

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Characterisation of an Interaction Involved in Viral Replication

Jolyon Claridge

Submitted in partial fulfilment of the requirements
of the degree of Doctor of Philosophy

Institute of Fundamental Sciences
Massey University

Abstract

Human rhinoviruses (HRVs) are a major cause of illness worldwide and, as members of *Picornaviridae*, are closely related to several other human and animal pathogens that exact a large medical and economic cost on society. Viral infections in general are particularly difficult to treat, as viruses co-opt many of the host's own biochemical pathways, making disabling the virus without harming the host very difficult. Carefully targeted strategies are required and detailed structural information is useful, both to identify new drug targets, and to fully understand interactions. One particular protein expressed by picornaviruses is 3C protease, which is responsible for post-translational processing of the viral capsid. This protease has a cysteine as its active site nucleophile, a functionality not found in eukaryotic proteases. The unusual active site makes 3C an attractive target for pharmaceuticals. Drugs that block the proteolytic action of 3C are currently in clinical trials. In addition to its proteolytic activity, 3C protease also has another function, that of an RNA binding protein. This activity has been shown to be required during replication of the viral RNA genome. In this study, the structure of 3C protease from HRV14 is investigated using NMR and other biophysical techniques. The structural information gained from these studies is used, along with data on 3C protease RNA-binding activity acquired using solution-state NMR and SAXS data, to elucidate a structure of the 3C-RNA complex. In addition, the dynamics of the free protein and of the protein in the presence of a specific inhibitor are investigated by solution-state NMR, and the potential role of dynamics in the function of the protein is explored. Finally, potential allosteric interaction between the RNA-binding and proteolytic functions of 3C is postulated, and further interactions of 3C and the 3C-RNA complex are discussed. It is hoped that a more complete understanding of 3C and its interactions will lead to more effective treatments for picornaviral infections in the future.

Acknowledgements

I would like to thank Dr. Steve Pascal for providing me with a grounding in NMR spectroscopy and for a great deal of helpful advice (mostly taken eventually...) about all aspects of the project. I also wish to thank Dr. Andrew Sutherland-Smith; Professors Geoff Jameson and Dave Harding and Dr. Gill Norris for many discussions and assistance over the years. I wish to thank Trevor Loo for a great deal of help with protein expression and purification when I was a new (and not so new) PhD. student. Dr. Pat Edwards for a phenomenal amount of help with all matters NMR. And also for the homebrew books! Dr. Stephen Headey for help with several aspects of the project and especially for teaching me practical NMR. Dr. David Libich for help with proof reading and for a great deal of input into the biophysical aspects of the project. The other members of the lab, especially Martin and Hari. The structural biology crowd at Massey especially Matt, Simon, Alice and Greg. My parents Gillian and Jolyon Claridge for their support and who can usually be relied upon to deliberately misunderstand my work in humorous and unpredictable ways. My sister Emma for everything. And finally, my wonderful and long-suffering wife Sarah who was briefly under the impression that a 3 year PhD lasted 3 years...

Glossary of Abbreviations

| | |
|--------------------|---|
| 1D | one-dimensional |
| 2D | two-dimensional |
| 3D | three-dimensional |
| 3C ^B | 3C protease in a previous study |
| 3CI ^B | 3C ^B with inhibitor |
| 3C ^{pro} | 3C protease |
| 3CI | 3C ^{pro} with inhibitor |
| 3CD | covalently linked 3C ^{pro} and 3D ^{pol} |
| 3CIR | 3CI with SLD |
| 3D ^{pol} | 3D polymerase |
| BB1 | β -barrel number 1 |
| BB2 | β -barrel number 2 |
| BMRB | Biological Magnetic Resonance Data Bank |
| BTM | bis-Tris propane / MES |
| CD | circular dichroism |
| CN1-3C | 1st C- and N-labelled sample of 3C produced |
| CPMG | Carr-Purcell-Meiboom-Gill |
| CSA | chemical shift anisotropy |
| CSI | chemical shift index |
| CSP | chemical shift perturbation |
| ddH ₂ O | distilled deionised water |
| DLS | dynamic Light Scattering |
| DTT | dithiothreitol |

