempts to observe direct binding of IAA to immobilised TIR1 ....................90
   3.3.4 TIR1-ASK1 immobilisation via antibody capture ............................................95
  3.4 Discussion ...............................................................................................................97

Chapter 4: A Novel in silico docking approach postulates mechanism of IAA perception and active compound differentiation for the TIR1 receptor ..........................................................................................................................99
  4.1 In silico docking algorithm .....................................................................................101
  4.2 Validation of AutoDock Vina docking algorithm ................................................101
4.3 The auxin binding site; how does auxin get there? – Development of novel docking application – TomoDock

4.4 TomoDock suggests a mechanism for excluding Tryptophan from the IAA binding site.

4.5 Discussion

Chapter 5: The development of an in vitro spatial auxin pharmacophore model

5.1 Compound selection and SAR library design

5.2 In vitro SPR method

5.3 In vivo root growth assays

5.4 SPR screening SAR summary and in vitro pharmacophore mapping

5.5 Conclusions

Chapter 6: Simply selective: A tetrazole analogue of Indole-3-acetic acid is an auxin and shows selectivity between SCF$^{\text{TIR1/AFB5}}$-Aux/IAA Co-receptors: TIR1 and AFB5

6.1 In silico design and validation of IAA tetrazole analogue

6.2 iMQ synthesis

6.3 iMQ activity assays
6.4 iMQ does not bind to AFB5 ................................................................. 191
6.5 iMQ causes weak agravitropism in primary roots and laterals .............. 193
6.6 iMQ possibly interferes with auxin uptake ........................................... 199
6.7 Discussion .......................................................................................... 202

Chapter 7: General discussion and conclusion ...................................... 205
  7.1 Observations of IAA binding to TIR1 with in silico docking .................. 207
  7.2 Pharmacophore mapping ..................................................................... 209
  7.3 Design of non-canonical auxins and receptor specificity ..................... 210
    7.3.1 Lead optimisation of iMQ .............................................................. 211
  7.4 Conclusion .......................................................................................... 212

References ................................................................................................. 213

Appendix A: SPR Methodology publication ........................................... 225
Appendix B: TomoDock methodology publication ................................. 251
Appendix C: TomoDock Supplementary Information .............................. 265
Appendix D: NMR and Mass spectra ....................................................... 271
List of Figures

Figure 1.1. IAA chemical structure.................................................................5
Figure 1.2 Other endogenous auxins ..........................................................6
Figure 1.3: Schematic representation of the cascade of events during auxin perception
by TIR1 ........................................................................................................11
Figure 1.4. The TIR1+ASK1 crystal structure..............................................13
Figure 1.5. Top view of TIR1 binding site ....................................................14
Figure 1.6. Phylogeny relationship between TIR1 and AFB homologues .........16
Figure 1.7 Kaethner’s auxin binding site model...........................................20
Figure 1.8. Observing Kaethner’s predictions in the auxin TIR1 crystal structure .....20
Figure 1.9. 2D bimolecular interactions between IAA and the co-receptor complex .22
Figure 1.10. 3D view of interactions between IAA and the co-receptor complex .....22
Figure 1.11 Currently available commercial auxinic herbicides ......................25
Figure 1.12 The Katekar auxin receptor binding model...............................27
Figure 2.1. TIR1 transfer vector ................................................................38
Figure 2.2. AFB5 transfer vector ..................................................................40
Figure 2.3. Western blot gel-to-membrane transfer step ................................44
Figure 2.4. Root growth assay seeds spotting and germination on media ..........51
Figure 2.5. Transferred seedlings on treated media, with markings for initial primary
root length ....................................................................................................53
Figure 2.6. On screen root growth measurements ........................................55
Figure 2.7. Basis of the SPR phenomenon .................................................60
Figure 2.8. Impact on the SPR angle with changes in mass at the surface ........61
Figure 2.9. SPR binding model on sensor surface ......................................62
Figure 2.10. Basic SPR experiment ............................................................62
Figure 2.11. Example of SPR ligand capture options ....................................64
Figure 2.12. CM5 Biacore chip schematics depicting carboxymethylated dextran
surface and electrostatic attraction surface conditions................................66
Figure 2.13. Carboxylic acid group activation to an NHS-ester ....................67
Figure 2.14. Overview of ligand immobilization on activated chip surface........67
Figure 2.15. SPR results processing ............................................................70
Figure 2.16. KNIME workflow for processing and clean up of SAR activity data....73
Figure 2.17. Control runs for the anti-auxin activity ................................................. 75
Figure 2.18. Outcomes of the anti-auxin assay............................................................ 76
Figure 3.1. Optimisation of TIR1 expression in T. ni Hi5 cells................................. 82
Figure 3.2. Time course expression profile for TIR1-ASK1 in insect cells ............. 83
Figure 3.3. Phase I of TIR1/ASK1 purification on a His affinity column.............. 85
Figure 3.4. Phase II of TIR1-ASK1 purification; TEV protease digest and FLAG
affinity column loading and washes ................................................................. 86
Figure 3.5. Phase III of TIR1-ASK1 purification with FLAG column elutions of
purified TIR1-ASK1 ............................................................................................. 87
Figure 3.6. pH scouting on CM5 chip with TIR1-ASK1 protein.............................. 89
Figure 3.7. TIR1-ASK1 amine coupling onto CM5 chip ........................................ 91
Figure 3.8. Single cycle kinetics for IAA vs. TIR1-ASK1 ....................................... 92
Figure 3.9. Baseline correction for Figure 3.8 ......................................................... 92
Figure 3.10. Repeated single cycle kinetics for IAA binding to TIR1-ASK1 .......... 94
Figure 3.11. Drifting baseline of TIR1-ASK1 post amine coupling......................... 94
Figure 3.12. Amine coupling of TIR1-ASK1 to CM7 chip ...................................... 95
Figure 3.13. Amine coupling capture of anti-FLAG antibody............................... 96
Figure 3.14. TIR1-ASK1 capture onto anti-FLAG antibody .................................. 97
Figure 4.1. Results of AutoDock Vina docking of IAA into TIR1. ......................... 103
Figure 4.2. Correlation plot comparing Docking vs. SPR rankings...................... 104
Figure 4.3. – The best-docked poses for IAA and tryptophan............................ 105
Figure 4.4 TomoDock grid box movement............................................................... 107
Figure 4.5. TomoDock binding transect for TIR1 .................................................... 109
Figure 4.6. Postulated mechanism of IAA movement into TIR1 binding pocket.... 111
Figure 4.7. Final 2 transitions of IAA into the binding pocket................................ 112
Figure 4.8 Docking energy profile of selected poses ............................................. 114
Figure 4.9. Rationale selection of poses from TomoDock experiments............... 115
Figure 4.10 LigPlots depicting molecular interactions between TIR1 and docked IAA
in slabs 2 & 3 ........................................................................................................ 116
Figure 4.11 LigPlots depicting molecular interactions between TIR1 and docked IAA
in slabs 4, 7, 8 and 9 .............................................................................................. 117
Figure 4.12 TomoDock results for tryptophan indicate a forbidden transition .... 119
Figure 4.13. IAA in relation to the hydrophobicity of the pocket when bound...... 122
Figure 5.1 The main focus scaffolds for our SAR compound library. .................. 126
Figure 5.2. Chronology of positive control RU levels in auxin screening assays. .... 129
Figure 5.3. Inter- an Intra-assay relative auxin activity data. ............................ 130
Figure 5.4. Indole scaffold SAR. .................................................................. 132
Figure 5.5 Relative SPR activities of the Indole SAR ..................................... 133
Figure 5.6. SAR SPR screening of Indole α-position antiauxins and the β-position
amino acid conjugates .............................................................................. 134
Figure 5.7. Evolution of Benzoic acid to 2,4-D and alkyl chain length at α-position.
.............................................................................................................. 135
Figure 5.8. SPR results of the evolution of Benzoic acid to 2,4-D and alkyl chain
length at α-position. ................................................................................. 136
Figure 5.9. SAR single step substitutions and analogues of 2,4-D..................... 139
Figure 5.10 SAR activities for 2,4-D single step substitutions ........................... 140
Figure 5.11 Benzoic acid SAR compounds .................................................... 142
Figure 5.12 Benzoic acid analogue auxin activities ......................................... 142
Figure 5.13 Napthalic acid SAR compounds .................................................. 143
Figure 5.14. Napthalic acid SAR compounds .................................................. 143
Figure 5.15 Auxin activities of napthalic acid scaffold derivatives .................... 143
Figure 5.16. Dose response profile for primary root growth measurements with
controls ....................................................................................................... 145
Figure 5.17. Normalised primary root growth inhibition profile ........................ 145
Figure 5.18. Impact of top 5 DMSO concentrations root growth and lateral root
density compared to untreated roots ........................................................ 146
Figure 5.19 Root growth phenotype observations .......................................... 147
Figure 5.20 Primary root growth dose response: Indole SAR series ............... 151
Figure 5.21 Primary root growth dose response: Phenoxy and benzoic SAR series. 152
Figure 5.22 Crystallographic compound reference structures of IAA, 2,4-D and 1-NAA in the TIR1 receptor .................................................. 154
Figure 5.23. Pharmacophore field points generated from Forge software. ......... 156
Figure 5.24. Alignments of Benzoic acid and Picloram ................................... 158
Figure 5.25 The average shape pharmacophore of active compounds for TIR1 SAR
study ......................................................................................................... 161
Figure 5.26. The activity cliff of shape pharmacophore for TIR1 ....................... 162
Figure 5.27 The inactivity of the 2- and 6- co-occupied positions of phenoxyacetic acid scaffolds versus the occupied active indolic position 4 .......................... 163
Figure 5.28. The hydrophobic pharmacophore for TIR1 ................................................. 166
Figure 5.29 Hydrophobicity pharmacophore in relation to TIR1 binding site residues. ......................................................................................................................... 167
Figure 5.30 The Electrostatic pharmacophore atlas of auxin .............................................. 169
Figure 5.31 A comparison of the electrostatics of indole and phenoxyacetic acid scaffolds. ........................................................................................................................................ 170
Figure 5.32 A comparison of the average shape and hydrophobicity pharmacophore. ........................................................................................................................................... 171
Figure 5.33 Shape and hydrophobicity trends for auxin activity ........................................... 173
Figure 6.1 Available space around the IAA carboxylic acid group in the TIR1 pocket. ................................................................................................................................. 180
Figure 6.2 Side-by-side comparison of IAA to iMQ in the TIR1 pocket ................................. 180
Figure 6.3 In silico docking for IAA and iMQ ................................................................. 182
Figure 6.4. iMQ characterisation data ..................................................................................... 184
Figure 6.5 iMQ activity assays ............................................................................................... 186
Figure 6.6 Whole root in-vivo DR5::GFP reporter assays ................................................. 187
Figure 6.7 Primary root growth inhibition assays ............................................................... 189
Figure 6.8 Lateral root densities for IAA and iMQ treated WT and tir1-1 Col-0 lines ........................................................................................................................................ 189
Figure 6.9. Dose response phenotypes for root growth assays with Col-0 WT................. 190
Figure 6.10. iMQ displays selectivity for TIR1, not AFB5 in SPR assays ......................... 191
Figure 6.11 Binding pocket amino acid sequence alignments, TIR1 vs. AFB5 ............... 192
Figure 6.12. In silico docking shows iMQ cannot bind to AFB5 ........................................ 192
Figure 6.13. Dose dependent primary root agravitropism with iMQ. ................................. 194
Figure 6.14 Root growth assay lighting and plate stacking positions in incubator ... 195
Figure 6.15. Position in the incubator did not affect bending in roots treated with iMQ. ........................................................................................................................................ 196
Figure 6.16. Gravitropic offset angle measurements .......................................................... 197
Figure 6.17. Gravitropic offset angle measurements in Col-0 WT and tir1-1 mutants treated with iMQ ........................................................................................................... 197
Figure 6.18. iMQ treated root gravitropism kinetics ......................................................... 198
Figure 6.19. Accumulation of labelled $[^3H]2,4$-D in AUX-1 accumulation functional assay ................................................................. 200
Figure 6.20. 1-NAA rescue of function assay ................................................................. 200
Figure 6.21 Rescue of bending phenotype with 1-NAA .................................................. 201
Figure 7.1 Lead optimisation candidates for iMQ .......................................................... 211
Figure D.1 Proton NMR of 3-[(1H-1,2,3,4-tetrazol-5-yl)methyl]-1H-indole .............. 273
Figure D.2 Mass spectrum of 3-[(1H-1,2,3,4-tetrazol-5-yl)methyl]-1H-indole ........... 274
List of Tables

Table 1.1 Summary of endogenous hormones and their roles ............................................. 4
Table 1.2 TIR1 crystal structures ......................................................................................... 13
Table 1.3 Percentage pairwise amino acid sequence similarities for the family of
Arabidopsis F-Box receptors ................................................................................................. 16
Table 2.1. Reference subtractions for SPR processing ......................................................... 71
Table 2.2 Reference subtractions for SPR processing ......................................................... 74
Table 4.1 Summary of SPR vs. Docking results for selected compounds ......................... 104
Table 4.2 TomoDock results of top three scoring IAA poses for each slab ................. 108
Table 4.3 TomoDock results of top three best scoring IAA poses for each slab ...... 114
Table 5.1. Primary root growth IC$_{50}$ inhibition and phenotyping for: indole,
phenoxyacetic acid, benzoic acid and napthalic scaffolds ............................................. 150
Table 5.2. SPR activity categorisation ................................................................................. 153
Table 6.1. Physiochemical property comparison between IAA and iMQ .................... 181
Table 6.2 Primary root IC$_{50}$ values for TIR1 knockout line ........................................... 188

List of Schemes

Scheme 6.1 Synthesis of iMQ .............................................................................................. 183

List of Videos (Playable from external media CD)

Video 1.1 ............................................................................................................................. 12
Video 4.1 ............................................................................................................................. 106
Video 4.2 ............................................................................................................................. 110
Acknowledgements

First and foremost, Alhamdulillah. (Praise be to Allah) for granting me the ability to pursue my dreams and embark on this journey.

I would like to thank my supervisor, Richard Napier for his invaluable guidance and tutelage throughout. His abounding encouragement kept me motivated even in the most challenging of times. I would also like to thank my co-supervisor Vilmos Fulop as well for his mentorship, advice and uplifting anecdotes.

It has been a pleasure to work with all members, past and present of the Napier and C10 structural biology labs. In particular, I’d like to thank my chaperon post-docs, Veselina Uzunova and Justyna Prusinska, for putting up with my antics, keeping me on track and helping with my thesis. A special thank you to Dr Charo Del Gino for his patient tutoring and explanations of mathematical, physical and culinary topics; especially the home made pizzas. Dr Andrew Marsh for his mentorship and assistance with synthesis. I would like to thank my advisory panel: Dr Stephen Jackson, Prof. David Hodgson and Dr Corinne Smith for their guidance, encouragement and support. A big thank you to Dr Ruth Schafer for assisting with establishing the root growth assays and an extended thank you to the remainder of the Schafer group for use of their incubator. It has also been a pleasure to collaborate with Dr Stephan Kepinski’s group, including Dr Martin Kieffer, Sigurd Ramans Harborough and Ryan Kaye. I also thank my fellow peers on the 2012 MIBTP DTP cohort, the Life Science support staff and members of the C10, C30 and phytobiology labs.

I am grateful to my parents and siblings; Nazrana, Haseena, Muhammad & Tasneem for their encouragement, patience and support. Not forgetting my [tolerable] brother-in-laws Mohsin and Yusuf either. The zimbos, Sumaiyah and Huz[tr]ai[tor]fah, I’m now “educated”. My beautiful niece Aisha, who has provided endless hours of [joyful] distraction and fun; you have all kept my spirits high and helped me throughout this journey.
Special thank you to: My housemate Karim-El-Halouï, for being a fantastic friend, the elder brother I never had and putting up with me for three and half years! Daniel Wilson-Nunn; a buddy [that I owe a lot of petrol money to], a shoulder to cry on and always there when the going got tough! To my wingmen and mentors Muzaffar Makhdoomi and Arsheed Sheikh. To Alvin Teo, for helping to proof read my work and inspiring scientific discussions; particularly on mini-preps! A heartfelt thanks you to Piraveen Gopalasingam for mentoring me during my mini project and making our reign of distraction in the C146 office a memorable experience. Tauseef Ahmed; for the many hours of gossip and trash talking, I became a doctor before you though! A special thank you to Tauseef Anjum for his everlasting friendship and hospitality, from Malawi all the way to Canada. To the Gibbet Hill Jamah crew: Jehad, Okba, Syed, Madikay, Sulaiman, and Ahmad for getting this printed! :P
Dedications

The thesis is dedicated to my late grandparents:

Mussa Siddique and Jenum Quareshy

Omar and Amina Mohammad
Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. The research described in this thesis has been carried out by myself under the supervision of Prof. Richard Napier. Furthermore it has been composed by myself and has not been submitted in any previous application for any degree.

We declare the contributed work by colleagues at the University of Leeds, Dr Martin Kieffer presented in Chapter 6 as Figure 6.5B and Ryan Kaye, presented in section Chapter 6 as Figure 6.18 as part of collaborations with an intention to co-publish.

We further declare the contributions from colleagues within the group at the University of Warwick: Dr Charo Gel Gino for assisting with the data fitting methodology in Section 2.5.5, and Prof. Richard Napier for work presented in Chapter 6 as Figure 6.19.

The author has published parts of the thesis:


Veselina V. Uzunova, Mussa Quareshy, Charo I. del Genio, Richard M. Napier Tomographic docking suggests the mechanism of auxin receptor TIR1 selectivity Open Biol. 2016 6 160139; DOI: 10.1098/rsob.160139. Published 19 October 2016 (Appendix B and supplementary information in Appendix C)


Abstract

Auxin (Indole-3-acetic acid; IAA) can be considered one of the most important hormones in plant development as it coordinates plant development through transcriptional regulation. For years the principal auxin receptor was sought and only relatively recently was it identified as an F-box protein known as TIR1, with five homologues; the auxin F-Box proteins (AFB1–5). The elucidation of a crystallographic structure elegantly demonstrated IAA’s integral role in forming a ternary complex between auxin F-box proteins and Aux/IAA proteins, revealing the mechanism of auxin perception. An exact chemical description of an auxin has been pursued for decades and researchers have always faced the challenge of factoring in the complex nature of multiple auxin-receptor classes, transport and metabolism.

My project aimed to develop a better understanding of auxin chemical specificity at the receptor level, focusing on TIR1 and one of its most distantly related homologues, AFB5. We employed a structure activity relationship study with a rational selection of compounds and have defined a pharmacophore for auxin activity for the TIR1 receptor. The thesis also describes a receptor-led, in silico rational drug design approach in which we replaced the carboxylic acid moiety with a tetrazole, giving a novel compound that demonstrated auxin-like activity both in vitro and in vivo. Furthermore this bioisosteric replacement serendipitously demonstrated a novel selectivity for TIR1, with no activity against AFB5. Preliminary in silico docking studies of the TIR1 binding site could not discern between active and in active compounds generating many false positive results, leading us to develop TomoDock; a novel in silico docking approach to study the entire TIR1 receptor binding pocket geometries. Results from TomoDock suggest that binding is a two phase process with active ligands firstly engaging with a niche, which orients them, allowing passage past and molecular filter region before interacting with the binding site. This process contributes to the mechanism of compound selection by the TIR1 pocket receptor.

A range of multidisciplinary approaches utilised in this project have allowed us to investigate and report many new insights on the mechanism of auxin perception. Such rational approaches may also be used in other drug discovery programmes alike to help researchers discern compound perception and selectivity.
## Abbreviations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>Angstrom ($10^{-10}$ meters)</td>
</tr>
<tr>
<td>Å. Thaliana</td>
<td>Arabidopsis Thaliana</td>
</tr>
<tr>
<td>ABP1</td>
<td>Auxin binding protein</td>
</tr>
<tr>
<td>AFB</td>
<td>Auxin F-Box</td>
</tr>
<tr>
<td>Ala</td>
<td>Alanine</td>
</tr>
<tr>
<td>ARF</td>
<td>Auxin response factor</td>
</tr>
<tr>
<td>Arg</td>
<td>Arginine</td>
</tr>
<tr>
<td>Asn</td>
<td>Asparagine</td>
</tr>
<tr>
<td>Asp</td>
<td>Aspartic acid</td>
</tr>
<tr>
<td>Aux/IAA</td>
<td>Auxin-IAA proteins</td>
</tr>
<tr>
<td>cryo-EM</td>
<td>Cryo-Electron Microscopy</td>
</tr>
<tr>
<td>Cys</td>
<td>Cysteine</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DTT</td>
<td>DL-Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EtOAc</td>
<td>Ethyl acetate</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>Gln</td>
<td>Glutamine</td>
</tr>
<tr>
<td>Glu</td>
<td>Glutamic acid</td>
</tr>
<tr>
<td>Gly</td>
<td>Glycine</td>
</tr>
<tr>
<td>HBS</td>
<td>Heps buffered saline</td>
</tr>
<tr>
<td>His</td>
<td>Histidine</td>
</tr>
<tr>
<td>IAA</td>
<td>Indole-3-acetic acid</td>
</tr>
<tr>
<td>IBA</td>
<td>Indole-3-butyric acid</td>
</tr>
<tr>
<td>Ile</td>
<td>Isoleucine</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilised metal affinity chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-thiogalactoside</td>
</tr>
<tr>
<td>IUPAC</td>
<td>International Union of Pure and Applied Chemistry</td>
</tr>
<tr>
<td>Leu</td>
<td>Leucine</td>
</tr>
<tr>
<td>Lys</td>
<td>Lysine</td>
</tr>
<tr>
<td>MBP</td>
<td>Maltose Binding Protein</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>Met</td>
<td>Methionine</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>MOI</td>
<td>Multiplicity of Infection</td>
</tr>
<tr>
<td>MS</td>
<td>Murashige &amp; Skoog</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>Optical Density at 600 nm</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>Pro</td>
<td>Proline</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>rpm</td>
<td>Rotations per minute</td>
</tr>
<tr>
<td>RU</td>
<td>Response Units</td>
</tr>
<tr>
<td>SAR</td>
<td>Structure Activity Relationship</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl sulphate</td>
</tr>
<tr>
<td>sec</td>
<td>Second</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris Buffered Saline</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine hydrochloride</td>
</tr>
<tr>
<td>TEV</td>
<td>Tobacco Etch Virus</td>
</tr>
<tr>
<td>THF</td>
<td>Tetrahydrofuran</td>
</tr>
<tr>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>TIF</td>
<td>Tagged Image File Format</td>
</tr>
<tr>
<td>TIR1</td>
<td>Transport Inhibitor Resistant 1</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>Trp</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>Val</td>
<td>Valine</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
<tr>
<td>μM</td>
<td>Micromolar</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction
1.1 Plant hormones

Despite the sessile and stationary disposition of plants, they are constantly sensing and adapting to their environment. They are able to respond to an array of responses including changes in light, water availability, nutritional sources and pathogen attack. Additional subtle adaptations at the root level, not obvious from aboveground, add to the complex sensing capabilities of plants; thus indicating their willingness to adapt or perish. The question then becomes how are they able to achieve this? The answer is partly down to the extensive hormone signalling pathways that define plant ontogeny (Depuydt & Hardtke, 2011; Gallavotti, 2013).

Plants produce an array of hormones (also known as phytohormones) with essential roles in growth and development and which control responses to environmental stresses such as drought, herbivory- and pathogen-attack (Westfall et al., 2013) and growth. Plant hormones are a group of naturally occurring, organic substances that are generally present at low concentrations and do not have any nutritional function (Davies, 2004; Woodward & Bartel, 2005a; Sauer et al., 2013). A summary of key plant phytohormones classes is shown in Table 1.

Each phytohormone class imposes its influence independently, however it is also known that there are overlapping, integrated and prerequisite dependencies that collectively impact on phases of plant development. Auxin is at the hub of numerous complex responses, but some pathways are non-redundant as shown in the observations on loss-of-function mutants. The overall hormone signalling and cross talk system is not yet fully understood and is a topical theme in current phytohormone research (Santner & Estelle, 2009; Jaillais & Chory, 2010; Depuydt & Hardtke, 2011).
<table>
<thead>
<tr>
<th>Structure</th>
<th>Hormone</th>
<th>Active form</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Auxins" /></td>
<td>Indole-3-acetic acid</td>
<td>Auxins</td>
<td>Cell Elongation, Cell Division, Apical dominance, Vascular differentiation, Tropistic Responses, Fruit ripening, Lateral Rooting, Branching</td>
</tr>
<tr>
<td><img src="image2" alt="Gibberellins" /></td>
<td>Gibberellin A1</td>
<td>Gibberellins</td>
<td>Stem Elongation, Seed Germination, Flowering, Floral Development</td>
</tr>
<tr>
<td><img src="image3" alt="Cytokinins" /></td>
<td>Zeatin</td>
<td>Cytokinins</td>
<td>Cell Division, Growth of Lateral buds, Delays leaf senescence</td>
</tr>
<tr>
<td><img src="image4" alt="Ethylene" /></td>
<td></td>
<td>Ethylene</td>
<td>Fruit Ripening, Seed Germination, Flower Opening, Release from dormancy</td>
</tr>
<tr>
<td><img src="image5" alt="Abscisic Acid" /></td>
<td></td>
<td>Abscisic Acid</td>
<td>Stomatal closure (water balance), Seed maturation, Germination, Seed storage protein synthesis</td>
</tr>
<tr>
<td><img src="image6" alt="Polyamines" /></td>
<td>Spermidine</td>
<td>Polyamines</td>
<td>Plant growth, Cell Division</td>
</tr>
<tr>
<td><img src="image7" alt="Brassinosteroids" /></td>
<td>Brassinolide</td>
<td>Brassinosteroids</td>
<td>Cell Division, Cell elongation, Vascular differentiation, Stress response</td>
</tr>
<tr>
<td><img src="image8" alt="Jasmonates" /></td>
<td>Jasmonic Acid</td>
<td>Jasmonates</td>
<td>Plant defences against insects &amp; other herbivores, Necrotrophic pathogen responses</td>
</tr>
<tr>
<td><img src="image9" alt="Benzoates" /></td>
<td>Salicylic Acid</td>
<td>Benzoates</td>
<td>Systemic acquired resistance, response to pathogens, enhanced flower longevity</td>
</tr>
<tr>
<td><img src="image10" alt="Strigolactones" /></td>
<td>(+)-5-Deoxystrigol</td>
<td>Strigolactones</td>
<td>Branch control, Plant architecture, Root growth</td>
</tr>
</tbody>
</table>

(Davies, 2004; Woodward & Bartel, 2005a; Mockaitis & Estelle, 2008; Hayashi et al., 2012; Ruiz Rosquete et al., 2012; Korasick et al., 2013; Sauer et al., 2013; Westfall et al., 2013)
1.2 Auxin/Indole-3-acetic acid (IAA) - the master regulator

Standing out amongst the phytohormones is auxin. The major endogenous auxin is known as (indole-3-acetic acid; IAA), but both names are used interchangeably in the literature. IAA is the hormone with the most diverse set of roles is considered one of (Table 1.1), thus earning it the name of master regulator of plant growth and development (Mockaitis & Estelle, 2008).

Auxin is derived from the Greek word “auxein” which means to enlarge or grow (Sauer et al., 2013). Early work on plant hormones by Charles Darwin hinted at some form of a signalling molecule that displayed positive phototropism. Only in 1934 did Kögl identify the active low molecular weight molecule - IAA, just 175 Daltons, consisting of an indole ring with an acetic acid group at the C3 position (Figure 1.1), (Kögl et al., 1934).

Figure 1.1. IAA chemical structure
IAA structure with IUPAC (International Union of Pure and Applied Chemistry) numbering convention.
1.3 Other auxins

Other endogenous compounds have been identified to have auxin activity in whole plants (Figure 1.2). 4-Chloro-IAA was discovered in pea plants as an auxin (Porter & Thimann, 1965), but its impact is not yet understood. Indole-3-butyric acid (IBA) (Zimmerman & Wilcoxon, 1935) is considered to be a storage precursor to IAA. In some commercial applications as an exogenous treatment such as for induction of adventitious roots it can be a more efficient auxin than IAA. This is perhaps due to IBA being more stable in vivo against catabolism and inactivation by conjugation. Phenylacetic acid (PAA) (Koepfli et al., 1938) has been found in a variety of plant species and is thought to play a role in root interactions with soil microorganisms (Davies, 2004; Sauer et al., 2013). Despite these other naturally occurring auxin-like compounds, IAA still reigns as the principle form of auxin.

![Chemical structures of 4-Chloro-IAA, IBA, and PAA](image-url)

**Figure 1.2 Other endogenous auxins**
Structures of other naturally occurring auxins.
1.4 Auxin homeostasis – regulating the master regulator

Auxin production is highly regulated in plants and thought to occur predominantly in the aerial parts of plants but it is also synthesised locally elsewhere in the plant (Ruiz Rosquete et al., 2012). Interplay between synthesis and inactivation mechanisms regulates auxin levels in plants with strict control during specified phases of growth, development and external responses.

Auxin biosynthesis can take two possible routes; tryptophan-dependant and independent pathways (Woodward & Bartel, 2005a; Korasick et al., 2013). Tryptophan as a precursor already contains the aromatic indole group of IAA and is subject to a series of enzymatic deamination, decarboxylation, oxidation and hydrolysis steps to yield IAA via intermediary compounds such as indole-3-acetaldehyde, indoleacetimide, and indole-3-pyruvic acid. Such is the prevalence of tryptophan that its pool size is three or more orders of magnitude greater than IAA. In addition, IBA which is an endogenous auxin precursor can undergo β-oxidation to yield IAA (Woodward & Bartel, 2005a; Ruiz Rosquete et al., 2012; Sauer et al., 2013).

Tryptophan independent pathways can also generate IAA, primarily from indole as a precursor. This phenomenon has been proven in the absence of key tryptophan-dependent enzymes, but a complete understanding of the mechanisms and intermediaries is yet to be elucidated (Ruiz Rosquete et al., 2012; Korasick et al., 2013).

Auxin inactivation is regulated through a series of mechanisms designed to be either reversible, or to permanently inactivate the hormone. Ester conjugation to glucose transforms IAA into a readily hydrolysable storage form. Amide conjugations to amino acids like Alanine/ Leucine/ Aspartic acid/ Glutamine/ Glycine/ Glutamic acid/ phenylalanine and tryptophan have been reported in a variety of roles ranging from storage, inactivation, and compartmentalisation. Auxin is also catabolised via oxidative degradation and can also be re-transformed into IBA, back to its precursor storage form (Korasick et al., 2013; Sauer et al., 2013; Enders & Strader, 2015).
1.5 Auxin transport – How does it get around?

Auxin is mainly synthesised in the aerial parts of the plants, but also in the meristematic region of the primary root. Auxin is transported via two main pathways; [i] the quickest, is via non-polar, passive, bulk transport route in the phloem and [ii] a slower, energy dependent cell-to-cell, polar transport method. Non-polar auxin transport progresses at 5 – 20cm/h with polar auxin transport at 5-20mm/h (Michniewicz et al., 2007). Polar auxin transport covers both long and short distances with the same underlying mechanisms for getting auxin into and out of cells.

The chemiosmotic model of transport describes polar auxin influx and efflux as well as the polarisation of these mechanisms to one face of the cell, to drive auxin transport in a unidirectional manner (Michniewicz et al., 2007). In the acidic pH 5.5 environment of the cell wall matrix, given that the pKa of IAA is calculated at 4.85 (Zazímalová et al., 2010), around 84% of the IAA is protonated as IAAH. In this uncharged state, IAA may diffuse passively across the plasma membrane into cells. Additional auxin uptake is facilitated by membrane localised Auxin Resistant 1 (AUX1) and Like AUX1 (LAX1) H⁺ symporters (Blakeslee et al., 2005; Sauer et al., 2013).

Once inside the cell at pH 7.0 – 7.5, IAA dissociates into an anionic IAA⁻ species and is confined inside the cell by this anion trap. Special efflux carriers identified as PIN-FORMED (PIN) family of proteins, driven by ATP, actively transport IAA⁻ out of the cell (Křeček et al., 2009), but the exact mechanism of this process is as yet undefined (Blakeslee et al., 2005). The PIN proteins localise at the plasma membrane (PM) and undergo rapid and constitutive cycling between the plasma membrane and endosomes. Auxin inhibits endocytosis which consequentially inhibits the internalization step of PIN constitutive cycling and increases the level of PINs at the PM (Paciorek et al., 2005). PIN localisation is regulated by phosphorylation with PINOID kinases apparently promoting apical localisation and basal localisation is seemingly determined by protein phosphatase 2A (Tromas & Perrot-Rechenmann, 2010).

The ATP-binding Cassette transporter subfamily B (ABCB) transporter proteins, also known as multidrug resistant proteins or P-glycoproteins (MDR-PGPs) (Verrier et al.,
have been reported to facilitate long distance auxin transport mechanisms (Reemmer & Murphy, 2014). ABCB19 has been shown to be involved in auxin efflux with the \textit{abcb19} knockout line displaying a similar phenotype to that of suboptimal auxin responses and/or transport impedance. As they have not been as well characterised as the PIN protein family, much more work is required to determine the role and mechanism of action for this efflux transporter class.
1.6 Auxin receptors – where does IAA bind?

1.6.1 F-box proteins - TIR1

The best-studied auxin receptor is the F-box protein Transport Inhibitor Resistant 1 (TIR1). It was discovered in genetic analysis studies (Ruegger et al., 1998) with mutations to the TIR1 gene, yielding a tir1-1 mutant line which displayed sensitivity to IAA, but was not impacted by auxin transport inhibitors. Thereafter it was known that auxin mediated a response through an Skp, Cullin, F-box (SCF) complex which is a type E3 ubiquitin ligase involving TIR1 as the F-box protein. It wasn’t until seminal work by two groups (Dharmasiri et al., 2005a) and (Kepinski & Leyser, 2005) who both independently showed that TIR1 was the major auxin receptor (Napier, 2014). They showed IAA bound to TIR1 and “Auxin/Indole-3-Acetic Acid” (Aux/IAA) proteins in a ternary co-receptor model, using protein immunoprecipitation assays where TIR1 and Aux/IAA would pull down each other in the presence of IAA.

1.6.1.1 Mechanism of auxin perception by TIR1 and Aux/IAA

TIR1 and Aux/IAA are part of an SCF complex involving Cullin 1 (CUL1), which is known to be the Scaffold protein of SCF complex and Arabidopsis S-phase Kinase-associated protein 1 (ASK1), linking the F-box protein to CUL1. RBX1 is a RING-H2 finger protein linked to CUL1, acting as a docking subunit for an E2 Ubiquitin conjugating enzyme (Ub-E2). The complex is abbreviated as SCF\textsuperscript{TIR1} (Mockaitis & Estelle 2008; Calderón Villalobos et al. 2012). (Figure 1.3)

At low auxin concentrations, Aux/IAA transcriptional repressor proteins, together with co-repressor (TOPLESS) proteins repress genes targeted by the Auxin Response Factor (ARF) transcriptional activators (Figure 1.3). As auxin concentrations rise, auxin binds to TIR1 and the Aux/IAA completes the SCF\textsuperscript{TIR1} co-receptor complex. The Aux/IAA protein is then marked for degradation by the E3 ubiquitin ligase activity and, subsequently, the Aux/IAA is degraded by 26S proteasomes. A reduction in Aux/IAA concentrations follows (Figure 1.3) and the ARFs are no longer repressed, allowing transcription to commence (Woodward & Bartel, 2005b; dos Santos Maraschin et al., 2009; Stefanowicz et al., 2015)
At low auxin concentrations, Aux/IAA transcriptional repressor proteins dimerise to repress transcription of genes targeted by auxin response factors (ARF), which are transcriptional activators.

The empty TIR1 receptor protein is part of an ubiquitin E3 ligase, the SCF\textsuperscript{TIR1} complex, based around the scaffold cullin protein.

At higher concentrations of auxin, auxin binds to the receptor TIR1 forming a binding site for Aux/IAA proteins. Auxin is the “molecular glue” between the two proteins in this co-receptor complex. Aux/IAA proteins bind via a degron motif in domain II (Kepinski & Leyser, 2005).

Once assembled, the E3 ligase activity rapidly ubiquitinates the Aux/IAA protein, marking it for degradation in the proteasome.

The resulting reduction in the Aux/IAA protein concentration leads to dissociation of the ARF-Aux/IAA dimers and de-repression of transcription (Woodward & Bartel, 2005a).

\textit{ASK1: Arabidopsis SKP1, Links F-box proteins to CULLIN.}
\textit{CULLIN; Scaffold protein of SCF complex.}
\textit{RBX1; RING-H2 finger protein, docking subunit for Ub-E2 Ligase.}
\textit{E2; Ubiquitin-conjugating protein.}

\textbf{Figure 1.3: Schematic representation of the cascade of events during auxin perception by TIR1}
1.6.1.2 TIR1 Crystal structures

Crystallographic elucidation of the structure (Tan et al., 2007) provided the evidence on how this co-receptor complex assembled with IAA as the molecular glue between TIR1 and Aux/IAA, consequently giving an insight into this fascinating perception mechanism of IAA. They solved the complex in its apo form and with IAA co-bound. They also obtained co-crystals with IAA, 1-NAA or 2,4-D, each with a 13 amino acid peptide representing the binding domain II degron form Aux/IAA7 (Table 1.2).

The crystal structure of TIR1 depicts TIR1 as a mushroom-like shape with 18 Leucine-Rich Repeats (LRRs) at the C-terminus and the F-Box domain interacting with an ASK-1 protein at the N-terminus (Figure 1.4). Figure 1.5 shows the binding pocket that forms between the Leucine-Rich Repeats α-helices 2, 12 and 14 from a bird’s eye view. The pocket is 18Å deep from the mouth to the binding site and 12Å at its widest. Video 1.1 (on attached CD media) shows a description of the TIR1 receptor structure with the mechanism of IAA binding and co-receptor complex assembly.

