

INVESTIGATING THE RECOGNITION  
AND INTERACTIONS OF NON-POLAR  
 $\alpha$  HELICES IN BIOLOGY

A THESIS SUBMITTED TO THE UNIVERSITY OF MANCHESTER  
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY  
IN THE FACULTY OF CHEMISTRY

2017

**James Alexander Baker**  
[orcid.org/0000-0003-0874-2298](https://orcid.org/0000-0003-0874-2298)

# Contents

Abstract	5
Abstract	6
Lay Abstract	7
Declaration	8
Copyright Statement	9
Acknowledgements	10

# List of Tables

# List of Figures

# The University of Manchester

James Alexander Baker

Doctor of Philosophy

Investigating the Recognition and Interactions of Non-Polar  $\alpha$  Helices in  
Biology

November 17, 2017

# Abstract

Non-polar helices figure prominently in structural biology, from the first protein structure (myoglobin) through trans-membrane segments, to current work on recognition of protein trafficking and quality control. Trans-membrane  $\alpha$  helix containing proteins make up around a quarter of all proteins, as well as two-thirds of drug targets, and contain some of the most critical proteins required for life as we know it. Yet they are fundamentally difficult to study experimentally. This is in part due to the very features that make them so biologically influential: their non-polar trans-membrane helix regions. What is missing in the current literature is a nuanced understanding of the complexities of the helix composition beyond a hydrophobic region of around 20 residues. Currently, it is known that the properties of trans-membrane protein  $\alpha$  helices underpin membrane protein insertion mechanisms.

By leveraging large datasets of trans-membrane proteins, this thesis is focused on characterising features of  $\alpha$  helices en masse, particularly regarding their topology, membrane-protein interactions, and intramembrane protein interactions.

In this thesis, I make the argument that there are different classifications of trans-membrane  $\alpha$  helices. These have markedly different evolutionary pressures, these different classes interact differently with the membrane, and each class serve the protein differently.

# Lay Abstract

The survival of each of our cells relies on a cellular barrier to separate themselves from the surrounding environment. This cellular skin can be thought of as the bag that contains all the important machinery required for normal cell function. The barrier works by being chemically very different to both the outside environment, and to the inside of the cell, which in both cases are mostly water. The membrane is fatty, and because of that repels water.

Proteins are the molecular machinery that form much of the cell structure and shape as well as carrying out many of the cell's routine tasks. Around a third of our genome codes for proteins that are permanently embedded in the membrane, but because these proteins are adapted for a life in the water repelling cell wall, they are very hard to study in laboratories which need to look at proteins in water.

In this thesis, we focus particularly on the parts of the protein that are embedded in the water repelling cellular skin. Traditionally, these regions are hard to study, because we must first remove them from the cellular wall, which causes problems since the embedded regions also repel water and this often causes them to stick to one another, making them hard to work with in a laboratory setting.

We analyse thousands of proteins to further our understanding of electrical charges in the embedded regions and find that negative charge on the outside of the cell has been evolutionarily selected across bacteria, animals, and plants. This is especially true for regions that specifically anchor the protein into the cellular wall. Where the embedded regions have additional function, for example ferrying something in or out of the cell, the negative charge "bias" can no longer be seen.

# Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.



# Copyright Statement

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s Policy on Presentation of Theses.

# Acknowledgements

I would like to thank all members of both the Eisenhaber research group, as well as the Curtis and Warwicker research group for discussion, but in particular Jim Warwicker, Frank Eisenhaber, Birgit Eisenhaber, and Wing-Cheong Wong for supervision and guidance during my research. I would also like to thank The University of Manchester and the A\*STAR Singapore Bioinformatics Institute for funding the project. Furthermore, I would like to extend my gratitude to the research group of Professor Stephen High.