| | |
|---------------|--|
| HRV | human rhinovirus |
| HRV-14 | human rhinovirus isotype 14 |
| HRV-2 | human rhinovirus isotype 2 |
| HSQC | heteronuclear single quantum coherence |
| IMAC | immobilised metal ion affinity chromatography |
| IPTG | isopropyl- β -D-thiogalactopyranoside |
| LB | Luria-Bertani broth |
| MES | (2-(<i>N</i> -morpholino)ethanesulphonic acid |
| NOE | nuclear overhauser effect |
| N2-3C | 2nd N-labelled sample of 3C produced |
| R_1 | longitudinal relaxation rate |
| $R_{1\rho}$ | relaxation rate under a transverse spinlock |
| R_2 | transverse relaxation rate |
| RBD | RNA-binding domain |
| RCI | random coil index |
| RDC | residual dipolar coupling |
| R_{ex} | contribution to relaxation from slow exchange processes |
| R_{ex}^{CC} | R_{ex} as determined from cross-correlation analysis |
| R_{ex}^{LS} | R_{ex} as determined from Lipari-Szabo model-free analysis |
| S^2 | order parameter |
| S_{ave}^2 | average order parameter |
| SDS-PAGE | sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| SLD | stemloop-d |
| T_1 | characteristic longitudinal relaxation time |
| T_2 | characteristic transverse relaxation time |
| TROSY | transverse relaxation optimised spectroscopy |

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Introduction | 2 |
| 1.2 | Picornaviruses | 3 |
| 1.3 | Clinical and Societal Significance | 4 |
| 1.4 | Picornaviral Phylogeny | 4 |
| 1.4.1 | Rhinoviruses | 5 |
| 1.5 | Picornaviral Capsid | 5 |
| 1.6 | Viral Genome | 9 |
| 1.7 | Picornaviral Life Cycle | 9 |
| 1.7.1 | Translation and protein processing | 10 |
| 1.7.2 | RNA replication | 12 |
| 1.7.3 | Virion packaging and cellular apoptosis | 14 |
| 1.8 | 3C Protease (3C ^{pro}) | 17 |
| 1.8.1 | RNA binding | 21 |
| 1.9 | Research Aims | 25 |
| 2 | Characterisation of 3C^{pro} in Solution | 27 |
| 2.1 | Introduction | 28 |
| 2.2 | Methods | 29 |
| 2.2.1 | Transformation of BL21 cod ⁺ <i>Escherichia coli</i> cells with HRV-14 3C | 29 |
| 2.2.2 | Unlabelled test expression and production of glycerol stocks of transformed <i>E. coli</i> | 29 |
| 2.2.3 | Glucose optimisation procedure | 29 |
| 2.2.4 | Trials of different affinity media | 30 |
| 2.2.5 | Cleavage of fusion protein and final polishing | 31 |
| 2.2.6 | Inhibitor synthesis | 31 |
| 2.2.7 | Flame spectroscopy | 32 |
| 2.2.8 | NMR sample preparation | 32 |
| 2.2.9 | NMR experiments | 35 |
| 2.2.10 | Structural modelling of 3CI | 36 |
| 2.3 | Results | 38 |
| 2.3.1 | Initial sample NMR | 38 |
| 2.3.2 | Assignment of apo 3C | 39 |
| 2.3.3 | Flame spectroscopy | 44 |
| 2.3.4 | Attempts to recover missing peaks | 44 |
| 2.3.5 | CSI | 46 |
| 2.3.6 | Inhibitor addition | 46 |
| 2.3.7 | Assignment of inhibited 3C | 51 |
| 2.4 | RDCs | 53 |
| 2.5 | Discussion | 54 |
| 2.5.1 | Sample production | 54 |
| 2.5.2 | Implications of structural model | 59 |