The crystallography revealed IP6 (inositol hexakisphosphate) (Figure 1.4) as a co-factor that was identified during the crystallisation and had not been intentionally added during the crystallisation process. It sits somewhat central to the ring of TIR1 LRRs, interfacing with a number of residues, mainly arginine residues.
Table 1.2 TIR1 crystal structures

<table>
<thead>
<tr>
<th>PDB structure</th>
<th>TIR1 - ASK1</th>
<th>Ligand</th>
<th>Aux/IAA degron</th>
<th>Resolution (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2P1M</td>
<td>Yes</td>
<td>-</td>
<td>No</td>
<td>1.80</td>
</tr>
<tr>
<td>2P1N</td>
<td>Yes</td>
<td>2,4-D</td>
<td>No</td>
<td>2.50</td>
</tr>
<tr>
<td>2P1O</td>
<td>Yes</td>
<td>1-NAA</td>
<td>Yes</td>
<td>1.90</td>
</tr>
<tr>
<td>2P1P</td>
<td>Yes</td>
<td>IAA</td>
<td>Yes</td>
<td>2.21</td>
</tr>
<tr>
<td>2P1Q</td>
<td>Yes</td>
<td>IAA</td>
<td>Yes</td>
<td>1.91</td>
</tr>
</tbody>
</table>

Figure 1.4. The TIR1+ASK1 crystal structure
A ribbon representation of the TIR1 (grey) + ASK1 (green) crystal structure. Key features are annotated with arrows. The degron (Aux/IAA), ligand and IP6 (inositol hexakisphosphate) are shown as sticks. Figure generated in PyMol.
Figure 1.5. Top view of TIR1 binding site
Right hand side: Top-down view of the auxin-binding site shown as a molecular surface with the relevant Leucine-Rich-Repeat regions labelled. The remainder of the receptor is represented as a ribbon. Right hand side: Magnification of the binding site surface display with IAA also shown. Figure generated in PyMol.
1.6.1.3 TIR1 homologues

In Arabidopsis there are five TIR1 homologues; AFB1- AFB5 (Auxin F-box Binding protein), clustered into three pairs (Figure 1.6) based on protein sequence similarity (Figure 1.3). AFB1, 2 and 3 have ~60-70% sequence identity with TIR1, whilst AFB4 and 5 have ~50% amino acid sequence identity (Table 1.3, using Clustal Omega) (Larkin et al., 2007; Sievers et al., 2011). It is intriguing as to why there are so many auxin receptors and why the need for such redundancy (Pierre-Jerome et al., 2013). Investigative work on TIR1 and AFB mutant lines has shown a range of impacts on auxin perception, with single mutations for all 6 F-box members having little or no impact on whole plant auxin responses (Dharmasiri et al., 2005b; Prigge et al., 2016), thus demonstrating auxin perception redundancy within the family. Interestingly, the loss of AFB5 confers resistance to the synthetic auxin picloram, whilst only being slightly resistant to IAA and other auxins (Walsh et al., 2006), with double afb4/afb5 knockout (Prigge et al., 2016) showing similar resistance indicating receptor-subclass specificity for AFB4 and AFB5. The tir1/afb2/afb3 triple and tir1/afb1/afb2/afb3 quadruple mutations have been reported as severely impaired, presenting high levels of arrested developmental phenotypes (Dharmasiri et al., 2005b; Mockaitis & Estelle, 2008). As yet there are no defined roles for individual family members other than auxin perception.

One explanation for how auxin is able to influence such a wide range of responses throughout plant growth and development is likely to be due to the multiplicity of F-box proteins and Aux/IAA proteins. In addition to the six auxin receptors, there are 29 Aux/IAA proteins that interact with 23 ARF proteins. These combinatory permutations add to the intricacy of auxin signalling and it remains a challenge to deduce the whole picture of IAA perception and subsequent downstream effects (Pierre-Jerome et al., 2013). This phenomenon has been studied in detail by various groups reveal varying levels of sensitivities to auxin concentrations between the F-box proteins and Aux/IAA proteins, thus suggesting that the concentration of auxin determines the co-receptor pairing to influence a range of responses (Calderón Villalobos et al., 2012; Pierre-Jerome et al., 2013; Shimizu-Mitao & Kakimoto, 2014).
Figure 1.6. Phylogeny relationship between TIR1 and AFB homologues
The phylogenetic tree is based on protein sequence similarities analysis of Arabidopsis auxin F-box proteins, created in Jalview 2.8.2 (Waterhouse et al., 2009) using the BLOSUM62 algorithm. The tree shows pairings (with NCBI accession numbers) of TIR1 (CAB87743) and AFB1 (Q9ZR12), AFB2 (XP_002875343) and AFB3 (NP_563915), and AFB4 (OAO97570) with AFB5 (AED95880).

Table 1.3 Percentage pairwise amino acid sequence similarities for the family of Arabidopsis F-Box receptors

<table>
<thead>
<tr>
<th>Receptor</th>
<th>TIR1</th>
<th>AFB1</th>
<th>AFB2</th>
<th>AFB3</th>
<th>AFB4</th>
<th>AFB5</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIR1</td>
<td>100%</td>
<td>69%</td>
<td>60%</td>
<td>59%</td>
<td>48%</td>
<td>50%</td>
</tr>
<tr>
<td>AFB1</td>
<td></td>
<td>56%</td>
<td>55%</td>
<td>47%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>AFB2</td>
<td></td>
<td></td>
<td>86%</td>
<td>50%</td>
<td>51%</td>
<td></td>
</tr>
<tr>
<td>AFB3</td>
<td></td>
<td></td>
<td></td>
<td>50%</td>
<td>51%</td>
<td></td>
</tr>
<tr>
<td>AFB4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>78%</td>
<td></td>
</tr>
<tr>
<td>AFB5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>
1.6.2 ABP1 receptor

ABP1 was originally thought to play a vital role in cell elongation and division in Arabidopsis embryogenesis. A T-DNA insertion resultant ABP1 null mutant allele was shown to be embryo lethal (Chen et al., 2001) and thereafter ABP1 was conventionally considered as being significant in auxin signalling. However, recent work refuted this principle, and reported ABP1 was not necessary for Arabidopsis development or auxin signalling (Gao et al., 2015). The authors generated two new null ABP1 alleles: abp1-c1 developed using CRISPR technology and a T-DNA insertion version abp1-TD. These null mutants were very similar to wild type Arabidopsis with no auxin resistance, no impact on development or physiology; moreover they were not embryo lethal. The observed embryo lethality in the first reported APB1 null alleles was later shown to be as a result of deleting a neighbouring embryogenesis-reliant BELAYA SMERT (BSM) gene and not ABP1 (Michalko et al., 2015; Dai et al., 2015). The BSM gene had been previously reported as embryo lethal in Arabidopsis (Babiychuk et al., 2011) but with no link to ABP1.

ABP1 has thus been considered an additional receptor candidate for auxin signalling post embryogenesis. It can be localised at the cell periphery, within the endoplasmic reticulum (ER) or at the plasma membrane/apoplast interface (Peer, 2013). It has been shown to be active in the presence of auxins and an ABP1 crystal structure co-crystallised with 1-NAA was reported (Woo et al., 2002). ABP1 found on the apoplastic side of the plasma membrane may play a role in perceiving auxin levels externally (Sauer et al., 2013).

1.6.3 SPK2a as a receptor

Another putative auxin receptor, SKP2a, is a component of SCF complexes, and can bind auxin and promote the degradation of the transcriptional repressor DPb (Dimerization Partners Of E2Fs b) to allow cell cycle progression (Jurado et al., 2010). The activity of SCF$^{SKP2}$ may provide a direct link between auxin and the cell cycle. SCF$^{SKP2}$ is part of the G1/S checkpoint during the cell cycle progression, where some transcription factors and proteins need to be degraded before the next phase can commence. Interestingly, the SKP2b homologue does not bind auxin despite an 80% sequence similarity. The selectivity was shown to be due to an amino acid difference
of a serine in SKP2b compared to a leucine at the corresponding position 128 in SKP2a. Arabidopsis \textit{skp2a} seedlings showed a slight increase of resistance to 2,4-D but not as pronounced as the \textit{tir1-1} mutant line. The double \textit{skp2a tir1-1} mutant showed an additive effect relative to the \textit{tir1-1} mutant. Comprehensive data on auxin binding to SKP2a comparable to auxin F-box proteins is lacking, as is a structure demonstrating the mechanism of binding. Thus, SKP2a’s role in auxin signalling is still yet undetermined (Peer, 2013; Napier, 2014).

However, not all auxin perception can be credited to the F-Box proteins, there are aspects of auxin responses that indicate the participation of additional receptors. Post auxin stimulus, there are responses [i] occurring within the first 5 minutes and [ii] in the locus of the plasma membrane. The former is too quick to be TIR1-mediated and the latter is contrary to TIR1’s location in the nucleus (Badescu & Napier, 2006). ABP1 was considered the auxin response mediator during this time, but the recent data noted above suggests that ABP1 is not necessary and it certainly lacks an elegantly understood mechanism comparable to TIR1 and its homologues. Whilst there is no depreciating the principal role of the auxin F-box proteins, it would appear all three receptor classes play a collective role in auxin signalling.
1.7 Kaethner’s auxin-binding site model

We studied the TIR1 binding pocket in order to establish the binding search space and gain a detailed understanding of the molecular interactions that govern IAA binding to TIR1. Until the elucidation of the TIR1 crystal structure, Kaethner’s model for the auxin-binding pocket (Kaethner, 1977) was considered the most plausible (Napier, 2001). The model, based on reported auxin activity at that time, predicted the properties of a binding site for IAA when neither the auxin signalling mechanism nor the auxin receptor crystal structure had been discovered. Kaethner’s model postulated a binding site pocket for IAA based on five regions (Figure 1.7). Electropositive regions (i) and (ii) interact with the carboxyl or other “acidic” group of the molecule. Region (iii) would be an adjacent electropositive region, whereas region (iv) is an electronegative region or the “floor” of the binding site for hydrogen bonding with the indole nitrogen and region (v) and a hydrophobic cleft interacting mainly with the benzene ring.

We show that the Kaethner’s model is very similar to the auxin-binding pocket in the crystal structure of TIR1. Figure 1.8A shows a mesh representation of hydrophobic regions (grey; TIR1, red; Aux/IAA) surrounding the indole ring. No hydrophobicity is seen around the carboxylic group when viewed from below (Figure 1.8B) but instead there is a region of polarity from a serine residue that is present adjacent to the indole nitrogen and an arginine residue in close proximity to the carboxylic group (Figure 1.8C & D). Remarkably, Kaethner’s foresight was accurate, despite no crystal structure or modern computational capabilities, and his model lacked only the 3D geometry of the binding pocket.
Figure 1.7 Kaethner’s auxin binding site model
The binding pocket by Kaethner showing the 5 regions (i, ii and iii - electropositive, iv - electronegative and v- hydrophobic. Image from Kaethner 1977.

Figure 1.8. Observing Kaethner’s predictions in the auxin TIR1 crystal structure.
Roman numeral labels correspond to those from Figure 1.7 for the predicted pharmacophore of the auxin receptor site. A; TIR1 hydrophobic interactions with IAA shown as a mesh enveloping the IAA molecule. Grey areas are TIR1 residues; red areas are Aux/IAA. B; Rotated view of A, looking up from the bottom of the pocket. The IAA is enveloped in hydrophobic residues, except where we see the carboxylic acid group (red) and the indolic nitrogen (blue). C; Hydrogen bond interactions (shown as black lines) between IAA and TIR1. The carboxylic acid residues interact with the TIR1 nitrogens (blue) and from the indole nitrogen to the TIR1 oxygen (Red). D; The polar residues at the base of the TIR1 binding pocket complete the site to complement the charged group on the IAA molecule. Figure generated in PyMol
1.8 IAA molecular interactions with the TIR1 binding pocket

LigPlot (Laskowski & Swindells, 2011) analysis (Figure 1.9) shows the molecular interactions between TIR1, IAA and Aux/IAA degron in a 2D representation. There are interactions between IAA and 11 neighbouring hydrophobic residues; two other hydrophobic interactions shown in highlighted circles are from the Aux/IAA peptide. Figure 1.10A displays a 3D representation of the LigPlot, highlighting the four residues with polar interactions with IAA (yellow) as shown in a colour cloud outline, which stabilise the IAA molecule in the binding site. We observe a hydrogen bond donated by the NH group of the auxin indole ring to a nearby carbonyl group from the backbone of a slightly twisted TIR1 β-strand of Leu439. The carboxylic acid of IAA forms hydrogen bonds to Arg403 and Ser438. Once bound at the bottom of the pocket, the indole ring of IAA (blue) serves as a platform for the tryptophan and proline residues (grey) from the Aux/IAA peptide (Figure 1.10B) to form a side-on π-stacking interaction.

The IAA molecule compliments the molecular interaction requisites for the binding site very well. It is evident that IAA fits efficiently into the pocket; the planar hydrophobic ring aligns with the hydrophobic lining (Figure 1.10A) and the carboxylic acid group orients toward a region of rich electropositive polar side chains (Figure 1.10B). For a compound to bind to TIR1, it should match the pose demonstrated by IAA, with the aromatic ring sitting flush to the hydrophobic residues and the carboxylic acid group orientated towards the bottom of the pocket in contact with the arginine residues.
Figure 1.9. 2D bimolecular interactions between IAA and the co-receptor complex
LigPlot-generated molecular interactions for IAA from 2P1Q and showing hydrophobic interactions as "eye lashes", hydrogen bonds as green lines with lengths in angstroms. Circled residues are those from the Aux/IAA degron.

Figure 1.10. 3D view of interactions between IAA and the co-receptor complex
A: 3D representation of a binding site showing residues with polar interactions to IAA as clouds. B: 3D view of hydrophobic interacting residues enclosing IAA in the TIR1 binding pocket. Tryptophan-5 and Proline-7 shown in red are from the Aux/IAA degron. Figure generated in PyMol.
1.9 Synthetic auxins as herbicides

In section 1.6.1.1, we detailed the mechanism of how auxin perception by TIR1 and Aux/IAA results in de-repression of auxin-regulated genes allowing transcription to proceed. Auxin is not transformed chemically when it interacts with TIR1, the binding is non-covalent and reversible and therefore is able to re-bind to TIR1 repeatedly until metabolised or transported away. In the case where homeostasis fails or excess auxin is applied externally, the result will be a surge of auxin-mediated signalling, uncontrollable growth and death (Walsh et al., 2011; Song, 2014).

Synthetic auxins are compounds capable of mimicking auxin. As such, synthetic auxins are employed as agricultural growth regulators, media supplements for plant propagation and as herbicidal agents against weeds (Grossmann, 2010). Herbicides are of considerable importance economically and in terms of food security in agriculture, horticulture and turf management. Weeds are considered biological pests in environments where they grow vigorously they can reduce crop yields considerably as they compete for light, water and nutritional resources. There have been reports of yield loses, varying for crops and locations, ranging from 30% to high 90’s. One such extreme case saw a 92% loss in Venezuela (1987) for cassava crops (Cobb & Reade, 2010). Weeds are responsible on average for between 8% - 13% or $75.6 Billion (1994) of all losses in global crop yields annually and weeds were by far the largest contributor to crop losses (34%), ahead of animal pests (18%) and pathogens (16%) between 2001-2003 (Oerke, 2006).

There has been reported resistance to all herbicide modes of action (Délye et al., 2013) with a recent review on herbicide resistance reporting a total of 404 unique cases of herbicide-resistant weeds globally, with ~11 reported cases a year on average (Heap, 2014). For example with synthetic auxin herbicides, there have been 30 species with reported resistance to 2,4-D, thus there is a need to discover novel compounds from existing herbicidal classes or even a novel mode of action. The 4-Hydroxyphenylpyruvate dioxygenase (4-HPPD) class of inhibitors targeting photosynthesis, were the last introduced mode of action and this was in the late 1970’s. (Grossmann, 2010; Heap, 2014)
Synthetic auxin-based weed killers selectively target dicotyledonous (dicot) plants over monocotyledonous (monocot) plants. The mechanistic basis of this distinct method of selection is yet to be discovered, and it is intriguing given that both monocots and dicots have conserved molecular mechanisms of auxin biosynthesis, metabolism, transport and signal-induction (McSteen, 2010). Nevertheless, this fortunate selectivity has been exploited in the protection of monocot crops such as rice, maize, wheat, and barley against dicot weeds. Additionally the same principle is employed for turf management of weeds for golf courses and parks.

Synthetic auxin weed-killers have been employed for many years, with the herbicidal activity of 2,4-D (2,4-dichlorophenoxy acetic acid) discovered in the early 1940’s at the same time as MCPA (2-methyl-4-chlorophenoxyacetic acid) (Cobb & Reade, 2010). Most commercial auxinic herbicides are based on three scaffolds: bicyclic aromatic rings, phenoxyacetic acids and benzoic acids, the last of which also include pyridine derivatives. These scaffolds have been the perennial basis of auxinic herbicides with an observed lack of diversity (Figure 1.11). Even the most recently discovered herbicides, Arylex™ Active (2005) (Haluxifen) and Rinskor™ Active (2010) (Epp et al., 2015) are pyridine derivatives scaffold. They do display some novelty with an aryl group off the 6-position, first demonstrated with the compound DAS534 (Walsh et al., 2006). This raises the question, are there other auxins waiting to be discovered and commercialised? (Enders & Strader, 2015)
Figure 1.11 Currently available commercial auxinic herbicides

The figure shows currently available commercial herbicides groups according to their chemical classes.
1.10 Structure activity relationship studies of auxin-like compounds

If we start to seek novel auxins we need to ask the question “What is an auxin?” It has no straightforward answer. We do know what an auxin can do in planta from our current understanding of endogenous and synthetic compounds. Yet, there is still no absolute “definition” for an auxin from a chemical point of view based on physicochemical characteristics like shape, electrostatics, functional groups and size. Auxin activity studies have been on-going for the past century (Napier, 2001). In that time plant assay systems have varied from Avena coleoptile growth, pea (Pisum sativum) IAA aspartate synthase activity, to the now established Arabidopsis thaliana plant model usually using the inhibition of primary root growth. Auxin receptors have been identified, characterised and crystallised to reveal details of molecular interactions, and we now know how IAA and other herbicides like 1-NAA and 2,4-D engage in the TIR1 binding site. We have a collection of compounds in use as herbicides that exploit the auxin-signalling pathway.

1.10.1 Prior structure activity relationship studies

The earliest attempts to define auxin activity was presented as a charge separation model (Thimann, 1963), suggesting a distance of 5.5 Å between the positive (indole N) and negative (acid) charges was necessary for auxin activity. This model could not account for all active auxins and was later superseded by a three-point attachment theory (Lehmann, 1978). Katekar collated and scored auxin activities from contemporaneous studies to develop a binding site model (Katekar, 1979). Katekar argued the three-point attachment theory could account for the activities of e.g. arylbenzoic acids and proposed an updated understating of auxin activity models. This was at a similar time to the auxin-binding site model by Kaethner (Kaethner, 1977) (Figure 1.8) Katekar’s model was based on assays covering indole, phenoxy, benzoic and naphthalene aromatic scaffold systems, all possessing carboxylic acids. He factored in electron density, steric bulk and chirality and inferred a 2-dimensional map as shown in Figure 1.12. Then, focussing on IAA, he deduced rules for carboxylic acid engagement, aromatic ring placement and influences on activity by substituents around the IAA molecule. It was a comprehensive study, presenting a working model and set the precedent for auxin structure activity investigations.
Figure 1.12 The Katekar auxin receptor binding model
Image from Katekar, 1979. **Top view:** Showing IAA with the carboxylic acid (COOH) group recognition domain, the α-carbon recognition domain and the aromatic (AR₁ and AR₂) binding areas. Surrounding regions a, b, c, d, e and f indicate substitution positions. Hashed angular lines under a and f show areas of steric obstruction. **Side view:** Showing the electrostatic regions of IAA in the "recognition conformation". The aromatic rings of the ligand align with electrophilic regions in the binding site.
Advances in physics and chemistry allowed structure activity relationships based on molecular orbital calculations (Farrimond et al., 1980), complementing the Katekar model. A systematic structure activity study on selected halogenated-IAA derivatives revealed differences between peas and wheat for indole-based auxins (Katekar & Geissler, 1983). They attributed this selectivity to the 5- and 6- positions on IAA (Figure 1.1), thus demonstrating some organism-based auxin specificity and showing the usefulness of detailed structure activity relationships studies.

The first more detailed 3D model of auxin activity aligned compounds to the crystallographic coordinates of IAA in both a Tilted (T) pose, equivalent to Kaethner’s recognition conformer, as well as Planar (P) pose equivalent to Kaethner’s modulation conformer (Tomić et al., 1998). They used force field simulations with five molecular probes: H$_2$O, NH$_2^+$, CH$_3$, O and DRY from GRID force fields software (Carosati et al., 2004), to generate surface interaction energy contours and producing a pharmacophore map of activity. However, the authors were not able to clearly distinguish active versus inactive compounds from these maps.

Other work on 3D auxin structure activity relationships revealed the reduction in auxin-like activity with the addition and elongation of an alkyl group at the C2 position of IAA attributed to hindering the free movement of acetic (CH$_2$-COOH) group of IAA (Antolić et al., 2003).

More recent work on *A. thaliana* and *Nicotiana tabacum* (BY-2) (Simon et al., 2013) looked at a variety of auxin-mediated responses at cellular, tissue, organ and whole plant levels against a small collection of IAA, 1-NAA and 2,4-D analogues (18 in total). Amongst the compounds, mono-substituted chloro-IAA derivatives at the 4, 5, or 6 positions showed similar or slightly less activity relative to IAA in primary root growth assays, and similar activity in *DR5::GFP* reporter assays. Despite the broad range of auxin activity assays, there is still no overall, accepted model of what defines auxin activity from a chemical point of view.

Most recently, a receptor-led structure activity relationship investigation studied a group of 58 compounds using surface plasmon resonance (SPR) (Lee et al., 2014). The authors tested these compounds for auxin-binding activity, observing the protein-
protein interaction of F-Box receptor proteins binding to the degron Aux/IAA. They studied the TIR1 and AFB5 receptors and observed distinct trends for e.g. aryl picolinate compounds, which showed an increased binding with AFB5 over TIR1, whilst a series of di-chlorinated phenoxyacetic acids showed the opposite preference. Amalgamating *in vitro* data on selected strong binding compounds with *in silico* chemometric analyses involving quantum chemical calculations and similarity indices, they derived a receptor-specific classification system. The classification associated compounds with van der Waals surfaces greater than IAA with a preference for AFB5 over TIR1.

All of the previous studies allowed for development of a series of models attempting to describe what an auxin is, based on data from on a whole plant level of assaying. However, some of these models are somewhat compromised to an unknown level as the perceived output data from such assays not only includes auxin perception, but also a combination of uptake, transport and metabolism. The availability of the auxin receptor crystal structure presents the opportunity to pursue a receptor focused SAR study. This approach lends itself to developing a model of what an auxin is at the receptor level, describing compound specificity and possibility of ‘auxins by design’.
1.11 Project approach

1.11.1 Detailed SAR analysis and assay selection

The aforementioned structure activity relationships studies (Section 1.10) contributed valuable information towards current understanding on what defines an auxin chemically. However, there is no complete, systematic and detailed study on auxin activity at the receptor level yet. Furthermore, we cannot compare data between the studies, as the model plant systems were different, so too were the assays and activity data. This presents an opportunity for us to address the lack of a thorough SAR investigation in a reductive, receptor focused in vitro assay with a biologically relevant in vivo assay.

We contemplated that the number of compounds versus the number of assays would be a determining factor for the workload and feasibility of completing the assay with suitable replicates. Thus, we proposed to investigate compounds using semi high throughput in vitro screening by SPR followed by low throughput in vivo root growth assays. These assays would enable us to observe auxin binding at the receptor level and further study their true auxin activity in live seedlings respectively. The compound library selection would focus on but be not limited to four main scaffolds: indole, phenoxyacetic acid, benzoic acid and naphthalic acid scaffolds.

Factoring in prior structure activity relationships study designs, we would observe auxin activity against TIR1 utilising a previously established SPR biophysical methodology (Lee et al., 2014). This in vitro assay would provide fast qualitative ‘yes or no’ answers, with compounds screened at 50 µM, far above the IC₅₀ of many active compounds but increasing the possibility of finding exploitable weaker hits. The binding for each compound would be standardised against IAA and the compounds would be ranked accordingly. Those compounds deemed as hits would be progressed onto further assays to determine kinetic binding constants and also in vivo root growth assays.

For whole plant assays, the Arabidopsis thaliana plant model will be used, similar to Simon et al., 2013. The in vivo assays would look at primary root growth inhibition and lateral root density as these traits were routinely reported in studies pertaining to
auxin activity. Our approach however would differ from the cited method with a widened dose range from 300 µM, with 14 subsequent 3-fold serial dilutions down to 20.9 pM. This is indeed excessive in terms of the highest and lowest doses, but we aimed to attain complete saturating profiles for dose responses that had not yet been demonstrated in the literature. In addition, we aimed to digitally capture in vivo data at 1200 dpi, obtaining high-resolution images to adequately count and measure phenotypic traits computationally. We would methodically assay ten roots for each dose and include an untreated control plate.

1.11.2 In silico docking approach

In silico protein-ligand docking is a relatively inexpensive and convenient tool in drug discovery, routinely employed in the repertoire of virtual screening tools assisting drug discovery (Alonso et al., 2006; Sousa et al., 2013; Grinter & Zou, 2014). It is an inexpensive technique requiring only standard computational hardware in addition to: [i] a target protein structure, usually derived from X-ray crystallographic data, cryo-electron microscopy (Cryo-EM) data or homology modelling and [ii] a library of small chemical molecules. Docking additionally requires a 3-dimensional cuboidal search space defined by x, y & z Cartesian coordinates suited to a known or hypothesised binding site on the target protein. (Alonso et al., 2006; Rognan, 2011).

Docking algorithms are based on two components: search algorithms and scoring functions; the former generates poses of the small molecules in the binding site of the protein and the latter identifies favourable poses for each compound which then ranks the compounds against one another in terms of binding affinity (Rognan, 2011).

Classical docking technique is based on rigid body docking, in which the receptor protein structure and the ligands were held rigid and the algorithm attempts to fit the ligand into the protein. This method has been superseded, now allowing flexibility for either the ligand or protein, or both simultaneously. The flexible ligand, rigid protein approach is predominantly used by current docking algorithms, delivering results in a suitable period of time, ranging from minutes to hours. These algorithms generate conformers of the small molecule and find the best scoring conformer for each
compound. A collection of compounds can be scored and thus ranked to identify the best potential binders (Cole et al., 2011).

For the scope of our work, we will initially employ a flexible ligand, rigid protein-docking algorithm. We selected the AutoDock Vina algorithm given the availability of local expertise with the algorithm, and it is free for academic use. Besides that, AutoDock Vina has been widely utilised in the literature for virtual screening work. In fact, AutoDock toolset was the most cited docking algorithm between 2001 and 2011 (Sousa et al., 2013).

1.11.3 Pharmacophore generation

We aimed to utilise a commercially available tool to build a pharmacophore map based on the results of the SAR. We also needed software to account for physiochemical properties of a large number of compounds and factor in activity data. This software tool would help us visualise the chemical space investigated and which regions are responsible for auxin activity. We used the commercially available FORGE software (Cheeseright et al., 2006) to help us summarise the SAR activity data. It employs similar principles to Tomic et al., 1998 with compound aligned to a known reference and calculated molecular descriptors summarised, but is far more advanced and is becoming widely validated in the pharmacological industries. The software outputs intuitive pharmacophore maps.
1.12 Project motivation

1. Non-canonical and ‘designer’ auxins that could also possibly drive assembly of the TIR1 and AFB5 complexes with Aux/IAA’s. Our current understanding of what makes an auxin (chemically) at the receptor level is as yet unrealised. The development of a receptor focused pharmacophore model could provide an intuition into the receptor chemical requisites for auxin activity and aid the design of novel auxin compounds.

2. TIR1 and AFBs bind IAA (and other auxins) in the absence of Aux/IAA co-receptor, but with low affinity (in the micro molar range). To date only, the formation of the co-receptor assembly has been measured (i.e. TIR1 + IAA + Aux/IAA). Direct observations’ of small molecule binding to the auxin receptors (i.e. IAA to TIR1) have not yet been demonstrated and would provide insightful observations into the first step of auxin perception.

3. The TIR1 and AFB5 pairing is amongst the most distantly related between members of the auxin F-box receptors. Whilst pyridine derivatives such as picolinates demonstrate preferential binding for AFB5 over TIR1, the reverse preferentiality is yet to be demonstrated in addition to total selectivity for the receptor pairing. Such compounds would lead towards the development of functional utility probes for better understanding of individual members of the F-box auxin receptor family.

4. Anti-auxins, which selectively prevent assembly of either TIR1 or AFB5 with Aux/IAA degrons, may be predicted. Non-selective anti-auxins have been designed and validated by Hayashi, based on the indole scaffold, as exemplified by the compound Auxinole (Hayashi et al., 2012). Physiological assays have suggested affinities in the low micro molar range. Anti-auxins based on other scaffolds are yet to be demonstrated and would prove useful probes in understanding the role such auxin perception antagonists.
1.13 Project aims

The overall objective is to better understand auxin perception at the receptor level, and understand what comprises auxin-like behaviour. This will help us gain an understanding of what kind of compounds can act as auxins, and to investigate further why there are redundant auxin receptors. We aim to address the project motivations as follows:

1. Compound- and receptor-focused structure activity relationship studies for the co-assembly of TIR1 and AFB5 receptors with Aux/IAA 7, aiming to:

   1.1. Generating a 3D pharmacophore map for TIR1 and AFB5 receptor sites, looking for differences to explain ligand preferences and probable functional roles. With further sub-aims of:
      1.1.1. Identifying auxins with non-canonical (e.g. other than carboxylic acids) functional groups.
      1.1.2. Identifying compounds that are selective between TIR1 and AFB5.

2. Observing the binding of auxin and other compounds to the TIR1 receptor by:
   2.1. Immobilising the TIR1 receptor in the context of the SPR biophysical technique.
   2.2. *In silico* docking into the TIR1 receptor binding pocket.

3. Using the SAR screening to discover novel anti-auxins. This would help us establish the role of compounds that do not demonstrate activity in our SPR activity assays. (This is not discussed in the thesis)
Chapter 2: Methods and Materials
2.1 Insect cell culture and protein expression

2.1.1 Adherent and suspension insect cell cultures

*Trichoplusia ni* (Tni. High Five) cells adapted to serum free media were used throughout as the host cell line. Tni. High Five cells stored in liquid nitrogen were reanimated in T25 adherent culture flasks with 5mL of Express Five™ SFM (Serum Free Media) media supplemented with glutamine at 28°C (the standard temperature for all insect culture from here onwards for suspension and adherent cultures). Upon achieving 90-100% confluence (surface coverage) the cells were passaged into T75 flasks. At optimal confluence the cells were suspended in the media (combined from 2 or more T75 flasks) and a sample counted in haemocytometer (Marienfeld, Neubauer-improved 0.1mm, catalogue number: 0640030) at 25x magnification under a light microscope. The cells were diluted to achieve a 40mL suspension culture with a cell density of 0.5x10^6 cells/ml, shaken at 115 rpm (rotations per minutes) in 100mL conical flasks, stopped with foam bungs and covered with foil. Cells were counted daily until they attained a density of 4x10^6 cells/mL; thereafter cells were passaged into 250mL cultures at 0.5x10^6 cells/ml and incubated until a density of 4x10^6 cells/mL ready for large scale expression, aliquots of cells were also passaged at 0.5x10^6 cells/mL to maintain suspension culture stocks. Techniques were adopted from (Wu et al., 1990; Taticek et al., 1997; Lynn, 2002; O’Connor & O’Driscoll, 2005; Drugmand et al., 2012; Zheng et al., 2014).

2.1.2 Generation of recombinant baculoviruses

Production of recombinant virus, selection, expression screening, and generation of high-titre viral stocks were done by Oxford Expression Technologies (Oxford, UK). Transfer vectors for codon-optimized, Arabidopsis *TIR1* (Figure 2.1) and *AFB5* genes (Figure 2.2) were engineered to give fusion proteins His-MBP-(TEV)-FLAG-TIR1 and His-MBP-(TEV)-FLAG-AFB5 respectively. These proteins were co-expressed with His-(TEV)-ASK1 encoded in the same transfer vector (Lee et al., 2014).
2.1.2.1 TIR1 expression vector

Figure 2.1. TIR1 transfer vector
Baculovirus transfer vector containing TIR1 under control of polyhedrin promoter & with 10xHis tag & MBP, also contains ASK1 under control of P10 promoter.
2.1.2.2  TIR1 amino acid sequence

**His**<sub>10</sub>-linker-MBP-TEV-FLAG-<i>A.thaliana</i>_TIR1

MSGHHHHHHHHHHHGDLVEEMAEKKLVIWINGDKYNGLAEVGKKFEK
DTGIKVTVEHPDKLEEKFPQQVAATGDGPDIIFWAHRDFGGYAQSGLLEAITPD
KAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLPNPPKTEEWIPAL
DKELKAKGKSLMFNLQEPYFTWPLIAADGGYAFKENGYDKDVGVDNA
GAKAGLTLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTS
KVNYGVTVLPTFKGQPSKPFVGVLSSAGINAAASPKNELAKEFLENYLLTEDGLE
AVNKDKPLGAVALKSYEELAKDPRIAATMENAQKGEIMPNPQMSAFWYA
VRTAVINAASGRQTVDÉALKĐAQTNSSSNSSSNSSNSNGIEGRG<ENSFLASHFQG
GAIAADYKDDDDKGLNMQKRIALSFPEEVEHVFSFIQLDGDNRVSLVCKS
WYEIERWCRRKVIGNCVASPATTVIRRFPKVRSVELKGPHTAFDNLVPGD
WGGYVYPWIEAMSSSYTWLEEIRKLKMVTDCCLELAISFSKFNKVLVSLSSC
EGFSTDGLAAIAATCRNLKELDLRESDDVDVSGWHLSHPDTYTSLVSNLISC
LASEVSFSALERLVTRCPNKSLKLNRAVPLEKLTLQRAPQEELGTTGYYT
AEVRPDVYSGLSVALSGCKELRCLSGFDWAVPAÝLPAVSYCSVSLTTLNLSY
ATVQSYDLVKLLCCPCLQLQLWRVLWVLDYIEDAGLEVLASTCKDLRELRFVPSEPFG
VMEPVNALTEQGLVSVSMGCPKLESVLYFCRQMTAAALITARNRNPMTRFR
LCIEPKAPDYTELELPDLGFGAIVEHCKDLRRISGGLTĐKVFYIATYAKK
MEMLSVAFAIGDSDLGMHỊVLSGCSLRLKLEIRDCPFDRNNDALLANASKLETM
RSLWMSSCSVSFGACKLLEQKMPKLNVEVIDERGAPDSRPESCPVERFYIYT
VAGPRFDMPGFVWNMDQDSTMRFSRQIITTNGL
2.1.2.3 AFB5 expression vector

Figure 2.2. AFB5 transfer vector

Baculovirus transfer vector containing truncated auxin receptor AFB5 under control of polyhedrin promoter & with 10xHis tag & MBP, also contains ASK1 under control of P10 promoter.
2.1.2.4  Truncated AFB5 Amino acid sequence

**His**<sub>10</sub>-linker-MBP-TEV-FLAG-*A.thaliana_truncated_AFB5*

* 70 additional amino acids relative to TIR1 at the N-terminus were truncated and are known not to impact the binding of compounds to AFB5.

```
MGSHHHHHHHHHGGDQLVEEMAEEGKLVIWINGDKGYNGLAEVGGKKEKDTGIKVTVEHPDKLEEKFPQQVAAATGDPDIIFWAHDFGGYAMQGLLAEITPD
KAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIYNKDLLPNPPKTWEIPAL
DKEKLKAGKSAQNLQEPYFTWPLIAADGGYAFKYENQKD1KVDVGVDNA
GAKAGLTFVLDLHIKNHMDTADTIIEAEAFNKGETAMTINGPWAWSNIDTS
KVNYGVTVLPTFKGQPSKFPFGVLSAGINAASPKNELAKEFLENYLLTDEGLE
AVNKDGPLGAVALKSYEEELAKDPRIAATMENAOQKGEIMPNDPQMMSAFWYA
VRTAVINAASGRQTVDEALKDAQTNSSSSSSSSSSSSSSSSSSSSG
GENLYFQGAIAiedyKdddkGLNMPDHVTENVLQFLDSRCDRNAASLVCKSWW
RVEALTRSEVFIGNCYALSPARLTQRKVRVSRLVLSGKPRFADFNLMPMDW
ANFAPWVSTMAQAAYPCLEKVDLKRMMFTDDLALLADSFPGFKEILVCCEG
FGTSISIVANKCRKLVDDLIESEVTDDEVWISCFFPEDVTCELSAFDCVEA
PINFKALEGVLVSPFLLKRLNRFLSVLVELHRLLLAPQLTSLGTGFSHDEE
PQSEEQPDYYAAAFRACKSVVCLSGFREMPEYLPAPCFLSLNSYNANIS
PDMPKPIILNCQKVFWALDSICDEGLQAVAATCKELRELIFPFDPREDSEG
PVSELGLQAISEGCRKSLISLYFCQRTMNAAVIAMSENCPETLTVRLCIMGHR
PDHVTGKPMDERGGFAGKKNKLTRLAVSGLLTDQAFRYMEGYGLVRGL
VAFAAGDSMALRHLEGCPRLQKLEIRDSPFEDVALRSGMHRYYMRFVW
MSACSLSGCCKCIDARAMPNLVVEIGSDDDDDDNRDYVETYLYMRSGLDPR
NDAPKVFVIL
```
2.1.2.5 ASK1 amino acid sequence

His$_6$–TEV-$A.thaliana$ _ASK1

MHHHHHHENLYFQGAMSAKKIVLSSGDGESFEVEEAVALESQTIAHMVEDDCVDNGVPLPNVTSLAKVIEYCKRHVEAAASKAEAVEGAAATSDDDLKAWDADFMDQATLFLILAANYLNKNNLDLTCQTVADMKGTKPEEIRTTFNIKNDFTEEEEEIVRRENQWAFAE
2.1.3 SDS-PAGE protocol

A total sample loading of 20uL per well was prepared with:

- 5µL of NuPAGE® LDS Sample Buffer (4X) (ThermoFisher Catalogue number: NP0007)
- 2µL of NuPAGE® Sample Reducing Agent (10X) (ThermoFisher Catalogue number: NP0009)
- Up to 13µL of a sample - lysate/protein etc.
- Should the sample volume be less than 13µL, Di-ionised water was used to make up the volume giving an overall total of 20µL.

The samples were then denatured at 95°C for 10 minutes, spun down at 13,000rpm for minutes and loaded into the wells. 5µL of MagicMark™ XP Western Protein Standard (Catalogue number: LC5603) was added to the left most well for reference.

15-well precast Tris-Glycine, 8-16% NuSep gradient gels (Generon, Cat# NG31-816) were used in accordance with the manufacturers specifications –180V for 45 minutes in a Novex® Tris-Glycine SDS Running Buffer (ThermoFisher Catalogue number: LC2675) in Bio-Rad Mini PROTEAN® Electrophoresis Cells. Once run, gels were removed from their cassettes, visualised in a Syngene GBOX under UV trans-illumination and images recorded. These gels were suitable to be progressed onto Western blotting development in 2.1.4.
2.1.4 Western blot development

Gels from 2.1.3 were transferred onto a PVDF (polyvinylidene difluoride) membrane and set up as portrayed in Figure 2.3. The PVDF membrane (Merck Millipore, Immobilon-FL, Catalogue #: IPFL00010) was activated in methanol for 30 seconds, Filter paper, 7 cm x 8.4 cm (ThermoFisher Catalogue number: 84783) and sponges were soaked in 1xNuPAGE® Transfer Buffer (Catalogue # NP00061) supplemented with 10% methanol. The cassettes were loaded into a Bio-Rad Mini Trans-Blot® Cell, filled with more transfer buffer and run at 120V for 1 hour.

![Western blot gel-to-membrane transfer step](image)

**Figure 2.3. Western blot gel-to-membrane transfer step**
Stacking order of Western blot cassette for protein transfer from SDS-PAGE gel onto PVDF membrane. Transfer occurs from “Black side” towards “Transparent side”.

The membrane was then removed and the following steps performed:

1. Blocked in TBS/Tween with 10% (w/v) skim milk (Marvel) powder for 1 hour on a shaker.
2. 3x 10 minute TBS/Tween washes.
3. Incubated with 1° antibody at 1:5000 (2µL in 10mL) in TBS/Tween with 5% (w/v) skim milk for 1 hour on a shaker.
4. 3x 10 minute TBS/Tween washes.
5. Incubated with 2° antibody at 1:5000 (2µL in 10mL) in TBS/Tween with 5% (w/v) skim milk for 1 hour on a shaker.
6. 3x 10 minute TBS/Tween washes
7. Wash media removed and placed in clean plastic container, mixed together 2mL of Luminol and 2mL of Peroxide solution from Immobilon Western Chemiluminescent HRP Substrate kit (Merk Millipore, catalogue #: WBKLS0500) and poured over membrane.
8. Membrane left to develop for 5 minutes with gentle shaking.
9. Membrane was imaged using Chemiluminescence mode on an ImageQuant LAS 4000 (GE Healthcare) with the following parameters:
   a. Exposure Type: Precision
   b. Exposure Time: Manual (between 15 to 120 seconds were investigated)
   c. Sensivity/Resolution: High Resolution
10. Images were saved as .tiff picture files.

1° antibody (either) =
   Monoclonal **Anti-polyHistidine** antibody produced in mouse (Sigma Aldrich, Catalogue #: H1029-2ML)
   Monoclonal **Anti-Maltose Binding Protein** antibody produced in mouse (Sigma Aldrich, Catalogue #: M6295-2ML)
   **ANTI-DYKDDDDD** (anti-FLAG) epitope tag monoclonal antibody produced in mouse (THERMO Fisher, Catalogue #: MA1-91878)

2° antibody
   * Anti-Mouse IgG (whole molecule)–Peroxidase antibody produced in sheep (Sigma A5906)
2.1.5 MOI titration estimation adherent cultures – with known virus titre
High Five cells in adherent flasks at 95-100% confluence (1x10^6 cells/mL) in 5mL of media were infected with recombinant baculovirus measure in qpfu/mL (Quantitative Plaque Forming Units / Millilitre) relating to 1x, 5x and 10x times the number of High 5 cells in the given volume (Multiplicity of Infection (MOI) of 1, 5 and 10 respectively). The flasks were incubated for 72 hours and cells were dislodged from the surface by gently striking/smacking on the fume hood surface. The suspension was transferred into a 15 mL falcon tube, centrifuged at 2000rpm for 2 minutes at 4°C. The pellets were immediately placed in an -80°C freezer. For each sample/time point the pellets were lysed in 100µL of CytoBuster, vortex mixed and spun at 13,000 rpm for 2 minutes at 4°C. 2µL of each supernatant was used in SDS-PAGE and Western blot carried out as in 2.1.3 and 2.1.4 respectively.

2.1.6 MOI titration estimation shaking cultures with known virus titre
The method was similar to 2.1.5 except that High Five cells were in suspension at (1x10^6 cells/mL) in a volume of 50mL and were infected with virus at various MOI’s, shaken at 115rpm. At various time points (24, 48, 72 and 96 hours) 5mL of the suspension was aliquoted into a 15 mL falcon tube, centrifuged at 2000rpm for 2 minutes at 4°C. The pellets were immediately placed in an -80°C freezer.

2.1.7 Optimal protein expression time using virus to cell volume for baculovirus with unknown titre
The same methodology was followed as in 2.1.5 for adherent or 2.1.6 for shaking cultures except that a range of virus volumes (increments of 5µL) was added to 5 mL (adherent cultures) or 50mL (shaking cultures) respectively. The cells were incubated as per their relevant methods and analysed accordingly.

2.1.8 Large scale insect cell baculovirus induced expression
Cells were infected at a density of 1x10^6 cells/mL in suspension cultures, with the optimal multiplicity of infection as determined in 2.1.5 and 2.1.6 and shaken at 115rpm. The cell pellets were harvested approximately 72 h after infection by centrifugation then stored at -80°C.
2.2 **TEV protease expression and purification**

### 2.2.1 Expression

An MBP-(TEV)-His-TEV construct, kindly donated by fellow colleagues was transformed into BL21-CodonPlus (DE3)-RIPL strain cells (Agilent Technologies) against 100µg/mL carbenicillin and chloramphenicol selection on LB plates. From the colonies glycerol stocks were created and stored at -80°C. The culture was grown in LB broth shaking cultures at 37°C with 100µg/mL carbenicillin and chloramphenicol to an OD$_{600}$ of 0.6. Expression was induced with 0.4mM IPTG and the culture shaken at 25°C for 18 hours. Thereafter, cells were harvested and resuspended using 5g/mL wet cell weight in CRB buffer and stored at -20°C.

### 2.2.2 Purification

Buffer definitions:

- **Cell resuspension buffer (CRB):** 1 x PBS, 0.3M NaCl.
- **Loading buffer (LB):** 20mM Tris-HCl, pH 7.5, 0.3M NaCl, 3mM DTT 20mM Imidazole, 20% Glycerol.
- **Wash buffer (WB):** 20mM Tris-HCl, pH 7.5, 0.3M NaCl, 3mM DTT, 50mM Imidazole, 20% Glycerol.
- **Elution buffer (EB):** 20mM Tris-HCl, pH 7.5, 0.3M NaCl, 3mM DTT 250mM Imidazole, 20% Glycerol.
- **Dialysis buffer (DB):** 20mM Tris-HCl, pH 7.5, 0.3M NaCl, 5mM DTT, 30% Glycerol.
- **Storage buffer (SB):** 20mM Tris-HCl, pH 7.5, 0.3M NaCl, 5mM DTT, 50% Glycerol.

Cells were thawed at room temperature, lysed in Bugbuster (Merck Millipore cat no. 70921), DNase I (Roche), 1x protease inhibitor tablet (cOmplete™ Protease Inhibitor Cocktail Tablets, Roche) for 1 hour at 4°C, sonicated with 3 x 25 pulses using a probe sonicator. The lysate was spun at 20,000 rpm for 20 minutes at 4°C. The supernatant was filtered with a 0.2µm syringe filters and injected onto a 5mL Hi-Trap column equilibrated with LB on an AKTA 9000 system, the column was washed with 10 column volumes of WB and then eluted with EB. The major peak was pooled and
dialyzed in DB (3 x 1L, changed every 3 hours). Thereafter the dialyzed sample was injected on the 5mL Hi-Trap column equilibrated with LB on the AKTA 9000 system; the column was washed with 10 column volumes of WB and then eluted with EB. Protein concentration was estimated with NanoDrop A$_{280nm}$. The purified enzyme was diluted to 0.5mg/mL in SB and stored at -80°C.

2.3 F-box protein purification

- **Buffer A**: pH 7.4, 20mM Tris-HCl, 200mM NaCl, 1mM EDTA, 50µM phytic acid and 1mM TCEP.
- **Buffer B**: pH 7.4, 20mM Tris-HCl, 200mM NaCl, 1mM EDTA, 50µM phytic acid, 1mM TCEP and 250mM Imidazole.
- **Buffer C**: pH7.4, 10mM HEPES, 150mM NaCl, 3mM EDTA, 50µM phytic acid, 1mM TCEP and 0.05% Tween 20 v/v.

2.3.1 Cell lysis and protein extraction

Frozen TIR1/ASK1 and AFB5/ASK1 pellets were thawed at room temperature and lysed for 40 minutes whilst rolling at 4°C (with the addition of the following according to the manufacturer’s instructions unless stated) in Cytobuster™ Lysis media (Invitrogen 5mL per each 250mL worth of cell pellet), supplemented with DNase I (Roche), protease inhibitors (cOmplete™ Protease Inhibitor Cocktail Tablets, Roche), 50µM phytic acid (Sigma) and 1mM reducing agent TCEP (Tris(2-carboxyethyl)phosphine hydrochloride – Sigma)). Thereafter the lysate was diluted into Buffer A and subjected to 3 x 30 seconds sonication (probe sonicator) before being centrifuged at 20,000 rpm and 4°C for 15 minutes. The supernatant was systematically filtered through 0.45µm (Whatman™ 25mm GD/X) and 0.2µm (Minisart. Cellulose Acetate) syringe filters respectively.

2.3.2 Tandem Affinity Protein Purification

The filtered lysate was loaded onto a gravity chromatography column with a 5mL bed of anti-His affinity resin (cOmplete His-Tag Purification Resin – Roche), washed with 10 column volumes of Buffer A and eluted with 50mL Buffer B. TEV protease was added to the eluate and the sample was kept rolling at 4°C overnight. The solution was
then loaded onto a gravity chromatography column with a 3mL anti-FLAG-affinity resin bed (ANTI-FLAG® M2 Affinity Gel - Sigma), washed with 10 column volumes of Buffer C and eluted with Buffer C containing 6mL of 3x-FLAG peptide (Sigma Aldrich) at 100µg/mL and stored on ice. The protein concentration was assayed by NanoDrop A$_{280nm}$ measurement.

2.4 Small molecule library set up

Compounds were commercially sourced from Sigma Aldrich (UK), Alfa Aesar (UK), Fischer Scientific (UK), Key Organics (UK), ChemBridge (California, USA) and OlChemIm (Olomouc, Czech Republic). We also acknowledge in kind donations from: Prof. Kenichiro Hayashi (Okayama University, Japan), Dr Matteus Bauer (MRC, Cambridge) and John Paul Evans (Syngenta, UK).

Compounds were weighed and dissolved in the appropriate volume of DMSO to achieve 100mM stock solutions and 10 mM working solutions. All compounds were catalogued with a location tag serving as an identifier and stored in the dark either at 4°C (10mM) or -20°C (100mM). Compound management was done with DataWarrior™ software.

2.5 General root growth assay methodology

All root growth assays were performed with Arabidopsis wild-type (WT) and $tir1-1$ (Ruegger et al., 1998) lines, all in Col-0 ecotype background.

2.5.1 Seed bulking

Seeds were stratified in 0.1% agarose solution for 48 hours at 4°C in opaque eppendorf tubes. The seeds were pipetted onto pre-soaked soil, placed in propagators and incubated under 12 hour day/night cycles, with temperatures set at 22°C for day and 18°C for night conditions in a greenhouse setting. After the emergence of cotyledons, excess plants were removed to leave a single plant per pot, further grown and watered regularly until flowering. After which, watering was ceased and the plants were placed in sealed in porous plastic bags and left to dry weeks before seeds were harvested.
2.5.2 Seedling germination

Col-0 WT seedlings were sterilised under chlorine gas (generated from bleach mixed 1:1 with concentrated hydrochloric acid) for 3 hours in a sealed container whilst in a vented fume hood. Square petri dishes (Greiner Bio One, 120x120x17mm) with 30 mL of \( \frac{1}{2} \) MS media (1.5% Phytoagar (Sigma Aldrich A7921) and 0.5% sucrose) were prepared. Once the media was cool, the treated seeds were suspended in 0.15% agarose and pipetted individually in 5 rows with 10 seeds per row in each plate, for a total of 5 plates. We spotted these seeds individually (Figure 2.4A), to ensure we could germinate individual seedlings (Figure 2.4B) for ease of transfer later in the protocol. This is in contrast to normal convention, where seeds are spotted in multiples, with the resulting seedlings clustered up and tricky to separate. Despite the arduous task of individually spotting each seed, we avoided having to impose undue stress on the roots. Had they germinated in groups, they would have required physical separation for transferring onto treated plates.

For each compound assayed 5 such plates were made. The plates were wrapped in cling film leaving a gap at the bottom of the plate to allow for air circulation. Plates were placed at 4°C for 48 hours in the dark before being moved to light conditions in growth incubators for 6 days. The Sanyo growth incubators were programmed with 12 hour/day night cycles with 22°C for day and 18°C for night conditions.
Figure 2.4. Root growth assay seeds spotting and germination on media

A: A photograph of *A. thaliana* Col-0 WT seeds spotted onto a plate in 5 x 10 grid prior to germination. Seeds were spotted individually whilst suspended in 0.15% agar.

B: Photograph of the germinated seeds after similar spotting as in Panel A. They were incubated for 2 days in the dark at 4°C, followed by 6 days in a growth incubator with 12 hour day/night cycles.
2.5.3 Seedling treatment with compounds

A series of plates with 15 serial dilutions for each compound were prepared commencing with 135µL of 100mM compound stock dissolved in DMSO added to 45mL of molten ½ MS media in a 50mL Falcon tube, 30mL of this was dispensed into plate 1 (300µM). To the remaining 15mL of treated media, 30mL of fresh ½ MS was added and mixed to give 100µM and 30mL poured into plate 2, the steps were repeated until 15 such serial dilutions were made. A control sample of 30mL untreated ½ MS media was also poured giving a total of 16 plates. 10 seedlings were randomly picked from the 5 plates and transferred onto each of the treated plates and the position of the primary root tip was marked (Figure 2.5). This was also repeated for the untreated plate. In total 160 seedlings were transferred this way. The plates were stacked up in a random manner held together with adhesive tape and covered with cling film except for the bottom lip of the plates. The stack of plates was placed, with the seedlings vertical, in the same growth incubator as for the germination for 6 days with the same day and night cycles. On the 6th day, the plates were scanned, using an HP PSC 2500 all in one device, at 1200dpi in colour mode and saved as uncompressed “.tif” files.
Figure 2.5. Transferred seedlings on treated media, with markings for initial primary root length

A petri dish with agar media onto which 10 seedlings have been placed along a horizontal guideline (black). The initial length of the primary root tip for each seedling has been marked with a black dash.
2.5.4 Primary root measurements

Primary root length from the marked point was measured in Image J (Schneider et al., 2012), a global scale of 475 pixels = 1 cm was set and roots measured using the freehand tool. The measurements were stored in a spreadsheet and plotted using Prism graphing software (version 7). IC50 values were fitted as described in section 2.5.5.

We trialled many automated root image measurement tools like RootTrace (Naeem et al., 2011), Ez-Rhizo (Armengaud et al., 2009) and SmartRoot (Lobet et al., 2011) but they were not suitable as they could not handle the image format (.tiff), image quality (1200 dpi) or could not detect and distinguish individual primary roots in the image especially where laterals overlapped between neighbouring plants. Thus the primary root lengths were measured manually on screen using a mouse to draw a segmented line between the marked spot (Figure 2.6) before treatment up to the tip of the primary root using the measurement tool in ImageJ. Granted it was a laborious process, but we could ensure that we were able to inspect all the seedlings at each concentration for all the compounds we tested.
Figure 2.6. On screen root growth measurements
Yellow dotted line = position of primary root tip before treatment. Yellow diamond line = final position of primary root tip after treatment. $\ell$ = measured primary root growth length after treatment. Panel A shows a seedling with no primary root growth after treatment. Panel B shows a seedling with uninhibited primary root growth and the measured section marked by $\ell$. Panel C is a magnification of part of the elongated primary root in B, showing a segmented line (with small squares for each segment) of the measured root as highlighted with yellow shading.
2.5.5 Root growth assay IC$_{50}$ determination

The primary root growth measurements were fitted to a logistic function (2.1), using the Levenberg–Marquardt algorithm, in QTI plot. The standard deviation of the data points was weighted into the algorithm.

$$f(x) = \frac{M}{1 + e^{-k(x-x_0)}}$$  \hspace{1cm} (2.1)

Where:
- $M$ = Maximal value on curve
- $e$ = natural logarithm
- $x_0$ = IC$_{50}$ value
- $k$ = Steepness of the curve

2.5.6 Lateral root density measurements

The number of laterals per seedling was counted and divided by the total length of the root (measured from the base of the hypocotyl to the primary root tip). The roots were measured as in section 2.5.4.
2.6 GFP confocal microscopy

Col-0 $DR5::GFP$ seeds from (Ottenschläger et al., 2003) were germinated in a similar manner to section 2.5.3 except that only 30 seedlings per plate, in 6 groups of 5 seedlings were spotted. Post germination, seedlings were treated $in situ$ with compound dissolved in ½ MS water solution. Treatments were between 30 to 120 minutes and seedlings were left in the dark for the duration. Once treated, the primary roots were cut below the hypocotyl to yield only the primary root, which was transferred into a Propidium Iodide (PI) (1:10000 in ½ MS solution) for 10 minutes. The seedlings were then placed onto a slide with water added to keep the moisture, covered with a glass slip and immersion oil applied to the slide and placed within the Leica LSM 880 imaging system with a 25x oil objective used. GFP was excited with a 488 nm laser line and detected between 499 nm and 544 nm. PI was excited with a 514 nm laser and detected between 598 nm and 720 nm.

2.7 Pull down assays

Pull down assays were performed in a similar manner to (Kepinski & Leyser, 2005).

2.8 Pharmacophore generation

The pharmacophore was generated in Forge, version [10.4.2], Cresset, Litlington, Cambridgeshire, UK, [2015]; http://www.cresset-group.com/forge/.
2.9 Synthetic of iMQ (3-[(1H-1,2,3,4-tetrazol-5-yl)methyl]-1H-indole)

Based on methodology from (Dolusici et al., 2011), 180mg of NaN₃ (Sodium Azide - Sigma Aldrich) was added to a stirring solution of 2.6mL of AlCl₃·THF (Sigma Aldrich) at 0°C in an ice bath and left for 2 hours in a 2-neck round bottom flask with an angled side neck kept under an inert N₂ atmosphere. 190 mg of IACN (Indole-3-acetonitrile - MP Biomedicals, LLC) powder was dissolved in 5mL THF and added via the side neck, and the residual powder in the neck was washed down with 2mL THF. The mixture was taken off ice and placed onto a heating mantle and left to reflux for 18 hours with a water condenser column attached. Teflon tape was used to seal the glassware connections. After 18 hours, the mixture was left to cool where we observed an off-white creamy precipitate. The aluminate precipitate was dissolved into 20mL of 1M citric acid, washed with 3x25mL of ethyl acetate (EtOAc), then 1x25mL water, followed by 1x25mL brine and 1x25mL water. The extract was then dried over MgSO₄ (Fischer Scientific). The excess solvent was evaporated off in a rotary evaporator, producing residual pale yellow oil. The crude material was dissolved in 5mL EtOAc and purified by flash chromatography on standard 60Å (Sigma Aldrich) silica gel with an EtOAc:MeOH (95:5) gradient set up. Fractions containing the product were pooled, concentrated with a rotatory evaporator and left to dry in a vacuum oven for 24 hours giving an off white powder in a 24% yield; 60mg based on expected weight of 250mg.

High resolution mass spec showed a peak at m/z = 200.0931 (Expected HRMS = 200.0931, ChemAxon calculated), indicative of the product with molecular formula of C₁₀H₁₀N₅, matching that of iMQ in positive ion mode with a mean error of -0.5ppm, as well as a species at 222.0750 indicative of iMQ species in salt complex with sodium C₁₀H₉N₅Na, mean error of -0.0.

¹H NMR assignments (300 MHz, DMSO-d₆) δ = ppm 4.36 (2 H, s), 6.93 - 7.01 (1 H, t), 7.05 - 7.12 (1 H, t), 7.23 - 7.27 (1 H, s), 7.31 - 7.39 (1 H, d), 7.39 - 7.44 (1 H, d), 10.78 - 11.21, (1 H, s), 15.46 - 16.28 (1H, s).
2.10 SPR methodology

2.10.1 Introduction

Surface Plasmon Resonance (SPR) is an optical phenomenon based on the resonance between plasmons excited on a surface by light under total reflection conditions and the evanescent wave of the light itself. It occurs at the interface of two different refractive indexes, in the case of Biacore, this is between a glass slide with a conductive thin film of gold on one side and a flowing sample solution on the other. At given combination of incident angle and wavelength (energy) of light, plasmons are excited in the gold film. This is a form of energy absorbance, distributing electric field intensity (the *evanescent wave*) across the gold interface as shown in Figure 2.7. The energy absorption results in a drop in the reflected light intensity for the given incident angle. This phenomenon is very sensitive to perturbations on and up to 150nm from the gold surface. Thus changes in solution viscosity, or adsorption to the interface, influence the reflective index, in turn impacting the SPR angle (Figure 2.8).
Figure 2.7. Basis of the SPR phenomenon
Schematic showing the Surface Plasmon Resonance (SPR) distributed across a gold surface with an attached surface matrix. The Evanescent field strength is depicted to the right of the gold layer. The SPR occurs as a result of light at a specific angle of incidence and wavelength refracting through a prism and reflecting off a gold monolayer. The reflected light shows a decreased intensity at the angle of incidence and is referred to as the SPR angle.
2.10.2 Binding interactions observed by SPR

One molecule (the ligand) is immobilized onto the gold surface (Figure 2.9) and yields a new baseline SPR angle (Figure 2.10 – baseline). Thereafter, a solution containing a complimentary binding partner (the analyte) is injected through the cell over the immobilized ligand. Binding leads to an increase in the mass at the SPR interface, changing the observable SPR angle in real time (Figure 2.8), and yielding a binding curve displayed on screen. The sensorgrams are presented in response units (RUs, Figure 2.10) from which the association rate of the complex formation ($K_a$) can be derived. During this period of time, the bound complex will also dissociate, but given the continuous presence of analyte in the flowing solution, another analyte may bind to free ligand as long as there is a continuous presence of analyte in the flowing solution.

Once the solution flowing across the interface no longer contains the analyte, the complexes will begin to dissociate (Figure 2.10), thus reducing the total mass on the interface and yielding a dissociation rate ($K_d$). Overall, both on and off rates describe the binding kinetics for the interaction and give the equilibrium dissociation constant ($K_D$) for the ligand-analyte pair.

With multiple analytes to investigate, the chip surface requires regeneration. A regeneration solution should efficiently disrupt the binding interaction between the
ligand and analyte, but not inflict any changes to the ligand, which may impact its subsequent ability to bind again. The regeneration solution should not also degrade the sensor surface so as to cause the immobilized ligand to detach.

Figure 2.9. SPR binding model on sensor surface
Cartoon representation of one binding partner (ligand) immobilized onto a sensor surface and the interaction with its binding partner (analyte). Figure from Biacore – Sensor Surface Handbook – BR-1005-71 AB (GE Healthcare).

Figure 2.10. Basic SPR experiment
Example binding profile (sensorogram) for two interacting partners in an SPR experiment. The segments below the curve describe the flowing solution passing over the immobilized ligand on the surface. Figure from Biacore – Sensor Surface Handbook – BR-1005-71 AB (GE Healthcare).
2.10.3 Ligand immobilization

SPR is a label-free technique where neither binding partner need be labelled, e.g., radioisotope or fluorophore labelling. Instead, fundamental to SPR is the immobilization of one of the binding partners onto the chip surface as ligand. There are many ways to achieve attachment, by covalent-, affinity tag- or capture partner-immobilization.

We describe four, but not exhaustive types of ligand capture techniques used in this work in Figure 2.11. Amide coupling covalently links the ligand to the surface through primary amine surface residues or nucleophilic groups forming an amide bond. Thiol coupling via disulphide is another covalent immobilization technique. Biotin capture requires the ligand to have a biotinylated linker group to bind to a streptavidin-coated surface; the biotin-streptavidin interaction is one of the strongest non-covalent interactions (Green, 1975) lending itself well to SPR immobilization. His-capture works for proteins with a poly-Histidine tag to co-ordinate with a Ni$^{2+}$ cation. Antibody capture relies on firstly capturing an antibody protein onto the surface via any of the aforementioned techniques and then subsequently capturing the ligand containing the complimentary antibody epitope.

Direct coupling through amide or thiol requires no prerequisite tags, thus allowing the use of an unmodified ligand. However, direct coupling does not guarantee all the captured ligand will orient in a uniform manner, given that any lysine or cysteine could be conjugated, thus rendering some of the ligand binding sites inaccessible. With Biotin, His and Anti-body capture techniques, predictable orientations will be achieved. These latter non-covalent techniques do require additional affinity tags, which need to be factored in during the construct design in the molecular-biology phase. On the other hand, for biotinylation coupling kits are commercially available, but these have the potential for compromising the ligand as noted for amine coupling to the chip surface. However, synthetic ligands like nucleic acids and short peptides can be purchased ready conjugated to biotin.
Amine capture

Biotin capture

His-Tag capture

Antibody Capture

Figure 2.11. Example of SPR ligand capture options
Amine coupling offers covalent attachment. Biotin and His- capture are alternative options requiring prerequisite affinity tags or epitopes. Antibody capture is a 2-step capture process, with an antibody followed by ligand capture sequence. Figure from Biacore – Sensor Surface Handbook – BR-1005-71 AB (GE Healthcare).
2.10.3.1 SPR chip surface chemistry and amine coupling

Throughout its time in the instrument, a continuous stream of buffer is passed over the gold chip (Figure 2.12A). Above a pH of 3.5 the dextran matrix (Shown in Figure 2.12B and magnified in 2.12 C, D, E) is negatively charged and it will concentrate positively charged ligands via electrostatic attractive forces. This process requires the pH of the buffer to be between 3.5 and the isoelectric point (pI) of the ligand to maintain a cationic protein surface on the latter (Figure 2.12D). In order to maximize ligand capture, a pH scouting protocol is used to identify the pH for maximum ligand attraction prior to activating the surface for covalent ligand capture. Biacore pH scouting is carried out in 10mM sodium acetate buffers in 0.5 unit increments.
Figure 2.12. CM5 Biacore chip schematics depicting carboxymethylated dextran surface and electrostatic attraction surface conditions
A = Schematic of a Biacore SPR CM5 sensor chip. B = Cartoon of gold layer with dextran matrix sitting on a glass layer. C = Shows no negative charge on the dextran matrix below a buffer pH of 3.5, therefore no electrostatic attraction to a positively charged ligand. D = When the pH of the ligand is above 3.5 but below the pI of the ligand, the dextran matrix will have a negative charge with the ligand positively charged and therefore both electrostatically attracted. E = When the pH of the buffer is above that of the pI of the ligand such that both the dextran matrix and the ligand are negatively charged and repel one another. Figure from Biacore – Sensor Surface Handbook – BR-1005-71 AB (GE Healthcare).

Amine coupling is based on amide bond formation between a carboxylic acid group on the chip and a primary amine on the ligand surface. In order to drive the formation of an amide bond, the OH of the carboxylic acid group must be activated into a better leaving group. This is achieved by conversion of the OH into a reactive succinimide ester group via EDC (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide) and then NHS (N-hydroxysuccinimide) additions to yield a stabilized but highly reactive NHS ester (Figure 2.13). Once activated, a ligand in favourable buffer conditions will be
conjugated to the carboxylic acid group and immobilized (Figure 2.14). Any remaining activated esters are reacted with ethanolamine.

**Figure 2.13. Carboxylic acid group activation to an NHS-ester**
A simplified carboxymethylated surface represented as a single COOH (I) group reacts with (II) EDC ((1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide)) to form an O-acylisourea intermediate (III) that further reacts with NHS (N-hydroxysuccinimide) (IV) to form a stabilised but reactive NHS-Ester and produce a urea by-product. Figure from Biacore – Sensor Surface Handbook – BR-1005-71 AB (GE Healthcare).

**Figure 2.14. Overview of ligand immobilization on activated chip surface**
The scheme depicts the activation of carboxylic acids (I) into reactive NHS-esters (II) and their subsequent conjugation to a primary amine on a ligand molecule (III). Figure from Biacore – Sensor Surface Handbook – BR-1005-71 AB (GE Healthcare).

### 2.10.3.2 Streptavidin – Biotin capture technique.
We describe this technique in our published methodology (Quareshy et al., 2016), refer to Appendix A (Method 3.1 and 3.2.).
2.11 Auxin activity screening

2.11.1 Assay run order and sampling
The auxin activity assay is based on the principles of the basic F-box protein activity assay described in Appendix A section 3.3. We use TIR1 as an example to describe the data processing, but the same method was applied to AFB5. The typical semi high-throughput activity assay consists of cycles (SPR runs) in the following order: (where “Protein” is a generic term used to describe an F-box receptor protein (TIR1 or AFB5, in complex with ASK1) being screened at that time).

3 x Buffer - Blank
Protein + IAA 1 (positive control)
Protein + DMSO 1 (negative control)
Protein + compounds 1-10
Protein + IAA 2 (positive control)
Protein + DMSO 2 (negative control)
Protein + compounds 11-20
Protein + IAA 3 (positive control)
Protein + DMSO 3 (negative control)

Repeat for every subsequent set of 10 compounds with final controls:

Protein + IAA
Protein + DMSO
3 x Buffer - Blank

2.11.2 Sample preparation
For 100 cycles 10mL of master mix was prepared with protein adjusted to achieve SPR binding response units (RU) of at least 300-800RU. For each experiment, the preparation required Protein + buffer + compound (50µM final concentration) to final volume of 100µL.
The relevant 10mM compound working stocks were allowed to thaw at room temperature before 1µL of each was dispensed into a 0.8 ml microvial followed by 199µL of the master mix giving 200µl final volume, sufficient to run duplicate
experiments for each compound. When 40 compounds were screened, for example, 4.5μL of DMSO (for the Negative control) and 4.5μL of a 10mM IAA stock (for the Positive control) respectively were dispensed into glass tubes followed by 895.5μL of the master mix.

2.11.3 Basis of the activity run
(With the TIR1 F-box protein as an example of a protein screened.) A pre-mixture of the TIR1 protein ± a compound was flowed over the degron peptide. Should a compound bind to the TIR1 protein as an auxin we then observe a binding response as the TIR1 + compound binds to the degron peptide as measured in Δ Response Units (RU) (Figure 2.15A red trace). If a compound does not bind to TIR1 as an auxin, it therefore will not enable TIR1 to bind to the degron and a non-specific binding response is observed (Figure 2.15A green trace).

2.11.4 SPR auxin activity assay reference subtraction
The data is “double reference subtracted” to factor in non-specific binding of the protein to the dextrose matrix and further non-specific binding to the Aux/IAA degron.

2.11.4.1 General methodology for reference subtraction
First referencing subtraction: As detailed in Appendix A section 3.3 we describe ligand immobilization on the two flow cells on the SPR SA chip; a control flow cell (Fc1) containing Biocytin and a sample flow cell (Fc2) containing the Aux/IAA degron. The Biacore software automatically subtracts Fc1 from Fc2 (Fc 2-1) to generate sensorgrams shown in Figure 2.15A. These curves are plotted on with units of ΔRU on the y-axis. Units of RU would be raw plasmons response values based on the immobilized ligand and buffer conditions at the chip surface.

We then normalize the data points for each curve (Appendix A, section 3.3; data processing step 6) so they have a baseline value of 0 ΔRU. The Red curve with TIR1 + IAA shows a binding profile, (Fc2-1) referenced and normalised. The control experiment of TIR1 + DMSO (in green) shows a non-specific binding response (Figure 2.15A). We implement the second reference subtraction manually in post-run processing to negate the non-specific binding of the TIR1+DMSO to the Aux/IAA
Figure 2.15. SPR results processing

Red = TIR1 + IAA control, Green = TIR1 + DMSO negative control and Blue = Buffer blank injection. **A**: SPR binding sensorgrams before background subtraction, showing the non-specific binding of just the TIR1 protein and DMSO to the degron. **B**: Sensorgrams after subtraction of non-specific binding.

This now shows the true binding response for the TIR1 + IAA.
2.11.4.2 Activity assay reference subtraction methodology

For each activity assay the second reference subtraction to remove the non specific binding and DMSO bulk shift was applied to all positive controls of TIR1 + IAA and compounds tested. During the assay, as described in Section 2.10, negative control experiments of TIR1 + DMSO were run at regular intervals. We typically reference subtracted the first positive control TIR1 + IAA experiment and the first 5 compounds against the first negative TIR1 + DMSO control. Compounds run in positions 6 – 15 and the second positive control of TIR1 + IAA were reference subtracted to the second TIR1 + DMSO control run. Compounds 6 – 10 would have been run before the control, whilst compounds 11 – 15 would have been run after the control. This pattern was repeated for every set of 10 compounds and succeeding TIR1 + IAA positive control subsequently. The final set of 5 compounds and final TIR1 + IAA positive control would have been referenced to the final negative control of TIR1 + DMSO.

Table 2.1. Reference subtractions for SPR processing

<table>
<thead>
<tr>
<th>Reference subtraction cycle</th>
<th>Sample cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIR1 + DMSO 1</td>
<td>TIR1 + IAA 1</td>
</tr>
<tr>
<td></td>
<td>Compound 1-5</td>
</tr>
<tr>
<td>TIR1 + DMSO 2</td>
<td>Compounds 6-10</td>
</tr>
<tr>
<td></td>
<td>TIR1 + IAA 2</td>
</tr>
<tr>
<td></td>
<td>Compound 11-15</td>
</tr>
</tbody>
</table>

Repeat for next 10 compounds

| TIR1 + DMSO (final)        | Preceding 5 compounds |
2.11.4.3 Data clean up in KNIME

The open source KNIME (Konstanz Information Miner) data manipulation platform was used to process the SPR data (Berthold et al., 2008). The data was exported from the Biacore software as a “.CSV” file and imported into the KNIME workflow shown in Figure 2.16. This workflow, removed all but one of the surplus and redundant x-axis columns generated for every SPR run. The remaining x-axis column was renamed to Time (seconds). The workflow also removed the TIR1 + DMSO negative control experiments. The maximal values for each TIR1 + IAA positive control and compound experiment were extracted and added as an additional row.
Figure 2.16. KNIME workflow for processing and clean up of SAR activity data

The Figure shows the KNIME workflow implemented in Version 2.12. The workflow imports a .CSV text file and removes all but the first y-axis for each experiment. The remaining y-axis column is renamed as “Time (seconds).” The workflow also removes all protein + DMSO negative control run experiments. From the data, the maximal value of each positive control Protein + IAA and compound experiment is derived and added as a new row. The data is then exported as a text file for further processing.
2.11.4.4 Relative activity determination

For each compound we calculate the relative response to IAA. The data was imported into an automated spreadsheet template, which would display the results of the relative activity for each compound with its duplicate. The calculations were set up on the following basis: during the assay, positive control experiments of TIR1 + IAA were run at regular intervals, the purpose of this was to check the protein was still active and to use it as reference for the relative activity. We typically referenced the first 5 compounds to the first positive control run. This was achieved by working out the percentage response of each compounds’ maximal ΔRU binding value relative to the relevant positive control. Compounds run in positions 6 – 15 were referenced to the second positive control of TIR1 + IAA. Compounds 6 – 10 would have been run before the control, whilst compounds 11 – 15 would have been run after the control. This procedure was repeated for every set of 10 compounds thereafter. The final set of 5 compounds would have been referenced to the final positive control of TIR1 + IAA. The data was collated into as master-record spreadsheet.

The methodology was applied to the following publication (Steenackers et al., 2016)

<table>
<thead>
<tr>
<th>Positive control reference</th>
<th>Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max ΔRU of: TIR1 + IAA 1</td>
<td>Compound 1-5</td>
</tr>
<tr>
<td>Max ΔRU of: TIR1 + IAA 2</td>
<td>Compounds 6-15</td>
</tr>
</tbody>
</table>

Repeat for next 10 compounds

| TIR1 + IAA (final) | Preceding 5 compounds |

Table 2.2 Reference subtractions for SPR processing
2.12 Anti-auxin activity

The anti-auxin activity measurement is similar to the typical auxin activity assay (section 2.11) and is processed and analysed in a similar manner. The only exception is that the assay identifies compounds that bind to TIR1 or AFB5 but prevent the complex from binding to the Aux/IAA degron. Anti-auxin compounds have been reported (Hayashi et al., 2008a, 2012). From this work, 2-(1H-indol-3-yl)-4-oxo-4-phenylbutanoic acid (PEO) was shown to be a potent anti-auxin.

The assay works on the principle that if a compound binds to TIR1 and does not subsequently bind to the Aux/IAA degron then it is likely to be an anti-auxin. Two control samples are used:

1) Protein + IAA (5µM) – sets the expected response units
2) Protein + IAA (5µM) + PEO (50µM) – known anti-auxin control
   (Protein = TIR1 or AFB5 in complex with ASK1)

Figure 2.17 shows the results of the control assays with the addition of PEO lowering the overall response. The PEO is present in 10-fold higher concentration than IAA and is sufficient to observe a significant reduction in activity.