CONTENTS

| | | |
|----------|--|------------|
| 3 | Dynamics of Inhibited and Uninhibited 3C | 63 |
| 3.1 | Introduction | 64 |
| 3.2 | Materials and Methods | 65 |
| 3.2.1 | Sample preparation | 65 |
| 3.2.2 | NMR-based dynamic analysis | 66 |
| 3.2.3 | Circular Dichroism (CD) | 69 |
| 3.2.4 | Dynamic Light Scattering (DLS) | 70 |
| 3.3 | Results | 70 |
| 3.3.1 | HSQC Spectra of 3C and 3CI | 70 |
| 3.3.2 | CD spectra of 3C and 3CI | 70 |
| 3.3.3 | DLS and SDS-PAGE analysis of 3C and 3CI | 73 |
| 3.4 | Dynamics Analysis | 73 |
| 3.5 | Discussion | 82 |
| 3.5.1 | 3CI relaxation | 82 |
| 3.5.2 | 3C relaxation and comparison with 3CI | 84 |
| 3.5.3 | RCI analysis of 3C and 3CI | 85 |
| 3.5.4 | Distinct dynamics states of 3C protease | 87 |
| 3.5.5 | Substrate-recognition loop mobility | 89 |
| 3.5.6 | Other active-site-associated dynamics | 89 |
| 3.5.7 | Mobility of RNA-binding surface | 90 |
| 3.6 | Conclusion | 93 |
| 4 | Interaction of 3C with Stemloop-D RNA | 95 |
| 4.1 | Introduction | 96 |
| 4.2 | Materials and Methods | 96 |
| 4.2.1 | NMR sample preparation | 96 |
| 4.2.2 | X-ray scattering sample preparation | 98 |
| 4.2.3 | NMR spectroscopy | 98 |
| 4.2.4 | Calculation of chemical shift indexing and perturbation | 98 |
| 4.2.5 | Docking with HADDOCK | 99 |
| 4.2.6 | Small-angle X-ray scattering data acquisition, reduction, and analysis | 99 |
| 4.2.7 | Shape restoration calculations from scattering data | 100 |
| 4.2.8 | Rigid-body refinement of models against the scattering data | 100 |
| 4.3 | Results | 101 |
| 4.3.1 | NMR data | 101 |
| 4.3.2 | Analysis of NMR results for use with SAXS refinement | 110 |
| 4.3.3 | SAXS data | 111 |
| 4.4 | Discussion | 116 |
| 5 | Conclusions and Further Work | 123 |
| 5.1 | Completed Work | 124 |
| 5.1.1 | Characterisation of 3C in solution | 124 |
| 5.1.2 | Dynamics of 3C ^{pro} in apo and holo forms | 124 |
| 5.1.3 | SAXS structure of complex | 124 |
| 5.2 | Contribution to Current Knowledge | 125 |
| 5.2.1 | Dynamics | 125 |
| 5.2.2 | Allostery | 125 |
| 5.2.3 | Allostery in single domain proteins | 126 |
| 5.2.4 | Possibilities of networks for 3C ^{pro} | 128 |
| 5.3 | Aims of the Project | 128 |
| 5.3.1 | Solution structure and Dynamics | 130 |
| 5.3.2 | RNA-binding surface mapping | 130 |
| 5.3.3 | Complex Structure | 130 |
| 5.3.4 | Instability of complex | 131 |

| | | |
|------------------|--|------------|
| 5.4 | Future Work | 131 |
| 5.4.1 | Conditions to get intermolecular NOEs | 131 |
| 5.4.2 | Dynamics of $3C^{pro}$ | 131 |
| 5.4.3 | Dynamics of complex | 131 |
| 5.4.4 | Crystallisation of complex | 132 |
| 5.4.5 | Higher order interactions | 132 |
| Appendix | | 149 |
| Appendix: | | |
| A | Chemical Shift Tables | 151 |
| A.1 | Chemical Shifts for 3C | 152 |
| A.2 | Chemical Shifts for 3CI | 160 |
| A.3 | Chemical Shifts for 3CIR | 168 |
| B | Modelfree 3CI | 177 |
| C | Relaxation Parameters for 3C and 3CI | 183 |
| D | R_{ex} Protocol | 185 |
| E | HADDOCK | 187 |
| E.1 | Introduction to HADDOCK | 188 |
| E.2 | Set-up | 188 |
| E.3 | Run | 189 |
| E.4 | Analysis | 190 |
| F | 3C Purification Protocol | 193 |
| F.1 | Two (long!) Day 3C purification protocol | 194 |
| G | Samples Produced | 195 |
| H | 3C Alignment | 197 |