![TIR1 anti-auxin assay control runs](image)

**Figure 2.17. Control runs for the anti-auxin activity**
The figure shows the basis of the anti-auxin activity assay. The blue dashed line represents the baseline control for the assay with the response from the protein + IAA (5µM) run. The green line represents the reduction in TIR1-IAA binding by a known anti-auxin compound like PEO, which reduces the binding response significantly compared to the baseline control.
We have developed an assay to quickly screen for compounds that act like PEO. In addition to the control runs mentioned before the samples being investigated are run as: Protein + IAA (5µM) + compound sample (50µM). Example results are shown in Figure 2.18 with Compound 1 (an active auxin) adding to the baseline signal (IAA Control) observed. This is the case because 5µM of IAA alone does not saturate the binding response. The presence of additional auxin-like compound will amplify the signal and give a higher binding response. Compound 2 demonstrates anti-auxin activity as it reduces the baseline signal, similarly to the PEO control. Compound 3 is not an anti-auxin like compound and does not impact the binding response giving a similar profile to that of the IAA control.

![TIR1 Anti-auxin assay](image)

**Figure 2.18. Outcomes of the anti-auxin assay**  
The results of the assay are shown with 3 examples. Compound 1 is an auxin-like compound amplifying the binding response relative to the IAA control (dashed lines). Compound 2 shows anti-auxin activity as it lowers the binding signal, whilst compound 3 is neither an auxin nor an anti-auxin, as it does not impact the binding response.
Chapter 3: IAA binding to TIR1
3.1 Introduction

In this chapter and the next we discuss our work towards observing direct binding of IAA [and other auxins] to the TIR1 receptor. To-date there is little work that looks at this interaction to derive binding or thermodynamic constants. The majority of in-vitro auxin F-box receptor studies observe the assembly of the co-receptor complex, providing an indirect measure of binding to TIR1, such as co-immunoprecipitation assays, radiolabelled binding assays and yeast-2-hybrid interaction assays (Kepinski & Leyser, 2005; Calderón Villalobos et al., 2012; Simon et al., 2013; Lee et al., 2014). There is a need to understand auxin binding to TIR1, as this is the fundamental underpinning of the auxin-signalling pathway.

There have been questions raised as to whether auxin F-box receptors can bind IAA alone without the presence of an Aux/IAA protein (Calderón Villalobos et al., 2012; Shimizu-Mitao & Kakimoto, 2014). This idea gives equal prominence to the F-Box and Aux/IAA protein partners in driving this co-receptor formation. In addition there is evidence that some of the 2 partners interact weakly in the absence of auxin (Calderón Villalobos et al., 2012).

By studying the crystal structures with IAA bound to TIR1 (2P1P) and TIR1 + IAA + Aux/IAA degron (2P1Q) (Tan et al., 2007) there were no significant conformational changes to the IAA molecule nor in the binding residues interacting with IAA. It is our opinion that IAA [or any other auxin] must first successfully bind to an auxin F-box receptor in a correct pose to allow the Aux/IAA to then bind. Furthermore the 2P1P crystal structure shows TIR1 crystallised with just IAA and leads us to postulate that TIR1 and the other F-box receptor members are capable of binding auxin molecules independently of Aux/IAAs.

It is our view that auxin is critical in facilitating the SCF complex formation by enriching the TIR1 receptor and thus expediting the interaction with Aux/IAA as is necessary for a quick auxin response to occur.
3.1.1 Techniques for studying small molecule binding

Well established techniques for studying small molecule direct binding to a protein target include Nuclear Magnetic Resonance (NMR)-based WaterLOGSY (Water Ligand Observed via Gradient Spectroscopy), Thermal Shift Assays (TSA), Isothermal Titration Calorimetry (ITC), X-ray crystallography, Mass Spectrometry (MS), Surface Plasmon Resonance (SPR) and in silico techniques including docking (Scott et al., 2012). There is no one perfect technique to study small molecule binding. We wished to utilise a technique with low sample demand, rapid sampling, local instrument availability and expertise.

Our approach to observe the direct binding of IAA to TIR1 was to firstly explore the use of SPR as means of observing small molecule binding to the TIR1 protein in vitro. Secondly we employed in silico docking (presented in chapter 4), as it required no additional instrumentation or expensive resources and had not been employed in the literature as a means to study auxin binding to TIR1. Recombinant expression and purification of TIR1 receptor protein from insect cells provided us material for in vitro studies whilst the crystal structures of TIR1 were readily available for docking.
3.2 Recombinant TIR1-ASK1 expression and purification

Consistent production and purification of recombinantly expressed TIR1 protein was achieved using a baculovirus insect cell expression system from literature based protocols (Calderón Villalobos et al., 2012; Lee et al., 2014). We further optimised their protein purification method and devised a 2-step affinity tag chromatography methodology, introducing a TEV protease cleavage step in-between to remove the large His-MBP affinity tags.

The *Trichoplusia ni* High Five cell line (referred to as “High Five”) was adapted into suspension cultures and infected with recombinant baculovirus containing TIR1 and ASK1 constructs. ASK1 was co-expressed based on the postulation that it aids the F-box proteins fold correctly and maintain their active shape (Tan et al., 2007). The proteins were under the control of separate baculovirus promoters during expression. Recombinant virus production was contracted to Oxford Expression Technologies (OET) to yield high titre virus ready to infect High Five cells.

The construct for TIR-ASK1 encodes proteins possessing the following tags and cleavage sites:

\[
\text{10 x His} - \text{MBP} - \text{TEV} - \text{FLAG} - \text{F-Box Protein} \\
\text{10 x His} - \text{TEV} - \text{ASK1}
\]

The 10xHis, MBP and FLAG affinity tags are for purification of the protein once expressed; MBP also serves as a folding and solubility scaffold.

3.2.1 Cell culture

Adaptation of the High Five cells into suspension was achieved after experimenting with various shaking speeds, for the current incubator, 115 rpm was ideal; fast enough to maintain suspension without cell precipitation or cell shearing (Wu et al., 1990; Zheng et al., 2014). We were able to achieve between 1 – 4 litres of culture per week whilst maintaining a stock culture for continuation. Cells double in number every 18-24 hours, indicative of healthy cells. Furthermore we achieved cell viabilities of >99% with trypan blue stain observations (not shown).
3.2.2 Virus to cell ratio determination

Test expression titrations identified the optimal virus to cell ratio for maximum protein expression (Figure 3.1) at when the virus titre was not established. The western blot was developed with anti-His antibody. The optimal volume of virus added for TIR1 protein production (~110 kDa) was shown to be 20µL of virus per 5mL of Hi5 cells at 1.0 x 10^6 cells/mL. There are bands at ~25 kDa and ~40 kDa. The former band refers to His-ASK1 protein, the latter is likely to be a suspected truncated His-MBP protein product.

Overall the purpose of such titrations is to establish the optimal set-up for large-scale expression. We then investigated the protein expression timeline. Figure 3.2 shows a time course for TIR1 expression at 12-hour intervals post infection. TIR1 production is evident after 24 hours but increases until 48 hours; we also observe a “truncated product” (approx. 40 kDa) after 24 hours. We selected 72 hours as the optimal harvest time.

![Figure 3.1. Optimisation of TIR1 expression in T. ni Hi5 cells](image)

Anti-His western blot for TIR1 virus titrations. 20µL was chosen as the optimum.
Figure 3.2. Time course expression profile for TIR1-ASK1 in insect cells
Western blot showing time course of TIR1-ASK1 protein expression at 12-hour intervals, developed with anti-His antibody. TIR1 expression commences at the 24-hour time.

3.2.3 Tandem Affinity Protein Purification

The cell lysate was loaded onto a His affinity column (Roche®). This column is employed on the same principles as traditional Immobilised Metal Affinity Chromatography (IMAC) (Hochuli et al., 1988; Hengen, 1995; Lichty et al., 2011) but the resin requires no recharging of the metal ions. Furthermore, reducing agents such as DTT (Dithiothreitol) or TCEP ((tris(2-carboxyethyl)phosphine)) and chelating agents such as EDTA (Ethylenediaminetetraacetic acid) were compatible with this column, which were advantageous to maintain protein stability.

In Figure 3.3A we track the His column purification from an SDS-PAGE gel showing total protein visualised under UV light. The western blot (Figure 3.3B) was developed with anti-FLAG antibody and shows some TIR1 protein leaking in FT1 to FT3, perhaps due to misfolded protein and occluded His epitope. The flow rate was maintained at ~1mL/min through 5mL of resin. Non-specific protein is washed off in W1 and there is no obvious protein in W2 (note, we collected up to W5 but the remaining fractions are not shown). TIR1 and ASK1 proteins are eluted in a buffer with 250mM imidazole in the first 20 mL. It is also evident other truncated
proteins/fragments (red arrow) also co-elute; hence a single affinity step was not adequate to obtain pure TIR1.

Material eluted from the His column was incubated with TEV protease. Cleavage is efficient, with total cleavage obtained in 2 hours (TEV 120 mins; Figure 3.4A), but this step can also be left overnight (TEV ON). We observe a change in the TIR1 protein mass from ~110kDa to ~70kDa corresponding to the mass of FLAG-TIR1 and accounts for the loss of the His-MBP tags of around 40kDa. The western blot is developed with anti-FLAG, thus only FLAG-TIR1 is detected (Figure 3.4B).

When the His column eluate is loaded onto a FLAG affinity column, the suspected-truncated product ~40kDa (Figure 3.4A) is lost. In addition, the TEV enzyme is washed away. Some TIR1 protein loss was also evident in 4th and 5th flow through fractions (FT4 and FT5). Elution from the FLAG column was monitored in 2mL fractions (Figure 3.5A). We observe protein in fractions Elution 2 (E2) up to (E6), with the majority in fractions E4 and E5. We observe only 2 bands, at the expected molecular weight for TIR1 (~70kDa) and ASK1 (~20kDa) proteins respectively. The “Excess elution” sample (Figure 3.5A) involved adding an additional 3 X FLAG peptide elution aliquot after E10 to check for leftover protein. The gel showed that all the protein was eluted at the first instance. The protein concentration was estimated using a NanoDrop at A$_{280}$nm.
Figure 3.3. Phase I of TIR1/ASK1 purification on a His affinity column
A. SDS PAGE gel showing His affinity chromatography profile, B: Western blot of A with anti-FLAG antibody. Column-volume 5mL. L (Lysate), FT (Flow through), W (Wash fractions) and E (Elution). Fractions were collected every 10 mL (2 column volumes). Flow rate was ~ 1mL/min by gravity flow.
Figure 3.4. Phase II of TIR1-ASK1 purification; TEV protease digest and FLAG affinity column loading and washes

A. SDS page gel of TEV digest and FLAG affinity. B. Western blot of A with anti-FLAG antibody. Tev # (Tev cleaving step, # = minutes), FT (Flow through) and W (Wash) fractions. Fractions were collected every 10 mL (2 column volumes). Flow rate was ~ 1mL/min by gravity flow.
Figure 3.5. Phase III of TIR1-ASK1 purification with FLAG column elutions of purified TIR1-ASK1
A. SDS-PAGE gel of TEV digest and FLAG affinity column elution fractions. B. Western blot of A with anti-FLAG antibody. E (elution), Excess Elution describes an additional column wash with elution buffer to check for any residual protein.
3.3 Observing direct binding of IAA to TIR1 by SPR

3.3.1 Biacore and introduction to SPR
Direct observation of IAA binding to the TIR1 receptor in the absence of degron peptide had yet to be achieved. In order to realise we decided to utilise the biophysical technique Surface Plasmon Resonance (SPR). The Biacore SPR technology from GE Healthcare is considered the gold standard and state of the art instrumentation. It offers a variety of immobilisation methods with a wide range of attachment surfaces and chemistries available. Their SPR detection sensitivity has improved greatly over time, and currently offers the ability to detect binding of small molecules down to 100 Da. The locally available Biacore T200 instrument was ideally suited for this project. We aimed to immobilise TIR1-ASK1 onto a chip surface and flow over IAA and other auxins to observe direct binding events.

3.3.2 TIR1-ASK1 immobilisation via amine coupling
With a purified TIR1-ASK1 complex, the results of the pH scouting are shown in Figure 3.6. The steepness of the curve after the 0-second mark indicates the accumulation of protein onto the surface. We observe no such accumulation for pH 5.5, but observe an increase in accumulation as the pH increases by 0.5 units. As explained in section 2.8.2, sufficient protein immobilisation by amine coupling is reliant on the protein possessing a global positive charge and is thus attracted to the negatively charged dextran on the chip. The conditions at pH 4.0 demonstrated maximal protein accumulation. This is in agreement with the pI for this complex, estimated around 5.2 (using ExPASy online calculator (Gasteiger et al., 2005), thus at pH 4.0 and 4.5 the protein is positively charged. The weaker associations for pH 5.0 and 5.5 are consistent, as these pHs environments are close to and above the pI of the protein complex, thus not capable of inducing a positive surface charge on the protein. We chose pH 4.5 conditions, as it is a less acidic and harsher condition for the protein.
Figure 3.6. pH scouting on CM5 chip with TIR1-ASK1 protein
The graph shows the association profiles of the TIR1-ASK1 protein between pH 4.0 and 5.5. The dashed magenta line represents pH 4.0, the best condition for accumulating protein as it shows the quickest association gradient.

TIR1-ASK1 was immobilised onto a CM5 chip surface in a pH 4.5 10mM sodium acetate buffer, using a manual injection mode (Figure 3.7). In this mode we controlled the loading time of the TIR1-ASK1 protein once the surface was activated. We achieved a total loading of 14,453 RU after the ethanolamine blocking. The loading equates to a surface concentration of 14ng/mm² based on the specification from GE.
The maximum theoretical observable RU response for IAA binding to TIR1 was modelled using equation 3.1, based on the known ligand and analyte molecular weights (MW) and the immobilised ligand RU, giving a binding capacity of 29 RU.

\[
\text{Analyte binding capacity (RU)} = \frac{\text{Analyte MW}}{\text{Ligand MW}} \times \text{Immobilised ligand level (RU)}
\]

(3.1)

Analyte MW (IAA) = 175Da
Ligand MW (TIR1 + ASK1) = 86357Da
Immobilised ligand = 14,453RU

\[
\frac{175}{86,357} \times 14,453 = 29RU
\]

3.3.3 Attempts to observe direct binding of IAA to immobilised TIR1

IAA binding to TIR1-ASK1 was carried out in Single Cycle Kinetic (SCK) mode, where 5 consecutive sample injections in order of increasing concentration of IAA were passed over the immobilised protein followed by a dissociation phase. This mode enables rapid kinetic screening for binding interactions, providing dose dependency data, affinity, association \(K_a\) and dissociation \(K_d\) values from a single run. For each run a control was run using buffer.

Unfortunately, the results from this experiment revealed a negative drifting baseline (Figure 3.8). The marked injection points in red correspond to the additions of IAA. The blank (in red) consistently drifts downward, and the same drift was seen for the IAA trace (blue). We do however, observe dose depending binding curves after the 12.5\(\mu\)M, 25\(\mu\)M and 50\(\mu\)M injections indicating IAA binding. The RU units drop significantly after the 25\(\mu\)M and 50\(\mu\)M injections, indicative of a bulk shift effect resulting from DMSO in which the stock solution of IAA was dissolved. Even after subtracting the buffer baseline from IAA, the drifting baseline is still obvious, and the DMSO bulk effects result in sudden dips before and after the injection.
Figure 3.7. TIR1-ASK1 amine coupling onto CM5 chip
A sensogram showing the response units (RU) for semi-automated amine coupling of the TIR1-ASK1 complex onto the surface of a Carboxymethylated dextran surface, activated by EDC/NHS chemistry. The reagent prep, injections and loading were all user specified, ad-hoc during the experiment.
Figure 3.8. Single cycle kinetics for IAA vs. TIR1-ASK1
Both traces show the RU units decreasing over time, suggesting that TIR1 or ASK1 (or both) is bleeding from the chip surface. After each injection of IAA (blue arrows) we observe small binding curves, but these signals are lost in the sudden steps in the sensorgrams, which are due to the presence of DMSO. The complimentary blank buffer baseline control is shown in red.

Figure 3.9. Baseline correction for Figure 3.8
A repeat of the SCK from Figure 3.8 with a baseline double correction shows a dose dependent binding curve profile after each compound injection (blue arrows). The response units are very low due to the continual loss of the immobilized protein.
Repeating the experiment, we added DMSO into the running buffer and were able to avoid the bulk-shifts observed (Figure 3.10). We were able to demonstrate dose dependent IAA binding to TIR1-ASK1 when correcting for baseline drift. It was not possible to get robust kinetic data with such baseline drift.

Diagnostic tests revealed no issue with the optics or the baseline for other SPR chips. We ultimately focused on the immobilisation step and discovered the immobilised TIR1-ASK1 protein was falling off immediately after the immobilisation at -1.75RU s\(^{-1}\) (Figure 3.11). Amine coupling does not selectively immobilise the complex through TIR1 or ASK1 exclusively, and possibly only one of the two was immobilised onto the dextran matrix.

We tested different batches of CM5 chips with no further success. We employed a CM7 chip with the aim of concentrating more protein onto the higher density dextran surface to increase the local concentration of TIR1-ASK1. Despite loading 17,977RU, this did not rectify the receding baseline drift (Figure 3.12) and consequently we abandoned direct-amine coupling techniques to focus on alternative questions.
Figure 3.10. Repeated single cycle kinetics for IAA binding to TIR1-ASK1
A repeat of the SCK with baseline correction shows a dose dependent binding profile after each injection (red arrows).

Figure 3.11. Drifting baseline of TIR1-ASK1 post amine coupling
The graph on the left is from Figure 3.7, with a magnified region of the graph on the right showing a decreasing baseline ($\Delta$RU) at a rate of approximately -1.75RU s$^{-1}$ hidden amongst the wide-ranging y-axis.
Figure 3.12. Amine coupling of TIR1-ASK1 to CM7 chip
The SPR sensorgram shows the amine coupling of TIR1-ASK1 onto a CM7 chip, achieving 17,977 RU units of immobilisation, but post-immobilisation, the magnification inset shows the recurrence of the baseline receding.

3.3.4 TIR1-ASK1 immobilisation via antibody capture
An anti-FLAG antibody was successfully immobilised in a pH 4.5 buffer, with no drifting baseline observed post-immobilisation (Figure 3.13), however the TIR1-ASK1 protein was not captured onto the immobilised antibody (Figure 3.14). TIR1 activity was established prior to the experiment, indicating the protein was active, suggesting that the FLAG tag was masked in the SPR buffer or the anti-FLAG antibody had deteriorated in the coupling reaction. We were also unsuccessful in capturing the complex by the His-capture method.
Figure 3.13. Amine coupling capture of anti-FLAG antibody
The graph shows the anti-FLAG antibody captured onto a CM5 chip, with the post-immobilisation baseline stable as presented in the magnified sector of the graph shown as an inset.
Figure 3.14. TIR1-ASK1 capture onto anti-FLAG antibody
The sensogram shows 2 sequential injections of TIR1-ASK1. We failed to capture the TIR1-ASK1 complex onto the anti-FLAG antibody on the CM5 chip surface as the protein complex dissociated off slowly after each injection.

3.4 Discussion
We have described our unsuccessful attempts to study auxin binding to TIR1 using SPR methodology. The technique was a promising prospect, hampered due to the non-covalent nature of the TIR1-ASK1 complex. We did observe IAA binding to TIR1 despite the drifting baseline but were not able to derive any meaningful information. In order to progress with SPR we postulate that a covalently bound TIR1 to ASK1 via cross-linking or a stabilised, ASK1 independent TIR1 analogue would be necessary. We are pursuing ITC and thermal shift assays as alternatives to study the binding of small molecules to TIR1 and other auxin F-box receptors. ITC in particular would be promising as a label- and moreover, immobilisation-free technique. (Ladbury, 2010; Scott et al., 2012).
Chapter 4: A Novel *in silico* docking approach postulates mechanism of IAA perception and active compound differentiation for the TIR1 receptor.
4.1 *In silico* docking algorithm

In this chapter we focus on *in silico* virtual screening as methodology for studying the binding of IAA to TIR1 employing a flexible ligand, rigid protein-docking algorithm to assess the method. Docking of IAA had been previously reported by (Hayashi et al., 2008b, 2012) in the design and successful development of anti-auxins against the TIR1 receptor. Work by (Calderón Villalobos et al., 2012) docked IAA and picloram structures into the TIR1 crystal structure as well as a homology model for AFB5. Apart from these works there has not been any other concerted *in silico* novel auxin discovery investigation reported.

4.2 Validation of AutoDock Vina docking algorithm

We validated the AutoDock Vina algorithm for TIR1 using the 2P1Q crystal structure. The algorithm was tasked with docking IAA into the binding region at the bottom of the TIR1 pocket, with the crystallographic IAA molecule excised and using a 10Å³ search space. The algorithm was able to find a pose matching that of the crystal structure. The best scoring docked pose matched the crystallographic pose very closely, almost superimposable onto the IAA molecule shown from various points of view in Figure 4.1. Repeating the docking with the 2P1P structure (lacking the Aux/IAA degron) gave a similar result with IAA docked in a superimposable manner to the crystal structure.

We extended the validation to assess if docking would serve as a predictive tool for auxin activity against TIR1. Previous work by (Lee et al., 2014) measured auxin-like activity for 58 compounds by SPR for TIR1 and AFB5 receptors. The range of compound activities ranging from >100% to 0%, were well suited to test the docking algorithm on its ability to predict auxin activity for TIR1 by means of correlation to the SPR results. An automated implementation of the AutoDock Vina algorithm (Price et al., 2014) docked a library of compounds iteratively into the receptor, generating a set of 10 poses and binding scores for each compound. We chose the highest docking score for each compound and ranked them to compare with the SPR results (Table 4.1).
There was a poor correlation of the ranked SPR data with that of the docking, (Figure 4.2). For IAA, the results compare well, as does Clopyralid, but for other known auxins like Fluroxypyr, 2,4-D, and 1-NAA we observed large differences between the rankings. Conversely, compounds like Tryptophan and Quinclorac were ranked higher by the docking experiments relative to their SPR activity.

We were intrigued, in particular by the tryptophan result from Table 4.1. Structurally similar to IAA (Figure 4.3) it was predicted to be a strong binder to TIR1, yet it has no auxin activity. It raised the question: How is Tryptophan prevented from getting to the TIR1 binding site given that it COULD fit in there as demonstrated by the docking?
Figure 4.1. Results of AutoDock Vina docking of IAA into TIR1. Green IAA = from crystal structure, White IAA = docked result from AutoDock Vina. TIR1 crystal structure residues shown as wire sticks. A = top down view of IAA crystal structure, B = with superimposed docked IAA C & D = side on view of panels A & B E & F = angled rotated view of panels A & B. Figure generated in PyMol
Table 4.1 Summary of SPR vs. Docking results for selected compounds

<table>
<thead>
<tr>
<th>Compound</th>
<th>TIR1 % Binding</th>
<th>Docking score (kcal/mol)</th>
<th>SPR Ranking</th>
<th>Docking Ranking</th>
<th>Ranking difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAA</td>
<td>100</td>
<td>-8.1</td>
<td>1st</td>
<td>3rd</td>
<td>-2</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>0</td>
<td>-7.9</td>
<td>37th</td>
<td>5th</td>
<td>32</td>
</tr>
<tr>
<td>4-OH-benzoic acid</td>
<td>3</td>
<td>-6.2</td>
<td>31st</td>
<td>48th</td>
<td>-17</td>
</tr>
<tr>
<td>Tryptamine</td>
<td>7</td>
<td>-6.9</td>
<td>26th</td>
<td>28th</td>
<td>-2</td>
</tr>
<tr>
<td>2,4-D</td>
<td>78</td>
<td>-6.7</td>
<td>4th</td>
<td>31st</td>
<td>-27</td>
</tr>
<tr>
<td>NAA</td>
<td>76</td>
<td>-7.4</td>
<td>5th</td>
<td>15th</td>
<td>-10</td>
</tr>
<tr>
<td>Fluroxypyr</td>
<td>80</td>
<td>-6.6</td>
<td>2nd</td>
<td>37th</td>
<td>-35</td>
</tr>
<tr>
<td>Phenylacetic acid</td>
<td>55</td>
<td>-6.6</td>
<td>9th</td>
<td>34th</td>
<td>-25</td>
</tr>
</tbody>
</table>

Figure 4.2. Correlation plot comparing Docking vs. SPR rankings.

No correlation was observed between SPR binding and in silico-docked results. Compounds selected for Table 4.1 are highlighted and labelled in red. Inset table shows Pearson r test.
parameters. The R-Squared value 0.0035 indicates no correlation. The red line indicates ideal positive correlation.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Top view</th>
<th>Side-view</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptophan</td>
<td><img src="image" alt="Top View" /></td>
<td><img src="image" alt="Side View" /></td>
</tr>
</tbody>
</table>

**Figure 4.3 – The best-docked poses for IAA and tryptophan.**

Best scoring poses of docked tryptophan (Green), relative to IAA (White) from top down and side views. Tryptophan was docked and appears to superimpose onto IAA, raising the question as to why it is not an active auxin, or anti-auxin. Figure generated in Chimera.
4.3 The auxin binding site; how does auxin get there? – Development of novel docking application – TomoDock

We postulated that IAA and other auxins must therefore successfully navigate the TIR1 pocket (Figure 4.4 A) in order to reach the binding site below in a preferential manner as dictated by the side chain groups along the way. In order to study this we adapted the docking protocol to “scan” the pocket from top to bottom. A collective effort with co-workers lead to a multidisciplinary (Biology, Chemistry, Computational Science and Mathematically) approach to study compound entry into the TIR1 pocket, for which we developed an application named “TomoDock”.

By analogy with medical tomography, our method scans down the binding pocket layer-by-layer. Using an automated series of docking experiments the docking box moves down 1Å and at each step the docking is repeated. For TIR1 we used a docking search-space of 18 Å³ a total of 15 times (Figure 4.4 D). The box dimensions are much larger than the binding pocket (Figure 4.4 B & D) to enable spatial freedom for the compound being docked. At each successive step the grid box expose residues closer to the pocket bottom, creating a pseudo-time simulation of the movement of a compound downward towards the bottom of the TIR1 pocket. TomoDock reveals the influence of residues as they appear in the search space, as illustrated graphically in Video 4.1 (on attached CD media).

For each step Autodock Vina generates a specified number of docked poses with the respective docking energy score. The poses are studied visually to elucidate a probable mechanism of transit from the top of the pocket to the bottom. The best-scored poses were considered first, but in some cases successive transitions were impossible. Therefore, in order to make a logical stepwise progression, the second best pose was selected as more appropriate in some cases.

We will present the initial work and proof of concept here and note that the methodology has been accepted for publication in Open Biology, explaining auxin activity for a set of indole molecules. The manuscript is attached in Appendix B.
Figure 4.4 TomoDock grid box movement
The pocket forming residues are shown as a white surface representation, with the remaining residues as sticks. The docking search space is shown as a cuboid with x-plane in Red, y-plane in Green and z-plane in Blue. Panel A shows the binding-pocket-forming residues represented as a surface from the front. Panel B is a rotated view with a birds-eye view looking at the top of the pocket with the docking-grid cube box in Blue (z axis) over the pocket opening. Panel C is similar to A but the non-pocket forming residues are hidden, and the grid cube box is at the first TomoDock position (Slab 1). Panel D shows the cube box after 15 successive 1Å movements towards and past the bottom of the pocket (Slab 15). Figure generated in AutoDock Tools.
4.3.1 TomoDock results

Our initial analysis focused on the 2P1P TIR1 crystal structure, with the IAA co-crystallised but lacking the Aux/IAA degron. We excised the co-crystallised IAA for the experiment. The top 3 best scoring TomoDock results (Table 4.2) demonstrate an energy landscape favouring the movement of the molecule down towards the pocket bottom, graphically represented in Figure 4.5. This was insightful and appears to be a logical approximation of IAA’s movement, given that IAA naturally binds TIR1 and would need a set of favourable transitions to reach the binding site. For some slabs the best scoring result shows a significant difference to the 2nd and 3rd best scores. This is evident at boxes 4 & 5, and 10 to 15 and is a result of the alternative poses found by the docking experiment.

For the best scoring function we observe 4 distinct phases of IAA transition between slabs: i] 1 to 4 ii] 5 to 7 iii] 8 & 9 and iv] 10 to 15. For slabs 1& 3 we observe the same scores between the 1st and 2nd results where the algorithm found 2 different poses; both equally plausible.

<table>
<thead>
<tr>
<th>Slab</th>
<th>Best docked result scores (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
</tr>
<tr>
<td>1</td>
<td>-4.1</td>
</tr>
<tr>
<td>2</td>
<td>-4.3</td>
</tr>
<tr>
<td>3</td>
<td>-4.5</td>
</tr>
<tr>
<td>4</td>
<td>-5.2</td>
</tr>
<tr>
<td>5</td>
<td>-5.1</td>
</tr>
<tr>
<td>6</td>
<td>-5.1</td>
</tr>
<tr>
<td>7</td>
<td>-5.1</td>
</tr>
<tr>
<td>8</td>
<td>-5.7</td>
</tr>
<tr>
<td>9</td>
<td>-5.9</td>
</tr>
<tr>
<td>10</td>
<td>-6.8</td>
</tr>
<tr>
<td>11</td>
<td>-6.8</td>
</tr>
<tr>
<td>12</td>
<td>-6.8</td>
</tr>
<tr>
<td>13</td>
<td>-6.8</td>
</tr>
<tr>
<td>14</td>
<td>-6.8</td>
</tr>
<tr>
<td>15</td>
<td>-6.8</td>
</tr>
</tbody>
</table>
Figure 4.5. TomoDock binding transect for TIR1
Results of the TomoDock experiment, observing IAA entry into the 2P1P binding site. We observe decreasing energy scores moving from the top of the pocket to the bottom. The decrease in energy indicates the bottom of the pocket is a more favourable region than the top and thus the approach the bottom is logical.
4.3.2 Postulated mechanism of IAA entry into TIR1

We studied poses for each slab, starting with the best scoring result for each, but also considering the next best scoring results as an alternative poses, in order to piece together a probable pathway. The proposed mechanism of IAA transitioning into the TIR1 binding site is illustrated in Video 4.2 (on attached CD media). We describe the rationale for the proposed mechanism in section 4.3.3, but we firstly describe it in a series of steps:

**Step 1.** IAA approaches the top of the pocket (slab 2- Figure 4.6 A) with the carboxylic acid (COOH) tilting towards a small niche in the TIR1 receptor, the indole nitrogen points towards the pocket bottom.

**Step 2.** The IAA COOH group moves a little further into the mouth of the niche (slab 3- Figure 4.6 B) with the indole-ring in the same orientation as in slab 2.

**Step 3.** IAA is much further into the niche, with the COOH fully enclosed (slab 4- Figure 4.6 C).

**Step 4.** IAA moves further down into the pocket, led by the indole ring towards the binding site (slab 7- Figure 4.6 D), with the COOH around the mouth of the niche.

**Step 5.** With the carboxylic acid still facing upwards, but now out of the niche, the indole ring moves further down (slab 8- Figure 4.6 E) towards the base. The indole ring is oriented perpendicular to that of the IAA crystallographic pose Figure 4.7A.

**Step 6.** A hemi-somersault movement of the molecule such that the indole ring now aligns with the IAA crystallographic pose (slab 9 Figure 4.6 F and from another point of view in Figure 4.7B).

**Step 7.** The carboxylic acid rotates around the CH₂ to face the bottom of the pocket. This final transition matches the IAA crystallographic pose (Figure 4.7C) where we observe both IAA models superimposed onto one another. It is also possible to arrive at this step directly from step 5 with step 6 as an intermediate transition pose.
Figure 4.6. Postulated mechanism of IAA movement into TIR1 binding pocket. The TIR1 receptor is shown in surface representation: Red = Oxygen, Blue = Nitrogen, Yellow = Sulphur carbon = beige. Panels A to F represent the docked poses from TomoDock at the respective slabs from the same point of view. In thin grey feature at the bottom of the pocket is the IAA crystallographic structure as a reference. Figure generated in Chimera.
Figure 4.7. Final 2 transitions of IAA into the binding pocket.
The same point of view is used for all 3 panels. The TIR1 receptor is shown in surface representation: Red = Oxygen, Blue = Nitrogen, Yellow = Sulphur, Carbon = Beige. Panel A shows the pose from slab 8 and how it flips over into (slab 9) in Panel B, this correctly aligns the aromatic rings to the crystal structure and requires only a tautomeric rotation around the α carbon to match the IAA crystal structure (light grey) almost identically.
4.3.3 Pose selection rational for IAA pathway into TIR1.

A summary of the mechanism steps, including the slabs and chosen poses are shown in Table 4.3. When overlaid onto the binding profile of the best scoring docked poses, we observe only small energy differences, and in no case was it necessary to select from lower than the second best scoring pose (Figure 4.8).

Slab 1 was not considered as the molecule was quite far away from the pocket (White arrow – Figure 4.9 A).

For slab 2, we chose the second best pose (Blue compound – Figure 4.9 B) over the best scoring pose (Green compound – Figure 4.9 B) for purposes of continuity, given the poses used in slab 3, then 4. The 2 poses in slab 2 have a small energy difference of 0.1kcal/mol between them and are mirror images of each other with their carboxylic acids meeting at the same point. Therefore we deemed it acceptable to choose the second best scoring pose.

Docked poses in slabs 5 and 6 were the same as slab 4 (Figure 4.9C), indicating no further residues introduced in these slabs affected IAA’s position and the molecule remained at this position (Figure 4.10), also signifying that this particular region may be a functionally-important part of the receptor. For slabs 7 and 8, we employed the same rationale as for slab 2 and chose the second best scoring poses for continuity of the mechanism.

The best scoring poses for slabs 11 to 15 were the same as for slab 10 and are shown to illustrate that the TomoDock transect has considered every feature relevant to binding, even below the base of the known site.
Table 4.3 TomoDock results of top three best scoring IAA poses for each slab

<table>
<thead>
<tr>
<th>Step</th>
<th>Slab</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-4.3</td>
<td>-4.2</td>
<td>-4.1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-4.5</td>
<td>-4.5</td>
<td>-4.3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-5.2</td>
<td>-4.5</td>
<td>-4.4</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>-5.1</td>
<td>-5</td>
<td>-4.9</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>-5.7</td>
<td>-5.6</td>
<td>-5.5</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>-5.9</td>
<td>-5.8</td>
<td>-5.8</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>-6.8</td>
<td>-5.9</td>
<td>-5.9</td>
</tr>
</tbody>
</table>

Energy profile of selected poses vs. best scoring poses

Figure 4.8 Docking energy profile of selected poses
The figure shows the docking energy profile for the proposed IAA mechanism (Red dashed line) overlaid into the energy profile for the best scoring docked poses. In choosing the second best scores for slabs 2, 7 and 8, we observe a small penalty to docking energy.
Figure 4.9. Rationale selection of poses from TomoDock experiments. The TIR1 receptor is shown in surface representation: Red = Oxygen, Blue = Nitrogen, Yellow = Sulphur, Carbon = Beige. Panel A: IAA in slab 1 is positioned (white arrow) far from the IAA binding pocket mouth, far away from the binding pocket mouth indicated by the yellow arrow. Panel B shows the top 2 scoring IAA poses in Slab 2. The second best pose (Blue) was picked over the best scoring pose (Green) for continuity of the mechanism pathway. Panel C shows superimposed, identical poses for slabs 4, 5 and 6. Figure generated in Chimera.
4.3.4 Molecular interaction analysis reveals key residues for IAA movement into TIR1.

We studied the molecular interactions between IAA and TIR1 amino acid residues with LigPlot analysis (Laskowski & Swindells, 2011) for slabs 2, 3, 4, 7, 8 & 9. Analysis of polar and hydrophobic interactions at each slab identified: Phe\textsubscript{78} and Phe\textsubscript{82} Cys\textsubscript{405} Lys\textsubscript{410}, Gly\textsubscript{441}, Ala\textsubscript{464}, Phe\textsubscript{465} and Arg\textsubscript{489} as key residues (Figures 4.10 and 4.11).

For interaction with the hydrophobic indole ring, Phe\textsubscript{465} is a key hydrophobic counterpart in slabs 2, 3 & 4 at the pocket of the mouth. Phe\textsubscript{78} and Phe\textsubscript{82} are important for slabs 7, 8 & 9, and Ala\textsubscript{464} plays a key role in all slabs. For polar contacts, Cys\textsubscript{405} stabilizes the indole nitrogen in slabs 2, 3 & 4. Lys\textsubscript{410} and Gly\textsubscript{441} backbone Carbonyl’s stabilize the COOH group of the IAA during the engagement in the small niche, at the mouth of the TIR1 pocket in slabs 3 and 4. Arg\textsubscript{489} provides polar stabilisation in slab 2 during the COOH group entry into the pocket and further stabilisation as the molecule navigates into the pocket bottom in slabs 7 and 8.

**Figure 4.10 LigPlots depicting molecular interactions between TIR1 and docked IAA in slabs 2 & 3**

Polar interactions are shown as dashed green lines with measured distances between atoms in Å. Eye-lash cartoons represent hydrophobic interactions with IAA in slab 2 (Panel A) and slab 3 (Panel B)
Figure 4.11 LigPlots depicting molecular interactions between TIR1 and docked IAA in slabs 4, 7, 8 and 9

Polar interactions are shown as dashed green lines with measured distances between atoms in Å. Eye-lash cartoons represent hydrophobic interactions with IAA in slab 4 (Panel A), slab 7 (Panel B), slab 8 (Panel C), slab 9 (Panel D)
4.4 TomoDock suggests a mechanism for excluding Tryptophan from the IAA binding site.

Tryptophan is not an auxin, yet docking it into the TIR1 binding site revealed it was able to adopt a similar pose to IAA and fit into the TIR1 receptor site. TomoDock has shown that tryptophan is perhaps not able to reach the binding site region as a possible result of steric constraints.

Tryptophan approaches TIR1 in a similar manner to IAA, with its polar group facing the niche (Figure 4.12 A). It then moves towards the binding site with the indole ring in a similar pose to step 4 for IAA engaging with the niche. However, the compound is unable to adopt a pose similar to that of IAA at step 5 (slab 8) or step 6 (slab 9). Instead, the docking algorithm finds a pose that satisfies the geometry of the binding site to accommodate tryptophan (Figure 4.12B). In slab 10, for step 7 of the pathway, tryptophan is docked in a pose comparable to IAA, but in transitioning from step 4 to step 7 no rational movements for tryptophan could be found to compare with the movement of IAA towards the binding site. The comparably larger amino acid group relative to the CH₂COOH of IAA possibly restricts the poses tryptophan could adopt in transitioning from step 4 to 7. The conclusion for tryptophan is that further progress towards the binding site is far less likely, resulting in the molecule leaving the receptor pocket. The transition from step 4 onwards can be considered critical in the passage of compounds to reach the binding site.
Figure 4.12 TomoDock results for tryptophan indicate a forbidden transition
The figure shows the critical stages of tryptophan movement into the pocket from 2 viewpoints (top and side). In going from step 4 to step 7, tryptophan is unable to transition in an energetically or geometrically feasible manner. The results for steps 5 and 6 were superimposable and offered no rational pathway between steps 4 to 7. Figure generated in Chimera.
4.5 Discussion

Our initial aim in utilising *in silico* docking was to screen compounds to identify novel auxin like molecules. The docking algorithm did dock IAA in the TIR1 binding site with the exact pose reported from the crystal structure (Tan et al., 2007), but we discovered it also docked non-auxin compounds like tryptophan. This result led us to question our presumption that the binding site at the base of the TIR1 pocket was the sole factor in auxin perception by TIR1. In developing TomoDock, we have been able to postulate a probable series of steps for IAA movement into the TIR1 receptor.

This sequence of steps in the proposed mechanism is plausible given the hydrophobic and hydrophilic features presented by residues that are present lining the sides of the pocket. TomoDock has shown IAA enters at the mouth of the pocket led by the COOH group in steps 1 to 3, where the negatively charged acid group neatly fits into the small niche, complementing the polar residues from backbone amides of Lys_{410} and Gly_{441} and the hydroxyl group of Ser_{440}. The niche helps orientate potential ligands, exposing the hydrophobic aromatic ring system to other features in the pocket in step 4, led by the indole ring for IAA. Below the niche is a band of hydrophobic residues that sits above the polar region of the pocket floor from residues Val_{463}, Phe_{79}, Phe_{82}, Leu_{378} and Phe_{352}. Given that the engagement niche has forced the aromatic ring system into this space, it is favourable for IAA to progress further in steps 5 to 7 and eventually adopt the final bound pose. In this pose, we observe the aromatic ring sits flush and in line with the hydrophobic residues band (Figure 4.13 A), and from an alternative view in Figure 4.13 B we see that the ring hovers above the pocket bottom consisting mainly of polar residues. Our proposed transition of the molecule into the binding pocket is supported by these observations and correlations to the hydrophobic mapping of the pocket.

For tryptophan, we have demonstrated that as a consequence of its bulkier side chain it is unable to reach the binding site even thought it could feasibly bind there. We envisage this methodology could also help us explain why other compounds are unable to bind TIR1 as auxins.

TomoDock has helped reveal key specific amino acids interacting with IAA as it traverses the receptor pocket down into the binding site. We are preparing analogues of TIR1, mutating Arg_{489} and Cys_{405} converting them to Alanine residues in single and
dual mutations. We postulate these are key amino acids enforcing the selective filtering of compounds determining which ones will progress to the binding site. We aim to observe the impact these mutations could have on auxin binding, postulating these two residues are crucial in determining if compounds can progress ultimately into the TIR1 binding site. They could possibly be acting as size-based selective filters for small molecules and thus regulate the specificity of auxin like compounds. It would be interesting to observe if the absence of one or both of these residues enables tryptophan to bind TIR1 as an auxin.

Previously reported molecular dynamic simulations (Hao & Yang, 2010) proposed Phe$^{82}$ and Phe$^{351}$ as key residues in the perception of auxin. Phe$^{82}$ was described to have undergone significant conformational change from unfavourable to favourable upon the binding of IAA, whilst Phe$^{351}$ underwent a similar change but upon the binding of Aux/IAA to the TIR1 + auxin complex. Phe$^{351}$ is located away from the auxin-binding site and was not identified in our TomoDock analysis. Accordantly, we also identified Phe$^{82}$ as an important residue in steps 4, 5 and 6 of the mechanism through hydrophobic interactions. However, Hao and Yang postulated a mechanism of auxin perception based on these two residues, with Phe$^{82}$ interacting with IAA in the first step. Whilst Phe$^{82}$ may indeed interact with IAA, TomoDock suggests that IAA other auxins are likely to encounter other residues further up the pocket prior to reaching this region.

TomoDock does not reinvent the docking process but rather it applies the principles of docking to study deep binding pockets like TIR1. Whilst the docking sequence is automated, subsequent analysis requires detailed study of the poses with their scoring functions in order derive a probable pathway. IAA is the benchmark compound to which other compounds can be compared as we have presented in Appendix B.
Figure 4.13. IAA in relation to the hydrophobicity of the pocket when bound. Blue = Hydrophilic regions. Black = Hydrophobic regions of the pocket. Panel A shows the pocket from a side on view, showing the polar COOH group in the foreground orientated towards the polar pocket bottom. In the background, we observe the aromatic rings aligned flush with the hydrophobic ring/band. Panel B is a top-down view of the receptor with crystallographic IAA. Figure generated in Chimera.
Chapter 5: The development of an *in vitro* spatial auxin pharmacophore model.
The chapter focuses on SAR investigations for auxin activity using SPR and root growth assays. We summarise the SAR results with the development of a spatial pharmacophore map of the TIR1 receptor.

5.1 Compound selection and SAR library design

We started with known active auxin scaffolds (shown in Figure 5.1) and identified regions around each molecule available for substitution, extension or branching. We aimed to collect and curate an SAR library with compounds selected for our methodical, iterative, compound-by-compound approach in order to understand what makes an auxin. Indolic-, phenoxy-, benzoic- and naphthalic chemical scaffolds were chosen, representing most families of auxinic herbicides. Other chemical scaffolds were added to the collection including non-aromatics.

We sought compounds with consistent substitutions for all sites around each scaffold, e.g. a methyl group at positions-1 to -7 to observe the implications of hydrophobicity. However we were limited by compound availability, and cost. We obtained a collection of single point substitutions around the indole ring, but it was not possible to collect a complete set for every position. For the phenoxy-acetic acids we were unable to acquire single substitutions at each position, and so instead we based our SAR around 2,4,-D as the foundation scaffold. We accumulated 317 compounds in total, including kind donations from: the MRC Fersht group (59 compounds), Syngenta (76 compounds) and Professor Ken-ichiro Hayashi (27 compounds). The remaining compounds had been collected from previous work in the group, or were purchased. The majority of the compounds (239) had carboxylic acid functionality. The other (71) compounds contained: nitro, amine, hydrazine, furan, triazole, tetrazole, hydroxyl, hydrotonin, thio-urea, acetonitrile, morpholine, thiomorpholine, sulphonyl and amide functionalities.
Figure 5.1 The main focus scaffolds for our SAR compound library.
The panel shows the scaffolds for our SAR compound library with IUPAC atom numbering convention for aromatic ring positions. The related side group positions are labelled as α and β in a non-conventional manner for simplicity. These numeral and alphabetical labels correspond to positions explored. Positions labelled such as “7a” are non-substitutable positions and were not investigated.
5.2 In vitro SPR method

Previous work in the group (Lee et al., 2014) developed a biophysical activity-screening platform using SPR. This assay exploits the co-receptor formation mechanism with a peptide sequence (degron) from the Aux/IAA7 protein immobilised onto the chip surface through a biotinylated tag. This degron sequence has been validated as sufficient to bind to TIR1 without the necessity for the entire Aux/IAA7 protein (Kepinski & Leyser, 2005; Tan et al., 2007; Pierre-Jerome et al., 2013; Lee et al., 2014)

5.2.1 Compound screening assay verification

Compound screening was performed using each compound at a concentration of 50µM, with a protein concentration selected to give RU units of binding between 300 to 1000 RU, so as to not saturate the binding response which is ~3000 ΔRU. Higher levels of binding wasted protein and resulted in unacceptable levels of non-specific binding. The aim of the screen was to profile compounds according to their activity relative to IAA established at 100%.

The immobilised biotinylated degron was found to have a limited lifetime, which added to the complexity to the logistics of screening. This degradation was unsystematic, sometimes surviving a screen of 40 compounds, other times up to 200 compounds, although deterioration was obvious during the course of an assay allowing intervention if necessary. Thus we chose to normalise test results against repeated runs with IAA, and negative controls lacking IAA, in all screens to avoid compromising the data with results arising from damaged chip surfaces.

We screened compounds in batches, with two control samples, i) Protein + IAA and ii) Protein + DMSO, at the start and end of each set of runs, and after every 10th compound. These controls not only served as assurance for the longevity of the protein and degron during the course of the experiment, but also as the reference points for the background subtraction as described in section 4.2.1. Given a screening could take up to 18 hours to assay 40 compounds in duplicate, with controls, some loss of quality of TIR1 was also sometimes observed during the course of the experiment. In order to stabilise the protein the reducing agent TCEP was included in
the buffer, the sample holding rack was kept at 10°C, and the tubes were capped to prevent oxidation. Under these conditions the protein always remained active throughout the assay, but its activity in terms of ΔRU would still reduce over time (Figure 5.2.). Nevertheless, the regular normalisation procedure outlined above, plus repetitions with separate batches of protein, randomised ordering etc. allowed us to extract reliable data.

In each assay any given compound was screened at 2 different time points, each binding normalised to the closest IAA positive control run. An example is given with (2-(2,4-dichlorophenoxy) pentanoic acid) referred to as M6 (Figure 5.3A) where the compound was run initially in the early part of the assay close to the second control, and repeated in the middle of the assay around the sixth control. When we look at the relative % binding to IAA (Figure 5.3B) we observe similar values. Compound M6 was screened in 2 further assays and the average activity (Figure 5.3C) shows M6 to have ~ 32% the activity of IAA (Figure 5.3D). Therefore, we were confident our methodology derives accurate and meaningful relative activities for compounds regardless of their run position.
Figure 5.2. Chronology of positive control RU levels in auxin screening assays.
Showing 9 sequential positive control SPR binding curves for TIR1 + IAA during the course of an activity assay. Inset graph is a magnification of the maximum binding response, showing the receding TIR1 activity as the assay progressed. The legend numbering indicates the order in which the controls were run.
Figure 5.3. Inter- an Intra-assay relative auxin activity data.
Panel A. Sensograms showing M6 to have auxin-like activity (blue line) and IAA (black lines) and the variation found within long screens illustrated with intra-assay replicates (dashed lines). Panel B shows a summary table of numerical values of from Panel A with the Δ RU values and the relative % binding compared to IAA from a single assay. Panel C shows the summary of the average activity for compound M6 from 3 distinct assays (with 2 replicates in each) giving n=6, shown graphically in Panel D.
5.2.2 SAR results – SPR screening

5.2.3 Indole ring scaffold activities

Figure 5.4 shows a selection of results from compounds based on the indole ring. The corresponding graphical results are shown in Figure 5.5. For example, the SPR results for modifications at position-1 with 1-Me-IAA give a relative activity of 8%. Methylation at position-2 (2-Me-IAA), decreases the activity to 57%, whilst a substitution with a nitrogen (indazole), further reduces activity to 23%.

The length of the aliphatic spacer at position-3 between the indole ring and COOH group is optimal with just a single CH$_2$ group (Group C compounds Figure 5.6). Removing the CH$_2$ (ICA) or elongation with 2 x CH$_2$ units (IBA), renders the compounds inactive, both at 1%. Only IPA showed some activity at 10% with an extra CH$_2$ unit compared to IAA. The absence of the COOH group abolished activity (3-Me-IAA) at 0.77%. If we explore the α–position, we see a reduction in activity (60%) with the addition of a methyl group for compound C1-IAA, but observe a recovery in activity to 90% with a longer ethyl group in C2-IAA (Group C* Figure 5.6).

Substitutions at the β-position with a variety of other functional groups (Group D Figure 5.6) with and without carboxylic acids did not support activity (<10%), except for indole-3-pyruvate (35%) and interestingly, a tetrazole group (18%). The tetrazole is a novel, non-canonical active auxin discussed further in Chapter 6. Indoxylacetate, does not poses a COOH group and was not an active compound.

Substituting a chlorine yielded a similarly if not more active auxin-like compound at positions-4 (120%), -5 (100%) & -6 (116%), where as a Fluorine and Bromine at group 5 also yielded very active compounds. Nitro groups yielded a less active compound at positions-4 (58%) and -6 (39%). Addition of an OH at position-5 reduces activity severely to 14%, but is recovered with when the OH is methylated (5-OMe-IAA) to 57%.

IAA anti-auxins with substituent changes at α (Figure 5.6A), designed to bind TIR1, but prevent the degron binding (Hayashi et al., 2012) showed no auxin activity (Figure 5.6B). Similarly, amino acids conjugated through the COOH group a the β-position
(Figure 5.6A), representing IAA metabolites, were not active auxins, all demonstrating <10% activity (Figure 5.6C).
Figure 5.5 Relative SPR activities of the Indole SAR

The panels show the results from the SPR screening of compounds, relative to IAA as part of the SAR investigation on the indole ring. The Panel labels refer to grouping of compounds in Figure 5.4. The dashed line at 10% relative activity represents the activity threshold with results relative to IAA.
Figure 5.6. SAR SPR screening of Indole α-position antiauxins and the β-position amino acid conjugates.

Panel A shows the structures of IAA anti-auxins at the α-position and IAA amino acid conjugates at the β-position, with their respective SPR relative activity results in panels B and C. The dashed line at 10% relative activity represents the activity threshold with results relative to IAA.
5.2.4 Phenoxyacetic acid as a scaffold for SAR

In this section we focused on the phenoxyacetic acid scaffold, although activities are still measured and presented as values relative to IAA. However we draw comparisons mainly to 2,4-D. Starting with a stepwise evolution of benzoic acid into 2,4-D (Figure 5.7A) we observe no auxin activity for any of the structures except 2,4-D. It would appear that 2,4-D is the ideal structure for this series of compounds and chlorines at position-2 are essential for activity.

In Figure 5.7B the addition and elongation of an alkyl group at the α-position shows an increase in activity with a methyl (RS Dichloroprop) and similarly RS-C2 2,4-D. Further elongation leads to a reduction in activity with only RS-C3 2,4-D demonstrating activity. The compounds in this series are racemic, but we also discuss a selection of stereospecific enantiomers further on.

![Figure 5.8. Evolution of Benzoic acid to 2,4-D and alkyl chain length at α-position.](image)

Panel A shows the structure-based evolution of Benzoic acid towards 2,4-D as part of the SAR study investigating 2,4-D activity. Panel B shows the addition and subsequent elongation of an alkyl group at the α-position, the compounds shown are racemic.
Figure 5.9. SPR results of the evolution of Benzoic acid to 2,4-D and alkyl chain length at α-position.

Panel A shows the SPR results for the evolution of Benzoic acid into 2,4-D with only 2,4-D showing activity. Panel B shows the increase in activity with the addition of an alkyl chain up to a maximum at RS C2-2,4-D, with a decrease in activity as the alkyl group elongates. The dashed line at 10% relative activity represents the activity threshold with results relative to IAA.
Focussing on the 2,4-D scaffold, we categorised a series of related compounds into 3 groups as shown in Figure 5.10 with their corresponding activities shown in Figure 5.11.

Our investigation into the substitution patterns of the chlorines in group (i) Figure 5.10A showed that compared to 2,4-D (27 % relative to IAA), the 2,5-D substitution isomer shows similar SPR activity at 26% whilst the 2,6-D isomer was not active (6%). Addition of another chlorine at position-6 with 2,4,6-T (1.90%) reduces the activity, and a tetra-substituted 2,3,4,5-Q (37%) compound shows higher activity compared to 2,4-D. The trisubstituted 2,4,5-T compound increases the activity considerably to 66% and forms the basis of group (ii) compounds.

Looking at group (ii) compounds, in replacing the aromatic carbon with nitrogen at position-6 of 2,4,5-T, retains aromaticity and yields a similarly active compound in Triclopyr (68%). Substituting with fluorine at the position-5 of Triclopyr and adding an amine at position-3 gave a more active compound in Fluroxypyr (110%). Focusing back on 2,4,5-T, compound T4, with an elongated CH$_2$ spacer group at the α-position increased the activity to 87%. The addition of a methyl group at the α-position adjacent to the carboxylic acid of 2,4,5-T gives Fenoprop, tested as a racemic mixture, with a higher activity of 80%.

Group (iii) compounds represent investigations into addition of alkyl groups to the 2,4-D scaffold. MCPA has a methyl at position-2 relative to 2,4-D showing an increase in activity to (45%). Increasing the number of methyl substituents on MCPA with a methyl at position-5 increases activity to 72%. The addition of a methyl at the α-position of MCPA, gives a highly active compound for the R-isomer (92%) but not the S-isomer (19%). A similar methyl addition to 2,4-D at the α-position, shows an increase in activity for R-Dichloroprop (73%), and is further increased with a longer alkyl group in R-Dichlorobut (91%). The corresponding S-isomers show low activity in S-Dichloroprop (14%) and S-Dichlorobut (20%).
Substitutions involving: hydroxyl, fluoro, bromo, alkene, nitro, methoxy and carboxylic acids at a variation of positions around the aromatic ring yielded inactive 2,4-D analogues, (<10% activity) (Figure 5.10B). It would appear that Chlorine atoms are best suited for auxin activity and in particular at positions -2 and -4 on the aromatic ring. We further observe that compounds with chlorines simultaneously occupying positions -2 and -6 show no auxin activity but variations of occupied -2, -3, -4 and -5 positions are active.
Figure 5.10 SAR single step substitutions and analogues of 2,4-D
Panel A shows 3 groups of compounds in a series of single step substitutions. Group i are halogen substitutions, Group ii are 2,4,5-T based analogues and Group iii are alkyl substitutions on the aromatic ring and the α-position.
Figure 5.11 SAR activities for 2,4-D single step substitutions.
The dashed line at 10% relative activity represents the minimum activity threshold with results relative to IAA.
5.2.5 Benzoic acid and Napthalic SPR activities

We showed benzoic acid itself was not an active auxin in section 5.4.2. The commercial auxinic herbicides derivatives of benzoic acid, TBA (11%), chloramben (26.15%) and Dicamba (14.83%) showed weak auxin activity for TIR1, whilst Clopyralid (1.38%), Aminopyralid (2.01%) and Picloram (4.4%) were not active (Group (i) Figure 5.12). This is consistent with previous findings (Walsh et al., 2006; Lee et al., 2014) where the picolinates bind AFB5 preferentially. Our screening of related compounds with various substitution patterns and substitutions in Figure 5.12 Group (ii) did not show any active compounds.

The transition from IAA to 1-NAA involves a ring expansion at position-1 and substitution of the indole nitrogen (Figure 5.4). 1-NAA demonstrates 65% activity relative to IAA. We studied a collection of 1-NAA analogues and started by looking at positional change of the CH$_2$-COOH group onto position-2 in 2-NAA, giving diminished activity (5.42%), but addition of an ether group between the napthyl ring and acetate group recovers activity in 2-NOA (18.11%). Elongating alkyl chains at the $\alpha$-position initially showed a slight increase in activity for C1-NAA compared to 1-NAA, but then progressively decreases from C2-NAA to C4-NAA. This positional change also gave active compounds for IAA and 2,4-D. When we study the $\beta$-position for the napthalic scaffold we observe no activity for amide conjugates such as NAA-Asp (0.87%) and NPA (1.42%). We also saw no activity for functional group replacements with either pair of methyl-hydrantoin (IMH) isomers. (Figure 5.15A). A bioisosteric replacement tetrazole group did not show any activity (2.62%), which was interesting given that iMQ (IAA tetrazole bioisostere) is active (Section 5.2.3).
Figure 5.12 Benzoic acid SAR compounds
Group i compounds are commercially available herbicides and group ii compounds were part of the SAR library based on benzoic acid structure, none of which showed auxin activity.

Figure 5.13 Benzoic acid analogue auxin activities. 
The graph shows the relative activity of commercial benzoic acid auxin analogues. The picolinate sub-class of compounds, Clopyralid, Aminopyralid and Picloram show no auxin activity for TIR1 but are established auxins. They are known ligands for the AFB5 receptor.
Figure 5.14. Napthalic acid SAR compounds. The panel shows the series of compounds based on the napthalic acid scaffold.

Figure 5.15 Auxin activities of napthalic acid scaffold derivatives. Panel A shows the auxin activity for the series of compounds focusing on the elongating alkyl chain at the α-position. Panel B shows the auxin activity for the positional and functional group substitutions.
5.3 *In vivo root growth assays*

The basis of the root growth assay was to ascertain the physiological impact at a whole plant level for a selection of compounds from the SPR activity screens. We measured primary root growth inhibition and increasing lateral root density traits.

### 5.3.1 Method verification

For IAA, we observed a sigmoidal dose response curve shown in Figure 5.16, with the four lowest concentrations plateauing at the average primary root lengths of the average untreated primary root length. To compare data across all compounds, we normalised against the average length of untreated control roots in each run, Figure 5.17. Control investigations on the effect of the highest top 5 DMSO concentrations found no obvious impact on primary root length (Figure 5.18A) or the lateral root density (Figure 5.18B).

Figure 5.19 shows an example with a non-active compound, indoxylacetate. We start to observe root growth inhibition at 1.23 µM, which compares to 5nM for IAA. Indoxylacetate does not however induce an increase in the lateral root density when compared to IAA. This suggests it is cytotoxic and the reduction in root growth is not auxin-controlled. This additional phenotype is key to identifying auxin-like compounds, especially less efficacious ones.
Figure 5.16. Dose response profile for primary root growth measurements with controls.
Primary root growth after treatment with IAA (blue solid line), with compound concentrations plotted in Log-scale. Dashed Red lines represent the maximum untreated length whilst green dashes represent the minimum untreated length. The Dotted Magenta line represents the average length of untreated roots.

Figure 5.17. Normalised primary root growth inhibition profile.
Primary root growth after treatment with IAA (blue solid line) normalised against the average length of untreated roots from Figure 5.7.
Figure 5.18. Impact of top 5 DMSO concentrations root growth and lateral root density compared to untreated roots.

DMSO treated and untreated seedlings physiological traits. A = Primary root lengths. B = Lateral root density. The data show that DMSO has no activity on the root growth assays.
Figure 5.19 Root growth phenotype observations

White line on each image indicates position of primary roots prior to compound treatment. The figure shows selected images of roots after 6 days of treatment at each concentration of the dose response. The comparison between an active auxin (IAA) and an inactive compound (Indoxyl acetate) shows a difference in the lateral root growth density between the 2 treatments. Although indoxyl acetate inhibits root growth at the top 4 concentrations, it does not induce lateral root growth, which is an additional induced phenotype trait of auxin activity.
5.3.2 SAR results – Primary root growth inhibition assay

IC$_{50}$ inhibition values were calculated from the in-vivo primary root growth assays for the compounds listed in Table 5.1. Dose response curves are shown in Figure 5.20 for the indole SAR series and Figure 5.21 for phenoxy and benzoic acid SAR series.

Compared to IAA (IC$_{50} = 41$ nM), we observe a drastic reduction in IC$_{50}$ with addition of a methyl group at indole SAR position-1, 1-Me-IAA (23 µM) and position-2, 2-Me-IAA (51 µM). We were unable to obtain a reliable model fit for the Indazole (estimated IC$_{50}$ between 7 to 37 µM), but its dose response shows a similar profile to 1-Me-IAA and 2-Me-IAA. All three modifications at positions-1 or -2 show no increase in lateral root induction and are merely cytotoxic at higher concentrations. These compounds demonstrated some weak auxin activity at the receptor level but not in vivo (Figure 5.20A). 1-NAA and Cl-IAA show a 5-fold reduction in the activity compared to IAA, around 200 nM each, and both induce lateral root formation. Root growth assays agree well with the SPR results.

At position-3 of the indole SAR (Figure 5.20B), ICA shows a very weak IC$_{50}$ inhibition at 138 µM, whilst IPA shows reasonable inhibition at 500 nM and IBA at 10 µM. IPA and IBA both show lateral root induction whilst ICA does not. IBA did not demonstrate activity in the SPR assays but shows some auxinic activity in vivo. This is likely due to metabolic processing which converts IBA to IAA (Korasick et al., 2013), and therefore we observed a partial IAA response.

Tryptophan with an inhibitory IC$_{50}$ of 102 µM and indole L-Lactic acid at 74 µM showed no auxin activity and no lateral root induction (Figure 5.20C). Despite indoxylacetate demonstrating an IC$_{50}$ of 15 µM, which is comparable to some weaker auxin compounds such as aminopyralid and picloram (both 5 µM), it did not induce laterals, suggesting it is cytotoxic to the seedlings. These results agree with the observed SPR activities against TIR1.

4-Cl-IAA has an IC$_{50}$ similar to IAA at 38nM, whilst 6-Cl-IAA has a 2-fold stronger inhibition at 19nM. 5- Chloro (188 nM), Bromo (189 nM) and Fluro (139 nM) analogues show a 4-fold reduction in primary root growth inhibition (Figure 5.20D & E). We observed lateral root induction for all halogenated IAA analogues and these
results correlate well with the SPR results. 5-OH-IAA demonstrated very weak inhibition with an IC$_{50}$ of 60 µM, whilst 5-OMe-IAA had an IC$_{50}$ of 7 µM (Figure 5.20F), although both induced lateral root formation. Halogen substitutions at positions -4, -5 and -6 are tolerable, even with bulky chlorine groups, however OH and an O-Me group did not show the same level of activity.

Focusing on the phenoxyacetic acid scaffold. 2,4-D demonstrated an IC$_{50}$ inhibition of 12 nM, almost 3 fold less than that of IAA. Precursor SAR compounds to 2,4-D showed weak primary root growth inhibition: BA (22 µM) and POAA (172 µM) with neither inducing lateral root growth. PAA, a reported endogenous auxin (Koepfli et al., 1938) had an IC$_{50}$ of 8 µM and showed an increase in lateral root growth formation (Figure 5.21A).

Relative to 2,4-D, the 2,5-D isomer was far less inhibitory with an IC$_{50}$ of 5 µM, but did show lateral root induction, whilst the 2,6-D isomer was even less active at 10 µM with no lateral root induction (Figure 5.22B). Despite 2,4,5-T showing better relative activity to 2,4-D in the SPR screening, we observe an IC$_{50}$ of 314 nM, Triclopyr shows an improvement by almost 2 fold at 187 nM, but the substitution of the chlorine at the 5 position to a fluorine and addition of an amino group at position-3 in Fluroxypyr leads to a 10 fold reduction in activity (Figure 5.21C). All three 2,4,5-T analogues induced lateral root growth.

The chiral isopropyl 2,4-D analogues at the $\alpha$-position, showed that the R isomers are more potent inhibitors than the S isomers with R-Dichloroprop (3 nM), R-Dichlorobut and R-Mecoprop both at 10 nM (Figure 5.21E). The corresponding S isomers show auxin activity with an increase in lateral root growth formation, but with weaker IC$_{50}$ values S-Mecoprop (1 µM), S-Dichlorobut (647 nM) and S-Dichloroprop (231 nM) (Figure 5.21F). The elongation of the alkyl chain on Dichloroprop yields a comparably potent auxin comparable to 2,4-D in R-Dichlorobut. The S-Dichlorobut is less active than its corresponding S-Dichloroprop isomer, but more active than S-Mecoprop. The in vivo results for these chiral compounds follow a similar trend to the SPR results in terms of R versus S isomers, but R-Mecoprop is less active compared to R-Dichloroprop.
Table 5.1. Primary root growth IC₅₀ inhibition and phenotyping for: indole, phenoxyacetic acid, benzoic acid and napthalic scaffolds

<table>
<thead>
<tr>
<th>Compound</th>
<th>IC₅₀</th>
<th>Error ±</th>
<th>Lateral root induction</th>
<th>SPR activity</th>
<th>Error ±</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAA</td>
<td>41 nM</td>
<td>7 nM</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>1-Me-IAA</td>
<td>23 µM</td>
<td>2 µM</td>
<td>No</td>
<td>8.37%</td>
<td>3.74%</td>
</tr>
<tr>
<td>2-Me-IAA</td>
<td>52 µM</td>
<td>7 µM</td>
<td>No</td>
<td>38.67%</td>
<td>9.26%</td>
</tr>
<tr>
<td><em>Indazole</em></td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>1-NAA</td>
<td>206 nM</td>
<td>15 nM</td>
<td>Yes</td>
<td>61.05%</td>
<td>5.59%</td>
</tr>
<tr>
<td>Cl-IAA</td>
<td>213 nM</td>
<td>54 nM</td>
<td>Yes</td>
<td>59.94%</td>
<td>1.52%</td>
</tr>
<tr>
<td>ICA</td>
<td>138 µM</td>
<td>12 µM</td>
<td>No</td>
<td>1.08%</td>
<td>0.37%</td>
</tr>
<tr>
<td>IPA</td>
<td>509 nm</td>
<td>59 nM</td>
<td>Yes</td>
<td>10.13%</td>
<td>5.03%</td>
</tr>
<tr>
<td>IBA</td>
<td>10 µM</td>
<td>1 µM</td>
<td>Yes</td>
<td>1.05%</td>
<td>0.78%</td>
</tr>
<tr>
<td>4-Cl-IAA</td>
<td>38 nM</td>
<td>4 nM</td>
<td>Yes</td>
<td>127.21%</td>
<td>7.56%</td>
</tr>
<tr>
<td>5-Cl-IAA</td>
<td>188 nM</td>
<td>14 nM</td>
<td>Yes</td>
<td>92.83%</td>
<td>5.83%</td>
</tr>
<tr>
<td>5-Br-IAA</td>
<td>189 nM</td>
<td>22 nM</td>
<td>Yes</td>
<td>91.76</td>
<td>7.36%</td>
</tr>
<tr>
<td>5-F-IAA</td>
<td>139 nM</td>
<td>45 nM</td>
<td>Yes</td>
<td>107.94%</td>
<td>7.51%</td>
</tr>
<tr>
<td>6-Cl-IAA</td>
<td>19 nM</td>
<td>2 nM</td>
<td>Yes</td>
<td>111.05%</td>
<td>5.94%</td>
</tr>
<tr>
<td>5-OH IAA</td>
<td>60 µM</td>
<td>7 µM</td>
<td>Yes</td>
<td>13.74%</td>
<td>1.48%</td>
</tr>
<tr>
<td>5-OME-IAA</td>
<td>7 µM</td>
<td>2 µM</td>
<td>Yes</td>
<td>57.18%</td>
<td>9.97%</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>102 µM</td>
<td>13 µM</td>
<td>No</td>
<td>0.47%</td>
<td>0.08%</td>
</tr>
<tr>
<td>Indole-Lactic acid</td>
<td>74 µM</td>
<td>8 µM</td>
<td>No</td>
<td>4.00%</td>
<td>2.78%</td>
</tr>
<tr>
<td>Indoxyl acetate</td>
<td>15 µM</td>
<td>2 µM</td>
<td>No</td>
<td>1.21%</td>
<td>1.75%</td>
</tr>
<tr>
<td>iMQ</td>
<td>46 µM</td>
<td>3 µM</td>
<td>Yes</td>
<td>17.60%</td>
<td>7.39%</td>
</tr>
<tr>
<td>2,4-D</td>
<td>12 nM</td>
<td>1 nM</td>
<td>No</td>
<td>27.13%</td>
<td>4.88%</td>
</tr>
<tr>
<td>BA</td>
<td>22 µM</td>
<td>3 µM</td>
<td>No</td>
<td>3.41%</td>
<td>2.45%</td>
</tr>
<tr>
<td>PAA</td>
<td>8 µM</td>
<td>1 µM</td>
<td>Yes</td>
<td>6.45%</td>
<td>3.01%</td>
</tr>
<tr>
<td>POAA</td>
<td>172 µM</td>
<td>9 µM</td>
<td>No</td>
<td>3.15%</td>
<td>2.68%</td>
</tr>
<tr>
<td>2,5-D</td>
<td>5 µM</td>
<td>1 µM</td>
<td>Yes</td>
<td>25.79%</td>
<td>9.15%</td>
</tr>
<tr>
<td>2,6-D</td>
<td>10 µM</td>
<td>1 µM</td>
<td>No</td>
<td>3.93%</td>
<td>2.62%</td>
</tr>
<tr>
<td>2,4,5-T</td>
<td>314 nM</td>
<td>19 nM</td>
<td>Yes</td>
<td>66.16%</td>
<td>9.93%</td>
</tr>
<tr>
<td>Triclopyr</td>
<td>187 nM</td>
<td>51 nM</td>
<td>Yes</td>
<td>68.10%</td>
<td>9.52%</td>
</tr>
<tr>
<td>Fluoroxypyr</td>
<td>2 µM</td>
<td>0.2 µM</td>
<td>Yes</td>
<td>110.33%</td>
<td>4.38%</td>
</tr>
<tr>
<td>R-Dichloroprop</td>
<td>3 nM</td>
<td>2 nM</td>
<td>Yes</td>
<td>73.25%</td>
<td>7.04%</td>
</tr>
<tr>
<td>S-Dichloroprop</td>
<td>231 nM</td>
<td>39 nM</td>
<td>Yes</td>
<td>12.11%</td>
<td>4.16%</td>
</tr>
<tr>
<td>R-Mecoprop</td>
<td>10 nM</td>
<td>2 nM</td>
<td>Yes</td>
<td>90.64%</td>
<td>6.63%</td>
</tr>
<tr>
<td>S-Mecoprop</td>
<td>1 µM</td>
<td>0.1 µM</td>
<td>Yes</td>
<td>17.67%</td>
<td>4.52%</td>
</tr>
<tr>
<td>R-Dichlorobut</td>
<td>10 nM</td>
<td>2 nM</td>
<td>Yes</td>
<td>91.05%</td>
<td>20.89%</td>
</tr>
<tr>
<td>S-Dichlorobut</td>
<td>647 nM</td>
<td>85 nM</td>
<td>Yes</td>
<td>20.57%</td>
<td>11.74%</td>
</tr>
<tr>
<td>Aminopyralid</td>
<td>5 µM</td>
<td>0.4 µM</td>
<td>Yes</td>
<td>2.01%</td>
<td>0.78%</td>
</tr>
<tr>
<td>Picloram</td>
<td>5 µM</td>
<td>0.9 µM</td>
<td>Yes</td>
<td>4.40%</td>
<td>0.82%</td>
</tr>
</tbody>
</table>

* Data could not be fitted confidently. The estimated IC₅₀ range is between 7 µM and 37 µM.
Primary root growth inhibition dose response: Indole SAR series

Figure 5.20 Primary root growth dose response: Indole SAR series.
The panels show collections of dose responses for primary root growth inhibition
Primary root growth inhibition dose response: Phenoxyacetic and Benzoic acid SAR series

Figure 5.21 Primary root growth dose response: Phenoxy and benzoic SAR series. The panels show collections of dose responses for primary root growth inhibition.
5.4 SPR screening SAR summary and in vitro pharmacophore mapping

We processed the results of the SAR study using Forge’s “Activity Atlas” feature to create pharmacophore maps. These maps help to summarise and draw conclusions from the SAR study, which is otherwise difficult to accomplish for such a large number of compounds. We describe the process of aligning the SAR compounds onto known reference structures, a prerequisite to processing the structures and their associated data in the Forge software. A series of pharmacophore maps were generated and we describe them in relation to IAA and the TIR1 receptor pocket.

5.4.1 Chemical structure import and alignment into Forge software

We categorised the activities from the SAR SPR screening into groups (Table 5.2), and linked these to the respective chemical structures in an “.sdf” database file format, which was imported into the Forge software as a training set. An activity variable was assigned based on the categories in the “Group” column with the “Weighting” values used as a ranking measure by the Forge software.

Table 5.2. SPR activity categorisation.

<table>
<thead>
<tr>
<th>Group</th>
<th>Range</th>
<th>Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$x \geq 100%$</td>
<td>7</td>
</tr>
<tr>
<td>B</td>
<td>$80% \geq x &lt;100%$</td>
<td>6</td>
</tr>
<tr>
<td>C</td>
<td>$60% \geq x &lt;80%$</td>
<td>5</td>
</tr>
<tr>
<td>D</td>
<td>$40% \geq x &lt;60%$</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>$20% \geq x &lt;80%$</td>
<td>3</td>
</tr>
<tr>
<td>F</td>
<td>$10% \geq x &lt;20%$</td>
<td>2</td>
</tr>
<tr>
<td>U</td>
<td>$10% &lt; x$</td>
<td>1</td>
</tr>
</tbody>
</table>

Poses of IAA, 1-NAA and 2,4-D were imported as reference structures from the coordinates of their crystallographic structures. Figure 5.22A shows the three-reference compounds individually and overlaid from 2 points of view. The aromatic rings superimpose in the same plane, and the carboxylic acids hinge perpendicular to the rings. Despite only having a single aromatic ring, the two chlorine atoms on 2,4-D provide sufficient volume (referred to as shape in the Forge software) to match with IAA and 1-NAA. The TIR1 receptor crystal structure was also imported and the aligned rings are shown sitting in the binding site (Figure 5.22 B).
Figure 5.22 Crystallographic compound reference structures of IAA, 2,4-D and 1-NAA in the TIR1 receptor.
Panel A: IAA, 2,4-D and 1-NAA co-crystallised individual and overlaid poses. Labels indicate IUPAC numbers with related side group positions labelled as α and β for simplicity. B. The reference structures are depicted in the TIR1 pocket, represented as a surface excluded volume. (Grey = Carbon, Blue = Nitrogen, Red = Oxygen and Yellow = Sulphur)
For each training set and reference compound, the software automatically calculated “field-points” (Cheeseright et al., 2006), which are maximal pharmacophore values of: charge (electropositive and electronegative), hydrophobicity and shape (van-der-Waals radii). Global surface representations of each pharmacophore feature are extrapolations of the field points. An example with IAA in Figure 5.23 shows the field points for each pharmacophore category from two different points of view. The electronegative field (Blue) points sit above and below the aromatic ring, with some localisation adjacent to the carboxylic acid group. Around the indole nitrogen we observe a small electropositive field (Red) in plane with the aromatic rings. The aromatic ring system and the CH$_2$ group are hydrophobic (Gold) with field points maximal at the centre of the heterocycle. Van-der-Waals points are observed around the molecule defining its shape and size (Yellow). For the remainder of this discussion, IAA will serve as a reference molecule to aid our description of the pharmacophore maps.
Figure 5.23. Pharmacophore field points generated from Forge software.
The panels show 4 pharmacophore features generated by the software for IAA. The field points are maximal values of each feature shown as spheres. The field points are extrapolated to create surface representations shown in the lower row of panels.
Each compound was aligned to one or more of the reference structures based on their common shape and field point similarity. The key requisites for alignments were: I) orientation of the carboxylic acid groups and II) alignment of the aromatic rings. The TIR1 binding site pocket was used as an excluded spatial region, for a more probable alignment and placement of the compound in the receptor site (Figure 5.24). The alignments are not docking experiments, but rather an overlay of chemical structures onto known reference poses. The alignments are energy minimized to ensure plausible geometrically and steric conformers of each molecule and that the compounds were aligned so that common features, like the core scaffolds and functional groups were superimposable or within close proximity.