List of Figures

| | | |
|------|---|-----|
| 1.1 | Picornaviral Phylogeny | 6 |
| 1.2 | Viral capsid and structural proteins | 8 |
| 1.3 | Types of picornaviral 5' UTR | 9 |
| 1.4 | Overview of picornaviral life-cycle | 10 |
| 1.5 | Polyprotein processing by viral proteases | 15 |
| 1.6 | Putative 3CD polymerisation during VPg uridylation | 16 |
| 1.7 | Picornavirus 3C ^{pro} Structures | 18 |
| 1.8 | Stereo Views of HRV 14 3C ^{pro} | 19 |
| 1.9 | Catalytic triad residues | 20 |
| 1.10 | Mutation data and electrostatics of 3C ^{pro} | 22 |
| 1.11 | SLD sequences | 23 |
| 1.12 | Surface representation of HRV14, coxsackie virus and reterovirus SLD | 24 |
| 2.1 | SDS-PAGE gel of samples from final 3C purification steps | 34 |
| 2.2 | 1D spectra of 3C | 38 |
| 2.3 | ¹⁵ N, ¹ H-HSQC-TROSY spectrum of 3C ^{pro} | 40 |
| 2.4 | 3C ^{pro} spectrum under Bjorndahl et al. buffer conditions | 42 |
| 2.5 | Plot of Residues with Missing Amide Peaks onto Structure of Human Rhinovirus 14 3C ^{pro} | 43 |
| 2.6 | Overlay of full-peak and low-peak spectra | 45 |
| 2.7 | ¹⁵ N, ¹ H-HSQC-TROSY of apo 3C in L-Arg and 50 mM L-Glu buffer | 47 |
| 2.8 | CSI Data for 3C, 3CI and PDB 2IN2 | 48 |
| 2.9 | JMC 99-9 (Also known as AG7088) and 98-3 specific 3C inhibitors | 49 |
| 2.10 | JMC 98-15 and 98-16 specific 3C inhibitors | 50 |
| 2.11 | Overlay of IPAP spectra | 52 |
| 2.12 | Distribution of RDCs | 53 |
| 2.13 | View of active site face of 3C ^{pro} showing acidic residues | 55 |
| 2.14 | Stereo view of residues with missing NH resonances in 3CI. | 58 |
| 2.15 | View of G154 | 59 |
| 2.16 | 3CI and 3CI ^B structural comparison | 61 |
| 3.1 | 3C and 3CI HSQC overlay | 71 |
| 3.2 | Plot of exchange-broadened residues and R_{ex} for 3C and 3CI | 72 |
| 3.3 | CD and DLS results for 3C and 3CI | 74 |
| 3.4 | Model-free and RCI-derived S^2 values for 3C and 3CI plotted onto pdb of 3CI | 76 |
| 3.5 | Relaxation parameters for 3C and 3CI | 77 |
| 3.6 | R_{ex}^{CC} values for 3C and R_{ex}^{CC} and R_{ex}^{LS} values for 3CI | 79 |
| 3.7 | S^2 values for 3C and 3CI | 81 |
| 3.8 | RCI-derived S^2 values for 3C, 3C ^B , 3CI and 3CI ^B | 86 |
| 3.9 | Average RCI-derived S^2 values broken down by secondary structure element | 88 |
| 3.10 | $R_{1\rho}$ -detected exchange in 3CI | 90 |
| 4.1 | Protonated 3CI complexed with excess 27mer | 102 |
| 4.2 | Overlay of 3CI and 3CIR spectra | 103 |

LIST OF FIGURES

| | | |
|------|---|-----|
| 4.3 | CSP from complex of protonated $3C^{pro}$ -SLD | 104 |
| 4.4 | CSI plots for 3C and 3CIR | 106 |
| 4.5 | Pecan Results | 107 |
| 4.6 | Chemical shift perturbation plot for 3CI vs 3CIR | 108 |
| 4.7 | $3C^{pro}$ chemical shift perturbation on binding SLD | 109 |
| 4.8 | $3C^{pro}$ peak intensity change on binding SLD | 109 |
| 4.9 | Scattering data for 3CI and 3CIR I | 112 |
| 4.10 | Scattering data for 3CI and 3CIR II | 112 |
| 4.11 | $3C^{pro}$ -SLD Interaction close-up | 114 |
| 4.12 | Interaction residues from Ohlenschläger and co-workers | 118 |
| 4.13 | Cartoon representation of SLD from coxsackie virus and HRV-14 | 118 |
| 4.14 | Complex showing SAXS envelope | 119 |
| 4.15 | Allosteric “communication pathway” | 121 |
| 5.1 | Allosteric Slice | 129 |
| E.1 | RMSD from lowest energy structure versus HADDOCK score | 191 |
| E.2 | Overlay of four lowest energy structures from cluster 4 | 192 |

List of Tables

| | | |
|-----|---|-----|
| 1.1 | Picornaviral Classification | 3 |
| 2.1 | Table of Glucose Optimisation Parameters | 31 |
| 2.2 | Residues with unobservable NH peaks in $^{15}\text{N},^1\text{H}$ -HSQC of 3CI. | 57 |
| 2.3 | Structure statistics of models | 60 |
| 3.1 | Residues with significant $R_{1\rho}$ values | 80 |
| A.1 | 3C chemical shifts | 152 |
| A.2 | 3CI chemical shifts | 160 |
| A.3 | 3CIR chemical shifts | 168 |
| B.1 | 3CI modelfree parameters | 178 |
| G.1 | Sample preparations with experiments carried out | 196 |