Indole based compounds were aligned exclusively to the IAA structure, as was the case for napthalic [to 1-NAA] and phenoxyacetic [to 2,4-D] reference compounds respectively to ensure the best alignments. There is no crystallographic reference structure for the benzoic acid scaffold, leading us to align it to the trio of known references. The result of the alignment is shown in Figure 5.24. We observe the benzoic acid aromatic ring aligns well with the reference aromatic rings, but the carboxylic group is positioned slightly further away from the references due to lack of a linker CH$_2$ group. The aligned benzoic acid structure was used as an additional reference for other benzoic and picolinate analogue compounds with picloram as an example (Figure 5.24). We were able to align 277 out of the 317 compounds, with the remainder of the compounds, [all inactive] not conforming to the alignment profile of reference structures.
Benzoic acid  
Picloram  

Top-view  

Side-on view  

Figure 5.24. Alignments of Benzoic acid and Picloram.
The left panels show the alignment of Benzoic acid (grey) to the three reference structures (thin sticks, IAA = Yellow, 1-NAA = slate and 2,4-D = Red). This was used as an additional reference for other benzoic acid and picolinate compounds, of which Picloram is shown on the right.
5.4.2 Pharmacophore mapping results

5.4.2.1 Shape of active compounds

We observe the average shape of compounds in Figure 5.25 A & B from 2 points of view. The pharmacophore shape compliments the active site well as the surface shape fills and fits into the cavity of the binding pocket (Figure 5.25 C & D). The average shape is very similar to IAA from Figure 5.22 and that of 1-NAA and 2,4-D. In conjunction with the average active shape, the software generated regions defined as “activity cliffs” of favourable (Green) and unfavourable (Purple) shape (Figure 5.25 E & F). An overlay of the cliffs with the average shape pharmacophore shows where substitutions exploring the shape could lead to an increase in activity (Figure 5.25 G & H).

The pharmacophore cliffs for shape are isolated and shown from various points of view in Figure 5.26, where we observe the presence of shape in region I is unfavourable. This overlaps with the fusion points of the aromatic rings of IAA (position 7a) and 1-NAA (position 8a), where we observe no substituents. Conversely, this region of un-favourability coincides with position-3 of the 2,4-D scaffold where we observe activity for compounds 2,3,4,5-Q and Fluroxypyr, which both have substituents at that position. Given that we only have only 2 active compounds versus many other inactive compounds with substituents or shape in that region, the model thus suggests that compounds are favoured when they have no additional volume in that region.

Adjacent to the α-position of IAA in region II, we observe favourable shape tolerance corresponding to where with the substituents from R-chiral isomers sit. Interestingly, the region occupied by S-isomers is not favoured. The S-isomers are active, therefore the model does not label this position as unfavoured, but given the significantly higher activity of the R isomers, the pharmacophore mapping indicates the R-space is a region of greater significance.

There is a region of favourable shape around the carboxylic acid group in region IV, surrounded with unfavourable shape in region III. Both these regions define a very particular shape identical to that of the carboxylic acid group. This is most likely
because the majority of active training compounds are carboxylic acids, with the only exception being iMQ. With respect to this particular region and our understanding of active auxins, the COOH functional group is still paramount.

The favourable shape in Region V surrounding the benzene ring of IAA is consistent with observed activity for substituents at positions -4, -5, & -6 on the indole scaffold, as well as positions -4 & -5 on the phenoxyacetic scaffold relative to 2,4-D. This region also corresponds to a pharmacophore shape activity cliff, meaning that this region can accommodate additional shape but it should preferably be hydrophobic in nature (see below).

We have observed an apparent conflict regarding substitution position 4 of the indole ring and position 6 of the phenoxyacetic scaffold. 4-Cl-IAA and 4-NO₂-IAA are active compounds, where as 2,6-D and 2,4,6-T are not (Figure 5.27A). Yet when aligned, they have a substituent in the same spatial position as shown with the alignment of 4-Cl-IAA and 2,6-D with a chlorine atom in the region of contention (Figure 5.27B). We surmise the inactivity is not due to the presence of a substituent in that region, but rather, it is due to the reduced freedom of rotation afforded to the COOH group for 2,6-D and 2,4,6-T compounds on the phenoxyacetic acid scaffold in this SAR study. Either side of the oxyacetic group, the large chlorines likely impose a steric presence, restricting overall molecule flexibility. Therefore, 2,6 di-substituted phenoxyacetic acid compounds are inactive auxin molecules, whereas compounds with a 2,4,5 substituent patterning are active auxins.
The Figure shows a summary of the average shape of active compounds (white surface) from 2 points of view in Panel A and B. IAA is also shown as a visual reference with positions numbered according to Figure 5.1. The same average shape with IAA is shown in relation to the TIR1 receptor site (Ribbon backbone, with thin stick residues in purple) in Panels C and D. Panels E and F show the activity cliffs, with regions II, IV & V in green leading to an increase in activity, whilst the presence of shape in the purple regions I & III leads to decreased activity. The average shape and shape-activity cliffs are overlaid in panels G and H to show which areas of the shape are pivotal for the activity.
Figure 5.26. The activity cliff of shape pharmacophore for TIR1.
Green = Regions of favoured shape, Purple= regions of unfavoured shape. The panel shows the activity cliffs from the SAR from various points of view. The labelled regions 1 to V are key areas of shape dictating auxin activity.
Figure 5.27 The inactivity of the 2- and 6- co-occupied positions of phenoxyacetic acid scaffolds versus the occupied active indolic position 4

Subscript letters under compound name indicates activity in blue, with decreasing activity going from A to F. The letter U in red represents inactive compounds. In yellow are IUPAC positions for the acetic and indolic scaffolds shown for reference. Panel A shows a selection of compounds with substituents at 2,4,5 and/or 6 position on the phenoxyacetic acid scaffold and position 4 on the indolic scaffold. 2,4,6-T and 2,6-D are inactive compounds and it would appear to be a result of having shape at the 6-position. Panel B shows the alignment of 2,6-D (dark grey) with 4-Cl-IAA from 2 points of view, where we observe they both have a chlorine substituent in the same region, yet the former is inactive. It is postulated, the 2,6-D and similarly 2,4,6-T have reduced conformational freedom about the CH₂-COOH group due to the large chlorines on either side.
5.4.2.2 Hydrophobicity in the Forge atlas

The average hydrophobicity surface of the active compounds is shown in Figure 5.28A & B with the encompassing space and spread over the binding site. The pharmacophore can accommodate two aromatic ring systems, with IAA shown as a reference in Figure 5.28C & D. When viewed from the side in the receptor site, the hydrophobicity is planar (Figure 5.28B) and neatly encompasses the IAA molecule in Figure 5.28D. We observe no hydrophobicity around the carboxylic acid β-position.

The hydrophobicity-activity cliffs show an unfavourable (purple) hydrophobic area around the indole nitrogen (I), with a small region of favourability (green) adjacent to it (II) (Figure 5.28 E & F). It would appear hydrophobicity is acceptable in Region I when we view the superimposition of the activity cliff with the average hydrophobicity. Region II is encompassed in the average hydrophobicity as well (Figure 5.28 G & H). Interestingly, region I of the shape pharmacophore overlaps here too. This is perhaps due to the nature of compounds studied by the model. Given that we have more inactive compounds than active compounds with hydrophobicity and shape in this region, the model is biased for this region. Furthermore, active compounds such as 1-NAA are also without polarity in this region. The data suggest that other heterocyclic compounds with polarity here would be interesting to study.

The more pronounced Region III favours hydrophobicity and extends off the α-position (Figure 5.28 G & H). Likewise, region IV also favours hydrophobic substituents around the -4, -5, -6 and -7 positions of IAA. Regions III and IV are exploitable with hydrophobic substituents as they extend beyond the region of the average shape, corresponding with the shape pharmacophore (5.4.2.1. above). The hydrophobicity pharmacophore correlates with the character of binding site residues in the TIR1 pocket in Figure 5.29 A & B. Leu_{378}, Phe_{380} and Cys_{405} are situated next to the α-position with Phe_{79}, Phe_{82} and Ala_{464} surrounding the aromatic rings. They are shown with their respective hydrophobic field surfaces flanking the hydrophobic pharmacophore. These residues and their positioning in the receptor site must therefore contribute favourably to the thermodynamics of binding.
We also show the polar residues in the binding pocket (Figure 5.29 C & D). Arg$_{489}$ is sandwiched between and slightly raised relative to the Phe$_{79}$ and Ala$_{464}$ hydrophobic residues. Forming part of the pocket bottom, Arg$_{403}$ & Arg$_{436}$ sit below the ring of hydrophobic residues and interact with the carboxylic acid group. In the region facing the indole nitrogen there is a Serine residue above (Ser$_{440}$) and below (Ser$_{438}$) the pharmacophore, with the amide backbone of Leu$_{439}$ in the same plane as indole nitrogen. There are no large polar or hydrophobic residues in the location corresponding with regions I and II (Figure 5.27 E & F), partially explaining why the indole, naphthalic and phenoxyacetic scaffolds are able to bind the receptor. The distribution of these residues, above and below the hydrophobic pharmacophore further determines the planar nature of the hydrophobic pharmacophore, restricting it to that region.

Hydrophobicity activity cliffs suggest that region III contributes towards higher activity. We tested compounds with alkyl groups in this region for the indole, naphthalic and phenoxyacetic scaffolds. In particular, the R-chiral isomers are highly active and occupy this space. In Figure 5.29 E & F we observe region III extends towards Leu$_{378}$ and Phe$_{380}$, and likely explains the higher activity of the R-isomers versus the S-isomers. There is an optimal tolerance for a 2-carbon branch at the $\alpha$-position and C2-2,4-D is previously unreported as an auxin. We have shown it binds the TIR1 receptor, is active in our root growth assays with the mechanism of action revealed by our pharmacophore model.

The cliff also suggests that the addition of hydrophobicity is favoured in region IV. (Figure 5.29 E & F). The aromatic side group of Phe$_{79}$ could establish a displaced $\pi$-$\pi$ interaction with an aromatic ring from an auxin satisfying the hydrophobic pharmacophore. Concurrently, Phe$_{82}$ could also form an edge-to-face $\pi$-$\pi$ interaction with the aromatic ring of an auxin. There is auxin activity with indolic 4-,5- & 6-substituted halogen compounds, which occupy this region of the model likely forming halogen-$\pi$ interactions (Bissantz et al., 2010). 5-OMe-IAA with the large methoxy substituent group was also active, consistent with the shape pharmacophore activity cliff in region V. 4-NO$_2$-IAA and 6-NO$_2$-IAA are less active compounds; fitting into the space available but their polar nature may explain their reduced activity.
Figure 5.28. The hydrophobic pharmacophore for TIR1.
Panels A and B show the average hydrophobic map around the aromatic indole ring of IAA, shown as a thin sticks representation. In panels C and D we observe the average hydrophobic pharmacophore in relation to the TIR1 binding site (Ribbon backbone, with thin stick residues in mauve, Blue = Nitrogen, Red = Carbon). Panel E and F show hydrophobic activity cliffs, with green representing favourable hydrophobicity and purple representing unfavourable hydrophobicity. Panel G and H are the superimpositions of panels A & E and B & F respectively showing the relationship between the cliffs and average hydrophobicity. Regions III and IV extend off the average hydrophobicity shape, indicating possible space for adding hydrophobic substituents in this region.
Figure 5.29 Hydrophobicity pharmacophore in relation to TIR1 binding site residues.
Panel A and B show the average hydrophobic pharmacophore (Gold surface representation) in relation to the surrounding [white labelled] hydrophobic residues, also with gold surface representation, flanking the pharmacophore. Panel C and D are similar to A & B but with polar arginine residues [as thick stick representation] labelled in blue. The arginines sit above, below and in front of the pharmacophore. Panels E and F show the activity cliffs, with favourable regions III and IV (Green surfaces) in close proximity to hydrophobic residues, explaining the trends observed for active compounds with hydrophobic residues in those regions. In purple is the region unfavoured for hydrophobicity, which faces polar residues.
5.4.2.3 Electrostatic pharmacophore

The atlas generated an average electrostatic pharmacophore of active compounds with only electropositivity shown (Figure 5.30 A & B). It proposes favourable electropositivity in the region where the carboxylic acid moiety lies, which is counterintuitive given the negative electrostatics of the carboxylic acid group. If we refer back to the example with IAA in Figure 5.23, there are field points of electronegativity around the carboxylic acid moiety, this is also true for the majority of compounds given auxins are exclusively carboxylic acids. Therefore, we would expect the pharmacophore mapping to at least show an average electronegative region around the carboxylic acid moiety.

As for the activity cliffs (Figure 5.30 C & D), the model suggests an increase in activity for substituents of negative electrostatic character around the carboxylic acid group region, as well as the -4 and -5 indole positions. The carboxylic acid group is a hydrogen bond acceptor and electronegative, agreeing with the activity cliff but conflicts with the average electrostatic pharmacophore. The activity cliff duly suggests positive electrostatics around the indole ring lead to an increase in activity. The nitrogen on the indole ring can act a hydrogen bond donor and is thus suitably placed in this region. However, the chlorine at the 2-position of 2,4-D also lies in this region, but it is electronegative.

A summary of the regions explored in positive electrostatics (Figure 5.30 E & F) and negative electrostatics (Figure 5.30 G & H) demonstrates the diverse nature of compounds analysed. On studying the different compound scaffolds we postulate the likely reason for the conflicting model results. Figure 5.31 shows regions of electropositivity and electronegativity for active indole and phenoxyacetic acid scaffolds. The electronegativity below the aromatic ring and around the carboxylic acid is common between both scaffolds, but above the aromatic ring we see some minor differences. The most prevalent difference between the two scaffolds is the distribution of the electropositivity. There is a region of electropositivity around the nitrogen of the indole ring whilst, the phenoxyacetic acid scaffolds have electropositivity around the α- and 6- positions. With such variation between these sets of active compounds in the current training set, there is no obvious trend with regards to the electrostatic pharmacophore output.
Figure 5.30 The Electrostatic pharmacophore atlas of auxin
Blue = electronegative regions. Red = electropositive regions. IAA (dark grey) is used as a reference in each panel. Panels A and B show the average electrostatic pharmacophore comprising entirely of electropositive values with no electronegative fields. Panel C and D show the electrostatic activity cliffs, which suggest more electronegativity/less electropositivity in the blue regions lead to an increase in activity, whilst in the red region, more electropositivity/less electronegativity leads to an increase. Panels E & F show the total spatial region explored from the training set in electropositivity, with panels G & H similarly showing the explored electronegative region.
Figure 5.31 A comparison of the electrostatics of indole and phenoxyacetic acid scaffolds.
The panel compares the electropositive and electronegative regions for two major auxinic scaffolds. They have similar electronegative profiles around the aromatic rings but differ in their electropositivity distributions. The electropositivity is located on opposite sides (except for Triclopyr) and is possibly the reason why the average electrostatic pharmacophore model presents a conflicted result.
5.4.2.4 Shape and hydrophobicity trends

The average hydrophobicity pharmacophore is planar and fits neatly into the average shape pharmacophore. Thus hydrophobicity in this plane leads to auxin activity. In the region indicated by a blue arrow in Figures 5.31D, E and F, we observe no hydrophobic pharmacophore. This is consistent with our discussion in 5.4.2.1 as this is where the carboxylic acid groups would be present in an active compound, corresponding to region IV from Figure 5.26. We would expect this to be electronegative, given that auxins are exclusively carboxylic acids or tetrazoles.

![Top-down view](image1)

**Figure 5.32 A comparison of the average shape and hydrophobicity pharmacophore.**
Panels A shows the overlay of the average hydrophobicity pharmacophore and average shape, indicating the majority of the shape for active auxins in this region is comprised of hydrophobic character. Panels B and C show these pharmacophores separately with both possessing identical shape profiles when viewed from above. Panel D shows the corresponding side view. There is a region of shape not filled by hydrophobicity, indicated by the blue arrow, where we would expect to see the polar electronegative groups.

We showcase a selection of compounds to discuss trends relating to the shape and hydrophobic pharmacophores in Figure 5.33. 2,4-D comparatively matches the shape of the larger aromatic rings of IAA and 1-NAA, sufficiently fulfilling the receptor shape pharmacophore. The large chlorine atoms compensate for the lack of a second aromatic ring by contributing favourable hydrophobicity. This is evident in the alignment of the reference compounds in Figure 5.22. With no halide substituents, phenoxyacetic acid (POAA) does not have auxin activity, nor does the
monochlorinated 4-chloro-phenoxyacetic acid (4-Cl-POAA). We were not able to source a monosubstitued 2-chloro equivalent, but predict it too would be inactive.

Where one of the chlorines is substituted for fluorine in 2-fluoro-4-chlorophenoxyacetic acid (2-F,4-Cl-D) there is a loss of activity, due to the smaller fluorine atom being unable to fulfil the shape requirement. Where we increase the size of the substituent at the 2-position with a vinyl group, 2-(4-chloro-2-ethenylphenoxy)acetic acid (2V,4-Cl-D), an analogue of MCPA, there is no activity, probably because the vinyl group is too large to fit into the receptor, failing the shape test (Figure 5.33).

2,4,5-trichlorophenoxyacetic acid (2,4,5-T) is more active than 2,4-D, which can be attributed to the additional chlorine atom increasing the overall size and hydrophobicity of the compound. 2,5-dichlorophenoxyacetic (2,5-D) acid is also similarly active against TIR1 relative to 2,4-D, satisfying the shape pharmacophore. 2,3,4,5-tetrachloro-phenoxyacetic acid (2,3,4,5-Q) is much more active than 2,4-D and indicates that 4 large chlorine groups can be accommodated around the phenoxyacetic acid ring. This series of compounds hints at the permissible space for substituted mono-aromatic ring systems and how their positioning can fulfil the shape pharmacophore in order to bind this region of TIR1 successfully.

Combinations of polar and hydrophobic substituents on the aromatic ring do not yield active auxins, illustrated by the activity cliff and average pharmacophore (Figure 5.32 A & D). Two selected examples of 2-methyl, 4-amino-phenoxyacetic acid (2-Me,4-NH₂-D) and 2,4-dichloro,5-hydroxy-phenoxyacetic acid (2,4-Cl,5-OH-T) posses polarity at the 4- and 5- positions respectively. This agrees with the activity cliff in that unfavoured polarity in close proximity to the hydrophobic results in decreased activity. Other functionalities at these positions such as nitro and carboxylic acids do have any auxin activity. The amino substituent of Fluroxypyr at position 3 does not conflict with the hydrophobicity activity cliff and is allowed, furthermore the smaller fluorine atom at position 5 does not impede activity as the compound already posses chlorines at 2- and 4-. positions.
Figure 5.33 Shape and hydrophobicity trends for auxin activity.

Subscript letters under compound names indicate activity, with decreasing activity going from A to F. The letter U in red represents inactive compounds. The white cloud overlaid onto each compound represents the average shape pharmacophore. For compounds based on the phenoxyacetic acid scaffold to be active, they should have chlorine at the 2- and 4-, or 2- and 5- positions. These 2 substitution patterns match the bicyclic shape of IAA and 1-NAA. Substituents are tolerable simultaneously or in any combination at the 2,3,4,5 and positions, but not 2- and 6- positions simultaneously. Polar groups at the 4 and 5 position lead to inactivity, but are allowed at the 3 positions.
5.5 Conclusions

Our SAR study has led to development of a novel 3D pharmacophore model for auxin activity in the context of the TIR1 receptor interacting with the Aux/IAA binding degron. Compared to the 2D binding site model based on IAA in Figure 1.7 (Kaethner, 1977), there is indeed a polar region, electropositive region facing the carboxylic acid group and a similarly polar region of the receptor around the indole nitrogen. As for the hydrophobic coverage, our pharmacophore model shows a wider coverage of the binding site relative to Kaethner’s model.

The model by Katekar (1979; Figure 1.12) based on whole plant SARs extended Kaethner’s predictions. The model proposed a generalised mapping of the auxin pharmacophore. Some of our findings are consistent with this model, especially the need to placate a specific shape within the receptor site with aromatic rings possessing the ideal shape to bind the receptor. Substituted mono-aromatic phenoxy acetic acids, with chlorine atoms at positions 2,3,4 and 5 were equally active, but chlorines at 2- and 6-positions was identified as unfavourable due to the steric hindrance imposed on the rotation of the carboxylic acid group. Katekar’s model also described the possibility of accommodating substituents around the aromatic rings, with chlorine atoms exhibiting similar or increased activity.

The process of generating the pharmacophore was reliant on the commercial software tool FORGE, made available in a timely manner to help us realise this objective. The process of aligning compounds and calculating pharmacophore descriptors has been previously reported (Tomić et al., 1998). However, the authors did not present a pharmacophore map, as they were unable to find any correlations with activity. They were perhaps limited due by the use of whole plant activity data. They instead generated series of similarity indices for each pharmacophore descriptor and classified compounds based on these.

Katekar’s model inaccurately described the pose of the carboxylic acid group, suggesting it would lie in the same plane as the aromatic ring in the TIR1 binding site. The pose suggested by Katekar was similar to the recognition conformation proposed by (Kaethner, 1977) except that the COOH group is rotated by 180°. In this regard
Kaethner’s modulation conformation, akin to the pose Tilted (T) pose by of Tomic et al., 1998 and also that of the bound IAA in the TIR1 crystallographic structure, predicted the correct conformer. Katekar’s model factored in space for alkyl groups to branch off the α-position, and this corresponds with the tolerances of our model, and rationalised these based on the prospect of favourable interactions with hydrophobic binding site residues in the area. We have shown that there is an optimal length tolerance for the alkyl group.

Despite being conceived prior to the identification of an auxin receptor and its co-receptor complex with Aux/IAA, both Kaethner’s and Katekar’s models were insightful. It is a testament to their foresight and deep understanding of auxin biology that their models have lasted so long. However, both Kaethner’s and Katekar’s auxin receptor models postulated a putative classical single receptor-ligand model for IAA binding, with no co-receptor.

Whilst their preceding models agree with our findings, we have further enriched the auxin activity model descriptor, providing a 3D context whilst also factoring in the receptor site features. The model revises and further aids our understanding of in vitro auxin binding. Our proposed model describes the shape, hydrophobic and electrostatic features contributing to auxin activity. The analysis has shown auxin activity is determined by combination of these features, most notably hydrophobicity and shape, which correlate to the TIR1 binding site.

As for the electrostatic pharmacophore, we are as yet unable to discern any obvious trends with the current training set. We can summarise from our atlas features of active auxins: they will posses a carboxylic acid group and aromatic ring system. Both are electronegative with the former functioning as a hydrogen bond acceptor, whilst the aromatic ring is crucial to auxin activity, probably helping determine binding pose. Further work will be required to understand the contributions and implications of both electropositive and electronegative pharmacophores.

With regards to auxin perception, the hydrophobic auxin pharmacophore within the TIR1 receptor is key to the Aux/IAA-receptor binding. Crystallography revealed that
tryptophan and proline from the degron of Aux/IAAs interfaced with bound auxin through hydrophobic interactions (Tan et al., 2007). IAA and auxin-like molecules are considered the “glue” between the 2 co-receptors, therefore it is befitting that the proposed widespread and planar hydrophobic pharmacophore seemingly acts as a platform for the Aux/IAA protein to bind favourably, influencing the interaction.

The next step is for these pharmacophore maps to be used to search chemical space for other novel auxin-like compounds, and not necessarily restricted to the 4 scaffolds in Figure 5.1. The discovery of novel scaffolds would help diversify our auxin compound portfolio, particularly applicable for the development of novel herbicidal agents.

With compound classes like the picolinates showing preferential binding to AFB5 over TIR1 it would be interesting to compare and contrast their pharmacophore models. As such, we are currently working on developing an equivalent pharmacophore for the AFB5 receptor and other remaining members of the auxin F-box receptor family.

We are also developing a global *in vivo* pharmacophore model from our whole plant root growth data. It will be interesting to compare and contrast this to the *receptor* model. We acknowledge the model will account for a multitude of processes including uptake, transport and metabolism, but it will enable us to understand what other chemical features compounds must possess in order to reach the TIR1 receptor. So far, between the *in vitro* and *in vivo* assays we have shown parallel trends for the hydrophobicity at positions -4, -5 and -6 around the aromatic indole as well as the potent activity of R-alkyl analogue of 2,4-D.
Chapter 6: Simply selective: A tetrazole analogue of Indole-3-acetic acid is an auxin and shows selectivity between $SCF^{TIR1/AFB5}$-Aux/IAA Co-receptors: TIR1 and AFB5.
6.1 *In silico design and validation of IAA tetrazole analogue*

In Chapter 5, our SAR investigations around the carboxylic acid functional group yielded a novel auxin active against the TIR1 receptor. We employed a classic medicinal chemistry approach by substituting the carboxylic acid for a tetrazole moiety, giving the compound; \(3-[(1H-1,2,3,4-tetrazol-5-yl)methyl]-1H-indole\), hereon referred to as iMQ. This functional group substitution was designed *in silico* to exploit the space available in the polar region of the TIR1 pocket. We measured distances to possible intra-atomic contacts from the carboxylic acid group of IAA bound in the TIR1 crystal structure to nearby residues. Distances of 3.47Å to Arg\(_{403}\) and 4.52Å to Arg\(_{436}\) signified some space to exploit in this region of the binding site (Figure 6.1).

The tetrazole functional group is a bioisostere of a carboxylic acid. Within the definition of bioisosteres, a carboxylic to tetrazole alteration is considered as non-classical due to the change in functional group rather than a single atom change. (Lima & Barreiro, 2005; Meanwell, 2011).

We modelled the tetrazole group in the TIR1 site using the structure builder feature in Chimera (Pettersen et al., 2004). Using IAA as a template, we excised the COOH group and replaced it with a manually constructed 5-membered tetrazole ring *in situ*. Figure 6.2 shows the side-by-side comparison of IAA with iMQ, the model supposing the indole rings to be superimposable. The tetrazole group would simply be an extension of the carboxylic acid group further down into the pocket, with shortened distances of 1.89Å to Arg\(_{403}\) and 2.34Å to Arg\(_{436}\). Furthermore, the proximity of Ser\(_{438}\), Arg\(_{403}\) and Arg\(_{436}\) residues were considered probable hydrogen bonding partners for iMQ (Bissantz et al., 2010).

Calculated physiochemical properties for IAA and iMQ using ChemAxon’s physiochemical calculators (Table 6.1) showed they have similar pKa’s, solubilities and number of rotatable bonds. There is also no difference in the major anionic species at pH 7.4, demonstrating a fundamental feature of bioisosteres in mimicking the original compound’s characteristics (Allen et al., 2012). The tetrazole group increases the total surface area of iMQ relative to IAA, and the polar surface area because of two additional hydrogen bond accepting groups. These differences are predicted to have an impact on biological activity, mobility and molecular interactions.
Figure 6.1 Available space around the IAA carboxylic acid group in the TIR1 pocket. Polar contacts (Green dashed lines) from the IAA carboxylic acid group (purple) to Arg\textsubscript{403} and Arg\textsubscript{436} indicate the exploitable space in this region of the binding pocket. Figure generated in Chimera.

Figure 6.2 Side-by-side comparison of IAA to iMQ in the TIR1 pocket
Panel A shows IAA in the binding site with Ser\textsubscript{438}, Arg\textsubscript{403} and Arg\textsubscript{436} shown as sticks (from docking or crystal structure?). Panel B shows iMQ model in the same binding site with the tetrazole group and is able to fit in, as it is a simple extension of the carboxylic acid group. Figure generated in Chimera.
Table 6.1. Physiochemical property comparison between IAA and iMQ

<table>
<thead>
<tr>
<th></th>
<th>IAA</th>
<th>iMQ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure</strong></td>
<td><img src="image" alt="IAA Structure" /></td>
<td><img src="image" alt="iMQ Structure" /></td>
</tr>
<tr>
<td><strong>Mass</strong></td>
<td>175.186</td>
<td>199.216</td>
</tr>
<tr>
<td><strong>Formula</strong></td>
<td>C_{10}H_{9}NO_{2}</td>
<td>C_{10}H_{9}N_{5}</td>
</tr>
<tr>
<td><strong>Atom Count</strong></td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td><strong>Acidic pKa</strong></td>
<td>4.66</td>
<td>4.84</td>
</tr>
<tr>
<td><strong>Major Species at pH 7.4</strong></td>
<td><img src="image" alt="IAA Major Species" /></td>
<td><img src="image" alt="iMQ Major Species" /></td>
</tr>
<tr>
<td><strong>Donor Count (pH 7.4)</strong></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Acceptor count (pH 7.4)</strong></td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><strong>cLogP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>([Octanol]/[Water])</td>
<td>1.1822</td>
<td>1.0278</td>
</tr>
<tr>
<td><strong>cLogS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Water solubility in mol/L pH=7.5 AT 25°C)</td>
<td>-2.119</td>
<td>-2.273</td>
</tr>
<tr>
<td><strong>Total Surface Area</strong></td>
<td>162</td>
<td>181</td>
</tr>
<tr>
<td><strong>Polar Surface Area</strong></td>
<td>53.09</td>
<td>70.25</td>
</tr>
<tr>
<td><strong>Rotatable Bonds</strong></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Aromatic Rings</strong></td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Our bioisostere predictions were backed up with *in silico* docking. Docking was performed using an automated docking script (Price et al., 2014) based on the Vina algorithm (Trott & Olson, 2010). iMQ and IAA were docked into an empty 2P1Q crystal structure. The indole ring superimposed as predicted above, (Figure 6.3) with only a slight difference in the alignments of the tetrazole and carboxylic acid. The calculated scoring function for IAA was -8.1 kcal/mol, -8.2 kcal/mol for iMQ.

![In silico docking for IAA and iMQ](image)

*Figure 6.3 In silico docking for IAA and iMQ*

Panels A, B and C; show docked IAA and iMQ structures align onto one another as viewed from different perspectives. Figure generated in Chimera.
6.2 iMQ synthesis

iMQ was prepared in via 1,3-dipolar cycloaddition (Wittenberger, 1994) of sodium azide (NaN$_3$) and 2-(1H-indol-3-yl)acetonitrile (Compound (1), Scheme 6.1) catalysed by AlCl$_3$ (Dolusić et al., 2011) as described in methodology 2.9. We obtained an off white powder, corresponding to a 25% yield by weight. (Figure 6.4A). A 200 Da species was observed by mass spectrometry in positive ion mode, confirming the presence of iMQ (uncharged mass 199 Da). Proton H$^1$ NMR spectroscopy peak assignments matched that of expected IAA tetrazole, similar to reported values by (Dolusić et al., 2011). The presence of the tetrazole moiety was further confirmed by a broad downfield peak around 16ppm definitive of the lone exchangeable proton on a tetrazole ring Figure 6.4B. Complete NMR and MS spectra are presented in Appendix D.

Scheme 6.1 Synthesis of iMQ.
Figure 6.4. iMQ characterisation data.
Panel A selectively displayed region of the H$^1$ NMR run in DMSO-d$_6$ of iMQ purified product showing a singlet peak for the indole proton at 11ppm and a broad peak at 16ppm, indicative of an exchangeable proton on a tetrazole. Panel C shows a high-resolution mass spectrum of iMQ in positive ion mode [MH]$^+$. The m/z of 200.9031 corresponds to a positively charged iMQ species with an additional proton.
6.3 iMQ activity assays

iMQ was active in our SPR assays, with a relative activity of 18% vs. IAA at 10 µM (Figure 6.5A), the iMQ + TIR1/ASK1 the SPR binding signal to the degron reached a plateau unlike IAA. Single cycle kinetics of iMQ showed a much weaker affinity of 300µM relative to IAA at 9µM, and iMQ demonstrates a quicker off-rate constant of 0.07 ± 0.00007 (1/s) relative to at 0.003 ± 0.000009 (1/s), 23-fold faster.

In order to check co-receptor complex formation with plant-expressed TIR1, a co-IP of FLAG-tagged TIR1 expressed in Nicotiana benthamiana with Aux/IAA peptide showed weak auxin-like characteristics for iMQ, compared to IAA, even at 100µM iMQ (Figure 6.5B). The weak pull down assay corroborates with the rapid off rate from the SPR experiments.

In order to evaluate iMQ activity as an auxin in living plants, 10-day-old A. thaliana DR5::GFP reporter seedlings were treated with 50µM iMQ or IAA by addition of the auxin solution to seedlings on agarose media for 2 hours, before transferring onto slides for observation under a confocal microscope. Untreated roots showed minimal background GFP signal, not affected by a control treatment with DMSO (0.05%). For both control treatments the background GFP signal was very minimal but observed on the range indicator (Figure 6.6). IAA treatment shows a characteristic increase in signal in the epidermis, steele and the lateral root cap (Overvoorde et al., 2010). iMQ treatment produced a GFP signal, localised to the nucleus of the cells. The IAA signal is of greater intensity and more generally spread entirely across cells. The reporter assay provided confidence that iMQ is behaving as an active auxin in vivo and proved uptake of the compound into the roots, as we observed the GFP signal in the appropriate root tissues.
Figure 6.5 iMQ activity assays
Panel A, SPR assay showing (50 µM) iMQ auxin-like activity (red trace), and (50 µM) IAA as a reference (blue trace), with TIR1. iMQ shows a saturating binding response with a rapid off rate. Panel B shows a western blot developed with anti-FLAG antibody of TIR1-FLAG protein and Aux/IAA7 degron co-immunoprecipitated, with increasing IAA and iMQ concentrations. iMQ shows a very weak pull-down efficiency compared to IAA but demonstrates auxin-like activity.
**Figure 6.6 Whole root *in-vivo DR5::GFP* reporter assays**

Channel overlay: GFP signal in Green with Propidium Iodide (PI) staining in Blue. Range indicator: Blue = background baseline, yellow arrows = meristem, white lines = PI staining signal and black to red signal in the root cells indicate GFP signal of increasing intensity. The series of panels show the GFP signal after 2 hours with very little GFP fluorescent signal observed for untreated and DMSO (0.05%) treated control, the background *DR5::GFP* signal is observed in range indicator mode for untreated and DMSO treated plants. There is a an obvious signal observed for IAA and iMQ with an auxin distribution pattern in the epidermis, steele and the lateral root cap cells.
In addition to activity at the cellular level, we conducted whole seedling root growth assays to evaluate the compound. iMQ displayed weaker auxin activity than IAA in primary root growth inhibition assays. We obtained an IC$_{50}$ value of 46,000nM (±3000), for iMQ compared to 41nM (±7) for IAA, (Table 6.2). We observe complete root growth inhibition at 300µM for iMQ compared with 11µM for IAA (Figure 6.7). We can attribute the weaker iMQ auxin activity in part to the lower affinity recorded for TIR1 binding, but also perhaps to root-uptake limitations. iMQ increased lateral root density (Figure 6.8), demonstrating additional phenotypic traits typical of auxins, but only at concentrations greater than 100µM, where as IAA induces this effect at 100nM (Figure 6.9).

In order to try and test and confirm that iMQ is selective for TIR1, we adopted the root growth assay with knockout lines for representative members of the TIR1 receptor family. For example, if iMQ worked only on TIR1 (and not AFBs), we would expect the tir1 mutant to be insensitive to iMQ. However, iMQ is sensitive to TIR1 knockout in Col-0 tir1-1 seedlings (Table 6.2), although sensitivity is reduced. We observe a ~2-fold reduction in efficacy giving an IC$_{50}$ of 90,000nM, with no total root growth inhibition, even at the highest concentration of 300µM (Figure 6.7). We also observe a significant reduction in the lateral root growth density in tir1-1 (Figure 6.8), whilst changes in IAA phenotypes are minimal for both traits in the TIR1 knockout line.

<table>
<thead>
<tr>
<th>Table 6.2 Primary root IC$_{50}$ values for TIR1 knockout line.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Line</strong></td>
</tr>
<tr>
<td>Col-0 WT</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td><strong>Compound</strong></td>
</tr>
<tr>
<td>IAA</td>
</tr>
<tr>
<td>iMQ</td>
</tr>
</tbody>
</table>
Figure 6.7 Primary root growth inhibition assays
Dose response curves for primary root growth inhibition assays for Col-0 WT (solid lines) and Col-0 tir1-1 mutants (dashed lines), IAA (Blue) and iMQ (Red).

Figure 6.8 Lateral root densities for IAA and iMQ treated WT and tir1-1 Col-0 lines
Top row of panels show the lateral root density dose response for wild type (WT) treatments of IAA (Panel A) and iMQ (Panel B), showing an increase in the lateral root density compared to untreated controls (dashed blue line). The bottom row of panels show lateral root densities for the TIR1 knockout tir1-1 line, treated with IAA (Panel C) and iMQ (Panel D), we observe a reduction in the lateral root density with iMQ whereas IAA shows no significant reduction.
Figure 6.9. Dose response phenotypes for root growth assays with Col-0 WT.
White horizontal dash represents the position of primary roots before treatment. The series of images compares the phenotypes observed for IAA and iMQ in the dose response series. IAA is a stronger auxin than iMQ.
6.4 iMQ does not bind to AFB5

It was noted that iMQ at 50 µM did not bind to the AFB5 receptor (Figure 6.10). This revealed iMQ was selective for TIR1 over AFB5. In order to investigate the basis of this selectivity, a binding site sequence alignment (Figure 6.11) was done between TIR1 and AFB5. It revealed two key changes. TIR1_{His78} is an arginine in AFB5 and TIR1_{Ser438} is an alanine in AFB5. Histidine to arginine represents an increase in residue size and polarity, whilst a serine to an alanine is a decrease in polarity and increase in hydrophobicity and removal of a key hydrogen bond acceptor.

Using a homology model built for AFB5 from (Irina et al., 2012) and in silico docking (Figure 6.12) it was revealed the “histidine to arginine” change as the most likely major contributing factor to the receptor selectivity. IAA docked into the AFB5 pocket in a favourable pose, such that the aromatic ring was flush with the hydrophobic residues and the carboxylic acid orientated towards the arginine residues at the bottom of the binding region, analogous to IAA binding in the crystal structure of TIR1. Docking of iMQ into the AFB5 pocket demonstrated the consequences increased steric bulk imposed by the arginine, preventing the tetrazole group from orienting towards the bottom of the pocket. As a result the tetrazole group faces upwards, and the indole ring is tilted with respect the pose for IAA, placing iMQ’s indole ring in an unfavourable pose.

Figure 6.10. iMQ displays selectivity for TIR1, not AFB5 in SPR assays. iMQ binds to TIR1 (Panel A, red trace), not with AFB5 (Panel B). Both panels show binding by IAA as a reference (blue trace).
Figure 6.11 Binding pocket amino acid sequence alignments, TIR1 vs. AFB5.
Protein sequence alignment (using Clustal 2.1) of TIR1 binding pocket residues, and AFB5. Key differences are highlighted: histidine (His78) in TIR1 is an arginine (R) in AFB5, and serine (S438) in TIR1 is an alanine (A) in AFB5.

Figure 6.12. *In silico* docking shows iMQ cannot bind to AFB5.
Panel A: IAA docked into AFB5 with the COOH group facing the two arginines on the pocket floor, and the aromatic ring aligned parallel to the pocket floor. Panel B: iMQ docked into AFB5 with the tetrazole moiety facing away from the arginines. Panels C and D show rotated view points to show iMQ cannot adopt the same pose as IAA due to the Arginine (Red) which would otherwise clash with the tetrazole moiety. Residues are represented as coloured sticks: Red: arginine^78, purple: arginine^340, and green: alanine^438. Figure generated in PyMol.
6.5 iMQ causes weak agravitropism in primary roots and laterals

Wild type seedlings treated with iMQ displayed unusual root bending between concentrations of 100µM and 33.33µM (Figure 6.13, right hand column). The roots appear to bend towards one direction and continue on that trajectory, there is no obvious preferred direction of bending and the amount of bending is variable across the roots. The effect is also seen in the tir1-1 line (Figure 6.13, left hand column).

Figure 6.14A is a photograph of a typical experimental set up, with 3 stacks of plates representing 3 compounds assayed for primary root growth elongation. Figure 6.14B, a schematic top down view of Figure 6.14A, showing the growth incubator lighting set up. Fluorescent light blubs face the plates and also from either side of the plates. In order to ensure that the agravitropic phenotype was not due to local growing conditions, we placed 9 identical plates at 33 µM iMQ, each with 5 Col-0 WT and 5 Col-0 tir1-1 mutant seedlings in positions shown (Figure 6.14A). We observed no obvious biased bending towards any particular light source, thus ruling out cabinet effects as a driver of the roots bending (Figure 6.15). In order to characterise the bending further, we measured the bending in terms of the offset angle (Figure 6.16).

Figure 6.16 depicts this angle for untreated roots and iMQ treated roots. There is no directionality for θ. A summary of the angles for Col-0 WT and Col-0 tir1-1 lines (Figure 6.17) show a statistically significant difference between treated (100µM) and untreated root angles in Col-0 WT. We also observe a difference at 33µM and untreated roots in the tir1-1 line. In a test of rate of response, iMQ treated Col-0 WT roots showed a dose dependent retardation in gravitropism kinetics (Figure 6.18). Post germination, seedlings were rotated 90° and their recovery measured over 6 hours. Untreated seedlings demonstrate a uniform return to vertical, whilst iMQ-treated roots do not display normal gravity sensing. At 50 µM iMQ the roots continue to grow agravitropically.
Figure 6.13. Dose dependent primary root agravitropism with iMQ.
Arrows indicate which roots are bending. The Figure shows unusual bending in the primary roots in Col-0 WT seedlings at 100µM and 33µM iMQ. The bending is also observed in tir1-1, indicating the effect is TIR1-independent.
Figure 6.14 Root growth assay lighting and plate stacking positions in incubator
LHS = Left-hand side, RHS = Right-hand side & MID = Middle row
Panel A is a photograph showing 3 stacks of plates in the growth incubator labelled according to their relative positions in the growth chamber. Each row contains a random stacking of 15 treated plates and 1 untreated control. Panel B is a top-down view schematic of the growth incubator, with the position of the lighting indicated. 9 positions in the growth chamber contained plates treated with iMQ and 5x Col-0 WT and 5x Col-0 tir1-1 seedlings.
Figure 6.15. Position in the incubator did not affect bending in roots treated with iMQ. The figure shows images of Col-0 WT and Col-0 tir1-1 seedlings showing no obvious differences in relation to their position in the growth chamber.
Figure 6.16. Gravitropic offset angle measurements.
Figure 6A (Left), shows the offset angle ($\theta$) measurement on un-treated Col-0 WT root, with a zoom (inset). Figure 6B (right), shows ($\theta$) for Col-0 WT roots treated with iMQ at 100$\mu$M.

Figure 6.17. Gravitropic offset angle measurements in Col-0 WT and tir1-1 mutants treated with iMQ
Offset angles are displayed in box plots for treated and untreated roots of Col-0 WT (Panel A) and Col-0 tir1-1 lines (Panel B). In the Col-0 WT line at 100$\mu$M iMQ treatment, a statistically significant difference was seen relative to the untreated control, based on one-way ANOVA compared to no treatment, $p$ value of <0.0001 at 95% confidence (range = -48.89 to -20.48).
Similarly for the Col-0 \textit{tir1-1} line we observe a statistically significant result at 33\(\mu\)M, \textit{p} value of \(<0.0004\) at 95% confidence (range = -33.23 to -10.20).

**Figure 6.18. iMQ treated root gravitropism kinetics.**

The figure shows gravitropism in seedling roots after being rotated 90\(^o\) and monitored for 6 hours. iMQ treated seedlings display a dose dependent retardation of gravitropism. At the highest concentration (50 \(\mu\)M iMQ), the roots growing agravitropically.
6.6 iMQ possibly interferes with auxin uptake

AUX-1 deficient roots have been reported to demonstrate severe agravitropic phenotypes (Marchant et al., 1999). We investigated the possibility of iMQ antagonising the membrane bound AUX-1 transporter, thus possibly giving rise to the root bending. *In vitro* auxin uptake assays were performed in BY2 tobacco cells similar to (Lankova et al., 2010), observing the accumulation of [\(^3\)H] labelled 2,4-D in the cells measured as radioactive count of disintegrations per minute (DPM). 2,4-D is a substrate for the AUX-1 carrier and addition of competing unlabelled substrate results in the reduced accumulation of 2,4-D within the cell. iMQ decreased 24-D accumulation in a dose dependent manner, with an estimated IC\(_{50}\) of 5 \(\mu\)M indicating it was competing with 2,4-D for the AUX-1 carrier (Figure 6.19).

Application of the synthetic auxin 1-NAA has been shown to recover gravitropism in AUX-1 deficient roots (Marchant et al., 1999). 1-NAA is thought to be membrane permeable, independent of uptake carrier proteins on the basis of its lipophilic nature (Marchant et al., 1999; Parry et al., 2001). Although this mechanism of entry has been challenged (Kell & Oliver, 2014) and suggests other types of transporters must facilitate compound movement across a membrane. Despite no clear mechanism for how 1-NAA moves into the cell independent of AUX-1, its recovery of function role is important nonetheless and is a useful utility.

In preliminary experiments, we sought to observe if 1-NAA could overcome the agravitropic response by supplementing 10nM of 1-NAA into the agarose media, in addition to 50\(\mu\)M iMQ in a rescue of function assay. There was a reduction in the measured offset bending angle (Figure 6.20) in addition to less bending when the roots are observed (Figure 6.21). 1-NAA treated roots show a recovery to gravitropism, indicating AUX-1 uptake had been affected with iMQ.
Figure 6.19. Accumulation of labelled [³H] 2,4-D in AUX-1 accumulation functional assay
The figure shows a dose dependent reduction in [³H] 2,4-D accumulation as the concentration of unlabelled iMQ increases, indicating iMQ is likely to be interfering with the AUX-1 channel activity.

Figure 6.20. 1-NAA rescue of function assay for iMQ induced 1° root bending
The boxplots show the measured root angles of various treatments. The combination of 1-NAA and iMQ shows a slight reduction in the bending compared to just iMQ. 1-NAA is able to slightly correct the bending as it can diffuse into plant cells independent of AUX-1 and thus subsequently correct against the agravitropism.
Figure 6.21 Rescue of bending phenotype with 1-NAA.
Arrows highlight root bending. The figure shows images of roots treated with iMQ and the addition of 1-NAA, which has reduced the bending phenotype on visual inspection. Corresponding controls show the bending is only observed on seedlings treated with iMQ.
6.7 Discussion

We have discovered the first receptor-selective auxin through in silico ligand design. Based on bioisostere chemistry we have described the activity of an indolic tetrazole (iMQ) as a functional auxin against the *A. thaliana* Col-0 WT model, with a variety of assays, both *in vitro* and *in vivo*. A tetrazole analogue for IAA had previously been synthesised (McManus & Herbst, 1959) and reported to have very weak auxin activity in *Avena coleoptile* shoot tip elongation assays (Hamilton et al., 1960). It is possible this work had categorised the tetrazole group as ineffective. As such, there has been no further auxin activity related work reported based on this functional group. Our work has also shown iMQ to be a very weak auxin, but we have discovered its ability to select between TIR1 and AFB5 receptors making it a useful probe to study auxins.

iMQ is a weak auxin, both *in vitro* and *in vivo*, which is in contrast to the *in silico* docking prediction that they would have similar binding affinities (with similar scoring functions) for the TIR1 receptor. The introduction of the tetrazole changes the physiochemical properties of the compound, with additional hydrogen bond acceptors. This increase in polarity compared to IAA could affect the uptake and transport of the compound *in planta*. Hence, poor transport could have contributed to the weaker activity in vivo. Furthermore, we have an indication that iMQ interferes with AUX-1 uptake carrier, thus impeding an alternative entry route into the cell. Additionally, as we have shown in chapter 4, active auxins must navigate their way into the deep binding site before docking at the base. Standard docking algorithms will not account for such details, and the scoring function assumes only the characteristics of docking at the final site. It is possible therefore that the scoring function is not a true representation of binding efficacy. Perhaps, iMQ is less efficient than IAA at traversing the TIR1 binding pocket and hence has weaker binding activity.

The selectivity of iMQ’s for TIR1 over AFB5 revealed insights into the two binding sites. Key residue changes in the binding site were shown to be key to this selectivity. The docking predicts iMQ to bind in an unfavourable pose that would likely prevent the AFB5 + iMQ complex from binding to the AUX/IAA degron. This corroborated well with the SPR experiment where we observed no binding for the putative AFB5 + iMQ complex to the degron. It is not established as to whether iMQ actually binds to
AFB5 at all. Our efforts to complement the TIR1 pull-down assays with AFB5 protein have not been successful yet, but would confirm our hypothesis on iMQ binding to AFB5. As we have not yet established a methodology to observe direct binding of compounds to auxin F-box receptors. We are continuing our efforts as described in chapter 3 to achieve this. It would be interesting to use site-directed mutagenesis to convert AFB5 into an iMQ binding protein, without changing any other residues in these similar proteins. This could reveal if it is perhaps indeed the binding residues contributing to the selectivity.

In tir1-1 mutants iMQ still displayed very weak auxin like activity. It is possible that it targets another member of the receptor family as well as TIR1. Most likely is AFB1, which phylogenetically pairs with IAA at a protein sequence level (Figure 1.6 and Table 1.3). As such, we are in the process of verifying if iMQ is a ligand for the remaining auxin receptors using similar assays. A consequence of its weaker auxin activity could also be due to the fact that it is a ligand for only 1 receptor, therefore less efficacious compared to IAA, which is the endogenous ligand for all the auxin receptors.

The unusual primary root bending indicated that iMQ might also affect other gravitropism-related processes. This was also independent of TIR1 and so was most like an off-target effect. Auxin uptake experiments show iMQ interacts with AUX-1, hindering its uptake functionality. Other known auxins are substrates for AUX-1 but do not show such interference, most obviously IAA and 2,4-D. AUX-1 activity is likely to be impeded by iMQ at high concentrations over a period of time, affecting the polar auxin distribution and therefore give rise to the observed bending. 1-naphthoxy-acetic acid (1-NOA) at around 10μM and 3-chloro-4-hydroxyphenylacetic acid (CHPAA) at 5μM originally discovered in an in vitro SAR study (Imhoff et al., 2000) have been reported as AUX-1 inhibitors, inducing similar agravitropic responses in Col-0 WT seedlings (Parry et al., 2001; Lankova et al., 2010). Gravitropism was rescued for seedlings exposed to 1-NOA and CHPAA with the addition of 1-NAA. Our initial rescue of function assay with supplemented 1-NAA also showed a slight return to gravitropism for iMQ treated seedlings. It is possible iMQ is acting in a similar manner to these antagonistic uptake compounds and would be interesting to observe the effect of these compounds in our root growth assays.
compared with iMQ. We also would like to investigate at what concentration of 1-NAA do we completely recover from the bending phenotype.

Despite iMQ’s interaction with AUX-1, the root bending is perhaps not as severe compared to AUX-1 deficient roots, which display curled-up roots and display a complete agravitropic response (Marchant et al., 1999). This is most likely because iMQ is also able to induce some auxin response thus countering some of the bending. As we have shown with DR5::GFP reporter assay and phenotypic observations that iMQ can traverse the root cells and induce auxin activity therein. iMQ’s pKa is similar to IAA, therefore it is capable of passive diffusion into the cell across the membrane. It appears that iMQ imparts a conflicted effect on the plant as it acts as an auxin yet prevents the additional mechanism that would aid its efficacy as an auxin. iMQ also presents an interesting starting point for an SAR study on the AUX-1 ligand specificity, to identify which other functional group replacements show a similar trend.

In addition to considering iMQ’s interactions with AUX-1 in the context of gravitropism, it would be worth establishing if iMQ interferes with statocytes. These are specialised cells and tissues, found in the columella cells of root cap region in a variety of plant species, which play a role in gravity sensing (Morita, 2010; Hashiguchi et al., 2013). They are also localised to the inflorescence stem in Arabidopsis. When primary roots are re-orientated, the starch-accumulating amyloplast sediments with respect to gravity within statocytes. This primary response leads to the re-localisation of PIN3 efflux carrier, which in turn leads to the re-distribution of auxin across the root cap into the elongation zone ultimately resulting in a gravitropic response (Baldwin et al., 2013). In addition to PIN3, we could also look at the impact on auxin efflux, with members of the PIN family including PIN1 and PIN2, similarly linked with gravitropism (Friml, 2003; Křeček et al., 2009; Petrásek & Friml, 2009; Gallavotti, 2013).
Chapter 7: General discussion and conclusion
A definitive chemical descriptor of auxins has been sought after for decades (Napier, 2001) with many models published (Kaethner, 1977; Katekar, 1979; Tomić et al., 1998) no absolute definition discerned as yet. This thesis looked at developing an auxin activity pharmacophore based specifically on the receptor TIR1, although the project did also study compound selectivity between the TIR1 and AFB5 receptors. In addition to the conclusions in each chapter, we present a further general discussion and propose future work.

7.1 Observations of IAA binding to TIR1 with in silico docking

A computational approach in Chapter 4 was intended to aid our efforts in probing chemical space for novel auxin compounds. IAA was docked, in a pose analogous to its bound crystallographic conformation. However, we observed that non-auxinic compounds were also docked with scoring functions similar to auxins section 4.2. In response, we developed TomoDock (Section 4.3), enabling us to consider residues along the deep binding pocket of TIR1 are likely to play a role in selecting for auxin-like compounds.

In our published work (Appendix B), we have identified two key selectivity regions in the TIR1 receptor pocket: The first, a niche towards the top of the TIR1 pocket that initially engages with the polar carboxylic acid group of compounds. The second, a filter that then determines if compounds are able to progress further towards the binding site or results in their exit from the pocket. The postulated interaction between a small molecule and these two regions sheds interesting and novel insight into the sequential multistep perception of auxin-like compounds. An analogous multistep binding mode of action has been reported previously for an antibody-antigen interaction (Lipschultz et al., 2002). The authors referred to the mechanism as a “molecular encounter followed by docking” involving a macroscopic conformation change between the binding partners.

TomoDock’s utility is not restricted to auxin F-box proteins. It could be applied to the study of e.g. the transit of small molecules down channel or transport proteins, or to any other similarly deep binding pockets such as the jasmonic acid hormone receptor,
COII of which there is a crystal structure (Sheard et al., 2010) which resembles that of IAA very closely and is also part of a ternary complex.

TomoDock could also be useful to study the recently discovered Nitrate-Regulated Auxin Transporter (NRT1.1) channel protein shown to transport (NO$_3^-$), but also involved in the uptake of auxin (Krouk et al., 2010) Its crystal structure was recently published (Parker & Newstead, 2014; Sun et al., 2014). Similarly, the established auxin transporters AUX-1 and PIN would be ideal candidates for TomoDock, but there are no reported crystal structures as yet. Despite this, homology modelling would be a useful initiation towards studying these channel proteins until reliable structural models are elucidated.

A reported structure for a GPCR (G-protein-coupled receptor) oncology target SMO (The Smoothened receptor) as part of the hedgehog (Hh) signal transduction network was shown to have deep binding pocket (Wang et al., 2014). The authors reported some small molecule antagonists bound very close to the top, whilst others were able to bind further down. The authors could not present an explanation for this but speculated at possible sub-sites along the pocket with particular compound specificities. This receptor candidate would be apt for future TomoDock analysis as well.

Crystallographic structures of AFB5 and the remaining auxin F-box receptors are yet to be produced. They will likely shed new information on binding site specifications for IAA and other auxin-like compounds. These crystallographic models would also be interesting candidate receptors for TomoDock pocket analysis and will likely explain some of the observed compound binding preferences, e.g. the pyridine based picolinate class of compounds preferentially binding AFB5 over TIR1 (Walsh et al., 2006; Prigge et al., 2016).
7.2 Pharmacophore mapping

Development of a receptor-based pharmacophore map in chapter 5 has been key towards developing an understanding of auxin perception for the TIR1 receptor. The results from our SAR investigations yielded a detailed 3D pharmacophore map, and we have shown how this correlates with the nature of the TIR1 binding site from Tan et al., 2007. Furthermore, the model offers a novel spatial perspective of what chemical features are required for auxin activity giving us a vital new tool for exploiting the binding site further. The pharmacophore mapping has now enabled us to develop further hypotheses for the binding site and apply them as mentioned below in the proposed lead optimisation of iMQ (Section 7.3.1).

It would be interesting to develop similar maps for the AFBs, hopefully giving us additional insights to their chemical specificity. Having such an insight may also help us understand the need for so many auxin receptors, suggesting if perhaps they have other as yet unknown signalling molecules as ligands and what functional role they may play. For example, the pea plant branching RMS2 (ramosus) gene has been reported to encode the AFB4/AFB5 gene in peas and is involved in the auxin-strigolactone regulation loop (Catherine Rameau et al., 2016, unpublished, presented at IPGSA 2016, Toronto and in correspondence).

Biological in vivo assays provided for many years a global output for activity of exogenously applied compounds. The development of similar, but separate spatial pharmacophore models for compound uptake, efflux and metabolism would enable a modular insight into these ancillary, but equally pertinent pharmacological processes in the general context of auxin activity. Better understanding of chemical features favoured at each step of a compound’s passage towards the receptor would greatly aid the rational design of novel auxins.
7.3 Design of non-canonical auxins and receptor specificity

Through a rational drug design approach in Chapter 6 we demonstrated that the tetrazole analogue of IAA, iMQ, is an auxin, albeit rather weak in vivo in Arabidopsis. It had been reported as a weak auxin as early as 1960 in Avena coleoptile shoot elongation assays (Hamilton et al., 1960) but has been overlooked ever since.

The tetrazole replacement is only one of many possible isosteres. Recent work by (Lassalas et al., 2016) evaluated the physiochemical properties of 34 carboxylic acid isosteric replacements including, sulphonamides, sulphonic acids and hydroaxmic acids. The authors reviewed the varying impact on solubility, lipophilicity and ionisation for each isostere, enabling one to rationally modulate the physiochemical properties of a compound for a desired increase in uptake, transport or binding site affinity. It would be interesting to investigate the auxin activity of these isosteres. This could help further diversify the auxin-chemical portfolio of non-canonical auxinic herbicides with the potential to aid in the fight to overcome herbicide resistance. In conjunction with the pharmacophore mapping, identifying isosteres of the aromatic rings also presents an additional avenue of investigation.

As such there is no clear explanation for the redundancy of auxin F-box receptors (Pierre-Jerome et al., 2013), knocking them out individually has very little impact on auxin perception (Dharmasiri et al., 2005b; Prigge et al., 2016). An application of receptor selective agonists like iMQ would be to help us separate and understand the individual activities controlled by members of the auxin F-Box receptor family without the need to mutate out the receptors.

One off-target effect of iMQ, agravitropism at high dose (Section 6.5 and 6.6) also presents an interesting prospect with regards to auxin uptake. The most recent putative model for the prediction of auxin uptake (Imhoff et al., 2000) considered only carboxylic acid compounds. iMQ presents itself as a candidate compound for a prospective new avenue of investigation into the chemical specificity of the AUX1 uptake carrier and.
7.3.1 Lead optimisation of iMQ

We propose a series of compounds for the lead optimisation of iMQ (Figure 7.1) with modifications arising from our pharmacophore model of the TIR1 binding site. Given the value we observed in adding an R-isomeric methyl or ethyl group at the α-position, and chlorines at positions: 4-, 5- and 6- our iMQ improvements would all carry these features. These modifications exploit the two favourable hydrophobic regions around the aromatic ring around positions 4 to 6, and the α-position relative to IAA. The addition of chlorines to the ring will not impact the pKa (4.84), whilst the addition of a methyl on the side chain may lower the pKa favourably to 4.75, and the pKa is further lowered to 4.73 with an ethyl group. Realising these molecules and assessing both binding by SPR and root growth inhibition in vivo will help validate the effectiveness of our pharmacophore model and demonstrate its utility.

Figure 7.1 Lead optimisation candidates for iMQ
The figure shows a series of proposed lead optimisation compounds for iMQ based on suggestions from the auxin pharmacophore model.
7.4 Conclusion

The goal to truly define what makes an auxin is not beyond reach but requires further concerted effort. The complex nature of auxin perception, with its multiple co-receptor combinations does not make this an easy undertaking. We have shown a rational approach to studying auxin at the receptor level can help lead us closer to this goal. Starting from the inside, process by process, we can work towards a far more comprehensive understanding of auxin perception.
References


Appendix A: SPR Methodology publication

Assaying auxin receptor activity using SPR assays with F-Box proteins and Aux-IAA degrons

Quareshy, M., Uzunova, V., Prusinska, J., and Napier, R.M

1. Introduction

The study of hormone-active proteins requires robust assays for binding. Most binding assays are developed for receptor candidates, but both enzymes and transport proteins also carry binding sites specific for their substrates. Historically, the basis of most analyses was radiolabelled ligand binding assays. More recently a range of biophysical techniques has offered new, label-free tools for binding analysis and whilst no single technique can address all experimental systems, surface plasmon resonance (SPR) has found widespread utility. It yields definitive kinetic data for binding interactions and may also be used for thermodynamics and high-throughput compound screening. SPR requires that one partner is immobilised on a chip surface which can be a limiting factor, but in most cases the variety of available chip surface chemistries allows binding reactions to be measured by immobilising either receptor or ligand. Ideally kinetics are recorded in both orientations.

The most revealing kinetic parameter sought from binding assays is the affinity of the site for its ligands. Let affinity be $K_D$, the dissociation constant for the interaction under the stated conditions. The units are Molar and $K_D$ is the concentration of ligand at which the receptor binding sites are at half occupancy. Radiolabel assays in various formats generally yield the affinity. For example, most binding analyses on Auxin-Binding Protein 1 (ABP1) used radiolabel displacement assays (e.g. Napier et al., 2002), competing radiolabel off the binding site with increasing concentrations of unlabelled auxin to give $K_D$ from a Scatchard analysis. However, label-free assays like SPR are able to provide additional valuable information. How much information depends somewhat on the availability of pure protein. For example, SPR typically requires micrograms of protein and can provide binding on- and off-rates as well as $K_D$, with the off-rate indicating the durability of the interaction - a useful feature in ligand screening. SPR can also be used for determining the thermodynamics of binding, measuring changes in free energy, enthalpy and entropy as well as affinity to reveal aspects of the mechanism of binding. We will explain how this experiment is run, but with more protein available (milligrams) isothermal titration calorimetry (ITC) will give more definitive thermodynamic values given that the ITC reaction is entirely in solution.

Many candidate auxin receptors have been listed (Venis and Napier, 1995; Napier et al., 2002) and two have been characterised extensively, ABP1 and Transport Inhibitor Resistant 1 (TIR1) which mediates auxin-inducible changes in transcription (Dharmasiri et al., 2005; Kepinski and Leyser, 2005). Both are soluble proteins and structures have been determined for both, empty and with auxin bound (Woo et al., 2002; Tan et al., 2007). However, crystallography offers only a snap-shot of a binding interaction taken under extreme conditions. We describe here an SPR assay developed for TIR1 (and its homologue AFB5; Calderon Villalobos et al., 2014). The assay not only gives us kinetic data on auxin binding, but it also provides a platform for novel auxin and anti-auxin discovery. On Fig. 1 we present a schematic of the complicated multi-component cascade of interactions involved in auxin perception in vivo. In the method we describe, a peptide, which includes the Aux/IAA degron motif, is immobilised on the SPR chip surface. The TIR1 will be the mobile partner flowing over the chip surface. Where
appropriate, auxin will be added to the TIR1 protein solution and allowed to bind before being passed over the chip (Fig. 2)

At low auxin concentrations, Aux/IAA transcriptional repressors dimerise with co-repressor proteins to repress transcription of genes targeted by auxin response factors (ARF), which are transcriptional activators.

The empty TIR1 receptor protein is part of a ubiquitin E3 ligase, the SCF\textsuperscript{TIR1} complex, based around the scaffold cullin protein.

At higher concentrations of auxin, auxin binds to the receptor TIR1 forming a binding site for Aux/IAA proteins. Auxin is the “molecular glue” between the two proteins in this co-receptor complex. Aux/IAA proteins bind via a degron motif in domain II (REF)

Once assembled, the E3 ligase activity rapidly ubiquitinates the Aux/IAA protein, marking it for degradation in the proteasome.

The resulting reduction in the Aux/IAA proteins leads to dissociation of the ARF-Aux/IAA dimers and de-repression of transcription (Woodward & Bartel, 2005a).

\textit{ASK1: Arabidopsis SKP1, Links F-box proteins to CULLIN. CULLIN; Scaffold protein of SCF complex. RBX1; RING-H2 finger protein, docking subunit for Ub-E2 Ligase. E2; Ubiquitin-conjugating protein}

\textit{Figure 1: Schematic representation of the cascade of events during auxin perception by TIR1.}
Figure 2: Experimental set-up for auxin binding assay based on the initial events in the TIR1 binding complex cascade.

1. Biotinylated AUX/IAA peptide sequence that binds to TIR1 in the presence of an auxin immobilized onto a Streptavidin surface chip.

2. TIR1/AFBS protein purified.

3. TIR1 or AFBS protein ± auxin flowed over peptide.

4. In the presence of an Auxin TIR1/AFBS binds to the peptide on the surface with the auxin acting as the molecular glue. The blue binding trace is observed.

5. In the absence of an Auxin and with a non-auxin compound, TIR1/AFBS does not bind to the peptide on the surface in each case and the red flat trace is observed.
2. Materials

Maintenance chip (GE Healthcare) and SA sensor chip (streptavidin-coated for biotinylated molecules) (GE Healthcare BR 1000-32)

Plastic vials 7 mm, 0.8 ml rounded polypropylene (GE Healthcare BR 1002-12) with rubber Caps, type 3 (GE Healthcare BR 1005-02)

Desorb kit consisting of BIAdesorb solutions 1 and 2 (GE Healthcare BR 1008-23)

Chip cleaning solutions: A) 1 M NaCl in 50 mM NaOH and B) 50 mM NaOH

Surfactant P20 10% v/v (GE Healthcare BR 1000-54)

SPR rinse solution, filtered (0.2 micron) and degassed: water, 0.05% v/v Surfactant P20

SPR running buffer, filtered and degassed: 10 mM HEPES pH 7.4, 150 mM NaCl, 3mM EDTA, 50 µM phytic acid, 1 mM TCEP, 0.05% v/v Surfactant P20 (See Note 1)

Biocytin, stock solution 1 mg/ml in water (Sigma Aldrich B4261)

Biotinylated peptide, stock solution 1 mg/ml in water. (Our standard is the sequence derived from the AXR3 (Aux/IAA7) degron, biot-QVVGWPPVRNYRK (ThermoFisher)

Indole-3-acetic acid (IAA). 100 mM as a stock solution in dimethylsulfoxide (DMSO) stored frozen at -20C, and as a working stock solution of 10 mM in DMSO.

TIR1/ASK1 purified protein in SPR running buffer

3. Methods

The methods discussed below relate to Biacore 2000 and 3000 instruments, but may be adapted for other SPR platforms.

The control software opens as shown below:

From the menu bar Command, Run and Tools are used most commonly. The software guides users through all the routines and gives instructions for sample placement and volume requirements etc. We recommend that new Biacore users take training on the instrument before your first experiment, but the instructions given by the software are comprehensive.

3.1 Chip preparation
Prior to coating a new sensor chip (or additional channels on a part-used chip), always run a Desorb routine to clean the fluid lines. Desorb is run using a maintenance chip.

1. Place the buffer inlet tubes into a bottle of SPR rinse solution.
2. Command -> Undock; wait until the undocking routine completes (the countdown time is shown, see below). Remove any other chip and place a maintenance chip in the holder. Command -> Dock; wait for docking to complete.

3. You will be asked if you want to Prime the system. Accept the Prime option. Read the notes in the dialogue box.

6. As in 2, undock to remove the maintenance chip and dock a new SA chip.
7. Replace SPR rinse solution with SPR running buffer. As above, select Prime. (See Note 2).
8. Perform a chip conditioning routine using Manual mode and cleaning solution A (See Note 3). Select Run -> Run sensorgam… A window entitled Detection opens and prompts a choice of flow cells. Choose the relevant cells and hit OK (See Note 4).
9. A Flow dialogue box appears. Input a flow rate of 20 ul/min. -> OK. You will be prompted to set up a .blr file where the results are saved (See Note 5). As soon as you hit Save the main screen refreshes (as shown) giving a read-out for each channel selected in 7 (See Note 6). The control buttons for manual injection are now active and a Command Queue dialogue box opens:

10. Set up three consecutive 1-minute injections of the cleaning solution (See Note 7). Select -> Inject; a dialog box opens.

Position asks you to nominate the rack position for the cleaning solution vial. Volume -> 20 ul is recommended. The volume entered determines the contact time for a given flow rate.

Consumption (the volume of conditioning solution used) is calculated. This includes system dead volumes plus the 20 ul wash volume.

Hit -> Start Injection places this operation in the Command Queue (shown). While the first command is executed (indicated by the cogwheel icon), add two identical injections to the queue.
11. After the last injection finishes, hit -> Stop Sensogram. The queue is cleaned, as a prompt informs you. Hit -> OK. An example sensorgram is shown in Fig. 3.

Figure 3: Manual sensorgram collected during chip surface cleaning. Three 1-minute injections are represented by three consecutive peaks. Data are shown for flow cell 1 only. You may also visualize sensorgrams for all channels simultaneously.
3.2 Coating a chip with degron peptide

The recommended experimental design is to use flow cell 1 as a reference. We will describe a method using flow cell 2 for biotinylated degron peptide, keeping flow cells 3 and 4 in reserve (Note 8). Coating is best run using the wizard, but may also be done with Manual Injections. You will need to consider surface coating density for kinetic experiments (See Note 9). Here we give conditions for high coating densities which are suitable for initial screening work.

1. Run -> Run Application Wizard… A new dialog box opens allowing you to select a wizard. Select -> Surface Preparation, -> Start.


4. Enter the Ligand Name for each cell to be coated (See Note 10).

5. Enter Injection Time and Flow rate as shown, -> Next.
6. The consumption of solutions is calculated and shown in the next box. Prepare the advised volumes of biocytin and biotinylated peptide at 0.05 mg/ml (stock diluted 1 to 20 in water) and place them into the rack positions indicated, -> Next.

7. Provide a filename when prompted, -> Save and the immobilization routine commences.
8. Sensorgrams are shown during the immobilisation run (Fig. 4).

Figure 4: SPR sensorgrams for an automated SA chip coating. For both flow cells it is clear that the signal does not return to baseline but reaches a new steady level showing that the surface was successfully coated. The sensorgram for peptide is typically ‘wavy’.

9. When finished, a Results box records the ligand bound for each cell/ligand in RUs (See Note 11).

<table>
<thead>
<tr>
<th>Sensor Chip SA</th>
<th>Response (bound)</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fc Ligand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Mectin</td>
<td>176.3</td>
<td>Injection time 5.0 min, flow 12 µl/min</td>
</tr>
<tr>
<td>2 AA7</td>
<td>189.3</td>
<td>Injection time 5.0 min, flow 12 µl/min</td>
</tr>
</tbody>
</table>

The bound RUs represent the differences between the baselines before and the ‘new baselines’ after each immobilization (Note: the report points are marked by an x at the data record points in Fig. 4. Cells 1 and 2 are now ready for use.
3.3 TIR1 Activity Assay

The purpose of the assay is to test protein for auxin binding activity. This assay is a simple and effective procedure to confirm successful purification of active TIR1. More analytical experiments are described below.

Sample preparation

TIR1 protein is co-expressed with ASK1 using the baculovirus system (Calderon Villalobos et al., 2013; Lee et al., 2014) and purified. In brief, cells 3 days post-infection with recombinant virus are harvested by centrifugation. Cells are lysed, the lysate cleared and proteins purified by nickel affinity chromatography. The eluted protein is incubated overnight with TEV protease (to remove His-MBP fusion tags) before affinity capture of TIR1 (and associated ASK1) by FLAG affinity chromatography. Elution is in SPR running buffer plus FLAG peptide.

1. In 3 labelled vials prepare the following mixes (See Note 12):

<table>
<thead>
<tr>
<th>Sample</th>
<th>Volumes (ul)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IAA*</td>
<td>DMSO</td>
<td>Running Buffer</td>
<td>TIR1*</td>
<td>Final Volume</td>
</tr>
<tr>
<td>Buffer Blank</td>
<td>0</td>
<td>0</td>
<td>400</td>
<td>0</td>
<td>400</td>
</tr>
<tr>
<td>TIR1 + IAA</td>
<td>0.5</td>
<td>0</td>
<td>49.5</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>TIR1 - IAA</td>
<td>0</td>
<td>0.5</td>
<td>49.5</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>

* From the 10 mM DMSO working stock this gives 50 uM IAA and DMSO is diluted to 0.5%.
** TIR1 concentration does not need to be measured because we are looking for yes/no answers.

2. Mix samples thoroughly by gentle vortexing, cap and keep on ice.

SPR setup

1. Cease Standby mode -> Stop.

2. If necessary, Dock your SA chip and Prime.
3. Run -> Run Application Wizard… Select -> Binding Analysis, -> Start.

4. Select -> Direct Binding, as we have already prepared the chip, -> Next.

5. Use Flow Cell(s) -> ‘2 with 1 as reference’.
6. Flow Rate -> 20 ul/min. The rest of the parameters are left as defaults, -> Next.

7. Ensure that Run Order is ‘As Entered’ (see Note 13)

8. You need to specify how many times each sample is assayed. In the column Repl. enter 1 against each sample except the buffer blank (See note 14), as shown. -> Next.
9. After each binding cycle the coated surfaces need to be cleared of bound analyte, a process referred to as regeneration. We use 50 mM NaOH. Use default values.
   Regeneration Method -> Single Injection,
   Flow Rate - 30 ul/min,
   Solution - 50 mM NaOH
   Injection Time - 30 s.
   -> Next.

10. Rack Positions for all sample tubes are pre-selected, but you can change them by drag-and-dropping onto the desired positions. Make sure vial contents correspond to positions. -> Next.
11. This is your last chance to check the vial contents, volumes and positions. Ensure the ‘Standby Flow After Run’ box is checked (this is the default) 
Click -> Start, and enter a filename when prompted.

12. At the end of the run you will see a data summary box with a set of tabs. Inspect these (especially the baseline drift report), but you will get a better picture of your data by opening the .blr file in BIAevaluation software.

13. The temperature of the chip and binding experiment will be 25 C as default.
Data processing

Data processing has several goals, including normalization, removal of anomalies (air bubble spikes etc.) and baseline adjustment. Later, additional manipulations are needed for successful kinetic analysis. Open the BIAevaluation package. This can also be started from inside BIACORE 2000 Control Software by hitting on the abacus button.

1. File -> Open then navigate to the results file.

2. A dialogue box will enquire which of the collected data you want to process and display. We want to see the binding minus control data, Flow cell (Fc)2 - Fc1.
Open -> Curves only
Values -> click on the ‘2 - 1’. The data files from all individual flow cells are deselected.
Curves -> select all the from the options, click -> OK.

3. You will see in the Project window the names, Ids and Sources of the data you are going to process.
4. Each curve is given an automatic name consisting of the file name of the results file, the cell selection (e.g. Fc+2-1) and the order in which they were run. To edit the name and display or change features of a binding curve, double left click on the respective title and work within the Curve Properties dialogue box.

The end result will appear as shown below:

5. Select the TIR1 + IAA and TIR1 - IAA traces and click on the graph icon button.
6. Normalize the curves on the pre-binding baseline. As shown below:
   Right click and drag over the initial baseline area marking out a region on which you want to normalize. Click -> Y-transform button.

7. A dialogue box opens with several choices of transformation. Select -> Zero at Average of Selection -> Replace Original. Both curves now start at 0 RU.

8. Cut out the spikes caused during regeneration and/or at the start and end of injections. Use right mouse to make a selection, as above;
   Click on Edit -> Cut. Data from both curves will be edited simultaneously.
9. Baseline corrections can be made as the final processing step, removing any bulk effects from solvents and/or non specific binding (in the example above there are almost none). Click on -> Y-Transform and select ‘Curve-Curve 2’ Select ‘TIR1 - IAA’ and -> Replace Original.

![Image of Y-Transform dialog box]

10. The data are now processed and clearly show TIR1 activity. In this case \( R_{max} \) is around 450 RU (See Note 15). This rapid assay confirms that the protein is suitable for compound screening or kinetic analysis. (See Notes 16, 17).

![Graph showing TIR1 activity]

- \( R_{max} \): Maximum response.
3.4 Compound screening assay

The screening assay described above is readily adapted for compound library screening. Purified TIR1 protein is divided and mixed with compounds, each diluted to 50 micromolar from 10 mM stocks dissolved in DMSO. Buffer blanks and controls with zero auxin are run in every set, and a number of replicates containing 50 micromolar IAA as positive control and to help normalise data between protein preparations. The activity of a compound is expressed as a percentage of IAA’s activity, taking the value from a report point taken just before the end of the association phase of the binding curve. Compound activities are classified as shown in Fig 5.

![Figure 5: Classification of binders based on % activity with respect to IAA. Data has undergone additional processing with Prism version 6.00f for Mac OS X, GraphPad Software, La Jolla California USA, www.graphpad.com](image)
3.5 Kinetics

In order to get values for affinity and rates of association and dissociation a kinetic experiment is needed. For kinetic work we need to ensure that no factor is limiting except one variable, the concentration of analyte. In particular, we need to ensure that there is no mass transport limitation (See Note 18), and we start by limiting coating of the chip surface (See Note 9) such that Rmax is 300 RU or lower (on Biacore 2000 and 3000 instruments). On the Biacore 2000 it is best to plan on using only flow cells 1 and 2.

1. For the actual experiment Run -> Run Application Wizard, -> Kinetic Analysis
2.

3. Provided that you have figured out optimal flow rates and surface coatings, go directly to Concentration Series. (See Note
Similarly to before for Assay Principle select Direct Binding. You will be prompted to fill in a concentration series about the chosen analyte molecule (in our case we let the concentration of IAA vary). The chosen flow rate is 25 ul/min. Select ‘2, Using 1 as reference’. Input the molar concentrations and the molecular weight.

Select the regeneration conditions and the rack positions in a similar way to the activity check assay. Check the summary. The wizard will then run the experiment and present you with a summary table of kinetic parameters, statistics etc. Kinetic analysis using the other flow cells remains possible in BIAevaluation.

Data processing for kinetic experiments shares some similarities with the one for protein activity check. Particularly important is the selection of curves to fit and create a working set as an Overlay plot. There are a number of kinetic models that can be applied to data. TIR1 has a single binding pocket for auxin and so we use the 1:1 Langmuir model (See Note 19).

Start the kinetic evaluation wizard by choosing Fit -> Kinetics Simultaneous ka/kd or using the toolbar button. The wizard will lead you through the procedure.

The first request is to process the curves as to remove regeneration spikes and to transform the y-axis as to obtain a zero baseline (analogous to the activity check processing).
The next request is via moving the cursors that appear to select portions of the curve that will correspond to association, steady-state and dissociation.

Hitting next will produce a summary table. The software is ready to fit.
We have displayed results of a fitting session with the residuals and the fitted parameters of the model.
Notes:

Note 1: The surfactant concentration is x10 higher than the suggested one in the ready-made buffer preparations from GE (0.005% v/v). We found out that whilst having no negative effects on binding, it does improve the smoothness of the signal and reduces bubble generation especially in instruments where fluidics are not brand new.

Note 2: Always Prime when exchanging the running solution, even between two batches of SPR running buffer. This exchanges the fluid in the pumps, flushes air bubbles and washes the chip.

Note 3: We strongly recommend that you condition the new chip before coating by running this extra cleaning step. It removes from the binding surface debris left from manufacture.

Note 4: The Detection dialogue box allows each cell to be selected independently. If you are cleaning flow cells 1 and 2 you select the option Fc 1-2. The reference subtraction - none, means that the signal from both cells will be displayed separately. The subtracted 2-1 signal is not helpful when cleaning. Sensograms will be recorded only from the flow cells you select.

Note 5: All files with extension .blr contain raw results and can be opened and post-processed with BIAevaluation software as will be shown.

Note 6: Notice the two flow cells of a brand new chip give different absolute response units (RU) when same fluid is moved over them. These units are arbitrary. When analyzing the data what matters is the response difference.

Note 7: The manual run can be adapted for other occasions when manual control over the process is preferable. Select combinations of commands via the buttons to create appropriate command queues.

Note 8: Heavy useage with flow cells left uncoated tends to allow debris to accumulate over them and impair subsequent coating. Whilst economical to hold open cells 3 and 4, remember that their performance may deteriorate if held for long periods. If you know what you want to coat 3 and 4 with, we recommend that all cells are coated at one time.

Note 9: You may vary surface coating density by diluting the biotinylated reagent or by reducing coating time. In practice with biotinylated reagents, the affinity between biotin and streptavidin is so high ($10^{-15}$ M) the better strategy is dilution. Our biotinylated peptide stocks at 1 mg/ml are diluted 200-fold for saturating chip surfaces, but diluted up to a further 1000-fold for kinetic experiments for which mass transport limitation (Note 18) needs to be avoided. You may also wish to use the Surface Preparation wizard and the facility for Target coating. In this you provide dilute biotinylated ligand and enter an RU value that you want to achieve. The software will then sip, wash and check as coating progresses, with short injections used to allow ligand binding up to the target value.

Note 10: The molecule captured on the chip is referred to as the ligand in Biacore terminology. The molecule in solution which binds to the immobilised ligand is referred to as the analyte.

Note 11: Although arbitrary, RUs scale directly with mass and you will see that far more peptide appears immobilised than biocytin. In fact, more moles of biocytin are immobilised than the peptide, probably because the small biocytin can access more available sites on the surface.
Note 12: For TIR1 + IAA and TIR1 - IAA we recommend pipetting the IAA in DMSO first onto the walls of the sample tube followed by the buffer; this will ensure better mixing.

Note 13: In more advanced analyses you may e.g. randomize sampling.

Note 14: We recommend that assay runs are preceded by at least 3 buffer injections to ensure the cells are purged and stabilised. The regeneration of the coated surfaces between assay cycles will remove some loosely associated ligand on the first runs and a few regeneration cycles settle the surface before accurate measurements are taken.

Note 15: $R_{\text{max}}$ is the value at the top of the binding hyperbola and represents full binding site occupancy. This value is important in kinetic experiments.

Note 16: The data can also be exported for processing in your favourite software. Select File then -> Export.

Note 17: Files from BIAevaluation are saved as .ble files.

Note 18: Mass transport limitation. If the rate of exchange of analyte from solution with the ligand on the surface is limiting, or binding is so advanced at the front of the flow cell such that the analyte concentration flowing over areas of the flow cell behind the front is depleted, the measured rates of binding will be affected. This happens, for example, when the coating density is too high or flow rates too slow. Even if you follow best practice with low coating densities we recommend that you check for mass transport limitation by evaluating association rates at different flow rates. The kinetic wizard offers this as an optional extra.

Note 19: The Kinetic Analysis dialogue box offers you the option of selecting two additional experiments, Mass Transport Limitation and Linked Reactions. If you have already checked that a flow rate of 30 ul/min will not limit the association rate, you do not need to run the mass transport option. However, if you have not checked, we recommend that this option is included. It will ask for more protein sample (positive control mix) and will run the binding at 5, 15.

Note 13: Once experiments are concluded the chip is undocked and the system is left in Standby mode with maintenance chip in place and SPR running solution circulating. In this condition the fluidics are kept as such for a maximum of 4 days (7 on newer Biacore instruments). Once undocked, store the SPR chip in a 50mL falcon tube and keep at 4°C.
Appendix B: TomoDock methodology

publication

Uzunova, V., Quareshy, M., Genio, C. I. del., Napier, R.M. Tomographic docking suggests the mechanism of auxin receptor TIR1 selectivity. Open Biology, accepted
Tomographic docking suggests the mechanism of auxin receptor TIR1 selectivity

Veselina V. Uznova, Mussa Quareshy, Charo I. del Genio, and Richard M. Napier*
School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
(Dated: August 26, 2016)

We study the binding of plant hormone IAA on its receptor TIR1 introducing a novel computational method that we call tomographic docking and that accounts for interactions occurring along the depth of the binding pocket. Our results suggest that selectivity is related to constraints that potential ligands encounter on their way from the surface of the protein to their final position at the pocket bottom. Tomographic docking helps develop specific hypotheses about ligand binding, distinguishing binders from non-binders, and suggests that binding is a three-step mechanism, consisting of engagement with a niche in the back wall of the pocket, interaction with a molecular filter which allows or precludes further descent of ligands, and binding on the pocket base. Only molecules that are able to descend the pocket and bind at its base allow the co-receptor IAA7 to bind on the complex, thus behaving as active auxins. Analyzing the interactions at different depths, our new method helps in identifying critical residues that constitute preferred future study targets and in the quest for safe and effective herbicides. Also, it has the potential to extend the utility of docking from ligand searches to the study of processes contributing to selectivity.

Keywords: tomographic docking, auxin, receptor selectivity, molecular filter

I. INTRODUCTION

The molecular recognition of specific small organic compounds by target proteins is of central importance in biology. Auxins are a particularly relevant class of small molecule plant hormones with considerable importance for growth and development. Both the naturally occurring indole-3-acetic acid (IAA) and synthetic auxins bind to the Transport Inhibitor Response 1 (TIR1) family of receptors. In turn, auxin binding to the receptor allows the co-receptor IAA7 to bind to the substrate receptor complex [1, 2]. Thus, one can say that auxins act as “molecular glues” between partners of the receptor system. The completion of this two-step mechanism triggers a cascade of events leading to changes in gene expression [3]. The macroscopic results of acute exposure to synthetic auxins are explosive, epinastic growth followed by plant death. Thus, such compounds have found widespread application in herbicidal formulations.

Further valuable features of auxin-based herbicides are a long history of safe use and their selective action against broad-leaved plants, making them preferred products for the control of weeds in cereal crops and turf [4]. However, rational design of novel biologically active molecules to influence the TIR1 receptor has proved challenging because the protein recognizes a diverse set of natural and synthetic ligands [2]. At the same time, TIR1 is highly selective. For instance, the native ligand IAA shares many structural features with its biosynthetic precursor, the indolic amino acid L-tryptophan (Trp), which, although ubiquitous and present at intracellular concentrations far in excess of that of IAA, does not elicit auxin responses [5].

A likely reason for this is that inactive compounds do not bind the receptor in the right location or with the right orientation, if at all, thus precluding assembly of the co-receptor complex. Then, knowledge of the mechanism of interaction for natural ligands is fundamental for the design of synthetic analogues. Computational methods for molecular docking have become standard tools in active compound design and discovery [6–8]. They allow a reduction of the search space, leading to targeted experimental binding assays, and they are widely used for ligand screening and identification of binding sites on bioactive targets [9, 10]. Specific examples of the application of molecular docking are the identification of a genetic cause of cancer drug resistance [11], ligand differentiation between human oestrogen receptors [12], rational drug design for neurodegenerative diseases [13], in silico screening of candidate therapeutic binding to mutated targets [14], and design of highly catalytic artificial metalloenzymes [15]. In all cases the binding site is considered as a single, holistic search space. In this article, we introduce a new approach to molecular docking, which we call tomographic docking, which we use to propose an explanation for the discrimination mechanism of small ligands by the TIR1 receptor.

One frequently overlooked aspect of the molecular recognition process is the depth of the protein binding site, which can extend significantly towards the protein core. Several computational tools exist that help describe and define pockets, tunnels, channels and pores, and some will identify the most likely high-affinity sites in the recess. Once defined, these are offered as binding sites for docking. While this approach is able to find a good candidate for the lowest-energy configuration of a given receptor-binder pair, it risks neglecting receptor features that will be encountered by the ligand on the approach to the best site. When this happens, it contributes to ligand misidentification and false positive results, both of which are recognized issues with docking experiments [16]. For example, AutoDock Vina [17, 18], which is currently one of the most popular docking platforms due to its speed, reliability and output accessibility, finds an apparently viable docking pose for Trp on TIR1, even though Trp is experimentally proven to be a non-binder. It is thus rea-
Figure 1: The deep binding pocket on TIR1. The binding site of TIR1 is not a shallow surface indentation but a deep pocket, shown in SURF representation in the left panel superimposed on the cartoon of the whole receptor (2P1P). The residues comprising the binding pocket, isolated and shown from a closer point in the right panel, are contributed by seven, non-sequential leucine-rich repeats [1]. The two reference residues Phe-351, in red, and Arg-403, in blue, indicate the mouth and the bottom of the pocket, respectively. Their distance, indicated by the arrow in the figure and corresponding to the pocket depth, is 16.5 Å. IAA is shown in CPK representation, and InsP6 is on the bottom left in bond representation. All 3d molecular visualizations were produced using VMD [21, 22].

II. METHODS
A. The target protein

The X-ray crystal structure of TIR1 is solved in three different binding states: unbound/empty (2P1M), in complex with the natural ligand IAA (2P1P, Fig.1), and assembled with both IAA and its co-receptor IAA7 (2P1Q). For our study, we use the unbound structure 2P1M as the closest approximation to what a free ligand interacts with. Superposition of 2P1M with 2P1P and 2P1Q suggests no significant conformational change is induced by ligand binding [1].

The crystallography data contain several associated biomolecules, as well as water. Thus, to prepare the docking input, we first processed the data using VMD [21, 22], excising water molecules and the co-expressed SCF

adaptor ASK1. AutoDockTools [23, 24] was then used to produce the final pdbqt input files. Note that TIR1 harbours inositol hexaphosphate (InsP6) as a second, probably structural, ligand. We left this in place because, unlike ASK1, its location is physically close to the bottom of the pocket. Nonetheless, our final results show that its position is still too far from the binding site to generate effective interactions with the ligand.

B. Ligands

The investigation focusses on IAA and Trp as TIR1 ligands because Trp, which is a precursor in the synthesis of IAA [25], has no auxin activity and no TIR1-binding activity [5], notwithstanding a significant structural similarity with IAA (Fig. S1). To validate the accuracy of our approach, we later extend the analysis to a few other compounds, comprising both binders and non-binders, with different degrees of structural similarity to IAA (Fig. S2). To prepare the ligands for docking, we first calculated their protonation state at pH 7.3 in wa-
Figure 2: Comparison of crystallographic data and “static” docking results. The docked position of IAA at the bottom of the pocket, in the left panel, matches the crystallographic one in the centre panel. The docked position of Trp at the bottom of the pocket, in the right panel, does not match the docking and crystallographic positions of IAA.

C. Numerical setup

To define the search space to be used in the docking algorithm, we observed SURF representations of TIR1 in both bound and unbound states, and noted that the binding site is not a shallow surface feature, but rather a deep binding pocket (see Fig. 1). In particular, we identified the constituents of the pocket to be a total of 43 amino acids in seven contiguous sets on the leucine-rich repeat loops, namely residues 77–84, 344–354, 377–381, 403–410, 436–441, 462–465 and 489–490. The pocket thus defined has a depth of 16.5 Å, as measured between Phe-351 at the mouth and Arg-403 at the bottom. To investigate the engagement process as the ligand moves into it towards the final binding site at the bottom, we defined an 18 Å×18 Å×18 Å cubic search space that moves from above the pocket mouth to below the bottom in steps of 1 Å. The search space at the first step includes Phe-351 at its bottom, and its motion is parallel to the principal axis of inertia of the receptor whose direction is along the pocket depth. At the last step, Arg-403 is completely included. Then, we performed independent numerical docking experiments at each step, building a sequence of results that provide information on the descent of the ligands into the deep pocket. For the actual simulations, we created a code that automates the tomographic scanning process by computing the geometry of the search space for any specified number of steps, search exhaustiveness, and set of ligands. The code, which we refer to as TomoDock, uses AutoDock Vina [17, 18] as docking engine, and produces tunable summaries of the results, as well as pdb files for further analysis and visualization. Note that with the choices detailed above, the search space is always larger than Vina’s cutoff threshold for the interactions, which is 8 Å. The standard TomoDock experiment was repeated 100 times, with search exhaustiveness of 16.

D. Experimental setup

Experimental evaluation of the numerical results was carried out using surface plasmon resonance (SPR) as a test of ligand binding to TIR1, and root growth assays for overall biological activity. We performed protein purification and set up the SPR experiments according to the protocols described in [5]. TIR1 was expressed in insect cell culture using a recombinant baculovirus. The construct contained sequences for three affinity tags, namely 6×His, maltose-binding protein (MBP) and FLAG. Initial purification using the His tag was followed by cleavage of His-MBP using TEV protease. After TEV removal and clean-up using FLAG chromatography, the purified protein was used for SPR assays by passing it over a streptavidin chip loaded with biotinylated IAA7 degron peptide. The SPR buffer was Heps-buffered saline with...
Figure 3: Best scores of docked poses along the transect for all compounds tested, namely IAA (blue circles), Trp (red squares), 1-NAA (magenta diamonds), 2,4-D (cyan triangles), IPA (orange triangles), ICA (green triangles), IBA (violet triangles) and IDA (olive hexagons). Each step represents the progression of the search space by 1 Å in the direction of the pocket bottom. Filled symbols indicate steps at which a significant change in depth or orientation of the docked pose occurs.

1 mM EDTA, 0.05% P20 and 1mM TCEP. Compounds to be tested were premixed with the protein to a final 50 µM concentration. Binding experiments were run at a flow rate of 20 µl/min using 3 minutes of injection time and 2.5 minutes of dissociation time. Data from a control channel (biocytin) and from a buffer-only run were subtracted from each sensorgram following the standard double reference subtraction protocol. To assay
III. RESULTS AND DISCUSSION

A. Conventional docking

We first ran a conventional, “static” docking experiment using AutoDock Vina using a cubic search space with an 18-Å side encompassing the whole pocket area. The results show that IAA docks at the pocket bottom and, even from a single run, the best docked position closely matches that of the ligand bound in the crystallised structure (Fig. 2). The indole ring is aligned parallel to the pocket bottom and is nested in a semi-circle of four non-polar residues, while the carboxylate anion orients itself towards a group of basic residues with which hydrogen bonds are made [1]. Despite the absence of activity for Trp as an auxin and no measurable affinity for binding, AutoDock Vina finds an apparently plausible docked position for it at the bottom of the pocket, although not with the same orientation as IAA (Fig. 2). This clearly shows that one cannot rely on a direct interpretation of docking results to identify binders, because a “cover-all” search space overlooks key features of the binding process and a more systemic approach is needed.

B. Tomographic docking

To study the transient interactions of IAA and Trp with the pocket as they move down into it, we performed tomographic docking experiments and analyzed each series of docked poses in detail, building a plausible binding pathway for both compounds over a transect of 15 Å. The docking process assigns a lower numerical score to better poses, representative of lower energy and more favourable binding. Thus, for each ligand we created a representative series of successive orientations choosing at each step the pose with the lowest score amongst all repetitions (Fig. 3). Note that the depth at which the ligand is positioned does not necessarily increase with step number. For instance, as described in greater detail below, neither the depth, nor the orientation of the docked pose of IAA changes between step 4 and step 7, indicating that the interactions relevant over these steps are dominated by the residues included at step 4, and that no further significant interactions are made until the ligand approaches residues deeper in the pocket than those at step 7. Later, we use these considerations to identify which residues are most likely to be responsible for the selection mechanism.

At step 1 both compounds are well out of the pocket, and at step 2 they are at the very edge of the pocket mouth. As the steps continue, the progressive inclusion of residues causes the docked position of IAA to undergo significant changes. At step 3, IAA has oriented itself with the carboxylic acid group in a niche at the back wall of the pocket (Fig. 4A). Then, for the next four steps, its scoring function, position and orientation remain constant, with the alignment of the indole perpendicular to the base of the pocket (Fig. 4B). Note that between step 3 and steps 4–7, IAA undergoes a small but significant rotation, which optimizes the perpendicularity of the indole-ring system with respect to the bottom of the pocket.

Considering Trp, its side-chain also becomes oriented towards the niche at step 3, and its docked depth and orientation do not change through step 6. However, unlike IAA, the orientation of the indole is parallel, not perpendicular, to the base of the pocket. This is probably due to the longer side chain and the extra rotational freedom, as well as to the proximity of the aromatic rings to residues distal to the niche (Fig. 4C). The next step for Trp, step 7, presents a somersault for the pose, with its polar side groups now pointing towards the pocket bottom (Fig. 4D).

From the pose in step 7, with the tail in the niche, IAA can proceed downwards, into the position observed at step 8, via a pivoting motion of the indole from the engagement niche (cf. Fig. 4B and Fig. 4E). Poses 8 and 9 for IAA are identical, with the side-chain continuing to point towards the niche, but not in it. Then, there appears to be a final transition as residues at the base of the pocket come into play, with poses 10 to 15 showing that IAA has flipped over from poses 8 and 9 (Fig. 4F), to a position that corresponds to that found in the crystal structure [1].

C. Binding mechanism

When the docking algorithm explores positions that include the pocket bottom, Trp is docked at the binding site. This indicates that, in principle, the final docking position is allowed. However, a detailed examination of the tomographic docking results suggests the presence of a barrier impeding the descent of the non-binder into the binding pocket, explaining why, in nature, Trp never reaches its bottom.

In the initial part of the pocket the tomographic docking identifies a region that we call the engagement niche formed by residues Lys-410, Ser-440, Gly-441, Ala-464 and Phe-465 (in violet in Fig. 5). This is the structure into which the binder orients its polar side-chain (Fig. 4B). We deem it one of the features with which potential binders need to interact in order to achieve an orientation that allows a subsequent transition to the
Figure 4: Progressive docking poses of IAA (blue) and Trp (red). The residues that form the engagement niche are highlighted in violet; those that we identify later as forming a molecular filter are highlighted in orange. (A) The position of IAA at step 3 features the side chain oriented towards the engagement niche. (B) The positions of IAA in steps 4 to 7 are superimposable, and show that the ring system is perpendicular to the bottom of the pocket. (C) The positions of Trp in steps 3 to 6 are superimposable. Its side-chain is oriented towards the niche, but the ring system is parallel, rather than perpendicular, to the pocket base. (D) At step 7, the side-chain of Trp is no longer in the niche, but points towards the pocket bottom. (E) At steps 8 and 9 (superimposable), IAA has moved towards the bottom of the pocket. (F) The positions of IAA at steps 10 to 15 are superimposable.

binding position. TomoDock results suggest that the particular orientation with the ring system perpendicular to the pocket base is likely to be an essential step in the selection process. For ligands to penetrate deeper, the aromatic rings must slice down while the polar tail, anchored in the engagement niche, acts as a pivot point (cf. Fig. 4B and Fig. 4E). After this motion, the rings of IAA are positioned at the bottom of the pocket, in the vicinity of a semi-circle of non-polar residues. The ligand then undergoes a slight rotation, allowing the polar carboxylic acid group to flip and engage with the polar residues at the pocket base.

To understand what blocks Trp from moving the same way as IAA, consider the results from steps 3 to 7. At step 3, Trp is docked with its polar tail in the engagement niche. However, we do not observe the perpendicular orientation of the aromatic system that we see in IAA (Fig. 4C). Note that its orientation and docking depth do not change through step 6 (Fig. 3). Then, at step 7, Trp assumes a new pose with the side-chain completely out of the niche, and pointing towards the pocket bottom (Fig. 4D). Such a geometry prevents the non-binder from moving further into the binding position via the same rotation that IAA performs, due to inappropriate orientation of the indole rings and of the side chain.

Note that the non-binder is allowed to dock further down the pocket from step 8 onwards (Fig. 3), because docking considers any position that is geometrically accessible, disregarding the motion a ligand would have to undertake in order to reach it. Also, to be active auxins
substrates need to bind in the correct orientation at the bottom of the pocket. A compound that can only interact and bind at the mouth of the pocket cannot have auxinic activity. Thus, for brevity we refer to compounds that are able to achieve an appropriate binding position simply as “binders”. Conversely, we refer to the compounds, like Trp, that cannot reach the pocket bottom as “non-binders”.

**D. Validation and experimental verification**

The striking difference between IAA and Trp revealed by tomographic docking indicates that an important role is played by the residues that become available at steps 3, 4, 7, 8 and 10 (filled symbols in Fig. 3). In particular, the TomoDock results suggest that they act as a molecular filter, promoting the correct orientation of IAA, and opposing it for Trp. To identify these residues, we built a table of the atoms newly included for interaction at each step, along with the residues they belong to (Table S1). Then, we considered the new entries at the steps indicated above, taking into account the number of atoms that interact, as well as their properties.

As a first example, consider Ser-438 (see Fig. 5). This residue enters the search space at step 10, which is the first step at which IAA assumes its final binding position. Ser-438 is physically located at the bottom of the pocket and, upon close inspection, the atoms that get included at step 10 are seen to form a highly polar group. Thus, we include it in the molecular filter, and consider it responsible for the correct orientation of IAA at the pocket bottom.

As a second example, consider Ser-440. This residue is structurally part of the engagement niche, and it enters the search space at step 6, with 3 atoms. However, neither IAA nor Trp changes its position at all over this step (see Fig. 4B and C). At step 7, where 4 more atoms of Ser-440 are considered, Trp exits the niche (Fig. 4D). One could thus consider Ser-440 partly responsible for this; however, at step 7 IAA maintains the same position as it has at step 6. Given the structural similarity of the two molecules, we believe this indicates that Ser-440 does not contribute actively to ligand filtering, particularly considering that a much better candidate for the observed effect on Trp exists, namely Leu-439.

The third example we discuss is Gly-441. This is a noteworthy residue, as it is not only structurally part of the engagement niche, but also because its mutation to aspartate yields the known tir1-2 mutant [27]. This residue gives its first big contribution to the search space at step 4. But by this step both IAA and Trp have already assumed positions that do not change for a few more steps. Thus, we do not consider this residue as an active player in the molecular filter. Substitution of the large polar side group of Asp for the small non-polar Gly could interfere with binding in many ways to explain the tir1-2 phenotype.

Performing this analysis on all viable residues shows that the filter is formed by Cys-405, Ser-438, Leu-439, Ser-462 and Arg-489 (in orange in Fig. 5). Of these, Arg-489 seems to be the residue that most significantly affects the orientation of the compounds with respect to the engagement niche, of which, however, it is not a part. Leu-439 and Ser-462 appear to block the descent of Trp and promote that of IAA, as they progressively enter the search space in steps 7–9. Finally, Cys-405 and Ser-438 are likely to be instrumental for IAA to assume the final binding position, since they start contributing significantly at steps 8 and 10, respectively. Note that all the residues in physical proximity of the ligand at the pocket bottom are likely to be involved in stabilizing docked auxins, but they are not necessarily part of the filtering mechanism.

To further exemplify our method, we apply it to six other potential binders (Fig. S2), namely indole-3-carboxylic acid (ICA), indole-3-propionic acid (IPA), indole-3-butyric acid (IBA), indol-3-yl acetate (IDA), 1-naphthaleneacetic acid (1-NAA) and 2,4-dichlorophenoxyacetic acid (2,4-D). The aliphatic side-chains of ICA, IPA and IBA differ in length from that of IAA, being one atom shorter, one atom longer and two atoms longer, respectively. Esterifying indole-3-ol with acetic acid gives IDA, while changing the indole system to a naphthalene double ring yields 1-NAA. Finally, we include 2,4-D as it is one of the most widely used herbicides in the world. For all compounds, tomographic docking predicts a plateau between steps 3 and 7 (Fig. 3). At these steps, all compounds are docked at the right depth to permit interaction with the engagement niche. Re-
Figure 6: Docking poses of 1-NAA (magenta), 2,4-D (cyan), IPA (orange), ICA (green), IBA (violet) and IDA (olive). (A) The positions of 1-NAA at steps 4 to 7 (superimposable) show the side-chain of the ligand engaged with the niche, and its ring system perpendicular to the pocket bottom. (B) In steps 5 to 8 (superimposable) the orientation of 2,4-D is the same as that of IAA and 1-NAA. (C) At steps 3 to 7, also IPA engages the niche with the ring system perpendicular to the pocket base. The positions at these steps are superimposable. (D) In steps 3 to 7 (superimposable), the tail of ICA never engages the niche. (E) The tail of IBA finds the niche in steps 3 to 7 (superimposable), but the ring system is misoriented. (F) At step 7, the side-chain of IDA is oriented towards the pocket bottom, in a pose reminiscent of that of Trp at the same depth.

Experimental confirmation of binding from SPR and root growth assays, (Fig. 7 and Tab. I), support the numerical predictions. In particular, SPR experiments indicate that Trp, ICA, IBA and IDA have no measurable binding activity at 50 µM, a concentration far in excess of the IC50 value of all active auxins (Tab. I). IPA and 2,4-D, instead, bind weakly compared to IAA and 1-NAA, with noticeably more rapid dissociation rates. Like IAA, 1-NAA is a strong ligand. Notice that, as mentioned above, substrate binding to the receptor is only the first step in the auxinic interaction, and the binding of the co-receptor IAA7 (in our case) to the substrate-receptor complex can only happen if the substrate is bound in the bottom of the receptor pocket, and in the correct orientation. Thus, SPR experiments offer a good method to validate the computational results: if a substrate binds...
incorrectly, it will impair or entirely preclude the binding of the receptor to the IAA7 co-receptor and produce no SPR signal. The relative effectiveness of each compound in root growth assays compares favourably with the SPR measurements. Trp, ICA, IBA and IDA inhibit root growth only at very high concentrations, where phytotoxicity sets in. IPA is seen to be a weak auxin. The root growth assays suggest that 2,4-D is the most active auxin, more active than the SPR data suggest. However, root growth assay activities depend on tissue and cellular transport, as well as on receptor binding of the compounds in question, and so IC50 values do not correspond exactly to in vitro binding values. Nonetheless, the assays are still a useful verification method, as one cannot observe root-growth inhibition for compounds that do not correctly bind to the receptor.

IV. CONCLUSIONS

We have studied the binding process of the plant growth hormone IAA onto its main receptor TIR1, to investigate its selectivity mechanism. For our study, we developed a novel tomographic docking approach suitable for investigating deep binding pockets in a series of pseudo-time steps. The method gradually changes the search space of a docking algorithm to allow one to consider sequential interactions of each potential ligand with pocket residues at increasing depths. This mimics what happens in nature when a small molecule descends into a binding cavity. In the present study we have considered the receptor structure as rigid, as is the case for most docking experiments. However tomographic docking can be adapted to allow for the flexibility of some side-chains of the receptor, and this will be the subject of future work. The tomographic method shows a plateau of scoring function values part-way down the pocket, indicating a region over which transient interactions are made en route to the docking site at the base.

Detailed study of the docking poses obtained for the natural ligand IAA and the related, but non-auxin Trp over this region points towards two features in the pocket responsible for selectivity. The first, an engagement niche in the back of the pocket, allows potential ligands to orient before subsequent motion towards the binding site. The second, a molecular filter, promotes the correct pose of the aromatic ring system for binders, necessary to access the pocket bottom. Tryptophan and a set of non-binders assume sub-optimal orientations, and are prohibited from onwards motion.

The identification of the residues that form the engagement niche and the molecular filter makes them a fundamental study subject for the rational design of novel auxin-based herbicides. They are critical for selectivity, and constitute preferred mutation targets for further experiments. One such mutant, tir1-2 (G411-Asp), is already known, and experiments have shown that the mutation has a small but measurable effect on the resistance of the plant to the auxin-transport inhibitor 2-carboxyphenyl-3-phenylpropane-1,2-dione (CPD) [27],

<table>
<thead>
<tr>
<th>Compound</th>
<th>IC50 (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,4-D</td>
<td>0.0118± 0.0009</td>
</tr>
<tr>
<td>IAA</td>
<td>0.041 ± 0.007</td>
</tr>
<tr>
<td>1-NAA</td>
<td>0.206 ± 0.015</td>
</tr>
<tr>
<td>IPA</td>
<td>0.51 ± 0.06</td>
</tr>
<tr>
<td>IBA</td>
<td>10.1 ± 1.2</td>
</tr>
<tr>
<td>ICA</td>
<td>10.4 ± 1.2</td>
</tr>
<tr>
<td>IDA</td>
<td>15 ± 2</td>
</tr>
<tr>
<td>Trp</td>
<td>102 ± 13</td>
</tr>
</tbody>
</table>

Table I: IC50 for primary root growth inhibition derived from root growth assays.
Although this non-conservative substitution may not help elucidate the role of residue 441 further. Experimental results from SPR and root growth assays performed on a set of active and inactive compounds are consistent with TomoDock results, confirming the validity of our method. The application of tomographic docking need not be limited to the analysis of auxin binding to TIR1. In fact, it can be used to examine also other members of the TIR/AFB auxin receptor family. Identifying similarities and differences between the interaction mechanisms in different receptors can play a key role in designing receptor-specific compounds, which are very useful in controlling herbicide resistance. In addition, the tomographic docking principle is general, and can be applied to any deep binding site. Thus, proteins such as those involved in the transport of small molecules, as well as enzymes and channel proteins, are all natural targets for tomographic docking investigation.

Data accessibility

The protein structures used in this study are available from the Protein Data Bank website at the URL http://www.rcsb.org/pdb/home/home.do The TomoDock code is available upon request.

Competing interests

We have no competing interests.

Authors’ contributions

VUV and CIDG performed the numerical simulations. MQ designed and performed all the SPR and bioassay experiments. CIDG implemented the TomoDock code. All authors developed the principle of tomographic docking, starting from an original idea of VUV. All authors analyzed data and results, and wrote the manuscript.

Funding

VUV and RMN acknowledge support from BBSRC via grant n. BB/L009366. MQ acknowledges support from BBSRC via a studentship awarded through the Midlands Integrative Biosciences Training Partnership (MIBTP).

References


[20] ImageJ http://imagej.net/


[22] VMD - Visual Molecular Dynamics http://www.ks.uiuc.edu/Research/vmd/


Appendix C: TomoDock Supplementary Information
Tomographic docking suggests the mechanism of auxin receptor TIR1 selectivity
Supplementary material

Veselina V. Uzunova, Mussa Quareshy, Charo I. del Genio, and Richard M. Napier
School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
(Dated: August 24, 2016)

Figure S1: Structures of IAA (on the left) and Trp (on the right), displaying bond orders and charges at pH=7.3. The insets show the respective skeletal structural formulae.
Figure S2: Structure of compounds used for validation. All representations show bond order and charges at pH=7.3, with skeletal structural formulae in the insets. Top-left: indole-3-carboxylic acid (ICA). Top-right: indole-3-propionic acid (IPA). Middle-left: indole-3-butyric acid (IBA). Middle-right: 3-indolyl acetate (IDA). Bottom-left: 1-naphthaleneacetic acid (1-NAA). Bottom-right: 2,4-dichlorophenoxyacetic acid (2,4-D).
<table>
<thead>
<tr>
<th>Residue</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIS-78</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>PHE-79</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PHE-82</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>LEU-84</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>VAL-321</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ARG-344</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>VAL-345</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PHE-346</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>PRO-347</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>PRO-350</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>PHE-351</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>MET-353</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GLU-361</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>LEU-378</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>TYR-379</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>PHE-380</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>CYS-381</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>GLN-383</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>MET-384</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PHE-402</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ARG-403</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>10</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>LEU-404</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>CYS-405</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>ILE-406</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>ILE-407</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>GLU-408</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>PRO-409</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>LYS-410</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>ALA-411</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PHE-424</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ARG-436</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>LEU-437</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>SER-438</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>LEU-439</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>SER-440</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>GLY-441</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>SER-462</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>VAL-463</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>ALA-464</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>PHE-465</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>ALA-466</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>ARG-489</td>
<td>9</td>
<td>13</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>ASP-490</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CYS-491</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>PRO-492</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SER-515</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>IHP-601</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table S1: Number of atoms belonging to each residue included in the search space at each step. The residues in violet are those belonging to the engagement niche; those in orange constitute the molecular filter. The steps across which the contribution of the molecular filter residues changes significantly are highlighted in bold italic typefont.
Appendix D: NMR and Mass spectra
Figure D.1 Proton NMR of 3-[(1H-1,2,3,4-tetrazol-5-yl)methyl]-1H-indole
Figure D.2 Mass spectrum of 3-[(1H-1,2,3,4-tetrazol-5-yl)methyl]-1H-indole

<table>
<thead>
<tr>
<th>Acquisition Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Type</td>
<td>ESI</td>
</tr>
<tr>
<td>Focus</td>
<td>Active</td>
</tr>
<tr>
<td>Scan Begin</td>
<td>50 m/z</td>
</tr>
<tr>
<td>Scan End</td>
<td>3000 m/z</td>
</tr>
<tr>
<td>Ion Polarity</td>
<td>Positive</td>
</tr>
<tr>
<td>Set Capillary</td>
<td>4500 V</td>
</tr>
<tr>
<td>Set End Plate Offset</td>
<td>-500 V</td>
</tr>
<tr>
<td>Set Collision Cell RF</td>
<td>1200.0 Vpp</td>
</tr>
<tr>
<td>Set Nebulizer</td>
<td>0.4 Bar</td>
</tr>
<tr>
<td>Set Dry Heater</td>
<td>180 °C</td>
</tr>
<tr>
<td>Set Dry Gas</td>
<td>4.0 l/min</td>
</tr>
<tr>
<td>Set Divert Valve</td>
<td>Waste</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meas. m/z</th>
<th>#</th>
<th>Formula</th>
<th>Score</th>
<th>m/z</th>
<th>err</th>
<th>Mean</th>
<th>mSig</th>
<th>rdb</th>
<th>e~</th>
<th>Conf</th>
<th>N-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>200.0932</td>
<td>1</td>
<td>C 10 H 10 N 5</td>
<td>100.00</td>
<td>200.0931</td>
<td>-0.6</td>
<td>-0.6</td>
<td>50.7</td>
<td>8.5</td>
<td>even</td>
<td>ck</td>
<td></td>
</tr>
<tr>
<td>222.0750</td>
<td>1</td>
<td>C 10 H 8 N 5 Na</td>
<td>100.00</td>
<td>222.0750</td>
<td>0.1</td>
<td>-0.0</td>
<td>48.9</td>
<td>8.5</td>
<td>even</td>
<td>ck</td>
<td></td>
</tr>
</tbody>
</table>

+MS, 0.3-0.5min #(16-20)